This is a repository copy of *Can accurate distance-specific emissions of nitrogen oxide emissions from cars be determined using remote sensing without measuring exhaust flowrate?*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/196533/

Version: Published Version

Article:

Can accurate distance-specific emissions of nitrogen oxide emissions from cars be determined using remote sensing without measuring exhaust flowrate?

Yoann Bernard [a,*], Jan Dornoff [a], David C. Carslaw [b]

[a] *The International Council on Clean Transportation (ICCT), Neue Promenade 6, 10178 Berlin, Germany*
[b] *Wolfson Atmospheric Chemistry Laboratories, University of York, York YO10 5DD, United Kingdom*

HIGHLIGHTS

- NO_X (g/km) can be estimated using RSD without measuring the exhaust flowrate.
- The conversion to distance-specific emissions does not introduce a systematic bias.
- At least 30 remote sensing measurements from a given vehicle's trip are recommended.
- Estimates do not significantly improve with more than 300 remote-sensing samples.
- The method may be used to screen car fleets and identify vehicles grossly emitting.

GRAPHICAL ABSTRACT
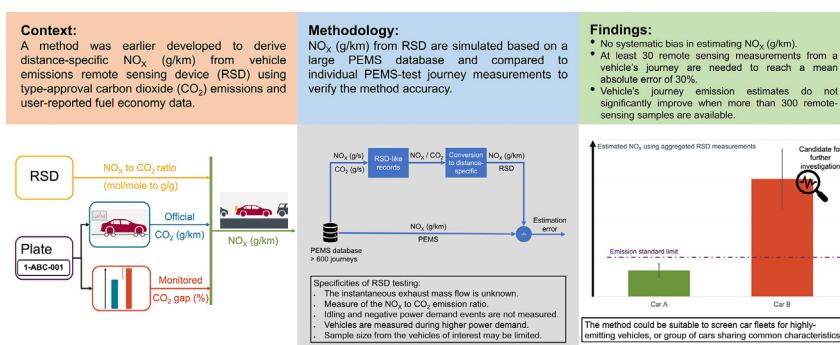
ABSTRACT

Portable Emission Measurement Systems (PEMS) are commonly used to measure absolute (mass per unit distance) emissions of a range of pollutants from road vehicles under real driving conditions. Because measuring large numbers of vehicles with PEMS is impractical, this paper investigates how vehicle emission remote sensing device (RSD) can supplement the use of PEMS. We simulate whether remote sensing measurements can accurately predict a vehicle's real-world distance-specific nitrogen oxides (NO_X) emissions using RSD without measuring its exhaust flow rate. The approach uses readily available type-approval carbon dioxide (CO_2) emission data together with average real-world divergences from studies based on user-reported fuel economy data. We find that at least 30 RS measurements from a given vehicle's journey are needed to reach a mean absolute error of 30% compared to a large reference data set of individual PEMS measurements. With that condition met, it is concluded that estimates agree well with actual NO_X emissions from cars and the applied method does not introduce a systematic bias. It is also found that the accuracy of estimates for distance-specific NO_X emissions does not significantly improve when more than 300 remote-sensing samples are available, with a mean absolute error converging to 23%. We conclude that this method could be used to screen large car fleets and identify vehicles or group of vehicles that are likely grossly exceeding air pollution standards.

* Corresponding author.
E-mail address: y.bernard@theicct.org (Y. Bernard).

1. Introduction

Nitrogen-oxides (NO_X) emissions from vehicles significantly contribute to air pollution responsible for a range of human health problems. Emissions of NO_X directly contribute to elevated ambient pollutant levels for nitrogen dioxide (NO_2) while enhancing the formation of ground-level ozone (O_3) and fine particulate matter ($PM_{2.5}$) (EEA, 2019). In Europe, NO_X-related air pollution from on-road diesel vehicles is responsible for about 28,400 premature deaths each year (Anenberg et al., 2017).

Almost three decades ago, emissions standards were introduced in Europe to limit vehicle emissions. However, studies have documented that the amount of NO_X measured in diesel vehicle exhaust in "real-world" operation, i.e. on the road, during typical driving conditions exceeded type-approval levels by many times (Franco et al., 2014). Until incorporating on-road testing in the type-approval procedure in September 2017, type-approval of passenger cars relied solely on tests under laboratory conditions. Taking advantage of the pre-defined conditions, combined with little enforcement actions in the European Union (EU), some auto manufacturers fitted their vehicles with "defeat devices" that resulted in laboratory compliance with air pollution standards, but multiple times exceedances on the road (Bernard et al., 2019).

In response, the European Commission devised the Real-Driving Emissions (RDE) procedure in an attempt to cap on-road NO_X and particulate number emissions. Following the European Commission's approach of stepwise introduction, RDE became first mandatory for the type approval of new passenger car models in September 2017. The introduction was completed in September 2020 since when RDE is mandatory for all new passenger cars and light commercial vehicles registered in the EU. The procedure entails fitting the vehicle with a Portable Emission Measurement System (PEMS) and driving for a predetermined time and distance (e.g. between 90 and 120 min) (Mock, 2017).

Since September 2020, the latest European Union type-approval framework regulation requires member states and the European Commission to conduct independent market surveillance activities to ensure that vehicles in use meet the emission limits (Regulation (EU) 2018/858 of the European Parliament and of the Council, 2018). However, vehicle sourcing, instrument calibration and the time required for driving collectively make PEMS testing costly and thus limited in regards to the number of vehicles that can be tested per day (Dallmann, 2018). In addition, the vehicle may detect the presence of a PEMS and based thereon switch to a mode producing lower tailpipe emissions. While this would constitute a defeat-device, the measurements could therefore still not reflect real-world emissions (Bernard et al., 2019).

The remote-sensing device (RSD) is an alternative means of measuring real-world vehicle emissions that is better suited for wide-ranging data collection in the field (Beaton et al., 1995). RSD and PEMS were found to show good correlations based on NO_X to CO_2 emission ratios from individual vehicles, and RSD therefore presents a promising technology for the screening of vehicle emissions (Gruening et al., 2019). In addition, the European Union recognizes remote sensing as one of the valid tools for providing information to type-approval authorities for the systematic selection of test candidate vehicles with potentially excessive emissions levels (Commission Regulation (EU) 2018/1832, 2018). RSD is set-up in public areas and is either situated by the side of the road, scanning across the road, or positioned somewhere above it. It does not require any contact with the vehicle and the vehicle cannot easily detect that it is being tested. Vehicles measured with RSD are typically not pre-selected, and present various maintenance states. Another key benefit of RSD is that it captures a whole range of driving and environmental conditions that affect vehicle emissions performance. For comparison, PEMS tested vehicles are checked against visible malfunctions prior to testing, and driven by professional drivers. The PEMS emission results of a given vehicle may also greatly vary depending on the selected route, the driving style, and the ambient condition encountered during the corresponding trip. Because RSD can measure emissions from many vehicles after being deployed, it is much less expensive per vehicle than PEMS.

However, RSDs have drawbacks. Whereas PEMS continuously measure NO_X emissions (and other species) on a given vehicle and trip, RSD can only provide a snapshot of a vehicle's emissions as it passes a given location. This means that it is necessary to obtain sufficient measurements of a specific vehicle or model-family/group-type to assess its emission levels. Using absorption spectroscopy, current RSDs only measure the emission ratios (e.g. NO_X) relative to CO_2 (a proxy for fuel consumption), i.e. it does not measure absolute concentration levels of each species (Burgard et al., 2006).

Besides, because RSD instruments are installed at a fixed location and measure vehicle emissions at one instant of its journey, results cannot be directly compared against regulatory emission limits that apply to the average emissions collected during entire drive cycles being up to about a hundred kilometers long. The emission limits are expressed in a distance-specific metric, i.e. grams of pollutant per kilometer. In this study we present a methodology to bridge these limitations. We verified the accuracy of the method using PEMS data as a benchmark.

2. Materials and methods

2.1. Estimation of distance-specific emission from remote-sensing

A simplified method has previously been developed to estimate average distance-specific NO_X emissions from RSD aggregated results (Bernard et al., 2018). Fig. 1 illustrates how this process works. Raw RSD NO_X-to-CO_2 emission ratios are converted to mass ratios using the respective molar masses for each gas, then averaged for a group of vehicles of interest. To convert this ratio to distance-specific NO_X emissions, the average distance specific real-world CO_2 emissions of the analyzed vehicles are needed. For this purpose, the vehicles' type-approval CO_2 emissions were adjusted upward according to real-word fuel-consumption figures, gathered from users to estimate actual, distance-specific CO_2 emissions (Tietge et al., 2017). Note that the results generated reflect the emissions level over an average journey, but not necessarily the instantaneous emissions rate that greatly fluctuate along the trip. Conversely, other experimented approaches have estimated the emissions rate for each RSD snapshots based on speed measurement and additional vehicles characteristics (e.g. mass, vehicle segment, etc.), and then simulated the distance-specific NO_X emissions on a specific journey (Davison et al., 2020).

For validation, we applied the simple method to a large dataset of Europe-wide remote-sensing measurements. The results were then compared with a separate PEMS dataset which showed a very close agreement for average distance-specific NO_X emissions of diesel Euro 5 and 6 vehicles (Bernard et al., 2018). However, a comparison at the individual vehicle and trip level was not feasible because the two data sets originated from different testing campaigns.

However, we show that the comparison between the remote-sensing prediction and PEMS measurement at the individual vehicle trip level is possible with the method described in the next section.

2.2. Method evaluation and source of PEMS dataset

A large PEMS dataset is used as a reference, containing both second-by-second emission mass rates and tailpipe concentrations. In a thought experiment, we simulate RSD measurements as if virtual RSD instruments were placed along-side the PEMS trip route. For varying numbers of NO_X-to-CO_2-ratio measurements, we then simulate how accurately the distance-specific NO_X emissions could be estimated compared to the PEMS trip result. This methodology is further broken down into steps, during which the associated uncertainty is analyzed, and seeks, in particular, to highlight the key differences between remote-sensing
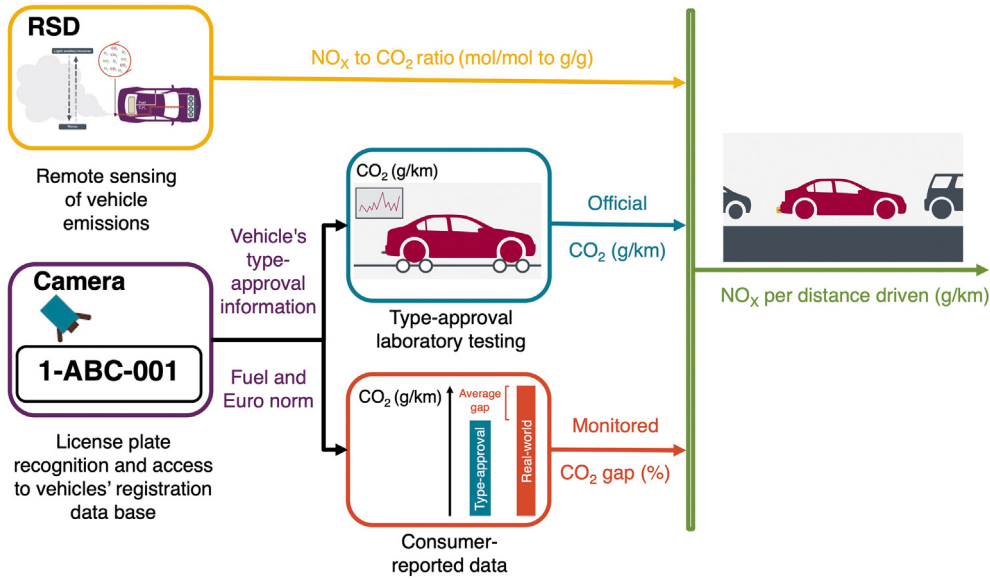
Fig. 1. Schematic of the estimation of distance-specific NO_X emissions based on remote-sensing measurements and representative, average CO_2 emissions.

estimates and measurements taken with PEMS. Finally, we reduce the simulated number of aggregated RSD measurements along the PEMS trip into fewer single measurements to quantify the effect on the estimation.

The PEMS measurements dataset consists of 672 on-road trips from 298 passenger cars (Table 1), each consisting of a mix of urban, rural and motorway driving. For each trip, emissions were sampled at a frequency of 1 Hz. A half of the data is from government's market surveillance investigations or non-governmental organizations (Baldino et al., 2017), and another half from data available on-request through the RDE monitoring phase. Two thirds (451) are vehicles using compression ignition (i.e. diesel) engines and one third (221) were positive ignition (e.g. gasoline, Liquid Petroleum Gas (LPG), Compressed Natural Gas (CNG) etc.). The majority of vehicles is certified with the Euro 6b standards (619), followed by 6d-TEMP (37), 6c (9) and Euro 5 (7). The total vehicle driven distance adds up to over 46,000 km.

2.3. Conversion of RSD-like snapshots into distance specific per-trip levels

Remote-sensing and PEMS measurement techniques share the same goal of measuring vehicle on-road emission performance but differ in many ways. PEMS equipment uses sensors mounted in, or at the back of the vehicle to measure tailpipe concentrations of specific compounds of interest (e.g. NO_X). An exhaust flow meter is used to calculate the mass flow rate of certain exhaust compounds while the GPS measures the vehicle speed.

For comparison with regulatory targets, NO_X and CO_2 emissions of Light-Duty Vehicles (LDV) need to be expressed in grams per kilometer driven. Distance-specific emissions are calculated as the integral of

emission mass flows divided by the cumulated distance covered in the journey from a 1 Hz data set (Eqs. 1 and 2).

$$dm_{NOX,pems} = \frac{\sum \dot{m}_{NOX}}{\sum v} \tag{1}$$

$$dm_{CO2,pems} = \frac{\sum \dot{m}_{CO2}}{\sum v} \tag{2}$$

In these equations, $\dot{m}$ is the instantaneous mass rate (e.g. gram per second), v represents the instantaneous vehicle speed (e.g. kilometer per second) and dm are the distance-specific mass emissions (gram per kilometer).

Remote sensing equipment is generally installed close to the road and analyzes about half a second-worth of the exhaust plume composition of a vehicle as it passes by the sensor. To characterize the emissions of a vehicle or a certain type of vehicle, it is necessary to gather several measurements. Due to the vehicle movement, remote-sensing can only measure the exhaust plume once it is some distance away from the tailpipe source and partially diluted in the ambient atmosphere. Consequently, remote-sensing systems can only report emission ratios (e.g. NO_X to CO_2).

In remote sensing measurements, congested traffic is typically not analyzed to maintain a clear gap between vehicles and avoid cross-contamination of the plumes. This means that idling emissions are not measured during remote sensing campaigns. Remote-sensing measurements tend to be made under some engine load to ensure a sufficient CO_2-concentration in the exhaust plume. For that purpose, a proxy of the power at the wheel is calculated using vehicle speed, acceleration and the road's gradient, known as "vehicle specific power" (VSP) and developed in former studies (Jiménez-Palacios, 1999). A positive VSP is typically a prerequisite for a valid RSD measurement, indicating the engine being under load (which is often the majority of an average journey) rather than coasting during deceleration. In practice, vehicle's decelerations are often excluded. For the same reason, the sites regularly selected have a slight uphill incline. That maximizes the chances of measuring vehicles under some engine load.

Therefore, vehicles tend to be measured at higher average VSP with remote-sensing than during an average PEMS trip, which is typically composed of urban, rural and motorway sections and includes idle and deceleration phases.

Table 1
Number of PEMS trips by combustion type, fuel type, and euro standard. LPG stands for liquid petroleum gas, CNG stands for compressed natural gas.

Combustion	Fuel	Euro standard			
		Euro 5	Euro 6b	Euro 6c	Euro 6d-TEMP
Compression ignition	Diesel	7	437	2	5
Positive ignition	Gasoline		175	7	29
	LPG		7		
	CNG				3
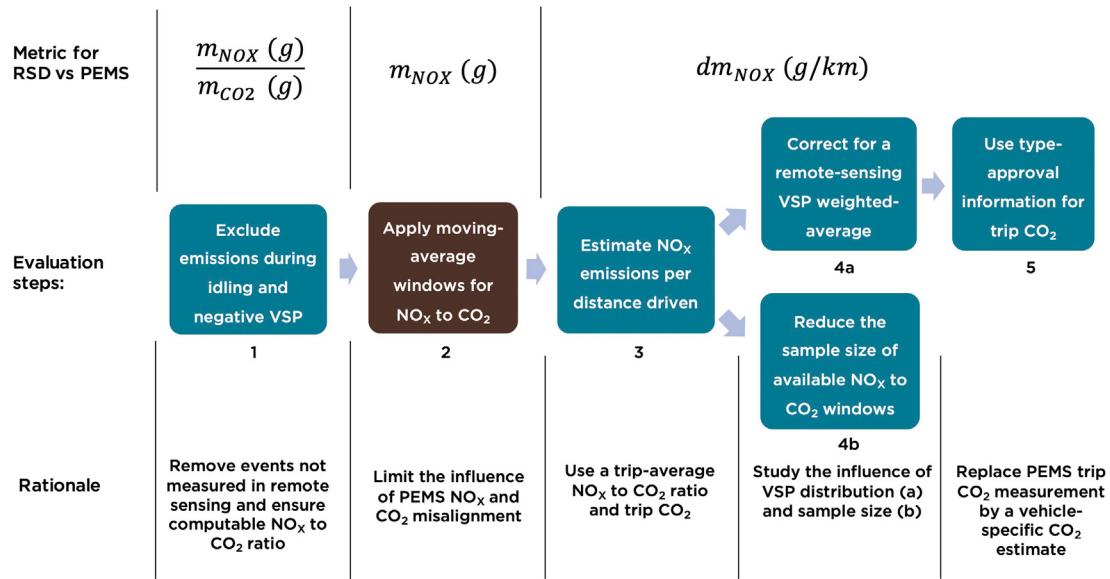
Fig. 2. Steps for evaluating uncertainty when estimating distance-specific nitrogen oxides (NO_X) emissions using remote-sensing. *m* stands for the trip-cumulated mass, *dm* for the trip distance-specific mass, CO_2 for carbon dioxide and VSP for vehicle-specific power.

The estimations of the virtual remote sensing measurements are calculated using the average NO_X to CO_2 emissions ratio during the trip, multiplied by distance-specific CO_2 emissions. To account for the described differences between remote-sensing and PEMS trip measurement, we adapted the method to establish this distance-specific CO_2 emission level. In steps 1 to 4 of the analysis (presented in Fig. 2), distance-specific CO_2 emissions are calculated from each PEMS trip-measurement. In step 5, we use the estimates of vehicle-specific real-world CO_2 emissions averages that are derived from their type-approval certification.

Potential sources of uncertainty within the methodology described in Fig. 2 are identified and detailed in each step below.

2.3.1. Step 1

Emissions at idle or during negative VSP are typically filtered out from remote sensing data as described previously. In this step we analyze the effect of excluding the PEMS trip data for these conditions on the accuracy of reproducing the trip NO_X to CO_2 mass ratio (Eq. 3).

$$\frac{m_{NOX,rsd}}{m_{CO2,rsd}} \approx \frac{m_{NOX,pems}}{m_{CO2,pems}} \quad (3)$$

where *m* is the trip-cumulated mass (e.g. gram) measured from the *pems* or recalculated for typical RSD VSPs.

2.3.2. Step 2

While an RSD measures the ratio of NO_X to CO_2 in the exhaust plume at a given time, a PEMS uses separate analyzers to measure NO_X and CO_2 which can require time alignment and thereby add inaccuracy in the calculated NO_X-to-CO_2 ratio. To simulate RSD measurements using PEMS data, it is therefore necessary to mitigate potential misalignment of NO_X and CO_2. For that purpose, a moving average-window method is adopted, discussed in more detail in the calculation section. We show the uncertainty introduced by this windowing process by comparing the accumulated NO_X mass flow determined by the PEMS with the moving-average NO_X mass flow calculated as the sum of NO_X-to-CO_2 emission ratio times instantaneous CO_2 mass flow for each window (Eq. 4).

$$m_{NOX,maw,rsd} = \sum (r_{maw} \times \dot{m}_{CO2,maw})_{rsd} \quad (4)$$

where *m* is the trip-cumulated mass, *r* is the NO_X-to-CO_2 emission ratio, *maw* indicates a calculation based on a moving average window, and *rsd* filtered for conditions compatible with RSD.

2.3.3. Step 3

We hypothesize that the ratio of trip NO_X mass to trip CO_2 mass can be estimated by the trip-mean of the instantaneous NO_X to CO_2 mass ratio measured by remote sensing (Eq. 5). We use that assumption and Eq. (2) to re-write Eq. (1) to estimate distance-specific NO_X emissions based on the average NO_X-to-CO_2 ratio of every window meeting remote sensing testing condition combined with distance-specific CO_2 emissions measured by PEMS (Eq. 6). For each trip, we compare the calculated NO_X emissions with those measured by PEMS. Note that this method does not calculate instantaneous NO_X emission flow rates (e.g. gram per second) for each measurement, unlike other methods proposed (Davison et al., 2020), or similar to how a PEMS would measure continuously. Instead, we aggregate multiple CO_2-specific remote-sensing measurements and scales a per-distance-driven unit.

$$\bar{r}_{maw,rsd} \approx \frac{dm_{NOX,pems}}{dm_{CO2,pems}} \quad (5)$$

$$dm_{NOX,rsd} \approx \bar{r}_{maw,rsd} \times dm_{CO2,pems} \quad (6)$$

where $\bar{r}$ is the trip-mean of the NO_X-to-CO_2 emission ratios, *maw* indicates a calculation based on a moving average window, and *rsd* means filtered for conditions compatible with RSD.

2.3.4. Step 4

a) As explained before, RSD measures vehicles at higher average VSP compared to PEMS testing. We investigate how this difference could create bias in the estimation of distance-specific NO_X emissions. In order to align the test conditions, the mean of the NO_X-to-CO_2 ratio in Eq. (6) was replaced by the weighted mean of the NO_X-to-CO_2 ratio using a VSP distribution commonly used for remote-sensing (detailed in the calculation section).

b) A PEMS measurement typically contains some thousand data points, while remote-sensing assesses a vehicles' performance with far fewer measurements. In this section, we investigate the influence of the number of windows (single data captures) considered on

the quality of the overall estimation. The sample sizes used range from 1 to 1000.

2.3.5. Step 5

For vehicles measured by RSD, technical information is obtained by screening the license plate, typically including the type-approval CO_2 values. We first investigated the impact of substituting the PEMS measured distance-specific CO_2 emissions (Step 3) for vehicle's real-world CO_2 emissions derived from the type-approval information (Eq. 7). The disparity between type-approval and real-world CO_2 emissions was calculated from a large dataset of consumer reporting information (Tietge et al., 2017). We perform the analysis on the subset of vehicles certified to Euro 6 emissions limits, for which the average real-world to NEDC CO_2 gap has peaked and plat around 2014–2016; coinciding with the early years of the Euro 6 standard and remained relatively constant since. The average gap was found to be +39% for diesel and +33% for gasoline Euro 6 passenger cars (Bernard et al., 2018). To account for differences between real-world and RDE testing, the equivalent average CO_2 emission gap between RDE-test and type-approval was determined for each fuel type using the PEMS database. Here, we show the effect of the CO_2 emission value on the NO_X estimate accuracy. It was however limited to vehicles for which the NEDC type-approval CO_2 values were available, due to the anticipated reduction of disparity under the new WLTP laboratory test.

$$dm_{NOX,rsd} \approx \overline{r}_{rsd,maw} \times dm_{CO2,type\ approval} \times (1 + gap/100) \quad (7)$$

where *gap* is the real-world to type-approval CO_2 divergence in percent.

2.4. Calculation

The PEMS database consists of measurements taken at 1 Hz including vehicle speed, exhaust mass flow, gas concentration and calculated emission mass rate (e.g. CO_2 and NO_X). The vehicles' gas concentration and mass rates are considered time-aligned with the vehicles' dynamics according to the RDE regulation. As discussed in Section 2.3, misalignment between the NO_X and CO_2 signals could skew results and potentially lead to non-realistic or non-computable ratios.

### 2.4.1. Choice of the window-based sampling method and duration

To mitigate this effect, we do not calculate the instantaneous NO_X to CO_2 ratios for each time step *i*, but instead apply a moving average on both signals and calculate the ratio for each window. While this reduces the effect of small time alignment errors, it also reduces the signal resolution and therefore introduces uncertainty, which we will quantify later in the results section.

For our analysis, we choose average moving windows with a width of three time steps (i.e. 3 s). We calculated mass ratios of NO_X to CO_2 within a window *i* using the mean emission ratio across time step $i-1$, i and $i+1$, for each compound and using corresponding molar mass (Eq. 8). This method is preferred to the mean of the ratios within each window because raw PEMS concentration can be used instead of PEMS calculated mass rates, which require a time alignment with the exhaust flow meter.

$$r_{maw[i]} = \frac{\overline{(c_{NOX} \times M_{NOX})}_{[i-1,i,i+1]}}{\overline{(c_{CO2} \times M_{CO2})}_{[i-1,i,i+1]}} \quad (8)$$

where *c* is the instantaneous concentration, and *M* the molar mass.

Fig. 3 is a short example obtained from a longer PEMS test, showing the vehicle speed, raw NO_X and CO_2 concentrations at 1 Hz. The raw NO_X to CO_2 ratio is displayed for indication as it does not always give meaningful information, such as during deceleration and at idling with little or close to zero emission flow rate. The lower part of the graph shows the NO_X to CO_2 raw ratio (yellow) and based on the moving average method at remote-sensing-like conditions (green).
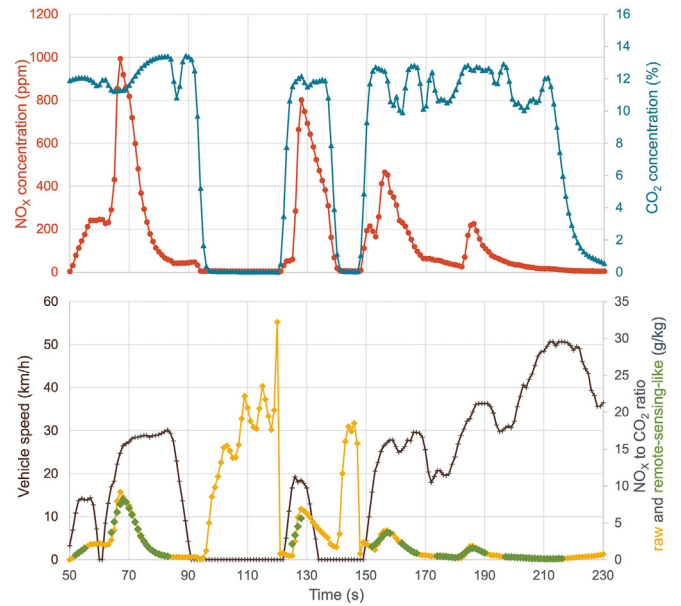
Fig. 3. Raw measurements of NO_X concentration (NO_X concentration (red line, top figure), CO_2 concentration (blue line, top figure), NO_X to CO_2 raw ratio (yellow line, bottom figure) and vehicle speed (black line, bottom figure) for a 180 s section of a PEMS trip. NO_X to CO_2 ratio of remote-sensing-like snapshot measurements are extracted for the analysis (green line, bottom figure).

Each green data point is considered for this analysis as a potential RSD measurement snapshot.

2.4.2. Typical RSD VSP distribution

In step 4 of the analysis, a VSP distribution commonly used for RSD testing is applied to calculate a weighted mean of the NO_X to CO_2 ratio. Fig. S1 of the supporting information appendix shows the selected distribution based on the CONOX data, measured across various European countries (Sjödin et al., 2018). For this analysis the share of measurements per bin is used to calculate a weighing factor for the PEMS emission data. The lower VSP boundary considered for the analysis is 0 kW/t and the highest bin contains VSP of 23 kW/t and above.

2.4.3. Statistical criteria and randomization function

The statistical criteria used in this analysis to express the correlations between the RSD estimations and PEMS are the coefficient of determination (R^2), calculated from a linear fit of first order, and the mean absolute relative error (MAREr). The latter criterion divides the absolute difference between the measurement and estimate by the emission levels of each trip. Then, the mean relative error is reported for all trips in percent.

In step 4b, a fixed starting point – or seed – for the random function is used to determine various sub-sample sizes of the windows. This enhances the result's replicability and comparability when varying the sample size.

3. Results

In this section, each distance-specific emission result of a PEMS-test journey is a data point. The correlation is assessed based on all available trips and evaluated from the coefficient of determination of the fit and mean absolute relative error criteria. As Table 1 showed, the PEMS dataset contains a variable share of PEMS trips by fuel type and euro standard. Euro 6b diesel and gasoline-powered vehicles represent the largest groups and are expected to primarily impact the upcoming results when vehicles of other fuel types and emission standards tend to be underrepresented.
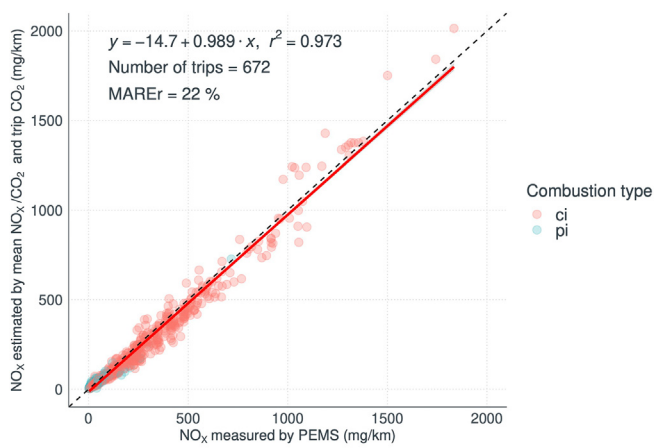
Fig. 4. Estimated distance-specific NO_X emissions using the mean of remote-sensing-like NO_X to CO_2 ratios and PEMS distance-specific CO_2 compared with distance-specific NO_X emissions measured by PEMS, i.e. based on NO_X concentration and exhaust flowrate. Red points indicate vehicles with compression ignition engine (ci), and blue points for positive ignition engine (pi).
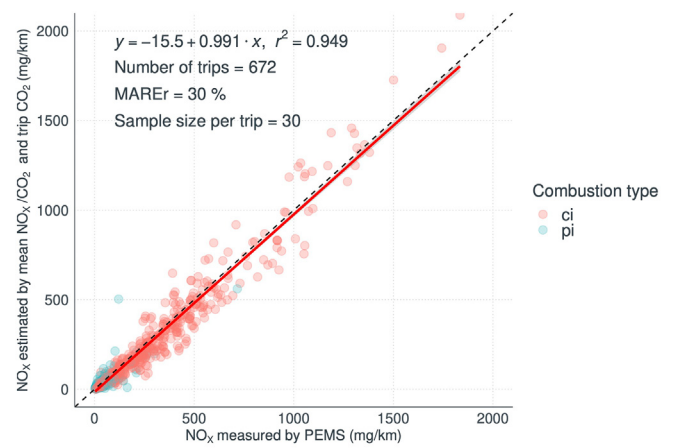
Fig. 5. Estimated distance-specific NO_X emissions based on a subset of 30 remote-sensing equivalent measurements per trip compared to PEMS emission measurements. Red points flag vehicles with compression ignition engine (ci), blue points for positive ignition engine (pi).

The effect of excluding emissions at idle and during negative VSP events on the ratio of trip NO_X to trip CO_2 is minimal for most trips as shown on Fig. S2 (step 1). The fitted line presents no bias, with a slope close to one, an R^2 above 0.99 and an average difference of 10%.

Applying the moving average window method on the filtered datasets (step 2) introduces on average an error of 8% when comparing the accumulated NO_X masses (Fig. S3) and a slope of 1.1 indicating emissions overestimation.

We estimate distance-specific NO_X emissions using the formula described in Eq. 6, i.e. multiplying the mean NO_X to CO_2 ratio by distance-specific CO_2 from the PEMS (step 3). Fig. 4 shows that uncertainty increases further to an average of 22%. However, no systematic bias is introduced when comparing the estimated distance-specific emissions with the PEMS measurement.

To account for the difference in engine load conditions between RSD and PEMS testing, the weighted mean emissions are calculated using the typical VSP distribution of RSD measurements (step 4a). Note that one trip is excluded due to not covering enough of the VSP bins. This step has little impact on the average error, as Fig. S4 shows, although the slope of the fit shows a tendency towards 6% higher estimated emissions.

We further analyze the effect of the sample size per trip in step 4b, using a subset of step 3 data. Unlike the previous steps, the calculation of the mean NO_X to CO_2 ratio for each trip is now based on a limited number of data points of each RDE trip. Fig. 5 shows the results for an equivalent of 30 remote-sensing records per trip. The coefficient of determination is around 0.95, with an average error estimate of 30%. The slope of the fitted line of 0.99 indicates the absence of bias. Results for sample sizes of 300, 10, 3 and 1 are detailed in Figs. S5, S6, S7 and S8 respectively. Although the correlation becomes less significant with samples below 30, the slope of the fitted line remains between the range of 1.03 and 0.97, signaling the absence of any significant bias independently of the number of records per trip.

Fig. 6 shows the results of a sensitivity study for a varying number of measurements per trips from 1 to 1000.

Statistical criteria, such as the coefficient of determination and mean absolute relative error, greatly improve with a growing sample size from 1 to 30. Between 30 and 300 measurements, both coefficient of determination and mean absolute relative error criteria are enhanced albeit at a slower rate relative to the higher number of remote-sensing measurements. Additional measurements do not translate to improved R^2 above 300, and the mean absolute relative distance-specific emission estimation error converges to a minimum of 23%.

In summary, this method can be used to determine a critical sample size by identifying a minimum R^2. With a sample size of at least 30 measurements, the R^2 coefficient reaches a value of 0.95, which can be considered as a fair level of variance accounted for by the linear model, with a MAREr of 30%. With more measurements, both R^2 and MAREr improve, resulting in 0.97 and 25%, respectively, for 200 measurements. The increase in the number of measurements above 300 offers only negligible improvements and would not justify the increased resources required for testing.

Finally, we study the practical, remote-sensing application, where absolute CO_2 emissions during a vehicle trip are not known but instead type-approval figures are typically identified from the registration data (step 5). To validate this method, all remote sensing-relevant measurements of each PEMS test are used (same as in step 4a); instead of the measured CO_2 emissions, the type approval values amplified by the real-world gap are used to calculate the distance specific NO_X emissions. It must be noted that this step focuses on Euro 6 vehicles only and that distance-specific emissions cannot be estimated for 19 trips due to the absence of NEDC type-approval CO_2 information. Fig. S9 compares the vehicles' anticipated average real-world CO_2 emissions with the measured values. The latter is offset by around −20 g/km, indicating that vehicles actually emitted less CO_2 during the analyzed RDE tests than under average real-world conditions.

The primary reason for this approach is that the on-road tests comply with the RDE protocol, during which a test can only be considered
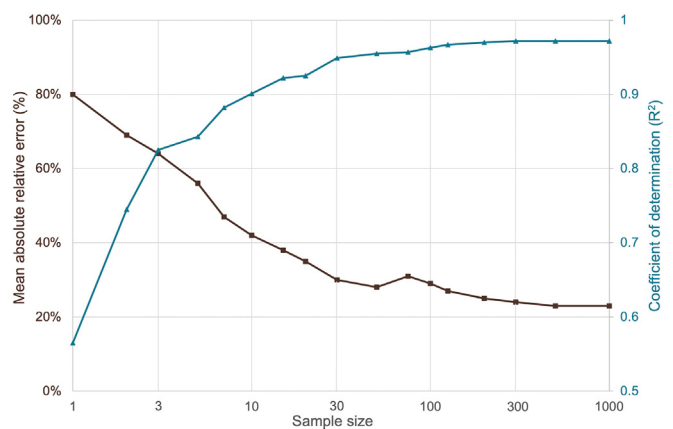
Fig. 6. Mean absolute relative error and coefficient of determination of estimated trip distance-specific NO_X emissions based on remote-sensing equivalent datapoints compared to PEMS emission measurements, depending on sample size.

valid if CO_2 levels are in a certain agreement with WLTP figures. WLTP type-approval CO_2 levels are expected to reduce the gap between NEDC type-approval and real-world CO_2 to a certain extent. These results support the hypothesis that a certain degree of gap remains between RDE and real-world CO_2 emissions.

This 20% increase is also visible in the linear fit between the estimated distance specific NO_X emissions, using the expected real-world CO_2 emissions, and the measurement, shown in Fig. S10. The average error increases from 20 to 27% compared to the results of step 4a.

We expect a better alignment of the slopes of the fitted line with the data obtained under conditions closer to average real-world driving (e.g. hillier terrain, more dynamic driving, etc.). Since the RDE measurements do not represent a sub-sample of real-world driving data it is reasonable to determine an average RDE to type approval CO_2 gap and multiply the type-approval CO_2 emissions with this value instead.

Table S1 presents the average CO_2 gap between RDE and type-approval CO_2 calculated from the PEMS dataset by fuel type for Euro 6 vehicles. This gap is 11–15 percentage-points lower than the real-world to type approval emission gap. The correlation between measured and estimated CO_2 emissions, using the RDE to type approval gap, is presented in Fig. S11. The mean absolute relative error for CO_2 has reduced from 15 to 10% by applying this correction.

The estimated average distance-specific CO_2 emissions under RDE are combined with averages of remote-sensing NO_X to CO_2 ratio to predict distance-specific NO_X estimation for each trip. Fig. S12 shows that the estimations are now close to the parity line, reflecting a better distribution around actual PEMS measurements, and the error has reduced from 27 to 23% when compared to Fig. S10. Using this method, vehicle average distance-specific NO_X emissions measured by PEMS are compared to simulated RSD levels in Fig. 7; on the left side for compression ignition and on the right side for positive ignition engine vehicles. The results are compared to the Euro 6 laboratory type-approval limits (dashed line), corresponding to 80 mg per kilometer, and 60 mg per kilometer respectively.

That figure demonstrates that the wide range of individual vehicles average emissions measured by PEMS are adequately captured by the methodology using RSD estimated distance-specific levels. The result suggests that with a sufficient number of measurements, the developed methodology is able to detect vehicles grossly emitting, e.g. with NO_X on-road emissions exceeding Euro 6 limits several times.

4. Discussion

Remote-sensing devices and PEMS are both used to measure on-road vehicle emissions. PEMS measures the vehicle's exhaust flowrate and NO_X
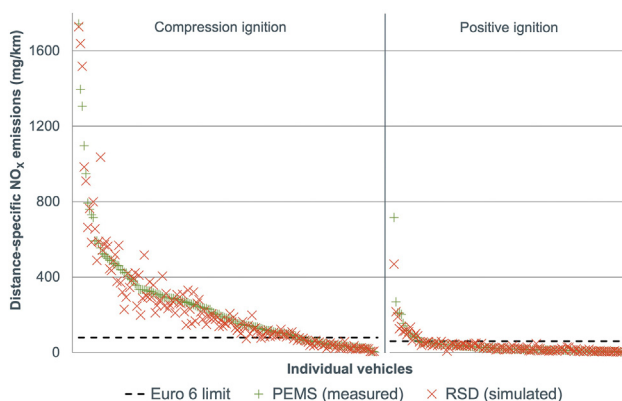
Fig. 7. Comparison of distance-specific NO_X emissions measured by PEMS and calculated for remote-sensing equivalent datapoints using vehicle-specific CO_2 type-approval information and average RDE gap. A pair of cross and plus markers represents results for one of the 289 Euro 6 vehicles. For vehicles driven multiple trips, we report their mean values. Vehicles are sorted from highest to lowest average NO_X emissions measured with PEMS by combustion engine type. The dashed lines represent the Euro 6 laboratory type-approval limits.

emission continuously during a whole trip, whereas RSD captures single snapshots of NO_X relative to CO_2 for vehicles passing a stationary RSD installation spot and under specific operating conditions of the vehicle. PEMS typically reports distance-specific emissions (e.g. gram per km) directly comparable with limits defined in emission standards for light-duty vehicles. However, PEMS results depend much on the individual tested vehicle, the selected route, the driving style, and the ambient conditions of the day and might be affected by the use of defeat devices. In the present paper, the PEMS data is considered as the reference technique at the individual trip level, which may differ from the vehicle's real-world emissions occurring in a wider range of operating conditions.

For RSD measurements, the raw output is CO_2-specific pollutant emissions (e.g. gram per kg of CO_2). RSD does not measure the exhaust flowrate. Thus, distance-specific NO_X emissions can only be estimated, for which we presented a method in this study. The method combines for an individual vehicle an average distance-specific CO_2 emissions value with average NO_X to CO_2 emission ratios determined from a subsample of PEMS measurements, simulating RSD snapshots. In a first step, average CO_2 levels measured during PEMS-test journeys are used. Compared to the distance specific NO_X emissions measured by the PEMS based on exhaust flowrate and NO_X concentrations, the results show that the method does not introduce a systematic bias. In a second step, the vehicle-specific CO_2 type-approval values are used as representative, average distance specific CO_2 value. Since average real-world CO_2 emissions are substantially higher than those determined during type-approval, the type-approval CO_2 emissions were amplified by average consumer reported real-world to type-approval fuel-consumption ratios. However, the CO_2 emissions of RDE tests are between those observed during real-world driving and type-approval, due to requirements set in the RDE regulation. The use of real-world CO_2 emissions resulted in an overestimation of distance-specific NO_X emissions from RSD compared to the PEMS measurements. Therefore, the ratio of PEMS CO_2 emissions levels and type approval value was derived from the data and used in the method instead. This resulted in a better correlation of the calculated distance specific NO_X emissions and the PEMS measurements. This shows the importance of using the appropriate distance-specific CO_2 estimate for a given situation, e.g. whether the objective is to estimate distance specific NO_X results for RDE tests, or for measurements made at real-world conditions. They also highlight that the method, by design, aims at predicting average vehicle emission levels rather than comprehensively covering trip-specific CO_2 variations.

While a PEMS measures contains thousands of data points, only a limited number of RSD measurements for a vehicle or vehicle type usually exist. Therefore, the effect of reducing the number of datapoints was analyzed. The analysis shows that a minimum of 30 remote sensing measurements is recommended to estimate distance-specific emissions with a satisfactory coefficient of determination ($R^2 > 0.95$) and a mean relative error of 30%. These margins seem to be compatible with market surveillance requirements, during which regulatory agencies aim to identify vehicle types with emission levels that significantly exceed the regulatory limits. The estimation considerably improves with greater sample size, but additional benefits become marginal above 300 measurements with a mean error stagnating at around 23%.

For this study, data points of PEMS-trip measurements are used to mimic remote-sensing measurements. In practice, RSD and PEMS emissions readings may differ due to their different measurement types of analyzers, with PEMS being considered the more accurate method. However, this topic is out of the scope of this analysis but suggests that the estimated error would be larger when using real RSD measurements.

These findings are pertinent to NO_X emissions only. Further research is required to understand if the method can be applied to other pollutants measured by RSD (e.g. carbon monoxide, hydrocarbons, particulate matter). We do not compare the accuracy of our approach with that of other potential analysis methods.

The results generated by the method described in this paper originate from multiple measurement snapshots of individual vehicles.

Long-term remote sensing data collection may manage to measure a small fraction of individual vehicles 30 times or more (Bernard et al., 2020). However, most vehicles are usually measured much less frequently. In Hong Kong, over a one-year continuous RS program, the average measurements per individual vehicle were only 2.2 times (Huang et al., 2020). This limits the use of remote sensing to estimate reliable distance-specific emissions factors for individual vehicles and thereby detecting high emitting vehicles present in the fleet. However, the results suggest that the method would be suited for larger vehicle groups sharing common characteristics (e.g. euro standard, fuel type, make, model, engine displacement, etc.) similarly to the concept of PEMS test families defined in the RDE regulation for the on-road type-approval of group of vehicles. Grouping remote-sensing measurements in larger families of vehicles allow reaching a critical sample size with less testing burden, as well as increasing fleet coverage, a point particularly important for market surveillance, while isolating some of the main causes of vehicles' emissions performance (Bernard et al., 2018).

5. Conclusions

This study shows that distance-specific NO_X emissions can be estimated with a fair accuracy using RSD without measuring the exhaust flowrate. While RSD measurements only provided NO_X to CO_2 emissions ratios, applying the presented method allows a direct comparison with emission limits expressed in gram per kilometer. Based on subsamples of PEMS-test data, the analysis reveals that 30–300 remote-sensing measurements are required for estimating NO_X emissions being sufficiently well in agreement with average PEMS-test results for a given vehicle or group of vehicles.

For market surveillance, RSD remains a complement to PEMS, albeit with some significant differences. RSD typically aggregates data from multiple vehicles of similar characteristics which leads on the one hand to a wide coverage of testing conditions (e.g. ambient temperature), but on the other hand to fewer measurements for individual vehicles when compared to e.g. PEMS-tests. This drawback is balanced by the cost and time efficiency of RSD, which makes it a suitable solution to rapidly monitor in-use emissions of large vehicle fleets. RSD measures on-road emissions in a contactless manner, and is therefore difficult to detect by a vehicle, which reduces the risk of defeat devices altering a vehicles' emissions in response to the awareness of being tested. RSD is also well suited to complement periodical technical inspection of exhaust pollutants to help finding individual high emitters based on few measurements, but only when compared to sufficiently high cutpoints. In particular, RSD is used for detecting owner's tampering such as the use of after-treatment emulators, that otherwise could have been removed or deactivated before a scheduled inspection. However, it remains essential to develop methods that can derive reliable emission factors (e.g. grams of pollutant per kilometer) that work with few measurements only.

CRediT authorship contribution statement

Yoann Bernard: Conceptualization, Methodology, Software, Formal analysis, Visualization, Project administration, Writing – original draft. **Jan Dornoff:** Resources, Data curation, Validation, Investigation, Writing – review & editing. **David C. Carslaw:** Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare no competing financial interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.151500.

References

Anenberg, S.C., Miller, J., Minjares, R., Du, L., Henze, D.K., Lacey, F., Malley, C.S., Emberson, L., Franco, V., Klimont, Z., Heyes, C., 2017. Impacts and mitigation of excess diesel-related NOx emissions in 11 major vehicle markets. Nature 545 (7655), 467–471. https://doi.org/10.1038/nature22086.

Baldino, C., Tietge, U., Muncrief, R., Bernard, Y., Mock, P., 2017. Road tested: Comparative overview of real-world versus type-approval NOX and CO2 emissions from diesel cars in Europe. International Council on Clean Transportation. https://theicct.org/publications/road-tested-comparative-overview-real-world-versus-type-approval-nox-and-co2-emissions.

Beaton, S.P., Bishop, G.A., Zhang, Y., Stedman, D.H., Ashbaugh, L.L., Lawson, D.R., 1995. On-road vehicle emissions: regulations, costs, and benefits. Science 268 (5213), 991–993. https://doi.org/10.1126/science.268.5213.991.

Bernard, Y., Dallmann, T., Tietge, U., Badshah, H., German, J., 2020. Development and application of a United States real-world vehicle emissions database. TRUE Initiative. https://theicct.org/publications/true-us-database-development-oct2020.

Bernard, Y., German, J., Kentroti, A., Muncrief, R., 2019. Catching defeat devices: how systematic vehicle testing can determine the presence of suspicious emissions control strategies. International council on clean. Transportation, 9–10. https://theicct.org/publications/detecting-defeat-devices-201906.

Bernard, Y., Tietge, U., German, J., Muncrief, R., 2018. Determination of real-world emissions from passenger vehicles using remote sensing data. TRUE Initiative., 5–16. https://theicct.org/publications/real-world-emissions-using-remote-sensing-data.

Burgard, D.A., Bishop, G.A., Stadtmuller, R.S., Dalton, T.R., Stedman, D.H., 2006. Spectroscopy applied to on-road mobile source emissions. Applied Spectroscopy 60 (5), 135A–148A doi:10/dkqjdm.

Dallmann, T., 2018. Use of remote-sensing technology for vehicle emissions monitoring and control. International Council on Clean Transportation. https://theicct.org/publications/remote-sensing-briefing-dec2018.

Davison, J., Bernard, Y., Borken-Kleefeld, J., Farren, N.J., Hausberger, S., Sjödin, Å., Tate, J.E., Vaughan, A.R., Carslaw, D.C., 2020. Distance-based emission factors from vehicle emission remote sensing measurements. Sci. Total Environ. 739, 139688. https://doi.org/10.1016/j.scitotenv.2020.139688.

EEA, 2019. Air Quality in Europe 2019 [Publication]. . https://www.eea.europa.eu/publications/air-quality-in-europe-2019.

Commission Regulation (EU) 2018/1832, 2018. of 5 November 2018 amending Directive 2007/46/EC of the European Parliament and of the Council, Commission Regulation (EC) No 692/2008 and Commission Regulation (EU) 2017/1151 for the purpose of improving the emission type approval tests and procedures for light passenger and commercial vehicles, including those for in-service conformity and real-driving emissions and introducing devices for monitoring the consumption of fuel and electric energy, Pub. L. No. 32018R1832, 301 OJ L. http://data.europa.eu/eli/reg/2018/1832/oj/eng.

Regulation (EU) 2018/858 of the European Parliament and of the Council, 2018. of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, amending Regulations (EC) No 715/2007 and (EC) No 595/2009 and repealing Directive 2007/46/EC, Pub. L. No. 32018R0858, 151 OJ L. http://data.europa.eu/eli/reg/2018/858/oj/eng.

Franco, V., Posada, F., German, J., Mock, P., 2014. Real-world exhaust emissions from modern diesel cars. International Council on Clean Transportation. https://theicct.org/publications/real-world-exhaust-emissions-modern-diesel-cars.

Gruening, C., Bonnel, P., Clairotte, M., Giechaskiel, B., Valverde Morales, V., Zardini, A., Carriero, M., 2019. Potential of Remote Sensing Devices (RSDs) to screen vehicle emissions (JRC117894). Publications Office of the European Union https://doi.org/10.2760/277092 10.2760/25928 (print).

Huang, Y., Yu, Y., Yam, Y., Zhou, J.L., Lei, C., Organ, B., Zhuang, Y., Mok, W., Chan, E.F.C., 2020. Statistical evaluation of on-road vehicle emissions measurement using a dual remote sensing technique. Environmental Pollution 267, 115456. https://doi.org/10.1016/j.envpol.2020.115456 Journal Article.

Jiménez-Palacios, J.L., 1999. Understanding and Quantifying Motor Vehicle Emissions With Vehicle Specific Power and TILDAS Remote Sensing. PhD ThesisMassachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/44505.

Mock, P., 2017. Real-Driving Emissions test procedure for exhaust gas pollutant emissions of cars and light commercial vehicles in Europe. International Council on Clean Transportation. https://theicct.org/publications/real-driving-emissions-test-procedure-exhaust-gas-pollutant-emissions-cars-and-light.

Sjödin, Å., Borken-Kleefeld, J., Carslaw, D., Tate, J., Alt, G.-M., De la Fuente, J., Bernard, Y., Tietge, U., McClintock, P., Gentala, R., Vescio, N., Hausberger, S., 2018. Real-driving emissions from diesel passenger cars measured by remote sensing and as compared with PEMS and chassis dynamometer measurements—CONOX Task 2 report. Federal Office for the Environment, Switzerland Journal Article https://www.ivl.se/download/18.2aa26978160972788071cd79/1529407789751/real-driving-emissions-from-diesel-passengers-cars-measured-by-remote-sensing-and-as-compared-with-pems-and-chassis-dynamometer-measurements-conox-task-2-r.pdf.

Tietge, U., Mock, P., German, J., Bandivadekar, A., Ligterink, N., 2017. From laboratory to road: A 2017 update of official and "real-world" fuel consumption and CO2 values for passenger cars in Europe. The International Council on Clean Transportation. http://theicct.org/publications/laboratory-road-2017-update.