doi:10.1038/nature22086

Impacts and mitigation of excess diesel NOx emissions in 11 major vehicle markets

This supplemental material provides additional detail on the emission scenario development, health and crop impact assessment methods, and results.

Supplementary Methods

1 Emission scenario development

The Baseline and Limits scenarios consider regulations that have been finalized (adopted into law) in each region and for each vehicle category. In addition to these scenarios that only consider adopted regulations, we developed realistic timelines for new regulations to evaluate the potential real-world NOx impacts of further action. These timelines are specific to each modeled region and vehicle category and consider the past timing, stringency and design of emissions regulations for new vehicles. Six regions – South Korea, Australia, India, Brazil, Russia, and Mexico – have largely followed the EU regulatory pathway for diesel vehicles with a lag time of one to five years. In contrast, the United States and Japan have developed their own emission control programs, and Canada has harmonized its standards with the United States. China has historically followed the European regulatory pathway, except with added fixes for urban bus emissions (Supplementary Table 1).

As a central part of this study, we developed a data set of real-world NOx emission factors for diesel light-duty vehicles (LDVs) – including passenger cars (PCs) and light commercial vehicles (LCVs) – and heavy-duty vehicles (HDVs) – including light-, medium-, and heavy-duty trucks, and buses – based on extensive emissions testing conducted in the U.S., the EU, China and Japan. These emission factors take into account varying real-world performance deviations with respect to the emission limits used to certify vehicles based on the results of laboratory testing (emissions certification and confirmatory in-use testing). The Euro 6/VI scenario uses the same emission factor adjustments as the Baseline scenario, with the only differences relating to the share of vehicle activity by emission control level in some regions (reflecting adoption of Euro 6/VI-equivalent standards). The Strong RDE scenario uses the same emission factor adjustments as the Baseline except they add stringent Real-Driving Emission (RDE) standards for LDVs. Lastly, the NextGen scenario uses the same emission factor adjustments as the Baseline scenario, except it adds next-generation standards for LDVs and HDVs.

1.1 Conversion of emission limits to distance-specific factors

In contrast to LDV emission limits, which are based on grams NO_X per unit of distance traveled (g/km and g/mi in the EU and US, respectively), HDV emission limits apply to truck and bus engines and are specified in grams NO_X per unit of engine work (g/kWh and g/bhp-hr). Thus, while no conversion is needed to compare LDV distance-specific emission factors to regulated emission limits, HDV brake-specific emission limits require a conversion to distance-specific emission factors to enable such a comparison. While there are substantial uncertainties involved in this conversion, it can nevertheless serve as a useful approach to put estimated real-world HDV emission factors (g/km) into context with regulated emission limits.

The following equations outline the method for estimating and applying the conversion factors, which are specific to each region, vehicle type and standard:

1) $Engine\ limit \left[\frac{grams\ NOx}{kWh}\right] * Conversion\ factor\ \left[\frac{kWh}{km}\right] = Emission\ factor\ [\frac{grams\ NOx}{km}]$

2) $Conversion\ factor\ \left[\frac{kWh}{km}\right] = \frac{Density\ of\ diesel\ \left[\frac{kg}{L}\right] * FC\ [\frac{L}{km}]}{BSFC\ [\frac{kg}{kWh}]}$

BSFC denotes brake-specific fuel consumption, a measure of engine thermal efficiency, and FC denotes average distance-specific vehicle fuel consumption, a measure of vehicle efficiency. BSFC applies to the engine as operating over the certification cycle, whereas FC applies to the average in-use fuel consumption of vehicles certified to a particular standard. This process of converting regulatory emission limits to distance-based emission factors is consistent with the recommended approach in Browning[78]. BSFC estimates were based on a literature review, which yielded inputs for the US and EU. Data were obtained from the U.S. EPA, West Virginia University, the German Federal Motor Transport Authority (KBA) and the CRC[79]. BSFC estimates for China, India, and Brazil were approximated by comparing engine efficiency estimates (using engine fuel maps available to the International Council on Clean Transportation, ICCT) to the US and EU over a standardized test cycle (FTP). Bus engines were assumed to have the same engine efficiency as tractor engines. To account for uncertainty in these estimates of BSFC, we added a 10% margin of error, which reflects the average variation found among estimates for a given vehicle type and region.

FC values were extracted from International Energy Agency's MoMo database (March 2016 version) and converted from liters gasoline-equivalent per 100 km (lge/100km) to liters diesel-equivalent per km (L/km) using a multiplier of (1/1.077)/100 (IEA Mobility Model). IEA's in-use fuel consumption estimates are calibrated such that bottom-up estimates of transportation fuel use match top-down energy consumption statistics. We account for uncertainty in these estimates by assuming an additional 20% margin of error. Since engine fuel consumption (BSFC) and in-use vehicle fuel consumption (FC) have opposite effects on the conversion factor (kWh/km), these conversion factors may be less sensitive to changes in one input to the extent that engine efficiency (over the regulated test cycle) and vehicle efficiency (over all in-use driving conditions) are correlated. However, since a perfect correlation is not guaranteed, the "low" and "high" conversion factor ranges are conservatively estimated so as to reflect the maximum variation in response to changes in these parameters (e.g. the upper range of BSFC and lower range of in-use FC). The uncertainty in equivalent distance-based NO_X emission factors (g/km) decreases as engine emission limits decrease. In other words, our confidence in the converted emission limits is higher for more stringent standards such as Euro VI or US 2010 than for earlier standards (e.g. Euro III).

1.2 HDV emission factors

1.2.1 EU HDV emission factors

As a starting point, the central NO_X emission factor estimates for diesel HDVs in the EU-28 were based on Emisia's Sibyl model, which draws its local air pollutant emission factors for each vehicle type, fuel type and technology (i.e. Euro certification level) from the COPERT software. COPERT is supported by the European Environment Agency and European Commission and is developed by Emisia SA. These emission factor estimates were selected as the starting point for this paper based on their consistency with remote sensing measurements[17,35] and other studies of real-world NO_X emissions in the EU[15,33]. Beyond the central estimates in this paper, we assume a 25% margin of error to account for variability in emission measurements and traffic composition[33].

Extended Data Figure 6 compares emission factor estimates of medium and heavy trucks and buses certified to Euro III, Euro IV, Euro V and Euro VI standards. While Euro IV and V regulations have tightened NO_X emission limits for HD engines compared to Euro III, independent emissions tests of Euro IV and V vehicles have found that these regulations have not translated into the expected real-world emission reductions[15,16,35]. Based on remote sensing measurements of almost 70,000 vehicles in the UK over model years 1985 to 2012, Carslaw and Rhys-Tyler[35] conclude that there is little evidence of real-world NO_X emission reductions from diesel vehicles compared to model years 1998 and earlier. Similarly, Carslaw *et al.*[17] find no significant reduction in urban NO_X emission factors from Euro III to Euro V for heavy-duty trucks and from Euro III to Euro IV for buses. Velders et al.[33] estimate real-world NO_X emission factors (g/km) for heavy-duty trucks based on real-world emissions measurements in the Netherlands, finding no significant reduction in urban driving conditions from Euro III to Euro V, but a roughly 50% reduction on motorways from Euro III to Euro IV/V.

The emission factor estimates applied in this paper–consistent with EMISIA[34] up through Euro V–reflect real-world emission reductions of Euro IV and Euro V vehicles in highway/motorway driving as well as the problems with excess NO_X emissions in urban driving conditions. For Euro VI vehicles, Emisia's estimates indicate an 80% reduction from Euro V, which is equivalent to the percent reduction in regulated NO_X emission limits; however, these resulting Euro VI estimates are substantially higher than indicated by chassis dynamometer tests of Euro VI vehicles conducted by VTT Technical Research Centre of Finland. Muncrief[15] reports on the VTT test results, finding that out of 55 tests–covering six buses, four tractors, and one rigid truck–only two data points (of 55) exceeded the Euro VI emission limit for the transient test, and then only slightly. To allow for the possibility that Euro VI vehicles emit more during cold start and low load conditions than captured by the duty cycles tested, lifetime average emissions of Euro VI buses (central estimates) are assumed to be 1.5 times the WHTC emission limit of 0.46 g/kWh, reflecting good compliance with confirmatory in-use emission testing. Since Euro VI trucks typically have a smaller share of vehicle activity in low load conditions, average emissions of Euro VI trucks are assumed to be equal to the WHTC emission limit of 0.46 g/kWh. These emission factors are between the average of test results reported in Muncrief[15] (0.13 g/kWh) and the emission factors extracted from Sibyl (EMISIA[34]).

1.2.2 China HDV emission factors

In China, baseline emission factors for Euro III and Euro IV were based on Yao et al.[36] for heavy trucks and Wu et al.[37] for buses. Emission factors for Euro V buses are drawn from Zhang et al.[38]. Emission factors for Euro V and Euro VI medium and heavy trucks and Euro VI buses were estimated using the percent reduction in real-world NO_X in the EU applied to the China-specific emission factor for the previous standard.

Overall, the five China-specific studies reviewed come to similar conclusions about the real-world NO_X emissions impacts of Euro III, Euro IV and Euro V-equivalent standards for HDDVs in China. In portable emissions measurement systems (PEMS) tests of more than 130 diesel HDVs in Beijing, Wu et al.[37] found no significant difference in NO_X emissions between Euro II, Euro III and Euro IV buses in typical urban driving conditions, and between Euro I, Euro II and Euro III trucks for most cases. The authors report mean and 95% confidence intervals for buses and trucks, which are reflected in our upper and lower emission factor estimates. Zhang et al.[38] conducted PEMS tests of two diesel Euro V buses equipped with selective catalytic reduction (SCR) systems, finding an average emission factor of 7.5 g/km over typical bus routes in Beijing. Guo et al.[73] conducted PEMS tests of 9 diesel buses certified to Euro III, Euro IV or Euro V standards and operating in Beijing. The authors estimate lower emission factors for Euro III and Euro V buses than the other studies but offer similar conclusions–namely, that Euro IV buses in China have not substantially reduced NO_X emissions, while Euro V buses have yielded NO_X benefits compared to Euro III. Since Euro VI-equivalent standards have not been implemented in China, emission factors for the proposed China VI standards reflect the same estimates as for Euro VI in the EU.

1.2.3 U.S. HDV emission factors

The central NO_X emission factor estimates for diesel heavy-duty trucks in the US were based on the US EPA's MOVES model, which estimates local air pollutant emission factors for each vehicle type, fuel type and model year (which correspond to emission standards). These estimates were validated against recent remote sensing measurements of exhaust emissions from in-use trucks in California conducted by the University of Denver[40]. For buses, average emission factors by certification level were extracted from West Virginia University's Integrated Bus Information System (IBIS), which maintains a transit bus emissions database with NO_X measurements from testing over 3,000 buses throughout the United States[41]. Since model years of tested vehicles range from 1990 to 2009, distance-weighted average NOx emission factors were computed for buses certified to EPA 1998, 2004, and 2007 standards. Since the database does not include vehicles certified to EPA 2010 standards, the emission factor for EPA 2010 buses was estimated assuming the same difference between IBIS and MOVES for EPA 2007 buses (a factor of 1.8).

Citing an absence of data for MY 2010 and later HDVs, US EPA assumed a ~90% reduction in NO_X emission rates corresponding to the drop in the regulated emission limit from 2.4 g/bhp-hr for EPA 2004 to 0.2 g/bhp-hr[39]. By comparison, remote sensing measurements from the University of Denver indicate that fuel-specific NO_X emissions of heavy-duty trucks decreased by 83% from MY 2004 to MY 2012[47]. Since additional PEMS testing data (*i.e.* from in-service conformity testing) is needed to establish a robust alternative estimate, this study applies the MOVES estimates for EPA 2010 trucks. In addition to the central estimates, for HDTs we assume a 25% margin of error to account for variability in emission measurements and traffic composition[33]. For heavy heavy-duty trucks (HHDTs) certified to EPA 2010, the estimated range of real-world multipliers (1.1 to 3.4) contains the factor of 3x (the 0.2 g/bhp-hr limit) measured for MY2014 HHDTs (n = 28)[80]. Lower and upper bound estimates for buses from EPA 1998 to EPA 2007 are based on 95% confidence intervals estimated from the IBIS dataset.

1.3 LDV emission factors

1.3.1 EU LDV emission factors

Compared to other regions and vehicle types, passenger cars in the EU are among the most studied with respect to real-world NO_X emissions. “Real-world” emission factor estimates for passenger cars certified from Euro 1 to Euro 5 are based on our review of the literature, including the results of PEMS, remote sensing, laboratory measurements and emission factor models. While real-world NO_X emissions of LCVs have been studied to a lesser extent, they have generally been found to emit at least 1.5 times the levels observed for passenger cars[45]. Since this multiplier generally corresponds to the difference in emission limits (for heavier LCV classes compared to passenger cars), average LCV emission factors are assumed in this analysis to be 1.5 times the level estimated for passenger cars for certification levels starting with Euro 4. For PCs and LCVs certified to Euro 3 and earlier, emission factors are aligned with COPERT under the expectation that it already reflects the results of earlier emissions testing.

Emission factors for Euro 6 diesel cars were estimated using ICCT’s diesel PEMS database. This database, which supported ICCT’s diesel PEMS metastudy[11], covers emissions from 32 Euro 6 diesel passenger cars over approximately 180 hours and 8,000 km of driving with second-by-second data resolution. To explore the statistical distribution of on-road NO_X emissions and their relation to specific driving conditions (e.g., cold-start events, or aggressive driving), second-by-second PEMS signals were binned using a moving averaging window concept (see Rubino et al.[81]).

The average NO_X emission factor for pre-RDE Euro 6 diesel passenger cars was directly estimated from ICCT’s PEMS database, resulting in a real-world multiplier of 5.7 times the Euro 6 emission limit of 80 mg/km. This estimate is in line with the findings of other PEMS studies (e.g.[76,82]). The subsequent evolution of real-world emission factors was modeled according to the specifications of the current RDE regulation as well as under a future strengthened program. These emission factors for diesel cars under various stages of RDE are described in detail in Miller and Franco[30]. Given the “in-flux” nature of many RDE requirements around the world (including the EU and China), several different policy pathways were evaluated without judgment as to which “will” happen. The Baseline RDE designation is consistent with the 1st and 2nd regulatory packages in the EU (“golden car” testing, conformity factor of 2.1, and narrow RDE test coverage), whereas Strong RDE reflects an idealized program (including cold-start provisions, in-service conformity testing, independent verification, expanded test boundaries, and lower conformity factors). These can be considered as upper and lower bounds for the effectiveness of provisions targeting real-world NOx compliance for LDVs.

The emission factors and uncertainty bounds applied in this analysis are consistent with the conclusions of recent real-world emissions studies in the EU[17,45,77]. Extended Data Figure 6 shows how the emission factors applied in this study relate to real-world emissions studies as well as other emission factor models: TRL[75] (2009) emission factors are shown to illustrate the expected performance of Euro 5 and Euro 6 before these regulations were implemented. Emisia[34] emission factors are based on COPERT, which is among the most widely recognized tools for conducting road transport emission inventories.

The following discussion gives additional detail on how our findings relate to other recent real-world emissions studies for LDVs in the EU. In a study conducted for the Dutch Ministry of Infrastructure and the Environment, TNO tested sixteen Euro 6 diesel passenger cars under laboratory and real-world driving conditions[77]. The authors found that while real-world NO_X emissions for most of the tested vehicles fell within a range of 400-800 mg/km, at least one vehicle met the regulatory limit of 80 mg/km in real-world conditions. The authors conclude that diesel cars could achieve the Euro 6 NO_X limit under real-world conditions if required to do so by a well-designed real-driving emissions regulation.

Ntziachristos et al.[45] undertake a comprehensive review of NO_X emission factors for Euro 5 and Euro 6 diesel passenger cars and light commercial vehicles, including both laboratory and real-world driving conditions. The authors find substantial variability in real-world NO_X emissions even after controlling for driving conditions (urban, rural and highway) and confirm the extent to which real-world NO_X emissions (as measured using PEMS) exceed regulatory emission limits. In addition, Ntziachristos et al.[45] made several conclusions which are taken into account in our study: LCVs emit about 1.5 times as much NO_X as PCs; emissions during conditions outside the operating range of the NEDC exceed those during normal driving; excess NO_X emissions are a problem for the majority of vehicles and multiple manufacturers; and the RDE regulation in its current form covers a wider range of operation conditions than the NEDC, but is still confined in its coverage.

In addition to the central estimates reported in this paper, we also quantify the uncertainties associated with real-world emissions of PCs and LCVs starting with Euro 5. Extended Data Figure 6 summarizes these upper and lower bounds in comparison to regulated emission limits. The range for Euro 5 is based on the range of average PEMS results from studies in France and the UK as reported in Ntziachristos et al.[45] The ranges for Euro 6 (pre-RDE) is based on a 95% confidence interval extracted from ICCT's PEMS database, with the lower bound adjusted to account for skewness in the distribution. Ranges in emission factors under the Baseline RDE and Strong RDE programs are derived from Monte Carlo simulation as reported in Miller and Franco[30].

Historical emissions of NOx and other species from vehicles in Europe, as well as the importance of a shift from gasoline to diesel vehicles for NOx emissions, have been explored in detail elsewhere (e.g. Crippa et al. [68]). Our study provides additional information by differentiating on-road diesel NOx emissions by real-world emissions versus certification limits for vehicles certified to each Euro standard, and projecting real-world emissions under more stringent policies in the future.

1.3.2 <u>China LDV emission factors</u>

While China has generally followed the European regulatory pathway for diesel LDVs up through the adopted Euro 5-equivalent standards, the recently announced China 6 program makes several key changes. Notably, China 6a and China 6b lower the NO_X emission limit for diesel cars to 60 mg/km in 2020 and 35 mg/km in 2023, respectively. Nonetheless, the RDE provisions of China 6b appear to be consistent with the "Baseline RDE" program in the EU (1st and 2nd packages adopted in May 2016), which is estimated to reduce diesel car NO_X from a factor of 5.7x the emission limit of 80 mg/km for Euro 6 (pre-RDE) to 4.0x with RDE (Miller and Franco[30]). This real-world emission factor estimate is substantially higher than the regulated conformity factor since ~70% of NO_X emissions result from defeat devices, poor calibrations, and extended operating conditions (among RDE-tested vehicles), while only ~30% of NO_X emissions arise from normal driving and cold-start conditions (the latter of which is incorporated into the RDE test in the 3rd regulatory package).

Under the China 6 RDE program, the greatest uncertainty relates to the share of vehicles that will continue to have poor calibrations or defeat devices (legal or otherwise). In the Baseline scenario, China 6b is assumed to reduce emissions under normal driving conditions (covered by the RDE test) in proportion to the EU limits (35 mg/km vs. 80 mg/km), bringing the normal driving emissions component from 118 mg/km in the EU to 51.5 mg/km in China. Other emission factor components (cold-start, extended driving, and defeat devices) are unchanged, leading to an average emission factor of 289 mg/km for China 6b. This represents an 11% reduction from the average NO_X emission factor for Europe (with Baseline RDE). The Strong RDE scenario considers the addition of in-service conformity testing, a tightened conformity factor, expanded boundaries of the RDE (including cold-start), monitoring of emissions using remote sensing, and publication of RDE test results to enable independent verification. With these provisions, China 6b (plus the aforementioned provisions) is estimated to achieve an average in-use emission factor of 63 mg/km.

1.3.3 U.S. LDV emission factors

Several studies have quantified the emissions and health impacts of the Volkswagen emissions scandal in the United States[26-28]. Each of these studies included the roughly 482,000 Volkswagen group vehicles with 2.0 liter engines and relied on the PEMS testing reported in Thompson et al.[48] for the average NOx emission factor of affected vehicles. Barrett et al.[26] apply a weighted average NOx emission factor of 0.93 g/km (95% CI: 0.33 to 1.53); Oldenkamp et al.[28] apply an emission factor of 0.9 g/km; and Holland et al.[27] apply separate NOx emission factors for vehicles equipped with lean NOx traps (LNT) and SCR technologies.

This study incorporates three categories of excess NOx from diesel LDVs in the US: emissions from the roughly 482,000 Volkswagen group vehicles with 2.0 liter engines; emissions from the additional 85,000 Volkswagen group vehicles with 3.0 liter engines; and emissions from passenger cars and light trucks unaffected by the Volkswagen emissions scandal, but which nonetheless may emit somewhat higher levels of real-world NOx than regulatory emission limits. The emission factor for LDVs certified to US Tier 2 standards is computed as a sales-weighted average of these three categories. Based on sales figures from the US Energy Information Administration, we estimate about 2.5 million LDVs were sold and certified to Tier 2 standards between 2004 and 2015 (US EIA[46]).

Adjustment factors are defined as a multiplier of the Tier 2 bin 5 emission limit, equivalent to 43 mg/km. The ranges of adjustment factors for Volkswagen group vehicles with 2.0 liter and 3.0 liter engines are generally consistent with the findings of Thompson et al.[48] and U.S. EPA[10], respectively. The central estimate for unaffected vehicles is based on Vehicle C (a BMW X5) in Thompson et al. (2014), and the range of adjustment factors varies from perfect compliance (i.e. a factor of 1) to about twice the regulated limit. The higher end of this range takes into account excess emissions that could result from a small share of driving in conditions such as the rural-uphill/downhill cycle tested in Thompson et al.[48] (i.e. 10 times the limit applied to about 5-10% of vehicle-km traveled).

Using remote sensing measurements, Bishop and Stedman[19] find that fuel-specific NOx emissions of diesel passenger cars have remained statistically unchanged since the progression from US Tier 1 to Tier 2 standards. Based on these data, which we obtained from the University of Denver website (http://www.feat.biochem.du.edu), vehicles manufactured by Volkswagen and Audi comprise more than 95% of the Tier 2 sample and have greater than 10 times the average NOx emissions of the Tier 2 diesel passenger cars from other manufacturers. Given the similarity in NOx emissions observed between Tier 1 vehicles and high-emitting Tier 2 Volkswagen group vehicles, we assume the same average emission factor for vehicles certified to Tier 1 standards as Volkswagen vehicles with 2.0 liter engines. This assumption results in a central estimate of 1.1 times the Tier 1 emission limit for "useful life" (equivalent to 780 mg/km after 10 years/100,000 miles), with a range of 0.8 to 1.4 times that limit. This emission limit including deterioration is appropriate for comparison, since most Tier 1 vehicles were about a decade old by 2015, the base year for this study.

While US LDV emissions in 2015 are determined by vehicles certified to Tier 1 and Tier 2 standards, emissions in 2040 will be determined primarily by vehicles certified to Tier 3 standards, which phase in from 2017 to 2025. Recent actions by California's Air Resources Board and US EPA indicate that future LD diesel NOx emissions will be much closer to regulatory emissions limits. These actions include ARB's newly-developed defeat device screening methods, which notably include the use of "special driving cycles and conditions that may reasonably be expected to be encountered in normal operation and use"[49]. As a result, average NOx emission factors for future vehicles certified to Tier 3 standards are estimated to be within 30% of the certification limit, equivalent to 14 mg/km. This estimate is based on the real-world multiplier of 1.27 for a Tier 2 diesel vehicle with good performance (Thompson et al.[48]), assuming NOx accounts for 57% of total non-methane organic gases (NMOG) + NOx emissions for Tier 3 vehicles. We assume the same range of adjustment factors for Tier 3 as for Tier 2 vehicles unaffected by the VW emissions scandal (1 to 2 times the Tier 3 limit). This assumption reflects the expectation that defeat device screening methods such as those recently adopted by ARB will be reasonably effective in achieving Tier 3 emission limits in the real world.

Extended Data Table 4 summarizes the emission factor estimates and equivalent adjustment factors applied for diesel LDVs in the US. The central estimate for Tier 2 vehicles—which includes those vehicles affected by the VW emissions scandal—still represents a 74% reduction from Tier 1 levels, reflecting the fact that most of the US diesel LDV fleet was not affected by the Volkswagen emission scandal. Tier 3 vehicles are expected to yield a 94% reduction from Tier 2 emission levels.

1.4 Emission factor limitations

It is important to note several limitations of the emission factors derived in this analysis. First, emission factors were derived from the best available emission testing data in each region. Since these emission testing studies used different methods and testing conditions (for example, remote sensing vs. PEMS testing, or measurement on a specific corridor vs. a specified route or drive cycle), there is not yet sufficient evidence to facilitate an apples-to-apples comparison of the real-world performance of specific standards (e.g. U.S. EPA 2010 and Euro VI).

Second, the emission inventory modeling in this study applies distance-based emission factors to estimates of vehicle activity by region, vehicle type, and emission control in each year. For HDVs, these distance-based emission factors (g/km) are not directly comparable to engine emission limits (g/kWh), since the conversion depends on estimated fuel efficiency. In the case of China, for example, heavy-duty trucks were assumed to have the same distance-based emission factors as in Europe in order to avoid introducing errors due to very limited data on in-use fuel efficiency in each region.

Third, given that HDV standards typically introduce fuel-specific emission limits (g/kWh) rather than distance-based limits, Next Gen standards were assumed to achieve a 90% reduction from region-specific emission levels under Euro VI-equivalent standards. Actual emissions performance could vary based on specified emission limits, test cycles, emission control system calibration, vehicle fuel efficiency, and driving conditions. Nevertheless, we assume that Next Gen standards build on the requirements for durability of emission controls and on-board diagnostic (OBD) systems that have proven successful in controlling real-world NOx emissions of Euro VI and U.S. 2010 HDVs. For example, Euro VI durability requirements for heavy trucks and buses were increased to 700,000 km/7 years—whichever comes first—compared to 500,000 km/7 years for Euro V. Euro VI standards also strengthen OBD requirements, requiring these systems to monitor the performance of emission related components, issue warnings and restrict vehicle operation in cases of insufficient urea or poor SCR performance, and provide access to OBD data to facilitate the repair and maintenance of vehicles.[83] Coupled with effective in-service conformity testing and defeat device screening procedures (the latter especially relevant for LDVs), such requirements are expected to reduce the incidence of high-emitting vehicles (with malfunctioning emission controls) and poor emission control calibrations. With sufficiently comprehensive compliance and enforcement provisions, Next Gen standards are expected to improve the application of vehicle emission control technologies—with respect to both technology uptake and effective calibration. With respect to the feasibility of Next Gen emission limits, U.S. EPA and ARB have already certified at least one compressed natural gas engine to the optional low NOx limit of 0.02 g/bhp-hr[84]; ARB estimates that multiple strategies could be used to achieve similar reductions for diesel engines.[85]

Lastly, in the West Virginia University (WVU) study of LDV diesels in the U.S., the BMW test vehicle with good performance had NOx emissions that were 1.27 times the Tier 2 manufacturer fleet average standard, which applies to the sales-weighted average of gasoline and diesel vehicles sold by each manufacturer. Since the upcoming Tier 3 standards also do not require diesels to meet the fleet average limit (as with Tier 2), we assumed that by 2025, new diesels emit at 1.27 times the Tier 3 fleet average limit. We used the NOx share of non-methane hydrocarbon (NMHC)+NOx for the test BMW to estimate the NOx share of the Tier 3 NMHC+NOx limit in 2025 (Thompson et al. [48]).

1.4.1 Comparison of emissions measurement methods

Each emissions measurement method—including remote sensing, PEMS testing, and chassis dynamometer testing—has advantages and disadvantages with respect to the development of fleet average, in-use emission factors. Since PEMS and chassis dynamometer testing are typically applied to a limited number of vehicles, a key challenge for these methods is whether the sample obtained is fully representative of the targeted fleet. For this reason, remote sensing can prove especially useful in determining whether the results of PEMS and chassis dynamometer testing match the emissions trends observed over much larger vehicle samples. Additional investigation into the distribution of NO_X emission factors for modern (Euro 6/VI-equivalent) light-duty vehicle and heavy-duty vehicle fleets could further refine the characterization of fleet average emission factors.

The principal challenge of using remote sensing data to inform real-world emission factors is that the driving conditions measured are not necessarily representative of all real-world driving conditions. For example, remote sensing measurement sites are typically selected with a positive (uphill) grade such as a freeway on-ramp in order to ensure vehicles are operating under hot-stabilized conditions (Bishop and Stedman[47]). Moreover, remote sensing measurements are typically reported in units of NO_X per unit fuel consumption or as a ratio of NO_X to CO_2; these results cannot easily be compared to distance-based emission factors (i.e. grams NO_X per km), since the vehicle's fuel efficiency would be needed to facilitate this conversion (i.e. kg fuel per km). An alternate approach to comparing PEMS and remote sensing measurements is to convert PEMS data to fuel-specific emission factors, since CO_2 measurements are normally collected along with NO_X emissions. Since remote sensing typically captures vehicles operated under hot-stabilized exhaust conditions, ideally the PEMS data for such comparisons would also be limited to hot-stabilized exhaust conditions. While such direct comparisons of PEMS and remote sensing data are not included in this study, these represent an opportunity for further investigation.

Some studies have identified a link between vehicle specific power (VSP)—a measure of engine load—and NO_X emission rates (expressed in grams per second), CO_2 emission rates are affected in a similar manner.[86] Since fuel-specific emission factors are a measure of NO_X emissions per unit fuel (or CO_2 if applying fuel carbon content), VSP has limited usefulness for comparisons of fuel-specific NO_X emission factors. In our analysis of heavy-duty remote sensing results obtained from the University of Denver, we found no clear relationship between fuel-specific NO_X emission factors and VSP.

While it may be possible to estimate distance-based emission factors from fuel-specific results by making assumptions about vehicle fuel efficiency, doing so could introduce substantial errors due to a likely mismatch in driving conditions between remote sensing measurements (i.e. a roadway with positive grade) and fuel efficiency estimates (i.e. average fuel efficiency of in-use vehicles). Instead of attempting such a conversion, we draw on remote sensing results to inform the relative effectiveness of progressive vehicle regulations (i.e. percentage reductions), for which there is no apparent bias other than a focus on hot-stabilized operating modes. In other words, whatever bias is introduced by the measurement conditions can be expected to apply—on average—to vehicles irrespective of their regulatory certification level.

1.4.2 Deterioration of emission controls

The Baseline scenario does not explicitly account for deterioration of emissions control equipment as vehicles age, although deterioration effects are likely captured to some extent in remote sensing studies that measure emissions of vehicles older than 5 years (i.e. Carslaw et al.[17]). For cars and light commercial vehicles, there is preliminary evidence that deterioration effects are lower for more recent emission standards (i.e. Euro 4) compared to earlier standards (i.e. Euro 2 and Euro 3): Chen and Borken-Kleefeld[18] suggest a deterioration factor of 0% for Euro 4 diesel cars over 80,000 km, compared to 22% and 10% for Euro 2 and Euro 3 diesel cars. These deterioration estimates are well within the uncertainty applied to our central emission factors for the estimation of emission inventories (typically 25% unless more robust quantification of uncertainty is possible, e.g. the standard deviation from ICCT's PEMS database for Euro 6 diesel cars).

In the EU, type approval requirements for the durability of emission control systems have increased from 80,000 km/5 years for Euro 3 to 100,000 km/5 years for Euro 4 and 160,000 km/5 years for Euro 5/6 vehicles; these durability requirements are significantly less stringent than U.S. Tier 2 standards, which define full useful vehicle life as more than 190,000 km (120,000 mi) or 10 years—whichever comes first. U.S. LDV Tier 3 standards, after which the next-generation scenario is modeled, extend these durability requirements to cover 240,000 km (150,000 mi) or 15 years. While further study is needed to ensure that the durability requirements of more recent standards do in fact reduce long-term deterioration rates, we expect deterioration effects to be lower for future policy scenarios (Euro 6/VI, Strong RDE, NextGen) than for the Baseline scenario—potentially resulting in greater NOx benefits than we estimated for 2040. Likewise, to the extent that deterioration effects are not fully captured in the Baseline scenario, our central estimate of "excess diesel NOx" is conservatively low; yet, since the effects of deterioration estimated in other work are within the range of uncertainty applied in our emission factors, further incorporating deterioration effects would not be expected to result in emissions higher than the upper bound of our emission inventory estimates.

1.4.3 Fuel economy improvements

Emission projections conservatively exclude any emissions reductions resulting from increasingly fuel-efficient engines. Fuel economy improvements are less likely to directly impact NOx emissions of LDVs compared with HDVs, since LDVs are regulated on the basis of emissions per unit of distance traveled (grams NOx per km or mi), whereas HDVs are regulated on the basis of emissions per unit of engine work (grams NOx per kWh or bhp-hr).

Among Euro 6 diesel cars, NOx emissions are primarily determined by the choice and calibration of emission control technologies. In a meta-analysis of PEMS testing data, larger vehicles were observed to emit lower levels of NOx on average than smaller vehicles over the Worldwide Harmonized Light Vehicles Test Cycle (WLTC): the authors attribute this difference in performance to the widespread SCR systems among larger vehicles as opposed to LNT that are more common among smaller vehicles (Yang et al.[87], p.14). While there may exist a trade-off to some extent between fuel economy and low NOx emissions for diesel vehicles (in that SCR systems can lead to a fuel economy benefit, whereas LNT systems can incur fuel economy penalties), the fuel economy effects of emission control technologies are small compared to other technologies designed to reduce fuel consumption (e.g. turbocharging and downsizing, advanced transmissions, hybridization).

For HDVs, future reductions in in-use vehicle fuel consumption (i.e. ‘FC’ in Equation 2) could effectively lower the converted (distance-based equivalent) emission limit for a given engine NOx limit (grams NOx/kWh) by reducing the conversion factor. Therefore, assuming HDVs continue to meet engine certification limits and the real-world adjustment factor does not increase, future fuel economy improvements can be expected to reduce tailpipe NOx in proportion to the reduction in fuel consumption. Since Euro VI/U.S. 2010 equivalent standards reduce new diesel HDV NOx emissions by approximately 90% or more compared to Euro V equivalent standards, such emission regulations are the primary policy driver of NOx emission trends from 2015–2040; however, in regions that have already implemented Euro VI/U.S. 2010 equivalent standards, substantial fuel economy improvements could result in NOx co-benefits that additional to the direct benefits of emissions regulations that are the focus of this study.

1.5 Emission inventory methods

The ICCT’s Global Transportation Roadmap Model (hereafter referred to as the “Roadmap model”) has been developed since 2012, applied in numerous global and regional emissions inventory studies, and validated against leading transportation models developed by the International Energy Agency and other research institutions (see Mishra et al.[88]). The Roadmap model takes into account varying age distributions, survival curves, and VKT degradation by region and vehicle type. HDV vehicles have longer average vehicle lifetimes, higher annual vehicle activity, and greater fuel consumption than LDVs, each of which contributes to their disproportionate share of NOx emissions.

For this study, we adapt the Roadmap model to develop the first detailed real-world NOx inventory for diesel light- and heavy-duty vehicles in 11 major markets covering nearly 80% of new diesel vehicle sales. These markets account for the majority of global NOx emissions from diesel vehicles and set the precedent for new vehicle regulations that are adopted elsewhere. Regions that import used vehicles can benefit indirectly from the regulations applied in leading markets. A key innovation of this study is that it relies on real-world NOx emission factors derived from an extensive literature review of emissions testing conducted in the US, EU, China and Japan. This review augments the de facto emission factor models in the US and EU with the results of PEMS, chassis testing and remote sensing studies covering thousands of passenger cars, light commercial vehicles, heavy-duty trucks and buses. Taking into account the differences and similarities in regulatory design across regions, we develop a detailed set of real-world NOx emission factors in each region that is differentiated by vehicle type and emission control level. To estimate current and future real-world NOx emissions, we combine these emission factors with existing Roadmap data and projections, including dates of implemented vehicle regulations, extensive historical data on diesel vehicle activity, sales and population, and validated projections of growth in vehicle activity through 2040.

A detailed discussion of the Roadmap emissions calculation methods can be found in Appendix II of Chambliss et al.[1]. A public version of the Roadmap model and documentation are available online at http://www.theicct.org/global-transportation-roadmap-model.

1.6 Emission inventory uncertainty

In a report to the European Commission, Kouridis et al.[89] assessed the uncertainty impacts of more than 50 inputs on resulting estimates of road vehicle emissions. Out of these inputs—which covered meteorological conditions, vehicle activity and traffic, urban and highway driving shares, speed profiles, and hot and cold emission factors—the authors found that for a country with good statistics on fuel consumption and vehicle stock, emission factors account for about 76% of the uncertainty in fleetwide NOx emissions inventory estimates. The authors also highlighted the effectiveness of calibrating to known fuel consumption statistics to reduce the uncertainty in emissions estimates.

In this study, we address the uncertainty present in our estimates of real-world NOx emission factors, as well as uncertainties associated with the conversion of engine-based emission limits for HDVs to distance-based estimates (when comparing real-world emissions to regulated emission limits). With respect to other sources of uncertainty—such as those arising from fuel consumption, activity by age and vehicle type, and driving modes—we do not perform a quantitative uncertainty evaluation. However, we note that the Roadmap model has been extensively validated with country-specific statistics and other international models, including parameters such as passenger and freight activity, vehicle-km traveled, fuel consumption and CO_2 emissions. At the global level, uncertainties in these parameters are typically less than 15%, though differences may be larger in regions with poorer data quality and availability.[90]

1.7 Review of emission factor studies

Supplementary Table 2 summarizes the list of studies reviewed that relate to real-world NOx emission factors. Studies of emissions testing apply a variety of methods, including PEMS, chassis dynamometer, and remote sensing. Bishop et al.[80] report on a new technique called On-Road Heavy-Duty Vehicle Emissions Monitoring System (OHMS), which consists of a drive-through tent outfitted with an exhaust plume capture and emissions measurement system. Rather than attempt to adjust the emission testing results of various studies and create a dataset with harmonized driving characteristics, we considered these differences qualitatively when comparing with the results of meta-analyses and established emission factor models.

2 Impact assessment methods

2.1 $PM_{2.5}$ and ozone concentrations

Country-level LDV and heavy-duty truck emissions are spatially allocated to 2.5’ x 2.5’ (arc minute) resolution using as spatial surrogates road network density (weighted 75%; obtained from Natural Earth (http://www.naturalearthdata.com/downloads/10m-cultural-vectors/roads/, Accessed Feb 6 2016) and population (weighted 25%) within each gridcell, following EPA's common practice for highway LDV and HDV emissions (more information at https://www.epa.gov/air-emissions-modeling, Accessed Feb 6 2016). For spatial allocation of emissions from buses, more weight was given to population (weighted 75%) considering the fact that buses are usually operated in urban or populated areas. Emissions were then aggregated to 2° x 2.5° resolution for input to the GEOS-Chem model.

Model emissions of NOx are emitted as the chemical species NO, which, given the rapid cycling of NO and NO_2, quickly establishes a NO/NO_2 ratio dependent on local physical and chemical conditions. The modeling results of our study reporting the impact of NOx emissions on ozone are thus nearly the same as would be reported if we emitted this much oxidized N as NO or NO_2.

Gridcell maximum annual average $PM_{2.5}$ and six-month average of the 1-hr daily maximum ozone concentrations are 249.7 μg/m^3 and 91.6 ppb for the Baseline 2015 and 249.8 μg/m^3 and 101.1 ppb for the Baseline 2040. Air pollution worsens from 2015 to 2040 in South Asia and much of Africa and Latin America, and improves in North America, Europe, and China, consistent with other studies (e.g.[59]).

2.2 $PM_{2.5}$-related health impact calculations

We obtained 1000 estimates of α, γ, and δ from the Integrated Exposure Response (IER) models for each health endpoint from the Institute for Health Metrics and Evaluation (IHME) website. Using these data, we then calculated 1000 age- and cause-specific relative risk estimates for each concentration level 0-300 μg/m^3 in 0.1 μg/m^3 increments, following Apte et al.[54]. We then created a lookup table of the mean of the 1000 relative risk estimates for each health endpoint, concentration level, and in the case of ischemic heart disease and stroke, 5-year age band. We use the annual average $PM_{2.5}$ concentration from the GEOS-Chem model to estimate $PM_{2.5}$-related premature deaths, consistent with Burnett et al.[55]

2.2.1 Future population projections

We use an existing 2030 population projection at 2.5' x 2.5' resolution which scales 2015 world population data from Columbia University's Center for International Earth Science Information Network (CIESIN) to 2030 using United Nations national population growth projections (http://www.un.org/esa/population/unpop.htm). We then developed country-wide growth rates to 2040 by calculating the ratios between U.N. population projections for 2040 and 2030 (2010 revision and applying them to each gridcell in the country. The growth rates range from 0.93-1.36 across the world with 86% of the countries having growth rates greater than 1 (i.e. increasing population). There are 39 countries/regions for which population projections are not available. The scaling factors for these countries were assigned as 1 (assuming no population growth or decrease from 2030 to 2040). Since these countries/regions are mostly small territories or less populated areas, the effect of ignoring population variation in these areas is marginal to the assessment of health impact globally. Countries/regions without individual population estimates are: American Samoa, Andorra, Anguilla, Antigua and Barbuda, Bermuda, British Virgin Islands, Cayman Islands, Commonwealth of Dominica, Cook Islands, Faeroe Islands, Gibraltar, Greenland, Guernsey, Isle of Man, Jersey, Kiribati, Liechtenstein, Marshall Islands, Monaco, Montserrat, Nauru, Niue, Norfolk Island, Northern Mariana Islands, Palau, Pitcairn, Saint Helena, Saint Kitts and Nevis, Saint Pierre and Miquelon, San Marino, Seychelles, Svalbard, Taiwan, Tokelau, Turks and Caicos Islands, Tuvalu, and Wallis and Futuna.

2.2.2 Future baseline mortality rate projections

For the chronic health endpoints, country-specific mortality rates for 186 countries were projected to 2040 using projections from the International Futures (IF) model version 7.15 base scenario (http://pardee.du.edu/access-ifs, Accessed March 30, 2016), which projects cause-specific and country-specific mortality rates to 2100. For each country and cause of death, we took the following steps: calculated 2015 and 2040 mortality rates by dividing total deaths by the total population, calculated rate adjustment factors by dividing the 2040 mortality rate by the 2015 mortality rate, multiplied the adjustment factors by the 2013 baseline mortality rates from the Institute for Health Metrics and Evaluation (IHME). As the IFs health endpoint categories are broader than the IHME health endpoints, we applied the IF cardiovascular ratio to the IHME ischemic heart disease (IHD) and stroke mortality rates, the IFs respiratory ratio to the IHME respiratory and chronic obstructive pulmonary disease

(COPD) mortality rates, and the IFs malignant neoplasms ratio to the IHME lung cancer mortality rate. This method assumes that each IHME health endpoint follows the same trajectory as the broader IFs health endpoint category. For acute lower respiratory infection (ALRI) for all countries, as well as the chronic diseases in countries not included in the IFs projections (limited to less populated countries such as island nations), present-day rates were held constant to 2040. The resulting 2040/2015 rate ratios are typically >1 (global mean = 1.3, 1.6, and 1.4 for cardiovascular, respiratory, and malignant neoplasms, respectively), indicating that incidence of these diseases rises over time, as is expected to occur as economies develop.

2.2.3 Uncertainty bounds for $PM_{2.5}$ and ozone mortality burdens

To calculate 95% confidence intervals for $PM_{2.5}$ mortality, we first calculated 1000 cause-specific relative risk estimates from the IER curves for each concentration level 0-300 μg/m^3 in 0.1 μg/m^3 increments. We then created three lookup tables of the mean and upper/lower 0.025 percentiles of the 1000 relative risk estimates for each concentration level, following the approach by Apte et al.[54] For each scenario, we matched the concentration level in each gridcell to the nearest 0.1 μg/m^3 concentration level in the lookup tables to find the mean and upper/lower 0.025 percentiles of the 1000 RR estimates, which were then used to estimate mortality. We then took the difference of the mortality estimates for two scenarios using the same mean and upper/lower 0.025 percentile of the 1000 RR estimates for each as the central estimate and 95% confidence bounds.

For ozone, we used the 95% confidence intervals in the relative risk estimate reported by Jerrett et al.[57] and assumed a normal distribution, following the approach taken by Silva et al.[3] We first calculated the burden of ozone exposure on premature mortality for each scenario, using either the mean or upper/lower 0.025 percentile of the relative risk reported by Jerrett et al.[57]

2.2.4 Comparison of baseline mortality burdens to literature values

We estimate that simulated $PM_{2.5}$ concentrations in the Baseline 2015 scenario result in 3.82 million premature deaths using 2015 population and baseline mortality rates, and 5.99 million using 2040 population and baseline mortality rates. These estimates are in line with other estimates in the literature (Supplementary Table 3). Our result for the 2015 Baseline scenario using 2015 population and baseline mortality rates are approximately 19% higher than the Global Burden of Disease 2010 study[54,56], which used the same Integrated Exposure Response functions though our study used updated baseline disease rates. Our result is approximately 10% lower than GBD 2015 estimate[21], and about 32% lower when we also use the updated GBD 2015 Integrated Exposure Response functions. Our result is also 16% higher than the 2010 estimate by Lelieveld et al.[2] and is 3% higher than the estimate by Anenberg et al.[58] Our result is 71% higher than that of Silva et al. (2016), which did not include child ALRI and quantified impacts of anthropogenic pollution only. Our result for the 2040 Baseline scenario using 2040 population and baseline mortality rates is 9% lower than the Lelieveld et al.[2] estimate for $PM_{2.5}$ in 2050. Lelieveld et al.[2] projected population and concentrations but not baseline mortality rates. The differences in results between these studies is due to a variety of methodology differences, including simulated concentrations, population, baseline mortality rates, and exposure-response functions.

Our estimates for ozone-related premature mortality are also in line with other estimates in the literature. For the Baseline 2015 scenario using 2015 population and baseline mortality rates and a low-concentration threshold (LCT) of 33.3 ppb, we estimate that ground-level ozone is associated with

216,000 premature respiratory deaths globally. This number is 15% lower than the GBD 2015 estimate[21], and is 56% lower than the values calculated by Anenberg et al.[58] and Silva et al.[3] A major difference in the latter two cases is the updated respiratory mortality rates used here are substantially lower than those used by the earlier studies. Our estimate for COPD only (a subset of all respiratory mortality in the core results) is 26% lower than the Forouzanfar et al.[21] estimate. Using a lower-concentration threshold of 41.9 ppb, the 5th percentile of the Jerrett et al.[57] concentrations and the upper end of the theoretical minimum risk level range used by Forouzanfar et al.[21], lowers our estimate by 28% to 155,000 premature deaths. We estimate that ozone mortality could increase to 500,000 in 2040, with most of the increase driven by population growth and changing baseline incidence rates (as indicated by the estimate of 474,000 premature deaths using the 2015 baseline concentrations with 2040 population and mortality rates).

2.2.5 Comparison of diesel NOx health impacts to literature values

As we isolate NOx emissions from diesel vehicles, and vehicles emit other pollutants such as black carbon and sulfur dioxide, it is important to consider the magnitude of our results in comparison to estimated impacts of all transportation emissions from other studies. However, drawing comparisons is complicated because of major methodological differences among the studies. We estimate that in 2015, on-road diesel NOx emissions were associated with 108,000 $PM_{2.5}$- and ozone-related premature deaths in these 11 markets, and 31,000 in China. Lelieveld et al.[2] estimated that land traffic was associated with 165,000 $PM_{2.5}$ and ozone deaths globally and 45,000 in China. Silva et al.[3] estimated that land transportation in 2005 was associated with 376,000 $PM_{2.5}$ and ozone deaths globally and 72,000 in China. While other studies estimated specifically on-road transportation[1,4], they were limited to $PM_{2.5}$ health impacts only. Comparing our estimate to others for China (global estimates are incomparable since we estimated health impacts for diesel emissions in the 11 markets only), we find that our estimate of the impact of on-road diesel NOx emissions in 2015 is 29-47% of the previous estimates, depending on the study[2-3]. Our estimate appears reasonably consistent with the Silva et al.[3] estimate given the magnitude of real-world diesel NOx emissions in China and baseline reductions in black carbon there. Our estimate appears high compared to the Lelieveld et al.[2] estimate, which was itself only ~60% of the Silva et al.[3] estimate. Major methodological differences in activity levels, emissions inventories, concentration-response functions, and baseline incidence rates preclude drawing conclusions from these comparisons.

For the future policy scenarios, we compare our results to those by Shindell et al.[5], who estimated the health, climate, and agricultural benefits of tighter vehicle emission standards globally. We compare results for implementation of Euro 6/VI standards in China, as that is the only country where Euro 6/VI standards are compared to a baseline in both studies which is also separately reported by both studies (we include Euro 6/VI in the baseline for India, and Shindell et al. [5] did not separate out Brazil). However, we found that implementing Euro 6/VI would avoid 88,000 $PM_{2.5}$ and ozone-related premature deaths in 2040 from NOx emission reductions, compared to the Shindell et al. [5] estimate of 61,000 premature deaths in 2030 from reductions of NOx, black carbon, sulfur dioxide, and other emissions. In addition to differences between assumptions in the emission scenarios (e.g. the proportion of Euro 4/IV and Euro 5/V vehicles on the road in the baseline scenario), our methods would produce higher estimates because we use the year 2040 (versus 2030), include ALRI, project baseline mortality rates to the future, and include the population aged 25-29 years. However, our methods could lead to lower estimates because we use the non-linear IERs (where Shindell et al. [5] extrapolated American Cancer Society study relative risk estimates linearly at high concentrations) and because we assessed only the impacts of NOx emission changes (where they included black carbon, sulfur dioxide, and others). These methodological differences along with others in vehicle activity levels and emission factors preclude our ability to draw conclusions about the magnitude of health benefits that could be achieved by the reductions in both real-world NOx emissions as well as other diesel vehicle emissions.

Supplementary Table 4 disaggregates total $PM_{2.5}$ and ozone-related premature deaths from NOx emissions by LDVs and HDVs, for both the impacts of excess diesel NOx and the benefits of policies to control diesel NOx emissions (shown graphically in Figure 4). We estimate that excess diesel NOx emissions from LDVs in the U.S. are associated with approximately 100 $PM_{2.5}$ and ozone-related premature deaths in 2015. This estimate is one-tenth of estimated impacts of excess diesel NOx from U.S. HDVs. Comparing our LDV result to previous studies examining the impact of the Volkswagen defeat device scandal in the U.S.[26-28], our estimate is on the same order of magnitude but higher for several reasons: 1) we include excess diesel NOx from LDVs unaffected by defeat devices, while the previous studies were limited to just those cars equipped with defeat devices; 2) we used the “proportional PAF” approach to estimate the influence of LDV and HDV excess NOx simultaneously, which has the effect of increasing the health impact result for LDVs compared with calculating only the LDV impact. Using the approach taken by the previous studies, where only LDVs were considered (and therefore estimated on the flatter portion of the Integrated Exposure Response curves compared with considering LDVs and HDVs simultaneously), we estimate that excess diesel NOx emissions from U.S. LDVs are associated with 56 premature $PM_{2.5}$- and ozone-related premature deaths in 2015. This estimate is remarkably similar to the previous estimates of 59[26], 46[27], and 59[28] premature $PM_{2.5}$- and ozone-related deaths associated with excess NOx from U.S. cars affected by defeat devices.

2.2.6 Years of life lost

To calculate years of life lost (YLL) associated with $PM_{2.5}$ and ozone mortality, we multiply the country- and disease-specific $PM_{2.5}$ and ozone mortality burden estimates for each scenario by the average YLL per death for that country and disease, following Anenberg et al.[58] Baseline YLL for ages under 5 years and over 25 years for 2015 were obtained from the Institute for Health Metrics and Evaluation Global Burden of Disease website. We then divided baseline YLL by baseline deaths for each country and disease to obtain a YLL per death estimate that could be multiplied by the $PM_{2.5}$ and ozone mortality results. The 2015 YLL/death values are used for both 2015 and 2040 analysis years, assuming that the YLL per death remains constant over time. Average YLL/death for the population age 25+ years across all countries is 17.9, 16.9, 17.0, 15.8, and 22.6 for ischemic heart disease, stroke, chronic respiratory disease, chronic obstructive pulmonary disease, and lung cancer, respectively. Global average YLL/death for the population under age 5 years for acute lower respiratory infections is 85.8. We calculate that $PM_{2.5}$ resulted in 70 million YLL in 2015 (3.5% lower than the value calculated by the Global Burden of Disease 2010 Study as reported by Apte et al.[54]). We also find that ozone resulted in 3.4 million YLL in 2015. For the Baseline in 2040, we estimate that $PM_{2.5}$ and ozone could result in 116 million and 8 million YLL, respectively.

2.2.7 $PM_{2.5}$ health benefits sensitivity analysis using Proportional PAF approach

As described in the Methods, our approach for calculating $PM_{2.5}$-related health benefits using the non-linear IER curves may underestimate the true benefits of diesel NOx emission policies by assuming that diesel NOx emission policies occur first, before any other air pollution control policies that may occur independently and concurrently. We therefore conducted a sensitivity analysis in which we applied the

"Proportional population attributable fraction (PAF) method" used by Chambliss et al.[1]. We apply the following equations in each gridcell, for each emission control scenario:

$$PM2.5\ mortality\ reduction = PM2.5\ mortality\ burden * \frac{(x_0 - x_s)}{x_0}$$

Where x_0 is the baseline $PM_{2.5}$ concentration in 2040 and x_s is the $PM_{2.5}$ concentration simulated for the emission control scenario.

2.3 Crop impacts

We estimate relative yield loss (RYL) and crop production loss (CPL) for three staple crops (maize, wheat, and soy) for each diesel NO_x emission scenario following van Dingenen et al.[62]. The methodology outlined in van Dingenen et al. [62] calculated the global impact of ground-level ozone on major staple crops, including wheat, maize and soy using concentration-based metrics. It has since been applied in other global assessments [e.g. 91], and was applied in this study's crop impact assessment to estimate RYL and CPL due to the ozone concentrations associated with each of the emission scenarios. Application of the methods requires gridded estimates of crop production (CP) and the start date of the growing season; these were available for the year 2000 in 1° x 1° grids globally derived as explained in van Dingenen et al. [62]. Gridded crop production estimates were derived by van Dingenen et al. [62] by distributing national, or state/province-level crop production estimates across a country using the Global Agro-Ecological Zones (GAEZv3) suitability index dataset for each crop (http://gaez.fao.org/Main.html). Results from the atmospheric modelling performed in this study provide ozone concentrations at a monthly time resolution. Therefore, the growing season start date was used to determine the 3 calendar months which were most representative of the growing season for each crop in each grid (i.e. the position of the actual growing season start date in relation to the mid-month date determined the month assigned as the first of the consecutive three-month average).

In line with the van Dingenen et al. [62] methodology, two types of concentration-based metrics were used to estimate ozone exposure. First, exposure was quantified as the growing season (3 month) average ozone concentration between 09:00 and 15:59 hours (referred to as the M7 metric) for wheat and rice, and between 08:00 and 19:59 hours (M12 metric) for soy and maize. The second exposure metric used was the sum of the positive differences between hourly average ozone concentrations and a threshold set at 40 ppb during daylight hours (08:00-19:59) across the relevant 3 month growing season for each crop; this is referred to as the AOT40 metric). M7/M12 and AOT40 were used by van Dingenen et al. (2009) to provide a measure of the variability in ozone crop impacts related to the different ozone concentration profiles. The M7/M12 ozone metrics both characterize growing season average concentrations, but do not capture changes in peak, episodic ozone concentrations. In contrast AOT40 metrics are more sensitive to changes in high ozone concentrations, but the cumulative nature of this metric means that it is more sensitive to inaccuracies in global ozone concentration modelling.[92] Metrics accounting for stomatal flux of ozone into the plant may be more biologically representative but are only available for wheat, potato and tomato and have not yet been parameterized adequately on a global scale.[93-95] For these reasons, and for comparability with previous global studies, we use both M7/M12 and AOT40 growing season exposure values and focus the resulting analysis on major agricultural regions experiencing large ozone concentration changes. These metrics provided a consistent relative change in crop impacts between scenarios, with differences in the absolute CPL resulting from differences in M7/M12 and AOT40-based exposure-response relationships, as outlined in van Dingenen et al. [62]. For each market and globally, CPL was represented as the average of the M7/M12 and AOT40 derived estimates, with variability characterized as the range between estimates. Monthly M7/M12, and AOT40 in 2 x 2.5° grids were calculated at canopy height (1m for wheat and soy; 2m for maize), and the metric values from the 2 x 2.5° grid which covered the majority of a 1 x 1° grid were selected as the M7/M12 and AOT40 concentrations for the smaller grid.

To calculate RYL, M7/M12 exposure-response relationships with the Weibull form shown in Equation 1 were used, with crop-specific parameters *a* and *b* derived from experimental exposure studies.[96-97] The RYL was calculated relative to the reference concentration specified in these exposure studies (Supplementary Table 6). For AOT40 linear exposure-response relationships (Equation 2) derived in Mills et al.[98] were applied. RYL was calculated using M7/M12 and AOT40 metrics for each grid, and it was assumed that there was no ozone-induced yield loss in those grids where the M7/M12 was below the reference concentrations, and when AOT40 = 0.

$$RYL = 1 - \exp[-MX/a^b]/\exp[-RC/a^b] \qquad \text{[Equation 1]}$$

Where MX = Ozone exposure metric (M7 for wheat and rice, and M12 for soy and maize)
RC = Reference concentration
a, b = Crop-specific parameters

$$RYL = cAOT40 \qquad \text{[Equation 2]}$$

Where AOT40 = Ozone exposure metric (Sum of daylight hourly ozone concentrations above 40 ppb during growing season)
c = Crop-specific parameter

For each grid, CPL was calculated using Equation 3 separately for the RYL estimated using the M7/M12 metrics and AOT40 metrics. For each country (including the 11 markets), these CPL values were aggregated to market-level and global CPL values. The market-level, and global crop production loss was expressed as a percentage of the theoretical crop production without ozone-induced crop loss (i.e. sum of total crop production and CPL). The average market and global CPL values from the M7/M12 and AOT40 estimates were calculated, and variability was characterized as the range between the M7/M12 and AOT40-derived CPL estimates.

$$CPL = \frac{RYL}{[1-RYL]} x\ CP \qquad \text{[Equation 3]}$$

For ozone-related crop yield impacts, despite relatively large uncertainties, there was a consistent decrease in the central estimates of crop production loss for all policy scenarios compared to the 2040 Baseline. For example, by 2040, next generation standards could result in an additional 1.8-2.8 million, 1.2-3.6 million, 0.93-0.95 million tonnes of maize, wheat, and soy production (at year 2000 crop production levels, based on 0.4-0.5%, 0.3-0.7%, and 0.6-0.7% average global avoided crop production loss). Regionally, we calculated greater avoided crop relative yield loss in Asia and Latin America compared to Europe and North America. For the major producing regions included in this study, the largest crop benefits from the four policy scenarios were estimated for maize and wheat in China and soy in Brazil.

2.4 Climate impacts

We calculated global radiative forcing (RF) of methane and ozone using regional RF efficiencies (mWm^{-2} per Tg of emission) from the multi-model study of Fry et al.[63], which are reported for North America, Europe, East Asia, and South America. Here we applied these efficiencies across the following four domains: (15° N – 90° N, 180° W – 50° W), (30°N – 90°N), 50°W – 60°E), (60°E – 180°E, 0° N – 90° N), and everywhere else. While most of the NO_X emissions reductions in this work occur in urban areas in the Northern Hemisphere, using these regional RF efficiencies allows us to capture the enhanced effect of NO_x emissions on ozone production and methane loss in regions with lower background concentrations in the southern hemisphere. Fry et al.[63] only reported standard deviations across the ensemble of model estimates for net RF of the combined impacts of ozone, methane, and sulfate aerosol. We apportioned this uncertainly equally, in quadrature, to ozone and methane to estimate a standard error of +/- 0.4 (North America, Europe, Asia) and +/- 1.4 (elsewhere) mWm^{-2} per Tg N yr^{-1}.

We calculated aerosol (nitrate, sulfate, and ammonia) RF from NO_X emissions using GEOS-Chem in conjunction with offline Mie theory calculations of aerosol optical properties and the LIDORT radiative transfer model to estimate the change in upward radiative flux from a pre-industrial atmosphere[64-66]. The adjoint of this model then tracks the impacts of an infinitesimal perturbation of the calculated upward flux backward in time to estimate the sensitivities of direct radiative forcing with respect to grid-cell changes in emissions ($mW\ m^{-2}$ per kg of emission). These sensitivities were calculated around the Baseline 2040 emissions and used to estimate the aerosol RF of all other emissions scenarios. We scaled the central estimates, lower, and upper bounds of direct aerosol RF based on model comparison to the ensemble of modeled RF by Myhre et al.[31], and included the impacts of aerosol cloud interactions following methods used by UNEP/WMO[67] by scaling the direct RF to the net effective RF.[99-100] The upper and lower bounds of the net aerosol impacts (95% CI's) were added in quadrature to the uncertainties in ozone and methane RF (twice the standard deviation across model ensembles, assuming these approximate standard error) to estimate the total RF uncertainty (95% CI) for each scenario relative to the baseline. Aerosol RFs have wide uncertainty bounds due to uncertainty about indirect radiative effects.[101]

Supplementary Tables

Supplementary Table 1. Region-specific policy implementation[1] timelines by diesel NOx scenario

Region	Type	Baseline[2]	Euro 6/VI	Strong RDE[3]	Next Gen
EU-28	Light	Euro 6 in 2014; adopted RDE phase-in 2017-2020	–	Strong RDE phase-in 2017-2020	Euro 7 in 2021
	Heavy	Euro VI in 2014	–	–	Euro VII in 2025
S. Korea	Light	Euro 6 in 2014	–	**1 year** after EU	**1 year** after EU
	Heavy	Euro VI in 2015	–	–	
Australia	Light	Euro 6 in 2018	–	**3 years** after EU	**3 years** after EU
	Heavy	Euro V in 2011	Euro VI in 2018	–	
India	Light	Bharat 6 in 2020[4]	–	**6 years** after EU	**6 years** after EU
	Heavy	Bharat VI in 2020	–	–	**4 years** after EU
Brazil	Light	Euro 4 in 2009	Euro 6 in 2018	**4 years** after EU	**4 years** after EU
	Heavy	Euro V in 2012	Euro VI in 2018	–	
Russia	Light	Euro 5 in 2016	Euro 6 in 2020	**5 years** after EU	**5 years** after EU
	Heavy	Euro V in 2016	Euro VI in 2020	–	
Mexico	Light	Euro 4 in 2009	Euro 6 in 2018	–	Tier 3 2021-2025
	Heavy	Euro IV in 2008	Euro VI in 2018	–	Same as U.S.
China	Light	China 5 in 2018[5]	China 6a in 2020; China 6b in 2023[6]	Strengthened RDE with China 6b	China 6b equivalent to U.S. Tier 3
	Heavy	China V in 2017	China VI in 2021	–	**2 years** after EU
United States	Light	Tier 3 phase-in 2017-2025	–	–	–
	Heavy	Tier 3[7]/EPA 2010	–	–	Next generation in 2025
Canada	Light	Harmonized with U.S.	–	–	–
	Heavy		–	–	Same as U.S.
Japan	Light	PNLTES[8] 2009	–	Same as EU	Same as EU
	Heavy	PNLTES 2016	–	–	Same as EU

Dash indicates no change from previous scenario. Grey fill indicates regions that have developed their own emission control programs from the ground up. RDE=Real Driving Emissions.

[1] Dates reflect the year of application to all sales and registrations of diesel vehicles. In most countries, this date is approximately one year later than the introduction of standards for new type approvals.
[2] Emission Limits scenario assumes same policy implementation dates as the Baseline.
[3] Strong RDE scenario applies to LDVs only.
[4] India's adoption of Bharat 6/VI for implementation in 2020 indicates potential for a shorter lag time for future EU standards.
[5] Nationwide dates are several years later than in certain provinces (e.g. Beijing, Shanghai, Guangzhou).
[6] Lower emission limits than Euro 6; assumes a real-driving emission program similar to Baseline RDE in the EU.
[7] US Tier 3 standards include complete Class 2b and 3 work trucks weighing up to 14,000 lbs.
[8] Post New Long Term Emission Standards. Source: TransportPolicy.net

Supplementary Table 2. Summary of emission factor studies reviewed.

Study	Region	Type of Testing	Sample size	Comments
Guo et al. (2014)[73]	Beijing	PEMS	9 (diesel buses)	Euro IV diesel buses usually emit the same NOx as Euro III
Wu et al. (2012)[37]	Beijing	PEMS	135	No significant NOx reduction from Euro II to Euro IV buses
Yao et al. (2015)[36]	Beijing	PEMS	18	Some China IV trucks emit > NOx than China III off-highway
Zhang et al. (2014)[38]	Beijing	PEMS	4 (diesel/hybrid)	Brake-specific NOx exceeds Euro V standard by 180%
Huo et al. (2012)[74]	China	PEMS	175	HDTs in China have similar NOx emissions (g/km) to EU
Liu et al. (2009)[102]	China	PEMS	75	HDT NOx emissions did not improve from Euro 0 to Euro III
Wu et al. (2015)[103]	China	PEMS	25	Poor SCR performance at low speeds and with limited urea
Chen and Borken-Kleefeld (2014)[104]	CHE	Remote sensing	114,500 (gasoline) 25,900 (diesel)	Positive speed and acceleration, uphill, hot-start; diesel cars emit 10-20x more NOx than gasoline cars of same MY
Chen and Borken-Kleefeld (2016)[18]	CHE	Remote sensing	62,000	No measured NOx deterioration for Euro 4 diesel cars, moderate deterioration (22%/10%) for Euro 2/3 diesel cars
Emisia (2016)[34]	EU	Modeled	Not applicable	EFs based on COPERT, which is used in official inventories
ERMES (2015)[105]	EU	Meta-analysis	Varies	Euro 5 diesel cars and LCVs emit ~0.8 and ~1.2 gNOx/km
Kadijk et al. (2015)[76]	EU	Chassis/PEMS	16	Euro 6 diesel cars emit ~500 mg/km and up to 8x the limit
Kouridis et al. (2010)[89]	EU	Modeled	Not applicable	EF uncertainty accounts for ~76% of NOx emission inventory variability in countries with good transport statistics
Ntziachristos et al. (2016)[45]	EU	Meta-analysis	Varies	Euro 6 diesel cars, Euro 5 LCVs emit 2x EFs in major models
Yang et al. (2015)[87]	EU	Chassis	73	Euro 6 diesel cars emit 4.9x NEDC when tested over WLTC
Velders et al. (2011)[33]	NLD (EU)	Modeled	Not applicable	Traffic composition and emission measurement uncertainty estimated at +/-25%
Beevers et al. (2012)[106]	UK	Remote sensing	74,614	Same dataset as Carslaw et al. (2011)[17]
Carslaw and Rhys-Tyler (2013)[35]	UK	Remote sensing	~70,000	Little evidence of NOx reduction from all diesel vehicle types
Carslaw et al. (2011)[17]	UK	Remote sensing	84,269	Urban-type gNOx/km estimates derived from UK gCO2/km
TRL (2009)[75]	UK	Meta-analysis	Varies	Testing data includes diesel cars only up to Euro 3
Bishop et al. (2013)[40]	CA (US)	Remote sensing	4293	MY2013 HHDTs emit 1.8x the 0.2 g/bhp-hr NOx limit
Bishop et al.	CA	OHMS	3088	MY2014 HHDTs emit 3x the 0.2 g/bhp-hr NOx

(2015)[80]	(US)			limit (n=28)
Barrett et al. (2015)[26]	US	Meta-analysis	2 (source study)	NOx EF by statistical distribution (Thompson et al., 2014)[48]
Bishop and Stedman (2008)[47]	US	Remote sensing	450–2000 per MY	Diesel and gasoline vehicles were analyzed together
Bishop and Stedman (2015)[19]	US	Remote sensing	2,593 (diesel)	Much higher real-world NOx for 2 L engine cars than trucks
Browning (1998)[78]	US	Meta-analysis	Varies	Recommends certification cycle BSFC and in-use FC for conversion of HDV emission limits.
Holland et al. (2016)[27]	US	Meta-analysis	2 (source study)	Highway, average urban NOx EFs (Thompson et al., 2014)[48]
Oldenkamp et al. (2016)[28]	US	Meta-analysis	2 (source study)	NOx EF average of PEMS tests (Thompson et al., 2014)[48]
US EPA (2015a)[10]	US	Not stated	Not stated	3.0 liter VW engines emit up to 9x Tier 2 NOx emission limit
US EPA (2015b)[39]	US	Assumption	Not applicable	Assume 50%/90% NOx reduction from EPA 2004–2007/2010
Zhai et al. (2008)[86]	US	PEMS	12	Bus NOx emission rates increase with vehicle specific power

Supplementary Table 3. Comparison of ambient $PM_{2.5}$ and ozone mortality burdens to other estimates in the literature.

Analysis	Annual $PM_{2.5}$-related premature deaths (millions)	Annual ozone-related premature deaths
GBD 2010 (Lim et al. 2012)[56], Apte et al. (2015)[54]	3.23	152,000 (COPD only)
GBD 2013 (Forouzanfar et al. 2015)[107]	2.93	217,000 (COPD only)
GBD 2015 (Forouzanfar et al. 2016)[21]	4.24	254,000 (COPD only)
Anenberg et al. (2010)[58] – anthropogenic only	3.7	700,000 (no threshold)
		490,000 (threshold=33ppb)
Lelieveld et al. (2015)[2] – 2010	3.3	-
Lelieveld et al. (2015)[2] – 2050	6.6	-
Silva et al. (2016)[3] – anthropogenic only	2.23	493,000
This study: Baseline 2015 + 2015 pop/mort	3.82 (IER 2010)	216,000 (threshold=33ppb)
	2.88 (IER 2015)	189,000 (COPD only, threshold=33ppb)
		155,000 (threshold=41.9ppb)
This study: Baseline 2015 + 2040 pop/mort	5.97	474,000 (threshold=33ppb)
This study: Baseline 2040 + 2040 pop/mort	5.99	500,000 (threshold=33ppb)

Supplementary Table 4. Regional sum of $PM_{2.5}$ and ozone-related premature deaths associated with a change in light-duty vehicle (LDV) versus heavy-duty vehicle (HDV) emissions in 2015 or 2040.

		Change due to excess NOx (2015)	Change due to future policies (2040)		
Region	Vehicle type	Baseline – Limits 2015	Euro 6/VI – Baseline 2040	Strong RDE – Euro 6/VI	NextGen – Strong RDE
Australia	LDV	0	0	0	0
	HDV	0	0	-	0
Brazil	LDV	0	-100	-100	0
	HDV	500	-3,900	-	-500
Canada	LDV	0	0	0	0
	HDV	100	100	-	-100
China	LDV	1,300	-11,200	-6,900	-1,600
	HDV	9,300	-76,500	-	-9,800
EU-28	LDV	6,900	0	-8,000	-2,900
	HDV	4,600	-700	-	-5,000
India	LDV	600	-100	-12,800	-2,900
	HDV	8,700	-700	-	-8,500
Japan	LDV	200	-100	-600	-100
	HDV	300	-600	-	-500
Mexico	LDV	0	0	-500	0
	HDV	100	-2,500	-	-500
Russia	LDV	300	-100	-300	-100
	HDV	500	-2,100	-	-500
South Korea	LDV	0	-100	-100	0
	HDV	200	-600	-	-200
United States	LDV	100	0	-100	0
	HDV	1,000	-400	-	-2,200
Rest of world	LDV	1,100	-500	-2,500	-700
	HDV	2,100	-3,900	-	-1,900

Supplementary Table 5. Regional $PM_{2.5}$- and ozone-related years of life lost (thousands) in 2015 and 2040.

		Baseline burden (all emission sources)		Change due to excess NOx	Change due to future policies relative to baseline in 2040 (scenario minus baseline)		
Region	Pollutant	Baseline 2015	Baseline 2040	Baseline – Limits 2015	Euro 6/VI	Strong RDE	NextGen
Australia	$PM_{2.5}$	1	1	0 (0)	0 (-5)	0 (-5)	0 (-5)
	Ozone	0	0	0 (100)	0	0	0
Brazil	$PM_{2.5}$	376	668	7 (2)	-69 (-10)	-71 (-11)	-80 (-12)
	Ozone	20	68	2 (10)	-16 (-23)	-16 (-24)	-18 (-26)
Canada	$PM_{2.5}$	85	99	1 (1)	-1 (-1)	-1 (-1)	-2 (-2)
	Ozone	6	9	0 (2)	0 (-2)	0 (-3)	-1 (-6)
China	$PM_{2.5}$	23,004	28,869	152 (1)	-1504 (-5)	-1621 (-6)	-1816 (-6)
	Ozone	1,166	2,341	27 (2)	-238 (-10)	-259 (-11)	-295 (-13)
India	$PM_{2.5}$	18,201	40,813	146 (1)	-16 (0)	-291 (-1)	-536 (-1)
	Ozone	1,348	4,089	52 (4)	-9 (0)	-81 (-2)	-147 (-4)
Japan	$PM_{2.5}$	354	307	5 (1)	-8 (-2)	-13 (-4)	-18 (-6)
	Ozone	29	42	1 (2)	-2 (-4)	-2 (-5)	-3 (-7)
Mexico	$PM_{2.5}$	108	142	2 (2)	-41 (-29)	-41 (-29)	-50 (-35)
	Ozone	16	45	0 (2)	-8 (-18)	-8 (-18)	-10 (-23)
Russia	$PM_{2.5}$	2,824	3,008	12 (0)	-35 (-1)	-39 (-1)	-48 (-2)
	Ozone	25	33	1 (3)	-1 (-4)	-2 (-5)	-2 (-7)
South Korea	$PM_{2.5}$	296	435	3 (1)	-10 (-2)	-12 (-3)	-15 (-4)
	Ozone	13	31	0 (3)	-1 (-3)	-1 (-4)	-2 (-6)
United States	$PM_{2.5}$	977	696	13 (1)	-7 (-1)	-8 (-1)	-39 (-6)
	Ozone	127	175	3 (2)	-3 (-2)	-3 (-2)	-10 (-5)
EU-28	$PM_{2.5}$	3,325	2,898	131 (4)	-8 (0)	-107 (-4)	-204 (-7)
	Ozone	109	122	10 (9)	-2 (-2)	-11 (-9)	-20 (-17)
Rest of world	$PM_{2.5}$	20,286	38,227	46 (0)	-77 (0)	-124 (0)	-172 (0)
	Ozone	551	1,277	10 (2)	-15 (-1)	-24 (-2)	-34 (-3)

Supplementary Table 6. Summary of parameters used to calculate crop impacts.

Crop type	*a* (standard error)	*b* (standard error)	Reference concentration (RC) (ppb)	Reference	*c*	Reference
Wheat	137 (6)	2.34 (0.41)	25	Lesser et al. 1990[97]	0.0163	Mills et al. 2007[98]
Maize	124 (2)	2.83 (0.23)	20	Lesser et al. 1990	0.00356	Mills et al. 2007
Soy	107 (3)	1.58 (0.16)	20	Lesser et al. 1990	0.0113	Mills et al. 2007

Supplementary References

78. Browning, L. (1998). Analysis of BSFCs and Calculation of Heavy-Duty Engine Emission Conversion Factors. Report prepared for the Motor Vehicle Emissions Laboratory at the U.S. Environmental Protection Agency. Retrieved from https://www3.epa.gov/otaq/models/mobile6/r02005.pdf
79. Khalek, I., Blanks, M., Merritt, P. (2013). Phase 2 of the Advanced Collaborative Emissions Study. Prepared for the Coordinating Research Council, Inc. Retrieved from https://energy.gov/eere/vehicles/downloads/advanced-collaborative-emissions-study-aces-phase-2-status-report
80. Bishop, G. A., Hottor-Raguindin, R., Stedman, D.H., McClintock, P., Theobald, E., Johnson, J.D., Lee, D.W., Zietsman, J., Misra, C. On-road heavy-duty vehicle emissions monitoring system. *Environ Sci Technol* **49**: 1639-1645 (2015).
81. Rubino, L., Bonnel, P., Hummel, R., Krasenbrink, A., Manfredi, U. (2007). Mobile measurement of pollutant emissions and fuel consumption of road vehicles in real-world driving situations using portable emission measurement systems (PEMS). Final report. European Commission Joint Research Centre, Ispra, Italy (61 pp. http://ec. europa. eu/clima/policies/transport/vehicles/docs/2007_pems_jrc_ 413194_en. pdf (accessed Mar. 30, 2016).
82. Ligterink, N., Kadijk, G., van Mensch, P., Hausberger, S., Rexeis, M. (2013). Investigations and real world emission performance of Euro 6 light-duty vehicles; TNO Report R11891; Netherlands Organization for Applied Scientific Research, Delft, The Netherlands; https://www.tno.nl/media/1969/investigations_emission_factors_euro_6_ld_vehicles_tno_2013.pdf.
83. Chambliss, S. and Bandivadekar, A. (2015). Accelerating progress from Euro 4/IV to Euro 6/VI vehicle emissions standards. International Council on Clean Transportation, Washington, DC. Retrieved from http://www.theicct.org/briefing-leapfrogging-to-euro-6-vi-mar2015
84. Air Resources Board, California (ARB) (2015). Draft Technology Assessment: Low Emission Natural Gas and Other Alternative Fuel Heavy-Duty Engines. Retrieved from https://www.arb.ca.gov/msprog/tech/techreport/ng_tech_report.pdf
85. Air Resources Board, California (ARB) (2015). Draft Technology Assessment: Lower NOx Heavy-Duty Diesel Engines. Retrieved from https://www.arb.ca.gov/msprog/tech/techreport/diesel_tech_report.pdf
86. Zhai, H., Frey, H., and Rouphail, N. A vehicle-specific power approach to speed- and facility-specific emissions estimates for diesel transit buses. *Environ Sci Technol* **42**: 7985-7991 (2008).
87. Yang, L., Franco, V., Mock, P., Kolke, R., Zhang, S., Wu, Y., German, J. Experimental assessment of NOx emissions from 73 Euro 6 diesel passenger cars. *Environ Sci Technol* **49**: 14409-14415 (2015).
88. Mishra, G.S., Fulton, L., Yeh, S., Kyle, P., McCollum, D.L., Miller, J., Cazzola, P. (2014). Summary of International Transport Energy Modeling Workshop. ITS, UC Davis, CA, USA (19 October 2014). Available at http://files.ctctcdn.com/0245d04e001/dbc3bf77-4ecb-4fb3-80e0-4ec52cd9ffc4.pdf.
89. Kouridis, C., Gkatzoflias, D., Kioutsioukis, I., Ntziachristos, L., Pastorello, C., Dilara, P. (2010). Uncertainty Estimates and Guidance for Road Transport Emission Calculations. European Commission Joint Research Centre, Institute for Environment and Sustainability. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/14202/1/uncertainty%20eur%20report%20final%20for%20print.pdf
90. Burton, A., Shao, Z., Miller, J., Façanha, C. (2013). Global Transportation Roadmap Model Comparative Analysis of Activity and Emissions. International Council on Clean Transportation, Washington, DC. Retrieved from http://www.theicct.org/info/assets/RoadmapV1/ComparativeAnalysis2013.pdf

91. Avnery, S., Mauzerall, D.L., Liu, J., Horowitz, L.W. Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage. *Atmos Environ* **45**: 2284-2296 (2011).
92. Tuovinen, J.P., Simpson, D., Emberson, L., Ashmore, M., Gerosa, G. Robustness of modelled ozone exposures and doses. *Environ Pollut* **146**: 578–586 (2007).
93. Mills, G., Hayes, F., Simpson, D., Emberson, L., Norris, D., Harmens, H., et al. Evidence of widespread effects of ozone on crops and (semi-)natural vegetation in Europe (1990-2006) in relation to AOT40-and flux-based risk maps. *Glob Chang Biol* **17**: 592–613 (2011).
94. Grantz, D.A. Diel trend in plant sensitivity to ozone: Implications for exposure- and flux-based ozone metrics. *Atmos Environ* **98**:571–580 (2014).
95. Hayes, F., Wagg, S., Mills, G., Wilkinson, S., Davies, W. Ozone effects in a drier climate: Implications for stomatal fluxes of reduced stomatal sensitivity to soil drying in a typical grassland species. *Glob Chang Biol* **18**: 948–959 (2012).
96. Adams, R.M., Glyer, J.D., Johnson, S.L., McCarl, B.A. A reassessment of the economic effects of ozone on united states agriculture. *Japca-the Journal of the Air & Waste Management Association* **39**:960-968 (1989).
97. Lesser, V.M., Rawlings, J.O., Spruill, S.E., Somerville, M.C. Ozone effects on agricultural crops - statistical methodologies and estimated dose-response relationships. *Crop Science* **30**:148-155 (1990).
98. Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., et al. A synthesis of AOT40-based response functions and critical levels of ozone for agricultural and horticultural crops. *Atmos Environ* **41**: 2630–2643 (2007).
99. Boucher, O., et al. (2013). Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
100. Lacey, F., and Henze, D. Global climate impacts of country-level primary carbonaceous aerosol from solid-fuel cookstove emissions. *Environmental Research Letters* **10**: 114003 (2015).
101. Chen, W.-T., Lee, Y.H., Adams, P.J., Nenes, A., Seinfeld, J.H. Will black carbon mitigation dampen aerosol indirect forcing? *Geophys Res Lett* **37**(9), doi: 10.1029/2010GL042886 (2010).
102. Liu, H., He, K.B., Lents, J.M., Wang, Q.D., Tolvett, S. Characteristics of Diesel Truck Emission in China Based on Portable Emissions Measurement Systems. *Environ Sci Technol* **43**: 9507-9511 (2009).
103. Wu, X. et al. Real-world emissions and fuel consumption of diesel buses and trucks in Macao; From on-road measurement to policy implications. *Atmos Environ* **120**: 393-403 (2015).
104. Chen, Y.C., and Borken-Kleefeld, J. Real-driving emissions from cars and light commercial vehicles Results from 13 years remote sensing at Zurich/CH. *Atmos Environ* **88**: 157-164 (2014).
105. European Research Group on Mobile Emission Sources (2015). Diesel light duty vehicle NOx emission factors. Retrieved from http://www.hbefa.net/e/pdf/ERMES_NOX_EF_V20151009.pdf
106. Beevers, S., Westmoreland, E., de Jong, M.C., Williams, M.L., Carslaw, D.C., Trends in NOx and NO2 emissions from road traffic in Great Britain. *Atmos Environ* **54**:107-116 (2012).
107. Forouzanfar, M.H., et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* **386**(10010): 2287–2323 (2015).