The use of remote sensing to enhance motor vehicle emission modelling in New Zealand

Prepared by NIWA for the Atmosphere and Climate Research Programme 6: Impacts of Air Pollution

May 2015

Authors/Contributors:
Robin Smit (University of Queensland)
Elizabeth Somervell

For any information regarding this report please contact:
Elizabeth Somervell
Air Quality Scientist
Urban Air Quality and Health
+64-9-375 2038
elizabeth.somervell@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
Private Bag 99940
Viaduct Harbour
Auckland 1010

Phone +64 9 375 2050

NIWA Client Report No:	AKL2015-012
Report date:	May 2015
NIWA Project:	CAAP1503

Contents

Tables

Figures

Reviewed by

Guy Coulson

……………………………………

Approved for release by

Ken Becker

……………………………

Executive summary

The use of remote sensing to measure vehicle emissions is not new, dating back to 1971. Remote sensing has been used extensively over the last two decades for various purposes around the world, including but not limited to identification of high-emitting vehicles, examination of on-road vehicle emissions distributions, trend analysis and model validation.

This study has scoped out and tested the best and most useful ways for validation of the New Zealand vehicle emission model, known as VEPM[1] using remote sensing data for CO, THC and NO_x. This was done through an examination of potential issues with the remote sensing data and ways to address them, as well as a discussion and demonstration of various (possible) ways to compare remote sensing data with VEPM emission factors.

Testing the overall accuracy of vehicle emission models is challenging, as the 'true' emission values in urban networks are unknown and cannot practically be determined by measurement. As a consequence, VEPM predictions and remote sensing data are *both* independent *estimates* of the true vehicle emissions in a road network.

If both VEPM and the remote sensing data show similar predictions of vehicle emissions, then this would increase the level of confidence that VEPM predictions are in fact accurate. If not, then VEPM model predictions and/or remote sensing estimates are not accurate. More validation work using other independent methods such as tunnel studies, near-road air quality sampling etc., and further analysis are required to determine where the (main) errors occur (specific traffic situations, specific vehicle classes, etc.).

Laboratory measurements ('bag' and 'modal', engine and chassis dynamometer) using drive cycles have remained the prominent empirical base for vehicle emission model development around the world, although this is changing now with increased use of PEMS (portable emissions measurement system). Other methods such as remote sensing, tunnel studies and on-road or near-road measurement and modelling have been commonly used for emission model validation purposes, and have contributed to an increased understanding of model accuracy and real-world emissions behaviour of vehicles.

There are, however, specific issues with and points of attention for all emission measurement methods. The strengths and weaknesses of *both* laboratory measurements, on which VEPM is based, and independent remote sensing measurements have therefore been examined. This is important to ensure that a valid comparison of methods is made and to better understand where potential differences may come from.

There are fundamental differences between emission estimates derived from remote sensing measurements and laboratory-based model predictions that need to be accounted for in a validation study.

[1] The Vehicle Emission Prediction Model. (http://air.nzta.govt.nz/vehicle-emissions-prediction-model)

Nine fundamental differences were identified and discussed in detail in this study:

- 1. Different sampling strategies
- 2. Different measurement techniques
- 3. Different determination of emission factors
- 4. Different levels of detail in the on-road fleet mix
- 5. Different spatial and temporal resolution regarding driving conditions
- 6. Different meteorological (test) conditions
- 7. Different fuel quality
- 8. Different impact of ageing effects
- 9. Different impact of vehicle loading

Vehicle emission models are often used beyond their intended purpose and capabilities, resulting in errors. This may partly be explained by lack of understanding of the model development process, the underlying empirical data and the exact definition of variables used by the model, and clear documentation of these matters.

VEPM's intended use is for emission estimation at road network level. The use of 'average speed' as the only variable to capture the impact of driving conditions on emission levels can introduce significant errors in the emission predictions with the same mean speed, but with different levels of speed fluctuation (i.e. different drive cycles or driving patterns). For local assessments (e.g. street level) substantial errors up to factor of four in emission predictions have been reported for speed-time traces with the same mean speed but with different levels of speed fluctuation.

More detailed vehicle emission models in addition to VEPM will be needed in the future for adequate emission predictions at the local level in New Zealand. This study has developed and presented a possible 'hybrid' approach to develop such as model, using remote sensing data.

A proof-of-concept hybrid model was created to illustrate how the rich New Zealand remote sensing database can be used to develop a genuine New Zealand emissions model, which can be used to 'validate' VEPM, but also allows for emission predictions at a significantly higher resolution in space and time than VEPM. The hybrid model links binned remote sensing data with CO_2 prediction algorithms through an appropriate vehicle classification scheme and use of a proxy variable for engine power (VSP).

We conclude that a validation study that takes into account the issues listed above will be capable of both validating the current VEPM model and allowing the development of a hybrid model specific to New Zealand that will give higher spatial and temporal resolution, as well as increasing confidence in predictions.

1 Introduction

Road traffic is an important global source of air pollution and greenhouse gas emissions and its significance is increasing. Around the World, air quality and greenhouse gas emission impacts of road traffic are commonly evaluated using transport and emission models and, in the case of air pollution, dispersion and exposure models. The scale of application of such models ranges from a single point near a road to entire urban or regional road networks, and even national or global scale.

Overseas road traffic emission inventory models such as COPERT in Europe and Australia and MOBILE/MOVES and EMFAC in the United States are well-known and often used in practice. To account for the unique New Zealand vehicle fleet characteristics, New Zealand has developed its own traffic emission inventory model called VEPM (Karr et al., 2008). The latest version of the NZ vehicle emission inventory model VEPM, version 5.1, was released in June 2013 (http://air.nzta.govt.nz/vehicle-emissions-prediction-model).

In VEPM mean emission factors are expressed as grams per vehicle kilometre of vehicle travel (g/veh.km). The emission factors are a function of average speed, where average speed is defined as the overall speed on a section of road or for an entire journey. As VEPM is derived from European models (COPERT, NIAE), VEPM is (indirectly) developed from – and reflects – overseas laboratory emission measurements on a sample of European vehicles involving various European drive cycles.

In comparison, COPERT Australia was developed specifically for Australian conditions as it was shown previously that direct application of COPERT and other models led to large prediction errors (Mellios et al., 2013; Ntziachristos et al., 2013; Smit and Ntziachristos, 2012; 2013). This was made possible due to a large database of empirical laboratory emission measurements of Australian vehicles. A similar database is not available in New Zealand. However, New Zealand has a large amount of useful remote sensing data that can be used to verify VEPM predictions or for model development, as will be discussed in this report.

An internal review of the VEPM background documentation (Karr et al, 2008) at NIWA by Robin Smit (Smit, 2012) concluded that the report is well-written and generally follows a logical process. However, given the many assumptions, intermediate computation steps (e.g. equivalency determination) and the use of a number of overseas models that underpin VEPM – which are inevitable due to a general lack of empirical NZ data – validation is essential to test the adequacy of VEPM model performance.

This requirement is not restricted to VEPM. As all models are simplifications of reality, assessment of prediction accuracy is essential for all. This is important as poor emission forecasting will cause poor policy decisions if left unchecked. There are various ways to (partially) validate motor vehicle emission models and they include tunnel studies, near-road concentration measurements, on-board emissions testing, laboratory emissions testing and remote sensing. However, all these methods have their own strengths and weaknesses (Smit, Ntziachristos and Boulter, 2010), and the weaknesses in particular must be explicitly considered in any validation study.

2 Objectives and Outcomes

The main objective of this project is to scope out and test the best and most useful ways for validation of the New Zealand vehicle emission model VEPM[2] using remote sensing data for CO, THC and NO_x.

Validation is defined as the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended use of the model using independent datasets.

This report presents the results for Phase 1 of the scoping project in which the focus is on method development and testing.

It is noted that the report contains much information, presented in a dense format, so a basic knowledge of vehicle emission assessment on the part of the reader is recommended.

The project will generate the following outcomes:

1. Discuss the New Zealand remote sensing data and VEPM, and their potential strengths and weaknesses.
2. Examine the relevance of potential issues with the RSD data and ways to address them.
3. Discuss and demonstrate various (possible) ways to compare remote sensing data with VEPM emission factors.
4. Promote the use of RSD data to improve vehicle emission modelling in New Zealand.

This report will discuss the strengths and weaknesses of both laboratory and remote sensing measurements. This is important to ensure that a valid comparison is made between VEPM and remote sensing data and to better understand where potential differences may come from.

Testing the overall accuracy of vehicle emission models is challenging, as the 'true' emission values in urban road networks are unknown and cannot practically be determined by measurement; i.e. this would require continuous emission measurement of all vehicles in the area and period concerned.

As a consequence, VEPM predictions and remote sensing data are both independent *estimates* of the true vehicle emissions in a road network. However, if both VEPM and the remote sensing data show similar predictions of vehicle emissions, then this would increase the level of confidence that VEPM predictions are accurate. If not, then either one or both VEPM predictions or remote sensing estimates are not accurate. In this case, more validation work, using other independent methods such as tunnel studies, near-road air quality sampling etc., and further analysis are required to determine where the (main) errors occur (specific traffic situations, specific vehicle classes, etc.).

[2] Vehicle Emission Prediction Model.

3 Differences between VEPM and remote sensing

Laboratory measurements ('bag' and 'modal', engine and chassis dynamometer) using drive cycles have remained the prominent empirical base for vehicle emission model development around the world (e.g. Smit et al., 2009).

Other methods such as remote sensing, tunnel studies and on-road or near-road measurement/modelling have been commonly used for emission model validation purposes, and have contributed to an increased understanding of model accuracy and real-world emissions behaviour of vehicles. There are, however, some issues with these approaches that complicate its direct use in emission models, such as a limited range of operating or traffic conditions (e.g. high speed driving conditions in tunnels). The strengths and weaknesses of remote sensing and laboratory vehicle emissions testing are explored in detail in this section.

The use of remote sensing to measure vehicle emissions is not new, dating back to 1971 when experiments were conducted in the US to construct a remote exhaust measurement system. Since then a number of remote sensing systems have been developed, in most cases with very similar operating principles (e.g. Bishop et al., 1989; Guo et al. 2007; Gala and De la Fuente, 2012). Disagreement between vehicle emissions model predictions and tunnel measurements in the US in the 1980s-1990s, as well as an initial interest to supplement or even replace routine emissions inspections, facilitated increased use of a remote sensing device (RSD) as an independent approach to measure real-world vehicle emissions (e.g. Eisinger and Wathern, 2008). Remote sensing has been used extensively over the last two decades or so for various purposes around the world. The following applications have been reported in order of occurrence:

- Identification of high-emitting vehicles ('maintenance and repair programs') or low-emitting vehicles ('clean screening'), and their emissions behaviour.
- Examination of on-road vehicle emissions distributions.
- Assessment of the impacts of ageing on vehicle emissions, inspection and maintenance programs and changes in vehicle emission control durability over time.
- Assessment of impacts of specific vehicle types, fuel types and emission control technologies[3] on vehicle emissions.
- Trend analysis of on-road fleet emissions.
- Validation of commonly used vehicle emissions models (e.g. COPERT, MOBILE).
- Comparison of remote sensing measurements with tunnel studies, near-road air quality studies and laboratory measurements using drive cycles.
- Use of remote sensing as a public vehicle emissions information system.
- Developing fuel-based emission inventories as an alternative to travel-based (VKT) emission inventories.
- Assessment of the impact of vehicle operating conditions on average pollutant-to-CO_2 concentration ratios.

[3] Including non-certification 'defeat' algorithms in the engine management system which lead to elevated real world emissions (see e.g. Jiménez et al., 2000; Burgard et al., 2006).

Comparing remote sensing data with VEPM predictions is not straight forward, and a number of issues need to be considered and addressed before a valid comparison can be made. These aspects are summarised and discussed in this section.

There are fundamental differences between emission estimates derived from remote sensing measurements and laboratory-based VEPM predictions that need to be accounted for in the validation study:

1. different sampling strategies;
2. different measurement techniques;
3. different determination of emission factors;
4. different levels of detail in the on-road fleet mix;
5. different spatial and temporal resolution regarding driving conditions;
6. different meteorological (test) conditions;
7. different fuel quality;
8. different impact of ageing effects; and
9. different impact of vehicle loading.

These points are critically reviewed and further discussed in this section. For the validation study it is important to know if both approaches introduce systematic errors (bias) in measured emission rates and to control, to the extent possible, for factors that (may) introduce bias.

3.1 Sampling strategy

Remote sensing and laboratory testing use a different approach to vehicle emissions data collection. Remote sensing typically collects ‘snapshot’ on-road emissions data from a relatively large sample of *moving* vehicles, whereas laboratory testing collects drive cycle (trip) based emissions data from a relatively small sample of both stationary and moving vehicles under controlled conditions.

3.1.1 Remote sensing

Measurement objectives[4] and prevailing site conditions determine location, method and time of sampling with remote sensing. For emission model validation, emission trend analysis or (fuel-based) emission inventory development, remote sensing data for *multiple* sites are required.[5] These sites need to be representative of the range of traffic and weather situations in (urban) areas, to the extent this is feasible. In contrast, high-emitter or ‘clean vehicle screening’ requires more narrowly defined measurement conditions, e.g. in terms of specific operational conditions such as hot engines and emission control systems and a limited range of engine loads.

[4] For instance, high-emitter identification or characterisation of on-road fleet emissions.
[5] It has been shown that site selection can strongly influence remote sensing measurements both in New Zealand (NZTA, 2014) and overseas (e.g. CRC, 2006). Hence, the use of RSD data from multiple sites is an essential requirement for an emission model validation study.

Remote sensing measurements are usually made during daytime in dry weather conditions at non-dusty single lane locations, where vehicles are under light to moderate acceleration conditions and on a positive road gradient with a warm engine and emission control system (hot running conditions). Remote sensing can record emission levels for a relatively large number of vehicles, but the system can produce a significant portion of invalid data, and may exclude relevant vehicle types and traffic conditions.

Invalid data

According to Bishop and Stedman (2008) the proportion of invalid measurements is typically 20-30%, but significantly lower capture rates have been reported. For instance, Rhys-Tyler and Bell (2012) reported 55% invalid measurements, and Mazzoleni et al. (2004b) even more than 70% invalid measurements.

Measurement issues occur in specific situations, for instance:

- Stationary vehicle drive conditions (idling).
- Vehicle speeds outside the range of 10 to 60 km/h.[6]
- Insufficient measurable CO_2 in the plume due to e.g. bicycles, trucks, decelerating vehicles (including engines that use 'fuel cut-off' strategies) or strong wind conditions that will disperse the plume rapidly[7].
- Measurements outside preset pre-specified concentration ranges or preset pollutant-to-CO_2 ratios.
- Failure to identify pre-car and post-car periods to detect the presence of a vehicle ('beam blocks') on busy roads where vehicles pass in rapid succession.[8]
- Failure to identify specific vehicle types, such as cars with caravans.[9]
- Heavy rain, snow or dusty roads, which inhibit light intensities and introduce additional measurement noise.

Excluding specific vehicle types

The RSD setup may not detect specific vehicle types, such as articulated trucks and buses. These heavy-duty vehicles (HDVs) have a disproportionate contribution to fleet emissions for several pollutants. As a consequence, accurate and comprehensive quantification of emissions from these vehicles is essential. However, remote sensing devices are normally set up at ground level with a sensing beam height of 20-45 cm. This setup is fine for light-duty vehicles (LDVs), but it can exclude a significant portion of heavy-duty vehicles with vertical exhausts or other oriented exhaust pipes.

This is the case in the US and Australia, but not in Europe where most trucks have ground level exhaust systems. The situation in New Zealand is a mixture of the two, and a portion of the on-road

[6] The viable instrument range is 10 to 120km/hr, but at high speeds vehicles may have sampling times that are significantly shorter than 0.5 seconds (50 data points at 100 Hz sampling frequency).
[7] There are generally more invalid readings in situations with lower emission levels (e.g. low or negative grades, cruising or decelerating vehicles).
[8] For instance, FEAT requires more than 0.5 seconds between unblock and block, otherwise it will 'restart'. Interrupted data collection could be caused by e.g. trailers.
[9] Cars with caravans may be measured accurately, but the caravan prevents number imaging, so the measurement cannot be allocated to a vehicle.

truck and bus with ground level exhausts fleet have been measured using remote sensing (Bluett, Kuschel and Unwin, 2010). It is recommended that any future remote sensing programs in New Zealand include trucks and buses with vertical exhausts.

This will impose a few new challenges to RSD program design. For instance, trailers will interfere with the measurements as they produce well-dispersed plumes causing an issue for measurement at the rear of the vehicle. As a consequence, modifications to the RSD setup are required. These include the use of multiple RSDs, the use of scaffolding to raise the detector height for HDVs with vertical exhaust configurations and lowering the measurement beams to read the plume under HDVs with low exhausts. Other modifications are specific vehicle detection equipment (manual or automated) and detection algorithms, longer scanning times (e.g. 1.0 instead of 0.5 seconds) and data post-processing procedures (e.g. Jiménez et al., 2000; ESP, 2010).

Excluding specific traffic conditions

Procedures are used to automatically filter invalid data from remote sensing measurements.[10] There is, however, a tendency for invalid readings in particular traffic situations (e.g. low engine power conditions, congested conditions), and hence smaller sample sizes, or no data for these conditions. As a consequence, there is a risk for biased emission factors if they are derived from filtered remote sensing data and applied to situations for which data are not or scarcely available.

To illustrate this, Jiménez et al. (2000) reported on a remote sensor technique with a long path length that was used across a four-lane high speed highway. Although this technique had its own challenges (wind effects, overlapping vehicles, etc.), the NO/CO_2 distribution was found to be substantially higher as compared with RSD measurements made at a location where vehicles just left a weighing station. The effect of 'defeat devices' was argued to be the main cause for the difference. Defeat devices optimise fuel economy at the expense of NO_x emissions and are typically activated after several minutes of high speed driving. So they were expected to be activated on the highway, but not at the other location. This is important because, whatever the technical reason, it shows that care is needed by translating 'urban' remote sensing measurements to freeway conditions.

Modification of driving behaviour

A potential source of bias is modification of 'normal' driving behaviour at RSD locations. This may include drivers releasing the accelerator and reduce speed when they see the RSD (e.g. Bishop et al., 2001; Kraan et al., 2012), which can affect emissions. For instance, it may cause short-lived HC emission spikes, which are not representative of the general low emission levels of modern vehicles. This could create a bias in measured HC/CO_2 ratios.

Emission types

A generally recommended remote sensing site is an uphill, curved off-ramp from a freeway. As a consequence, remote sensing is generally assumed to exclude evaporative HC emissions or cold start emissions, although there are some exceptions. For instance, Sjödin and Lenner (1995) measured emissions at four sites and found that CO and HC emissions are substantially elevated at urban sites with an expected significant proportion of cold start vehicles.

[10] For instance, the FEAT system uses preset minimum CO_2 concentration values and specific CO, propane and NO concentration values. These are combined into pollutant to CO_2 ratio limits that are used to flag invalid measurements.

Few researchers have attempted to determine which specific vehicles operate in cold start mode. Monateri et al. (2004) report on the (manual) use of an infrared camera to measure the heat signature of passing vehicles (exhausts, tyres, underbody) and determine if a vehicle is in cold start mode. CRC (2006) used a less thorough approach and simply excluded specific hours of RSD data for sites where "*5% of newer vehicles had excessive HC emissions*".

Cold start emissions, and CO and HC in particular, are of increasing relative importance, due to strong reduction in hot running emissions on a per vehicle kilometre basis (Smit and Ntziachristos, 2013). As a consequence, it is important for a validation study to examine the (expected) proportion of cold start vehicles in the remote sensing measurements for each location, and to determine if the measured emissions are likely to be hot running emissions, or not. This can be done, for instance, through analysis of license plate information and determination of the distance and likely travel time between registration address and RSD location. If travel time exceeds a few minutes it can be assumed that the vehicle has negligible cold start emissions.[11]

Although several remote sensing studies explicitly state that remote sensing only measures exhaust emissions (e.g. Pokharel et al., 2002; Mazzoleni et al., 2004a; 2004b; Bishop et al., 2012; Fujita et al., 2012), this is incorrect. In fact, St. Denis and Roeschen (2012) actually *used* remote sensing to identify vehicles with high evaporative emissions (e.g. fuel leaks), which was verified with a number of evaporative SHED tests showing that evaporative emissions are measured with RSDs. It is unclear, however, to what extent evaporative running loss and resting loss VOC emissions[12] are captured in the 'exhaust plume'. But it can be expected that HC emissions measured with remote sensing are higher compared to hot running exhaust emission factors in VEPM, as they will also include some evaporative emissions.

Finally, although remote sensing has not been used extensively (yet) to determine PM emissions from motor vehicles, it is possible that non-exhaust PM emissions are to some extent included in the measurement.

3.1.2 Laboratory

Laboratory measurements are usually conducted in a highly (quality) controlled environment, which includes predefined and repeatable driving behaviour (drive cycles), and explicit measurement of most relevant types of emissions. This includes hot running, cold start, warm start, evaporative diurnal and evaporative hot soak emissions, but typically excludes running loss emissions and emissions of non-exhaust particulate matter due to tyre, brake and road wear.

Laboratory measurements can be influenced and biased by e.g. high emission events (Smit, 2013a). This can result in clipped emission traces in e.g. modal testing, but also affect the more robust bag sampling measurements. This is evidenced by some remarks made in the NISE2 study where about 400 petrol cars were tested (Orbital, 2009):

[11] There are different factors that contribute to cold start emissions and they differ in magnitude as well as duration of impact. At an ambient temperature of about 20°C, hot running conditions should generally be achieved for all relevant vehicle components (engine, transmission, catalyst) within 15 minutes of driving. However, catalyst 'light-off' conditions and tight control of the air-to-fuel ratio, which together largely determine the magnitude of cold start emissions, will be achieved much faster than this, i.e. typically within a minute of engine start for modern vehicles and a maximum of a few minutes for older technology vehicles (Smit and Ntziachristos, 2013).

[12] Evaporative vehicle emissions are emissions that emanate from the fuel system. Diurnal and hot soak emissions are not captured with remote sensing as these emissions are generated when vehicles are parked. This is less clear for running loss emissions, which are a result of heating of the fuel system during vehicle operation and resting-losses, which are due to permeation of plastic/rubber materials in fuel system (tanks, lines, fittings) and liquid leaks.

- "*$PM_{2.5}$ emissions on the first phase of the ADR37 test were the highest recorded on this project, before the particulate sampling system stopped due to overloading*", and
- "*THC and CO emissions were at levels too high to be recorded for some test phases*".

So laboratory measurements, and the models developed from them, can exclude high-emitters from the emissions database when the measurement equipment cannot handle excessive emission levels. This aspect is compounded by the relatively small sample sizes typical for laboratory testing due to costs, with an associated risk of inadequately reflecting the proportion of high-emitters (with highly variable emission levels) that occur in the on-road fleet.

There are more subtle factors that may create bias, but they are difficult to quantify. An example is the representativeness of gear shift behaviour of vehicle operators in laboratory conditions to gear shift behaviour in on-road vehicle fleets. Another example is how well dynamometer settings and loading algorithms reflect on-road driving conditions.

The laboratory sample may also not (entirely) reflect the composition of the in-service vehicle fleet due to issues with sourcing test vehicles and achieving truly random vehicle samples. For instance, it may be geared towards newer vehicles (e.g. government fleets, rental vehicles).

3.1.3 Conclusion

It is not possible to make a statement about which measurement method leads to more accurate vehicle emission predictions because both methods have particular strengths and particular challenges that could introduce bias in emission predictions based on these measurements. Rather, the comparison, and possible combination, of different measurement methods (laboratory and remote sensing, but also on-board, tunnel, near road, etc.) should be the focus of developing accurate vehicle emission models.

3.2 Measurement techniques

Remote sensing uses different measurement methods as compared with standard laboratory testing on which VEPM is based, and these differences need to be considered when a validation study is conducted.

3.2.1 Remote sensing

Remote sensing system uses the principle that the majority of gases will absorb light at particular wavelengths. It measures on-road emissions by absorbance of ultraviolet (UV) and infrared (IR) light across an open (optical) path using wave-length specific detectors for different air pollutants. The remote sensing device (RSD) consists of an IR component for detecting CO, CO_2 and HC, together with an UV spectrometer for measuring NO. Remote sensing systems have evolved over time and have become more accurate and less sensitive to interference (e.g. Popp et al., 1999; Sjödin and Andréasson, 2000; Bishop et al., 2001; Mazzoleni et al., 2004b). More recently, other pollutants can be measured with RSD such as PM (smoke), NH_3, NO_2 and SO_2 (e.g. Jiménez et al., 2000; Mazzoleni et al., 2004a; 2004b; Sjödin and Jerksjö, 2008; Carslaw and Rhys-Tyler, 2013). A typical setup is shown in Figure 3-1 and Figure 3-2.

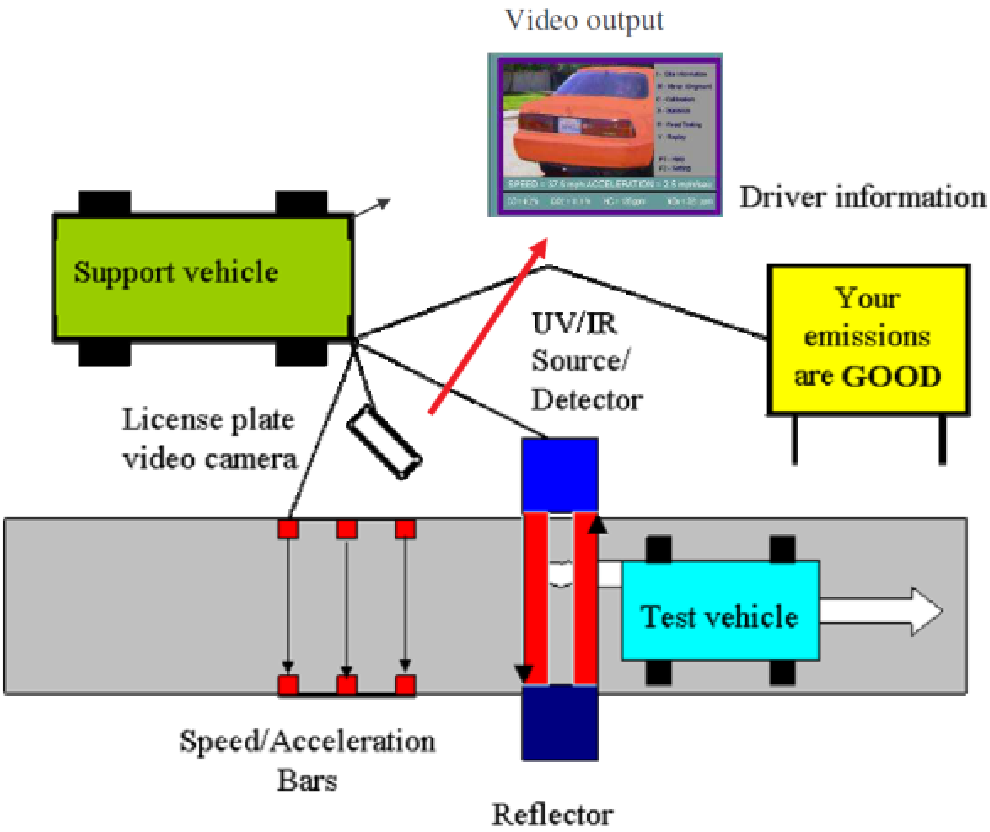

Figure 3-1: Example of a remote sensing measurement setup schematic. (Source: NIWA, 2008a)

Figure 3-2: Photo of remote sensing measurement setup including speed/acceleration bar and a digital camera. (Source: Gala and De la Fuente, 2012)

The source/detector unit is positioned on one side of a single-lane road, with a corner cube reflector on the opposite side. Beams of IR and UV light are passed across the roadway into the corner cube reflector and returned to the detection unit. A reference detector monitors a portion of the spectrum where no exhaust gas compounds are absorbed to correct measured pollutant signals for fluctuations in transmitted intensity due to e.g. particulate matter or fluctuations in source intensity. Vehicles are identified through video or camera imaging of license plates, and driving behaviour is quantified by upstream measurement of speed and acceleration. License plate information is matched with vehicle registration data to obtain specific vehicle information such as vehicle type, year of manufacture, fuel type, and so on.

The remote sensor used by NIWA (ESP4000EN, ESP Inc. Denver, USA) reports concentration levels in the exhaust gas, i.e. CO_2 (%), CO (%), PM (mg/m^3), HC (ppm) and NO (ppm), corrected for water and excess oxygen not used in the combustion process. However, the exhaust plume path length and the density of the observed plume are highly variable from vehicle to vehicle and are a function of the height of the vehicle's exhaust pipe, wind direction and speed, and turbulence behind the vehicle, amongst other factors. The RSD can therefore only reliably measure ratios of CO, HC and NO to CO_2. These ratios are assumed to be constant in a particular vehicle's exhaust plume. Due to the short time period involved in the measurement[13], this is expected to hold true for reactive species such as NO as well. However, RSD data are naturally noisy and sufficiently large sample sizes are required to obtain significant results.

The remote sensing measurements include a 'background correction' by subtracting the concentration measurements just before the 'beam block'[14] (detecting vehicle presence) from the concentration measurements just after the 'beam block'. It is unclear, however, to what extent mixing of emission plumes from different vehicles or residual plume interference affects overall remote sensing measurements.

In this respect, Stephens and Cadle (1991) examined the extent to which remnants of plumes from previous vehicles affect accuracy. These researchers analysed and classified their remote sensing data as a function of time between measurements and whether a previous car was a high or low emitter of CO. The results are shown in Figure 3-3.

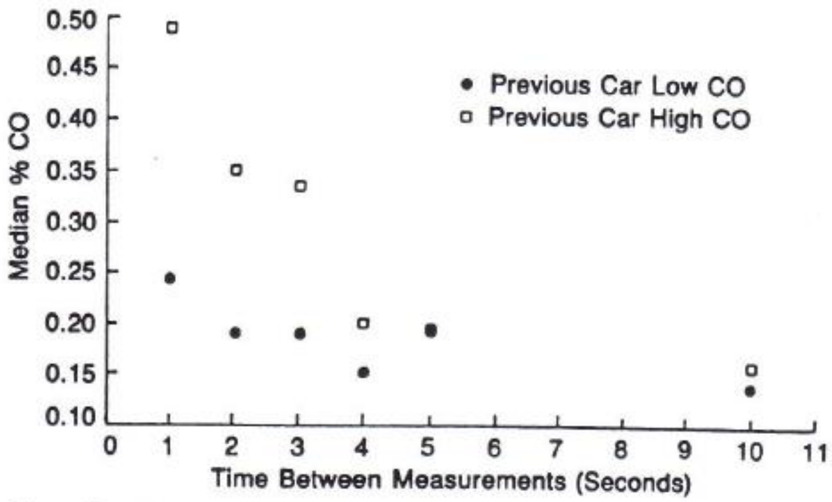

Figure 7. The median percent CO measured for all cars following behind either (1) high CO emitting vehicles (more than 5 percent CO) or (2) low CO emitting vehicles (less than 1 percent CO), as a function of car separation time.

Figure 3-3: Examination of residual plume interference. (Source: Stephens and Cadle, 1991)

[13] Each ratio reported is the mean of half a second of 10 ms interval measurements, giving a sample size of approximately 50 values for each reported ratio.
[14] For instance, 0.2 seconds or 20 data points at 100 Hz sampling frequency (e.g. Bishop and Stedman, 1996).

Their results show that the median CO measured in exhaust plumes is significantly higher when previous cars have high CO emissions (> 5%) and time intervals are less than 4 seconds. A later study (Mazzoleni et al., 2004b) remarked that 'carryover effects' when a clean vehicle follows a dirty vehicle were "not evident in the data", but did not further elaborate how this was determined.

Remote sensing measurements are subjected to automated error checking and estimation algorithms with preset error limits. An example is a set of non-linearity testing algorithms aiming to detect mixing of two plumes of (significantly) different concentrations, and reject these measurements. For instance, Bishop et al. (1989) fitted a linear regression model where CO_2 concentration is the predictor variable and the pollutant concentration (CO) the response variable. The slope was then used to determine the CO/CO_2 ratio, but the slope standard deviation was not allowed to exceed ±20%, and a minimum CO_2 concentration was required.

Nevertheless, residual plume interference may still introduce some bias in remote sensing measurements by inclusion of small amounts of residual emissions, but the relevance of this effect is unclear. It is therefore recommended that the New Zealand RSD data are examined to see if a similar effect can be observed as in Figure 3-3.

Instrument calibration is performed by comparing measured ratios with those in a puff of certified gas containing a mixture of e.g. CO, NO, propane and CO_2. Calibration should be performed regularly to account for variations changes in meteorology and instrument sensitivity. Reported calibration frequencies vary from hourly (Carslaw et al., 2011b; Borken-Kleefeld and Chen, 2015) to twice daily (Burgard et al., 2006; CRC, 2007).

As compared with the controlled environment of laboratory measurements, it is clearly more challenging for remote sensing to measure vehicle emissions accurately. For instance, exhaust emissions are diluted rapidly and the RSD needs to strike the right balance between measuring emissions from all on-road vehicles and preventing incorrect or inaccurate measurements due to low signal-to-noise ratios.

RSD measurements can also be sensitive to equipment vibration affecting beam alignment due to moving vehicles shaking the roadway or producing air pulses (e.g. Bishop and Stedman, 1996). For instance, Mazzoleni et al. (2004b) suggested a positive offset/bias in remote sensing PM measurements is caused by this vibration. Stephens and Cadle (1991) also mentioned possible bias of infrared emissions from hot exhausts which will affect CO and CO_2 emission measurements. An experiment with the infrared source off detected infrared emissions from 8% of the vehicles.

Given the challenging testing environment, it is essential to perform thorough quality checks on remote sensing data before it is used in the validation study. This is in addition to the internal and automated filtering and verification procedures already used by the RSD. This can be done in different ways. For instance by comparing emission measurements for each site over time and between sites (e.g. binned by vehicle year of manufacture, also see discussion in section 3.2.3 regarding offsets), or as was shown in Figure 3-3, to ensure measurement consistency and identify any detector problems that may have occurred for specific pollutants.

3.2.2 Laboratory

Figure 3-4 shows an example of a typical laboratory vehicle emission measurement setup.

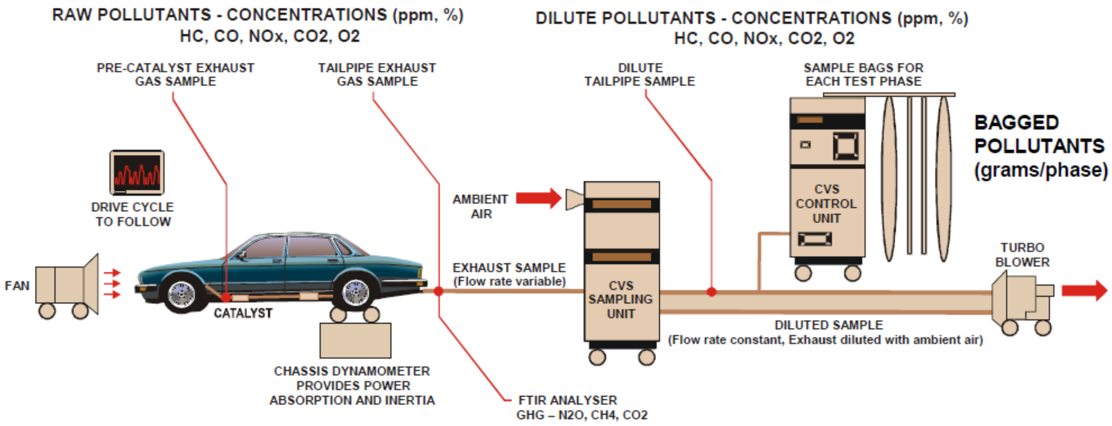

Figure 3-4: Example of Laboratory Measurement Setup. (Source: Orbital, 2009)

Laboratory emissions are measured with standard pollutant analysers. Exhaust emissions are typically sampled from a dilute proportion of tailpipe emissions using a CVS (Constant Volume Sampling) system. This sample is stored in a sample bag. The bag usually samples the emissions for a complete phase of the test (i.e. drive cycle or parts thereof). These results are background corrected for ambient concentration levels.

Exhaust emissions results can also be measured as “modal” data. In this case, the emissions are typically sampled raw at a 1-10 Hz measurement frequency using exhaust gas analysers with high range detectors. No compensation is made for ambient levels for these raw readings. The integration of these data should give values similar to the level measured in the bag, and this is often used as a data quality verification step (Smit, 2013a).

Laboratory measurements typically use FID (Flame Ionisation Detector) for THC measurements, Chemiluminescence for NO_x (NO, NO_2), a gravimetric filter method for PM and NDIR (Non-Dispersive Infra-Red) for CO and CO_2 measurements. Laboratory grade analysers tend to use expensive high precision systems. Lower-cost systems are also available and used for e.g. emission test and repair programs, and whilst they are of the same type, they will generally have less accuracy, precision, repeatability and stability.

Laboratory measurements have their own challenges and issues. For instance:

- PM measurements may be affected by sampling issues such as loss, condensation, and other factors rendering the result unreliable.
- FID is regarded as the most accurate equipment for detecting hydrocarbons. They are very sensitive to HCs, and the FID response tends to be reasonably linear across a wide range of concentrations. FIDs are also insensitive to H_2O, CO_2, SO_2, CO and NO_x, preventing interference issues. However, FID will not adequately detect all hydrocarbon species present in vehicle exhaust, which is an artefact. The molecular structure and composition of the gas, as well as detector design limitations, will affect the FID response. For instance, hydrocarbons that contain other functional groups (e.g. carbonyl, alcohol, amines) are not ionized as effectively as compared with pure hydrocarbons, and will give a weaker or no

signal. FID responses are commonly assumed to be proportional to molecular weight, but no correction is applied for this "non-linearity" in the FID response, i.e. the compound-specific response factors are all set to unity (De Saint Laumer et al., 2010). Given the substantial variability that is reported for FID response factors (Tong and Karasek, 1984), not correcting FID responses can significantly affect the accuracy of reported vehicle emission levels.

So it is important to realise that even controlled high quality laboratory measurements are not equivalent to 'true' vehicle emission values, but are estimates of them.

3.2.3 Comparison of laboratory and remote sensing vehicle emission testing

A literature review was conducted to find studies that compared remote sensing with dynamometer based vehicle emissions testing under different test conditions. Older studies are less relevant for the VEPM validation study, as remote sensing equipment has continued to evolve and improve over time. On the other hand, fleet emissions have substantially reduced as time progressed, posing new challenges for remote sensing measurements. So the level of agreement between remote sensing and laboratory tests may or may not have been better in the past. Older studies are included here for the sake of completeness.

It is noted that the RSD results will to some extent also depend on the actual device that is used. For example, Stephens and Cadle (1991) compared two different remote sensing instruments and found a mean instrument-to-instrument ratio of 0.88 and a small offset for measured CO/CO_2 ratios. Later studies also reported differences between different remote sensing techniques (Mazzoleni et al., 2004b).

<u>Pre-2000 comparison studies</u>

- Bishop et al. (1989) compared remote sensing data for CO, expressed as grams CO per gallon of fuel, with results of FTP dynamometer tests and reported a reasonable correlation coefficient of 0.81.

- Sjödin et al. (1997) compared remote sensing results for CO and HC with idle tests and dynamometer tests and concluded that remote sensing may be a useful tool to identify high-emitting cars on the road that no longer comply with (dynamometer based) legislative standards, but that correlation with the idle test was very poor. Poor correlation with idle tests have also been reported in later studies (Mazzoleni et al., 2004a).

- TRL (1998) directly compared RSD measurements with dynamometer testing, but the focus was on the ability to detect high-emitting vehicles, and not the accuracy of RSD measurements. The data presented in this report show that CO, HC and NO RSD measurements generally follow the trend of dynamometer measurements, but that there is a substantial amount of scatter. An example is shown below for CO (Figure 3-5). It is noted that a direct comparison is not possible due to the difference in units (g/km versus %CO):

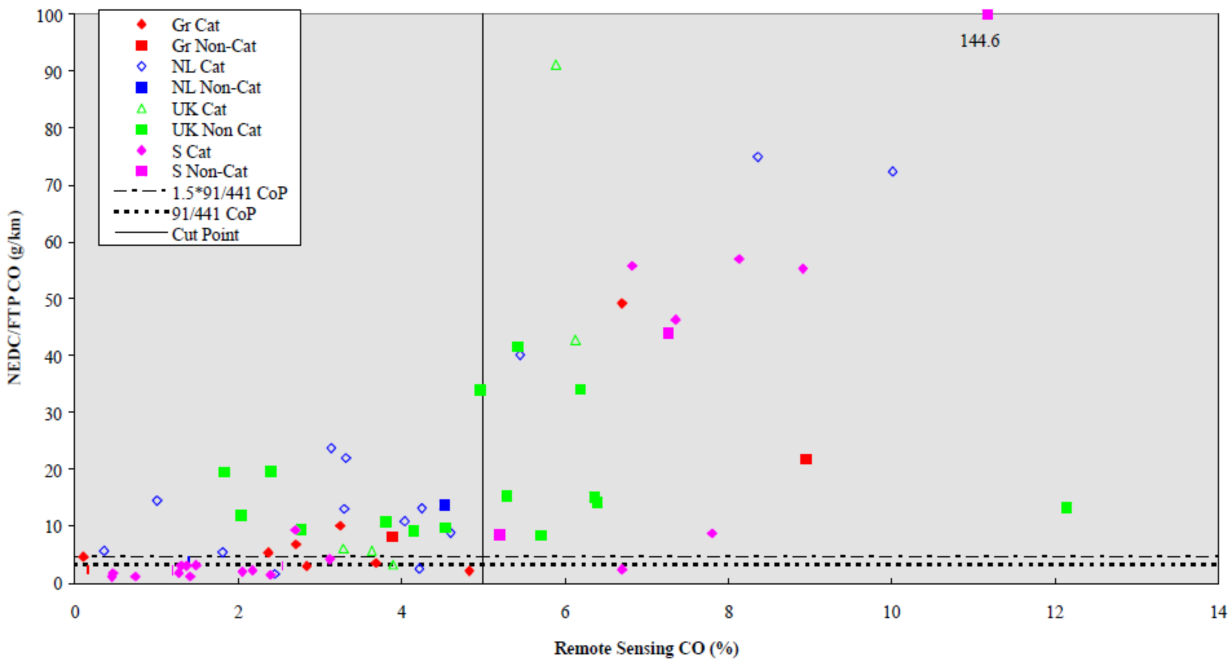

Figure 3-5: Dynamometer versus RSD test results (CO). (Source: TRL, 1998)

Post-2000 comparison studies

Yanowitz, McCormick and Graboski (2000) compared remote sensing with the results of tunnel studies and dynamometer tests in the US for HDVs and concluded that there is reasonable comparability of HDV emission factors (g/l fuel) between the three methods. Taking data from Table 8 in this paper, differences between remote sensing and dynamometer testing are calculated[15] to vary from -28% to +65% for NO_x and from +5% to almost a factor of 2 higher for CO. The variability in RSD emission results was larger as compared with the tunnel study data, which is probably due (to some extent) to the more narrowly defined traffic conditions in tunnels.

Pokharel *et al*. (2000) compared remote sensing CO, THC and NO data with dynamometer I&M data (IM240) *aggregated* to vehicle model year and using g/kg fuel as the comparison variable, and found good correlation[16] (R^2 typically > 0.95).

[15] (RSD-LAB) ÷ LAB
[16] Note: these authors actually refer to the coefficient of determination, not (linear) correlation.

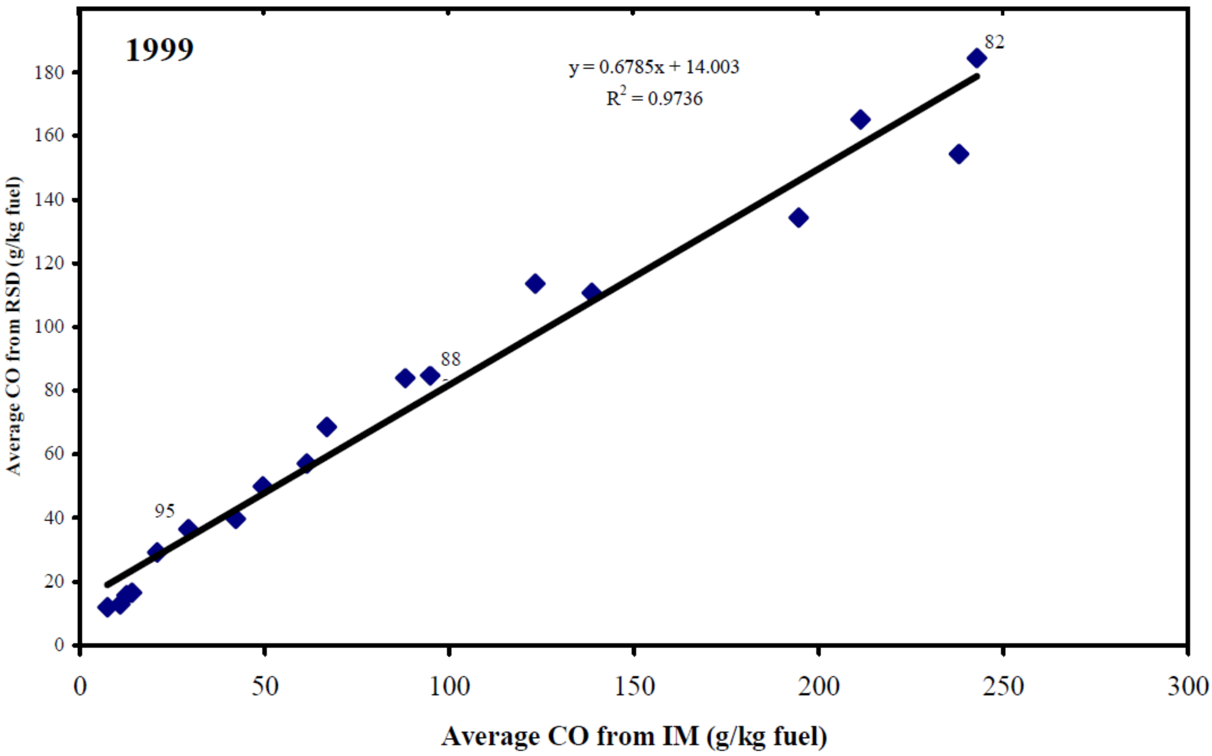

Figure 3-6: Example of correlation plot for CO between IM240 and RSD for one year. (Source: Pokharel et al, 2000).

This implies that remote sensing is able to accurately capture trends and changes in vehicle emissions due to fleet turnover. There are, however, a few interesting points that emerge from this comparison:

- The *slope* of the linear regression line is significantly below unity for CO (about 0.6) and HC (about 0.4). This implies that remote sensing results tend to be substantially lower than those measured with laboratory tests. Pokharel *et al*. (2000) hypothesized that this may be due to differences in vehicle engine load (i.e. driving conditions and grade). For HC the differences were explained with the 'Singer and Harley Factor', which is discussed later, to correct for the difference with FID measurements. The slope for NO was close to unity, but dynamometer NO_x was directly compared with RSD NO, introducing a small error.

- There are significant intercepts for the linear regression lines, in particular for HC and NO. This suggests that very low dynamometer measurements are not replicated in the remote sensing data. The question is if:
  - 1) this offset also shows up in the NZ remote sensing data *and* is <u>constant</u> for all vehicle results, and, if so,
  - 2) this offset is an actual remote sensing artefact or due to other factors such as driving conditions?

Some caution is needed with these results as Pokharel *et al*. (2000) do not appear to have taken all necessary preparation steps needed for a valid comparison (e.g. expressing data as NO_2-eq). CRC (2003; 2006; 2007), however, does discuss and use a method to remove an artificial offset in the RSD

HC measurements due to an '*optical misalignment*'.[17] The adjustment is computed as the lowest value of either the mode or mean of the newest model year vehicles, which effectively assumes that these vehicles emit negligible HC emissions. CRC (2006) also used an offset for RSD CO results for a specific data subset. It is recommended that the NIWA RSD data are examined to determine if a similar offset exists in the New Zealand data.

Elder et al. (2011) compared remote sensing data collected at 6 sites in Auckland for 4 cars with laboratory measurements using the IM240 drive cycle, as well as steady-state tests. These cars were run multiple times (up to 31) past the RSD. A considerable amount of scatter was observed within each VSP range for each specific vehicle for both remote sensing and laboratory measurements, with often higher variability in the remote sensing data than in the laboratory data. Although the repeatability at the same RSD site was good, agreement between laboratory and remote sensing was often assessed as being 'poor'. There are, however, a few comments that can be made here.

- These researchers compared concentrations (ppm, %) and not pollutant-to-CO_2 ratios or fuel-based emission factors, which would have been a more appropriate statistic.
- It is unclear to what extent (cold) starts affected the remote sensing results as there is no information about vehicle preparation.[18]

In any case, the results show that remote sensing data for individual vehicles exhibit a high level of uncertainty. This is an issue for identification of high-emitters, which would require multiple RSD measurements, but not so much for the validation study, as long as the vehicle sample size is sufficiently large and aggregated data statistics are used.

More aggregated comparisons have been made between laboratory test and remote sensing results. For instance, Rhys-Tyler and Bell (2012) compared laboratory measurements over the 11 km long NEDC drive cycle for Euro 2, 3 and 4 petrol and diesel cars with VSP-weighted remote sensing data (expressed as gram pollutant per kg fuel), reflecting the VSP frequency distribution of the NEDC drive cycle. They found large differences up to a factor of 3, but acknowledged that there are various significant caveats in the comparison:

- The RSD only measures NO and not NO_x.
- The laboratory tests reflect new vehicles, whereas the remote sensing includes a mix of vehicle ages.
- An unknown proportion of cold start vehicles in the remote sensing data, whereas all NEDC tests include a cold start.

An important point for the validation study is that vehicle emissions are subject to large levels of variability, which is independent of the measurement method used (laboratory, remote sensing, on-board, etc.). Previous studies (e.g. Supnithadnaporn *et al*., 2011) have shown that the main source of variability is the actual (test) vehicle and that the only way to reduce this variability is to conduct multiple tests per vehicle. This high level of variability is no longer "visible" in VEPM predictions as regression modelling has smoothed the relationship between e.g. average speed and emission rates. This could give the impression that RSD measurements are highly uncertain in comparison with

[17] Bishop and Stedman (1996) reported that the wavelength used to measure HC suffers from some interference from gas/particulate phase water, which is particularly evident for cold vehicles in cold climates. This may be one of the reasons for this artefact.
[18] Laboratory measurements were taken after the remote sensing tests, so vehicles would have been in hot running mode during laboratory measurements, but may be in cold start conditions during remote sensing tests.

VEPM predictions. But it is important to realise that the empirical database on which VEPM is based also contained similar levels of uncertainty.

The comparison studies indicate that substantial differences between remote sensing and laboratory measurements can be expected at individual vehicle level, but that aggregated emissions data will yield reasonable agreement. It is now interesting to see if comparison of remote sensing with other measurement techniques supports this conclusion.

3.2.4 Comparison of remote sensing with other methods

On-board measurement methods have been compared with remote sensing:

Lawson et al. (1990) compared the results from on-board measurements (CO, HC, CO_2) with remote sensing measurements for a single petrol car with a three-way catalyst and found good agreement, stability and reproducibility of the two measurements systems. The ratio of means (tailpipe/remote sensing) for all 34 comparison measurements was 1.05 with an accuracy of ±10%. The correlation for CO was quite similar as the results for NO reported by Kraan et al. (2012) in Figure 3-7.

Kraan et al. (2012) compared the results from on-board measurements (using Portable Emissions Measurement System or PEMS) with remote sensing measurements for a single heavy duty truck driving at different speeds and accelerations past the RSD and found “reasonable correspondence” between the two methods as shown in Figure 3-7.

There appear to be no significant issues with the RSD offset, but the slope suggests that RSD measures higher NO/CO_2 ratios than PEMS at higher NO concentrations.

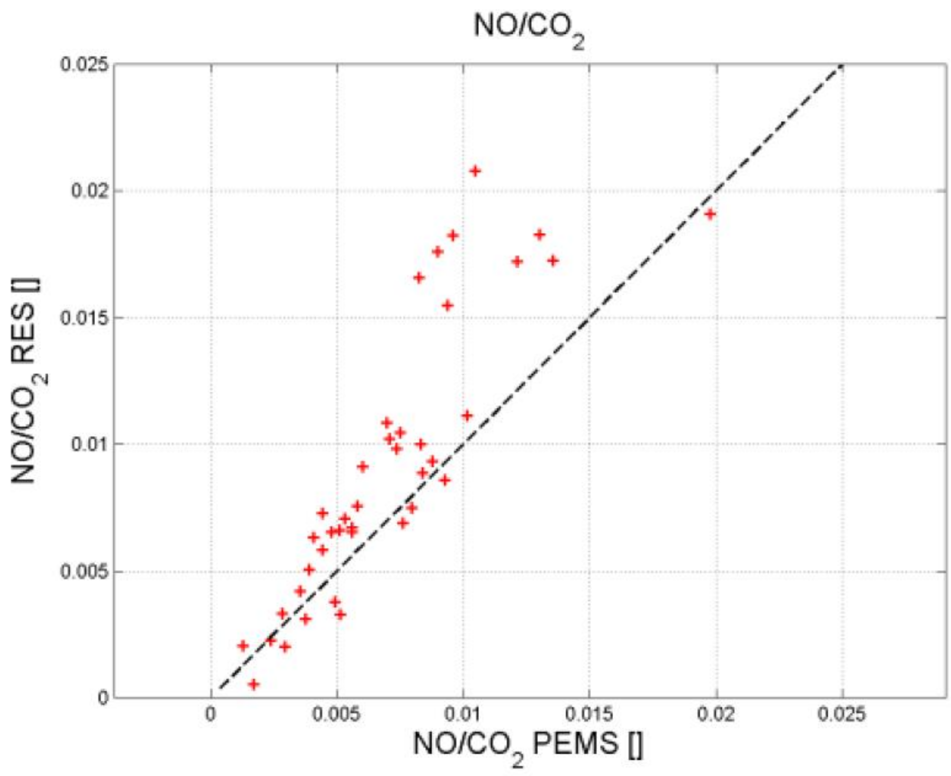

Figure 3-7: RSD versus PEMS for NO. (Source: Kraan et al., 2012).

A few studies compared remote sensing specifically with tunnel measurements. This comparison is of interest as a strength of both methods is the quantification of on-road fleet emissions, including high emitting vehicles, although the spatial scale is different.

- Bishop et al. (2012) found reasonable agreement between fuel-based emission factors (g/kg fuel) determined with RSD and tunnel measurements, i.e. the average remote sensing CO, HC, and NO_x measurements were 9% lower, 41% higher, and 24% higher than the tunnel measurements, respectively.
- Fujita et al. (2012) also concluded that fleet-averaged CO and NO_x emission factors from a tunnel during weekends (low number of trucks which were not measured with RSD) agreed reasonably well with emission factors derived from RSD measurements, i.e. CO was 8% lower and NO_x was 33% as compared with the tunnel measurements.

The on-board and tunnel comparison studies indicate a better agreement with remote sensing measurements than comparison with laboratory measurements, even at an individual vehicle level. This provides further confidence that use of independent RSD data to 'validate' VEPM predictions is a reasonable approach.

3.2.5 Uncertainties around low concentration values

Motor vehicles become progressively cleaner over time. Accurate detection of low emission levels poses challenges to any emission measurement device.

Smit and Bluett (2011) compared emission distributions of laboratory and remote sensing measurements, which show that a substantial proportion of THC (~ 20%) and NO (~ 20%) RSD measurements are recorded as zero or negative emission levels[19], whereas laboratory measurements show low, but not often zero, emission levels. Interestingly, this is not the case for CO where both laboratory and RSD measurements show a significant proportion of zero (or slightly negative) emission levels (~10-30%).

RSD has a lower sensitivity and more noise as compared with laboratory measurements, so accurate detection of low concentration levels could be a significant issue because the majority of modern vehicles typically exhibit low emission levels. It is also important for adequately quantifying vehicle emissions in the future, which will likely be dominated by vehicles with very low emission rates.

To investigate the potential impact of inaccurate measurement of low concentration values with RSD, modal (second-by-second) emissions for 8 medium ADR79/00 and ADR79/01 petrol cars (equivalent to Euro 2 and Euro 3) were collected. These vehicles were tested in a laboratory over the real-world Australian CUEDC-P drive cycle, which is depicted in Figure 3-8.

[19] True zero emission plumes will result in 50% negative values and 50% positive values in remote sensing measurements. Hence, negative RSD readings should be retained.

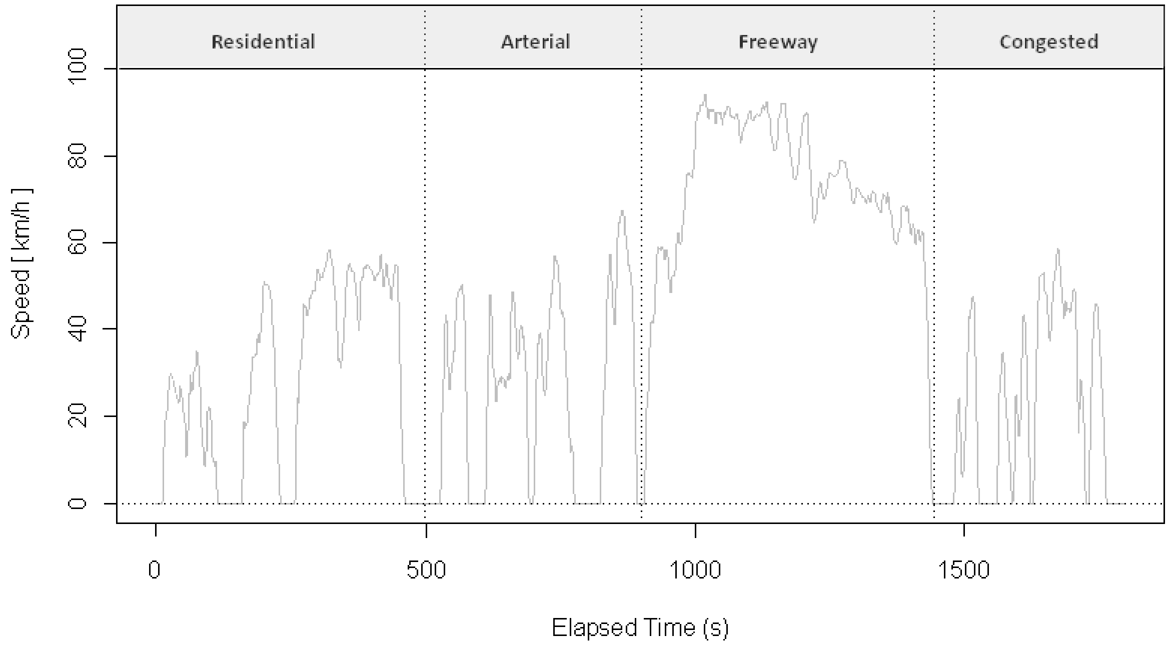

Figure 3-8: Australian CUEDC-P Drive cycle.

A database with approximately 23,500 second-by-second measured CO, HC, NO_x and CO_2 emission rates (g/s) were created for this purpose. The second-by-second pollutant-to-CO_2 ratios were also computed and added to the database. The ratios are highly skewed as is shown in Figure 3-9.

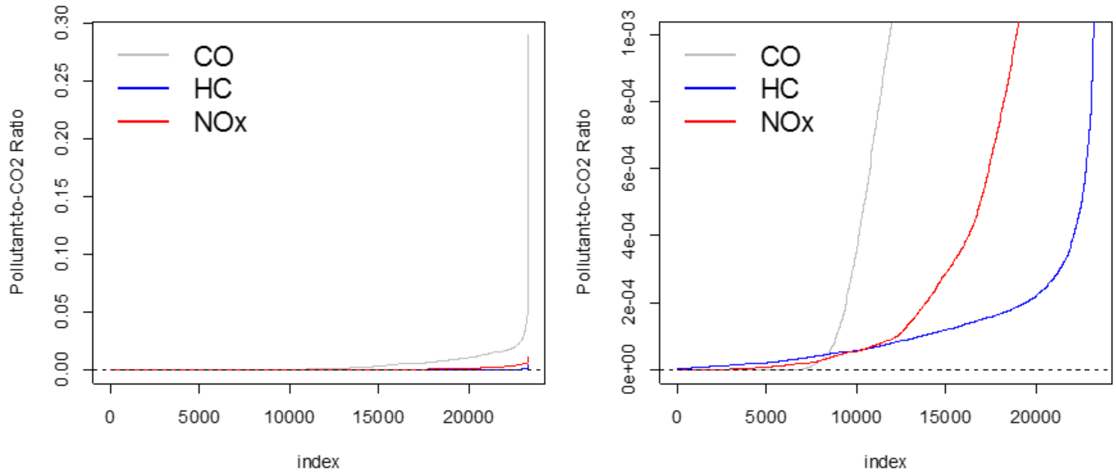

Figure 3-9: Modal Pollutant-to-CO2 Ratio Distributions for 8 Medium Petrol Cars (Euro 2/3).

To investigate the impact of inaccurate measurements at low concentration levels on mean pollutant-to-CO_2 ratios, different proportions of the modal laboratory emissions test data for CO, HC and NO_x were replaced with zero values, while CO_2 values were retained. All measured emission rates below pre-specified percentile values (10%, 20%, 30%, 40% and 50%) were set to zero and mean ratios were recomputed for each of the six VSP bins used by Smit and Bluett (2011). The percentage error ($e_{i,p}$), was computed as follows:

$$e_{i,p,x} = 100\,(R_{i,p,x} - R_{i,p,0}) \div R_{i,p,0} \qquad (1)$$

$R_{i,p,x}$ represents the modified mean pollutant-to-CO_2 ratio for VSP bin *i*, pollutant *p* and percentile value x below which all emission rates were artificially set to zero. $R_{i,p,0}$ represents the original laboratory data. The results are shown in Table 3-1.

Table 3-1: Computed errors in mean pollutant ratios as function of zero threshold values (CO/CO_2, HC/CO_2, NO_x/CO_2) by VSP bin.

	10% Zero	20% Zero	30% Zero	40% Zero	50% Zero
VSP < 0	0%, 0%, 0%	0%, -2%, 0%	0%, -5%, 0%	0%, -10%, -1%	-2%, -16%, -3%
0 ≤ VSP < 5	0%, -1%, 0%	0%, -2%, 0%	0%, -5%, -1%	0%, -8%, -1%	-2%, -14%, -3%
5 ≤ VSP < 10	0%, -1%, 0%	0%, -2%, 0%	0%, -4%, 0%	0%, -8%, -1%	-1%, -13%, -2%
10 ≤ VSP < 20	0%, 0%, 0%	0%, -1%, 0%	0%, -3%, 0%	0%, -5%, 0%	-1%, -9%, -1%
20 ≤ VSP < 40	0%, 0%, 0%	0%, 0%, 0%	0%, -1%, 0%	0%, -3%, 0%	-1%, -5%, -1%
VSP ≥ 40	0%, 0%, 0%	0%, 0%, 0%	0%, -1%, 0%	0%, -1%, 0%	0%, -2%, 0%

This simulation indicates that the impact of zero versus low emission measurements on the mean pollutant ratios is small and typically generates an (underestimation) error within 5%. Only after 40% or more of the “lower end” laboratory test data are replaced with zero values will the error in the mean value be higher than 5%, and then only for HC/CO_2 ratios.

The lower errors in mean CO and NO_x ratios are caused by their modal emission distributions (Figure 3-9). Measured CO and NO_x emission rates both have a significant proportion of zero or close to zero values in the laboratory emissions data. This is not the case for THC where sorted pollutant ratios increase more steadily from the start. As a consequence, the impact of replacing low pollutant ratios for THC with zero values, only has a significant impact on mean values for this pollutant.

3.2.6 Specific issues with NOx and HC

There are a number of incompatibility issues between laboratory and remote sensing techniques with NO and THC measurements in particular, which require further data preparation steps. These steps will be discussed further in Section 6.2. This will bring the RSD data and laboratory based emission factors on a more even footing in terms of definitions and differences in measurement methods and equipment.

NO_x definition

Most remote sensing studies have used RSDs that measure NO only, whereas laboratory measurements relate to the sum of NO and NO_2 (i.e. NO_x, expressed as NO_2 equivalents). Direct use of NO data from RSD would therefore result in an underestimation (bias) of RSD NO_x emission rates (e.g. Bishop et al., 2001; Ekström et al., 2004). Therefore, two steps are needed to properly define NO_x in the validation study:

1. Apply a correction factor to the RSD data to account for primary NO_2 emissions.
2. Express the RSD data as NO_2 equivalents.

With respect to the first point, Smit and Bluett (2011) used a scaling factor of 1.1, which assumes an average NO_2 to NO_x mass ratio of 0.10. However, there is a trend of increasing NO_2 to NO_x mass ratios, and this is particularly the case for diesel vehicles where NO_2 to NO_x mass ratios as high as 0.75 have been reported (Gense et al., 2006). For a validation study a plausible range of scaling factors should be used for different vehicle technology classes. An example is given below, reflecting RSD data collected by Sjödin and Jerksjö (2008):

- Petrol cars exhibit low NO_2/NO_x shares in the order of 2-5%, which are independent of emissions standard (pre-Euro to Euro 4).
- Diesel cars show increasing NO_2/NO_x shares as a function of emissions standard, i.e. 14% (Euro2), 47% (Euro 3) and 55% (Euro 4).
- Diesel trucks show increasing NO_2/NO_x shares as function of emissions standard, i.e. 25-30% (Euro II), 30-52% (Euro III), 48% (Euro IV), 38% (Euro V).
- CNG buses have a reported a value of 6%.

Similar data can be found elsewhere (e.g. EEA, 2007; Kousoulidou et al., 2008; Carslaw et al., 2011a). For the validation study it is suggested that a range of primary NO_2 proportions are computed for each site depending on the local fleet composition, which is accurately determined with the remote sensing equipment.

NO_x conversion

Another aspect that may affect RSD NO measurements is the extent to which NO is converted to NO_2 in the exhaust plume (and hence not detected). This would clearly depend on environmental factors such as ozone concentration, mixing characteristics and the speed of NO to NO_2 conversion. No useful quantitative information could be found to assess the significance of this.

HC emissions - FID versus NDIR

Remote sensing measures total HC emissions using NDIR, which is commonly calibrated with propane, although others such as hexane have been used (e.g. Ekström et al., 2004). However, a large and complex range of hydrocarbons is emitted in vehicle exhaust, and each of these specific hydrocarbons has different infrared absorption strengths within the infrared region used in remote sensing. So each specific hydrocarbon will incur a certain measurement error, with the exception of propane for which the instrument has been calibrated.

Stephens *et al*. (1996) conclude that "*the wide range of different response factors for different HC compounds represents a serious problem for NDIR-based measurements of vehicle exhaust. The NDIR approach currently has poor accuracy for quantitative determinations of exhaust HC concentration*". RSD determines around 50% of the HC mass compared to flame ionisation detection (FID) techniques used in laboratory measurements (Singer et al., 1998). Therefore a 'scaling factor' of about a factor of two is commonly applied to the RSD data (e.g. Pokharel et al., 2002; Kuhns et al., 2004).

However, given the sensitivity of NDIR to the mixture of hydrocarbons in the exhaust plume, the use of a single factor of 2 obscures this uncertainty. This is, to some extent, acknowledged by Bishop and Stedman (1996) who state that "*RSD results for individual vehicles cannot be expected to correlate perfectly with FID*". They also mention that these errors appear to average out for large fleets.

Therefore, for the validation study fleet level HC emission results should be used, as well as a plausible range of 'FID correction factors'. Stephens *et al*. (1996) conclude that an adjustment factor could range from 1.5 to 4.4, and this range could be adopted in the validation study. However, this range is large and it reflects the results for gas samples that do not resemble the complex HC mixture in vehicle exhaust.

Singer et al. (1998) developed a multivariate linear regression model to estimate NDIR/FID response factors for individual VOCs for remote sensors with a 3.40 µm filter. These researchers then applied this model to speciated fleet-averaged VOC emission measurements from various tunnels in the US and estimated an average scaling factor for vehicles using conventional petrol and reformulated petrol and found fleet averaged scaling factors varying between 1.9 and 2.3. For conventional (US) petrol a scaling factor of 2.2 is recommended. The scaling factor is sensitive to the filter used in the remote sensing equipment. A 3.45 µm filter reduced the fleet averaged scaling factors by about 15% to 1.7 and 2.0.

There are a few issues with using these factors directly in New Zealand:

- The measured VOC emission speciation reflects tunnel conditions with typically high speed free-flow traffic conditions, which can differ significantly from e.g. urban traffic conditions where remote sensing measurements are normally carried out.
- The VOC speciation reflects US fuels and the US on-road fleet, which will differ from the New Zealand fleet.
- The tunnel measurements were carried out in the 1990s and, therefore, reflect an old fleet with a significantly different average VOC profile, as compared with current fleets.

It is recommended that the availability of New Zealand VOC emission profiles for the on-road fleet is examined (tunnel, near-road air quality measurements). If this information is available, a New Zealand specific NDIR/FID scaling factor can be computed and used, provided that a 3.4 µm filter is used in the RSD used in New Zealand. Alternatively, VEPM can be used to create a fleet-averaged VOC profile and estimate a HC scaling factor for New Zealand conditions for use in the validation study. A less time-consuming approach can be to assume a feasible range in a sensitivity analysis as will be discussed later, or to simply consider that the error is small and negligible with respect to overall errors.

Hydrogen-to-carbon ratio

The hydrogen-to-carbon or H/C ratio in the exhaust gas varies with vehicle technology and fuel type. In laboratory tests, the conversion of ppmC FID readings to THC emission factors require an assumption with respect to the H/C ratio in the exhaust gas. Typically, $CH_{1.85}$ is assumed for both petrol and diesel. This H/C ratio reflects the H/C ratio of the fuel (e.g. EC, 1999), which is assumed to be equivalent to the H/C ratio in the exhaust gas.

As will be seen later, remote sensing uses a slightly different assumption of $CH_{2.00}$ for non-oxygenated petrol and diesel fuels (e.g. Bishop, 2011). However, this variation in hydrogen content does not have a large impact on the molecular weight of HC, and therefore the HC emission factor, due to small molar mass of hydrogen as compared with carbon.

Nevertheless, it is interesting to examine how accurate the $CH_{1.85}$ assumption is in relation to the actual H/C ratio in vehicle exhaust. Ye et al. (1997) report a detailed breakdown of the composition of unleaded (Australian) petrol. Using this detailed information a H/C ratio of 1.8 is computed. To further examine the H/C ratio in *exhaust*, laboratory emissions data from 21 light-duty Australian vehicles (DEWHA, 2008) operating on both petrol and E10 were examined. This study conducted high quality and comprehensive measurements for methane, C_2-C_{12} VOCs, carbonyls (mainly formaldehyde, acetaldehyde and acetone), alcohols (mainly ethanol and methanol) and 1,3-butadiene.

The emissions data were collated to create an average hydrocarbon profile for 91 organic compounds. These data were then used in this study to compute the H/C ratio for each vehicle category. The results are shown in Table 3-2.

Table 3-2: Computed H/C ratios in Exhaust for Australian Light-Duty Vehicles.

Petrol ADR37/01 (~ Euro 1)	Petrol ADR79/00 (Euro 2)	Petrol ADR79/01 (Euro 3)	E10 ADR37/01 (~ Euro 1)	E10 ADR79/00 (Euro 2)	E10 ADR79/01 (Euro 3)
2.03	2.23	2.08	2.05	2.27	2.08

It can be seen that there is some variation in the H/C ratio depending on the vehicle class and fuel. A typical H/C ratio appears to be 2.10, which means a 2% increase in molecular weight as compared with $CH_{1.85}$ and a similar increase in the corresponding VEPM THC emission factors. These changes are small and could be ignored from the perspective of expected overall accuracy. However, for consistency reasons, it is recommended that an accurate fleet average H/C ratio is determined for the New Zealand fleet and used to modify both VEPM and remote sensing results (by fuel type).

3.3 Different determination of emission factors

Remote sensing uses different calculations to determine emission factors as compared with standard laboratory testing on which VEPM is based, and these differences need to be considered when a validation study is conducted.

3.3.1 Remote sensing

RSD data are converted into emission factors in grams per litre or kg of fuel burned (EF^*, g/kg fuel) using a chemical mass (carbon) balance approach. Bishop (2011) presents the derivation of the FEAT[20] combustion equations for CO, HC and NO. In a more simplified form, the approach uses the following (unbalanced) chemical reaction equation:

$$CH_2 + air \rightarrow CO + 2\ C_3H_6 + CO_2 \qquad (2)$$

A few comments can be made here:

- Exhaust gas is assumed to have a similar composition as propane (C_3H_8), but a multiple of petrol of diesel fuel (CH_2) has been used to simplify the calculation, introducing a small discrepancy.
- The multiplication factor of 2 for C_3H_6 reflects the FID calibration factor.
- The equation represents stoichiometric combustion, and therefore does not properly reflect lean burn combustion conditions in diesel engines.
- Particulate matter may be a significant carbon sink for e.g. diesel vehicles and is not included in the FEAT equation, which introduces some error in the carbon mass balance for these vehicles.

[20] Fuel Efficiency Automobile Test. The system derived its name from the use of the measured CO/CO_2 ratio as an indicator for stoichiometric combustion and hence optimised fuel efficiency (Lawson et al., 1990).

Pollutant-to-CO_2 ratios can be used for both petrol and diesel engines. Therefore, combining information from Pokharel et al. (2000; 2002) and Bishop (2011), the following generic equation is used for conversion of RSD measurements to fuel-based emission factors (EF*, g/kg fuel):

$$EF^* = M_{pol}\left(\frac{\alpha\, r_{pol/co2}}{r_{co/co2} + 1 + 3\,\delta\, r_{hc/co2}}\right) \times \beta \tag{3}$$

Here,

- M_{pol} represents the molar mass of the pollutant (CO, HC, NO), which for CO, NO and HC is 28, 30 and 44 g/mol (as propane), respectively.
- The term within brackets is derived from the carbon balance in equation 1 and it computes the molar ratio of the selected pollutant and carbon, using the measured pollutant-to-CO_2 ratios (*r*) for CO, HC and CO_2. Note that it does not include all 'carbon sinks' such as particulate matter, introducing a small error, as was mentioned before. The significance of this error is a function of PM emission levels (mass basis), which are generally insignificant for petrol and LPG vehicles, but are elevated for (old) diesel HDVs.
- The constants α and δ both represent the NDIR/FID 'scaling factor'. Pokharel et al. (2002) used a value of 2.2, whereas Bishop (2011) used a value of 2.0. The difference is that α only has a value of 2.0 or 2.2 when EF* is computed for HC, but a value of 1 when computed for CO or NO. As δ is part of the carbon mass balance it is used for each pollutant. Note that this calibration factor depends on the calibration gas used to measure THC concentrations. For instance, Guo et al. (2007) used UV-DOAS (calibrated with 1,3-butadiene) and used a calibration factor of 3.6.
- The constant β represents the amount of carbon (mol C) per kg of fuel. It has a reported value of 71.4, which reflects an assumed H/C ratio of $CH_{2.00}$ for petrol and diesel. Bishop (2011) uses a slightly different factor of 71.7 mol C /kg fuel ($CH_{1.94}$), so there is some variation in RSD publications. Using a detailed breakdown of the composition of unleaded Australian petrol fuel (Ye et al., 1997) a value of 72.4 mol carbon per kg of fuel is computed. Using equation 3 and assuming a possible range of *fuel* H/C ratios from 1.70 to 2.30 (72.9 and 69.9 mol carbon per kg of fuel, respectively) shows that fuel-based emission factors decrease by 4% when moving from a H/C ratio of 1.70 to 2.30.
- Both petrol and diesel fuel have similar H/C ratios, so equation 3 can be used for both fuel types. However, other fuels like CNG require modified equations.

To obtain vehicle emission factors (EF) in units of grams per km, EF* needs to be multiplied with (mean) fuel consumption rate ($\bar{f}$, kg/km):

$$EF = EF^* \times \bar{f} \tag{4}$$

In conclusion, fuel-based emission factors (g/kg fuel) are estimated by making assumptions regarding the carbon content of the fuel and an appropriate 'NDIR to FID' scaling factor for HC. Distance-based emission factors (g/km) are then estimated by making assumptions about representative fuel consumption rates.

It has been acknowledged that, ideally, instantaneous fuel consumption rates at the time of measurement should be combined with the RSD measurements for each individual vehicle to reflect the proper spatial and temporal scale (e.g. CRC, 2007). In fact, a new hybrid model is discussed in section 5 that does this. In the absence of high resolution and vehicle-specific fuel consumption data, vehicle class averaged fuel consumption rates over longer (trip) distances can be used, but they will only provide an approximate estimate of fleet averaged emission factors, expressed as g/km. A simpler way has been used recently, which is to directly combine pollutant-to-CO_2 ratios with CO_2 emission predictions (e.g. Carslaw et al., 2011b; Kraan et al., 2012).

No correction is made for the effects of humidity on NO emissions. Very few remote sensing studies have actually measured humidity (e.g. Burgard et al., 2006), and even less studies have corrected remote sensing measurements for humidity (Bishop et al., 2001).

3.3.2 Laboratory

The following equation is used to convert laboratory (CVS) test results into vehicle emission factors (EF, g/km):

$$EF = \frac{V \times \delta_{pol} \times k_{no} \times C}{d} \tag{5}$$

Here V is the total volume (l/test) of diluted exhaust gas flow in standard conditions (273 K, 101.33 kPa), δ_{pol} is the pollutant density in standard conditions (g/l), k_{no} represents the humidity correction (for NO only), C is the background-corrected (mean) concentration (ppm) and d is the total distance driven over the test cycle (km). The pollutant density for HC is based on $CH_{1.85}$.

3.3.3 Conclusions

Laboratory emission test results are converted into distance-based emission factors (g/km). Remote sensing test results can be converted into the same units, but this requires assumptions regarding the carbon content of the fuel, an appropriate NDIR to FID scaling factor and representative fuel consumption rates.

Both laboratory and remote sensing computations are corrected for 'background concentration', although this is done in different ways. The methods effectively use a different definition of the hydrocarbon mixture in the exhaust gas, which is illustrated by the different values used for the H/C ratio. A humidity correction for NO is included in the conversion of laboratory test results to emission factors, but not in the RSD calculations. There is a clear difference in temporal and spatial resolution of the measurements (half a second versus a drive cycle test phase).

For a validation study it is recommended that:

1. The H/C ratio is harmonised; an accurate fleet average H/C ratio is determined for the New Zealand fleet and used to modify both VEPM and remote sensing results (by fuel type).
2. A range of α values are used in the conversion of RSD measurements to emission factors as part of a sensitivity analysis, as will be discussed in section 6.
3. A β value for NZ conditions is computed reflecting the local fuel composition at the time of RSD measurements.

4. A plausible range of mean vehicle class-specific fuel consumption rates (kg/km) is used as determined with VEPM.

5. A NO_x humidity correction factor is used to correct either the RSD or VEPM data.

6. The difference in temporal/spatial resolution is accounted for. This will be discussed in section 3.5.

7. A new approach is explored to combine remote sensing results with high resolution predictions of fuel consumption or CO_2 emissions (g/s). Refer to section 5.

3.4 Fleet composition: difference in level of detail

Vehicle design characteristics significantly impact on emissions and fuel consumption. A vehicle classification scheme is normally used in emission modelling to take differences in vehicle design characteristics into account. Given the large number of vehicle design characteristics, an almost infinite number of vehicle classes could be defined. However, the level of detail in vehicle classification should be comprehensive (i.e. include all important classes), as well as practical (i.e. level of detail that matches available data on fleet or traffic composition). In addition, a vehicle classification scheme should be up-to-date.

In practice, vehicle technology classes are usually defined in models by a taking into account a limited number of factors such as “main vehicle type” (e.g. passenger car, light-commercial vehicle, motorcycle, articulated trucks, buses, etc.), “fuel type” (e.g. diesel, petrol, LPG) and “emission standards” or “year of manufacture”. However, more detailed vehicle classification schemes are sometimes used. For instance, the VERSIT+ model also considers vehicle weight, fuel injection technology, emission reduction technology and type of transmission and the CMEM model also uses power-to-weight ratio, mileage, model year, after-treatment technology and emitter type (normal, high-emitter) as classification variables.

The choice for classification variables generally appears to reflect an established but arbitrary decision process, although in some cases it may actually involve statistical analysis to group vehicles according to their emission characteristics. In this respect, NIWA (2008a) confirmed that vehicle age, fuel type, speed, acceleration and measurement location were the most important predictor variables for a large remote sensing database of Australian vehicles. So it is clear that these variables need to be included in the validation study.

Emission models can be “incomplete” because they predict emissions for specific vehicle categories only (e.g. passenger cars), or because they are outdated (e.g. based on test data that do not reflect the latest developments in vehicle technology). Use of incomplete models introduces errors. It effectively restricts emission prediction to a specific part of a traffic stream, and additional models would be needed to estimate total traffic emissions. So, from a practical point of view, models with a comprehensive vehicle classification scheme are most useful for model users.

Remote sensing provides both detailed emissions and vehicle information for the majority of vehicles it measures. This is in contrast with VEPM where emission factors are available for specific vehicle classes only. As a consequence, the remote sensing data needs to be *aggregated* to develop emission factors that correspond to the VEPM vehicle classification, before a comparison can be made. RSD data provides vehicle emission factors for specific vehicle classes (e.g. small Euro 2 petrol passenger car). This facilitates a detailed comparison with VEPM predictions, and specific determination of which vehicle classes differ most.

But before this can be done the representativeness of the RSD sites needs to be examined and confirmed. In other words the combined set of RSD locations cannot *automatically* be assumed to be representative of the overall (VKT weighted) fleet composition (or even particular segments of it, excluding trucks with e.g. vertical exhausts) and the wide range of driving conditions encountered in the real-world. One or a few RSD sites will likely not be sufficient in this respect and this may introduce potentially significant bias in emission factors, were they directly derived from RSD results.

As a result, careful examination and a description of the RSD locations is required in the validation study to make qualified comments on aspects that impact on fleet composition such as demographic factors, which will influence vehicle age and make/model distributions, but also the proportion of vehicles with cold or warming up engines and emission control systems. Although it may be difficult to confirm that the RSD data are representative of the VKT-weighted average fleet composition, it is recommended that the RSD data be used to examine the spatial and temporal variation in both fleet composition and measured emission rates. This includes an examination of the occurrence of repeat RSD measurements of the same vehicle.

3.5 Driving conditions: difference in spatial and temporal resolution

It is not possible to directly compare VEPM emission factors (g/veh.km) that are derived from measurements over various drive cycles with remote sensing data collected at fixed roadside locations. The reason for this is that remote sensing techniques measure concentrations and (approximately) associated driving conditions at a particular point on the road (generally under slight acceleration) at a high frequency for about half a second, whereas drive cycles used in laboratory testing typically reflect 10 to 30 minutes of continuous driving over a range of driving conditions to simulate a journey.

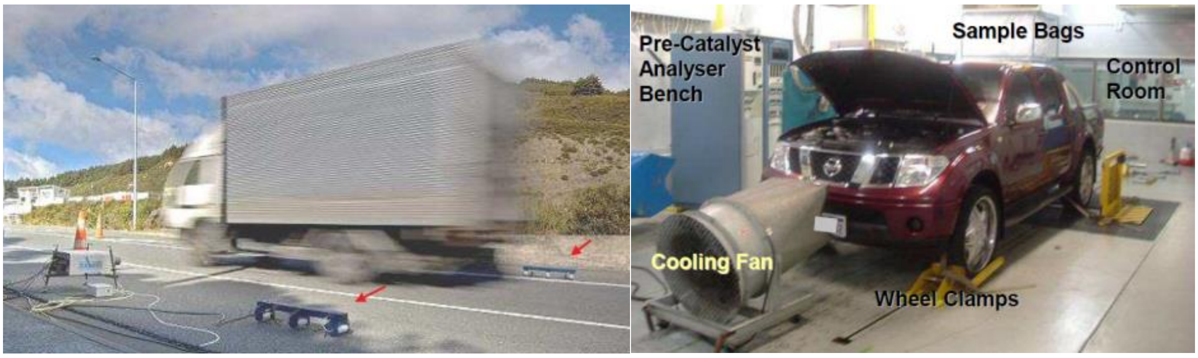

Figure 3-10: On-Road versus Laboratory measurements.

3.5.1 Remote sensing

Two alternative approaches to fixed site remote sensing measurement are worth noting here:

1. Wang et al. (2012) reported on the use of a mobile remote sensing (DOAS) technique to measure tropospheric vertical column densities (mole/cm^2) of NO_2. A vehicle was equipped with remote sensing and GPS equipment and driven over a predefined route through Shanghai, China. The measurements were then used to estimate NO_2 and NO_x emission fluxes (ton/h), which were compared with emission inventory predictions. This is an innovative approach using remote sensing, which overcomes the spatial limitations of the technique.

2. A number of international studies[21] have reported on the development and application of a 'moving platform plume-chasing' approach to determine fuel-based emission factors for selected target vehicles. These methods do not use remote sensing, but sample air from a sampling duct installed between e.g. two passenger windows. However, similar mobile systems may be developed using remote sensing in the future.

Nevertheless, for the situation in New Zealand, VEPM reflects drive cycle averaged measurements and the roadside measurements provide a snapshot. A proper comparison of remote sensing and VEPM predictions requires a method to account for the *difference in spatial and temporal resolution*. To make the two data sets more comparable, variables are required that quantify 'driving conditions' and link the two independent databases.

Candidate variables are 'vehicle-specific power' (VSP) or vehicle speed and acceleration.

- VSP is commonly used as an explanatory variable in emission predictions, particularly in the USA. Similar but more detailed power-based algorithms are used around the world to simulate vehicle emissions at a high spatial and temporal resolution, e.g. in the US (Barth et al. 2000), Europe (Hausberger et al., 2003) and Australia (Smit, 2014). VSP (kW) is estimated from a second-by-second speed profile with the formula:

$$\mathrm{VSP} = M\,v\,\left(1.1a + 9.81\sin(\arctan(g)) + 0.132\right) + 0.000302\,M\,v^3 \tag{6}$$

  Here, M is vehicle mass (tonne), v is the instantaneous speed (m/s), *a* is the instantaneous acceleration or deceleration (m/s^2) and *g* is the terrain gradient (%). These input data are collected at RSD sites. It is noted that vehicle mass can be *estimated* using vehicle registration information such as gross vehicle mass (GVM) or tare mass. However, often this is not done and VSP is expressed as kW/tonne.

- Speed-acceleration (v-a) bins can also be used as classification variables for driving conditions. In fact, the first generation modal (second-by-second) vehicle emission models used v-a to simulate the impact of driving conditions on emission factors (e.g. Kent et al., 1982; St. Denis and Winer, 1994; Joumard et al., 1995). The combination of speed and acceleration may provide a better classification of driving conditions and a better distinction in pollutant emission ratios.

Although VSP has been extensively used in remote sensing studies to quantify driving conditions and validate RSD measurements, some studies have suggested that VSP is not a significant factor with respect to measured emission levels (e.g. Kuhns et al., 2004; Sjödin and Jerksjö, 2008), which contradicts the findings reported in other studies (e.g. Smit and Bluett, 2011).

To examine how well VSP or v-a classify vehicle emissions, modal (second-by-second) laboratory emission data for 8 medium ADR79/00 and ADR79/01 petrol cars (equivalent to Euro 2 and Euro 3) were collected (refer to section 3.2.5). An empirical emissions database of about 23,500 seconds of data was created with the following variables: speed, acceleration, VSP and the ratios of CO/CO_2, HC/CO_2 and NO_x/CO_2.

The VSP values were then classified into the six VSP bins (A-F) used in Smit and Bluett (2011). These bins and their boundaries were selected to achieve a logical distinction in driving behaviour. They are

[21] For instance, the On-road Plume Chasing and Analysis System (OPCAS) in Hong Kong (Ning et al., 2012) and 'zero emission (vehicle) Mobile Measurement Platform' (MMP, Park et al., 2011) and similar work (Kittelson, 2006) in the USA.

defined as VSP < 0 kW, 0 ≤ VSP < 5 kW, 5 ≤ VSP < 10 kW, 10 ≤ VSP < 20 kW, 20 ≤ VSP < 40 kW, VSP ≥ 40 kW. In addition, 11 speed (km/h) and 16 acceleration (m/s^2) bins were defined.

Both VSP class and speed-acceleration bin, were then computed for each second of emissions and driving data. Figure 3-11 shows a summary table for the 23,500 seconds of binned data. It shows for which v-a bins driving conditions and emissions were measured, i.e. blank cells mean 'no data'. In addition, the 'dominant' VSP class is shown for each v-a bin, meaning that the majority of measurements within a particular v-a bin are allocated to the designated VSP class. It can be seen that the six VSP classes tend to cluster in certain areas of the v-a matrix, as would be expected as VSP is computed using instantaneous acceleration and speed – refer to Equation (6). However, this is not a perfect fit, as different VSP bins can occur within specific v-a cells.

Instantaneous speed (km/h)	Acceleration (m/s^2)															
	≤ -3.0	-3.0 to -2.5	-2.5 to -2.0	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to -0.1	-0.1 to +0.0	+0.0 to +0.1	+0.1 to +0.5	+0.5 to +1.0	+1.0 to +1.5	+1.5 to +2.0	+2.0 to +2.5	+2.5 to +3.0	≥ +3
≥ 100																
90-100							C	D	D	E						
80-90					A	A	A	D	D	E	E	F				
70-80						A	B	C	D	D	E					
60-70						A	A	C	C	D	E					
50-60					A	A	A	B	C	D	E	E				
40-50					A	A	A	B	C	D	D	E	F			
30-40				A	A	A	A	B	B	C	D	E	E			
20-30				A	A	A	A	B	B	C	D	D	E	E		
10-20				A	A	A	A	B	B	B	C	D	D	D		
0-10				A	A	A	A	B	B	B	B	B	C			

Figure 3-11: Speed-acceleration matrix and designated dominant VSP classification.

The average pollutant-to-CO_2 ratios were computed for each bin, i.e. 6 values for the VSP method and 176 values for the speed-acceleration method. For each speed-acceleration bin, the mean percent difference in pollutant-to-CO_2 ratios between the speed-acceleration and VSP method was then computed using 1 Hz data, i.e.

$$d_{i,p} = 100 \sum_{1}^{n} \left(R_{i,p,vsp,x} - R_{i,p,va,x} / R_{i,p,va,x} \right) \div n \tag{7}$$

Here $d_{i,p}$ represents the computed mean percent difference for v-a bin i and pollutant p, x denotes the x^{th} observation for this specific v-a bin i (1, …, n), $R_{i,p,va,x}$ represents the x^{th} pollutant-to-CO_2 ratio for v-a bin i and pollutant p and $R_{i,p,va,x}$ represents the same measurement for the VSP method. The detailed results are shown in Appendix B for each pollutant. The computed mean difference between the two methods is typically within ± 50%, but larger differences do occur. The mean absolute difference (MAD) over all v-a bins is 29%, 21% and 69% for CO, HC and NO_x respectively.

It is noted that these results represent driving in a wide range of operating conditions, some of which may not typically be observed in remote sensing. For instance about 20% of the (laboratory) data points represent idling conditions, which is not measured by remote sensing. Large errors for

NO_x/CO_2 ratios were observed in these low speed operating conditions, which explains the large MAD value for this pollutant.[22]

Figure 3-12 presents the empirical cumulative distributions for both classification methods. The plots show the probability that the pollutant-to-CO_2 ratio takes a value of less than or equal to 'x'. The x-axis presents the allowable domain for the probability function. Since the y-axis represents cumulative probability, it must fall between zero and one. It increases from zero to one as we go from left to right on the horizontal axis.

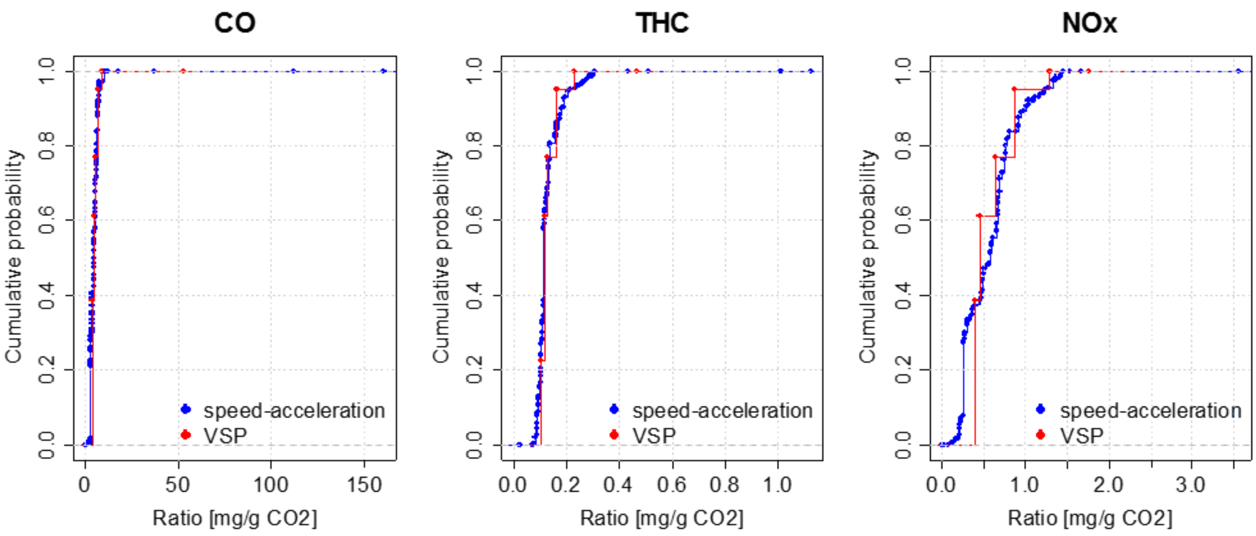

Figure 3-12: Cumulative probability distributions for v-a and VSP 6 bin classification.

These results show the strength of the VSP approach. The VSP approach uses only 6 bins, whereas the v-a classification uses 176 bins. So despite the small number of bins, VSP is able to account for a large portion of the variation in the emissions data.

Frey et al. (2003) determined the best explanatory variable for an on-board and laboratory emission database using a technique called Hierarchical Tree-Based Regression. A range of variables were included in this analysis such as VSP, speed, acceleration, engine capacity and model year. It was found that VSP was consistently identified as the most important variable for each pollutant and 14 VSP bins were selected as the best "driving mode" definition. The number of bins used by Frey et al. (2003) is higher compared with the 6 bins used by Smit and Bluett (2011), which suggests that a larger number of VSP bins could be considered for the validation study. To test this, the analysis was repeated but now with the fourteen bin VSP classification method. Figure 3-13 visualizes the results.

[22] In reality, driving behaviour does not have a uniform distribution over the v-a bins and would be concentrated in specific regions of the v-a matrix. So weighted mean differences and MAD values can be computed, but this is beyond the scope of this project.

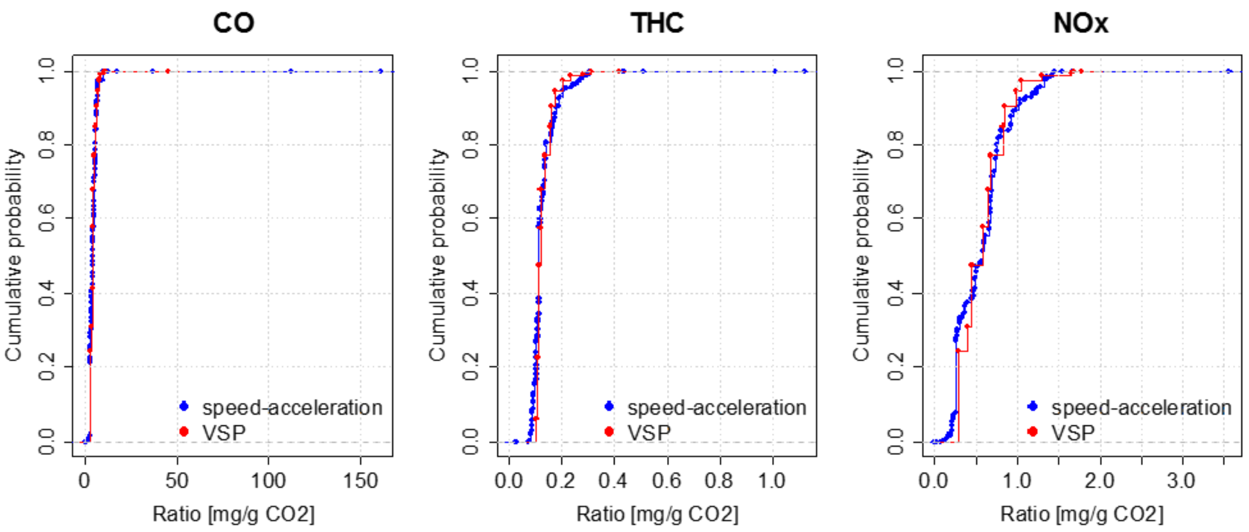

Figure 3-13: Cumulative probability distributions for v-a and VSP 14 bin classification.

To determine the statistical significance of the differences between the methods it is necessary to consider both the variability between the groups (i.e. "driving behaviour classification methods"), which have been analysed so far, as well as the variability in the emissions data. The F-test can be used to test the null hypothesis that the three classification methods (VSP6, VSP14, v-a) are not significantly different from each other[23]. Therefore the mean squared errors (MSEs) were computed for each pollutant and classification method, and the F-statistics were computed as the ratio of the MSE values for the two VSP methods 1 to the MSE value for the most detailed v-a method. The results are shown in Table 3-3.

Table 3-3: MSE and F-Test p-values for Three Classification Methods.

	MSE			p-value F test	
Variable	VSP 6 Bin	VSP 14 Bin	v-a	VSP6/v-a	VSP14/v-a
CO/CO_2 ratio (mg/g)	73.98	74.02	72.29	0.41	0.44
HC/CO_2 ratio (mg/g)	0.042	0.042	0.041	0.41	0.43
NOx/CO_2 ratio (mg/g)	1.13	1.12	1.09	0.40	0.42

It is clear that when the variability in the emissions data is considered, the three methods are not significantly different ($p > 0.05$). This is evident from the similar MSE values for all pollutants considered. This means that the simplest method (VSP6) could be used. The most detailed v-a approach does not provide significant benefits with respect to classification of driving behaviour when the variability in vehicle emissions is considered. There are two other issues with the speed-acceleration approach:

1. It does not explicitly include vehicle weight and road gradient in the classification, whereas VSP does. Both vehicle weight and road grade are known to significantly affect vehicle emission levels. So a speed-acceleration approach requires that measurement sites with

[23] The F-statistic is the ratio of the between-group variability (e.g. according to the v-a classification) to the within-group variability (reflecting the emissions data). The F statistic will be large if the between-group variability is large relative to the within-group variability. In other words it represents a good fit to the data.

similar road grade and vehicle weight distributions are grouped before the emissions analysis is carried out. A benefit of the VSP approach is that the data from multiple sites can be pooled together for analysis.

2. There is a trade-off between the number of bins and the complexity and usefulness of the emissions analysis. The number of measurements can become too small for particular bins, affecting the accuracy of the mean emission factors for these bins.

A 6 bin or 14 bin VSP classification method is therefore recommended for the validation study. As will be discussed later, there are different ways to take driving conditions into account in the validation study. This includes analysis of VSP distributions in drive cycles and comparison with remote sensing VSP data.

Although VSP is the best approach to quantify driving conditions at RSD locations, there are still a number of unresolved issues, which are discussed further below. These limitations cannot be addressed, but they add uncertainty to the quantification of driving conditions at the time of measurement.

Exclusion of relevant aspects

VSP does not include all relevant aspects of driving conditions/driving behaviour:

- It does not account for additional loads due to e.g. road curvature.
- It does not account for the impacts of gear shift behaviour.[24]
- It does not account for 'driving history' such as changes in engine power seconds *before* the measurement took place.

Time alignment errors

Speed and acceleration are measured at a specific distance before the remote sensing measurements take place. Assuming a vehicle speed of 60 km/h and a typical distance between the speed and remote sensing measurement of 1 to 5 meters, the time gap between emission and speed measurements is about 0.1-0.3 seconds.

However, engine-out emissions take up to a few seconds to travel from the engine to the exhaust pipe. For instance, transport time for a diesel and petrol car varies between 1.3 or 6.6 seconds for low engine load (idling) conditions for a diesel and petrol car, respectively, to 0.1 seconds for high engine load conditions for both cars (Weilenmann, Soltic and Atjay, 2003). As a consequence exhaust transport time is typically larger than the estimated time gap between emission and speed measurements. This means that RSD measures (engine-out) emissions for driving behaviour that range from a few meters *after* the speed measurement up to say 50 m *before* the speed measurement (assuming a 3 second delay). These results will be different for other vehicle speeds and vehicle specifications (including engine size and exhaust system volume), but it is clear that there will be time alignment errors between VSP and remote sensing measurements. These errors will be less relevant for sites with homogeneous driving conditions (e.g. constant acceleration).

[24] Gear shift behaviour affects instantaneous vehicle emissions. This is particular the case for (heavily loaded) HDVs, which undergo many gear shift changes with rapid engine transitions from effectively 'no load' to 'maximum load'.

Errors in measured acceleration

The time intervals for computation of speed and acceleration are quite short. For the New Zealand remote sensing data a one meter long speed bar is used where the lasers are placed about 0.5 m apart (Figure 3-14).

4). It is noted that this is a different system than what has been reported in other studies (e.g. Jiménez et al., 1999; Ekström et al., 2004; CRC, 2007; Carslaw et al., 2013). For instance, Sjödin et al. (1997) used two pairs of pneumatic tubes with a 10 m distance between the pairs and a 5 m distance between the tubes.

Assuming a vehicle speed of 60 km/h, the time interval between the speed measurements is about 0.3 seconds (car) to 0.7 seconds (bus). This is significantly shorter than time intervals used to compute instantaneous acceleration from dynamometer data to get robust and realistic values, e.g. 1-3 seconds (e.g. Smit, 2014). So it is possible that these short time intervals produce noisy and unrealistic acceleration values.

Figure 3-14: Speed bar (red circle) used for NZ remote sensing measurements.

This seems to be the case for several RSD studies. For instance, CRC (2003) reports that accelerations within the range of -13 to +14 mph/s (-5.8 to +6.3 m/s^2) measured with the FEAT RSD are considered to be valid accelerations. Chen and Borken-Kleefeld (2014) also show very high acceleration levels, as is shown in Figure 3-15. Elder et al. (2011, Fig. 2) also reported high acceleration rates measured with RSD in New Zealand of up to almost 5 m/s^2.

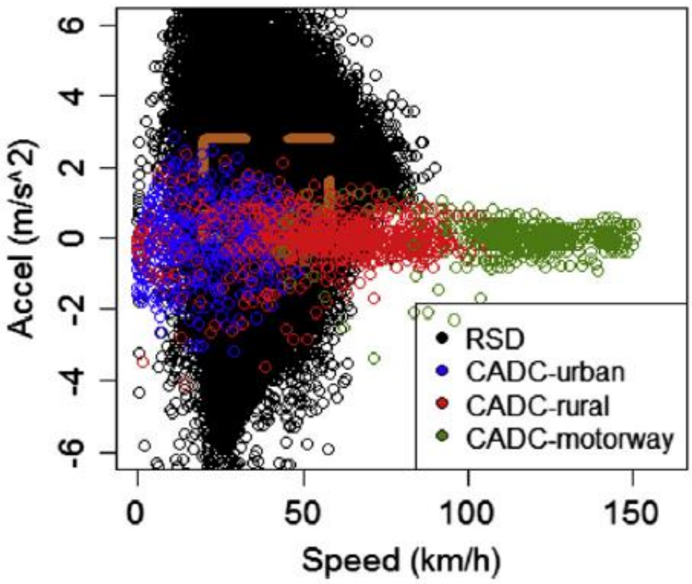

Figure 3-15: Speed-acceleration data measured with RSD and comparison with CADC drive cycles. (Chen and Borken Kleefeld, 2014).

Typical maximum real-world accelerations for LDVs are expected to be about 3 m/s^2 at low speeds, with increasingly lower maximum values at higher speeds. This is in line with some remote sensing studies. For instance, Guo et al. (2007) report measured accelerations ranging between -2.2 and +2.1 m/s^2. Interestingly, these workers note in their paper that they used three lasers for speed/acceleration measurements, rather than the conventional two, for greater accuracy. It is unclear what is causing these high acceleration values as the accuracy of acceleration measurements has been reported to be better than 0.25 m/s^2 (e.g. Sjödin et al.,1997). It is recommended that the speed-acceleration distributions for New Zealand are further examined.

3.5.2 Laboratory

VEPM is based on emission testing studies using chassis or engine dynamometers. Although these tests generally use high quality equipment and are conducted under highly controlled conditions, there are still several challenges to produce emission results that are representative for real-world driving conditions.

It is clear that 'real-world' drive cycles must be used for emission factor development, rather than standard drive cycles such as the Eurotest and US FTP cycles (used in the Australian Design Rules or ADRs), which substantially underestimate emissions due to relatively low speed and acceleration levels. VEPM reflects emission measurements over real-world drive cycles. It is, however, not possible to determine if these cycles adequately and accurately represent New Zealand driving behaviour, given the large variety of European drive cycles that have been used in the development of COPERT for different vehicle classes (e.g. Smit, Brown and Chan, 2008).

But even if 'real-world' cycles are used in laboratory testing, there are remaining issues. As drive cycles are of limited duration to keep costs within acceptable limits, specific driving conditions may be excluded.[25] Real-world factors can also be excluded, such as road gradient effects, air conditioning

[25] An example is the Australian real-world CUEDC-P, which has instantaneous speeds up to 94 km/h, whereas freeway driving occurs at higher speeds.

use and 'real world' variation in driving styles (including gear shift behaviour). The quality of the dynamometer and dynamometer settings also impact on the emission results. Dynamometer system configuration[26] and specifications[27] vary and affect how well on-road driving conditions are replicated in the laboratory. In addition, the extent to which vehicle specific parameters are taken into account is important.[28]

Given that VEPM is based on overseas models, which in turn are based on a variety of overseas vehicle test programs and drive cycles, this information cannot readily be retrieved and verified.

3.6 Meteorological conditions: difference in test conditions

Meteorological variables such as ambient temperature, air density and humidity have a direct impact on combustion processes and thus vehicle emissions. In addition, these variables affect air-conditioning use, which affects emissions because the air-conditioning compressor requires a significant amount of energy.

- Ambient temperature will fluctuate substantially in time. Vehicle emissions are a function of ambient temperature, particularly CO and THC emissions, and especially during cold starts. In general, exhaust emissions increase gradually with decreasing temperature from about 25 °C. Below about 10 °C the emissions increase dramatically in a non-linear fashion. Laboratory tests are typically conducted at temperatures between 20-30 ºC. RSD measurements reflect actual ambient temperatures. As a consequence, temperature correction needs to be applied to either VEPM predictions or remote sensing data. It is noted that VEPM uses temperature correction factors as an optional input (-10 to 30^{o}C).
- Laboratory measurements are normally corrected for humidity to standard test conditions (i.e. 10.71 g H_2O/kg dry air, corresponding to 61% RH at 23 ºC and 101.3 kPa pressure). NO_x emissions are a function of humidity where an increase in ambient humidity lowers NO_x emissions, and vice versa. The differences can be quite substantial (e.g. 70-160%), so this could cause a significant difference with RSD measurements. For the validation study, the RSD results need to be normalised for humidity.
- Air density affects engine operation and emissions in older vehicles, but not significantly in modern vehicles where electronic fuel injection systems automatically compensate for changes in air density. The issue of air density is an issue for RSD measurements at high altitudes, so this variable is not expected to be relevant for the validation study and can be excluded.
- The use of air-conditioning has the potential to substantially increase vehicle emissions as it increases the load placed on the engine. Increases in emissions of light-duty vehicles by more than one order of magnitude have been reported in the past, but lower values have also been reported. Air-conditioning is normally not incorporated in laboratory vehicle emission testing. VEPM has no air-conditioning correction factors, but air-conditioning impacts will be reflected in the RSD results. So a correction of either the VEPM or RSD results is recommended. However, the feasibility of this is unclear. Simulation of the effects of air-conditioning on traffic emissions requires not only consideration of emission effects, but also

[26] e.g. hydraulic or electrical power absorption unit.
[27] e.g. base inertia, response time, power absorption capability, motoring capabilities, permissible axle loading.
[28] For instance, dynamometer settings can be based on coast-down test results, general settings as specified by legislation or based on empirical formulae.

air conditioning use (which is a function of local climate and time of year) and penetration of air conditioners in the on-road fleet.

Assuming that meteorological information is available for the RSD measurements, their impacts on either VEPM predictions or the RSD measurements need to be quantified and included in the validation study.

3.7 Fuel quality: difference in test conditions

The composition and quality of fuel has a significant impact on emissions. Commercial fuels change over time. Examples are the phasing out of lead from petrol and the ongoing reduction of sulphur content in fuels. VEPM is to a large extent based on overseas models, and therefore reflects the quality and composition of the fuel used in overseas testing programmes, at the time of measurement. The RSD data will reflect the local NZ fuel quality and composition, and will therefore be more accurate in this respect. VEPM emission predictions need to be corrected for commercial NZ fuels, to the extent that this is possible, if this has not already been done in the development of VEPM.

3.8 Ageing effects: difference in test conditions

Mileage has a significant and unavoidable effect on vehicle emissions. Even with reasonable maintenance, vehicle emissions increase with mileage and age due to deterioration of engine and emission-control components. Emission deterioration is intrinsically related to the durability of emission control devices (including OBD), in-service I/M programs, driving conditions and driving behaviour.

The remote sensing data reflect snapshots in time. Fleet-wide ageing effects can be captured by relating RSD emissions to vehicle model year and by conducting remote sensing campaigns in subsequent years and, ideally, multiple sites (e.g. Sjödin and Andréasson, 2000; Lau et al., 2012; NZTA, 2014).

In contrast, the laboratory based VEPM predictions are ‘fixed’ and reflect the ageing effects at the time of measurements only. Ageing effects are normally simulated in emission models by generic mileage correction algorithms, which are based on limited data. In fact, the accuracy of those algorithms in VEPM requires some attention. For instance, Borken-Kleefeld and Chen (2015) compared deterioration rates for petrol cars, derived from 13 years’ of remote sensing data at a single site in Switzerland, with those used in COPERT and HBEFA. They found substantial differences in deterioration rates and, importantly, estimated a relative impact on emission predictions of up to a factor of 3 for specific vehicle classes. It is therefore recommended that deterioration rates derived from New Zealand remote sensing data are also compared with those used in VEPM. This should include an analysis of consistency between the results for multiple RSD sites (e.g. NZTA, 2014).

3.9 Vehicle loading: difference in test conditions

Vehicle load can have a significant impact on vehicle emissions. For instance, towing a caravan or trailer can greatly increase engine load and therefore emissions. Like engine operating conditions (e.g. ‘cold start’), actual vehicle loading is not measured with the RSD, although a few studies reported the use of a weighing station before remote sensing measurements were carried out (Bishop et al., 2001; Burgard et al., 2006).

There appears to be no practical way to include ‘vehicle loading’ into the validation study due to a lack of information in this respect. For instance, there are no data on vehicle load distributions in on-road fleets. This information is useful input for the development of more accurate emission inventories. The lack of vehicle weight information appears to be less of an issue for the actual remote sensing measurements. This is because the pollutant-to-CO_2 ratios are less sensitive to vehicle loading (e.g. Mazzoleni et al., 2004b). For instance, Burgard et al. (2006) did not find a clear trend of CO and NO_x fuel-based emission factors with truck weight.

Vehicle loading is recorded during laboratory testing, but the extent to which dynamometer loading corresponds to real-world driving is another factor to consider here, as was discussed in section 3.5.2. This comes (in part) down to the quality of the dynamometer and dynamometer settings. Given that VEPM is based on overseas models, which in turn are based on a variety of overseas vehicle test programs, this information cannot readily be retrieved and verified.

4 Intended use of VEPM

Another source of error is application of models beyond their intended use or design capabilities. This has relevance for the validation study because, as we have seen, RSD measurements are highly localised, whereas, as will be discussed in this section, VEPM's intended use is for emission estimation at road network level.

This difference cannot be resolved, but it can (to some extent) be accounted for in the validation study through the use of more robust and stable comparison statistics (fuel-based emission factors) and explicit consideration of drive cycles and driving behaviour at RSD sites.

Vehicle emission models can be applied at different scales, depending on the modelling objectives. Three main modelling objectives may be distinguished:

1. verify compliance with air quality standards;
5. development of emission inventories; and
6. evaluation of transport policies (scenario testing).

Assessment may be conducted at the local, regional, national or even global scale. For instance, it may involve local air quality impacts due to new road projects or the implementation of (new) traffic management measures (e.g. lower speed limits, traffic signal coordination, metering signals). Assessment of regional air quality are commonly based on emission inventories and may involve modelling of key pollutants such as NO_x and PM_{10}, but is often directed towards the analysis and prediction of photochemical smog levels. National emission inventories are needed to verify compliance with international agreements (e.g. national emission ceilings, greenhouse gas emissions).

Models are often used beyond their intended purposes and capabilities, resulting in errors. This may partly be explained by lack of understanding of the model development process, the underlying empirical data and the exact definition of variables used by the model, and clear documentation of these matters.

For instance, "speed" can be defined in several ways (e.g. space mean speed, time mean speed, running speed, travel speed, instantaneous speed), but the correct speed definition for average speed emission models like VEPM would correspond to "travel speed", which is defined as: the overall speed between two points of "sufficient" distance, i.e. distance corresponding to drive cycles used in emission model development, including all delays, for either an individual vehicle or a traffic stream. So the length of the drive cycles used for emissions testing and subsequent model development is an important aspect for consideration. Figure 4-1 shows that drive cycles on which well-known average speed models were based, have a cycle length that varies between 130 m to 98 km, and a median value of 6 km. Shorter drive cycles typically occur at lower more congested mean speeds, whereas longer cycles tend to occur at higher mean speeds.

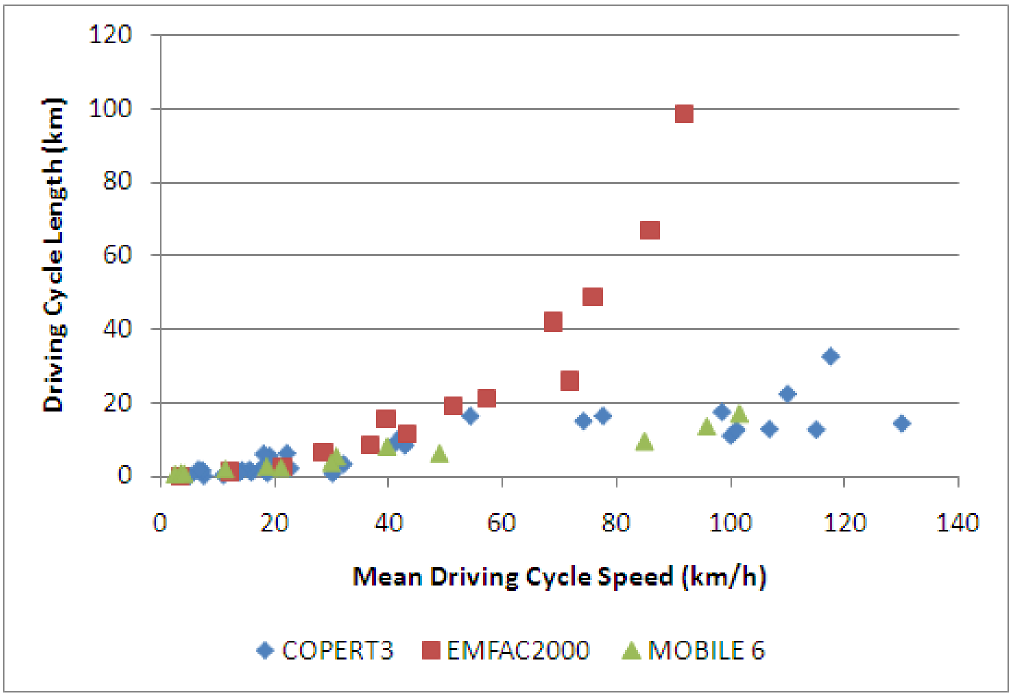

Figure 4-1: Drive cycle Distance versus Average Speed. (Source: Smit et al., 2009).

This has implications for the appropriate spatial resolution at which average speed models such as VEPM should be applied. For instance, a drive cycle with an average speed of 70 km/h that represents a journey through a road network may involve driving on residential, arterial and freeways and would be expected to have significantly different driving characteristics and emissions as compared with a driving pattern with the same average speed for a 100 m stretch of arterial road with free-flowing driving conditions.

As a consequence, the correct spatial application of average speed models such as COPERT (and therefore VEPM) is likely to be at large or small network level. Indeed, this was the reason to explicitly develop emission factors for COPERT Australia for 100 m driving segments (Smit and Ntziachristos, 2012). The COPERT (Europe) model documentation suggests, for example:

COPERT III:

> *"COPERT III has mainly been developed for the compilation of national annual emission inventories."*
>
> *"The proposed methodology can be used with a sufficient degree of certainty at a higher resolution too, i.e. for the compilation of urban emission inventories with a spatial resolution of 1x1 km^2 and a temporal resolution of 1 hour."*
>
> *"The application of the methodology can be applied at maximum spatial resolution on a street network (e.g. highways) or on a whole urban area (e.g. small city)."*

COPERT IV:

> *"The proposed formulas should only be used with average travelling speed and by no means can be considered an accurate approach when only instant speed values are available."*
>
> *"Emission factors can be considered representative of emission performance with constant speed only at high velocities (>100 km/h) when, in general, speed fluctuation is relatively low."*
>
> *"The emission factors should not be applied in cases in which the driving pattern differs too much from what is common, e.g. in traffic calming areas."*
>
> *"The emission factors proposed are aggregated emission factors, averaged over a large number of drive cycles, therefore not necessarily representative of the instantaneous emissions of vehicles driven under actual conditions."*

The use of 'average speed' as the only variable to capture the impact of driving conditions on emission levels can introduce significant errors in the emission predictions with the same mean speed, but with different levels of speed fluctuation (i.e. different drive cycles or driving patterns). Errors in emission predictions of up to a factor of four have been reported for speed-time traces with the same mean speed but with different levels of speed fluctuation.

This is illustrated in Figure 4-2. It shows the predictions of NO_x emission factors by both COPERT IV (solid line) and a more complex model (VERSIT+, grey dots) for a variety of driving conditions (drive cycles). The more complex model uses multiple variables to predict emissions (e.g. average speed, number of stops, level of speed fluctuation, etc.)

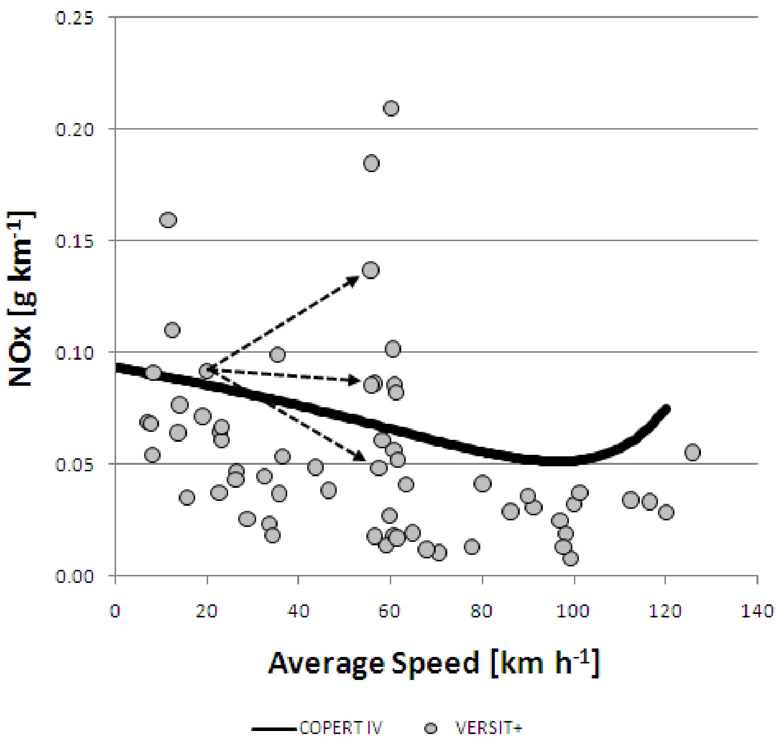

Figure 4-2: NOx Emission Factors as Predicted by COPERT IV and a more Complex Model. (Source: Smit, 2008b)

It can be seen that for traffic situations with different dynamics but similar average speeds the more complex model predicts different emission factors, whereas COPERT IV predicts the same emission factors for all these situations. Suppose now that the implementation of a particular local traffic management measure (e.g. improved signal timing) has smoothed the flow of traffic (i.e. reduced dynamics) and has increased the average speed from 20 to about 55 km/h. For this particular situation, COPERT IV would always predict a decrease in emissions. In contrast, the more complex model could predict either a decrease or an increase in emissions, or even no effect, depending on input driving pattern data in both the reference and the new traffic situation. This effect is shown by the dashed arrows in Figure 4-2.

Clearly, for policy makers and transport planners the correct direction and magnitude of these effects is necessary information for imposing effective and cost-effective measures in order to improve on local air quality and reduce greenhouse gas emissions. More detailed models in addition to VEPM will be needed in the future for adequate emission predictions at the local level. Section 5 will discuss a possible approach to develop such as model, using remote sensing data.

5 A new hybrid emission model based on RSD

Remote sensing has been extensively used around the world, including the US (Jiménez et al., 2000; Bishop and Stedman, 2008), Europe (Carslaw et al., 2013; Chen and Borken-Kleefeld, 2014), Asia (Chan and Ning, 2005; Lau et al., 2012; Yam, 2012), Australia (NIWA, 2008b; Smit and Bluett, 2011) and New Zealand (Bluett and Fisher, 2004; 2005) for different purposes, but rarely for actual emission factor development. There are a few exceptions:

- Chan and Ning (2005) fitted linear polynomial regression equations to remote sensing data using the pollutant-to-CO_2 ratios as the response variables and instantaneous speed and acceleration as the predictor variables. These ratio predictions were then used in combination with diesel fuel density, carbon content of diesel fuel, molar mass of the fuel to convert the ratio to fuel-based emission factors (g/l) for Hong Kong. A mathematical relationship between instantaneous fuel use and speed (l/km) was sourced from another study and combined with the remote sensing fuel based emission factors for CO, HC and NO to compute composite emission factors (g/km) for diesel vehicles.
- Lau et al. (2012) combined RSD 'emission indices' (g pollutant per litre of fuel) directly with fuel consumption regression equations (litre of fuel per km) using speed and acceleration that were developed from dynamometer (one petrol car) and on-board emission measurements (diesel and LPG vehicles) in Hong Kong. Although the paper is not clear on this, it appears that the regression equations provide second-by-second estimates of fuel consumption.

These studies used their models to compare emissions at specific instantaneous speeds, but did not use them in combination with drive cycles to develop emission factors for specific traffic situations. Some other researchers have done this. Smit and Bluett (2011) developed statistical models to compute an 'on-road emission distribution' correction factor for laboratory based emission factors using remote sensing emissions data. A real-world Australian drive cycle was then used to quantify representative driving conditions (percentage of time spent in each VSP bin). Rhys-Tyler and Bell (2012) used remote sensing data, expressed as grams pollutant per kg fuel, in combination with the VSP frequency distribution (less than -10 to larger than + 20 kW) of the NEDC drive cycle to calculate a weighted emission factor.

Previous investigations have shown that vehicle emission models need to reflect local fleet composition and driving characteristics to provide adequate vehicle emission predictions. Large errors have been reported when overseas models are directly applied to local conditions without calibration (e.g. Smit and McBroom, 2009), because these models do not reflect local fuels, climate, fleet composition and driving conditions. Indeed, this was the main reason for the development of a dedicated Australian version of the COPERT software, known as 'COPERT Australia', using extensive emission measurements of Australian vehicles (Smit and Ntziachristos, 2012).

The remote sensing database is the only substantial emissions database available in New Zealand. It thus makes sense to develop an emission prediction model that incorporates these data. This however is not a trivial task given the significant number of differences between laboratory measurements, the common base for emission models, and remote sensing measurements, as was discussed in previous sections in this report.

This section discusses a new approach that combines power based emissions modelling with remote sensing measurements to develop a **hybrid micro-scale emissions model**. The model allows for detailed second-by-second emission simulation and time-efficient assessment of the impact of various traffic conditions on emissions. The hybrid model presents an alternative way of predicting vehicle emissions in New Zealand and the results can be used in a comparison study with VEPM.

5.1 Hybrid model structure

This section presents a proof-of-concept of the hybrid emission model. This illustrates how the rich New Zealand remote sensing database can be used to develop a genuine New Zealand emissions model, which can be used to 'validate' VEPM, but also allows for emission predictions at a significantly higher resolution in space and time than VEPM.

The key ingredients for the hybrid model are:

- Remote sensing data
- CO_2 emission algorithms
- Relationship between air pollutants and CO_2
- Input drive cycles

As Australian data were readily available for all these four aspects, these were used for the proof-of-concept. However, New Zealand remote sensing and drive cycle data can be used instead of and replace the Australian data to develop a New Zealand hybrid remote sensing model, as will be discussed later.

<u>Remote sensing data</u>

An extract of a large remote sensing database was used to develop quantitative relationships between air pollutant and CO_2 emissions (Smit and Bluett, 2008). The extract contains approximately 1,000 remote sensing measurements that were made in Perth, Australia, at five monitoring sites on relatively high capacity urban arterial roadways. The extract contains information on date and time of measurement, instantaneous vehicle speed, instantaneous vehicle acceleration, VSP, concentration ratios of CO, HC and NO to CO_2, license plate number, vehicle make and model, year of manufacture, number of cylinders and tare mass. Actual on-road vehicle mass was estimated for each vehicle by adding 200 kg to reported tare mass. VSP can therefore be expressed as kW (equation 6), and not the conventional kW/tonne. VSP expressed as kW is required as input to the CO_2 emission algorithms. The data reflects measurements for two vehicle classes (model years 1998-2003):

1. Small petrol passenger cars with an engine capacity of less than 2 litres. Typical vehicles are BMW 318, Ford Focus, Holden Barina, Holden Astra, Honda Civic, Hyundai Getz, Hyundai Excel, Mazda 323, Mitsubishi Lancer, Renault Clio, Toyota Echo and Toyota Corolla.

2. Large petrol SUVs with an engine capacity typically larger than 3 litres. Typical vehicles are Jeep Cherokee, Landrover Discovery, Mitsubishi Pajero, Nissan Patrol, Toyota Kluger, Toyota Landcruiser and Toyota Prado.

CO_2 prediction algorithms are developed for those vehicle classes.

Modelling of CO_2 emissions

A hierarchy of vehicle emission models can be distinguished based on the level of complexity and types of application. These include 'average-speed' models (e.g. VEPM, COPERT), where emission rates (g/veh.km) are a function of mean traveling speed, 'traffic-situation' models (e.g. **HBEFA**, see e.g. INFRAS, 2004), where emission factors (g/veh.km) correspond to particular traffic situations (e.g. 'stop-and-go-driving', 'freeflow') and 'modal' models (e.g. **PHEM**, see e.g. Hausberger et al., 2009; **CMEM**, see e.g. Barth et al., 2000; **PΔP**, see e.g. Smit, 2014), where emission factors (g/s or g/ driving mode) correspond to specific engine or vehicle operating conditions.

As remote sensing measurements are effectively snapshots in time, the appropriate time scale for a model using these measurements is 1 Hz or less. This means that modal (second-by-second) vehicle emission models reflect the most appropriate scale for the hybrid model.

Pollutant-to-CO_2 ratios are the most appropriate way of presenting and analysing remote sensing data, but these ratios do not show significant differences in fuel consumption (and hence CO_2 emissions) between different vehicles. It is therefore logical to stratify the on-road fleet in relatively homogeneous segments with respect to the main vehicle parameters that affect fuel consumption and CO_2 emissions.

Mellios et al. (2011) reported that the best performing CO_2 emission models for light-duty vehicles include the variables vehicle mass, rated engine power or engine capacity, fuel type and a variable representing engine power. The hybrid model should include these variables through an appropriate vehicle classification scheme and use a proxy variable for engine power such as VSP.

The CO_2 emission algorithms in the hybrid model follow a similar approach as the PΔP model (Smit, 2013b; 2014), although a few modelling aspects are necessarily simplified to allow for a proper combination with remote sensing data. The model uses a similar vehicle classification as COPERT Australia, which is based on the combination of fuel type, main vehicle type and vehicle emission standard. CO_2 emission algorithms are developed for the two vehicle classes discussed before, i.e. a small petrol passenger car and a large petrol SUV.

Empirical CO_2 emissions data were sourced from the Second National In-Service Emissions (NISE2) study (RTA, 2009) to develop CO_2 emission algorithms for the hybrid model. NISE2 provide vehicle emissions test data in both cold start and hot running conditions for 409 Australian light-duty petrol vehicles. Tests were conducted on a second-by-second ('modal') and aggregate ('bag') basis.

The exhaust emission database contains tests which were collected in a vehicle emissions testing laboratory over a 30-minute real-world drive cycle (speed-time trace) called the 'CUEDC-P' (Composite Urban Emission Drive Cycle for Petrol vehicles). This cycle was developed from Australian driving pattern data collected in the field. It consists of four phases, or sub-cycles, representing 'Residential', 'Arterial', 'Freeway' and 'Congested' driving conditions. Figure 3-8 shows the CUEDC-P drive cycle.

For each vehicle class a representative vehicle was selected for which verified and time-aligned modal CO_2 emissions tests are available. The vehicles are 'representative' as their average fuel consumption is similar to the average value for all vehicles in that vehicle class (± 5%). Table 5-1 shows vehicle information for the selected vehicles.

Table 5-1: Vehicle information for selected representative vehicles.

Vehicle Class	Make/Model	GVM (kg)	Rated Power (kW)	Engine Capacity (l)	Number of Cylinders (-)
PC-S petrol	Toyota Corolla 2000	1507	85	1.8	4
SUV-L petrol	Mitsubishi Pajero 2000	2799	140	3.5	6

Acceleration (a_t, m/s^2) is computed as:

$$a_t = \begin{cases} (v_2 - v_1) \div (t_2 - t_1) & t = 1 \\ (v_n - v_{n-1}) \div (t_n - t_{n-1}) & t = n \end{cases} \tag{8}$$

where v_t represents instantaneous vehicle speed (m/s) at time *t*, which varies from the cycle start time t = 1 to the cycle end time t = n. VSP (kW) is then computed for each time step using vehicle test mass, vehicle speed and vehicle acceleration as follows:

$$VSP_t = M\, v_t \left(1.1 a_t + 9.81 \sin(\arctan(g_t)) + 0.132\right) + 0.000302 M\, v_t^3 \tag{9}$$

Here, M is vehicle test mass (tonne) and g_t is the road gradient (%) at time *t*. Note that road gradient is zero for the laboratory tests. Figure 5-1 shows an example of a speed-acceleration scatter plot and two histograms for acceleration and VSP for the small petrol passenger car. It is clear that all VSP predictions are substantially below rated engine power for this vehicle.

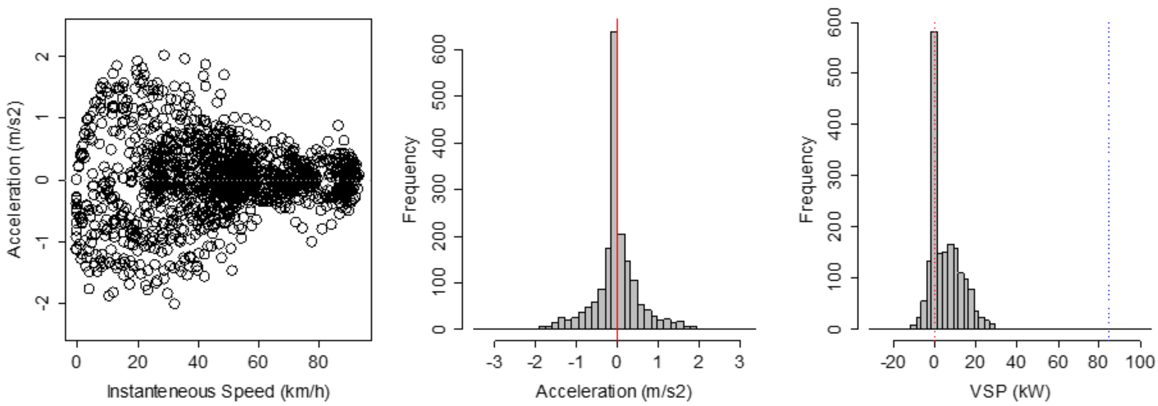

Figure 5-1: Speed, acceleration and VSP distributions in modal data for the small petrol passenger car (blue dotted vertical line shows rated engine power).

The empirical data were split into two datasets (called 'verification' and 'validation') to enable both types of assessment.[29] The empirical 'verification' dataset was fitted to the following equation to simulate the measured CO_2 emission rate ($e_{t,CO2}$, g/s):

$$e_{t,CO2} = \begin{cases} \alpha & v_t = 0 \\ \beta_0 + \beta_1 VSP_t + \beta_2 VSP_t^2 + \varepsilon & v_t > 0 \end{cases} \qquad (10)$$

For idling conditions (v_t = 0 km/h) a constant average value (g/s) is used. This value is determined for each vehicle by taking the average value of all instantaneous emission rates in idling conditions. For non-stationary driving conditions (moving vehicle) a polynomial regression function was fitted using the ordinary least-squares method, where β_0, β_1, β_2 represent the regression coefficients. Residual analysis was used to verify that the assumptions of the regression analysis were not violated. An example of graphical evaluation of model performance is shown for the small petrol passenger car in Figure 5-2. The plots show examination of goodness-of-fit, a normal probability plot and a residual analysis plot to verify normality and homoscedasticity of error terms (Neter et al., 1996). The results for the other vehicle are similar.

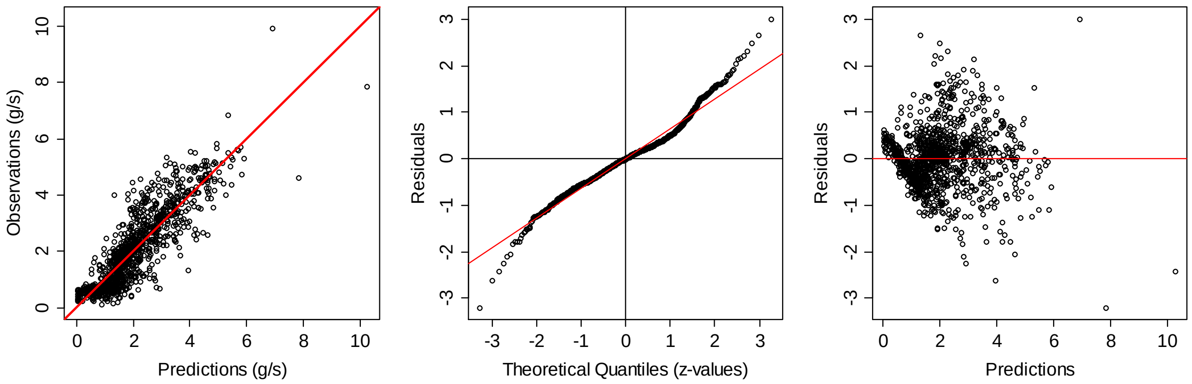

Figure 5-2: CO_2 model fitting and residual analysis for the small petrol passenger car.

The simple polynomial regression function provides a reasonable fit with coefficients of determination (R^2) varying from 0.79-0.80 and normalised RMSE[30] values varying from 5-6% (verification), and a slightly higher value of 8% for the validation data set, as is shown in Table 5-2. The empirical data were also fitted to higher order polynomial functions, but this did not lead to significantly improved emission functions, so the most parsimonious second-order polynomial model is retained.[31]

[29] Model verification assesses how well a model predicts the data on which it is based, whereas model validation assesses how well a model predicts with respect to independent data.
[30] Root-Mean-Square-Error (RMSE) is a frequently used measure of the differences between predictions and observations. It aggregates the second-by-second errors into a single measure of predictive power. Normalized RMSE is used to make RMSE scale-independent and it is computed by dividing RMSE by the range of observed values.
[31] A parsimonious model is a model that accomplishes a desired level of prediction accuracy with as few predictor variables as possible.

Table 5-2: CO_2 model parameters and performance statistics.

Vehicle Class	a	b_0	b_1	b_2	φ	NRMSE verification	R^2 verification	NRMSE validation	R^2 validation
PC-S petrol	0.75	1.24	0.121	0.0030	0.98	6%	0.80	8%	0.80
SUV-L petrol	1.08	2.10	0.122	0.0021	1.01	5%	0.79	8%	0.81

It is noted, however, that more accurate models can be developed from the laboratory emissions data. For instance, the PΔP model (Smit, 2013b) performs significantly better, with R^2 values varying from 0.93-0.94 and normalised RMSE values varying from 2-4% using the same empirical data.

The main differences between the simple VSP model and the PΔP model is the use of more complex algorithms in the PΔP model to quantify instantaneous engine power[32], as well as inclusion of an additional variable that quantifies the change in engine power over the last 3 seconds of driving (Smit, 2014). The PΔP model also uses time-series models to account for the fact that CO_2 measurements are dependent in time (autocorrelation effects).

However, the hybrid model needs to balance the vehicle driving information that is available from the remote sensing data with the level of detail and associated accuracy that is possible in CO_2 algorithms. The second-order polynomial VSP function aims to strike the right balance in this respect.

It is instructive to show time-series plots of predicted and observed CO_2 emissions and the speed-time profile used during emissions testing. Figure 5-3 and Figure 5-4 show the results for the small petrol car and the large petrol SUV. The black line represents the observations and the red dots represent the predictions.

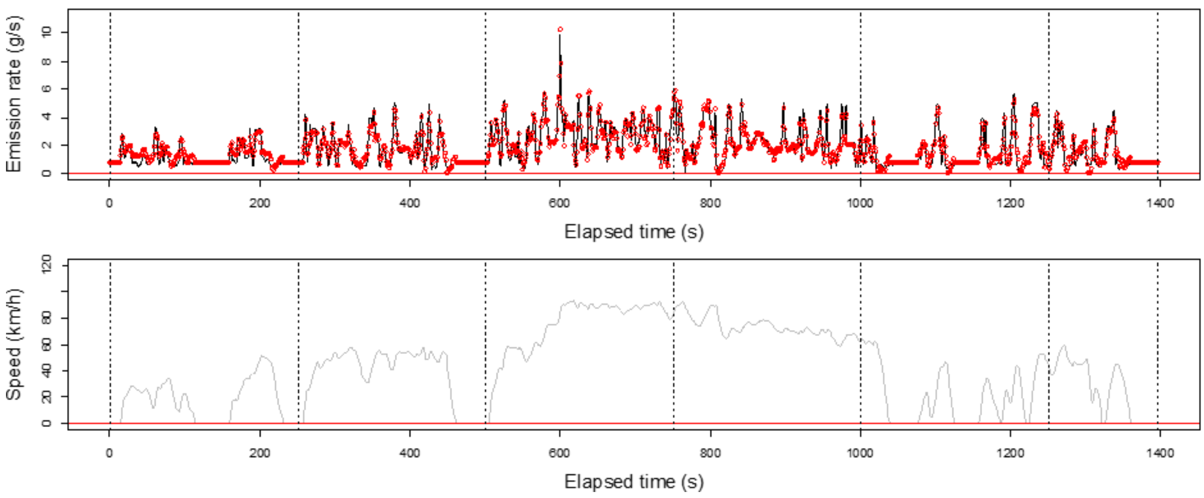

Figure 5-3: Predicted (red dots) and observed (black line) CO_2 emission rates for the small petrol passenger car and speed-time plot (model verification).

[32] For instance, explicit consideration of rolling resistance coefficients, frontal area and aerodynamic drag coefficient.

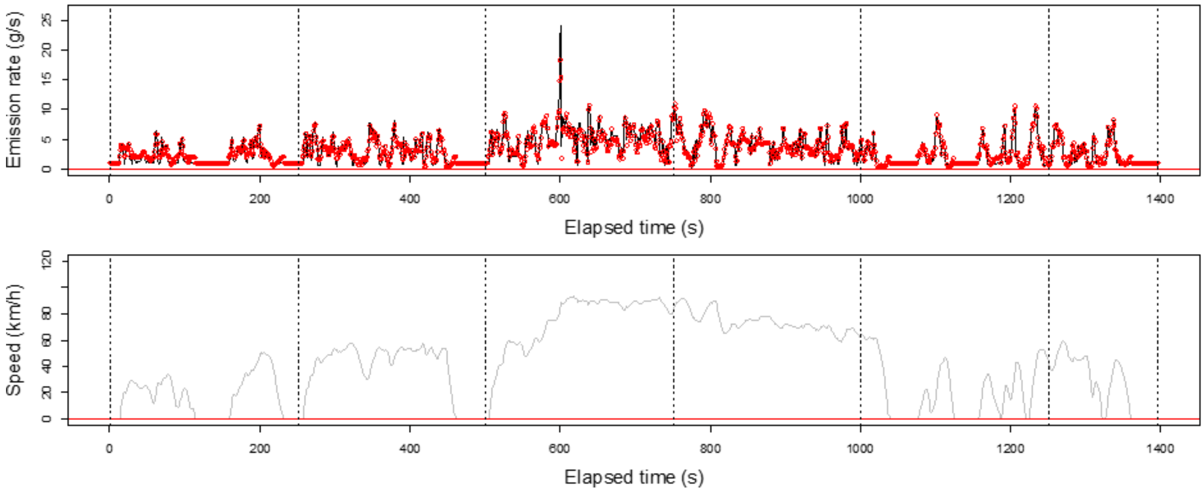

Figure 5-4: Predicted (red dots) and observed (black line) CO_2 emission rates for the large petrol SUV and speed-time plot (model verification).

The emission algorithms were then used to predict CO_2 emissions for the validation dataset, which was not used in model development. The time-series plots are shown in Figure 5-5 and Figure 5-6. It is clear that the model performs well with regard to the validation dataset.

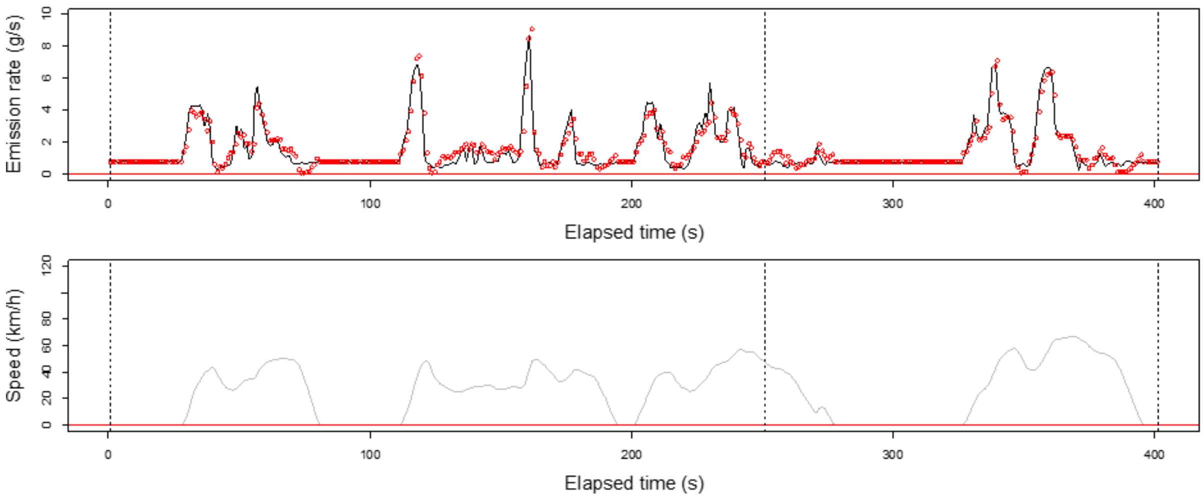

Figure 5-5: Predicted (red dots) and observed (black line) CO_2 emission rates for the small petrol passenger car and speed-time plot (model validation).

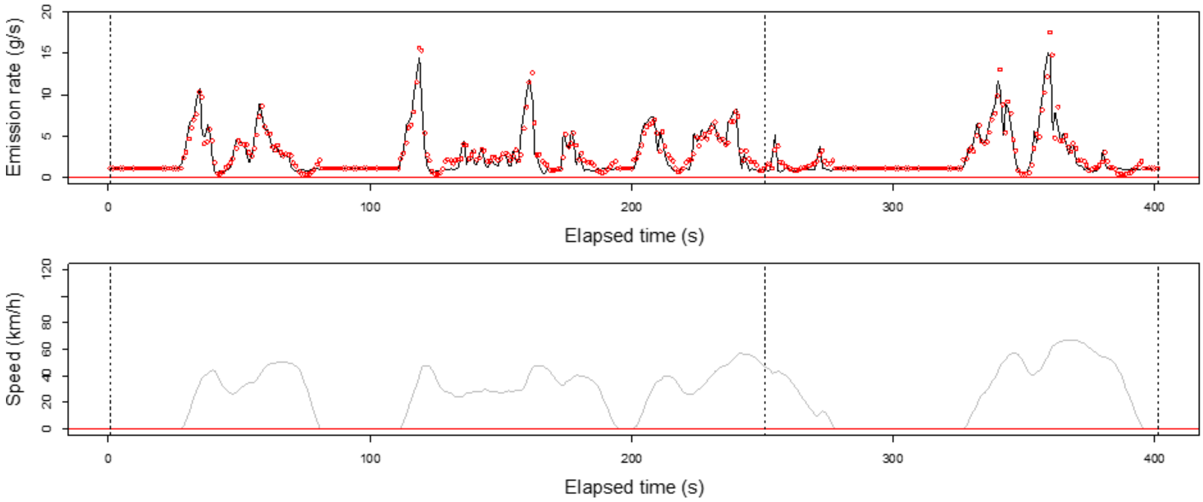

Figure 5-6: Predicted (red dots) and observed (black line) CO_2 emission rates for the large petrol SUV and speed-time plot (model validation).

A comparison between the model validation and model verification results with respect to model performance (R^2, NRMSE) is included in Table 5-2. With respect to prediction errors (NRMSE) and goodness-of-fit (R^2) the validation and verification show similar results, which demonstrates that the emission algorithms are robust with respect to second-by-second prediction performance.

Model prediction bias is calculated as the percent difference between total predicted (P) and total observed (O) CO_2 emissions (g) for the entire speed-time profile, i.e. $100 \times$ (P-O)/O. For the verification dataset the bias is zero percent, as would be expected. The validation dataset indicates that model bias varies from 2% to 8%, suggesting a model tendency to slightly overpredict total CO_2 emissions in urban conditions.

It is noted that spatial/temporal aggregation of second-by-second predictions will reduce prediction errors. This is illustrated in Figure 5-7 for the small petrol car for four different spatial resolutions, i.e. 100 m segments, 500 m segments, 1000 m segments and the CUEDC-P subcycles "Congested", "Residential" and "Freeway" (refer to Figure 3-8).

8). It is clear that model performance improves and prediction errors are reduced with decreasing spatial resolution.

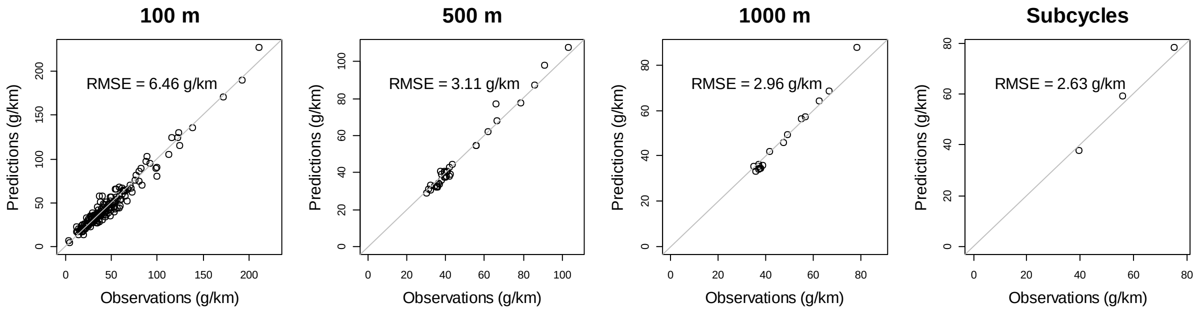

Figure 5-7: Impact of aggregation on CO_2 emission prediction error for the small petrol passenger car.

This is relevant information for application of the hybrid model, and for proper comparison with VEPM. For instance, it seems reasonable to use longer driving segments as input to the hybrid model, such as subcycles or stop-go-stop segments, to better align with the presumed spatial and temporal resolution of VEPM predictions.

A few final steps are required to complete the hybrid CO_2 emission prediction algorithms:

- It is important that total drive cycle emissions for the vehicles used in model development match those of the average values of similar vehicles in the empirical database. A calibration factor φ is therefore introduced and computed as the ratio of total cycle emissions (g) for the vehicles used in model development to average total cycle emissions of all tested vehicles of a particular vehicle class, in the same test conditions (CUEDC-P drive cycle). All CO_2 predictions are multiplied with φ. Calibration factor values are presented in Table 5-2.
- A few operational boundaries are applied to the emission simulation. Firstly, instantaneous VSP values cannot exceed 110% of the minimum and maximum values encountered during the emission tests. Secondly, emission rates are capped at a maximum measured value times a factor of 1.5.

Relationship between air pollutants and CO_2

The hybrid model links remote sensing data with CO_2 prediction algorithms through an appropriate vehicle classification scheme and use of a proxy variable for engine power (VSP).

Individual remote sensing measurements are highly variable, and very high pollutant-to-CO_2 ratios occur for specific measurements, often without a clear relationship to VSP. On the one hand this highly variable emissions behaviour distorts the relationship between VSP and emission ratio. On the other hand these data are valid outliers and they should be retained to ensure accurate mean ratios.

A simple approach was therefore adopted to establish a quantitative relationship between pollutant-to-CO_2 ratio and VSP. First, measurements with extreme and unrealistic acceleration values were removed, i.e. only measurements with accelerations within -6 and +3.5 m/s^2 were retained.

The data were then binned into predefined VSP bins and mean pollutant-to-CO_2 ratios were computed for each bin. Two options were discussed in section 3.5, i.e. a 6 bin or 14 bin VSP classification method. It was concluded that the two methods are not significantly different ($p > 0.05$) when the variability in the emissions data is considered, which means that the simplest method (VSP6) can be used.

Figure 5-8 shows box-and-whisker plots for both binning methods. A box-and-whisker plot is an exploratory graphic used to show the distribution of a dataset at a glance. The plots suggest that the VSP14 method generates a quantitative relationship between mean pollutant ratio and VSP that is ‘jumpy’ as high ratios may or may not occur in a particular bin. The VSP6 method is less affected by this erratic emissions behaviour because the bins are larger and, as a consequence, the likelihood of extreme values occurring within a bin is higher.

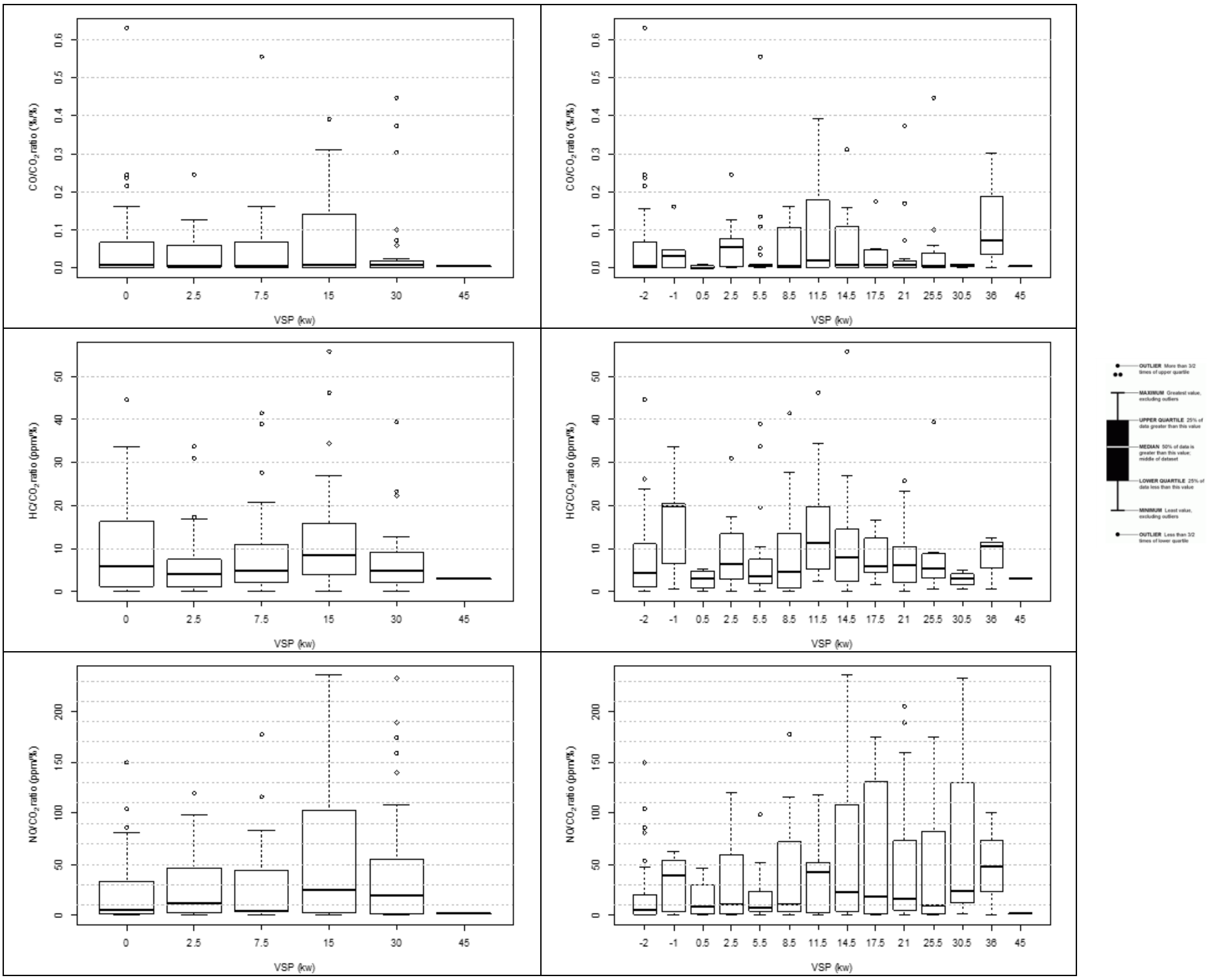

Figure 5-8: Box-and-whisker plots: mean pollutant ratios versus VSP bin for the small petrol passenger car using two methods (VSP6 left charts, VSP14 right charts).

The 6 bin VSP approach was used to compute average pollutant-to-CO_2 ratios. The results are shown Figure 5-9. The red bar indicates a sample size of one measurement.

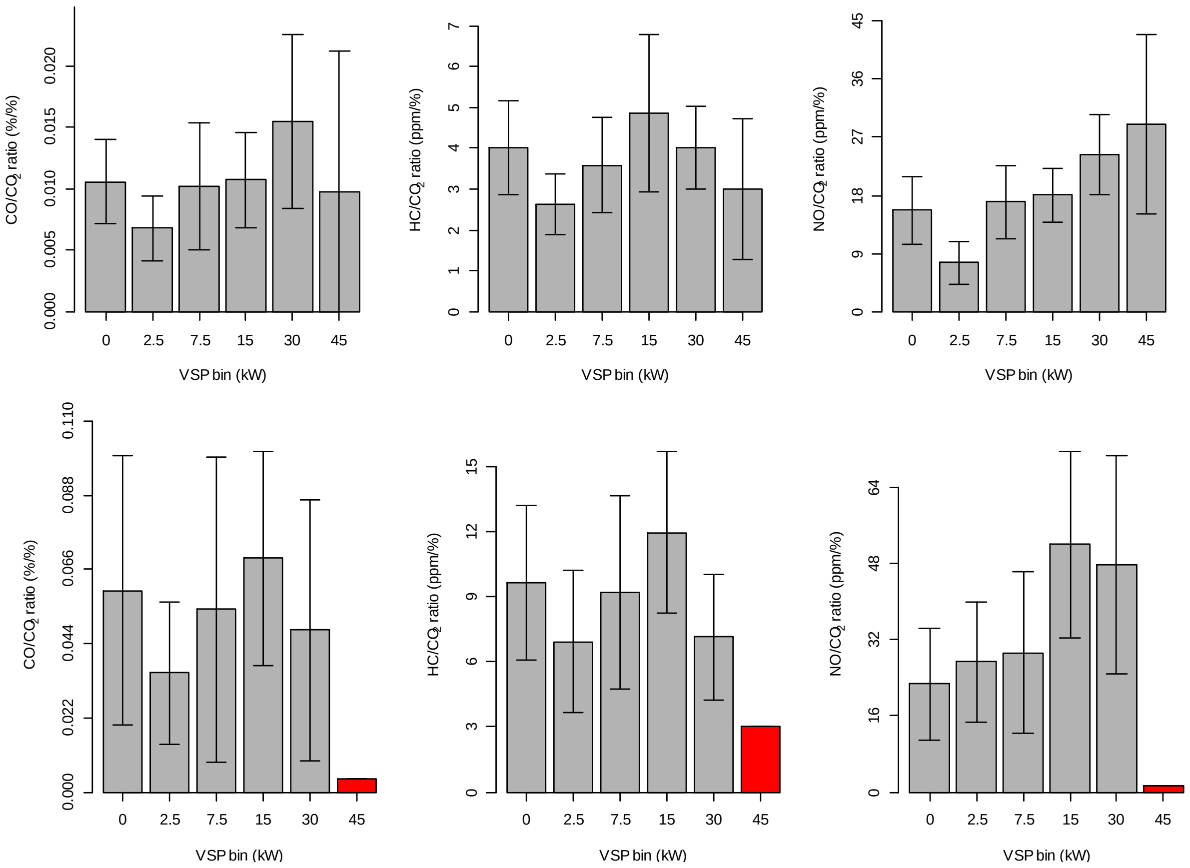

Figure 5-9: Mean pollutant ratios versus VSP bin for the small petrol passenger car (top charts) and the large petrol SUV (bottom charts), including 95% confidence intervals of the mean.

The 95% confidence intervals are also shown, and it is evident that the mean ratios are uncertain. The accuracy of the mean ratios can be improved by including more remote sensing measurements and increasing the sample size for each bin, which should be feasible as the sample size for the data extract has been relatively small.[33] The accuracy of the model predictions can be increased by aggregating second-by-second model predictions, as was discussed before.

[33] The sample sizes for VSP bins 1-6 are as follows for the small petrol car (n = 134, 91, 104, 239, 174 and 14) and for the large petrol SUV (n = 37, 29, 27, 42, 33 and 1).

Final hybrid model structure

The full hybrid model structure is as follows:

$$e_{t,CO2,j} = \begin{cases} a_j & v_t = 0 \\ b_{0,j} + b_{1,j} VSP_{t,j} + b_{2,j} VSP_{t,j}^2 & v_t > 0 \end{cases} \tag{11}$$

$$VSP_{t,j} = M_j \, v_t \left(1.1 a_t + 9.81 \sin(\arctan(g_t)) + 0.132\right) + 0.000302 M_j \, v_t^3 \tag{12}$$

$$a_t = \begin{cases} (v_2 - v_1) \div (t_2 - t_1) & t = 1 \\ (v_n - v_{n-1}) \div (t_n - t_{n-1}) & t = n \end{cases} \tag{13}$$

$$e_{t,p,j} = e_{t,CO2,j} \, \varphi_j \, R_{t,p,VSP6(i),j} \, \frac{M_p}{M_{CO2}} \, u_p \left(1 + \eta_{j,NO2}\right) \tag{14}$$

Where:

- $e_{t,CO2,j}$ is the predicted second-by-second CO_2 emission rate for vehicle class *j* (g/s),
- a_j presents the idle CO_2 emission rate for vehicle class *j* (g/s) (Table 5-2),
- $b_{0,j}$, $b_{1,j}$ and $b_{2,j}$ represent the fitted regression coefficients for vehicle class *j* (g/s, g/s.kW and g/s.kW^2, respectively) (Table 5-2),
- $VSP_{t,j}$ is vehicle specific power for vehicle class *j* (kW) and it represents an estimate of second-by-second engine power (kW) at time *t*,
- M_j is a representative on-road vehicle mass of vehicle class *j* (tonne),
- g_t is the road gradient (%) at time *t*,
- a_t represents second-by-second acceleration (m/s^2) at time *t*,
- v_t represents instantaneous vehicle speed (m/s) at time *t*,
- $e_{t,p,j}$ is the predicted emission rate (g/s) at time *t* for pollutant *p* (CO, HC, NO_x) and for vehicle class *j*,
- φ_j is the vehicle class calibration factor for CO_2 for vehicle class *j* (Table 5-2),
- $R_{t,p,VSP6(i)}$ is the mean pollutant-to-CO_2 ratio for VSP6 bin *i* at time *t* for pollutant *p* (CO, HC, NO) and for vehicle class *j*, as measured with remote sensing,
- M_p is the molar mass of pollutant *p*, which is 28.0, 44.1 and 46.0 g/mol for CO, HC and NO_x, respectively,
- M_{CO2} is the molar mass of CO_2, which is 44.1 g/mol,
- U_p is the unit conversion factor for pollutant *p*, which is 1 for CO (%/%) and 0.0001 for HC and NO (ppm/%), and
- $\eta_{j,NO2}$ is the proportion of exhaust NO_2 in NO_x emissions for vehicle class *j* that is not measured with remote sensing (this is assumed to be 0.04 for light-duty petrol vehicles).

Note that the hybrid model uses remote sensing measurements of NO, adds direct NO_2, and predicts NO_x emission rates, expressed as NO_2-equivalents.

5.2 Hybrid model application - drive cycles

The hybrid model predicts emissions at a high resolution and requires 1 Hz speed-time data and a selection of the appropriate vehicle class. This information can be obtained from various sources, including microscopic transport models, on-road GPS measurements and drive cycles.

As an example of a possible application, various Australian drive cycles were combined to create a single input file to the model. These drive cycles are the CUEDC-P (Composite Urban Emission Drive Cycle for Petrol vehicles, Orbital, 2005), six CUEDC-Ds (Composite Urban Emission Drive Cycles for Diesel vehicles, Brown et al., 1999), the AUC (Australian Urban Cycle, Watson, 1995 and Watson and Trayford, 1999), the MPC (Melbourne Peak Cycle, Watson et al., 1982) and the PDC (Perth Drive Cycle, Kenworthy et al., 1983 and Lyons et al., 1986). The drive cycles were broken up into 68 stop-go-stop segments, i.e. so-called 'microtrips'. An example of a segmented drive cycle is shown in Figure 5-10.

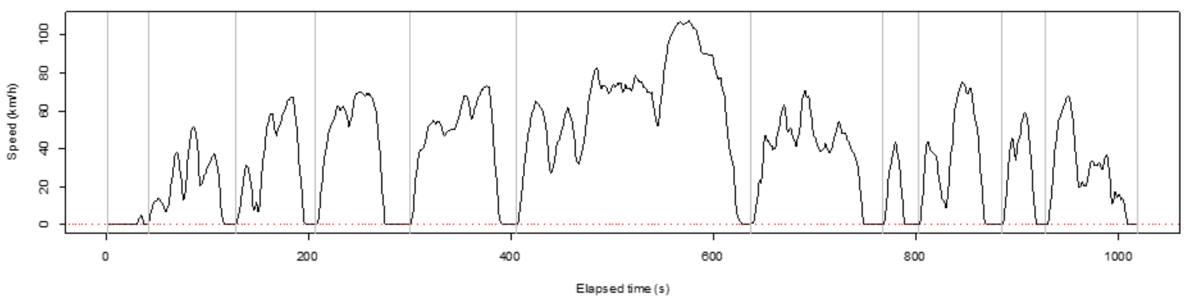

Figure 5-10: Example of dividing a drive cycle into microtrips.

Figure 5-11 (next page) shows an overview of all the microtrips that were generated. Microtrips with a length of 100 m or more were used as input to the hybrid model. The hybrid model was run to compute total emissions and emission factors (g/km) for each vehicle class and air pollutant. An example of an output table is shown in Table 5-3. These emission factors can be compared with emission factors used in VEPM.

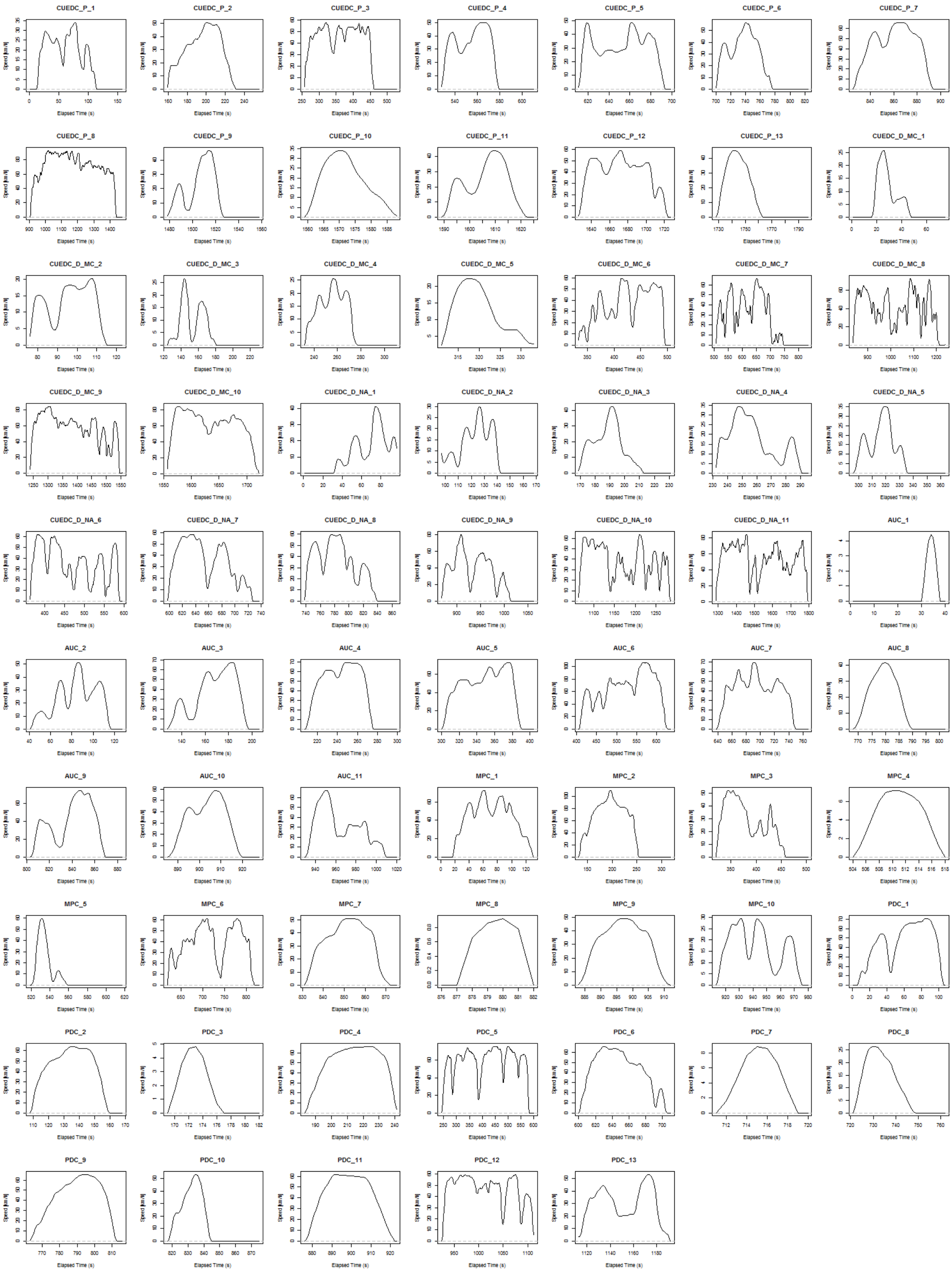

Figure 5-11: Microtrips used as input to the hybrid model.

Table 5-3: Example of output results table (CO, small petrol passenger car).

cycle	tt_sec	dist_km	av_spd_kmh	emis_gkm	vsp_less_0	vsp_0_5	vsp_5_10	vsp_10_20	vsp_20_40	vsp_more_40
CUEDC_P_1	158	0.59	13.5	1.67	41	102	13	2	0	0
CUEDC_P_2	98	0.63	23.2	1.14	19	60	16	3	0	0
CUEDC_P_3	271	2.62	34.8	0.91	55	132	57	27	0	0
CUEDC_P_4	83	0.50	21.7	1.41	17	45	11	9	1	0
CUEDC_P_5	89	0.72	29.3	1.43	33	40	3	6	7	0
CUEDC_P_6	125	0.72	20.6	1.34	33	66	13	13	0	0
CUEDC_P_7	80	0.85	38.3	1.29	26	21	16	10	7	0
CUEDC_P_8	573	10.78	67.7	0.77	70	162	174	157	10	0
CUEDC_P_9	81	0.31	13.6	1.91	16	48	11	6	0	0
CUEDC_P_10	30	0.15	18.0	1.87	17	5	5	3	0	0
CUEDC_P_11	37	0.21	20.6	1.99	16	11	3	5	2	0
CUEDC_P_12	102	1.10	38.7	1.01	30	39	15	16	2	0
CUEDC_P_13	70	0.27	13.7	1.82	18	41	3	8	0	0
CUEDC_D_MC_1	75	0.10	4.9	3.69	11	60	3	1	0	0
CUEDC_D_MC_2	48	0.14	10.9	2.13	13	32	3	0	0	0
CUEDC_D_MC_3	108	0.15	5.1	3.84	20	81	7	0	0	0
CUEDC_D_MC_4	79	0.19	8.6	2.33	17	58	4	0	0	0
CUEDC_D_MC_6	173	1.66	34.6	1.24	50	69	29	16	9	0
CUEDC_D_MC_7	327	2.29	25.3	1.42	94	151	23	47	12	0
CUEDC_D_MC_8	405	4.41	39.2	1.22	136	125	52	66	25	1
CUEDC_D_MC_9	317	5.01	56.9	0.99	81	78	81	51	26	0
CUEDC_D_MC_10	168	2.87	61.4	0.98	55	39	33	26	15	0
CUEDC_D_NA_1	96	0.32	11.8	2.30	25	60	6	2	3	0
CUEDC_D_NA_2	72	0.19	9.4	2.48	19	44	7	2	0	0
CUEDC_D_NA_3	63	0.23	13.2	2.11	20	34	5	3	1	0
CUEDC_D_NA_4	64	0.29	16.1	1.78	27	26	7	4	0	0
CUEDC_D_NA_5	68	0.18	9.7	2.42	15	43	5	5	0	0
CUEDC_D_NA_6	233	2.27	35.1	1.19	68	96	27	31	11	0
CUEDC_D_NA_7	141	1.29	33.0	1.08	42	46	35	18	0	0
CUEDC_D_NA_8	129	1.03	28.7	1.45	43	52	9	16	9	0
CUEDC_D_NA_9	196	1.61	29.5	1.43	59	77	31	18	11	0
CUEDC_D_NA_10	224	2.38	38.3	1.33	85	46	38	38	17	0
CUEDC_D_NA_11	509	7.71	54.5	0.91	131	112	150	101	15	0
AUC_2	87	0.51	20.9	1.80	26	36	13	8	4	0
AUC_3	79	0.70	31.7	1.53	23	24	8	18	6	0
AUC_4	93	0.97	37.4	1.15	17	41	14	17	4	0
AUC_5	105	1.24	42.6	1.06	21	39	16	27	2	0
AUC_6	232	4.06	63.1	1.14	69	44	38	43	31	7
AUC_7	131	1.31	36.0	1.27	49	45	12	16	8	1
AUC_8	35	0.14	14.7	2.10	9	16	2	8	0	0
AUC_9	82	0.75	32.7	1.82	30	26	5	10	11	0
AUC_10	44	0.33	27.0	1.93	13	15	4	7	5	0
AUC_11	92	0.69	27.1	1.55	24	42	6	15	5	0
MPC_1	131	1.39	38.2	1.33	47	34	15	24	11	0
MPC_2	188	2.37	45.4	1.20	32	87	15	34	17	3
MPC_3	184	1.14	22.2	1.24	51	106	15	11	1	0
MPC_5	99	0.23	8.5	3.98	20	71	0	2	6	0
MPC_6	213	2.20	37.2	1.00	69	70	48	24	2	0
MPC_7	45	0.38	30.4	1.24	15	14	7	9	0	0
MPC_9	30	0.25	30.2	1.58	11	7	4	6	2	0
MPC_10	68	0.30	15.8	1.74	23	34	8	3	0	0
PDC_1	107	1.27	42.7	1.15	26	22	29	25	5	0
PDC_2	61	0.66	38.8	1.07	13	20	12	16	0	0
PDC_4	59	0.81	49.6	1.06	14	7	27	8	3	0
PDC_5	358	5.64	56.7	0.91	61	105	99	80	13	0
PDC_6	111	1.27	41.1	0.91	31	44	18	17	1	0
PDC_8	42	0.12	10.0	2.41	16	21	5	0	0	0
PDC_9	54	0.61	40.6	1.10	13	14	11	16	0	0
PDC_10	59	0.23	13.8	2.16	9	37	5	7	1	0
PDC_11	48	0.54	40.2	1.29	15	19	2	6	6	0
PDC_12	189	2.44	46.5	0.98	44	70	43	21	11	0
PDC_13	80	0.64	28.9	1.34	30	24	14	10	2	0

Figure 5-12 shows the emission factor results versus average (microtrip) speed for all pollutants in scatter plots. The plots also show the hot running emission factors that are predicted by COPERT Australia for Australian light-duty vehicles (black dotted line), which includes petrol, diesel and LPG cars, SUVs and LCVs. Similar plots can be created using VEPM emission factors instead of COPERT Australia.

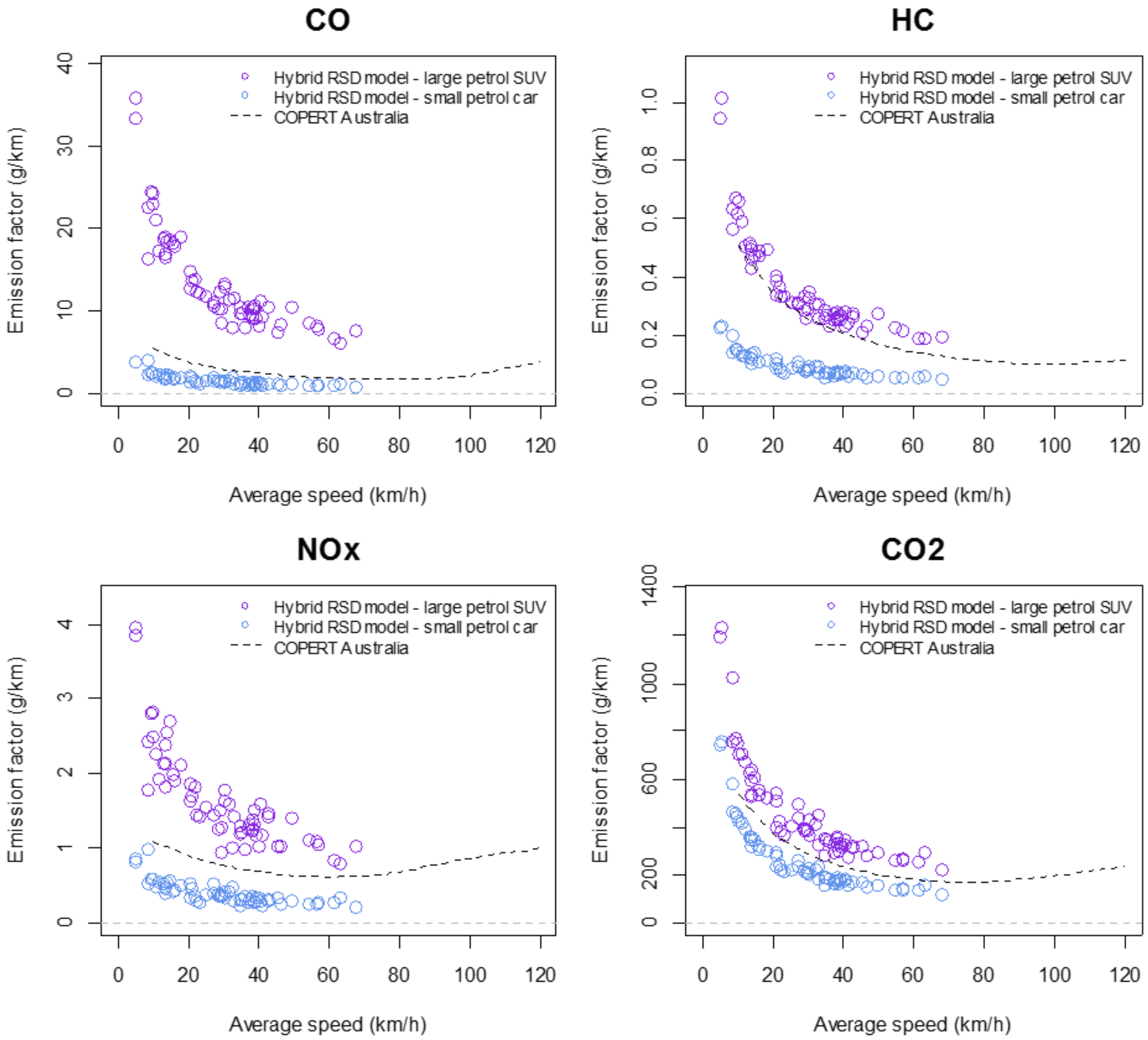

Figure 5-12: Emission factors predicted with the hybrid model as a function of average speed, and COPERT Australia predictions for the LDV fleet.

It is interesting to note that the light-duty fleet averaged emission factors for CO, HC and NO_x predicted by COPERT Australia appear to be reasonably consistent with the predictions made by the hybrid model for two specific vehicle classes. This indicates that the hybrid model produces reasonable results.

5.3 A hybrid model for New Zealand

The main benefit of the hybrid model is that it will make use of a large database of New Zealand specific remote sensing data. There is, however, also additional functionality as compared with VEPM, because the hybrid model can make emission predictions at a high resolution in time and space.

The hybrid model can be used to compute average vehicle emission factors (g/km) for various drive cycles or specific driving patterns of any length. A driving pattern is a second-by-second speed-time profile that has been recorded (measured) on the road or has been generated by a (microscopic) traffic model. It is rapidly becoming easier to collect speed-time data in the field by anyone who travels on public transport or by car. An example is a smart phone app called ATLAS II[34], which has been used in New Zealand (Safi et al., 2015). The app allows for quicker and less involved data collection efforts and records speed in time.

There are a number of developments that are expected to lead to an increased use of high resolution models:

- There is an increasing focus on the reduction of population exposure to air pollution and (health) risk. As a consequence, it will be important to know exactly which parts of the population are exposed to relatively high air pollution levels (e.g. near busy roads), what the level of impact is, and when this occurs. This type of assessment requires a fine spatial and temporal allocation of vehicle emissions in study areas, which can be achieved with the hybrid model.
- There is increasing interest in the effects of local traffic conditions on traffic emissions, fuel consumption and exposure to air pollution. It is essential to know if local traffic measures (signal settings, roundabout versus traffic light, dynamic speed limits, etc.) adversely affect or improve air pollution and greenhouse gas emissions. Sensitive models are therefore needed to accurately predict the correct direction and magnitude of these effects as this kind of measure typically generates relatively small but still significant impacts.

Development of a New Zealand hybrid remote sensing model would require the following steps:

- Analysis of the New Zealand on-road fleet with respect to variables that are known to govern fuel consumption and CO_2 emissions (fuel type, vehicle mass, rated engine power/engine capacity)
- Determination of an appropriate vehicle classification for the New Zealand fleet achieving relatively homogeneous segments with respect to the vehicle parameters that govern fuel consumption and CO_2 emissions.
- Development of CO_2 emission algorithms for each New Zealand vehicle class using laboratory measurements from e.g. Australia or other countries for representative vehicles, in the absence of New Zealand data.

[34] ATLAS II (Advanced Travel Logging Application for Smartphones II) is a smart phone app, which is downloadable from the AppStore. It aims to efficiently collect travel survey data. The app runs continuously in the background and uses advanced battery optimisation algorithms without interference with normal phone usage. The phone receives and processes real-time GSM and GPS signals to detect and record the device’s location on a second-by-second basis. Recorded data includes a timestamp, longitude, latitude and instantaneous speed. The app user can upload and view the trips on a map and download the travel data in a spreadsheet format.

- Compute average pollutant-to-CO_2 ratios for each VSP bin and for each vehicle class using remote sensing data collected in New Zealand.
- Collect New Zealand on-road driving data, for instance using mobile phone apps such as ATLAS II.

A New Zealand hybrid model will provide an alternative and independent method to predict vehicle emissions in New Zealand, and predictions can be directly compared with VEPM, as a 'validation' exercise. Any differences between the models will provide policy makers with information regarding the possible range of air quality impacts due to motor vehicles. If predictions by the two model turn out to be quite similar, then this this will increase confidence in vehicle emission predictions and scenario modelling for the New Zealand fleet.

6 Recommendations for the validation study

This section makes recommendations for the VEPM validation study.

6.1 Literature review and possible approaches

Seventeen international vehicle emission model validation studies using remote sensing were found and reviewed. Appendix A presents an overview of these studies. Depending on the specific case (pollutant, vehicle class, emission model), both small discrepancies (i.e. good agreement) and large discrepancies have been reported.

For instance, Kuhns et al. (2004) compared MOBILE6/PART5 to RSD data and reported differences varying between a factor +3 or –2 for CO and HC, and even a factor of 5 for PM. Large differences between US models and RSD based emission factors were also found in more recent US studies (e.g. Fujita et al., 2012). Sjödin and Jerksjö (2008) concluded that the comparison with RSD data shows that three European models (COPERT / HBEFA / ARTEMIS) can perform well, but can also perform very poorly (differences up to a factor of 23), depending on the pollutant, vehicle class and model. Beevers et al. (2012) compared trends in NO_x emission factors (g/km) determined using remote sensing data with those used in HBEFA and the UK NAEI and found reasonable agreement for petrol cars, but large differences for diesel cars.

So it is clear that no generic conclusions can be drawn from the literature. As a consequence, a validation study comparing New Zealand remote sensing data with VEPM will be required to examine the accuracy of VEPM.

In general, the literature review shows that comparisons between remote sensing and model predictions are made on a gram pollutant per kg fuel basis after selecting the appropriate traffic conditions in the emission models for hot running conditions. However, the studies do not consider meteorological conditions, and ignore cold start, evaporative emissions ('running loss' and resting loss HC emissions) and non-exhaust emissions (PM). In addition, selection of corresponding 'traffic conditions' in the emission model is a somewhat 'rubbery' step in the model validation that is not recognized as such in the literature. In other words, the selected corresponding traffic situation in the emission model is either an arbitrary process (e.g. choosing a particular traffic situation such as 'urban distributor road, 50 km/h speed limit, free-flow conditions'), or the use of measured average speeds by the RSD to select a corresponding emission factor in an average speed model like VEPM does not take into account the large difference in spatial resolution (few kilometres versus a few meters, see section 3.5).

In the next sections, a more thorough approach to validation of VEPM is proposed to ensure a comparison with remote sensing data is put on an even footing as much as possible. Nevertheless, it is recommended to also explore comparison of VEPM predictions with other independent data such as near-road air quality measurements or tunnel studies (e.g. Bluett and Fisher, 2005). Each measurement technique, whether it is laboratory emissions testing, remote sensing or a tunnel study, has its own strengths and weaknesses, but together they should provide a reasonably robust validation of VEPM.[35]

[35] For more info on this refer to Smit, Ntziachristos and Boulter (2010).

6.2 Summary of preparation steps for the validation study

It is clear from the previous discussions that there are several factors that impact on remote sensing and/or the basis for VEPM (laboratory emission tests). This section briefly outlines the necessary preparation steps that are required before the validation study can be conducted.

It is not possible to make a statement about which measurement method leads to more accurate vehicle emission predictions because both methods have particular strengths and particular challenges that could introduce bias in emission predictions based on these measurements. To ensure a valid comparison, the remote sensing measurements and VEPM predictions need to be modified or corrected, to the extent that this is possible.

Table 6-1 provides an overview of issues and challenges with remote sensing and laboratory vehicle emission measurements, effectively summarising the main points discussed in section 0. It also recommends specific actions to verify and possibly control for significant factors in the validation study.

An initial assessment of the relevance/importance of the individual factors is also included in the table within […]. This assessment is based on expert judgement only, and uses three qualitative (colour coded) levels: 'low', 'medium' and 'high'. However, it is recommended that a more in-depth sensitivity analysis is conducted to adequately quantify the impacts of these factors on the outcomes from a validation study, as will be discussed in section 6.4.

Table 6-1: Overview and factor control in the VEPM validation study.

Factor	Can be Controlled?	How?	Challenges and Remaining Issues
Sampling strategy (section 3.1)	Partially for remote sensing	VEPM: - add running loss and resting loss emissions to hot running emission factors [c] [medium (HC only)] - vary predictions with a plausible range of start emission impacts [c] [high] Remote sensing: - use of multiple RSD sites [a] [high] - cold start: analysis of vehicle registration address database and compare with distance to RSD location [c] [high] *Note on future options:* - employ long-path RSD on multilane freeways [a] [medium] - modified RSD setup to capture vehicles with vertical exhausts [b] [high]	Laboratory: - small sample size - excluding high emission events - underrepresenting high emitters - vehicle sourcing - representative gear shift behaviour - representative dynamometer loads Remote sensing: - large portion of invalid data - excluding specific traffic conditions[a] - excluding vehicle types e.g. trucks [b] - only moving vehicles - modification of 'normal' drive behaviour - extent of inclusion evaporative and cold start emissions is unknown [c]
Measurement techniques (section 3.2)	Partially for remote sensing	Remote sensing: 1 - quality verification and examination of NZ data [d,e,f,g]: • potential plume interference (section 3.2.1) [medium] • ensure measurement consistency for multiple sites [high] • identify and correct for potential measurement offsets (section 3.2.3) [high] 2 - NO data to represent NO_x (section 3.2.6) [h] • include direct NO_2 emissions [high] • express NO_x as NO_2-equivalents - develop NZ specific FID scaling factor for HC (section 3.2.6) [i] [high]	Laboratory: - FID response factors - PM measurement issues - Lack of details regarding underlying EU test programs Remote sensing: - low signal to noise ratios [d] - sensitive to equipment vibration [e] - residual plume interference [f] - artificial offset (optical misalignment) [g] - accurate detection of low emissions - only measuring NO [h] - NO to NO_2 conversion in plume - NDIR/FID compatibility [i]

Table 6-1 (Continued): Overview and factor control in the VEPM validation study.

Factor	Can be Controlled?	How?	Challenges and Remaining Issues
Emission factor computation (section 3.3)	Mostly	Remote sensing: - harmonised H/C ratio (section 3.3.3) [low] - develop NZ specific fuel carbon content (section 3.3.3) (j) [low] - develop hybrid modal – remote sensing model (l) [high] - correct for humidity and ambient temperature (section 3.3.3 and 3.6) (m) [medium (CO/HC)/high (NOx)] VEPM: - derive vehicle class specific fuel consumption rates for combination with RSD data, refer to equation 4 (k) [high] - correct for humidity and ambient temperature [36] (section 3.3.3 and 3.6) (m) [medium (CO/HC)/high (NOx)]	Remote sensing: - PM not included in combustion equations - assumed carbon content of the fuel (j) - assumed fuel consumption rates (k) - incompatible resolution for assumed fuel consumption rates (l) - no correction for humidity or ambient temperature (m)
Fleet composition (section 3.4)	Yes	Remote sensing: - aggregate remote sensing data to same vehicle technology classes as defined in VEPM (n) [high] - use detailed remote sensing information to examine variability in local fleet composition and impact on emission levels (n) [high]	Laboratory: - vehicle sample bias (n) Remote sensing: - vehicle sample bias (n) - representativeness of RSD sites
Driving conditions (section 3.5)	To some extent	Remote sensing: - develop hybrid modal – remote sensing model (p) [high] - examine speed-acceleration distributions in NZ data (q) [high] VEPM: - analysis of VSP distributions in drive cycles and comparison with remote sensing VSP data, VSP correction if required (o) [high] Note on future options: - resolve issues with acceleration measurements (a) [high]	Laboratory: - adequately accounting for real-world driving behaviour - unclear if drive cycles reflect New Zealand driving behaviour (o) - exclusion of factors such as road grade and aircon use Remote sensing: - incompatible spatial and temporal resolution with laboratory tests (p) - VSP does not account for all engine load factors, gear shift behaviour and history effects - time alignment errors - unexplained errors in acceleration measurements (q)

[36] This means either normalise RSD to 25°C or use ambient temperature correction in VEPM.

Table 6-1 (Continued): Overview and factor control in the VEPM validation study.

Factor	Can be Controlled?	How?	Challenges and Remaining Issues
Meteorological conditions (section 3.6)	Yes	Remote sensing: - correct for humidity and ambient temperature (r) [medium (CO/HC)/high (NOx)] - air density effects can be ignored for the NZ sites [low] VEPM: - correct for humidity and ambient temperature (r) [medium (CO/HC)/high (NOx)] - vary predictions with a plausible range of air conditioning use and impacts (s) [medium]	Laboratory: - typically measured at standard test conditions - exclusion of effects of aircon use Remote sensing: - meteorological data during remote sensing testing required [37] (r) - information on in-fleet aircon use and emission impacts (s)
Fuel quality (section 3.7)	Yes	VEPM - correct predictions to reflect commercial New Zealand fuels (t) [medium]	Laboratory: - reflects fuels used in overseas testing programs (t)
Ageing effects (section 3.8)	Yes	Remote sensing: - verify mileage correction algorithm with deterioration rates determined with RSD data over multiple years (u) [high] VEPM: - ensure mean vehicle class age corresponds with remote sensing data [high]	Laboratory: - reliability of mileage correction factors derived from laboratory data (u)
Vehicle loading (section 3.9)	No	Remote sensing: - use fuel based emission factors or pollutant-to-CO_2 ratios which are less affected by vehicle loading (v) [medium] VEPM: - vary predictions with a plausible range of vehicle loads (v) [medium]	Laboratory: - adequately accounting for real-world vehicle load distributions Remote sensing: - lack of information regarding in-fleet vehicle loads (v)

[37] If not available, these data can retrospectively be obtained from meteorological data measured at air quality monitoring stations close to the RSD sites.

6.3 Internal versus external errors

Emission predictions are affected by both internal and external errors. *Internal errors* are associated with the model itself (e.g. emission factors). *External errors* are associated with the errors in the input.

It has been shown that external errors are at least as important as internal errors (Smit, 2008a). Validation data for which traffic volume, traffic composition and driving behaviour are known (tunnel, remote sensing, laboratory and on-board measurements) tend to quantify internal errors. In contrast, studies that combine emission and dispersion models (ambient concentration, mass-balance) and/or traffic flow models (mass-balance) tend to quantify both internal and external errors of all models concerned (transport, emission and dispersion). As a consequence, these studies validate a 'modelling chain' and do not directly assess the prediction errors of traffic emission models. This complicates the explanation of the discrepancies between predictions and observations. For instance, errors due to dispersion modelling may offset or amplify modelling errors, but with an unknown direction and magnitude.

However, given the detailed vehicle information (type, fuel, emission standard, technology) that is recorded with a RSD, it is also possible to compare emission model input data (and their impact on emission predictions) that are commonly provided by other sources such as traffic models and national statistics ('generic' fleet composition). This will allow separate examination of internal and external errors in a New Zealand validation study. Given the importance of external errors it is recommended that these are examined separately in the project, i.e. start with examination of internal errors and then proceed with external errors.

6.4 Sensitivity testing

It is clear from previous sections that a number of preparation steps are required before RSD data can be compared with VEPM predictions. After taking those steps, several assumptions still underpin the RSD results, each one with their own inherent uncertainty. These assumptions may or may not have a significant impact on the RSD results.

To take this uncertainty into account in the validation study and to address concerns that are raised with respect to the use of remote sensing as a validation method, it is suggested that a sensitivity analysis is conducted. Sensitivity analysis (SA) can be used to examine the uncertainty in the predictions (VEPM) *and/or* measurements (RSD) of traffic emissions. For instance, VEPM can be used to quantify the impacts of cold starts and different vehicle load distributions. Similarly, a range rather than a single value (e.g. α, β, H/C ratio) can be used in the computation of emission factors from the remote sensing data (refer to equations 3 and 4).

There are different SA methods, but mathematical SA is well-suited to quantitatively assess the sensitivity of a model output to the (possible) range of variation of an input (Smit, 2008a). For instance a nominal range sensitivity analysis (conditional NRSA) is suitable.[38]

The sensitivity of the model/measurement data to the underlying assumptions can then be quantified in terms of distributions of emission values through simulation.

[38] NRSA is applicable to deterministic models and typically evaluates the effect of model outputs exerted by individually varying only one of the model inputs (OAT), while holding all other inputs at constant values. Conditional NRSA conditions the sensitivity on specific sets of input values ("situation"). These inputs are varied across their entire range of plausible values (two extreme values), which are derived from either test data, expert judgement or literature review. For each situation the impact on the results is then evaluated.

These emission distributions represent a _range_ of VEPM predictions and RSD measurements (rather than point estimates), which can be compared.

6.5 Suggested statistics and graphic presentations

The following plots and statistics are suggested for graphical and statistical analysis:

Initial Checks

- Compare VSP time distributions between cycles used for VEPM and RSD. Use a statistical test such as the Kolmogorov-Smirnov test to determine if they are significantly different.
- Multiply each VSP bin with average fuel-based emission factors to compute VSP-weighted average emission factors for both sets, and employ statistical testing to determine if they are significantly different.
- Use VSP distributions to select corresponding average speed ranges in VEPM.
- Examine the differences in emission ratios (g/g CO_2) and emission factors (g/kg fuel) between RSD sites, examine the robustness of these ratios and relate those differences to VSP and fleet composition (e.g. regression analysis).

Analysis options – comparison of model with measurement

- Permutations:
    - emission ratios: g pollutant/g CO_2
    - emission factor types: g pollutant/kg fuel or g/km
    - different levels of aggregation (e.g. model year, vehicle class, aggregated)
    - prediction and measurement *ranges* (min, mean/median, max) instead of point estimates – both RSD and VEPM
- Emission inventory comparison and calibration: fuel-based RSD emission inventory versus traffic model VEPM emission inventory.
- Present scatter plots of mean RSD versus VEPM results, where each data point is the average emission factor value for a particular model year: determine regression trend line and statistics, focus on correlation, slope and intercept (e.g. Pokharel *et al*., 2000, Fig. 1).
- Examine emission factor distributions (ECDF), compare relevant statistics (mean, median, 99^{th} percentile etc. – refer to e.g. Smit and Bluett, 2011) and use statistical tests to determine if they are significantly different. Another option for emission factor distributions is emission factors on the y-axis and percentile bins on the x-axis.

- Collect the drive cycles used in the (UK) NAEI databases, which underlie VEPM, and compute the proportion of time spent in each VSP bin. Compute the mean RSD ratio for each individual cycle and plot them against average (cycle) speed, then fit a mean ratio prediction algorithm to the data, convert to g/km and compare with the speed-based algorithms as used in VEPM.
- Explore the feasibility of an RSD-based correction factor.
- Develop a hybrid micro-scale emission model based on remote sensing for New Zealand. Apply the model and compare detailed hybrid model predictions with VEPM.

7 Description of the Remote Sensing database

7.1 Description of the campaigns

NIWA has consolidated the data from four Auckland based RSD campaigns in 2003, 2005, 2009 and 2011 into a database. The focus of these campaigns has been on the light duty vehicle fleet.

Other campaigns have taken place including Auckland in 2004 for buses and trucks, Wellington in 2005 and 2006, Brisbane in 2006 and Perth and Sydney in 2007. These data are not currently included in the database.

7.2 Instrumentation

NIWA used the RSD3000 in the 2003 campaign and then purchased the RSD4000^{EN} model, also called in some documentation RSD4600. Before the 2011 campaign the software was updated to 'nextgen', which did not change any of the underlying assumptions or values that calculation of emissions are based upon.

7.3 Sites of measurement

15 sites were used for the 2003 campaign; only four were continued in the later campaigns. Many of the sites from the original 2003 campaign were rendered unsuitable for remote sensing as they had been either upgraded to more than one lane, installed with ramp metering signals which interrupt the traffic flow, or no longer existed. Sites that are in similar areas to those lost have been brought online as replacements. The aim has always been to include as wide a distribution of sites across Auckland as possible, both geographically and in terms of surrounding land use and potential vehicle mix. Figure 7.1 shows the seven sites used in the 2011 and 2009 campaigns. They were all used in 2005 along with 16 others.

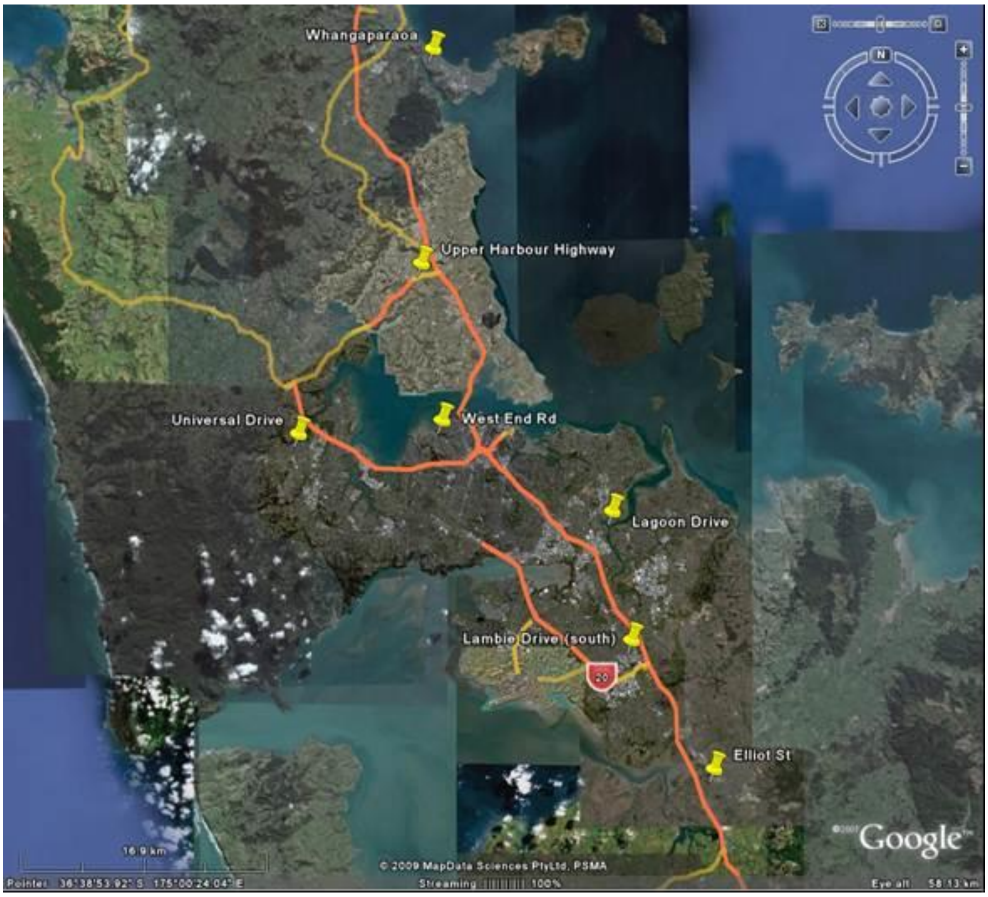

Figure 7-1: Remote Sensing Sites in the 2011 NIWA Remote Sensing Campaign.

The 2011 campaign consisted of 7 sites that have all been used in previous campaigns. Table 7.1 from the 2011 report compares them over the years.

Table 7-1: Remote Sensing Sites in the 2011 NIWA Remote Sensing Campaign.

Site No	Site Name	Site Code	2011	2009	2005	2003
1*	Lagoon Dr	AUC2	4,045	4,437	7,785	3,884
2*	Lambie Dr (S)	MAN2	930	1,339	4,295	2,379
3	Universal Dr	WAI5	2,052	5,385	2,545	n/a
4	West End Rd	AUC8	1,133	1,066	2,555	n/a
5	Whangaparaoa Rd	ROD3	9,213	3,826	3,850	n/a
6	Elliot St (W)	PAP1	1,349	1,342	1,367	1,447
7	Upper Harbour Highway (W)	NOR5	5,660	5,558	2,992	1,937
Total valid readings			24,382	22,953	25,389	9,647
Total individual vehicles*			20,895	21,383	23,310	9,338

* Note some vehicles went through the remote sensor more than once – in one case 67 times – and therefore the number of individual vehicles captured is lower than the number of valid readings. The results presented in the following sections show the number of individual vehicles.

What is not shown in the 2011 report is a comparison of the capture rates at these sites, that is, the number of valid readings to total readings obtained. This can be done by collating the data in all four campaign reports, as shown in Table 7-2. For a reading to be valid it must firstly have valid RSD readings, meaning the exhaust plume was measurable, and it must also have a readable number plate captured by the imaging, in order to link those readings with an individual vehicle.

Table 7-2: Capture rates at sites over the four campaigns.

Site No	Site Name	Site Code	2011	2009		2005	2003
1	Lagoon Dr	AUC2	48%	54%		60%	78%
2	Lambie Dr (S)	MAN2	37%	45%		72%	73%
3	Universal Dr	WAI5	49%	50%		55%	n/a
4	West End Rd	AUC8	47%	57%		87%	n/a
5	Whangaparaoa Rd	ROD3	67%	47%		57%	n/a
6	Elliot St (W)	PAP1	60%	51%		76%	86%
7	Upper Harbour Highway (W)	NOR5	56%	63%		73%	90%
Mean Capture rate			**52%**	**52%**		**69%**	**82%**

The capture rates show a decline over the campaigns. Potential reasons for this include changes in the road layout at the sites, which make it more likely a vehicle will pass through the RSD while not actively accelerating. Over the campaigns there have been changes in the layout of the instrumentation at some sites to try to combat these effects or increase the capture rate. The fleet composition recorded at these sites has not changed substantially, so an increase in unreadable vehicles will not be a major factor. Changes to the instrumentation, either through upgrades, servicing or degradation of the instrument over time appear an unlikely cause, according to the developers, ESP. The importance of the decline in capture rates depends upon whether any particular types of vehicle are captured less frequently than they used to be, thus introducing a sampling bias over time.

7.4 The Motochek database

Valid readings (vehicles with valid measurements and a readable number plate image) are collated and the number plates are sent to the NZTA’s Motochek service, which provides confidential information about each vehicle and owner. Table 7-3 shows the information provided by Motochek that is incorporated into the RSD database. This information is crucial in allowing the emissions to be grouped by vehicle class, age, mileage and fuel type. It could also be used to potentially filter out cold start emissions, from cars measured too close to their registered address, or brand new cars still being driven in.

Table 7-3: database field acquired from Motochek for every vehicle surveyed. (From the 2011 report)

Motochek database field	Description of data
Make	Company which manufactured the vehicle
Model	
Year of manufacture	
Body style	Saloon, hatchback, station wagon, utility, light van, flat deck truck, heavy bus/service coach etc.
Main colour	
Engine capacity	Cc
Engine power	kW
Vehicle type	Passenger car/van, goods van/truck/utility, motorcycle, bus, trailer/caravan, tractor etc.
Purpose of vehicle use	Private passenger, taxi, commercial passenger transport, licensed goods, other (standard) goods, ambulance, fire brigade, diplomatic etc.
Fuel type	Petrol, diesel, LPG, CNG, other
Country of origin	Country where vehicle was manufactured
WOF expires	Warrant of fitness expiry date
Registration status	Active, cancelled or lapsed
Country of first registration	Country where vehicle was first registered
Gross vehicle mass	kg
TARE weight	kg
Odometer reading	km or miles
Plate type	Standard, trade, personalised, investment, diplomatic or crown
Ownership	Private (male or female), company, fleet or lease
Subject to RUC	Subject to road user charges

It should be noted that the information Motochek provides introduces its own uncertainties. For instance low odometer readings will mostly be new cars but may include old cars that have 'clocked' their odometer and have travelled more than one million km. Registered address details may not be actually where the vehicle is parked when not in use.

7.5 The structure of the RSD database

The RSD data is stored as a Microsoft Access database. The structure is that the data itself are kept in two large tables: one contains all the data from the RSD instrumentation itself, the other contains all the data retrieved from the Moto-check database.

'Surrounding' these two core tables are a number of look-up tables that list out all the variables and codes used within them. The screen grab (Figure 7-2) shows the main tables and dependent tables and the list at the side includes the look-up tables.

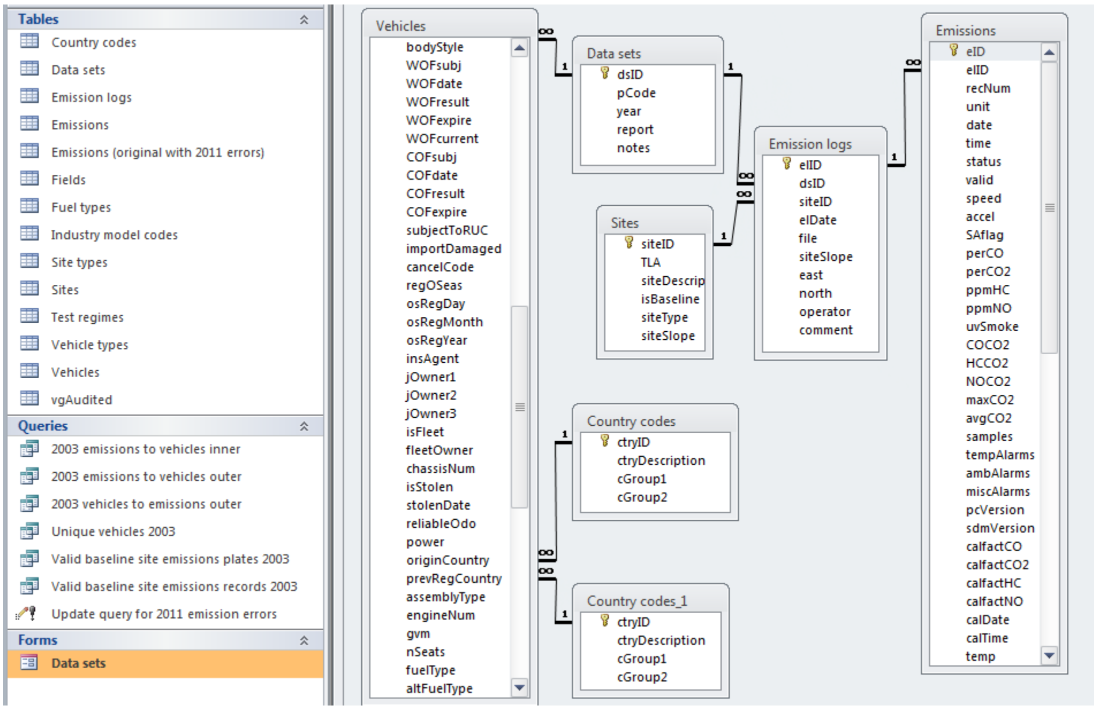

Figure 7-2: Structure of the RSD database.

7.6 QA procedures to clean up the database for use

All the raw data from the RSD are stored within the database. Although the user is free to create modified databases with only subsets of the raw data (i.e., QA'd data) if they wish, and interrogate the database in Access, the current tool for QA and analysis is the R statistics package. A series of precursor scripts have been developed by Martin Unwin of NIWA to load the database into R and then strip out unusable data (see Appendix 3 for these steps). The end result is a set of dataframes that are ready to be analysed. (A minor restriction to this approach is that the function that first reads the Access database into R only works with the 32 bit version of R, not 64 bit.)

8 Conclusions

The use of remote sensing to measure vehicle emissions is not new, dating back to 1971. Disagreement between vehicle emissions models and tunnel measurements in the 1980s-1990s contributed to increased use of remote sensing devices (RSDs) as an independent approach to measure real-world vehicle emissions. Remote sensing has been used extensively over the last two decades for various purposes around the world, including but not limited to identification of high-emitting vehicles, examination of on-road vehicle emissions distributions and trend analysis.

This study has scoped out and tested the best and most useful ways for validation of the New Zealand vehicle emission model VEPM[39] using remote sensing data for CO, THC and NO_x. This was done through an examination of potential issues with the remote sensing data and ways to address them, as well as a discussion and demonstration of various (possible) ways to compare remote sensing data with VEPM emission factors.

Testing the overall accuracy of vehicle emission models is challenging, as the 'true' emission values in urban networks are unknown and cannot practically be determined by measurement. As a consequence, VEPM predictions and remote sensing data are *both* independent *estimates* of the true vehicle emissions in a road network.

If both VEPM and the remote sensing data show similar predictions of vehicle emissions, then this would increase the level of confidence that VEPM predictions are in fact accurate. If not, then one or both the VEPM model predictions and remote sensing estimates are not accurate. More validation work using other independent methods such as tunnel studies, near-road air quality sampling etc., and further analysis are required to determine where the (main) errors occur (specific traffic situations, specific vehicle classes, etc.).

Laboratory measurements ('bag' and 'modal', engine and chassis dynamometer) using drive cycles have remained the prominent empirical base for vehicle emission model development around the world. Other methods such as remote sensing, tunnel studies and on-road or near-road modelling have been commonly used for emission model validation purposes, and have contributed to an increased understanding of model accuracy and real-world emissions behaviour of vehicles.

There are specific issues with and points of attention for all emission measurement methods. The strengths and weaknesses of *both* laboratory measurements, on which VEPM is based, and independent remote sensing measurements have therefore been examined. This is important to ensure that a valid comparison of methods is made and to better understand where potential differences may come from.

There are fundamental differences between emission estimates derived from remote sensing measurements and laboratory-based model predictions that need to be accounted for in the validation study. Nine fundamental differences were identified and discussed in this study:

[39] Vehicle Emission Prediction Model.

- **1. Different sampling strategies**

  Remote sensing typically collects ‘snapshot’ on-road emissions data from a relatively large sample of moving vehicles, whereas laboratory testing collects drive cycle (trip) based emissions data from a relatively small sample of both stationary and moving vehicles under controlled conditions.

  For proper emission model validation, emission trend analysis or emission inventory development, remote sensing data for *multiple* sites are required to properly reflect the real-world variation in weather, road features, local fleet mix and driving behaviour.

  Remote sensing can record emission levels for a relatively large number of vehicles, but the system can produce a significant portion of invalid data (20-70%). In addition, the equipment setup and weather conditions can modify driving behaviour and result in exclusion of relevant vehicle types (e.g. trucks, cars with caravans) and traffic conditions (e.g. congested conditions, multi-lane highways).

  Laboratory measurements can exclude high emission events when the high quality measurement equipment cannot handle excessive emission levels. More subtle factors that may create bias, but they are difficult to quantify, are the representativeness of gear shift behaviour of vehicle operators in laboratory conditions and how well dynamometer settings and loading algorithms reflect on-road driving conditions.

  The usual Remote Sensing configuration used in New Zealand does not capture trucks and buses with vertical exhausts.It is recommended that any future remote sensing programs in New Zealand are able to include these vehicles.

  Given the increasing relevance of cold start emissions it is important to examine the (expected) proportion of cold start vehicles in the remote sensing measurements for each location, and to determine if the measured emissions are likely to be hot running emissions, or not (e.g. using license plate information and analysis of travel time to RSD site).

  Although remote sensing studies often state that only exhaust emissions are measured, this is incorrect, as evaporative emissions are included. It is unclear, however, to what extent evaporative running loss and resting loss hydrocarbon (HC) emissions are captured in the ‘exhaust plume’.

- **2. Different measurement techniques**

  Remote sensing uses different measurement methods compared to laboratory testing on which VEPM is based, and these differences need to be considered when a validation study is conducted.

  Whereas laboratory emissions are measured with standard pollutant analysers (FID, gravimetric filter method, chemiluminescence, NDIR), remote sensing system uses the principle that the majority of gases will absorb light at particular wavelengths (UV/IR).

Compared to the controlled environment of laboratory measurements, it is more challenging for remote sensing to measure vehicle emissions accurately. For instance, RSD measurements can be sensitive to equipment vibration caused by passing vehicles. There are also specific artefacts such as the reported artificial offset in the RSD HC measurements due to an '*optical misalignment*'. It is therefore essential to perform thorough quality checks on remote sensing data before it is used in the validation study, in addition to the internal and automated data verification procedures already used by the RSD.

The density of the observed exhaust plume and path length are highly variable and are a function of the height of the vehicle's exhaust pipe, wind direction and speed, and turbulence behind the vehicle, amongst other factors. The RSD can therefore only reliably measure ratios of CO, HC and NO to CO_2. These ratios are assumed to be constant in a particular vehicle's exhaust plume. However, RSD data are naturally noisy and sufficiently large sample sizes are required to obtain significant results.

The remote sensing measurements include a 'background correction' by subtracting the concentration measurements just before the 'beam block' (detecting vehicle presence) from the concentration measurements just after the 'beam block'. It is unclear, however, to what extent residual plume interference affects overall remote sensing measurements. It is therefore recommended that the New Zealand RSD data are examined to see if this effect can be detected.

Laboratory measurements have their own challenges and issues. For instance, FID will not adequately detect all hydrocarbon species present in vehicle exhaust, which will create an artefact.

International comparison studies indicate that substantial differences between remote sensing and laboratory measurements can be expected at individual vehicle level, but that aggregated emissions data will yield reasonable agreement.

Interestingly, comparison of remote sensing results with on-board and tunnel studies indicate a better agreement with remote sensing measurements than comparisons with laboratory measurements, even at an individual vehicle level. This provides further confidence that use of independent RSD data to 'validate' VEPM predictions is a reasonable approach.

Potential issues with low sensitivity of RSD at low concentration values were examined using modal laboratory emissions test data. This work indicates that the impact of zero versus low emission measurements on the mean pollutant ratios is small and typically generates an (underestimation) error within 5%.

There are a number of incompatibility issues between laboratory and remote sensing techniques, with NO and THC measurements in particular, which require further data preparation steps. These include accounting for direct NO_2 emissions, expressing NO_x emissions as NO_2 equivalents, determination of an accurate fleet average H/C ratio in the exhaust gas for the New Zealand fleet and computation of an appropriate HC scaling factor e.g. using New Zealand VOC emission profiles.

- **3. Different determination of emission factors**

  Remote sensing uses different calculations to determine emission factors as compared to standard laboratory testing on which VEPM is based, and these differences need to be considered when a validation study is conducted.

  Laboratory emission test results are directly converted into distance-based emission factors (g/km).

  RSD data are converted into emission factors expressed as grams per litre or kg of fuel burned using a chemical mass (carbon) balance approach, which excludes PM, and by making assumptions regarding the carbon content of the fuel and an appropriate 'NDIR to FID' scaling factor for HC. Distance-based emission factors (g/km) are then estimated by making assumptions about representative fuel consumption rates.

  Both laboratory and remote sensing computations are corrected for 'background concentration', although this is done in different ways.

  A humidity correction for NO is included in the conversion of laboratory test results to emission factors, but not in the RSD calculations.

  There is a clear difference in temporal and spatial resolution of the measurements (half a second versus a drive cycle test phase).

- **4. Different levels of detail in the on-road fleet mix**

  Remote sensing provides both detailed emissions and vehicle information for the majority of vehicles it measures. This is in contrast with emission models such as VEPM where emission factors are available for specific vehicle classes only.

  As a consequence, the remote sensing data needs to be *aggregated* to develop emission factors that correspond to the VEPM vehicle classification, before a comparison can be made.

  The representativeness of the RSD sites needs to be examined and confirmed. In other words the combined set of RSD locations cannot *automatically* be assumed to be representative of the overall fleet composition, or even particular segments of it, excluding trucks with e.g. vertical exhausts, and the wide range of driving conditions encountered in the real-world.

- **5. Different spatial and temporal resolution regarding driving conditions**

It is not possible to directly compare VEPM emission factors (g/ km) that are derived from measurements over various drive cycles with remote sensing data collected at fixed roadside locations. Remote sensing techniques measure concentrations and (approximately) associated driving conditions at a particular point on the road (generally under slight acceleration) at a high frequency for about half a second, whereas drive cycles used in laboratory testing typically reflect 10 to 30 minutes of continuous driving over a range of driving conditions (effectively a 'journey').

A proper comparison of remote sensing and VEPM predictions requires a method to account for the difference in spatial and temporal resolution. To make the two data sets (more) comparable, variables are required that quantify 'driving conditions' and link the two independent databases. Two candidate variables were examined in this study: 'vehicle-specific power' (VSP) or vehicle speed and acceleration.

A statistical analysis on a subset of modal laboratory emissions data suggests that three classification methods, i.e. 6 VSP bins, 14 VSP bins and 176 speed-acceleration (v-a) bins, are not significantly different ($p > 0.05$), when the inherent variability in the emissions data is considered. This means that the simplest method (6 bin VSP) could be used. The detailed v-a approach does not provide significant accuracy benefits and generates specific issues such as exclusion of road gradient and vehicle mass as prediction variables.

Although VSP is the best approach to quantify driving conditions at RSD locations, unresolved issues remain, including incomplete quantification of driving behaviour (e.g. gear shifting, driving history), variable time-alignment errors and errors in measured acceleration levels (i.e. a significant number of noisy and unrealistic acceleration values). These issues cannot be addressed or resolved, but they add uncertainty to the quantification of driving conditions at the time of measurement.

VEPM is based on emission testing studies using chassis or engine dynamometers. Although these tests generally use high quality equipment and are conducted under highly controlled conditions, there are challenges remain to produce emission results that are representative for real-world driving conditions. Issues for consideration are, for instance, the use of real-world drive cycles and how well they represent New Zealand driving behaviour, exclusion of factors such as road gradient and air conditioning use and the quality of the dynamometer and dynamometer settings. Given that VEPM is based on overseas models, which in turn are based on a variety of overseas vehicle test programs and drive cycles, this information cannot readily be retrieved and verified.

- **6. Different meteorological (test) conditions**

  Meteorological variables such as ambient temperature, air density and humidity have a direct impact on vehicle emissions. In addition, these variables affect air-conditioning use, which affects emissions because air-conditioning requires a significant amount of energy.

  Laboratory tests are typically conducted at specific temperatures and humidity. RSD measurements reflect actual ambient temperatures and humidity. As a consequence, temperature correction needs to be applied to either VEPM predictions or remote sensing data.

  Air-conditioning is normally not incorporated in laboratory vehicle emission testing. VEPM has no air-conditioning correction factors, but air-conditioning impacts will be reflected in the RSD results. So a correction of either the VEPM or RSD results is recommended, but the feasibility of this is unclear as it requires accurate information regarding the penetration of air conditioners in the on-road fleet and its actual use.

- **7. Different fuel quality**

  The composition and quality of fuel has a significant impact on emissions.

  VEPM is to a large extent based on overseas models, and therefore reflects the quality and composition of the fuel used in overseas testing programmes, at the time of measurement. The RSD data will reflect the local NZ fuel quality and composition, and will be more accurate in this respect.

  VEPM emission predictions need to be corrected for commercial NZ fuels, to the extent that this is possible, if this has not already been done in the development of VEPM.

- **8. Different impact of ageing effects**

  Even with reasonable maintenance, vehicle emissions increase with mileage and age due to deterioration of engine and emission-control components. Emission deterioration is intrinsically related to the durability of emission control devices (including OBD), in-service I/M programs, driving conditions and driving behaviour.

  The remote sensing data reflect snapshots in time. Fleet-wide ageing effects can be captured by relating RSD emissions to vehicle model year and by conducting remote sensing campaigns in subsequent years and, ideally, multiple sites.

  Laboratory based VEPM predictions are ‘fixed’ and reflect the ageing effects at the time of measurements only. Ageing effects are normally simulated in emission models by generic mileage correction algorithms, which are based on limited data. Recent research indicates that the accuracy of those algorithms in vehicle emission models is low and requires attention.

- **9. Different impact of vehicle loading**

  Vehicle load can have a significant impact on vehicle emissions.

  Actual vehicle loading is not measured with the RSD, although a few studies reported the use of a weighing station before remote sensing measurements were carried out.

  Vehicle loading is recorded during laboratory testing, but the extent to which dynamometer loading corresponds to real-world driving is unclear.

  There appears to be no practical way to include 'vehicle loading' into the validation study due to a lack of information in this respect.

  The lack of vehicle weight information appears to be less of an issue for the actual remote sensing measurements. This is because the pollutant-to-CO_2 ratios are less sensitive to vehicle loading.

Vehicle emission models are often used beyond their intended purposes and capabilities, resulting in errors. This may partly be explained by lack of understanding of the model development process, the underlying empirical data and the exact definition of variables used by the model, and clear documentation of these matters.

VEPM's intended use is for emission estimation at road network level. The use of 'average speed' as the only variable to capture the impact of driving conditions on emission levels can introduce significant errors in the emission predictions with the same mean speed, but with different levels of speed fluctuation (i.e. different drive cycles or driving patterns). For local assessments (e.g. street level) substantial errors up to factor of four in emission predictions have been reported for speed-time traces with the same mean speed but with different levels of speed fluctuation.

More detailed vehicle emission models in addition to VEPM will be needed in the future for adequate emission predictions at the local level in New Zealand. This study has developed and presented a possible approach to develop such a model, using remote sensing data. A proof-of-concept hybrid model was created to illustrate how the rich New Zealand remote sensing database can be used to develop a genuine New Zealand emissions model, which can be used to 'validate' VEPM, but also allows for emission predictions at a significantly higher resolution in space and time than VEPM.

The hybrid model links binned remote sensing data with CO_2 prediction algorithms through an appropriate vehicle classification scheme and use of a proxy variable for engine power (VSP).

A New Zealand hybrid model will provide an alternative and independent method to predict vehicle emissions in New Zealand, and predictions can be directly compared with VEPM, as a 'validation' exercise. Any differences between the models will provide policy makers with information regarding the possible range of air quality impacts due to motor vehicles. If predictions by the two model turn out to be quite similar, then this this will increase confidence in vehicle emission predictions and scenario modelling for the New Zealand fleet.

9 References

Barlow, T.J., 1998. The Inspection of In-Use Cars in order to Attain Minimum Emissions of Pollutants and Optimum Energy Efficiency – Remote Sensing, TRL, Crowthorne, UK.

Barth, M.; An, F.; Younglove, T.; Scora, R.; Levine, C.; Ross, M., Wenzel, T., 2000. Development of a Comprehensive Modal Emissions Model, National Corporate Highway Research Project. Transportation Research Board.

Beevers, S.D., Westmoreland, E., De Jong, M.C., Williams, M.L., Carslaw, D.C., 2012. Trends in NO_x and NO_2 emissions from road traffic in Great Britain, Atmospheric Environment, 54, 107-116.

Bishop, G.A., 2011. FEAT Equations for CO, HC and NO, http://www.feat.biochem.du.edu/assets/reports/FEAT_Math_II.pdf

Bishop, G.A., Starkey, J.R., Ihlenfeldt, A., Williams, W.J., Stedman, D.H., 1989. IR long-path photometry: a remote sensing tool for automobile emissions, Analytical Chemistry, 61, 671-677.

Bishop, G.A., Stedman, D.H., 1996. Measuring the emissions of passing cars, Acc. Chem. Res., 29, 489-495.

Bishop, G.A., Stedman, D.H., Hutton, R.B., Bohren, L., Lacey, N., 2000. Drive-by motor vehicle emissions: immediate feedback in reducing air pollution, Environ. Sci. Techno., 34, 1110-1116.

Bishop, G.A., Morris, J.A., Stedman, D.H., Cohen, L.H., Countess, R.J., Countess, S.J., Maly, P., Scherer, S., 2001. The effect of altitude on heavy-duty diesel truck on-road emissions, Environmental Science and Technology, 35, 1574-1578.

Bishop, G.A., Stedman, D.H., 2008. A decade of on-road emissions measurements, Environ. Sci. Technol., 42, 1651-1656.

Bishop, G.A., Schuchmann, B.G., Stedman, D.H., Lawson, D.R., 2012. Multispecies remote sensing measurements of vehicle emissions on Sherman Way in Van Nuys, California, J. Air & Waste Manage. Assoc., 62 (10), 1127–1133.

Bluett, J., Fisher, G., 2004. What is being discharged from the tailpipe? Modelling versus measurements, 27th Australasian Transport Research Forum (ATRF), Adelaide, 29 September – 1 October 2004.

Bluett, J., Fisher, G., 2005. Validation of a vehicle emission model using on-road emission measurements, Proceedings of the 17th International Clean Air and Environment Conference, Hobart, 2-5 May 2005, Published by the Clean Air Society of Australia and New Zealand, ISBN 0957850395.

Bluett, J., Kuschel, G., Unwin, M., 2010. Remote Sensing of Vehicle Emissions: Light and Heavy Duty Diesel Vehicles, Auckland Regional Council Technical Report 2010/063.

Borken-Kleefeld, J., 2013. Guidance note about on-road vehicle emissions remote sensing, July 2013, http://www.theicct.org/sites/default/files/publications/ RSD_Guidance_BorKlee.pdf.

Borken-Kleefeld, J., Kupianen, K., 2012. Teleconference, 5 September 2012, IIASA.

Borken-Kleefeld, J., Chen, Y., 2015. New emission deterioration rates for gasoline cars – results from long term measurements, Atmospheric Environment, 101, 58-64.

Brown, S., Bryett, C., Mowle, M., 1999. In-Service Emissions Performance - Drive Cycles, National Environment Protection Council Service Corporation (NEPC), Australia, ISBN 0 642 32323 2.

Burgard, D.A., Bishop, G.A., Stedman, D.H., Gessner, V.H., Daeschlein, C., 2006. Remote sensing of in-use heavy-duty diesel trucks, Environmental Science & Technology, 40 (22), 6938–6942.

Carslaw, D., Beevers, S., Westermoreland, E. Williams, M., Tate, J., Murrells, T., Stedman, J., Li, J., Grice, S., Kent, A., Tsagatakis, I., 2011. Trends in NO_x and NO_2 emissions and ambient measurements in the UK, report prepared for Defra, 3rd March 2011a.

Carslaw, D.C., Beevers, S.D., Tate, J.E., Westmoreland, E.J., Williams, M.L., 2011b. Recent evidence concerning higher NO_x emissions from passenger cars and light duty vehicles, Atmospheric Environment, 45 (39), 7053-7063.

Carslaw, D.C., Rhys-Tyler, G., 2013. New insights from comprehensive on-road measurements of NO_x, NO_2 and NH_3 from vehicle emission remote sensing in London, UK, Atmospheric Environment, 81, 339-347.

Carslaw, D.C., Williams, M.L., Tate, J.E., Beevers, S.D., 2013. The importance of high vehicle power for passenger car emissions, Atmospheric Environment, 68, 8-16.

Chan, T.L., Ning, Z., 2005. On-road remote sensing of diesel vehicle emission measurement and emission factors estimation for Hong Kong, Atmospheric Environment, 39, 6843-6856.

Chen, Y., Borken-Kleefeld, J., 2014. Real-driving emissions from cars and light commercial vehicles - Results from 13 years remote sensing at Zurich/CH, Atmospheric Environment, 88, 157-164.

CRC, 2003. On-road remote sensing of automobile emissions in the Denver area: Year 4, January 2003. Prepared by Burgard, D.A., Bishop, G.A., Williams, M.J., Stedman, D.H., 2003, University of Denver, Denver, CO, July 2003.

CRC, 2006. Analysis of Remote Sensing Data to Determine Deterioration Rates for OBDII Equipped Vehicles, Coordinating Research Council (CRC), CRC Report No. E-23-8, September 2006.

CRC, 2007. On-Road Remote Sensing of Automobile Emissions in the Chicago Area: Year 7, September 2006, Coordinating Research Council (CRC), CRC Report No. E-23-9, February 2007.

De Saint-Laumer, J.Y., Cicchetti, E., Merle, P., Egger, J., Chaintreau, A., 2010. Quantification in gas chromatography: prediction of flame ionization detector response factors from combustion enthalpies and molecular structures, Anal. Chem., 82, 6457-6462.

DEWHA, 2008. Evaluating the Health Impacts of Ethanol Blend Petrol, Commonwealth of Australia, Department of Environment, Water, Heritage and the Arts (DEWHA), Final Report KW48/17/F3.3F, June 2008.

EC, 1999. Commission Directive 1999/100/EC: Adapting to technical progress Council Directive 80/1268/EEC relating to the carbon dioxide emissions and the fuel consumption of motor vehicles, L 334/36, 28.12.1999, http://ec.europa.eu/ enterprise/sectors/automotive/documents/directives/motor-vehicles/index_en.htm

EEA, 2007. EMEP/CORINAIR Emission Inventory Guidebook - Road Transport, 23 August 2007, B710-1, http://www.eea.europa.eu/publications/EMEPCORINAIR5.

Elder, S., Graham, M., Jones, K., Raine, R., 2011. Vehicle Emissions Remote Sensing Campaign: Control Vehicle Study, UniServices Report 30793.001, August 2011, University of Auckland, New Zealand.

Eisinger, D.S., 2005. Evaluating Inspection and Maintenance programs: a policy-making framework, J. Air & Waste Manage. Assoc., 55 (2), 147-162.

Eisinger, D.S., Wathern, P., 2008. Policy evolution and clean air: the case of US motor vehicle inspection and maintenance, Transpn. Res.-D, 13, 359-368.

Ekström, M., Sjödin, Å, Andréasson, K., 2004. Evaluation of the COPERT III emission model with on-road optical remote sensing measurements, Atmospheric Environment, 28, 6631-6641.

ESP, 2010. Smoke Factor Measurements with Remote Sensing Device Technology – Recommended Practice, Environmental Systems Products (ESP), November 2010.

Frey, H.C., Unal, A., Chen, J., Li, S., 2003. Evaluation and recommendation of a modal method for modelling vehicle emissions,12th International Emission Inventory Conference – “Emission Inventories - Applying New Technologies”, San Diego, April 29 - May 1, 2003.

Fujita, E.M., Campbell, D.E., Zielinska, B., Chow, J.C., Lindhjem, C.E., DenBleyker, A., Bishop, G.A., Schuchmann, B.E., Stedman, D.H., Lawson, D.R., 2012. Comparison of the MOVES2010a, MOBILE6.2, and EMFAC2007 mobile source emission models with on-road traffic tunnel and remote sensing measurements, J. Air & Waste Manage. Assoc., 62 (10), 1134-1149.

Gala, F. De la Fuente, F., 2012. TRAFFIC 3E: an online platform to manage real traffic emissions, 19th Transport and Air Pollution Conference, Thessaloniki, Greece, 26 - 27 November 2012.

Guenther, P.L., Bishop, G.A., Peterson, J.E., Stedman, D.H, 1994. Emissions from 200 000 vehicles: a remote sensing study, The Science of the Total Environment, 146/147, 297-302.

Guo, H., Zhang, Q., Shi, Y., Wang, D., 2007. On-road remote sensing measurements and fuel-based motor vehicle emissions inventory in Hangzhou, China, Atmospheric Environment, 41, 3095-3107.

Hausberger, S.; Rodler, J.; Sturm, P., Rexeis, M., 2003. Emission factors for heavy-duty vehicles and validation by tunnel measurements, Atmospheric Environment, 37, 5237-5245.

Hausberger, S., Rexeis, M., Zallinger, M., Luz, R., 2009. Emission Factors from the Model PHEM for the HBEFA Version 3. Graz University of Technology, Institute for Internal Combustion Engines and Thermodynamics, Report Nr. I-20/2009 Haus-Em 33/08/679, 07.12.2009.

INFRAS, 2004. HBEFA 2.1, Handbuch Emissionsfaktoren, Keller M., de Haan P, et. al., Dokumentation (in German), 18. Aug. 2004, www.hbefa.net.

Jiménez, J.L., Koplow, M.D., Nelson, D.D., Zahniser, M.S., Schmidt, S.E., 1999. Characterization of on-road vehicle NO emissions by a TIDLAS remote sensor, J. Air & Waste Manage. Assoc., 49, 463-470.

Jiménez, J.L., McCrae, G.J., Nelson, D.D., Zahniser, M.S., Kolb, C.E., 2000. Remote sensing of NO and NO_2 emissions from heavy-duty diesel trucks using tunable diode lasers, Environmental Science & Technology, 34 (12), 2380-2387.

Jones, K., Elder, S., Raine, R., Graham, M., 2011. Vehicle Emissions Prediction Model Version 4.0, University of Auckland, Project No. 20659.002.

Joumard, R., Jost, P., Hickman, J., 1995. Influence of instantaneous speed and acceleration on hot passenger car emissions and fuel consumption, SAE Technical Paper Series, Paper No. 950928, 207-214.

Kent, J.H., Post, K., Tomlin, J.A., 1982. Fuel consumption and emission modelling in traffic links, Proceedings of SAE-A/ARRB 2nd Conference on Traffic Energy and Emissions, Paper No. 82140, 10.1-10.20.

Kenworthy, J.R., Newman, P.W.G., Lyons, T.J., 1983. A Driving Cycle for Perth, Report to National Energy Research Development and Demonstration Program, NERDDP/EG/83/129.

Kittelson, D.B., 2006. Ultrafine Particle Emission & Control Strategies, South Coast Air Quality Management District Conference on Ultrafine Particles: The Science, Technology, and Policy Issues, Wilshire Grand Hotel, Los Angeles, April 30 – May 2, 2006.

Ko, Y.W., Cho, C.H., 2006. Characterization of large fleets of vehicle exhaust emissions in middle Taiwan by remote sensing, Science of The Total Environment, 354, 75-82.

Kousoulidou, M., Ntziachristos, L., Mellios, G., Samaras, Z., 2008. Road-transport emission projections to 2020 in European urban environments, Atmospheric Environment, 42, 7465–7475.

Kraan, T.C., Baarbe, H.L., Eijk, A.R.A, Stelwagen, U., Vonk, W.A., 2012. Consistency tests of remote sensing for vehicle exhaust emissions, 19th Transport and Air Pollution Conference, Thessaloniki, Greece, 26 - 27 November 2012.

Kuhns, H.D., Mazzoleni, C., Moosmüller, H., Nikolic, D., Keislar, R.E., Barber, P.W., Li, Z., Etyemezian, V., Watson, J.G., 2004. Remote sensing of PM, NO, CO and HC emission factors for on-road gasoline and diesel engine vehicles in Las Vegas, NV, Science of the Total Environment, 322, 123-137.

Kuschel, G., Bluett, J., Unwin, M., 2012. Trends in Light Duty Vehicle Emissions - 2003 to 2011, Auckland Council Technical Report TR2012/032.

Lau, J., Hung, W.T., Cheung, C.S., 2012. Observation of increases in emission from modern vehicles over time in Hong Kong using remote sensing, Environmental Pollution, 163, 14-23.

Lawson, D.R., Groblicki, P.J., Stedman, D.H., Bishop, G.A., Guenther, P.L., 1990. Emissions from in-use motor vehicles in Los Angeles: a pilot study of remote sensing and the inspection and maintenance programs, J. Air & Waste Manage. Assoc., 40 (8), 1096-1105.

Li, B., Chu, J., 2013. Remote sensing measurement and analysis for alcohol-gasoline vehicle, Information Technology Journal, 12 (20), 5547-5552.

Lyons, T.J., Kenworthy, J.R., Austin, P.I., Newman, P.W.G., 1986.The development of a driving cycle for fuel consumption and emissions evaluation, Transpn. Res.-A, 20A (6), 447-462.

Mazzoleni, C., Moosmüller, H., Kuhns, H. D., Keislar, R.E., Barber, P.W., Nikolic, D., Nussbaum, N.J., Watson, J.G., 2004a. Correlation between automotive CO, HC, NO, and PM emission factors from on-road remote sensing: Implications for Inspection and Maintenance Programs, Transportation Research Part D: Transport and Environment, 9 (6), 477-496.

Mazzoleni, C., Kuhns, H.D., Moosmüller, H., Keislar, R.E., Barber, P.W., Robinson,N.F., Watson, J.G., Nikolic, D., 2004b. On-road vehicle particulate matter and gaseous emission distributions in Las Vegas, Nevada, compared with other areas, J. Air & Waste Manage. Assoc., 54 (6), 711–726.

Mellios, G., Hausberger, S., Keller, M., Samaras, C., Ntziachristos, L., Dilara, P., Fontaras, G., 2011. Parameterisation of fuel consumption and CO_2 emissions of passenger cars and light commercial vehicles for modelling purposes, JRC Scientific and Technical reports, European Union.

Mellios, G., Smit, R., Ntziachristos, L., 2013. Evaporative emissions: developing Australian emission algorithms, Proceedings of the CASANZ Conference, Sydney, 7-11 September 2013.

Monateri, A. M., Stedman, D. H., Bishop, G. A., 2004. Infrared thermal imaging of automobiles, Poster presented at the 14th CRC On-Road Vehicle Emissions Workshop, San Diego, March 2004, http://www.feat.biochem.du.edu/reports.html.

Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W., 1996, Applied Linear Statistical Models, 4th Ed., McGraw-Hill, Chicago, IL, ISBN 0 256 11736 5.

Ning, Z., Wubulihairen, M., Yang, F., 2012. PM, NO_x and butane emissions from on-road vehicle fleets in Hong Kong and their implications on emission control policy, Atmospheric Environment, 61, 265-274.

NIWA, 2008a. Assessing Vehicle Air Pollution Emissions, Prepared by J. Bluett, K. Dey and G. Fisher, Report CHC-2008-001.

NIWA, 2008b. On-Road Vehicle Emissions Monitoring - Sydney, NIWA Client Report: CHC2008-023, March 2008, NIWA Project: NASU07905, Jeff Bluett, Prepared for Department of Environment and Climate Change (New South Wales).

Ntziachristos, L., Samaras, C., Smit, R., Tooker, T., Mellios, G., 2013. Air pollutant and greenhouse gas road transport inventory using COPERT Australia, Proceedings of the CASANZ Conference, Sydney, 7-11 September 2013.

NZTA, 2014. The Comparison of Modelled and On-Road Light Duty vehicle Emission Factors, New Zealand Transport Agency (NZTA), prepared by J. Bluett and M. Unwin (Golder Associates), Report No. 1378104194_003_R_REV0, January 2014.

Orbital, 2005. NISE 2 – Contract 2: Drive Cycle and Short Test Development, Orbital Australia, for Department of the Environment and Heritage, Canberra, September 2005.

Orbital, 2009. Second National In-Service Emissions Study (NISE2) Light Duty Petrol Vehicle Emissions Testing, by Roads and Traffic Authority of NSW, RTA.07.2828.0309, March 2009.

Park, S.S., Kozawa, K., Fruin, S., Mara, S., Hsu, Y., Jakober, C., Winer, A., Herner, J., Emission factors for high-emitting vehicles based on on-road measurements of individual vehicle exhaust with a mobile measurement platform, J. Air & Waste Manage. Assoc., 61, 1046–1056.

Pokharel, S.S., Stedman, D.H., Bishop, G.A., 2000. RSD versus IM240 Fleet Average Correlations, 10th CRC On-Road Vehicle Emissions Workshop, March 2000, University of Denver, USA.

Pokharel, S.S., Bishop, G.A., Stedman, D.H., 2002. An on-road motor vehicle emissions inventory for Denver: an efficient alternative to modelling, Atmospheric Environment, 36, 5177-5184.

Popp, P.J., Bishop, G.A., Stedman, D.H., 1999. Development of a high-speed ultraviolet spectrometer for remote sensing of mobile source nitric oxide emissions, J. Air & Waste Manage. Assoc., 49, 1463-1468.

Rexeis, M., TU Graz, Austria, pers.comm., 18 November 2014.

Rhys-Tyler, G.A., Bell, M.C., 2012. Toward reconciling instantaneous roadside measurements of light-duty vehicle exhaust emissions with type approval drive cycles, Environ. Sci. Technol., 46, 10532-10538.

RTA 2009, Second National In-Service Emissions Study (NISE2) Light Duty Petrol Vehicle Emissions Testing, Roads and Traffic Authority of NSW, RTA.07.2828.0309, March 2009.

Safi, H., Assemi, B., Mesbah, M., Ferreira, L., Hickman, M., 2015. Design and implementation of a smartphone-based system for personal travel survey: case study from New Zealand, Transportation Research Board 94th Annual Meeting, Washington DC, US.

Singer, B.C., Harley, R.A., Littlejohn, D., Ho, J., Vo, T., 1998. Scaling of infrared sensor hydrocarbon measurements for motor vehicle emission inventory calculations, Environmental Science & Technology, 32 (21), 3241-3248.

Sjödin, Å, Lenner, M., 1995. On-road measurements of single vehicle pollutant emissions, speed and acceleration for large fleets of vehicles in different traffic environments, The Science of the Total Environment, 169, 157-165.

Sjödin, Å, Andréasson, K., Wallin, M., Lenner, M., Wilhelmsson, H., 1997. Identification of high-emitting catalyst cars on the road by means of remote sensing, Int. J. of Vehicle Design, 18 (3/4), 326-339.

Sjödin, Å, Andréasson, K., 2000. Multi-year remote sensing measurements of gasoline light-duty vehicle emissions on a freeway ramp, Atmospheric Environment, 34, 4657-4665.

Sjödin, Å, Jerksjö, M., 2008. Evaluation of European Road transport emission models against on-road emission data as measured by optical remote sensing, 17th International Transport and Air Pollution Conference, Graz.

Smit, R., 2008a. Errors in model predictions of NO_x traffic emissions at road level – Impacts of input data quality, in Air Pollution XVI, 116, eds. Brebbia, C.A., Longhurst, J.W.S., WIT Press, ISBN 978 1 84564 127 6, 255-269.

Smit, R., 2008b. Developments in road traffic emission and fuel consumption modelling: some recent experiences from Europe, Clean Air and Environmental Quality, 42 (3), 18-21.

Smit, R., Brown, A.L., Chan, Y.C., 2008. Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow?, Environmental Modelling & Software, 23 (10), 1262-1270.

Smit, R., McBroom, J., 2009. Development of a new high resolution traffic emissions and fuel consumption model, Road and Transport Research, 18 (4), 3-13.

Smit, R., Dia, H., Morawska, L., 2009. Road Traffic Emission and Fuel Consumption Modelling: Trends, New Developments and Future Challenges, in "Traffic Related Air Pollution and Internal Combustion Engines", Demidov, S. and Bonnet, J. (Eds.), Nova Publishers, U.S.A..

Smit, R., Ntziachristos, L, Boulter, P., 2010. Validation of road vehicle and traffic emission models — a review and meta-analysis, Atmospheric Environment, 44 (25), 2943-2953.

Smit, R., Bluett, J., 2011. A new method to compare vehicle emissions measured by remote sensing and laboratory testing: High-emitters and potential implications for emission inventories, Science of the Total Environment, 409, 2626–2634.

Smit, R., 2012. Review Notes for NIWA, 6-7 December 2012.

Smit, R., Ntziachristos, L., 2013. Cold start emission modelling for the Australian petrol fleet, Air Quality and Climate Change, 47 (3).

Smit, R., Ntziachristos, L., 2012. COPERT Australia: Developing Improved Average Speed Vehicle Emission Algorithms for the Australian Fleet, 19th International Transport and Air Pollution Conference, Thessaloniki, Greece, 26-27 November 2012.

Smit, R., 2013a. A procedure to verify large modal vehicle emissions databases, Proceedings of the CASANZ Conference, Sydney, 7-11 September 2013.

Smit, R., 2013b. Development and performance of a new vehicle emissions and fuel consumption software (PΔP) with a high resolution in time and space, Atmospheric Pollution Research, 4, 336-345.

Smit, R., 2014. P∆P: A simulation tool for vehicle emissions and fuel consumption software with a high resolution in time and space, Vehicle Technology Engineer, SAE Australasia, July 2014, 17-21.

St. Denis, M.J., Winer, A.M., 1994. Prediction of on-road emissions and comparison of modelled on-road emissions to the FTP emissions, The Emission Inventory: Perception and Reality, J. Air & Waste Manage. Assoc., 495-505.

St.Denis, M., Roeschen, J. 2012. Evaporative emissions reductions using remote sensing, Presented at the I/M Solutions, 20 May 2012, Sacramento, USA. http://obdclearinghouse.com/index.php?body=get_file&id=1566.

Stephens, R.D., Cadle, S.H., 1991. Remote sensing measurements of carbon monoxide emissions from on-road vehicles, J. Air & Waste Manage. Assoc., 41 (1), 39-46.

Stephens, R.D., Mulawa, P.A., Giles, M.T., Kennedy, K.G., Groblicki, P.J., Cadle, S.H., Knapp, K.T., 1996. An experimental evaluation of remote sensing-based hydrocarbon measurements: a comparison to FID measurements, J. Air & Waste Manage. Assoc., 46, 148-158.

Supnithadnaporn, A., Noonan, D.S., Samoylov, A., Rodgers, M.O., 2011. Estimated validity and reliability of on-board diagnostics for older vehicles: comparison with remote sensing observations, J. Air & Waste Manage. Assoc., 61, 996-1004.

Tong, H.Y., Karasek, F.W., 1984. Flame ionization detector response factors for compound classes in quantitative analysis of complex organic mixtures, Anal. Chem., 56, 2124-218.

Wang, S., Zhou, B., Wang, Z., Yang, S., Hao, N., Valks, P., Trautmann, T., Chen, L., 2012. Remote sensing of NO_2 emission from the central urban area of Shanghai (China) using the mobile DOAS technique, Journal Of Geophysical Research, 117, 1-14.

Watson, H.C., Milkins, E.E., Braunsteins, J., 1982. The development of the Melbourne Peak Cycle, Proceedings of SAE-A/ARRB 2nd Conference on Traffic Energy and Emissions, Paper No. 82148, 18.1-18.21.

Watson, H.C., 1995. Effects of a wide range of drive cycles on the emissions from vehicles of three levels of technology, SAE International, Warrendale, USA, paper no. 950221.

Watson, H.C., Trayford, R., 1999. A comparison between in-service experience and deterministic modelling of fleet fuel consumption, SAE International, Warrendale, USA, paper no. 99088.

Weilenmann, M., Soltic, P., Atjay, D., 2003. Describing and compensating gas transport dynamics for accurate instantaneous emission measurement, Atmospheric Environment, 37, 5137-5145.

Yam, Y.S., 2012. Emission control of in-use petrol and LPG vehicles in Hong Kong using remote sensing and transient emission testing, Presented at the Motor Vehicle Emissions Control Workshop, 5 December 2012, Hong Kong Polytechnic University, Hong Kong, http://www.cse.polyu.edu.hk/~activi/MoVE2012/ Presentation%20materials/Session%203.1.pdf.

Yanowitz, J., McCormick, R.L., Graboski, M.S., 2000. In-use emissions from heavy-duty diesel vehicles, Environmental Science & Technology, 34 (5), 729-740.

Ye, Y., Galbally, I.E., Weeks, I.A., 1997. Emission of 1,3-butadiene from petrol driven motor vehicles, Atmospheric Environment, 31, 8, 1157-1165.

Zhang, Y.Z., Stedman, D.H., Bishop, G.A., Guenther, P.L., Beaton, S.P., Peterson, J.E., 1993. On-road hydrocarbon remote sensing in the Denver area, Environ. Sci. Technol., 27, 1885-1891.

Zhang, Y.Z., Stedman, D.H., Bishop, G.A., Guenther, P.L., Beaton, S.P., 1995. Worldwide on-road vehicle exhaust emissions study by remote sensing, Environ. Sci. Technol., 29 (9), 2286-2294.

Appendix A Remote sensing validation studies

Reference	Country	Site Description	No. of Sites	Specific Assumptions	Equipment	Time Period	Pollutants Measured	Sample size	Model Validation	Comparison Variable	Comments Model
Sjodin et al 1997	Sweden	Main Road, straight, slight inclination, 70 km/h speed limit, constant cruise speed at 45 km/h	1	Cold start not relevant due to location of site	RSD 1000, pressure-induction tubes (10 m apart) for speed and acceleration	8 weekdays	CO, HC, CO_2	8,470 valid readings	-	-	-
Pokharel et al, 2002	USA	Denver area	8	Start and Evap Omitted	FEAT	-	CO, HC, NO	-	MOBILE 6	gram pollutant / kg of fuel	This study compares fuel-based and travel model based estimates, but leaves out start and evaporative emissions
Ekström et al, 2004	Sweden	One-way streets, slightly uphill	3	Cold start not relevant due to location of site, NO equals NOx	AccuScan RSD 3000	Several months	CO, HC, NO_x	~ 20,000 valid readings	COPERT 3	gram pollutant / kg of fuel	Presumably "Hot Running" only (not specified in paper), for COPERT3 average speed is equivalent to average speed RSD for each location, for ARTEMIS/HBEFA assumptions made on applicable traffic situation
Bluett and Fisher, 2004; 2005	New Zealand	Auckland	16	-	FEAT	2003	CO, HC, NO_x, particles	~ 35,000 vehicles	NZ Transport Emission Rate database	g/km	Fuel consumption estimates based on Australian fuel consumption guide (2004 paper) or NZTER data (2005 paper)
Kuhns et al, 2004	USA	Variable, speeds ranging from 44-72 km/h, road grade ranging from -0.7 to +3.7 degrees	10		LIDAR (PM), VERSS (RSD 3000), laser beams for speed and acceleration, video camera for license plate records	~ month	CO, HC, CO_2, NO, PM	~ 60,000 vehicles	MOBILE6 and PART5	gram pollutant / kg of fuel	Hot running only, (evaporative and start emissions excluded), fleet composition based on registered census data, selection of one traffic situation "Freeway at 64 km/h" as being closest to average speed in RSD database, conversion of g/VKT to g/kg fuel using fuel economy factors
Mazzoleni et al., 2004b	USA	Single lane sites in Las Vegas metropolitan areas	5	Start, Non-Exhaust and Evap Omitted	DRI VERSS	24 days	CO, HC, NO, PM	40,245/15,220 valid measurements for CO/HC/NO and PM respectively	PART5	gram pollutant / kg of fuel	Hot running only, non-exhaust PM and evaporative not considered.
Guo et al, 2007	China	Variable, speeds ranging from 8-95 km/h, road grade ranging from 0 to +2.2 degrees	5	Cold start negligible	INSPECTOR IV, 3 laser beams for speed and acceleration, video camera for license plate records, measurement of meteorology	1-2 months	CO, HC, CO_2, NO	~ 47,000 valid readings	IVE	ton/year	Fuel-based emission inventory with remote sensing compared with VKT based emission inventory with IVE
NIWA, 2008b	Australia	Various locations in	15	-	RSD 4000	May 2006 – April	CO, HC, CO_2, NO, smoke	~53,000 vehicles	Various	g/km	-

		three major cities				2007					
Sjödin and Jerksjö, 2008	Sweden	Variable, city and freeway with speed limits of 30 or 70 km/h, average speeds ranging from 29-58 km/h	4	Cold start negligible	FEAT		NO, NO_2, NO_x, HC, CO_2, NH_3,	15,590 vehicles	COPERT 4, HBEFA 2.1, ARTEMIS 0.4d	gram pollutant / kg of fuel	Presumably “Hot Running” only (not specified in paper), for COPERT4 average speed is equivalent to average speed RSD for each location, for ARTEMIS/HBEFA assumptions made on applicable traffic situation
Smit and Bluett, 2011	Australia	High capacity urban roads	5	Humidity and ambient temperature not considered	-	-	CO, HC, NO_x	~ 10,000 valid readings	Laboratory NISE2	gram pollutant / kg of fuel	-
Carslaw et al., 2011a	UK	Various	-	-	FEAT	-	CO, HC, NO_x	72,000 valid readings	COPERT4, HBEFA 3.1	g/km	“Hot Running” only, for COPERT4 average speed is equivalent to average speed RSD for each location, for ARTEMIS/HBEFA assumptions made on applicable traffic situation
Carslaw et al., 2011b	UK	Typical urban conditions	-	-	AccuScan RSD-4600	2007-2010	NO, (NO_2 estimated)	-	UK NAEI, HBEFA 3.1	g/km	Results only presented for NO_x.
Beevers et al., 2012	UK	Single lane road with 30-50 km/h speed	1	Cold start emissions assumed to be zero	RSD4600	2007-2010	NO, (NO_2 estimated)	74,614 vehicles sampled	HBEFA, UK NAEI	Total emissions	Results only presented for NO_x.
Fujita et al., 2012	USA	Road leading to tunnel	1	Cold start emissions assumed to be zero	-	2010 (one week)	CO, NO_x, THC	13,000 measurements	MOVES, MOBILE, EMFAC	gram pollutant / kg of fuel	RSD data only measured for LDVs so weekend results were used with lower proportions of HDVs in the tunnel
Kraan et al., 2012	Netherlands	Typical urban roads	5	-	-	-	NO	1,950 valid readings	VERSIT+	g/km	Results only presented for NO.
Li and Chu, 2013	China	Typical urban roads, uphill gradient of 0-2º, average speed of 10-18 km/h	5	-	RSD3000	April 2011 – April 2012	CO, HC, CO_2, NO_x	8,537 valid readings	IVE	g/l	-
Chen and Borken-Kleefeld., 2014	Switzerland	AADT of 5000 vehicles, uphill gradient of 9º, average speed of 45 km/h	1	Cold start emissions assumed to be zero	-	2000-2012	NO (NO_2 estimated)	128,000 valid readings	PHEM, HBEFA	gram pollutant / kg of fuel	NO_2 not measured but estimated to compute NO_x emissions.

Appendix B Difference in v-a and VSP approach

Table 9-1: Mean differences in CO/CO_2 emission ratios between the v-a and VSP6 bin methods for each speed-acceleration bin.

Difference	<= -3.0	-3.0 to -2.5	-2.5 to -2.0	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to -0.1	-0.1 to +0.0	+0.0 to +0.1	+0.1 to +0.5	+0.5 to +1.0	+1.0 to +1.5	+1.5 to +2.0	+2.0 to +2.5	+2.5 to +3.0	>= +3
≥ 100	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
90-100	NA	NA	NA	NA	NA	NA	-9%	17%	15%	22%	NA	NA	NA	NA	NA	NA
80-90	NA	NA	NA	NA	NA	-40%	-10%	4%	-5%	-22%	-50%	NA	NA	NA	NA	NA
70-80	NA	NA	NA	NA	NA	-31%	-10%	17%	10%	64%	29%	NA	NA	NA	NA	NA
60-70	NA	NA	NA	NA	NA	50%	24%	25%	53%	110%	27%	NA	NA	NA	NA	NA
50-60	NA	NA	NA	NA	-61%	-13%	-28%	-35%	-20%	-9%	-1%	127%	NA	NA	NA	NA
40-50	NA	NA	NA	-59%	-40%	-8%	-17%	-18%	-25%	-10%	-5%	48%	96%	NA	NA	NA
30-40	NA	NA	NA	-12%	-41%	6%	-7%	-31%	-42%	-23%	-9%	5%	57%	NA	NA	NA
20-30	NA	NA	NA	10%	48%	28%	13%	-20%	-26%	-11%	-8%	-10%	4%	15%	NA	NA
10-20	NA	NA	NA	20%	47%	23%	57%	-44%	-66%	-44%	-18%	-15%	7%	26%	NA	NA
0-10	NA	NA	NA	126%	25%	13%	19%	27%	88%	12%	44%	5%	13%	16%	NA	NA

Table 9-2: Mean differences in HC/CO_2 emission ratios between the v-a and VSP6 bin methods for each speed-acceleration bin.

Difference	<= -3.0	-3.0 to -2.5	-2.5 to -2.0	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to -0.1	-0.1 to +0.0	+0.0 to +0.1	+0.1 to +0.5	+0.5 to +1.0	+1.0 to +1.5	+1.5 to +2.0	+2.0 to +2.5	+2.5 to +3.0	>= +3
≥ 100	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
90-100	NA	NA	NA	NA	NA	NA	-11%	5%	22%	11%	NA	NA	NA	NA	NA	NA
80-90	NA	NA	NA	NA	NA	5%	29%	37%	23%	5%	-45%	NA	NA	NA	NA	NA
70-80	NA	NA	NA	NA	NA	6%	25%	54%	56%	47%	9%	NA	NA	NA	NA	NA
60-70	NA	NA	NA	NA	NA	35%	14%	42%	52%	54%	14%	NA	NA	NA	NA	NA
50-60	NA	NA	NA	NA	-37%	-3%	3%	12%	3%	12%	-7%	37%	NA	NA	NA	NA
40-50	NA	NA	NA	-16%	-13%	0%	-5%	0%	6%	2%	-2%	-6%	-10%	NA	NA	NA
30-40	NA	NA	NA	-2%	-19%	2%	-4%	-16%	-29%	-20%	-7%	-11%	-20%	NA	NA	NA
20-30	NA	NA	NA	16%	22%	1%	-1%	-12%	-30%	-23%	-30%	-45%	-29%	10%	NA	NA
10-20	NA	NA	NA	20%	18%	-10%	-19%	-38%	-56%	-52%	-55%	-47%	-41%	-60%	NA	NA
0-10	NA	NA	NA	39%	11%	4%	10%	2%	-4%	-8%	-9%	-26%	-36%	27%	NA	NA

Table 9-3: Mean differences in NO_x/CO_2 emission ratios between the v-a and VSP6 bin methods for each speed-acceleration bin.

Difference	<= -3.0	-3.0 to -2.5	-2.5 to -2.0	-2.0 to -1.5	-1.5 to -1.0	-1.0 to -0.5	-0.5 to -0.1	-0.1 to +0.0	+0.0 to +0.1	+0.1 to +0.5	+0.5 to +1.0	+1.0 to +1.5	+1.5 to +2.0	+2.0 to +2.5	+2.5 to +3.0	>= +3
≥ 100	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
90-100	NA	NA	NA	NA	NA	NA	-57%	-37%	-38%	-22%	NA	NA	NA	NA	NA	NA
80-90	NA	NA	NA	NA	NA	-61%	-50%	-17%	-14%	-17%	12%	NA	NA	NA	NA	NA
70-80	NA	NA	NA	NA	NA	-55%	-44%	-18%	25%	21%	50%	NA	NA	NA	NA	NA
60-70	NA	NA	NA	NA	NA	-44%	-36%	-20%	10%	10%	12%	NA	NA	NA	NA	NA
50-60	NA	NA	NA	NA	-59%	-26%	-26%	-37%	-16%	12%	5%	-8%	NA	NA	NA	NA
40-50	NA	NA	NA	-22%	-22%	-4%	-11%	-20%	-31%	1%	2%	0%	-8%	NA	NA	NA
30-40	NA	NA	NA	89%	-4%	0%	16%	-15%	-37%	-5%	20%	4%	-22%	NA	NA	NA
20-30	NA	NA	NA	80%	69%	59%	41%	11%	-24%	3%	9%	-12%	-8%	72%	NA	NA
10-20	NA	NA	NA	124%	97%	109%	59%	16%	-23%	13%	88%	94%	223%	614%	NA	NA
0-10	NA	NA	NA	138%	99%	111%	154%	58%	-2%	96%	42%	45%	405%	1505%	NA	NA

Appendix C Steps in data processing R script

The scripts can be found in the file ARC11504_data_sets_V2.r and sequentially:

1. Read the tables in the Access database into R
2. Create a 'Vehicles' dataframe based on a desired subset of valid data
   - Convert odometer readings in miles to kilometres
   - Remove invalid records where:
     - Number plate is "NA"
     - Weight is below700kg (and above 3 500 if looking at LDVs)
     - Engine capacity is below 500cc and above 10 000cc
   - Consolidate gender of owner
   - Consolidate fuel types down to "Petrol", "Diesel", "Other"
   - Bin year of manufacturer
   - Create Vehicle age based on date of record and year of manufacturer: bin results.
   - Bin Odometer readings
   - Relabel "NA" in previous country of registration as "NZL": Consolidate to "NZN", "JPN", "Other"
   - Determine test regime for Japanese Vehicles where it was not supplied
   - Tidy up test regime for NZ vehicles
   - Check for duplicate records and set them to "NA"
   - Write a text file of resulting QA'd vehicles if required
3. Create an 'Emissions' dataframe
   - Check emissions have a valid plate number
   - Check 2005-2011 data have valid flags for CO, CO2, HC and NO
   - Check 2005-2011 data have valid flags for speed and acceleration
   - Check 2003 data have valid flags for CO, CO2, HC and NO. For the sites MAN1 and MAN2 keep records flagged as invalid for NO, as it was not recorded at these sites.
   - Merge emissions data with emissions logs to obtain site details
4. Merge vehicles and emissions data frames
5. Add variable to record differences between emissions date and odometer date
6. Check for and eliminate duplicate vehicles by choosing the reading closest to the odometer date
7. Add speed and acceleration
8. Add VSP and only keep records for where VSP is between 0 and 40: bin
9. Remove extreme outliers for HC > 30 000, NO > 9 000 and uvSmoke < -5 & > 5