# Air quality in Enclosed Railway Stations: quantifying the impact of diesel trains through deployment of multi-site measurement and random forest modelling

**Citation for published version:**
Font, A, Tremper, AH, Lin, C, Priestman, M, Marsh, D, Woods, M, Heal, MR & Green, DC 2020, 'Air quality in Enclosed Railway Stations: quantifying the impact of diesel trains through deployment of multi-site measurement and random forest modelling', *Environmental Pollution*, vol. 262, 114284. https://doi.org/10.1016/j.envpol.2020.114284

**Digital Object Identifier (DOI):**
10.1016/j.envpol.2020.114284

**Link:**
Link to publication record in Edinburgh Research Explorer

**Document Version:**
Peer reviewed version

**Published In:**
Environmental Pollution

# Air quality in Enclosed Railway Stations: quantifying the impact of diesel trains through deployment of multi-site measurement and random forest modelling

Anna Font[1], Anja H. Tremper[1], Chun Lin[2], Max Priestman[1], Daniel Marsh[1], Michael Woods[3], Mathew R. Heal[2] and David C. Green[1]

[1] MRC PHE Centre for Environment and Health, King's College London, 150 Stamford St London, SE1 9NH, UK

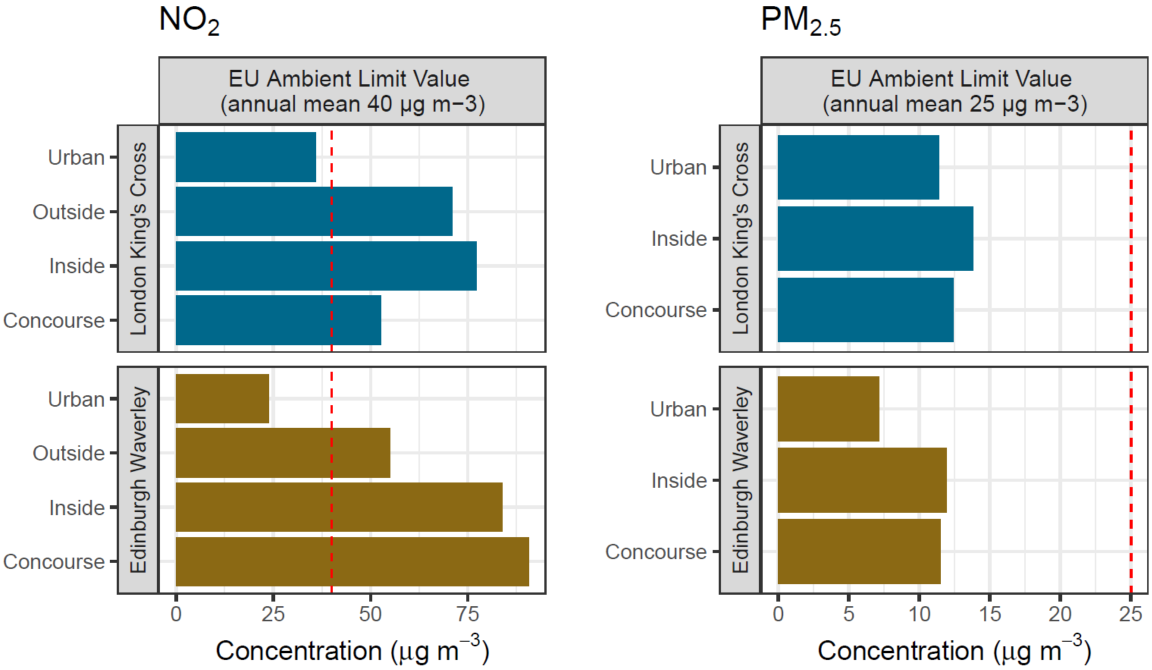
[2] School of Chemistry, University of Edinburgh, Joseph Black building, David Brewster Road, Edinburgh EH9 3FJ, UK

[3] RSSB, The Helicon, 1 South Place, London EC2M 2RB, UK

## Abstract

Concentrations of the air pollutants ($NO_2$ and particulate matter) were measured for several months and at multiple locations inside and outside two enclosed railway stations in the United Kingdom – Edinburgh Waverly (EDB) and London King's Cross (KGX) – which, respectively, had at the time 59% and 18% of their train services powered by diesel engines. Average concentrations of $NO_2$ were above the 40 µg $m^{-3}$ annual limit value outside the stations and were further elevated inside, especially at EDB. Concentrations of $PM_{2.5}$ inside the stations were 30-40% higher at EDB than outside and up to 20% higher at KGX. Concentrations of both $NO_2$ and $PM_{2.5}$ were highest closer to the platforms, especially those with a higher frequency of diesel services. A random-forest regression model was used to quantify the impact of numbers of different types of diesel trains on measured concentrations allowing prediction of the impact of individual diesel-powered rolling stock.

## Abstract Art



**Keywords**: Diesel exhaust, diesel trains, enclosed railway stations, random forest

**Capsule**: Diesel emissions in two UK enclosed railway stations

# Introduction

Rail is usually considered a green mode of transport compared with road and air in terms of its relative impact on climate change (Givoni et al., 2009). However, rail services, particularly those operated by diesel-powered trains, also emit air pollutants: in the European Union (EU-27) diesel trains were estimated to contribute 2.0%, 2.8% and 2.5%, respectively, of mobile sources of nitrogen oxides, particulate matter <2.5 µm in diameter ($PM_{2.5}$) and black carbon in 2005 (Borken-kleefeld and Ntziachristos, 2012). Diesel emissions are widely considered to be harmful to human health; in 2012 the World Health Organisation classified diesel engine exhaust as carcinogenic (WHO-IARC, 2012). In the United Kingdom (UK), although electrification of the rail network is expanding, only 34% of the routes are electrified (Department for Transport, 2017) and the railway industry used around 700 million litres of diesel to run passenger and freight services (Office of Road and Rail, 2017).

The concentration of air pollutants in enclosed railway stations is partly influenced by the outdoor air drawn inside plus all the contributions from internal sources; these include $NO_x$ and particles from the exhaust of diesel-powered trains; particles generated by the wear of trains (e.g. wheels, brakes); and $NO_x$ and particles from cooking in food outlets (Chong et al., 2015). A number of studies have reported measurements of air quality in subway systems, for example in Stockholm (Johansson and Johansson, 2003), Helsinki (Aarnio et al., 2005), Seoul (Kim et al., 2008), New York (Vilcassim et al., 2014), Athens (Barmparesos et al., 2016), Rome(Perrino et al., 2015) and Barcelona (Martins et al., 2015; Querol et al., 2012), but these have focused on $PM_{10}$ and $PM_{2.5}$ due to the predominance of wear emissions in the absence of diesel trains on these networks. Fewer studies have been conducted in ground-level railway environments. In the UK air quality has been measured at London Paddington (Chong et al., 2015) and Birmingham New Street (Hickman et al., 2018) stations, but these were based on short campaigns (less than 7 days and 10 weeks, respectively) so assessment of concentrations relative to long-term limit values or against rolling stock characteristics was difficult. As there is currently no legislation regulating public exposure to the indoor concentrations of air pollutants, most studies in both subway and railway environments compared measured concentrations to limit values in outdoor air.

The aim of this study was to characterize the impact of diesel-powered train emissions on concentrations of nitrogen dioxide ($NO_2$) and $PM_{2.5}$ and $PM_{10}$ inside two enclosed stations in the UK, Edinburgh Waverley and London King's Cross. Measurements were made for several months at multiple locations inside and also outside each station. A specific objective was to highlight the type of rolling stock that most influenced the concentrations measured inside, via a data mining approach. Decision trees are a common data mining approach, and have previously been applied in the railway industry, for example to improve rail network velocity (Li et al., 2014) and to evaluate service quality (de Oña et al., 2014; 2016). However, whilst they are simple to implement and interpretation is straightforward, the prediction accuracy can be low (James et al., 2013). Random-forest (RF) is a machine-learning algorithm that can be used for classification or regression and represents an improvement in prediction accuracy compared to decision trees. RF produces multiple trees which are then combined to yield a single consensus prediction at the expense of some loss in interpretation (James et al., 2013). RF presents several advantages: it is a simple non-linear regression model that requires few parameters to be chosen; it is robust to parameter specifications; it can handle high-order interactions among predictive variables; and it is robust to over-fitting (Faganeli Pucer and Štrumbelj, 2018). RF has been popularized in many areas in recent years, including in air quality applications, such as predicting $PM_{2.5}$ concentrations from satellite imagery (Huang et al., 2018; Brokamp et al., 2018) and removing meteorological confounding in pollutant concentrations (Grange and Carslaw, 2019) for trend estimations (Faganeli Pucer and Štrumbelj, 2018); but not previously in the context presented here.

# Methods

## Experimental campaigns and measurements

Monitoring was carried out at two UK railway stations: Edinburgh Waverley (EDB) and London King's Cross (KGX). These are both large enclosed railway stations (without active ventilation) directly managed by Network Rail (rather than by train operating companies), with a large number of train movements a day but contrasting proportions of diesel-powered services. Averaged over the period August – December 2018, Edinburgh Waverley had 828 trains day$^{-1}$, of which 59% were scheduled to run on diesel, whilst London Kings' Cross had 420 trains day$^{-1}$ of which 18% were scheduled to run on diesel. Most of the diesel-powered trains in Edinburgh Waverley were Sprinter Diesel Multiple Units (DMUs) (Class 15X and Class 17X) (83%) followed by High Speed Trains (HSTs) (6%), 220/221 (Voyagers) (5%); and diesel locomotive or locomotive-hauled trains (5%). At London King's Cross, around 62% of the diesel stock were HSTs and 33% were Class 180 Adelante (a diesel-hydraulic multiple-unit passenger train). EDB and KGX are both large enclosed stations of similar size. Plans of both stations are shown in Figure S1 and S2, respectively. EDB has both terminus and through tracks, with the primary openings for the through tracks at either end of the station, aligned with the main wind direction (south-west to north-east direction). There are two additional openings, the vehicular/pedestrian access ramps located north and south of ED3N and ED3 monitoring locations shown in Figure S1. There ramps are used by delivery vans and lorries but these predominantly occur at night, during station closure periods. KGX is a terminal station with the tracks 0 – 8 aligned in a north to south direction and housed under a double arched glazed roof (Figure S2). Platform 0, whilst under the main station roof, is partially enclosed with a lower roof. Platforms 9 to 11 are separated from the other set of tracks and positioned at an angle to the main station. These two areas are linked by a semi-circular departure concourse area (Figure S2). The primary external opening is where the trains enter and exit at the north end of the station. Other significant openings are created by the station access doors to the south side of the station.

The measurements of $NO_2$, $PM_{10}$ and $PM_{2.5}$ were undertaken from May to November 2018 at EDB and from August to December 2018 at KGX. The simultaneous multisite measurement of $NO_2$ was made using Palmes-type passive diffusive tubes (PDTs) (Palmes et al., 1976). These were exposed between 2–4 weeks at 8 locations inside each station, 3 locations outside EDB station, 2 outside KGX, and at one urban background site. The specific locations of the $NO_2$ measurements inside and outside of both railway stations are shown in the Supplementary Information (Table S1–S2; Figure S1–S2). PDTs were deployed in triplicate for every exposure period. All triplicates showed good measurement consistency with an average intra-site coefficient of variation of 4.9% at EDB (range: 3.1-7.6%) and 3.6% at KGX (range: 1.6-6.9%). The PDTs at the urban background sites were co-located with a reference chemiluminescence instrument traceable to national metrological standards. In addition, to the PDT $NO_2$ data, hourly $NO_2$ concentrations were also measured inside each station with a reference chemiluminescence instrument (ENVEA, Environment AC31M, Poissy, France) traceable to national metrological standards. These measurements were made for a period of 8 weeks at the ED4-OfficeDepot at EDB and for 6 weeks at the LO1-Platform0/1 location at KGX. The location at EDB was not close enough to a platform to permit specific analysis of relationships between $NO_2$ and rolling stock characteristics. Extension of reference $NO_2$ measurements at other locations within the stations were not possible due to power and space restrictions.

$PM_{10}$ and $PM_{2.5}$ were measured using an Osiris Airborne Particle Monitor (Turnkey Instruments Ltd., Cheshire, UK) concurrently at four of the inside locations in each station, and at the urban background site (Figure S1; Figure S2). The Osiris instruments measure the particles suspended in the air in four fractions (total suspended particles, $PM_{10}$, $PM_{2.5}$ and $PM_1$) by means of the light they scatter and from now on are referred to as optical particle counters (OPCs). Co-location of these monitors with a reference instrument (TEOM-FDMS, Thermo Scientific, Waltham, MA, US) was undertaken at the start and at the end of the measurement campaigns at the Marylebone Road national network monitoring station in central London (51°31’21’’N; 0°9’16.56’’W). The roadside location for the co-location was chosen due to the presence of small particles coming from vehicular diesel combustion, similar to the particular mix in the railway stations. Loss of volatile PM due to the heated inlet of the Osiris was corrected using a volatile correction model approach (Green et al., 2009) and corrected measurements from OPCs correlated well to the reference concentrations (Figure S4) so no further correction was needed. Further details of the correction method are in the Supplementary Information. The 15-minute resolution Osiris $PM_{10}$ and $PM_{2.5}$ data were aggregated to hourly means.

The timetables for the numbers of different type of trains operating in each station were obtained from www.realtimetrains.co.uk from July 2018 onwards. Railway industry representatives provided updated information where there were some mismatches between the rolling stock categories reported on the website and the actual trains in use.

Hourly outdoor wind (speed and direction), temperature, pressure and relative humidity data were obtained from the NOAA ISD network using the R-package *worldmet* (Carslaw, 2019) for Edinburgh and London City airports, located at 10 km and 12.3 km from their respective railway stations (Figure S1-S2).

## Statistical analysis

The station increment above the urban background concentration was used to quantify the contribution of internal sources, similar to the approach described by Lenschow (2001). This assumes that there is a background concentration in the station similar to the urban-wide background. This approach may present large uncertainty when using it especially at high time resolution (e.g. hourly) because one or both of the urban or station background measurements may be transiently affected by localised variations. In this work, station increment in concentration variables are denoted by ‘Δ’.

RF regression models were built to reproduce the hourly concentration in $PM_{2.5}$ and $NO_2$, and also in $\Delta PM_{2.5}$ and $\Delta NO_2$, as the dependent variables, respectively. Multiple explanatory variables were included in the RF models including information about train numbers and rolling stock, and meteorological conditions. The selection criteria to choose the explanatory variables was based on the trend in the hourly $PM_{2.5}$ (and $\Delta PM_{2.5}$) versus the explanatory variable. The trend was evaluated by means of the Siegel’s Repeated Median Estimator. This is a nonparametric approach to linear regression that is robust to outliers in the dependent variable. All possible slopes between each point and the others is computed and the slope estimator is the median of these slopes. Only those explanatory variables that had a statistically significant slope were included in the RF model.

To avoid co-linearity, which can potentially lead to the wrong identification of the relevant predictors in the statistical model (Dormann et al., 2013), several RF models were built avoiding explanatory variables with correlation $R > 0.7$. Each RF was built using 500 trees and its performance evaluated by means of the correlation coefficient and the mean-square-error (MSE).

For each RF regression model, partial dependence plots representing the marginal effect of the explanatory variables on the predicted outcome were produced. These quantify the relationship between the dependent variable and the explanatory variable and were used to quantify the impact of numbers of diesel trains and rolling stock types on the dependent variable. The slope of the reduced-major-axis (RMA) regression fit to the partial dependence plots was used to identify the rolling stock that should be prioritized for an emission reduction activity.

Different levels of significance were considered in all the statistical procedures: $p<0.001$ (coded as ***); $p<0.01$ (**); $p<0.05$ (*); and $p<0.1$ (+).

# Results and discussion

## Overall concentrations

Table 1 presents a summary of the PDT $NO_2$ and Osiris $PM_{10}$ and $PM_{2.5}$ concentrations for each location at each station. At EDB, PDT $NO_2$ concentrations differed significantly between inside the station, outside the station and at the background site, with average concentrations across the measurement campaign for all locations of a given type of 86.5 µg $m^{-3}$, 55.0 µg $m^{-3}$ and 23.8 µg $m^{-3}$ (the reader is pointed to see the graphical abstract). Location ED14-Platform 14 had the highest $NO_2$ concentration averaged across all the exposure periods: 103.1 ± 7.8 µg $m^{-3}$ (± 1 standard deviation). This location was close to several terminating railway lines. Other trackside measurement locations had slightly lower concentrations: ED2-Waverley steps (91.3 ± 4.4 µg $m^{-3}$; i.e. 11% less) and ED1-Platform 11 (77.3 ± 3.6 µg $m^{-3}$; 25% less). Sites on the main concourse (ED3 and ED3N) had higher $NO_2$ concentrations (89.7 and 94.7 µg $m^{-3}$, respectively) than some other trackside sites. The concourse area is a somewhat enclosed area, bounded by two platforms and the main building which can lead up to the build-up of pollutants. Also, these sites are adjacent to the north access ramp into the station from Waverley Bridge (Figure S1). Sites ED4 and ED4S had the lowest $NO_2$ concentrations inside the station, consistent with these sites being the furthest from the busiest platforms.

At KGX, average $NO_2$ concentrations inside (71.4 µg $m^{-3}$) and outside (71.0 µg $m^{-3}$) the station were similar, but both were significantly higher than the urban background (36.0 µg $m^{-3}$). The highest $NO_2$ concentrations were measured at sites closest to the main cluster of tracks (sites LO1, PL2/3, PL4, LO2, PL6/7 on Platforms 0–8), with an average of 78.3 ± 7.1 µg $m^{-3}$, whereas the lowest concentrations were measured at sites on the concourse and the mezzanine, with an average of 52.7 ± 0.7 µg $m^{-3}$ (32.7% less). Site LO3-Platform 9 had concentrations of $NO_2$ in between these two groupings (66.8 ± 4.9 µg $m^{-3}$) consistent with this location being within the platform area but with fewer tracks nearby.

Comparing stations, the average $NO_2$ concentrations at KGX were lower than at EDB (71.4 and 86.5 µg $m^{-3}$, respectively) despite urban background $NO_2$ concentrations being higher in London (29.7 µg $m^{-3}$) than in Edinburgh (13.7 µg $m^{-3}$). The mean station increment in $NO_2$ ($\Delta NO_2$) at EBD was 1.7 times higher than that measured at KGX (72.8 µg $m^{-3}$ and 41.7 µg $m^{-3}$, respectively). The higher $\Delta NO_2$ inside EDB is fully consistent with the factor 6 times greater numbers of diesel trains in EDB (~490 trains $day^{-1}$) compared to KGX (~80 trains $day^{-1}$). Average $NO_2$ concentrations inside both stations exceeded the EU annual limit value of 40 µg $m^{-3}$ set for outdoor air.

Breaches of the 200 µg m$^{-3}$ hourly limit value for $NO_2$ were assessed for the periods when hourly data were available. There were no breaches at EDB (May-June 2018; $N = 2160$ hours) but at KGX, where the measurements were closer to the platform, there were 29 breaches (Aug-Oct 2018; $N = 1566$). The lack of breaches of the hourly limit value of $NO_2$ at EDB is probably due to the distance between the measurement location and any track line.

Temporal correlation coefficients between the PDT $NO_2$ concentrations at EDB and the exposure-averaged $NO_2$ concentration from the reference analysers at the urban background sites was low and lacked significance ($r$ ranged from -0.2 to 0.5) (data not shown). This supports the interpretation that the $NO_2$ inside the station were dominated by strong local sources independent of the general background meteorology. The similar correlation analysis at KGX also showed no statistical significance but as the number of temporal $NO_2$ observations was limited to only 4, this result may not be robust.

On average, concentrations of particulate matter were similar in both stations (Table 1). For $PM_{10}$, average concentrations ranged from 17 to 25 µg m$^{-3}$ across the four inside locations at EDB, and from 18 to 30 µg m$^{-3}$ for the four inside locations at KGX. For $PM_{2.5}$, concentrations ranged from ~10 to 15 µg m$^{-3}$ at both stations. However, background concentrations of both $PM_{10}$ and $PM_{2.5}$ were slightly higher in London (15 and 12 µg m$^{-3}$, respectively) than they were in Edinburgh (10 and 7 µg m$^{-3}$). Overall, the campaign-average concentrations of both $PM_{10}$ and $PM_{2.5}$ at all locations in both stations were below their respective EU annual limit values of 40 and 25 µg m$^{-3}$, but $PM_{2.5}$ concentrations were above the WHO air quality guideline concentration of 10 µg m$^{-3}$. Substantial hour-to-hour variability and some highly elevated concentrations were observed at some locations. The highest $PM_{10}$ concentrations were measured in EDB with a maximum hourly concentration of 804 µg m$^{-3}$ measured at ED1-Platform 11. The maximum $PM_{10}$ concentration at KGX was only a quarter of that observed at EDB (170 µg m$^{-3}$ at LO2-Platform 4). The highest $PM_{2.5}$ hourly concentrations were measured at ED1-Platform 11 and ED2-Waverley Steps (117 µg m$^{-3}$) whereas at the other locations in Edinburgh, the maximum concentrations were only half this value. At KGX, the maximum $PM_{2.5}$ hourly concentration measured at LO3-Platform 9 (110 µg m$^{-3}$) was twice that observed at the other monitoring locations.

*Table 1. Summary statistics for the 4-weekly PDT $NO_2$ concentrations and the hourly Osiris $PM_{10}$ and $PM_{2.5}$ concentrations for all locations inside and outside Edinburgh Waverley and London King's Cross. N indicates the number of exposure periods for the $NO_2$ PDT measurements or the number of hours of available data with PM measurements. All concentrations are in μg $m^{-3}$.*

| | **$NO_2$** | | | | **$PM_{10}$** | | | **$PM_{2.5}$** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Station** | Code | Mean (±1 s.d.) | Range | *N* | Mean (± 1 s.d.) | Range | *N* | Mean (± 1 s.d.) | Range | *N* |
| **Edinburgh Waverley (EDB)** | ED1 | 77.3 (±3.6) | 73.3 - 82.5 | 8 | 24.2 (±28.5) | 1.6 – 804 | 2841 | 11.9 (±8.2) | 0.8 – 117 | 2841 |
| | ED14 | 103 (±7.8) | 91.1 - 114 | 8 | -- | -- | -- | -- | -- | -- |
| | ED2 | 91.3 (±4.4) | 85.8 - 99.1 | 8 | 17.0 (±10.9) | 1.4 – 180 | 4363 | 11.7 (±8.3) | 0 – 117 | 4363 |
| | ED3 | 89.7 (±8.0) | 77.3 - 99.1 | 8 | 25.3 (±27.9) | 1.9 – 335 | 2996 | 11.5 (±7.4) | 0.8 - 64.7 | 2996 |
| | ED3N | 94.7 (±7.6) | 80.7 - 101 | 8 | -- | -- | -- | -- | -- | -- |
| | ED3W | 87.6 (±6.6) | 76.6 - 95.0 | 8 | -- | -- | -- | -- | -- | -- |
| | ED4 | 72.4 (±6.8) | 61.7 - 82.0 | 8 | 18.1 (±13.1) | 1.2 – 222 | 3783 | 9.9 (±5.4) | 0.5 – 60.0 | 3783 |
| | ED4S | 75.8 (±6.6) | 66.3 - 85.7 | 8 | -- | -- | -- | -- | -- | -- |
| | PS | 48.9 (±11.1) | 41.6 - 62.0 | 8 | -- | -- | -- | -- | -- | -- |
| | WB | 59.9 (±8.6) | 47.3 - 82.4 | 8 | -- | -- | -- | -- | -- | -- |
| | MS | 56.2 (±6.9) | 45.4 - 67.6 | 8 | -- | -- | -- | -- | -- | -- |
| | ED5 | 23.8 (±4.2) | 19.7 - 30.9 | 8 | 10.0 (±6.4) | 0.3 - 78.3 | 4642 | 7.2 (±4.9) | 0.2 – 62.9 | 4642 |
| **London King's Cross (KGX)** | LO1 | 71.4 (±11.1) | 54.9 - 78.3 | 4 | 18.6 (±8.0) | 2.1 - 85.0 | 2662 | 14.5 (±6.8) | 1.8 – 60.2 | 2662 |
| | PL2/3 | 87.1 (±5.3) | 79.8 - 92.5 | 4 | -- | -- | -- | -- | -- | -- |
| | PL2/3N | 70.5 (± --) | 70.5 - 70.5 | 1 | -- | -- | -- | -- | -- | -- |
| | PL4 | 75.4 (±4.6) | 70.3 - 79.9 | 4 | -- | -- | -- | -- | -- | -- |
| | LO2 | 79.1 (±6.4) | 69.8 - 84.4 | 4 | 30.3 (±11.7) | 5.7 – 170 | 2658 | 13.6 (±6.3) | 2.3 – 47.1 | 2658 |
| | PL6/7 | 86.0 (±6.3) | 76.8 - 90.8 | 4 | -- | -- | -- | -- | -- | -- |
| | LO3 | 66.8 (±4.9) | 60.7 - 71.7 | 4 | 18.2 (±10.1) | 2.9 – 141 | 1679 | 11.5 (±7.3) | 1.7 – 110 | 1679 |
| | CC | 53.2 (±5.6) | 45.2 - 57.8 | 4 | -- | -- | -- | -- | -- | -- |
| | LO4 | 52.2 (±4.6) | 45.9 - 56.2 | 4 | 20.2 (±7.8) | 2.6 – 74 | 2657 | 12.7 (±6.2) | 1.8 – 51.3 | 2657 |
| | FC17 | 75.9 (±8.3) | 64.6 - 82.3 | 4 | -- | -- | -- | -- | -- | -- |
| | FC2 | 66.1 (±8.4) | 55.3 - 73.3 | 4 | -- | -- | -- | -- | -- | -- |
| | IS6 | 36.0 (±5.2) | 30.6 - 41.0 | 3 | -- | -- | -- | -- | -- | -- |
| | KX8 | -- | -- | -- | 15.4 (±7.5) | 2.4 – 72.7 | 2709 | 12 (±6.5) | 1.2 – 53.4 | 2709 |

*-- measurements not made at that location*

*s.d. standard deviation*

Each of the measurement locations within the stations had a different degree of influence of diesel fumes as indicated by the increments in coarse and fine fractions. The coarse fraction increment ($\Delta PM_{10-2.5}$) dominated the $PM_{10}$ increment at all measurement locations inside the station (>70%), except at LO1-Platform0/1 at KGX, where 79% of $\Delta PM_{10}$ was in the fine fraction, and at ED2-Waverley steps at EDB, where 65% of $\Delta PM_{10}$ was in the fine fraction. The latter locations are therefore interpreted as the locations in the stations with greater influence from diesel fumes. This is also shown in Figure 1, where scatter plots of the $\Delta PM_{10}$ and $\Delta PM_{2.5}$ data averaged over each $NO_2$ PDT exposure period against the corresponding location $\Delta NO_2$ concentration are displayed. Both stations are combined. Correlation was considerably stronger between $\Delta PM_{2.5}$ and $\Delta NO_2$ ($R^2 = 0.54$, $p < 0.001$) than between $\Delta PM_{10}$ and $\Delta NO_2$ ($R^2 = 0.12$, $p < 0.05$). This suggests that internal $PM_{2.5}$ and $NO_2$ shared common source(s), i.e. exhaust emissions from diesel trains. The relationship between $\Delta PM_{10}$ vs $\Delta NO_2$ was likely not as strong due to the more diverse sources of coarse PM not relating to directly to $NO_2$ emissions (e.g. wheel and rail wear, resuspension, construction, people). Figure 1B also shows that measurements from KGX were lower both in $\Delta NO_2$ and $\Delta PM_{2.5}$ which is entirely consistent with the lower number of diesel trains at KGX compared to EDB.

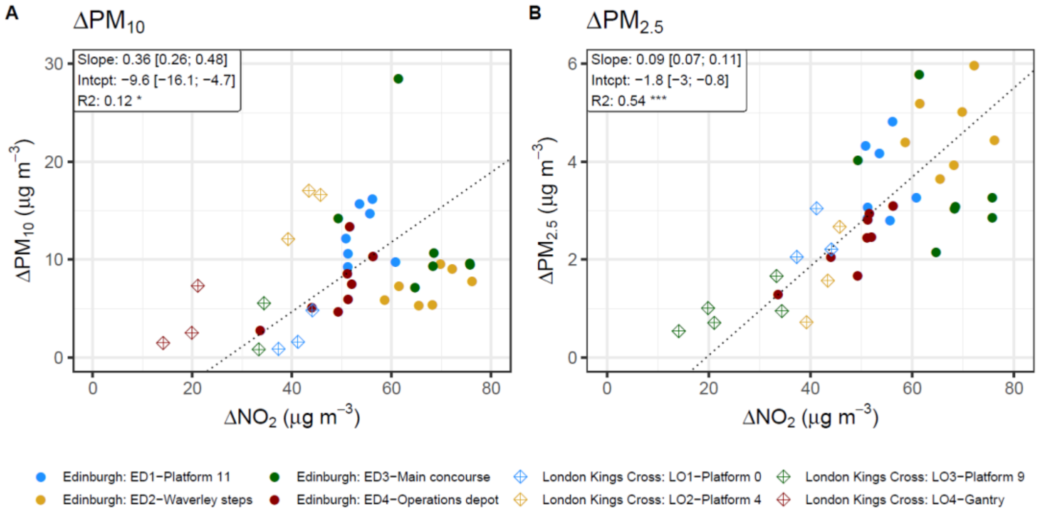


*Figure 1: Increments in $PM_{10}$ (A) and increments in $PM_{2.5}$ (B) versus increments in $NO_2$ in Edinburgh Waverley and London King's Cross stations. Each point corresponds to an exposure time for an $NO_2$ PDT measurement. Dotted straight lines denote the reduced-major-axis regression lines.*

## Influence of trains on air pollutant increments

The $\Delta NO_2$ and $\Delta PM_{2.5}$ variables at the trackside locations showed good correlations with the number of diesel train services at the adjacent platform during the measurement period: $R^2 = 0.72$ ($p < 0.001$) (Figure 2A) and $R^2 = 0.47$ ($p < 0.01$) (Figure 2C), respectively, with the significant positive slopes indicating an increase of increments with an increasing number of diesel trans. However, $\Delta PM_{10}$ showed no such correlation ($R^2 = 0.005$, Figure 2B). When correlating absolute concentrations against the number of diesel trains, the correlation for $NO_2$ was lower ($R^2 = 0.51$; $p < 0.001$) and $PM_{2.5}$ showed no correlation ($R^2 = 0.16$; $p > 0.1$) (Figure S5). This further supports the conclusion that the inside-station increments of these two pollutants is strongly associated with a common source of diesel-train emissions, but also indicates that whilst the diesel trains are the dominant source for within-station $NO_2$ (as noted earlier) they are less important as a within-station source for $PM_{2.5}$ for which general background concentrations are an important factor. This latter point is also consistent with the substantially lower increments above urban background for $PM_{2.5}$ than for $NO_2$. Other possible indoor sources of $PM_{2.5}$ that might influence the station variability might include cooking aerosols from the station food stalls and secondary organic aerosols that might form in the station environment. Neither $\Delta NO_2$ nor $\Delta PM_{10}$ were correlated with the number of electric trains (Figures 2D and 2E, respectively). The statistically significant negative correlation of $\Delta PM_{2.5}$ with the number of electric trains ($R^2$ = 0.70; $p$ < 0.001, Figure 2F), and non-significant negative relationship of $\Delta NO_2$ with number of electric trains, is likely due to electric trains displacing diesel trains in the timetabled slots rather than any other impact on reducing the concentrations.

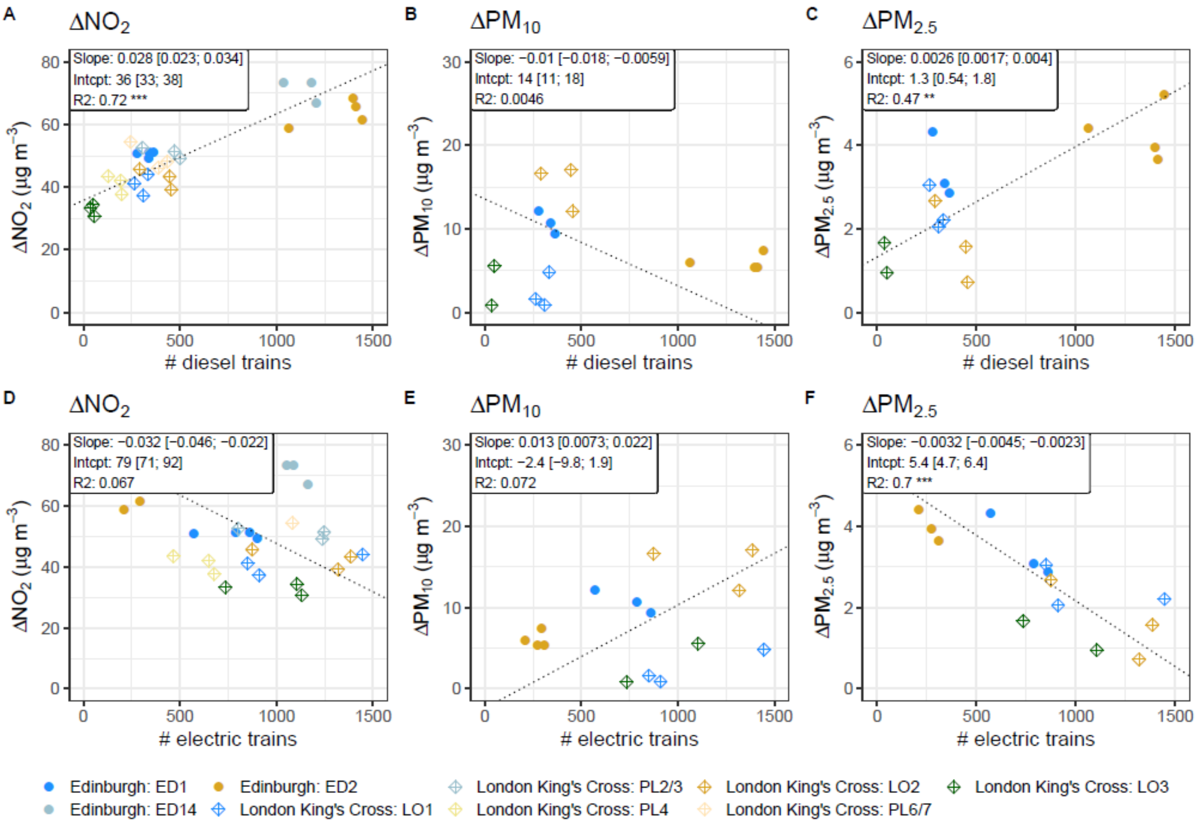


*Figure 2: Relation between increments in $NO_2$, $PM_{10}$ and $PM_{2.5}$ and the number of diesel or electric trains on the adjacent platform during the measurement period. Each point corresponds to the exposure times for $NO_2$ PDT. Data from 20 July. Dotted straight lines denote the reduced-major-axis regression lines.*

## Temporal variability

The hourly $PM_{2.5}$ data available at five trackside locations in each station (2 at EDB and 3 at KGX) permits comparison between average diurnal cycles of $\Delta PM_{2.5}$ and rail stock movements (Figure 3). Diurnal patterns in both differed between both stations. The highest $\Delta PM_{2.5}$ of >10 µg $m^{-3}$ on average was measured in the early afternoon at ED2-WaverleySteps (Figure 3C), but the hourly variability in $\Delta PM_{2.5}$ did not correlate well with the frequency of diesel trains ($R^2$ = 0.14) (Figure 3A). This was also the case at ED1–Platform11, for which $R^2$ = 0.009. At KGX, $\Delta PM_{2.5}$ increased from 0 to ~4 µg $m^{-3}$ in the early morning at both LO-Platform0/1 and LO2-Platform4/5, coincidental with the presence of diesel trains (Figure 3F). An association between $\Delta PM_{2.5}$ and diesel trains was particularly pronounced at LO3-Platform9 after 10.00 when the number of diesel trains substantially increased and $\Delta PM_{2.5}$ increased from 0 to ~12 µg $m^{-3}$. The mean hourly variation in $\Delta PM_{2.5}$ at the KGX trackside locations showed moderate to high correlations with the mean hourly variation in the number of diesel trains: $R^2$ = 0.45 (LO1-Platform0/1), $R^2$ = 0.61 (LO2-Platform4/5) and $R^2$ = 0.47 (LO3-Platform9). The mean hourly $\Delta NO_2$ at KGX LO1-Platform0/1 measured during the six-week period showed a better correlation with the number of diesel trains ($R^2$ = 0.79, Figure 3G) than $\Delta PM_{2.5}$.

Meteorological variables also varied depending on the hour of the day. Warmer, windier and drier conditions were observed during the central hours of the day (Figure 3D, 3H). The mean hourly variation in $\Delta PM_{2.5}$ showed good correlations ($R^2 > 0.54$) with the mean hourly variation in temperature and wind speed at ED2-Waverley steps, LO1-Platform0/1 and LO2-Platform4/5. $\Delta PM_{2.5}$ showed a negative correlation to the relative humidity with higher concentration with drier conditions at the same locations. $\Delta PM_{2.5}$ at ED1-Platform10 and LO3-Platform9 was not correlated to any of the meteorological variables tested except for LO3-Platform9 and pressure.

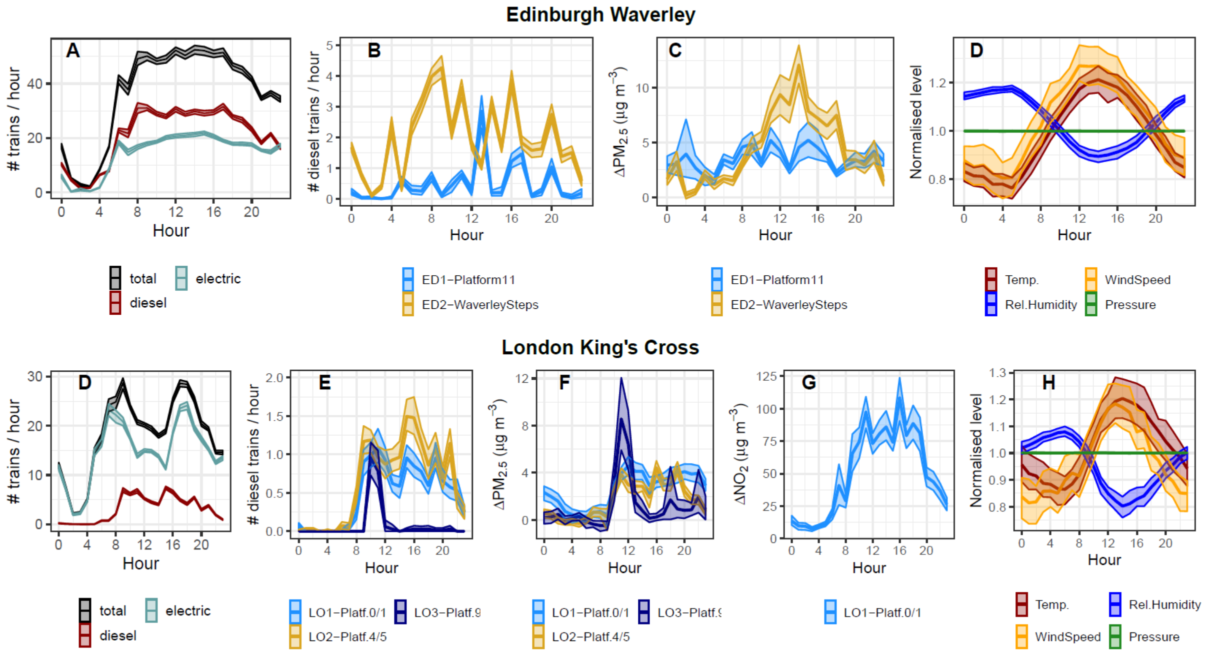


*Figure 3. Average number of total, diesel and electric trains per hour at Edinburgh Waverley and London King's Cross; number of diesel trains and increment in $PM_{2.5}$ ($\Delta PM_{2.5}$) concentrations in the tracksides with available measurements; and mean hourly variation of the meteorological conditions (temperature, relative humidity, wind speed and pressure) as measured outside the stations (normalised levels).*

The lack of correlation between $\Delta PM_{2.5}$ and the number of diesel trains in EDB may be explained by the configuration of the station, which although fully roofed does have through tracks and is therefore open at the two ends with the main track lines aligned with the main wind direction (south-west to north-east). This might enhance the dispersion of diesel fumes. The correlation between $\Delta PM_{2.5}$ and meteorological parameters such as temperature and wind speed at ED2-WaverleySteps is explained by this measurement location being situated a few metres above the tracks; vertical movement of the diesel plumes to this location are therefore enhanced at higher temperatures and wind speeds. Conversely, at KGX, the main tracksides were perpendicular to the main wind direction, and the station was fully closed at one side (terminal station).

## Random-forest modelling for $PM_{2.5}$

A regression random-forest model to predict hourly concentrations of $PM_{2.5}$ was built for each station. The explanatory variables used in each model were selected based on the Siegel repeated medians, selecting those showing a significant regression on the hourly $PM_{2.5}$ concentrations (Table S3–S4). At EDB, the effect of the diesel rolling stock and the meteorological conditions was different on the $PM_{2.5}$ measured at ED1-Platform11 compared to the $PM_{2.5}$ at ED2-WaverleySteps. Furthermore, ED1-Platform11 had a low data capture. Therefore, only the data from ED2-WaverleySteps was used to build the RF regression model at EDB. At KGX, data from LO1-

Platforms0/1 and LO2-Platforms4/5 were combined to build the RF regression model because both locations showed similar trends between variables (Table S4).

The performance of the RF models for ED2-WaverleySteps for hourly PM$_{2.5}$ concentrations was moderate, with $R^2$ ~0.50 and large RMSE of 4.6–4.8 μg m$^{-3}$ (Table S5). The most influential explanatory variables in all models were the concentration of PM$_{2.5}$ in the urban background, temperature and wind direction (Figure S8-S10). This indicates that the PM$_{2.5}$ measured at ED2-WaverleySteps was predominantly explained by the ambient background concentration (consistent with inference from other analyses of the data), with influence also from the transport of diesel emissions to the measurement site (in turn dependent on both the wind direction, controlling advection of emissions from other platforms; and temperature, controlling turbulent transport of emissions from the trains to the measurement location). The number of diesel trains at other platforms had greater importance than the diesel trains at the adjacent platform (Figure S9). For the model considering the type of rolling stock adjacent to the platform (model#2), these were the variables with the least importance (Figure S10) and the order of association was Sprinters > Diesel locomotives > Voyager > HST.

Background PM$_{2.5}$ and meteorological conditions are independent of rail management activities and therefore cannot be explicitly controlled within the station. Focusing on those variables that can be actively controlled inside the station, the partial dependency plots shown in Figure 4 indicate that at ED2-WaverleySteps, reducing the number of diesel trains at the platform adjacent to the monitoring site would lead to the largest reduction in PM$_{2.5}$ concentrations at the measurement location (0.25 μg m$^{-3}$ per diesel train on average) and that the rolling stock associated with the greatest reduction are Sprinters (reduction of 0.18 μg m$^{-3}$ per train on average). The reduction of diesel locomotive/locomotives hauled trains would also be associated with a reduction of PM$_{2.5}$ concentrations measured at ED2-WaverleySteps by 0.18 μg m$^{-3}$ train$^{-1}$ (Figure 4).

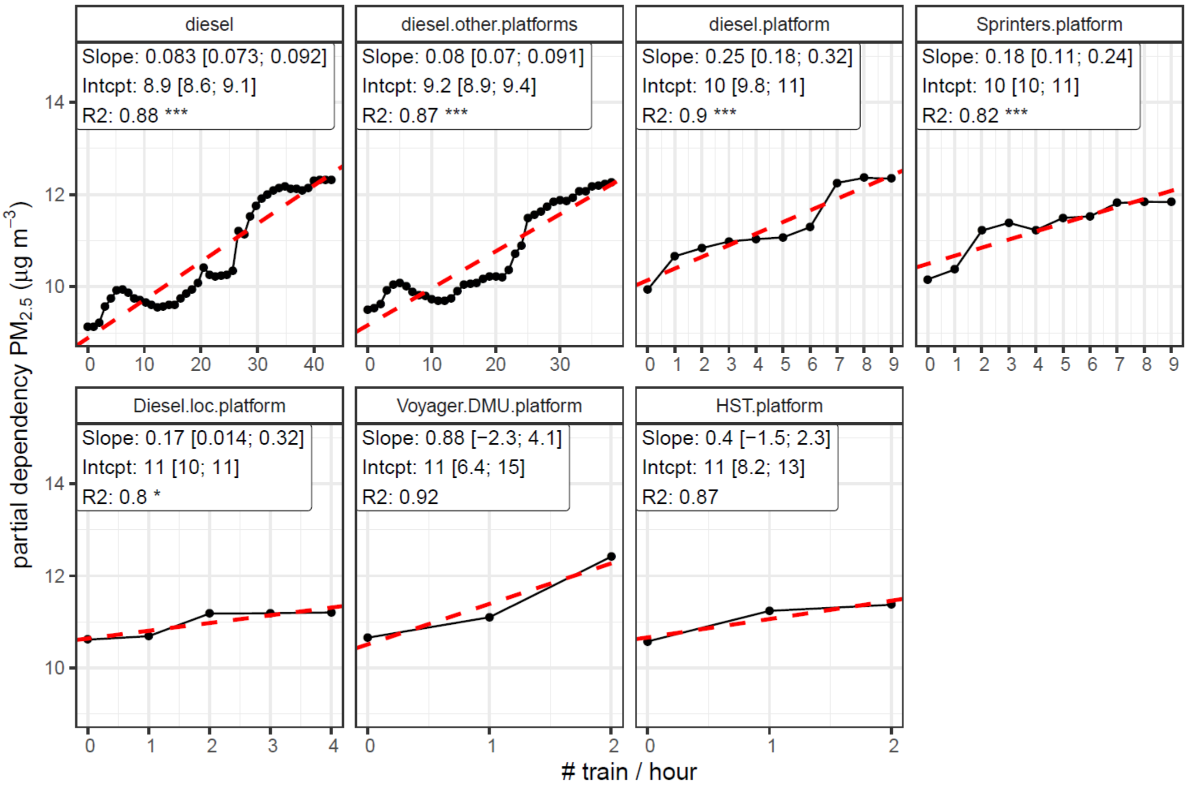


*Figure 4. Partial dependency of hourly PM$_{2.5}$ concentrations at ED2-WaverleySteps on the numbers and types of diesel trains per hour. The red dashed lines are RMA regression fits.*

The performance of the RF models to predict the $PM_{2.5}$ concentrations at KGX was good, with $R^2$ ~0.80 and low RMSE (2.7–2.9 µg $m^{-3}$) (Table S6). The $PM_{2.5}$ background concentration and the wind direction were the most important variables in all models (Figure S12-Figure S14). The importance of variables related to the rolling stock appear in the middle of the ranks and they were ordered as diesel trains at other platforms > diesel trains at the platform, and Class 180 > HST.

Figure 5 shows that, for $PM_{2.5}$ at KGX, the partial dependencies for diesel trains, diesel trains at the platform and diesel trains at other platforms all increased as number of trains increased from 0 to 4 per hour, but then levelled off as the number of trains increased further. One possible explanation may be a reduction of the idling time when increasing the frequency of trains per hour as the actual time that each train individually remains in the station is reduced. This levelling off in partial dependency with number of diesel trains was not observed in the equivalent partial dependencies at ED2-WaverleySteps (Figure 4. Partial dependency of hourly PM2.5 concentrations at ED2-WaverleySteps on the numbers and types of diesel trains per hour. For the rolling stock next to the monitoring sites, $PM_{2.5}$ increased linearly as the frequency increased. RMA regression was calculated for train frequencies up to 4 services per hour and all showed good correlations ($R^2 > 0.71$) that were statistically significant ($p < 0.1$) (Figure 5). Decreasing the number of diesel trains at platforms 0-8 by one per hour was associated with a decrease in $PM_{2.5}$ of 0.57 µg $m^{-3}$ on average; and reducing the number of diesel trains at other platforms was more effective than reducing the number of diesel trains next to the measurement site. This is explained by the fact that emissions from all platforms contribute to the levels in the area between platforms 0-8. Reducing the number of Class 180 trains was associated with a reduction in $PM_{2.5}$ concentration of 0.40 µg $m^{-3}$ per train on average, whilst reducing HSTs was associated with a reduction of 0.29 µg $m^{-3}$ per train on average (Figure 5).

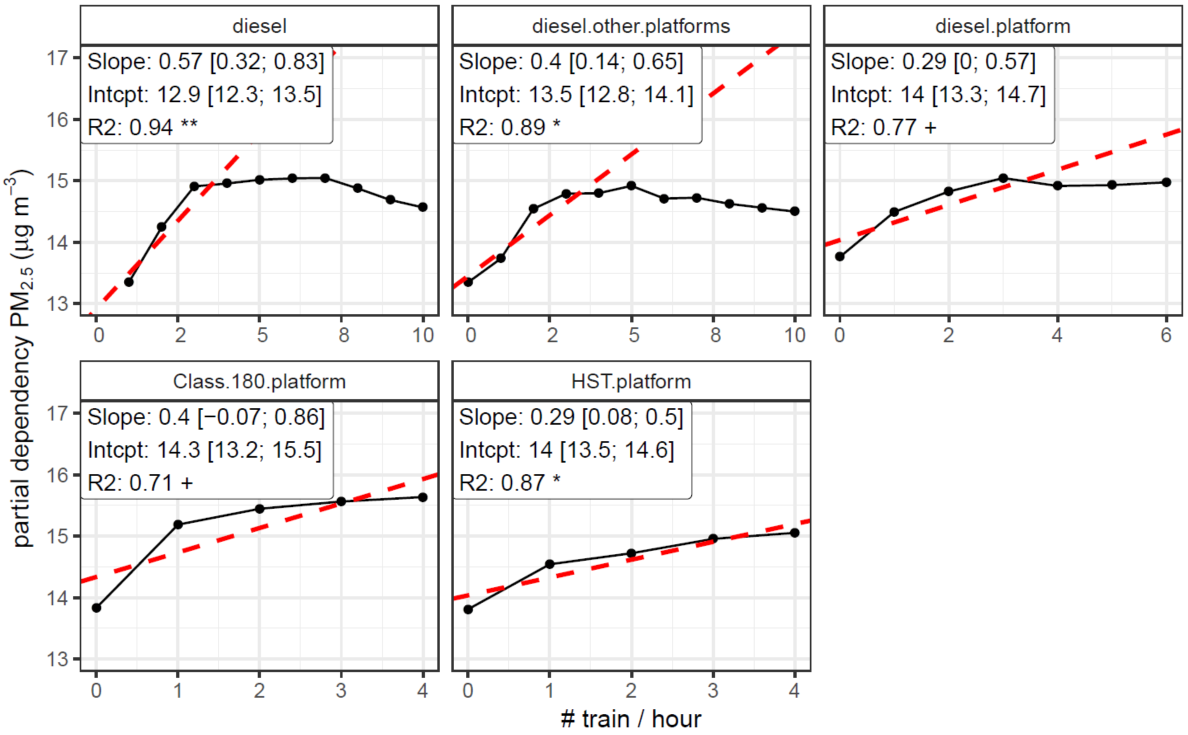


*Figure 5. Partial dependency of hourly $PM_{2.5}$ concentrations at London King's Cross on the numbers and types of diesel trains per hour. The red dashed lines are RMA regression fits to the data up to and including 4 trains per hour.*

## Random-forest modelling for $\Delta PM_{2.5}$

The RF regression modelling was also applied to hourly $\Delta PM_{2.5}$ for ED2-WaverlySteps at EDB and for LO1/LO2 at KGX. However, the performance of these models was lower than those for $PM_{2.5}$: $R^2$ = 0.43–0.50 (EDB) and 0.23–0.28 (KGX); and RMSE = 4.4–4.7 µg $m^{-3}$ (EDB) and 2.9–3.0 µg $m^{-3}$ (KGX). The lower performance of the statistical model for $\Delta PM_{2.5}$ compared with the statistically significant interpretations for $\Delta PM_{2.5}$ when using longer period averaging (as shown in Figure 2) is probably due to the unsuitability of the incremental approach for very short time periods, i.e. for hourly data. One issue is that ambient background concentrations may be impacted by localised sources or dispersion affects in the short-term and therefore not always be representative of the background concentrations in the stations. Furthermore, the discrete train information (counts of different train types per hour) does not fully describe the emissions from these trains, merely their presence. The model for $\Delta PM_{2.5}$ might be further improved by the inclusion of train idling information.

## Random-forest modelling for $NO_2$ and $\Delta NO_2$

Random-forest modelling was also undertaken for the six-week period with high-time resolved (hourly) data for $NO_2$ at LO1-Platform0/1 at KGX. The same model formulations as per $PM_{2.5}$ were implemented. Overall, the models predicted both $NO_2$ and $\Delta NO_2$ moderately ($R^2$ = 0.48 – 0.52) and with large RSME (31.9 – 33.2 µg $m^{-3}$). However, as for the RF analyses on $PM_{2.5}$ measurements, the partial dependencies for the $NO_2$ concentrations also indicated that Class 180 was associated with larger $NO_2$ concentrations at LO1-Platform 0/1 (23.7 µg $m^{-3}$ per train) compared to HST trains (8.6 µg $m^{-3}$ per train) (Figure 6) and therefore its replacement or emissions management should be prioritised.

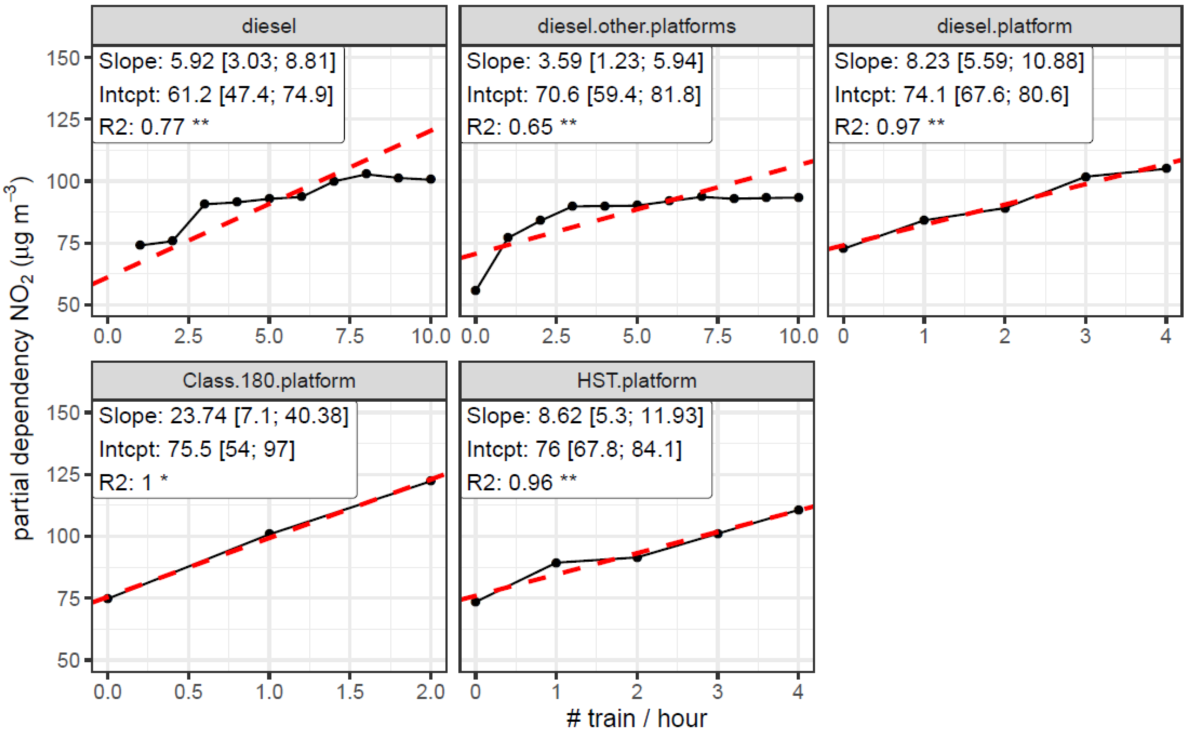


*Figure 6. Partial dependency of hourly $NO_2$ concentrations at LO1-Platform 0/1 at London King's Cross on the numbers and types of diesel trains per hour.*

# Conclusions

This study demonstrated that whilst 4-week averaged pollutant measurements allowed a focus on the internal sources and factors influencing the pollutant incremental concentrations independent of hour-to-hour variability, long-term averaging obscures the useful insight that can be derived from hourly correlations between pollutants, train movement and individual train types. Overall, this study has provided clear evidence that diesel-powered trains increase concentrations of $NO_2$ and $PM_{2.5}$ in enclosed stations to levels that exceed WHO guidelines for their concentrations in ambient air. In particular, the diesel-powered rolling stock types contributing most to $PM_{2.5}$ levels within both stations were identified. However, this study did not have enough information to discern how much of their contribution was due to their absolute emissions or because of the way those particular trains operated in the station, for example increased idling time or position of the engine along the platform when stationary. Other studies have observed that diesel-powered trains also lead to increased air pollutant concentrations within the passenger carriage (Andersen et al., 2019; Jeong et al., 2017). Their replacement with cleaner powered trains is therefore encouraged to reduce exposure both in the station and on board.

# Acknowledgments

This research was supported by RSSB under grant reference T1221. The authors would like to thank technical and practical assistance during the project by Aqeel Janjua and James Wright at RSSB. We also thank Russell Preece from Virgin Trains (now Avanti West Coast) for providing information regarding timetable schedules and rolling stock.

# References

Aarnio, P., Yli-Tuomi, T., Kousa, A., Mäkelä, T., Hirsikko, A., Hämeri, K., Räisänen, M., Hillamo, R., Koskentalo, T., Jantunen, M., 2005. The concentrations and composition of and exposure to fine particles (PM2.5) in the Helsinki subway system. Atmos. Environ. 39, 5059–5066. https://doi.org/10.1016/j.atmosenv.2005.05.012

Andersen, M.H.G., Johannesson, S., Fonseca, A.S., Clausen, P.A., Saber, A.T., Roursgaard, M., Loeschner, K., Koponen, I.K., Loft, S., Vogel, U., Møller, P., 2019. Exposure to Air Pollution inside Electric and Diesel-Powered Passenger Trains. Environ. Sci. Technol. acs.est.8b06980. https://doi.org/10.1021/acs.est.8b06980

Barmparesos, N., D. Assimakopoulos, V., Niki Assimakopoulos, M., Tsairidi, E., 2016. Particulate matter levels and comfort conditions in the trains and platforms of the Athens underground metro. AIMS Environ. Sci. 3, 199–219. https://doi.org/10.3934/environsci.2016.2.199

Borken-kleefeld, J., Ntziachristos, L., 2012. The potential for further controls of emissions from mobile sources in Europe.

Brokamp, C., Jandarov, R., Hossain, M., Ryan, P., 2018. Predicting Daily Urban Fine Particulate Matter Concentrations Using a Random Forest Model. Environ. Sci. Technol. 52, 4173–4179. https://doi.org/10.1021/acs.est.7b05381

Carslaw, D., 2019. Package ‘ worldmet .’

Chong, U., Swanson, J.J., Boies, A.M., 2015. Air Quality in London Paddington Air Quality in London Paddington Train Station. Environ. Res. Lett. 10.

de Oña, J., de Oña, R., López, G., 2016. Transit service quality analysis using cluster analysis and decision trees: a step forward to personalized marketing in public transportation. Transportation (Amst). 43, 725–747. https://doi.org/10.1007/s11116-015-9615-0

de Oña, R., Eboli, L., Mazzulla, G., 2014. Key Factors Affecting Rail Service Quality in the Northern Italy: a Decision Tree Approach. Transport 29, 75–83. https://doi.org/10.3846/16484142.2014.898216

Department for Transport, 2017. Rail factsheet: 2017 - GOV.UK. 29 Novemb. 2017 5, 1–6.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Faganeli Pucer, J., Štrumbelj, E., 2018. Impact of changes in climate on air pollution in Slovenia between 2002 and 2017. Environ. Pollut. 242, 398–406. https://doi.org/10.1016/j.envpol.2018.06.084

Givoni, M., Brand, C., Watkiss, P., 2009. Are railways “climate friendly”? Built Environ. 35, 70–86. https://doi.org/10.2148/benv.35.1.70

Grange, S.K., Carslaw, D.C., 2019. Using meteorological normalisation to detect interventions in air quality time series. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2018.10.344

Green, D.C., Fuller, G.W., Baker, T., 2009. Development and validation of the volatile correction model for PM10 - An empirical method for adjusting TEOM measurements for their loss of volatile particulate matter. Atmos. Environ. 43, 2132–2141. https://doi.org/10.1016/j.atmosenv.2009.01.024

Hickman, A., Baker, C., Cai, X., Delgado-Saborit, J., Thornes, J., 2018. Evaluation of air quality at the Birmingham New Street Railway Station. Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit 232, 1864–1878. https://doi.org/10.1177/0954409717752180

Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Gu, D., Liu, Y., 2018. Predicting monthly high-resolution PM2.5 concentrations with random forest model in the North China Plain. Environ. Pollut. 242, 675–683. https://doi.org/10.1016/j.envpol.2018.07.016

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning - with applications in R 426. https://doi.org/10.1007/978-1-4614-7138-7

Jeong, C.-H., Traub, A., Evans, G.J., 2017. Exposure to ultrafine particles and black carbon in diesel-powered commuter trains. Atmos. Environ. 155, 46–52. https://doi.org/10.1016/J.ATMOSENV.2017.02.015

Johansson, C., Johansson, P., 2003. Particulate matter in the underground of Stockholm. Atmos. Environ. 37, 3–9. https://doi.org/10.1016/S1352-2310(02)00833-6

Kim, K.Y., Kim, Y.S., Roh, Y.M., Lee, C.M., Kim, C.N., 2008. Spatial distribution of particulate matter (PM10 and PM2.5) in Seoul Metropolitan Subway stations. J. Hazard. Mater. 154, 440–443. https://doi.org/10.1016/j.jhazmat.2007.10.042

Lenschow, P., 2001. Some ideas about the sources of PM10. Atmos. Environ. 35, 23–33. https://doi.org/10.1016/S1352-2310(01)00122-4

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., Hampapur, A., 2014. Improving rail network velocity: A machine learning approach to predictive maintenance. Transp. Res. Part C Emerg. Technol. 45, 17–26. https://doi.org/10.1016/j.trc.2014.04.013

Martins, V., Moreno, T., Minguillón, M.C., Drooge, B.L. Van, Querol, X., 2015. Chemical composition and source apportionment of PM2.5 in subway stations of Barcelona , Spain 315760.

Office of Road and Rail, 2017. Rail infrastructure, assets and environmental 2016-17 Annual Statistical Release 21.

Palmes, E.D., Gunnison, A.F., Dimattio, J., Tomczyk, C., 1976. Personal sampler for nitrogen dioxide. Am. Ind. Hyg. Assoc. J. https://doi.org/10.1080/0002889768507522

Perrino, C., Marcovecchio, F., Tofful, L., Canepari, S., 2015. Particulate matter concentration and chemical composition in the metro system of Rome, Italy. Environ. Sci. Pollut. Res. Int. 757. https://doi.org/10.1007/s11356-014-4019-9

Querol, X., Moreno, T., Karanasiou, A., Reche, C., Alastuey, A., Viana, M., Font, O., Gil, J., De Miguel, E., Capdevila, M., 2012. Variability of levels and composition of PM10 and PM2.5 in the Barcelona metro system. Atmos. Chem. Phys. 12, 5055–5076. https://doi.org/10.5194/acp-12-5055-2012

Vilcassim, M.J.R., Thurston, G.D., Peltier, R.E., Gordon, T., 2014. Black Carbon and Particulate Matter (PM2.5) Concentrations in New York City's Subway Stations. Environ. Sci. Technol. 48, 14738–45. https://doi.org/10.1021/es504295h

WHO-IARC, 2012. The diesel exhaust in miners study: A nested case-control study of lung cancer and diesel exhaust. Int. Agency Res. Cancer - Press Release 213. https://doi.org/10.1093/jnci/djs034