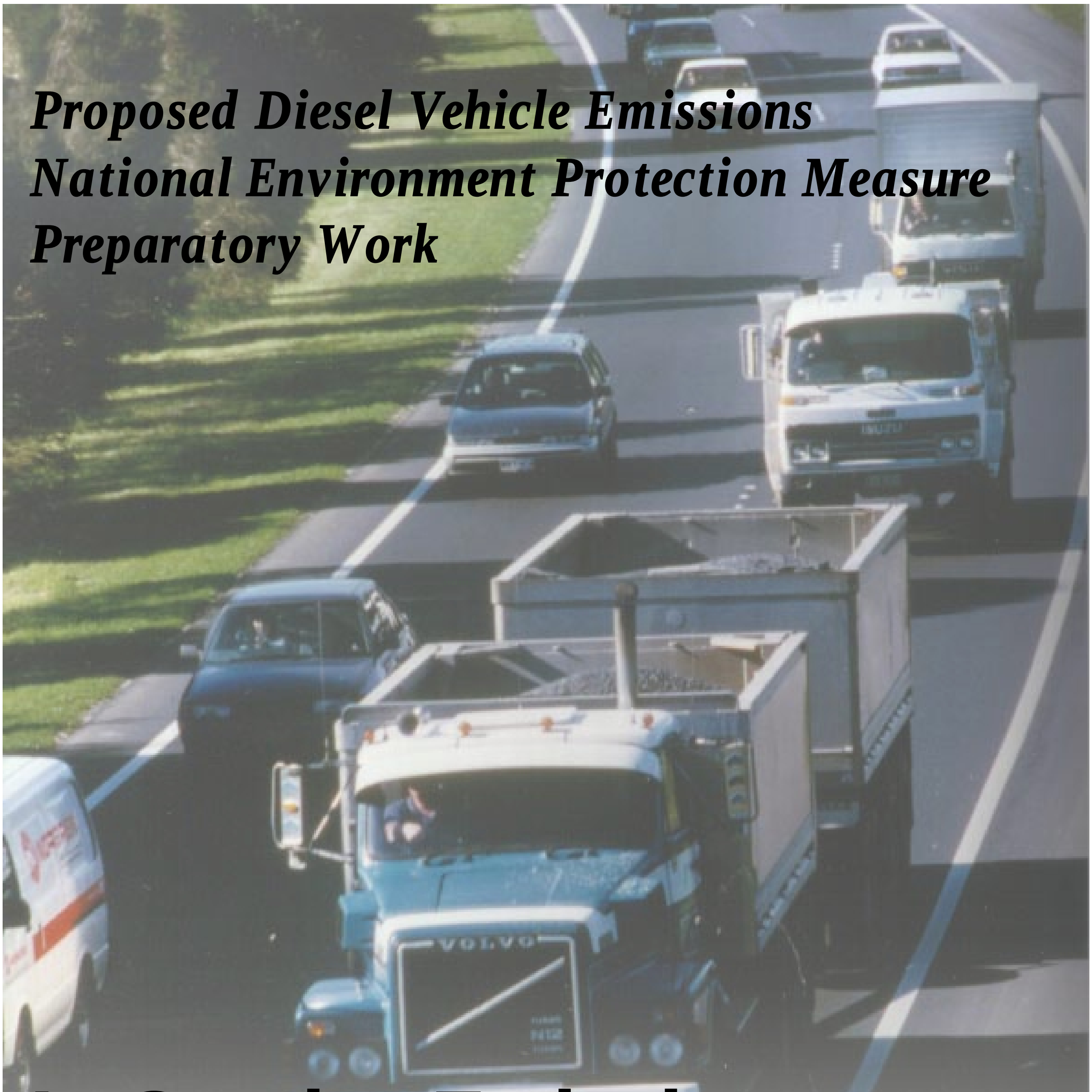
Proposed Diesel Vehicle Emissions National Environment Protection Measure Preparatory Work



In-Service Emissions Performance - Drive Cycles

Volume 2

March 1999

NEPC

NATIONAL ENVIRONMENT PROTECTION COUNCIL

Prepared for the

National Environment Protection Council

by

Steve Brown, Chris Bryett and Mike Mowle

NSW ENVIRONMENT PROTECTION AUTHORITY
MOTOR VEHICLE TESTING UNIT

PO Box 29
LIDCOMBE NSW 2141
Telephone: (02) 9795 5000
Facsimile: (02) 9643 2466

ISBN 0 642 32323 2

A suite of projects have been developed during the preparatory work for a proposed Diesel Emissions National Environment Protection Measure. These projects are:

The Australian Diesel Fleet Existing Vehicle Characteristics and the Modelling of Transport Demand, Vehicle Populations and Emissions

In-Service Emissions Performance - Phase 1: Urban Drive Cycle Development

In-Service Emissions Performance - Phase 2: Vehicle Testing

In-Service Certification Correlation Studies

A Review of Dynamometer Correlations, In-Service Emissions and Engine Deterioration

In-Service Emissions Testing – Pilot Study, Fault Identification and Effect of Maintenance

Major funding for these projects has been provided by Environment Australia. The other contributing agencies are the Department of Transport and Regional Services, NSW Roads Traffic Authority and the National Road Transport Commission.

Electronic copies of these documents are available from:

National Environment Protection Council Service Corporation
Level 5, 81 Flinders Street
Adelaide SA 5000

Telephone: (08) 8419 1200
Facsimile: (08) 8224 0912

These documents are also available online: http://www.nepc.gov.au

CONTENTS OF VOLUME II

Proposed Diesel Vehicle Emissions
National Environment Protection Measure
Preparatory Work

In-Service Emissions Performance Phase 1: Urban Drive Cycle Development

Volume 2

Attachment 1

CSIRO Mathematical and Information Sciences (CMIS) – The Identification of Typical Drive Cycles for Diesel Vehicles

NEPC
NATIONAL ENVIRONMENT PROTECTION COUNCIL

March 1999

Identification of Typical Drive Cycles for Diesel Vehicles: Final Report

by

Steven Davies, Ross Sparks and John Donnelly

CSIRO Mathematical and Information Sciences
Macquarie University Campus
Locked Bag 17, North Ryde NSW 1670

Report No. CMIS 99/21
February, 1999

Prepared for
NSW Environment Protection Authority and
National Environment Protection Council Service Corporation

CSIRO

CONTENTS

A CD accompanies this report and contains electronic versions of all documentation, including Excel (*.xls) and text (*.txt) files for the UEDCs and CUEDCs.

1 EXECUTIVE SUMMARY

For data collected from 17 diesel vehicles by the EPA, typical Urban Emissions Drive Cycles (UEDCs) were obtained for each of six ADR categories – MC, NA, NB, ME, NC and NCH, where the standard ADR category NC has been subdivided into two weight categories 12 to 25t (referred to as NC in this report) and over 25t (NCH in this report) – over four road conditions: congested, residential/minor, arterial and freeway/highway.

The following summary statistics were calculated for each road condition in each ADR category:

- Distance travelled (km)
- Time spent (hr)
- Average speed (km/hr)
- Idle time (hr)
- Idle time (%)

Typical UEDCs for each ADR category were determined by pooling the data for the appropriate vehicles and combining representative *microtrips* to fit within time intervals specified by the EPA for each road condition.

Cluster analyses were performed to investigate if groupings of vehicles indicated by the data were different to those nominated as the ADR categories. Different groups of vehicles were obtained for each of the four different road conditions.

Cluster analyses were also done on the six ADR categories to identify possible groupings of the categories. Different groups of categories were obtained for each road condition. Given these results, the small numbers of vehicles, the complex interaction and confounding effects of topography, road conditions, traffic conditions, drivers, driver skill and vehicle types, it is not recommended that a single Composite Urban Emissions Drive Cycle be used to represent all vehicles.

Composite Urban Emissions Drive Cycles (CUEDCs) were then obtained for each ADR category by combining representative microtrips to fit within time intervals specified by the EPA for each road condition. Piecewise linear segments were added to reflect the change in mode from idle, to acceleration, cruising or deceleration.

The following summary statistics were calculated for the UEDCs and the CUEDCs:

- Distance travelled (km)
- Time spent (sec)
- Average speed (km/hr)
- Idle time (sec)
- Idle time (%)
- Maximum speed (km/hr)
- Number of periods of acceleration

2 INTRODUCTION

The NSW Environment Protection Authority (EPA) contracted CSIRO Mathematical and Information Sciences (CMIS) to assist with the statistical analysis of data collected for *Project 2 Phase 1: Drive Cycles,* which is part of a major project coordinated by the National Environment Protection Council (NEPC) Service Corporation aimed at developing a Diesel National Environment Protection Measure or DNEPM.

In this project CMIS has collaborated closely with the EPA to categorise the driving cycles for 17 diesel vehicles representing six vehicle types and weight categories, that is, five Australian Design Rule (ADR) categories MC, NA, NB, ME and NC. The EPA requested that the ADR category NC be split into the two categories NC and NCH to represent the weight classes 12 to 25t and > 25t respectively. So in this report NC refers to vehicles in the weight class 12 to 25t and NCH refers to vehicles in the heaviest weight class (>25t). These categories are referred to as six ADR categories in this report, although the fifth and sixth categories (NC and NCH) are actually one ADR category (NC).

3 OBJECTIVES

CMIS was given a brief to characterise, for vehicles in each ADR category, the driving patterns occurring in the urban driving profiles of speed versus time across different road conditions. That is, an Urban Emissions Drive Cycle (UEDC) typifying diesel vehicle driving conditions was to be developed for each ADR category and approximately within the following time intervals specified by the EPA:

Congested	Res/Minor	Arterial	Fwy/Hwy	Total (min)
6	18	18	18	60

The UEDCs were to be combined where possible over the ADR categories to determine one or more Composite Urban Emissions Drive Cycles (CUEDCs) approximating the following time intervals specified by the EPA:

Congested	Res/Minor	Arterial	Fwy/Hwy	Total (min)
5–6	7–8	6–7	8–9	26–30

4 THE DATA

Data for 17 vehicles from the Sydney fleet were collected by the EPA and supplied to CSIRO for analysis. Each vehicle (in sequence) had been fitted with sensors for road and engine speed and was equipped with a data logger. The vehicles were chosen to represent nominated ADR vehicle/weight categories and major use/route categories.

The number of vehicles was limited to that which could reasonably be accessed within the budget constraints. A study based on a larger number of vehicles and covering more Australian cities may yield more comprehensive results.

The vehicles were driven for several business cycles, using specified routes (eg buses), or routes representative of the appropriate mix of driving conditions. A very small number of road conditions were covered by some vehicles, e.g. buses in freeway/highway conditions.

There were a number of errors identified in the data for some vehicles (see section 4.1). Data were screened for errors and only valid data have been used in our analyses. For some of the vehicles, the RPM measurements were unreliable or not available.

Details on the vehicles used in the study, including ADR category, vehicle type, weight category, presence of RPM measurements, type of vehicle, time of data collection, total distance travelled and total duration of measurement are given in Appendix A.

4.1 Data cleaning

For Vehicle 1 (Pajero) it appeared that the EPA instruments would not record accurate measurements for speed greater than 107 km/h under freeway conditions so speed was extrapolated from the engine rpm where appropriate.

Spiky or noisy data were obtained for Vehicle 10 (Mercedes Bus) so another run was done. Again there were problems with spiky data and many missing values in the results (apparently due to instrument malfunction) for the second day (14^{th} October) so these were excluded from the analysis.

Short gaps in many profiles caused problems with our analyses. In general, we replaced missing data with a missing data indicator (NA). In gaps of one second the missing value was replaced with the average of neighbouring measurements.

Determining RPM from Data Logger Digital Signal

We reviewed and revised the linear fit between RPM (from dash) and data logger digital signal for Vehicle 4 (Ford Transit panel van), omitting the first measurement, to obtain the calibration relationship for predicting RPM values. The revised relationship was:

$$RPM = 399.43 + 2.9834 \times Signal$$

4.2 Test data

To obtain a test dataset for validation of the classification procedures an extra vehicle (Vehicle 8, a rigid truck in the 3.5 to 12 t weight category, ADR NB) was driven by the same EPA staff member on roads typical of each of the four road conditions (Congested, Residential/ Minor, Arterial or Freeway/Highway).

5 DEVELOPMENT OF URBAN EMISSIONS DRIVE CYCLES

Representative cycles were required for each of four road conditions: Congested, Residential/Minor, Arterial or Freeway/Highway. The types of road conditions were not recorded throughout each profile for each vehicle, although supporting log sheets did contain incidental information for some of the vehicles, especially some of the buses where the particular route was known.

So it was necessary to develop a procedure which would predict the various road conditions experienced during each run. The approaches described below were used to develop a classification procedure for each vehicle. We have based all methodologies on the notion of a *microtrip* and the data from each vehicle were segmented into microtrips. A *microtrip* is similar to a "sequence" as defined and used by Andre *et al*. (1995).

5.1 Definition of a microtrip

A *microtrip* is defined as the excursion between two successive time points at which the vehicle was stationary. By convention, a period of rest is at the beginning of a microtrip rather than at the end. Thus a typical microtrip is a period of rest followed by periods of acceleration, cruising, deceleration until the vehicle is at rest again, whereupon the next microtrip starts.

There were two aims of the statistical analysis: firstly, to identify the representative driving patterns of different vehicles in various urban driving conditions, and secondly, to compare the driving patterns of different vehicles. Achieving both aims will assist the EPA in defining a small number of vehicle tests which, subject to engineering and time constraints, will best typify the driving patterns within Australian cities.

5.2 Classification methods

Classification Method 1

Information published by Carnovale *et al* (1997) was used to define important predictors of road condition and the values of their respective cut-points for distinguishing transition between different road conditions.

Carnovale *et al* (1997) provide the following descriptions of traffic conditions on different road categories –

Freeway/Highway:	Major roads with relatively high average speeds (say in excess of 40km/hr), and low congestion levels (say less than 5% idle time).
Arterial:	Major roads with moderate average speeds (say 20 – 40 km/hr), and moderate congestion levels (say 20% idle time).
Residential/Minor:	Secondary roads with moderate average speeds (say 20 – 40 km/hr) and negligible congestion.

From this we infer –

Congested: Any roads where average speeds are limited by traffic congestion to, say, less than 20 km/hr.

These conditions are tabulated as follows –

	Congested	**Res/Minor**	**Arterial**	**Fwy/Hwy**
Average speed	<20	20 – 40	20 – 40	>40
Idle time	0 – 100%	negligible (<5%)	approx 20%	<5%

These descriptors were assumed to be the important predictors of road condition and their values above were used as the initial cut-points for distinguishing transition between different road conditions.

Using these empirical rules we attempted to classify every microtrip for one of the vehicles (Vehicle 12). The table above failed to classify a number of microtrips, as they fell outside the above ranges of idle time and speed. Also, there was a very large number of congested microtrips.

From examination of the microtrips we found –
- categorisation was difficult, since very many microtrips had characteristics of others
- even at high (freeway/highway) average speeds, there were congested elements, and idle times of 15%+
- in many sequences having characteristics of minor roads, there were significant idle periods
- the difference between minor and arterial sequences was difficult to see
- in most obviously congested sequences, % idle time was quite short

Some adjustment to the cut-points was required to categorise all the data and to assign reasonable splits between categories. This included –
- adoption of a definition of congestion, which was arbitrarily assigned as average speed below 15 km/hr, or microtrips of less than 20 seconds duration.
- assignment of 15 km/hr as the minimum average speed for 'freeflow' arterial and minor road sequences.
- removal of the limitation on idle time for freeway/highway trips.

The revised empirical rules are given in the following table:

	Congested	**Res/Minor**	**Arterial**	**Fwy/Hwy**
Average speed	< 15 km/hr	15-40 km/hr	15-40 km/hr	> 40 km/hr
Idle time	0-100%	< 10%	> 10%	0-100%

Although the rules are complete (all microtrips were classified), there were some microtrips with excessively long idle periods for which none of the categories seemed appropriate. A new category, *Wait*, was defined to accommodate these microtrips. This category was intended for long periods (> 150 sec) where the vehicle was at rest. Reasons for these periods may include warming the engine or loading and unloading with the engine on.

Aside: all microtrips falling into the *Wait* category were excluded from the determination of UEDCs and CUEDCs, but summary statistics for these categories were calculated and are reported in section 6.

In order to validate the expert rules, a test vehicle (Vehicle 8) was dispatched specifically to be driven in each of the four driving conditions (the other vehicles undertook their normal daily trips). Some data for this vehicle were culled by the EPA to increase the accuracy of the nominal driving conditions.

The empirical expert rules were applied to the microtrips for Vehicle 8 and the results are given in the following classification table:

	Congested	Res/Minor	Arterial	Fwy/Hwy	TOTAL
Congested	56	9	15	0	80
Res/Minor	1	8	2	3	14
Arterial	26	9	25	9	69
Fwy/Hwy	11	1	7	15	34

The row labels indicate "nominal" road conditions and the column labels indicate the classified or predicted road conditions. So for example, out of the total of 80 microtrips for data collected from road conditions nominated as Congested, 56 were classified as Congested, 9 as Res/Minor, 15 as Arterial and 0 as Fwy/Hwy, giving a success rate of 56/80 or approximately 70%.

The overall "nominal" classification rate was (56+8+25+15)/197 = 53% (or 47% misclassified). The term "nominal" is used here since, although the general driving conditions of Vehicle 8 were recorded, there may have been short periods of one road condition (congestion say) during driving in other road conditions which were not taken into account. Therefore 53% should be viewed as a minimum overall classification rate.

Misclassifications occur where microtrips have some features pertinent to one category and some features pertinent to another. The method of selecting representative microtrips (as described in section 7) endeavours to select microtrips with features distinct to one road condition, and are therefore less likely to be misclassified.

This is a fairly low overall classification rate for many applications. However, for the present application, that of finding the most representative microtrips for each driving condition, it may be adequate. How robust the choice of representative microtrips is to misclassifications is an important consideration, but the resources of this project are too limited to investigate this question.

Classification Method 2

Using classification tree methodology (see Breiman *et al* 1984 or Clark and Pregibon, 1992) we generated a classification tree which would best predict road conditions given the set of predicting variables. A number of statistics of the Vehicle 8 microtrips were supplied as predicting variables to determine a tree-based classifier using the statistical package S-PLUS (1996). These 13 statistics (or predicting variables) included duration of microtrips, average speed and number of periods while in each of the four modes idling, cruising, accelerating or decelerating (3 x 4 = 12 variables), as well as the overall average speed. Of these variables, the tree-based classifier chose idle time and average speed as being the two most important variables, these being the same variables used for Method 1. For a tree with four nodes specified the tree-based classifier produced the following classification rules:

	Congested	Res/Minor	Arterial	Fwy/Hwy
Average speed	< 31 km/hr	31-46 km/hr	31-46 km/hr	> 46 km/hr
Idle time	0-100%	< 20%	> 20%	0-100%

The resulting classifications of each microtrip using the tree-based classifier are given in the table below, where the row and column labels are as defined previously:

	Congested	Res/Minor	Arterial	Fwy/Hwy	TOTAL
Congested	76	3	1	0	80
Res/Minor	3	11	0	0	14
Arterial	44	12	9	4	69
Hwy/Fwy	16	2	3	13	34

That is, the classification rate for congested roads was 76/80 = 95% while the overall classification rate was (76+11+9+13)/197 = 55%, an expected improvement on Method 1 since the classifier was constructed from the same data used for validation.

Classification Method 3

Finally, to classify all microtrips it was agreed that the expert classification rules from Method 1 should be used, with the following modifications: (i) the average speed threshold between Freeway/Highway driving and the other categories to be 45 km/hr instead of 40 km/hr, and
(ii) 15% was used as the cut-point for proportion of idle time to distinguish Residential/Minor and Arterial road conditions. Note that the supporting rule that all microtrips less than 20 seconds duration be classified as Congested road conditions was redundant because the average speed was less than 15 km/hr for all such microtrips.

Method 3 was preferred for the following reasons: (i) more realistic cut-points, based on the observed data, (ii) improved classification rate over Method 1 for Residential/Minor road conditions, (iii) improved classification rate over Method 2 for Arterial road conditions, and
(iv) less chance of over representation of Congested road conditions.

The classification rules used throughout the remainder of the analysis were then:

	Congested	Res/Minor	Arterial	Fwy/Hwy
Average speed	< 15 km/hr	15-45 km/hr	15-45 km/hr	> 45 km/hr

Idle time	0-100%	< 15%	> 15%	0-100%

This approach is illustrated in the figures contained in Appendix B where the microtrips for Vehicles 8 and 12 are superimposed over the cut points for the different road conditions represented by vertical and horizontal lines. For Vehicle 8, each microtrip is represented by a coloured letter, each letter identifying each nominal (observed) road condition. For example, microtrips for congested road conditions represented by a blue "C" fall predominantly within the rectangular area of 0 to 15 km/hr. The horizontal line at 15% idle represents a supporting cut-point to separate Arterial (above the line) microtrips from Res/Minor (below the line) microtrips.

The resulting classifications of each microtrip using the revised empirical classifier are given in the table below, where the row and column labels are as defined previously:

	Congested	**Res/Minor**	**Arterial**	**Fwy/Hwy**	**TOTAL**
Congested	56	12	12	0	80
Res/Minor	1	12	1	0	14
Arterial	26	14	24	5	69
Hwy/Fwy	11	3	7	13	34

The classification rate for congested roads was 56/80 = 70%, while the overall classification rate was (56+12+24+13)/197 = 53%. The classification rate for residential/minor was 12/14 = 86%.

Classification of an individual microtrip did not take into account neighbouring microtrips. Thus it was often the case that a sequence of microtrips had frequently changing classified road conditions, whereas in reality a vehicle would stay in particular road conditions for longer periods.

6 SUMMARY STATISTICS

Initially, exploratory data analysis tools were used to gain a better understanding of typical urban driving patterns for the vehicles. Graphs of speed versus time were produced and important summary statistics were calculated. Summary statistics for each road condition within each ADR category are presented in the following tables. Statistics for the Wait category defined previously are also given.

6.1 Summary statistics of microtrips

The following summary statistics were calculated on microtrips:

- Number of microtrips
- Median microtrip distance (m)
- Median microtrip duration (sec)
- Median microtrip average speed (km/hr)
- Median microtrip idle time (sec)
- Median microtrip idle time (%)
- Median microtrip maximum speed (km/hr)
- Median number of periods of acceleration

A period of acceleration was defined as a period of increasing speed during which the actual acceleration exceeded 2 km/hr/sec some time in the period, not necessarily over the whole period.
For a number of microtrips, especially those classified as Congested, there were no periods of acceleration using this definition. This definition and similar definitions for periods of deceleration and cruising are used for the fitting of piecewise linear segments to the CUEDCs in section 13.

Summary statistics of microtrips for ADR category MC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	537	137	163	103	3
Median microtrip distance (m)	2.4	702.9	740.6	2539.3	0
Median microtrip duration (sec)	9	73	96	182	206
Median microtrip av. speed (km/hr)	1.2	33.7	28.1	50.0	0
Median microtrip idle time (sec)	4	4	31	10	206
Median microtrip idle time (%)	50.0	6.1	30.0	5.0	100
Median microtrip maximum speed (km/hr)	3.6	60.0	64.0	78.0	0
Median nr of periods of acceleration	0	2	2	3	0

Summary statistics of microtrips for ADR category NA

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	913	182	164	37	12
Median microtrip distance (m)	2.6	463.8	504.2	2174.4	0
Median microtrip duration (sec)	11	64.5	75	146	214.5
Median microtrip av. speed (km/hr)	1.1	28.6	23.1	48.6	0
Median microtrip idle time (sec)	5	4	22	7	214.5
Median microtrip idle time (%)	58.3	6.6	27.6	5.5	100
Median microtrip maximum speed (km/hr)	4.6	51.8	56.7	76.3	0
Median nr of periods of acceleration	0	2	2	2	0

Summary statistics of microtrips for ADR category NB

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	1130	226	176	77	2
Median microtrip distance (m)	3.3	687.8	681.0	2765.6	0.1
Median microtrip duration (sec)	11	81.5	96.5	202	1694.5
Median microtrip av. speed (km/hr)	1.3	29.7	26.1	51.1	0
Median microtrip idle time (sec)	4	4	28.5	8	1694.5
Median microtrip idle time (%)	52.9	6.0	27.7	3.7	100
Median microtrip maximum speed (km/hr)	4.6	52.2	60.5	75.1	0.3
Median nr of periods of acceleration	0	2	1	3	0

Summary statistics of microtrips for ADR category ME

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	1470	538	631	48	57
Median microtrip distance (m)	6.0	578.1	365.5	2013.8	0
Median microtrip duration (sec)	19	69.5	59	138	246
Median microtrip av. speed (km/hr)	1.4	29.8	22.0	48.9	0
Median microtrip idle time (sec)	10	4	16	7	246
Median microtrip idle time (%)	56.2	7.1	25.0	5.0	100
Median microtrip maximum speed (km/hr)	7.5	50.3	47.8	76.8	0
Median nr of periods of acceleration	0	2	1	2	0

Summary statistics of microtrips for ADR category NC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	2219	350	202	122	40
Median microtrip distance (m)	0.9	867.0	786.9	3585.6	0
Median microtrip duration (sec)	7	103.5	111.5	250	279
Median microtrip av. speed (km/hr)	0.8	30.7	25.2	51.1	0
Median microtrip idle time (sec)	3	5	30	9	279
Median microtrip idle time (%)	50.0	4.9	25.6	3.8	100
Median microtrip maximum speed (km/hr)	2.2	51.9	55.3	76.4	0
Median nr of periods of acceleration	0	2	1	2	0

Summary statistics of microtrips for ADR category NCH

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait
Nr of microtrips	1662	274	147	67	51
Median microtrip distance (m)	2.6	878.3	855.1	4380.0	0
Median microtrip duration (sec)	11	109	125	292	268
Median microtrip av. speed (km/hr)	1.1	30.4	24.1	52.4	0
Median microtrip idle time (sec)	3	4.5	33	9	268
Median microtrip idle time (%)	50.0	4.5	25.7	4.2	100
Median microtrip maximum speed (km/hr)	3.0	54.2	56.0	82.1	0
Median nr of periods of acceleration	0	2	2	2	0

6.2 Summary statistics for each road condition in each ADR category

Summary statistics were calculated for each road condition within each ADR category as follows:

- Distance travelled (km)
- Time spent (hr)
- Average speed (km/hr)
- Idle time (hr)
- Idle time (%)
- Number of periods of acceleration

Summary statistics for ADR category MC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	18.0	133.4	152.1	505.2	0	808.6
Time Spent (hr)	3.2	3.8	5.2	8.1	0.2	20.4
Average Speed (km/hr)	5.6	35.5	29.5	62.5	0	39.7
Idle time (hr)	1.9	0.2	1.7	0.4	0.2	4.4
Idle time (%)	59.9	5.8	32.4	4.9	100	21.4
Nr of periods of acceleration	181	402	356	445	0	1384

Summary statistics for ADR category NA

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	37.8	135.8	119.8	89.1	0	382.5
Time Spent (hr)	6.5	4.2	4.4	1.7	0.8	17.5
Average Speed (km/hr)	5.8	32.4	27.4	52.7	0	21.9
Idle time (hr)	4.1	0.3	1.3	0.1	0.8	6.6
Idle time (%)	63.4	6.5	30.7	6.2	100	37.7
Nr of periods of acceleration	393	495	369	102	0	1359

Summary statistics for ADR category NB

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	42.4	244.4	156.4	508.6	0	951.8
Time Spent (hr)	7.6	7.1	5.5	7.7	0.9	28.8
Average Speed (km/hr)	5.6	34.4	28.6	66.4	0	33.0
Idle time (hr)	4.3	0.4	1.6	0.3	0.9	7.5
Idle time (%)	55.8	5.8	29.4	3.8	100	26.0
Nr of periods of acceleration	436	640	316	260	0	1652

Summary statistics for ADR category ME

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	78.5	427.1	292.4	133.1	0	931.1
Time Spent (hr)	12.9	13.5	12.4	2.6	4.4	45.7
Average Speed (km/hr)	6.1	31.6	23.6	51.1	0	20.4
Idle time (hr)	8.1	0.9	3.5	0.1	4.4	16.9
Idle time (%)	62.4	6.4	28.5	4.7	100	37.0
Nr of periods of acceleration	765	1315	981	113	0	3174

Summary statistics for ADR category NC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	43.7	433.8	216.6	709	0	1403.1
Time Spent (hr)	11.4	13.1	7.7	12.7	4.6	49.5
Average Speed (km/hr)	3.8	33.0	28.2	55.8	0	28.4
Idle time (hr)	6.7	0.7	2.1	0.5	4.6	14.6
Idle time (%)	58.8	5.3	27.9	3.8	100	29.6
Nr of periods of acceleration	356	841	369	365	0	1931

Summary statistics for ADR category NCH

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Wait	Overall
Distance Travelled (km)	62.1	332.4	143.2	561.6	0	1099.3
Time Spent (hr)	14.0	10.5	5.5	8.7	4.8	43.5
Average Speed (km/hr)	4.4	31.7	26.2	64.4	0	25.3
Idle time (hr)	7.8	0.6	1.5	0.4	4.8	15.1
Idle time (%)	55.7	5.3	27.9	4.4	100	34.7
Nr of periods of acceleration	408	734	334	263	0	1739

7 REPRESENTATIVE MICROTRIPS

In order to find representative microtrips which could be used to form the UEDCs, it was first necessary to define a representative microtrip. In the absence of a standard definition or procedure, the following definition was used:

"The representative microtrip of a set of microtrips is that microtrip which spends time at speeds and accelerations in similar proportions to the entire set of microtrips."

In mathematical terms, this is the microtrip which minimises the distance between its empirical distribution function (a function of speed and acceleration) and the pooled empirical distribution function (edf) from all microtrips. Many statistics can be determined in terms of the edf. In theory, choosing microtrips with a similar edf to that of the pooled microtrips ensures that the chosen microtrips have similar statistics to the pooled microtrips.

The particular distance measure used here was the integrated absolute difference between the two cumulative distribution functions. This criterion is based on a Kolmogorov-Smirnov distance measure or score given by

$$\iint |\hat{F}_{m:R}(s,a) - \hat{F}_R(s,a)| \, ds \; da$$

where m identifies a particular microtrip, R is the road category, a is the given acceleration, s is the speed and $\hat{F}$ is an estimator for the distribution function.

The particular estimator used was based on a binned estimate of the density function, $\hat{f}$, thus

$$\hat{F}(s,a) = \int \hat{f}(x,y)dxdy$$

where the limits for the double integral are from $\{-\infty, -\infty\}$ to $\{s, a\}$.

Binning reduces the amount of computational effort while maintaining a level of robustness of results, since it is used to compare distribution functions. The bins were the product of separate partitions of speed and acceleration, similar to the acceleration/deceleration tables in section 10. Each partition split the range up into 10 intervals, each interval containing a similar number of observations. Thus the density estimator used was:

$$\hat{f}(x,y) = \text{number of seconds in the bin containing (x, y) /}$$
$$\text{[total number of seconds x area of bin containing (x, y)]}$$

The individual $\hat{f}_R$ for the different road categories are graphed in Appendices C and D. These graphs show the different proportions of time spent at various speeds and accelerations and provide a comparison between the four road types.

This method, albeit with a minor modification, was used to obtain a number of representative microtrips for each of the four road conditions for each of the vehicles within each ADR category.

This modification was a weighting used to downweight periods spent at rest, as these tended to dominate in the choice of microtrip; microtrips were being selected almost exclusively on length of idle period rather than on the remainder of the driving sequence.

The representative microtrips obtained for each vehicle are displayed in Appendix C. Representative microtrips for each ADR category were then obtained by pooling the data for the vehicles in each category. Representative microtrips for each ADR category are presented in Appendix D.

Each set of graphs in Appendix C or D consists of the following:

i. a plot of the density estimates
ii. a plot of the Kolmogorov-Smirnov scores for the microtrips
iii. the first fifteen representative microtrips for Congested roads and the first five representative microtrips for Residential/Minor, Arterial and Freeway/Highway road conditions.

8 URBAN EMISSIONS DRIVE CYCLES FOR ADR CATEGORIES

To form the UEDC for each ADR category we selected the appropriate number of representative microtrips to be joined, starting with the most representative microtrip, which fulfilled the duration requirements specified by the EPA. For congested road conditions, representative microtrips were of relatively short duration, so 15 representative microtrips are displayed. For all other road conditions the first five representative microtrips are shown in the Appendices, although more than five are sometimes required to build a UEDC .

The duration (in minutes) of each component of the cycle for each road condition followed by the microtrip identifiers in parentheses and the total duration of the UEDC (in minutes), are shown in the following table.

ADR category	Congested	Res/Minor	Arterial	Fwy/Hwy	UEDC Duration (min)
MC	5.55 (1-6)	17.57 (1-6)	16.73 (1-4)	18.37 (1,2,3,5)	58.22
NA	5.57 (1-5)	17.98 (1-8)	17.58 (1-6)	19.02 (1-5)	60.15
NB	6.28 (1-7)	17.28 (1-4)	18.18 (1-7)	17.47 (1-3)	59.21
ME	6.35 (1-5)	17.85 (1-6)	17.97 (1-9)	17.48 (1-3)	59.65
NC	5.47 (1-4)	17.97 (1,3,4)	18.05 (1-6)	16.05 (1,2)	57.54
NCH	6.07 (1-3)	17.17 (1-4)	18.28 (1-6)	20.23 (3,4,5)	61.75

UEDCs do not necessarily contain the most representative microtrips for two reasons.

(i) In some cases it was necessary to include less representative microtrips in order to fit the desired duration.

(ii) The microtrips of UEDCs comprise a small number of actual microtrips from the vehicles in the study. The data from some of these microtrips was poor and so less representative microtrips, but with better quality data, were used.

For example, to form the UEDC for ADR category NCH we joined the first three most representative microtrips for Congested, the first four representative microtrips for Res/Minor, the first six representative microtrips for Arterial and for Freeway/Highway, we used the 3rd, 4th and 5th most representative microtrips because the first was too long (30.72 minutes) and the speed measurements for the second were spiky (see Appendix D).

The resulting UEDCs for each ADR category are presented in Appendix E. UEDCs and gear information are supplied on the accompanying CD as a text file uedc.txt and as a Microsoft Excel file uedc.xls. In each file there are five columns labelled as ADR,Road,Seconds,Speed,Gear . Column 1 contains the ADR category, column 2 contains the road condition, column 3 the time in seconds from zero, column 4 the speed in km/hr and column 5 contains the predicted gear. Column 5 will contain integers 0,1,2,etc. or NA. Zero is used to identify points which could not be accurately classified (see section 14). NA is used to distinguish those microtrips for which no gear classification was done, eg because the vehicle had an automatic gearbox.

9 SUMMARY STATISTICS FOR UEDCs

Summary statistics calculated for the UEDC for each ADR category are presented in the following tables:

Summary statistics for UEDC for ADR category MC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	10.5	9.0	18.0	38.2
Time Spent (sec)	333	1054	1004	1102	3493
Average Speed (km/hr)	7.1	35.9	32.3	58.9	39.4
Idle time (sec)	148	41	256	23	468
Idle time (%)	44.4	3.9	25.5	2.1	13.4
Maximum speed (km/hr)	29.0	67.3	81.8	91.4	91.4
Nr of periods of acceleration	7	31	22	18	78

Summary statistics for UEDC for ADR category NA

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.2	9.9	9.5	16.5	37.1
Time Spent (sec)	334	1079	1055	1141	3609
Average Speed (km/hr)	13.0	33.0	32.5	52.2	37.1
Idle time (sec)	87	97	208	53	445
Idle time (%)	26.0	9.0	19.7	4.6	12.3
Maximum speed (km/hr)	45.3	71.1	82.1	86.9	86.9
Nr of periods of acceleration	12	33	28	19	92

Summary statistics for UEDC for ADR category NB

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.1	10.6	8.9	20.7	41.3
Time Spent (sec)	377	1037	1091	1048	3553
Average Speed (km/hr)	10.6	36.8	29.5	71.0	41.9
Idle time (sec)	50	26	300	27	403
Idle time (%)	13.3	2.5	27.5	2.6	11.3
Maximum speed (km/hr)	30.9	68.7	73.7	97.3	97.3
Nr of periods of acceleration	11	31	17	9	68

Summary statistics for UEDC for ADR category ME

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.1	9.5	7.3	14.9	32.9
Time Spent (sec)	381	1071	1078	1049	3579
Average Speed (km/hr)	10.8	32.0	24.5	51.3	33.1
Idle time (sec)	128	54	280	54	516
Idle time (%)	33.6	5.0	26.0	5.1	14.4
Maximum speed (km/hr)	36.4	62.6	63.3	85.0	85.0
Nr of periods of acceleration	11	32	29	12	84

Summary statistics for UEDC for ADR category NC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	10.2	9.1	15.0	35.0
Time Spent (sec)	328	1078	1083	963	3452
Average Speed (km/hr)	8.0	34.1	30.2	56.0	36.5
Idle time (sec)	73	24	257	16	370
Idle time (%)	22.3	2.2	23.7	1.7	10.7
Maximum speed (km/hr)	28.2	59.9	69.3	80.8	80.8
Nr of periods of acceleration	4	16	16	8	44

Summary statistics for UEDC for ADR category NCH

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	8.6	8.2	21.7	39.2
Time Spent (sec)	364	1030	1097	1214	3705
Average Speed (km/hr)	6.7	30.2	26.9	64.4	38.1
Idle time (sec)	94	75	270	47	486
Idle time (%)	25.8	7.3	24.6	3.9	13.1
Maximum speed (km/hr)	33.0	62.9	68.5	99.2	99.2
Nr of periods of acceleration	6	21	15	7	49

10 CLUSTER ANALYSIS OF VEHICLES FOR EACH ROAD CONDITION

Hierarchical Cluster Analysis was used to explore which vehicles fall into groups as suggested by the data. The distance measure used in the cluster analyses was the same as that used for identifying representative microtrips. That is, for each road condition, cluster analysis was performed on a 17 x 17 similarity matrix consisting of all pairwise distance measures between the 17 vehicles. Note that this distance measure takes into account both acceleration and speed.

As for the distance measure used in determining representative microtrips, all idle periods were ignored.

Results of the cluster analyses shown in Appendix F indicate no consistent clustering or grouping of vehicles according to the ADR categories across all road conditions.

11 COMBINING ADR CATEGORIES – CLUSTER ANALYSIS

Hierarchical Cluster Analysis was used to explore which ADR categories, if any, could be combined in order to reduce the number of tests. The distance measure used in the cluster analyses was the same as that used for identifying representative microtrips. For each road condition, cluster analysis was performed on a 6 x 6 similarity matrix consisting of all pairwise distance measures between the six ADR categories. This distance measure takes into account both acceleration and speed.

As for the distance measure used in determining representative microtrips, all idle periods were ignored.

Results of the cluster analyses shown in Appendix G indicate no consistent clustering or grouping of subsets of ADR categories across all road conditions, although NC and NCH were grouped together for all road conditions except Freeway/Highway.

Hence we can not recommend that a composite UEDC be obtained for groups of ADR categories

12 COMPOSITE URBAN EMISSIONS DRIVE CYCLES

Since the results of the cluster analyses did not support the notion of combining the ADR categories into smaller groups, Composite Urban Emissions Drive Cycles were obtained for each ADR category according to time intervals specified by the EPA.

The duration (in minutes) of each component of the composite cycle for each road condition followed by the microtrip identifiers in parentheses and the total duration of the CUEDC (in minutes), are shown in the following table. As for the UEDCs, the CUEDCs do not necessarily contain the most representative microtrips to satisfy the required durations.

ADR category	Congested	Res/Minor	Arterial	Fwy/Hwy	CUEDC Duration (min)
MC	5.55 (1-6)	6.87 (1,2)	7.80 (1)	8.48 (1,2)	28.70
NA	5.57 (1-5)	8.40 (1-3)	7.45 (1,2)	8.48 (1)	29.90
NB	5.32 (1-6)	6.75 (1,2)	6.50 (1,2)	9.85 (2)	28.42
ME	5.37 (1-4)	8.43 (1,2)	7.25 (1-4)	6.90 (1)	27.95
NC	5.47 (1-4)	8.48 (2,3)	7.18 (1-3)	8.80 (2,3)	29.93
NCH	6.07 (1-3)	7.95 (2,3)	7.40 (1,2)	6.50 (3)	27.92

For example, to form the CUEDC for ADR category ME we joined the first four most representative microtrips for Congested, the first two representative microtrips for Res/Minor, the first four representative microtrips for Arterial and the first representative microtrip for Freeway/Highway (see Appendix H).

Several approaches to fitting piecewise linear segments to each CUEDC were investigated, including fitting piecewise linear B-splines (see de Boor, 1978 or Cheney and Kincaid, 1985) and using multivariate adaptive regression splines or MARS (see Friedman, 1991 or Friedman and Silverman, 1989).

Because of problems encountered with determination of the location of the knots (change points) in the spline approach, piecewise linear segments have been superimposed over the CUEDCs using an automatic method which restricts change points to points identified as a change in mode from idle to acceleration to cruising to deceleration. The resulting superimposed piecewise linear segments are shown in Appendix H for each ADR category.

CUEDCs and gear information are supplied on the accompanying CD as a text file cuedc.txt and as a Microsoft Excel file cuedc.xls. In each file there are six columns labelled as ADR,Road,Seconds,Speed,Linear,Gear. Column 1 contains the ADR category, column 2 contains the road condition, column 3 the time in seconds from zero, column 4 the speed in km/hr, and column 5 the speed for the piecewise linear segments and column 6 the predicted gear information. As for the UEDCs, column 6 will contain integers 0,1,2,etc. or NA. Zero is used to identify points which could not be accurately classified (see section 14). NA is used to distinguish those microtrips for which no gear classification was done, eg because the vehicle had an automatic gearbox.

13 SUMMARY STATISTICS FOR CUEDCs

Summary statistics calculated for the Composite Urban Emission Drive Cycle for each ADR category are given in the following tables:

Summary statistics for CUEDC for ADR category MC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	4.0	4.4	7.9	16.9
Time Spent (sec)	333	412	468	509	1722
Average Speed (km/hr)	7.1	34.5	33.9	55.7	35.3
Idle time (sec)	148	14	89	37	288
Idle time (%)	44.4	3.4	19.0	7.3	16.7
Maximum speed (km/hr)	29.0	66.7	72.8	85.3	85.3
Nr of periods of acceleration	7	14	11	8	40

Summary statistics for CUEDC for ADR category NA

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.2	4.6	4.0	7.7	17.5
Time Spent (sec)	334	504	447	509	1794
Average Speed (km/hr)	13.0	32.8	32.0	54.5	35.1
Idle time (sec)	87	51	79	3	220
Idle time (%)	26.0	10.1	17.7	0.6	12.3
Maximum speed (km/hr)	45.3	62.4	82.1	85.0	85
Nr of periods of acceleration	12	13	16	7	48

Summary statistics for CUEDC for ADR category NB

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.0	4.0	3.2	12.4	20.6
Time Spent (sec)	319	405	390	591	1705
Average Speed (km/hr)	10.9	35.8	29.9	75.3	43.5
Idle time (sec)	28	13	109	3	153
Idle time (%)	8.8	3.2	27.9	0.5	9.0
Maximum speed (km/hr)	30.9	68.7	73.7	91.6	91.6
Nr of periods of acceleration	9	15	8	3	35

Summary statistics for CUEDC for ADR category ME

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	1.0	4.7	2.9	5.8	14.4
Time Spent (sec)	322	506	435	414	1677
Average Speed (km/hr)	11.1	33.3	23.7	50.8	30.9
Idle time (sec)	102	6	122	22	252
Idle time (%)	31.7	1.2	28.0	5.3	15
Maximum speed (km/hr)	36.4	60.7	63.3	85.0	85
Nr of periods of acceleration	9	15	12	4	40

Summary statistics for CUEDC for ADR category NC

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	4.6	3.8	8.0	17
Time Spent (sec)	328	509	431	528	1796
Average Speed (km/hr)	8.0	32.4	31.5	54.2	34.1
Idle time (sec)	73	28	78	29	208
Idle time (%)	22.3	5.5	18.1	5.5	11.6
Maximum speed (km/hr)	28.2	57.2	63.3	82.9	82.9
Nr of periods of acceleration	4	9	9	4	26

Summary statistics for CUEDC for ADR category NCH

Statistic	Congested	Res/Minor	Arterial	Fwy/Hwy	Overall
Distance Travelled (km)	0.7	4.3	3.5	7.1	15.5
Time Spent (sec)	364	477	444	390	1675
Average Speed (km/hr)	6.7	32.6	28.1	65.2	33.4
Idle time (sec)	94	8	91	12	205
Idle time (%)	25.8	1.7	20.5	3.1	12.2
Maximum speed (km/hr)	33.0	60.3	59.0	96.1	96.1
Nr of periods of acceleration	6	9	7	2	24

14 DESCRIBING DRIVING CHARACTERISTICS WITH RPM AND SPEED MEASUREMENTS

In order to obtain gear information to accompany the CUEDCs, the following simple model was assumed for predicting speed from rpm when the particular gear is known:

$$rpm = a \times speed + error$$

where a is a regression parameter for a particular gear, each gear having a different value for a, and *error* is taken as normally distributed with mean zero and variance s^2_{gear}. The only problem in our situation is that the gear the vehicle is in for each measurement is unknown, thus it is treated as missing information. Our method for dealing with this missing information involves maximum likelihood estimation of the model parameters. This is described below and parameter estimates are summarised in subsequent tables.

14.1 Estimation of the models

The first step in our analysis used speed and rpm values to classify which gear vehicles were in at each second measured. Then acceleration characteristics are displayed for each gear of the vehicle.

The speed and rpm values were used to classify what gear the vehicles were in for each time point measured. This was carried out using the EM algorithm (Dempster, Laird and Rubin, 1977) where the gear is treated as missing. Although adjacent rpm and speed measures were correlated (autocorrelation) was evidence, this correlation was ignored and the data treated the observations as independent. This has the effect of reducing the efficiency of the estimates slightly but reduces the complexity of the analysis considerably. Starting values for these missing values were determined by eye. First we constructed a scatterplot of rpm versus speed (eg see Appendix I, Figure 1).

In the scatterplot of the data we drew by eye lines which approximately corresponded to the five gears. We then used these lines to classify each observation to a gear, and if the speed and rpm value was not close to a gear then we classified it as a transition point. Any point not classified to a gear was given zero weight in the next step of estimating model parameters. Multiple classifications were always assigned to the lowest gear. The maximum likelihood estimate of the model parameters given these classified gears was then calculated. These estimates were then used to re-classify the gears more accurately by using prediction intervals to classify the points. We repeated the procedure in an iterative fashion until the parameter estimates converged. The final estimates gave the estimated model, and a means of accurately classifying the measured values to a gear.

The EM algorithm could only fit the number of gears that we started with in our initial estimate. Therefore, not all gears were fitted. We were not too concerned about this because if gears are indistinguishable by eye, then effectively for our classification of phases these should be treated as identical gears with very similar diesel emission characteristics.

For the vehicles with manual gearboxes (omitting test Vehicle 8) and reliable RPM measurements (Vehicles 1, 2, 4, 7, 11, 13, 14, 16 and 17) the above maximum likelihood method was used to estimate a for each gear. Note that RPM measurements were collected for Vehicle 3 (Ford Transit) and Vehicle 9 (Scania bus) but these vehicles were omitted as they had automatic gearboxes.

During the testing program it emerged that several vehicles did not have a suitable engine speed signal which could be accurately logged or did not have an engine speed sensor fitted. Furthermore, certain heavier vehicles had a large number of gears, not all of which were used in succession during on-road acceleration or deceleration. Attempts to classify gears using the model described above did not always accurately predict the number of forward gears fitted to the vehicle. As this analysis was not considered to be fundamental to the construction of the drive cycles, results for one vehicle (Vehicle 1, Pajero) are presented here to illustrate the procedure. Results for all analyses are available from the authors.

14.2 Results - prediction of gears for Pajero

About 27% of the time is spent between gears or stationary in no particular gear. The gear that the vehicle spent most of its time in is fourth gear (nearly 28%).

Gear	**1st or reverse**	**2nd**	**3rd**	**4th**	**5th**
a	138.1750	77.6494	48.2012	34.7669	26.4115
Proportion of time spent in this gear	0.0459	0.1028	0.1607	0.2792	0.1335

14.3 Results - acceleration characteristics for Pajero

In this section we examine the relationship between RPM, Speed and Acceleration within each gear.

Gear 1: The relationships are presented in Appendix I, Figure 2 for gear 1. Notice that whenever the Pajero is accelerating then the relationship between these variables appears linear, but at near zero acceleration this linear relationship is not as good. A linear relationship between acceleration and speed or rpm is linear whenever the Pajero is not decelerating. As expected speed seems to reduce when decelerating but rpm tends towards a minimum.

Gear 2: The relationships are presented in Appendix I, Figure 3 for gear 2. The relationship between speed and rpm appears linear, but the linear relationship is not as good for acceleration with speed or rpm. An approximate but weak linear relationship between acceleration and speed/rpm is evident. Note that speed seems to be low when decelerating and rpm tends towards a minimum (as expected).

Gear 3: The relationships are presented in Appendix I, Figure 4 for gear 3. The relationship between speed and rpm appears linear, but the linear relationship is not as good for acceleration with the other variables. An approximate but weak linear relationship between acceleration and speed or rpm is evident.

Note that the slope of this linear relationship changes as the Pajero moves from acceleration to deceleration.

Gear 4: The relationships are presented in Appendix I, Figure 5 for gear 4. The relationship between speed and rpm appears linear, but the linear relationship is not as good for acceleration with the other variables. When decelerating, there appears to be an approximate linear relationship between it and speed, and it and rpm. However, acceleration is not related to either speed or rpm.

Gear 5: The relationships are presented in Appendix I, Figure 6 for gear 5. The relationship between speed and rpm appears linear, but the linear relationship is not as good for acceleration with the other variables. When decelerating, there appears to be an approximate linear relationship between it and speed, and it and rpm. However, acceleration is not related to either speed or rpm.

RPM less than 820: The relationships are presented in Appendix I, Figure 7 for low rpm values. Deceleration is related to speed in the natural way (higher deceleration at higher speeds). Acceleration here is not common, but we would assume that it would only occur when descending a steep hill.

14.4 Correspondence analyses

For the vehicles where no reliable RPM measurements were available (Vehicles 3, 5, 6, 9, 10, 12 and 15) correspondence analyses were done to search for meaningful clusters or groups. RPM values were measured for Vehicle 8 but no correspondence analysis was done on the validation test data.

A two-dimensional correspondence analysis tries to find the two dimensional space that explains the most information about the counts matrix. It is the two dimensional space that best fits the full space defined by the counts matrix in the weighted least squares sense. The dominant features in the counts matrix are generally displayed or explained by these two major axes.

The horizontal axis corresponds to the first major axis. This axis defines the dimensional direction (within the full dimensional space) that describes a maximum amount of the information in the counts matrix. The vertical axis is the axis orthogonal to the horizontal axis that describes the second most information in the counts matrix.

Again, as this analysis was not considered to be fundamental to the construction of the drive cycles, results for one vehicle (Vehicle 5, MAN Euro-II Bus) are presented here to illustrate the procedure. Results for all analyses are available from the authors.

This bus spent 40.5% of its time stationary, but spent 9.38, 11.24, 12.93, 16.06, 8.24, 1.44, 0.25, and 0.01 % of its time travelling between 0-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70 and 70-80 km/hr respectively. Also, it spent 24.50% of its time accelerating at a rate greater than 0.6 km/hr/sec, and this is 41.2% of its travel time.

When it travelled between 0 to 50 km/hr it spent 37.96% of its time accelerating at more than 0.6 km/hr/sec. See following table.

Vehicle 5: Man Euro-II Bus

	Acceleration(+)/Deceleration(-) (km/hr/sec)														
Speed km/hr	**-20 to -5**	**-5 to -2**	**-2 to -1**	**-1 to -0.6**	**-0.6 to -0.3**	**-0.3 to -0.1**	**-0.1 to 0**	**0 0**	**0 to 0.1**	**0.1 to 0.3**	**0.3 to 0.6**	**0.6 to 1**	**1 to 2**	**2 to 5**	**5 to 20**
0	8	489	560	289	280	385	9	16536	11	220	169	206	475	277	19
0-10	90	1012	572	264	224	185	59	142	81	166	197	242	589	796	1
10-20	233	1236	512	251	217	154	73	9	91	164	228	262	596	1435	77
20-30	151	1208	661	290	275	188	126	2	122	206	309	408	815	1555	56
30-40	63	675	808	606	657	540	289	5	273	552	708	688	1089	914	46
40-50	16	183	368	352	372	329	177	1	192	346	472	443	562	228	17
50-60	2	32	78	67	61	48	26	0	27	38	82	94	113	36	4
60-70	0	1	12	13	11	15	3	0	3	11	20	14	16	2	0
70->	0	0	0	2	0	0	0	0	1	0	1	0	2	0	0

The results of correspondence analysis for Vehicle 5 (Man Euro-II Bus) are shown in Appendix J, Figure 1 where the first two dimensions explain 97.5% of the variation.

15 CONCLUSIONS

Given profiles of speed versus time collected by the EPA for 17 diesel vehicles drawn from the Sydney diesel fleet, we have applied innovative algorithms to determine Composite Urban Emissions Drive Cycles for each of six ADR categories. We have based our approach on the definition of a microtrip, which has been previously defined as a "sequence" in the literature (see Andre *et al*.,1995).

16 REFERENCES

Andre, M., Hickman, A.J., Hassel, D. and Joumard, R. (1995) Driving cycles for emission measurements under European conditions. Institut National de Recherche sur les Transports et leur securite, Bron, France, Report 950926.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984) *Classification and Regression Trees*. Belmont, Ca.: Wadsworth.

Carnovale, F., Tilly, K., Stuart, A., Carvalho, C., Summers, M. and Eriksen, P. (1997) *Metropolitan Air Quality Study - Air Emissions Inventory*. Published by NSW Environment Protection Authority. ISBN 0 7130 3833 9.

Cheney, W. and Kincaid, D. (1985) *Numerical Mathematics and Computing*, Second edition. New York: Brooks/Cole Publishing Co.

Clark, L.A. and Pregibon, D. (1992) *Tree-Based Models*. In *Statistical Models in S*. Ed. By John M. Chambers and Trevor J. Hastie. Pacific Grove, Ca.: Wadsworth & Brooks/Cole. Chapter 9, pp 377-419.

de Boor, C. (1978) *A Practical Guide to Splines*. Berlin: Springer Verlag.

Dempster, A.P. Laird, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society,* Series B, **39**, 1-38.

Friedman, J.H. (1991) Multivariate adaptive regression splines (with discussion). *Annals of Statistics*, **19**, 1-141.

Friedman, J.H. and Silverman, B.W. (1989) Flexible parsimonious smoothing and additive modeling (with discussion). *Technometrics*, **31**, 3-39.

S-PLUS (1996) *Version 3.4 for Unix Supplement*. Data Analysis Products Division, MathSoft, Seattle.

17 CONTENTS OF APPENDICES

Appendix	Contents
A	Summary of vehicle information
B	Illustration of Classification Rule for Microtrips
C	Representative microtrips for each vehicle
D	Representative microtrips for each ADR category
E	Urban Emissions Drive Cycle for each ADR category
F	Cluster analysis of vehicles for each road condition
G	Cluster analysis of ADR categories for each road condition
H	Composite Urban Emissions Drive Cycle for each ADR category
I	Scatterplots of Speed vs RPM
J	Correspondence analysis