



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

H.M. Ngwangwa

Road surface profile monitoring based on vehicle response and artificial neural network simulation

Harry Magadhlela Ngwangwa

A thesis submitted in partial fulfilment of the requirements for the
degree

Doctor of Philosophy

in the Department of Mechanical and Aeronautical Engineering
in the Faculty of Engineering, Built Environment and Information
Technology

at the

University of Pretoria

2014

© University of Pretoria



Road surface profile monitoring based on vehicle response and artificial neural network simulation

by

Harry Magadhlela Ngwangwa

Supervisor : Prof. P.S. Heyns
Department : Mechanical and Aeronautical Engineering
Degree : PhD (Mechanics)

Summary

Road damage identification is still largely based on visual inspection methods and profilometer data. Visual inspection methods heavily rely on expert knowledge which is often very subjective. They also result in traffic flow interference due to the need for redirection of traffic to alternative routes during inspection. In addition to this, accurate high-speed profilometers, such as scanning vehicles, are extremely expensive often requiring strong economic justifications for their acquisition. The low-cost profilometers are very slow, typically operating at or less than walking speeds, causing their use to be labour-intensive if applied to large networks.

This study aims at developing a road damage identification methodology for both paved and unpaved roads based on modelling the road-vehicle interaction system with an artificial neural network. The artificial neural network is created and trained with vehicle acceleration data as inputs and road profiles as targets. Then the trained neural network is consequently used for reconstruction of road profiles upon simulating it with vertical vehicle accelerations. The simulation process is very fast and can often be completed in a very short time thus making it possible to implement the methodology in real-time.

The pioneering work for this study was carried out on mine haul roads by Thompson and Visser (2003) where mine haul road management systems were developed and the need for the development



of real-time maintenance and management systems was identified. Thompson, Visser, Miller and Lowe (2003) attempted to lay the foundations of such a real-time management system by employing a vehicle vibration signature analysis technique, but the output was purely qualitative since it could not estimate the extent of the severity of road damage. Later, Hugo, Heyns, Thompson and Visser (2008) developed a methodology for reconstructing the road profiles via an inversion of the vehicle numerical model. Though the approach offered practically acceptable approximations of road damage, it was onerous since it involved complex system characterisation. Consequently the use of black box models, such as an artificial neural network where system characteristics would not be required, was deemed appropriate.

In this thesis, the possibility of using vehicle responses and ANNs to assist in road maintenance management is investigated. Similar to the study by Hugo (2005), the adopted methodology is two-fold: Firstly, can vehicle response data be used to identify specific road defects, in terms of location size? Thus recognition of defect type is outside the scope of the current study. Secondly, can road roughness be qualitatively assessed using the well established roughness classification procedures? Three case studies have been used to investigate and accordingly demonstrate the feasibility of using neural networks for monitoring of road surface conditions. The study is undertaken in three main phases. The first phase consists of development of the methodology and demonstration of its feasibility by using a numerical model which is simulated with artificially generated road profiles. The road profiles are synthetically generated using a theoretical random roughness generating function and half-sinusoids to represent potholes and bumps. Although the operating conditions are changed by varying the vehicle speeds, the payloads and by adding noise at different levels, the methodology yields very high levels of fitting accuracy and correlation between the reconstructed and actual road profiles.

In the second phase, the developed methodology is demonstrated on an experimental Land Rover Defender 110, which is used for research in vehicle dynamics by the Vehicle Dynamics Group at the University of Pretoria. The Land Rover is permanently instrumented with sensors, has carefully controlled suspension properties and, during the investigation, is driven along specially constructed and accurately measured roads at Gerotek (a vehicle-testing facility located close to Pretoria in South Africa). The vehicle's suspension characteristics are switched between ride and handling modes, and the vehicle speeds are also varied between 14.5 km/h and 54 km/h. Different layouts of trapezoidal bumps and Belgian paving are used in the test. This study provides an opportunity to investigate further the performance of the methodology with different vehicle suspensions under controlled experimental conditions. The ANNs are trained by using vertical vehicle accelerations calculated using a numerical model of the Land Rover, simulated for variously altered versions of the test road profiles. The findings show that the quality of the reconstructed bumps is superior to that of the Belgian paving. The neural network performance for the Belgian paving is found to be better represented by their displacement spectral densities (DSDs) than the raw road profiles themselves.

In the third phase, the methodology is demonstrated and implemented on mine haul vehicles where the road surfaces are not as accurately constructed and measured as those at Gerotek. This is to replicate the actual situation in public rural roads and mine haul roads condition monitoring where



accurate profilometers are both inapplicable and unavailable for profile measurement. In addition, there is far less control over the operation of the vehicle than in the experimental Land Rover, and no numerical vehicle models are available. For these reasons, the training data are selected from the test data only. The greatest challenge with this selection is the scarcity of data due to the many unknown random operating conditions. A substantial number of tests and measurements would have had to be performed if the acquired data were to be completely representative of the vehicle's real dynamic behaviour under every possible operating condition, which unfortunately is not the case during the test due to limited vehicle availability. Furthermore, though vehicle control is a key input into the vehicle-road interaction system, there is no simple way to measure it quantitatively. The inconsistencies in the vehicle control over different road conditions introduce drastic variations in the quality of the measured data, in this way affecting the representativeness of the underlying vehicle dynamics. However, these challenges are what make this investigation unique. Despite the challenges, it is found that a large proportion of the tests yield bias errors in the reconstructed profiles of less than 25 % with correlations higher than 50 %.

Therefore through these studies, it is shown that the methodology is practically feasible. The most critical factor though is a properly trained neural network. The training process is like a calibration process and it takes a much longer time than the simulation process, therefore it is recommended that for online implementations, the training should be completed off-line.

Keywords

Artificial neural networks, Non-linear AutoRegressive with eXogenous inputs (NARX), radial basis function neural networks, road profiles, PSD roughness classification, International Roughness Index (IRI), road management systems (PMS), road damage identification, road roughness, inverse modelling, dynamic tyre forces.



Acknowledgements

Foremost, I thank God for giving me the enablement to carry through even when things were seemingly impossible. My wife Mbilire has been extremely supportive. Our kids, Akuzike and Nonhlanhla (Nalesi) were at an age where they could not understand my busy schedules but she managed to communicate that to them in a way they could understand.

Then, a word of gratitude to the following people and institutions whose contributions made this work possible:

- My mentor and promoter Professor Stephan Heyns for his support and technical guidance.
- Professor Grant Keeble Kululanga, for building the initial momentum for this work and providing every necessary support in the initial stages of the research.
- Mr Kobus Labuschagne of the Council for Scientific and Industrial Research (CSIR) (South Africa). The support that he mobilised for this work in its inception was extremely momentous and can never be forgotten.
- Professor Schalk Els of Vehicle Systems Dynamics, University of Pretoria: for the time to accommodate me and guide through vehicle dynamics theory at this level. Messrs. Abraham Breytenbach and Carl Martin Becker deserve my thanks for the information and data they provided for the research.
- Messrs. Hennie Klopper, Mohamed Shiraaz Amade, and Fanus van der Westhuizen all of Anglo-American: Their work can never get by unnoticed. It forms a very important part of this work and they made it possible to carry out the most challenging test at the mines.
- Dr Theo Heyns, for his time and dedication during the field tests and assistance in the neural network modelling.
- Dr Thembalani Sithebe and Professor Moses Strydom, my former and current managers at University of South Africa, Department of Mechanical and Industrial Engineering. They both gave me time to focus on my studies. I remain indebted to them.
- National Research Foundation (SA) for the financial support they provided at the beginning of this project.



Table of Contents

Summary	i
Keywords	iii
Acknowledgements	iv
List of symbols	x
Abbreviations	xiii
List of publications based on this work	xviii
Chapter 1	1
1.1 Introduction	1
1.2 Study objectives	1
1.3 Road deterioration and need for maintenance	2
1.4 Current road management and maintenance systems	3
1.5 The developed methodology	4
1.6 Scope of work	4
1.7 Thesis layout	5
Chapter 2	7
Literature review	7
2.1 Introduction	7
2.2 Real-time road management systems (RTRMS)	8
2.3 ANN-based systems	9
2.3.1 ANNs as diagnostic tools	9
2.3.2 ANNs as prognostic tools	14
2.3.3 ANNs as aids for road maintenance recommendations	21
2.4 Concluding remarks on ANN-based methodologies	23
Chapter 3	25
Road-vehicle interaction	25
3.1 Vehicle ride dynamics	25
3.2 Transmissibility	26
3.3 Determination of road profiles and dynamic tyre forces	31
3.3.1 Road profile calculation	31
3.3.2 Calculation of dynamic tyre forces	34
3.4 Road surface topography	36
3.4.1 Representation of road surface profiles	36
3.4.2 International roughness index (IRI)	39
3.4.3 Summary of important road-vehicle interaction system characteristics	41



3.5	Conclusions	43
Chapter 4	44
Theory of ANN modelling for road-vehicle interaction	44
4.1	Introduction to inverse problem solution	44
4.2	Artificial neural networks (ANN)	45
4.2.1	Definition and architecture	45
4.2.2	ANN in function approximation	48
4.2.3	Back-propagation	49
4.2.4	Levenberg-Marquardt	54
4.2.5	ANN generalisation	57
4.3	ANN architectures and formulations.....	58
4.3.1	Static feed-forward neural network (FFNN).....	58
4.3.2	The nonlinear autoregressive network with exogenous inputs (NARX).....	59
4.3.3	Radial basis function neural networks (RBFNN)	60
4.3.3.1	RBFNN training.....	61
4.4	Comparison of network performances	62
4.5	Model validation	66
4.5.1	Model's root mean square error (RMSE)	66
4.5.2	Correlation tests	67
4.6	Conclusions	67
Chapter 5	68
Development of a generic methodology	68
5.1	Data acquisition.....	69
5.1.1	Vehicle preparation.....	71
5.1.2	Data acquisition system	72
5.1.3	Sensors	72
5.1.4	Road preparation and profiling	73
5.1.5	Test procedure.....	73
5.2	Data preparation	74
5.2.1	Data alignment	74
5.2.2	Removing trends	74
5.2.3	Filtering.....	75
5.2.4	Data normalization.....	75
5.2.5	Decorrelation.....	75
5.2.6	Covariance equalisation	75



5.3	ANN identification	76
5.3.1	Creation of neural network	76
5.3.2	Network training	77
5.4	ANN simulation	78
5.5	Road condition identification	78
5.5.1	Reconstructed road profiles	79
5.5.2	Roughness classification	79
5.6	Methodology verification via case studies	79
5.6.1	First case study: Numerically-generated data	79
5.6.2	Second case study: Experimental Land Rover	80
5.6.3	Third case study: Field tests on mine vehicles	80
5.7	Conclusions	81
Chapter 6	82
Numerical case study	82
6.1	Introduction	82
6.2	Customisation of the proposed methodology to this case study	83
6.3	Theory	87
6.3.1	Vehicle modelling	87
6.3.2	Bayesian-regularized NARX neural network	89
6.4	Description of the vehicle dynamics	90
6.5	Numerical studies	92
6.5.1	General road roughness identification	94
6.5.2	Defect identification under different speeds	96
6.5.3	Identification under changing/growing defects	99
6.5.4	Identification under noisy conditions	101
6.5.5	Identification under different vehicle weight	103
6.6	Conclusions	105
Chapter 7	106
Experimental case study	106
7.1	Introduction	106
7.2	Land Rover properties, suspension and numerical model	107
7.2.1	Geometry and inertial properties	107
7.2.2	The tyre	108
7.2.3	The suspension forces	108
7.2.4	The vehicle model	110



7.3	Nominal road profiles.....	112
7.3.1	Trapezoidal bumps	113
7.3.2	Belgian paving	114
7.4	Generating the training data	115
7.5	ANN architecture and training	117
7.6	Response measurements.....	119
7.7	Simulations, results and discussions	121
7.7.1	Discrete obstacles	121
7.7.2	Belgian paving	125
7.8	Conclusions	128
Chapter 8	130
Field tests on mine vehicles	130
8.1	Introduction	130
8.2	Materials and methods	131
8.2.1	The test vehicles and roads	131
8.2.2	Vehicle instrumentation	133
8.2.3	Road preparation	133
8.2.4	Road profiling and roughness classification	134
8.2.5	Vehicle operation and response data measurements.....	136
8.2.6	Data pre-processing	137
8.2.7	ANN identification and training	137
8.3	Measured data and selection of training data	139
8.3.1	Organisation of the measured data.....	139
8.3.2	Selection of training data	141
8.4	Results and discussions	143
8.4.1.	Haul truck tests	143
8.4.2.	Small utility vehicle for underground mining.....	148
8.4.3	Summary of test results.....	151
8.5	Conclusions	152
Chapter 9	154
Conclusions	154
9.1	Introduction	154
9.2	Detection, quantification and location of road damage as it occurs in real time ...	155
9.3	Better understanding on the application of ANN to function approximation problems	155
9.4	Influence of road and vehicle operating factors affect the identification process...	155



9.5	Influence of vehicle suspension characteristics affect the identification process ...	156
9.6	Summary of key findings from the case studies	156
9.6.1	Numerical case study	156
9.6.2	Experimental case study	157
9.6.3	Mine vehicle test	157
Chapter 10	158
Recommendations for further study	158
References	159
Appendices	176
Appendix A	176
Appendix B	178



List of symbols

When a symbol is shown as small and capitalised, the capitalised denotes amplitude.

b_k^n	Bias for neuron k in layer n
\mathbf{C}	Generalised Damping constant matrix
C_s	Suspension damping constant
C_t	Tyre damping constant
D	Aggregate weighted force
$D(t)$	System transfer function expressed as a differential operator
$E(\)$	The expected value
f	Frequency, Hz
\mathbf{f}, \mathbf{F}	Vector of generalised forces
$f(\cdot), g(\cdot), h(\cdot)$	Neuron activation functions
F_g	Gravitational forces
f_s, F_s	Force on the sprung mass
f_u, F_u	Force on the unsprung mass
\mathbf{H}	Hessian matrix
\mathbf{I}	Identity matrix
\mathbf{K}	Generalised Spring stiffness matrix
K_s	Suspension stiffness constant
K_t	Tyre stiffness constant



j	Imaginary part
\mathbf{g}_i	Function of displacement and velocities imposed on each tyre
\mathbf{M}	Generalised Mass matrix
M_s	Sprung mass
M_u	Unsprung mass
$p(t)$	Neural network input as function of time
$S_U(\kappa)$	Displacement power spectral density
\mathbf{z}, \mathbf{Z}	Vector of generalised displacements
$\mathbf{z}_r, \mathbf{Z}_r$	Road profile displacement input.
z_s, Z_s	Sprung mass displacement
z_u, Z_u	Unsprung mass displacement
\mathbf{S}	Transformation forces for suspension forces
t_i, \mathbf{T}	Neural network targets
\mathbf{T}	Transformation forces for suspension forces
u_i, v_i	First and second hidden layer neuron i output
v	Velocity, m/s
w_{ik}^n	Weight connecting neuron i to neuron k in layer n
$y(t)$	Neural network output as function of time
z_s, Z_s	Mass sprung displacement
z_u, Z_u	Unsprung mass displacement



ε, e	Symbol for error
η_{ut}	Displacements of the tyre attachment points on the unsprung mass
η_{us}	Displacement of the suspension attachment point on the unsprung mass
η_{ss}	Displacement of the suspension attachment point on the sprung mass
κ, κ_0	Wavenumber in cycles/m, Datum wavenumber
ω	Angular frequency
λ	Regularizing parameter
σ	Width parameter of radial basis function



Abbreviations

3D	Three dimensional
3DRT	3D Radon Transform
4S ₄	Four state semi-active suspension
AAE	Average Absolute Error
AASHTO	American Association of State Highway and Transportation Officials
AASHTO TP-62	Standard Method of Test for Determining Dynamic Modulus of Hot-Mix Asphalt Concrete Mixtures
A-In	All inputs considered
ANN	Artificial Neural Networks
ASA	Anti-Striping Agents
BLW	Band-Limited White
CBR	California Bearing Ratio
Corr.	Correlation
CPU	Computer processing unit
DAP	Distributed Array of Processors
DBP	Deflection Basin Parameter
DOF(s)	Degree(s)-of-freedom
DSD	Displacement Spectral Density
DSTI	Directorate for Science, Technology and Industry
DTPS	Dynamic Tire Pressure Sensor
ENOP	Effective number of training parameters



ELP	Elastic Layered Program
EU	European Union
FDOT	Florida Department of Transportation
FAA	Federal Aviation Administration
FEA/M	Finite Element Analysis/Method
FFNN	Feed-Forward Neural Network
FFT	Fast Fourier Transform
FWD	Falling Weight Deflectometer
GEP	Gene Expression Programming
GNSS	Global Navigation Satellite System
GP	Genetic Programming
GDP	Gross Domestic Product
GPR	Ground Penetrating Radar
GVM	Gross Vehicle Mass
HDD	Hard Disk Drive
HDM	Highway Development and Management
HWD	Heavy Weight Deflectometer
IRF	International Road Fund
ISO	International Organization for Standardization
ITS	Indirect Tensile Strength
JPCP	Jointed Plain Concrete Roads



KDOT-POS	Kansas Department of Transportation Project Optimisation System
LPP	Linear Pitch Plane
LSSVM	Least Squares Support Vector Machine
LTPP	Long-Term Road Performance
LVR	Low Volume Roads
MAE	Mean Absolute Error
MARE	Mean Absolute Relative Error
MC-HARP	Monte-Carlo Hierarchical Adaptive Random Partitioning
MEPDG	Mechanistic-Empirical Road Design Guide
MLP	Multi-Layer Perceptron
MLR	Multivariate Linear Regression
MPI	Maintenance Priority Index
M&R	Maintenance and Repair
NAPTF	National Airport Road Test Facility
NARX	Nonlinear Auto-Regression with eXogenous inputs
NDT	Non-Destructive Test
NSV	Network Survey Vehicle
OECD	Organization for Economic Co-operation and Development
OSDOT	Oregon State Department of Transportation
PALSAR	Phase Array Type L-band Synthetic Aperture Radar
PCI	Road Condition Index



PCR	Road Condition Rating
PMS	Road Management Systems
PSD	Power Spectral Density
PSI	Road Serviceability Index
RAP	Recycled Asphalt Road
RBFNN	Radial Basis Function Neural Network
RMSE	Root Mean Square Error
R-O	Road Only
RTRMS	Real-Time Road Management Systems
RTRRMS	Response-Type Road Roughness Measuring Systems
SCE	Shuffled Complex Evolution
SHRP	Strategic Highway Research Program
SIF	Stress Intensity Factor(s)
NSR	Noise-Signal Ratios
SPS	Specific Road Study
SSE	Sum of Squared Errors
SSW	Sum of Squared Weights
SVM	Support Vector Machines
TSR	Tensile Strength Ratio
UK	United Kingdom
UML	Unified Modelling Language



VCI	Visual Condition Index
VLSI	Very-Large-Scale-Integrated
VIMS	Vehicle Information Management Systems
WB	Wheelbase



List of publications based on this work

Articles in refereed and peer-reviewed journals

1. Ngwangwa, H.M., Heyns, P.S. (2014). Application of an ANN-based methodology for road surface condition identification on mining vehicles and roads. *Journal of Terramechanics* 53: pp 59-74. ISSN 0022 – 4898.
2. Ngwangwa, H.M., Heyns, P.S., Breytenbach, H.G.A., Els, P.S. (2014). Reconstruction of road defects and road roughness classification using Artificial Neural Networks simulation and vehicle dynamic responses: Application to experimental data. *Journal of Terramechanics*, 53: pp 1 - 18.
3. Ngwangwa, H.M., Heyns, P.S., Labuschagne, F.J.J., Kululanga, G.K. (2010). Reconstruction of road defects and road roughness classification using vehicle responses with artificial neural networks simulation. *Journal of Terramechanics*, 47: pp 97–111.

Conference Proceedings

1. Heyns, P.S., Ngwangwa, H.M., Heyns, T., Van der Westhuizen, S. (2012). e-Monitoring for haul road maintenance in mining applications. *e-Maintenance 2012*. Luleå Sweden 12-14 December 2012.
2. Heyns, P.S., Stander, C.J., Heyns, T., Wang, K. and Ngwangwa, H.M. (2012). Vibration based condition monitoring under fluctuating load and speed conditions. *18th World Conference on Non-destructive Testing*. Durban, South Africa, 16 – 20 April 2012.
3. Ngwangwa, H.M., Heyns, P.S., Labuschagne, F.J.J., Kululanga, G.K. (2008). Monitoring of heavy vehicle dynamic response for road condition monitoring: An overview of the web based artificial intelligence technique. *Southern African Transport Conference (SATC2008)*, Pretoria, 7-10 July 2008.
4. Ngwangwa, H.M., Heyns, P.S. (2011). Comparative evaluation in the performance of NARX and GRNN on Road profile reconstruction. *Second African Conference on Computational Mechanics – An International Conference – AfriCOMP11*, January 5 – 8, 2011, CapeTown.

Book Chapters

1. Ngwangwa, H.M., Heyns, P.S., Breytenbach, H.G.A., Els, P.S. (2011). Validating road profile reconstruction methodology using ANN simulation on experimental data. *Optical Measurements, Modeling, and Metrology*, Volume 5, T. Proulx (ed.), The Society for Experimental Mechanics, Inc. 2011.



Chapter 1

Introduction

1.1 Introduction

This study is aimed at developing a road damage identification methodology based on road profile reconstruction using artificial neural networks (ANNs) and vehicle responses. ANNs are initially trained with vertical vehicle accelerations (as inputs) and road profiles (as targets) obtained from a number of representative road and driving conditions. Then vertical vehicle accelerations from 'unseen' roads are applied to the trained ANN to reconstruct corresponding road profiles. This methodology is expected to provide an inexpensive and faster procedure for quantitative assessment of road surface conditions. The results from the methodology can either be implemented directly on real-time road management systems (RTRMS) or can be utilized as part of the input into more comprehensive road management systems.

This work is motivated by a study on mine haul roads by Thompson, Visser, Miller and Lowe (2003) where they use ANNs to assist in qualitative identification of defects for specific mine haul roads. They realize that there is a need for the ability to analyse, recognize and interpret various forms of the same defect signature and also a need for the use of multi-sensor data to isolate road surface defects from different sources. Part of the response to this need is carried out by Hugo (2005) who develops a methodology for reconstruction of road profiles by inversion of vehicle ride model. The methodology is successfully applied to a mine haul road; however its reliance on extensive system characterisation limits its practical feasibility. System characterisation necessitates that the road maintenance manager designs and conducts some technically demanding laboratory test procedures to obtain the vehicle system characteristics, if they are not available from the vehicle manufacturers. Besides, the calculation of the road profiles involves demanding numerical integration routines. Thus the application of ANN as presented in this study is deemed rather more practically prudent, since it only requires the acquisition of vehicle dynamic responses (in the form of vertical accelerations). Though the use of ANNs may not provide enough physical insight into the behaviour of the vehicle-road interaction system, it nevertheless offers significant advantages by eliminating the need for the characterisation and calculation of an inverse model. The approach has a further advantage that once the neural network is trained; the process of simulating the ANN is much less computationally demanding and can easily be implemented for real-time road condition monitoring.

1.2 Study objectives

The main goal of the study is to investigate the possibility of using vehicle response data and ANNs for reconstruction of road profiles thereby aiding road maintenance management. This overall goal is broken down into four specific objectives namely:



- Detect, quantify and locate road damage as it occurs in real time.
- Enhance better understanding on the application of ANN to function approximation problems.
- Determine how various road and vehicle operating factors affect the identification process.
- Determine how vehicle suspension characteristics affect the identification process.

1.3 Road deterioration and need for maintenance

New paved roads deteriorate very slowly and almost imperceptibly in the first 10 to 15 years of their life and then deteriorate much more rapidly unless timely maintenance is undertaken (Paterson, 1987). Thus all over the world, road agencies have to deal with aging road transport infrastructure that have to be kept productive through sustained efforts of maintenance. This is further complicated by the fact that traffic volumes and axle loadings are increasing annually due to the increased demands of modern business. For example, surveys carried out in the UK by the Audit Commission (Atkinson, 1990) reveal that there is a growing trend towards usage of larger and heavier vehicles due to the shift from traditional rail transport to road transport. The study also shows that spending on local roads fell by 5% whereas traffic increased by 20%. These changes threaten to shorten the service life of roads and necessitate an increase in the cost and frequency of maintenance needed to keep the roads serviceable.

Significant recognition of the importance of maintenance in road management systems only started to emerge from the 1980s when most governments of the developed world felt some budgetary pressures to match with the rate of road deterioration caused by the growing volumes of traffic on the roads (Peshkin, Hoerner, and Zimmerman, 2004). There was increased emphasis on road preservation and preventive maintenance concepts and programs. Governments and road agencies began to evaluate the cost-effectiveness of their maintenance techniques. Harral and Faiz (1988) identified that the loss of approximately US\$ 45 billion in road infrastructure in the 1970s and 1980s could have been averted by spending just US\$ 12 billion on preventive maintenance. Postponement of road maintenance resulted in high direct and indirect costs. If defects are neglected, an entire road Section may fail completely, requiring full reconstruction at three times or more the cost, on average, of maintaining the existing road infrastructure. The South African National Road Agency Limited (SANRAL) estimates that repair costs rise to six times maintenance costs after three years of neglect and to 18 times after five years of neglect (Burningham and Stankevich, 2005). Neglected roads steadily become more difficult to use, resulting in increased vehicle operating costs (more frequent vehicle repairs, more fuel use as well as reluctance by transport operators to use the road). This imposes a heavy burden on the economy as passenger and freight services are curtailed, there is a consequent loss of economic and social development opportunities.

Road deterioration in general, covers a wide spectrum of faults that include degradation of drainage systems, retaining walls, road markings and road shoulders, but this study focusses on road surface deterioration. The major distress modes in road surface are fatigue cracking, permanent deformation (rutting), shrinkage cracking, disintegration and loss of skid resistance (AASHTO, 2001). Fatigue cracking and permanent deformation are usually associated with repeated loading; shrinkage cracking is associated with temperature variations; disintegration might be due to traffic, environmental or



impact loading; and the loss of skid resistance is usually associated with traffic effects and seasonal changes.

Road surface deterioration impinges upon the overall road functionality by lowering surface integrity, structural capacity, ride quality and surface friction. Thus vehicle and/or human safety cannot be guaranteed on such roads. Ogden (1997) reports that in Britain, about 15 % of all vehicles involved in injury accidents have skidded prior to collision and these accidents account for 41 % of vehicle user deaths and 30 % of serious injuries. It is further noted that improved skid resistance of roads could have prevented up to 70 % of these cases. On economic effects, quality delivery and normal traffic flow are negatively affected on deteriorated roads. Chesher and Harrison (1987) allege that such roads lead to increased running costs per unit output per vehicle in the short term. Long term effects depend on the vehicle operators' adjustments in response to the changes in the road conditions.

Although the need for maintenance is widely recognized, it is still not getting adequately done. Many countries spend just between 20 and 50 percent of what they should be spending on maintenance of their road network (Burningham and Stankevich, 2005). Heggie (2003) estimates that less than half of the required expenditure to prevent further deterioration is being met and that the required increases are on average 0.85 per cent of GDP. On average the quality of main road networks is extremely low, with not less than 25 per cent of paved and 33 per cent of unpaved roads in poor condition. The situation is even worse for feeder road networks, with about 90 per cent of roads in poor condition, thereby severely penalizing remotely located agricultural, industrial and mining production.

1.4 Current road management and maintenance systems

Current practices in road maintenance depend on several factors like, type of national economy, type and grade of road network, specific area of location, and political expediency. Most developing economies are still employing breakdown road maintenance even on their principal highways owing to lack of vital road management tools, like decision-making tools (Mushule and Kerali, 2001). Road maintenance is carried out on the subjective authority of some very experienced road engineers plus support from political officials. In most cases, this results in delayed maintenance action such that the road no longer requires mere maintenance and repair (M&R) actions but full rehabilitation for it to remain cost-efficient. In developed economies, some variations of preventive maintenance are in use especially on principal road networks but feeder road networks and road networks servicing some agricultural, manufacturing and mining hauling operations located in remote areas, largely remain unsatisfactorily maintained (Brushett, 2005).

The Highway Development and Management (HDM) model was developed to optimize the economic benefits of road maintenance. However, it cannot be applied directly to any local operating conditions for the first time without some sort of calibration (Thube, 2012). Secondly, most developing countries view the HDM model largely as the World Bank's tool for evaluating the viability of road maintenance, rehabilitation and construction projects. Thus it is not as popular as conventional visual



inspections and use of expert assessments and recommendations, especially at project level management. Even where the HDM model is popular, it still requires the road condition parameters obtained from a good road inventory, which implies that it can potentially benefit from a system such as the one developed in this study that provides valuable road data like road profiles and surface roughness classes.

1.5 The developed methodology

This study develops a road damage identification methodology in which ANNs are trained with vertical vehicle accelerations (inputs) and road profiles (targets) with the intention that the trained network can later be used for reconstructing road profiles from different driving operations. The study seeks to capitalize on the popularization of vehicle information systems, where sensors are increasingly being mounted on vehicles for assessing vehicle performance and the structural integrity of suspension. González, O'Brien, Li and Cashell (2008) argue that despite major improvements in recent years in the quality of road-profiling equipment, profiling devices remain generally expensive to purchase, their use of time is inefficient and their operation is highly specialized. Hugo, Heyns, Thompson and Visser (2008) note that the management and scheduling of maintenance on mine haul roads have not been widely reported in the literature and that the management of such roads is often inadequate, leading either to over-maintenance or a failure to recognize significant road deterioration.

The whole methodology is broken down into five main parts: data acquisition, data preparation, ANN identification, ANN simulation and road condition identification. Data acquisition involves the selection and preparation of the vehicles, roads and measurement equipment used during the test as well as detailed description of the testing procedures. Data preparation comprises various data pre-processing techniques prior to ANN training. ANN identification involves network creation, training, validation and testing. The training data are divided into three groups where one part is used for training and the other two parts are used for validating and testing so that the network achieves good generalization over new and 'unseen' inputs. During ANN simulation, the trained ANN is simulated with vehicle accelerations (similarly pre-processed as for the training data). In road condition identification, either statistic values calculated from the reconstructed road profiles or road roughness classes determined through plots of displacement spectral density (DSD) are used. It is recommended in this study that for discrete faults, statistic values such as RMS values should be used, while DSD classification should be used in identifying surface condition of generally degraded random rough surfaces. Comprehensive discussions on each part are presented in the Chapter 5.

1.6 Scope of work

The developed methodology can be applicable to both paved and unpaved roads. This is evidenced by successful applications to a Land Rover and small utility mine vehicle which are tested on both paved and unpaved roads and the haul truck which is tested on the unpaved roads only.



The methodology cannot be used for accurate road profile inspection of random rough road surfaces where the required measurement resolution may be within a few millimeters. It can however be used for fairly good estimation of profile heights for discrete faults such as road bumps and potholes. The methodology has been successfully applied to a number of characteristically similar vehicles.

1.7 Thesis layout

The thesis is organised as shown in the flowchart in Figure 1.1. A brief discussion of each of the parts and blocks in the flowchart follows.

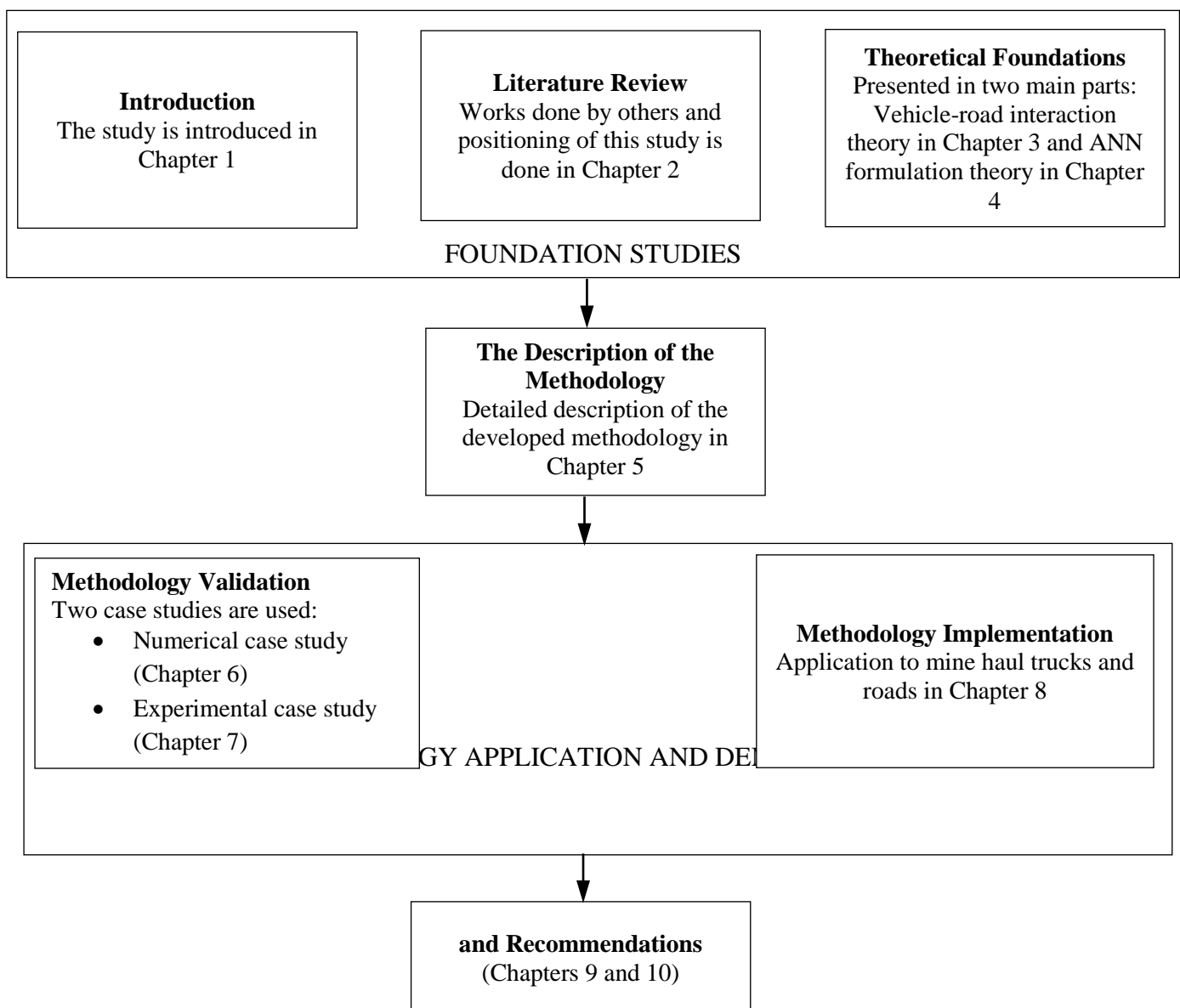


Figure 1.1. Organisation of the thesis.



The first part of the thesis covering Chapters 1 to 4 forms the foundational framework. Chapter 1 introduces the study by presenting the motivation, objectives and scope for the work. The Chapter discusses the adverse effects of road deterioration and benefits of road maintenance. The current practice in road maintenance is presented and its shortcomings are highlighted. The methodology is also briefly presented. In Chapter 2, a detailed review of the literature is presented. The literature categorises current ANN-based road damage identification methodologies into three main groups: methods that focus on damage detection, methods that focus on damage prediction and those that use ANNs as aids to maintenance decisions.

Theoretical foundations are developed and presented in Chapters 3 and 4. Road-vehicle interaction theory is presented in Chapter 3. The Chapter discusses vehicle ride dynamics using a quarter car model and formulates a procedure for calculating road profiles by inverting the dynamic equations. The Chapter further compares the road-unsprung mass and road-sprung mass transmissibility ratios to reinforce the usage of unsprung mass accelerations for this work. The Chapter also discusses the calculation of dynamic tyre forces. Chapter 4 presents the theoretical concepts governing ANN creation, training and simulation. An MLP as well as the back-propagation training scheme are presented. This is extended to Levenberg-Marquardt routine which is popularly used in function approximation problems. Three different ANN models are presented and compared in terms of their performances in reconstructing common types of waveforms in order to motivate the choice for the types of neural networks that have been used in the study.

The second part of the thesis is the development of the methodology. Chapter 5 reports a detailed account of the developed methodology as an integral unit. It gives a detailed discussion of each of its five main parts and how they relate with each other in the overall process of road damage identification.

The third part comprises application and demonstrations of the methodology using three different case studies. The results from these case studies form the central part of the publications from this study. The first case study is a purely numerical demonstration where the effects of a number of factors on the performance of the methodology are examined (Chapter 6). The second case study is an experimental case study which validates the key findings from the numerical case study and further establishes the different performances over discrete obstacles and random rough surfaces and help to redefine how the developed methodology can be applied in dealing with these different road defects. The case study also investigates the problem of different vehicle suspensions. The third case study is presented in Chapter 8 and reports on the findings of real life application of the methodology on two mining sites. It is applied on a small utility mine vehicle at one mining site and a mine haul truck at another site. The methodology is consequently implemented on the mine haul trucks. However, this thesis does not include the implementation procedures due to proprietary issues. The thesis is concluded in Chapter 9 and recommended topics for future work are dedicated to Chapter 10.



Chapter 2

Literature review

2.1 Introduction

There is a direct link between road condition and vehicle operating costs, and effective preservation of road condition through good road maintenance practices contributes positively towards reliable transport at reduced cost. The selection of an appropriate maintenance strategy is dependent on the correct identification of road condition. The condition and pattern of deterioration of a road surface are characteristically dependent on whether the road is paved or unpaved. Researchers have developed a number of different methodologies to assist in determining the appropriate timing for road maintenance for both paved and unpaved roads. This Chapter reviews these works without differentiating between the methods applied to the two road types in order to focus on the subject matter of ANN application to road identification.

In spite of all this research effort, most road agencies are still dependent on the conventional methods: routine, periodic or ad hoc maintenance methods, all of which are heavily dependent on visual inspections and subjective panel ratings (Thompson and Visser, 2003). This over-reliance on visual inspections and panel ratings has a number of widely published shortcomings. The entire exercise is labour-intensive, tedious, and largely subjective, often resulting in (1) a high degree of variability and systematic error among personnel, (2) difficulties in transferring expertise, and (3) inconsistent evaluation.

Ideally, a system that can capture road condition data during normal vehicle operation, relay such information to a centralised database management system and apply the data to a road identification scheme as required can offset such difficulties. This is specifically beneficial to organizations that operate vehicles as well service their own roads (privately owned roads) such as mining and farming organizations. It can also address the difficulties of fragility often faced with the standard road profiling equipment. The sensors can typically be mounted on the vehicle and data acquisition can proceed without interrupting normal traffic flow or worry of redistribution of traffic to other routes. Such a system offers the most practical way for the development of RTRMS. The current proliferation in the development of vehicle information management systems (VIMS), where sensors are increasingly being mounted on the vehicle for monitoring of vehicle vital signs, makes the development of RTRMS feasible.

There are a number of studies that have investigated the development of such systems, but this Chapter reviews only those that are based on artificial neural networks.

2.2 Real-time road management systems (RTRMS)

As already mentioned in this Chapter, RTRMS can offset most of the disadvantages of traditional methods in road maintenance management that are often based on visual inspections and expert panel ratings. It becomes even more relevant with the trend where budgets assigned to construction of new transport infrastructure are being reduced in most economies including some of the developing economies. For example, a recent study by the IRF Research Council (2007) reports that public investment in transport infrastructure in the European Union (EU) fell from 1.5% of GDP in the 1980s to less than 1% in 2004. Thus the importance of procedures that ensure prolonged lives of the existing infrastructure becomes more and more evident.

There are a number of attempts at development of RTRMS reported in literature. For the purpose of this study, existing RTRMS have been classified into two groups (as shown in Figure 2.1):

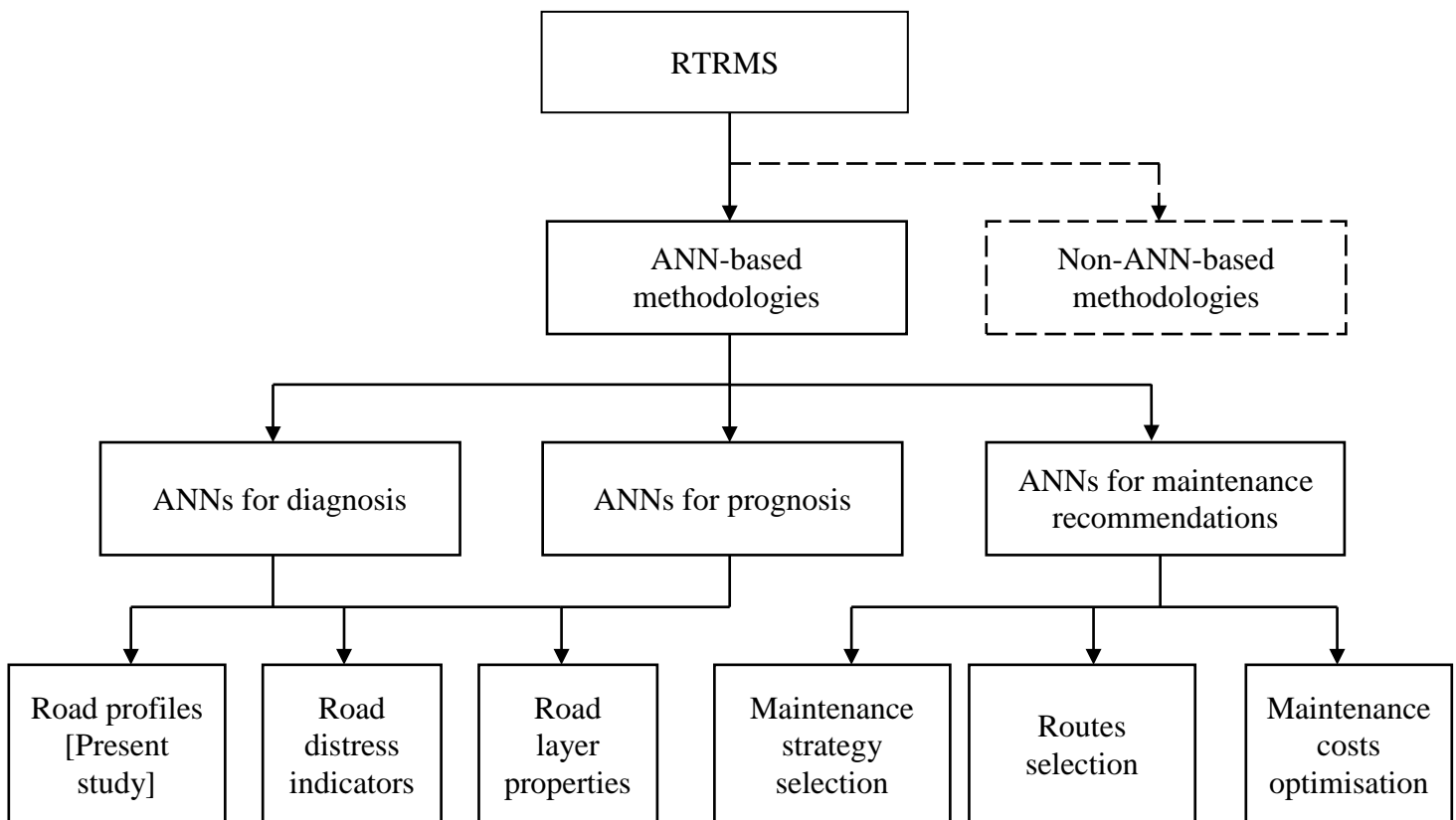


Figure 2.1. Organisation of the literature review.

ANN-based systems and non-ANN-based systems. Owing to their relevance to the present study, ANN-based systems are the focus of this review, and they are further classified according to their functionality as ANNs for diagnosis (estimation of present road condition parameters), ANNs for prognosis (prediction of future road condition parameters including service life) and ANNs for maintenance decisions. ANNs for diagnosis and prognosis have been largely organised into two groups of studies: those that use derived road distress indicators such as IRI and PCI and those that use evaluation of road layer properties such as layer modulus. A few other studies under the diagnostic



functionality also use ANNs for reconstruction of road profiles to which the present study belongs. Studies where ANNs are used as aids for maintenance decisions are organised into three separate groups, namely: maintenance strategy selection, route selection and maintenance costs optimisation. Figure 2.1 summarises how the literature is presented in this thesis.

2.3 ANN-based systems

Studies have shown that applications of ANNs to road maintenance management have some benefits with respect to development of real-time systems. This is due to the fact that once a neural network is developed and trained; the consequent simulation process is rapid and not as involving as the procedures implemented in most conventional systems which often involve either optimisation or iterative search algorithms. The studies reviewed in this Chapter show that ANN-based procedures are faster than the mechanistic- and FE-based models, more accurate than the statistical/regression models, and more robust and consistent than expert-knowledge-based systems. As already mentioned, ANN-based methodologies have been applied to real-time road management systems (RTRMS) at three main levels: estimation of current road condition (diagnosis), prediction of future road condition (prognosis), and assessment of maintenance needs as well as selection of maintenance strategies. The first two applications are consistent with provision of road condition monitoring functionality. The following Sections present each of these categories of ANN applications to RTRMS guided by the structure given in Figure 2.1.

2.3.1 ANNs as diagnostic tools

This Section presents studies where ANNs are used to estimate prevailing road conditions given certain measured operational/environmental conditions. The reported studies have been broadly classified into three groups, as methods based on: ANN-estimated road distress indicators/statistics, ANN-reconstructed profiles and ANN-estimated road layer properties. There are however other studies that do not fit into any of these categories and they have been presented at the end under a separate Section entitled “Others”. This also applies to studies under prognosis in Section 2.3.2.

2.3.1.1 *Estimation of road distress indicators/statistics*

The main rationale behind the development of these methods is either to automate the distress identification procedure which is highly manual and subjective due to its over-reliance on expert knowledge and interpretation, or to find alternatives for more computationally demanding identification schemes. Investigation of neural networks as replacements for subjective expert ratings constitutes most of the earliest studies. For example, Eldin and Senouci (1995) use a feed-forward neural network as a substitute and possible replacement for a more rigorous mathematical model developed by the Oregon State Department of Transportation (OSDOT). The OSDOT model consists of a long, stepwise calculating algorithm based on cracking, and rutting indices on alligator, transverse cracks, block cracks, patching, bleeding, and rutting distresses. Their results indicate that



the ANN outperforms the OSDOT model in estimating subjective condition ratings performed by the maintenance personnel for similar portions of two interstate highways.

Banan and Hjelmsted (1996) show that a Monte-Carlo Hierarchical Adaptive Random Partitioning (MC-HARP) algorithm (a type of neural network) outperforms the well-established formula developed by the American Association of State Highway Officials (AASHO) in estimating road durability of road Sections. This is later confirmed by Terzi (2007) who achieve similar results in modeling the Road Serviceability Index (PSI) of flexible roads. He alleges that the proposed approach can be easily and realistically used to replace more complicated models for solving complex problems. Heiler and McNeil (1997) propose the ANN as a replacement for the onerous manual segmented procedure for road condition assessment using GPR data. Van der Gryp, Bredenhann, Henderson and Rohde (1998) use a one-hidden-layer feed-forward neural network to estimate the Visual Condition Index (VCI) which is normally calculated from visual road condition data and requires the selection of a weighting factor for each type of distress.

Attoh-Okine (2001) demonstrates the use of ANN self-organizing maps for the grouping of road condition variables in developing road performance models to evaluate road conditions on the basis of road distresses. Xiong (2001) uses analysis of sort functions of ANN and the diagnostic technique of road surface distress, to perform diagnostic function on roads. A most recent study is reported by Zakeri, Nejad, Fahimifar, Torshizi and Zarandi (2013) who replace expert rating with a system based on wavelet modulus and 3D Radon Transform (3DRT) for road distress detection and classification. The features and parameters of the peaks are finally used for training and testing the ANN classifier. The proposed approach yields very good performance and the researchers conclude that this high performance demonstrates the advantages of this method in correct classification of road cracking.

Though visual inspection is still a popular method of acquiring road surface condition data, it is highly unreliable due to its proneness for inconsistency from one data collector to another. Thus, as part of the initiative for automation of road damage identification procedure, road images have been collected and applied to ANN models. Kaseko et al. (1991-1994) use MLP neural networks to identify transverse, longitudinal, alligator, and block cracks through image processing on road images for the U.S. Strategic Highway Research Program. Later, Nallamothu and Wang (1996) use the radial basis neural network on the same images to identify similar cracks and they find that 85 % of the patterns are identified correctly. Garrick, Kalikiri and Achenie (1994) develop an automated system using the pattern recognition capabilities of ANNs to characterize common road distress features like cracks, patches and potholes.

Tillotson, Snaith and Huang (1998) also apply a massively-parallel computer system called a Distributed Array of Processors (DAP) in which 1024 processors work simultaneously on a digitized road image to determine whether or not there are cracks in an image in real-time. They show that they are able to achieve about 80% success. Recently, Mokhtarzade and Zoj (2007) apply the approach to high-resolution satellite images. Xu, Ma, Liu and Niu (2008) exploit the self-studying feature of ANNs to identify surface cracks from real road images. By converting cracking recognition to the



cracking probability judgment for every sub-block image, cracking trend is calculated, and a method for revising the neural network output is shown to increase accuracy of identification. Yu and Salari (2011) use ANNs to develop an efficient and more economical approach for road distress inspection from laser image data. García-Balboa, Reinoso-Gordo, and Ariza-López (2012) use ANN to develop a method and tool for automatically evaluating road quality. It is found that the ANN achieves over 70% correlation with expert scores for all the test cases.

A number of recent studies have also compared performances between one type of neural network and another or neural network models and regression models. For example, Fajardo (2011) compares the performance of Support Vector Machines (SVM) and ANN as tools for determining the visual condition of roads on traffic data. The results show that, with data of similar dimensions to the one used in creating it, the ANN performs better classification than SVM. Shahnazari, Tutunchian, Mashayekhi and Amini (2012) show that an ANN-based model is more precise than a GP-based model in estimating PCI. Chandra, Sekhar, Bharti and Kangadurai (2013) compare the performances of an artificial neural network with linear and nonlinear regression models by investigating the relationships between road roughness and distress parameters such as potholes, raveling, rut depth, cracked areas, and patch work. The road distress data collected on four national highways in India using a network survey vehicle (NSV) are used to develop linear and nonlinear regression models between roughness and distress parameters. Analysis of variance of these models indicates that the nonlinear model is a better estimator than the linear model. R^2 value, root mean square error (RMSE), and mean absolute relative error (MARE) also yield similar results. Then an artificial neural network with five input nodes, 15 hidden nodes, and one output node is developed and trained with 90% of the data and tested with remaining 10% data. When the performance of the ANN model is compared with that of linear and nonlinear regression models, it is found that its mean absolute error (MAE) is around 18% less than that for the linear model and 11% less than that for the nonlinear model, and its MARE values are 12.5% lower than both linear and nonlinear regression models which implies that the ANN model performs much better in estimating the road roughness from the given set of distress parameters.

Therefore, through the foregoing studies, ANNs show much greater possibilities in the development of automated condition monitoring process and identification of road damage than the conventional methods. Since, in this application, ANNs are primarily used as pattern classifiers the road distresses are merely identified qualitatively and average statistic values are typically used. Quantification of the extents of particular defects, such as depths of cracks or ruts, can be very hard without implementation of additional data and/or advanced data processing methods, which can result in increase in the computational overload.

2.3.1.2 Reconstruction of road profiles

This work is specifically applicable to identification of road roughness and the ANNs are used as function approximators. Roughness identification is not only performed qualitatively but also



quantitatively by calculating road profile heights at specific points along the road. The calculated road profiles can be used for simulation of numerical vehicle models.

Literature has a lot of road damage assessment systems based on measured or model-generated road profiles however; there is not much that is reported on measurement processes or simulations whose operation is governed by ANN modeling. Besides, the present study, there are two other groups of workers that have used this approach. At the same time as the initial publication of this study (Ngwangwa, et al., 2008) results of a similar study were reported by González, et al., 2008. They use simulated data to demonstrate the feasibility of the methodology. Much later, another group of workers (Yousefzadeh, Azadi and Soltani, 2010) confirm these findings by applying ANNs to ADAMS-simulated data.

These systems have the advantage of being able to provide quantitative assessments of the identified road defects. It is also easy to investigate road-dependent effects using these systems. However, the methodologies cannot be used for identification of cracks and slipperiness unless road image data are integrated into the identification scheme.

2.3.1.3 Estimation of road layer properties

The conventional method for determining the condition of road layer moduli involves time consuming back-calculation procedures which are carried out for each road layer. Thus the primary goal of using ANNs is to alleviate the computational demands. Gucunski, Abdallah and Nazarian (2000) develop an ANN for back-calculation of road properties in real time from test data. The network is developed using synthetic dispersion curves from a numerical simulation of the test. Saltan, Tiğdemir and Kardeş (2002) successfully apply ANNs in calculation of road layer thickness with greater improvements in computational time and accuracy. Sakhaeifar, Underwood, Kim, Puccinelli and Jackson (2011) use ANNs to estimate the dynamic modulus of HMA layers in long-term road performance (LTPP) test Sections. A large national data set that covers a substantial range of potential input conditions is utilized to train and validate the ANNs. The data consist of mixture dynamic moduli measured with two test protocols: the asphalt mixture performance tester and AASHTO TP-62, under different aging conditions. It is found that approximately 30% to 50% of all LTPP layers contain enough information for reliable moduli estimations. Saltan, Uz and Aktas (2013) use synthetically derived deflections from a typical flexible road to estimate asphaltic concrete layer's elastic modulus, Poisson's ratio and thickness. Furthermore, ANN is successfully used to determine layer elastic modulus, Poisson's ratio and layer thickness.

Sometimes ANNs have been used in conjunction with the conventional mechanistic-models in order to enhance understanding of the underlying road physical behavior. Meier, Alexander and Freeman (1997) combine ANNs and conventional layer moduli back-calculating programs to estimate layer moduli for flexible roads. A suite of neural networks, covering a range of flexible road structures, are trained using numerically generated data. When tested on 110 experimental data sets the approach



yields accurate and much faster solutions than the conventional algorithms. Xu, Ranjithan and Kim (2002) examine the sensitivity of various deflection basin parameters (DBPs) to layer conditions. The DBPs identified from the sensitivity study are used in developing new relationships between the selected condition indicators and FWD deflections by applying regression and ANN techniques.

Xie, Guo and Cong (2007) present a methodology for processing dynamic FWD measurements using DBPs and ANNs to estimate road layer moduli and condition. A two-dimensional dynamic FEA is formulated from the deflection information by using FORTRAN. This work is an extension of a study done by Kim, Lee and Ranjithan (1999) who use ABAQUS for dynamic FEA. In these methods, two DBPs (Base Damage Index and Shape Factor) are initially determined to calculate the subgrade modulus, then the estimated subgrade modulus plus other parameters are used as input variables to a trained ANN for estimation of upper layers' moduli. Field test data are used to validate the feasibility of the technique.

Saltan and Sezgin (2007) use ANN to model the unbound material behavior of sub-base layer from experimental data and FEM. Experimental deflection data groups from NDT are used to show the capability of the approach. Kim and Gopalakrishnan (2010) use DBPs and a hybrid Shuffled Complex Evolution (SCE)-ANN for characterisation of structural condition of rubblized concrete roads using NDT deflection measurements.

In order to enhance the computational capabilities of ANNs, some studies have investigated the combined use of ANNs with genetic algorithms (GAs). Saltan et al. (2003 – 2005) show that using ANN and GEP provides significant reduction in computational time and further conclude that the approach can be used to solve complex optimization problems. Rakesh, Jain, Reddy and Reddy (2006) combine ANNs and GAs to compute surface deflections in which elastic moduli and thicknesses of road layers are used as inputs. Pekcan, Tutumluer and Ghaboussi (2009) use an approach based on a set of algorithms derived from the use of ANNs and GAs to determine layer properties from NDT road testing and evaluation. It is capable of back-calculating properties of road layers using results of the FWD test. Saltan and Terzi (2008) use a cross-validation technique to improve ANN performance in modeling the deflection bowl in the presence of insufficient training data.

There are other unique applications of ANNs in estimating road layer properties such as automation of data reduction from seismic testing (Mallick, Bradley and Nazarian, 2006). The seismic testing method is developed to determine the moduli of reclaimed layers in thin surface hot-mix asphalt (HMA) roads. Sixteen Maine Department of Transportation projects with foamed asphalt and plant-mixed recycled asphalt road (RAP) reclaimed layers are tested according to this method using a portable seismic property analyzer.

Ceylan, Gopalakrishnan, Bayrak and Guclu (2013) apply ANNs to develop robust road layer moduli back-calculation algorithms that can tolerate noise or inaccuracies in the FWD deflection data



collected in the field. Though, the methods have largely been demonstrated on flexible asphalt, they are equally applicable on relatively rigid Portland Cement Concrete (PCC) roads. Ceylan, Gopalakrishnan and Bayrak (2008) apply the technique to rigid airfield roads subjected to full-scale dynamic traffic testing using simulated new generation aircraft gears. HWD tests are routinely conducted on three PCC test items at the Federal Aviation Administration's (FAA) National Airport Road Test Facility (NAPTF) to verify the uniformity of the test road structures and to measure road responses during full-scale traffic testing. Kim, Gopalakrishnan and Ceylan (2012) use ANNs to estimate rubblized PCC layer coefficient and modulus.

These systems have the advantages that they can identify longitudinal as well transverse pavement distresses. The detection can also take place in the initial phase of defect growth. The disadvantage is that the processes of acquiring the road properties are either through measurements or numerical simulations, both of which are time consuming and often requiring very specialized skills.

2.3.1.4 Others

Kongrattanaprasert, Nomura, Kamakura and Ueda (2009) propose a new method for automatically detecting the states of the road surface from tyre noises of vehicles. The methods are based on an FFT analysis, an ANN, and the mathematical theory of evidence. The proposed classification is carried out in sets of multiple neural networks using the learning vector quantization networks. The outcomes of the networks are then integrated using the voting decision making scheme. The classification results show very high levels of accuracy. Also, Kamimura (2002) applies a new information theoretic competitive learning model for actual road classification problem. The experimental results confirm that the new method can improve classification performance quite significantly.

2.3.2 ANNs as prognostic tools

In this Section, methods that use ANNs for predicting future road deterioration are presented. The available literature shows that these methods fall into two groups: those that rely on prediction of road distress indicators and those that rely on prediction of road layer properties. No studies were found at the compilation of this thesis that have used ANN-generated road profiles in predicting future road deterioration, despite the observation that the word 'prediction' in some studies has been mistakenly used to imply 'estimation' or 'approximation' without the necessary connotation of a future event.



2.3.2.1 Prediction of road distress indicators/statistics

Most of the early studies show that ANNs were mainly used for prediction of roughness-related distresses. La Torre, Domenichini and Darter (1998) employ an MLP to predict the IRI of a flexible road Section for 4 years from year of analysis. Overall short- and long-term performances on the training set look promising though the validation set needs further investigation. Huang and Moore (1997) use multiple linear regression and two ANN models to predict roughness distress level probability for bituminous roads. The models are based on historical road condition data and specific project-level data concerning road structural characteristics, traffic, and climatic conditions. Li, Wei and Kurt (1997) use a multi-back-propagation neural network model to predict roughness distress level probability for bituminous road as defined by the KDOT project optimization system (KDOT-POS).

Later, Roberts and Attoh-Okine (1998) apply a self-organizing quadratic function ANNs to the same KDOT's PMS database. The results show a significant improvement in the predictive performance of the new neural network. Then Attoh-Okine (1999) investigates the effects of learning rate and momentum term on back-propagation trained neural network to the prediction of flexible road performance. It is concluded that, for the specific data set and ANN architecture used, extremely low learning rates and momentum terms do not give satisfactory results. It is also established that the learning rate and momentum term, and validation data can be used to identify when over-learning is taking place in a training set.

Najjar and Felker (2003) use dynamic ANN approach to develop a reliable and accurate time-dependent roughness prediction model for newly constructed Kansas Jointed Plain Concrete Roads (JPCP). In order to achieve that objective, relevant data are obtained from the historical Kansas road condition database. The database includes construction and materials data as well as other inventory items such as traffic and climatic related data. Utilizing a two-stage training approach and through numerous training, testing, and validation processes, a three-layer (19-10-1) time-dependent ANN based roughness prediction model is developed. The developed model produces output values that are very close to the measured IRI values. Accordingly, it is able to project the time-dependent roughness behavior with reasonably high accuracy.

There are also isolated ANN applications for prediction of other most common road distresses such as joint faulting in jointed concrete roads (Saghafi, Hassani, Noori and Bustos, 2009), where ANNs are demonstrated to yield more accurate predictions than MLR. Zeghal (2008) presents ANN-based approach to predict thermal cracking of asphalt concrete in Canadian climatic conditions. Several ANN models are trained and tested using simple parameters such as road structure, material properties and environmental conditions as neural network inputs. Results show that ANNs effectively predict low severity thermal cracking of LTPP sites.



ANNs have also been successfully applied for prediction of rut depth. Braban-Ledoux and Sundin (2000) investigate the use of ANNs for predicting rut depth on a short term basis. Abiola, Owolabi, Sadiq and Aiyedun (2012) also develop reliable and time-dependent ANN based rut depth prediction model for Lagos-Ibadan Expressway. The model incorporates relevant variables such as road distresses, road layer thickness, road roughness, cumulative equivalent single axle load, sub grade California Bearing Ratio (CBR) and overlay asphalt concrete characteristics. The results show that the forecasting accuracy of the 11-24-1 network architecture is higher than other tested architecture in terms of both average absolute error (AAE) and RMSE.

Shalaby and Reggin (2007) predict visual survey scores using a resilient back-propagation algorithm from laser-measurements of rut depth and IRI and they find that only moderate accuracy is achieved and that additional data elements are required to improve the predictive ability of the model. Abd El-Hakim and El-Badawy (2013) develop an ANN model to predict the IRI for JPCP Sections. The inputs for this model are obtained from the LTPP database and include: initial IRI value, road age, transverse cracking, percent joints spalled, flexible and rigid patching areas, total joint faulting, freezing index, and percent subgrade passing No. 200 U.S. sieve. It is the same data and inputs that are used for the development of the Mechanistic-Empirical road Design Guide (MEPDG) IRI model for JPCP. The data include a total of 184 IRI measurements. The results of this study show that using the same input variables, the ANN model yields higher prediction accuracy than the MEPDG model. In addition, the bias in the predicted IRI values using the ANN model is significantly lower compared to the MEPDG regression model.

Since the early 2000s, more and more studies show a growing interest towards the development of unified road distress prediction techniques where ANNs are used for prediction of combinations of different distress types rather than just one type of distress. Yang, Lu, Gunaratne and Xiang (2003) report the implementation of an overall ANN-based methodology for road condition prediction. They develop three ANN models to predict the three key indices: crack rating, ride rating, and rut rating, used by the FDOT for road evaluation. The ANN models for each index are trained and tested by using the FDOT road condition database. In addition to the three key indices, FDOT uses PCR, which is the minimum of the three key indices. PCR is forecast with a combination of the three ANN models. Results of the research suggest that the ANN models are more accurate than the traditional regression models.

Sirvio and Hollmén (2008) introduce two hybrid methods for forecasting road roughness and rutting. It is found that Markovian models outperform ANN models and that roughness can be forecast more accurately than rutting. These results are consistent with the findings of Yang, Lu, Gunaratne and Dietrich (2006) who upon comparing the prediction capabilities of the same models on the Florida Department of Transportation's (FDOT) road condition data, they show that ANNs have a tendency toward over-predicting crack deterioration.

Thube (2012) summarizes the implementation of a road condition prediction methodology using the ANN to forecast cracking, raveling, rutting and roughness for Low Volume Roads (LVR) in India.



Road inventory data, as well as six cycles of road performance data that include distresses, subgrade characterisation and traffic data, are collected from 61 in-service LVR road Sections over a 3 year period. ANN models with different architectures are trained and tested to select the optimum ANN model. The results suggest that ANN models satisfactorily forecast future individual distresses. The performance of the selected ANN models is also found comparable to the calibrated HDM-4 models.

2.3.2.2 Prediction of road layer properties

In prognosis the emphasis is on prediction of future road layer moduli rather than just estimating their present values. Some of the studies have extended their predictive capacity to include determination of remaining service life of the road. However most of the earliest works focus on investigation of ANNs for predicting road layer modulus and the feasibility of using them as replacements for conventional methods.

Meier and Rix (1994) use two back-propagation neural networks to back-calculate road moduli for three-layer flexible road profiles. One network is trained using ideal deflection basins while the other is trained using deflection basins to which random noise is added to simulate measurement errors. It is observed that training with noise-contaminated data improves the predictive capacity of the ANN model. The network performance is demonstrated on experimental data measured from two road test Sections of Strategic Highway Research Program (SHRP).

Tutumluer and Seyhan (1998) use a feed-forward back-propagation neural network for layer moduli prediction when trained with two tri-axial stresses, measured vertical deformation, and two aggregate properties as input variables. The output variables are the horizontal and shear moduli for which the target values are derived and computed from test results. The ANN models predict the two moduli, with mean errors of less than 3% compared with those computed by using experimental stresses and strains.

Kim, Xu and Kim (1999) present a layer moduli back-calculation procedure using transient measurements from both FWD test and surface wave test. This algorithm employs numerical solutions of a multi-layered half-space based on Hankel transforms as a forward model and ANNs for the inversion process. Two ANNs are used in series; a depth to a stiff layer is predicted first and used as input for the second ANN to be used for the prediction of layer moduli. Dispersion analysis is performed on FWD transient deflections and surface wave test measurements using the Short Kernel Method. Conventional back-calculation is also performed on FWD deflections using the MODULUS 5.0 program. It is concluded that the dispersion-based back-calculation method is sensitive to changes in upper layer condition and results in less variable sub-surface layer moduli and more accurate predictions of depth than the conventional back-calculation. Ceylan, Tutumluier and Barenberg (1999) use trained similar technique to predict stresses and deflections in jointed concrete airfield roads serving the Boeing B-777 aircraft. It is observed that the ANN model is able to approximate FEA results much faster (in less than 1s) and with accuracies within 0.5 percent error from the FEA results.



Thus it is concluded that ANNs can be used to replace more demanding mechanistic models such as the elastic layered programs (ELPs), which are used in road design procedures for the analysis of jointed concrete roads.

More recently, Ceylan, Guclu, Tutumluer and Thompson (2005) and Bayrak and Ceylan (2008) also demonstrate the better speed and accuracy of ANN models. Abdallah, Yuan and Nazarian (2001) show that ANNs can also produce more robust and consistent predictions than conventional methods.

As noted in Section 2.3.1.3, ANNs can be used in conjunction with mechanistic models to provide some physical insight of the deterioration of road layer moduli. Ceylan and Gopalakrishnan (2007) demonstrate this further by using ANN-based structural models for prediction of critical strains based on FWD deflection data. Later, Ceylan, Gopalakrishnan, Kim, Schwartz and Li (2013) develop two response surface modeling approaches: MLR and ANN for evaluation of AASHTO MEPDG input sensitivities across the problem domain. The AASHTO MEPDG provides road analysis and performance predictions for various what-if scenarios. The response surface modeling approaches based on ANNs not only provide robust and accurate representations of the complex relationships between MEPDG inputs and distress outputs but also captured the variation in sensitivity across the problem domain. The normalized sensitivity index for the design limit proposed in the study provides practical interpretation of sensitivity by relating a given percentage change in an MEPDG design input to the corresponding percentage change in predicted distress relative to its design limit value.

Other studies have used ANNs for prediction of remaining service life through evaluation of crack propagation in roads. Examples of such workers include, Abdallah, Ferregut, Nazarian and Lucero (1999) who develop an ANN-based software program to predict the remaining life of flexible roads due to either rutting or fatigue cracking. The inputs to the software are the best estimate of the thickness of the layers, the deflection basin measured with FWD, and optionally, the extent of damage at the time of the FWD test. Miradi and Molenaar (2006) also develop ANN models for porous asphalt (PA) lifespan prediction. The study uses 102 PA road Sections obtained from a database in which data are collected over a total period of ten years from 1991 to 2000. Sharma and Das (2008) use ANN in predicting both the elastic moduli of asphalt road and remaining life without back-calculating the layer moduli.

Xiao, Amirkhanian and Juang (2009) use an ANN in predicting the fatigue life of rubblized asphalt concrete mixtures containing reclaimed asphalt road (RAP). It is concluded that ANN-based techniques are more effective in predicting fatigue life of the modified mixtures than traditional statistical-based prediction models. Tsai, Lytton and Lee (2010) calibrate mechanistic-based models to field data to produce a design process for predicting reflection cracks. In the mechanistic computations, material properties and fracture-related stress intensity factors (SIF) are generated using efficient ANN algorithms. Ceylan, Gopalakrishnan and Lytton (2011) use ANN to model the stress intensity factor (SIF) as cracks grow upward through a HMA overlay as a result of both load and thermal effects with and without reinforcing interlayers. It is observed that the use of ANN significantly reduces the computational overload.



Literature also reports applications of ANN for prognosis using other road properties and in the presence of material nonlinearities. Gopalakrishnan and Thompson (2004) develop a tool for back-calculating non-linear road layer moduli from HWD data using ANN. Zhang, Li and Zhang (2007) apply ANN in predicting unconfined compressive strength of field test data. They report very high prediction accuracies. Pekcan, Tutumluer and Thompson (2008) are able to predict the road layer moduli and critical road responses using ANNs in the presence of nonlinear layer modulus behavior such as stress hardening for granular materials, stress softening for fine grained soils, and lime stabilization on road responses.

Gopalakrishnan (2008) develop an ANN-based approach for back-calculation of road moduli based on HWD test data in the analysis of airport flexible roads subjected to new generation aircraft (NGA). Two medium-strength subgrade flexible test Sections, at the National Airport Road Test Facility (NAPTF), are modeled using FEA program, which can take into account the non-linear stress-dependent behavior of road geomaterials. Gandhi, Xiao and Amirkhanian (2009) develop a series of ANN models to predict the indirect tensile strength (ITS) and tensile strength ratio (TSR) of various mixtures considering five input variables such as asphalt binder source, aggregate source, anti-stripping agents (ASA), conditioning duration, and asphalt binder content. The results indicate that ANN-based models are effective in predicting the ITS and TSR values of mixtures regardless of the test conditions and can easily be implemented in a spreadsheet, thus making it easy to apply.

Far, Underwood, Ranjithan, Kim and Jackson (2009) develop ANN-based models for predicting the dynamic modulus of HMA mixtures on LTPP test Sections. Park, Kweon and Lee (2009) use ANN models to predict the resilient modulus of granular subgrade soils and sub-base materials from basic material properties and in-situ conditions related to stresses. The trained ANN models are successfully applied to the actual HWD test data acquired at the NAPTF to predict the asphalt concrete moduli and non-linear subgrade moduli of the medium-strength subgrade flexible test Sections. Gopalakrishnan and Papadopoulos (2011) use ANN model to derive a decision rule for the inverse prediction of non-linear road layer moduli from Non-Destructive Test (NDT) deflection data.

The growing number of AI tools has also necessitated investigations into comparative performances among themselves and/or benchmarking them against well-established regression models. Ceylan, Schwartz, Kim and Gopalakrishnan (2009) discuss the accuracy and robustness of the various ANN-based predictive models for estimating HMA inputs for mechanistic-empirical road design guide. They show that ANN-based HMA models, having similar input variables, exhibit significantly better overall prediction accuracy, better local accuracy at high and low temperature extremes, less prediction bias, and better balance between temperature and mixture influences than their regression-based counterparts. The ANN models as a group also produce the best agreement between predicted rutting and alligator cracking computed using predicted versus measured HMA values for a typical road scenario.



Taddesse and Mork (2009) present the development of flexible road roughness progression models, expressed in terms of IRI, from PMS data of the Ethiopian road network, using MLR and ANN techniques. A comparative study also demonstrates that the ANN models outperform the MLR models. Miradi, Molenaar, Van De Ven and Molenaar (2009) compare the predictive performances of support vector machines (SVM) and ANN on stiffness of CTBs from deflection measurements. The structures consist of an asphalt top layer, a cement treated base and subgrade. The results show that SVMs produce better results than ANNs. However, an extra validation using 100 new data points shows better quality for ANN than SVMs. It is concluded that both ANN and SVM are powerful tools for accurate prediction of stiffness of CTB.

Bianchini and Bandini (2010) show that neuro-fuzzy models are superior to linear regression models in terms of accuracy in prediction of performance of flexible roads using the parameters routinely collected by agencies to characterize the condition of an existing road. It is further shown that the neuro-fuzzy models possess better generalization capability. Oh and Barham (2011) compare ANN-based methodology and a conventional regression model to predict the responses of HMA. In order to ensure the generality of the final neural network, the input/output training patterns are taken from different conditions. Chik and Aljanabi (2013) show that a ten-fold cross-validation ANN model gives higher efficiency during both training and testing than a non-cross-validation ANN model in prediction of settlement under embankment load. They use actual angle of internal friction, spacing ratio, cylindrical ratio and height of the embankment as the input parameters, while the settlement ratio is the main output.

2.3.2.3 Others

Yildirim and Uzmay (2001) investigate the variation of vertical vibrations of vehicles using a Radial Basis Function Neural Network (RBFNN). Venayagamoorthy and Allopi (2007) apply ANN-based technique for the prediction of bearing capacity of road structures under different loading conditions. Heimann, Bouzid and Trabelsi (2006) present a technique where road surface condition is detected based on multi-sensor data fusion for an improved preconditioning of automotive control systems. It involves the use of different measuring systems in three levels: the environment description, the slip based statistical slippery recognition and the reactive friction detection and adaptation. The optical and acoustic sensors' measurements form the inputs of a pre-processing block, where a specific frequency and statistical analysis is implemented. The road profile is simultaneously captured with a laser device and attendant roughness descriptors are computed. The regression between descriptors and grip is obtained by an ANN, which is used for prognostic purposes after learning.

Beguin, Bastin and Wertz (2005) present a method for predicting the presence or absence of ice on the road. The method is based on a Least Squares Support Vector Machine (LSSVM) applied to data from the road in Wallonia (Belgium). It is shown that including a prediction of the air temperature given by a meteorological center in the model helps having better accuracy in this application, 95% accuracy is achieved for a 3 hours prediction horizon, and 92% for 6, 12 and 24 hours horizon.



2.3.3 ANNs as aids for road maintenance recommendations

Artificial neural networks have also been reported as aids to decision making in road maintenance management. Literature reveals that the neural networks have been applied both at project and network-level for selection of appropriate maintenance strategies, optimization of maintenance costs and in prioritization of routes for M&R action.

2.3.3.1 *Selection of appropriate maintenance strategy*

Traditionally engineers select maintenance strategies based on their experience and past maintenance data. This Section presents methods in which ANNs are used for selection of appropriate strategy. There is no single indicator or property that is used to perform this function. The networks can be trained with subjective visual inspection data or model generated data or even experimental data by using elaborate measuring equipment.

Taha and Hanna (1995) utilize a hybrid evolutionary-learning system using gradient descent learning as well as genetic algorithms (GAs) to determine network weights for selecting the optimal maintenance strategy for flexible roads. This ANN misclassifies only 6 out of 100 unseen cases with an error rate of 0.024. They conclude by stating that the critical issue to the effectiveness of ANNs is the optimisation of the global error function and also that ANN performances can be improved by applying GAs. Goh (1997) presents an integrated neural network and expert system that facilitates decision-making for asphalt road maintenance. The software consists of a rule-based, knowledge-based system and user interface, a database system and a multilayer neural network. The integrated system enables massive amounts of data to be handled systematically and assists the researchers to make rational decisions regarding road rehabilitation.

Alsugair and Al-Qudrah (1998) apply ANN for recommending appropriate M&R actions. Each training set consists of a road condition represented by deduct values for each distress present in the road and the corresponding recommended M&R action. Road condition data used in this study are obtained from comprehensive visual inspection data conducted on the Riyadh road network in Saudi Arabia. The associated M&R actions are obtained based on consulting human expertise and M&R actions recommended by PMDSS software. They show that ANN is appropriate for implementation in identifying appropriate M&R actions. Abdelrahim and George (2000) develop a genetic adaptive neural network training algorithm with a single hidden layer and sigmoid squashing function to choose an appropriate road maintenance strategy. A set of examples is derived from experts' judgments with a total of 144 cases randomly divided into "in-sample" and "out-of-sample" data for training and testing purposes, respectively. The trained network successfully predicts 83% of the test cases. The remaining 17% of cases are one or two levels away from the expert judgments.

There are more recent works such as Bosurgi and Trifirò (2005) who propose a hybrid approach for optimal management of road road maintenance for a pluri-annual planning period. The proposed



procedure is used for optimal management of rehabilitation interventions in road friction. In particular, a sideways force coefficient prediction model is created through an appropriate neural network to characterize the road conditions. A genetic algorithm is used to tackle the combinatorial nature of the network-level maintenance programming. The procedure is successfully applied to a motorway network of Eastern Sicily in Italy. Abo-Hashema (2009) applies ANN to recommend road overlay thickness by learning from mechanistic-empirical overlay design cases. The study involves using a neural network with expert knowledge data collected through a questionnaire. This is aimed at enhancing the capability of selecting more appropriate strategies for repair of road distress, thus to ensure optimal road service performance. Gupta, Kumar and Rastogi (2011) use statistical analysis tools and ANN to develop models for analysis of road condition and development of Maintenance Priority Index (MPI) from investigations carried out on 18 Sections of low volume roads in Uttarakhand and Uttar Pradesh states of India. Statistical performance indicators and logical relationships between input and output parameters are used to select the best fit model.

Shaaban and Nabwey (2012) present a new approach to the rough-set theory in a PMS database that enables road engineers to discover the shortest subsets of condition attributes having quality equal to the general quality of defined characteristics in the information system, to assess and describe road conditions, and to derive decision rules for rehabilitation and reconstruction of the roads. The best algorithm of defined attributes in the information system is determined by invoking an ANN and the results compared with rough-set performance. The findings are that rough-set theory performs better and has much stronger operational capability than the ANN both in identifying the effective parameters for the severity evaluation of typical distresses in roads and in decision-making for selecting the type of repair.

2.3.3.2 Routes selection

Selection of road Sections or networks that are suitable for maintenance can be a very tough task involving complex route/network optimization procedures. ANNs can be used to perform such tasks automatically without requiring as much effort. Hajek, Haas, Chong and Phang (1987) compare the rule-based system and ANN in selecting and recommending routing and sealing treatments of roads. It is shown that the ANN technique is faster and easier to implement. However after analysing in-depth the pros and cons of each technique the authors conclude by recommending combined use of both techniques. Flintsch, Zaniewski, and Delton (1996) use a neural network as part of an automatic procedure to screen and recommend roadway Sections for road preservation at the Arizona Department of Transportation (ADOT). The results show that the ANN learns 100 % of the training patterns and 76 % of the test patterns and this is considered acceptable performance taking into account that not all factors influencing the selection of the preservation projects can be traced back to the PMS database.

Flintsch and Zaniewski (1997) develop an automatic project recommendation procedure and implement it in a user-friendly, modular computer program. This automatic system is expected to reduce considerably the effort required to develop the preservation programs. The trained ANN



prepares a list of candidate Sections, using the criteria learned from past selection and the current condition of all road Section. This preliminary list of candidate Section is further analyzed by a project recommendation procedure. This procedure recommends a preservation treatment, assigns a priority rating to each Section in the list, and sorts the projects by priority. Funding is assigned to the highest-priority Sections within each roadway group until the budget recommendation provided by the network optimization process is reached.

2.3.3.3 Optimization of road maintenance costs

Optimisation of road maintenance costs is an important aspect of road maintenance management to which ANNs have also been successfully applied. Only two examples of recent studies have been reported in this Section, but it should not be considered exhaustive. Kargah-Ostadi, Stoffels and Tabatabaee (2010) develop a method for assessing changes in the IRI over time through ANNs pattern recognition, using information from the Specific Road Study (SPS)-5 asphalt concrete rehabilitation experiment extracted from FHWA's Long-Term Road Performance database. The developed model can be used to predict and compare road roughness variation trends after various rehabilitation alternatives. This case study illustrates the implementation of the roughness model along with life-cycle cost analysis in making future road rehabilitation recommendations.

Shehab and Meisami-Fard (2013) present a cost-estimating system for rubblized asphalt road rehabilitation projects which utilize information collected from 44 projects and apply ANN to perform its task. This tool is believed to be helpful in many road road applications, such as preparing accurate budget estimates and life-cost analyses, as well as managing financial resources in limited budget environments.

2.4 Concluding remarks on ANN-based methodologies

In this Chapter, ANN-based methodologies applied to RTRMS have been presented. The need for development of these methodologies has been spurred by increased costs of maintenance incurred with current road maintenance systems (Thompson and Visser, 2003).

The reviewed literature has been classified into three main categories of ANN applications to RTRMS, namely: ANNs as tools for fault diagnosis, ANNs as tools for fault prognosis and ANNs as aids to road maintenance recommendations. The first two applications provide a good basis for the development of a real-time road condition monitoring system where different road distresses are inventorised in an online database easily accessible by the maintenance managers for application to maintenance systems. The third application is also divided into three parts: selection of road strategy, selection of routes and optimisation of maintenance costs. It has been pointed out that once ANNs are created and trained, they can be simulated in a very short time often a few seconds, which makes them ideal for fast online processing. This has been confirmed by most of the reviewed studies where it has been reported that ANNs yield much faster results than the conventional mechanistic, FE models and



optimisation schemes. ANNs have also been noted to yield more accurate and consistent results than regression models. And they have further been used with road image data to automate the road condition monitoring process thereby replacing subjective expert panel ratings.

Studies that apply ANNs as road condition monitoring systems (diagnosis and prognosis) are further classified according to the types of quantities that are estimated by the neural network model. There are two common types of quantities that are used by both applications: road distress parameters and road layer properties. Studies using ANNs as road profiles estimators are observed to be scarce and only occur under the first area of application (diagnosis). The use of road distress indicators is noted to provide variety in terms of different types of distresses that can potentially be identified but with a limitation that it can only yield qualitative assessments. Quantitative assessments cannot be added to these systems without a corresponding computational penalty. The use of road layer properties is also noted to provide variety in terms of number of distress types with the added benefit that the identified defects can be interpreted in terms of the road properties. These systems have also been noted to provide more reliable determination of remaining service life of roads. Their main shortcoming is the complexity of the processes used to acquire data for training the neural network. The application of ANN-reconstructed road profiles is noted to provide accurate and reliable quantitative assessments of road roughness with a shortcoming that it is limited to roughness identification. So, why should there be any further development in this application?

It is true that a number of road condition parameters play a role in road maintenance management, namely roughness, rutting, texture, skid resistance, relationship between friction and texture, cracking and strength, however, road roughness emerges as a very crucial property of road condition in any economic evaluation of design and maintenance standards for roads, as well as in any functional evaluation of standards desired by road users (Paterson, 1987). Besides, various researchers (La Torre, Ballerini, and Di Volo, 2002; Hassan and McManus, 2001; Liu, Gazis and Kennedy, 1999; Potter, Hannay, Cairney and Markarov, 1992) have showed that the widely used panel ratings are highly correlated with road roughness. Paterson (1987) further reports that the cost of operating vehicles and transporting goods rises as road roughness increases, and that small improvements in roughness can yield high economic returns given that the total operating costs of all vehicles on a roadway outweigh the agency costs of maintaining the road by typically ten- to twenty-fold. Hence the need for development of methodologies that focus on the identification of roughness-related road defects.



Chapter 3

Road-vehicle interaction

This Chapter presents the relevant theory that governs the interaction of the vehicle and the road using quarter-car model with linear suspension characteristics (Sections 3.1 and 3.2). This provides a very good background towards understanding how road undulations are transmitted through the vehicle system. Then, in Section 3.3, an analytical procedure is formulated on how road undulations can be reconstructed from given vehicle vibrations (vehicle axle and body accelerations). However this procedure relies on the availability of system characteristics and can become computationally intractable if the suspension system has nonlinearities.

In order to afford the reader an understanding of road surface roughness classification, Section 3.4 presents two of the most popular classification methods i.e. the International Roughness Index (IRI) and the ISO Displacement Spectral Density (DSD) classification methods. The two methods are briefly discussed and qualitatively compared to clarify why the ISO DSD has been preferred in this work. Other important factors such as: road surface probability distribution, magnitudes of dynamic tyre forces, frequencies of vehicle ride vibrations, suspension nonlinearities, wheelbase filtering, cross-coupling between suspensions, payload distribution, structural vibration, roll motion influence, tyre non-uniformities, tyre envelopment and spatial repeatability are also summarised in this Section. The Chapter is concluded in Section 3.5.

3.1 Vehicle ride dynamics

A moving vehicle exerts forces on the road surface; and the road, in return, exerts back forces into the vehicle's tyre tread and, through its suspension into the vehicle body. Whereas these forces may be useful in that they provide traction, cornering and braking to the vehicle; they also impact both the vehicle and road structures negatively. The vehicle experiences wear and tear in the tyres, suspension struts and its body structure, while the road surface undergoes degradation which appears in the form of increased roughness, material loss from the surface and rutting, in the case of unpaved roads; and cracking, ravelling, potholing, degraded skid resistance, rutting and increased roughness for paved roads. The magnitudes of these forces depend on road surface roughness, and on vehicle's own geometry, inertial, tyre and suspension characteristics and its travelling speed.

Road surface condition is particularly viewed as a key factor to these forces (Paterson, 1987; DSTI, 1998). Thus the knowledge of road profiles can assist in determining the forces and even assist in optimising suspension design and determination of remaining fatigue life of the pavement.

In this study, vertical vehicle responses (particularly accelerations) are used to reconstruct road surface profiles. In this Chapter, this problem is treated as an inverse ride dynamics problem. Gillespie (1992) formulates the forward vehicle ride dynamics problem as

$$\mathbf{M}\ddot{\mathbf{z}} + \mathbf{C}\dot{\mathbf{z}} + \mathbf{K}\mathbf{z} = \mathbf{f} \quad (3.1)$$

where



\mathbf{M} is a matrix containing the system mass and inertial properties, \mathbf{K} is a system stiffness matrix, \mathbf{C} is a system damping matrix, \mathbf{f} is a vector of generalised forces and \mathbf{z} is a vector of generalised displacements.

Equation 3.1 is a second order differential equation in the generalised displacements \mathbf{z} which can be solved through any of the classical methods or by a numerical integration scheme. In this problem, the generalised forces \mathbf{f} are often functions of road surface profiles, \mathbf{z}_r . If the road profiles are assumed harmonic in nature, equation 3.1 may be expressed in terms of force and response magnitudes as (Rao, Jian, Mohamad, Gang, and Zuo, 2008):

$$\mathbf{Z}_r = \left(\frac{-\omega^2 \mathbf{M} + j\omega \mathbf{C} + \mathbf{K}}{j\omega \mathbf{C}_t + \mathbf{K}_t} \right) \mathbf{Z} \quad (3.2)$$

where

\mathbf{K}_t and \mathbf{C}_t are augmented tyre stiffness and damping matrices, respectively; \mathbf{Z}_r and \mathbf{Z} are the amplitudes of road profile displacement input and vehicle response, respectively; and ω is the angular frequency (rad/s) at which the profile is traversed by the vehicle.

Equation 3.2 is better presented in terms of the ratio of the amplitude quantities as (Rao, et al., 2008):

$$\frac{\mathbf{Z}}{\mathbf{Z}_r} = \sqrt{\frac{\mathbf{K}_t^2 + (\omega \mathbf{C}_t)^2}{(\mathbf{K} - \omega^2 \mathbf{M})^2 + (\omega \mathbf{C})^2}} \quad (3.3)$$

Equation 3.3 can be used to determine the ratio between the road displacement input and the vehicle response quantities. This ratio is referred to as the transmissibility ratio and it is very important in this study since it represents how the excitation input influences the vertical vehicle responses. A discussion of the transmissibility ratio is presented in Section 3.2 after which an analytical procedure for the determination of road profiles and dynamic tyre forces is presented in Section 3.3. Road surface profile characterisation and classification are presented in Section 3.4. Finally the Chapter is concluded in Section 3.5 with highlights of some of the important factors in road-vehicle interaction.

3.2 Transmissibility

Transmissibility is the non-dimensional ratio of response amplitude to excitation amplitude for a system in steady-state forced vibration (Gillespie, 1992). There are three main excitation inputs to a vehicle system that influence its ride dynamics (vertical vehicle dynamic behaviour), namely: road input, on-board forces (forces acting directly on the sprung mass such as the exhaust and driveline excitations), and the forces acting directly on the wheel axle such as those due to tyre non-uniformities. Transmissibility ratio is used in this study to investigate the influence of each of these forces on the acquired accelerations and to determine the relative applicability of sprung mass and unsprung mass accelerations for use in road profile reconstruction.

Transmissibility is primarily governed by tyre and the suspension characteristics. The tyre-suspension system provides some sort of a mechanical filter to these excitation inputs. Thus the excitations may

be amplified at certain frequency ranges while being attenuated at other frequency ranges. In this study, a quarter car model in Figure 3.1 is used to determine and understand the relationships that exist between the motions of the two lumped masses in response to the road displacement input z_r , the force on the sprung mass F_s , and the force on the unsprung mass F_u .

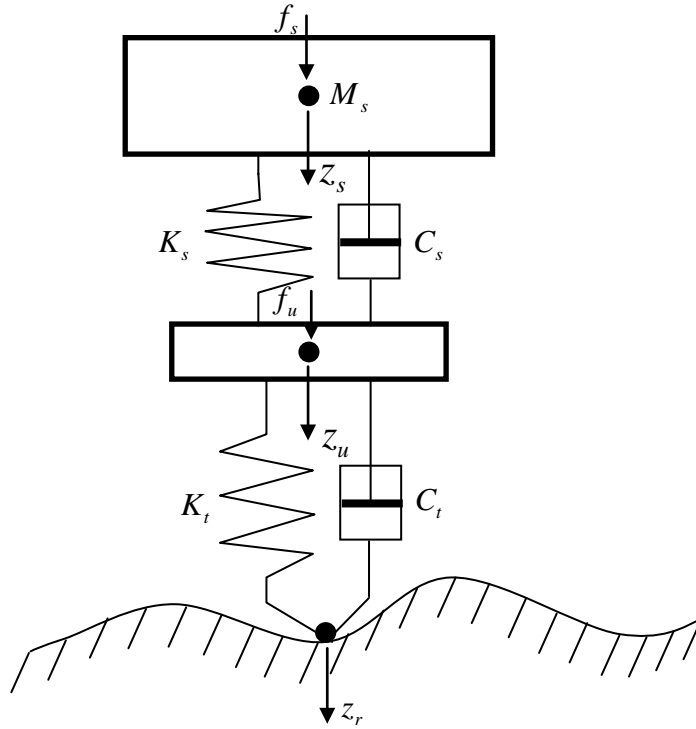


Figure 3.1. Quarter car model

Gillespie (1992) gives the relationships for the sprung mass only, tyre damping is neglected and the system is vibrating from an equilibrium position. That was sufficient because ride was the central theme in that study, but where the investigation is centred on inference of road surface profiles from given response data, it can be quite beneficial to consider the relationships for the unsprung mass as well. This is not to imply that sprung mass responses should not be used at all. There are several practical situations where measuring directly on the unsprung mass system is infeasible. In this study, these relationships are examined comparing the transmissibility ratios for the sprung and unsprung masses' accelerations (i.e. \ddot{z}_s and \ddot{z}_u , respectively) in response to the road displacement input z_r . The transmissibility ratios between the sprung and unsprung mass displacements and the road displacement input can be derived from equations of motion for each of the masses in Figure 3.1. Equations (3.4) and (3.5) represent the equations of motion for the sprung and unsprung masses respectively.

$$M_s \ddot{z}_s + C_s \dot{z}_s + K_s z_s = C_s \dot{z}_u + K_s z_u + f_s \quad (3.4)$$

$$M_u \ddot{z}_u + (C_s + C_t) \dot{z}_u + (K_s + K_t) z_u = C_s \dot{z}_s + K_s z_s + C_t \dot{z}_r + K_t z_r + f_u \quad (3.5)$$

If the road input and forces are assumed to be harmonic, then equations (3.4) and (3.5) can be expressed in their complex notations as equations (3.6) and (3.7), respectively. Note that the vehicle



displacements, road input displacements and forces are not expressed in terms of their magnitude quantities only as denoted by their capital letters.

$$(K_s - \omega^2 M_s + j\omega C_s)Z_s = (K_s + j\omega C_s)Z_u + F_s \quad (3.6)$$

$$(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))Z_u = (K_s + j\omega C_s)Z_s + (K_t + j\omega C_t)Z_r + F_u \quad (3.7)$$

When deriving the transmissibility of a particular response and excitation, it is assumed that all the other excitations are zero. Thus in deriving the transmissibility of sprung mass response z_s in response to road excitation input z_r , the external forces applied directly on the sprung F_s such as excitations due to driveline, wind loads, or forces arising from loose or moving payloads. Unsprung F_u masses are not considered and the resulting transmissibility is given by

$$Z_u = \frac{(K_s - \omega^2 M_s + j\omega C_s)}{K_s + j\omega C_s} Z_s$$

$$\left(\left((K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t)) \cdot \frac{(K_s - \omega^2 M_s + j\omega C_s)}{K_s + j\omega C_s} \right) - (K_s + j\omega C_s) \right) Z_s = (K_t + j\omega C_t) Z_r$$

$$\frac{Z_s}{Z_r} = \frac{(K_t + j\omega C_t)(K_s + j\omega C_s)}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.8)$$

A similar equation as in Rao et al. (2008) may result if tyre damping is neglected in equation (3.8). The expressions for the other transmissibility ratios are calculated as follows:

$$\frac{Z_u}{Z_r} = \frac{(K_t + j\omega C_t)(K_s - \omega^2 M_s + j\omega C_s)}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.9)$$

$$\frac{\omega^2 Z_s}{F_s / M_s} = \frac{\omega^2 M_s (K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.10)$$

$$\frac{\omega^2 Z_u}{F_s / M_s} = \frac{\omega^2 M_s (K_s + j\omega C_s)}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.11)$$

$$\frac{\omega^2 Z_s}{F_u / M_u} = \frac{\omega^2 M_u (K_s + j\omega C_s)}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.12)$$

$$\frac{\omega^2 Z_u}{F_u / M_u} = \frac{\omega^2 M_u (K_s - \omega^2 M_s + j\omega C_s)}{(K_s + K_t - \omega^2 M_u + j\omega(C_s + C_t))(K_s - \omega^2 M_s + j\omega C_s) - (K_s + j\omega C_s)^2} \quad (3.13)$$

Notice that in order to determine transmissibility ratios for the force inputs, the resulting displacement/force relationship, which does not represent a non-dimensional ratio but system flexibility/compliance expressed in metres/newton (m/N), is accordingly transformed into non-dimensional ratios between acceleration magnitudes by multiplying the displacement by ω^2 and



dividing the force by the respective mass as shown in equations 3.10 – 3.13. However, space limitations in the plot area resulted in the legends being conveniently shortened.

Since the basic dynamic system is the same, the system dynamics as represented by the denominator expression is the same throughout; hence any differences in the transmissibility ratios are due to the numerator expressions. Figure 3.2 plots the transmissibility ratios over temporal frequencies for soft suspension spring (typical for ride comfort setting) and stiff suspension springs (typical for handling mode setting). Figure 3.3 shows how the ratios vary over different spring stiffness. The tyre stiffness is the same for both suspension settings and the suspension system is assumed to behave linearly.

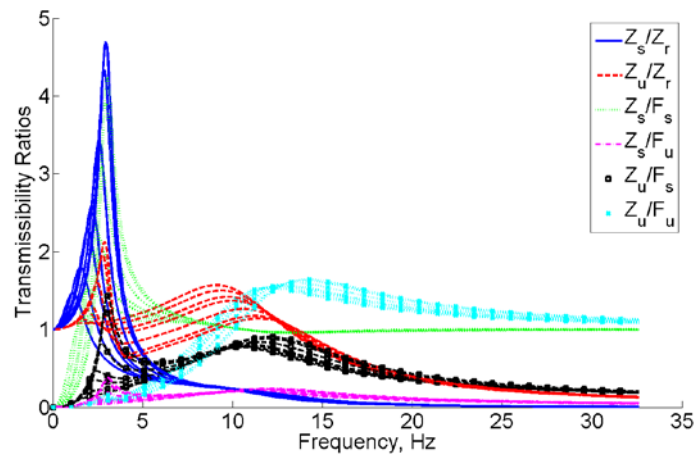
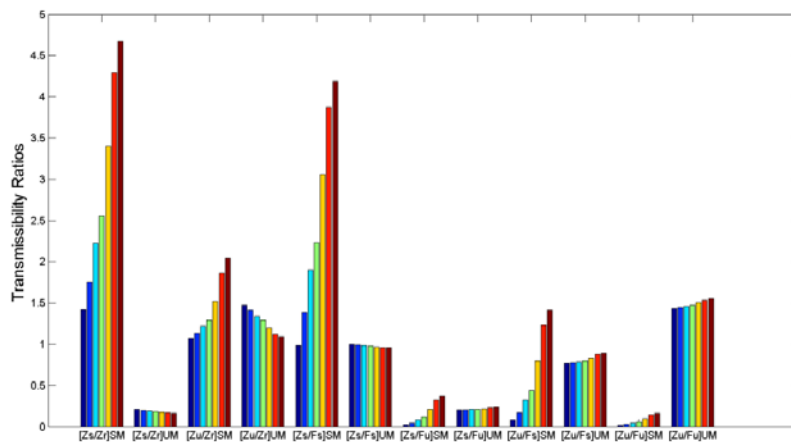


Figure 3.2. Transmissibility ratios for a range of suspension spring rates from 28 – 220 kN/m. The legends have been conveniently shortened but here are their actual labels:

$$Z_s/F_s \Rightarrow \frac{\omega^2 Z_s}{(F_s/M_s)}; Z_s/F_u \Rightarrow \frac{\omega^2 Z_s}{(F_u/M_u)}; Z_u/F_s \Rightarrow \frac{\omega^2 Z_u}{(F_s/M_s)} \text{ and } Z_u/F_u \Rightarrow \frac{\omega^2 Z_u}{(F_u/M_u)}.$$



1

Figure 3.3. Summary of ratios calculated at the sprung mass and unsprung mass resonances for each transmissibility. Labels on x-axis corresponding to forces have similar annotations to Figure 3.2.



Table 3.1. Transmissibility ratios for the road displacement inputs and forces applied directly on the sprung and unsprung masses

Transmissibility ratio	Sprung Mass Resonance	Unsprung Mass Resonance	Remark
$\frac{\ddot{Z}_s}{\ddot{Z}_r}$	2.9	0.2	Figure 3.2 shows that road inputs are amplified at sprung mass bounce resonance but heavily attenuated from the unsprung mass resonance onwards. This implies that estimation of higher frequency content in the reconstructed profiles when sprung mass accelerations are used will be poor. Figure 3.3 shows that the amplification of the road profiles at the sprung mass resonance increases with suspension stiffness but there is no significant variation in the response around wheel hop due to changes in spring stiffness.
$\frac{\ddot{Z}_u}{\ddot{Z}_r}$	1.4	1.2	Figure 3.2 shows that road inputs are amplified at both sprung mass and unsprung mass bounce natural frequencies. Thus estimation of road profiles using unsprung mass accelerations can be relatively more accurate than for sprung mass accelerations especially for higher frequency content in the reconstructed profiles given that road profiles are the sole inputs. Figure 3.5 shows that there are minor increases in transmissibility around the sprung mass resonance with increased suspension stiffness and minor decrease around the wheel hop with the increase.
$\frac{\ddot{Z}_s}{F_s / M_s}$	2.5	1.0	The estimation of body forces F_s using accelerations from the sprung mass can be very accurate. Figure 3.3 shows an appreciable amount of amplification of the forces at lower frequencies especially with stiffer suspension springs. At higher frequencies, the forces are transmitted without loss and the transmissibility does not vary with changes in suspension stiffness.
$\frac{\ddot{Z}_s}{F_u / M_u}$	0.2	0.2	Reconstruction of forces applied directly to the unsprung mass can be carried out with a lot of difficulty using sprung mass accelerations only. Though the transmissibility ratio improves with increased suspension stiffness around sprung mass



			resonance, it is not high enough to guarantee satisfactory results.
$\frac{\ddot{Z}_u}{F_s / M_s}$	0.6	0.8	On-board forces can be reconstructed from given unsprung mass accelerations but with some difficulties as compared to using sprung mass accelerations. There are minor penalties in the road profiles at both frequencies. Figure 3.5 shows that it is extremely hard to estimate low frequency content when the suspension stiffness is too soft.
$\frac{\ddot{Z}_u}{F_u / M_u}$	0.1	1.5	Estimation of forces acting on the unsprung masses can only be accurate for higher frequency content beyond wheel hop. The lower frequency content is severely penalised.

3.3 Determination of road profiles and dynamic tyre forces

The key purpose of this study is to determine road profiles from given vehicle responses. But it is also realised that dynamic tyre forces are an integral part of the problem since they provide a transition into the study of vehicle structure fatigue and suspension integrity which are the main concerns of mechanical engineers. Thus Section 3.3.1 presents the formulation of an analytical road profile reconstruction procedure applied to a quarter car model and used to demonstrate some of the main ideas presented in Section 3.2. Then Section 3.3.2 presents the calculation dynamic tyre forces.

3.3.1 Road profile calculation

As mentioned in Section 3.1, the problem of determining road profiles is treated in this Chapter as an inverse problem to the normal ride dynamics problem expressed by equation 3.1. The solution presented in equation 3.2 is applicable to inputs of harmonic nature. In this Section, a technique for computing the road profiles given a vehicle model and the accelerations acquired from either the sprung or unsprung mass or even from both.

The technique is summarised as follows:

- **STEP 1:** Prepare and process the given accelerations
- **STEP 2:** Compute velocities and displacements from the available accelerations.
- If only unsprung mass accelerations are available, proceed as follows:
 - **STEP 3:** Apply them to the suspension stiffness and damping to calculate the absolute reaction of the suspension onto the unsprung mass thus,



$$\mathbf{f}_{(s \rightarrow u)abs} = K_s(\mathbf{z}_u) + C_s(\dot{\mathbf{z}}_u)$$

where

\mathbf{z}_u denote generalised displacement vector of the unsprung mass, K_s and C_s are functions representing suspension stiffness and damping, and $\mathbf{f}_{(s \rightarrow u)abs}$ denote generalised vector of reaction forces from the unsprung mass onto the system comprising the sprung mass only.

- **STEP 4:** Write and solve the equation for the sprung mass only; thereby solving for the generalised displacements \mathbf{z}_s of the sprung mass only, thus

$$M_s(\ddot{\mathbf{z}}_s) + C_s(\dot{\mathbf{z}}_s) + K_s(\mathbf{z}_s) = \mathbf{f}_{(s \rightarrow u)abs}$$

where

M_s denotes the sprung mass inertial function

- **STEP 5:** Apply \mathbf{z}_s to the road-unsprung mass equation to solve for the road input vertical displacement/profile \mathbf{z}_r by using the generalised vector of forces acting on the road-unsprung mass system as $\mathbf{f}_{(u)abs} = M_u(\ddot{\mathbf{z}}_u) + K_s(\mathbf{z}_s, \mathbf{z}_u) + C_s(\dot{\mathbf{z}}_s, \dot{\mathbf{z}}_u) + K_t(\mathbf{z}_u) + C_s(\dot{\mathbf{z}}_u)$ in the equation,

$$K_t(\mathbf{z}_r) + C_t(\dot{\mathbf{z}}_r) = \mathbf{f}_{(u)abs}$$

where

K_t and C_t are the tyre stiffness and damping functions respectively, and M_u is the unsprung mass inertial function.

- If accelerations for the sprung mass only are available, perform the following,
 - **STEP 3:** Solve for the generalised displacements of the unsprung mass \mathbf{z}_u in the system comprising the sprung mass. Thus the generalised vector of forces being represented by $\mathbf{f}_{(s)abs} = M_s(\ddot{\mathbf{z}}_s) + K_s(\mathbf{z}_s) + C_s(\dot{\mathbf{z}}_s)$ in the equation,

$$C_s(\dot{\mathbf{z}}_u) + K_s(\mathbf{z}_u) = \mathbf{f}_{(s)abs}$$
 - Proceed as in Step 5 above.
- If both sprung and unsprung mass accelerations are available, Step 5 is carried out directly.



3.3.1.1 Effects of forces acting directly on the sprung and unsprung masses

The results of applying the above procedure using unsprung mass accelerations only (Figure 3.4) and sprung mass accelerations (Figure 3.5) are presented in this Section. In the graphs, A-In implies that forces acting directly on the sprung and unsprung masses were included while R-O implies that only road input excitation was considered. Firstly, it is noted that road profiles reconstructed by using unsprung mass accelerations only (Figure 3.4) are more accurate than those obtained from using sprung mass accelerations only (Figure 3.5). For this vehicle, the error is almost double as much but the extent of this ratio between the errors is dependent on the type of vehicle and its suspension, and therefore cannot be generalised over all vehicles.

It is also noted that when only road inputs (R-O) are considered, the errors are higher than when all forces (A-In) are taken into account. The differences are quite significant and therefore it is recommended here that when these tests are carried out, necessary precautions must be taken to minimise the influence of these forces. Part of the solution lies in the DSD plots. There are some higher spectral energy near the spatial frequencies 0.2 cycles/m and 2 cycles/m which represent the frequencies at which the applied forces on the sprung mass and unsprung mass respectively were applied. Thus it is possible to use the DSD plots to determine the frequencies that can be safely filtered out so that the effects of these forces can be minimised especially in cases where it was obvious that they participated in the excitation of the vehicle.

In this case, the applied forces were chosen such that they could cause static deflections of 10 mm in the suspension (for forces acting directly on the sprung mass) and 10 mm in the tyre springs (for forces acting directly on the unsprung mass). However, when a vehicle is properly serviced, these forces cannot be permitted to cause such high deflections. Thus the errors due to neglect of these forces can actually be lower than 5% which may be practically permissible.

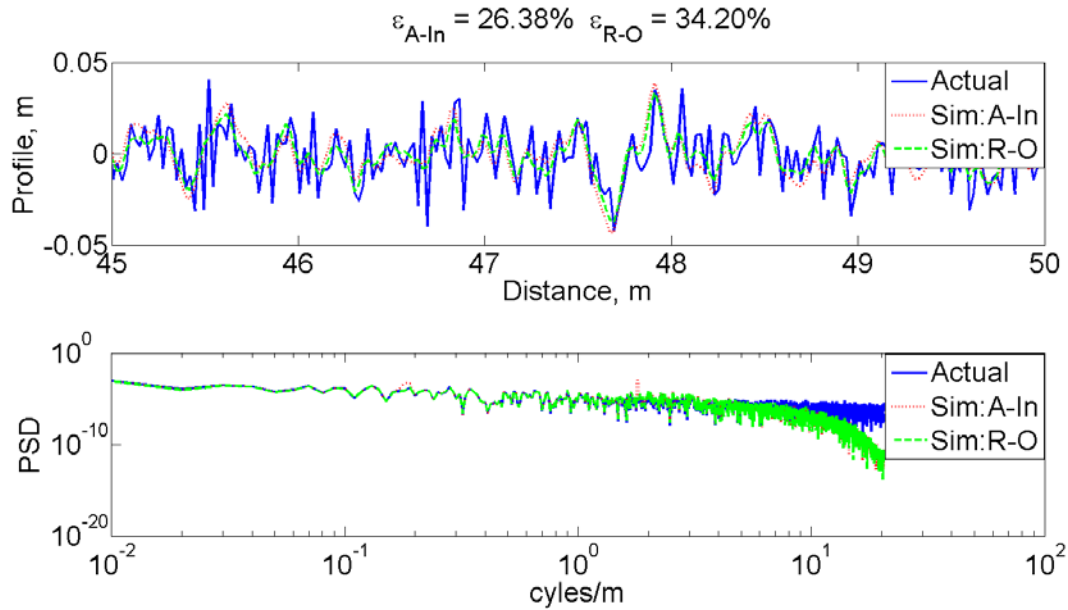


Figure 3.4. Comparing errors in profile reconstruction between pure road excited accelerations and contaminated acceleration when only unsprung mass accelerations are used.

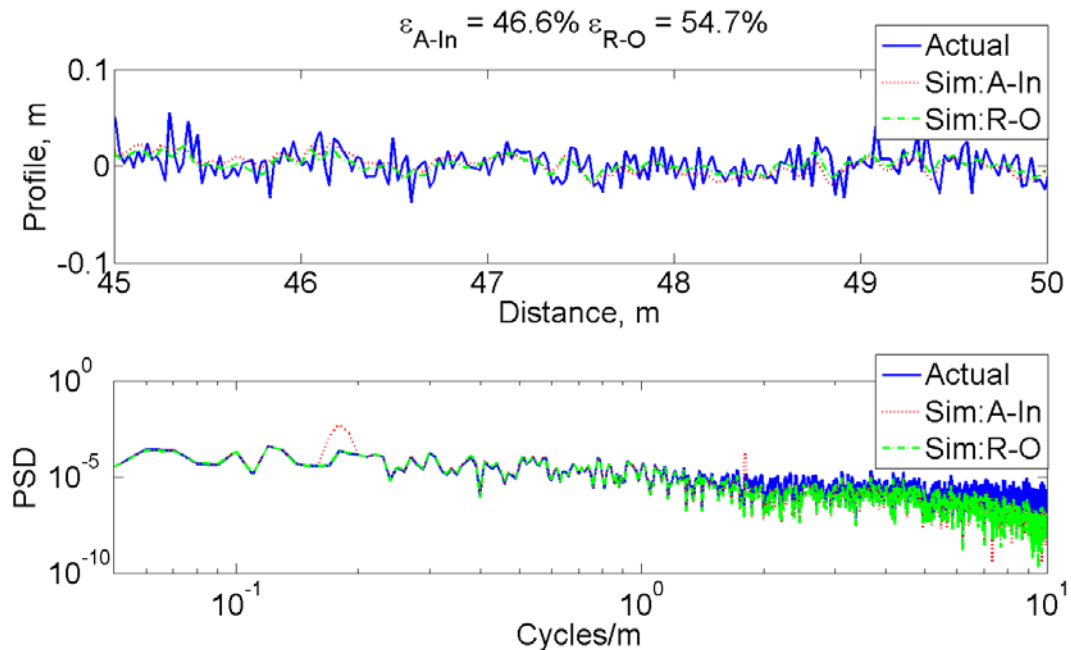


Figure 3.5. Comparing errors in profile reconstruction between pure road excited accelerations and contaminated acceleration when only sprung mass accelerations are used.

3.3.2 Calculation of dynamic tyre forces

A study by the Organization for Economic Co-operation and Development (DSTI, 1998) highlight the importance of dynamic tyre forces in prediction of road damage. The OECD study confirmed these



dynamic wheel forces depend on the suspension type, the profile of the road road and the speed of the vehicle. On reasonably smooth roads, for example, it is reported that steel suspensions produce impact factors of 1.3 while air suspensions produce impact factors between 1.1 and 1.15 for bogie axles of semi-trailer with negligible differences between the impact factors for single axles of rigid trucks or tractors (DSTI, 1998). On rough roads, however, it is reported that the steel suspensions produce impact factors of 1.4 – 1.5 while air suspensions produce impact factors of 1.2 for the bogie axles. In summary, stiffer suspensions cause more road damage than softer suspensions and the rate of wear increases with increased road surface roughness.

There are numerous techniques in the literature for the determination of these dynamic tyre forces but in this study, the method proposed by Cebon (1999) is adopted. Cebon (1999) presents two solution procedures, one specific for linear vehicle models while the other allows modelling of nonlinear suspension elements. The first method uses the Dynamic Load Coefficient (DLC) as a measure for dynamic tyre forces in which the DLC is obtained from the RMS dynamic load. In the second procedure, Cebon (1999) reformulates the equation of motion as a generalised multi-body vehicle model given by $\mathbf{M}\ddot{\mathbf{z}} = \mathbf{S}\mathbf{F}_s + \mathbf{T}\mathbf{F}_t + \mathbf{F}_g$ where \mathbf{F}_s are the generalised suspension forces, \mathbf{F}_t are the generalised tyre forces and \mathbf{F}_g are the gravitational forces and \mathbf{S} and \mathbf{T} are transformation matrices obtained from simple coordinate transformations. Then the tyre force (\mathbf{F}_t) and suspension force (\mathbf{F}_s) vectors are calculated from

$$\mathbf{F}_t = g_1(\mathbf{z}_r - \eta_{ut}) + g_2(\dot{\mathbf{z}}_r - \dot{\eta}_{ut}) \quad (3.14)$$

where

$g_i \dots i = 1, 2, \dots$ are functions of the displacements and velocities imposed on each tyre element; \mathbf{z}_r is the vector of road displacement input; η_{ut} denotes the displacement of the tyre attachment points on the unsprung mass; and

$$\mathbf{F}_s = g_3(\eta_{us} - \eta_{ss}) + g_4(\dot{\eta}_{us} - \dot{\eta}_{ss}) \quad (3.15)$$

where

η_{us} is the displacement of the suspension attachment points on the unsprung masses and η_{ss} is the displacement of the suspension attachment points on the sprung masses.

Then the road damage prediction may be calculated by using the method proposed by Pesterev and co-workers (2004; 2005) which is based on calculating the aggregate weighted force D_k^n given by



$$D_k^n = \sum_{j=1}^{N_a} \left(\frac{F_{t(jk)}}{a_j} \right)^n, \quad k = 1, 2, 3 \dots N_s \quad (3.16)$$

where $F_{t(jk)}$ is the force applied by the axle j to location k on the road, N_a is the number of axles on the vehicles, N_s is the number of the points along the road, a_j is the nominal tyre contact area for the axle j , n depends on the type of damage being considered, typically $n = 1$ for rutting while $n = 4$ for cracking. The criterion takes into account the accumulation of damage at various tyre contact points along the road as each vehicle axle passes and allows corrections to be made for different tyre configurations.

3.4 Road surface topography

3.4.1 Representation of road surface profiles

Johannesson, Podgórski and Rychlik (2014) state that road variability consists of three main components: topography (slowly changing landscapes), road roughness (road surface unevenness) and road texture (the high variability components). They further say that for fatigue applications, the road roughness is the one that is most relevant. Actually road roughness input is the principal source of vertical input to the vehicle exciting bounce and pitch motions. Another factor is that topography may be more pronounced in unpaved courses while road texture is more pronounced in paved courses, and often one assumes that landscape changes with wavelengths longer than 100 m cannot affect vehicle vertical dynamics while wavelengths shorter than 10 cm associated with road texture can be filtered out by the tyres' envelopment effect. Thus roughness acts as a key vertical displacement input to the wheels, exciting ride vibrations on both paved and unpaved road surfaces.

In order to simulate vehicle vibration behaviour for the purposes of calculating dynamic tyre forces or assessing ride performance, it is necessary to estimate a road roughness input (Cebon, 1999). Literature shows that typical road surfaces may be considered as realisations of homogeneous and isotropic two-dimensional Gaussian random processes and that a single direct spectral density function provides a road surface description that is sufficient for a vehicle response analysis (Andren, 2006; Bendat and Piersol, 2000; Dodds and Robson, 1973). Results of extensive studies of roads in four European countries by La Barre, Forbes and Andrew (1969) show that PSDs can reliably be used to represent road roughness. The isolated irregularities such as potholes or bumps can be treated separately. Thus it is generally found that the amplitude/wavelength characteristics of roads can be characterized by the spectral density of the profile height. However when the PSDs are calculated with a constant bandwidth method, their representation in a log-log plot gives heavy undulations at higher frequencies due to noise. Therefore ISO 8608 standard (ISO 8608: 1995E, 1995) recommends a further representation using smoothed form which is obtained by averaging the PSDs in the following manner: octave bands from the lowest calculated frequency (excluding zero) to a centre frequency of 0.0312 cycles/m; third-octave bands from the last octave band up to a centre frequency of 0.25 cycles/m; and twelfth-octave bands up to the highest calculated frequency. For simulation



purposes, it is more convenient to describe the road profile as a constant-velocity PSD where the smoothed form is fitted by a straight line in the spatial frequency range 0.011 cycles/m to 2.83 cycles/m.

This road roughness classification identifies eight levels ranging from Class A to Class H in increasing order of roughness. In the ISO classification, the relationships between the displacement power spectral density $S_u(\kappa)$ and the wave number κ for different classes of road roughness may be approximated by (ISO 8608: 1995E, 1995; Andr n, 2006)

$$S_u(\kappa) = S_u(\kappa_0) \left(\frac{\kappa}{\kappa_0} \right)^{-n} \quad (3.17)$$

where $S_u(\kappa)$ is expressed in $\text{m}^3 \text{cycle}^{-1}$, κ is expressed in cycles m^{-1} , $S_u(\kappa_0)$ is the displacement power spectral density at the datum or cut-off wave number κ_0 , which is equal to 0.1 cycles/m. With the constant n equal to 2.0. Equation (3.17) can be used to generate plots of $S_u(\kappa)$ as a function of the wave number ratio κ/κ_0 as shown in Figure 3. where for convenience roughness classes up to E are shown.

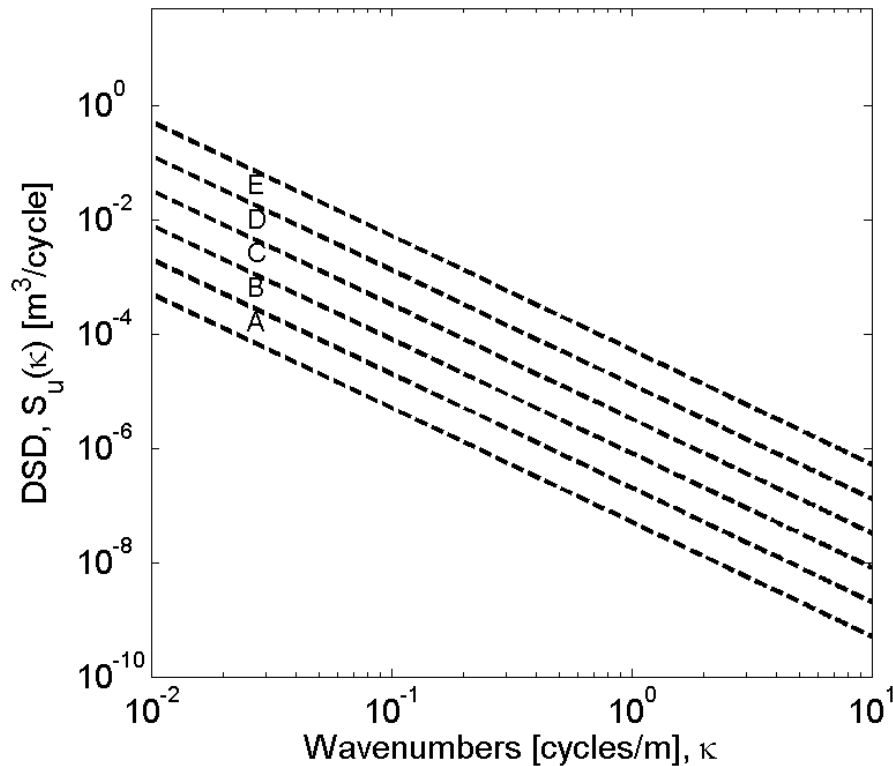


Figure 3.7 ISO classification for road roughness spectral densities.



Table 3.2 displays the values of the displacement spectral densities in equation 3.17 and gives a short description for each region in the graph (ISO 8608: 1995E, 1995). Obviously road roughness DSD lying below the lower limit for class A is a perfectly smooth road while that lying above the upper limit of E is extremely poor.

Table 3.2 Degree of road roughness using displacement spectral densities

Road Class	Displacement Spectral Density $S_U(\kappa_0) [10^{-6} m^3 / cycle]$	Description
A	up to 32	Very Good
B	32 – 128	Good
C	128 – 512	Average
D	512 – 2048	Poor
E	2048 – 8192	Very Poor
F	8192 – 32768	
G	32768 – 131072	
H	131072 - beyond	

The displacement spectral density as given in equation (3.17) is usually calculated from the measurement of surface roughness described by vertical ordinates at equally spaced points along the road. However, in the absence of such measurements, pseudo-random profiles can be generated to fit those spectral densities. Cebon (1999) presents a formula for generating a one-dimensional random profile as:

$$z_r = \sum_{k=0}^{N-1} \sqrt{S_k} e^{i\left(\theta_k + \frac{2\pi kr}{N}\right)}, \quad r = 0, 1, 2, \dots, (N-1) \quad (3.18)$$

where $S_k = (2\pi/N\Delta)S_{11}(\gamma k)$ and $S_{11}(\gamma k)$ is the target spectral density, $\gamma k = 2\pi k/N\Delta$ represents wave number in rad/m , Δ is the distance interval between successive ordinates of the surface profile, and θ_k is a set of independent random phase angles uniformly distributed between 0 and 2π .



The DSDs shown in Figure 3.7 are plotted for wavenumbers between 0.01 and 10 cycle/m. The frequency (Hz) is related to the wavenumber (*cycles/m*) by the relationship $f = v\kappa$, where v is the vehicle velocity in *m/s*. Thus, for a vehicle speed range between 20 and 80 *km/h* corresponds to frequency ranges between 0.009 and 35.4 Hz which are enough to cover interest ride vibration spectrum.

In practice, road surfaces may also come in the form of corrugations, cobbles, stoniness and localised defects such as pot-holes and bumps. Corrugations specifically result from the effects of wheeled equipment or after the wet seasons on country unpaved courses. Stoniness is a function of the surfacing gravel primarily, or its consequent loss thereby exposing the base layer and is also evident in unpaved roads that cut through rocky or hilly areas where there are a substantial amount of loose stones along the road. Discrete obstacles such as potholes and bumps are common to both paved and unpaved roads that are old and poorly maintained. These defects cannot be accurately represented in terms of DSD due to the averaging effect of the underlying algorithm in equation (3.17). If these discrete obstacles are not much, the resulting DSD may still be a good representation of the general road roughness over the road length under investigation.

3.4.2 International roughness index (IRI)

The International Roughness Index (IRI) is a popular measure of road roughness. It is based on a quarter-car model travelling at a speed of 80 km/h and calculates the suspension deflection of a simulated mechanical system with a response similar to a passenger car (Sayers, Gillespie and Paterson, 1986). The simulated suspension motion is accumulated and divided by the total distance travelled to give the units similar to slope (m/km) (Sayers and Karamihas, 1998). The algorithm that is used to calculate the IRI effectively filters out road wavelengths that do not impact much on the road users when the vehicle is travelling at a speed of 80 km/h. These non-critical wavelengths were calculated as those lying outside 1.3 m to 30 m range (Sayers, Gillespie, Paterson, 1986; COTO, 2007). Hence IRI values are occasionally evaluated for lengths less than 30 m; COTO (2007) recommends that IRI values be averaged over 100 m Sections. A detailed computation procedure for IRI can be found in the work by Sayers et al. (1986).

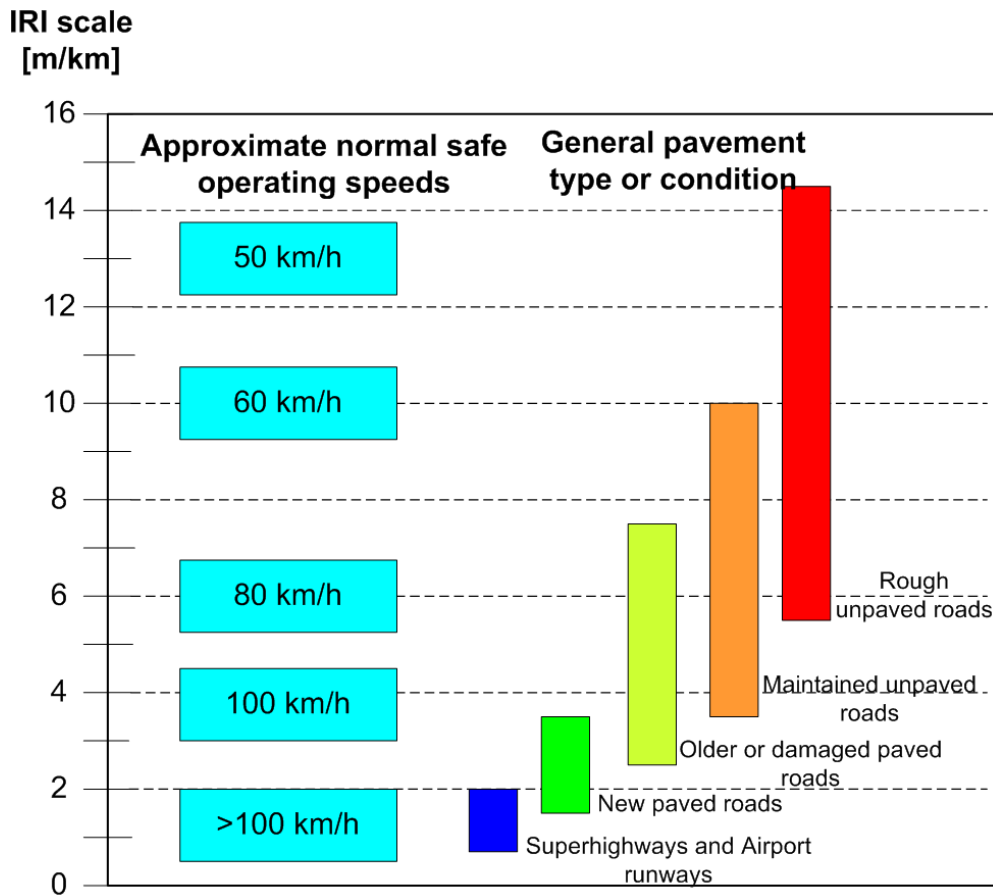


Figure 3.8 IRI scale and its interpretation adopted from COTO, 2007.

As shown in Figure 3.8 the IRI generally ranges from 0 to 16 with good condition paved roads lying between 1.5 and 3.5 while unpaved roads in good condition generally lie between 3.5 and 10 IRI (Sayers and Karamihas, 1998; COTO, 2007).

When compared to Figure 3.7 for ISO DSD scale, the IRI scale returns a single-valued statistic which seems to render itself easy for interpretation as to the approximate condition of the road surface and what level of vehicle speeds are safe over that surface. IRI can be more easily used for making road maintenance decisions than the ISO DSD scale. Its computation is also relatively cheaper than that for ISO DSD. However there are a number of other critical factors that make it less attractive than the ISO DSD to road analysts. IRI cannot be used as a road input for vehicle simulations. Thus there is no way of directly computing road profiles from these statistic values once the original measured profile is lost or made inaccessible by some intricate online condition monitoring systems. The IRI value is an averaged statistic over the entire road Section being monitored; hence it may spread the effect of localised defects over a very long road Section thereby making them wrongly insignificant and it may also be hard to extract any road-dependent defects, such as determining which wavelengths are the most problematic on the road. This can also make it hard to identify any localised defects since the averages can only be made for Sections longer than 30 m.



3.4.3 Summary of important road-vehicle interaction system characteristics

The following factors are noted by Cebon (1999) as necessary for understanding the characteristics of dynamic tyre forces. Since dynamic tyre forces are functions of road profiles, they are equally useful for understanding of road profiles. Only tyre envelopment is not included in Cebon's listing.

Probability distribution

The probability distribution of the road surface is generally observed to be broad-band, normally distributed realisations for continuous flexible or rigid roads. This is expected to be the case even in unpaved courses though it might not be the case in rigid roads with periodic joints between concrete slabs.

Magnitudes

The amplitudes of dynamic tyre forces depend on the road surface roughness and speed of the vehicle, its configuration, geometry and mass distribution, as well as the properties of the suspensions and tyres.

Frequencies

The vehicle ride vibrations fall into two distinct frequency ranges: sprung mass bounce, pitch and roll vibration modes in the range of 1.5 – 4 Hz and unsprung mass bounce and roll modes, and load-sharing suspension pitch modes between 8 and 15 Hz. Various experimental and theoretical studies have shown that the lower frequency sprung mass modes usually dominate the dynamic tyre forces on highways, except for vehicles with axle group suspensions with poorly damped bogie pitching modes.

Nonlinearity effects

Interleaf friction in leaf-spring suspensions can cause the natural frequencies of vehicles to be dependent on the amplitude of vibration and hence the roughness of the road surface. For high friction levels and smooth roads, interleaf friction can effectively lock the suspension so that the vehicle vibrates on its tyres with no relative suspension deflection. This can lead to lightly damped motion at frequencies of 3 – 4 Hz with little energy dissipation by the suspension and consequently high dynamic tyre forces.

Wheelbase filtering

The dominant frequencies of dynamic tyre forces can change significantly with this phenomenon. Despite the road surface containing some components of vibration, vehicle geometry can result in relative attenuation of some frequency components. These effects depend on the spacings between axles and the speed of the vehicle. As the vehicle traverses the road its rear wheels see same road excitation as its front wheels but only delayed by a time equivalent to the vehicle's wheelbase divided by the speed of travel. When a road wavelength equals the vehicle's wheelbase, the vehicle does not



experience any pitch motion but bounce. On the other hand, if its wheelbase is half a particular wavelength, pitch is present while bounce is not there. These ideas apply to the multiples of such wavelengths as well.

Such ideas are important during post processing of the road profile data. Since only bounce is important in the measured data, wheelbase filtering is implemented by applying stop bands around the frequencies given by

$$f = \frac{n.v}{WB} \quad (3.19)$$

where v is the velocity of the vehicle in m/s, WB is the wheelbase in m, and $n = 1/2, 3/2, 5/2, \dots$

Cross-coupling between suspensions

The tyre forces generated by an axle depend on the design of the whole vehicle not solely on the suspension of that axle. By implication therefore, quarter car models of articulated vehicles are not accurate for prediction of the road damage done by a suspension.

Payload distribution

The important low frequency bouncing and pitching modes of vibration depend on the moments of inertia of the vehicle and therefore on the distribution of payload.

Structural vibration

Suspension systems are located near to the nodes of first bending vibration mode as a result motion imposed on the axles by frame bending is small relative to motion caused by major sprung mass and unsprung mass rigid body resonances.

Roll motion influence

Under normal highway operating conditions of roughness and speed, roll-plane excitation of vehicles is small relative to pitch-plane excitation for frequencies below about 5 Hz. Thus sprung mass roll resonances which typically occur below 3 Hz are not significantly excited and hence do not contribute to the dynamic tyre forces.

Tyre non-uniformities

Radial run-outs, circumferential stiffness variations and mass out-of-balance can sometimes lead to measurable dynamic tyre force variation at the wheel rotation frequency (6 – 8 Hz) and higher harmonics. These forces can only be significant where they contribute forces that are comparable in magnitude to those caused by road input such as on relatively smooth roads.



Tyre Envelopment

The vehicle does not experience any excitation resulting from the road waviness whose wavelengths are shorter than the width of the tyre's contact patch. Consider that the contact patch has a width of a [m], then the road profile data may be low pass filtered below the following frequency,

$$f = \frac{v}{a}.$$

Spatial repeatability

There are a number of studies that have shown that, for any given testing speed, the wheel load time histories generated by a particular heavy vehicle are repeated closely on successive runs over a given stretch of road.

General remarks on tyres and suspensions

Lower suspension stiffness reduces tyre forces. Four-spring tandem suspensions were generally found to generate smaller dynamic tyre forces than walking beams. Torsion bar and air-suspensions generated the lowest loads. Lower tyre pressures usually result in reduced wheel loads. Wide based super single tyres were found to generate slightly lower dynamic tyre forces than dual tyres. Radial ply tyres are slightly preferable to bias ply tyres.

3.5 Conclusions

The analytical procedure for reconstruction of road profiles presented in this Chapter yields reasonably good accuracies in the presence of filtering effects at the higher frequency end. This approach has the advantage that it enhances physical insight since vehicle characteristics are used in the process and can easily be studied in a sensitivity analysis. There are however, some notable challenges in the application of this procedure to real vehicle systems:

- It requires a numerical model which implies that the vehicle system should be characterised. Such characterisation tests are highly specialised laboratory experimental procedures involving sophisticated equipment and highly trained personnel. Besides, some of the tests may be rather intrusive on the structural integrity of certain parts of the vehicle system thus the vehicle owners may not grant permission for that exercise to be carried out on their vehicles.
- Accelerations have to be numerically integrated twice to calculate displacements and once to calculate the velocities.
- It requires two numerical procedures to be carried out which might be computationally expensive.



Chapter 4

Theory of ANN modelling for road-vehicle interaction

This Chapter presents the theory governing the development of the artificial neural network (ANN) model for solution of function approximation such as the one being developed in this study. The Chapter starts by presenting the inverse problem solution in Section 4.1 as comprising three different types: simulation, deconvolution and system identification. ANN modelling is essentially a system identification problem which is accomplished during its training phase after which it becomes a simulation procedure. The formulation of the ANN model is carried out in Section 4.2 including the back-propagation and the Levenberg-Marquardt optimisation procedures. Section 4.3 presents the architectures of three common types of ANN models used for function approximation, namely, Feed-Forward Neural network (FFNN), Non-linear AutoRegressive with eXogenous inputs network (NARX), and Radial Basis Function Neural Network (RBFNN). The performances of these three networks are compared and the results are presented in Section 4.4. The results show that the NARX network outperforms the other two networks when operating at its optimal solution. Section 4.5 presents two ways by which the model performances have been assessed in this study. The Chapter is concluded in Section 4.6.

4.1 Introduction to inverse problem solution

If one considers a single-input-single-output (SISO) system with the input and output sequences denoted as $p(t)$ and $y(t)$ in continuous time representation, then the system dynamics may be represented by $D(t)$ so that $y(t) = D(t)p(t)$ (Worden and Tomlinson, 2001; Ljung, 1999). In principle given any two members of these quantities it is possible to estimate the third member. Worden and Tomlinson (2001) classify three different problems that arise based on this formulation namely, simulation, deconvolution and system identification. Simulation is when the input $p(t)$ and the system function $D(t)$ are given and it is required to solve for the system output $y(t)$. This is termed a forward problem and there is a substantial amount of literature that addresses such types of problems. Worden and Tomlinson (2001) describe both deconvolution and system identification as inverse problems but distinguish them in that deconvolution is the first type while system identification is the second type of the inverse problems. In the first type (deconvolution), the system output $y(t)$ and function $D(t)$ are given but it is required to estimate the input $p(t)$. In the second type (system identification), the system input $p(t)$ and output $y(t)$ are given but the system function $D(t)$ is the quantity to be estimated.

In the vehicle ride dynamics problem, road profiles form part of the forcing inputs to the vehicle system while vertical vehicle responses (e.g. accelerations) are part of the outputs. Thus as already mentioned in Chapter 3, the current study involves solution of an inverse problem and what was demonstrated in Chapter 3 is an example of deconvolution. The vehicle system function was assumed to be given in the form of a quarter car model from which both sprung mass and unsprung mass



accelerations were obtained. As demonstrated in that Chapter, the actual process of deconvolution is a lot more involved than the literal mathematical representation: $p(t) = D^{-1}(t)y(t)$ since such direct inverse for the system function $D(t)$ may not always exist. The computational difficulties in employing deconvolution multiply with existence of nonlinearities in the system. This is where system identification offers some advantages in that an inverse system function is obtained by attempting to solve a reverse problem where $y(t)$ form system inputs and $p(t)$ are system targets after which simulation is performed to solve for $p(t)$ given new set of $y(t)$. Ljung (1999) has a comprehensive coverage of system identification techniques applied to time invariant systems. However, artificial neural networks are particularly popular for solution of problems with high nonlinearities in the system. This is the approach adopted in this study. The biggest challenge though is how to make this so-realised ‘inverse’ model representative of real system dynamics in the presence of so many different operating conditions.

In Section 4.2, artificial neural networks are introduced in terms of how they are used to solve function approximation problems and some of the popular neural networks in function approximation are presented and discussed in Section 4.3. Their relative computational performances are also presented after which two neural networks are selected for further application to the different case studies. Finally, Section 4.4 presents the two procedures that have been used to validate neural network model performance in this study.

4.2 Artificial neural networks (ANN)

4.2.1 Definition and architecture

Haykin (2009) defines an artificial network as a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the human brain in that it acquires ‘knowledge’ through a learning process and that ‘knowledge’ is stored in interconnection weights just like the brain’s synaptic weights. The learning process involves adjusting the interconnection weights until they meet a certain goal. Figure 4.1 shows a clearly labelled diagram of a typical artificial neural network model.

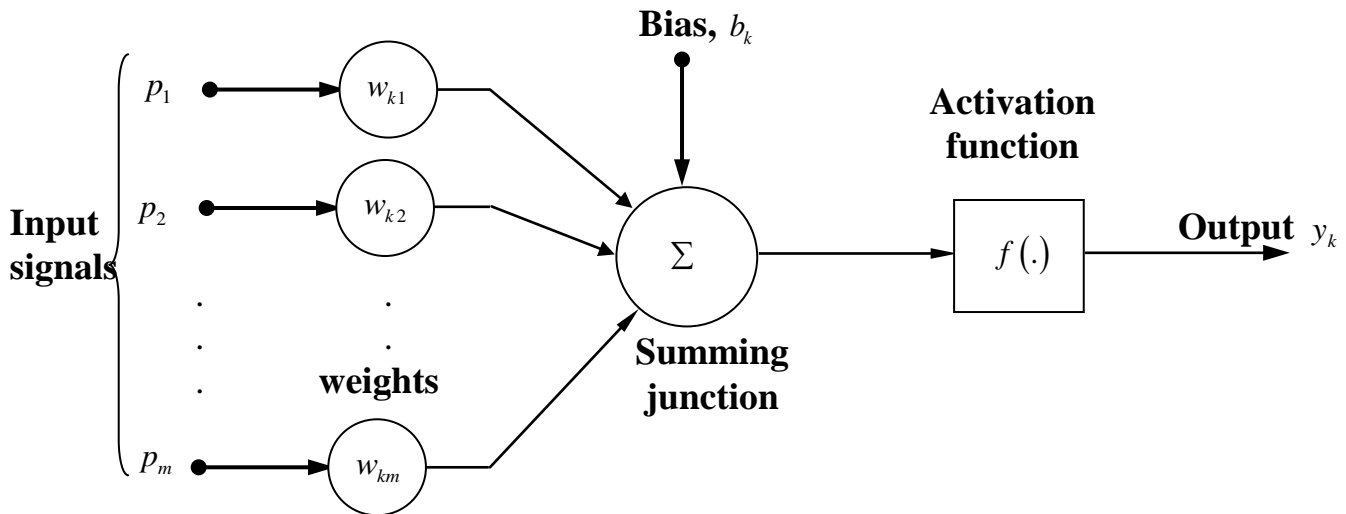


Figure 4.1. Neural network model showing basic elements. **Conclusions**

In Figure 4.1, the weights are used for ‘weighting’ the inputs and may lie in the range that includes negative as well as positive values. An activation function moderates the amplitude of the output. There are three main types of these functions namely, the threshold function, the linear function and the sigmoid function. The threshold function produces outputs that are either 1 or zero depending on the range of the weighted input values. It is largely used in pattern classification problems. The sigmoid function is by far the most common activation function and gets its name from its s-shape that is perfectly blended between two t-ordinates. It has two main variations: logistic sigmoid function which characterised by s-shaped fit between zero and +1, and the tangent sigmoid (hyperbolic) function which assumes values between ± 1.0 . Though by design, the range is actually ± 1.7159 to allow for the squashing effect of the sigmoid function. The linear function reproduces the inputs with or without a scale. When it reproduces the inputs without scaling them, it is termed pure-linear function. This is the most common form of the linear functions.

Artificial neural networks are mainly used for two purposes: function approximation/regression and pattern classification. In this application, the ANN is used for function approximation. Haykin (2009) summarises the following useful properties and capabilities offered by artificial neural networks:

Nonlinearity

ANN can model underlying nonlinear system behaviour by making its activation functions nonlinear. This nonlinearity is distributed throughout the network.



Input-Output mapping

ANN constructs an input-output mapping for any problem using available input-output data which it uses as a learning example.

Adaptivity

ANN has a built-in capability to adapt their synaptic weights to changes in the surrounding environment. Neural networks can be retrained to deal with minor changes in the operating environmental conditions or designed to change its synaptic weights in real time.

Evidential response

In pattern classification problems it is possible to define the confidence in the selection made which implies that ANN can be used to identify spurious/ambiguous patterns when they arise and thereby assist in improving its classification performance.

Contextual information

Every neuron in the network is potentially affected by the global activity of all other neurons in the network. Consequently, contextual information is dealt with naturally by a neural network.

Fault tolerance

In principle, a neural network implemented in a hardware component, exhibits graceful degradation in performance rather than catastrophic failure.

Very-large-scale-integrated (VLSI) implementability

The fast nature in processing via its massively parallel processor makes it well-suited for implementation using VLSI technology. One particular beneficial virtue of VLSI is that it provides a means of capturing truly complex behaviour in a highly hierarchical fashion (Haykin, 2009).

Uniformity of analysis and design

Neural networks enjoy universality as information processors.

Neurobiological analogy

The analogy with neurobiology implies that researchers can borrow ideas from each other to develop new techniques.



4.2.2 ANN in function approximation

As already mentioned, ANNs are used in this study, for approximation of road profiles given vertical vehicle responses. It is thus imperative that regression capabilities of ANNs and ways of improving them by ensuring good generalization are discussed. The first question though to be addressed is: what is the minimum number of hidden layers that are required for multilayer perceptron to produce a good approximation of the desired target (Haykin, 2009)? The universal approximation theorem states that a single hidden layer is sufficient for a multilayer perceptron to produce a good enough approximation to a given training set represented by the set of inputs p_1, \dots, p_s and a desired target $f(p_1, \dots, p_s)$ (Haykin, 2009).

ANNs are trained in two main ways: Supervised training where they learn through examples of input and output data, and unsupervised training where the neural networks learn the functional mapping from the statistical properties and probability distribution of the input data only. In this study, ANNs learn by supervised training. Therefore in the following Section two of the most popular ANN learning routines have been presented, namely the pure back-propagation and the Levenberg-Marquardt routines.

4.2.3 Back-propagation

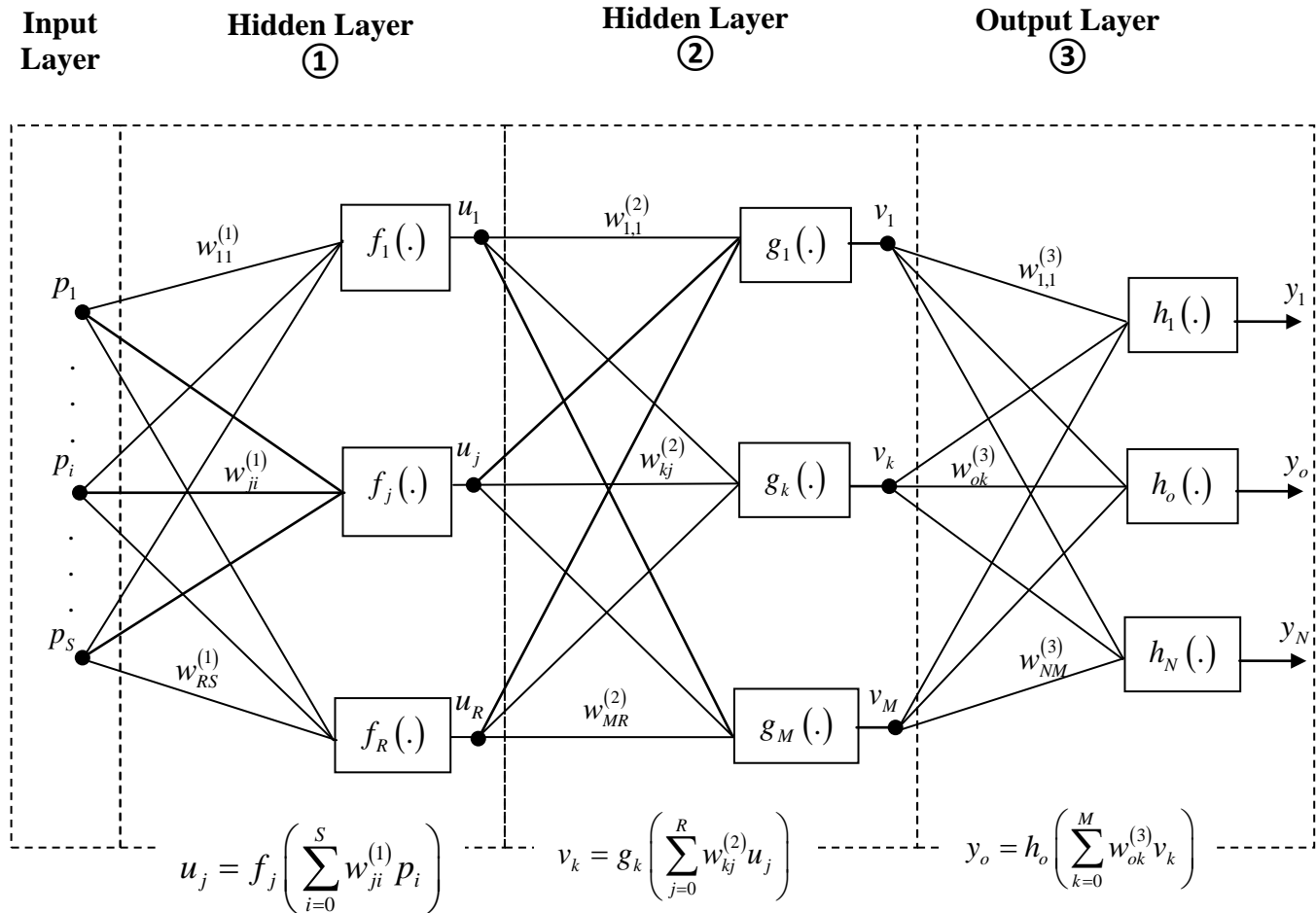


Figure 4.2. General architecture of a multilayer perceptron (MLP).

Back-propagation is a process of calculating how much contribution to the ensemble averaged error has been made by each of the elements of the weight matrices in the network. This clearly involves calculating partial derivatives of the ensemble averaged error of the network with respect to each of the elements of the weight matrices. Multilayer perceptron (MLP) architecture in Figure 4.2 is used to illustrate the formulation of the back-propagation procedure. The architecture consists of a total number of S inputs (labelled p) and N outputs (labelled y). It has two hidden layers and though the neuron activation functions are denoted differently, it is done to simply avoid confusion later on during the demonstrations of the back-propagation, but they essentially utilise the same sigmoid function. The output layer neurons carry a linear activation function which is like a signal scaling function.



If t_o denotes the target for the network node producing y_o then the ensemble averaged error over the batch being trained can be given by

$$\varepsilon_o = \frac{1}{2Q} \sum_{q=1}^Q (t_o - y_o)^2 \quad (4.4)$$

where Q is the total number of samples in the batch being used for training (Haykin, 2009).

Rewriting equation (4.1) by unwrapping all the functional terms contained in the network output y_o results in the ensemble averaged error being expressed as

$$\varepsilon_o = \frac{1}{2Q} \sum_{q=1}^Q \left(t_o - h_o \left(\sum_{k=0}^M w_{ok}^{(3)} g_k \left(\sum_{j=0}^R w_{kj}^{(2)} f_j \left(\sum_{i=0}^S w_{ji}^{(1)} p_i \right) \right) \right) \right)^2 \quad (4.5)$$

The partial derivatives of ε_o with respect to the weight elements are given as follows

$$\frac{\partial \varepsilon_o}{\partial w_{ok}^{(3)}} = \frac{-1}{Q} \sum_{q=1}^Q \left(t_o - h_o \left(\sum_{k=0}^M w_{ok}^{(3)} v_k \right) \right) \times \frac{\partial h_o \left(\sum_{k=0}^M w_{ok}^{(3)} v_k \right)}{\partial w_{ok}^{(3)}} \quad (4.6)$$

$$\frac{\partial \varepsilon_o}{\partial w_{ok}^{(3)}} = \underbrace{\frac{-1}{Q} \sum_{q=1}^Q \left(t_o - h_o \left(\sum_{k=0}^M w_{ok}^{(3)} v_k \right) \right)}_{\text{Average error at Output Layer} = e_o} \times \sum_{k=0}^M v_k \cdot h_o' \left(\sum_{k=0}^M w_{ok}^{(3)} \right) \quad (4.7)$$

$$\frac{\partial \varepsilon_o}{\partial w_{ok}^{(3)}} = -e_o \times h_o' \left(\sum_{k=0}^M w_{ok}^{(3)} \right) \times \sum_{k=0}^M v_k \quad (4.8)$$

In this study, the activation function h_o is actually a constant because the output layer neuron is a linear function, hence the result in equation (4.5) is modified accordingly and yields.

$$\frac{\partial \varepsilon_o}{\partial w_{ok}^{(3)}} = -e_o \times h \times \sum_{k=0}^M v_k \quad (4.9)$$

where h is a constant.

$$\frac{\partial \varepsilon_o}{\partial w_{kj}^{(2)}} = \underbrace{\frac{-1}{Q} \sum_{q=1}^Q \left(t_o - h_o \left(\sum_{k=0}^M w_{ok}^{(3)} \left(g_k \left(\sum_{j=0}^R w_{kj}^{(2)} u_j \right) \right) \right) \right)}_{\text{Average error at Output Layer} = e_o} \times h_o' \left(\sum_{k=0}^M w_{ok}^{(3)} v_k \right) \times g_k' \left(\sum_{j=0}^R w_{kj}^{(2)} u_j \right) \quad (4.10)$$

$$\frac{\partial \varepsilon_o}{\partial w_{kj}^{(2)}} = -e_o \cdot h \sum_{k=0}^M v_k g_k' \left(\sum_{j=0}^R w_{kj}^{(2)} \right) \cdot \sum_{j=0}^R u_j \quad (4.11)$$

$$\frac{\partial \varepsilon_o}{\partial w_{kj}^{(2)}} = \frac{\partial \varepsilon_o}{\partial w_{ok}^{(3)}} \times g_k' \left(\sum_{j=0}^R w_{kj}^{(2)} \right) \times \sum_{j=0}^R u_j \quad (4.12)$$



$$\frac{\partial \varepsilon_o}{\partial w_{ji}^{(1)}} = \frac{-1}{Q} \sum_{q=1}^Q \underbrace{\left(t_o - h_o \left(\sum_{k=0}^M w_{ok}^{(3)} g_k \left(\sum_{j=0}^R w_{kj}^{(2)} f_j \left(\sum_{i=0}^S w_{ji}^{(1)} p_i \right) \right) \right) \right)}_{\text{Average error at Output Layer} = e_o} \times h_o' \left(\sum_{k=0}^M w_{ok}^{(3)} v_k \right) \times \dots \quad (4.13)$$

$$g_k' \left(\sum_{j=0}^R w_{kj}^{(2)} u_j \right) \times f_j' \left(\sum_{i=0}^S w_{ji}^{(1)} p_i \right)$$

$$\frac{\partial \varepsilon_o}{\partial w_{ji}^{(1)}} = -e_o \cdot h \sum_{k=0}^M v_k g_k' \left(\sum_{j=0}^R w_{kj}^{(2)} \right) \sum_{j=0}^R u_j f_j' \left(\sum_{i=0}^S w_{ji}^{(1)} \right) \cdot \sum_{i=0}^S p_i \quad (4.14)$$

$$\frac{\partial \varepsilon_o}{\partial w_{ji}^{(1)}} = \frac{\partial \varepsilon_o}{\partial w_{kj}^{(2)}} \times f_j' \left(\sum_{i=0}^S w_{ji}^{(1)} \right) \times \sum_{i=0}^S p_i \quad (4.15)$$

As can be noted from equations (4.3) to (4.12) the back-propagation procedure can also be used to compute the first order partial derivatives of output error function ε_o at a specific output neuron with respect to a particular network weight in any layer using manipulation of first order partial derivatives only. This is very useful in the computation of the Hessian matrix for the Levenberg-Marquardt optimisation procedure presented in Section 4.2.4 (Bishop, 1995; Haykin, 2009).

Ways of improving the performance of back-propagation algorithms

Haykin (2009) presents a number of ways on how to make back-propagation algorithms perform better. Since back-propagation forms the heart of neural network modelling, these ways have been summarised in this Section as applicable to a function approximation problem.

1. Sequential vs batch learning

Back-propagation algorithms are designed to learn the input-output mapping in two modes: sequential (stochastic) and batch learning. In sequential learning the neural network is presented with the input-output pair for a single sample point after which the weights are updated before a next pair is presented. These input-output pairs are selected randomly from the training data sample in order to ensure that the examples do not come from the same class. The neural network weights are updated sample-point-by-sample-point. In batch learning mode, the neural network is presented with a number of sample points that represent the overall input-output mapping. Thus the weights can only be updated after the neural network learns the overall trend. This is computationally slower than the sequential learning and is also said to suffer from singularity problems in computing the Jacobian especially with large data samples with high levels of redundancy.



2. Maximising information content

A good training sample data should contain as much information as possible which is representative of the actual system under study (LeCun, 1993 cited by Haykin, 2009). This can be achieved by using examples that are radically distinct from one another and yield the largest training errors. It is not desirable to use examples that correlate with one another. The examples with the largest errors ensure that more of the weight space is searched.

3. Choice of sigmoid activation functions

The neural network training process benefits from the use of sigmoid activation functions that are odd functions of their arguments. This is a condition that is satisfied by the hyperbolic function (Haykin 2009).

4. Preprocessing of training data

The target values should be processed in such a way that they fit within the range of values normally spanned by the hyperbolic function. Normally the tangent sigmoid function is designed to span the range ± 1.7159 , thus the target values can be processed so that their values fall in the range ± 1.0 (Figure 4.3), otherwise the back-propagation algorithm tends to drive the network weights and biases to very large values, thereby slowing down the learning process by driving the hidden neurons into saturation (Haykin, 2009).

The input variable data should be equally normalised as well as removing its mean value. In order to accelerate the training process the data set should be uncorrelated and their covariances set to be approximately equal to ensure that the different weights in the network learn at approximately the same speed (Haykin, 2009). These points can be summarised in four steps involving: normalisation, detrending for mean values, decorrelation, and covariance equalization.

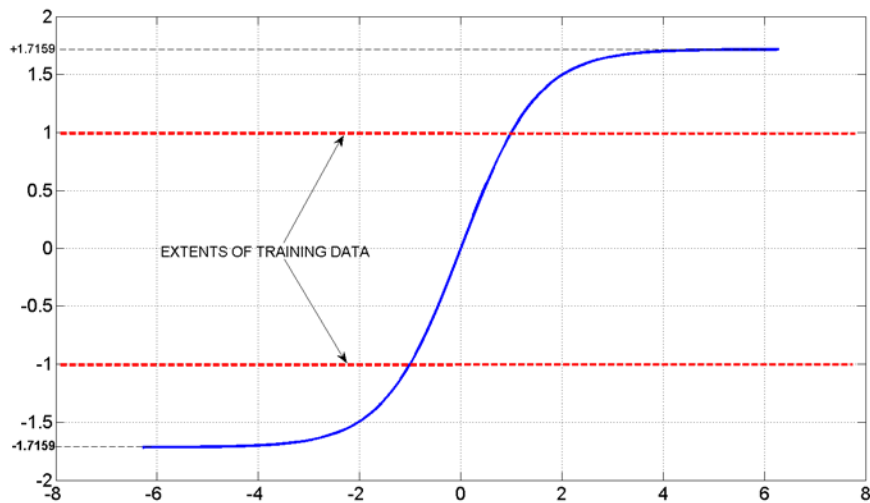


Figure 4.3. Illustration of recommended training data extents for a hyperbolic tangent activation function.

5. *Choice of initial values of the weights*

The initial values of the weights determine the starting point of the learning procedure. If initial weights are chosen very close to the optimal weights, the learning process can be very fast and computational less demanding. When the weights are assigned large initial values, it is likely that the neurons in the network will be driven into saturation (Haykin, 2009). If this happens the local gradients in the back-propagation algorithm assume small values, which in turn will cause the learning process to slow down (Haykin, 2009). On the other hand if the weights are assigned small initial values, the back-propagation may operate on a very flat area around the origin of the error surface where it can get stuck. Thus, the proper choice of the initialisation lies somewhere between the two extremes. Haykin (2009) recommends that the initial weights should be selected in such a way that they have a mean of zero and a variance equal to the reciprocal of the number of weights in the neuron.

6. *Use of prior information about the system dynamics*

It is recommended that any other prior information available about the system should be included in the learning process to facilitate the process. Haykin (2009) states that information such as invariance properties, symmetries, or any other knowledge about the system function can be used to advantage.



7. Learning rates

A learning rate is a factor that controls the size of weight and bias changes during learning of the training algorithm. This parameter is applied to the gradients which define the rate of change of the weights with respect to the neuron's output error. The last layers usually have larger local gradients than the layers at the front end of the neural network. Hence, Haykin (2009) recommends that the learning-rate parameter should be assigned smaller values in the last layers than in the front layers of the multilayer perceptron. Neurons with many inputs should have a smaller learning-rate parameter than neurons with few inputs so as to maintain a similar learning time for all neurons in the network. LeCun, 1993 (cited by Haykin, 2009) recommends that the learning-rate parameter should be inversely proportional to the square root of the number of weights connected to that neuron.

4.2.4 Levenberg-Marquardt

This is one of the learning procedures where the supervised learning problem is viewed as a numerical optimisation problem. The error function is optimised with respect to the network weights. Using the Taylor series expansion we can express the error surface as (Haykin, 2009)

$$\varepsilon(\mathbf{w} + \Delta\mathbf{w}) = \varepsilon(\mathbf{w}) + \mathbf{g}^T \Delta\mathbf{w} + \frac{1}{2} \Delta\mathbf{w}^T \mathbf{H} \Delta\mathbf{w} + (\text{third- and higher-order terms}) \quad (4.13)$$

where \mathbf{g} is the local gradient vector, defined by

$$\mathbf{g} = \frac{\partial \varepsilon(\mathbf{w})}{\partial \mathbf{w}} \quad (4.14)$$

The matrix \mathbf{H} is the local Hessian representing "curvature" of the error performance surface, defined by

$$\mathbf{H} = \frac{\partial^2 \varepsilon(\mathbf{w})}{\partial \mathbf{w}^2} \quad (4.15)$$

In the steepest descent method as represented by the back-propagation algorithm, the adjustment $\Delta\mathbf{w}$ applied to the network weights is defined by

$$\Delta\mathbf{w} = -\eta\mathbf{g} \quad (4.16)$$

where η is the fixed learning-rate parameter. In effect the steepest descent method operates on the basis of a linear approximation of the cost function in the local neighbourhood of the operating point. This has the advantage that it is simple since it only uses first order approximation but it suffers from slow rate of convergence due to negligence of higher order effects. The inclusion of the momentum term in the update equation for weights is a crude attempt at using second-order information about the error surface. However it makes the training process more delicate to manage by adding one more item to the list of parameters that have to be tuned by the designer.



Haykin (2009) states that in order to produce a significant improvement in the convergence performance of an MLP, higher-order information have to be used in the training process. A quadratic approximation of the error surface around the operating point can be employed to achieve that. Thus the optimum value of the adjustment $\Delta \mathbf{w}$ applied to the weights is given by (Haykin, 2009)

$$\Delta \mathbf{w}^* = \mathbf{H}^{-1} \mathbf{g} \quad (4.17)$$

where

\mathbf{H}^{-1} is the inverse of the Hessian matrix assuming that it exists. Equation (4.17) is the essence of Newton's method and it converges in one iteration if the error function is quadratic. However, Haykin (2009) states that the practical application of the Newton method is hampered by the following three factors:

- It requires calculation of the inverse Hessian, which can be computational expensive.
- The Hessian matrix must be nonsingular. This is a very hard condition due to the nature of supervised learning problems which are often ill-conditioned.
- When the error function is non-quadratic, there is no guarantee that the Newton method will converge which makes it unsuitable for the training of a multilayer perceptron.

The Levenberg-Marquardt method therefore, modifies equation (4.17) in such a way that it forces the supervised learning problem to become well-conditioned throughout the training process by the expression:

$$\Delta \mathbf{w}^* = [\mathbf{H} + \lambda \mathbf{I}]^{-1} \mathbf{g} \quad (4.18)$$

where

\mathbf{I} is the identity matrix of similar dimension to \mathbf{H} and λ is the regularizing parameter that forces the sum matrix to be positive definite. If a multilayer perceptron with a single output neuron is considered, the error/cost function can be expressed as

$$\varepsilon(\mathbf{w}) = \frac{1}{2N} \sum_{i=1}^N [y(i) - F(\mathbf{p}(i); \mathbf{w})]^2 \quad (4.19)$$

where the training sample is given by $\{\mathbf{p}(i); y(i)\}_{i=1}^N$ and $F(\mathbf{p}(i); \mathbf{w})$ is the approximating function realised by the neural network and is given as a function of the inputs \mathbf{p} at an instant i and weights \mathbf{w} . The gradient and Hessian of the error function are given by

$$\mathbf{g} = -\frac{1}{N} \sum_{i=1}^N [y(i) - F(\mathbf{p}(i); \mathbf{w})] \times \frac{\partial F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}} \quad (4.20)$$



$$\mathbf{H} = -\frac{1}{N} \sum_{i=1}^N [y(i) - F(\mathbf{p}(i); \mathbf{w})] \times \frac{\partial^2 F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}^2} + \dots$$
$$\frac{1}{N} \sum_{i=1}^N \left(\frac{\partial F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}} \right) \left(\frac{\partial F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}} \right)^T \quad (4.21)$$

In order to mitigate the computational cost of evaluating the Hessian in equation (4.21), the first term is neglected and the Hessian is approximated by the second term only, which comprises first order terms only

$$\mathbf{H} \approx \frac{1}{N} \sum_{i=1}^N \left(\frac{\partial F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}} \right) \left(\frac{\partial F(\mathbf{p}(i); \mathbf{w})}{\partial \mathbf{w}} \right)^T \quad (4.22)$$

These first order partial derivatives can be efficiently computed using the back-propagation procedure. This second order partial derivative approximation is termed the outer product approximation of the Hessian matrix and its use is justified when the Levenberg-Marquardt is operating in the neighbourhood of the local or global minimum. The regularising parameter λ plays a critical role in the way the Levenberg-Marquardt algorithm functions. If λ is assigned a value equal to zero, equation (4.18) reduces to a Newton formula; if a large value is assigned, the algorithm behaves like a gradient descent method. This switching functionality exploits the different capabilities between the Newton method and the gradient descent method. When the algorithm is operating near the optimum point, the Newton method causes it to converge rapidly towards the optimal point while the gradient descent method ensures that it does attain convergence through a proper selection of step-size parameter though at a very slow rate.



4.2.5 ANN generalisation

Generalisation is when a neural network is able to produce a correct input-output mapping even when the input was not the one used during its training process (Haykin, 2009). The goal of neural network training is not to learn the training data exactly and fail to generalise its response over fresh data extracted from similar set. Any properly designed ANN model should be able to replicate its performance, with comparable accuracy, over data that has not been seen before. Generalisation is influenced by three factors: the size and representativeness of the training sample to the operating environment; the architecture of the neural network; and the physical complexity of the problem (Haykin, 2009).

The training sample should be as representative as possible to the prevailing operating conditions. The data should cover as much as possible of the space of the data being generated by the process under investigation. In other words, it is necessary to ensure that the training data contain all of the important spectral properties typical of data generated by the system under all key operating conditions. This obviously requires a proper prescription of the operating conditions with respect to function approximation problems since they are often too dependent on these conditions.

In polynomial curve fitting, it is observed that when the polynomial coefficients are too few, the polynomial is too inflexible so that it fails to give good approximations over new data, on the other hand, when the coefficients are too many, the polynomial fits too much of the noise on the data that was used to build the polynomial. This is also experienced in designing of neural network models. Therefore, in order to achieve good network generalisation, there is need to find a way of controlling the complexity of the network architecture. In the case of artificial neural networks this is achieved by changing the number of adaptive parameters in the network (Bishop, 1995). One way to address this problem is to compare a range of models with different complexities and then select the one that seems to yield optimal performance. The other more automatic way is through cross-validation. The training set is partitioned into three data sets: estimation set, validation set and test set. The main rationale is to validate the model on a different data set from the one that is used in training it and then, in order to avoid over-fitting the validation data set, the network is further tested on another data set. This is a more practical approach and it is the one which has been employed in this study. Haykin (2009) recommends this approach for designing large networks with good generalisation.

Problem complexity results in loss of quality in the network generalisation with similar training data and network architecture. Besides, there is no guarantee that any two similar physical problems at different levels of complexities though being modelled by similar neural networks will experience similar quality in network generalisation. This is where ANN design has sometimes been argued to be a combination of science and art.

In Section 4.3, three different neural networks are presented. Their choice has been largely due to their popularity in function approximation problems. Later in Section 4.4 their computational performances are compared upon applying them to road profile reconstruction of a number of selected road defects.

4.3 ANN architectures and formulations

In this Section, the architectures and mathematical formulations of three neural network models commonly used in function approximation are presented. The first is the static feed-forward neural network (FFNN) with a traditional MLP architecture. It rather acts as a control or reference in this investigation. The second network model is the autoregression with exogenous inputs (NARX) network. This network has feed-back of output onto its input layer. The third is the radial basis function neural network (RBFNN) which is based on exact interpolation method using radial distance between the weight parameters and the inputs.

4.3.1 Static feed-forward neural network (FFNN)

This neural network was used in this work simply as a means of testing the extent to which time dependency modelling in the dynamic neural network impact the accuracy of the resulting solution. When using feedback of the network output there is a computational overhead that is encountered and it increases with an increase in number of delays and neurons in the hidden layer. The FFNN being presented here does not suffer from that problem though it is realised that for the same quality of performance the FFNN requires a larger number of neurons than the NARX in the hidden layer. Its architecture is given in Figure 4.4.

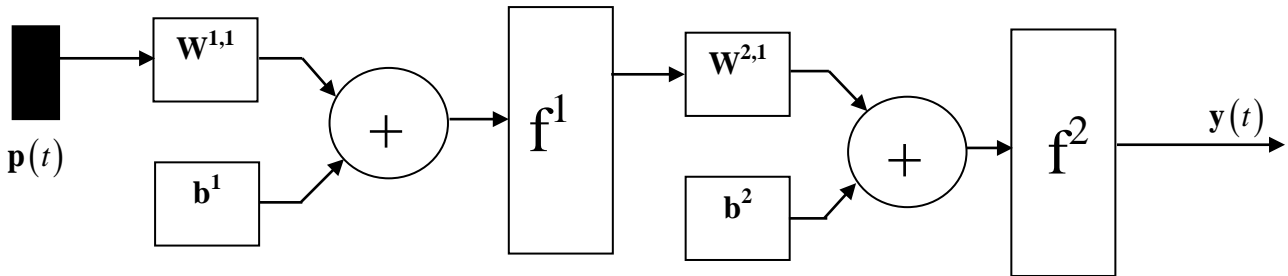


Figure 4.4. Architecture of the FFNN.

The hidden neural network transfer function f^1 is a hyperbolic function and output network transfer function f^2 is the pure-linear function. Thus the FFNN model can be represented by

$$y(t) = f^2 \left[\underbrace{\mathbf{W}^{2,1} \times f^1 \left[\overbrace{\mathbf{W}^{1,1} \times \mathbf{p}(t) + b^1}^{\text{Hidden Layer}} \right] + b^2}_{\text{Output Layer}} \right]$$

$$y(t) = \mathbf{W}^{2,1} \times \tanh(\mathbf{W}^{1,1} \times \mathbf{p}(t) + b^1) + b^2 \quad (4.23)$$



For sequential training, the training data are presented as cell arrays (not as matrices). Assuming that three main different operating conditions have been identified, the data may be packaged in the format:

$$\begin{aligned} \mathbf{p}\{i\} &= [p_{tr1} \quad p_{tr2} \quad p_{tr3}] \\ \mathbf{y}\{i\} &= [y_{tr1} \quad y_{tr2} \quad y_{tr3}] \end{aligned} \quad (4.24)$$

where i represents a training epoch and the superscripts $tr1$, $tr2$ and $tr3$ identify data selected from each of the three operating conditions. The curly brackets simply denote that data are packaged as cell arrays, a notation that is adopted from the MATLAB coding which has been used in this study. In order to ensure network generalisation across these different operating conditions, different noise levels expressed as fractions of input and output amplitudes can be added to the training data to derive validation and testing data sets. This might be useful where data are scarce as will be demonstrated in one of the case studies later. Thus the validation and test data sets may be generated in a similar manner in the following manner:

$$\begin{aligned} \mathbf{p}_{v|t}\{i\} &= [p_{tr1} + \rho_1 \quad p_{tr2} + \rho_2 \quad p_{tr3} + \rho_3] \\ \mathbf{y}_{v|t}\{i\} &= [y_{tr1} \quad y_{tr2} \quad y_{tr3}] \end{aligned} \quad (4.25)$$

where $\rho_n = c \cdot p_{trn}$; ($n = 1, 2, 3$) and $0 < c \leq 0.1$ and the subscript s/t represent either validation or test data sets.

4.3.2 The nonlinear autoregressive network with exogenous inputs (NARX)

The NARX network is a recurrent dynamic network, with feedback connections enclosing several layers of the network (Mathworks, Inc. 2007). The feedback connections make the model structure considerably complex but at the same time offer very useful flexibility. It is expressed as

$$y(t) = f\left(y(t-1), y(t-2), \dots, y(t-n_y), p(t-1), p(t-2), \dots, p(t-n_p)\right) \quad (4.26)$$

where the next value of the dependent output signal $y(t)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal (Mathworks, Inc. 2007). An architecture of the NARX network is given in Figure 4.5. The mathematical representation of this model is given as

$$\mathbf{y}(t) = f^2\left(\mathbf{W}^{2,1} f^1\left(\mathbf{W}^{1,1} \mathbf{p}_{TDL} + \mathbf{W}^{1,3} \mathbf{y}_{TDL} + \mathbf{b}^1\right) + \mathbf{b}^2\right) \quad (4.27)$$

where \mathbf{p}_{TDL} and \mathbf{y}_{TDL} represent the time delay lines from the inputs $\{p(t), p(t-1), \dots, p(t-n_p)\}$ and network outputs $\{y(t-1), y(t-2), \dots, y(t-n_y)\}$, respectively. In this study, the activation functions f^1 and f^2 are the tangent sigmoid and pure linear functions, respectively. Thus equation (4.27) can be re-written as

$$\mathbf{y}(t) = \mathbf{W}^{2,1} \tanh\left(\mathbf{W}^{1,1} \mathbf{p}_{TDL} + \mathbf{W}^{1,3} \mathbf{y}_{TDL} + \mathbf{b}^1\right) + \mathbf{b}^2 \quad (4.28)$$

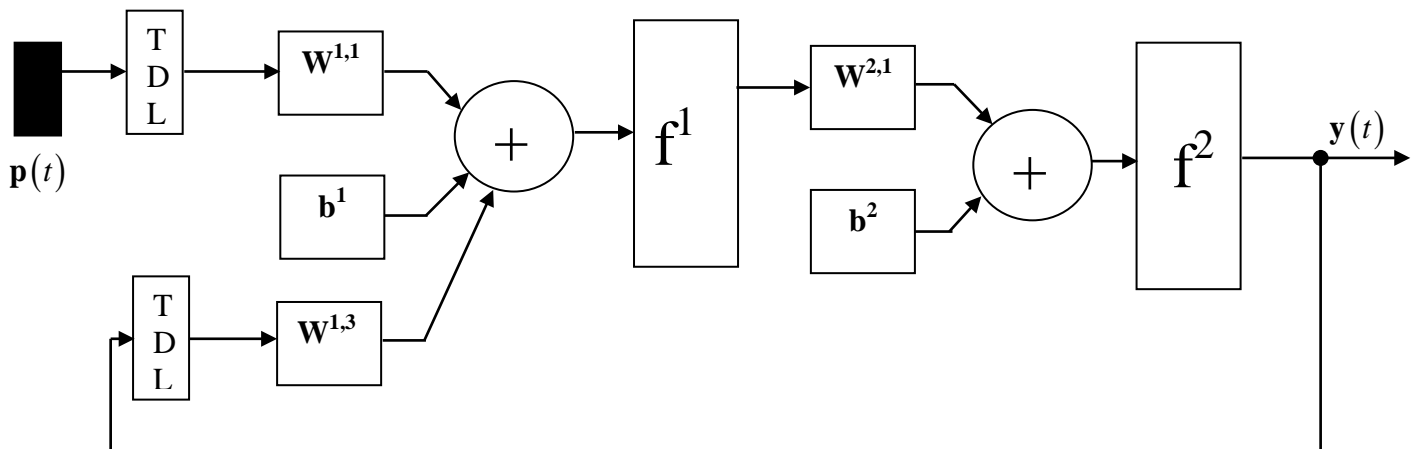


Figure 4.4. NARX architecture

The NARX network is largely used for function approximation and has found many applications in prediction of next values, nonlinear filtering and modelling of nonlinear dynamic systems.

4.3.3 Radial basis function neural networks (RBFNN)

An RBFNN determines the activation of a hidden unit by the distance between the input vector and a prototype vector. This is different from the multi-layer perceptron networks where the hidden units are activated by a non-linear function of the scalar product of the input vector and a weight vector. RBFNN can be trained much faster than MLP networks in two-stages: firstly, the parameters governing the basis functions are determined using relatively fast, unsupervised methods (i.e. methods which use the input data only), and secondly, the final-layer weights are determined in a procedure that requires the solution of a linear problem which is extremely fast (Bishop, 1995).

Radial basis function methods are based on techniques for performing exact interpolation of a set of data points in a multi-dimensional space (Bishop, 1995 citing Powell, 1987). The exact interpolation problem requires every input vector to be mapped exactly onto the corresponding target vector (Bishop, 1995). As proposed by Powell (1987), the radial basis function approach introduces a set of basis functions for each data point which take the form $f(\|\mathbf{x} - \mathbf{x}^n\|)$ where $f(\cdot)$ is some non-linear function. However, the resulting radial basis function yields an undesirable result as far as generalisation is concerned because it provides an interpolation function which passes through every data point. Thus in designing the RBFNN certain modifications are made to the original radial basis function formulation. Broomhead and Lowe (1988) and Moody and Darken (1989) propose a smooth interpolating function in which the number of basis functions is determined by the complexity of the mapping and not by the size of the data set. Bishop (1995) summarises the proposed modifications as follows:

1. The number of basis functions typically much less than the number of data points.

2. The centres of the basis functions are not determined by the input data vectors but their determination becomes part of the training process.
3. Each basis function is assigned its own width parameter σ by the training process. The width parameter is not fixed for all basis functions as in exact interpolation.
4. Bias parameters are included in the linear sum.

4.3.3.1 RBFNN training

The training process of the RBFNN can be understood better with a representation of its architecture shown in Figure 4.6. The RBFNN mapping as represented by this figure can be expressed as:

$$\mathbf{y} = f^2 \left(\mathbf{W}^{2,1} f^1 \left(\left\| \mathbf{p} - \mathbf{W}^{1,1} \right\| \cdot \mathbf{b}^1 \right) + \mathbf{b}^2 \right) \quad (4.29)$$

where the radial basis function is given by $f^1(\cdot) = \exp\left(-\frac{\left\| \mathbf{p} - \mathbf{W}^{1,1} \right\| \cdot \mathbf{b}^1}{2\sigma}\right)$ and the output function is a pure linear function $f^2(x) = x$, thus equation (4.29) can be rewritten as

$$\mathbf{y} = \mathbf{W}^{2,1} \exp\left(-\frac{\left\| \mathbf{p} - \mathbf{W}^{1,1} \right\| \cdot \mathbf{b}^1}{2\sigma}\right) + \mathbf{b}^2 \quad (4.30)$$

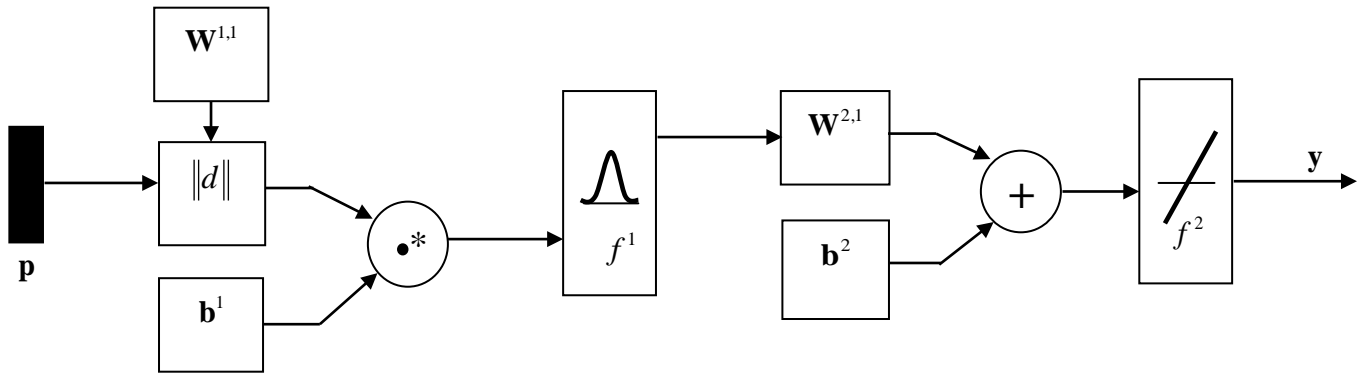


Figure 4.6. Architecture of radial basis function neural network.

If the bias parameters are absorbed into the weights then the above form can be represented in a more compact format as

$$\mathbf{y} = \mathbf{W}^{2,1} \exp\left(-\frac{\left\| \mathbf{p} - \mathbf{W}^{1,1} \right\|}{2\sigma}\right) \quad (4.31)$$

A key aspect of RBFNN is the distinction between the roles of the first and second layers of weights. In the first stage of training, the input data set \mathbf{p} alone is used to determine the weight parameters



$\mathbf{W}^{1,1}$ and width parameters σ using unsupervised training techniques (Bishop, 1995). The basis functions are then kept fixed while the second-layer weights are found in the second phase of training. Thus the second phase of training then involves solving equation (4.31) with the term containing the basis function simplified into a matrix $\boldsymbol{\phi}$ thus (Bishop, 1995)

$$\mathbf{y} = \mathbf{W}^{2,1}\boldsymbol{\phi} \quad (4.32)$$

The weights $\mathbf{W}^{2,1}$ can then be determined through the linear equations that result from optimisation of the weights by minimising the error function between the targets and the network outputs:

$$\frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{W}^{2,1}} = \frac{1}{2}(\mathbf{W}^{2,1}\boldsymbol{\phi} - \mathbf{T})^2$$

$$\frac{\partial \boldsymbol{\varepsilon}}{\partial \mathbf{W}^{2,1}} = (\mathbf{W}^{2,1}\boldsymbol{\phi} - \mathbf{T}) \times \boldsymbol{\phi}, \quad \text{setting this equal to zero, we have}$$

$$\boldsymbol{\phi}^T \boldsymbol{\phi} \mathbf{W}^{2,1} = \boldsymbol{\phi}^T \mathbf{T}$$

$$\mathbf{W}^{2,1(\tau)} = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1} \boldsymbol{\phi}^T \mathbf{T} = \boldsymbol{\Phi}^\dagger \mathbf{T} \quad (4.33)$$

where the superscript T denotes the transpose of the attendant matrix and the bold-type \mathbf{T} is the matrix containing the targets. $\boldsymbol{\Phi}^\dagger$ is termed the pseudo inverse (Bishop, 1995). Since equations (4.33) are linear, their solution is extremely fast. This situation only applies when the output layer neurons are activated by linear functions but when the activation functions are non-linear then the determination of second layer weights is no longer a linear problem, non-linear optimisation of the weights will be required.

4.4 Comparison of network performances

In this Section, the computational performances of the three neural networks in Section 4.3 are compared in terms of training errors and training time. A quarter car model is used to calculate the accelerations while the road profiles are numerically generated in MATLAB. All these neural networks have unsprung mass acceleration as input with the road profile as the output. The performances of these neural networks have been compared by fitting their outputs to five different arbitrarily selected waveforms.

In Figure 4.7, the NARX and FFNN networks automatically stopped after 100 epochs (an epoch is a learning iteration) taking training durations of 3.5 s and 3.8 s respectively. The FFNN yields a slightly larger error than the NARX network and a large proportion of that error is around the bump itself while the NARX network error largely occurs at the beginning due to initial effects. If this portion is neglected, the error can actually be much smaller than 4 % because the NARX network exactly fit the bump. The RBFNN on the other hand has the largest error much of which comes from around the bump itself. The only advantage is that it takes less than one second to train.

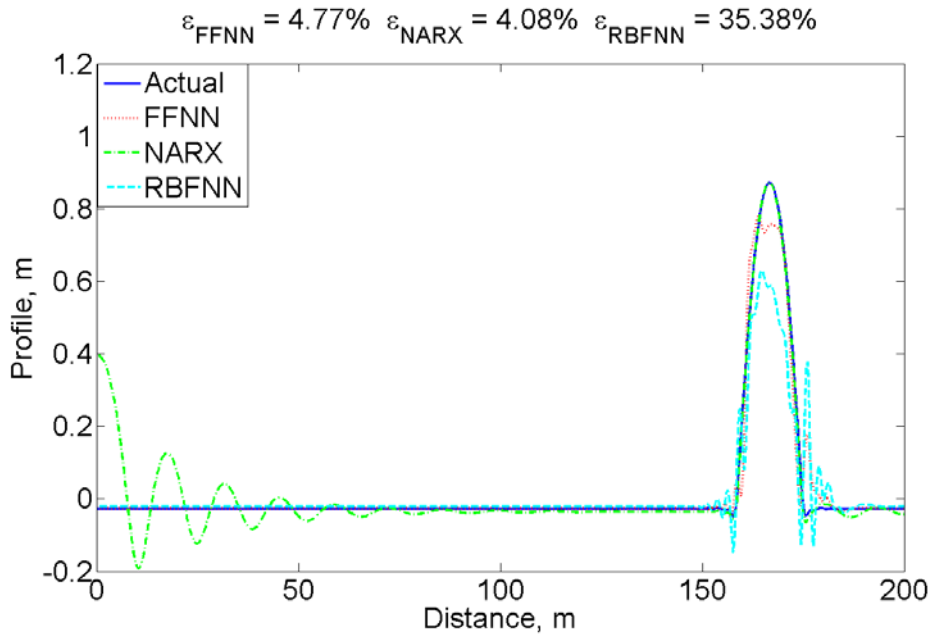


Figure 4.5. Comparing the performances of FFNN, NARX and RBFNN over a half-sine bump reconstruction.

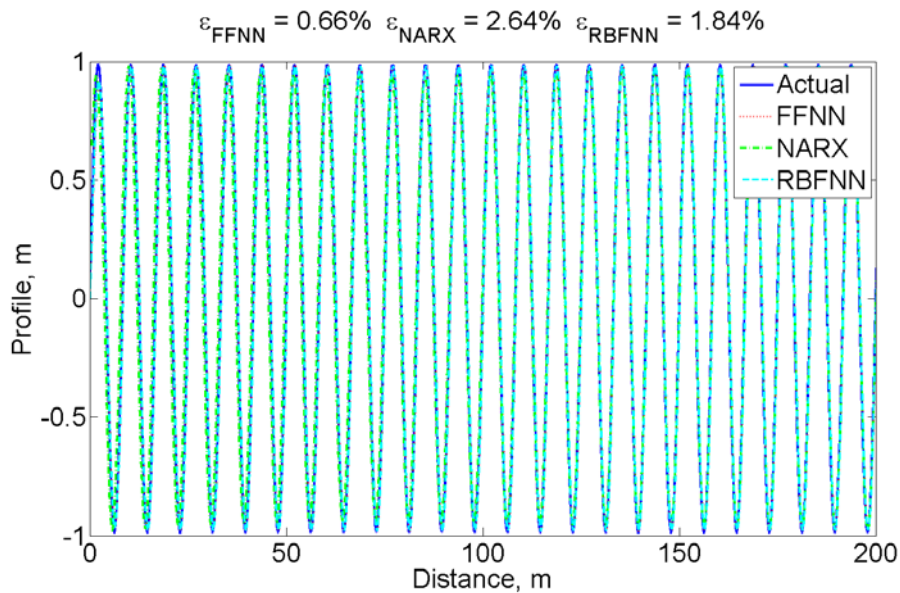


Figure 4.6. Comparing the performances of FFNN, NARX and RBFNN over a harmonic waveform reconstruction.

In Figure 4.8, all the three networks perform very well in reconstructing the harmonic waveform. Smooth transitions and constant rate of change of curvature within the harmonic curves make these waveforms easy to learn by all the neural networks. However, the FFNN has the best performance of the three networks. The minor amplitude attenuation as observed on the RBFNN reconstructed waveform was caused by the use of a large width parameter σ in order to achieve better generalisation over new test data. The error on the NARX network was noted to be due to incorrect



prescription of delays both on the inputs and outputs. FFNN and RBFNN do not delays modelled into their architecture.

The performances of the networks over sine sweep waveform are shown in Figure 4.9. The FFNN shows very good performance especially in learning the changing frequencies at the start of the sweep waveform with some errors resulting from amplitude clipping. The RBFNN yields severely attenuated amplitudes at the start of the waveform but learn the sweeping effect very well. The NARX network performs very poorly at the start with peaks being over-amplified. However, all the reconstructed waveforms settle to very accurate waveforms during steady state. Therefore, the largest proportion of the errors indicated in Figure 4.9 come from the transient effects (in the initial time steps).

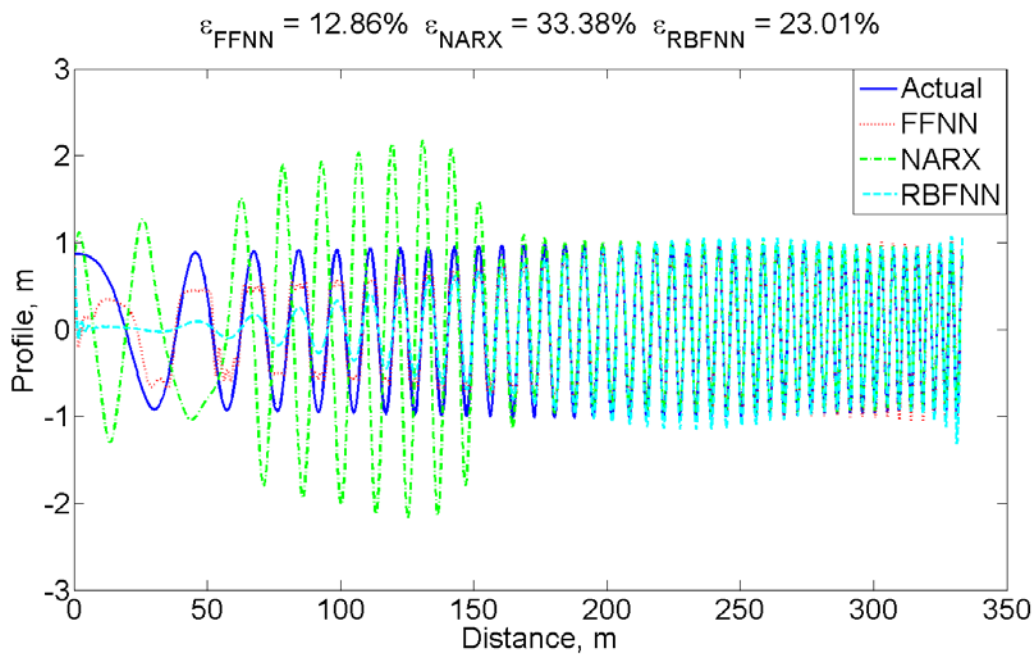


Figure 4.7. Comparing the performances of FFNN, NARX and RBFNN over a sine sweep.

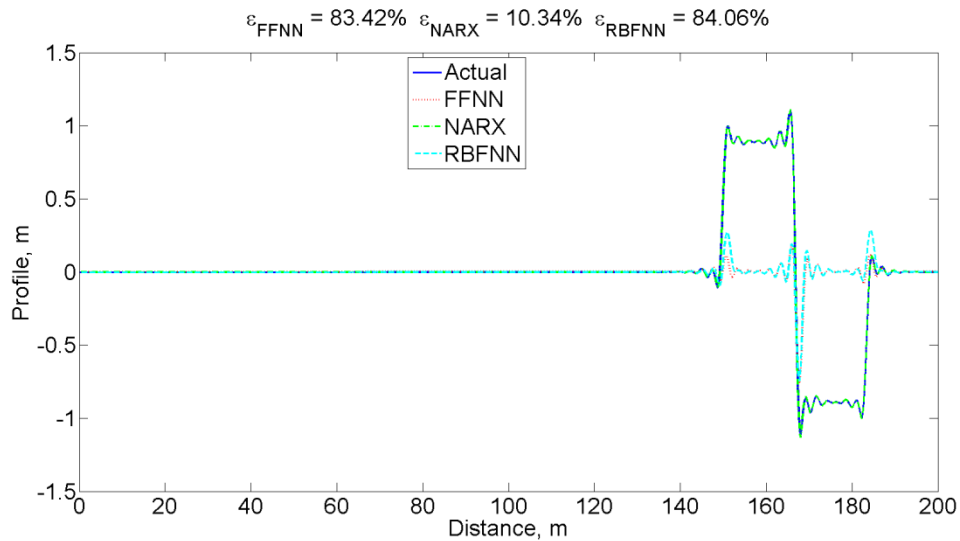


Figure 4.8. Comparing the performances of FFNN, NARX and RBFNN over a square waveform.

The results in Figure 4.10 are slightly surprising results since the errors for the FFNN and RBFNN were not expected to be that high. Only the NARX network produces a correct result while the other networks simply show undulations at the corners but remain flat for the most part of the waveform. Multiple training sessions with different training parameters did not improve the results. A similar result was observed when the square waveform was extended to a random square waveform as shown in Figure 4.11.

Despite these deficiencies, both the FFNN and RBFNN showed a greater sense of stability during training and would almost always achieve the same results on successive trials as compared to the NARX that often exhibited high degree of randomness in converging to an optimal solution. It is however noted that when the NARX converges to an optimal solution, its solution is the most accurate of the three. RBFNN performed very nearly like FFNN except its advantages on speed of training. In this study therefore, the NARX network has been preferred with only one case study where the FFNN was applied.

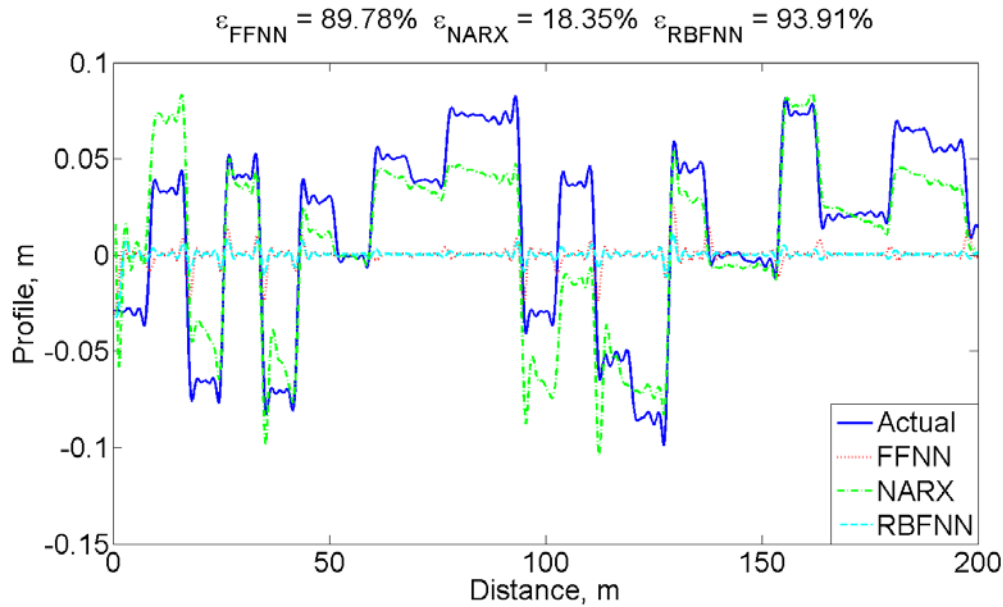


Figure 4.9. Comparing the performances of FFNN, NARX and RBFNN over a random-square waveform.

4.5 Model validation

It is important to know if the model has successfully captured the system dynamics so that it will provide good predictions of the system output for different input excitations, or if it has simply memorized the fit to the data for which it was trained. In this work, two methods of checking model's validity are employed: error in model's predicted output and correlation tests.

4.5.1 Model's root mean square error (RMSE)

The estimated outputs are compared with the measured outputs with good agreement being a necessary condition for accepting the model. Their amplitudes are compared using the RMSE which is given by

$$RMSE = \sqrt{\frac{\sum_{i=1}^N t_i^2 - \sum_{i=1}^N y_i^2}{2N}} \quad (4.34)$$

where t is the target and y is the network output. RMSE errors less than 30 per cent of the target RMS value can be considered accurate enough for this sort of work especially when dealing with data from real vehicle testing.



4.5.2 Correlation tests

The correlation function ϕ_{yt} between two sequences of data y and t is defined as

$$\phi_{yt} = E(y_i t_{i+k}) = \frac{1}{N-k} \sum_{i=1}^{N-k} y_i t_{i+k} \quad (4.35)$$

Correlation values above 0.65 can be practically satisfactory.

4.6 Conclusions

The theory governing the modelling of ANN applicable to function approximation has been presented. Three common neural networks used in function approximation have been compared and the results show that the NARX network outperforms the FFNN and RBFNN networks when operating at its optimal solution. However, the NARX network has exhibited instabilities in its performance when operating away from its optimal solution. Compared to FFNN and RBFNN, the NARX network also seems to suffer from convergence problems. In those cases and if an optimal solution cannot be obtained it is advisable to use FFNN or the RBFNN which demonstrate higher levels of stability despite the loss in accuracy in the fitted curve.

Chapter 5

Development of a generic methodology

This Chapter presents the methodology in terms of all the stages from data acquisition to road roughness classification. The Chapter also presents brief descriptions of the case studies on which the methodology has been demonstrated.

This profile reconstruction methodology is based on a technique developed for mine haul road maintenance by Hugo et al. (2008). However, this methodology eliminates the requirements for conducting onerous system characterisation procedure and the need for deconvolution of system's ride dynamics by simply training a neural network to learn the vehicle's input-output relationship with the intention that it can later be used to approximate the road profile upon being simulated with vertical vehicle accelerations. As stated in Chapter 4, this is a system identification process where the inverse model is derived from reversing the roles between the vehicle's actual inputs and outputs as shown in Figure 5.1.

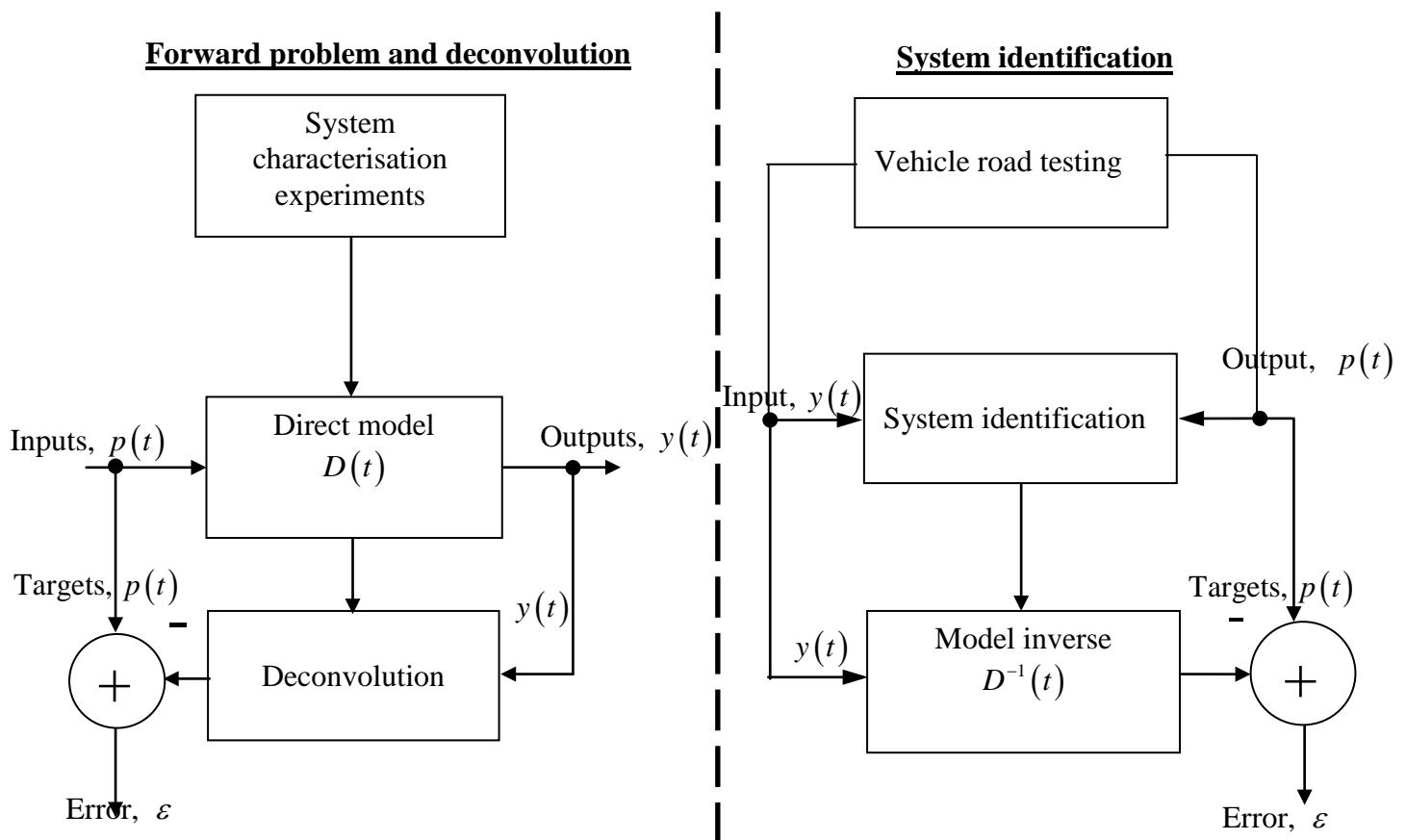


Figure 5.1. Illustration of formulation of system identification as compared to forward problem.



In effect, the analyst is capable of deriving the inverse without worrying much about the theoretical rigours of the underlying model and non-linear dynamics in the vehicle system. A particular advantage of the network is that it can model system non-linearity very accurately. In this study, vehicle body or axle accelerations are used as inputs to the neural network while the road profile elevations are used as the network targets. It is however, recommended that, wherever it is possible, the vertical accelerations should be from the axle. There are however a number of problems that are encountered in practice with this recommendation:

- The axle sub-assembly design may not accommodate any mounting of accelerometers due to space limitations,
- Instrumenting the axle may compromise its structural integrity due to a need for machining (i.e. grinding and drilling) the surface where the accelerometer can be mounted. The car owners cannot be in favour of this, and
- The harsh vibration environment on the axles may result in the loss and damage of the sensors and attached cables.

Figure 5.2 presents the general outline of the activities that are involved in carrying out the proposed road profile reconstruction methodology. It comprises the following five main parts: data acquisition, data preparation, ANN identification, ANN simulation and road condition identification. Comprehensive discussions on each part are presented in the Sections that follow.

5.1 Data acquisition

There are a number of ways in which data can be generated: use of numerical simulations, through laboratory prototypes, laboratory testing of real vehicle systems using synthetically generated data, and by measurement in actual vehicle road testing. In this study, all data are generated only through numerical simulations and field testing. Irrespective of the process used in acquiring the data, the desire for integrity in the acquired data places specific demands on any process used in generating the data. In the vehicle ride dynamics, for example, the frequency range between 0 and 25 Hz is very important and should sufficiently be covered by the acquired data; important features on the road should be clearly indicated in the acquired data; vehicle speeds and other operating conditions should be properly and accurately documented during data capturing process; tyre and suspension conditions should be checked and maintained at prescribed states; and the vehicle's transient conditions should be closely monitored.

There is a general demand for rigorous record keeping during data acquisition. Table 5.1 shows some of the information that should be recorded together with the measured data. This list is not exhaustive - one may include more documentation of the test. Since ANNs are black-box (or blind-models) careful record keeping and a good inventory of the proceedings throughout the testing period can assist in interpreting the results more accurately. It is also very helpful to capture data while the vehicle is in brake-stop mode so that it can be used later as a control during analysis.

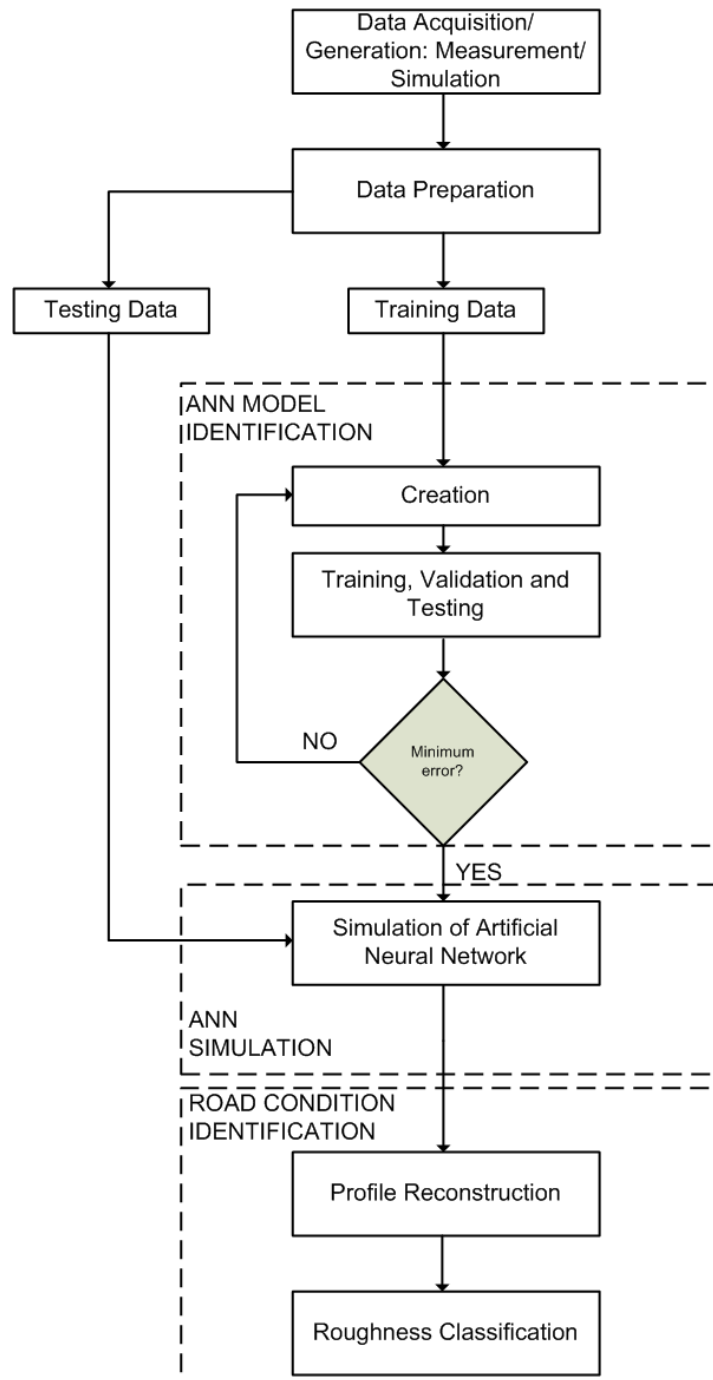


Figure 5.2. General outline of the methodology for road condition monitoring using ANN simulation.



Table 5.2. Vehicle information and sensor positions

	Measurement parameter	Description
1	Vehicle Designation	e.g. ToyotaCamry_2005_vvti
2	Vehicle Condition	e.g. Fully-Serviced 3 months ago
3	Tyre and wheel condition (including type)	
4	Tyre pressure	
5	Suspension condition	
6	GVM	
7	Wheelbase	
8	Track width	
9	Accelerometers	
10	Positions of accelerometers from Reference point, preferably located in driver's cabin as shown in Figure 5.3 below.	
11	Position of GPS from Reference point	
12	Position of Infra-red position sensor from Reference point	

For numerical experiments, data should be generated using accurate and validated models representing the important dynamic characteristics of the vehicle under investigation. Full-car models may be recommended especially where the results of the analysis are to influence vehicle and/or road management decision. The accuracy of the tyre model is also very crucial and issues such as tyre envelopment, wheelbase filtering, roll and pitch dynamics, and suspension friction do affect the final results (Hugo, 2005).

5.1.1 Vehicle preparation

Figure 5.3 shows the key sensors and their locations on the generic vehicle system. Whereas it is very easy to calculate accelerations for all the indicated locations for a numerical simulation experiment, actual vehicle tests may impose some challenges on one or two of the locations, often limiting the measurements to be carried out only at a few points, or even worse, at positions on the vehicle body

itself. This may make the resulting ANN model lose some important input-output mapping information. However, it is very important that the available locations must be maintained throughout the testing procedure for consistency. Any devices that can assist in aligning the captured data later on during analysis are strongly recommended. This may include use of position sensors for example that can mark the beginning and end of an important road feature. Although this information may not be used in the actual neural network, it is very useful during pre-processing of data.

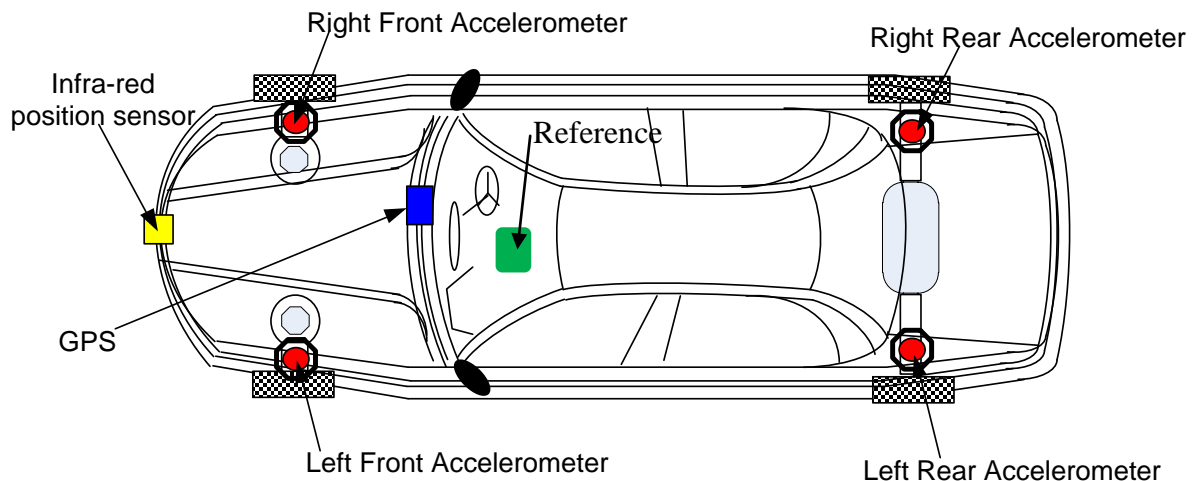


Figure 5.10. Typical data pick-up points.

5.1.2 Data acquisition system

In the numerical experiment the data are directly stored in files on the computer in any format which is compatible with the software to be used for analysis. In field tests, it is preferred to have a data logger, a computer system and other portable storage devices. In this study, a SoMat eDAQ-lite data acquisition system has been used as a data logger. The system has a leading-edge signal conditioning and a capacity to perform a broad range of on-board data processing, triggering, intelligent data storage and complex computations (HBM, Online). Its intelligent data storage enables it to collect relevant measurement data in condensed and manageable form. A laptop was used to set up and monitor the data acquisition process through the SoMat InField software.

5.1.3 Sensors

In this study, Crossbow 4- and 6-g/V tri-axial accelerometers are used due to their good response definition at low frequencies. Important vehicle ride vibrations are said to occur below 25 Hz. A GPS sensor is also extremely useful for vehicle's location around the tracks. The vehicle's speed is also measured using either the shaft encoder on the drive shaft or calculated from the measured time on the the eDAQ-lite system and the measured distances on the road. The infra-red position sensor for control of measurement start-up and finish points is also used in this study. This is necessary for alignment of acceleration data with road profile data.

5.1.4 Road preparation and profiling

Figure 5.4 shows an example of how a road may be prepared for measurement over a 100 m Section only. The yellow reflective tapes are used as light reflectors back onto the infra-red position sensor as shown in Figure 5.3.

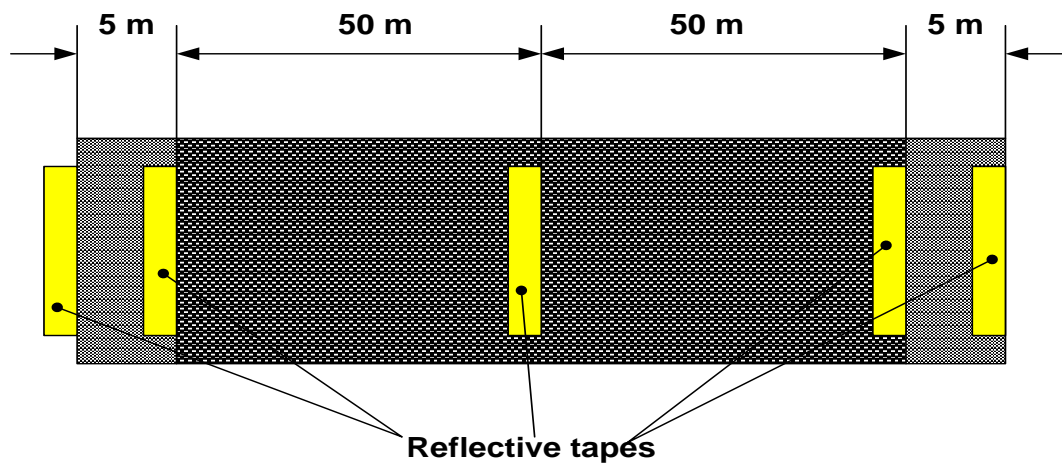


Figure 5.4. Example of road preparation for measurement over 100 m length.

Road profiles are measured using any available profiling method. The following information is very crucial whichever apparatus is used:

1. A reference elevation,
2. A height relative to the reference, and
3. Longitudinal distance.

5.1.5 Test procedure

The following guidelines can be followed during testing:

1. With a properly instrumented vehicle, as shown in Figure 5.2, drive over the prepared road Section starting from a distance of about 10 m away from the position of the first reflective tape. Start the measurement 2-5 m away (with respect to the front wheels) from the first reflective tape and stop it about 2-5 m (with respect to the rear wheels) after the last reflective tape.
2. Carry out a minimum of three tests in alternating runs according to the different speeds at which the test are carried out (for example if the vehicle is tested at 20 km/h and 50 km/h, ensure that the tests are alternated between these two speeds). This is done to eliminate any systematic errors.



5.2 Data preparation

It is hard to have freshly acquired data in the exact format that is amenable for neural network training. Therefore, data preparation involves a number of processes that help to get data ready for neural network training. Some of these processes are already mentioned in Chapter 4 (Section 4.4), but they are discussed here in some detail. Since the final desired output from any neural network training process is a set of matrices of input and layer weights and biases with characteristically similar outputs as a physical numerical model of the system, the training data should ideally contain all important information that represent every important characteristic of the modelled system.

In the following Sections various data preparation procedures are presented. At the same time, all this is limited by the computational aspects during the training procedure. At its core, the neural network training procedure, as applied to function approximation (regression) problems, involves minimizing a sum-of-squares error function between the target and the simulated output with respect to the network's weight and bias parameters.

5.2.1 Data alignment

This is relevant to situations where the neural network inputs and targets were acquired using two different systems and there were no automatic triggers that would align the two sets of data during the acquisition procedure itself. This was the case in this study. The vehicle accelerations (network inputs) were acquired through the eDAQ-lite system, while the road profiles (network targets) were variously measured by poles, ropes and rules and the 'can-can' machine. Thus before applying them to neural network training, they had to be aligned perfectly so that they were representative of the actual system's time response characteristics.

Yellow reflectors pasted on the road surface, as shown in Figure 5.4, were used as an aid to data alignment. The instance of the yellow reflectors was marked by a spiky peak of the infra-red position sensors' voltage. Thus it was fairly easy to match these positions against the actual distances as calculated from the time and vehicle speed. The spacing between the laser position sensor and the accelerometer should be taken into account.

5.2.2 Removing trends

This might not be absolutely necessary in all cases but it is fairly useful in this study. The road profile elevations are calculated relative to each other and due to the highly approximate nature of the measurement procedure, fairly smooth road Sections were taken to be zero millimetres high. However, the accelerometers would still record some bias plus oscillations at those Sections due to mixed influences of higher frequency road undulations and on-board sources. The bias would cause



some problems in the computation of the derivatives by the back-propagation routine during neural network training. This is achieved by removal of any trends from the acceleration data.

5.2.3 Filtering

This is mainly performed on the acceleration data. In this study, the acceleration data were captured by the eDAQ-lite at much higher sampling frequencies (mostly at 1 kHz). Since important ride vehicle dynamics occur below 25 Hz, it might be prudent to apply a low-pass filter with a cut off at 50 Hz. A zero-phase forward and reverse digital filtering routine was applied in this study. After filtering in the forward direction, the filtered sequence is then reversed and run back through the filter; thus the final filter output is the time reverse of the output of the second filtering operation. The result has precisely zero phase distortion and magnitude modified by the square of the filter's magnitude response (Mathworks Inc., 2007). Care is taken to minimize start-up and ending transients by tapering the data before applied to the filters.

5.2.4 Data normalization

The nonlinear sigmoid activation function is designed to span the range ± 1.7159 , thus the target values can be processed so that their values fall in the range ± 1.0 (Haykin, 2009). If the values are allowed to exceed these limits, then the neural network is driven into saturation.

5.2.5 Decorrelation

If any training data are highly correlated there is high likelihood that the training process will not converge to weights that yield the targets because the problem has a very large number of possible solutions. Haykin (2009) suggests that this can be dealt with by removing the correlation from the data so that they become uncorrelated. This can be achieved by performing match filtering to the data or carrying out a principal component analysis. MATLAB uses `processpca.m` to perform this function. Another method that has been used in this study is to simply ignore all data that correlates with any of the other data that are currently being used for training the network. The advantage is that it is very easy to implement and not as computationally demanding as the decorrelation procedures described earlier. The only disadvantage is that it does not deal with autocorrelation.

5.2.6 Covariance equalisation

If any sets of training data are not varying together, then a situation arises during training where respective weights learn at different speeds. This may result in network outputs that are either persistently sub-optimal when network generalisation is employed or over-fitting the training data which further result in much longer training durations. Haykin (2009) recommends that the training



data's covariances be equalised in order to eliminate the problem. In this study, the data have been rescaled by normalizing the mean and standard deviation of the training set. The MATLAB function `mapstd.m` can be used to do that since it normalizes the inputs and targets so that they will have zero mean and unity standard deviation. However that will change the range of values of network outputs and can only be compared with the real road profile heights after transformation into the original range.

5.3 ANN identification

In Chapter 4 (Section 4.4), the performances of three different commonly-used neural network models in function approximation are presented and compared. NARX and FFNN networks are chosen for further application to this study due to their accuracies.

The identification of an artificial neural network fit for a given task goes further than mere creation of a neural network model. The number of different neural network models that can be created for any given problem is very large. One can only choose a particular network model depending on the nature of the problem, computational demands and largely past experience or *a priori* knowledge and/or recommendations by more experienced ANN users. In this study, the choice was dictated by the nature of the problem and recommendations by others provided through literature study.

Vehicle-road interaction is both dynamic and nonlinear. Hence a dynamic and nonlinear neural network was deemed necessary. Besides, based on a number of preliminary numerical results presented in the Chapter 4, a feed-forward dynamic neural network FFNN and a nonlinear autoregressive network with exogenous inputs (NARX) neural network are adopted in this study. However, the NARX network has been used predominantly due to its high flexibility and adaptability.

As shown in Figure 5.2, this phase has two main parts: network creation and neural network training (which includes validation and testing). The following Sections discuss each of these stages.

5.3.1 Creation of neural network

In this study, the neural networks are created using MATLAB (R2007b) Neural Networks Toolbox (Mathworks Inc., 2007). The neural networks toolbox in MATLAB contains a number of well-developed functions for creation, training and simulation of neural networks. In order to create the NARX network, the function called `newnarxsp.m` is used in MATLAB. This function creates a NARX network with a series parallel architecture, where the target data are presented to the training algorithm and form part of the inputs and the training process is executed in a normal feed-forward manner without feeding back the network outputs. The function requires the following arguments: inputs, targets, input and output delays, sizes of hidden layers, layer activation functions, training algorithm and other training parameters.

If one wishes to prescribe other training parameters than default, one simply has to add a few more lines of code below the `newnarxsp` function call line. For example, if one wishes to define a new training goal ε_{new} , the following line can be added:

```
net.trainParam.goal =  $\varepsilon_{new}$ 
```

However the created network expects the availability of targets even during network simulation, which is not the case in real life simulations. In order to address this problem, the network is further transformed into a parallel network which feeds back its own output according to the prescribed output delays. Figure 5.5 illustrates this procedure.

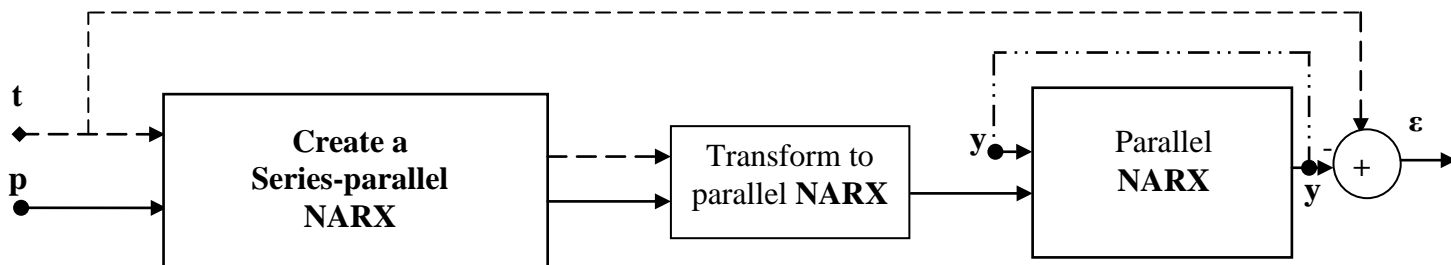


Figure 5.5. Parallel NARX network creation process.

5.3.2 Network training

During the creation of the artificial neural network, the weight and bias matrices are often populated with initial values generated by default function in `initlay.m`. One may choose to reinitialise the network with the function: `net = init(net)`.

The training procedure is performed upon invoking a command `train.m` as shown in the following syntax:

```
net = train(net, P, T, Pi, Ti)
```

where `Pi` and `Ti` are the matrices containing previous input and target values, respectively, according to the prescribed number of input and output delays. Besides, the quality of pre-processing of the training data, the manner in which the input data `P` is structured and presented to the network also influences the training procedure. If `P` is presented as normal matrix, the network trains the network in batch mode while presentation in a cell array, inputs are treated as sequential and training proceeds sequentially and the order of the inputs is very important. In addition, if the training data are presented in distinct groups such as, $[P, T] = [\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_k, t_k\}, \dots, \{p_N, t_N\}]$, the training algorithm uses the first pair for training the network, the second pair or any other randomly picked pair may be used to validate how well the network generalizes and another pair for testing the network.



The network continues to adjust the weights and biases as long as errors on both the validation and test data do not meet the desired goal. This process implements generalisation by cross-validation and is evident by three training performance curves during the training process. As already pointed out in Chapter 4, this automatically avoids the problem of overfitting, which plagues many optimization and learning algorithms (Mathworks Inc., 2007). One can monitor how well the error function is reduced on all these three curves. If unsatisfactory performance is displayed, the network might be adjusted by modifying the training parameters or reinitialising it. This can proceed until satisfactory performance is achieved. It is very hard to achieve a good performing neural network model at first attempt and one has to keep modifying the network and its training parameters to arrive at a good enough model.

5.4 ANN simulation

If the performance function has been minimized to desired level, the neural network can be stored and used later for simulation purposes on freshly acquired inputs. The output from this stage provides input for the final stage of road condition identification.

The good thing about the simulation process is that it is not as computationally demanding as the training process since it merely involves multiplication of matrices. Thus it is possible to implement this model in an online road condition monitoring system. Since most of the online programs are written in lower programming languages, such as c-program, the simulation process can be carried out at very high computational speed. In this work, this was done by a consultant and cannot be included in this thesis for proprietary reasons.

5.5 Road condition identification

Road condition identification per se is a very broad term and covers a wide spectrum of road condition criteria. In this study, road condition is limited to condition of longitudinal road surface profile i.e. its roughness in the longitudinal sense. There are two ways in which that has been done in this study: firstly in the form of raw profiles and secondly in the form of displacement spectral densities as shown in Chapter 3.



5.5.1 Reconstructed road profiles

The ANN simulated road profiles from Chapter 3 (Section 3.4) are plotted against distance along the road surface. This is displayed over the GPS coordinates data and gives indications of what is actually taking place over specific road Sections on a road map. Some roughness limits in terms of profile elevation RMS values can be benchmarked and if they are exceeded, appropriate triggers may be accordingly sounded to carry out a specified maintenance action.

The use of raw profiles is found to be very suitable for discrete obstacle type of road degradation such as pot-holes, bumps, stones etc. The degradation in the form of random roughness over the road surface may require a further step described in Section 5.5.2.

5.5.2 Roughness classification

This stage involves the calculation of road displacement spectral densities. The resulting DSDs are subsequently overlaid on ISO DSD scale (ISO 8608: 1995E, 1995) to give indication of road surface roughness level. This is found to be quite useful for random rough terrains. More detailed determination of the DSDs is presented in Chapter 4.

5.6 Methodology verification via case studies

The proposed methodology has been verified through three case studies involving: numerically-generated data, real vehicle with a full-car model driven over accurately profiled roads and real mine vehicles tested in their normal operating environments.

5.6.1 First case study: Numerically-generated data

The case study numerically demonstrates the capabilities of the proposed methodology in the presence of noise, changing vehicle mass, changing vehicle speeds and road defects. In order to avoid crowding out understanding of the methodology, a simple linear pitch plane model is employed. Initially, road profiles from known roughness classes were applied to a physical model to calculate vehicle responses. The calculated responses and road profiles are used to train an ANN. In this way, the network renders corresponding road profiles on the availability of fresh data on model responses.



5.6.2 Second case study: Experimental Land Rover

This case study reports the findings on applying the proposed methodology on measured data obtained from a Land Rover Defender 110 which was operated over well-constructed and accurately-profiled discrete obstacles and Belgian paving at a vehicle testing facility (Gerotek) in Pretoria, South Africa. In Section 5.6.1, it was found that the methodology was able to reconstruct road profiles and their associated defects within very good levels of fitting accuracy and correlation using data calculated from a numerical model. A NARX network was trained in a series-parallel framework. When compared to the parallel framework, the series-parallel framework offered the advantage of fast training but had a shortcoming in that it required feed-forward of true road profiles. In this study, the true profiles are not available and the test data are obtained from field measurements. Training data are numerically generated by making minor adjustments to the real measured profiles and applying them to a full vehicle model of the Land Rover. This is done to avoid using the same road profile and acceleration data for training and testing or validating the neural network. A static feed-forward neural network (FFNN) is trained and consequently tested on the real measured data. The results show very good correlations over both the discrete obstacles and the Belgian paving. The random nature of the Belgian paving necessitated correlations to be made using their DSDs as well as evaluations of RMS error percent values of the raw road profiles. The use of DSDs is considered to be of much more practical value than the road profiles since they can easily be interpreted into road roughness measures by plotting them over an internationally recognized standard roughness scale.

5.6.3 Third case study: Field tests on mine vehicles

The proposed methodology is applied to two different vehicles in their normal operating environments at two mining sites. An ultra-heavy haul truck used for hauling operations in surface mining and a small utility underground mine vehicle were utilised in the current investigation. Unlike previous studies where numerical models were available and road surfaces were accurately profiled with profilometers, in this study, that was not the case in order to replicate the real mine road management situation. The results show that the methodology performed very well in reconstructing discrete faults such as bumps, depressions or potholes but, owing to the inevitable randomness of the testing conditions, these conditions could not fit the fine undulations present on the arbitrary random rough surface. These are better represented by the DSDs of the road surfaces. Accordingly, it is recommended that the proposed methodology can be applied to road condition identification in two ways: firstly, employ raw road profile matching to detect, locate and quantify any existing discrete road faults/features and secondly, utilise DSD plots to classify their general roughness level.



5.7 Conclusions

The Chapter has outlined the main stages in the development of the methodology. General guidelines on how to execute each stage in the methodology are stipulated. Each of the case studies that follow have adhered to this methodology. There are application specific differences with respect to data acquisition and preparation among the case studies, but the fundamental steps remain similar. Its adaptability to different applications without drastic changes to the fundamental stages makes the methodology attractive and programmable for use on many different kinds of terrain and vehicle combinations.



Chapter 6

Numerical case study

6.1 Introduction

This Chapter demonstrates the developed methodology using assessments performed on a linear pitch-plane vehicle model simulated with synthetically-generated road profiles. González, O'Brien, Li and Cashell (2008) argue that despite major improvements in recent years in the quality of road-profiling equipment, these devices remain generally expensive to purchase, their use with time is inefficient and their operation specialized. They demonstrate numerically the applicability of vehicle acceleration measurements for classification of road roughness. This methodology may provide an initial and inexpensive assessment of the uneven condition of the road surface.

Thompson et al. (2003) realized that there was a need for the ability to analyse, recognize and interpret various forms of the same defect signature and also a need to apply multi-sensor data fusion to isolate faults from different sources. Kang, et al. (2007) used ANNs merely for classifying the severity of faults on unpaved roads by means of DSDs but did not apply the ANNs to profile reconstruction and defect recognition.

This Chapter demonstrates that it is possible to use ANNs to reconstruct specific road defects and classify overall road damage within bounds of very reasonable accuracy by using the accelerations measured on the vehicle. This, as previously stated, eliminates the requirement for the extensive system characterisation used by Hugo et al. (2008).

The case study includes evaluations of the widely used road roughness indicator, the IRI, merely for correlation purposes. The IRI is an intricate scale represented by statistical values which describe overall roughness levels ranging from airport runways to rough unpaved roads. Therefore the indicator may be useful where the engineer is primarily interested in interpreting and comparing the overall properties of road surface roughness (in the road's longitudinal direction), but might be a poor parameter when analysing the effects that depend on the road surface as experienced by the road user (Andren, 2006). For example, IRI may not indicate the status of roughness over different ranges of wave numbers.

The Chapter is organised as follows. Section 6.2 describes how the proposed methodology has been applied to this particular investigation. Relevant theoretical concepts are presented in Section 6.3. The Section details how vehicle modelling, generation and classification of road profiles and formulation



of Bayesian-regularised NARX network have been done. Section 6.4 describes the dynamic behaviour of the vehicle under investigation. The findings on different case scenarios are presented in Section 6.5 Finally, the Chapter is concluded in Section 6.6.

6.2 Customisation of the proposed methodology to this case study

The methodology utilises procedures quite similar to normal system identification techniques as presented by Ljung, 1999. However, in this work, neural networks are used which have unique advantages, in that they can handle non-linearity problems that are difficult for most of the other inverse models calculated from direct linear models. In this study, sprung mass acceleration \ddot{z}_s , front axle acceleration \ddot{z}_{uf} and driver's upper torso acceleration \ddot{z}_1 , were variously used as network inputs while the road profile $z_r(t) = z_{rf}$ was the network target.

Figure 6.1 presents a variation of the proposed methodology as applied to the present case study. It comprises all the four main parts of the proposed methodology as given in Chapter 5 with more detail on the type of inputs from one stage to another. The following discussion covers each stage with reference to Figure 6.1 and some mathematical expressions to follow in Section 6.3.

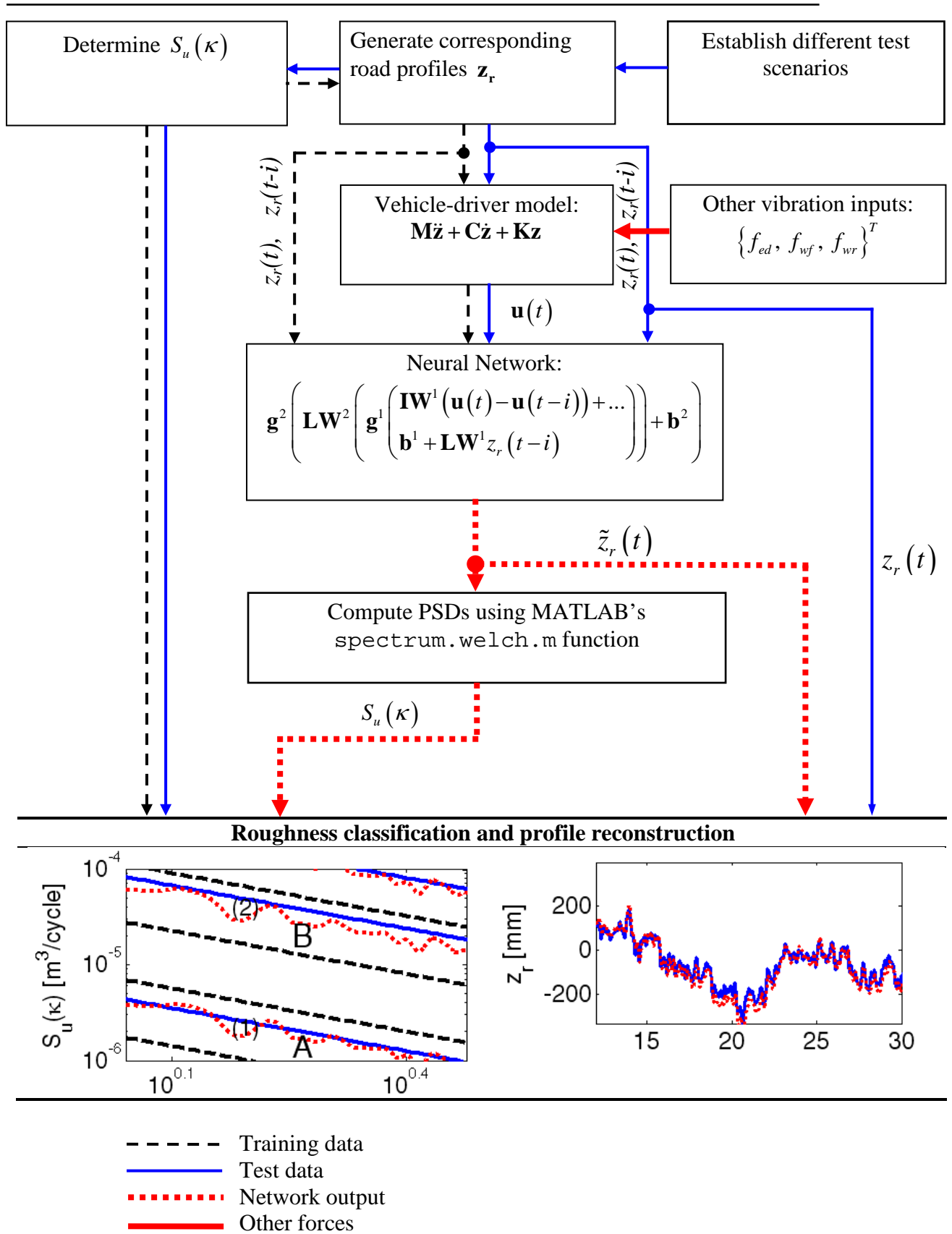


Figure 6.1. Identification procedure employed in the numerical verification.

The ANN training data are obtained from the simulations of LPP vehicle model of Figure 6.2 and synthetically-generated road profiles. The random roughness in the generated road profiles is modelled by equation (3.18) in Chapter 3 while discrete faults are modelled by half-sine waveforms. In this study, the road profile data used for ANN training are generated from DSD classes A to H (Cebon, 1999; ISO 8608: 1995E, 1995). The two-line approach in Cebon (1999) has been used because the generated profiles cover spatial frequencies higher than the validity limit of 2.83 cycles/m for the straight-line approach (Andren, 2006; ISO 8608: 1995E, 1995). In this study, a committee of neural networks was trained for each class using the data generated from its lower and upper bound displacement PSDs, so that a particular network structure from the committee was realized for each of the eight ISO PSD classes. Therefore, during simulation, the algorithm searched for a network structure that yielded the minimum mean square error value to be used for reconstructing the road profile. The ISO PSD class to which the network structure belonged, gave the roughness class of the road under test.

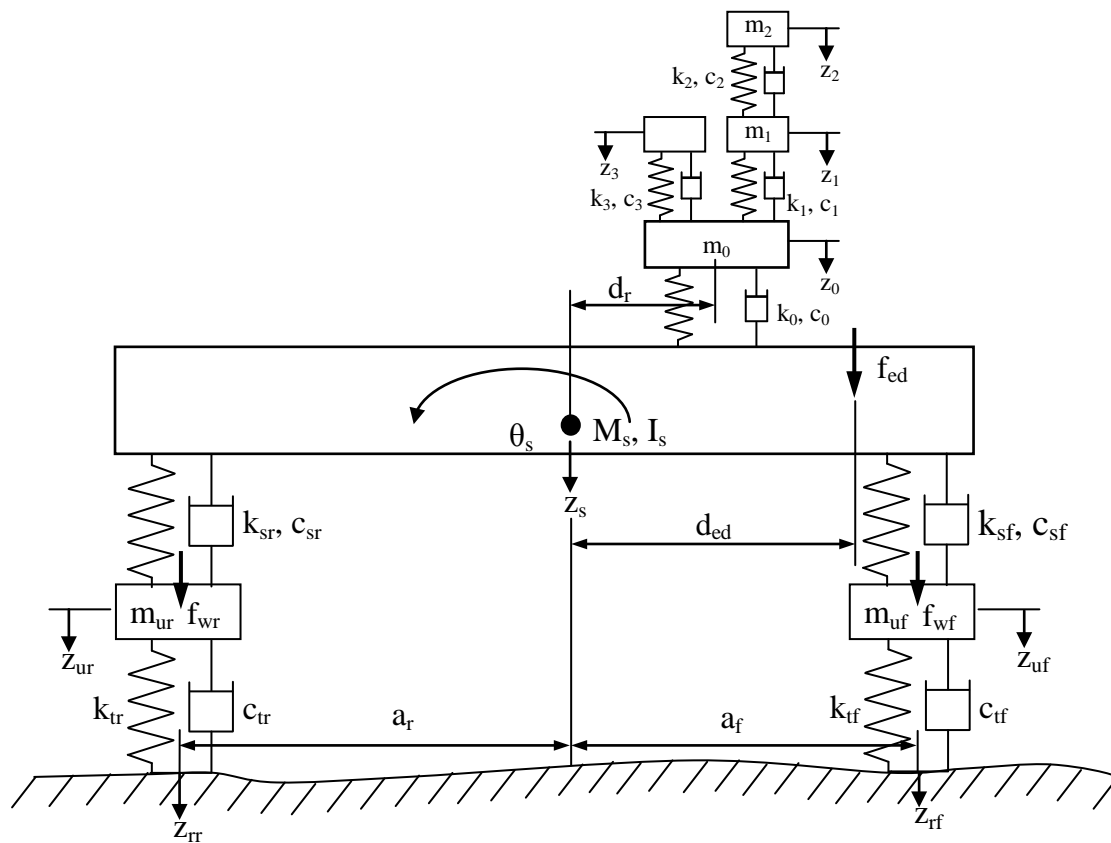
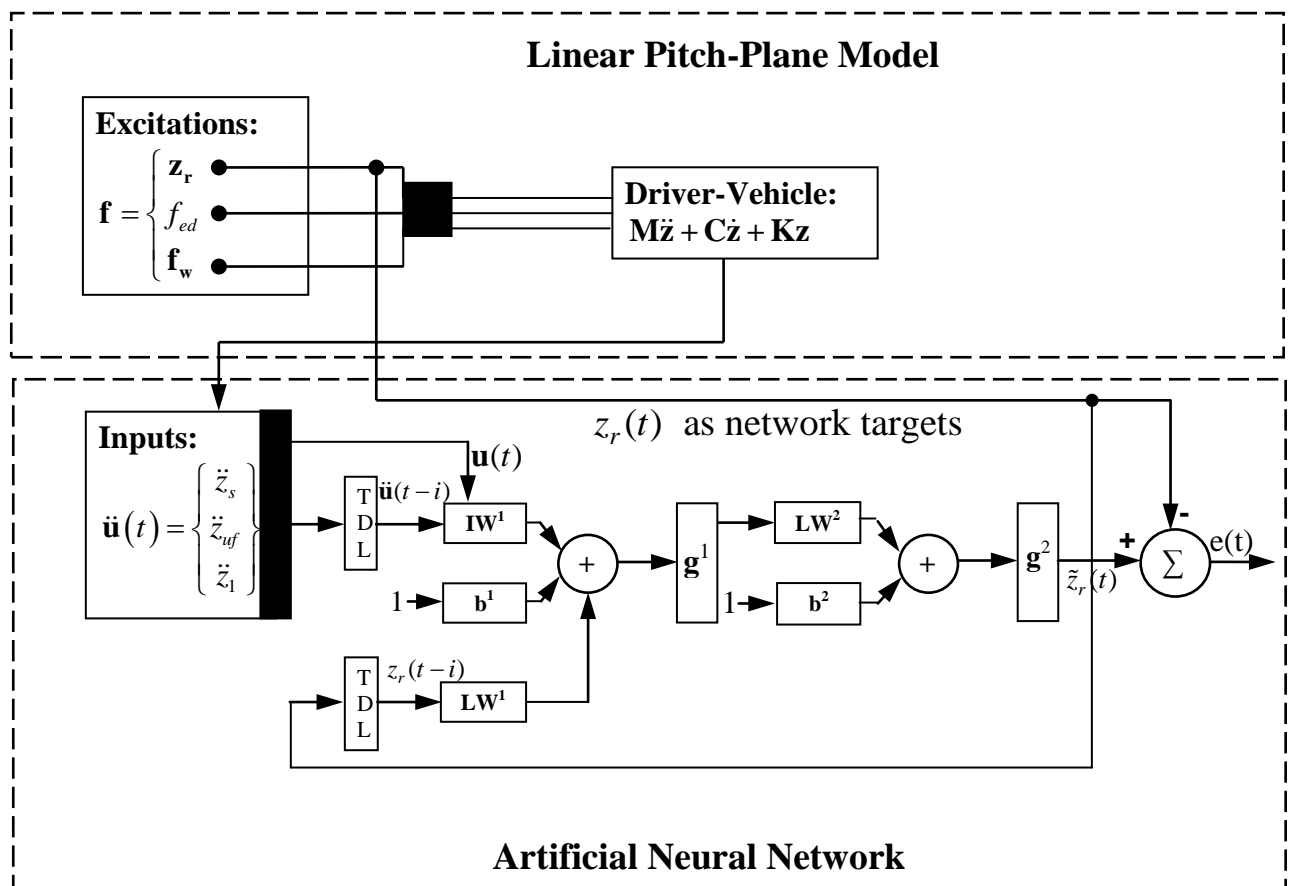


Figure 6.2. A linear pitch-plane (LPP) vehicle model with a seated-driver model.

In order to account for vehicle vibrations due to extraneous sources, engine/driveline forces f_{ed} and tyre/wheel assembly non-uniformity forces $\mathbf{f}_w = \{f_{wf}, f_{wr}\}^T$ are also applied accordingly as random noise to the LPP model.

The neural network was created using MATLAB's `newnarxsp.m` in the Neural Networks Toolbox (Mathworks Inc., 2007). The network target $z_r(t)$ was fed back via tapped delay lines to form part of network inputs $\mathbf{u}(t) = \{\ddot{z}_s, \ddot{z}_{uf}, \ddot{z}_1\}^T$ as shown in Figure 6.3. The time lags between network inputs and outputs were modelled by allowing for 15 delays on both $\mathbf{u}(t)$ and $z_r(t)$. The network was trained by the `trainbr.m` algorithm discussed in Section 6.3.2.



b: biases; **IW:** input weights; **LW:** layer weights and **TDL:** time delay lines

Figure 6.3. A series-parallel architecture of the NARX model used for road profile reconstruction.

Different DSDs drawn from the roughness classes that were represented in the test were generated and their corresponding road profiles calculated. These were used for validating and testing the network. The network was subsequently simulated and its output $\tilde{z}_r(t)$, was plotted over the actual profile $z_r(t)$, to validate the network. Then road defects were created by adding sinusoidal functions to the actual profiles. The defects were altered by adding simple changes in geometries to account for



increase in damage. The mean square errors and correlation coefficients were used in determining the performance of the network, and they are reported in Section 6.5. The network showed that it never memorized the relationships because it did not over-fit the data but was able to reproduce all the prominent curves in the profiles.

The identification process was performed in two stages. First the network outputs (reconstructed profiles) were matched with the actual road profiles (targets). Secondly, DSDs were calculated using MATLAB's `spectrum.welch.m` function (Mathworks Inc., 2007). The function used the Hamming window with a segment length of 256 points. The resulting DSDs were then plotted over the ISO PSD classification scale to determine the roughness level of the road.

6.3 Theory

6.3.1 Vehicle modelling

For demonstrating the concept proposed in this work, the entire system was modelled by the LPP 8DOF model comprising 4DOFs associated with the vehicle, 1DOF associated with the driver seat and a further 3DOFs for the seated driver. The linear LPP model was chosen because, bounce and pitch motions are more dominant than roll motion at all frequencies. For most vehicles, resonances in roll occur at lower frequencies (typically between 0.5 and 1 Hz) than resonances in bounce (Gillespie, 1992). However, at these low frequencies the PSD ratio of roll displacement input to vertical excitation from the roadway is usually below 0.1. At high frequencies the PSD ratio approaches unity because the left and right wheel tracks tend to become more uncorrelated. However the vehicle is less responsive to roll input at these frequencies.

The driver seat mass is combined with the mass of the seated driver's lower torso, into a single mass m_0 in the resulting equation of motion. The seat's foam cushion is modelled by a spring and a damping constant, k_0 and c_0 (Figure 6.2) whose values are dependent upon static pre-loads (Wei and Griffin, 1998). Table 6.1 shows the characteristics of a 1DOF model of the seat with preload values from 300 to 800 N.

The 3DOF seated-driver model is based on ISO 5982:2001 (2001) which does not attach any direct relationships to the body segments. The main advantage of this model is its relative simplicity and linearity, while taking into account human body mass as a parameter of the model (Stein and Múčka, 2003).



Table 6.3. Static pre-load dependent parameters of the SDOF cushioned seat (from Wei and Griffin, 1998)

Pre-load (N)	Damping Constant, c_0 (N s / m)	Spring Constant, k_0 (N / m)
300	260	42 300
400	270	44 121
500	276	50 210
600	280	59 300
700	285	68 000
800	293	73 000

The mass of the seated driver was estimated by the standard assumption that approximately three-quarters of the total seated person's mass acts on the vehicle cabin floor via the seat and the rest acts via the legs and possibly the backrest (Stein and Múčka, 2003).

The vibrations are small for most heavy-vehicle driving conditions, so the geometry can be assumed to be linear (Cebon, 1999). The equation of motion of a linear vehicle model (with linear geometry and linear springs and dampers representing tyre and suspension elements) can be written in matrix form as (Gillespie, 1992)

$$\mathbf{M}\ddot{\mathbf{z}} + \mathbf{C}\dot{\mathbf{z}} + \mathbf{K}\mathbf{z} = \mathbf{f}(t) \quad (6.1)$$

where \mathbf{M} is a matrix representing vehicle body mass, \mathbf{C} is a matrix of vehicle damping, \mathbf{K} is the stiffness matrix of the vehicle, \mathbf{z} is a vector of vehicle dynamic responses and $\mathbf{f}(t)$ is a vector of force inputs acting on the vehicle. It was assumed that the pitch angle was so small that kinematical motions at the front axle, rear axle, engine mount and driver's seat might be approximated by the following equations, respectively

$$z_{sf} = z_s - a_f \theta_s \quad (6.2)$$

$$z_{sr} = z_s + a_r \theta_s \quad (6.3)$$

$$z_{sed} = z_s - d_{ed} \theta_s \quad (6.4)$$

$$z_{sd} = z_s - d_r \theta_s \quad (6.5)$$



Equation (6.1) was applied to each DOF in Figure 6.2 and the respective relationships appear in the Appendix A. The property matrices \mathbf{M} , \mathbf{C} and \mathbf{K} are represented by equations. (A.9) to (A.11). Then equation (6.1) was recast into multivariable state space equations given by

$$\dot{\mathbf{Z}} = \begin{Bmatrix} \hat{\mathbf{0}} \\ \mathbf{M}^{-1}\mathbf{f} \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \mathbf{Z} \quad (6.6)$$

where \mathbf{I} and $\mathbf{0}$ represent identity and null matrices with similar dimensions as the property matrices, $\hat{\mathbf{0}}$ is a null column vector with a length similar to the forcing vector \mathbf{f} and \mathbf{Z} is a vector containing the states (displacement and velocity vectors). Equation (6.6) was solved using a Runge-Kutta fourth-order integration routine in MATLAB 7.

6.3.2 Bayesian-regularized NARX neural network

Recently ANN type models have been popularly used to model non-linear dynamics due to their high adaptability to various non-linear systems, the ready availability of tools and the proliferation of computer algorithms (Wong and Worden, 2007). This Section describes how the NARX network was used for reconstruction of road profiles from calculated vertical vehicle accelerations. Although the vehicle model used in this study was linear, the parameters of interest are not related in a simple linear manner. It would therefore be inappropriate to use linear system identification models.

The theory governing the operation of the NARX network is presented in Chapter 4, Section 4.3.2. Here, only the regularisation part is presented.

In order to enable the network to perform as well on novel inputs as on training set inputs, the NARX network utilized Bayesian regularization. Typical training only involves reducing the sum of squared errors, whereas regularization includes minimizing the sum of squares of the network weights to achieve optimal network performance. Therefore the objective function in regularization is written as

$$\boldsymbol{\varepsilon} = \boldsymbol{\beta}\boldsymbol{\varepsilon}_D + \boldsymbol{\alpha}\boldsymbol{\varepsilon}_W \quad (6.7)$$

where $\boldsymbol{\varepsilon}_D$ is the sum of squared errors, $\boldsymbol{\varepsilon}_W$ is the sum of squares of the network weights, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are parameters of the objective function (Hagan and Menhaj, 1994; Bishop, 1995), compromising between fitting the data and producing a smooth network response. In Bayesian regularization this is performed by searching for the optimal weight that maximizes the posterior probability which is equivalent to minimizing the regularized objective function (Bishop, 1995).



6.4 Description of the vehicle dynamics

The LPP model of Figure 6.2 is now used to numerically demonstrate the procedure for reconstructing the road profiles, using ANN simulation. The model takes into account pitch motion which forms an important part of overall heavy-vehicle vertical vibratory response during normal road operation (Gillespie, 1992).

An LPP model of the haul truck which had been characterised by Hugo (2005) was developed. Hugo (2005) determined the tyre characteristics *in situ* whereas those for the hydro-pneumatic struts were determined by fitting the ideal gas law and the energy equation to experimental observations on the nitrogen and oil in the struts. Tyre damping c_t was assumed to be negligible. Although the struts exhibited non-linear characteristics, they were conveniently linearized around their equilibrium states. Table 6.2 shows the model parameters used in the LPP model.

Table 6.2. Parameters of LPP models extracted from Hugo (2005)

Model	Mass (kg) and Inertia (kg m ²)				Damping Constants (Ns/m)		Stiffness Constants (N/m)			
					c_{sf}	c_{sr}				
LPP	M_s	I_s	m_{uf}	m_{ur}	c_{sf}	c_{sr}	k_{sf}	k_{sr}	k_{tf}	k_{tr}
	65790	263160	8578	8578	40×10^3	46152	2.6×10^6	3×10^6	2.7×10^6	5.4×10^6

The distances of the front a_f and rear a_r axles from the centre of gravity of the vehicle were 3.7 and 1.8 m respectively. The driver's seat was located 0.2 m away from the vehicle's centre of gravity (cog) towards the front axle. The engine and driveline forces were assumed to act at 2.5 m away from the centre of gravity on the front axle side. Tables 6.1 and 6.3 show the model parameters, assuming that the total standing mass of the driver was 70 kg.



Table 6.3. Parameters of the driver model based on ISO 5982:2001 (2001) as adapted from Stein and Múčka (2003)

i	m_i kg	c_i (Ns/m)	k_i (N/m)
0	2	---	---
1	6	387	9 999
2	2	234	34 400
3	42.5	1390	36 200

When the models were subjected to unit step road inputs, the sprung mass responses exhibited fading sinusoidal vibrations at the frequency of 4.1 Hz, corresponding to axle hop resonance, and at 0.95 Hz, corresponding to the sprung mass bounce resonance (Figures. 6.4(a) and (b)). The first sinusoidal motion dominated the vibration of the vehicle immediately after both axles had passed the step and died out within a short while after that. The second part of the motion persisted much longer.

Figures. 6.4(a) and (b) furthermore show that the rear suspension deflection was in phase and vibrated at the same frequency with the vehicle body vibration, while the front suspension deflection was at a slightly higher frequency being dominated by pitch motion at a frequency of 1.4 Hz. This was expected since the moment arm of the front suspension from the centre of gravity was more than twice as long as that for the rear suspension. The influence of pitch motion on the vibration of the truck is represented by the differences between quarter-car model results (Figures. 6.4(a) and (c)) and LPP model results (Figures. 6.4(b) and (d)). The differences in the phases and magnitudes of displacements between the two model results are significant and justify the preference for the LPP model.

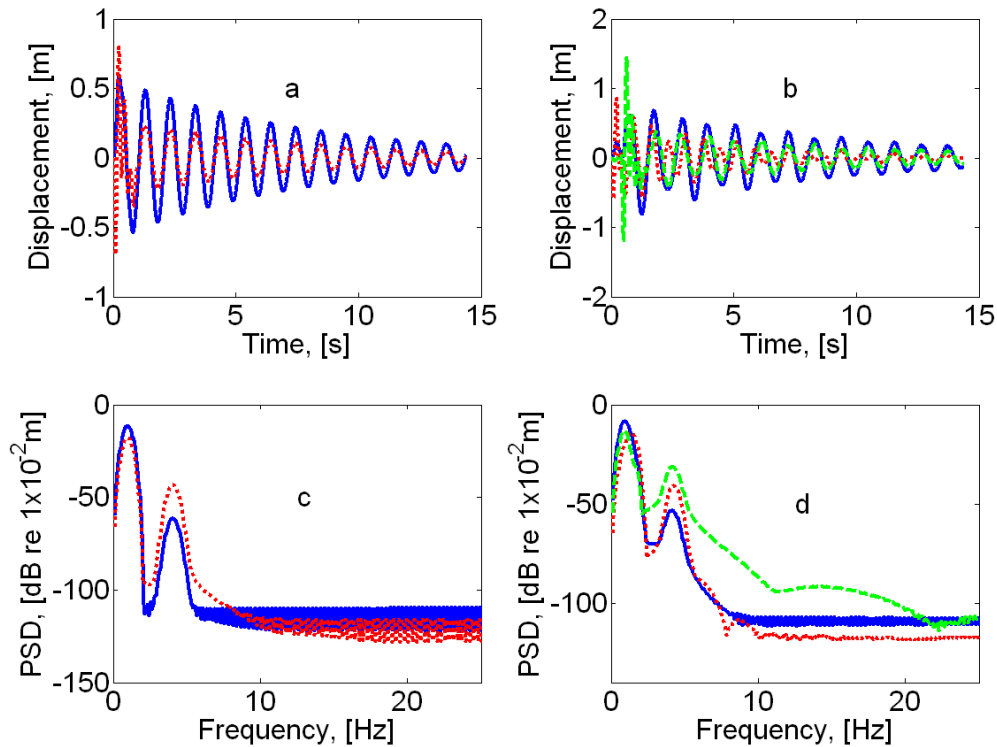


Figure 6.4. Vehicle responses to unit step road input for quarter car (QC) model displacements (a) and PSDs (c), sprung mass responses (solid) and unsprung mass responses (dotted), for LPP (b) and (d) with displacements of the rear suspension (dashed) and front suspension (dotted).

6.5 Numerical studies

The proposed methodology is demonstrated for different vehicle speeds for cases of generally varying roughness grades, emerging and growing defects and for different vehicle payloads and various levels of noise. In all these cases the NARX network was composed of twenty neurons with tangent sigmoid (**tansig.m**) transfer functions in the hidden layer and one neuron having a linear (**purelin.m**) transfer function in the output layer. The network used a training function **trainbr.m** which updates the weights and biases according to Bayesian regularisation in equation. (6.7).

The computational efficiencies of the network for various combinations of inputs and delays were computed and are presented in Table 6.4. The results show that the differences in errors and correlations were influenced more by the number of input delays than by the number of network inputs. Though the sprung mass acceleration \ddot{z}_s with 8 delays shows the best performance, it exhibited too much inflexibility in the presence of road surface imperfections, like speed humps or potholes. The use of front unsprung mass acceleration \ddot{z}_{uf} with 15 delays might provide an alternative efficient solution computationally, but the practical difficulties of mounting accelerometers on the



vehicle axle frustrate its application. Hence the sprung mass acceleration \ddot{z}_s with 15 delays was chosen owing to its relatively better stability, comparative computational efficiency, and ease of practical implementation. Furthermore, for this ultra-heavy truck with massive and stiff suspension system, the transmissibility of the road input onto the vehicle body is quite high so that the use of axle accelerations may not yield any significant differences. So a single-input single-output, 1-20-1 NARX network was created, which had to yield road profile, $\tilde{z}_r(t) = z_{rf}$ as its output.

Table 6.4. Summary of NARX network computational efficiency for different numbers of input vectors and delays

Inputs (Fig. 2)	No. Delays	ENOP n/N	SSE (%)	SSW (%)	Overall Corr. (%)	Duration (s)
\ddot{z}_s	8	60/361	11.68	8.37	97.76	713
\ddot{z}_s	15	116/641	11.04	9.52	97.52	1418
\ddot{z}_{uf}	15	121/641	10.97	11.39	97.89	1410
$\{\ddot{z}_s, \ddot{z}_{uf}\}^T$	15	151/941	10.50	10.93	96.96	2735
$\{\ddot{z}_s, \ddot{z}_{uf}, \ddot{z}_1\}^T$	8	95/681	11.20	10.76	97.23	1569
$\{\ddot{z}_s, \ddot{z}_{uf}, \ddot{z}_1\}^T$	15	163/1241	10.37	11.16	96.73	4655

ENOP – Effective number of training parameters represented by n while N is the total number of training parameters available network

SSE – sum of squared errors

SSW – sum of squared weights

Corr. – Correlation

The length of the road under study was 512 m and was sampled similarly for all velocities at 4 samples per m. The results include evaluations of the correlation coefficient R and the error ε . The correlation coefficient was calculated by using MATLAB's `postreg.m` function, and the error ε was computed by using RMSE values between the reconstructed and actual profiles (Chapter 4: Section 4.5).



The following Sections contain an investigation of different case scenarios. Section 6.5.1 addresses the general problem of reconstructing a road profile of any roughness class and the ability to identify its class within the ISO PSD classification scale. Section 6.5.2 deals with the problem of reconstructing and identifying emerging defects over the road surface. The problems of applying the methodology under changing or growing defects are presented in Section 6.5.3. The methodology is applied to different noise levels and vehicle payloads in Sections 6.5.4 and 6.5.5 respectively.

6.5.1 General road roughness identification

The aim of this case study was to show whether the methodology could be applied to identifying roads that typically fall into different roughness classes. Four roads belonging to classes A, C, D and F were generated. The resulting profiles were applied to the vehicle model to determine the responses. The vehicle responses were then applied to the NARX network previously trained with road profiles lying on class boundaries to reconstruct the corresponding road profiles.

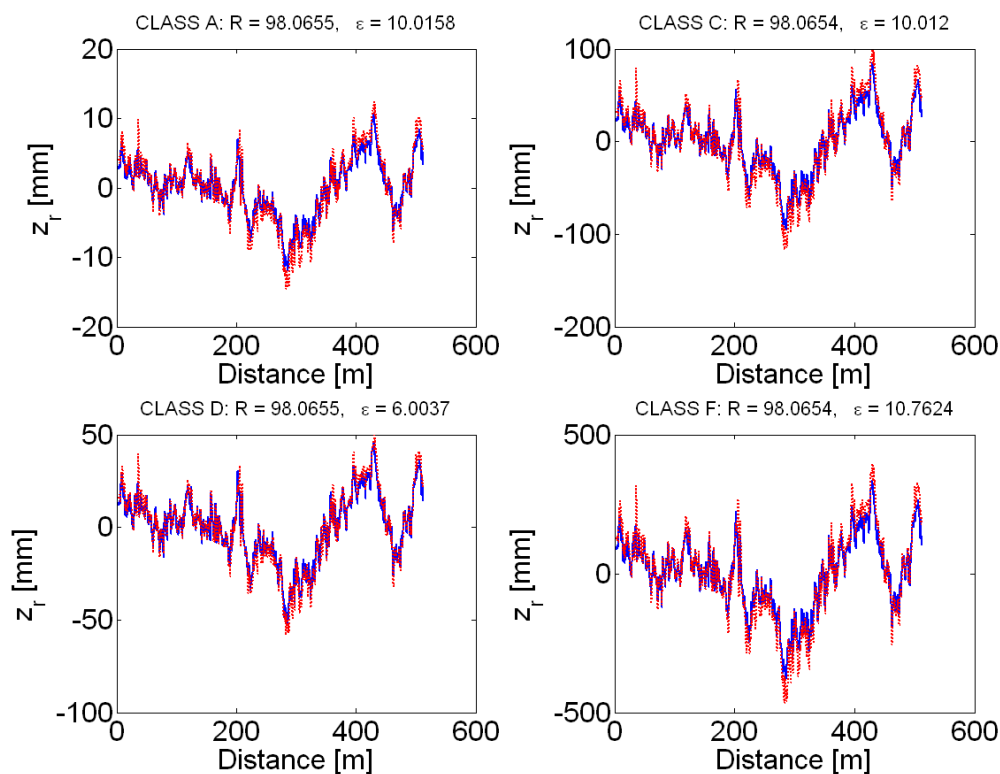


Figure 6.5. Actual (solid) vs. reconstructed profiles (dotted) for different road roughness classes.

Figure 6.5 plots the actual profiles over reconstructed profiles for the four test roads. The figure shows that actual and reconstructed profiles correlated uniformly at 98.1% with a reconstruction error around 10% attaining a maximum of 11% for test profile F. These network performance indicators (correlation coefficients and error percentages) were averaged over a set of ten tests. The errors resulted from network failure to fit corners in the road profiles perfectly. In a previous study



(Ngwangwa, Heyns, Labuschagne and Kululanga, 2008) it was noted that the errors were also influenced by the distance between the displacement PSD of the profiles used during network training and those used in the test. These errors can be reduced by generating more training profiles between any two class boundaries, however, that causes network training to become onerous. Alternatively, training with Gaussian noise at different levels of noise was shown to improve network performance although that resulted in increasing the roughness class and required some sort of correction for the added noise.

In Figure 6.6, the reconstructed profiles were correctly classified into their respective roughness grades A, C, D and F as can be observed from their good correspondence with the actual PSDs (dotted line). These roughness classes are actually interpreted on the ISO PSD scale as very good, average, poor and extremely poor respectively. The IRI scale (Sayers and Karamihas, 1998) identifies the same test roads by the values 2.5, 10.0, 19.9, and 79.8 representing the regions covered by new roads, maintained unpaved roads or damaged roads with frequent shallow depressions some deep, rough unpaved roads with erosion gulleys and deep depressions. The test road corresponding to ISO PSD class F or IRI value of 79.8 was used here merely for theoretical purposes because it lies outside the practical limits of allowable road severity.

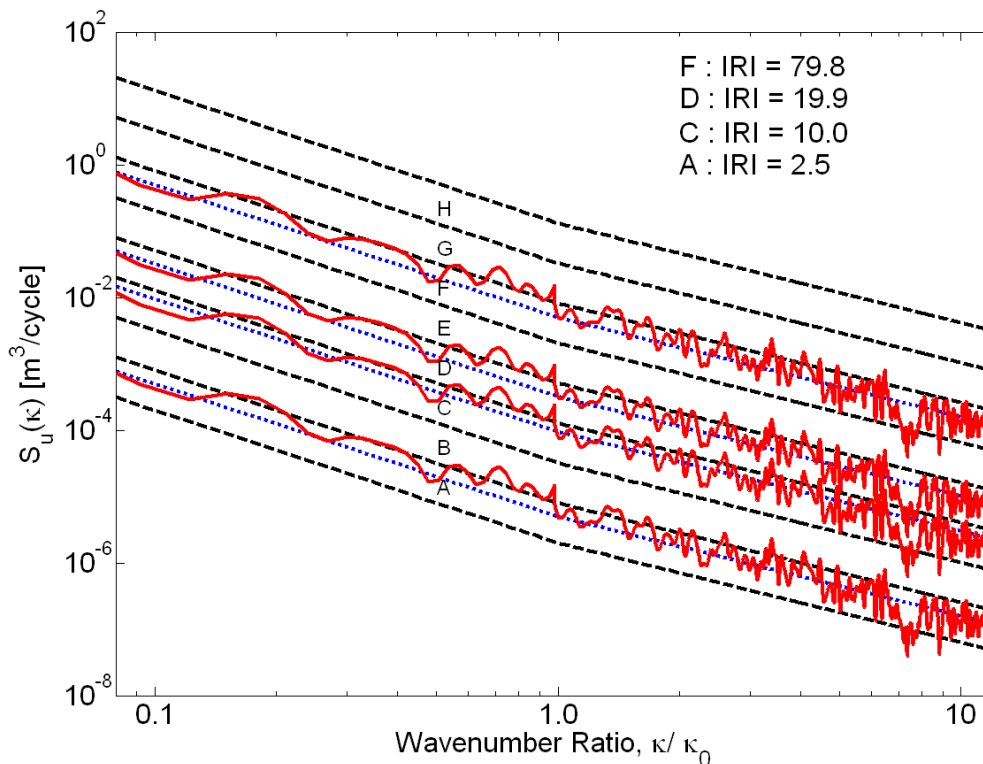


Figure 6.6. Classification by ISO PSD and IRI (in m per km) on the reconstructed profiles showing class boundaries (dashed), actual test road PSDs (dotted) and PSDs of reconstructed profiles (solid).



6.5.2 Defect identification under different speeds

The purpose of this case study was to show the ability of the methodology in identifying localized surface irregularities in the longitudinal direction of the road at different vehicle speeds. The pothole had a width of 2 m and was located along the road at a distance of 50 m from the start. It had an amplitude of 0.012 m expressed as a sinusoidal function of distance x , so that the defect profile can be given by

$$z_{def} = \begin{cases} 0.012 \sin \frac{2\pi x}{\lambda} & x \geq x_1 \text{ and } x \leq x_2 \\ 0 & \text{otherwise} \end{cases} \quad (6.8)$$

where x represents the distance along the test road, λ is the defect wavelength, and z_{def} is the profile's vertical ordinate at distance x . Therefore the road profile to be reconstructed was given by the summation of z_r as given in Chapter 3 and of Eq. (6.10).

$$z_{rx} = z_r + z_{def} \quad (6.9)$$

However, the use of PSD classification restricted the defects, making them periodic in nature. The identification of the pothole was examined for the vehicle speeds of 20, 30, 50 and 80 km h⁻¹. Figs. 6.7(a), (b), (c) and (d) show results of reconstructing the same pothole when the vehicle traverses it at the speeds given above. The correlation values and error percentages on top of the graphs represent network performance over the entire road length.

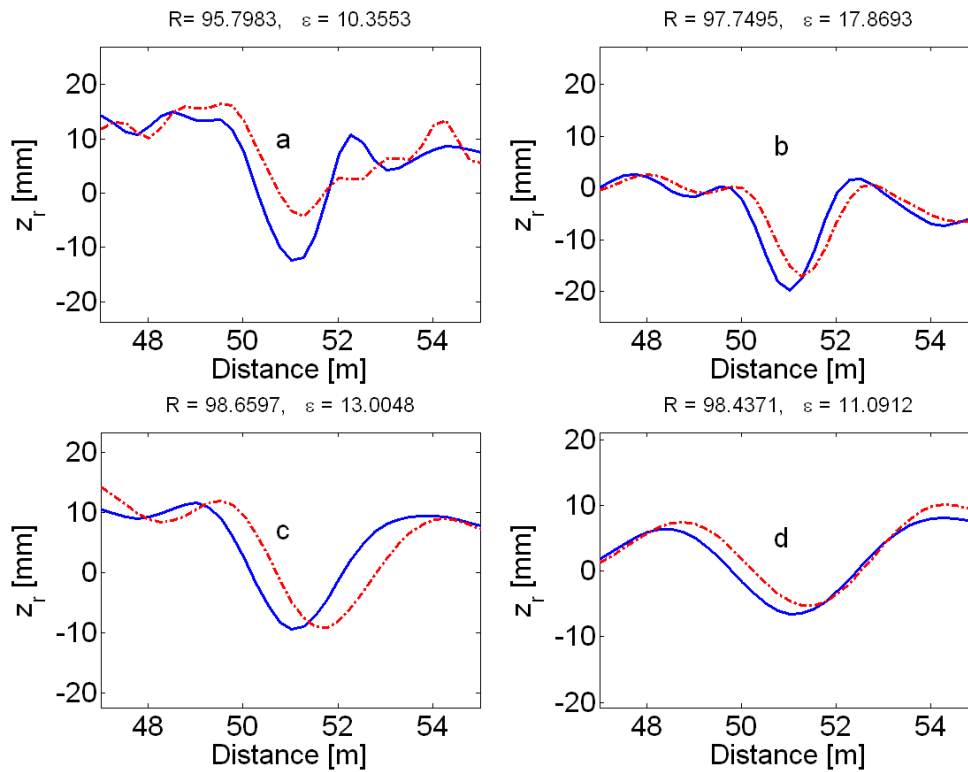


Figure 6.7. Reconstruction at different vehicle speeds: (a) 20 km/h, (b) 30 km/h, (c) 50 km/h, and (d) 80 km/h with actual profiles (solid) and reconstructed profiles (dash-dot).

The pothole was correctly located between 50 and 52 m at all vehicle speeds. The minor shift of 0.2 m in the reconstructed profile might have been caused by an inability to model network delays perfectly, given the inefficiencies of a trial and error procedure. However, it may be argued that for a heavy vehicle travelling at speeds above 20 km/h, such a small delay, corresponding to a maximum time delay of 0.036s, may be insignificant. Besides, the benefits obtained by achieving a zero delay fit by a trial and error procedure could not justify the time and effort required to achieve such a goal.

Figs. 6.8 and 6.9 show the profile reconstructions and classifications for three different damage scenarios at each of the specified vehicle speeds. The different damage scenarios are applied to a class A road. The results in Figure 6.8 show very good correspondence between the actual and reconstructed profiles with a bit of random performance around some of the sharp turns, in that the reconstructed profiles either overshoot or smoothen out. This network behaviour makes it difficult to distinguish among different degrees of roughness within any particular roughness class. In Figure 6.9, for example, it can be observed that for the same underlying roughness, the road class either lies close to the lower or upper boundary of class A. The higher PSDs between wave number ratios of 0.4 and 1.1 correspond to surface defects with various wavelengths. IRI classification shows that the test roads are typically new roads but fails to reflect the small surface imperfections.

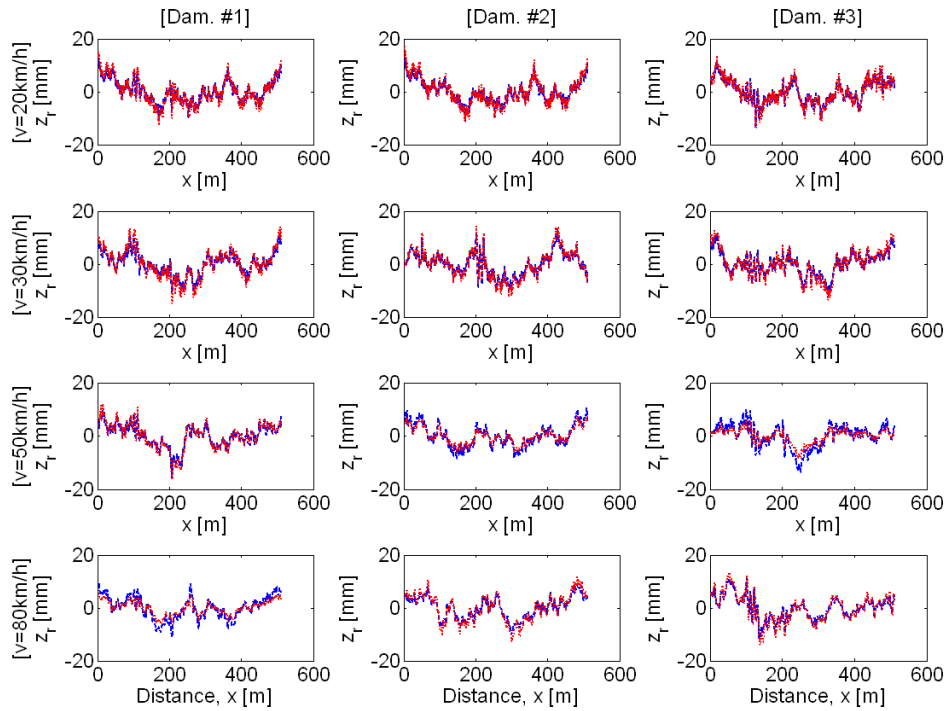


Figure 6.8. Network reconstructions at different speeds for three selected stages of road damage with actual profiles (dashed) over reconstructed profiles (dotted).

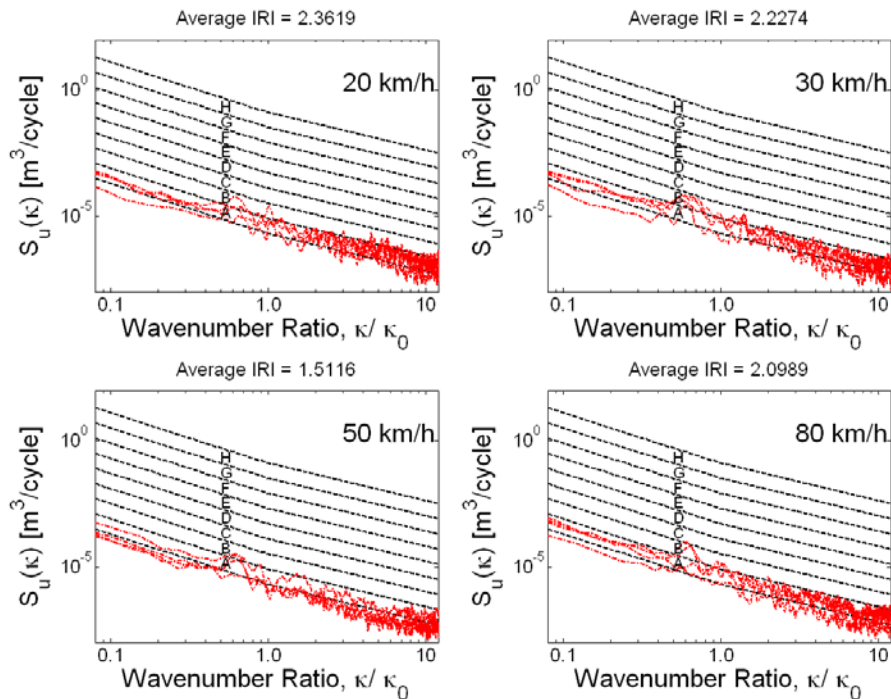


Figure 6.9. Identification under different vehicle speeds for growing defects, IRI values in m per km.



6.5.3 Identification under changing/growing defects

In this case study, the underlying rationale was to show the neural network's capabilities under changing/growing defects. The defect changes are represented in terms of damage scenarios Dam. # N_i , where N_i denotes 1, 2, 3, and 4. These damage scenarios are distinguished by the lengths of deteriorated road Sections and assumed wavelengths of the defects. Figure 6.10 shows that the reconstructed profiles are above 96% level of correlation with errors under 20% for all the damage scenarios. In both graphs (a) and (b) the dash-dotted lines trend average network performance is taken over the four different vehicle speeds.

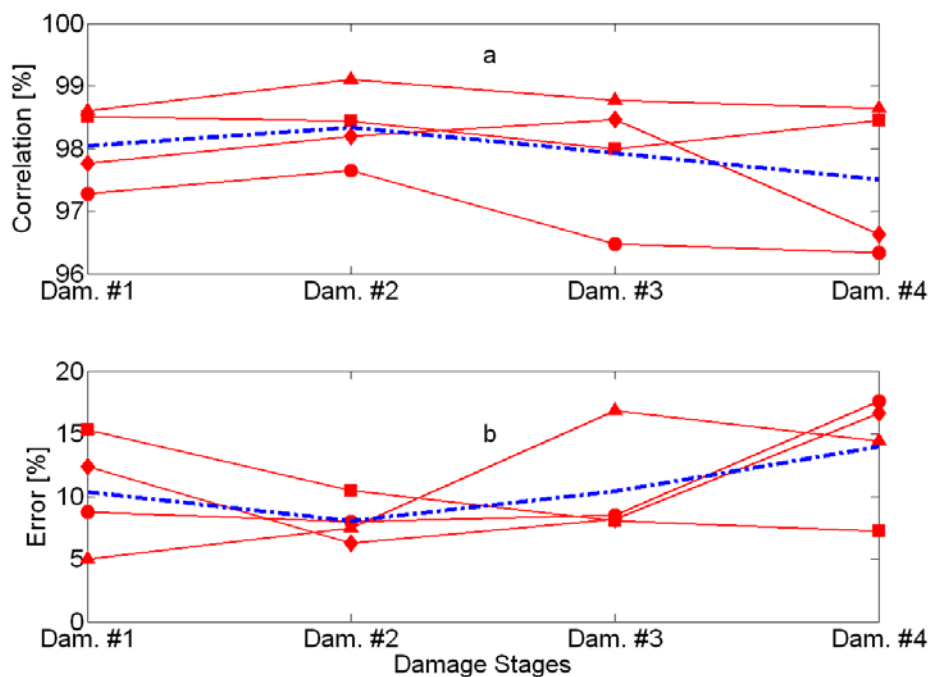


Figure 6.10. Average correlation and error percentages for the different damage levels: 20 km/h (—●—), 30 km/h (—◆—), 50 km/h (—▲—) and 80 km/h (—■—)

Figure 6.11 shows average network performance over all the damage scenarios at each vehicle speed. The profiles are reconstructed at an average correlation level of above 97% with errors lying between 10 and 11%. These results show that the level of damage does not necessarily have any significant influence on network performance at any given vehicle speed. Errors resulted from the usual network shortcomings as previously pointed out and the difficulties encountered in accurately estimating the necessary number of delays in the feedback loop of the NARX network.

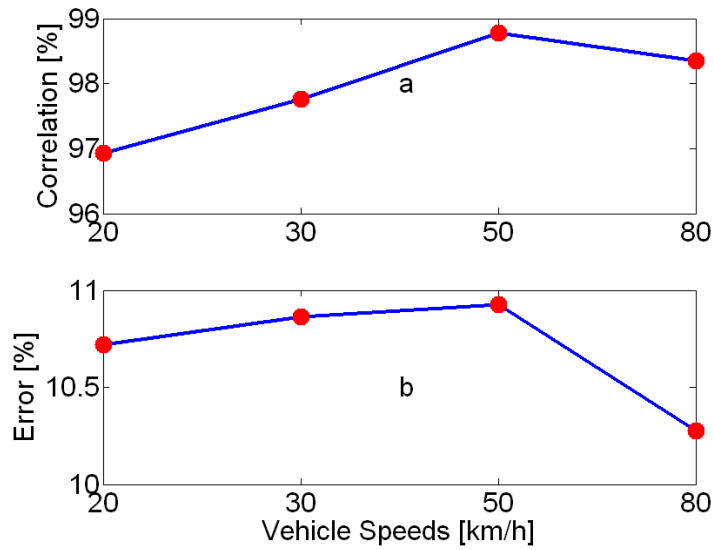


Figure 6.11. Average correlation coefficients and error percentages for each vehicle's speed.

Figure 6.12 shows the DSDs with the corresponding IRI values in m per km. The general level in the values of IRI indicates that the road is in very good condition and in terms of the condition of roads; it implies that the road lies between new roads to older roads (Sayers and Karamihas, 1998). This interpretation agrees with the classification using the ISO PSDs where the road roughness, especially at higher wave numbers, belongs to class A.

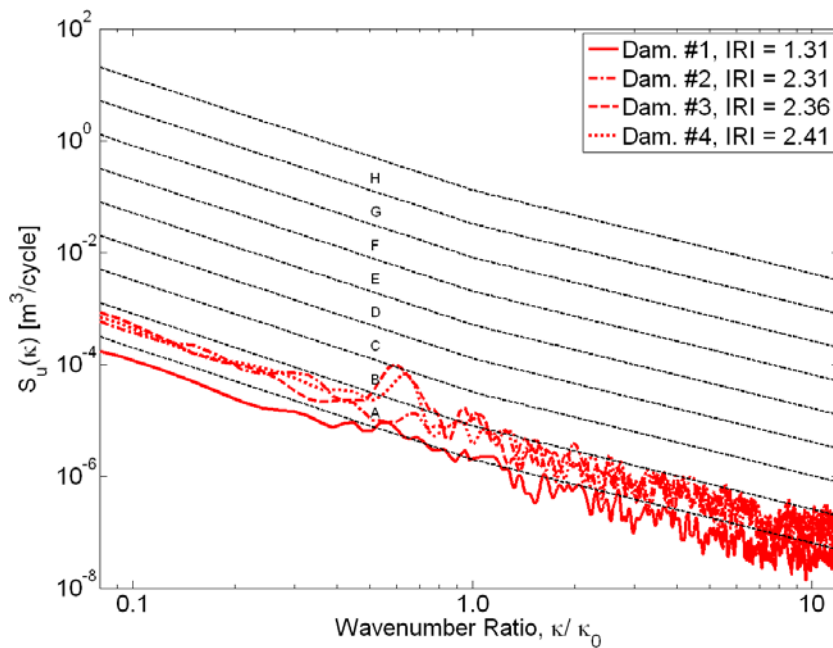


Figure 6.12. DSDs with IRI values for ride interpretation in m per km.



The DSDs show that for Dam. #1 the road lies in class A at lower wave numbers and slightly under class A at very high wave numbers. Dam. #2 largely lies within class A, though at lower wave numbers, it lies very close to its upper boundary. In Dams. #3 and #4 the DSD lies in class C around a wave number ratio of 0.6. However, that information is so obscure in the IRI classifications of 2.36 and 2.41, both of which may typically be interpreted as new roads having near-perfect conditions though not absolutely perfect. This implies that the surface irregularities have to grow beyond these levels in order for the IRI classification to signal emergence of surface imperfections that need attention. Typically IRI values in excess 3.5 may be considered significant see Figure 3.8.

The analysis helps to underscore the advantage of ISO PSD classification over IRI classification. The only shortcoming of the ISO PSD classification being that the road defects should be estimated as periodic functions. The more interesting result in this analysis consists in that both classifications agree on the underlying roughness class. Thus it is possible to use the two methods simultaneously to enhance understanding of the nature of road damage.

6.5.4 Identification under noisy conditions

In this case study, Band-Limited White (BLW) Noise was applied to a class A road having a surface irregularity. The added BLW Noise lay in the frequency range 1 – 25 Hz with noise-signal ratios (NSR) between 0 and 0.5. Fig. 6.13 shows that the error in reconstructing the road surface profiles grows with noise level and it might be necessary for practical purposes to keep the NSR below 0.2. Fig. 6.14 shows ISO PSD and IRI classifications at 0.2 NSR level for different vehicle speeds. IRI classifies the road around an average roughness level of 2.5 indicative of new roads without surface imperfections, although the picture is slightly different at the lowest vehicle speed of 20 km/h where it approaches a domain of new roads with surface defects. ISO PSD lines largely lie within class A with slight upward shift in lower wave numbers which are emphasized around the wave number ratios 0.4 – 1.1 corresponding to the surface defects' wavelengths.

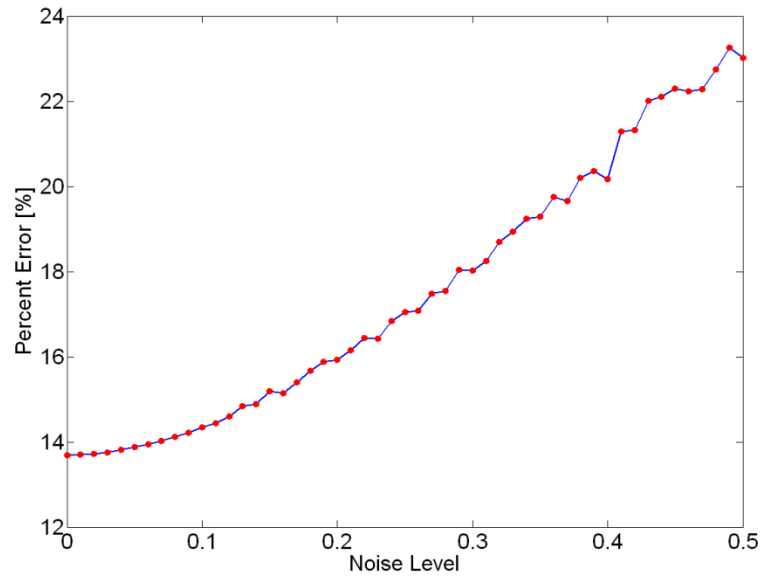


Figure 6.13. Errors at each noise-signal ratio (NSR) averaged over 10 test runs when trained without noise.

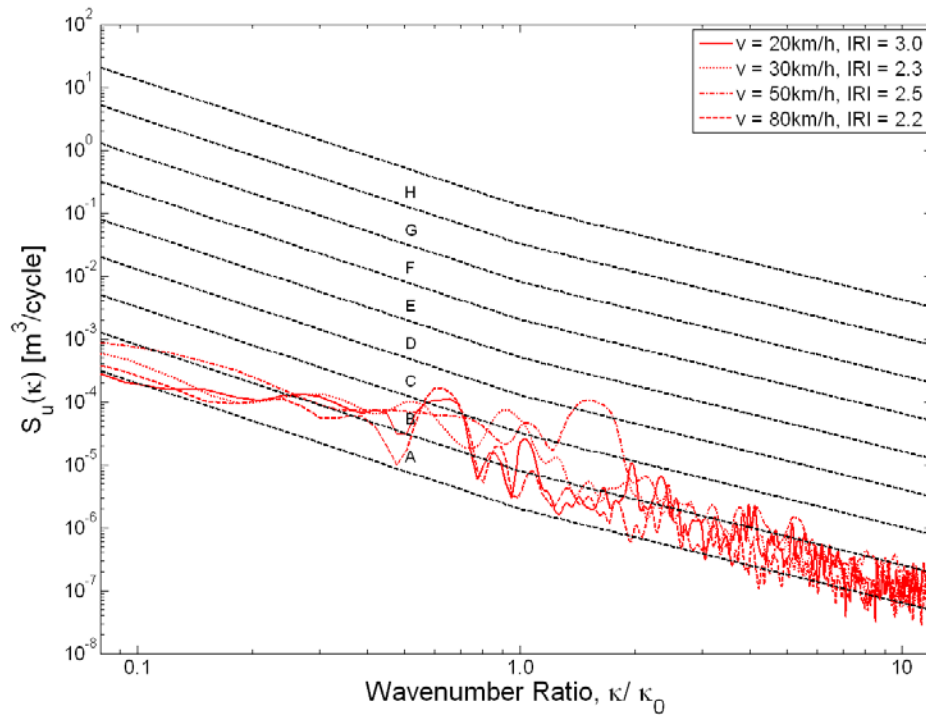


Figure 6.14. ISO PSD classification at 0.2 NSR for different vehicle speeds, IRI values in m per km.



It is further noted that profiles with smaller NSR are well classified into class A whereas those with larger NSR tend to shift upwards into higher roughness classes particularly in lower wave numbers. Theoretically, that is expected because small levels of noise make the NARX network behave like a regularised network. Bishop (1995) notes that during training, provided the noise amplitude is small, so that the neglect of the higher-order terms in the Taylor expansion is valid, the minimization of the sum-of-squares error with noise added to the input data is equivalent to the minimization of the regularized sum-of-squares error without the addition of noise.

6.5.5 Identification under different vehicle weight

The gross weights of vehicles vary significantly during operation owing to changes in payloads. It is assumed that suspension characteristics remain generally invariant over a larger proportion of the vehicle's operational period. Therefore this case study examines how the methodology performs under vehicle weight changes for different speeds. The weights are changed by adding to the gross vehicle weight (GVW) proportions of its own weight between 0 and 0.5. The vehicle was assumed to operate along a lower-boundary class E rough road at a speed of 20 km/h.

The results in Fig. 6.15 show that profiles are accurately reconstructed with errors around 20% at greater than 94% correlation level. The road roughness is also accurately classified in terms of both ISO PSD and IRI classifications as shown in Fig. 6.16. The IRI classification of 22.66 implies a rough unpaved road which corresponds to ISO PSD's lower-boundary E road. The results show that changes in vehicle weight do not significantly affect the performance of the network, at least within practical limits of payload variation.

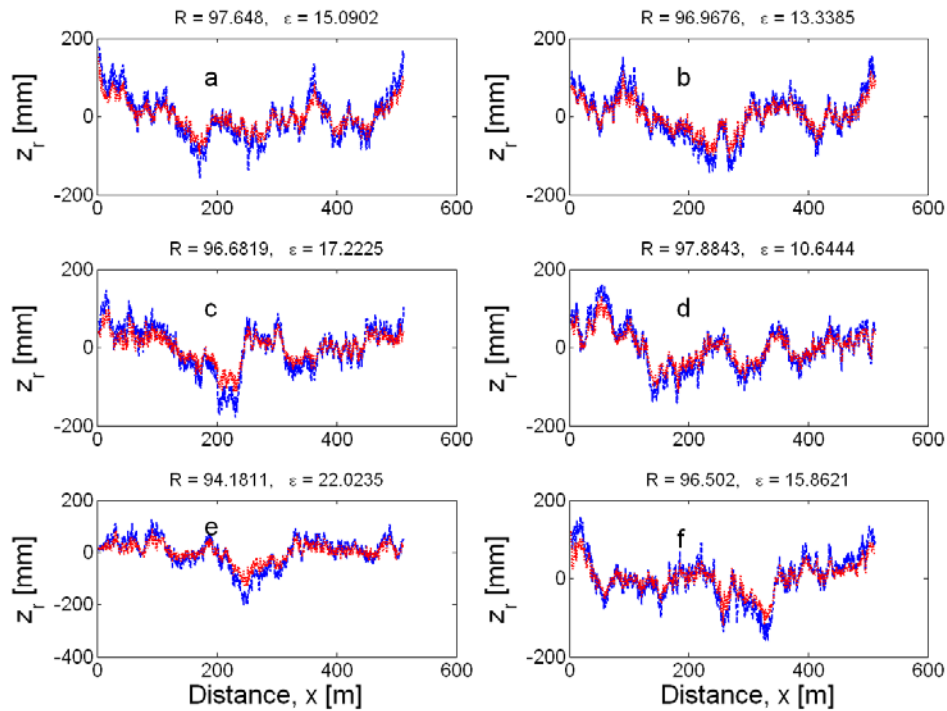


Figure 6.15. Actual (solid) Reconstructed (dotted) profiles in different payloads expressed as proportions of normal gross vehicle mass (GVM): (a) no added payload (b) 0.1 GVM (c) 0.2 GVM (d) 0.3 GVM (e) 0.4 GVM (f) 0.5 GVM.

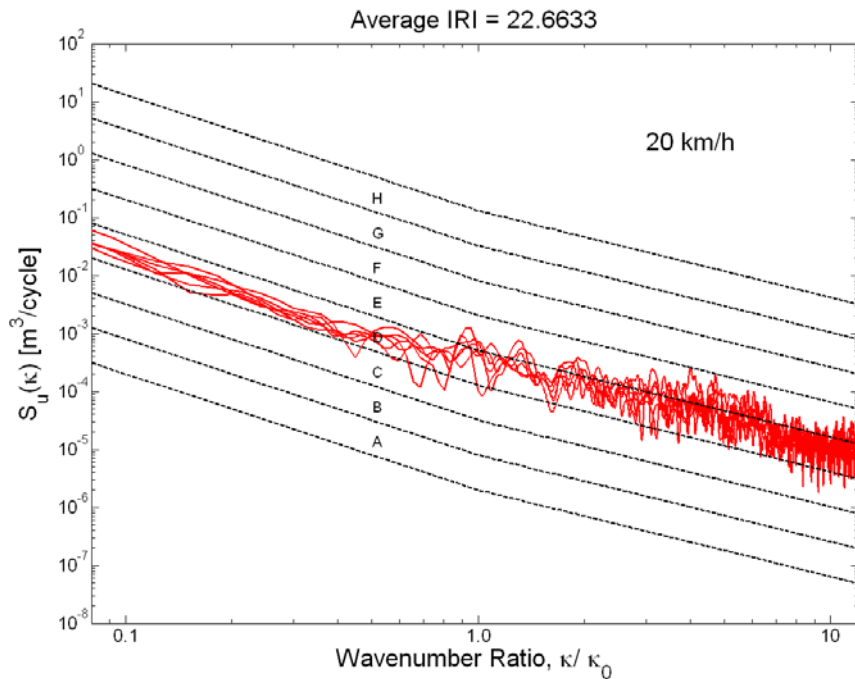


Figure 6.16. ISO PSD classification for the different vehicle payloads in Figure 6.15, IRI values in m per km.



6.6 Conclusions

This Chapter is aimed at numerically demonstrating the proposed road damage identification methodology. The use of neural networks has the primary advantage that it does not require excessive system characterisation. It simply requires the road profile data representative of various degrees of road roughness as network targets, and vertical accelerations measured on the vehicle system as network inputs. In this particular study, owing to the nature of the vehicle suspension system, the use of accelerations other than those measured on the sprung mass offered no significant advantages. However, it is realised that determining the optimal number and type of network inputs might be application-specific and is a subject of further study in this work.

The second advantage of the neural network based methodology is that it requires relatively fewer analytical skills to create the network than the parametric model. The current developments in technical computing have ensured that most neural network models become standard and implemented in technical software e.g. MATLAB, whereas the development of physical parametric models still demands a great deal of technical skills, even in virtual-computing environments. Besides, the physical parametric models require the calculation of the inverse models in order to determine the road profiles from measured vehicle accelerations, which may also be quite challenging.

The performance of the methodology has been demonstrated for varying road roughness levels, emerging surface defects, and changing or growing surface defects under varying conditions of noise, vehicle payload and speed through evaluations of correlation coefficients and mean square errors. The network performance does not exhibit any clear dependencies on the different conditions under study except that under extremely high noisy conditions, the performance deteriorated significantly.

The methodology was observed to correlate well with IRI which is a widely used roughness indicator. The study reinforces the combined use of IRI and ISO PSD classification to enhance road surface condition interpretation. More promising is the fact that a practical test case inherently includes the filtering effects of tyre envelopment, which were not considered in this numerical study. Therefore some of the requirements for stringent accuracy in this numerical study might be superfluous in a practical situation. Accordingly the proposed methodology aptly provides a methodology for approximating and making reliable initial judgements of road surface damage.



Chapter 7

Experimental case study

7.1 Introduction

This Chapter builds on the investigation of Chapter 6 where the methodology was demonstrated on a numerical case study. In that Chapter, an eight degree-of-freedom (8-DOF) linear pitch-plane model with four degrees of freedom representing vehicle motions and additional four degrees for seat and driver vibrations, was employed for the investigation. Eight road roughness classes having well-known DSDs on the ISO PSD classification scale were applied to the model to calculate corresponding sprung mass and axle accelerations. The accelerations, serving as inputs, and the corresponding road profiles, serving as targets, were applied to a Bayesian regularised NARX network with 20 hidden neurons and one output neuron (Ngwangwa, Heyns, Labuschagne and Kululanga, 2010).

The vehicle model was subsequently simulated with different sets of road profiles which were generated by a random road surface function to calculate corresponding sprung mass and axle accelerations. The accelerations were accordingly applied to the trained NARX network to obtain the road profiles. The simulated road profiles were compared with the actual road profiles. These evaluations were performed for different road roughness conditions, different vehicle speeds, growing road defects, noisy conditions, and different vehicle payload conditions. It was noted that the methodology was able to reconstruct the road profiles to within an error margin of 20 % with a minimum correlation of 94 % (Ngwangwa, et al., 2010).

In this Chapter, the methodology is applied to field test data using an experimental Land Rover Defender 110. This vehicle has been extensively used by the Vehicle Dynamics Group at the University of Pretoria and has a great deal of instrumentation permanently mounted on it (Els, 2006). It has been adapted for suspension tests to allow for suspension settings to be safely switched between ride comfort and handling modes, a system that is known as four state semi-active suspension ($4S_4$) (Els, 2006). The vehicle response measurements used in this paper were carried out by Breytenbach and Els (2011) while the road profiles were measured by Becker and Els (2014) using a 'can-can' machine. The tests were performed at Gerotek, a vehicle testing infrastructure in Pretoria, South Africa. The tests were conducted over discrete obstacles and Belgian paving with the driving conditions changing between different combinations of two speeds and suspension settings from one test run to another. These would typically yield four possible test scenarios, namely, ride comfort mode at low speed, ride comfort mode at high speed, handling mode at low speed and handling mode at high speed though some of these combinations were not performed over the discrete obstacles.



One of the most crucial issues with the use of neural networks in non-linear regression or function approximation is its generalization capabilities. The data used in this study was obtained to validate a mathematical model of the Land Rover and since most evaluations of vehicle vertical dynamics are generally conducted over discrete obstacles and random road surfaces (Breytenbach, 2009), there was no need, in that study, for more road profile data other than the discrete obstacles and Belgian paving. However, in this study, such scarcity in the measured profile data poses a challenge as far as ensuring network generalization is concerned. This has been addressed in this Chapter by generating slightly altered versions of the actual measured road profiles which are applied to a full-vehicle 7DOF of the Land Rover as formulated by Breytenbach (2009). The numerically generated data is used for training the neural network while the measured data is used for validating and testing the neural network.

The following Section presents the properties of the Land Rover, its suspension forces and the numerical model which is used in generating the accelerations that serve as inputs when training the neural network.

7.2 Land Rover properties, suspension and numerical model

7.2.1 Geometry and inertial properties

The Land Rover geometry and inertial properties used in the study were characterized by Uys, Els, Thoresson, Voigt, and Combrinck (2006) in a laboratory environment. The determination of the mass moments of inertia and the location of the vehicle's centre of gravity were of particular interest and were carried out by using the rotational vibration of a rigid body about a pivoting point, with a restoring force provided by a spring. Uys et al. (2006) developed an ADAMS model of the Land Rover based on these properties which was later validated with field results by Els (2006). The details of these measurements and the procedures that were employed can be found in Uys et al. (2006). Table 7.1 summarises these vehicle inertial properties and geometry.



Table 7.1 Summary of vehicle inertial properties and geometry.

Land Rover Property and Dimensions	Quantity
Sprung mass	1734 kg
Sprung mass pitch moment of inertia	2440 kg.m ²
Sprung mass roll moment of inertia	688 kg.m ²
Front unsprung mass	229 kg
Front unsprung mass roll moment of inertia	33.1 kg.m ²
Rear unsprung mass	229 kg
Rear unsprung mass roll moment of inertia	33.1 kg.m ²
Centre of mass longitudinal position from front tire centre	1.395 m
Centre of mass height	1.19 m
Wheel track width	1.5 m
Suspension track width	1.1 m

7.2.2 The tyre

The tyre is the most important component of the vehicle model as it acts as the interface between the vehicle and the road (Hugo, Heyns, Thompson and Visser, 2008; Breytenbach, 2009; Gillespie, 1992). Ideally it must maintain traction on the road as it absorbs the road irregularities and because of the range of functions the tyre must perform, it is also equally difficult to model. In most vehicle simulations, the tyre imposes the largest limitation to achieving correlation with measured results (Hugo et al., 2008). In this study the tyre is modelled as a point follower with a parallel linear spring and damper of 250000 N/m and 1200 Ns/m respectively (Breytenbach, 2009).

7.2.3 The suspension forces

Owing to its complex fluid dynamics and need to optimize vehicle performance between ride comfort and handling modes, the suspension dynamics of the $4S_4$ could be oversimplified if it is assumed that suspension forces are directly proportional to suspension displacements and velocities (Els, 2006; Breytenbach, 2009). There are several mathematical models proposed in the literature for complex



damper behaviour but this study adopts the approach employed by Breytenbach (2009) (also see Breytenbach and Els, 2011) where empirical models are employed. The suspension forces are considered to comprise five main contributing components namely: hydro-pneumatic spring forces, hydraulic damping force, friction force in the $4S_4$ suspension struts, bushing force in the vehicle trailing arms and bump stop forces (Breytenbach, 2009). The Land Rover's hydro-pneumatic spring is modelled as a polytropic gas compression process with constant area in equation (7.1) (Breytenbach, 2009),

$$F_s = p_{stat} A \left(\frac{x_{stat}}{x} \right)^{n_p} \quad (7.1)$$

where F_s is the force in the hydro-pneumatic spring; p_{stat} is static pressure (2.167 MPa); A is the bore cross-section area ($1.963 \times 10^{-3} \text{ m}^2$ equivalent to 50 mm diameter bore); x_{stat} is the static displacement (0.0509 m for handling mode and 0.255 m for ride comfort mode); x is the hydro-pneumatic spring displacement; and n_p is the polytropic gas constant equal to unity. The polytropic constant of unity represents an isothermal process and it was found that it adequately predicted the spring force in most tests (Breytenbach, 2009).

The other four forces in the $4S_4$ suspension are nonlinear and could not be modelled as linear functions as is often assumed. They were modelled empirically using MATLAB's interpolation function `ppval.m` (Mathworks Inc., 2007). Figure 7.1 shows the behaviour of these forces over some typical operating suspension displacements and velocities.

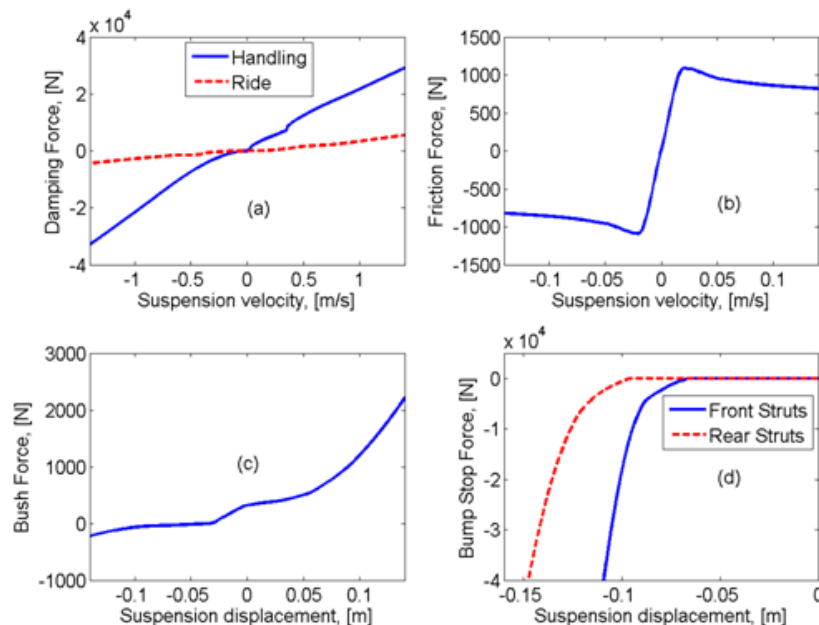


Figure 7.1. The $4S_4$ suspension forces: (a) hydraulic damping forces for ride comfort (dashed) and handling mode (solid) (b) Friction forces (c) Bushing forces and (d) Bump stop forces for front struts (solid) and for rear struts (dashed).



Figure 7.1 (a) shows larger damping forces in the handling mode than in the ride comfort mode at all suspension velocities. This behaviour is uniform in all the suspension struts, though in practice, small variations within manufacturing tolerances have led to varying characteristics (Breytenbach, 2009). Figure 7.1 (b) shows that friction forces within the $4S_4$ suspension are lower than the damping forces but its limiting static value is higher than the dynamic friction forces. The value of 0.02 m/s is ideally too large for stick-slip velocity. This is caused by the participation of viscous damping effects on damper blocks in the valve assembly of the laboratory suspension strut. This fact was verified by another laboratory suspension strut which had no damper blocks in the valve assembly and it exhibited near-zero threshold velocity in the stick-slip transition region (Breytenbach, 2009). Since the suspension strut on the actual Land Rover under test has damper blocks, the former friction characteristic was utilized. Stribeck curves were consequently fitted to the measured ride and handling friction curves and averaged over the velocity range ± 0.05 m/s (Breytenbach, 2009; Breytenbach and Els, 2011). Thus Figure 7.1(b) represents a generalised friction curve which is a result of averaging ride and handling friction curves. However during the actual simulations, specific ride and handling friction curves were employed.

The bushing forces in the trailing arms are also much lower than the damping forces for suspension displacements between -0.1 m and +0.05 m, yet increase almost exponentially outside these limits (Figure 7.1 (c)). By design, the rear suspension struts have greater working space than front struts as shown in Figure 7.1 (d).

7.2.4 The vehicle model

The development of the full-car model of the Land Rover Defender 110 and validation against measured data was carried out by Breytenbach (2009). The model comprises translational motion in the vertical direction and rotational motions in the roll and pitch directions for the sprung mass as well as translational motion in the vertical direction and rotational motion in the roll direction for each of the unsprung masses (front and rear solid axles).

The full development of the model can be found in reference (Breytenbach, 2009). In this study, only the seven equations of motion for each degree of freedom are presented and briefly discussed for clarity.

The equation of motion for the sprung mass can be represented by equation (7.2).

$$m_b \ddot{z}_b = m_b g - F_{srr} - F_{slr} - F_{slf} - F_{srf} \quad (7.2)$$

where m_b is the vehicle body mass, \ddot{z}_b is the vertical body acceleration measured at its centre of gravity and F_{srr} , F_{slr} , F_{srf} and F_{slf} are the suspension forces measured on the struts at the right rear, left rear, right front and left front positions. Force of gravity on the sprung mass $m_b g$ is included to



model sprung mass's free body motion in the vertical direction since occasionally the tires lose contact with the ground.

The forces in the suspension struts were modelled empirically using MATLAB's interpolation function `ppval.m` (Mathworks Inc., 2007). This function constructs a polynomial that fits the best line to the experimental data. The total suspension force in each strut is therefore determined by adding the spring forces, damper forces, friction forces, bump stop forces, and bushing forces. The force equation at the front axle is given by equation (7.3)

$$\begin{aligned} & m_{uf} \ddot{z}_{uf} + k_{trf} z_{uf} + k_{trf} c \theta + c_{trf} \dot{z}_{uf} + c_{trf} c \dot{\theta} + k_{tlf} z_{uf} - k_{tlf} c \theta + c_{tlf} \dot{z}_{uf} - c_{tlf} c \dot{\theta} \dots \\ & = k_{trf} z_{rrf} + c_{trf} \dot{z}_{rrf} - k_{trf} \delta_{stf} + k_{tlf} z_{rlf} + c_{tlf} \dot{z}_{rlf} - k_{tlf} \delta_{stf} + m_{uf} g + F_{srf} + F_{slf} \end{aligned} \quad (7.3)$$

where, m_{uf} is the mass of the front solid axle; z_{uf} , \dot{z}_{uf} , and \ddot{z}_{uf} are the displacement, velocity and acceleration of the axle's centre of gravity; k_{trf} and k_{tlf} are the right and left front tire stiffnesses; c is the wheel track half width; c_{trf} and c_{tlf} are the right and left front tire stiffnesses; θ and $\dot{\theta}$ are the roll angle and velocity of the front axle; δ_{st} is the tire static deflection which is assumed 10 mm in this study; and z_{rrf} , \dot{z}_{rrf} , z_{rlf} and \dot{z}_{rlf} are the profile heights and their time rates of change on the right and left front axle.

In equation (7.3), besides the suspension forces on the left and right hand sides, the axle experiences the gravitational forces due to its own mass $m_{uf} g$; the forces from the road input on the tires and the static forces due to tire deflection δ_{stf} . Similarly the force equation of motion on the rear axle is given by,

$$\begin{aligned} & m_{ur} \ddot{z}_{ur} + k_{trr} z_{ur} + k_{trr} c \phi + c_{trr} \dot{z}_{ur} + c_{trr} c \dot{\phi} + k_{tlr} z_{ur} - k_{tlr} c \phi + c_{tlr} \dot{z}_{ur} - c_{tlr} c \dot{\phi} \dots \\ & = k_{trr} z_{rrr} + c_{trr} \dot{z}_{rrr} - k_{trr} \delta_{str} + k_{tlr} z_{rlr} + c_{tlr} \dot{z}_{rlr} - k_{tlr} \delta_{str} + m_{ur} g + F_{srr} + F_{slr} \end{aligned} \quad (7.4)$$

where the quantities are similar to those already defined in equation (7.3) but, in this case, they refer to the rear axle.

The moment equations on the sprung mass in the roll and pitch directions are given by equations (7.5) and (7.6), respectively.

$$I_{XX} \ddot{\alpha} = (F_{slr} + F_{slf})d - (F_{srr} + F_{srf})d \quad (7.5)$$

$$I_{YY} \ddot{\beta} = (F_{slf} + F_{srf})b - (F_{srr} + F_{slr})a \quad (7.6)$$

where I_{XX} and I_{YY} are the moments of inertia of the vehicle body in the roll (α) and pitch (β) directions respectively; d is half the distance between left and right front suspension struts; a and b are the distances of the rear and front axles from the vehicle's centre of gravity respectively.



The moment equations on the front and rear axles are given by equations (7.7) and (7.8), respectively.

$$\begin{aligned} I_{XX_a} \ddot{\theta} + ck_{tlf} z_{uf} + c^2 k_{tlf} \theta - cc_{tlf} \dot{z}_{uf} + c^2 c_{tlf} \dot{\theta} + ck_{trf} z_{uf} + c^2 k_{trf} \theta + cc_{trf} \dot{z}_{uf} + c^2 c_{trf} \dot{\theta} \dots \\ = -ck_{tlf} z_{rlf} - cc_{tlf} \dot{z}_{rlf} + ck_{tlf} \delta_{stf} + ck_{trf} z_{rrf} + cc_{trf} \dot{z}_{rrf} - ck_{trf} \delta_{stf} + d(F_{stf} - F_{slf}) \end{aligned} \quad (7.7)$$

$$\begin{aligned} I_{XX_a} \ddot{\phi} - ck_{tlr} z_{ur} + c^2 k_{tlr} \phi - cc_{tlr} \dot{z}_{ur} + c^2 c_{tlr} \dot{\phi} + ck_{trr} z_{ur} + c^2 k_{trr} \phi + cc_{trr} \dot{z}_{ur} + c^2 c_{trr} \dot{\phi} \dots \\ = -ck_{tlr} z_{rlr} - cc_{tlr} \dot{z}_{rlr} + ck_{tlr} \delta_{str} + ck_{trr} z_{rrr} + cc_{trr} \dot{z}_{rrr} - ck_{trr} \delta_{str} + d(F_{srr} - F_{slr}) \end{aligned} \quad (7.8)$$

where I_{XX_a} is the mass moment of inertia of each axle.

The equations may now be rewritten in the matrix form

$$\mathbf{M}\ddot{\mathbf{z}} + \mathbf{C}\dot{\mathbf{z}} + \mathbf{K}\mathbf{z} = \mathbf{f} \quad (7.9)$$

where the matrices and vectors in equation (7.9) are given in References (Breytenbach, 2009; Breytenbach and Els, 2011).

Equation (7.9) is a stiff differential equation owing to the existence of the nonlinear suspension forces and the resulting nature of its stiffness and damping matrices. Hence it is solved by using the low-order routine `ode23s.m` implemented in MATLAB, which is suited to such stiff differential equations (Mathworks Inc., 2007). The standard ODE solver `ode45.m` in MATLAB encounters convergence problems especially in the presence of higher damping, friction and bushing forces. This fact was also noted by Breytenbach (2009).

7.3 Nominal road profiles

In the previous Chapter, half-sine waves were used to represent road bumps and random road profiles were generated in a purely mathematical way using the well-known one-dimensional random profile function represented by equation (3.18) in Section 3.4.1. In this study, two kinds of road profiles are used: trapezoidal-shaped bump profiles and Belgian paving. The unique feature is that these profiles are based on real measured road surfaces (Becker, 2008; Becker and Els, 2014). The bump profiles were made of trapezoidally shaped steel blocks whose dimensions could be obtained using a one meter rule. The distances along the road were measured using measuring tapes.

The measurements over the Belgian paving were more complicated and therefore required a better profilometer. The profile measures were taken by Becker (2008) using three different methods: ‘can-can’ machine, photogrammetry and laser scanner, which are shown in his work to be accurately correlated (Becker, 2008; Becker and Els, 2014). It is however reported that the ‘can-can’ machine provided a better profile measuring tool due to its ease in setting up, operation and data pre-processing. The ‘can-can machine’ is a light weight, right-angled triangular structure with a wheel at each of its corners, designed in such a way that one wheel is in front and is used for steering the



apparatus and controlling its speed during profiling, whereas the other two wheels are at the rear carrying the measuring beam. The rear beam is 4.5 m long and it carries 30 pivoting arms positioned and spaced to cover a 3 m wide road Section. It is reported that when a ‘can-can’ machine is properly used it is capable of achieving profiling errors of less than a few millimetres (Becker, 2008; Becker and Els, 2014).

This accuracy is very satisfactory for the purposes of this study where small bumps may be neglected due to the tire filtering effect (Becker, 2008; Becker and Els, 2014). The vehicle tires tend to cushion the vehicle from the impacts of surface undulations whose wavelengths are less than its contact patch width. A detailed description of the profiling procedure and the equipment used can be found in Becker's work (Becker, 2008; Becker and Els, 2014).

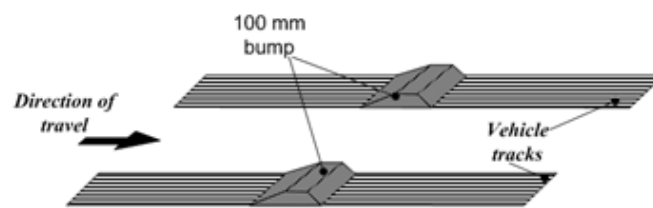
7.3.1 Trapezoidal bumps

The trapezoidal bumps are used here to represent discrete-obstacle type of defects such as potholes, bumps and stones (Breytenbach, 2009). Though the real-life discrete obstacles on the road may not often be in that shape, trapezoidal bumps are popularly used in vehicle model validation tests (Becker, 2008; Becker and Els, 2014; Letherwood and Gunter, 2001) due to their ramp-type of rise and fall thereby imposing mild harshness to the vehicle structure in driving. Two trapezoidal bumps of different sizes (small bump with 100 mm height and large bump having a height of 150 mm) were placed along the tracks in three different layouts as shown in Figure 7.2.

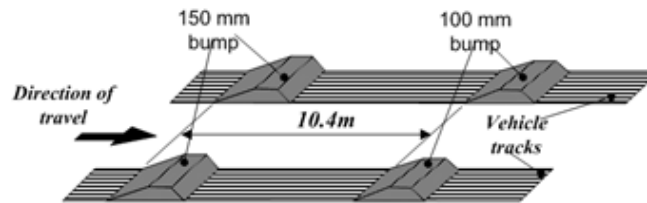
The bump layouts are shown in the picture in Figure 7.2 (a) with the Land Rover traversing the bumps. The spacing between the bumps is depicted in Figure 7.2(b) to (d). The first layout had two smaller trapezoidal bumps placed symmetrically along each wheel track. This was ideal for testing the vehicle's vertical motion. The second layout, Figure 7.2 (c) had the large bumps placed symmetrically along each wheel track followed by the small bumps similarly laid out and positioned at a distance of 10.4 m from the large bumps in the direction of travel. This layout was ideal for testing of both vertical and pitch motions. The third layout Figure 7.2 (d) comprised one large bump on the left hand side of the wheel track followed by the small bump at a distance of 10.4 m placed on the right hand side of the wheel track. This layout was ideal for testing vehicle's vertical, pitch and roll motions.



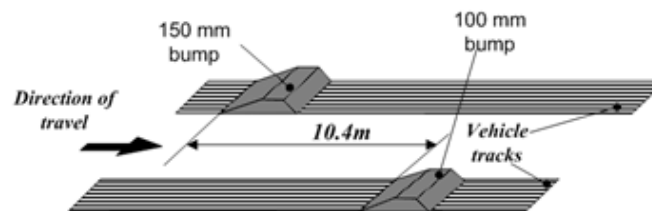
(a)



(b)



(c)



(d)

Figure 7.2. Bump layouts in picture (a) and in schematic form showing the distances in: (b) symmetric small bump (c) symmetric large-then-small bump (d) asymmetric large-then-small bump.

7.3.2 Belgian paving

The Belgian paving is often used to evaluate vehicle dynamics typically experienced when traversing random rough terrains (Becker, 2008; Becker and Els, 2014; Letherwood and Gunter, 2001). Its surface is made up of cobbles which are firmly cemented together in the road. Figure 7.3 shows the Belgian paving in picture (a), its DSD in (b) and the Land Rover traversing its stretch in (c) at Gerotek. Becker (2008) noted that the peak at a spatial frequency of 6 cycles/meter in frame (b)

corresponds to a wavelength of 167 mm which is the average length of the cobbles (a) in the direction of vehicle travel.

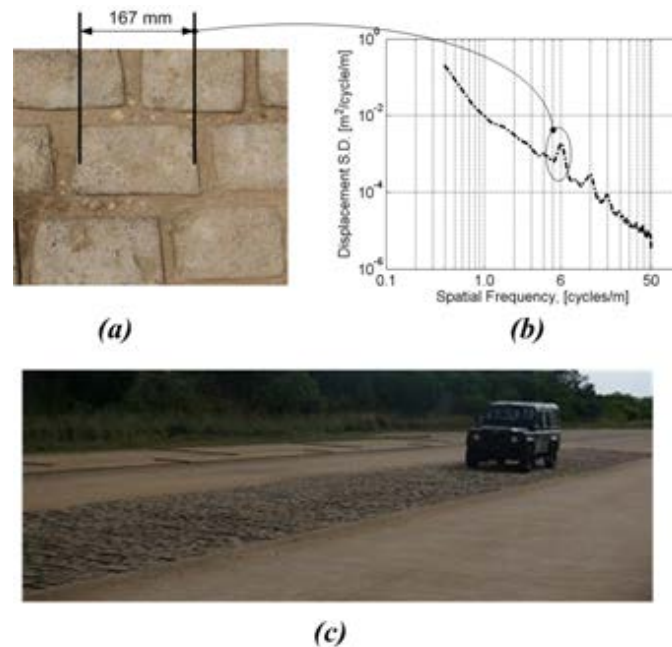


Figure 7.3. Belgian paving (a) in picture (b) Displacement Spectral Density plot (c) vehicle on the Belgian paving.

7.4 Generating the training data

The training data is generated by applying slightly adjusted road profiles to the numerical model of equation (7.9) in Section 7.2.4. Ideally, the training data should be as representative as possible in order to achieve good network generalization over the testing domain. In this study, the network is to be tested on two geometrically different road surfaces and for computational expediency the alterations are applied to these two profiles.

For the bumps, the geometrical shapes remain trapezoidal, while the heights (h) and average lengths (x) shown in Figure 7.4 are resized by a factor of 0.1 and the spacing between the bumps as shown in layouts (c) and (d) shown in Figure 7.2 are reduced from 10.4 m to 8 m. Thus the dimensions used in the training algorithm are $(x \pm 0.1x) m$ and $(h \pm 0.1h) m$ for the average widths and heights respectively. This only yielded geometrically similar trapezoidal bumps with equal corresponding angles, which of course rather simplified the test for network's generalization capabilities over different bump profiles. It is however recommended that in order to achieve better network generalization capabilities over different geometries of bumps, it would be necessary to alter the geometrical shapes and their internal angles besides merely rescaling the heights and widths of the bump blocks.

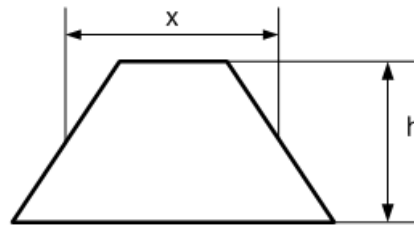


Figure 7.4. Trapezoidal bump showing height (h) and average length (x).

For the Belgian paving, a Gaussian noise with a noise-signal ratio of 0.1 and a 1-m delay are added to the spatial profile. The RMS of the Belgian profile is calculated and a noise level equivalent to 10% of the original spatial profiles' RMS is added to the Belgian paving.

These altered road profiles are applied to the vehicle model under the case scenarios tabulated in Table 7.2. For each case scenario in Table 7.2, a matrix consisting of accelerations calculated on the front right, rear left and rear right sides of the vehicle body are stored. The accelerations are calculated on corresponding positions of the model to those where accelerometers were positioned on the actual Land Rover. Due to some technical difficulty during the experiments, the accelerations on the left front could not be measured. The stored accelerations are processed to remove any constant trends and low-pass filtered with a cut-off frequency of 25 Hz to cover the important frequency range for vehicle ride vibrations. The accelerations (inputs) with their corresponding road profiles (targets) provide the training data for the neural network.

Table 7.2 Summary of road profiles and conditions for generating training data

Case Scenarios	Profile type	Vehicle suspension mode	Vehicle speed (km/h)
1	Symmetric small bump	Handling	14.5
2	Symmetric large-then-small bump	Handling	14.5
3	Asymmetric large-then-small bump	Ride	14.5
4	Asymmetric large-then-small bump	Handling	14.5
5	Belgian paving	Ride	14.5
6	Belgian paving	Handling	14.5
7	Belgian paving	Ride	54
8	Belgian paving	Handling	54

As shown in Table 7.2, there are four case scenarios each under discrete obstacles and under the Belgian paving. It was possible to increase the number of cases for the discrete obstacles according to combinations of suspension settings and vehicle speeds, but that has been intentionally avoided in this study to reduce the resulting computational cost during network training and simulation. In cases of unavailability of *a priori* knowledge of the test demands, it might be necessary to include as many conditions as possible when generating the training data.

However, the use of such simulated data implied that the performance of the neural network in this thesis is largely influenced by the accuracy of the numerical model itself and the quality of the training process. If the training process does not induce much larger errors it might be easier to make some inferences on how the physical parameters of the vehicle system itself impact the final results from the neural network otherwise such inferences may remain hidden and hard to fathom.

7.5 ANN architecture and training

A feed-forward neural network with 50 tan-sigmoid neurons in the hidden layer and two linear neurons in the output layer is used. The network is trained with simulated data from the vehicle model for each of the eight case scenarios presented in Table 7.2. The architecture of the static feed-forward neural network employed in this study is shown in Figure 7.5. This is a departure from the dynamic layer-recurrent NARX used in the previous Chapter (Ngwangwa et al., 2010). The reason for the change is that the NARX network requires feed-forward from true road profiles which unfortunately may not be available in real test situations.

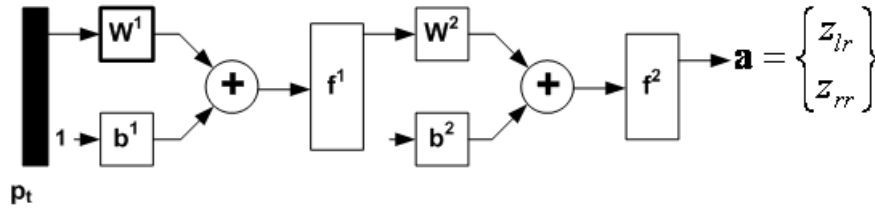


Figure 7.5. Static feed-forward ANN architecture.

The input \mathbf{p}_t contains three sets of sprung mass accelerations with 3964 data points. This two-layer neural network can be mathematically represented by (Mathworks Inc., 2007; Bishop, 1995)

$$\mathbf{a} = \mathbf{f}^2 \left(\mathbf{W}^2 \mathbf{f}^1 \left(\mathbf{W}^1 \mathbf{p}_t + \mathbf{b}^1 \right) + \mathbf{b}^2 \right) \quad (7.10)$$

where $\mathbf{p}_t = \left[\ddot{\mathbf{z}}_{srf}(x) \quad \ddot{\mathbf{z}}_{slr}(x) \quad \ddot{\mathbf{z}}_{srr}(x) \right]^T$ is a matrix of sprung mass simulated accelerations for the training process on right front, left rear and right rear sides of the vehicle body; \mathbf{W}^1 and \mathbf{b}^1 are a 50×3 weighting matrix and a 50-element column bias vector respectively; \mathbf{f}^1 comprises 50 sigmoid



activation function (**tansig.m**), in the hidden layer (Mathworks Inc., 2007) where $\text{tansig}(x) = 2/(1 + \exp(-2x)) - 1$; while \mathbf{W}^2 and \mathbf{b}^2 are 2×50 weight matrix and a 2-element column bias vector respectively being acted on by 2-linear activation functions operating in parallel in \mathbf{f}^2 . The activation in MATLAB function, \mathbf{f}^2 is implemented by function **purelin.m** (Mathworks Inc., 2007): $\text{purelin}(x) = x$. The network output $\mathbf{a} = [z_l \quad z_r]^T$ represent the road profiles on the left and right wheel tracks respectively. If \mathbf{f}^2 is dropped the network output is represented by equation (7.11) so that the network output may be rewritten as

$$\mathbf{a} = \mathbf{W}^2 \mathbf{f}^1 (\mathbf{W}^1 \mathbf{p}_i + \mathbf{b}^1) + \mathbf{b}^2 \quad (7.11)$$

Roll velocities are not used despite their importance as reported by Breytenbach (2009) due to the fact that the numerical model's accuracy in predicting roll velocities is not good. However the application of accelerations from both sides of the vehicle is deemed satisfactory to capture the roll effect. The Land Rover had accelerometers instrumented on the vehicle body only at the rear and above the right front suspension strut. This may impose a limitation to the performance of the technique since a position on the chassis does not experience as much excitation as that on the axle. However, in Chapter 6 where all data were generated from the numerical model it was relatively easy to use axle accelerations, and achieve excellent matches between simulated and actual profiles. This presents a further challenge to the application of the methodology in the present set up and, of course, in the real test situation.

In summary, there are a number of important issues to be noted when evaluating the performance of the methodology in the present application against its former application in previous study (Ngwangwa et al., 2010):

1. A static feed-forward neural network with 50 neurons in the hidden layer and two neurons in the output layer is employed as opposed to the dynamic recurrent NARX network that was employed in the previous Chapter. This is because the NARX network required a feed-forward of true road profiles which are not available in the current application.
2. The network is trained with simulated data generated from altered versions of the real measured road profiles and later tested on measured data. This is done to avoid using the same data in both training and testing the neural network; hence providing a real test to the methodology.
3. The acceleration data used in this investigation are calculated (training data) and measured (testing data) on the vehicle body and not on the vehicle axle. This might provide a limitation on the quality of correlation between simulated and actual road profiles due to the suspension's isolation characteristic.
4. The network is trained with data at two different speeds and two different suspension settings.
5. The neural network uses the Levenberg-Marquardt algorithm as its learning function. This requires that weights and biases are continuous functions.



6. The accuracy of the calculated accelerations which are used for training the network is limited by the numerical model's deficiencies in predicting roll velocities.

7.6 Response measurements

All the vehicle tests were conducted by Breytenbach and Els (2011) on Gerotek's suspension track for wheeled vehicles. These tracks are used to perform repeatable and comparative suspension tests under simulated conditions in order to monitor the structural integrity of body structure, body mountings, suspension, axles, steering, chassis and driveline. Besides, the tests assist in determining and monitoring specific properties, e.g. suspension, steering and structural characteristics.

An eDAQ-lite data acquisition system was used as a data logger (given in Chapter 5). The Crossbow $\pm 4g$ tri-axial accelerometers with model number CXL04GP3, were used because of their good response definition at low frequencies, given that important vehicle ride dynamics fall below a frequency of 25 Hz. A sampling frequency of 1 kHz was used with a linear roll-off anti-aliasing filter set at 333 Hz. In the test, several quantities were measured, but for the purpose of this study, only the vertical accelerations and vehicle speeds are of interest. Table 3 shows the parameters of interest. The vehicle speeds were measured by three different methods to ensure repeatability. Right front accelerometers were located on the vehicle strut mount 120 mm away from the right front axle centre line, while the rear accelerometers were mounted 620 mm away from the rear axle centre line. It was practically difficult to mount the accelerometers directly on axles near wheel centres. Some of the surfaces on which the vehicle was tested were very rough and mounting the accelerometers directly on the axles without a firmer attachment would compromise the accelerometers themselves. On the other hand, the firmer attachments would necessitate that the accelerometer mounting blocks be bolted into the axle members which would also compromise its structural integrity. But if conditions permit, it is strongly recommended that the accelerations be measured on the vehicle axles.



Table 7.3 Summary of measured quantities on the Land Rover

Parameter	Transducer
Time	eDAQ-lite built-in
Vehicle speed	VBOX GPS, eDAQ-lite GPS, and Proximity probe measuring drive shaft speed
Left rear (LR) vertical acceleration	Crossbow triaxial-accelerometer ($\pm 4g$) (CXL04GP3)
Right rear (RR) vertical acceleration	Crossbow triaxial-accelerometer ($\pm 4g$) (CXL04GP3)
Right front (RF) vertical acceleration	Crossbow triaxial-accelerometer ($\pm 4g$) (CXL04GP3)

The tests were carried out in random order to reduce the effects of systematic errors in the measurements. The Land Rover was carefully maintained at a constant speed by driving the diesel powered Land Rover against the engine governor in gear (Breytenbach, 2009). Breytenbach (2009) further showed that the tests were repeatable for tests conducted under each case scenario of Table 7.2. Thus, for testing the neural network in this study, it is satisfactory to select only one test run for each test case. The test data are presented in Table 7.4.

Table 7.4 Summary of test data

Test No.	Case Scenario	Test Description	Suspension Setting	Speed (km/h)
27	1	Bump course, layout (a)	Handling	14.5 (low range, 1st gear)
22	2	Bump course, layout (b)	Ride comfort	14.5 (low range, 1st gear)
33	3	Bump course, layout (c)	Ride comfort	14.5 (low range, 1st gear)
34	4	Bump course, layout (c)	Handling	14.5 (low range, 1st gear)
10	5	Belgian paving	Ride comfort	14.5 (low range, 1st gear)
12	6	Belgian paving	Handling	14.5 (low range, 1st gear)
11	7	Belgian paving	Ride comfort	54 (low range, 4th gear)
13	8	Belgian paving	Handling	54 (low range, 4th gear)



Each test in Table 4 belongs to a particular case scenario presented in Table 7.2. It can be noted that no data was available for the higher speed on the bump course, due to the severity of the test bump at 54 km/h.

7.7 Simulations, results and discussions

This Section discusses the performance of the network on the real test data. The use of model simulated data for training the network have significant advantages during research and development where models are often used to enhance understanding of underlying complex physical phenomenon and/or optimize system behaviour and performance. In the present case however, simulated data have been used to avoid the problem of having to use same data for training and testing or validating the neural network. In Sections 7.7.1 and 7.7.2 the performance of the network over discrete obstacles and Belgian paving respectively, is presented.

7.7.1 Discrete obstacles

The discrete obstacles are evaluated according to the performance of the methodology over the four bump case scenarios (Table 7.2 and Table 7.4), namely: symmetric small bumps in handling mode at vehicle speed of 14.5 km/h; symmetric large-then-small bumps in ride comfort mode at 14.5 km/h; asymmetric large-then-small bumps in ride comfort mode at 14.5 km/h; and asymmetric large-then-small bumps in handling mode at 14.5 km/h. The bump tests were all carried out at low vehicle speed, to avoid the excessive tire dynamics that may accompany such speeds upon traversing a hard bump such as the one under study. There was also the difficulty of maintaining the vehicle speed when traversing the bump.

For the symmetric small bump, no data is available for the vehicle suspension in ride comfort mode. It has been mentioned previously, that the second bump test scenario provides an ideal situation for testing the vehicle's pitch behaviour. This is true if the distance between the front and rear bumps does not coincide with the vehicle's wheelbase. The third and fourth bump test scenarios excite the vehicle's roll motion and it was carried out for both ride comfort and handling modes. This scenario provides the toughest test in this study since the numerical model does not accurately simulate the vehicle roll dynamics due to its inability to accurately capture the friction in the suspension (Breytenbach, 2009).

The actual bump geometries and layouts are given in Figure 7.2. The symmetric small bump layout is shown in Figure 7.2(b) and has a height of 0.1 m. The neural network is simulated with the accelerations measured in Test No. 27 in Table 7.4 to observe its outputs. The reconstructed profiles are shown in Figure 7.6(a) and (b). The solid lines represent the actual profiles while the dashed lines show the reconstructed profiles. The neural network correctly reconstructs the bumps up to a height of



about 0.08 m after which the simulated profiles exhibit some waviness. The waviness also occurs around the base where the bumps start and end. This network behaviour at and around corners can be expected due to the difficulties associated with numerical integration at and in the neighbourhood of such discontinuities. The cause for the differences in the quality of the reconstructed bump between the left wheel track in Figure 7.6(a) and right wheel track in Figure 7.6(b) is not clear at this stage. The minor attenuation in the amplitudes may have been caused by the tire enveloping effect that is not properly taken into consideration in simulation model by the point-follower tire model as well as the averaging effect of the numerical filter applied to the simulated profile. Besides, the point-follower tire model with a parallel linear spring and damper does not accurately represent the tire spring and its non-linear effect wheel hop as well as changes in tire pressure, thereby contributing to some error on the steady-state part of the discrete bumps. Furthermore, this tire model does not accurately simulate the change in tire/ground contact point. However, these shortcomings are tire model-dependent which can be improved if more comprehensive models were used.

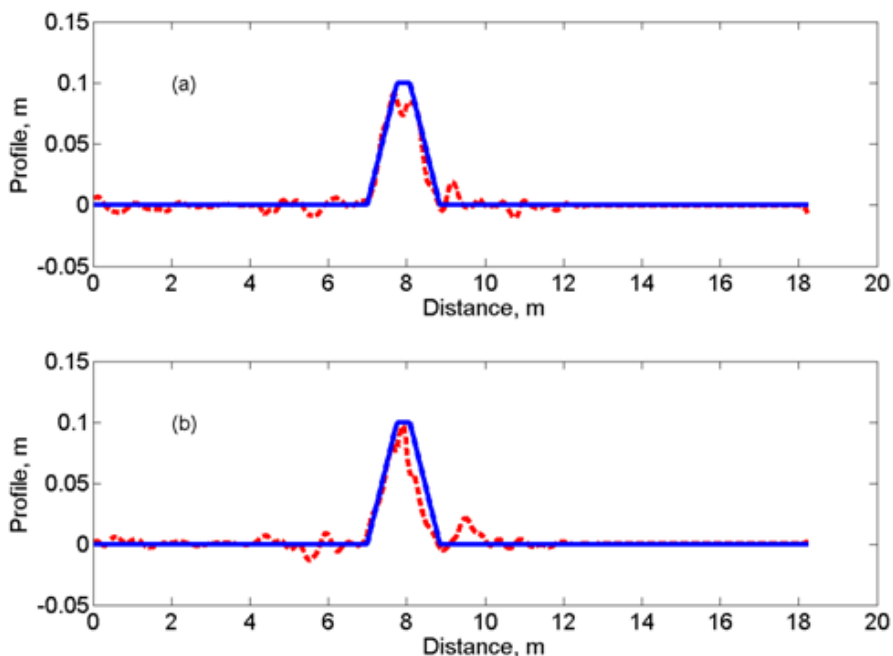


Figure 7.6. Simulated bumps (dashed) correlated with actual bumps (solid) on the left hand track (a) and right hand track (b) for Case Scenario 1 for Test No 27 in Table 4 (Handling mode at 14.5 km/h).

Figure 7.7 shows the results of the network simulations over symmetric large-and-then-small bumps at the vehicle speed of 14.5 km/h when the vehicle suspension was set to ride comfort mode as in Case Scenario 2 Test No. 22 in Table 7.4. The left hand wheel track is represented by Figure 7.7(a) and the right hand wheel track is shown in Figure 7.7(b). The reconstructed bumps correlate very closely with the actual profiles on both sides of the wheel tracks. Moderate undulations occur at the corners especially at the bases of the bumps. However, the quality of the reconstructed profiles in terms of heights and general curve fitting is much superior to the Case Scenario 1. Breytenbach (2009) noted that the larger number of uncertainties associated with the harder suspension setting in



handling mode, affected more adversely the quality of the correlation between simulated and actual vehicle responses than in the more compliant ride comfort mode.

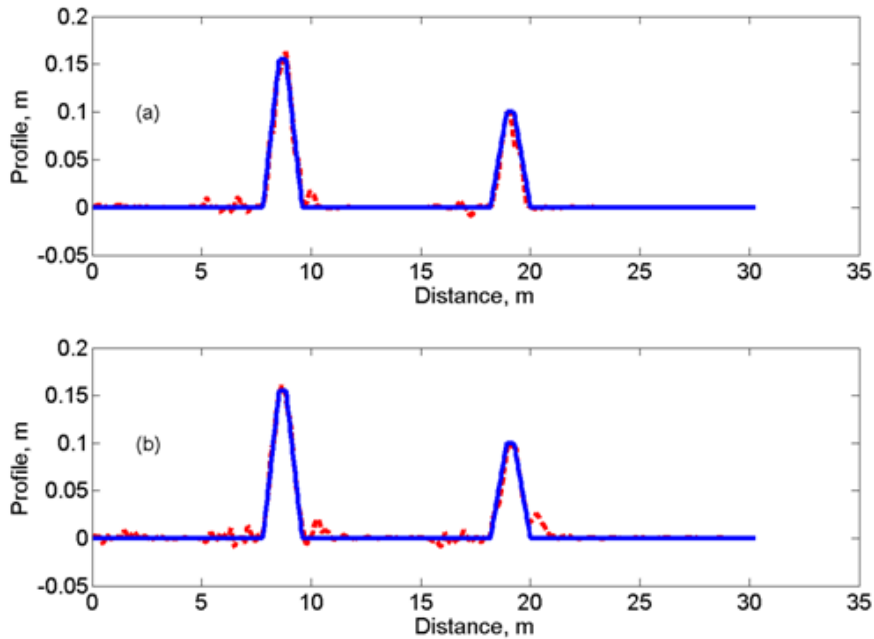


Figure 7.7. Simulated bumps (dashed) correlated with actual bumps (solid) on the left hand track (a) and right hand track (b) for Case Scenario 2 Test No. 22 in Table 4 (Ride comfort mode at 14.5 km/h).

The performance of the network over asymmetric large-and-then-small bumps are presented in Figure 7.8 for Case Scenario 3 and in Figure 7.9 for Case Scenario 4. The reconstructed bumps on the left hand wheel track exhibit better correlation than on the right hand wheel track for both test cases. Though this has been observed in the previous test cases, it is clearly more magnified in these two cases by the fact that the underlying numerical model fails to model roll motion with sufficient accuracy as to allow its attendant neural network to reproduce any form of motion with acceptable level of accuracy, in which roll plays a significant part. Breytenbach (2009) noted that the model over-predicted roll velocity and investigated the causes of this error through sensitivity analyses. It was concluded that roll stiffness in suspension bushings, which was not considered in the model, might be culpable for the over-prediction of roll velocities. Frame flexibility was also considered to be a contributor but not necessarily the main cause. It is, however, not clear at this stage why these effects seem to affect the ride comfort mode case in Figure 7.8 more than handling mode case in Figure 7.9, especially on the right hand wheel tracks. The reconstructed bump in ride comfort mode (Figure 7.8(a)) does not suffer from a similar energy spill over at the location corresponding to the location of the right hand bump as is observed in Figure 7.9(a) for handling mode. This might be due to relatively more accurate modelling in the ride comfort mode as compared to more demanding handling mode.

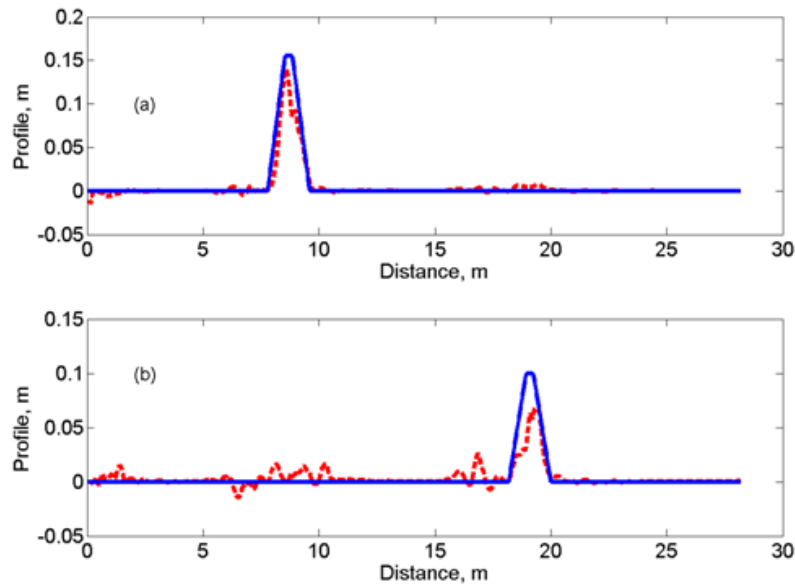


Figure 7.8. Simulated bumps (dashed) correlated with actual bumps (solid) on the left hand track (a) and right hand track (b) for Case Scenario 3 Test No. 33 in Table 7.4 (Ride comfort mode at 14.5 km/h).

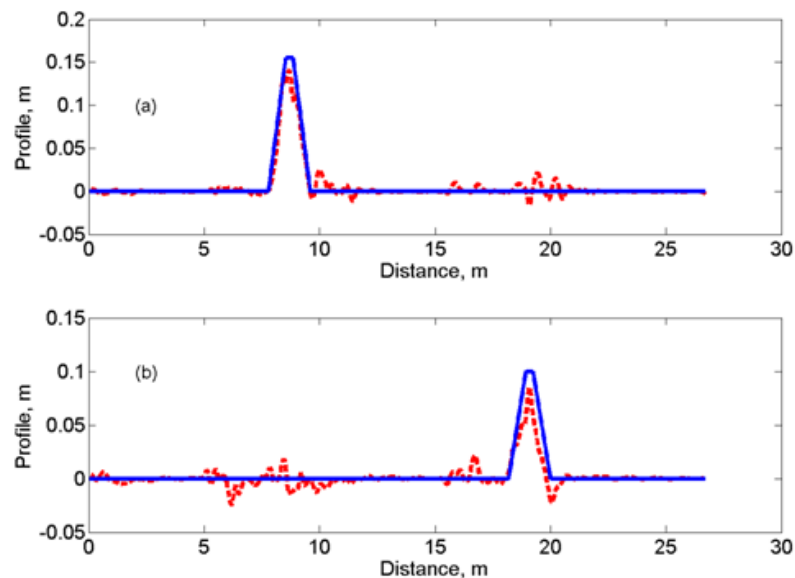


Figure 7.9. Simulated bumps (dashed) correlated with actual bumps (solid) on the left hand track (a) and right hand track (b) for Case Scenario 4 Test No. 34 in Table 7.4 (Handling mode at 14.5 km/h).

In this Section the performance of the network over different layouts of the discrete obstacles have been presented and discussed. It is observed that the quality of the correlation between the reconstructed and actual bumps is generally better in ride comfort mode than in handling mode. This



is a property that the neural network adapted from the numerical model that is used to generate the training data. The causes for the differences in the quality of correlations between the left hand and right hand wheel tracks are not clear. There could be a number of factors that may influence this anomaly including vehicle manoeuvre over the bump, inaccuracies in representation of measurement points and slight suspension strut manufacturing disparities.

There are also some errors observed on the steady-state part of the bumps. This may be due to the tyre being modelled as a point follower model with a parallel linear spring and damper. The non-rolling dynamic tire stiffness and damping properties were experimentally determined by using a pendulum type tire tester at 200 kPa inflation pressure for a range of vertical loads over tires up to 1600 kg. As a result, a linearized tire stiffness of 250 kN/m and linearized damping constant of 1.2 kN.s/m were recommended for vertical loads of tires below 700 kg. Thus the tire spring and its non-linear effect wheel hop plus the changes in tire pressure during the actual test are not properly represented by the model thereby contributing to some error on the steady part on discrete bumps. Also the change in tire/ground contact point is not properly simulated by the model.

However the fact that the methodology is still able to yield very good results in the presence of all these uncertainties makes it more practically feasible. The methodology is able to locate the bumps where they occur and is able to estimate the sizes of bumps with maximum discrepancies as shown in Figures 7.8(b) and 7.9(b) that might be acceptable for road condition monitoring. It is expected that if the measured data were to be used for training the network some of these offending factors might be eliminated and the methodology would presumably yield better results.

7.7.2 Belgian paving

In the case of Belgian paving tests, it makes more practical sense to correlate their statistical values and DSDs rather than the raw profiles themselves. Heyns, Heyns and De Villiers (2012) noted that data scarcity may often render profile estimation difficult and as such they proposed the use of a metric calculated from vehicle accelerations. Specifically, Heyns, et al. (2012) developed a technique for cost-effective condition monitoring of mine haul roads based on speed normalised response type road roughness measuring systems (RTRRMS) by using a Bayesian framework. Basically they used a Bayesian framework to extract speed normalised RMS acceleration from measured vehicle vibration data. The DSDs are used in this study due to their ability to be interpreted as a measure of road roughness (ISO 8608:1995E, 1995; Andr n, 2006). It is also relatively easy to translate DSDs into IRIs which themselves are popular measures of road roughness. A similar approach was employed by Kang, Lee and Goo (2007) except that they used the DSDs computed from profilometer measured data which might be quite demanding. The approach proposed in this study takes advantage of the current trends in vehicle information systems where vehicles are increasingly being mounted with sensors for optimization of suspension systems.

Figure 7.10 shows simulated raw profiles plotted over the actual measured profiles. The plots show a limited road Section, for clarity purposes only, since the actual training and testing processes were carried out over the entire 100 m length.

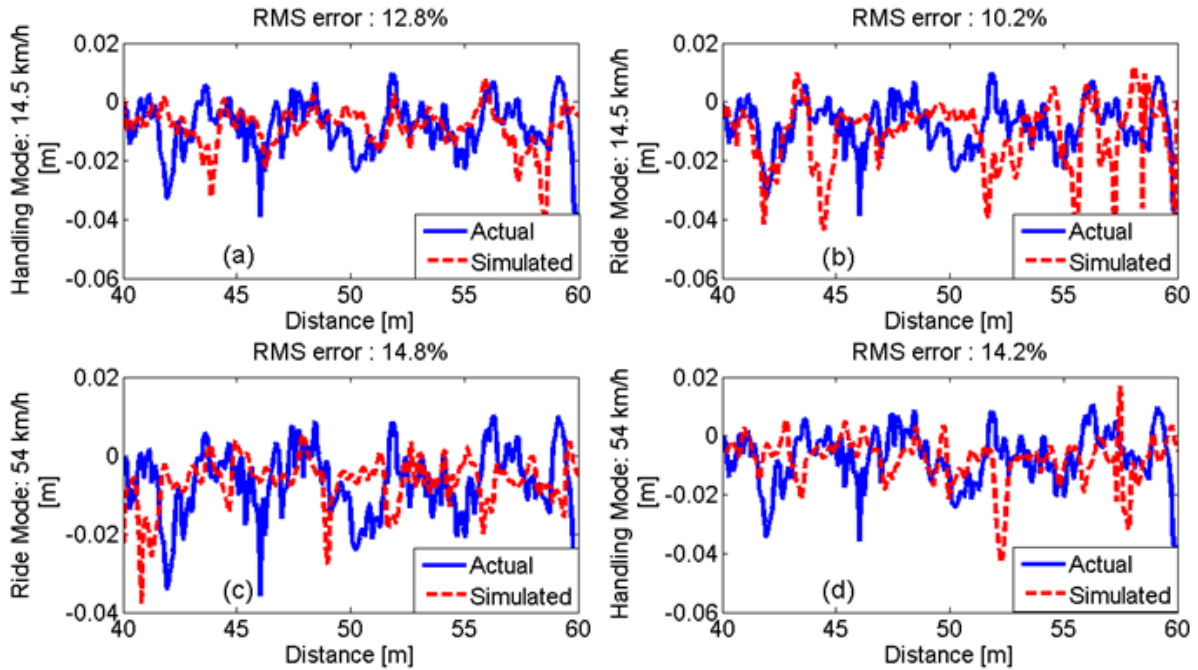


Figure 7.10. Simulated profiles (dashed) correlated with Actual profiles (solid) for near-optimal and stable training process.

In this analysis, the RMS error percentage was used to demonstrate correlation between the simulated and the actual road profiles. The RMS error percentages are indicated on top of each plot in Figure 7.10. The errors are calculated by

$$\varepsilon_{rms} = \frac{|y_{rms} - a_{rms}|}{y_{rms}} \times 100 \% \quad (7.12)$$

where $y_{rms} = \sqrt{\sum_{i=1}^N (y(x)^2) / N}$; $a_{rms} = \sqrt{\sum_{i=1}^M (a(x)^2) / M}$; $y(x)$ and $a(x)$ are the measured and simulated profile elevations, respectively, at a distance of x metres along the road; N is the total number of sample points in the measured profile; and M is the total number of sample points in the simulated profile.

The errors in RMS estimations in Figure 7.10 are below 15 % with ride comfort mode exhibiting both the best correlation of 10.2 % RMS error at the lower vehicle speed of 14.5km/h (Figure 7.10(b)), and the worst correlation of 14.8 % RMS error at the higher speed of 54 km/h (Figure 7.10(c)). The performance of the ANN in handling mode lies between these two extremes at 12.8 % for lower vehicle speed (Figure 7.10(a)) and at 14.2 % for higher vehicle speed (Figure 7.10(d)). It is observed that the extents of the RMS errors are much higher than those obtained from the numerical model itself, and also, that the relative performance of the ANN model over the Belgian paving did not



always agree with the accuracies as obtained from the numerical model. For example there is a ripple effect on the RMS error percentages and the higher vehicle speed does not always yield better correlations than the lower vehicle speed as was the case with the numerical model.

This might be attributed to the rather perverse interplay between the training data and the numerical model itself. The target data for training the ANN was measured from a 100 m long Belgian paving and that could not be long enough to contain all the necessary low frequency information. Furthermore, the numerical model that was used to generate the input data (vehicle body accelerations) had two main deficiencies: first, the implementation of point-follower tire model made the model prone to inducing appreciable levels of higher frequency noise in the calculated accelerations, and secondly, the disregard of wheel-base filtering overestimated spectral energies at some of the frequencies. The combined effects of these two factors negatively affected the stability and convergence abilities of the ANN.

Figure 7.11 shows the DSDs for the ANN simulated and actual (measured) road profiles plotted over each other for each combination of vehicle suspension mode and speed on a road roughness scale. The results show that the DSDs are very well correlated above a wavenumber of 0.2 cycles/m (corresponding to a wavelength of 5 m). The DSDs on the lower frequency portion are relatively poor due to insufficient low frequency data content in the 100 m long road. The roads are consistently classified as D in the mid-frequency range between 0.2 to 3.0 cycles/m and C to A in the higher frequency ranges. The exception exists at and around the wavelengths corresponding to the cobble width of 0.17 m where the DSDs are amplified.

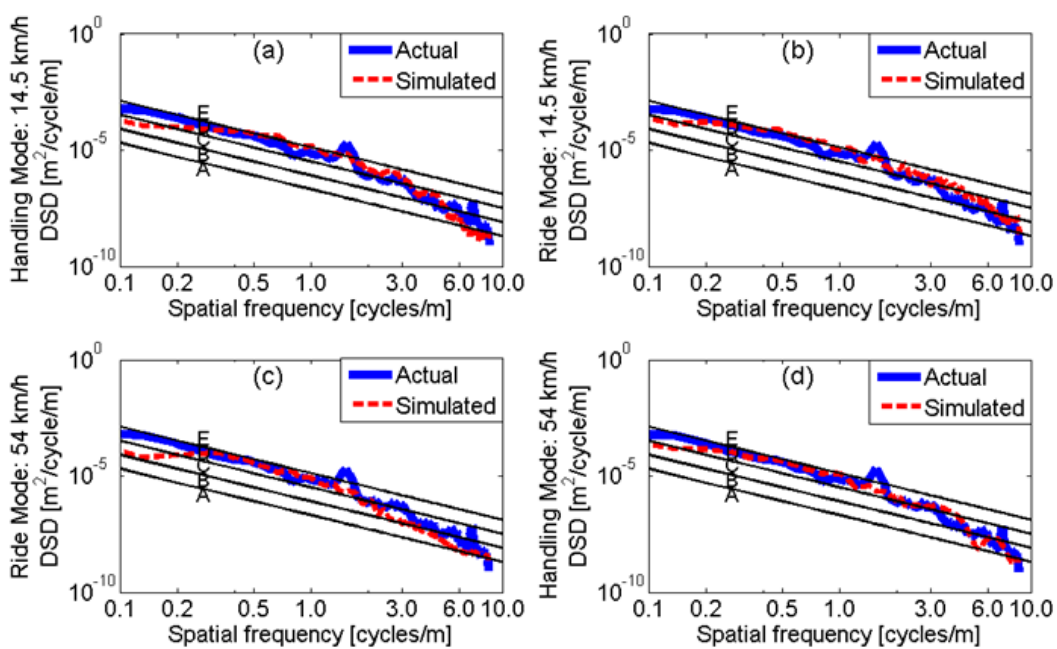


Figure 7.11. Simulated DSDs (dashed) correlated with Actual DSDs (solid) for near-optimal and stable training process.



The ANN training and simulation processes were carried out on a Dell XPS L502X laptop with Intel Core i7 CPU at 2.2GHz having a random access memory (RAM) of 8GB and running on a 64-bit Windows. It has a hard drive disk (HDD) storage of 750 GB. The total duration in training and simulation of the ANN was measured as 65s out of which 2s was for simulation. Thus for a pre-trained ANN the short duration in simulation can benefit an online condition monitoring system.

7.8 Conclusions

In this Chapter a methodology for estimating road profiles and roughness classes has been demonstrated by simulating an artificial neural network on measured data. The neural network was trained with data calculated by a numerical model of the Land Rover. This was deemed necessary due to lack of sufficient data to be used for both training and testing the neural network. The training data was obtained by applying altered versions of the real road profiles to the numerical model. This presented an opportunity to test the network with data that were different from the data that was used to train it and hence test its generalization capabilities.

Eight different test cases were drawn from a combination of road profiles, vehicle suspension modes and speeds. The road profiles comprised discrete obstacles, from which three different layouts were constructed, and the Belgian paving. The vehicle suspension was adjusted between ride comfort and handling modes while the vehicle was either driven at 14.5 km/h or 54 km/h. For all bump layouts, the vehicle speed was kept at the lower speed of 14.5 km/h.

The methodology has been applied differently to the two road profiles without necessarily departing from the central theme of the study of providing a means for road condition monitoring. The ANN reconstructed profiles were correlated with actual measured profiles for both discrete obstacles and the Belgian paving. However in the discrete obstacles, only overlay plots of the simulated and actual bump profiles were found satisfactory, while for the Belgian paving, overlay plots as well as RMS error percentages were used to provide a measure of accuracy in correlation. The overlay plots of discrete obstacles show very good correlations between ANN simulated bumps and actual bumps for all cases under investigation. The results showed superior quality in the reconstructed bumps for ride comfort as compared to handling mode. However, the overall network performance over bumps was observed to be marred by the poor roll motion estimation capabilities of the numerical model as well as its use of the point-follower tire model. This shortcoming shrinks the validity area but it is only model-dependent. If, on the other hand, a more comprehensive vehicle model is used, the validity area is effectively enlarged.

The Belgian paving however presented its own peculiar challenges due to combined effects caused by the deficiencies in the training data and the numerical model that was used to generate the data. Thus ANN training was carried out in the absence of sufficient training data. As a result, three typically different cases were identified from the simulated results (see Appendix B): first case, the ANN could achieve near-optimal state where the RMS error percentages were all below 15%; second case, the ANN could only manage to converge to an overly sub-optimal state where the RMS error percentages



were largely between 15 and 25 %; and third case, the ANN could be trapped in an unstable region and the resulting RMS error percentages were largely above 25 %.

For the first case (with stable and near-optimal training process), the simulated profiles matched the actual profiles very well and the evaluated RMS error limit of 15 % lies within the error margin of 20 % as per the findings of Chapter 6. The simulated profile DSDs showed perfect correlation with the actual profile DSDs for all frequencies above 0.2 cycles/m. In the second case, the ANN would either yield very good correlations over the entire range of frequencies above 0.2 cycles/m as in the first case, or it would yield poor estimations of the spectral densities at frequencies higher than 1.8 cycles/m. Besides converging to sub-optimal weights that might yield underestimation of some of the outputs; neural network generalization might have enforced high frequency filtering on the simulated profiles. The third case was characterized by spiky spectral densities at the higher frequency end particularly beyond 2 cycles/m.

However, these two problems are not viewed as absolutely debilitating to the application of the methodology to road condition monitoring. The findings in this study show that the methodology has a lot of potential. Firstly, the methodology is able to locate the bumps where they occur and is able to estimate the sizes of bumps with practically acceptable maximum discrepancies. Secondly, the consistency with which accurate estimations of the DSDs for the Belgian paving, have been made in the frequency range between 0.2 cycles/m and 1.8 cycles/m, for all three cases, is an extremely encouraging result. Though such stable frequency ranges may vary from one vehicle-road test scenario to another, it might yet be useful to identify such frequency ranges for any given specific test scenario so that practically useful road roughness estimations could be made irrespective of the state of the training process. In this study, the errors would have been alleviated if the attendant numerical model was made more accurate by replacing the point follower tire model with a more accurate tire model. Then low-frequency data scarcity due to short Belgian paving would have been dealt with by generating the training data from much longer road lengths (600 – 1000 m) that closely resembled the Belgian paving.



Chapter 8

Field tests on mine vehicles

8.1 Introduction

Current management techniques for the maintenance of mine haul roads, such as ad hoc blading, scheduled blading and maintenance management systems, have shortcomings in complex mining environments (Thompson and Visser, 2003; Thompson et al., 2003; Hugo et al., 2008). Too little road maintenance leads to excessive cost in the operation and maintenance of vehicles, whereas excessive road maintenance leads to greater cost but does little to reduce the cost of the operation and maintenance of vehicles (Hugo et al., 2008). Haul roads are subjected to variable traffic volumes, vehicle types and payloads. The standard systems for the management of haul-road maintenance are, in general, poorly suited to dealing with these complex and dynamic environments. Model-based road classification techniques have a greater potential to be very powerful, especially in an environment where details of the road characteristics are required (Heyns et al., 2012).

This case study completes a series of investigations on the application of the proposed ANN-based road damage identification methodology. The details of the findings from the other case studies are presented in Chapters 6 and 7, and both show strong agreement that the methodology is feasible though it still remains to be tested in more practical situations. Therefore in this Chapter, the methodology is tested on two mine vehicles operating in their normal environment with very little control on their operation and where the roads were poorly constructed, to replicate real damage scenarios, and they are not as accurately profiled as in the previous cases. Besides, no numerical models of the vehicles are available; thus training data can only be selected from the test data. The greatest challenge with this selection was the scarcity of data due to the many unknown random operating conditions. A substantial number of tests and measurements would have had to be performed if the acquired data were to be completely representative of the vehicle's real dynamic behaviour under every available operating condition. Furthermore, though vehicle control is a key input into the vehicle-road interaction system, there is no simple way to measure it quantitatively. The inconsistencies in the vehicle control over different road conditions introduced drastic variations in the quality of the measured data, in this way affecting the representativeness of the underlying vehicle dynamics. However, these challenges are what have made the present investigation unique.

Section 8.2 details the materials and procedures that have been used and followed in this investigation. It describes the vehicles and the roads in detail: the ways in which the vehicles were instrumented and the accelerations were measured and processed; and how the neural networks were created and trained. The organisation of the measured data is presented in Section 8.3. Section 8.4 presents and discusses the results and Section 8.5 draws the conclusions and recommendations.



8.2 Materials and methods

The section is aimed at describing the vehicles, roads, measurement tools and procedures that were used in the investigation. In this study, the choice of the vehicles and roads was based on the availability of the vehicles and roads for the duration of the planned duration; the operational representativeness of the actual vehicle-road system; and the possibility of introducing artificial defects to the roads with minimal interference in the hauling operations. The vehicles were instrumented and the roads profiled in preparation for the measurement of the accelerations. After measuring the vehicle accelerations, the data were pre-processed for ANN training and simulation. Finally, the classes of road roughness were determined from the ANN-simulated road profiles.

8.2.1 The test vehicles and roads

The ultra-heavy mine haul truck (Figure 8.1(a)) is commonly used in surface hauling and the small utility underground vehicle (Figure 8.1(b)) in underground operations. The haul truck's wheel-base measures about 5.7 m with a front height of 6.2 m and a tire diameter of 3.75 m (Figure 8.1(a)). It has four nitrogen-over-oil (hydro-pneumatic) suspension struts, each mounted above each wheel axle in front and linked at the rear by a trailing arm. This particular haul truck has a load carrying capacity of 300 tonnes.

The small utility underground vehicle has a gross vehicle mass (GVM) of about 0.5 tonnes with a tire diameter of 0.7 m. It has a wheel-base of 2.85 m and a vehicle width of 1.75 m (Figure 8.1(b)). It is below the average human height in order to allow for easy underground operation. It has a very low centre of gravity and does not have the bounce-pitch coupling problems that the haul truck has.



Figure 8.1. Haul truck (a) and small utility vehicle (b).

The haul truck was tested on a typical haul road where three different forms of artificial defects were constructed. The defects were constructed by using a road grader. The small utility vehicle was tested on four different types of roads: a paved track with speed bumps, a gravel track with depressions, a



coal-ash compacted smooth track with a brick bump and an underground track. All these test roads were selected from the existing road networks at the mine sites.

8.2.2 Vehicle instrumentation

The two vehicles were instrumented differently. The haul truck had two Crossbow tri-axial accelerometers, each mounted on either side of the front spindle. They were powered by a 12-V battery which was affixed to the front bumper of the truck. An eDAQ-lite system was used as a data logger and speed was calculated from the measured time and distances. A Panasonic Toughbook computer was used for data display and monitoring. The data was electronically transferred through the Toughbook computer to external storage. The infra-red position probe was used to identify the positions of the yellow reflectors that had been pasted along the road to mark the positions of the defects.

The small utility underground vehicle was instrumented with three Crossbow tri-axial accelerometers, two on the rear axle and the third on the front axle. A position probe was mounted on the front bumper to pick up the positions of the yellow reflectors positioned along the road. The vehicle speed was calculated from the measured rotational speed of the engine shaft and the given wheel-engine speed ratio. The engine speed was measured by a shaft encoder. An eDAQ-lite data logger and a Toughbook computer were also used for data acquisition.

8.2.3 Road preparation

In the haul road tests, three different defects were constructed along the road. The position of each defect was accurately marked by pasting a single line of yellow strips at the start and double lines of yellow strips at the end of the defect. Then the truck was driven in the opposite direction to traverse the defects in a return mode. Though the measurements were taken over the entire road length, the data for training and simulation was extracted from each defect length only.

In the tests of the small utility underground vehicle, four different tracks were used: the first, was a 24-m long gravel track; the second, was a 24-m paved track; the third was a 10-m coal-ash compacted smooth track where a brick bump was placed in the middle; and the fourth, was the underground track. This Chapter reports the results obtained for these two vehicles over the first two tracks only. Heyns et al. (2012; 2012) report the findings for the same vehicles over the brick bump and underground roads. Yellow strips were similarly used to mark the depressions along the gravel Section and the two speed bumps on the paved road. The two ends of the test road Sections were also clearly marked by the yellow strips.

8.2.4 Road profiling and roughness classification

In this study, the roads were not profiled by using standard road profilometers. Rules, poles and strings were used for measuring the profiles in both tests. A string was tied between two poles erected at the ends of a road Section. The heights of the string from the ground were measured with a metre rule at different spacings, depending on the condition of the road surface. In both tests, the roads were sampled at 1 m except over the bumps and other important features where they were sampled at 0.25 m. Figure 8.2 shows the simplified profiling set-up that was employed in this study.

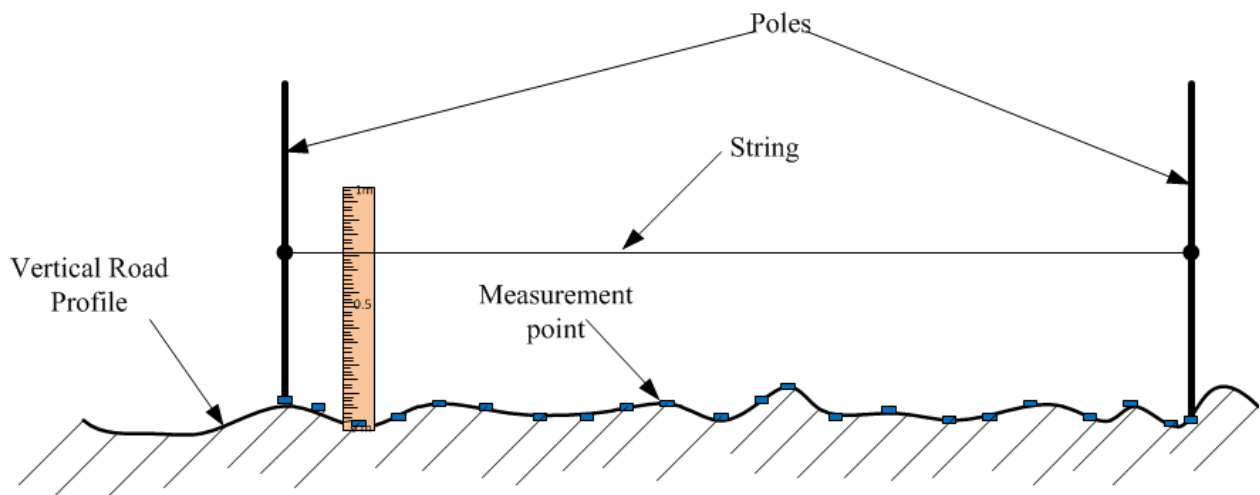


Figure 8.2. Simple road profile measurement.

The string in Figure 8.2 was regarded as a single variable. The height of the string at one pole was benchmarked as a reference point, hence that height was subtracted from each of the string heights at the other points along the road. The resulting values represent the deviations from the ground point at the selected pole. In order to determine the road profile, these deviations were simply sign-reversed. The profile data was stored as a data file that would be accessible in MATLAB for further processing. The other data that was recorded and saved, included the measurement points along the road and the positions of the yellow reflectors. This procedure has the main advantage of being simple and easy to perform without the need for specialised staff and equipment, and it is in line with most practices in the industry. However, it is understood by the investigators that it was performed at the expense of the accuracy of the measured profile.

Figures 8.3 and 8.4 show averaged road profiles between the left and right wheel tracks, and their corresponding DSDs for the test on the haul truck and the test on the small utility vehicle. In the haul truck test, the DSD plots show that the waviness in the roads was dominated by spatial frequencies from 0.1 to 1 cycles/m corresponding to wavelengths from 1 to 10 m. Accordingly, when the truck was travelling at the lowest nominal speed of 8 km/h, the road was capable of generating frequencies of excitation between 0.22 and 2.2 Hz, and at the highest nominal speed of 34 km/h, the excitation frequencies ranged from 0.94 to 9.4 Hz.

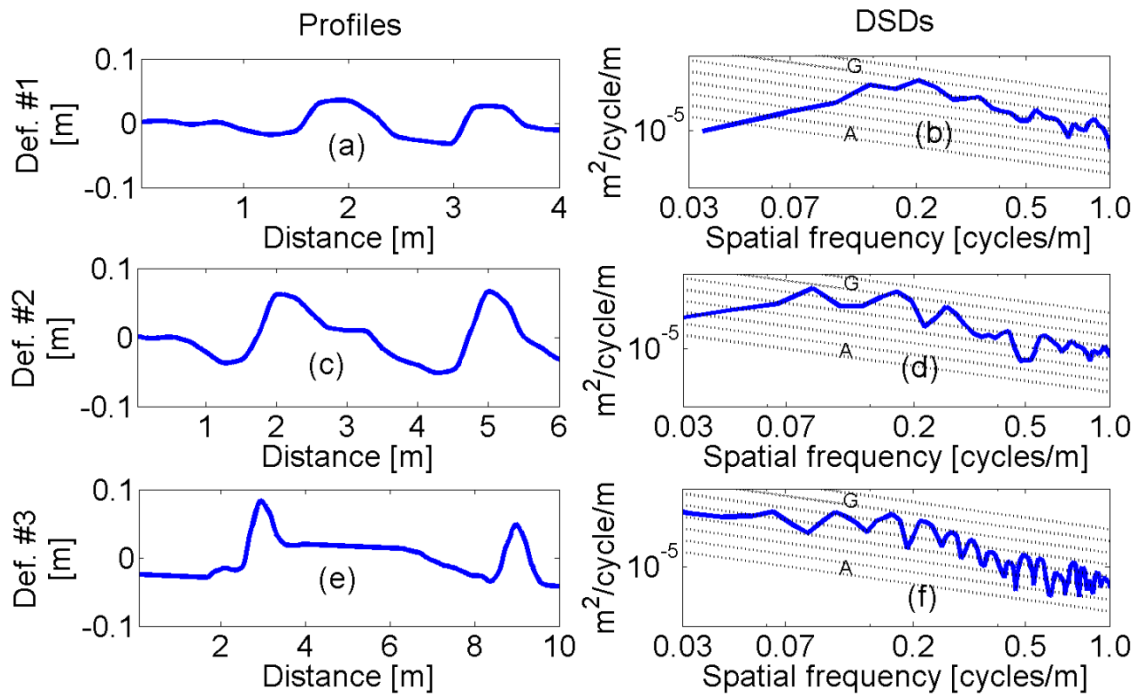


Figure 8.3. Averaged road profiles and PSDs in (a) and (b) for Defect 1, in (c) and (d) for Defect 2 and in (e) and (f) for Defect 3 during the mine haul truck test.

In the small utility vehicle test, Figure 8.4 shows the averaged gravel road profile in (a), its DSD in (b), the averaged paved road profile in (c) and its DSD in (d). By inspection, these tracks present a wider range of excitation frequencies. The depressions over the gravel track were due to surface degradation though the bumps over the paved track were actually speed bumps which had been constructed to check the vehicle speeds near the fuel pumps. Therefore these tracks have spatial frequencies in the entire range, as shown in Figure 8.4(b) and (d). When the vehicle was travelling at the nominal lowest speed of 6 km/h and the highest nominal speed of 12 km/h, the corresponding temporal frequency ranges were: 0.25 to 8.3 Hz and 0.5 to 16.7 Hz. Once again, the lower vehicle speeds did not sufficiently cover the unsprung mass frequencies which were expected to lie between 10 and 15 Hz.

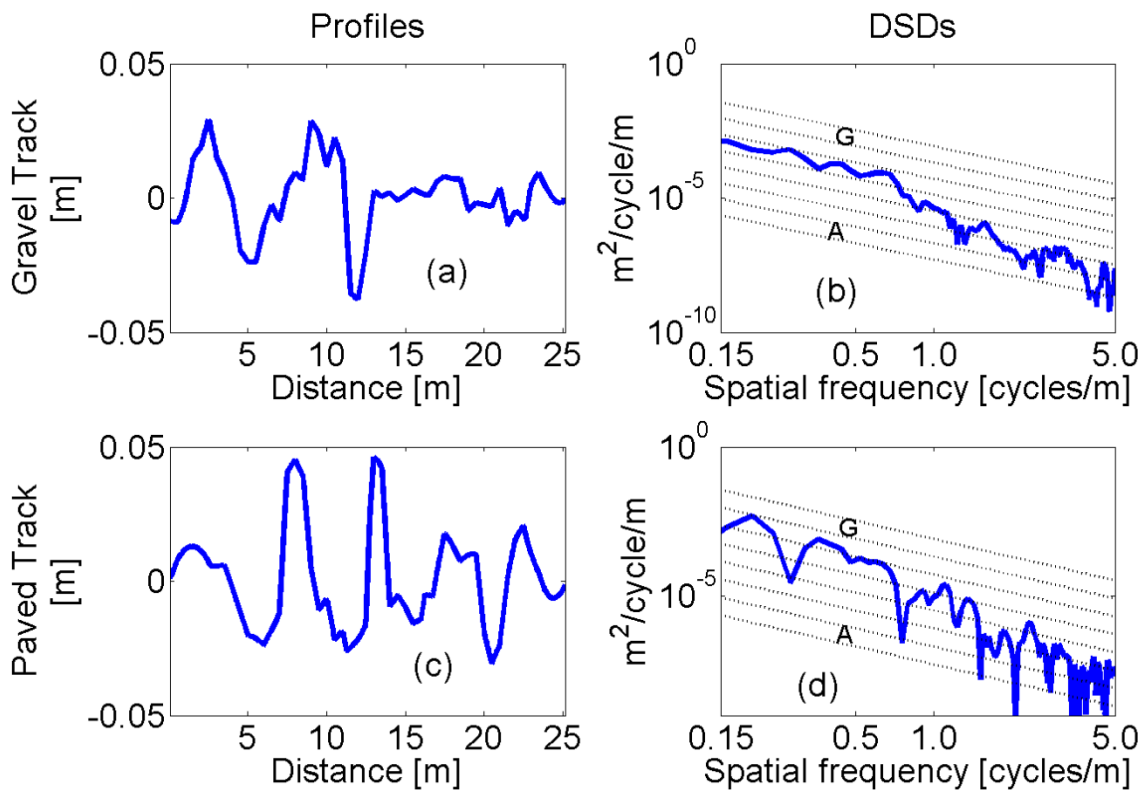


Figure 8.4. Road profiles and PSDs for the gravel road in (a) and (b); and for the paved road in (c) and (d).

8.2.5 Vehicle operation and response data measurements

For each of the two different tests, one driver was chosen to operate the vehicle throughout the testing period, to ensure some level of consistency in driving behaviour. The haul truck was tested in both an unloaded and a fully-loaded state. It was driven to and fro over the defects, in this way creating three more defects with reversed geometries. By contrast, the small utility vehicle was driven in one direction only, in a cyclic manner. The intention was that the haul truck should be driven at the nominal speeds of 8, 12, 18, 28 and 34 km/h and that the speeds should remain constant throughout any particular test cycle. In the same way, the small utility vehicle was intended to be driven at the nominal speeds of 6, 8, 10 and 12 km/h. However, subsequent analyses showed that there had been lapses during a particular test run where speed fluctuations were observed, and also across different test runs where it was observed that there had been no adherence to the given nominal speeds. For the haul truck, it was noted that any such loss of control during a particular test run led to the excitation of the pitch motion when the truck was travelling at speeds lower than 20 km/h. The haul truck showed no significant effects at the higher nominal speeds.

In both tests, the accelerations were measured on the axles as stated in Section 8.2.2. The accelerations were captured at a sampling frequency of 400 Hz with an upper cut-off point at a



frequency of 250 Hz. The data was recorded on a SoMat eDAQ-lite data-acquisition system connected to the Toughbook computer used for monitoring the captured data. The integrity of the dc-coupled Crossbow accelerometers was verified on a high-frequency actuator and found to be highly accurate over the frequency range of interest (Heyns et al., 2012).

8.2.6 Data pre-processing

The captured data were processed in MATLAB. As shown, the vehicle axle accelerations and road profile data were measured at different times, by different processes and at different sampling frequencies. This implies that there was a need for data alignment, filtering and resampling before the data could be used for ANN training. The positions of the yellow reflectors were used for aligning the eDAQ-lite data with the road profile data. This was done manually in MATLAB by using the `ginput.m` function (Mathworks Inc., 2007). This function gathers the co-ordinates of a point on a graph through a mouse input. Therefore the eDAQ-lite data corresponding to the stored road profile data was extracted by picking up the points on the eDAQ-lite data that marked the yellow reflectors at the START and END of each type of track. An allowance was made for the relative distances between the positions of the infra-red sensor and the response pick-up points in order to align the START and END points perfectly. Owing to the inevitable changes in vehicle travelling speeds, it was also necessary to consider aligning the other yellow reflectors near the defects.

After alignment, the eDAQ-lite data as well as the road profile data was resampled at 100 Hz and constant and linear trends in the data were removed. The acceleration data was low-pass-filtered at a cut-off frequency of 6 Hz for the haul truck and 30 Hz for the small utility vehicle. These cut-off frequencies covered the important vehicle frequencies and road wavelengths. However, the data had to be scaled to be applied to the ANN so that the inputs and targets fell within the same range. In both tests, the data was rescaled to the range [-1 1] by using `mapminmax.m` in MATLAB (Mathworks Inc., 2007).

8.2.7 ANN identification and training

The NARX network presented in Chapter 4 is used in this Chapter. In this application, the series-parallel NARX model (`newnarxsp.m`) was only used during training and then converted into a parallel configuration by using the MATLAB function `sp2narx.m` for simulation purposes. The resulting model does not require the actual road profiles to be fed forward during simulation. The ANN for the haul truck test is a 7-5-2 network, which implies that it has seven inputs with five tan-sigmoid neurons in the hidden layer and two linear neurons in the output layer. The seven inputs comprise: accelerations measured on the right-front and left-front axles, the calculated velocities and displacements, and the vehicle travelling speed. The velocities and displacements were further treated for low frequency drift by removing quadratic trends from the displacements and linear trends from the velocities. Ideally, the inclusion of the velocities and displacements as inputs do not affect the quality of the correlation between simulated and actual profiles, but these so-treated displacements



and velocities were observed to give more stability to the training process in terms of number of iterations taken to converge to the targets. This is, however, an observation that requires further investigation and cannot be considered conclusive. In order to cater for different payload scenarios, the accelerations were further rescaled according to the ratio of the unloaded truck mass to approximate the fully loaded truck mass.

A 10-10-2 network was used for the small utility vehicle test, where the ten inputs comprise: accelerations measured on the right-rear, left-rear, and left-front axles; the numerically calculated velocities and displacements (similar to the haul truck case); and the vehicle travelling speed. There is no need for rescaling the accelerations due to payloads, since all the tests performed in the unloaded state. It was found that three delays in the input and feed-forward output were satisfactory for both neural networks. The number of neurons in the hidden layer was obtained through trial-and-error testing whereas the choice of tan-sigmoid activation functions was purely dependent on their superior performance in regression problems (Mathworks Inc., 2007; Bishop, 1995; Hagan and Menhaj, 1994). Since the network was intended to yield a profile for each wheel track, two neurons in the output layer were specified. The Levenberg-Marquardt algorithm described in Chapter 4 was used as a training function. It was chosen because of its superiority in solving regression problems and its computational efficiency, since it avoids the more costly evaluation of the Hessian matrix (Hagan and Menhaj, 1994).

For any test in which the number of permutations of unique test scenarios is R , and the sampling points in each test is N , the input (\mathbf{p}) and target (\mathbf{t}) training data is organised as cell arrays in the form:

$$\begin{aligned} \mathbf{p} &= \{[\mathbf{x}_1(1) \ \cdots \ \mathbf{x}_R(1)] \ \cdots \ [\mathbf{x}_1(N) \ \cdots \ \mathbf{x}_R(N)]\} \\ \mathbf{t} &= \{[\mathbf{y}_1(1) \ \cdots \ \mathbf{y}_R(1)] \ \cdots \ [\mathbf{y}_1(N) \ \cdots \ \mathbf{y}_R(N)]\} \end{aligned} \quad (8.1)$$

where $\mathbf{x}_i(k)$ and $\mathbf{y}_i(k)$ are column vectors containing input and output elements respectively, for the i th test scenario. Therefore in the case of the haul truck test, they are represented by,

$$\mathbf{x}_i(k) = [\ddot{z}_{urf}(k) \ \ddot{z}_{ulf}(k) \ \dot{z}_{urf}(k) \ \dot{z}_{ulf}(k) \ z_{urf}(k) \ z_{ulf}(k) \ v(k)]_i^T \text{ and,}$$

$$\mathbf{y}_i(k) = [z_{rr} \ z_{rl}]_i^T ;$$

where $\ddot{z}_{urf}(k)$, $\ddot{z}_{ulf}(k)$, $\dot{z}_{urf}(k)$, $\dot{z}_{ulf}(k)$, $z_{urf}(k)$, $z_{ulf}(k)$ and $v(k)$ are the accelerations, velocities and displacements measured and calculated on the right-front and left-front axles, and the vehicle speed at the k th sample point; and z_{rr} and z_{rl} are the road profile heights at the right and left wheel tracks respectively. The superscript T denotes the transpose of the relevant matrix or vector. The input column for the small utility vehicle has ten elements, as previously mentioned in this Section.

The network is prescribed to train for a total number of 100 epochs but generalisation stops the training process much earlier at around 25 epochs. Generalisation is achieved through the use of a performance function, **msereg.m** (Mathworks Inc., 2007), that minimises the sum of the square



errors of the weights and biases (Mathworks Inc., 2007; Bishop, 1995). It is said that this performance function produces smaller weights and biases in the network, and forces the network response to be smoother and less likely to overfit the training data (Bishop, 1995). The performance ratio is set to 0.5, which implies giving equal weight to the mean square errors and the mean square weights.

During simulation, “unseen” inputs are applied to the trained neural network models. These inputs are processed in a manner similar to that for the training data. The network yields normalised profile heights which are converted back into their actual values. The roughness classes represented by DSDs are subsequently computed from the transformed road profiles as described in Chapter 3.

8.3 Measured data and selection of training data

This Section presents the measured data from both tests. Section 8.3.1 presents the way that the data was organized and Section 8.3.2 presents the method employed for selection of training data.

8.3.1 Organisation of the measured data

The measured data for the haul truck test is classified as shown in Table 8.1 part (A). There are eleven test cycles, of which seven are for the unloaded haul truck. Test cycles (2) and (3) do not have data for the return test runs because the haul truck was not driven over the defects on return. The letters L, H and V appended to the test run numbers represent low, high and variable speed ranges, respectively. A test run is considered as being conducted at a low speed range if all its elements in the truck velocity vector are below 20 km/h and vice versa for a test run in the high speed range. As test runs 3, 9 and 10 have elements in the velocity vector belonging to both ranges, however, they are labelled as variable speed test runs.

During fully loaded truck testing, the speed ranges alternated between low and high ranges from test cycle (8) to (11). Test cycle (8) was discarded because the results were spurious. The analysis of the data from this test cycle shows that the acquired accelerations did not indicate the existence of the any of the defects shown in test cycle (10), which was conducted under very similar conditions.



Table 8.1. Summary of how measured data was organised

(A) HAUL TRUCK TEST											
	UNLOADED TRUCK							FULLY LOADED TRUCK			
	TEST CYCLE NOS.										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
FORWARD											
Defect 1	1L	7L	10V	13L	19L	25H	31H	37L	43H	49L	55H
Defect 2	2L	8L	11L	14L	20L	26H	32H	38L	44H	50L	56H
Defect 3	3V	9V	12L	15L	21L	27H	33H	39L	45H	51L	57H
RETURN											
Defect 3	4L	---	---	16L	22L	28H	34H	40L	46H	52L	58H
Defect 2	5L	---	---	17L	23L	29H	35H	41L	47H	53L	59H
Defect 1	6L	---	---	18L	24L	30H	36H	42L	48H	54L	60H
AVERAGE											
SPEED (km/h)	16	9	8	12	11	28	35	---	28	9	25
(B) SMALL UTILITY VEHICLE TEST											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)		
Gravel Road	1	3	5	7	9	11	13	15	17		
Paved Road	2	4	6	8	10	12	14	16			
AVERAGE											
SPEED (km/h)	6	8	11	12	12	10	10	12	12		

The remaining ten test cycles had 24 different test scenarios derived from different combinations of truck load condition (unloaded or fully loaded); defect type (defect 1, defect 2 or defect 3); speed range (low or high); and travelling direction (forward or return). The condition of different truck loads is addressed by “weighting” the measured accelerations with the two different ratios derived from the



respective approximate weights of the truck during an unloaded and fully loaded state. This reduced the number of possible test scenarios from 24 to 12. Ideally the neural network should therefore be trained with 12 different sets of data, each representing a different test scenario. However, that holds true only if those 12 test scenarios are found to contain data that is linearly independent or not highly correlated.

Table 8.1 part (B) summarizes the test data for the small utility vehicle test. The presented data is from nine test cycles, each cycle comprising travelling first over the gravel road and then over the paved road, giving a total of 17 test runs, since the last cycle was not completed over the paved road. The nominal vehicle speeds are shown below each test cycle. Each test track is 24 m long. Since all the speeds are clustered within the same range (from 8 km/h to 12 km/h) and the vehicle is unloaded throughout these tests, there are only two obvious test scenarios, i.e. tests over the gravel road and tests over the paved road.

8.3.2 Selection of training data

The selection of training data is a crucial exercise owing to the desire to minimize computer training time in the presence of multiple candidates, as is the case for the haul truck test. The small utility vehicle test does not present as great a challenge because there are ideally only two obvious cases from which the training data can be selected, given that the vehicle speeds differ only marginally. Accordingly, only the haul truck test data is discussed in this Section.

The ANN may require 12 different sets of training data in order to achieve generalization over all possible different scenarios for the haul truck test if these test scenarios comprise data that is linearly independent. A number of different tools may be used to check linear the dependencies in data, but in this study, the Pearson's correlation coefficient implemented in MATLAB (Mathworks Inc., 2007) by the function `corr.m` was employed. In addition to the correlation coefficient matrix which indicates correlation between two column vectors, `corr.m` also returns a matrix of p-values for testing the hypothesis of no correlation against the alternative that there is a non-zero correlation (Mathworks Inc., 2007). Each element in the matrix is the p-value for the corresponding element in the correlation coefficient matrix. Any two given data sets are considered insignificantly correlated if an element in the p-value matrix is greater than 0.05.



Table 8.2. Number of training data combinations as determined by the use of p-values ≥ 0.05

Training Data Id.	1L	31H	8L	26H	15L	27H	16L	28H	17L	29H	18L	30H
1L		0.08	0.57		0.61	0.06						
31H	0.08					0.26	0.63	0.99	0.87			
8L	0.57											0.45
26H						0.32	0.37		0.99			
15L	0.61					0.06			0.12	0.76		0.25
27H	0.06	0.26		0.32	0.06		0.88	0.10	0.20	0.29	0.94	0.33
16L		0.63		0.37		0.06			0.27			0.96
28H		0.99				0.10					0.07	
17L		0.87		0.99	0.12	0.20	0.27				0.40	
29H					0.76	0.29						
18L						0.94		0.07	0.40			
30H			0.45		0.25	0.33	0.96					
<i>No.</i>	4	5	2	3	6	10	6	4	7	3	3	5
<i>Uncorr.</i>												

The possible training data combinations are shown in Table 8.2 where the last row shows the number of uncorrelated data sets (abbreviated as *No. Uncorr.*) for a given data set. In order to minimise the computational overheads during training, it is recommended that combinations should be chosen that have a few number of data sets such as: {8L, 1L, 30H}, {26H, 27H, 16L, 17L}, {28H, 31H, 27H, 18L}, {29H, 15L, 27H}, or {18L, 27H, 28H, 17L}. It was furthermore observed that the fewer the number of uncorrelated data sets (*No. Uncorr.*), the greater the influence of that data set in the combination. For example, in {26H, 27H, 16L, 17L}, 26H has the highest influence with only three uncorrelated data sets whereas 27H has the least influence with a total of 10 uncorrelated data sets. When using this combination, therefore, data sets 27 and 17L (with a total of 7 uncorrelated data sets) may be dropped with a minimal risk of losing ANN performance. Therefore the results presented in Section 8.4 are based on the training data combination {26H, 16L}.



8.4 Results and discussions

A total of 67 test results are presented here with 50 from haul truck tests and 17 from small utility vehicle tests. The accuracy of the simulation results for these two types of tests are presented in terms of the root mean square error (RMSE) and correlation coefficient (R), as presented in Chapter 4: Section 4.5, expressed here as percentages. The results of the haul truck tests are presented in Section 8.4.1 and Section 8.4.2 presents the results of the small utility vehicle tests.

8.4.1. Haul truck tests

The results for haul truck testing in Table 3 show that, except for a very few cases, bias errors lie below 25% and a good proportion of the test cases have their correlation coefficients above 50%. These are very encouraging results, considering the size of the truck and the shapes of the defects under investigation. The truck size and that of its tires make it almost practically impossible to detect the presence of small discrete obstacles or high-frequency undulations. This truck attenuates all axle accelerations at frequencies higher than its wheel hop, which in a similar truck studied by Hugo et al. (2008) is reported to be in the range between 3 – 4 Hz. At the same time, the content of the low frequencies in the measured accelerations is limited by the lengths of the test profiles themselves, which are: 4 m long for defect 1, 6 m long for defect 2 and 10 m long for defect 3.

For example, when the truck is travelling at the highest nominal speed of 34 km/h, it will be sensitive to a minimum road roughness wavelength of $\lambda_{\min} = \frac{v}{f} = \frac{34}{(3.6 \times 4)} = 2.4\text{ m}$, whereas at the lowest nominal speed of 8 km/h, the minimum detectable roughness wavelength is, $\lambda_{\min} = \frac{8}{(3.6 \times 4)} = 0.6\text{ m}$. This implies that at the high truck speeds only a few wavelengths of surface roughness from 2.4 m to 4 m for defect 1, 2.4 m to 6 m for defect 2, and 2.4 m to 10 m for defect 3, effectively contribute to the axle accelerations, with the result that the reconstructed profiles are relatively smooth. At the low truck speeds, the wavelengths start from 0.6 m which implies the participation of high frequencies in the axle accelerations, with the result that the reconstructed profiles have some undulations.

The results in Table 8.3 show that the neural network is capable of learning the geometry of defect 2 more accurately than of defect 1 and defect 3. As the space between the two main bumps in defect 1 does not allow for a complete settling of responses from the first bump, the transient responses from the first bump have more impact on the responses over the second bump. This problem may be aggravated by the difficulties encountered with controlling the operating conditions when the haul truck was traversing these two bumps. In relation to defect 3, defect 2 has a more compliant shape with smoothly blended transitions between adjacent curves. Defect 3 has rather more abrupt changes in gradients that may be more susceptible to instability during numerical integration. It is very



encouraging to note from the results of the training data, however, that the bias error for defect 3 actually compares very favourably with that for defect 2.

The results in Table 8.3 show the following order of accuracy in estimating the three defects: defect 2 in forward run, defect 2 in return run, defect 3 in return run, defect 1 in forward run, defect 3 in forward run and defect 1 in return run. The ones in bold-type correspond to the data that was used for training the neural network. The correlation coefficient results for defect 1 in the return run can benefit from re-training the ANN with a representative “defect 1 in return run” test data. Unfortunately this was observed to have so adversely affected the simulation results of the other defects that, despite an improvement in the performance of the ANN on defect 1, there were significant decreases in its performance on the other two defects. This is an aspect of generalization that requires further investigation.

Table 8.3. Summary of results showing defect type and fitting errors as measured by the RMSE and the correlation coefficient (R)

Defect 1			Defect 2			Defect 3		
Test Run No.	RMSE (%)	Correlation Coeff., R (%)	Test Run No.	RMSE (%)	Goodness of Fit, R (%)	Test Run No.	RMSE (%)	Correlation Coeff., R (%)
FORWARD								
1L	22.2	57.1	2L	8.9	65.2	12L	33	-44.9
7L	5.5	63	8L	11.4	58.7	15L	27.6	-31.2
13L	15.4	78.2	11L	8	56.6	21L	3.1	25.4
19L	19.1	63.3	14L	11.6	67.1	27H	3.8	61.1
25H	22.9	-36.9	20L	6.2	63.6	33H	10.4	62.4
31H	9.8	-10.1	26H	3.3	89.1	45H	65.2	40.2
43H	9.1	25.2	32H	11.4	85.9	51L	36.8	15.3
55H	3.9	47.2	44H	23.1	71.8	57H	24.9	36
---	---	---	50H	19.9	74.2	---	---	---
---	---	---	56H	23.9	85	---	---	---
RETURN								
6L	5.6	29.7	5L	13.3	37.6	4L	14.5	13.4



18L	21.6	0.5	17L	15.6	66.3	16L	4.3	51.3
24L	7.5	13.7	23L	6.9	63.3	22L	20.2	57.1
30H	10.9	9.6	29H	2.6	68.6	28H	14.5	57.9
36H	13.3	-6.8	35H	5	60.4	34H	8.9	59.1
48H	15.8	11.9	47H	20.1	66.3	46H	34.8	59.2
54L	12.5	-18.1	53L	22.8	53.7	52L	1.8	-3.9
60H	6.2	43.9	59H	22.7	77.1	58H	73.7	18.7

The following three Sections present the ANN simulation results for each defect. The test runs are identified simply by their numbers prefixed by the symbol “#” and the letters L, H and V have been dropped for simplicity.

8.4.1.1 Defect 1

Figures 8.5(a) and (b) show the correlation plots of the actual and reconstructed profiles in the forward and return runs respectively. In Figure 8.5(a), Test runs #1 and #7 represent the road profiles reconstructed from low-speed truck data whereas Test runs #25 and #31 have been reconstructed from high-speed truck data. The road profiles from the high-speed truck data appear smoothed out due to the filtering effects at high speeds. In Figure 8.5(b), Test runs #6 and #18 for the return runs represent the simulations from the trucks at low speed whereas Test runs #36 and #60 represent those from the trucks at high speed. All the low-speed results have some undulations which can be attributed to the participation of short wavelengths in exciting the axle accelerations. Test run #6 is flattened out over the second bump, presumably because the truck followed a different wheel track during the earliest stages of the test when the truck operator was still becoming familiar with the requirements of the test. The undulation over the second bump in Test run #60 was most probably caused by the superposition of pitch and accelerations from the increased sprung mass reaction on the axle on the measured axle accelerations. The sprung mass bounce and pitch for the haul truck were coupled and observed to be excited when the vehicle was fully loaded. However, the results show that the two prominent bumps of the defect were correctly identified in both the forward and return runs. The slight errors in the locations of the bumps were caused by a combination of data alignment and resampling disparities prior to the ANN training and simulation.

Figure 8.5(c) shows the roughness classification for these reconstructed profiles compared to that of the actual road profile. All the test runs yielded similar roughness classifications between classes D and E for the spatial frequency range from 0.15 cycles/m to 0.4 cycles/m, corresponding to wavelengths between 2.5 m and 6.7 m. In the roughness classification plot, the dotted lines represent



different roughness classes from A to G. In all the roughness plots presented in this thesis, only classes A and G are labelled for the sake of convenience.

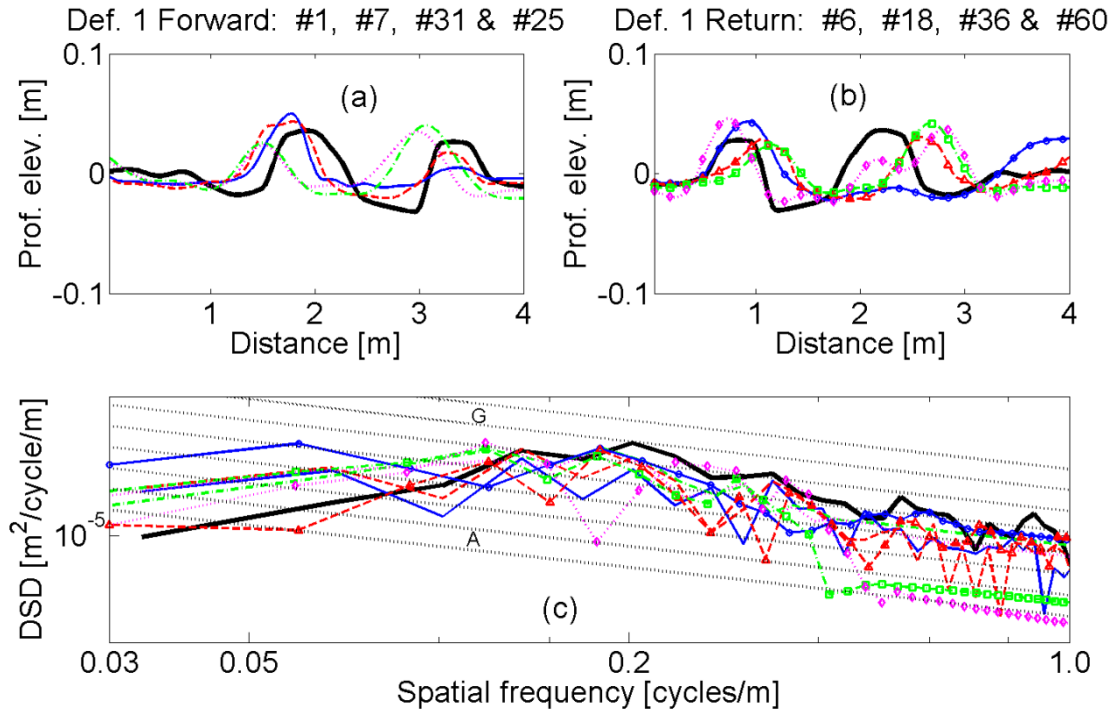


Figure 8.5. Correlations over Defect #1 in forward run (a), in return run (b) and their corresponding DSDs (c): **—** Actual, **—** #1, **- - -** #7,

- · - #31, **· · ·** #25, **—○—** #6, **- ▲ -** #18, **- □ -** #36, **· ◆ ·** #60.

8.4.1.2 Defect 2

Figures 8.6(a) and (b) present the ANN simulation results for defect 2 in the forward and return runs, respectively. The results show an excellent correlation, particularly for Test runs #32 and #29, and all the other test runs correctly identify the two main bumps in this defect though there are variations in how they are able to reconstruct the intermediate saddle. Test runs #32 and #29 are both for an unloaded truck travelling at high speeds. The roughness classifications in Figure 8.6(c) show that defect 2 largely lies in the roughness classes between D and F. There is generally an excellent correlation in the roughness classifications from spatial frequency of 0.5 to 0.15 cycles/m.

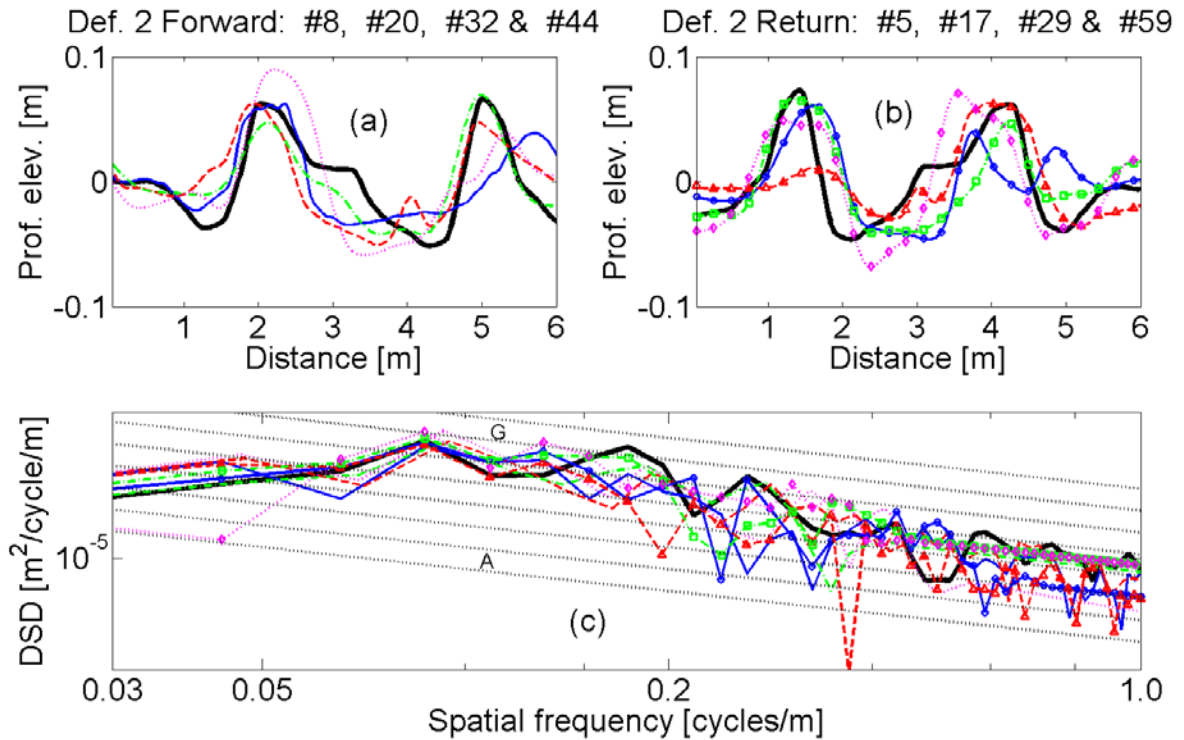


Figure 8.6. Correlations over Defect #2 in forward run (a), in return run (b) and their corresponding DSDs (c): — Actual, — #8, - - #20, — · — #32, ····· #44, —○— #5, -▲- #17, —□·— #29, ···◇··· #59.

8.4.1.3 Defect 3

In Figures 8.7(a) and (b), Test runs #27 and #28 show an excellent correlation with the actual profiles, especially in identifying and locating the two prominent bumps in the profile. Both of these test runs are for an unloaded truck travelling at high speeds. The intermediate portion between the two bumps poses a challenge, in that the flat Section and abrupt gradient changes introduce instabilities in the training algorithm, thus yielding a substantial number of undulations. The problem is actually aggravated in the unloaded truck, at low truck speeds in the return run, where the reconstructed profiles show very poor correlations. However, the road roughness classification in Figure 8.7(c) presents well-correlated DSDs of spatial frequency between 0.5 and 0.15 cycles/m for the actual and reconstructed profiles, irrespective of the poor correlations for the raw profiles themselves.

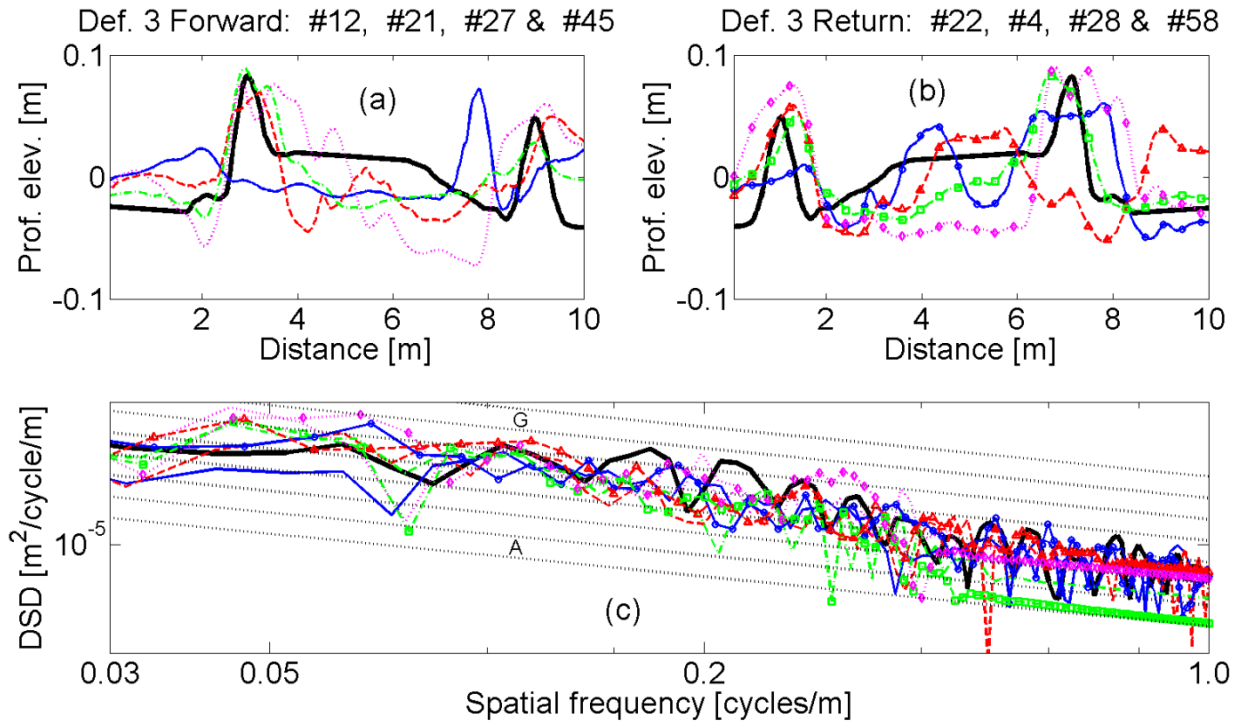


Figure 8.7. Correlations over Defect #3 in forward run (a), in return run (b) and their corresponding DSDs (c): — Actual, — #12, - - #21, — · — #27, ····· #45, —○— #22, -▲- #4, —□·— #28, ···◇··· #58.

8.4.2. Small utility vehicle for underground mining

This vehicle and its tires are much smaller in size than the haul truck and therefore it does not pose similar problems to the haul truck, yet it has its own unique challenges. Its tire size and contact patch area make it responsive to shorter road roughness wavelengths. In addition, its unsprung mass resonance is much higher than for the haul truck. Consequently, unlike the haul truck where DSDs in the higher frequency ranges are largely underestimated by the ANN, the DSDs for the small utility vehicle are overestimated due to the errors introduced by the crude road profiling procedure as well as by the presence of high frequency noise in the ANN reconstructed profile. Consider the lowest speed of 6 km/h with a vehicle whose wheel hop frequency is 15 Hz. This vehicle is potentially sensitive to a minimum road roughness wavelength $\lambda_{\min} = \frac{6}{(3.6 \times 15)} = 0.11\text{m} = 110\text{mm}$ and at the highest speed of 12 km/h, the minimum road roughness wavelength is 220 mm. The shortest sample spacing during actual profiling is 250 mm, which automatically excludes the shorter wavelengths and therefore fails to account for the higher frequency content in the measured axle accelerations. Though this problem is dealt with by resampling, aligning and filtering the measured accelerations and road profiles, it is difficult to achieve perfect alignment and filtering without roll-off effects.



The results in Table 8.4 show that only 6 out of 17 test runs have bias errors above 25% and 3 out of 17 test runs have correlation coefficients below 50%. The results show that the ANN generally performs better over the gravel track than the paved track in this test. Although the results for the training data show that the paved track has a much lower RMSE at 2.3%, there are four test runs with RMSE above 25% compared to only two test runs for the gravel track. The performance of the ANN around depressions and bumps contributes significantly to the differences in performance between the two tracks. As mentioned earlier, abrupt changes in gradients at and around certain profile Sections make the ANN unstable, thus generating an augmented transient behaviour that cannot be sufficiently accounted for by the attendant profile geometry. In this test, the paved track profile had such sudden changes in profiles, especially around its bumps and depressions. Moreover, the paved track's harder surface made it harsher, especially at sufficiently high speeds, than the more flexible and tractable gravel track, which was wet on the day of the test. Bold-type in Table 8.4 corresponds to data that was used for neural network training.

Table 8.4. Summary of results for the small utility vehicle.

Gravel Road Section			Paved Road Section		
Test Run No.	RMSE (%)	Correlation Coeff. R (%)	Test Run No.	RMSE (%)	Correlation Coeff. R (%)
1	36.5	42.3	2	34.8	55.4
3	27.2	56.7	4	14.5	64.4
5	13.6	53.5	6	14.2	54.6
7	10.7	66.4	8	27.1	68.6
9	10.5	75.8	10	27	69.7
11	20.2	65.3	12	29.1	39.5
13	3	69.1	14	2.3	33.5
15	2.2	66.9	16	6.4	50.5
17	5.8	75.3	---	---	---

8.4.2.1 Gravel road

The results in Figure 8.8(a) show a good correlation between the actual and reconstructed profiles, except for Test run #1 where the ANN locates depressions where there are no depressions. It is not known what might have caused this error. Test runs #3, #9 and #17 follow the actual profile very well, locating all the important depressions, even though Test run #3 underestimates the depressions and bumps. This underestimation may be due to the relatively lower speed at 8 km/h of the small utility



vehicle. This led to a corresponding underestimation of the DSDs, especially on the lower frequency end, as shown in Figure 8.8(b). The higher frequency end from 2 to 5 cycles/m (wavelengths from 0.2 m to 0.5 m) is overestimated by the ANN due to the disparities between the measured profile and acceleration data sampling rates. However, the DSDs correlate very well between 0.4 and 2.0 cycles/m spatial frequencies for all the test runs.

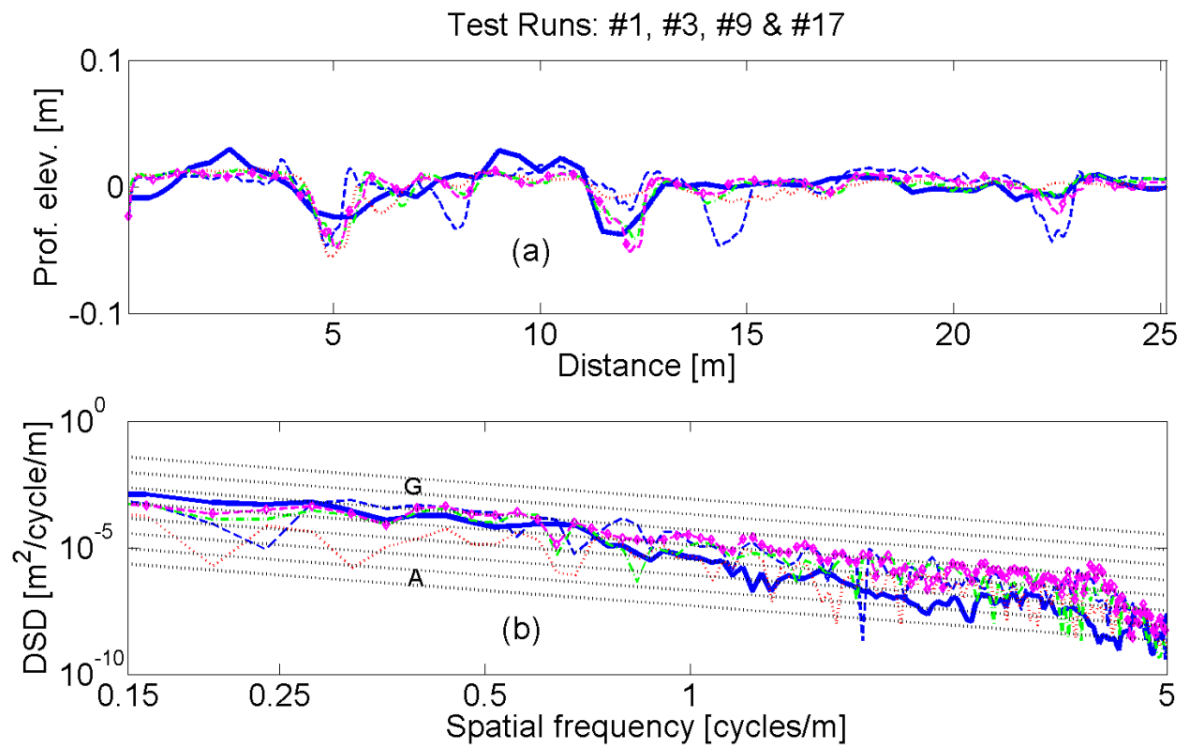


Figure 8.8. Comparison of actual with reconstructed profiles for Gravel Road Section

(a) their corresponding DSDs (b): **— Actual, - - #1, ····· #3, — · — #9, ·····◇···· #17.**

8.4.2.2 Paved road

In Figure 8.9(a), Test runs #2 and #4 show a good correlation with the actual profiles whereas Test runs #12 and #16 have extra bumps over a depression before the first bump. Both of these test runs are at slightly higher speeds and the ANN tends to generate unstable results due to the nature of the profile in this region. As is the case with the gravel road tests, the DSDs are overestimated at the higher frequency end in Figure 8.9(b) but show very good correspondence within the same spatial frequency ranges from 0.25 to 2.0 cycles/m which correspond to wavelengths of between 0.5 m and 4.0 m.

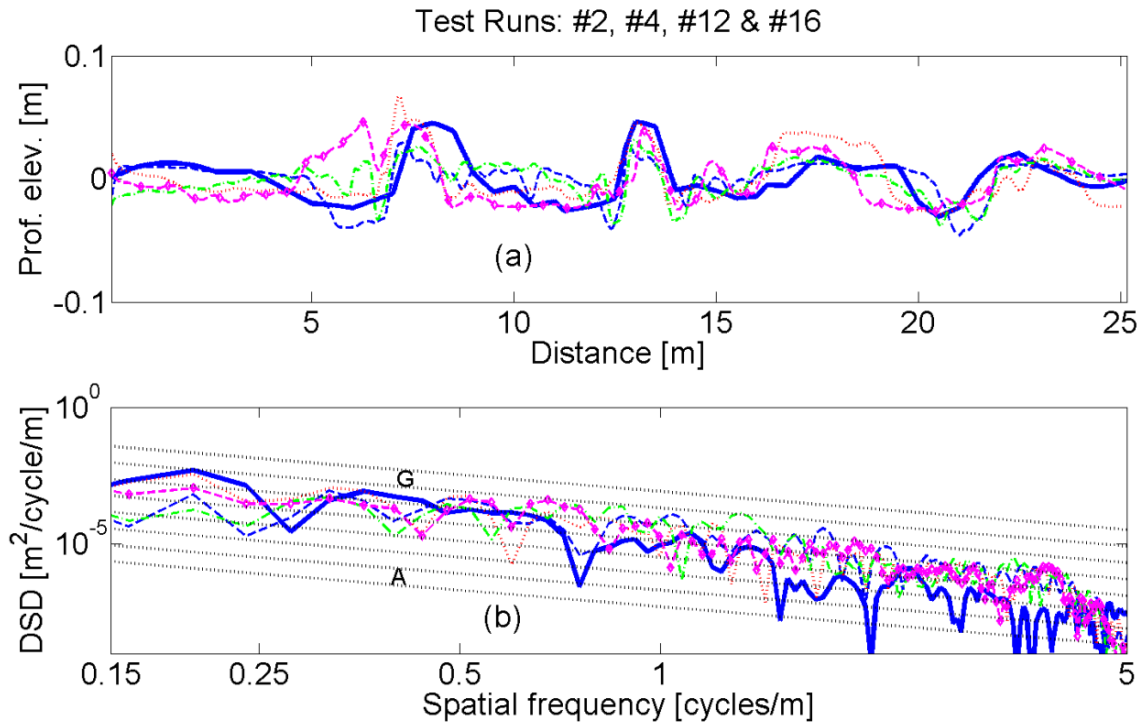


Figure 8.9. Comparison of actual with reconstructed profiles for Paved Road Section (a) their corresponding DSDs (b): — Actual, - - #2, #4, - . - #12,◇..... #16.

8.4.3 Summary of test results

The results presented in this Chapter validate the feasibility of the proposed methodology. The following points summarize the important findings obtained from the two tests.

- The ANNs yield good profile correlations, particularly with respect to identifying the prominent defects (i.e. bumps and depressions), on all road profiles for both tests. For the same ANN and vehicle type, the performance varies with the vehicle operating speeds and the geometry of the defects.
- In both tests, the ANNs are observed to perform better at high vehicle operating speeds than at low operating speeds. The reason is that, at the high speeds, enough dynamic energy is imparted to the vehicle structure while automatically eliminating the participation of non-essential short wavelengths from the test road profiles.
- The quality of the estimated DSDs is very good within the spatial frequencies between 0.15 and 0.5 cycles/m (corresponding to wavelengths of 2.0 to 6.7 m) for all defects in the haul truck tests and from 0.25 to 2 cycles/m (corresponding to wavelengths of 0.5 to 4.0 m) for the small utility vehicle tests. These ranges of wavelengths are noted to be influenced by the size of the vehicle and its tires. Smaller vehicles and tires give relatively better DSD definition in the shorter wavelengths (or high frequency) ranges.



- The profile geometry influences the performance of the ANN. A gentler profile with perfectly blended geometry allows a better ANN performance than profile geometry with abrupt curvature changes. For the haul truck test, defect 2, which is observed to be gentler than defect 3, yields relatively better correlations. For the small utility vehicle, the less aggressive gravel track yields relatively better correlations than the harsher paved track.
- The correlation method used in selecting the training data-sets has worked very well in this application, but the ANN performance for defect 1 in the return run, in terms of correlation coefficients, imply that further refinement is required.

8.5 Conclusions

A methodology for road profile reconstruction and road roughness identification has been applied successfully to two vehicles at different mine sites, where the vehicles were tested in their normal operating environment. This Chapter sought to concretize the findings of Chapter 6 for the numerical vehicle model using numerically generated road profiles, and later, Chapter 7 for an experimental vehicle with adjustable suspension to which accurately measured road profiles were applied. In the investigation of the numerical model, road profiles were reconstructed to very high levels of fitting accuracy under all simulation conditions. By contrast, in the investigation of the experimental vehicle, its performance was noted to be affected by two main factors, namely the scarcity of the training data and consistency in following similar wheel tracks. Despite such difficulties, this investigation still benefitted from the available numerical model of the vehicle and the accurately measured profiles. Accordingly, the present study has the following unique problems:

1. The non-availability of vehicle numerical models so that the training data does not benefit from data that can be easily generated from the model.
2. Lack of control or minimal control over the operating conditions in the normal working environment of the test vehicles.
3. Lack of accurate road profiles. The profiles are measured by a procedure which is very crude yet practically sound.
4. Application to two characteristically different vehicles.

In view of these challenges, the findings summarized in Section 8.4.3 are highly encouraging and indicate that the methodology holds promise for practical application to road condition monitoring systems. Quite a large proportion of the tests yielded bias errors of less than 25% with correlation levels higher than 50%. Generally the methodology provides two fronts in the monitoring of road condition: firstly, detecting and locating any prominent road profile features such as imminent potholes or bumps, and secondly identifying the general roughness condition of the road network, which is particularly useful where the road surface is becoming degraded by increased random roughness rather than by the size of discrete obstacles.

Caution should be taken regarding certain key issues when applying this methodology to a practical situation:

1. Vehicle speeds should be sufficiently high to allow for the participation of prominent road roughness wavelengths in the excitation of measured accelerations.
2. The profiles and measured accelerations should allow for easier alignment.



3. In order to assist with making a reliable decision about maintenance, the final results should come from averaged simulations so that some errors can be reduced.
4. Profiles should be smoothed out by running a type of moving average filter over the measured road profiles so that any abrupt profile changes can be eliminated.



Chapter 9

Conclusions

9.1 Introduction

The study is aimed at developing a methodology for identifying road damage through artificial neural networks simulation. The use of artificial neural networks is perceived to offer two main advantages. The first advantage is that it does not require excessive vehicle and/or road characterisation. Secondly (and related to the first one), it requires relatively fewer analytical skills to create the network than parametric models. Besides, developments in technical computing have ensured that most neural network models become standard and implemented in technical software e.g. MATLAB, whereas the development of physical parametric models still demands a great deal of technical skills, even in virtual-computing environments. The physical parametric models may require the calculation of the inverse models in order to determine the road profiles from measured vehicle responses, which can also be very rigorous on their own. Artificial neural networks only require input-output data that can reliably be used to establish a causal relationship for the system under study.

In this study, vertical vehicle accelerations and travelling speeds are used as inputs while road profiles are used as outputs. There are a number of other factors which affect the vehicle-road interaction system, such as excitations from vehicle body (powertrain vibrations) and wheel sub-assembly, but it has been shown that when these are kept minimal through careful vehicle control and adherence to vehicle's maintenance standards, their effects can be safely neglected. In addition to this it has been demonstrated through a study of transmissibility ratios on a mechanistic inverse model that vertical unsprung mass accelerations yield better results than vertical sprung mass accelerations.

The performance of the methodology has been demonstrated through three case studies: numerical investigation (in Chapter 6), experimental vehicle study (in Chapter 7) and field testing on two mine vehicles (in Chapter 8). The results from these studies generally show that the methodology can be used for road damage identification. However there are some specific differences in its performance from the numerical case study to the other two case studies. These differences can largely be attributed to the inherent differences between a purely numerical and more practical investigations. For example, the accuracies obtained during the numerical case studies are much higher than those in the other two case studies. There is some fuzziness over relative accuracies between high and slow vehicle speeds as well. This can be attributed to the differences in the mechanical properties of the vehicles used for the numerical study and the other two studies. Despite these details, the developed methodology holds a lot of promise for application to public unpaved roads or mine haul roads where most of the modern road maintenance management methods are not very well developed and implemented.



From the main purpose of the study, one may extract four key objectives that must be examined if they have been satisfied in the present study, namely:

- Detect, quantify and locate road damage as it occurs in real time
- Enhanced understanding on the application of ANN to function approximation problems
- Determine how various road and vehicle operating factors affect the identification process
- Determine how vehicle suspension characteristics affect the identification process.

A discussion on each of these objectives is presented in the following Sections.

9.2 Detection, quantification and location of road damage as it occurs in real time

This is the central focus of the study and has been demonstrated in every case study. The performance of the methodology has been demonstrated on how accurate the road elevations are reconstructed. For discrete obstacles such as potholes or bumps, their profile heights and locations along the road are assessed. In the case of random rough road surfaces, detection and quantification have been measured through ISO DSD scale benchmark.

It has not been demonstrated how this methodology can be implemented online but it is relatively easy to implement after a neural network is trained. The simulation procedure involves less-demanding multiplications of matrices as compared to iterative optimisation procedures during network training. Thus neural network training can be done offline using data from key typical vehicle operating conditions and once that has been completed, the optimal weight matrices can be used for simulation process online.

9.3 Better understanding on the application of ANN to function approximation problems

It has been discussed in Chapter 4 how the ANN is used in solving function approximation problems by minimising the error function between the network outputs and the targets with respect to network weight parameters. This has been demonstrated for purely back-propagation FFNN network, Levenberg-Marquardt optimisation-based NARX network and the radial basis function network. All these networks perform well with a slight advantage in terms of accuracy obtained from the NARX network. The radial basis function was shown to have advantages over the other networks with respect to speed of convergence.

9.4 Influence of road and vehicle operating factors affect the identification process

A number of operating factors such as the travelling speeds, sprung mass and unsprung mass excitations have been randomly investigated. Though the numerical study did not yield any clear dependencies on vehicle speeds, the analysis on the experimental vehicle showed better accuracies at



a higher vehicle speed (54 km/h) than the lower vehicle speed (15 km/h). This was attributed to the participation of more frequencies in the vehicle vibration at the higher speed than the lower speed. However, marginal effects were not thoroughly investigated. On the effects of the sprung mass excitations, it was found that these forces can affect the vibrations on both the sprung mass and unsprung mass quite significantly at all frequencies in the range 0 – 25 Hz, with much higher impact on the sprung mass than on the unsprung mass. While the forces acting directly on the unsprung mass, can affect the unsprung mass accelerations significantly only at higher frequencies from just below the wheel hop frequency up to slightly beyond it. It has however been argued in this work that these excitations can be kept in check by adherence to vehicle maintenance standards and careful vehicle operation where they are observed to have almost negligible effect.

9.5 Influence of vehicle suspension characteristics affect the identification process

This was specifically investigated through the experimental case study where it was possible to switch suspension settings modes between ride comfort and handling modes. The results on discrete obstacles showed superior quality in the reconstructed bumps for ride comfort as compared to handling mode. However, the overall network performance over bumps was observed to be marred by the poor roll motion estimation capabilities of the numerical model as well as its use of the point-follower tire model. This shortcoming shrinks the validity area but it is only model-dependent. If, on the other hand, a more comprehensive vehicle model is used, the validity area is effectively enlarged.

9.6 Summary of key findings from the case studies

9.6.1 Numerical case study

- The performance of the methodology was demonstrated for varying roughness grades, emerging surface defects, and changing or growing surface defects under varying conditions of noise, vehicle payload and speed through evaluations of correlation coefficients and mean square errors. The network performance did not exhibit any clear dependencies on the different conditions under study except that under harsh noisy conditions, the performance deteriorated significantly.
- The methodology was observed to correlate well with IRI which is a widely used roughness indicator. The study somewhat reinforced the combined use of IRI and ISO PSD classification to enhance road surface condition interpretation.



9.6.2 Experimental case study

- The ANN reconstructed profiles were correlated with actual measured profiles for both discrete obstacles and the Belgian paving. However in the discrete obstacles, only overlay plots of the simulated and actual bump profiles were found satisfactory, while for the Belgian paving, overlay plots as well as RMS error percentages were used to provide a measure of accuracy in correlation. The overlay plots of discrete obstacles show very good correlations between ANN simulated bumps and actual bumps for all cases under investigation.
- The Belgian paving however was problematic due to combined effects caused by the deficiencies in the training data and the numerical model that was used to generate the data. Thus ANN training was carried out in the absence of sufficient training data. The best result came from the near-optimal performance of the training procedure which could only be achieved through trial-and-error iterations.
- The methodology classified the road consistently and accurately within a certain frequency range (0.2 to 1.8 cycles/m).

9.6.3 Mine vehicle test

This investigation was carried out with so many limitations: no vehicle models available, profiles were crudely measured, minimal control of the vehicle operation, and two different vehicles at two sites. A large proportion of the tests yielded bias errors of less than 25% with correlation levels higher than 50%.

The following precautions were drawn from the conduct of this test:

1. Vehicle speeds should be sufficiently high to allow for the participation of prominent road roughness wavelengths in the excitation of measured accelerations.
2. The profiles and measured accelerations should allow for easier alignment.
3. In order to assist with making a reliable decision about maintenance, the final results should come from averaged simulations so that some errors can be reduced.
4. Profiles should be smoothed out by running a type of moving average filter over the measured road profiles so that any abrupt profile changes can be eliminated.



Chapter 10

Recommendations for further study

The following points have not been carefully investigated and concluded in this study and therefore are proposed for further investigation:

10.1 Online implementation with an automated training procedure: In the present application of the developed procedure, neural network training is carried out offline. This would also include the automation of the training data preprocessing techniques.

10.2 Methodology standardisation over different vehicles and operating conditions: This is obviously part of the implementation. This methodology has been applied to a fleet of characteristically similar vehicles for which the neural network is trained. Thus standardisation would enable application of the methodology to be extended across characteristically different vehicles.

10.3 Integrated road damage identification methodology where direct mechanistic inverse and ANN routines can be used interchangeably: This is merely to aid in understanding how neural networks operate. A mechanistic model is very easy to interpret so if they are used together with the neural network the analysts may benefit from this characteristic to interpret the operation of the system.

10.4 Thorough influences of vehicle speed to the neural network performance: It has not been established so clearly how vehicle speeds affect this methodology so there is need for this to be pursued further.

10.5 Investigate use of more accurate vehicle and tyre models in the generation of training data: It is expected that if more accurate vehicle and tyre models are used, the methodology can yield much more improved results.

10.6 Investigate the way to achieve network convergence in a more stable manner rather than perform a number of trial-and-error guesses: The NARX network is particularly observed to behave in an unstable manner in that it does not always converge to optimal weights with every training session.

10.7 Investigate the entire width of the road not limited to the wheel path only.

10.8 ANN in the thesis use sigmoid functions. However, these may not always be the best functions. Later research will focus on the use of functional networks.



References

Abdallah, I., Yuan, D. and Nazarian, S. (2001). Integrating seismic and deflection methods to estimate road moduli. *Transportation Research Record* Issue 1755: pp 43-50.

Abd El-Hakim, R. and El-Badawy, S. (2013). International roughness index prediction for rigid roads: An Artificial Neural Network Application. *8th International Conference on Road and Airfield Road Technology*, ICPT 2013, 14 - 18 July 2013; Taipei; Taiwan: pp 854-86.

Abdelrahim, A.M. and George, K.P. (2000). Artificial neural network for enhancing selection of road maintenance strategy. *Transportation Research Record* 1699: pp 16-22.

Abiola, O.S., Owolabi, A.O., Sadiq, O.M., Aiyedun, P.O. (2012). Application of dynamic artificial neural network for modeling ruts depth for Lagos-Ibadan Expressway, Nigeria. *ARPJ Journal of Engineering and Applied Sciences* 7(8): pp 987-991.

Abo-Hashema, M. (2009). Artificial neural network approach for overlay design of flexible roads. *International Arab Journal of Information Technology* 6(2): pp 204-212.

Alsugair, A.M. and Al-Qudrah, A.A. (1998). Artificial neural network approach for road maintenance. *Journal of Computing in Civil Engineering* 12(4): pp 249-255.

American Association of State Highway and Transportation Officials (AASHTO) (2001). *Road management guide*. AASHTO: Washington D.C.

Andr n P. (2006) Power spectral density approximations of longitudinal road profiles. *International Journal of Vehicle Design* 40(1-3): pp 2 – 14.

Atkinson K. (1990). Introduction. In the *Highway maintenance handbook*. Edited by Atkinson, K. Thomas Telford Ltd: London.

Attoh-Okine, N.O. (1999). Analysis of learning rate and momentum term in back-propagation neural network algorithm trained to predict road performance. *Advances in engineering software* 30(4): pp 291-302.



Banan, M.R. and Hjelmstad, K.D. (1996). Neural networks and AASHO road test. *Journal of Transportation Engineering* 122(5): pp 358 – 366.

Basheer I.A. and Najjar Y.M. (1996). A neural network-based distress model for Kansas JPCP longitudinal joints. *Intelligent Engineering Systems Through Artificial Neural Networks*, 6: pp 983-987.

Bayrak, M.B. and Ceylan, H. (2008). Neural network-based approach for analysis of rigid road systems using deflection data. *Transportation Research Record* Issue 2068: pp 61-70.

Becker C.M., Els P. (2014). Profiling of rough terrain. *International Journal of Vehicle Design* 64(2/3/4): pp 240 – 261.

Becker C.M. (2008). Profiling rough terrain. MEng Dissertation, Department of Mechanical and Aeronautical Engineering, University of Pretoria, South Africa. <http://upetd.up.ac.za/thesis/available/etd-11262009-171410/unrestricted/00front.pdf>

Beguin, S., Bastin, G. and Wertz, V. (2005). Automatic prediction of ICY conditions on roads using a LS-SVM classifier. *16th Triennial World Congress of International Federation of Automatic Control*, 16, 3 – 8 July 2005, 2005, Prague, Czech Republic: pp 37-42.

Bendat, J.S. and Piersol, A.G. (2000). *Random data: analysis and measurement procedures*. 3rd Edition, John Wiley & Sons, Inc. New York.

Bianchini, A. and Bandini, P. (2010). Prediction of road performance through neuro-fuzzy reasoning. *Computer-Aided Civil and Infrastructure Engineering*, 25 (1): pp 39 – 54.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press. Oxford.

Bosurgi, G. and Trifirò, F. (2005). A hybrid approach for road maintenance management of an Italian motorway. *Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering*: 12 - 15 July 2005, Cancun, Mexico: pp 725-733.

Breytenbach B., Els P.S. (2011). Optimal vehicle suspension characteristics for structural fatigue life. *Journal of Terramechanics* 48: pp 397–408.



Breytenbach H.G.A. (2009). Optimal vehicle suspension characteristics for increased structural fatigue life. MEng Dissertation, Department of Mechanical and Aeronautical Engineering, University of Pretoria, South Africa. <http://upetd.up.ac.za/thesis/available/etd-09172010-190101/unrestricted/dissertation.pdf>.

Broomhead, D.S. and Lowe, D. (1988). Multivariate functional interpolation and adaptive networks. *Complex Systems 2*: pp 321 – 355.

Brushett, S. (2005). Experience in reforms of road maintenance financing and management in Sub-Saharan Africa. *Transport and Communications Bulletin for Asia and the Pacific*, Number 75, pp. 25-43.

Burningham, S. and Stankevich, N. (2005). Why road maintenance is important and how to get it done. Transport Note No. TRN-4, World Bank, Washington D.C.

Cebon D. (1999). *Handbook of vehicle-road interaction*, Swets & Zeitlinger Publishers, Lisse.

Ceylan, H. and Gopalakrishnan, K. (2007). Neural networks based models for mechanistic-empirical design of rubblized concrete roads. *Geotechnical Special Publication*, Issue 169.

Ceylan, H., Gopalakrishnan, K., Bayrak, M.B. (2008). Neural networks based concrete airfield road layer moduli back-calculation. *Civil Engineering and Environmental Systems 25*(3): pp 185-199.

Ceylan, H., Gopalakrishnan, K., Bayrak, M.B. and Guclu, A. (2013). Noise-tolerant inverse analysis models for nondestructive evaluation of transportation infrastructure systems using neural networks. *Nondestructive Testing and Evaluation 28*(3): pp 233-251.

Ceylan, H., Gopalakrishnan, K., Kim, S., Schwartz, C.W. and Li, R. (2013). Global sensitivity analysis of jointed plain concrete road mechanistic-empirical performance predictions. *Transportation Research Record*, Issue 2367: pp 113-122.

Ceylan, H., Gopalakrishnan, K. and Lytton, R.L. (2011). Neural Networks modeling of stress growth in asphalt overlays due to load and thermal effects during reflection cracking. *Journal of Materials in Civil Engineering 23*(3): pp 221-229.



Ceylan, H., Guclu, A., Tutumluer, E. and Thompson, M.R. (2005). Back-calculation of full-depth asphalt road layer moduli considering nonlinear stress-dependent subgrade behavior. *International Journal of Road Engineering* 6(3): pp 171-182.

Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). Accuracy of predictive models for dynamic modulus of hot-mix asphalt. *Journal of Materials in Civil Engineering* 21(6): pp 286-293.

Ceylan, H., Tutumluer, E. and Barenberg, E.J. (1999). Artificial neural networks for analyzing concrete airfield roads serving the Boeing B-777 aircraft. *Transportation Research Record*, 1684: pp 110-117.

Chandra, S., Sekhar, C.R., Bharti, A.K. and Kangadurai, B. (2013). Relationship between road roughness and distress parameters for Indian highways. *Journal of Transportation Engineering* 139(5): pp 467-475.

Chen, C. and Flintsch, G.W. (2007). Fuzzy logic road maintenance and rehabilitation triggering approach for probabilistic life-cycle cost analysis. *Transportation Research Board 86th Annual Meeting*, paper No. 07-2979.

Chik, Z. and Aljanabi, Q.A. (2013). Intelligent prediction of settlement ratio for soft clay with stone columns using embankment improvement techniques. *Neural Computing and Applications*. pp 1-10.

Chesher, A. and Harrison, R. (1987). Vehicle operating costs: Evidence from developing countries. A World Bank Publication: The Highway Design and Maintenance Series, The Johns Hopkins University Press: Washington D.C.

Committee of Transport Officials(COTO) (2007). South Africa. Guidelines for Network Level Measurement of Road Roughness Version 1.0 <http://www.nra.co.za/content/Blogin/TMH10.pdf>.

Directorate for Science, Technology and Industry (DSTI). (1998). *Scientific Expert Group IR6 on the Dynamic Interaction between Vehicles and Infrastructure Experiment (DIVINE Project)*: Technical report. DSTI/DOT/RTR/IR6(98)1/FINAL (unclassified). OECD, France.

Dodds, C.J. and Robson, J.D. (1973). The description of road surface roughness. *Journal of Sound and Vibration* 31(2). pp 177 – 183.



Eldin, N. N. and Senouci, A. B. (1995). A road condition rating model using backpropagation neural network, *Microcomputers in Civil Engineering* 10 (6): pp 433 – 441.

Els, P.S. (2006). The ride comfort vs. handling compromise for off-road vehicle. PhD Thesis, Department of Mechanical and Aeronautical Engineering, University of Pretoria, South Africa. <http://upetd.up.ac.za/thesis/submitted/etd-07152008-102911/unrestricted/00front.pdf>.

Fajardo, E.L.J. (2011). Road visual condition rating assessment using artificial neural network and support vector machine. Proceedings - AMS 2011: Asia Modelling Symposium 2011 - 5th Asia International Conference on Mathematical Modelling and Computer Simulation, Article number 5961234: pp 18-22.

Far, M.S.S., Underwood, B.S., Ranjithan, S.R., Kim, Y.R. and Jackson, N. (2009). Application of artificial neural networks for estimating dynamic modulus of asphalt concrete. *Transportation Research Record*. Issue 2127: pp 173-186.

Flintsch, G.W. and Zaniewski, J.P. (1997). Expert project recommendation procedure for Arizona department of transportation's road management system. *Transportation Research Record* Issue 1592: pp 26-34.

Flintsch, G.W., Zaniewski, J.P. and Delton, J. (1996). Artificial neural network for selecting road rehabilitation project. *Transportation Research Record* Issue 1524: pp 185-193.

Gandhi, T., Xiao, F. and Amir Khanian, S.N. (2009). Estimating indirect tensile strength of mixtures containing anti-stripping agents using an artificial neural network approach. *International Journal of Road Research and Technology* 2(1): pp 1-12.

García-Balboa, J., Reinoso-Gordo, J. and Arzia-López, F. (2012). Automated assessment of road generalization results by means of an artificial neural network. *G.I. Science and Remote Sensing*, 49(4): pp 558-596.

Garrick, N.W., Kalikiri, V.K. and Achenie, L.E.K. (1994). Artificial neural networks for road evaluation. *Proceedings of the Artificial Neural Networks in Engineering Conference (ANNIE'94)*, Volume 4, 13 - 16 November 1994. St. Louis, MO, USA: pp 461-466.

Gillespie, T.D. (1992). Fundamentals of vehicle dynamics. Society of Automotive Engineers, Inc. Warrendale.



Goh, A.T.C. (1997). A hybrid neural network based road management system. *Road and Transport Research* 6(4): pp 62-71.

González, A., O'Brien, E.J., Li, Y.-Y., and Cashell, K. (2008). The use of vehicle acceleration measurements to estimate road roughness. *Vehicle System Dynamics*, 46(6): pp 483-499.

Gopalakrishnan, K. (2008). Evaluation of accelerated deterioration in NAPTF flexible test roads. *Journal of Zhejiang University: Science A* 9(9): pp 1157-1166.

Gopalakrishnan K., Agrawal A., Ceylan H., Kim S., and Choudhary A. (2013). Knowledge discovery and data mining in road inverse analysis. *Transport*, 28(1): pp 1-10.

Gopalakrishnan, K., and Papadopoulos, H. (2011). Reliable road back-calculation with confidence estimation. *Scientia Iranica*, 18(6): pp 1214-1221.

Gopalakrishnan, K. and Thompson, M.R. (2004). Back-calculation of airport flexible road non-linear moduli using artificial neural networks. *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2004 Volume 2*: pp 652-657.

Gucunski, N., Abdallah, I.N., Nazarian, S. (2000). ANN back-calculation of road profiles from the SASW test. *Proceedings of Sessions of Geo-Denver 2000 - Road Subgrade, Unbound Materials, and Nondestructive Testing*, GSP 98 Volume 286, 5 - 8 August 2000, Denver, CO; United States: pp 31-50,

Gupta A., Kumar P. and Rastogi R. (2011). Road deterioration and maintenance model for low volume roads. *International Journal of Road Research and Technology*, 4(4): pp195-202.

Hagan MT, Menhaj MB. (1994). Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5(6): pp 989-993.

Hajek, J.J., Haas, R., Chong, G.J. and Phang, W.A. (1987). ROSE: A knowledge-based expert system for routing and sealing. *Proceedings of the 2nd North American Road Conference*. Toronto, Canada, pp 2.301 – 2.341.

Harral C. and Faiz A. (1988). *Road Deterioration in Developing Countries - Causes and Remedies*. A World Bank Policy Study. Washington, DC: World Bank.



Hassan, R.A. and McManus, K.J. (2001). Truck ride number: a new management tool. 20th ARRB Conference, 19 – 21 March 2001, Melbourne, Victoria, Australia.

Haykin, S. (2009). *Neural networks and learning machines*. Pearson Education, Inc. New Jersey, USA.

HBM (Online). HBM SoMat eDaqlite modular rugged mobile data acquisition system. Accessed from: <http://www.hbm.com/en/menu/products/measurement-electronics-software/rugged-data-acquisition/somat-edaqlite/>

Date: 18/04/2014.

Heggie I.G., (2003). Commercializing management and financing of roads in developing and transition countries, *Transport Reviews*, 23:2, 139 – 160.

Heiler, M., and McNeil, S. (1997). Interpreting ground penetrating radar road data. *Proceedings of the Speciality Conference on Infrastructure Condition Assessment: Art, Science, Practice*. 25 - 27 August 1997, Boston, MA, USA: pp 208-217.

Heimann, B., Bouzid, N., and Trabelsi, A. (2006). Road-wheel interaction in vehicles. – A mechatronic view of friction, *2006 IEEE International Conference on Mechatronics*, ICM2006, Article No. 4018346, 3-5 July, pp 137-143.

Heyns T., Heyns, P.S., De Villiers J.P. (2012) A method for real-time condition monitoring of haul roads based on Bayesian parameter estimation. *Journal of Terramechanics* 49: pp103 – 113.

Huang, Y., and Moore, R.K. (1997). Roughness level probability prediction using artificial neural networks. *Transportation Research Record* 1592: pp 89 – 97.

Hugo D. (2005). *Haul road defect identification and condition assessment using measured truck response*, MEng Thesis, Department of Mechanical and Aeronautical Engineering, University of Pretoria, South Africa.

Hugo D., Heyns P.S., Thompson R.J., Visser A.T. (2008). Haul road defect identification and condition assessment using measured truck response, *Journal of Terramechanics*, 45, pp79 – 88.

International Organization for Standardization ISO 8608: 1995(E), (1995). Mechanical vibration – Road surface profiles – Reporting of measured data.



International Organization for Standardization ISO 5982:2001, Mechanical Vibration and Shock – Range of idealized values to characterize seated-body biodynamic response under vertical vibration, 2001, Geneva.

International Road Fund (IRF) Research Council (2007). *The Socio-Economic benefits of roads in Europe*. November 2007 Edition.

Available at: http://www.ptcarretera.es/assets/files/the_socio_economic_benefits_of_roads_in_europe.pdf. [Accessed: 10 April 2014].

Johannesson, P., Podgórski, K. and Rychlik, I. (2014). Modelling roughness of road profiles on parallel tracks using roughness indicators. *Chalmers Publication Library: Matematiska Vetenskaper*. Preprint: 4: ISSN: 1652-9715. <http://www.math.chalmers.se/Math/Research/Preprints/2014/4.pdf>.

Kamimura, R. (2003). Information-theoretic competitive learning with inverse Euclidean distance output units. *Neural Processing Letters* 18(3): pp 163-184.

Kang D.-K., Lee S.-H., Goo S.-H. (2007). Development of standardization and management system for the severity of unpaved test courses. *Sensors* 7: pp 2004-2027, <http://www.mdpi.org/sensors/papers/s7092004.pdf>.

Kargah-Ostadi, N., Stoffels, S.M. and Tabatabaee, N. (2010). Network-level road roughness prediction model for rehabilitation recommendations. *Transportation Research Record* 2155: pp 124-133.

Kaseko, M.S., Lo, Z.-P. and Ritchie, S.G. (1994). Comparison of traditional and neural classifiers for road-crack detection. *Journal of Transportation Engineering*, 20(4), pp 552 – 569.

Kaseko M.S. and Ritchie S.G. (1993). A neural network-based methodology for road crack detection and classification. *Transportation Research Part C*, 1(4), pp 275 – 291.

Kaseko, M.S. and Ritchie, S.G. (1991). Road image processing using neural networks. Proceedings of the 2nd International Conference on Application of Advanced Technologies in Transportation Engineering, Code 15135, 18 - 21 August 1991, Minneapolis USA.

Kim, S. and Gopalakrishnan, K. (2010). Sustainable rehabilitation of deteriorated concrete highways: Condition assessment using Shuffled Complex Evolution (SCE) global optimization approach.



Sustainable and Resilient Critical Infrastructure Systems: Simulation, Modeling, and Intelligent Engineering: pp 249-265.

Kim, S., Gopalakrishnan, K. and Ceylan, H. (2012). Structural Characterisation of Iowa's Rubblized PCC Roads. *Journal of Transportation Engineering* 138(4): pp 406-413.

Kim, Y.R., Xu, B. and Kim, Y. (1999). A new back-calculation procedure based on dispersion analysis of FWD time-history deflections and surface wave measurements using artificial neural networks. *ASTM Special Technical Publication Issue 1375*: pp 297-312

Kong, L., Song, J., Yan, Y., and Shen, J. (2007). Algorithm of feature recognition of road roughness for anti-lock brake system. *Jixie Gongcheng Xuebao/Chinese Journal of Mechanical Engineering* 43(11): pp 229-234.

Kongrattanaprasert, W., Nomura, H., Kamakura, T., and Ueda, K. (2009). Application of neural network analysis to automatic detection of road surface conditions utilizing tire noise from vehicles. *ICCAS-SICE 2009 - ICROS-SICE International Joint Conference 2009*, Article number 5335122: pp 2354 – 2358.

La Torre, F., Ballerini, L. And Di Volo, N. (2002). Correlation between longitudinal roughness and user perception in urban areas. *Transportation Research Record* 1806(1): pp 131 – 139.

La Torre, F., Domenichini, L. and Darter, M.I. (1998). Roughness prediction model based on the artificial neural network approach. *4th International Conference on Managing Roads Vol. 2*

LeCun, Y. (1993). *Efficient learning and second-order methods*. A Tutorial at NIPS 93, Denver.

Letherwood M.D., Gunter D.D. (2001). Ground vehicle modelling and simulation of military vehicles using high performance computing, *Parallel Computing* 27: pp 109-140.

Li, Q., Wei, H. and Kurt, C. E. (1997). Road distress level prediction using multi-BPANNS techniques. *Proceedings of the 1997 4th Congress on Computing in Civil Engineering*; 16 - 18 June 1997, Philadelphia, PA, USA: pp 49-56.

Liu, C., Gazis, D.C. and Kennedy, T.W. (1999). Human judgement and analytical derivation of ride quality. *Transportation Science*, No. 3: pp 290 – 297.



Ljung, L. (1999). *System identification: Theory for the user*. Prentice Hall PTR, New Jersey, USA.

Lombard P. and Coetzer L. (2007). Estimating the impact of rural-road investments on socio-economic development. *International Seminar on Sustainable Road Financing & Investment*. 16-20 April 2007, Arusha, Tanzania.

Available at: <http://www.org/resources/documents/1026,Paul-Lombard-Session-3.pdf>

[Accessed: 10 April 2014]

Mallick, R.B., Bradley, J.E., Nazarian, S. (2006). In-place determination of stiffness of subsurface reclaimed layers in thin surface hot-mix asphalt roads: Fast non-destructive test and analysis method. *Transportation Research Record* Issue 1949: pp 11-19.

Mathworks Inc., (2007). MATLAB Neural Network Toolbox Help Tutorial. Available online at <http://filer.case.edu/pjt9/b378s10/nnet.pdf> accessed on 26/06/2013.

Meier, R.W., Alexander, D.R. and Freeman, R.B. (1997). Using artificial neural networks as a forward approach to back-calculation. *Transportation Research Record* Issue 1570: pp 126-133.

Meier, R.W. and Rix, G.J. (1994). Back-calculation of flexible road moduli using artificial neural networks. *Transportation Research Record*. Issue 1448: pp 75-82.

Miradi, M. and Molenaar, A.A.A. (2006). Application of Artificial Neural Network (ANN) to PA lifespan: Forecasting models. *Conference Proceedings 2006 IEEE International Conference on Neural Networks*, Article Number 1716604, 16 - 21 July 2006, Vancouver, BC; Canada: pp 3679-3685.

Miradi, M., Molenaar, A.A.A., Van De Ven, M.F.C. and Molenaar, S. (2009). Back-calculation of the stiffnesses of cement treated base courses using artificial intelligence. *Proceedings of the 8th International Conference on the Bearing Capacity of Roads, Railways and Airfields* Volume 1: pp 633-642.

Mokhtarzade, M. and Zoej, M.J.V. (2007). Road detection from high-resolution satellite images using artificial neural networks. *International Journal of Applied Earth Observation and Geoinformation* 9(1): pp 32-40.

Moody, J. and Darken, C.J. (1989). Fast learning in networks of locally-tuned processing units. *Neural Computation* 1(2): pp 281 – 294.



Mushule N.K. and Kerali H.R. (2001). Implementation of new highway management tools in developing countries: A case study of Tanzania. *In Transportation Research Record 1769*, 01-0277, pp 51-60, TRB, National Research Council, National Academy Press. Washington, D.C.

Najjar, Y. and Felker, V. (2006). Characterizing the roughness progression on Kansas Roads: A sequential neuronet approach. *GeoCongress 2006: Geotechnical Engineering in the Information Technology Age*. Volume 2006, 26 February - 1 March 2006; Atlanta, GA; United States: page 127.

Nallamotheu, S. and Wang, K.C.P. (1996). Experimenting with recognition accelerator for road distress identification. *Transportation Research Record Issue 1536*: pp 130-135.

Ngwangwa HM, Heyns PS, Labuschagne FJJ, Kululanga GK. (2008). Monitoring of heavy vehicle dynamic response for road condition monitoring: An overview of the web based artificial intelligence technique, *Proceedings of the Southern African Transport Conference (SATC2008)*, 7-10 July 2008, Pretoria.

Ogden K.W. (1997). *Safer roads: A guide to road safety engineering*. Avebury Technical: Cambridge, England.

Oh, I. and Barham, W. (2011). The application of artificial neural network for the prediction of the deformation performance of hot-mix asphalt. *Proceedings of the Congress on Computing in Civil Engineering*: pp 227-233.

Park, H.I., Kweon, G.C. and Lee, S.R. (2009). Prediction of resilient modulus of granular subgrade soils and sub-base materials using artificial neural network. *Road Materials and Road Design* 10(3): pp 647-665.

Paterson W.D.O. (1987). Road deterioration and maintenance effects: Models for planning and management. A World Bank Publication: The Highway Design and Maintenance Series, The Johns Hopkins University Press: Washington D.C.

Pekcan, O., Tutumluer, E. and Ghaboussi, J. (2009). SOFTSYS for back-calculation of full-depth asphalt road layer moduli. *Proceedings of the 8th International Conference on the Bearing Capacity of Roads, Railways and Airfields BCR2A'09*; Volume 1, 29 June - 2 July 2009, Champaign, IL; United States: pp 679-687.



Pekcan, O., Tutumluer, E. and Thompson, M.R. (2008). Artificial Neural Network based back-calculation of conventional flexible roads on lime stabilized. *12th International Conference on Computer Methods and Advances in Geomechanics*, Volume 3, 6 October 2008, Goa; India.: pp 1647-1654.

Peshkin D.G., Hoerner T.E, and Zimmerman K.A., (2004). *Optimal Timing of Road Preventive Maintenance Treatment Applications*. National Cooperative Highway Research Program Report 523. Transportation Research Board National Research: Washington D.C.

Pesterev, A.V., Bergman, L.A. and Tan, C.A. (2004). A novel approach to the calculation of pothole-induced contact forces in MDOF vehicle models. *Journal of Sound and Vibration* 275(1-2). pp 127 – 149.

Pesterev, A.V., Bergman, L.A., Tan, C.A. and Yang, B. (2005). Application of the pothole DAF method to vehicles traversing periodic roadway irregularities. *Journal of Sound and Vibration* 279(3-5). pp 843 – 855.

Potter, D., Hannay, R., Cairney, P.T. and Markarov, A. (1992). An investigation of car users' perceptions of the quality of roads, *Australian Research Board: Road and Transport Research* 1(1): pp 6.

Powell, M.J.D. (1987). Radial basis functions for multivariate interpolation: a review. In book: *Algorithms for approximation*. Clarendon Press, NY, USA. ISBN: 0-19-853612-7: pp 143 – 167.

Rakesh, N., Jain, A.K., Reddy, M.A., and Reddy, K.S. (2006). Artificial neural networks – Genetic algorithm based model for back-calculation of road layer moduli. *International Journal of Road Engineering* 7(3): pp 221-230.

Rao, D.V., Jian, P., Mohamad, Q.S., Gang, S., and Zuo, S. (2008). Road vehicle dynamics. SAE International.

Roberts, C.A. and Attoh-Okine, N.O. (1998). A comparative analysis of two artificial neural networks using road performance prediction. *Computer Aided Civil and Infrastructure Engineering* 13(5): pp 339 – 348.

Rostow W.W. (1962). *Process of economic growth*. Norton & Company, Incorporated, W.W. ISBN-10:0393001768.



Saghafi, B., Hassani, A., Noori, R., and Bustos, M.G. (2009). Artificial neural networks and regression analysis for predicting faulting in jointed concrete roads considering base condition. *International Journal of Road Research and Technology* 2(1): pp 20-25.

Sakhaeifar, M.S., Underwood, B.S., Kim, Y.R., Puccinelli, J. and Jackson, N. (2010). Development of artificial neural network predictive models for populating dynamic moduli of long-term road performance Sections. *Transportation Research Record* 2181: pp 88-97.

Saltan, M. and Sezgin, H. (2007). Hybrid neural network and finite element modeling of sub-base layer material properties in flexible roads. *Materials and Design* 28(5): pp 1725-1730.

Saltan, M. and Terzi, S. (2004). Backcalculation of road layer parameters using artificial neural networks. *Indian Journal of Engineering and Materials Sciences* 11(1): pp 38-42.

Saltan, M. and Terzi, S. (2005). Comparative analysis of using artificial neural networks (ANN) and gene expression programming (GEP) in back-calculation of road layer thickness. *Indian Journal of Engineering and Materials Sciences* 12(1): pp 42-50.

Saltan, M. and Terzi, S. (2008). Modeling deflection basin using artificial neural networks with cross-validation technique in back-calculating flexible road layer moduli. *Advances in Engineering Software* 39(7): pp 588-592.

Saltan, M., Tiğdemir, M. and Karaşahin, M. (2002). Artificial neural network application for flexible road thickness modeling. *Turkish Journal of Engineering and Environmental Sciences* 26(3): pp 243-248.

Saltan, M., Uz, V.E. and Aktas, B. (2013). Artificial neural networks-based back-calculation of the structural properties of a typical flexible road. *Neural Computing and Applications* 23(6): pp 1703-1710.

Sayers, M.W, Gillespie, T.D., Paterson, W.D.O. (1986). Guidelines for Conducting and Calibrating Road Roughness Measurements. World Bank Technical Paper Number 46. The World Bank, Washington, D.C.

Sayers, M.W., Karamihas, S.M. (1998). *The little book of profiling: Basic information about measuring and interpreting road profiles*. University of Michigan.



Shaaban, S.M., Nabwey, H.A. (2012) Rehabilitation and reconstruction of asphalt road decision making based on rough set theory. *12th International Conference on Computational Science and Its Applications*, ICCSA 2012; 18 -21 June 2012. Salvador de Bahia; Brazil.

Shahnazari, H., Tutunchian, M.A., Mashayekhi, M. and Amini, A.A. (2012). Application of soft computing for prediction of road condition index. *Journal of Transportation Engineering* 138(12): pp 1495-1506.

Shalaby, A. and Reggin, A. (2007). Optimization of data collection needs for manual and automated network-level road condition ratings based on transverse variability and neural networks. *Canadian Journal of Civil Engineering* 34(2): pp 139-146.

Sharma, S. and Das, A. (2008). Back-calculation of road layer moduli from falling weight deflectometer data using an artificial neural network. *Canadian Journal of Civil Engineering* 35(1): pp 57-66.

Shehab, T. and Meisami-Fard, I. (2013). Cost-estimating model for rubblized asphalt road rehabilitation projects. *Journal of Infrastructure Systems* 19(4): pp 496-502.

Sirvio, K. and Hollmén, J. (2008). Spatio-temporal road condition forecasting with Markov chains and artificial neural networks. *3rd International Workshop on Hybrid Artificial Intelligence Systems*, HAIS 2008; Burgos; Spain; 24 - 26 September 2008; pp 204-211.

Stein GJ, Múčka P. (2003). Theoretical investigation of a linear planar model of a passenger car with seated people. *Proc. IMechE Part D: J. Automobile Engineering* Vol. 217: pp 257 – 268.

Sundin, S. and Braban-Ledoux, C. (2001). Artificial intelligence-based decision support technologies in road management. *Computer-Aided Civil and Infrastructure Engineering* 16(2): pp 143-157.

Taddesse, E. and Mork, H. (2009). Roughness progression models by regression and artificial neural network techniques. *Proceedings of the 8th International Conference on the Bearing Capacity of Roads, Railways and Airfields*. Volume 1: pp 521-530.

Taha, M.A. and Hanna, A.S. (1995). Evolutionary neural network model for the selection of road maintenance strategy. *Transportation Research Record* 1497: pp 70 – 76.

Terzi, S. (2007). Modeling the road serviceability ratio of flexible highway roads by artificial neural networks. *Construction and Building Materials* 21(3): pp 590-593.



Thompson R.J., Visser A.T. (2003). Mine haul road maintenance management systems. *The South African Institute of Mining and Metallurgy*: pp 303–312.

Thompson R.J., Visser A.T., Miller R.E., Lowe N.T. (2003). Development of real-time mine road maintenance management system using haul truck and road vibration signature analysis. *Transportation Research Record*, 1819A: pp305 – 312.

Thube, D.T. (2012). Artificial Neural Network (ANN) based road deterioration models for low volume roads in India. *International Journal of Road Research and Technology* 5(2): pp 115-120.

Tillotson, H.T., Snaith, M.S., Huang, Y. (1998). Detecting cracks by image analysis on a parallel computer. Proceedings of the 1998 International Computing Congress on Computing in Civil Engineering, 18 - 21 October 1998, Boston, MA, USA: pp 11-23.

Tsai, F.-L., Lytton, R.L. and Lee, S. (2010). Prediction of reflection cracking in hot-mix asphalt overlays. *Transportation Research Record* 2155: pp 43-54.

Tutumluer, E. and Seyhan, U. (1998). Neural network modeling of anisotropic aggregate behavior from repeated load tri-axial tests. *Transportation Research Record*. Issue 1615: pp 86-93.

Uys P.E., Els P.S., Thoresson M.J., Voigt K.G., Combrinck W.C. (2006). Experimental determination of moments of inertia for an off-road vehicle in a regular engineering laboratory. *International Journal of Mechanical Engineering Education* 34(4): pp 291-314.

Van der Gryp, A., Bredenhann, S.J., Henderson, M.G. and Rohde, G.T. (1998). Determining the visual condition index of flexible roads using artificial neural networks, *4th International Conference on Managing Roads*, Durban, South Africa.

Wei L, Griffin MJ. (1998). The prediction of seat transmissibility from measures of seat impedance. *Journal of Sound and Vibration* 214(1): pp 121-137.

Wong, C.X., Worden, K. (2007). Generalised NARX shunting neural network modelling of friction. *Mechanical Systems and Signal Processing* 21: pp 553-572.

Worden, K., Tomlinson, G.R. (2001). *Nonlinearity in structural dynamics: Detection, identification and modelling*. Institute of Physics Publishing, Bristol, UK.



Xiao, F., Amirkhani, S. and Juang, H.C. (2009). Prediction of fatigue life of rubberized asphalt concrete mixtures containing reclaimed asphalt road using artificial neural networks. *Journal of Materials in Civil Engineering* 21(6): pp 253-261.

Xie, H., Guo, Z., and Cong, L. (2007). Artificial neural network-based evaluation of layer modulus of asphalt road. *Tongji Daxue Xuebao/Journal of Tongji University* 35(8): pp 1044-1047.

Xiong, H.-J. (2001). Neural network diagnostic methods for road surface distress. *Gongku Jiaotong Keji/Journal of Highway and Transportation Research and Development*. 18(1): pp 10-12.

Xu, G., Ma, J., Liu, F. and Niu, X. (2008). Automatic recognition of road surface crack based on BP neural network. *2008 International Conference on Computer and Electrical Engineering*, Article Number 4740938, 20 - 22 December 2008, Phuket; Thailand: pp 19-22.

Xu, B., Ranjithan, S.R. and Kim, Y.R. (2002). New relationships between falling weight deflectometer deflections and asphalt road layer condition indicators. *Transportation Research Record*. Issue 1806: pp 48-56.

Yang, J., Lu, J.J., Gunaratne, M. and Xiang, Q. (2003). Forecasting overall road condition with neural networks: Application on Florida Highway Network. *Transportation Research Record*, 1853: pp 3-12.

Yildirim, Ş. and Uzmay, I. (2001). Statistical analysis of vehicles' vibration due to road roughness using Radial Basis artificial Neural Network. *Applied Artificial Intelligence* 15(4): pp 419 – 427.

Yousefzadeh, M., Azadi, S., and Soltani, A (2010). Road profile estimation using neural network algorithm. *Journal of Mechanical Science and Technology*, 24(3): pp 743 – 754.

Yu, X. and Salari, E. (2011). Road pothole detection and severity measurement using laser imaging. *IEEE International Conference on Electro Information Technology 2011*, Article Number 5978573, 15 - 17 May 2011, Mankato, MN; United States.

Zakeri, H, Nejad, F.M., Fahimifar, A., Torshizi, A.D. and Zarandi, M.H.F. (2013). A multi-stage expert system for classification of road cracking. *Proceedings of the 2013 Joint IFSA World Congress and NAFIPS Annual Meeting, IFSA/NAFIPS 2013*, Article Number 6608558,24 - 28 June 2013, Edmonton, AB; Canada; pp 1125-1130.



Zeghal, M. (2008). Thermal cracking prediction using artificial neural network. Road cracking: Mechanisms, Modeling, Detection, Testing and Case Histories; *6th RILEM International Conference on Cracking in Roads*, 16 - 18 June 2008; Chicago, IL; United States: pp 379-386.

Zhang, X.-D., Li, D.-W. and Zhang, L. (2007). Experimental investigation and simulation for road surface base material in aeolian area. *Liaoning Gongcheng Jishu Daxue Xuebao (Ziran Kexue Ban)/Journal of Liaoning Technical University (Natural Science Edition)* 26(3): pp 366-368.



Appendices

Appendix A

The dynamic equations of motion for the system in Figure 6.2 may be derived by considering the free-body diagrams. The system has a total of 8 DOFs comprising: sprung mass vertical displacement z_s , sprung mass pitch θ_s , front axle vertical displacement z_{uf} , rear axle vertical displacement z_{ur} , seat displacement z_0 , and displacements on the driver's body (z_1, z_2, z_3). In the following, dynamic equations are formulated for each mass.

The formulation of sprung mass motions requires the kinematic relations in equations (6.2) to (6.5) presented under Section 6.3. So the dynamics of the sprung mass can be written as

$$M_s \ddot{z}_s + k_{sf} (z_{sf} - z_{uf}) + c_{sf} (\dot{z}_{sf} - \dot{z}_{uf}) + k_{sr} (z_{sr} - z_{ur}) + c_{sr} (\dot{z}_{sr} - \dot{z}_{ur}) + k_0 (z_{sd} - z_0) + c_0 (\dot{z}_{sd} - \dot{z}_0) = f_{ed} \quad (\text{A.1})$$

$$I_s \ddot{\theta}_s - a_f k_{sf} (z_{sf} - z_{uf}) - a_f c_{sf} (\dot{z}_{sf} - \dot{z}_{uf}) + a_r k_{sr} (z_{sr} - z_{ur}) + a_r c_{sr} (\dot{z}_{sr} - \dot{z}_{ur}) - d_r k_0 (z_{sd} - z_0) - d_r c_0 (\dot{z}_{sd} - \dot{z}_0) = -d_{ed} f_{ed} \quad (\text{A.2})$$

For the front and rear axles, the equations of motion are

$$m_{uf} \ddot{z}_{uf} + k_{sf} (z_{uf} - z_{sf}) + c_{sf} (\dot{z}_{uf} - \dot{z}_{sf}) + k_{tf} (z_{uf} - z_{rf}) + c_{tf} (\dot{z}_{uf} - \dot{z}_{rf}) = f_{wf} \quad (\text{A.3})$$

$$m_{ur} \ddot{z}_{ur} + k_{sr} (z_{ur} - z_{sr}) + c_{sr} (\dot{z}_{ur} - \dot{z}_{sr}) + k_{tr} (z_{ur} - z_{rr}) + c_{tr} (\dot{z}_{ur} - \dot{z}_{rr}) = f_{wr} \quad (\text{A.4})$$

The equations for the seat and driver are

$$m_0 \ddot{z}_0 + k_0 (z_0 - z_{sd}) + c_0 (\dot{z}_0 - \dot{z}_{sd}) + k_1 (z_0 - z_1) + c_1 (\dot{z}_0 - \dot{z}_1) + k_3 (z_0 - z_3) + c_3 (\dot{z}_0 - \dot{z}_3) = 0 \quad (\text{A.5})$$

$$m_1 \ddot{z}_1 + k_1 (z_1 - z_0) + c_1 (\dot{z}_1 - \dot{z}_0) + k_2 (z_1 - z_2) + c_2 (\dot{z}_1 - \dot{z}_2) = 0 \quad (\text{A.6})$$

$$m_2 \ddot{z}_2 + k_2 (z_2 - z_1) + c_2 (\dot{z}_2 - \dot{z}_1) = 0 \quad (\text{A.7})$$

$$m_3 \ddot{z}_3 + k_3 (z_3 - z_0) + c_3 (\dot{z}_3 - \dot{z}_0) = 0 \quad (\text{A.8})$$



Upon making the necessary substitutions for the kinematic equations in the relevant equations above, and grouping like terms according to degrees of freedoms and their derivatives, the above equations can be rewritten in the compact matrix form of equation (6.1), where the property matrices and force vector are given by equations (A.9) to (A.11).

$$\mathbf{M} = \text{diag} \left(\left[M_s \quad I_s \quad m_{uf} \quad m_{ur} \quad m_0 \quad m_1 \quad m_2 \quad m_3 \right] \right) \quad (\text{A.9})$$

$$\mathbf{K} = \begin{bmatrix} k_{zs1} & k_{\theta s1} & k_{zuf1} & k_{zur1} & k_{z0,1} & & & & \\ k_{zs2} & k_{\theta s2} & k_{zuf2} & k_{zur2} & k_{z0,2} & & & & \\ k_{zs3} & k_{\theta s3} & k_{zuf3} & & & & & & \\ k_{zs4} & k_{\theta s4} & & k_{zur4} & & & & & \\ k_{zs5} & k_{\theta s5} & & & k_{z0,5} & k_{z1,5} & & & k_{z3,5} \\ k_{zs6} & k_{\theta s6} & & & k_{z0,6} & k_{z1,6} & k_{z2,6} & & \\ & & & & & k_{z1,7} & k_{z2,7} & & \\ & & & & k_{z0,8} & & & & k_{z3,8} \end{bmatrix} \quad (\text{A.10})$$

where all elements with zero values are not indicated and the matrix elements are given by the following relationships:

$$\begin{aligned} k_{zs1} &= k_{sf} + k_{sr} + k_0, & k_{\theta s1} &= -k_{sf}a_f + k_{sr}a_r - k_0d_r \\ k_{zuf1} &= -k_{sf}, & k_{zur1} &= -k_{sr}, & k_{z0,1} &= -k_0 \\ k_{zs2} &= -k_{sf}a_f + k_{sr}a_r - k_0d_r, & k_{\theta s2} &= k_{sf}a_f^2 + k_{sr}a_r^2 + k_0d_r^2 \\ k_{zuf2} &= k_{sf}a_f, & k_{zur2} &= -k_{sr}a_r, & k_{z0,2} &= k_0d_r \\ c_{zs3} &= -c_{sf}, & c_{\theta s3} &= c_{sf}a_f, & c_{zuf3} &= c_{sf} + c_{tf} \\ k_{zs4} &= -k_{sr}, & k_{\theta s4} &= -k_{sr}a_r, & k_{zur4} &= k_{sr} + k_{tf} \\ k_{zs5} &= -k_0, & k_{\theta s5} &= k_0d_r, & k_{z0,5} &= k_0 + k_1 + k_3 \\ k_{z1,5} &= -k_1, & k_{z3,5} &= -k_3 \\ k_{z0,6} &= -k_1, & k_{z1,6} &= k_1 + k_2, & k_{z2,6} &= -k_2 \\ k_{z1,7} &= -k_2, & k_{z2,7} &= k_2, & k_{z0,8} &= -k_3, & k_{z3,8} &= k_3 \end{aligned}$$

The corresponding equations for the damping matrix \mathbf{C} and its elements are similar.

$$\mathbf{f} = [F_1 \quad F_2 \quad F_3 \quad F_4 \quad 0 \quad 0 \quad 0 \quad 0]^T \quad (\text{A.11})$$

where

$$F_1 = f_{ed}, \quad F_2 = -d_{ed}f_{ed}, \quad F_3 = c_{if}\dot{z}_{rf} + k_{if}z_{rf} + f_{wf}, \quad F_4 = c_{tr}\dot{z}_{rr} + k_{tr}z_{rr} + f_{wr}$$



Appendix B

It was very hard to consistently converge to the same or approximately similar weights and more so, to a stable region during different training sessions. In order to deal with this problem in this study, several training sessions were run and typical results are plotted in Figures B.1 – B.6. The specific impacts of each vehicle suspension mode are not clear owing to the relatively higher magnitudes of the RMS errors introduced by the training process itself.

The results shown in Figure 7.10 represent the case of near-optimal performance for the ANN training process. However two more different cases were identified. The first case is where the neural network converged to an overly sub-optimal state where some of the resulting RMS error percentages lay between 15 % and 25 %. In the second case the neural network could not achieve any convergence but it would simply be trapped in an unstable region where the resulting weights yielded impractically high output values (RMS error percentages above 25 %). The overly sub-optimal performance was typified by the results as shown in Figures B.1 – B.4 while Figures B.5 and B.6 represent the unstable performance.

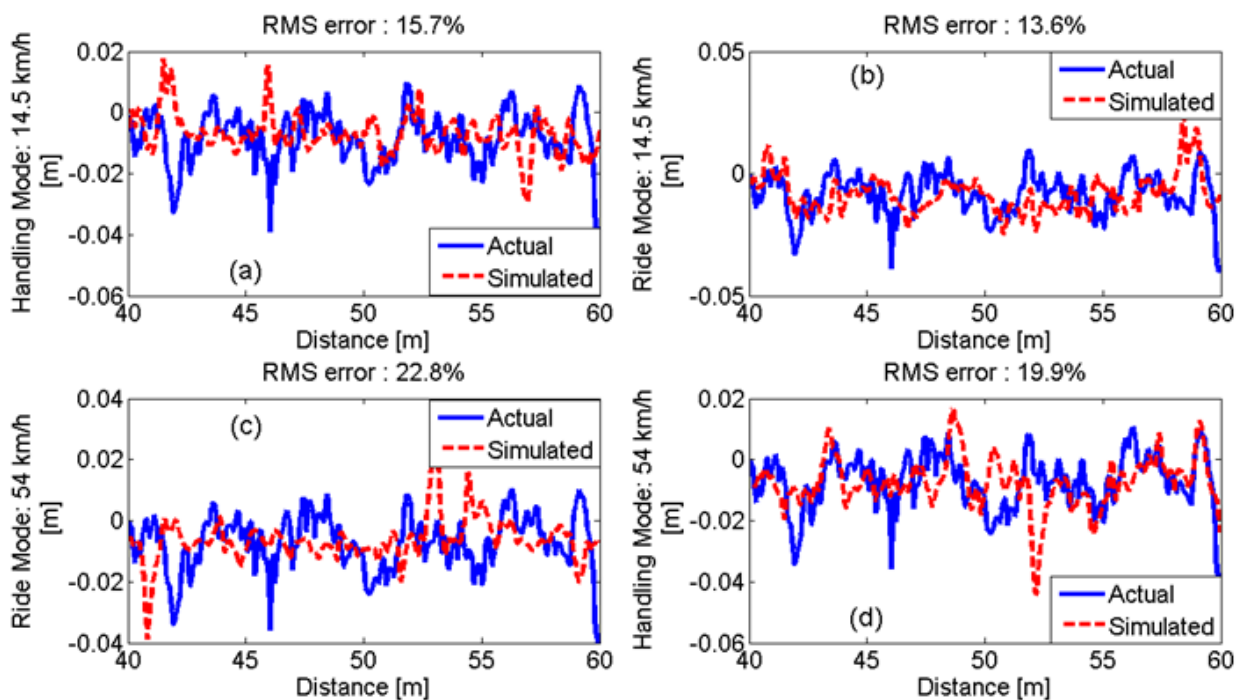


Figure B.1. Simulated profiles (dashed) correlated with Actual profiles (solid) for sub-optimal and stable training process.



Figure B.1 shows results that are similar in nature to Figure 7.10 except the differences in the magnitude of the RMS error percentages. The simulated profiles are better correlated at lower vehicle speeds (Figures B.1(a) and B.1(b)) than higher vehicle speeds (Figures B.1(c) and B.1(d)) for both vehicle suspension modes. The ANN exhibits much better performance for ride comfort mode (Figure B.1(b)) than handling mode (Figure B.1(a)) at lower vehicle speed while, at higher vehicle speed, better ANN performance was obtained from handling mode (Figure B.1(d)) than ride comfort mode (Figure B.1(c)).

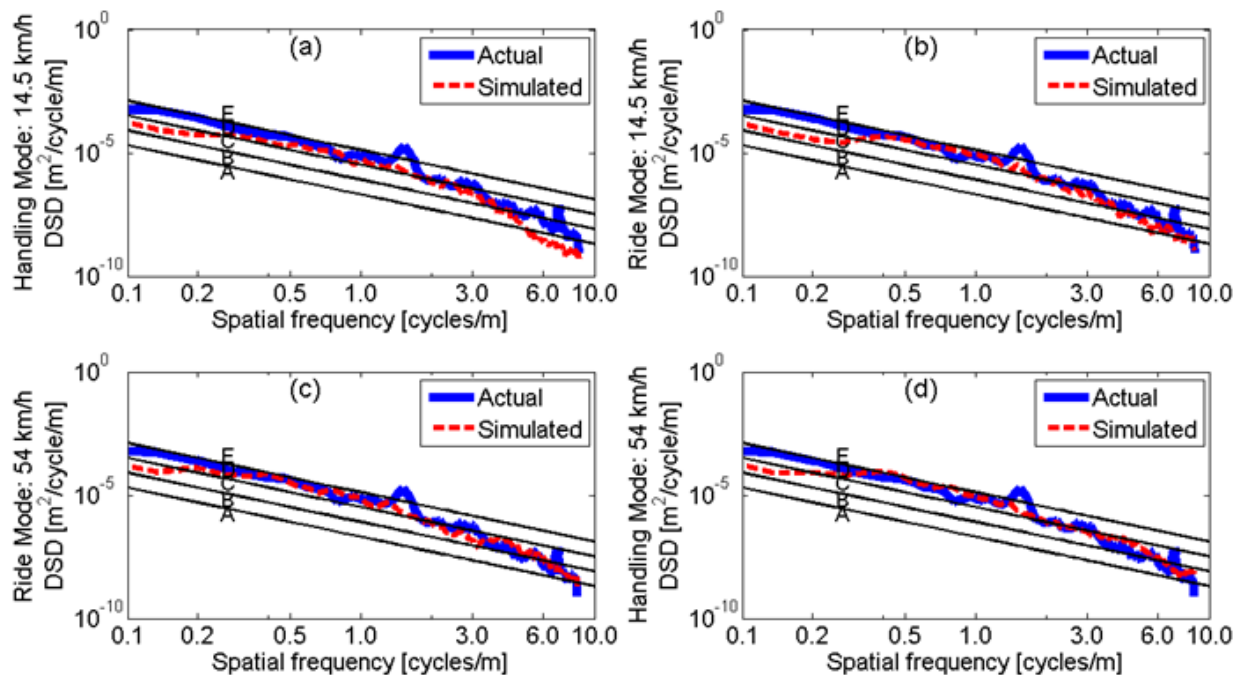


Figure B.2. Simulated DSDs (dashed) correlated with Actual DSDs (solid) for sub-optimal and stable training process.

Despite overly sub-optimal performance in the training process, the DSDs for the simulated profiles in Figure B.2 are very well correlated with the DSDs for actual profiles. There are insignificant differences from the DSDs plotted in Figure 7.11. Similar low frequency errors as shown in Figure 7.11 and already discussed in Section 7.7.2 occur in Figure B.2. This is where the application of ANN and DSD evaluations to road maintenance decision making could be potentially useful. The DSDs are not too sensitive to errors introduced by small amount of sub-optimality in the training process of the ANN.

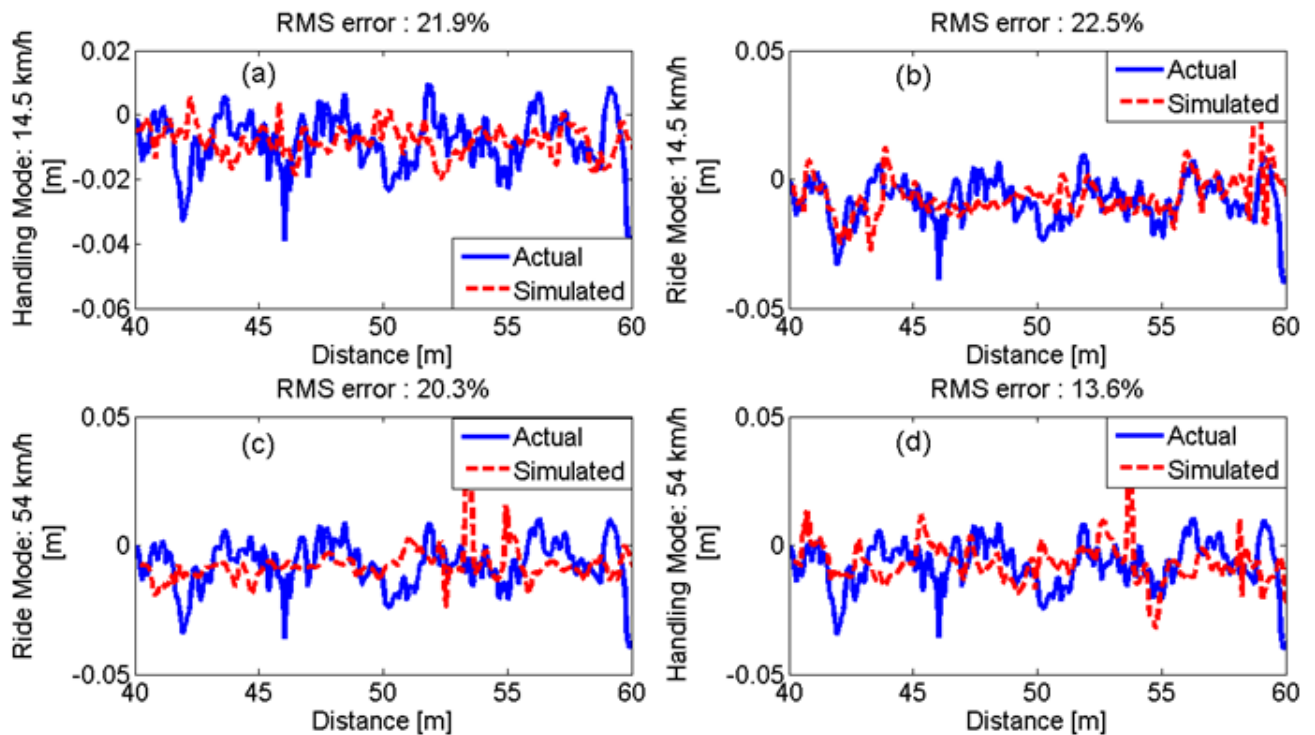


Figure B.3. Simulated profiles (dashed) correlated with Actual profiles (solid) for sub-optimal and stable training process.

Figure B.3 shows different ANN performance from that shown in Figures 7.10 and B.1. The neural network produces better correlated results at higher vehicle speed (Figure B.3(c) and B.3(d)) than at lower vehicle speed (Figure B.3(a) and B.3(b)) and the neural network yields better correlations in handling mode (Figure B.3(a) and B.3 (d)) than ride comfort mode (Figure B.3(b) and B.3(c)) at both vehicle speeds. With the exception of the latter observation, this seems to agree well with the performance of the numerical model itself but as pointed in Section 7.7.2 it might be hard to fathom specific vehicle suspension influences given the magnitude of error induced by the training process itself.

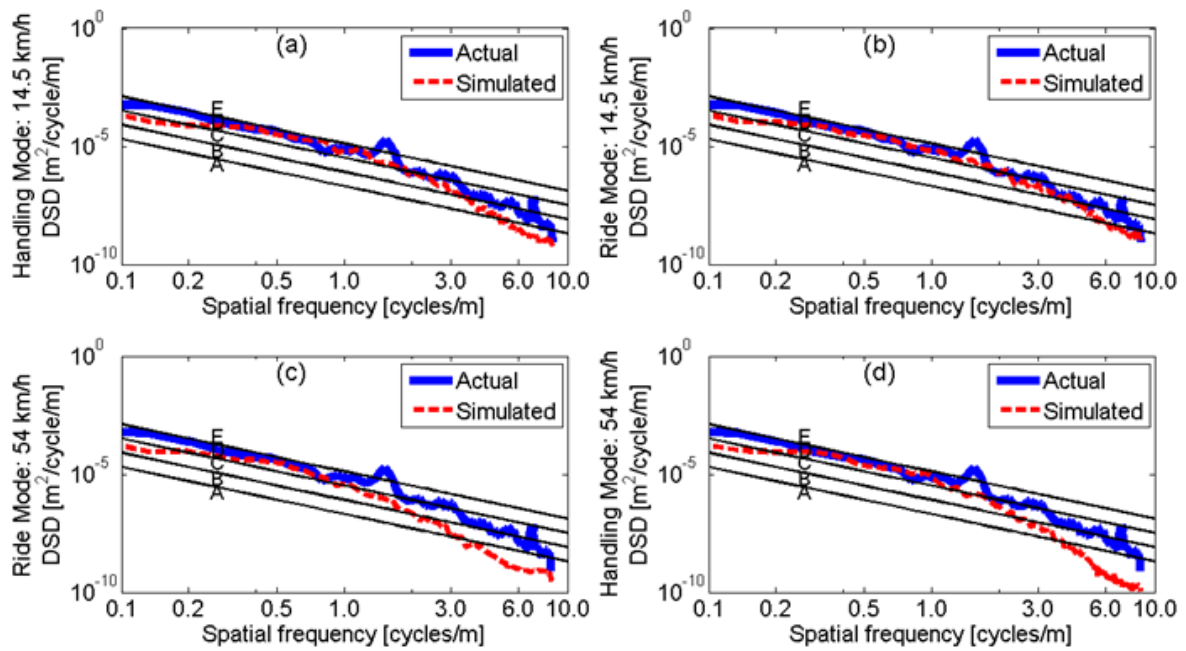


Figure B.4. Simulated DSDs (dashed) correlated with Actual DSDs (solid) for sub-optimal and stable training process.

Figure B.4 exhibits a potential shortcoming of the DSD analysis. Despite the better raw profile correlation at higher vehicle speeds as shown in Figure B.3, the resulting DSDs in Figures B.4(c) and B.4(d) are not as well correlated at higher frequencies (above 1.8 cycles/m). However one can still use the information from mid-frequency range between 0.2 cycles/m to 1.8 cycles/m and then apply a linear extrapolator to the higher frequencies to estimate their corresponding spectral densities.

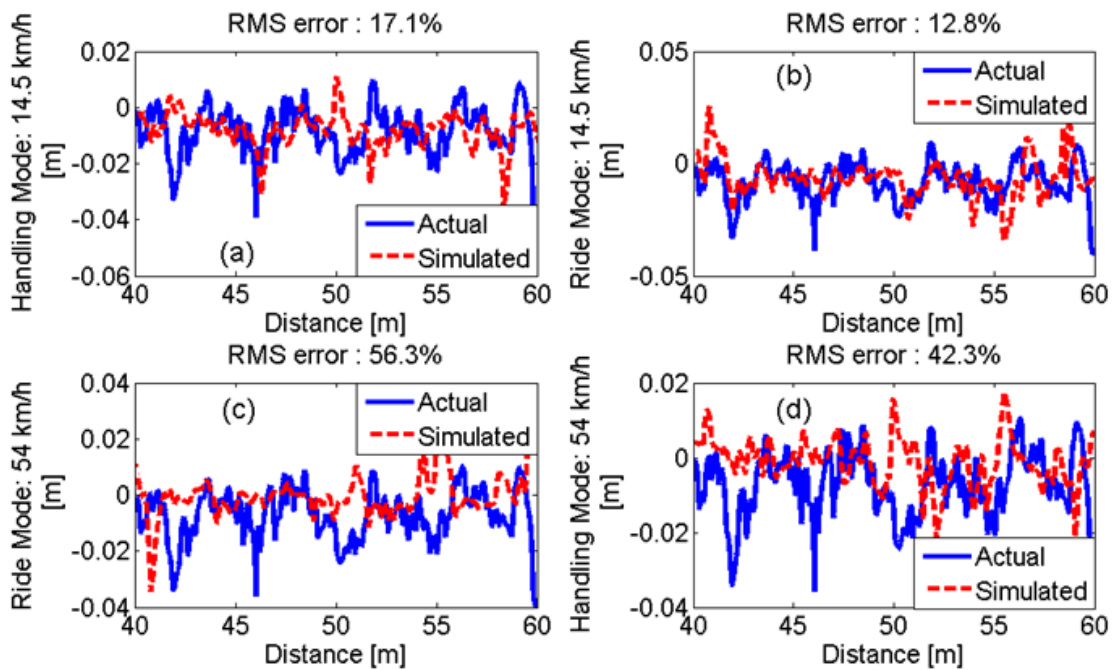


Figure B.5. Simulated profiles (dashed) correlated with Actual profiles (solid) for unstable training process.

Substantial RMS error percentages occurred at higher vehicle speeds for the unstable training process shown in Figure B.5. The ANN performed reasonably well for ride comfort mode at lower speed only (Figure B.5(b)). DSD plots show that the problem affected certain high frequencies (Figures B.6(c) and B.6(d)). The range between 0.2 and 1.8 cycles/m remain unaffected by this problem such that it is also possible to apply the linear extrapolator, as in the previous case, for estimation of the spectral densities at the higher frequencies.

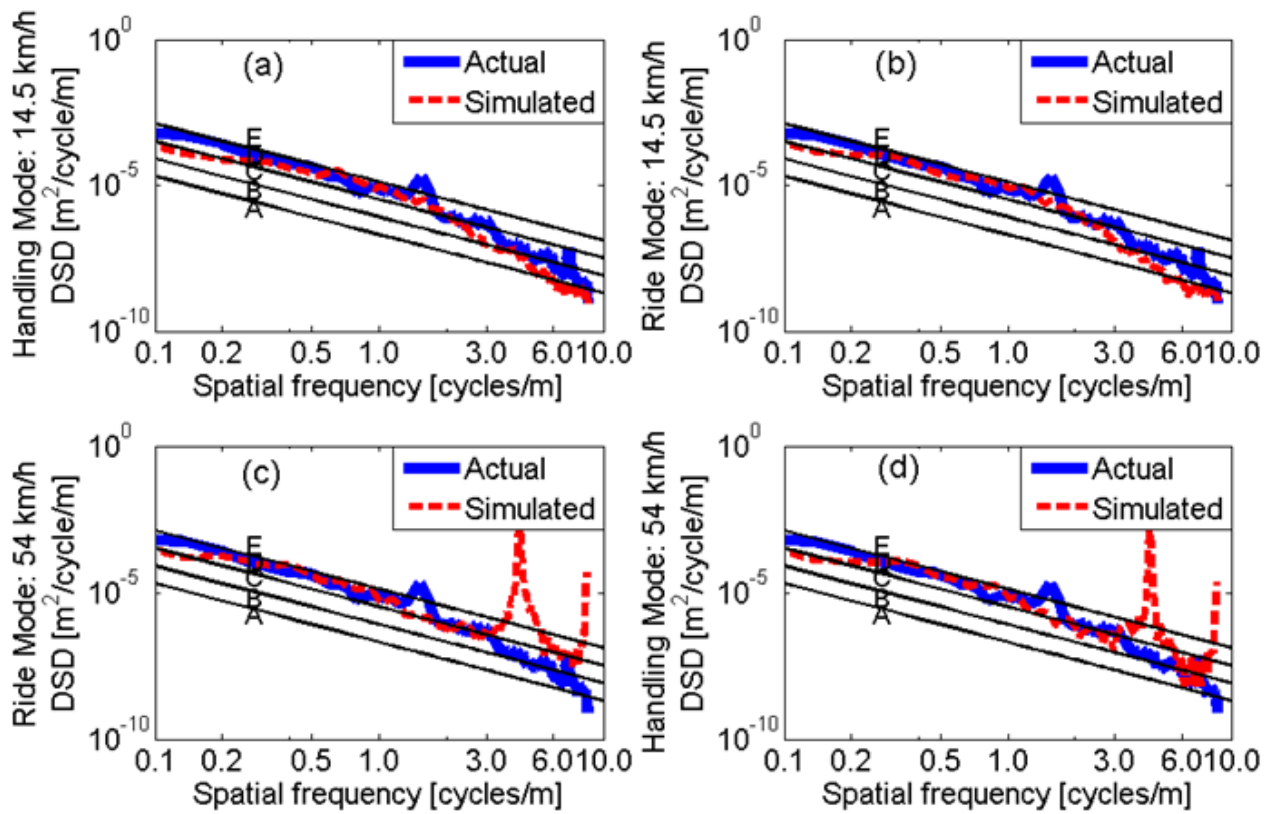


Figure B.6. Simulated DSDs (dashed) correlated with Actual DSDs (solid) for unstable training process.