

University of London
Imperial College of Science, Technology and Medicine
Department of Civil & Environmental Engineering

Real-world variability, modelling and mitigation of road transport emissions

Clémence Marie Anne Le Cornec

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Civil & Environmental Engineering of the University of London and
the Diploma of Imperial College, 1th February 2021

Declaration of originality

All the work presented in this thesis is that of the author. Any work or contribution made by others is referenced appropriately.

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND). Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

[ENGLISH]

Outdoor air pollution is considered the largest single environmental health risk and is estimated to cause 4.2 million deaths every year. Despite the major vehicle emissions reduction achieved over the past two decades, road transport remains a major source of air pollutants such as nitrogen oxides (NO_x), contributing 39% of the total EU-28 NO_x emissions in 2017. The regular exceedances of the annual mean concentration limit for NO_2 , particularly in urban areas, have been largely attributed to the discrepancies between type approval limits and real-world driving emissions as well as the fitting of defeat devices on diesel vehicles. In order to design effective air quality mitigation strategies, it is therefore crucial to improve our understanding of and ability to model real-world driving emissions. Based on the largest recorded Portable Emissions Measurement System (PEMS) dataset, which comprised 287 Euro 5 and Euro 6 diesel and petrol vehicles, this PhD thesis aims to fill these gaps by providing new emissions models based on an extensive dataset.

It is demonstrated in this thesis that while physical parameters such as vehicle weight or engine size did not show any correlation with real-world NO_x emissions, external parameters, particularly driving dynamicity, are directly correlated with real-world driving emissions. The effect of driving dynamicity on real-world emissions is shown to decrease with the successive regulations (Euro standard), indicating a general improvement of aftertreatment systems. The first model presented is an aggregated emission model, while the second emission model is an instantaneous emission model but both directly account for driving dynamicity, although the way they do differ significantly. Both models are reliable and accurate, presenting relative error in prediction smaller than 20%. Additionally, this PhD thesis also intends to gain insights on the real-world impact of an air quality mitigation strategy on the emissions and local air quality. Application of the developed models to the assessment of the impact of a traffic intervention on air quality demonstrated that although the chosen mitigation strategy had locally a measurable impact on emissions and air quality, this impact was small compared to the variations in pollutant concentrations induced by the meteorological conditions. The ease of use of both models, as well as their wide range of applicability make them ideal operational tools for policy makers aiming to build emission inventories or evaluate emissions mitigation strategies.

La pollution de l'air est considérée comme l'un des plus grands risques environnementaux actuels et est estimée directement responsable du décès de 4,2 millions de personnes dans le monde chaque année. Malgré les immenses progrès réalisés les deux dernières décennies en termes d'émissions, le secteur routier reste la principale source de polluants atmosphériques tels que les oxydes d'azote (NO_x), représentant 39% des émissions de NO_x de l'UE-28 en 2017. Les dépassements réguliers de la concentration limite annuelle de NO_2 , particulièrement prévalant en milieu urbain, ont été largement attribuée aux écarts entre les valeurs d'émissions homologuées et les valeurs effectivement mesurées en conditions réelles, ainsi qu'à l'installation de logiciels capable de contourner la procédure d'homologation sur les véhicules de type diesel. Afin d'élaborer des stratégies efficaces de lutte contre la pollution de l'air, il est donc crucial d'améliorer notre compréhension et notre capacité à modéliser les émissions en conditions réelles de conduite. Basée sur le plus grand set de données de type Portable Emissions Measurement System (PEMS) enregistré à ce jour, cette thèse de doctorat vise à combler ces lacunes, en procurant notamment de nouveaux modèles d'émissions reposant sur un set de données statistique significatif.

Il est démontré dans cette thèse de doctorat que tandis que les paramètres physiques tels que le poids du véhicule ou la taille du moteur ne présentent que peu ou pas de corrélation avec les émissions de NO_x , les facteurs externes tels que le style de conduite sont eux directement liés aux émissions de NO_x en conditions réelles de conduite. Il est notamment montré que l'impact du style de conduite sur les émissions en conditions réelles tend à diminuer avec les régulations successives (standard Euro), indiquant une amélioration générale des systèmes d'après traitement. Le premier modèle d'émissions présenté est un modèle agrégé, tandis que le second est un modèle instantané mais chacun d'eux prend en compte directement l'influence du style de conduite, quoique d'une façon fondamentalement différente. Les deux modèles présentent des erreurs relatives de prédiction toujours inférieures à 20% et peuvent donc être considérés comme fiables et précis. De façon complémentaire, cette thèse de doctorat a également pour objectif de quantifier l'impact réel d'une stratégie de lutte contre la pollution de l'air sur les émissions et la qualité de l'air. L'application des modèles développés dans cette thèse permet de démontrer que bien que la stratégie choisie ait un effet mesurable tant sur les émissions que sur la qualité de l'air locale, cet effet se révélait faible en comparaison avec les variations de

concentrations directement induites par les conditions météorologiques. La facilité d'utilisation des deux modèles développés, ainsi que leur large champ d'application en fait des outils idéaux pour les décideurs politiques aspirant à bâtir des inventaires d'émissions ou à évaluer l'efficacité réelle de stratégies de lutte contre la pollution de l'air.

Acknowledgements

I would like first to express my thanks to my supervisors, Dr. Marc Stettler and Dr. Maarten van Reeuwijk, who helped me to develop and pursue my ideas. A huge thank you to Nick Molden at Emissions Analytics for giving access to his fantastic data. Thanks a lot to Dr. Anna Schroeder for the collaboration on the MAGIC project, and to Dr. Tom Grylls for the collaboration on the SIRANE-uDALES project, working together made my PhD certainly more pleasant. Finally, a special thank also to Sarah Willis for her continuous support with all the administrative side of a PhD.

A huge thank you to Csaba, Nils, Nina, Alessandra, Petrina, He-in, Abdul, Qiming, Roger, Renos, Eduardo and Jose from the Centre for Transport Study for their friendship, support and the nice Thursdays evenings online games in these times of pandemic. I would also like to thank everyone from the Academy Fencing Club, especially Thibault (and Caroline), Béatrice, Aladdin, Hugo, Amalia, Josh, the two James, Lizelle, Femi, Eva, Rachel and Jonathan. It has been great to fence and compete alongside with you and you definitely made my journey towards the completion of this PhD as well as my stay in London much nicer! A big thank you to my very old friend Etienne, as well as to all my close friends back home in Switzerland: Margot, Anne-Lise, Céline, Alex, Victoria, Sébastien, Alice, Eva and Coline. Trust me when I say that you all contributed in a large extend to this thesis. Thank you my dear friends.

Invaluable to my efforts throughout was the unwavering support of my family. Il m'est impossible de remercier assez mon oncle, ma tante et mes quatre cousins, Richard, Alex, Axel, Noah, Alan et Quim, qui m'ont accueilli chez eux durant ma première année d'étude à l'EPFL. Sans vous et votre accueil, cette thèse n'existerait très probablement pas aujourd'hui. Je vous en suis (et vous en serai toujours) extrêmement reconnaissante. Un immense merci également aux deux meilleures petites soeurs de la planète: Raphaëlle et Agathe. Merci pour vos encouragements, votre soutien constant et vos taquineries, même dans les pires moments, sans oublier, bien sûr, nos nombreux échanges sur Radio Potins et les mauvais films du samedi soir 20h30. Durant le long chemin de l'enfance à l'âge adulte, nous avons eu l'immense chance de rencontrer Renée, qui nous a partiellement élevées. Ma gratitude pour ta présence aimante et attentionnée va au-delà des mots. Finalement, je suis aussi extrêmement reconnaissante envers mes parents, qui ont encouragé ma curiosité d'aussi loin que remontent mes souvenirs et ont passé de très

nombreuses heures au square pour me permettre de dépenser mon énergie débordante. Papa et Maman, merci pour tout. A tous et toutes, merci pour ce qui fait de vous ce que vous êtes.

Finally, no-one has endured quite so much as my partner, Romain, who has been a great source of comfort and stability through the ups and (many) downs of the past three years. Merci mon amour pour ton soutien inconditionnel, et tout particulièrement pour ton aide ces deux derniers mois. J'ignore sincèrement comment j'aurais fait cela sans toi. L'accomplissement de cette thèse de doctorat est sans aucun doute un succès joint. Maintenant que cette thèse est enfin terminée, je me réjouis énormément de continuer à vivre plus paisiblement la suite de notre longue aventure pour rentrer en Suisse. Sans toi, ma vie serait beaucoup moins belle. C'est peu dire. Je t'aime.

Dedication

To my Catalan great-grand-mother, Teresa Doñate i Cuevas, who did not get the chance to go to school, to my Austrian grand-mother, Margarete Penz, who had to stop school at an early age because of her sex, and more largely to every woman across the globe who did not have the chance to access education.

‘La recherche de la vérité doit être le but de notre activité; c’est la seule fin qui soit digne d’elle. Sans doute nous devons d’abord nous efforcer de soulager les souffrances humaines, mais pourquoi? Ne pas souffrir, c’est un idéal négatif et qui serait plus sûrement atteint par l’anéantissement du monde. Si nous voulons de plus en plus affranchir l’homme des soucis matériels, c’est pour qu’il puisse employer sa liberté reconquise à l’étude et à la contemplation de la vérité.’

La Valeur de la Science - Henri Poincaré

Contents

Declaration of originality	i
Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 Background	1
1.2 Research aims and objectives	4
1.3 Structure of the thesis	5
1.4 Contributions	7
1.5 Publications	9
2 Traffic related air pollution	10
2.1 Introduction	10
2.2 Emissions	11
2.2.1 Sources of emissions	11
2.2.2 Formation of emissions as a results of combustion	12
2.2.3 Emissions regulations	22

2.2.4	Modelling of emissions	30
2.3	Pollutant dispersion	37
2.3.1	Dispersion in urban environments	37
2.3.2	Pollutant dispersion modelling	38
2.4	Air quality	40
2.4.1	Air quality regulations in the UK	41
2.4.2	Air quality regulations in the EU	42
2.4.3	Pollutants impacts on health, ecosystems and economies	43
2.5	Mitigation strategies to air quality issues	44
3	REVS: Road transport Emissions and Variance Simulator	47
3.1	Introduction	48
3.2	Methods	50
3.2.1	Dataset description	50
3.2.2	Processing of the raw data	52
3.2.3	The COPERT5 emission model	54
3.2.4	Extraction of segments	55
3.3	Statistical analysis of the data	55
3.4	Results	57
3.4.1	Real-world driving NO _x emissions	58
3.4.2	Impact of average speed and driving dynamicity on NO _x emissions	61
3.5	REVS	63
3.5.1	Description of the REVS model	63

3.5.2	Emission maps for the COPERT5, the REVS model and the measured data	66
3.5.3	Comparison of the models performance against the dataset	68
3.5.4	Validation of the REVS model for a fleet mix	69
3.6	Example of the impact of VSP	69
3.7	Conclusion	75
4	Modelling of instantaneous NO_x emissions from diesel vehicles	78
4.1	Introduction	79
4.2	Methods	81
4.2.1	Data description	81
4.2.2	Clustering	83
4.2.3	Calibration of existing emission models	84
4.2.4	Neural network model description	85
4.2.5	Dealing with trips that do not belong to a cluster	86
4.2.6	Code availability	87
4.3	Results	89
4.3.1	Clustering analysis	89
4.3.2	Emission maps for the LT, the NLR and the MLP models	96
4.3.3	Comparison of models performance against the dataset	99
4.3.4	Case study and comparison with COPERT5	103
4.4	Conclusion	106

5	Impact of a traffic intervention on the local air quality	109
5.1	Introduction	110
5.2	Methods	112
5.2.1	Field study description	112
5.2.2	Instrumentation	115
5.2.3	Meteorological conditions	116
5.2.4	Modelling of the field study	118
5.3	Results	123
5.3.1	NO ₂ concentrations	124
5.3.2	Traffic conditions	128
5.3.3	Comparison of the NLR and REVS models	129
5.3.4	Spatial distribution in emissions	132
5.3.5	Operational air dispersion simulation	133
5.3.6	Comparison of the SIRANE predicted pedestrian average NO ₂ concentrations with the sensors measurements	137
5.4	Conclusion	140
6	Conclusions	143
6.1	Main findings and contributions	143
6.2	Limitations and future work	145
6.3	Final thoughts	147
A	REVS: Road transport Emissions and Variance Simulator	148
A.1	Datasets description	148

A.1.1	Description of the EA2019 dataset	148
A.1.2	Description of the EA-2016 dataset	151
A.1.3	Description of the DfT dataset	151
A.1.4	Description of the DVSA dataset	152
A.2	Processing of the raw data	152
A.2.1	EA2019	152
A.2.2	EA2016	155
A.2.3	DfT	155
A.2.4	Extraction of the trips	156
A.2.5	Bootstrapping	157
A.2.6	VSP operating modes	158
A.2.7	Comparison of the typical VSP distribution in London to the VSP distribution of the dataset	159
A.3	Computing the average and standard deviation of a sum of different distributions	160
A.4	REVS model	162
A.4.1	Model coefficients	162
A.4.2	Performance of the REVS model per driving environment and comparison with COPERT5	164
A.5	Procedure to apply the REVS model	166
B	Modelling of instantaneous NO_x emissions from diesel vehicles	167
B.1	Dynamic Time Warping	167
B.2	K-means++	168
B.3	Calinski-Harabatz and Davies-Bouldin indices	168

B.4	Artificial Neural Networks	169
B.5	Multilayers perceptron	170
B.6	Results	171
B.6.1	Intra-vehicle real-world driving variability	171
B.6.2	NLR ₁ coefficients	174
B.6.3	Clustering	174
B.6.4	Robustness	215
B.6.5	Multilayer perceptron	215
B.6.6	Influence of time history	219
B.6.7	Case study	219
C	Impact of a traffic intervention on the local air quality	225
C.1	Instrumentation specification	225
C.2	LIDAR average wind speed and wind direction	226
C.3	Background concentration per day	226
C.4	Distribution of the mean NO _x emissions and its standard deviation per street . .	227
C.5	Operational air dispersion modelling	227
	Bibliography	227

List of Tables

2.1	US Environmental Protection Agency - Tier 3 FTP emission standards for LDVS, LDTS and MDPVS	24
2.2	European emissions standards for Light-Duty diesel vehicles	26
2.3	European emissions standards for Heavy-Duty diesel vehicles at steady-state . .	27
2.4	European emissions standards for Heavy-Duty diesel vehicles in transient mode .	27
2.5	Ambient air quality directive limit values for a selected set of air pollutants . . .	43
3.1	Test routes dynamic characteristics EA2019.	52
3.2	Characteristics of the fleet used in the case study	72
4.1	Dynamic characteristics of the test segment	82
4.2	Summary of the models used in the study	88
4.3	Mean and standard deviation of the FB and NMSE statistical indices for the 60 unseen tests used for the case study	106
5.1	Cameras traffic counts for the two weeks of field study between 12 am and 4 pm. Data for the 17 th and the 25 th of September have been largely lost and are therefore not exploitable to get reliable traffic counts.	116
5.2	Distribution of fuel type and engine size used to estimate the instantaneous NO _x emission using the NLR model.	121

5.3	Average background NO ₂ concentrations, NO ₂ concentrations and average measured above-background NO ₂ concentrations for the CAPS (CI) and the M200E (BL) over the two weeks of field study	127
5.4	SIRANE predicted NO ₂ concentrations at pedestrian level	136
A.1	Vehicles characteristics EA2019 dataset	148
A.2	Raw processing of the tests of EA2019 dataset	152
A.3	VSP operating modes	158
A.4	Coefficients of the REVS model for diesel vehicles, per Euro standard	163
A.5	Coefficients of the REVS model for petrol vehicles, per Euro standard	163
B.1	Dynamic and emitting characteristics of tests from the same vehicle falling in different clusters	173
B.2	Mean, standard deviation and coefficient of variation (CV) of the NO _x emission time serie for each vehicles whose tests are classified in different clusters	175
B.3	NLR ₁ coefficients during deceleration	211
B.4	NLR ₁ coefficients during acceleration	211
B.5	Evolution of the mean and the standard deviation of the predicted NO _x emission factor, the fractional bias and the normalised mean squared error for the three performed runs.	217
C.1	Comparison of the technical specifications of the CAPS and the M200E	226
C.2	LIDAR average wind speed and wind direction from 12am to 4pm during the two weeks of field study. Days where the 96s cycle was implemented are displayed in bold font while days where the 48s cycle was implemented are displayed in normal font.	226
C.3	Background concentrations per day	227

C.4	Predicted canyon average NO ₂ concentration in ppb. Emissions were calculated using the NLR model.	229
C.5	Predicted canyon average NO concentration in ppb. Emissions were calculated using the NLR model.	229
C.6	Predicted canyon average O ₃ concentration in ppb. Emissions were calculated using the NLR model.	230
C.7	Predicted canyon average NO ₂ concentration in ppb. Emissions were calculated using the REVS model.	230
C.8	Predicted canyon average NO concentration in ppb. Emissions were calculated using the REVS model.	231
C.9	Predicted canyon average O ₃ concentration in ppb. Emissions were calculated using the REVS model.	231

List of Figures

3.1	Illustration of the procedure followed to develop the model	58
3.2	Comparison of the cumulative distributions of the NO _x emission factor split per driving environment and vehicle technology category	59
3.3	Joint distribution of the percentile of average speed and the percentile of NO _x emission factor in urban driving conditions.	62
3.4	Joint distribution of the percentile of VSP and the percentile of NO _x emission factor in urban driving conditions.	64
3.5	Comparison of the measured and predicted NO _x emission factors in urban driving conditions split by vehicle technology category	67
3.6	Comparison of the measured (x-axis) and predicted (y-axis) mean (a, b) and standard deviation (c) for all the vehicle technology categories for speed between 10 and 130 km/h aggregated per bin of 10 km/h and VSP between -2 and 6 kW/t in urban driving conditions, 4 and 12 kW/t in rural driving conditions and 6 and 18 kW/t in motorway driving conditions per bin of 2 kW/t	70
3.7	Comparison of the ratio of $\bar{e}_M$ to $\bar{e}_C$ (top) and the ratio of σ_M to σ_C (bottom) as a function of the number of vehicles considered	71
3.8	Distribution of the real NO _x emission factor and comparison with the predictions from the REVS and the COPERT5 model (left) and distribution of the average VSP of the fleet (right)	74

4.1	Vehicle manufacturer, NO _x aftertreatment technology, engine size, vehicle weight and NO _x emission factor per cluster	91
4.2	Mean measured instantaneous NO _x emissions as a function of the distance travelled for the main clusters	93
4.3	Real-world variation of the NO _x emission factor for the same vehicle when the tests of these vehicles fall in different clusters	94
4.4	Comparison of the emission maps for the main clusters for the LT ₁ , NLR ₁ and MLP ₁ models	98
4.5	Comparison of the measured and predicted emission factors for LT, NLR and MLP models	100
4.6	Boxplots of the measured and predicted emission factors (a), boxplots of the Fractional Bias (b) and boxplots of the Normalised Mean Squared Error (c) for the LT, NLR and MLP models	102
4.7	Aggregated emission maps of the unseen tests: (a) are the measurements, (b) the NLR ₁ model, (c) the MLP ₁ model and (d) the COPERT 5 model.	104
4.8	Aggregated measured and modelled emissions as the function of the distance travelled: (a) are the instantaneous emissions and (b) are the cumulative emissions.	105
5.1	Map of the field study area	113
5.2	Description of the traffic light timings for the two interventions.	114
5.3	Comparison of the distribution in wind speed and wind direction for the two weeks of field study measured by the anemometer located on London Road. . . .	117
5.4	Distribution of the background NO ₂ concentration and the above-background NO ₂ concentrations for the CAPS (CI) and the M200E (BL) over the two weeks of field study	126
5.5	Comparison of the impact of the two traffic interventions on the traffic conditions	130

5.6	Comparison of the predicted emissions from the REVS and the NLR models, split by vehicle category.	131
5.7	Comparison of the spatial distribution of the NLR predicted NO _x emissions induced by the two traffic interventions	133
5.8	Comparison of the SIRANE predicted canyon averaged (a) NO concentrations, (b) NO ₂ concentrations and (c) O ₃ concentrations	134
5.9	Comparison of the SIRANE predicted pedestrian NO ₂ concentrations with the CAPS (CI) and the M200E (BL) measured NO ₂ concentrations	138
A.1	Comparison of the minimum number of trips needed per vehicle to capture the true standard deviation with an error lower than 20%	156
A.2	Comparison of the typical VSP distribution in London, based on the TfL car cycles, in different traffic flow conditions, with the measured VSP over the entire dataset	159
A.3	Comparison of the measured and predicted NO _x emission factors in rural driving conditions split by vehicle technology category	164
A.4	Comparison of the measured and predicted NO _x emission factors in motorway driving conditions split by vehicle technology category	165
A.5	Illustration of the procedure to follow to apply the model to a case study.	166
B.1	Illustration of a single artificial neuron	170
B.2	Illustration of a fully-connected multilayer perceptron with 2-hidden layers. Note that the layers do not have necessarily the same dimension, i.e. $n_0 \neq n_1 \neq n_2 \neq m$	171
B.3	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT	176
B.4	Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.	177

B.5	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Citroen Grand C4 Picasso	178
B.6	Instantaneous speed as a function of the distance travelled as a function of the distance travelled for the Citroen Grand C4 Picasso.	179
B.7	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Citroen DS5 DSTYLE BLUEHDI	180
B.8	Instantaneous speed as a function of the distance travelled for the Citroen DS5 DSTYLE BLUEHDI.	181
B.9	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 116D ED PLUS	182
B.10	Instantaneous speed as a function of the distance travelled for the BMW 116D ED PLUS.	183
B.11	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 220d SE184	
B.12	Instantaneous speed as a function of the distance travelled for the BMW 220d SE.185	
B.13	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Volkswagen Passat	186
B.14	Instantaneous speed as a function of the distance travelled for the Volkswagen Passat.	187
B.15	Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Ford S-Max TDCi Duratorq Titanium X	188
B.16	Instantaneous speed as a function of the distance travelled for the Ford S-Max TDCi Duratorq Titanium X.	189

B.17 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Golf SV SE TDi	190
B.18 Instantaneous speed as a function of the distance travelled for the Golf SV SE TDi.	191
B.19 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Volvo S60 R-DESIGN LUX NAV D3	192
B.20 Instantaneous speed as a function of the distance travelled for the Volvo S60 R-DESIGN LUX NAV D3.	193
B.21 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW X3 XDRIVE20D M SPORT AUTO	194
B.22 Instantaneous speed as a function of the distance travelled for the BMW X3 XDRIVE20D M SPORT AUTO.	195
B.23 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Mercedes C200 AMG Line BlueTec	196
B.24 Instantaneous speed as a function of the distance travelled for the Mercedes C200 AMG Line BlueTec.	197
B.25 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Peugeot 308	198
B.26 Instantaneous speed as a function of the distance travelled for the Peugeot 308. .	199
B.27 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT	200
B.28 Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.	201

B.29 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Nissan Qashqai	202
B.30 Instantaneous speed as a function of the distance travelled for the Nissan Qashqai.	203
B.31 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the Audi Q3 . . .	204
B.32 Instantaneous speed as a function of the distance travelled for the Audi Q3. . . .	205
B.33 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT	206
B.34 Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.	207
B.35 Instantaneous NO _x emissions as a function of the distance travelled and residuals of the two time series as a function of the distance travelled for the BMW 218D LUXURY	208
B.36 Instantaneous speed as a function of the distance travelled for the BMW 218D LUXURY.	209
B.37 Distribution of the 12 coefficients of the NLR ₁ model for each main cluster for 20 runs of training, with 10% of the datafiles randomly taken out as validation set.	210
B.38 Vehicle manufacturer, NO _x aftertreatment technology, CO, CO ₂ and NO _x emis- sion factors and f _{NO₂} per cluster	212
B.39 Vehicle manufacturer, NO _x aftertreatment technology, engine size, vehicle weight and CO ₂ emission factor per cluster	214
B.40 Aggregated instantaneous CO ₂ emissions for the main clusters	215
B.41 Vehicle manufacturer, NO _x aftertreatment technology, engine size, vehicle weight and NO _x and CO emission factors per cluster	216
B.42 Aggregated instantaneous CO emissions for the main clusters	217

B.43 Map of the standard deviation for the 400 runs and residuals between the mean MLP emission map and the mean NLR model emission map	218
B.44 Evolution of the relative error distribution for the MLP model as a function of the time history considered for the three training runs	220
B.45 Aggregated speed as a function of the distance travelled	221
B.46 Aggregated instantaneous NO _x emission factor as a function of the distance travelled.	222
B.47 Residuals between the aggregated measured NO _x emissions and the predicted NO _x emissions from the NLR ₁ , the MLP ₁ and COPERT5 models	224
C.1 Distribution of the mean NO _x emissions and its standard deviation per street . .	228

Chapter 1

Introduction

1.1 Background

More than half of the global population now lives in urban areas and this proportion is expected to continue to increase in the future [United Nations, 2019]. There is a large consensus on the fact that one of the greatest challenge with respect to this rapid urbanisation is air quality. Even though the common stereotype directly relates unsafe air quality to the pictures of the fog episodes repeatedly occurring in Beijing or Mumbai, air quality in European large urban areas is also in a state of crisis. The European Environmental Agency estimates that 467, 000 premature deaths in Europe can be directly attributed to air pollution [EEA, 2019], while the World Health Organisation attributes 4.2 million premature deaths worldwide to air pollution [World Health Organisation, 2016]. The cost of air pollution to European economies was quantified to US\$1.6 trillion (about 10% of the European GDP) in 2014 [Salvi and Brickman, 2015]. In the UK, the situation is not different: the Royal College of Physicians estimated that around 40,000 premature deaths could be directly attributed to outdoor air pollution [Royal College of Physician, 2016]. Bad air quality is also perceived as a threat by the general public, with a YouGov poll conducted in 2016 indicating that Londoners believed air pollution to be the “greatest health threat” facing the city [Cecil, 2016]. Finally, in these times of a global pandemic, new evidence shows that poor air quality is also statistically correlated to more

severe respiratory symptoms in some patients contracting COVID-19 [Copat et al., 2020].

In most western European countries including the UK, the most problematic issue relating to air quality is the concentration of NO_2 (nitrogen dioxide) in urban environment, particularly at roadside locations. In fact, the UK consistently fails to meet the European regulatory targets for NO_2 concentrations and has even been fined in 2018 by the European Court of Justice for this reason. One of the main causes of these high NO_2 concentrations, especially in urban environments, is road transport, which represents about one third of the total nitrogen oxides (NO_x) emissions in the UK [EEA, 2019]. The European Environment Agency even estimates that the contribution of road transport to population exposure to ambient NO_2 concentrations is substantially higher, especially in urban environments, where emissions are typically produced close to the ground and distributed across densely populated areas [EEA, 2019].

Exceedances of the annual mean concentration limit for NO_2 has been largely attributed to the discrepancies between type approval limits and real-world driving emissions (RDE) of diesel vehicles [Weiss et al., 2012, Fontaras et al., 2017b, O’Driscoll et al., 2016, O’Driscoll et al., 2018]. The difference between real world and type approval limits was thought to be due to the type approval being conducted in a laboratory and the test cycle not being representative of the real-world [Williams and Carslaw, 2011]. The NEDC (New European Driving Cycle, used for type approval in Europe until 2018) has been much criticised for being too smooth and covering only a little fraction of the engine operating range thus allowing manufacturers to optimise emissions performance for specific and unrealistic engine operating range [Weiss et al., 2012, Sileghem et al., 2014]. In addition to this, some manufacturers including Volkswagen illegally installed software able to cheat the type approval tests on some of their diesel vehicles [Contag et al., 2017]. The combination of these defeat devices to the non-realistic type approval procedure directly resulted in the poor air quality experienced nowadays in large European cities [Jonson et al., 2017].

To mitigate this issue, the Real-World Driving Emissions (RDE) test procedure was introduced for new vehicles from September 2017 onwards [European Commission, 2018a]. The RDE test procedure uses a Portable Emission Measurement System (PEMS), allowing vehicles to be

tested on a wide range of operating and environmental conditions, including but not limited to different levels of congestion, driving style or meteorological conditions. Repeated tests of the same vehicle exhibit subsequently a greater level of variability than laboratory tests where these parameters are controlled, making PEMS testing more representative of real-world situations [Weiss et al., 2011]. The introduction of the RDE test procedure was expected to drastically reduce the discrepancies between real-world driving and type-approval emissions [Weiss et al., 2012, Degraeuwe et al., 2016].

Improvement of type approval is therefore expected in the long run to lead to a general decrease in emissions and an improvement of air quality, particularly in urban areas. However, in large cities, rapid compliance with air quality regulations requires the implementation of additional mitigation strategies. The challenging issue is therefore to produce accurate emissions models that, combined with atmospheric dispersion models, are able not only to assess the impacts of road transport on the local air quality but also to assess the evolution of emissions over the years and evaluate the effectiveness of a specific mitigation strategy. The development of such models typically requires a good understanding of the factors affecting real-world driving emissions.

We therefore devote this PhD thesis to the study of real-world driving NO_x emissions, the development of real-world emission models and their application to the evaluation of mitigation strategies. The three investigative chapters focus on three topics: 1) the impact of driving dynamics on emissions and the development of an aggregated real-world NO_x emission model, 2) the intra and inter-vehicle real-world variability in emissions and the development of an instantaneous NO_x emission model and 3) building on the two previous chapters and on data collected during a field study, the analysis of the impact of two traffic interventions on both the emissions and the local air quality.

To the best of our knowledge, there is no large spatial and temporal real-world driving emissions study, as most PEMS studies, which typically offer both good spatial and temporal resolution, analyse less than 10 vehicles [Weiss et al., 2012, Franco et al., 2014, Zhang et al., 2016, Gallus et al., 2016, O’Driscoll et al., 2016, Gallus et al., 2017, Demuynck et al., 2017, O’Driscoll et al.,

2018]. Remote sensing studies, on the other hand, typically collect real-world data for thousands of vehicles but at the cost of a very reduced spatial and temporal resolution [Carslaw and Rhys-Tyler, 2013, Carslaw et al., 2019]. The results of most PEMS data studies are therefore hard to generalise, while the conclusions of most remote sensing studies can be location-specific [Carslaw and Rhys-Tyler, 2013, Carslaw et al., 2019]. The first two investigative chapters aim to fill this gap. New and previously unused PEMS data sources make it possible to provide an extensive analysis of the factors affecting real-world driving emissions, with a focus on urban environment. This knowledge can be used in a second step to develop accurate real-world instantaneous or aggregated emissions models that can, in turn, be used to design and evaluate air quality mitigation strategies. Building on the two previous chapters, our last investigative chapter focuses on evaluating the impact of a traffic intervention on both emissions and local air quality. It aims to provide both theoretical and empirical insights on the efficacy of a local scale traffic intervention (single junction on a busy road), which is in contrast with most studies that typically adopt one or the other approach.

1.2 Research aims and objectives

This research project aims to improve the understanding of the physical and non-physical variables affecting real-world emissions, to improve real-world vehicle emissions models and to provide theoretical and empirical insights about the real-world impact of an air quality mitigation strategy on both emissions and the local air quality.

Although each chapter of the thesis has its own individual and clearly defined set of research questions, a number of overarching problem statements can be formulated for the PhD research as a whole. Let us begin with a detailed statement of the research questions on real-world vehicle emissions:

What are the factors affecting directly real-world driving NO_x emissions? Are internal or external factors the most important and to which extend? How do the influence of these factors evolve with the successive regulations?

Real-world emissions of each vehicle within a certain fleet are therefore impacted by a certain number of factors, but the extent of the impact of each factor varies from one vehicle to the other. In practice however, most air quality studies only consider the average emissions of a fleet [Oxley et al., 2009, Soulhac et al., 2012]. Nonetheless, an increasing number of studies emphasise the importance of also considering the variability in real-world driving emissions [ODriscoll, 2017], hence raising the following questions:

How can we quantify the real-world variability in emissions? Knowing that only a part of the fleet is sampled, what is the influence of sampling on the measured real-world variability in emissions?

The identification and the quantification of the factors affecting real-world emissions and its variability can be integrated into the development of emission models:

How can we integrate the factors directly impacting real-world driving NO_x emissions in emission models? What is the improvement in terms of accuracy of such a model? What are the limitations of the developed models?

Finally, this research project aims to analyse the performance of the developed emissions models when applied to a real case field study aiming to evaluate the impact of a traffic intervention on local air quality. Research questions directly raised are:

What are the impact of a traffic intervention on the local emissions? Is there a measurable and statistically significant impact of the traffic intervention on the measured local pollutants concentrations? What are the causes of the measured differences?

1.3 Structure of the thesis

The remaining parts of this thesis are organised as follow:

Chapter 2 introduces the key concepts behind this thesis. In particular, it provides an in-depth review of the literature on emissions from on road vehicles and air quality mitigation strategies.

It also provides highlights on pollutant dispersion, air quality regulations and the impact of poor air quality on health, ecosystems and economies. Although these latest topics are not the direct focus of this thesis, they are part of the essential broader picture.

Chapter 3 presents an extensive analysis of the impact of driving dynamicity on real-world NO_x emissions, based on the largest PEMS dataset currently recorded in the literature. It lays down the statistical foundations necessary to analyse PEMS tests acquired in different conditions and presents a new aggregated NO_x emission model able to predict not only the average NO_x emission factor but also its standard deviation for different segments of the vehicle fleet.

Chapter 4 uses Machine Learning techniques to gain insights into real-world NO_x emissions of Euro 5 and Euro 6 diesel vehicles. It develops and compares three instantaneous real-world NO_x emissions models. The analysis of the three proposed models is assessed on both the available data and COPERT5, an average-speed emission model used by local authorities in 22 out of the EU-27 countries.

Chapter 5 is built on the two previous chapters and analyses the impact of a change in traffic light timings at a single junction on both the local emissions and the local air quality. It compares the modelled impact of a traffic intervention on emissions to the pollutant concentrations effectively measured at the side of the road. Analysis of the traffic provides insights on the effective impact of the aforementioned traffic intervention on the local traffic and the reasons why differences in pollutant concentrations are modelled and measured.

Chapter 6 presents the conclusions and summarises the research objectives met throughout of each the chapters of this thesis. It also identifies the limitations of the proposed approaches and presents suggestions for future work and research continuity.

Appendices present further descriptions of relevant information to this thesis:

Appendix A presents the Supplementary Information (SI) associated with Chapter 3.

Appendix B presents the Supplementary Information (SI) associated with Chapter 4.

Appendix C presents the Supplementary Information (SI) associated with Chapter 5.

1.4 Contributions

This thesis contributes substantially to the literature on real-world vehicle emissions analysis and modelling. It also provides significant insights on the real-world impact of a traffic intervention on the local air quality.

1. To the best of our knowledge, the analysis presented in Chapter 3 is based on the largest PEMS dataset recorded, which comprised 287 Euro 5 and Euro 6 diesel and petrol vehicles. It lays down the foundations to derive reliable fleet statistics from unbalanced PEMS tests and consequently calibrate reliable fleet emissions models.
2. The analysis conducted in Chapter 3 is the first of its kind that demonstrates the impact of driving dynamicity on real-world driving NO_x emissions. Earlier studies, reviewed in Section 2.2.2.4, were conducted on considerably smaller datasets and did not analyse the evolution of the impact of driving dynamicity with the successive Euro standards. This thesis contributes with an innovative method to compare the relative influence of driving dynamicity on real-world NO_x emissions from vehicles belonging to the same vehicle technology category with different baseline emissions.
3. Chapter 3 also provides a new aggregated NO_x emission model that accounts directly for the impact of driving dynamics on real-world NO_x emissions. This proposed model goes beyond the capabilities presented by most aggregated emissions models including COPERT, as it predicts not only the average NO_x emission factor but also its standard deviation. The model can also be applied to derive the average NO_x emission factor and its standard deviation of a fleet.
4. In Chapter 4, this thesis uncovers new insights on the internal and external factors directly affecting NO_x emissions of diesel vehicles. In particular, it quantifies the significant intra-vehicle variability in emissions in real-world driving conditions, demonstrating that, on the same road, with the same driver, in the same traffic conditions, a factor 2 difference between repeated test of the same vehicle is not unusual. Chapter 4 also shows that

physical factors such as vehicle weight or vehicle engine size are not correlated with real-world NO_x emissions.

5. To the best of our knowledge, the innovative clustering analysis presented in Chapter 4 is the first of its kind to be used not only to analyse real-world driving emissions but also to drastically reduce the number of emission models needed to accurately capture instantaneous NO_x emissions. The comparison between different types of emissions models carried out in Chapter 4 reveals that if only the instantaneous speed and acceleration are known, an artificial neural network model is not able to perform better than a non-linear regression and that the latter should be preferred.
6. Finally, we devote Chapter 5 to the analysis of the real-world impact of a change in traffic light timings at a single junction on local air quality. The analysis presented is based on data collected during a two-weeks field study conducted in September 2019. The study is conducted both theoretically, using simulations, and empirically, using the data collected, going beyond most existing studies. The data collected and the simulations conducted allowed us to: 1) analyse how the traffic interventions impacted the traffic flow and road transport NO_x emissions and 2) quantify the variability in NO_2 concentration attributable to the traffic interventions, the meteorological conditions and the location of the sensors with respect to the main source of emissions. It is demonstrated that while the lengthening of the traffic lights cycle could explain change in pedestrian level NO_2 concentration comprised between -0.06 and 1.8 ppb, most of the variability in measured concentrations can be directly attributed to the meteorological conditions (up to 13.9 ppb) and the location of the sensors with respect to the main source of traffic. This chapter provides very interesting insights on how the optimisation of traffic light timings could help mitigating the impact of road traffic on pedestrian exposure.

The list of key contributions is further detailed in the concluding sections of Chapters 3 to 5.

1.5 Publications

A preliminary work on which Chapter 5 is partially based was published as:

Tom Grylls, Clémence M.A. Le Cornec, Pietro Salizzoni, Lionel Soulhac, Marc E.J. Stettler and Maarten van Reeuwijk (2019). Evaluation of an operational air quality model using large-eddy simulation. *Atmospheric Environment*, 3, 100041.

The research outcomes of Chapter 4 were published as:

Clémence M. A. Le Cornec, Nick Molden, Maarten van Reeuwijk and Marc E. J. Stettler (2020). Modelling of instantaneous emissions from diesel vehicles with a special focus on NO_x: Insights from machine learning techniques. *Science of The Total Environment*, 737, 139625.

An article in partnership with the Data Learning Meetings at Imperial College London was submitted to the Journal of Computational Sciences as:

Rossella Arcucci, Caterina Buizza, Cesar Quilodran Casas, Philip Nadler, Julian Mack, Stefano Marrone, Zainab Titus, Clémence Le Cornec, Evelyn Haylen, Tolga Hasan Dur, Luis Baca Ruiz, Claire Heaney, Julio Amador Diaz Lopez (2021). Data Learning: Integrating Data Assimilation and Machine Learning. *Journal of Computational Sciences*. [Under review]

The research outcomes of Chapter 3 were submitted to Science of the Total Environment as:

Clémence M. A. Le Cornec, Maarten van Reeuwijk, Nick Molden and Marc E. J. Stettler (2021). REVS: Road transport Emissions and Variance Simulator. Submitted to Science of the Total Environment.

Chapter 2

Traffic related air pollution

2.1 Introduction

Ambient air pollution is a global threat, with numerous adverse impacts on human health and ecosystems. As a consequence of economic growth, pollutant emissions and concentrations have dramatically increased worldwide over the past century. In Europe, air quality remains poor in many areas, despite the introduction of more stringent legislation [EEA, 2019]. Effective actions to mitigate air pollution and its impacts require a good understanding of its causes, how pollutants are synthesised and emitted, how pollutants are dispersed and transformed in the atmosphere and finally how these pollutants interact and influence both humans and ecosystems [Bickel and Friedrich, 2005]. For the sake of clarity, this “impact pathway approach” will be used in the remainder of this chapter to introduce the key concepts of this thesis.

This chapter has the following structure: Section 2.2 gives an extensive overview on vehicles emissions, describing their formation as a result of combustion, the existing regulations and the current state of the art on emissions modelling. Section 2.3 then gives an overview on pollutant dispersion in urban environments, while Section 2.4 discusses the current air quality regulations and the impacts of air pollutants on human health and ecosystems. Finally, Section 2.5 presents an overview of air quality mitigation strategies.

2.2 Emissions

2.2.1 Sources of emissions

Air pollutants can be classified as primary or secondary. Primary pollutants are emitted directly into the atmosphere, while secondary pollutants are synthesised in the atmosphere from primary pollutants through physico-chemical reactions [EEA, 2019]. Air pollutants can be produced naturally, as a result of anthropogenic activities or be of mixed-origin, depending on the sources of their precursors. Key primary pollutants include particulate matter (PM), black carbon (BC), nitrogen oxides (NO_x) and carbon monoxide (CO) among others. Key secondary pollutants include PM, NO_2 and volatile organic compounds (VOCs), among others [EEA, 2019].

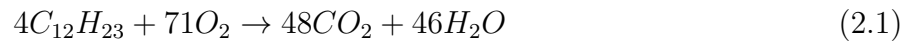
The European Commission distinguishes between six main sectors contributing to air pollution: 1) transport, split into road transport and non-road transport (rail, maritime, air), 2) commercial, institutional and households, 3) energy production and distribution, 4) industry, 5) agriculture and 6) waste (landfill and waste incineration) [EEA, 2019]. Each of these sectors contributes in different proportions to the primary and secondary air pollutants cited above. Quantifying the contribution of each of these sectors to ambient air pollution in Europe is beyond the scope of this thesis, and the interested reader is referred to [EEA, 2019] for a detailed description. This thesis focuses on air pollution caused by road transport and consequently does not discuss the impact or contributions of other sources.

In Europe, the road transport sector was the largest emitter NO_x (39%) and the second largest emitter of BC (28%), CO (19%), primary $\text{PM}_{2.5}$ (11%) and Pb (20%) in 2017 [EEA, 2019]. It is estimated that the contribution of road transport to population exposure to air pollution in urban areas, where emissions are typically produced close to the ground, is considerably higher [EEA, 2019]. In London, road transport is estimated to be the largest NO_x emitter, accounting for 47% of the total NO_x emitted in 2012. This aspect is reflected by the high pollutants concentration typically encountered in cities: in the UK, [Harrison et al., 2012] reported $\text{PM}_{2.5}$ urban background concentrations 3-6 $\mu\text{g}/\text{m}^3$ above the rural background (10-11 $\mu\text{g}/\text{m}^3$) while

PM_{2.5} concentrations along the kerbside reach the EU limit value 25 $\mu\text{g}/\text{m}^3$ in some locations [Harrison et al., 2012]. In Spain, the high NO₂ concentrations recorded in 14 Spanish cities were directly linked to 6,085 excess deaths, about three times higher than the excess deaths attributable to PM₁₀ concentrations [Linares et al., 2018].

2.2.2 Formation of emissions as a results of combustion

Combustion is defined as a “*rapid oxidation generating heat, or both light and heat*” [Turns, 1996]. Combustion transforms the energy stored in chemical bonds to heat that can be used in a variety of ways. One example are the internal combustion engines, on which our transportation system relies heavily. Combustion engines use the oxygen present into the air to transform fuel-based chemical energy into heat that is then converted into mechanical work. During the combustion process, the heat released into the engine working medium induces a rise in pressure, which, by exploiting expansion, can be converted into piston motion [Mollenhauer et al., 2010]. Hydrocarbons oxidise in three steps. The first step consists in breaking-up carbon-carbon bonds in the fuel molecule to form alkenes (hydrocarbons with double carbon bonds) and hydrogen. During the second step, the molecule is further oxidised to form CO and hydrogen. In the last step, CO is oxidised further to form CO₂. Most of the energy is released during this third step [Ferguson and Kirkpatrick, 2015]. In the case of a complete and ideal combustion of fuel (for which C₁₂H₂₃ is considered as a representative molecule) with pure oxygen, only carbon dioxide and combustion water are produced in addition of the desired heat (Equation 2.1):



The combustion processes in both spark-ignition (petrol) and compression-ignition (diesel) engines are extremely complex and depend on the type of fuel and the amount of air used in the combustion process [Ferguson and Kirkpatrick, 2015].

Petrol and diesel fuels are primarily obtained by distillation of petroleum crude oil, a process during which the various hydrocarbon fractions contained in the crude oil are physically sep-

arated using their differences in boiling point [Ferguson and Kirkpatrick, 2015]. The various fractions extracted are then chemically processed into fuels and other products. Petrol fuel is a blend of hydrocarbons distillates with a range of boiling points between 25°-225°C while diesel fuel is a blend of hydrocarbon distillates with a range of boiling points ranging between 180°C and 360°C. Some of the differences between petrol and diesel exhausts are related to the difference in chemical composition. Diesel fuels have approximately an 8% greater energy density by volume than petrol fuels which, coupled to the diesel engines improved pumping efficiency [Ferguson and Kirkpatrick, 2015], is the reason why diesel vehicles historically presented lower CO₂ emissions than petrol.

The rest of the section details the formation of exhaust emissions in both diesel and petrol engines, highlights the differences between these two types of fuels and describes the common exhaust gas aftertreatments used to reduce the harmful emissions from both petrol and diesel engines.

2.2.2.1 Diesel engines

The combustion process in a diesel engine is initiated by the autoignition of the fuel that is injected into the combustion chamber, where air was previously introduced and highly compressed, raising its temperature above 250°C [Mollenhauer et al., 2010]. This mechanism means that the mixing of air and fuel occurs directly inside the chamber. Diesel engines operate under a wide range of air/fuel ratios (λ), though nearly always higher (i.e. more air) than the stoichiometric ratio and are said to run “lean”. The throttle, that controls the amount of air injected in a petrol engine, is therefore not required in a diesel engine and running lean allows to reduce the amount of fuel needed to produce a certain power output, resulting in lower fuel consumption and CO₂ emissions. In practice, the internal mixture formation and the delayed injection of the fuel into the combustion chambers lead to non-uniform air/fuel gradients and temperature distribution in the chamber. NO_x are formed in the zones with excess air and high temperatures ($> 1500^\circ\text{C}$), while air deficient zones produce soot particulates and carbon monoxide [Mollenhauer et al., 2010]. Unburnt hydrocarbons are generated in low

temperatures zones. Sulphur compounds, aldehydes, cyanides, ammonia, special hydrocarbons (benzene) and polycyclic aromatic hydrocarbons (phenatrene, pyrene, fluorene) can also be formed [Mollenhauer et al., 2010]. The chemical composition of the untreated exhaust gas depends significantly on the engine operating conditions (fuel injection pressure, injection timing, exhaust gas recirculation and engine load).

2.2.2.2 Petrol engines

Petrol engines are commonly known as spark ignition engines and can be divided into two categories: conventional Port Fuel Injection (PFI) engines, where the air is mixed with fuel before the intake stroke and the Gasoline Direct Injection (GDI) engines, where the fuel is directly injected into the combustion chamber.

In a PFI petrol engine, the fuel is injected into the intake manifold where it mixes with air. The mixture is then sucked down into the cylinders, where it is ignited by an external energy source (i.e. a spark) [Ferguson and Kirkpatrick, 2015]. PFI engines operate at an air/fuel ratio close to stoichiometry, resulting in lower particules formation compared to diesel engines.

In a GDI engine, the fuel can be distributed according to two modes depending on the load and engine speed: homogeneous mode and stratified-charge mode [Çelik and Ozdalyan, 2010]. In homogeneous mode, which is typically used for acceleration, full loads and high engine speed, a stoichiometric or slightly rich air/fuel mixture is injected in the combustion chamber during the intake stroke [Çelik and Ozdalyan, 2010]. In stratified-charge mode, which is used for light loads, at constant or low speeds, where no acceleration is required, the fuel is highly pressurised (30-130 bars) and directly injected into the combustion chamber of the cylinder shortly before the ignition [Ferguson and Kirkpatrick, 2015]. In stratified-charge mode, GDIs operate under lean conditions, which allows to reduce both fuel consumption and CO₂ emissions, at the expense of increased particulates emissions and NO_x caused by the inhomogeneous air/fuel mixing. Compared to PFI engines, GDIs were found to result in fuel consumption savings between 3 and 14.5% [Chan et al., 2012, Saliba et al., 2017], while presenting up to two times higher PM, PN and ultrafine particules emissions [Liang et al., 2013, Wang et al.,

2014, Saliba et al., 2017]. New after-treatment technologies, such as Gasoline Particulate Filters (GPFs) have been found to successfully reduce both PM and PN emissions from GDIs [Saliba et al., 2017].

2.2.2.3 Aftertreatment technologies

This section describes the most common aftertreatment technologies used to reduce tailpipe emissions in modern petrol and diesel engines.

Three-way catalysts

Three-way catalysts (TWCs), invented in 1950, were the first aftertreatment system installed on passenger cars in 1975. They now equipped all petrol passenger vehicles. Noble-metal catalysts (platinum, palladium and rhodium), spread over a 3D honeycomb structure, provide active sites for reactions that oxidise CO and unburnt hydrocarbons to CO₂ and H₂O while simultaneously reducing NO to N₂ [Turns, 1996, Ferguson and Kirkpatrick, 2015]. The conversion efficiency of TWCs at fully warm conditions is about 98-99% for CO and 95% for HC. To achieve high rates of conversion, the exhaust gas composition through the TWC should be close to the stoichiometric ratio. If the exhaust is too lean, NO is not reduced and if the exhaust is too rich, CO and HC are not oxidised. The engine must therefore operate in a narrow range about the stoichiometric fuel-air ratio [Turns, 1996, Ferguson and Kirkpatrick, 2015].

Diesel Oxidation Catalyst

Diesel engines operate lean, far above the stoichiometric ratio, and therefore cannot use a traditional TWC system to reduce CO, HC and NO. Diesel Oxidation Catalysts (DOCs) were the first catalytic converters installed on diesel vehicles as standard in the 1970s. Similarly to TWCs, precious metal coating (palladium and platinum) spread over a 3D honeycomb structure are used to oxidise CO and unburnt hydrocarbons into the harmless gases CO₂ and H₂O. It also partially oxidises NO to NO₂, which is a conducive step for nitrogen oxide reduction, especially when the vehicle is fitted with an SCR system and improves passive particulate filter regeneration. Finally, the DOC can be supplied with HC and CO by running the engine rich

to generate the heat necessary for particulate filter regeneration [Mollenhauer et al., 2010].

Particulate filters

From Euro 5 onwards, all diesel vehicles are equipped with a Diesel Particulate Filter (DPF) to reduce PM emissions. Depending on the filtration mechanisms, the DPF can be classified as a wall-flow or a flow-through filter. Flow-through filters present better performance in terms of fuel consumption than wall-flow filters but at the cost of a decreased filtration efficiency (below 50%) [Neeft et al., 1996]. Wall-flow filters, on the other hand, are able to reach filtration efficiency higher than 95% [Bergmann et al., 2009, Mollenhauer et al., 2010]. In practice, almost only wall-flow filters are fitted on diesel vehicles, which ensure the fulfillment of the regulations. Wall-flow filters are honeycomb monoliths made of ceramic materials where every other channel is plugged at each end to force the exhaust flow through the porous filter wall [Lapuerta et al., 2012]. Over time, the deposition of particles on the filter increases the flow restriction, creating an exhaust back-pressure, which negatively affects the performance of the engine [Mollenhauer et al., 2010]. When the exhaust back-pressure increases to an unacceptable level, a regeneration process (active or passive) must be induced in order to burn the accumulated soot and restore the filter efficiency. During an active regeneration process, the exhaust gas temperature is artificially increased to 600°C by injecting fuel in the exhaust pipe before the DOC, resulting in the oxidation of the soot accumulated on the filter by O_2 [Reşitoğlu et al., 2015]. Depending on the design, filter regeneration is usually necessary after 300-800 km [Mollenhauer et al., 2010]. During a passive regeneration process, PM is oxidised by NO_2 at temperature ranging between 250-450°C (which are frequently reached in commercial vehicles applications and in car applications during motorway driving) and therefore does not require the use of additional fuel [Schejbal et al., 2010, Reşitoğlu et al., 2015]. On average, it is estimated that the use of a DPF increases CO_2 emissions by 2-5% because of: 1) an increase in back-pressure requiring additional mechanical work to be overcome and 2) the fuel necessary to trigger an active regeneration event [Lapuerta et al., 2012].

Exhaust Gas Recirculation

Exhaust Gas Recirculation (EGR) is an effective strategy to reduce NO_x emissions from in-

ternal combustion engines. It consists in recycling a fraction of the exhaust gas at the engine intake, which lowers the combustion temperature and O_2 concentration and consequently, NO_x formation [Ferguson and Kirkpatrick, 2015]. The exhaust gas recirculation rate is defined as the proportion of the intake mixture that is recycled and can be up to 50% [Abd-Alla, 2002, Mollenhauer et al., 2010]:

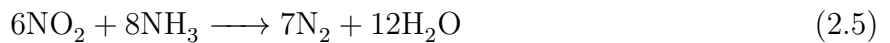
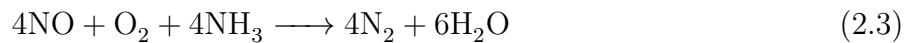
$$EGR(\%) = 100 * \frac{m_{EGR}}{m_{EGR} + m_{air} + m_{fuel}} \quad (2.2)$$

The use of EGR incurs penalties in terms of both PM emissions and fuel consumption [Abd-Alla, 2002]. The limitation of the oxygen needed to oxidise soot is the primary reason for higher PM emissions, while slower combustion and thus the shift of the position of the center of combustion toward retardation leads to an increase in fuel consumption [Mollenhauer et al., 2010]. The increase in soot formation caused by EGR can lead to carbon deposits that adversely affect long-term engine durability [Abd-Alla, 2002].

All diesel passenger vehicles from Euro 5 standard onward are equipped with an EGR system to reduce NO_x formation. An EGR system is typically not sufficient to meet Euro 6 type approval limits and most Euro 6 diesel deploy additional technologies, such as Selective Catalytic Reduction and Lean NO_x Trap, to remove NO_x from the exhaust gas after its formation.

Selective Catalytic Reduction

Selective Catalytic Reduction (SCR) system, initially developed in the 60s for large firing plants, uses liquid NH_3 sprayed into the exhaust stream to reduce NO_x to nitrogen as the flow passes through a catalytic converter [Mollenhauer et al., 2010, Ferguson and Kirkpatrick, 2015]. The SCR reactions are given by:

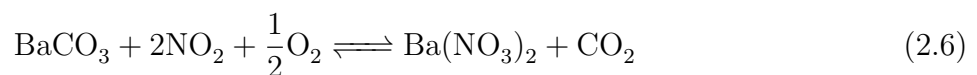


The first two reactions are dominant in most cases. A significant quantity of NH_3 is needed,

about 0.3-1% of the fuel quantity [Mollenhauer et al., 2010]. Due to the toxicity of ammonia, the use of ammonia injection systems is limited to large marine engines. For vehicular applications, a mixture of 32.5% aqueous urea solution (called the Diesel Exhaust Fluid or DEF) is used instead [Ferguson and Kirkpatrick, 2015]. The DEF is injected in before the SCR catalyst and is hydrolysed into ammonia and CO₂ at temperatures above 250°C. The ammonia produced is then adsorbed in the catalytic converter, a honeycombe structure of precious metals (Platinum, Rhodium), where it is available for the SCR reactions. High reduction rate of NO_x to N₂ by NH₃ can be obtained at temperatures between 180-450°C (250°C and above to allow for the hydrolysis of urea) [Mollenhauer et al., 2010]. At temperature below 300°C, the reaction mainly proceeds through Eq. 2.4 and therefore an oxidising catalytic converter (such as the DOC) situated upstream from the SCR catalyst and the urea injection point can be used to increase the ratio of NO₂ to NO_x to around 50% and boost the reaction [Ciardelli et al., 2007]. One of the criticisms of SCR systems is the sensitivity to the amount of DEF injected. In particular, unconverted NH₃ can be released into the environment (ammonia slip) if the quantity of urea injected exceeds the catalytic converter's adsorption capacity or if incomplete hydrolysis of the aqueous urea solution occurs.

NO_x Storage Catalyst

NO_x Storage Catalyst (also called Lean NO_x Trap or LNT) allows NO_x reduction without external additional agent. A LNT system works sequentially with a first loading phase of approximately 30 to 300 seconds in which NO_x from the lean exhaust gas is stored as nitrate in the catalytic converter's storage component followed by a regeneration phase of approximately 2 to 10 seconds during which the stored NO_x is reduced to N₂ by a fuel-rich desorption process (similar to DPF regeneration) [Mollenhauer et al., 2010, Ferguson and Kirkpatrick, 2015]. LNT are typically coated with oxides and carbonates of alkali and alkaline earth metals (such as Ba, Sr, Ca) that show a high propensity to form stable but chemically reversible bonds with nitrogen oxides. The trapping is most effective if exhaust NO is first oxidised to NO₂ [Harold, 2012], before reacting with the storage compounds in the coating and O₂ to form nitrate:



Storage is only optimal at exhaust gas temperature between 200-250°C and 350-450°C due to kinetic limitations encounter at low temperatures and unstable formation of nitrate above 450°C leading to NO_x thermal discharge [Mollenhauer et al., 2010, Harold, 2012]. The NO_x storage continues until all the accessible storage sites are filled, leading to a NO_x breakthrough [Harold, 2012]. Regeneration is then performed to restore the storage components by discharging the stored nitrogen oxides and converting them to N₂ and CO₂. The discharge of the stored nitrogen oxides requires rich exhaust gas conditions, with CO and HC acting as reductants to reform the carbonate originally present [Harold, 2012]:



Finally, similarly to three-way catalysts, a rhodium or platinum coating is used to reduce nitrogen oxides to N₂ and oxidise CO to CO₂:



LNTs present two main issues: thermal aging and sulphur poisoning [Choi et al., 2007]. Sulfur compounds in the fuel and lubricating oil are oxidised to sulphur dioxide (SO₂) during combustion, which presents an higher bonding affinity and strength to the LNT storage compounds (BaCO₃) than nitrate [Mollenhauer et al., 2010]. As sulfates are not removed during normal regeneration, they accumulate with time and block the formation and storage of nitrates [Choi et al., 2007]. Desulfation is typically needed after 500 to 2,500 km and requires to heat the catalyst above 650°C for more than 5 minutes [Mollenhauer et al., 2010]. Such high temperatures can result in thermal aging with permanent performance loss [Choi et al., 2007].

Comparison of NO_x reduction systems

LNT and SCR systems present a multitude of significant differences and the choice of the system fitted on a vehicle depends on the legal requirements and boundary conditions. At present, SCR systems present better NO_x removal than LNT systems [Myung et al., 2017], which is a fundamental criterion to achieve compliance under stringent regulation. However,

one of the main issue of SCR systems is the requirement for an external agent (e.g. AdBlue) to be stored in the vehicle and regularly replenish [Johnson, 2008]. LNT systems, on the other hand, do not require any external agent but increase the fuel consumption by 2-4%, making them unattractive for large engines [Johnson, 2008]. Because of the fixed-cost of an on-board urea system, LNTs are usually cheaper than SCRs for passenger vehicles with small engines. It is generally admitted that when their NO_x reduction is comparable, SCR systems present a lower cost than a LNT system [Mollenhauer et al., 2010].

2.2.2.4 Factors impacting emissions

Numerous external factors are known to have a direct impact on vehicle emissions, including driving style, congestion, road grade, external temperature and cold starts.

A cold start refers to an attempt to start a vehicle engine when it is cold relative to its normal operating temperature ($< 30^\circ\text{C}$) [André and Joumard, 2005]. During the first minutes of operation, the catalytic converter operates at temperatures below its optimum (situated around $300\text{-}400^\circ\text{C}$), preventing the efficient removal of HC, CO, PM and NO_x . For instance, [Martin et al., 2017] show that during cold-starts Tier 2 vehicles (manufactured between 2004 and 2016) present on average HC emissions 11.5 times higher, NO emissions 12.5 times higher and CO emissions 10.3 times higher than during a 20-minutes hot start. The magnitude of the emissions produced during a cold-start event is largely influenced by the external ambient temperature [Weilenmann et al., 2009]. [Varella et al., 2017] found that for a decrease in external temperature from 10 to 5°C , NO_x emissions from a diesel engine increase by up to 30%. A larger study conducted by [Varella et al., 2019] shows that the CO_2 emission factor is on average 25% higher during a cold start, while the NO_x emission factor increased by 55%. Overall, the impact of a cold starts on emissions is more pronounced for petrol than diesel vehicles [Weilenmann et al., 2005, Weilenmann et al., 2009, Varella et al., 2019]. Previous studies estimated the cold start period to span over the first 1 to 5 km of urban driving [André and Joumard, 2005, Favez et al., 2009]. For short trips (< 5 km), which are known to represent about half of all journeys in the UK [de Nazelle et al., 2010], cold start emissions are therefore significant and can last for the

entire duration of the trip.

Driving style or driver variability can be defined as the differences in vehicle operating conditions between drivers for the same trip (e.g. same road, similar level of congestion, same vehicle) and may include variations in duration, frequency or intensity of cruising, accelerations and decelerations [Holmén and Niemeier, 1998]. Over the past 20 years and with the recent development of the so-called “eco-driving”, the influence of driving style on vehicle exhaust emissions increasingly became a topic of interest for both the general public and the research community. An early study conducted by [Jackson et al., 2006] where 20 drivers were asked to drive the same vehicle along the same route resulted in a wide range of tailpipe emissions (74 s.d. 108 ppm for NO_x in urban driving conditions). Unfortunately no analysis of driving dynamicity was conducted alongside the analysis of tailpipe emissions and therefore only limited conclusions can be withdrawn. A similar study performed by [Frey et al., 2008] found that average emissions differences between drivers when stratified by route, travel direction, vehicle and time of day ranged between 4 to 5% for CO₂, 9 to 11% for HC, 16 to 18% for NO_x and 102 to 114% for CO. These differences are linked to variation in the speed profile and accelerations. [Ericsson, 2001] draw similar conclusions by showing that harsh accelerations ($> 1.5 \text{ m/s}^2$) result in a significant increase in HC, NO_x and CO₂ emissions. According to this study, other parameters impacting fuel consumption and therefore CO₂ emissions are the proportion of time spent idling, the relative positive acceleration (RPA) and the proportion of time at speeds $< 15 \text{ km/h}$ and between 50 and 70 km/h (where vehicles are the most fuel efficient). [Luján et al., 2018] (respectively [Prakash and Bodisco, 2019]) confirm the results from [Ericsson, 2001] for a Euro 6 (respectively a Euro 5) vehicle and demonstrate that NO_x and NO₂ emissions are directly linked to abrupt accelerations, which typically happen during urban driving. The high CO₂ emissions measured during urban driving are related to the inefficient operation of the engine due to long period of idling. Contrary to NO_x emissions, [Luján et al., 2018] show that HC and CO emissions are mainly driven by motorway driving.

Road grade is commonly admitted to impact directly on exhaust emissions. [Boriboonsomsin and Barth, 2009] demonstrated that driving on a flat road resulted in a fuel economy (and therefore the CO₂ emissions reduction) comprised between 15 to 20% for the flat road compared

to the hilly road. [Gallus et al., 2017] investigated the effect of road grade on CO_2 and NO_x emissions for one Euro 5 and one Euro 6 diesel light-duty vehicles running at cruising conditions during on-road tests and showed that emissions increase by up to 66% to 81% for CO_2 and 94 to 115% for NO_x for a 5% increase in road grade. Similar results were found by other studies [Sentoff et al., 2015, Varella et al., 2019].

2.2.3 Emissions regulations

Exhaust gas is emitted as a result of the combustion of fuels (e.g. diesel, petrol, coal). It is mainly composed of dinitrogen N_2 , water vapor H_2O and carbon dioxide CO_2 , which are harmless to human health. A small fraction of the exhaust (around 0.1%) is composed of undesirable and toxic components including carbon monoxide (CO) arising from incomplete combustion, hydrocarbons (HC) that are a byproduct of unburnt fuel, nitrogen oxides (NO_x) formed under excessive combustion temperatures and particulate matter (PM) arising from a lack of oxygen during the combustion process [Mollenhauer et al., 2010].

In order to achieve air quality standard and protect human health, exhaust emissions are regulated by emission standards, which define the acceptable level for exhaust emissions of the new vehicles sold on the market. The USA was the first country to impose emission standards in the early 1960s [Mollenhauer et al., 2010]. Japan and Europe followed about 10 years later. Regulations are mainly focused on limiting PM, NO_x , CO, HC and CO_2 emissions. Historically, emission tests were performed on a chassis-dynamometer over standardised driving cycles that aimed to represent a wide range of driving conditions. The Volkswagen scandal of 2015 attracted the attention on the growing discrepancy between type-approval limits and real-world driving emissions for both NO_x and CO_2 , triggering the development of more stringent type-approval procedures in Europe, alongside with legal actions both in Europe and in the USA. Production should conform to legislation and manufacturers have to demonstrate compliance with emissions limits over the entire useful lifetime of the vehicle (160,000 km in Europe, 193,000 km in the USA) [Mollenhauer et al., 2010].

2.2.3.1 American emission standards

Emissions limits in the USA are defined by the Environmental Protection Agency (EPA) but states are free to apply stricter limits (such as California).

The emission limits in the USA are defined as tiers, Tier 3 being the most recent (phased-in from 2017 to 2025, it aligns with the LEV III Californian emission limits, see Table 2.1). Each tier contains several bins (eight bins in total, 1 being the strictest and 8 the least stringent), that defines sets of limits [Mollenhauer et al., 2010]. Manufacturers are free to choose between several bins but must reach an average fleet NO_x emissions of 30 mg/mile by 2025. In practice, it greatly reduces the manufacturers possibilities as bin 5 must be reached on average. One particularity of the American system is that the same emission limits apply to diesel, petrol or alternative fuel vehicles. Tier 3 divides the vehicles into four categories: light-duty vehicles (up to 3855 kg), light-duty trucks (between 3,855 kg and 4,535 kg), medium-duty passengers vehicles (between 4,535 kg and 27,215 kg) and heavy-duty vehicles (from 4,535 kg to 27,215 kg) [United States Environmental Protection Agency, 2016c]. The vehicles should meet the Tier 3 requirements under three tests, namely the FTP-75 test procedure which is composed of a cold start transient phase, a stabilized phase and a hot soak phase, the US06 procedure which aims at representing more aggressive high-speed driving and the SC03 which aims at representing the use of air conditioning. Finally the Highway Fuel Economy Test Cycle (HWFET) is used to determine the highway fuel efficiency. The city driving fuel efficiency is computed directly from the FTP-75 [United States Environmental Protection Agency, 2016b]. Since 2006 manufacturers must ensure that their vehicles meet the legal requirements in terms of emissions for a useful life period, which depends on the vehicle category. In addition, The manufacturers have to reached a $\text{NMOG} + \text{NO}_x$ fleet average of 30 mg/mi by 2025. The fleet average limit includes all the vehicles categories (contrary to Tier 2 which allowed previously less stringent limits for heavy-duty vehicles). There are no fleet average limitations for PM. Finally the Tier 3 regulation limits the amount of sulphur into gasoline fuel to 10 ppm [United States Environmental Protection Agency, 2016b].

Table 2.1: US Environmental Protection Agency - Tier 3 FTP emission standards for LDVS, LDTS and MDPVS

	NMOG ¹ + NO _x [mg/mi]	PM [mg/mi]	CO [g/mi]	HCHO ² [mg/mi]
Bin 160	160	3	4.2	4
Bin 125	125	3	2.1	4
Bin 70	70	3	1.7	4
Bin 50	50	3	1.7	4
Bin 30	30	3	1.0	4
Bin 20	20	3	1.0	4
Bin 0	0	3	0	0

¹ NMOG refers to Non-Methane Organic Gases

² HCHO refers to formaldehyde, a human carcinogen

2.2.3.2 European emission standards

Vehicle emissions limits in the EU are regulated by the Euro emission standards as defined in the EU framework for the type approval of passenger vehicles, light-duty and heavy-duty vehicles (Directive 2007/46/EC) and in the Regulation 715/2007/EC that specifies limit values. Legislation defining the emissions limit is approved by the European Parliament and the European Council. The implementation of the legislation defining the testing procedures is reviewed by the technical Committee for Motor Vehicles (TCMV) where all Members states are represented. Type-approval of a vehicle can be conducted by any national technical service in accordance with the EU legislation and the manufacturers are free to choose in which country they want to perform the type-approval of their vehicles [Caudet and Pietila, 2015b]. Once approved in one Member State, a vehicle can be registered and sold throughout the EU [Caudet and Pietila, 2015b].

The emission limits in the EU are defined as Euro standards, Euro 6 being the most recent (phased-in from September 2014 to September 2021). Following the increase of CO concentrations to health-threatening levels in urban areas in the 1950s, the first European regulations regarding emissions from passenger vehicles was introduced in 1970 and 1975 (ECE-R 15-00 and 15-01) [Wurzel and Wurzel, 2002]. NO_x was first regulated in 1977 (ECE-R 15-02). These limit values were gradually strengthened (Directives 78/665/EEC, 83/351/EEC and 88/76/EEC) and in 1988 the first limit value for particulate matter emissions was introduced. Since 1988,

more stringent regulations, the Euro standards, were enacted. The European Regulations divides the vehicles into five categories: passengers vehicles (up to 5,000 kg), light commercial vehicles I (lower than 1,305 kg), light commercial vehicles II (between 1,305 kg and 1,760 kg), light commercial vehicles III (between 1,760 kg and 3,500 kg) and heavy-duty vehicles (above 3,500 kg). The current Euro standards define limits for carbon monoxide (CO), non-methane hydrocarbon (NMHC), hydrocarbons and oxides of nitrogen (THC + NO_x), Particulate Matter (PM), Particulate Number (PN) and oxides of nitrogen (NO_x). Contrary to the USA legislation, petrol and diesel vehicles are subject to different limit values for some pollutants. The vehicles should meet the Euro 6 requirements under two tests, namely the laboratory-based Worldwide harmonised Light vehicle Procedure (WLTP) and the real-world Real Driving Emission (RDE) Procedure. The WLTP procedure, and its corresponding test cycle (WLTC), replaced the New European Driving Cycle (NEDC) procedure in September 2018. The NEDC was effective since 2000 and included urban and extra-urban segments, with a cold start but was heavily criticized for not representing real-world emissions especially for NO_x [Weiss et al., 2011, Degraeuwe and Weiss, 2017] leading to its abandonment. Although capturing better emissions than NEDC, the WLTP was demonstrated to lead to significantly lower emissions and fuel consumption than during real-driving events [May et al., 2014]. To overcome this issue, RDE testing was introduced via amendments between March 2016 and June 2017 with a phase-in for all vehicles by September 2019. The RDE test is performed with a Portable Emissions Monitoring System (PEMS) during real vehicle operation on road. The route must include three segments of at least 16 km: urban (< 60 km/h), rural (60-90 km/h) and motorway (> 90 km/h). RDE emission limits are defined relatively to Euro 6 emission limits via a conformity factor. The manufacturers must therefore ensure since September 2017 that the differences in emissions between laboratory-testing and RDE-testing is lower than 110% for new models (conformity factor of 2.1). The conformity factor has been lowered to 1.5 (50%) for new vehicles in September 2020 [Caudet and Pietila, 2015a].

European emission standards for heavy-duty vehicles are commonly referred to as Euro I to Euro VI. While the Euro V and earlier emission standards only applied to vehicles presenting a maximum laden mass above 3,500 kg and equipped with either compression ignition engines or

Table 2.2: European emissions standards for Light-Duty diesel vehicles

	PM [g/km]	PN [/km]	CO [g/km]	HC + NO _x [g/km]	NO _x [g/km]
Euro 1 (1993)	0.14	-	-	0.97	-
Euro 2 (1997)	0.08	-	1	0.70	-
Euro 3 (2001)	0.05	-	0.64	0.56	0.50
Euro 4 (2006)	0.025	-	0.50	0.30	0.25
Euro 5 (2011)	0.005	$6.0 \cdot 10^{11}$	0.50	0.23	0.18
Euro 6 (2015)	0.005	$6.0 \cdot 10^{11}$	0.50	0.17	0.08

positive ignition natural gas, the Euro VI standards apply to all vehicles with a reference mass exceeding 2,610 kg regardless of the engine type. Since the introduction of the Euro III standards, heavy-duty vehicles have been required to meet legal requirements both in steady-state and transient modes. Similarly to light-duty vehicles, the regulatory emission test cycles have been changed several times. Since the Euro VI standards, the World Harmonized Stationary Cycle (WHSC) and the World Harmonized Transient Cycle (WHTC) replaced the European Stationary Cycle (ESC) and the European Transient Cycle (ETC). WHSC and WHTC engine operating points were shifted toward lower engine speed and load compared to ESC and ETC, to better represent real-driving conditions [Myung et al., 2013]. WHSC is a hot start ramped steady-state test cycle with sequences of specific speed and torque at each mode and ramps between the modes. WHTC has both cold and hot start requirements, with several monitoring segments, with a total duration of 30 minutes. In addition to the aforementioned laboratory-based tests, the Euro VI regulation introduced off-cycle and in-conformity testing. Off-cycle testing consists of two different procedures: the US EPA Not-To-Exceed (NTE) engine testing over the WNTC cycle and an RDE procedure. Contrary to the classical test cycles, the WNTC procedure does not consist of a specific driving cycle of any specific length (time or distance) but rather involves testing the performance of a vehicle at all engine speeds and loads expected to be encountered during normal vehicle operation (also called the NTE zone). Emissions are averaged over a minimum time of thirty seconds and then compared to the applicable NTE emission limits.

Table 2.3: European emissions standards for Heavy-Duty diesel vehicles at steady-state

	PM [g/kWh]	PN [/kWh]	CO [g/kWh]	HC [g/kWh]	NO _x [g/kWh]
Euro I (1993)	0.612	-	4.5	1.1	8.0
Euro II (1997)	0.25	-	4.0	1.1	7.0
Euro III (2001)	0.10	-	2.1	0.66	5.0
Euro IV (2006)	0.02	-	1.5	0.46	3.5
Euro V (2011)	0.02	-	1.5	0.23	2.0
Euro VI (2015)	0.01	$8.0 \cdot 10^{11}$	1.5	0.13	0.40

Table 2.4: European emissions standards for Heavy-Duty diesel vehicles in transient mode

	PM [g/kWh]	PN [/kWh]	CO [g/kWh]	NMHC [g/kWh]	NO _x [g/kWh]
Euro III (2001)	0.16	-	5.45	0.78	5.0
Euro IV (2006)	0.03	-	4.0	0.55	3.5
Euro V (2011)	0.03	-	4.0	0.55	2.0
Euro VI (2015)	0.01	$6.0 \cdot 10^{11}$	4.0	0.16	0.46

2.2.3.3 CO₂ fleet target

Unlike exhaust air pollutants regulations, which apply to individual vehicle models, the CO₂ legislation sets a fleet target per manufacturer. The current European legislation was enacted in 2009 (EC 443/2009). It succeeded to the voluntary agreement between the European Commission and the manufacturers associations of 1998, which aimed to cut down the CO₂ emissions to 140 g/km by 2008-2009. The voluntary agreement resulted in an insufficient progress, with an average fleet CO₂ emissions of 158 g/km in 2007 [Fontaras and Samaras, 2010, Thiel et al., 2014]. The current CO₂ passenger vehicle legislation sets that the volume weighted of new vehicle sales CO₂ emissions should not exceed 130 g/km in 2015 and 95 g/km in 2021. In order to account for the different vehicle size, the EU legislation has defined a limit curve as a function of the mass of the vehicle. As a consequence, each manufacturer has a slightly different CO₂ emissions target, depending on their respective lines of vehicles. If the average CO₂ emissions of a manufacturer's fleet exceeds its target, the manufacturer has to pay an excess emission premium for each vehicle registered (95 euros per grams of excess CO₂ per km since 2019) [European Commission, 2017].

2.2.3.4 Discrepancies between type-approval and real-world driving emissions

Despite the introduction of more stringent regulations through the successive Euro standards, NO₂ concentrations in European cities have failed to decrease as expected. In 2015, the Volkswagen emission scandal highlighted the large discrepancies in emissions, particularly of NO_x and CO₂, measured under laboratory type approval versus under real-world driving conditions. The scandal triggered the development of an abundant research on the real-world variability in emissions.

[Baldino et al., 2017] compared official laboratory tests and on-road NO_x emissions of 145 vehicles and showed that the gap between NO_x measurements in laboratory tests and real-world driving conditions had been continuously increasing: for Euro 3 vehicles, the difference was only of a factor of 2, for Euro 5 vehicles the difference increased to a factor of 3-4 while the discrepancy for the first Euro 6 vehicles, introduced between 2012 and 2015, reached a factor of 3 to 7 higher than the legal limit [Franco et al., 2014, O’Driscoll et al., 2016, O’Driscoll et al., 2018]. Three main reasons were found to explain such a discrepancies: 1) the use of defeat devices, which are inactivated by the on-board computer on the road, and this, even if vehicles are driven under certification-like conditions [Degraeuwe and Weiss, 2017], 2) the so-called “thermal windows”, which reduces the efficiency of the emissions aftertreatment systems under driving conditions which are not covered by type-approval [Degraeuwe and Weiss, 2017] and 3) the unrealistic low dynamics and narrow temperature ranges of the NEDC test [Degraeuwe and Weiss, 2017].

Large discrepancies were also found for CO₂ emissions: [Fontaras et al., 2017b] shows that although official-laboratory measurements indicate a gradual decrease in fuel consumption and CO₂ emissions of the European fleet, the gap between type-approval and real-world driving conditions has been continuously increasing with time, with real-world CO₂ emissions 12% higher than type-approval in 2005 to real-world CO₂ emissions almost 45% higher than type-approval in 2015. A study commissioned by the French ministry of the environment [CERAM, 2016] even showed that the vast majority of the reported CO₂ emissions cannot be reproduced under laboratory test conditions and that a reproduction of the certification test produces CO₂

emissions that are on average 15% higher than the reported value. The European Joint Research Group confirmed these results in a report released in 2017 [Pavlovic et al., 2017]. In addition of an experimental study of the discrepancies between type-approval and real-world driving fuel consumption (mean and median gaps of 45% and 35% compared to type-approval respectively), they also demonstrate that fuel consumption and CO₂ emissions from non-compliant RDE trips are systematically higher than for compliant RDE trips. The RDE framework will therefore only provide only a conservative estimate of the average fuel consumption and CO₂ emissions of a vehicle.

Similar conclusions were driven from studies conducted in India [Lairenlakpam et al., 2018, Mahesh et al., 2018], China [Hu et al., 2012, Zhang et al., 2014] and Korea [Cha et al., 2019].

In Europe, the replacement of the NEDC type approval procedure by the Worldwide harmonised Light-duty Test Procedure (WLTP) and the additional Real World Driving (RDE) emission type approval is expected to help reducing the aforementioned discrepancies. [Fontaras et al., 2017a] demonstrates that the WLTP procedure allows to reduce the gap between CO₂ type-approval and real-world driving emissions to 10-15% for three vehicles and it is commonly admitted that the WLTP procedure better represents real-world driving conditions than the NEDC [Tsokolis et al., 2016, Tsiakmakis et al., 2017, Pavlovic et al., 2018]. RDE type approval is also expected to bring major improvements in the emission performance of passenger and light-duty vehicles [Demuynck et al., 2019] but the RDE boundary conditions in terms of temperature ranges, driving style, frequency and distribution of accelerations and trip elevation gain might be too restrictive for the conditions typically happening on European roads [Triantafyllopoulos et al., 2019, Varella et al., 2019]. If the emission aftertreatment technology is specifically calibrated to achieve low emissions within the RDE boundaries, elevated emissions might occur in situations that are outside the RDE boundaries. [Triantafyllopoulos et al., 2019] show for three diesel vehicles that driving the vehicle dynamically, at conditions outside the RDE boundaries results in an increase in NO_x emissions by up to 1300%. They underline that this could be a critical weakness for populations living on hilly areas or close to specific elements affecting the traffic such as highway entrance ramps or traffic lights.

2.2.4 Modelling of emissions

Accurate estimation of traffic emissions is crucial to assess the role of road transport in air quality issues and its related health impacts. Estimated emissions are used to predict background and roadside concentrations of pollutants and subsequently to evaluate the impacts of potential mitigation strategies. Emissions models are becoming more complex and detailed with time, reflecting the change in vehicles types, fuel types or the number of pollutants considered. In the 1980s, only a few pollutants for hot-running conditions petrol vehicles were estimated, based on the certification data [Smit et al., 2010] while current emissions models can typically account (up to a certain extend and depending on the scale of the model) for fuel type, engine size, driving style, congestion, road grade and cold start. The size of the emission test database used to develop emission factors has also considerably increased, from less than 100 for pre-1990 vehicles to more than thousands of vehicles nowadays [Smit et al., 2010]. All emissions models account for the various factors affecting emissions but the way in which they do it differ significantly from one model to the other. [Smit et al., 2010] proposed a classification framework of the different emission models, based on: 1. the way the driving behaviour is incorporated into the model and 2. the model complexity depending on the amount of input data required. They classify emission models into 5 categories: 1. Average-speed models, 2. Traffic-situation models, 3. Traffic-variable models, 4. Cycle-variable models and 5. Modal models. In this thesis, we stick to these commonly used denominations, to which we add a sixth category, the black box models. For the sake of clarity, we provide a brief description of each of these emission models hereafter.

2.2.4.1 Average-speed models

Average-speed models are macroscopic, aggregated models that allow to derive emission factors as a function of the mean travelling speed only [Smit et al., 2010]. In the European Union, the average-speed model COPERT (COmputer Program to calculate Emissions from Road Transport), is officially used by councils and local authorities in 22 out of the 27 countries [Ntziachristos et al., 2009]. The COPERT model, developed on a large sample of vehicles tested

on a chassis-dynamometer over several speed profiles, directly relates the mean emission level of each individual vehicle to its mean velocity [Ntziachristos et al., 2009]. It distinguishes between several vehicle categories (passenger cars, light-duty vehicles, heavy-duty vehicles, buses and motorcycles) and inside each category subdivide again the model depending on the Euro standard, engine size and fuel type [Ntziachristos et al., 2009]. Its use is limited to trips longer than 300 meters [O’Driscoll et al., 2016]. Comparison with real-world data shows significant deviation with reality [Kousoulidou et al., 2010, O’Driscoll et al., 2016]. [Kousoulidou et al., 2010] shows that for a Euro 5 diesel vehicle the discrepancies between the PEMS measurements and the COPERT4 model are comprised between 225 and 231% for CO, 61 et 80% for NO_x, -32 and -36% for HC and -4 to -7% for FC. Interestingly, they also show that for NO_x and CO, the discrepancies observed seemed to be directly correlated with resolution considered, which was confirmed later by [Papadopoulos et al., 2018]. A larger study conducted by [O’Driscoll et al., 2016] on 39 diesel vehicles show that PEMS measurements over a trip composed of both urban and motorway segments are on average 1.6 times higher for NO_x and 2.5 times higher for NO₂ than COPERT4.11 predictions. Average-speed emissions models are attractive by their simplicity to use and their low computational cost. However, they are unable to account directly for a significant number of parameters known to have a direct influence on emissions such as the level of congestion, driving behaviour or external temperature. They also do not provide a detailed spatial resolution of emissions, which limits the resolution of dispersion models that estimate the effect of emissions on air pollutants concentrations.

2.2.4.2 Traffic-situation models

Traffic-situation models allow to derive emissions factors as a function of the traffic conditions (e.g. stop-and-go driving, free-flow-motorway driving) [Smit et al., 2010]. In the European Union, the traffic-situation model HBEFA (HandBook of Emissions FActors for road transport), a database providing emissions factors for all existing road vehicles, is used in a wide range of applications including national inventory reports in Switzerland, Germany, Austria and Sweden. It distinguishes between several vehicle categories and vehicle parameters such as the type of engine, the Euro standard, the aftertreatment system or the vehicle mass and provides emissions

factors for various road gradient and traffic conditions [Kühlwein and Friedrich, 2000]. HBEFA is regularly updated to better cover the evolution in vehicle fleet through time (last update in November 2019). It is typically applied at the city or regional level [Fontaras et al., 2014]. HBEFA simplicity makes it appealing, despite a limited accuracy and a high sensitivity to the aggregation level chosen. [Tsanakas et al., 2017] show that HBEFA emissions factors are heavily influenced by the aggregation period chosen and that the long analysis periods typically used in static models can result in an underestimation between 12% and 24% for NO_x, CO₂, PM, CO and HC for an aggregation period of 3 hours compared to 15 minutes. The HBEFA model assumes a mileage of 50'000 km, which is not representative for a real fleet. The ARTEMIS (Assessment and Reliability of Transport Emission Models and Inventory Systems), developed between 2000 and 2003 as a part of a large European research program, is also an example of a traffic-situation model [Boulter and McCrae, 2007]. The current HBEFA model integrates the ARTEMIS model, which is in practice not used anymore as a standalone [Andre et al., 2014].

2.2.4.3 Traffic-variable models

Traffic-variable models allow to derive emission factors as a function of traffic-flow variables (e.g. average speed, queue length, traffic density, signal settings) [Smit et al., 2010]. Although emission models belonging to this category are currently not used in research, it is worth mentioning as an example the Traffic Emissions and Energy (TEE) model. TEE is an adjusted-average speed model that reconstructs the speed profile along a road by considering the traffic density and the traffic management system [Negrenti, 1999]. The behaviour is divided into two different modes: 1. the free-flow far from intersections mode and 2. the intersections mode. The free-flow mode depends on the traffic density, which determines driving dynamics (e.g. accelerations, decelerations). The intersection mode is mostly a function of the traffic management systems regulating the intersection. The emissions are finally estimated from the traffic density, the average speed and the length of the green traffic light stage. TEE explicitly accounts for congestion [Smit et al., 2008], but trip parameters are specific to the time and traffic conditions, limiting greatly the possibility to generalise and the model was not updated for Euro 4 and above standards [Saharidis and Konstantzos, 2018].

2.2.4.4 Cycle-variable models

Cycle-variable models allow the derivation of emission factors as a function of various driving cycle variables (e.g. idle time, average speed, positive kinetic energy) at high resolution (seconds to minutes) [Smit et al., 2010]. Example of these models include the meso-scale US EPA Mobile Emissions Assessment System for Urban and Regional Evaluation (MEASURE) model and the dutch VERSIT+ model.

MEASURE estimates emissions using a multivariate regression model accounting for the vehicle operating conditions (idle, cruise, acceleration, deceleration), the vehicle characteristics (engine size, model year, etc.) and road specificities (road grade, traffic flow, etc.) [Bachman et al., 2000]. The spatial resolution of the output is the grid cell. Compared to the average-speed model MEASURE, [Fomunung et al., 2001] reports improvement on average of 1.6 for CO, 1.8 for HC and 1.4 for NO_x. MOBILE does not account for any vehicle or fuel standards promulgated after 2004 or the very large amounts of new emission data that were later incorporated in the multi-scale modal model MOVES [United States Environmental Protection Agency, 2016a] (discussed in the next section).

VERSIT+ is a multivariate regression model able to predict emission factors for 246 vehicle classes, which are defined in terms of vehicle fuel type, vehicle type, Euro standard, fuel injection technology, aftertreatment system and type of transmission [Smit et al., 2007]. The initial model was based on more than 12,000 chassis-dynamometer tests over 126 cycles covering a wide range of speed-time profiles [Smit et al., 2007] and is regularly updated. The validation of the model provided by [Smit et al., 2007] reported satisfactory correlation coefficients between measured and predicted emission factors: 0.91-0.97 for NO_x emission factors for diesel vehicles and biases of $\pm 15\%$. Large discrepancies were reported by [Nowak and Pielecha, 2017]: VERSIT+ was found to overestimate NO_x emissions by more than a factor of 2, HC by 17%, CO by 133% and CO₂ by 17% for one Petrol Euro 4 passenger vehicle driven in RDE conditions. One main limitation of VERSIT+ is the requirement for a detailed driving pattern (e.g. average speed, relative positive acceleration, number of stops per km) that might not be always available [Saharidis and Konstantzos, 2018].

2.2.4.5 Modal models

Modal models allow to derive emission factors as a function of the engine or vehicle operating conditions (acceleration, deceleration, idling), at high resolution (seconds) [Smit et al., 2010]. Examples of these models include the US EPA multi-scale MOtor Vehicle Emissions Simulator (MOVES), the Passenger Car and Heavy Duty Emission Model (PHEM) developed by TU Graz and the Comprehensive Modal Emission Model (CMEM) developed by the University of California Riverside.

Modal models mentioned above are based on the concept of operating modes, described as a function of speed and Vehicle-Specific Power (VSP). VSP, which is defined as the engine power output per vehicle unit mass and was shown to be correlated with emissions [Zhai et al., 2008], is defined as:

$$\text{VSP} = \frac{2500 + 1.08(R_0v + R_1v^2 + 0.5C_dA\rho v^3)}{1000m} + 1.08v(1.04a + gs) \quad (2.9)$$

where VSP is given in kW/t, v is the vehicle instantaneous speed in m/s, a is the vehicle instantaneous acceleration in m/s², m is the mass of the vehicle in tons, R_0 and R_1 are the road load coefficients in N and N/(m/s) respectively, C_d is the dimensionless aerodynamic drag coefficient, A the frontal surface area in m², ρ is the density of air in kg/m³, g is the acceleration due to gravity in m/s² and s is the road gradient in terms of vertical rise per horizontal distance.

MOVES provides estimates of emissions from mobile and non-mobile emissions sources using a discrete binning approach accounting for vehicle types, time periods, geographical areas, pollutants, vehicle operating characteristics, and road types, at scales ranging from a single intersection to a the entire country [US EPA, 2005]. MOVES is based on the assumption that pollutants are emitted by sources via one or more processes (also called activities). The bins that differentiate activity according to the vehicle characteristics that significantly affect emissions (e.g. fuel type, engine size, regulatory class) are labelled as “Source Bins”, while the “Operating Mode Bins” differentiate the emissions based on driving behaviour characteristics (e.g. speed and VSP) [Vallamsundar and Lin, 2011]. After distribution of the total activity

into different bins, the final emissions are computed as a unique combination of sources and operating bins. Correction factors are applied to account for factors such as the external temperatures or the use of air conditioning [Vallamsundar and Lin, 2011]. MOVES default emission rates are based on millions of chassis-dyno vehicles tests [US EPA, 2005]. Only very few comparisons with PEMS measurements were found in the literature. It is worth citing the study from [Wyatt, 2017], which reports a mean absolute percentage error of 7.9% for CO₂ emission factor for 48 tests for a Euro 4 petrol passenger vehicle. It must be noticed that the MOVES emission model was designed specifically to represent the U.S fleet and emission standards and that application of the model to a European context represents a considerable challenge [United States Environmental Protection Agency, 2019].

PHEM estimates the instantaneous engine power and speed from the vehicle speed and the road grade and in a second step interpolates tailpipe emissions from engine emission maps [Hausberger, 2003]. The model accounts for the driving resistance, transmission losses and gear shift and applies corrections for cold start and transient regimes. Initially developed only using chassis-dyno data, for which the engine torque is given, PHEM also includes PEMS data since 2017 [Matzer et al., 2017]. Since torque and engine speed are usually not provided by the PEMS instrument, [Matzer et al., 2017] proposes a methodology to derive engine power from the measured CO₂ mass flow and engine speed from generic engine efficiency maps. Compared to road tunnel experiment, an early version of PHEM presented discrepancies of less than 0.1 g/km for NO_x emission factors [Hausberger et al., 2003]. Validation of the new methodology to derive engine power and speed by [Matzer et al., 2017] show larger discrepancies with real-world measurements: -8% to 1% for the CO₂ emission factor, 3 to 25% for the NO_x emission factor and -16% to 33% for the CO emission factor for four diesel Euro 5 and Euro 6 vehicles.

CMEM estimates instantaneous emissions based on a parametrised analytical representation depending on power-demand and vehicle physical parameters [Scora and Barth, 2006]. Prediction of emissions is broken down into different modules that correspond to the different physical phenomena associated with the production of emissions and an analytical representation is provided for each module as a function of various parameters representing the characteristics of the process [Scora and Barth, 2006]. The model distinguishes between 28 light duty and 3

heavy duty vehicle categories, defined by the type of fuel, the engine size and vehicle weight, the accumulated mileage, the power-to-height ratio, emission standard and after-treatment technologies. Although former validation of the model with respect to bag data on the FTC, MEC and US06 driving cycle presented correlations larger than 0.71 for NO_x, CO, CO₂ and HC for 26 vehicles [Scora and Barth, 2006], it should be noticed that the very high degree of parametrisation significantly complicates the modelling process and can lead to accumulation of errors [Saharidis and Konstantzos, 2018]. Similarly to MOVES, the CMEM model has been developed in the U.S, where the fleet composition greatly differs from Europe, limiting its application elsewhere.

2.2.4.6 Black box models

Black box data-driven models have become popular in recent years due to the increasing complexity. characterising real-world driving emissions and the availability of larger real-world driving datasets. The formation of emissions is a highly complex non-linear problem that depends on both local environmental variables and engine specificities [Alonso et al., 2007], which traditional physical, chemical, empirical and semi-empirical models often fail to accurately capture [Panis et al., 2006]. Artificial Neural Networks (ANNs), which have shown to be effective in solving a wide range of non-linear problems, appear to be very appealing to model vehicle emissions. The published literature on the topic stays limited to superficial proofs of concept: while initial works focused on the potential of ANNs to predict fuel consumption [Thompson et al., 2000, Parlak et al., 2006] or particulate matter [de Lucas et al., 2001], recent publications also consider NO_x, CO and HC alongside with CO₂ and PM [Traver et al., 1999, Kesgin, 2004, Alonso et al., 2007, Johri and Filipi, 2014, Sun et al., 2020]. A drawback of these studies was that only one to four vehicles, tested on a chassis-dynamometer for standardised type-approval tests (US06, NEDC), were used to perform the network training and validation. The models show reasonable performance but are specific to a small set of vehicles operating in specific conditions and are extremely unlikely to generalise to real-world driving conditions. Finally, a non-negligible number of published studies did not understand well the theoretical concepts behind ANNs and consequently present theoretical mistakes (e.g. [Obodeh

and Ajuwa, 2009, Jaikumar et al., 2017]).

2.3 Pollutant dispersion

The dispersion of pollutants in a surrounding medium is a result of transport, mixing and chemical reactions induced by the flow field over a wide range of horizontal length scales. The upward dispersion is controlled by the atmospheric stability (temperature stratification) that directly influences the intensity of thermal turbulence (buoyancy) and the mixing height (depth of the atmospheric mixed layer) [Oke, 1978]. The horizontal dispersion is directly related to the wind field, which governs the distance of downwind transport and, jointly with the surface roughness (effect of surface texture on the wind field), determines the intensity of mechanical turbulence [Oke, 1978]. During the dispersion process, pollutants can be subjected to various physical or chemical transformation, depending on meteorological characteristics such as the intensity of solar radiation, the air temperature or the presence of other pollutants [Oke, 1978]. Finally, pollutants can be removed from the atmosphere by gravitational settling, wet or dry deposition [Oke, 1978].

2.3.1 Dispersion in urban environments

The presence of obstacles such as buildings, roadway or trees, which are characteristics of a urban environment, modifies drastically the surface shape as well as the radiative, thermal and aerodynamic properties of a region. As a result of the interaction of the Planetary Boundary Layer (PBL) with these features, a Urban Boundary Layer (UBL) forms above cities. The UBL consists of three sublayers: (1) the Urban Canopy Layer (UCL) situated below the average building height, (2) the Roughness SubLayer (RSL) situated above the UCL and where the flow is dynamically influenced by the roughness elements and (3) the Inertial SubLayer (ISL) that is situated above the RSL. The flow through urban canopies is heavily influenced by various parameters including but not limited to building morphology, direction and speed of the wind, local forcing. In practice, simplified morphologies such as the idealised street canyon are used to

develop an understanding of the physics and processes within the UCL and RSL and assess the importance of these flow dynamics on urban air quality. Within a street canyon, the flow field depends directly on the geometry and in particular the aspect ratio, defined as the ratio of the canyon width (W) to the canyon height (H). [Oke, 1978] identified three typical flow regimes dependant on the aspect ratio of the canyon: skimming flow ($H/W > 0.25$), wake interference ($0.17 < H/W < 0.25$) and isolated roughness flow ($H/W < 0.17$). These flow regimes are characterised by the degree of interaction between the vortex generated behind the windward obstacle and the leeward obstacle [Baik and Kim, 1999]. In a skimming flow, the proximity of the obstacles limits the penetration of the flow to weak eddies in the gaps, almost as if the obstacles were continuous. In a wake interference flow, the obstacles wakes interfere with each other, causing an enhanced turbulence. In an isolated roughness flow, the wakes created by the obstacles dissipate before interfering with each other [Gordon et al., 2004]. Within the canyon, the concentration field is highly nonuniform. The canyon vortex advects traffic emissions towards the leeward side of the canyon, with the highest concentrations situated at pedestrian-level in the centre and leeward side of the canyon. Once the pollutants reach the top of the UCL, a portion escapes from the street canyon and the other portion enters back the street canyon [Xie et al., 2003]. Flux of pollutants from the street canyon to the overlying atmospheric flow is driven by the turbulent dynamics of the shear layer at the top of the canyon [Salizzoni et al., 2009].

2.3.2 Pollutant dispersion modelling

Dispersion models use mathematical equations describing the atmosphere, the dispersion, the physico-chemical processes affecting a plume of pollutant to derive the concentration field [Holmes and Morawska, 2006]. A wide range of those models exist that consider the problem of urban pollution dispersion at different levels of abstraction, with an inherent trade off between the provided resolution and accuracy against the computational costs. Operational models simplify the highly complex flow field occurring within and above the urban canopy using parametrisation and typically associate this representation with detailed data (e.g. meteorological data).

logical conditions, background concentrations) and a physico-chemical model. Computational Fluid Dynamics (CFD) models resolve the concentration field by solving the Navier-Stokes equations using finite differences and finite volumes in every direction in space. They are able to provide detailed information on the relevant flow at every point in the domain, under well-controlled conditions. While many CFD models exist, the two most popular approaches are Reynold-Averaged Navier-Stokes simulations (RANS) and Large Eddy Simulations (LES). Traditional RANS models assume a non-convective transport in turbulent flows and models the entire range of eddies length scale, which is made at the expense of a strong model dependency and a decreased reliability of the results. On the other hand, LES models directly solve the turbulence, only ignoring the smallest scales, which contain less turbulent energy than the larger scales. As a consequence, LES models are for modelling the unsteady three-dimensional flow field to the highest possible resolutions, but at the expense of a higher simulation complexity and a much higher computational time. One operational dispersion model, SIRANE, was used in this thesis and described hereafter.

2.3.2.1 SIRANE

SIRANE is an operational urban dispersion model developed by the Ecole Centrale de Lyon [Soulhac et al., 2011]. The domain is divided into two sub-domains: the urban canopy and the external atmosphere. The urban canopy is described as a series of connected boxes. The pollutant concentration in each street is assumed to be homogeneous and is computed as a mass balance between convective transfer along the street, turbulent transfer at roof level, chemical transformation occurring inside the canyon and the dispersion at intersections. The pollutant dispersion above the canopy is represented by a Gaussian puff model. The mesoscale meteorological properties (e.g. boundary layer height, surface heat flux) are modelled through parametrisations based on observed data (e.g. wind speed, wind direction, cloud cover, rainfall rate). The chemical equilibrium between NO, NO₂ and O₃ is assumed to follow the null cycle:





where J_{NO_2} is the photolysis rate for NO_2 , k_2 is the $\text{O}^\bullet + \text{O}_2$ reaction rate and k_3 is the $\text{NO} + \text{O}_3$ reaction rate. The reaction between the oxygen and the oxygen radical happens instantaneously and a photostationary steady-state is achieved when (Leighton relationship):

$$J_{\text{NO}_2}[\text{NO}_2] = k_3[\text{NO}][\text{O}_3] \quad (2.13)$$

SIRANE assumes that pollutants reach a photostationary equilibrium within each canyon over an hourly time step. Within the model, the passive dispersion of the NO_x concentration field is calculated within and above the street canyon and in a second step the resulting concentration field and the NO , NO_2 and O_3 background concentrations are used to solve the Leighton relationship. The model has been validated against real measurements in Lyon, France [Soulhac et al., 2017]. For a more detailed description of SIRANE, the interested reader is referred to [Soulhac et al., 2011].

2.4 Air quality

Domestication of fire during the Palaeolithic and the subsequent burning of wood for cooking and heating is considered as the first anthropogenic source of air pollution. In early cities and settlements, the absence of a proper sewer and garbage collection systems, coupled with the domestic use of wood, contributed to significantly deteriorate air quality [Makra and Brimblecombe, 2004]. As urbanisation progressed, air quality in urban environments decreased further, with descriptions of the darkening of Roman buildings from smoke by Horace (65 BC - 8 AD) or description urban air as “*stuffy, smoky, polluted, obscure and foggy*” by Maimonides (1135-1204 AD). Early regulations in terms of air quality can be found in the Talmud, which prohibits the implantation of tanneries or threshing floors closer than 50 cubits from a town walls. Around 4 BC, the Roman Senate introduced a law stating that “*Polluting the air is not allowed*” (*Aerem*

corrumpere non licet), and several regulations with regard to the minimal distances between polluting premises and other houses were adopted [Makra and Brimblecombe, 2004]. During the Medieval times, buildings in cities became higher, contributing to create narrow street canyon that retained smoke and vapour. In England, the earliest documented air quality incident date from 1257, when the Queen Eleanor was reportedly forced to leave Nottingham due to the stench of sea-coal smoke. In medieval London, pollution resulting from coal burning was regarded as a matter of sufficient importance to prohibit the use of powder coal in 1273 and set up a commission to investigate on the problem in 1285 [Brimblecombe and Makra, 2005]. In 1306, the lack of improvement of air quality prompted King Edward I to proclaim a ban on the use of sea-coal in London. Evidences suggest however that coal burning was being used again less than 25 years later. By the end of the 18th century, the advent of steam-powered factories resulted in a further increase of the smoke emitted. Authorities in Manchester tried to address nuisances from engines via the promulgation of Acts and the creation of police commissioners but with only little success. In London, the Smoke Nuisance Abatement Acts of 1853 did not appear to have had more success (rather surprisingly for modern people, a lot of discussion on this act focused not on the health risks of smoke but on the expense of laundry caused by the dirty air of London). During the 19th century, fogs in London became thicker, more frequent and of different colour than those in the past, a combination of both factories and domestic emissions. In December 1873, a particularly dense coal smoke fog was reported to an increase in mortality of about 700 people, a combination of accidents due to the lack of visibility and diseases directly imputable to the fog. The worst of these fogs was The Great Smog of 1952 [Bell and Davis, 2001], which catalysed the adoption of a new regulation.

2.4.1 Air quality regulations in the UK

Concerns about acute health impacts related to heavy smog motivated the parliament to ask for a stronger regulation on smokes. Although the ministry of housing was not convinced that the law should go beyond the smoke-clauses of the Public Health Act from 1936, the Clean Air Act (CAA) was presented in 1954 and adopted in 1956 [Brimblecombe, 2006]. The CAA

addressed first domestic emissions (coal burning is gradually banned and replaced by fuel) and had more than 40 clauses compared to the six of the Public Health Act of 1936 [Brimblecombe, 2006]. The CAA was extended in 1968 and 1993 and allowed local authorities to choose areas where smoke cannot be emitted, fuels only allowed if their sulphur content was less than 2% and were smoke-less [Brimblecombe, 2006]. Industrial emissions were tackled more recently through the Environmental Protection Act in 1990 [Beattie et al., 2001]. Local government adopted a Local Air Pollution Control (LAPC) schema and the ancestor of the actual Environmental Agencies (the Agency for England and Wales, the SEPA for Scotland) became responsible for supervising the more polluting processes [Beattie et al., 2001]. The Environmental Act of 1995 required the preparation of a national air quality strategy with detailed objectives for specific pollutants to be achieved by local authorities through local air quality management (LAQM) [Department of the Environment Great Britain (Transport and the Regions.), 2000]. The first Air Quality Strategy was adopted in 1997 and replaced in 2000 by the Air Quality Strategy for England, Scotland, Wales and Northern Ireland (AQSESWN) [Department of the Environment Great Britain (Transport and the Regions.), 2007]. The local governments are responsible for reaching the common goals in terms of air quality listed in the AQSESWN.

2.4.2 Air quality regulations in the EU

As a Member State of the EU (until 2021), the UK must also fulfil the European regulations regarding air quality. The EU Ambient Air Quality Directive (2008/50/EC) sets legally binding Air Quality Limit Values limiting the concentrations of SO₂, NO₂, PM₁₀, PM_{2.5}, Lead, Benzene and CO in the ambient air (see Table 2.5). The directive also requires Member States to monitor air quality inside their territories and implement air quality plans to improve air quality where the standards are not met and maintain it where the air quality is good.

In addition to the Ambient Air Quality Directive, the National Emissions Ceiling Directive (2016/2284/EU) sets national emission reduction commitments for Member States for NO_x, non-methane volatile organic compounds, SO₂, NH₃ and PM_{2.5}.

The UK is so far failing to comply with the hourly mean NO₂ limit.

Table 2.5: Ambient air quality directive limit values for a selected set of air pollutants

Pollutant	Concentration	Averaging period	Permitted exceedences per year
CO	10 mg/m ³	Max. daily	n/a
NO ₂	200 µg/m ³	1 hour	18
SO ₂	350 µg/m ³	1 hour	24
PM _{2.5}	25 µg/m ³	1 year	n/a
PM ₁₀	50 µg/m ³	24 hours	35
	40 µg/m ³	1 year	n/a

2.4.3 Pollutants impacts on health, ecosystems and economies

Ambient air pollution is associated with numerous adverse health and environmental impacts. The effect of poor air quality on public health is estimated to be the largest single environmental health risk by the WHO causing 4.2 millions death annually (7.6% of all death in 2016), with more than 90% of the global population living in areas that exceed WHO guidelines limits [World Health Organisation, 2016]. In Europe, air quality remains poor in many areas, with concerns specifically related to concentrations of Particulate Matter (PM), NO₂ and ground-level O₃ [EEA, 2019]. The cost of air pollution to European economies was estimated to 1.6 US\$ trillion in 2014, a result of premature death, loss in productivity and health care costs [Salvi and Brickman, 2015]. Ambient PM concentrations have consistently shown the highest correlation to human mortality and morbidity. PM penetrates deeply into the human respiratory system and is causally linked to a wide range of health impacts, including, but not limited to, reduced heart variability, raised blood pressure, blood clotting and cancers [Monks et al., 2012, COMEAP, 2017, EEA, 2019]. Even at PM concentrations below the WHO guidelines, toxic effects on human health were identified [EEA, 2019]. Currently, there is a very limited understanding of which particle properties, such as their size or the presence of specific chemical substances adsorbed at the surface of the particle, are the most responsible for their adverse health effects [Giechaskiel et al., 2014]. Black Carbon (BC), a constituent of PM, is a climate short-live forcer, that contributes directly to global warming [EEA, 2019]. Other constituents of PM, such as ammonium (NH₄⁺), nitrate (NO₃⁻) and sulfate (SO₄²⁻), are correlated to the acidification and eutrophication of natural ecosystems [Monks et al., 2012]. Exposure to NO₂ received a growing interest over the past 10 years and was demonstrated to

be associated with an increased risk of cardiovascular diseases, decreased lung functionalities, aggravated asthma, low birth weights and premature deaths [Vineis et al., 2006, Andersen et al., 2012, EEA, 2019]. NO_2 and O_3 have also been linked with the degradation of natural ecosystems and the reduction in biodiversity [EEA, 2019]. Finally, CO_2 emissions play a central role in global warming and represent a rising concern world-wide [Alfieri et al., 2015, Huang et al., 2017, Klutse et al., 2018].

2.5 Mitigation strategies to air quality issues

A various number of mitigation strategies exist to control air pollution in urban areas. During peaks of pollution, city councils usually adopt strategies ranging from the ban of the most polluting vehicles and limitation of the trucks from the city center, authorisation to drive private vehicles only one day over two depending on the plate number and speed limit reductions [Zhou et al., 2010, Madireddy et al., 2011]. Outside times with pollution episodes, strategies to push people to waive their private vehicle in favour of sustainable mobility or public transportation are commonly implemented. One example is the London congestion traffic charge, which charges all vehicles entering the very center of London between 8:00 and 18:00, Monday to Friday and lead initially to a decrease in NO_x emissions of 12% and to a decrease in PM_{10} emissions of 11.9% within the congestion charging zone [Beevers and Carslaw, 2005]. Later evaluations of the congestion charge scheme efficiency showed that the initial benefits eroded with time, due to the limited size of the area affected by the charge and changes in the network [Beevers et al., 2016]. One other popular strategy consists in the implementation of a charging system aiming at penalising pollutant vehicles entering a polluted area. It could be argued that a fair charging system should differentiate between high and low emitting vehicles [Santos and Fraser, 2006] but currently the distinction is typically based on the emission standard as attested by the manufacturer during type approval, even though it has been demonstrated to not represent well emissions in real driving conditions [Weiss et al., 2012, Ligterink et al., 2013, Franco et al., 2014, O'Driscoll et al., 2016, O'Driscoll et al., 2018]. One example is the implementation in central London of the Low Emission Zone (LEZ) in 2012, aiming to tackle emissions from heavy

diesel vehicles, only based on type-approval, which resulted in undetectable changes in PM_{10} and NO_2 concentrations [Beevers et al., 2016]. The London Ultra Low Emission Zone (ULEZ), which impacts all categories of vehicles, extend over the entire Greater London area and came into effect in April 2019, resulted in a reduction of 44% of NO_2 road side concentrations [GLA, 2020]. Finally, one latest strategy consists in improving the emissions standards of buses and coaches, by replacing gradually the fleet and by retrofitting the old diesel vehicles with low- NO_2 Selective Continuous Regeneration Trap, which systems have been shown to reduce primary emissions of NO_2 by 61% and NO_x by 45% under real-driving conditions in London compared to buses only fitted with Continuous Regeneration Trap [Carslaw et al., 2015]. The technology has shown some promising results in decreasing local NO_x and NO_2 concentrations by 6% and 20% respectively during experiments conducted on Putney High Street, London [Barratt and Carslaw, 2014].

Research is currently actively on-going to develop and quantify the efficiency of new mitigation strategies. Optimisation of the traffic management system (in particular traffic lights) to reduce emissions is of increasing interest in the scientific community. Traffic lights have in fact a significant impact on both congestion and emissions as they interrupt the traffic flow and increase acceleration, deceleration and idling [Zhang et al., 2013]. As such, bad signal timing is a well known factor for congestion and increased emissions [Unal et al., 2003, Peters et al., 2009, He and Hou, 2012, Madireddy et al., 2011]. The potential of swarm optimisation and genetic algorithms was explored to find an optimised cycle program [He and Hou, 2012, Putha et al., 2012, Zhang et al., 2013, Lv et al., 2013, Olivera et al., 2015, Péres et al., 2018] but the limitations of these algorithms, in particular the fact that they do not scale well with complexity, limit their application to non-realistic scenarios. More recently, the advent of Reinforcement Learning (RL) algorithms opened new opportunities for traffic control systems optimisation. Reinforcement learning is a class of machine learning algorithms where an agent tries to maximise its perceived cumulative reward by developing a control policy through successive interactions with its environment [Genders and Razavi, 2016]. The use of RL to optimise traffic control has been investigated since the mid-1990s and is currently of increasing

interest among the scientific community. Numerous published works report very promising results [Genders and Razavi, 2016, Gao et al., 2017, Mousavi et al., 2017, Wei et al., 2018, Liang et al., 2019, Higuera et al., 2019, Chu et al., 2019]. All the aforementioned studies were purely theoretical and none of them performed a real-world evaluation of their solutions.

Chapter 3

REVS: Road transport Emissions and Variance Simulator

This chapter presents a statistically consistent analysis of 500 meters segments NO_x emissions and establishes an aggregated NO_x emission model that relies on simple inputs of vehicle average speed and VSP. Particular attention was paid to (1) develop a statistical approach to extract reliable fleet statistics from PEMS tests of variable size acquired in different conditions, (2) analyse the temporal evolution of real-world NO_x emissions, (3) determine the extent of the impact of driving dynamicity on real-world NO_x emissions, (4) develop a driving dynamicity based aggregated emission model and (5) assess the performance of the new model with respect to the data and COPERT5. The results presented in this chapter have been submitted as:

Clémence M. A. Le Cornec, Maarten van Reeuwijk, Nick Molden, and Marc E. J. Stettler (2021). REVS: Road transport Emissions and Variance Simulator. Science of the Total Environment.

The chapter has the following structure. As an introduction, Section 3.1 presents the existing research gaps and the objectives of this study. Section 3.2.1 details the data sources and any preprocessing performed prior to the study. Section 3.3 presents the statistical reasoning that supports all the results from the chapter. Section 3.4 discusses the results of the statistical

analysis, and Section 3.5 presents and assesses the performance of the developed aggregated NO_x emission model with respect the data and to other existing models. Section 3.6 presents an example of application of the REVS model. Finally, Section 3.7 summarises the main findings and discusses how the presented work could be used.

3.1 Introduction

Exceedances of the annual mean concentration limit for NO_2 has been largely attributed to the discrepancies between type approval limits and real-world driving emissions (RDE) and the fitting of cheating devices on diesel vehicles [Weiss et al., 2012, Fontaras et al., 2017b, O’Driscoll et al., 2016, O’Driscoll et al., 2018]. To mitigate this issue, the Real-World Driving Emissions (RDE) test procedure was introduced for new vehicles from September 2017 onwards [European Commission, 2018b]. The RDE test procedure uses Portable Emission Measurement System (PEMS), allowing vehicles to be tested across a wide range of operating and environmental conditions, including but not limited to different levels of congestion, driving style or meteorological conditions. Repeated tests of the same vehicle exhibit subsequently a greater level of variability than laboratory tests where these parameters are controlled, making PEMS testing more representative of real-world situations [Weiss et al., 2011]. The introduction of the RDE test procedure was therefore expected to drastically reduce the discrepancies between real-world driving and type-approval emissions [Weiss et al., 2012, Degraeuwe et al., 2016].

Improvement of type approval is therefore expected to lead to a general decrease in emissions in the long run and an improvement of air quality, particularly in urban areas. However, in large cities, where poor air quality is generally the norm, the need for rapid compliance with regulations requires the implementation of additional mitigation strategies. The challenging issue is therefore to produce accurate emissions models that, combined with atmospheric dispersion models, are able not only to assess the impacts of road transport on the local air quality but also to assess the evolution of emissions over the years and evaluate the effectiveness of a specific mitigation strategy. The development of such models requires a good understanding of

the factors affecting real-world driving emissions.

A number of studies have focused on quantifying and understanding real-world driving variability in emissions using PEMS data [Weiss et al., 2012, Franco et al., 2014, Zhang et al., 2016, Gallus et al., 2016, O’Driscoll et al., 2016, Gallus et al., 2017, Demuynck et al., 2017, O’Driscoll et al., 2018]. However, apart from the study conducted by [O’Driscoll et al., 2018], most of the sample sizes in the literature were limited to a few vehicles only, which might not be representative of the overall fleet. Similarly, the influence of driving style on vehicle exhaust emissions has been widely investigated over the past 20 years, but studies were typically conducted on a limited dataset and the definition of driving dynamicity varies from one study to the other, making them difficult to compare. An early study from [Ericsson, 2001] showed that harsh accelerations ($> 1.5 \text{ m/s}^2$) result in a significant increase in HC, NO_x and CO₂ emissions. [Luján et al., 2018, Prakash and Bodisco, 2019] confirm the results from [Ericsson, 2001] for a Euro 6 and a Euro 5 vehicles and demonstrate that NO_x and NO₂ emissions are directly linked to abrupt accelerations, which typically happen during urban driving. To the best of our knowledge, no studies have quantified the relationship between driving dynamics and emissions on a large PEMS dataset nor studied its evolution over the years.

All emissions models account for the various factors affecting emissions but the way in which they do it differ significantly from one model to the other. [Smit et al., 2010] proposed a classification framework of the different emission models, based on: (i) the way the driving behaviour is incorporated into the model and (ii) the model complexity depending on the amount of input data required. This study focuses specifically on average-speed models, which are particularly attractive for their simplicity to use and their low computational cost. COPERT, an average-speed emission model widely used for planning by local authorities in 22 of the EU-28 countries, provides an estimate of the average NO_x emission factor per vehicle category (Euro standard, fuel type and engine size). One of its main limitations is its lack of estimate of the variability in emissions around the mean, which affects dispersion processes. Moreover, the use of average speed only to estimate emissions does not allow to capture the variability induced by different driving dynamics when they result in the same average-speed. Several studies have compared real-world emissions data (PEMS and remote sensing) with predictions from

the average-speed emission model COPERT4 and concluded that COPERT4 underestimates drastically NO_x emissions [Sjödin and Jerksjö, 2008, Borge et al., 2012, O’Driscoll et al., 2016]. Similar work has not been conducted for COPERT5, which was released in 2017.

3.2 Methods

As mentioned earlier, most of the studies on real-world driving emissions were conducted on datasets with a very limited size, and are not able to draw conclusions on the evolution of real-world driving emissions following the introduction of the RDE legislation. Moreover, COPERT does not account for driving dynamics nor provides an estimate of the spread in emissions. These observations provided the motivation to: (i) conduct an analysis on the evolution of real-world driving emissions, (ii) quantify the performance of COPERT5, (iii) analyse the influence of driving behaviour on real-world emissions and (iv) present a new and simple emission model allowing to not only predict the mean emission factor of a fleet as a function of the average speed of a trip but also to provide an estimate of the spread in emissions. This section firstly gives a description of the dataset used in this study in Section 3.2.1. The methodology to process the PEMS raw data is presented in Section 3.2.2. Section 3.2.3 then describes the COPERT model. Finally, Section 3.2.4 explains how trips were extracted from the PEMS time series.

3.2.1 Dataset description

Four different datasets were used in this study. The first dataset was provided by Emissions Analytics and was already described in the study from [O’Driscoll et al., 2018]. It is hereafter referred to as EA2016. The second dataset was also provided by Emissions Analytics and had never been used in any study. It is hereafter referred to as EA2019. The third and the fourth datasets are publicly available datasets from the UK Department for Transport (DfT) Vehicle Emissions Testing Program in 2016 and the UK Driver and Vehicle Standards Agency (DVSA) Vehicle Market Surveillance Unit Program results from 2017 and 2018. These datasets are hereafter referred to as DfT and DVSA.

For the sake of clarity and brevity, we provide in the main text only a detailed description the EA-2019 dataset. A more detailed description of the EA2016, DfT and DVSA datasets can be found in Section A.1 Appendix A.

3.2.1.1 Description of the EA2019 dataset

A Portable Emissions Measurement System (PEMS) was used to measure real-world driving emissions of 24 diesel, 1 diesel hybrid, 39 petrol and 1 petrol hybrid passengers vehicles between May and December 2018 around Stuttgart (Germany) and High Wycombe (United Kingdom) using a SEMTECH-DS instruments (Sensors Inc.). The 65 tested vehicles were made by 22 different manufacturers (Audi, BMW, Citroen, Fiat, Ford, Honda, Hyundai, Kia, Land Rover, Mazda, Mercedes-Benz, Mini, Mitsubishi, Nissan, Opel/Vauxhall, Peugeot, Porsche, Renault, Skoda, Ssangyong, Suzuki, Volkswagen and Volvo). With one exception, all the vehicles were manufactured between 2017 and 2019 and certified Euro 6b, Euro 6c or Euro 6d-TEMP. These vehicles were manufactured around and after the enactment of the RDE regulation and in the remaining of the study they are referred to as post-RDE vehicles. In the rest of this article these vehicles are referred to as post-RDE vehicles. Twenty-seven of the petrol engines were Gasoline Direct Injection (GDI) while the other 13 were Port Fuel Injection (PFI). All the petrol engines were fitted with a Three Way Catalyst (TWC). The diesel vehicles had a mixture of NO_x aftertreatment technologies: 1 used EGR only (the Euro 5 vehicle), 4 used EGR + LNT, 18 used EGR + SCR and 3 were fitted with EGR + LNT + SCR. Vehicles in the dataset had relatively low start mileage with an average of 9,657 (s.d. 21,519) km. Therefore, deterioration of the emissions control system as a result of ageing, usually observed after 50,000 km [Borken-Kleefeld and Chen, 2015], is not a factor. Engine sizes ranged between 998 and 4395 cm³, with petrol vehicles presenting on average smaller engine size than the diesels: 1611 (s.d. 750) cm³ for the petrol vehicles compared to 2154 (s.d. 555) cm³ for the diesel vehicles. More information can be found in Table A.1 in Appendix A, which summarizes the technical characteristics of all the tested vehicles. The average dynamics characteristics of the routes used for testing are described in Table 3.1.

Table 3.1: Test routes dynamic characteristics EA2019.

Parameter	Stuttgart (Ge)	High Wycombe (UK)
Number of tests [-]	60	19
Duration [min]	241.9 ± 75.1	271.5 ± 86.6
Distance [km]	180.5 ± 55.7	217.0 ± 64.8
Average speed [km/h]	44.1 ± 7.1	48.3 ± 3.6
% of idle [%]	27.2 ± 11.5	20.1 ± 3.5
Relative Positive Acceleration (RPA) [m/s^2]	0.11 ± 0.02	0.11 ± 0.01
95 th percentile of the product of RPA and speed [m^2/s^3]	68.1 ± 15.8	56.0 ± 13.3
Kinetic intensity [km^{-1}]	0.44 ± 0.08	0.38 ± 0.07
RDE-compliant for dynamic boundaries	NO	NO

3.2.2 Processing of the raw data

The RDE regulation establishes a strict framework with regard to the processing of raw PEMS data for type approval [Official Journal of the European Union, 2017]. It is composed of a speed correction and a GPS altitude smoothing procedures, as well as a specific definition of acceleration and guidelines to convert species volumetric flow rate to mass flow rate. For the sake consistency and comparison, the RDE framework was used to process the PEMS raw data of the four datasets included in this study. We make the reader attentive to the fact that the tests considered in this study are not RDE-compliant and do not respect the respective share of urban-rural-motorway driving nor the maximum duration as set out by the RDE regulation. Nevertheless, they still represent a very valuable source of information.

The RDE regulation requires the application of the T4253 Hanning filter to the GPS speed signal if the acceleration resolution, defined as the minimum acceleration greater than zero, is greater than 0.01 m/s^2 . The T4253 Hanning filter is a non-linear data smoothing procedure that consists in repeating odd-span running medians. We refer the interested reader to [Velleman, 1980] for a more detailed description of the filter and its mathematical properties. This procedure was required for the EA2016 and EA2019 datasets but not for the DfT and DVSA datasets.

After the correction of the speed signal, the instantaneous acceleration can be computed as:

$$a(t) = \frac{v(t + \Delta t) - v(t - \Delta t)}{2\Delta t} \quad (3.1)$$

where $v(t)$ is the instantaneous corrected speed at time t and Δt is the sampling frequency.

In a third step, the GPS altitude is corrected for any missing data using linear interpolation. A comparison against a topographic map is then performed and any discrepancy between the measured and the real altitude larger than 40 meters is corrected by replacing the measured altitude by the real one. Topographical altitude was obtained from the GPS visualiser website that is based on the NASA Shuttle Radar Topography Mission (SRTM) elevation data. The measured and real altitude are in good agreement with each other, and only a few points per test had to be corrected. Finally, the signal is further corrected if the difference in altitude between two consecutive measurements is higher than an instantaneous speed dependent threshold. Upon the completion of the procedure, the altitude signal is considered as valid. This procedure was conducted for all the datasets, except the DVSA dataset for which neither the altitude signal nor the route coordinates are given. For the next steps of the data processing, we assumed a null elevation for all the tests of the DVSA dataset.

A detailed description of the preprocessing steps carried out for each dataset is available in Section A.2 of Appendix A.

The DfT and DVSA datasets provided species concentrations as volumetric flow rates instead of mass flow rate. Volumetric flow rates were converted to mass flow rates following the methodology, described in the [Official Journal of the European Union, 2017]:

$$\dot{m}_{\text{gas}} = u_{\text{gas}} c_{\text{gas}} q_{\text{exhaust}} \quad (3.2)$$

where $\dot{m}_{\text{gas}}$ is the mass flow rate of the gaseous component in g/s, u_{gas} is the ratio of the density of the exhaust component gas to the overall density of the exhaust (0.001586 for NO_x for diesel exhaust gas and 0.001587 for NO_x for petrol exhaust gas), c_{gas} is the measured wet concentration of the gaseous exhaust component in ppm and q_{exhaust} is the measured exhaust mass flow rate in kg/s. Similarly to the RDE legislation, we did not apply any correction for ambient humidity.

The Vehicle Specific Power (VSP) is an instantaneous measure of the vehicle power per unit of

mass. VSP allows to evaluate driving dynamics in real-time, by opposition with the RPA and $v \cdot a_{\text{pos}}$ metrics. It is a function of vehicle speed, acceleration and road grade, and is correlated with emissions [Zhai et al., 2008]. We use the same formulation as in [Davison et al., 2020]:

$$\text{VSP} = \frac{2500 + 1.08(R_0 v + R_1 v^2 + 0.5 C_d A \rho v^3)}{1000m} + 1.08v(1.04a + gs) \quad (3.3)$$

where VSP is given in kW/t, v is the vehicle instantaneous speed in m/s, a is the vehicle instantaneous acceleration in m/s^2 , m is the mass of the vehicle in tons, R_0 and R_1 are the road load coefficients in N and $\text{N}/(\text{m/s})$ respectively, C_d is the dimensionless aerodynamic drag coefficient, A the frontal surface area in m^2 , ρ is the density of air in kg/m^3 , g is the acceleration due to gravity in m/s^2 and s is the road gradient in terms of vertical rise per horizontal distance. We applied the same decision tree based vehicle segmentation as in [Davison et al., 2020] and used the coefficients R_0 , R_1 and $C_d A$ they provide. The distribution of VSP over the entire dataset is in agreement with the typical VSP distributions for London (see Figure A.2 in Section A.2.7 of Appendix A).

3.2.3 The COPERT5 emission model

COPERT is an average-speed emission model developed by the European Environment Agency and widely used in modelling studies in 22 out of the 28 European state members. The model was initially designed to produce emissions inventories (i.e. to assess the exhaust emissions from a set of vehicles or trips) for the most important pollutants emitted by road vehicles, namely exhaust PM, NO_x , CO, CH_4 and VOCs. COPERT establishes continuous speed functions to associate emission factors with travel speed by combining the results of tests with different average speed [Lejri and Leclercq, 2020]. These functions are derived for each pollutant and each vehicle class (e.g. passenger cars, light duty vehicles, buses and heavy duty vehicles) and technology (diesel or petrol, Euro 1 to Euro 6, with considerations regarding the aftertreatment system). The coefficients of the COPERT functions were fitted on driving cycles of various length, and the model assumes that the established relationship remains valid at the scale of a vehicle flow [Lejri and Leclercq, 2020]. In this study, the focus is on NO_x emissions from Euro

5 and Euro 6 diesel and petrol vehicles, for which COPERT5 derive the following expression for NO_xEF [g/km]:

$$\text{NO}_x\text{EF} = \begin{cases} (1 - RL) \frac{\alpha \bar{v}^2 + \beta \bar{v} + \gamma + \delta / \bar{v}}{\epsilon \bar{v}^2 + \zeta \bar{v} + \eta} & \text{if COPERT5 Diesel Euro 6} \\ (1 - RL) \frac{\alpha + \gamma \bar{v} + \epsilon \bar{v}^2 + \zeta / \bar{v}}{1 + \beta \bar{v} + \delta \bar{v}^2} & \text{otherwise} \end{cases} \quad (3.4)$$

where RL , α , β , γ , δ , ϵ , ζ , η are the empirical coefficients and $\bar{v}$ is the average speed of the trip in km/h.

3.2.4 Extraction of segments

In most of the analysis conducted in this study, the tests were broken down into non-overlapping segments presenting similar driving conditions (i.e. urban, rural and motorway). We distinguish between urban, rural and motorway driving based on the measured instantaneous speed. Urban segments are defined by an instantaneous speed comprised between 0 and 50 km/h, rural trips have instantaneous speed between 50 and 90 km/h and motorway segments have instantaneous speed above 90 km/h. We extracted segments of 500 meters and recorded the following quantities: average speed, average VSP, NO_x emission factor, vehicle ID, driving environment. Except where explicitly stated otherwise, the term trip always refers to these 500 meters segments. The number of trips extracted per vehicle was comprised between 7 and 126 in urban driving conditions, between 6 and 186 in rural driving conditions and between 8 and 233 in motorway driving conditions. A total of 19,496 urban trips, 14,093 rural trips and 21,094 motorway trips were extracted.

3.3 Statistical analysis of the data

The aim of the statistical analysis is to provide an estimate of the expected NO_x emission factors and its variance per vehicle category m , which will be denoted $\bar{e}_{V,m}(\vec{f})$ and $\sigma_{V,m}^2(\vec{f})$,

respectively. Here, $\vec{f} = [v, \text{VSP}]$ denotes the parameter vector, v is the average vehicle speed over the segment and VSP is the vehicle specific power.

The joint probability distribution function (PDF) $P_i(e, \vec{f})$ for a vehicle i where e is the emission factor and $\vec{f}$ the parameter vector is related to the PDFs of the individual tests $p_i(\vec{f})$ via:

$$p_i(\vec{f}) = \int P_i(e, \vec{f}) de. \quad (3.5)$$

The joint probability distribution function $P_{V,m}(e, \vec{f})$ for vehicle technology class m can be constructed by the superposition of all individual vehicle-type PDFs via:

$$P_{V,m}(e, \vec{f}) = \sum_{i \in V_m} \tilde{w}_i P_i(e, \vec{f}), \quad (3.6)$$

where V_m is the set of vehicle types belonging to vehicle category m and $\tilde{w}_i$ is the relative contribution of vehicle type i to vehicle technology category m . As weightings are region-dependent and not all vehicles types are represented in the dataset, an equal weighting $\tilde{w}_i = 1/N_m$ was given to all the vehicle types, where N_m is the total number of vehicle types in the vehicle category. With this choice, Equation (3.6) simply becomes the average of the vehicle emission PDFs in the vehicle category. Integrating (3.6) over e results in:

$$p_{V,m}(\vec{f}) = \frac{1}{N_m} \sum_{i \in V_m} p_i(\vec{f}), \quad (3.7)$$

where $p_{V,m}(\vec{f})$ is the joint PDF of $\vec{f}$ for vehicle technology class m .

In terms of $P_i(e, \vec{f})$, the expected vehicle emission factor $\bar{e}_i$ and its variance σ_i^2 are conditional probabilities of the form:

$$\bar{e}_i(\vec{f}) \equiv \frac{\int P_i(e, \vec{f}) e de}{\int P_i(e, \vec{f}) de}, \quad \sigma_i^2(\vec{f}) \equiv \frac{\int P_i(e, \vec{f}) (e - \bar{e}_i(\vec{f}))^2 de}{\int P_i(e, \vec{f}) de}. \quad (3.8)$$

Similarly, the expected value and variance for the vehicle technology class m are defined as:

$$\bar{e}_{V,m}(\vec{f}) \equiv \frac{\int P_{V,m}(e, \vec{f}) e de}{\int P_{V,m}(e, \vec{f}) de}, \quad \sigma_{V,m}^2(\vec{f}) \equiv \frac{\int P_{V,m}(e, \vec{f}) (e - \bar{e}_{V,m}(\vec{f}))^2 de}{\int P_{V,m}(e, \vec{f}) de}. \quad (3.9)$$

Substituting (3.6) into (3.9) and making use of (3.8) results in:

$$\bar{e}_{V,m}(\vec{f}) = \frac{1}{N_m} \sum_{i \in V_m} \frac{p_i(\vec{f})}{p_{V,m}(\vec{f})} \bar{e}_i(\vec{f}), \quad (3.10)$$

$$\sigma_{V,m}^2(\vec{f}) = \frac{1}{N_m} \sum_{i \in V_m} \frac{p_i(\vec{f})}{p_{V,m}(\vec{f})} \left(\sigma_i^2(\vec{f}) + (\bar{e}_i(\vec{f}) - \bar{e}_{V,m}(\vec{f}))^2 \right), \quad (3.11)$$

where V_m is the set of vehicle types in the database belonging to vehicle category m , N_m is the number of vehicle types in category m , $p_i(\vec{f})$ is the joint PDF of $\vec{f}$ for vehicle type i , and $p_{V,m}(\vec{f}) = N_m^{-1} \sum_{i \in V_m} p_i(\vec{f})$ is the joint PDF of $\vec{f}$ for vehicle technology category m .

To eliminate any potential bias due to an individual vehicle being under-sampled (i.e. not having enough trips), we empirically computed the minimum number of trips necessary to reconstruct an individual vehicle pdf. We refer the interested reader to Section A.2.4 in Appendix A for a more detailed description. We concluded that at least 10 non-overlapping trips were needed to accurately reconstruct an individual vehicle pdf. We consequently eliminated 2 vehicles out of 282, 8 vehicles out of 282 and 3 vehicles out of 282 in urban, rural and motorway driving conditions respectively. The full statistical procedure is illustrated in Figure 3.1.

Finally, we used a bootstrapping procedure to evaluate the uncertainty due to both the sampling of the individual vehicle (i.e. what if another set of trips had been extracted for each individual vehicle) and the sampling of the fleet (what if the test fleet had been constituted of another set of vehicles). We refer the interested reader to Section A.2.5 of Appendix A for a brief overview on bootstrapping.

3.4 Results

This Section analyses the PEMS measurements with two main point of focus: (i) an analysis of the evolution of the emissions with the introduction of the RDE regulation, split per driving

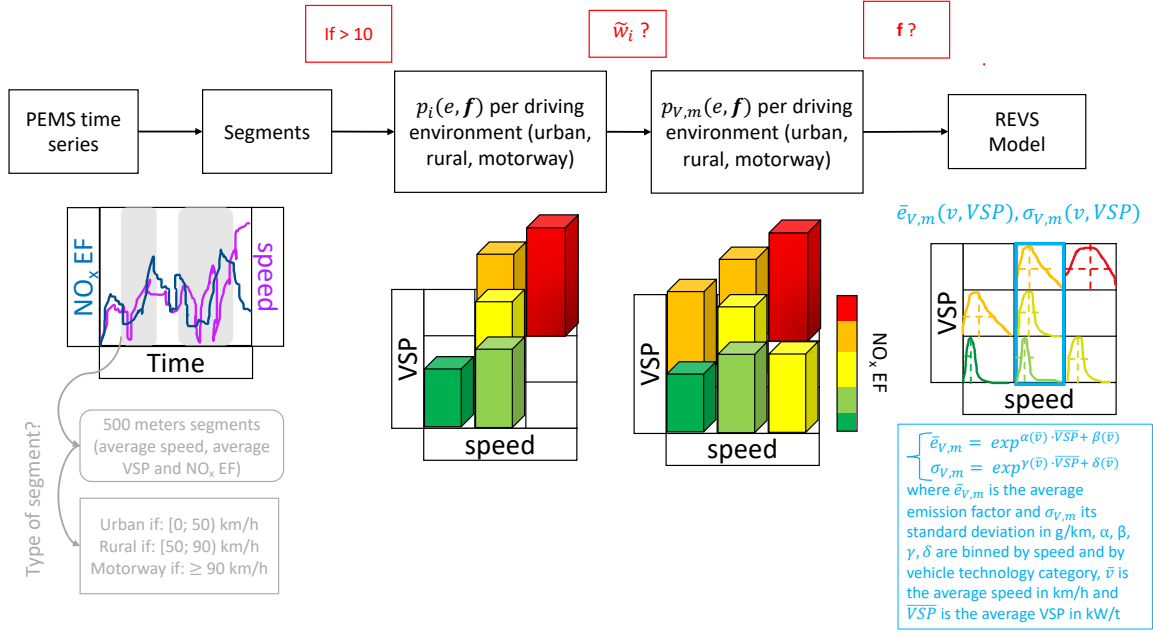


Figure 3.1: Illustration of the procedure followed to develop the model. It consists in 4 main steps: 1. the extract of the trips for every vehicle, 2. the construction of a pdf for every vehicle per driving environment, 3. the construction of the joint pdf per vehicle technology category and 4. the fitting of the model. Note: for the sake of clarity, not all the details are presented in this Figure.

environment, is carried out, (ii) the impact of average speed on NO_x emissions is discussed and (iii) the impact of driving dynamics on NO_x emissions and its evolution with the successive legislation is evaluated. Results specific to rural and motorway driving conditions are available in Appendix A and are only briefly discussed here. Except where stated otherwise, results presented in this section are specific to urban driving conditions.

3.4.1 Real-world driving NO_x emissions

Figure 3.2 displays a comparison of the cumulative distribution of the measured NO_x emission factors split per driving environment (urban, rural, motorway) and vehicle technology category (diesel or petrol, Euro 5, Euro 6 pre-RDE and Euro 6 post-RDE) integrated over $\vec{f}$ ($\int_{\vec{f}} \int_e p_{v,m} d\vec{e} d\vec{f}$, where $\vec{f} = [v, VSP]$). The dashed red lines represent the Euro standard emission limits. The average cumulative distribution is depicted by grey dashed lines, while the 95% confidence interval is depicted by grey shaded areas around the average cumulative distribution. The 95% confidence interval was calculated by conducting a 500-iteration bootstrap procedure

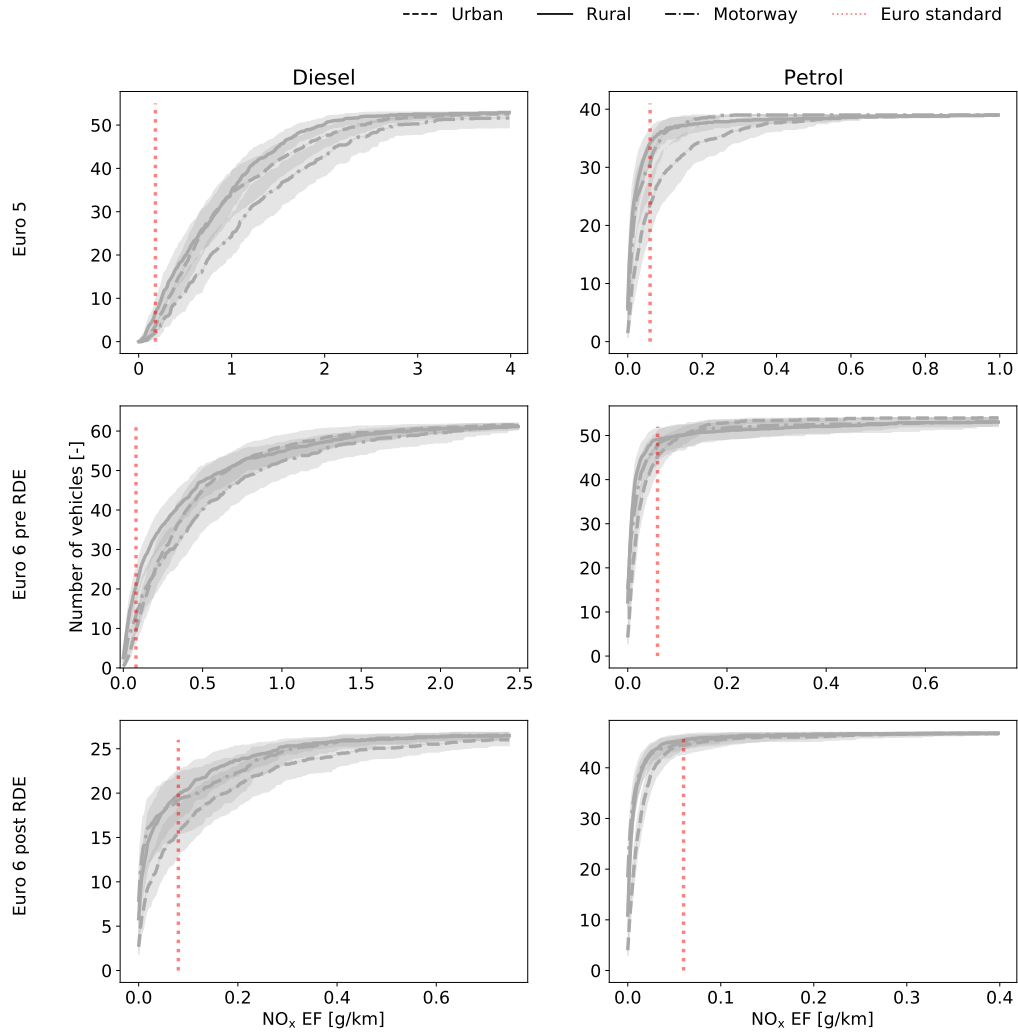


Figure 3.2: Comparison of the cumulative distributions of the NO_x emission factor split per driving environment and vehicle technology category. Diesel (respectively petrol) vehicle emissions are depicted on the left (respectively right) column, classified from top to bottom by successive Euro standard. The shaded grey areas represent the 95% confidence interval around the mean NO_x emission factor distribution, represented by dashed grey lines. The dashed red lines represent the Euro standard emission limits.

on the vehicles constituting the test fleet, and therefore quantifies the uncertainty associated to the sampling of the fleet, as explained in Section 3.2.4. The current dataset allows the direct evaluation of whether or not the prevailing type-approval regulation has been successful in ensuring low-emitting vehicles in real-world driving conditions. It can be first observed from Figure 3.2 that the mean and the spread of NO_x emission factors from diesel vehicles decrease with the successive regulations. However, this positive finding should not mask that the majority of the diesel vehicles considered in this study have local emission factors significantly higher than allowed by the European legislation. In urban driving conditions, on average only 3.4 out of the 53 Euro 5, 8.3 out of the 62 Euro 6 pre-RDE and 15.5 out of the 27 Euro 6 post-RDE diesel vehicle have real-world driving NO_x emission factors respecting the prevailing regulation at the time of their release on the market. On the other hand, the data presented in Figure 3.2 shows that a larger fraction if not the vast majority of the petrol vehicles present satisfactory average real-world NO_x emission factors. In urban driving conditions, on average 22.5 of the 39 Euro 5, 44.6 of the 54 Euro 6 pre-RDE and 44.5 of the 47 Euro 6 post-RDE petrol vehicles present real-world driving NO_x emission factor lower than the European legal limit. Finally, the 95% confidence interval of the urban, rural and motorway emissions are overlapping for all the vehicle technology categories. Therefore, it can be concluded that emitting behaviour in different driving environment cannot be statistically differentiated. The maximum difference in mean NO_x emission factor between two driving environments at 50% of the dataset does not exceed 65% (0.37 g/km) for diesel Euro 5 vehicles, 49% (0.17 g/km) for diesel Euro 6 pre-RDE vehicles, 13% (0.042 g/km) for diesel Euro 6 post-RDE vehicles, 20% (0.032 g/km) for petrol Euro 5 vehicles, 20% (0.012 g/km) for petrol Euro 6 pre-RDE vehicles and 20% (0.0064 g/km) for petrol Euro 6 post-RDE vehicles. On average though, and apart from Diesel Euro 5 and Diesel Euro 6 pre-RDE vehicles, the highest emissions are recorded in urban driving conditions. From a policy maker's perspective, this result highlights the relevance of considering urban emissions separately from rural or motorway emissions.

3.4.2 Impact of average speed and driving dynamicity on NO_x emissions

The relative impact of driving dynamicity on vehicle emissions is difficult to assess and quantify due to the important variability in the level of emissions within a vehicle technology category. To offset this issue, and allow the comparison of the relative influence of vehicle dynamicity on emissions between different vehicles belonging to the same vehicle technology category, we followed the subsequent procedure once all the segments for each vehicle are extracted: (i) for each segment, compute the mean VSP and NO_x emission factor associated with each operating mode (a definition of operating mode is available in Section A.2.6 of Appendix A) and (ii) compute the percentile corresponding to each datapoint relatively to the entire distribution for the average speed, the VSP and the NO_x emission factor. The results are then aggregated per vehicle technology category, and the correlation coefficient between the percentile of VSP and the percentile of NO_x emission factor is used as a proxy for the impact of driving dynamicity on NO_x emissions. Similarly to Section 3.3, a relative weight of 1 was attributed to each vehicle within the dataset.

3.4.2.1 Impact of average speed on NO_x emissions

Figure 3.3 displays the joint distribution of the percentile of average speed and the percentile of NO_x emission factor in urban driving conditions, split by vehicle technology category. It can be observed from Figure 3.3 that the variability in NO_x emission factor induced by the variability in average speed appears to be almost non-existent. The correlation coefficient between the percentile of average speed and the percentile of NO_x emission factor is ranging between -0.33 and -0.20, confirming the visual evidence of Figure 3.3. This result is also consistent with the study from [ODriscoll, 2017]. Overall, it provides compelling evidence that NO_x emissions are only weakly correlated with average speed and that trying to predict NO_x emissions only based on the average speed is an inaccurate oversimplification.

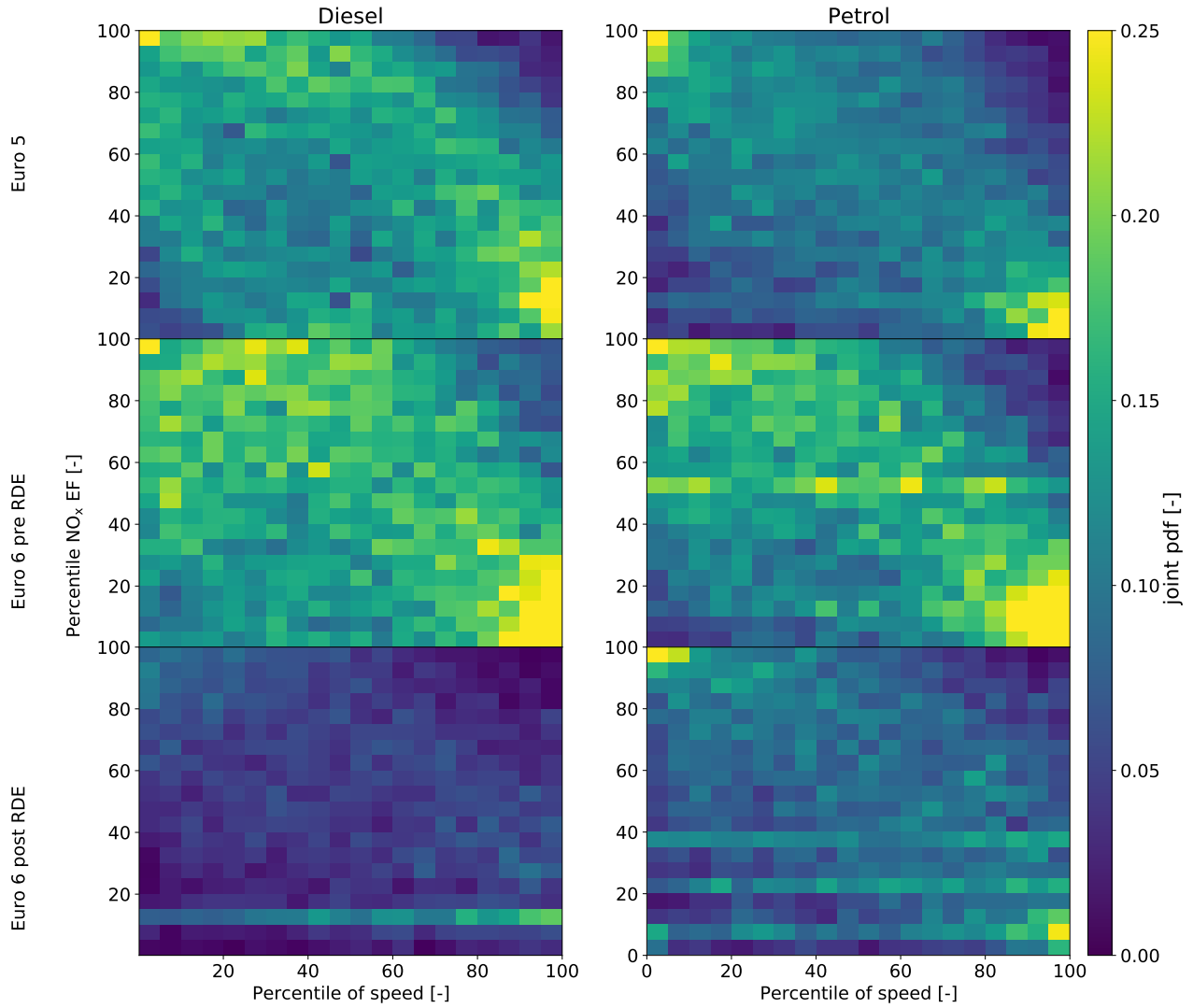


Figure 3.3: Joint distribution of the percentile of average speed and the percentile of NO_x emission factor in urban driving conditions. Results are split by vehicle technology.

3.4.2.2 Impact of driving dynamics on NO_x emissions

Figure 3.4 displays the joint distribution of the percentile of VSP and the percentile of NO_x emission factor in urban driving conditions, split by vehicle technology category. It can be observed from Figure 3.4 that the impact of driving dynamicity on emissions varies depending on the vehicle technology category. An increase in VSP is strongly correlated to an increase in emissions for diesel Euro 5 and diesel Euro 6 pre-RDE vehicle, with a correlation coefficient of 0.70 and 0.47 respectively. The correlation coefficient between the percentile of VSP and the percentile of NO_x emission factor decreases to 0.30 for diesel Euro 6 post-RDE vehicles, indicating a significant decrease of the impact of driving dynamicity on NO_x emissions. Despite being less sensitive to driving dynamicity overall, similar conclusions can be drawn for petrol vehicles: the correlation coefficient between the percentile of VSP and the percentile of NO_x emission factor is of 0.35 for Petrol Euro 5 vehicles, reaches 0.41 for Petrol Euro 6 pre-RDE and decreases to 0.23 for Petrol Euro 6 post-RDE. For this latest category, driving dynamicity appears to have a very limited if any impact on NO_x emissions. Overall, it is noteworthy that the dependency between VSP and NO_x emissions decreases with the successive regulations, indicating that the aftertreatment systems are becoming better at coping with a large variety of driving style. If this trend is confirmed, the impact driving dynamicity on NO_x emissions might become unimportant in future vehicles. Until then, we argue that driving dynamicity is a factor that should be directly accounted for in emission models.

3.5 REVS

3.5.1 Description of the REVS model

In this section, we describe a Road transport Emissions and Variance Simulator (REVS) that aims to predict the average emission factor of a fleet and its associated standard deviation. As more and more studies emphasise the importance of the real-world variability in emissions for policy-making perspective [ODriscoll, 2017], we argue that this aspect should be accounted for.

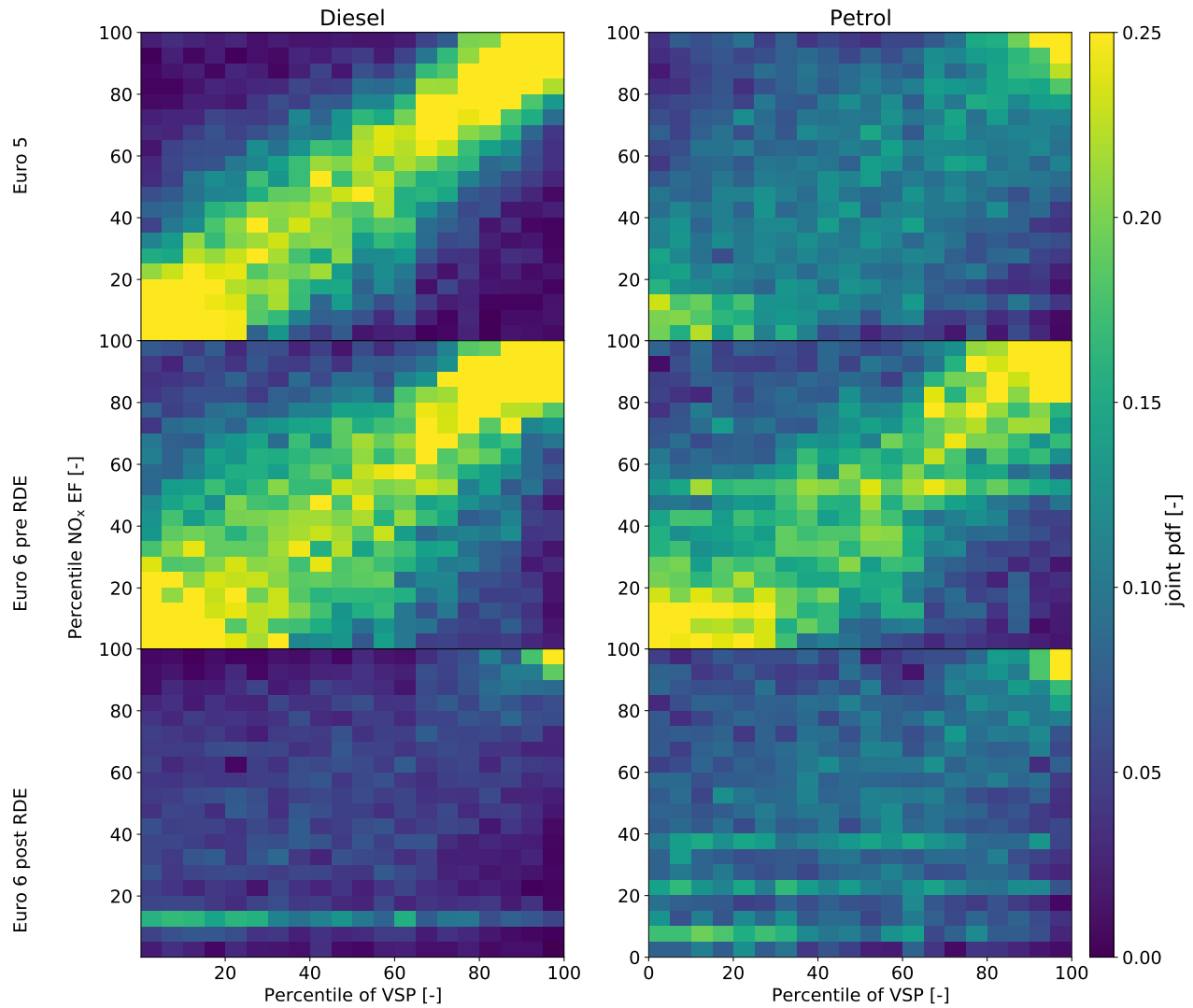


Figure 3.4: Joint distribution of the percentile of VSP and the percentile of NO_x emission factor in urban driving conditions. Results are split by vehicle technology.

Moreover, the use of average speed only to estimate emissions does not allow to capture the variability induced by different driving dynamics when they result in the same average speed. Numerous studies have highlighted the importance driving dynamics on emissions [Ericsson, 2001, Luján et al., 2018, Prakash and Bodisco, 2019] and we believe that neglecting this parameter can lead to important discrepancies between the estimated and the measured emissions. Based on our large dataset, we seek to provide a parametrisation of the NO_x emission factor as a function of speed and VSP of the form:

$$\bar{e}_{\text{REVS}}(\bar{v}, \overline{VSP}) = \exp(\alpha(\bar{v}) \cdot \overline{VSP} + \beta(\bar{v})) \quad (3.12)$$

$$\sigma_{\text{REVS}}(\bar{v}, \overline{VSP}) = \exp(\gamma(\bar{v}) \cdot \overline{VSP} + \delta(\bar{v})) \quad (3.13)$$

where $\bar{e}_{\text{REVS}}$ is the average NO_x emission factor in g/km, σ_{REVS} is the standard deviation around the average NO_x emission factor in g/km, α , β , γ and δ are the model parameters, $\bar{v}$ is the average speed in km/h and $\overline{VSP}$ is the average vehicle specific power in kW/t. We used a range of speed comprised between 10 km/h and 130 km/h, with bins of speed of 10 km/h. The range of VSP considered was comprised between -2 and 6 kW/t for urban driving conditions, between 4 and 12 kW/t for rural driving conditions and between 6 and 18 kW/t for motorway driving conditions, with bins of VSP of 2 kW/t. When the joint distribution of speed and VSP contained less than 10 trips, it was not used during the fitting procedure. We attract the reader attention that due to the way the data was collected, it was no possible to determine the coefficients for the petrol Euro 5 vehicles in rural driving conditions. In order to find the optimal parameters α , β , γ and δ , a jackknife procedure was conducted by randomly removing 20% of the data from the fitting set. This procedure was conducted 1000 times and the parameters corresponding to the maximum likelihood were chosen as the optimal parameters. All the computations were performed using the function `minimize` of the Python library `Scipy` [Virtanen et al., 2020] and the Python library `chainConsumer` [Hinton, 2016]. The REVS model implemented in this work is freely available, with examples of use [Le Cornec, 2021]. The datasets provided by Emissions Analytics cannot be disclosed and we therefore do not share the extracted trips.

Finally, equations 3.10 and 3.11 can be applied to derive the average emission factor and its standard deviation of a defined fleet, based on average speed and VSP dependent emission factors and standard deviations per vehicle technology category.

3.5.2 Emission maps for the COPERT5, the REVS model and the measured data

Figure 3.5 displays a comparison of the measured and predicted NO_x emission factor in urban driving conditions split by vehicle technology category for the two compared models. The average NO_x emission factor predicted by COPERT5 (respectively the REVS model) is depicted by a dashed red (respectively blue) line. The predicted standard deviation by the REVS model is depicted by blue shaded areas around the average. The distributions of the measured NO_x emission factor are constituted of all the extracted trips with the corresponding binning in speed and VSP. We make the reader attentive to the fact that the data presented in Figure 3.5 does not correspond to the data presented in Figure 3.4: in Figure 3.4 each vehicle is analysed separately in order to assess the relative importance of driving dynamicity on emissions while on Figure 3.5 all the vehicles within one vehicle technology category are aggregated together. Since the baseline level of emissions varies greatly within one vehicle technology category, the correlation between VSP and NO_x emission factor visually appears in agreement although visually slightly weaker than in Figure 3.4. Figures A.3 and A.4 in Appendix A display a similar comparison but for rural and motorway driving conditions respectively.

Not surprisingly, COPERT5 is not able to capture the variability in real-world driving emissions induced by different driving dynamicity when they result in the same average speed. It can also be observed from Figure 3.5 that COPERT5 poorly predicts the mean NO_x emission factor for diesel vehicles: COPERT5 presents average absolute discrepancies of 31.4% for Euro 5, 70.8% for Euro 6 pre-RDE and 380.6% for Euro 6 post-RDE diesel vehicles. For petrol vehicle, discrepancies between the measured mean NO_x emission factors and the prediction from both models are significantly lower in absolute terms but stay high relatively: COPERT5 presents absolute discrepancies of on average 53.0% for Euro 5, 30.2% for Euro 6 pre-RDE and 146.9%

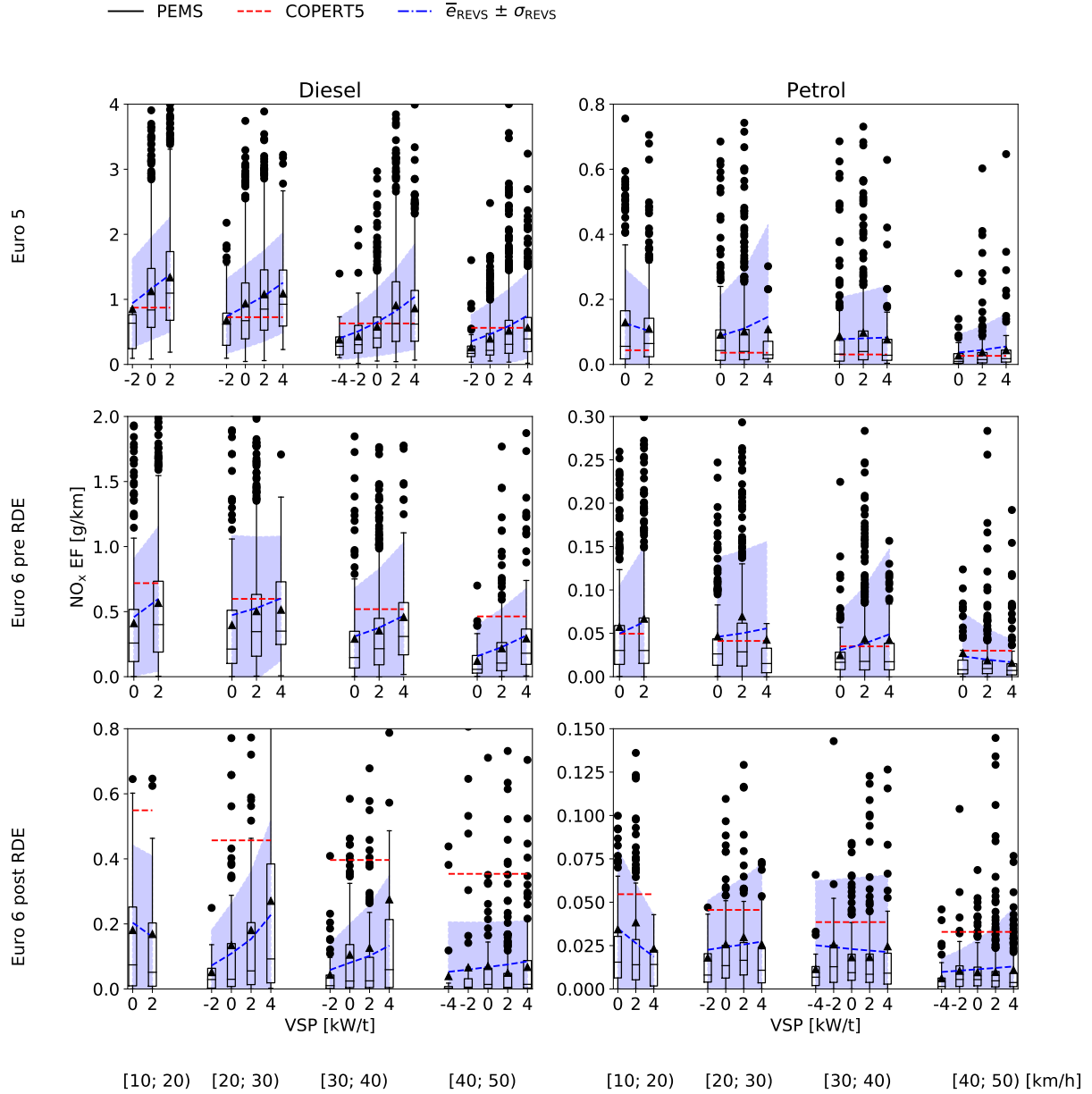


Figure 3.5: Comparison of the measured and predicted NO_x emission factors in urban driving conditions split by vehicle technology category. The average NO_x emission factor predicted by COPERT5 (respectively the REVS model) is depicted by a dashed red (respectively blue) line. The predicted standard deviation by the REVS model is depicted by blue shaded areas around the average. The average speed of the trip is comprised between 10 and 50 km/h aggregated per bin of 10 km/h. The average VSP of the trip is comprised between -2 and 6 kW/t aggregated per bin of 2 kW/t (only the lower bin is indicated on the Figure). The distributions of the measured NO_x emission factor are constituted of all the extracted trips with the corresponding binning in speed and VSP. Note: Some outliers have been cropped and distributions are displayed only then more than 10 trips per bin are available.

for Euro 6 post-RDE petrol vehicles. We attribute these discrepancies to 2 main factors: (i) COPERT is calibrated, at least partially, on driving cycles that might not be representative of real-world driving conditions [Lejri et al., 2018] and (ii) the calibration of COPERT5 for the new generation of vehicles for which testing is not available is based on assumptions that might prove erroneous. To compare, the REVS model present significantly lower absolute discrepancies with the measured average NO_x emission factors: these are on average of 13.5% for diesel Euro 5, 10.4% for diesel Euro 6 pre-RDE, 24.2% for diesel Euro 6 post-RDE, 15.7% for petrol Euro 5, 13.8% for petrol Euro 6 pre-RDE and 24.3% for petrol Euro 6 post-RDE. We further discuss the general predictive skills of the REVS model in the next Section.

3.5.3 Comparison of the models performance against the dataset

The general predictive skills of the REVS model are assessed in Figure 3.6 through the comparison of the predicted and measured mean NO_x emission factor (a, b) and its standard deviation (c) as a function of the number of trips per bin of speed and VSP (colour). Tables S3 and S4 in Appendix A give all the coefficients of the REVS model per vehicle technology category. Not surprisingly, COPERT5 model (in red) is not able to capture the real-world variability in emissions. In contrast, the REVS offers satisfactory performance in predicting the mean NO_x emission factor as a function of VSP: 56.9% (respectively 86.5%) of the predictions present a relative error smaller than 20% (respectively 50%). The performance of the prediction of the standard deviation is similar: 59.2% (respectively 84.5%) of the predictions present a relative error smaller than 20% (respectively 50%). Considering that the REVS model is a simple two-parameter non-linear function, this performance can be considered as satisfactory. Nevertheless, it is also clearly evident from Figure 3.6 that some predictions are inevitably inaccurate. This could be explained by 2 main factors: (i) the variable number of datapoints per bin of speed and VSP (i.e. errors are larger where the number of datapoints is small) and (ii) the standard deviation directly depends on the number of datapoints per bin of speed and VSP and greatly varies from one bin to the other. In fact, the discrepancies between the REVS model and the measurements are the highest where the number of datapoints per bin of speed is small (typically

between 50 and 70 km/h and above 130 km/h). Unfortunately, this aspect is not something that can be solved with the dataset currently available. Despite this, the REVS model appears to be robust and accurate.

3.5.4 Validation of the REVS model for a fleet mix

This Section validates REVS for a fleet containing a mix of vehicle belonging to the six vehicle technology categories described earlier. The accuracy of the model is investigated against a fleet constituted of 20 to 70 passenger vehicles with the following fleet mix: 30% of petrol Euro 5, 15% of petrol Euro 6 pre-RDE, 15% of petrol Euro 6 post-RDE, 20% of diesel Euro 5, 10% of diesel Euro 6 pre-RDE and 10% of diesel Euro 6 post-RDE. Figure 3.7 displays a comparison of the ratio of the calculated average (respectively the calculated standard deviation) NO_x emission factor to the measured average (respectively the measured standard deviation) NO_x emission factor for both REVS and COPERT5, split by average speed. It can be observed that both the ratio of $\bar{e}_M$ to $\bar{e}_{\text{REVS}}$ and the ratio of σ_{TRUE} to σ_{REVS} are close to 1 and all the predictions present a relative error smaller than 25%. It can be noticed that the predicted standard deviation present larger relative absolute error than the predicted mean, but the absolute error stays below 0.018 g/km. Moreover, it can be noticed from Figure 3.7 that the accuracy of the predicted standard deviation is not dependent on the size of the fleet but presents a speed-dependent bias. This bias is more pronounced at speeds where the fewest number of trips were extracted (50 to 60 km/h, 90 to 100 km/h). Overall, this result indicates that it is possible to accurately and easily infer on the mean NO_x emission factor and its standard deviation for a group of vehicles of different vehicle technology categories only based on the mean and the standard deviation of the separated vehicle technology categories.

3.6 Example of the impact of VSP

In this section, the performance of the REVS model is demonstrated by a case study, and compared to the COPERT5 model. The evaluation of the two models was conducted against

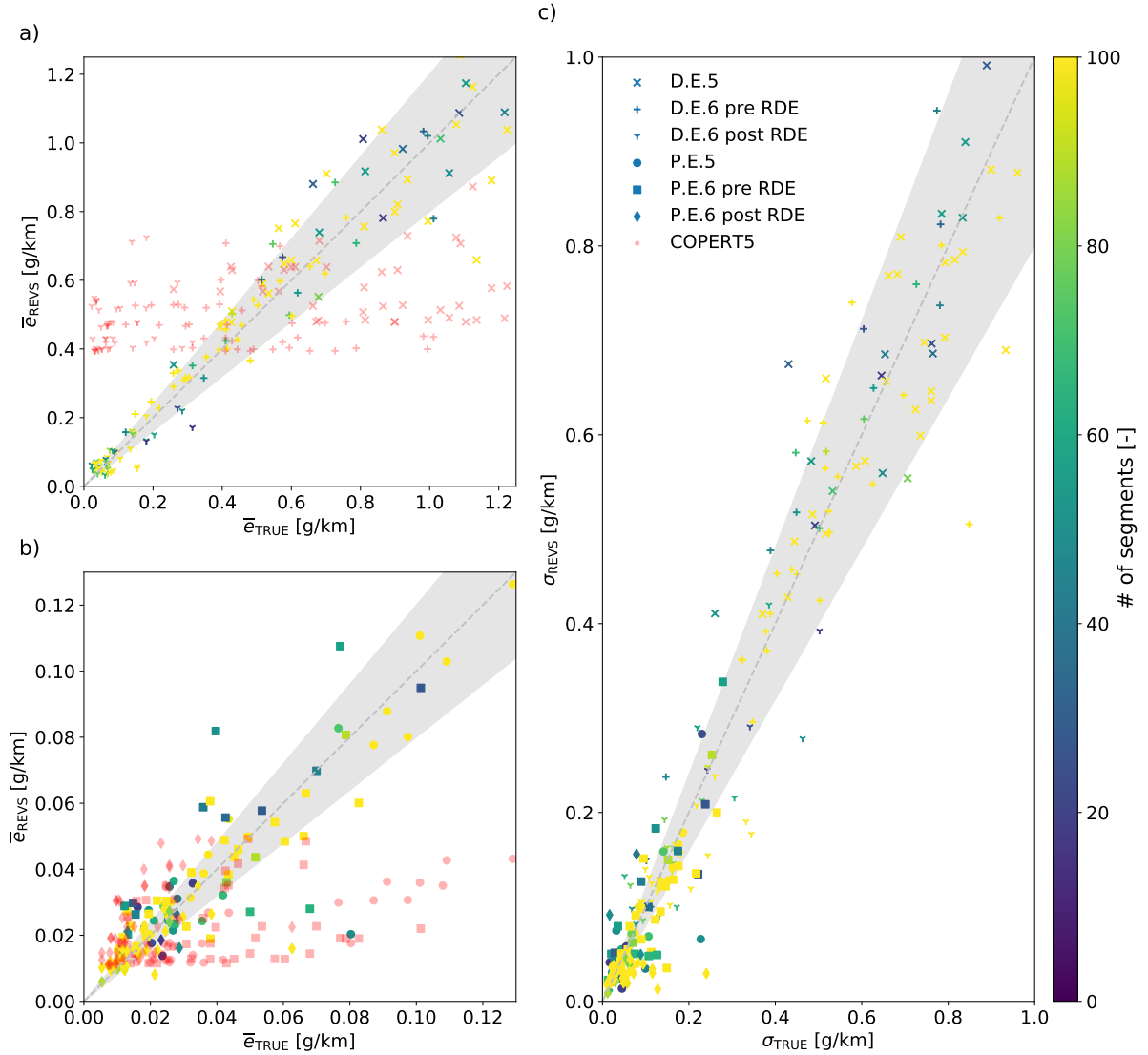


Figure 3.6: Comparison of the measured (x-axis) and predicted (y-axis) mean (a, b) and standard deviation (c) for all the vehicle technology categories for speed between 10 and 130 km/h aggregated per bin of 10 km/h and VSP between -2 and 6 kW/t in urban driving conditions, 4 and 12 kW/t in rural driving conditions and 6 and 18 kW/t in motorway driving conditions per bin of 2 kW/t. The marker represents the vehicle technology category. The predicted average NO_x emission factor by COPERT5 is depicted in light red. The silver dashed line represents the identity line and the light silver envelop represents the 20% error around the identity line. The number of trips extracted per datapoint is depicted with a viridis colour scheme. Note: datapoints are displayed only when more than 10 trips per bin is available.

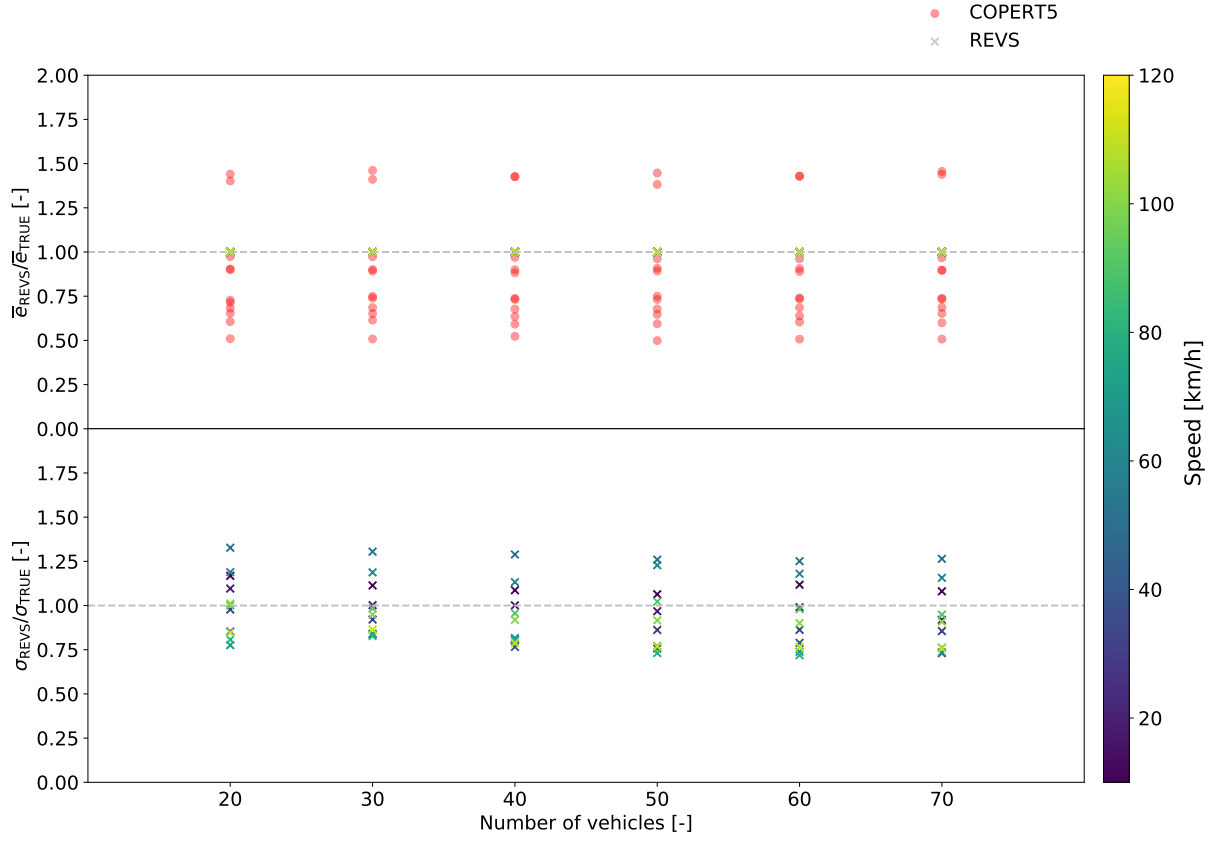


Figure 3.7: Comparison of the ratio of $\bar{e}_M$ to $\bar{e}_C$ (top) and the ratio of σ_M to σ_C (bottom) as a function of the number of vehicles considered. The fleet is constituted of: 30% Petrol Euro 5, 15% Petrol Euro 6 pre-RDE, 15% Petrol Euro 6 post-RDE, 20% Diesel Euro 5, 10% Diesel Euro 6 pre-RDE and 10% Diesel Euro 6 post-RDE vehicles. The predicted NO_x emission factors from COPERT5 are depicted in red. The silver dashed line represents the identity line. The average speed of each datapoint is depicted with a viridis colour scheme.

Table 3.2: Characteristics of the fleet used in the case study

P.E.5	P.E.6 pre-RDE	P.E.6 post-RDE	D.E.5	D.E.6 pre-RDE	D.E.6 post-RDE
9	6	6	3	5	9

a hypothetical scenario aiming to analyse the impact of driving dynamicity on emissions. A diagram summarising the steps to apply the REVS model is available in Section A.5 of Appendix A.

London Road (Southwark, London, United Kingdom), a street of 500 meters, situated within the London Ultra Low Emission Zone (ULEZ), was examined. The London ULEZ was introduced on the 8th of April 2019 in Central London (Covent Garden, Bloomsbury, City of London and Southwark) and operates 24 hours a day, every day of the year [GLA, 2020]. Vehicles must meet strict emission standards to drive in the ULEZ area: Euro 4 and above for petrol passenger vehicles and vans, Euro 6 and above for diesel passenger vehicles and vans, Euro VI for lorries, buses and coaches. In January 2020, after the first 10 months of operation, the compliance rate with the ULEZ standard was of 79% in a 24 hour period [GLA, 2020].

The case study aims to compare the outcome of an increase in driving dynamicity of some vehicles on the NO_x emissions occurring on a single street for the current London ULEZ fleet. This study is limited to passenger vehicles and does not provide any insight or model for other road users (buses, trucks, vans, motorcycles). The case study aims to represent an off-peak, non-congested traffic situation to match the conditions in which the datasets were acquired. A total of 38 passenger vehicles, which corresponds to the traffic flow over 10 minutes on London Road in the aforementioned traffic conditions, was retained [Department for Transport, 2019]. In order to derive the hourly traffic flow from the daily traffic counts provided, it was assumed that the majority of the traffic occurs between 5am and 11pm and the average daily traffic counts were simply divided by 18. The fleet mix was assumed to follow the NAEI fleet projection for the ULEZ in 2020 [National Atmospheric Emissions Inventory, 2020]. The detailed characteristics of the fleet are presented in Table 3.2.

Finally, two scenarios were compared:

1. The speed is limited to 40 km/h, and the average speed of the vehicles is comprised between 30 and 40 km/h. The traffic flow is smooth, and the average VSP of each vehicle is comprised between -2 and 2 kW/t.
2. The speed is limited to 40 km/h, and the average speed of the vehicles is comprised between 30 and 40 km/h. Some vehicles are driving aggressively, and the average VSP of each vehicle is comprised between -8 and 8 kW/t.

In order to estimate the real distribution of the NO_x emission factor, 5,000 random samples corresponding to the scenarios specifications were drawn from the dataset. More specifically, for the 2020 fleet, every random sample corresponds the average NO_x emission factor of 9 petrol Euro 5, 6 petrol Euro 6 pre-RDE, 6 petrol Euro 6 post-RDE, 3 diesel Euro 5, 5 diesel Euro 6 pre-RDE and 9 diesel Euro 6 post-RDE trips. All the trips were selected randomly within the scenarios specifications. Once the real distribution of the NO_x emission factor for the scenarios is built, the real average NO_x emission factor and its associated standard deviation can be easily estimated. For every vehicle technology category, the average speed and the average VSP of all the trips selected to build the distribution were used as inputs to the REVS model and COPERT5. The REVS model and COPERT5 were subsequently applied, as described in Sections 3.2.3 and 3.5. In a second step, the statistics described in Sections 3.3 and A.3 (in Appendix A) were applied to derive the mean NO_x emission factor and, for the REVS model, its associated standard deviation, for the entire fleet.

Figure 3.8 displays the distribution of the real NO_x emission factor and offers a comparison with the predictions from both the REVS and the COPERT5 models. It can be observed in Figure 3.8 that the second scenario results in an increase in emissions by about 71.1% (0.089 g/km) compared to the first scenario. Not surprisingly, this indicates that an increase in driving dynamicity, will have a detrimental effect on the local air quality. This results particularly highlights the importance of accounting for the impact of driving dynamicity on emissions when designing an air quality mitigation strategy. Since the effect of driving dynamicity has been demonstrated to decrease with the successive regulations, it is to expect that the deterioration induced by the second scenario will decrease over time. We make the reader attentive to the

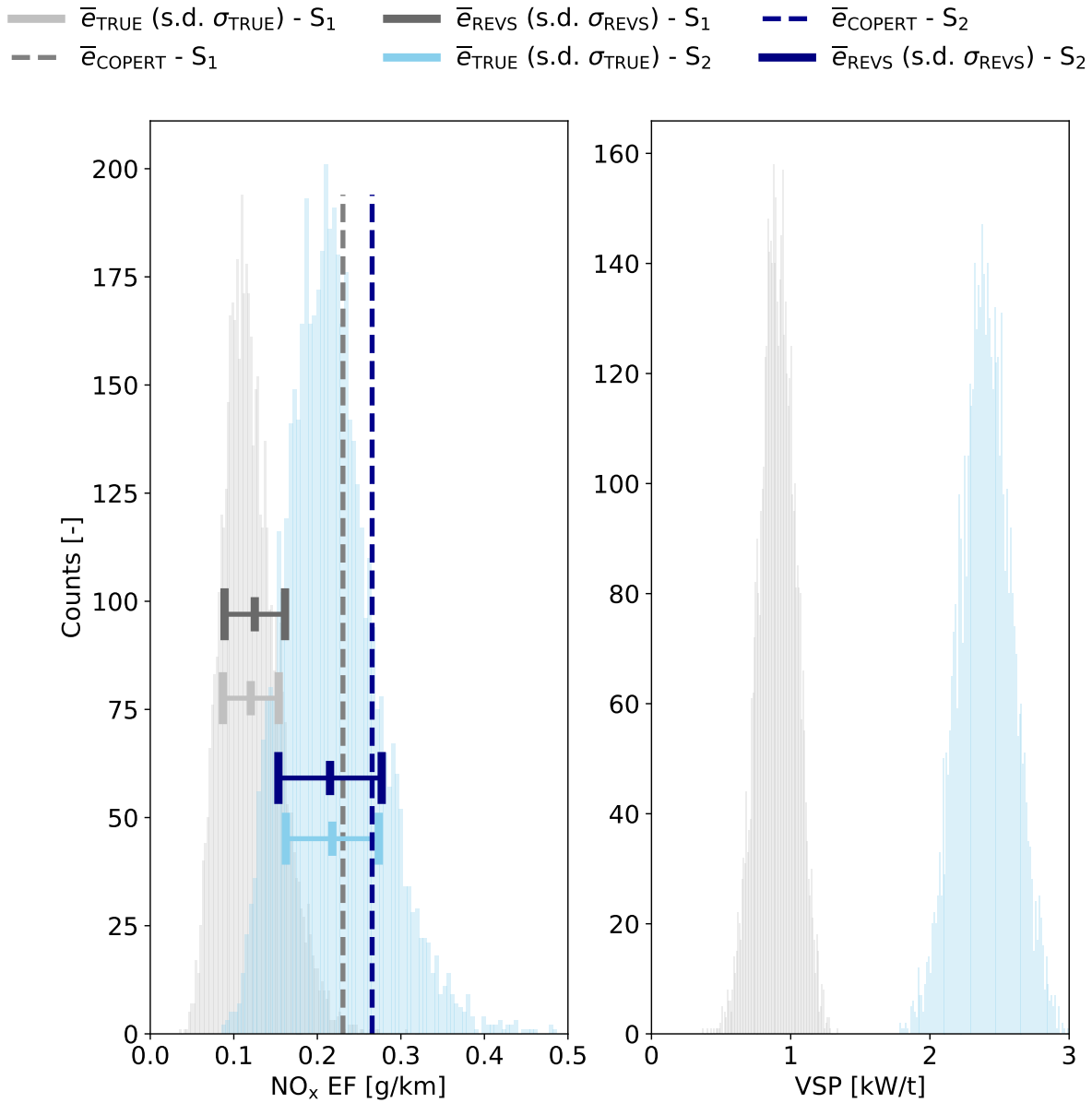


Figure 3.8: Distribution of the real NO_x emission factor and comparison with the predictions from the REVS and the COPERT5 model (left) and distribution of the average VSP of the fleet (right). Scenario 1 is depicted in grey, while scenario 2 is depicted in blue. The vertical dashed lines corresponds to the COPERT5 prediction. The intervals represent the average NO_x emission factor plus or minus one standard deviation. The real mean NO_x emission factors and associated standard deviation are depicted in light blue and light grey. The predicted mean NO_x emission factors and associated standard deviation are depicted in dark blue and dark grey.

fact that the model was developed with the data available at the time. Without knowledge on real-world emissions of vehicles that will be manufactured in 5 to 10 years, any conclusion on the emissions of a future fleet must be taken carefully. Moreover, should the model be used in the future, regular updates will be necessary.

It is interesting to notice from Figure 3.8 that COPERT5 tends to overestimate the average NO_x emission factor by more than 20%, which is significant. We believe that this is mainly due to the systematic overestimation of the emissions from post-RDE vehicles. Finally REVS presents a good accuracy in the prediction, with a relative error on the average NO_x emission factor smaller than 5%. The standard deviation tends to be slightly overestimated but by no more than 9.5% (0.0054 g/km), which can be considered as satisfactory.

3.7 Conclusion

In this study, an analysis of a large PEMS dataset comprising more than 300 Euro 5 and Euro 6 vehicles was conducted. The large size of the dataset allowed us to 1) study the evolution of the NO_x emissions with the introduction of the successive legislation, 2) analyse and quantify the impact of driving dynamics on NO_x emissions and 3) provide an aggregated speed-average and VSP-average NO_x emission model that could be used in future studies to estimate the NO_x emissions from a fleet of flexible size.

Analysis of the data revealed that the successive European regulations were successful at decreasing real-world NO_x emissions, although not to the extent expected. While it is evident that an increasing number of vehicles have satisfactory real-world emission levels with respect to their Euro standard, a significant fraction of them still present NO_x emission factors significantly higher than stated by the legislation. Within each vehicle technology category, the largest variability in NO_x emission factors occurs in urban and motorway conditions. Added to the known intra-vehicle real world variability [Le Cornec et al., 2020], it raises the question of how to design a type-approval procedure able to guarantee that the legal emissions limits are achieved most of the time under driving conditions representative of real-world driving.

We provided evidence suggesting that driving dynamics largely impacts NO_x emissions of both diesel and petrol Euro 5 and Euro 6 pre-RDE vehicles but has a more moderate impact on the NO_x emissions of other vehicle technology categories. Moreover, it appears that the relative influence of driving dynamics on NO_x emissions tends to decrease with the introduction of the successive regulations. We believe that these observations should be accounted for and integrated into the emission models, motivating the development of the new proposed emission model.

Analysis and comparison of proposed model against real data and COPERT5 shows that: 1) with respect to prediction of NO_x emissions, the COPERT5 model is not able to capture the variability induced by different driving dynamicity when they result in the same average-speed, 2) COPERT5 largely overestimates the emissions from post-RDE vehicles, 3) the new model is able to accurately predict the mean NO_x emission factor using a simple two-parameter non-linear function of the average VSP of the trip, binned by average speed; 4) the REVS model is able to accurately predict the standard deviation of the NO_x emission factor using the same simple parametrisation than for the mean NO_x emission factor. Our results show that a reliable estimate of the mean and standard deviation of the NO_x emission factor of a specific fleet can be easily derived from the estimates of the mean and the standard deviation of the NO_x emission factor provided by the REVS model for the individual vehicle technology categories. Finally, one main limitation of the REVS model is its current unavailability for non-passenger road users such as HGVs and buses, which represent a large fraction of the NO_x emissions. Once the data will be available, future work should focus in solving this shortcoming.

This study is the first to demonstrate on a large dataset that it is possible to accurately predict the mean and the standard deviation of the NO_x emission factor only knowing the mean NO_x emission factor and its standard deviation per vehicle technology category. These could be estimated with a reasonable accuracy using the new REVS model. The model presented rely on simple inputs such as speed, acceleration and slope, which could be extracted from existing sources including traffic cameras, vehicle tracking devices and topographic maps and therefore deployed immediately. We believe that the better modelling of NO_x emissions achieved with the methodology presented in this study will tremendously benefit to policy makers in their

attempt to implement air quality mitigation strategies.

Chapter 4

Modelling of instantaneous emissions from diesel vehicles with a special focus on NO_x : Insights from machine learning techniques

This chapter presents a machine learning based-analysis of instantaneous NO_x emissions in real-world driving conditions and establishes an instantaneous NO_x emission model that rely on simple inputs of vehicle speed and acceleration. Particular attention is paid to (1) clustering vehicles with similar emission performance, (2) analyse the obtained clustering, (3) the development of a new neural network based instantaneous emission model and (4) the assessment of the new model with respect to the data and to other existing instantaneous NO_x emission models. The results presented in this chapter have been published in:

Clémence M. A. Le Cornec, Nick Molden, Maarten van Reeuwijk and Marc E. J. Stettler (2020). Modelling of instantaneous emissions from diesel vehicles with a special focus on NO_x : Insights from machine learning techniques. *Science of The Total Environment*, 737, 139625.

The chapter has the following structure. As an introduction, Section 4.1 presents the existing research gaps and the objectives of this study. Section 4.2.1 details the data sources and any preprocessing performed prior to the study. Section 4.2 presents the methodology, first introducing how dynamic time warping and conventional clustering techniques were applied to the data, and then explaining how a neural network able to predict instantaneous NO_x emissions was developed and compared to existing instantaneous emission models. Section 4.3.1 discusses the results of the clustering analysis, and Sections 4.3.2 and 4.3.3 assess the performance of the developed instantaneous NO_x emission model with respect the data and to other existing models. Finally, Section 4.4 summarises the main findings and discusses how the presented work could be used.

4.1 Introduction

Emissions models, combined with atmospheric dispersion models, are required to assess the impacts of road transport on the local air quality and implement successful mitigation strategies. While aggregated emissions models operate at a trip- or link-based level, instantaneous emissions models aim to provide a precise description of vehicle emission behaviour by relating vehicle emission rates to vehicle operating conditions during a series of short time steps, usually seconds. They typically provide the resolution needed to evaluate the very local impact of a traffic intervention, at the cost of an increase in computational power and data requirement.

Most of the current emissions models are based on physical, chemical, empirical or semi-empirical assumptions but these classical approaches present limited performance when used to model instantaneous NO_x emissions [Panis et al., 2006]. Existing emissions models achieve limited performance to estimate NO_x emissions in real-world driving conditions [Smit et al., 2010]. [O’Driscoll et al., 2016] demonstrate that COPERT4.11 underestimates NO_x emissions in real-driving conditions by a factor 1.6 compared to real measurements. PHEM overestimates NO_x emissions in real-driving conditions by up to 25% [Matzer et al., 2017]. Finally, [Ntziachristos et al., 2016] show that VERSIT+ emission factors underestimate NO_x emissions by

a factor up to 1.5 compared to emissions in real-world driving conditions. The lack of understanding of the highly non-linear processes behind real-world emissions partially explains the limited performance of the existing emission models. Improvements to real-time estimation of NO_x emissions is therefore vital to the assessment of road transport responsibility in air quality issues and will only be possible by improving our understanding of the processes governing real-world emissions. With the continuous increase in computational power and data availability, unsupervised machine learning techniques, which are able to successfully find previously unknown patterns in datasets, could help to improve emission models. Similarly, Artificial Neural Networks (ANNs), which have shown their effectiveness in solving a wide range of non-linear problems, could help to improve emissions models.

To the best of our knowledge, no studies have used unsupervised machine learning techniques in order to acquire a better understanding of the processes underlying real-world instantaneous NO_x emissions and ultimately improve emission models. However, ANNs have been used to predict fuel consumption [Parlak et al., 2006, Thompson et al., 2000] or particulate emissions [de Lucas et al., 2001]. A drawback of these studies was that only a small amount of data, acquired on a chassis-dynamometer for standardised type-approval tests (US06, NEDC), was available to perform the network training and validation. In addition, the networks' architectures developed were simple, consisting of a multi-layer perceptron of one or two hidden layers. The models show reasonable performance but are specific to a small set of vehicles operated in specific conditions and are extremely unlikely to generalise to real-world driving conditions. Only a limited number of studies have explored the potential of ANNs to predict NO_x emissions from road vehicles (e.g. [Alonso et al., 2007, Kesgin, 2004, Obodeh and Ajuwa, 2009, Atkinson et al., 1998, Johri and Filipi, 2014, Traver et al., 1999]). In all these studies, the data available to develop these models was limited to as small number of chassis-dynamometer tests and comprised only a small set of parameters. Subsequently, the network architectures consisted of one hidden layer multilayer perceptron, which is appropriate in the context of a small dataset where the variability is limited. However, a more complex architecture is likely to be required for a larger dataset comprising tests from several different vehicles in real-world driving.

4.2 Methods

As mentioned earlier, there is no single emission model that meets all criteria of accuracy and simplicity required for policy-making purposes. Current general instantaneous emission models such as look-up tables [Manzie et al., 2007, Hatzopoulou and Miller, 2010] or non-linear regression models [Panis et al., 2006] are unable to reach the accuracy required [Hagena et al., 2006, Panis et al., 2006] while vehicle-specific models present the desired accuracy but would require every vehicle on the market to be tested in real-world driving conditions, which is currently unrealistic. This observation provided the motivation to: (i) perform a clustering analysis on the dataset in order to group together vehicles presenting similar emitting behaviours in similar operating conditions, which would ultimately help reduce the number of models needed to reach the desired accuracy; (ii) conduct an analysis on the results of the clustering to identify external variables able to explain the results of the clustering; (iii) update the existing instantaneous emission models based on the clusters previously defined; (iv) use Artificial Neural Networks to develop a new instantaneous emission model based on the clusters previously defined; (v) unify the clusters-based models into general models; (vi) quantify the performance and the uncertainties of each model.

This section firstly gives a description of the dataset used in this study in Section 4.2.1. The methodology for clustering is presented in Section 4.2.2. Section 4.2.3 then describes the methodology used to update existing instantaneous emission models and finally a description of the model developed in this study is provided in subsection 4.2.4.

4.2.1 Data description

The dataset used in this study was supplied by Emission Analytics and was previously described by [O’Driscoll et al., 2018]. A Portable Emissions Measurements System (PEMS) was used to measure real world driving emissions of 36 Euro 5 and 34 Euro 6 diesel passenger vehicles in the Greater London area between 2012 and 2016. The vehicles were made by 19 different manufacturers (Ford, Nissan, Mercedes, Citroen, Kia, Peugeot, Skoda, Volkswagen, Toyota,

Table 4.1: Dynamic characteristics of the test segment

Parameter	Value
Duration [min]	27.55 ± 6.56
Distance [km]	10.30 ± 0.095
Average speed [km/h]	23.47 ± 2.50
Kinetic intensity ¹ [km^{-1}]	2.12 ± 0.32
% of idle [%]	20.75 ± 5.75
Relative Positive Acceleration (RPA) [m/s^2]	0.19 ± 0.03
95 th percentile of the product of RPA and speed	12.49 ± 3.99

¹ Defined in Equation 19 in [O’Keefe et al., 2007] as:

$$KI = \frac{\sum_{j=1}^{N-1} \text{positive}(1/2(v_{j+1}^2 - v_j^2) + g(h_{j+1} - h_j))}{\sum_{j=1}^{N-1} \overline{v_{j,j+1}^3} \cdot \Delta t_{j,j+1}} \quad (4.1)$$

where v_j is the instantaneous speed at time j , g is the gravitational constant, h_j is the instantaneous altitude at time j , $\Delta t_{j,j+1}$ is the time step between time t and $t+1$ and $\overline{v_{j,j+1}^3}$ is defined as:

$$\overline{v_{j,j+1}^3} = \frac{v_{j+1}^3 + v_{j+1}^2 v_j + v_{j+1} v_j^2 + v_j^3}{4} \quad (4.2)$$

Audi, Land Rover, Hyundai, Mitsubitshi, Ssangyong, BMW, Fiat, Volvo, Mazda, Vauxhall). All these vehicles were fitted with a Diesel Oxidation Catalyst (DOC), a Diesel Particulate Filter (DPF) and an Exhaust Gas Recirculation (EGR) system. The Euro 6 vehicles had a mixture of NO_x aftertreatment systems, 7 used EGR only, 16 used EGR + Lean NO_x Trap (LNT) and 11 used EGR + Selective Catalytic Reduction (SCR). We refer the interested reader to [O’Driscoll et al., 2018] for a more detailed description of the data collection process. For the sake of clarity, we also indicate that the dataset used in this study is a fraction of the EA2016 dataset presented in Chapter 3.

All the vehicles were tested on the same route and for our analysis we have extracted a geographically consistent segment of the route using the instantaneous coordinates as measured by the embedded GPS. This sub-section was in an urban area and had a length of 10.3 km. All the vehicles in the dataset travelled this journey one to three times and instantaneous speed, coordinates, external temperature and NO_x emissions were recorded. A total of 110 segments were extracted: 54 for Euro 5 vehicles and 56 segments for Euro 6 vehicles. The average dynamic characteristics of the segments of are given in Table 4.1.

4.2.2 Clustering

Clustering is the process of grouping a set of objects into clusters so that objects within clusters are similar to each other and dissimilar to objects in other clusters [Park and Jun, 2009]. An important step in most clustering algorithms is the selection of a distance metric, which determines how the similarity between two elements is calculated [Velmurugan, 2012]. This step is particularly challenging when working with time-series because a standard data point to data point distance metric is not appropriate due to potential variations in speed and length between two time-series that are compared. Dynamic Time Warping (DTW) is an algorithm designed to solve the aforementioned issue and estimate the distance between two time-series that may vary in speed and length. For a detailed description of the DTW algorithm, the interested reader is referred to Section B.1 of Appendix B.

In this study, DTW was applied to the instantaneous measured NO_x emissions of the 110 test segments previously extracted. The distance d between two tests was quantified by the integral of the time-warped absolute NO_x emissions over the test route. By applying this procedure to each test pair, an emission distance matrix d_{ij} was constructed; this matrix is square, hollow and symmetric, each row containing the computed distances from one specific test to all the others. Two tests presenting similar emitting behaviours (namely there are similarities in the structure of the time series of NO_x emissions across the route) will consequently be close to each other and similarly far from other tests. The distance matrix can therefore be used to cluster the dataset. The k-means++ algorithm, a popular unsupervised clustering algorithm, was identified as a suitable technique for this task. For a more detailed description of the k-means++ algorithm, the interested reader is referred to Section B.2 of Appendix B. The k-means++ algorithm, using a Euclidean distance evaluation, was applied to the previously defined distance matrix for 2 to 20 clusters. The optimal number of clusters was finally estimated using the Davies-Bouldin and the Calinsky-Harabatz indices. For more details about the the Davies-Bouldin and the Calinsky-Harabatz indices, the interested reader is referred to Section B.3 of Appendix B.

The resulting clusters were subsequently analysed with respect to emission factors, vehicle

aftertreatment, vehicle engine size and vehicle weight in order to identify a physical variable able to explain the clusters. In addition, the defined clusters were used in the next steps of this study to update existing instantaneous emission models and develop a new emission model. For the sake of completeness, the clustering analysis was also performed for CO and CO₂ for which results are discussed in the main text and figures are available in Appendix B.

4.2.3 Calibration of existing emission models

Clustering based on the NO_x enables the identification of groups of vehicles presenting similar emitting behaviours. These clusters can then be used to adjust existing emission models whose previous general formulation limited the accuracy of the instantaneous NO_x prediction. In this study, we consider an instantaneous speed and acceleration look-up table (hereafter referred to as the LT model) and the non-linear regression developed by [Panis et al., 2006] (hereafter referred to as the NLR model, described in Section B.6.2 in Appendix B). As the aim of this study is to demonstrate the potential of machine learning techniques applied to vehicle emissions modelling, the analysis has not been conducted on other instantaneous emission models. One look-up table with bins of 2 m/s for speed and 0.5 m/s² for acceleration was developed per cluster. The instantaneous mean NO_x emission was computed for each cell of the look-up table. Similarly, the twelve parameters of the NLR model were refitted for each cluster.

In order to evaluate the robustness of the methodology, bootstrapping was conducted by randomly withdrawing 20 tests out of the 110 composing the original dataset before the determination of the coefficients of the emission models (i.e. before computing the look-up table and before refitting the twelve parameters of the non-linear regression). This 20 tests belonged to the clusters containing more than one test (later referred to as main clusters), selected in a way that 20% of the tests in each cluster is used as validation set. For clusters containing only one test, 20% of the test was randomly selected and used as validation set. This step was performed 400 times to ensure that the results were repeatable. Three of the runs (3 x 20 tests) were aggregated together and used to assess the performance of the models with respect to real data. The whole 400 bootstraps were used to study the convergence properties of the models.

We make the reader attentive to the fact that the clustering was performed only once, before any further analysis, on all the tests.

4.2.4 Neural network model description

Artificial Neural Networks were identified as a potentially effective method to create new emission models with improved performance compared to existing emission models. The interested reader is referred to Section B.4 of Appendix B for a brief non-exhaustive introduction to Artificial Neural Networks. One important feature that should be highlighted is the absence of a methodology to determine the optimal architecture and optimization parameters for specific problems. Architectures and parameters must therefore be tested in an iterative search to improve the results, until a trade-off between the time spent and the desired accuracy is reached.

The model presented in this study is a Multi-Layer Perceptron (MLP) consisting of five layers of 100, 75, 50, 25 and 1 neurons respectively, with ELUs activation functions and a dropout rate of 0.1. The Root Mean Squared Error (RMSE) was chosen as loss function as the only aim is to match the real emission values. The smaller the RMSE, the better the agreement between the prediction and the real value. The initial learning rate λ was set to 0.0001, which is a standard value. The training algorithm is Adam [Kingma and Ba, 2014], which was demonstrated in a previous study to present improved performance and convergence compared to Adagrad, Stochastic Gradient Descent (SGD), RMSProp and Adadelta [Kingma and Ba, 2014]. The model was developed using **TensorFlow 1.6** [Abadi et al., 2015] and **Python 3.6** (Python Software Foundation).

In order to allow a fair comparison between the MLP model and the LT and NLR models, only the instantaneous speed and acceleration at each time step were used as inputs to the MLP model. The same randomly selected 3 x 20 tests described in Section 4.2.3 were withdrawn from the training set. These three runs were used to estimate the accuracy and the robustness of the methodology, with respect to the unseen data on the three runs. In order to study the convergence properties of the MLP model, the bootstrapping procedure was performed 400

times, with 20% of the tests in each cluster randomly removed at each run. The resulting mean emission map was compared with the LT and the NLR models emission maps. Since instantaneous emissions at time t are a complex non-linear process that might depend not only on the operating conditions at time t but also on the operating conditions a few seconds before, the impact of including some time history data was explored. Both models were therefore provided with not only the instantaneous speed and instantaneous acceleration at time t but also the instantaneous speed and instantaneous acceleration from time $t - X$ to t , where X belong to $[0; 10]$, resulting in 10 different models. However, and contrary to what was expected, including time history data did not help improving the accuracy of the MLP model (Figure B.44 in Appendix B).

4.2.5 Dealing with trips that do not belong to a cluster

For tests that do not belong to one of the main clusters, two options were compared for each model (i.e. LT, NLR and MLP). The first option (hereafter referred to as LT_1 , NLR_1 and MLP_1) consisted of developing a specific model for the test. The second option diverges slightly for the MLP model compared to the LT and NLR models. For the LT and NLR models (hereafter referred to as LT_2 and NLR_2), the final predictions were interpolated from the predictions of the clusters using an inverse distance weighting function. The distance between the tests and the centroids of the clusters was computed using the DTW algorithm.

The instantaneous emissions for the LT_2 and NLR_2 models are therefore computed as:

$$E_N = \begin{cases} E_i & \text{if } N \in S_i \\ \frac{\sum_i E_i / \overline{d_{iN}}}{\sum_i 1 / \overline{d_{iN}}} & \text{if } N \notin \bigcup_{i=1}^M S_i \end{cases} \quad (4.3)$$

where N is the cluster to which the evaluated test belongs to, S_i represents the set vehicles belonging to cluster i , M is the total number of clusters, $\overline{d_{iN}}$ represents the distance between cluster N to the mean of cluster i , and E_i denotes the predicted instantaneous emissions of the cluster i for LT_1 or NLR_1 model.

For the MLP model (hereafter referred to as MLP_2), one unique model for all the main clusters and individual tests not belonging to one of the main cluster was defined. The MLP_2 model was provided with not only the instantaneous speed and acceleration but also the distances of the segment to all the main clusters' centroids. A complete description of the six models and their inputs is available in Table 4.2.

The accuracy and robustness of the MLP_1 model was compared with the LT_1 and the NLR_1 models. The same approach was conducted to compare the accuracy and robustness of the MLP_2 model with the LT_2 and the NLR_2 models. A cross-comparison of the two types of models was also performed.

4.2.6 Code availability

The Python implementation of the fast DTW of [Salvador and Chan, 2007] was used. Instructions to install the package can be found in [Salvator, Stan and Chan, Philip, 2019]. The python implementation of the k-means++ algorithm of the scikit-learn package, developed by [Pedregosa et al., 2011], was used.

Finally, the MLP model implemented in this work is freely available [Le Cornec, Clemence M. A, 2019]. The dataset described in this study is the property of Emissions Analytics and cannot be disclosed. Therefore, to test the code, the interested reader can use the freely available dataset from the UK Department for Transport, [Department for Transport, 2016]. We verified our code using the euro6-renault-megane-1461.xlsx file, available in the Euro 6 vehicle data zip file. This file was used both as training and testing dataset to ensure that the code was running properly.

Table 4.2: Summary of the models used in the study

Model	Description	Inputs	Number of models
LT_1	Updated look-up table model. One look-up table per cluster is developed. If a test does not belong to one of the main cluster, an individual look-up table is developed.	$v(t), a(t)$	Number of main clusters + Number of tests outside the main clusters
LT_2	Updated look-up table model. One look-up table per cluster is developed. If a test does not belong to one of the main cluster, the prediction is interpolated from the main clusters on a distance basis.	$v(t), a(t)$, distances to the main clusters	Number of main clusters
NLR_1	Updated Int Panis regression model. One regression model per cluster is developed. If a test does not belong to one of the main cluster, an individual regression model is developed.	$v(t), a(t)$	Number of main clusters + Number of tests outside the main clusters
NLR_2	Updated Int Panis regression model. One regression model per cluster is developed. If a test does not belong to one of the main cluster, the coefficients of the regression are interpolated from the main clusters on a distance basis.	$v(t), a(t)$, distances to the main clusters	Number of main clusters
MLP_1	Multi-layer perceptron model. One MLP per cluster is developed. If a test does not belong to one of the main cluster, an individual MLP is developed.	$v(t), a(t)$	Number of main clusters + Number of tests outside the main clusters
MLP_2	General multi-layer perceptron model.	$v(t), a(t)$, distances to the main clusters	1

4.3 Results

The present study was conducted with two main points of focus. The results of the clustering are analysed: (i) analysis of the main clusters (clusters containing more than 3 tests) is provided in terms of driving dynamics and physical characteristics (vehicle weight, engine size, NO_x aftertreatment system) of the vehicles in each cluster; then (ii) special attention is given to the analysis of tests of the same vehicle classified in different clusters. The performance of the six models (LT₁, LT₂, NLR₁, NLR₂, MLP₁ and MLP₂) are analysed, with a focus on the analysis of the importance of spatial variability in emissions.

Quantitative analysis of the driving dynamics was performed using the average speed, the Relative Positive Acceleration (RPA, minimum metric of the dynamic boundary conditions in (EU) 2016/646), the 95th percentile of $v \cdot a_{\text{pos}}$ ($v \cdot a_{\text{pos}}[95]$, the maximum metric of the dynamic boundary conditions in (EU) 2016/646) and the percentage of idling (percentage of the test where the engine was running while the vehicle was not in motion). Quantitative analysis of the performance of the three studied models is performed using the Normalised Mean Square Error (NMSE, a metric quantifying the overall deviation or scatter between the two sets of results) and Fractional Bias (FB, a metric quantifying mean bias and therefore identifying potential sources of systematic error) [Chang and Hanna, 2004].

4.3.1 Clustering analysis

4.3.1.1 Analysis of the main clusters

The best partition is achieved with 17 clusters. Six clusters (hereafter referred to as the main clusters) regroup 97 tests, which represent 88% of the dataset. The eleven other clusters regroup tests of vehicles presenting an emitting behaviour that differs significantly from those of the main clusters. Figure 4.1 presents the main clusters on the x-axis with respect to (a) the average NO_x emission factor over the test segment, (b) the engine size, (c) the vehicle weight. Each manufacturer is depicted with a different marker, while the aftertreatment technology is

displayed in red when the vehicle is equipped with an EGR system only, in blue when the vehicle is equipped with an LNT system in addition to the EGR system and in green when the vehicle is equipped with an SCR system in addition to the EGR system. Finally, the distribution of the NO_x emission factors within each cluster is shown in (d). Figure 4.1a) shows that the NO_x emission factors inside each cluster vary much less than between the clusters, which is expected since the clustering was carried out and based on NO_x emissions. This is further demonstrated in Figure 4.1d). Moreover, it can be observed that neither the manufacturer, the aftertreatment system, the engine size nor the vehicle weight fully explain these clusters even if some partial conclusions can be drawn. Clusters 4 and 10 contain the tests with the lowest emission factors, presenting cluster average emissions factors close to the Euro 6 and Euro 5 limits, respectively. With one exception, cluster 4 contains only vehicles equipped with LNT or SCR systems, which is in agreement with the assertion from [Weiss et al., 2012] that vehicles equipped with EGR systems are not able to reach Euro 6 type approval limits in real-driving conditions. Cluster 10 contains four vehicles each equipped with an EGR system only, but these four vehicles have a significantly smaller engine compared to the rest of the vehicles in this cluster. At the other side of the spectrum, vehicles in cluster 8, which present the highest emission factors distribution are vehicles only equipped with an EGR system. A closer inspection of Figure 4.1 indicates that in general, vehicles equipped with the same aftertreatment system from the same manufacturer (which uses the same family of engines) tend to be grouped together (with one exception, vehicles from Volkswagen equipped with an LNT system are grouped in cluster 4 while most of the vehicles from Volkswagen equipped with an EGR system only are grouped in cluster 5 and 11). The engine optimisation strategies of the different manufacturers are probably one of the main reasons for the discrepancies between the vehicles. Moreover, the overlap of clusters 4 and 10 in terms of NO_x emission factors can be explained by their different emission patterns, as evidenced by Figure 4.2, which presents the mean measured instantaneous NO_x emissions as a function of the distance travelled. This is in agreement with the grouping performed during the clustering and emphasizes that two vehicles presenting a similar emission factor do not necessarily present the same emitting behaviour.

For an analysis of the main clusters in terms of CO , CO_2 and f_{NO_2} , see Section B.6.3 of Appendix

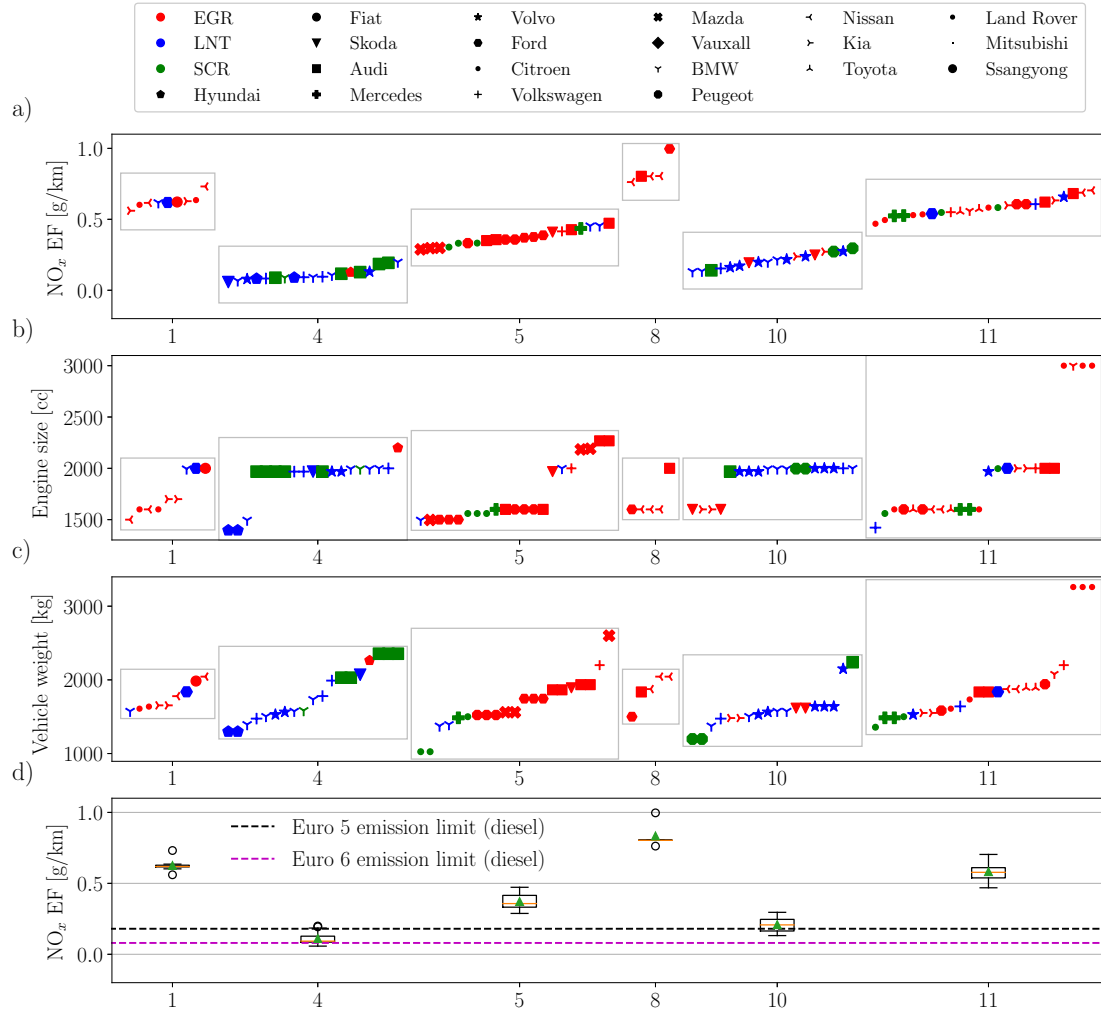


Figure 4.1: (a) Vehicle manufacturer, NO_x aftertreatment technology and emission factor of each vehicle in the cluster, (b) Vehicle manufacturer, NO_x aftertreatment technology and engine size of each vehicle in the cluster, (c) Vehicle manufacturer, NO_x aftertreatment technology and vehicle weight of each vehicle in the cluster and (d) Boxplots of emission factors per cluster. The whiskers of the boxplots represent the lower and upper quartile respectively. The median of the data is represented by an orange line, while the mean is depicted with a green triangle. Clusters containing less than three tests are not depicted here.

B. CO and CO_2 emissions are uncorrelated with vehicular NO_x emissions, which indicates that specific knowledge on CO and CO_2 emissions is unlikely to give any insights on NO_x emissions.

4.3.1.2 CO and CO_2 clustering

A clustering analysis was also performed on CO and CO_2 . For a description of the approach and a detailed analysis of the results, see Section B.6.3 of Appendix B. It is demonstrated that the variability in CO_2 emissions is significantly less marked than the variability in NO_x emissions. From a modelling perspective, performing a clustering analysis in order to improve the CO_2 prediction accuracy is likely to be unnecessary. It is also demonstrated that the real-world variability in CO emissions among the vehicles is high, which indicates that, from a modelling perspective, an *a priori* clustering analysis would likely improve the accuracy of the CO prediction.

4.3.1.3 Intra-vehicle real-world driving variation

Further analysis reveals that for half of the vehicles that were tested more than once (17 out of 33), the tests are classified in different clusters. Table B.1 in Appendix B summarises all the vehicles concerned and gives, in addition to the emission factor and ambient temperature for each of the tests, driving dynamics metrics such as the percentage of idling, the kinetic intensity, the RPA and the 95th percentile of v^*_{apos} . All the tests are compliant with the EU 2016/646 directive for RPA (which means that none of them were driven “too passively”) while eight of the tests are not compliant with the EU 2016/646 directive for v^*_{apos} [95] (which means that these tests were driven “too aggressively” and would not be RDE compliant). Most of the tests present somewhat lower RPA compared to the average real world RPA presented by [Pelkmans and Debal, 2006] ($\approx 0.3 \text{ m/s}^2$ at 25 km/h) and therefore the driving style can globally be described as “smooth”. Figure 4.3.a) (exact numbers available in Table B.1 in Appendix B) demonstrates that the emission factors for the same vehicle can vary by up to a factor of 7 (vehicle 16) while variations up to a factor of 2 are not unusual (vehicles 1 and 9). Ten vehicles out of the 17 present variations of the NO_x emission factor by more than a

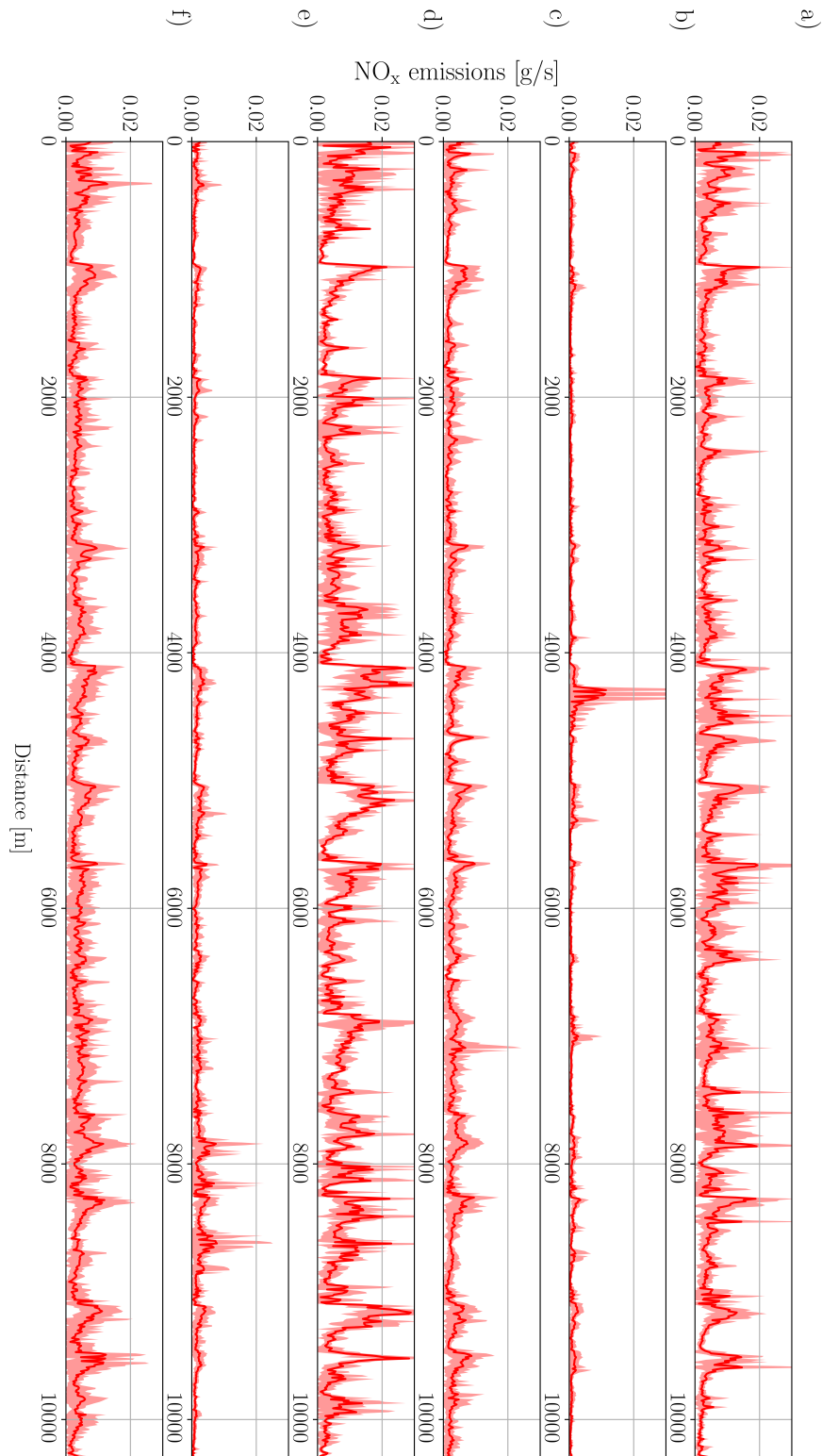


Figure 4.2: Mean measured instantaneous NO_x emissions as a function of the distance travelled for the main clusters (a) Cluster 1, (b) Cluster 4, (c) Cluster 5, (d) Cluster 8, (e) Cluster 10 and (f) Cluster 11. The standard deviation is depicted by the light red envelope.

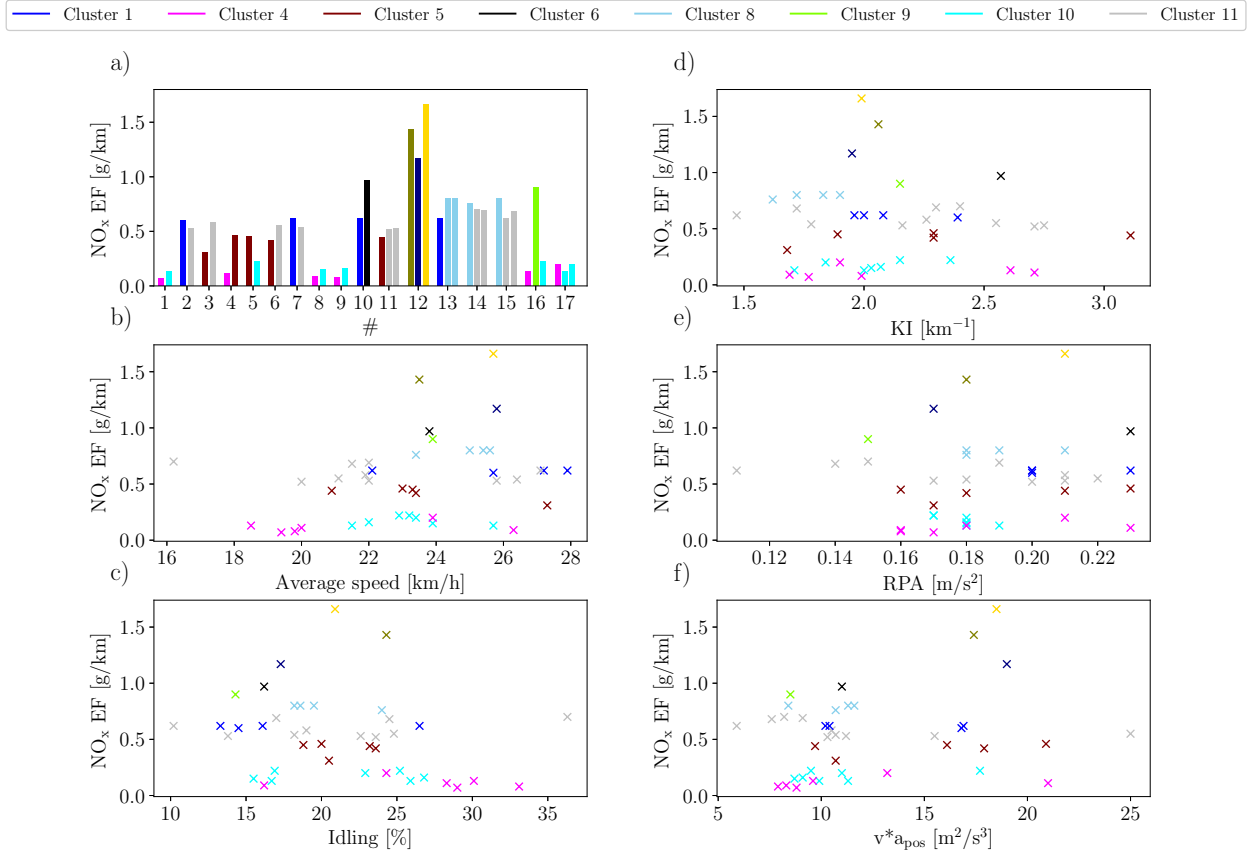


Figure 4.3: Real-world variation of the NO_x emission factor for the same vehicle when the tests of these vehicles fall in different clusters, as a function of a) the vehicle number, b) the segment average speed, c) the percentage of idling, d) the kinetic intensity, e) the relative positive acceleration and f) the 95th percentile of v^*a_{pos} . The cluster number is represented by the colour of the bar or the colour of the point.

factor 1.4, while the remaining 7 vehicles present variations of the NO_x emission factor smaller than 30%, which could be explained by the normal real-world variability among the tests (traffic variations in particular). Therefore, the overall variations in emission factors can be significantly large. Figures B.3 to B.36 in Appendix B show that for the ten vehicles presenting significant variations in the emission factor, the successive tests present significantly different time-series, while for the seven other vehicles, significant differences are not noticeable at first glance. However, even for those seven vehicles, it can be observed from Table B.2 in Appendix B that the tests classified in the same cluster present more similar mean, standard deviation and coefficient of variation than tests of the same vehicle classified in different clusters.

Table B.1 shows that the ambient temperature does not seem to be the explanation for such discrepancies between successive tests of the same vehicle as it does not vary among tests of the same vehicle. Figure 4.3 shows that variations in percentage of idling (all the other driving dynamic metrics remaining constant) do not present a systematic effect, even for vehicles from the same manufacturer (vehicle 4 presents a lower emission factor for a greater percentage of idling while vehicle 10 presents the exact opposite effect). With the exception of one vehicle (vehicle 14), tests of the same vehicle presenting a lower RPA also present lower emission factors (vehicles 3, 6 and 15). However, a similar trend with $v \cdot a_{\text{pos}}[95]$ cannot be found (higher $v \cdot a_{\text{pos}}[95]$ does not always result in higher emission factors). Previous studies found contradictory outcomes on the effect of driving dynamics: while [Gallus et al., 2017] presented results indicating that aggressive driving leads to an increase in NO_x emissions, [Franco et al., 2014] showed that it is difficult to link an increase in NO_x emissions to driving dynamicity. The data presented in Figure 4.3 shows clearly that no single dynamic metrics is able to explain the variation in the measured NO_x emissions factor. The current dataset does not allow us to isolate with certainty a parameter able to explain these intra-vehicle real-world driving emitting variations and further studies should be conducted in order to understand the underlying phenomena.

In conclusion, the performed clustering analysis enables the grouping together of vehicles presenting similar emitting behaviours. However, it is not possible to find a simple relationship that is able to explain the clusters based only on physical parameters (vehicle weight, engine

size, etc.), even though some patterns can be identified. Manufacturer calibration strategies are likely to be able to explain a significant part of the observed variations. Interestingly, a vehicle does not always present the same emitting behaviour, even on the same road. Driving dynamics seem to influence emissions but this is unlikely to be the only factor and further work is necessary to obtain a better understanding of the underlying phenomena.

4.3.2 Emission maps for the LT, the NLR and the MLP models

Figure 4.4 displays a comparison of the emission maps as a function of vehicle speed and acceleration in urban driving conditions for the three studied models. The emission maps for the three models are an average over 400 bootstraps. It should be noticed that the range of speed and acceleration displayed in Figure 4.4 is significantly larger than typical driving conditions found in the dataset. The white dashed lines represent the fifth and ninety-fifth percentiles of the speed and acceleration distribution in the available dataset (this range is hereafter referred to as 90% dataset range). Predictions outside this range are depicted for comparison purposes but should not be considered as robust or accurate.

It can be observed from Figure 4.4 that the emission map of the LT_1 model, which is a simple aggregation of the individual datapoints, presents a relatively rough profile characterised by sharp transitions from high to low emissions at relatively similar speed and acceleration. These sharp transitions are due to a lack of data in some bins and although some variability inherent to the vehicle or driving style might occasionally lead to such discrepancies, it is not representative of real world driving emissions (e.g. emission maps from [O’Driscoll et al., 2016]). All the three models highlight that high emissions occur mainly during acceleration events, which is consistent with the literature [Gallus et al., 2017]. The emission maps of the NLR_1 model and the MLP_1 model present extremely similar patterns for all the six main clusters, especially in the 90% dataset range. Once trained, the MLP_1 model tends to converge to a formulation that is almost identical to the NLR_1 model. The general accuracy of the 3 models was computed as the weighted average of the joint probability distribution of speed and acceleration:

$$\bar{e} = \int_{a_{\min}}^{a_{\max}} \int_{v_{\min}}^{v_{\max}} e f_{v,a} dv da \quad (4.4)$$

where $f_{v,a}$ is the joint probability distribution of speed and acceleration and e for each cluster is defined as:

$$e = \frac{\sigma_E}{\bar{E}} \quad (4.5)$$

where σ_E is the standard deviation for each bin over the 400 bootstraps and $\bar{E}$ is the weighted average emission for each bin over the 400 bootstraps, over all clusters, based on trips in each cluster.

For the LT_1 model, $\bar{e}$ is equal to 0.80, while for the MLP_1 and NLR_1 model, it is equal to 0.45 and 0.39 respectively. Not surprisingly, $\bar{e}$ is the highest for the LT model that shows the poorest performances and the highest standard deviation per bin, which can be explained by the lack of data in some bins. $\bar{e}$ is slightly higher for the MLP model compared to the NLR model, which can be explained by the nature of both models: the NLR model is deterministic and therefore converges almost systematically towards the same coefficients, whereas, to prevent the overfitting of the neural network, a dropout rate of 0.1 was used, introducing some stochasticity to each MLP training run, which increases slightly the final standard deviation and therefore $\bar{e}$.

The near equivalence between the NLR_1 and MLP_1 models suggests that with the limited set of inputs available (instantaneous vehicle speed and acceleration) a better representation cannot be achieved. If only instantaneous speed and acceleration are available and if it is possible to assign the vehicle to one cluster, we recommend therefore opting for simplicity and using the NLR_1 model rather than a neural network. Although further work is necessary to confirm this, neural networks are likely to be an appealing tool if a more detailed set of inputs is available (e.g. engine load, engine speed).

For an analysis of the convergence of the MLP_1 model, see Section B.6.5 in Appendix B. It is shown that for all the pixels within the 90% dataset range, the MLP_1 model converges to a unique solution (standard deviation almost equal to zero) within the 90% dataset range. Outside this range, the lack of data makes the model less stable, due to difficulties to converge to a single result.

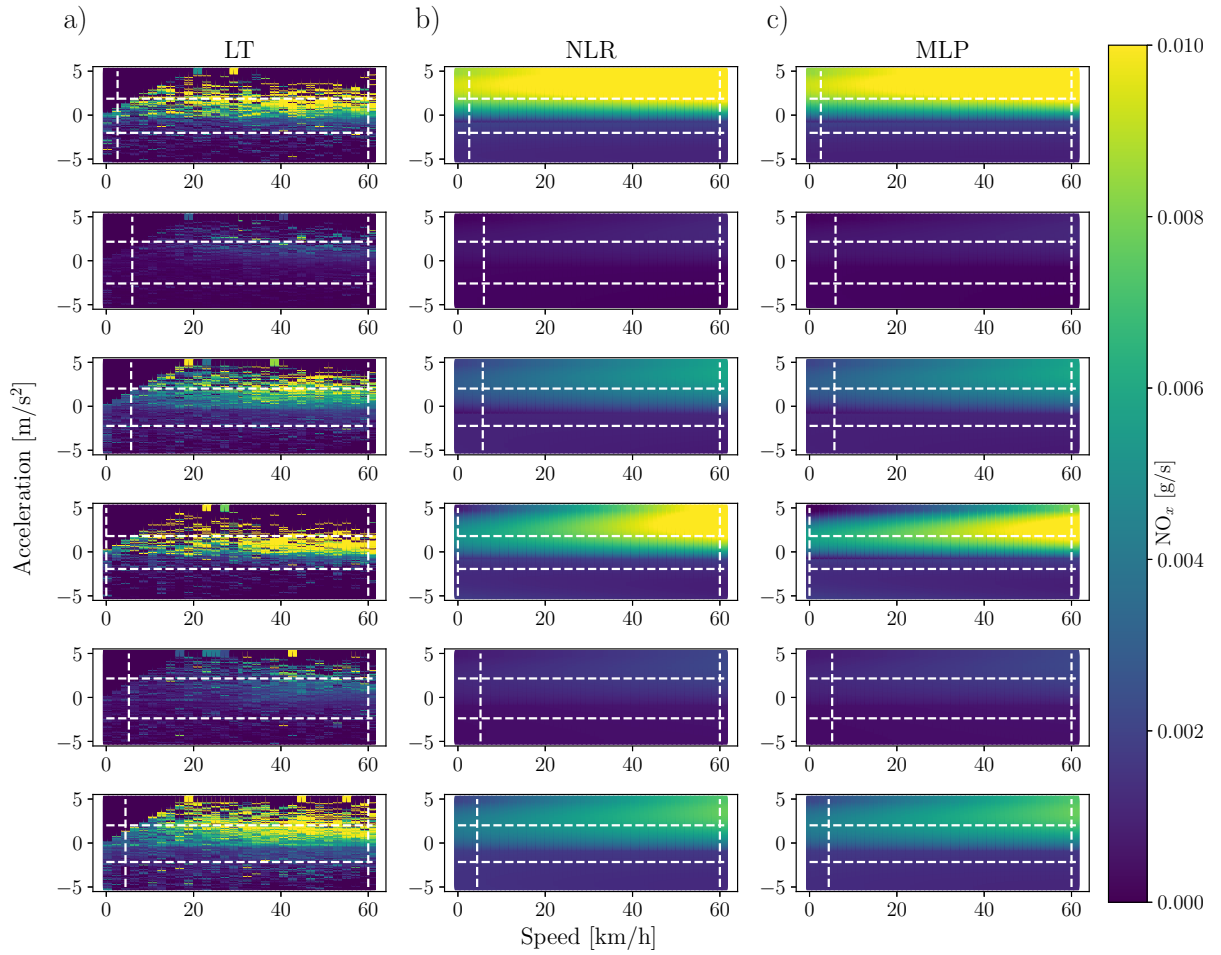


Figure 4.4: Comparison of the emission maps for the main clusters for the three models; (a) LT_1 , (b) NLR_1 and (c) MLP_1 models. White dash lines represent the 5th and 95th percentiles. Clusters are presented in numerical order (1, 4, 5, 8, 10, 11) from the top to the bottom of the graph.

4.3.3 Comparison of models performance against the dataset

The general predictive skills of each of the three instantaneous emissions models are assessed in Figure 4.5 through the comparison of the measured and predicted emission factors for the previously defined 3 x 20 unseen tests. Although it can be argued that the emission factor does not fully represent the emitting behaviour since it does not represent the variability in instantaneous emissions, it is still a good indicator of the average performance of a model. Results for the generalised models (LT_2 , NLR_2 , MLP_2) are depicted in green, while results for the calibrated models (LT_1 , NLR_1 , MLP_1) are depicted in magenta.

It can be observed in Figure 4.5 that the LT_1 model (Figure 4.5.a) systematically underestimates the emission factors by more than 20%, which is significant. In particular, it does not appear capable of predicting emissions factors over 0.2 g/km, thereby severely underpredicting the emissions of the worst polluters. On the other hand, the NLR_1 model (Figure 4.5.b) and the MLP_1 model (Figure 4.5.c) show similar performance with relative errors below 20% (with one exception). Considering that only the vehicle speed and acceleration measured by an embedded GPS were used as inputs to the models, this performance can be considered as satisfactory. However, it is extremely likely that improved performance would be achieved with a wider range of inputs (e.g. engine speed). One interesting feature to notice is the ability of the NLR_1 model to capture the emission factor, implying that meaningfully grouping vehicles together allows surprisingly simple emission models to be used. For accuracy reasons, we recommend to always prefer the NLR_1 and the MLP_1 models to the LT_1 model.

It is also clearly evident in Figure 4.5.a that the LT_2 and the NLR_2 models are unable to correctly predict the emission factors for vehicles not in one of the main clusters. In fact, all the tests presenting a relative error significantly larger than 20% in Figure 4.5.b do not belong to any of the main clusters and the emission factors were extrapolated as presented in Equation 4.3. This reflects the fact that a simple inverse distance weighting function does not allow the generalisation of the LT_1 and the NLR_1 models. It can therefore be concluded that the relationship between the models of the different clusters is non-linear and cannot be easily inferred. If a test does not fall in any of the main clusters, the development of an individual

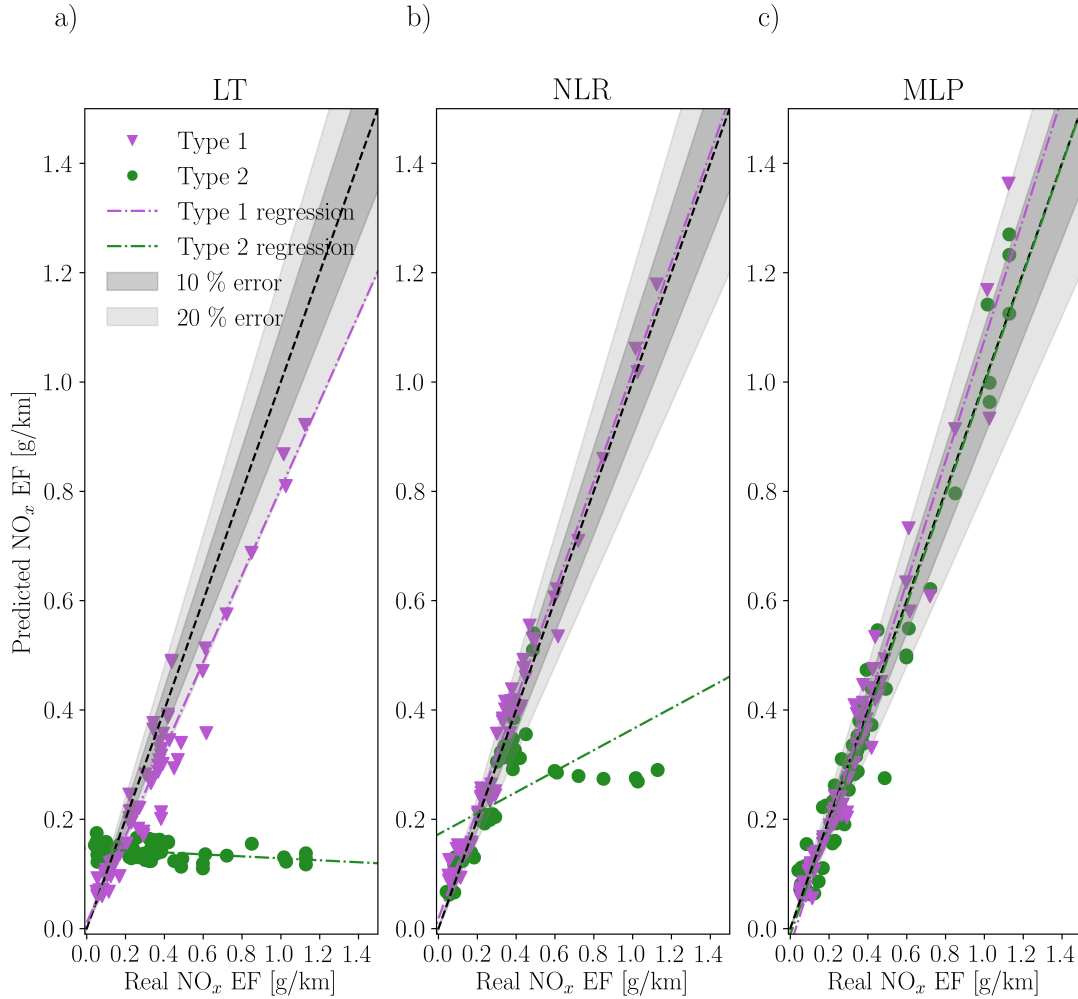


Figure 4.5: Comparison of the measured and predicted emission factors for three instantaneous emission models, (a) a look-up table, (b) the non-linear regression model and (c) the MLP developed in this study. The results for the LT_1 , NLR_1 and MLP_1 models (or clustered models) are depicted with magenta dots while the results for the LT_2 , NLR_2 and MLP_2 models (or general models) are depicted with green triangles. The black dashed line represents the identity line.

model should be preferred. In contrast to the LT_2 and the NLR_2 models, the MLP_2 model, depicted in green in Figure 4.5.c performs well in predicting the emission factor, within the range of a 20% relative error. This demonstrates that an MLP model, given the distances to the centroids of the main clusters in addition to the instantaneous speed and acceleration, is able to build a complex non-linear generalised model, at the cost of a minimal decrease in accuracy compared to a model directly calibrated for each cluster.

The good performance of the NLR_1 , MLP_1 and MLP_2 as well as the poor performance of the LT_1 , LT_2 and NLR_2 models described in Figure 4.5 are also reflected in Figure 4.6, which shows boxplots of the measured and predicted emissions factors, fractional bias (FB) and Normalised Mean Squared Error (NMSE). It can be observed in Figure 4.6.a that the LT_1 and LT_2 models present relatively high fractional bias and normalised mean squared errors, reflecting not only the systematic underestimation of the average emissions but also the inability to capture peaks in emission. The NLR_1 and NLR_2 models achieve a higher level of accuracy compared to the LT models. Figure 4.6.b shows that the fractional bias is close to zero for both cases, demonstrating that the NLR_1 and NLR_2 models are not systematically biased. The higher fractional bias of the NLR_2 model reflects its inability to predict emissions for tests that do not belong to any of the main clusters. Although lower than for the LT_1 and LT_2 models, the normalised mean squared error is around 0.3, which shows the difficulty of capturing the peaks in emission. The MLP_1 and MLP_2 models perform more accurately compared to the LT_1 and LT_2 models. Both the MLP_1 and the MLP_2 models perform well; they do not show any systematic error (as reflected by the fractional bias) but are not able to capture the peaks in emission (as reflected by the normalised mean squared error). None of the models are able to accurately capture all the peaks in emissions. However, it should be noted that those peaks are extremely challenging to capture even when variables like engine speed, gear ratio or EGR rate are known. Vehicle speed and acceleration do not characterise those peaks sufficiently to allow them to be predicted accurately. Nevertheless, these peaks generally account for only a small fraction of the total emissions and although they can be important locally, we consider that not capturing these features represents only a minor drawback. The robustness of the results through the runs is assessed in Table B.3 in Appendix B. Results are almost identical through

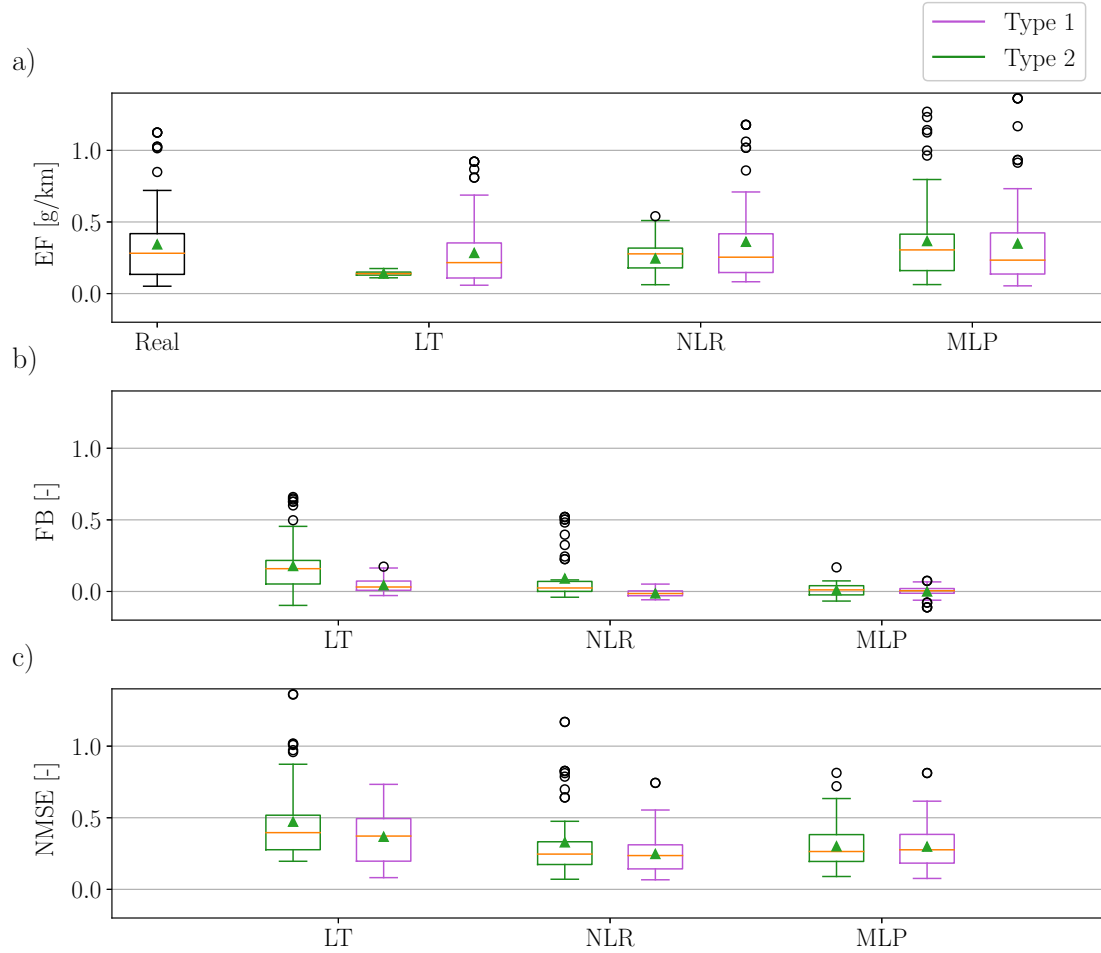


Figure 4.6: Boxplots of the measured and predicted emission factors (a), boxplots of the Fractional Bias (b) and boxplots of the Normalised Mean Squared Error (c) for the three studied instantaneous models. The results for the LT_1 , NLR_1 and MLP_1 models (or clustered models) are depicted in magenta while the results for the LT_2 , NLR_2 and MLP_2 models (or general models) are depicted in green. The median of the data is represented by an orange line while the mean is depicted with a green triangle.

the runs and can therefore be considered as robust.

4.3.4 Case study and comparison with COPERT5

Finally, the performance of the NLR_1 and MLP_1 models were assessed on a case study, and compared to the COPERT5 model. The evaluation of the three models was conducted on the previously defined 3 x 20 unseen tests to eliminate any potential bias in favour of the NLR_1 and MLP_1 models that would result from using the same dataset to develop and test the model. The NLR_1 and MLP_1 models were applied as described in sections 4.2.3 and 4.2.4 to compute the instantaneous NO_x emissions. COPERT5, which is widely used in European countries, is an emission model estimating emissions based on the average-speed along a link. COPERT5 performance increases as road link increases above 400 m [Samaras et al., 2014]. In this analysis, road links were taken as uniform 1 km lengths, similarly to [O'Driscoll et al., 2016]. The average speed at each point was computed as the rolling average speed of the 1 km link centred on that specific data point. We note that this is not the usual way to use COPERT5. However, this is motivated by the wish to compare fairly this average-speed model with the two instantaneous emission models adapted or developed for this study. The first and last 500 meters of each test route were therefore discarded. The specific NO_x emission factor for each data point was computed as described in the [National Atmospheric Emissions Inventory, 2020]. The instantaneous emissions were finally computed by multiplying the obtained emission factor by the travelled distance. The measured and estimated instantaneous emissions for the 3 x 20 tests were aggregated at a 10 meter level, based on the computed travelled distance. In other words, all the measured or predicted points between 0 and 10, 10 and 20, ..., meters in the direction of travel were aggregated together.

Figures 4.7 and 4.8 present a comparison of the NLR_1 , the MLP_1 and COPERT5 models predicted instantaneous emissions aggregated at a 10 meters level for all the 3 x 20 unseen tests, as described earlier. It can be observed in Figures 4.7 and 4.8 that the agreement between the measurements and the predictions from both the NLR_1 model and the MLP_1 model is good overall. Both models give very similar predictions, which confirms results presented in Section

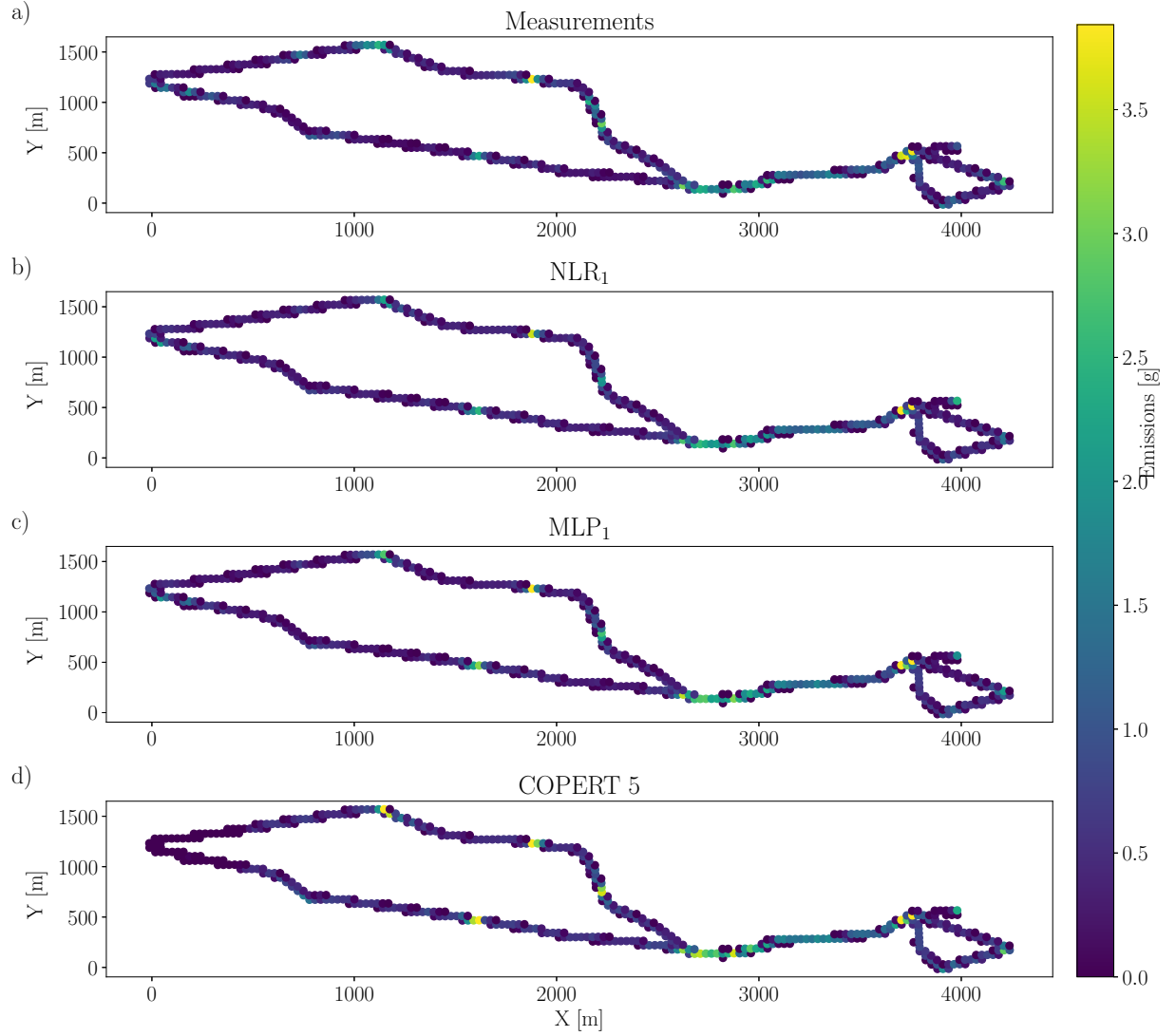


Figure 4.7: Aggregated emission maps of the unseen tests: (a) are the measurements, (b) the NLR_1 model, (c) the MLP_1 model and (d) the COPERT 5 model.

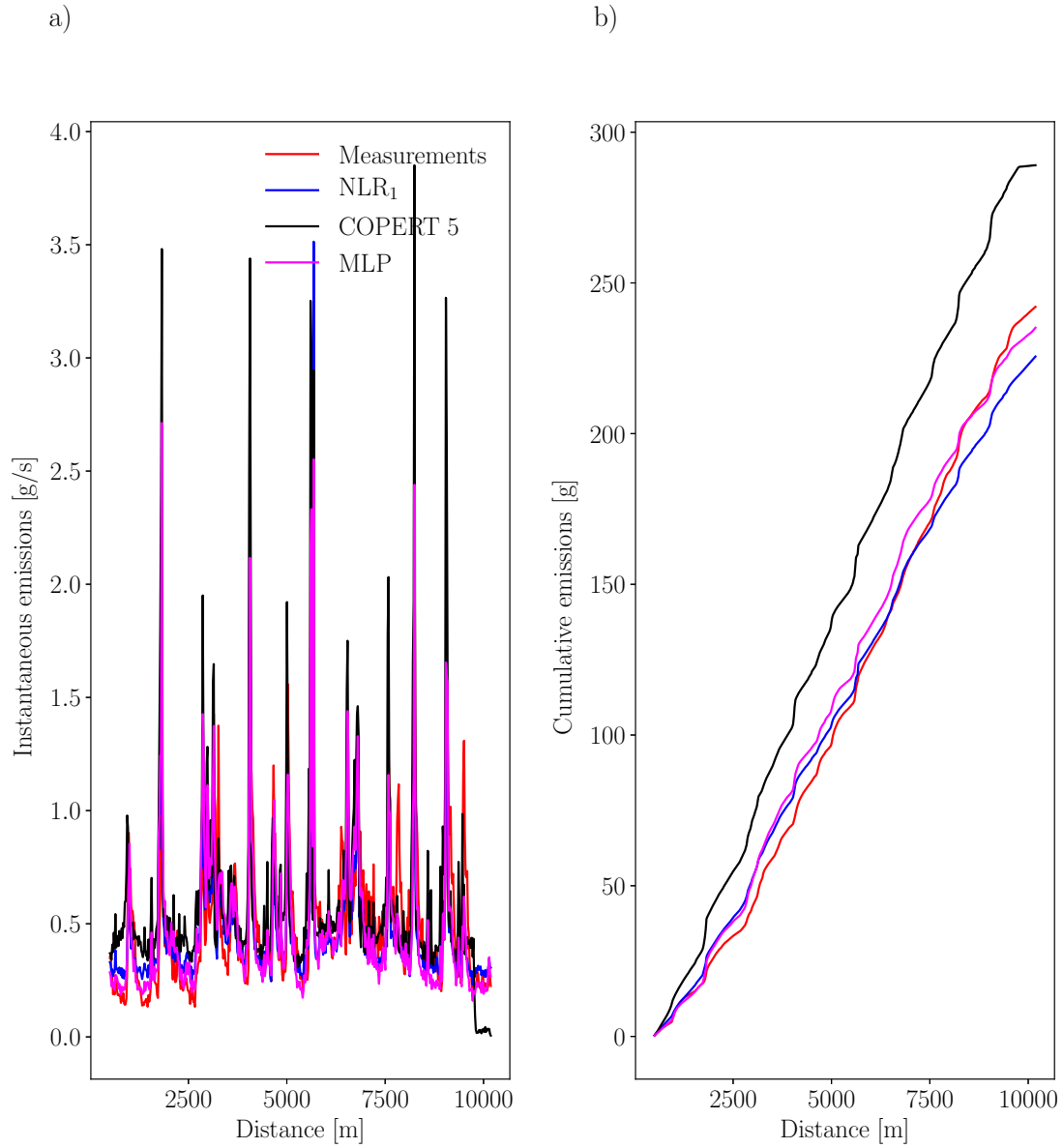


Figure 4.8: Aggregated measured and modelled emissions as the function of the distance travelled: (a) are the instantaneous emissions and (b) are the cumulative emissions.

Table 4.3: Mean and standard deviation of the FB and NMSE statistical indices for the 60 unseen tests used for the case study

	NLR_1	MLP_1	COPERT5
EF error [%]	-1.90 (s.d. 19.13)	-2.49 (s.d. 15.13)	27.26 (s.d.) 91.59
FB [-]	0.01 (s.d. 0.03)	0.007 (s.d. 0.03)	0.05 (s.d. 0.11)
NMSE [-]	0.27 (s.d. 0.13)	0.31 (s.d. 0.14)	0.38 (s.d. 0.14)

4.3.2. Although the NLR_1 model misses some of the peaks in emission, the overall trends are well modelled and the relative error in total emission prediction is only of -6.8%. The MLP_1 model also modelled very well the overall trends and underestimates only by 2.9% the total NO_x emissions. The COPERT5 model predictions also show a satisfactory agreement with measurements, despite presenting larger residuals than the NLR_1 and the MLP_1 models (see Section B.6.7 in B). A closer inspection at Figures 4.7 and 4.8 also shows the lower accuracy presented by the COPERT5 model compared to the MLP_1 and the NLR_1 models. Overall, the COPERT5 model overestimates the total emissions by 19.4%.

It is worth emphasising that the use of a rolling average to compute the average speed used to evaluate the emissions is not performed during a typical application of COPERT5 but might represent an improvement of the model at an aggregate level. However, as reflected in Table 4.3, which presents the mean and standard deviation of the relative error in emission factor prediction, the FB and NMSE over the 3 x 20 unseen tests, the individual instantaneous NO_x emissions predicted by the COPERT5 model clearly presents unsatisfactory results compared to the NLR_1 and the MLP_1 models. COPERT5 should therefore not be used to evaluate the performance of a specific vehicle but can provide a reasonable estimate of the total emissions of a fleet over a portion of the network, which corresponds to its normal usage.

4.4 Conclusion

In this study, machine learning techniques were applied to the large dataset of PEMS measurements in urban driving conditions described in [O’Driscoll et al., 2016]. A clustering analysis yielded (i) an update of existing emission models (a look-up table model and a non-linear regression) and (ii) the development of a new emission model (a multi-layer perceptron). A

comparison between the updated emission models and the new emission model is provided.

The clustering analysis enabled the grouping of vehicles presenting similar emitting behaviours and allowed us to develop accurate instantaneous NO_x emissions models that sit in-between vehicle-specific and average-fleet models in terms of complexity. Analysis of the clusters reveals that: (i) no obvious internal variable (such as engine size, vehicle weight, vehicle manufacturer or NO_x aftertreatment technology) or external variable (external temperature or driving dynamic) is able to explain the clustering and could be used to predict, *a priori*, the cluster in which a specific non tested vehicle would fall; (ii) different tests of the same vehicle along the same test route by a trained driver can be assigned to different clusters. When the discrepancies in emission factor between the repeated tests classified in different clusters (7 out of 17 vehicles) was smaller than 20-30%, no fundamental difference between the time-series was noticed and real-world variability constitutes the main explanation for these differences. Higher discrepancies in emission factors were correlated with visible differences, localised on specific parts of the tests, between the time-series of the repeated tests.

Analysis and comparison of the updated emission models with the new MLP model shows that: (i) the clustering works well in producing instantaneous NO_x emissions models with a reduced complexity compared to vehicle-specific models; (ii) a relatively limited set of inputs (vehicle instantaneous speed and acceleration) can accurately predict NO_x emissions, even though the peaks in emissions can be missed; (iii) the NLR_1 and the MLP_1 models present similar performance, and given that the NLR_1 model is less computationally expensive, we recommend it as a preferred choice (coefficients are given in the SI); (iv) for non-clustered vehicles, it is either needed to calibrate a specific model as a linear interpolation from the other clusters was proven to not work or to use directly the MLP_2 model. If a more detailed set of inputs is available (i.e. engine speed, EGR rate), neural networks are likely to present better performances than a simple non-linear regression and are worth exploring. We emphasise that a model is only valid under the conditions for which it was designed and tested, which means that extrapolating outside this range will likely provide wrong or biased estimates. The main challenge in operational use of the presented models is their reliance on a prior clustering analysis. Future work should focus on resolving this issue and is likely to require evidence and

knowledge of emissions characteristics of engine families shared across manufacturers. In the future, the outputs of RDE testing, which aims to better represent real-world driving conditions, could be used for to adapt the models.

Advanced emissions models that are accurate and fast enough to be used in real-time will enable a range of intelligent transport systems services including real-time traffic management and signal control optimisation to minimise emissions. This study is the first to demonstrate feasibility and apply machine learning techniques to a large dataset of real-world vehicle NO_x emissions. The new models that we presented rely on simple inputs of vehicle speed and acceleration, which could be extracted from existing sources including traffic cameras and vehicle tracking devices and therefore deployed immediately.

Chapter 5

Impact of a traffic intervention on the local air quality

This chapter presents an analysis of the real-world impact of a traffic intervention on both the local emissions and air quality. It relies upon data collected during a two-weeks field study on a busy junction on London Road (Southwark, London, United Kingdom) between the 16th and the 27th of September 2019. This chapter investigates with traffic microsimulations, emissions modelling and operational air dispersion simulation how two distinct traffic interventions impact on the traffic conditions, on the vehicle NO_x emissions and on the pedestrian average NO₂ concentrations. These results are compared to the measured NO₂ concentrations, and the impact of the traffic interventions on the NO₂ concentrations is assessed separately from the meteorological conditions.

This chapter has not been submitted yet for journal publication, but a paper is current in preparation for submission, based on the results presented hereafter.

The analysis of the LIDAR data, which is mentioned for the sake of completeness, has been conducted by Matthew Williams and Prof. Janet Barlow from the University of Reading (UK). The analysis of the camera data has been conducted by Dr. Anna Schroeder from the University of Cambridge (UK). The analysis of the street anemometer data has been conducted by Dr. Shiwei Fang from the University of Cambridge (UK).

This chapter has the following structure. As an introduction, Section 5.1 presents the existing research gaps and the objectives of this study. Section 5.2.1 describes the field campaign while Section 5.2.2 presents any preprocessing performed prior to this study. Section 5.2.4 details the successive steps and assumptions made to simulate the impact of both traffic interventions on local air quality. Section 5.3 first discusses the variability in the measured pollutant concentrations as a function of the meteorological conditions, the location of the sensors and the traffic interventions. It then analyses the impact of both traffic interventions on the traffic conditions and NO_x emissions. A comparison of the REVS model with respect to the NLR model is conducted in Section 5.3.3. Section 5.3.6 presents and discusses the outcomes of the simulations with respect to the measured NO_2 concentrations. Finally, Section 5.4 summarises the main findings and discusses the implications of the presented work.

5.1 Introduction

The effectiveness of an intervention that aims to improve air quality is usually evaluated by modelling studies prior to its implementation. Among the multiple interventions that have been explored, optimisation of the traffic lights, which are known to heavily impact both emissions and congestion, is of particular interest. Traffic lights disrupt the traffic flow, causing an increase in vehicle driving dynamicity [Zhang et al., 2013] that can raise emissions to a detrimental level [Unal et al., 2003, Peters et al., 2009, He and Hou, 2012, Madireddy et al., 2011]. The increase in computational power has allowed the exploration of the potential of swarm optimisation, genetic algorithms and reinforcement learning algorithms for traffic control systems optimisation accounting for emissions [Olivera et al., 2015, Zhang et al., 2013, Péres et al., 2018], despite most studies being primarily concerned with congestion. Although reporting very promising results, all the aforementioned studies focused on idealised or simplified junctions, did not monitor the evolution of the spatial and temporal distribution in emissions and typically focus on demonstrating that the planned intervention will lead to a decrease in the strength of emission sources, at a reasonable cost in terms of congestion. Although any reduction in emissions is desirable in an air quality perspective, the intrinsic relationship between

meteorological conditions and air quality [Monks et al., 2009] signifies that a local reduction in emissions might not have the expected outcome in terms of air quality.

It is extremely hard to detect and quantify the impact of an intervention on the ambient air quality, which is known to be primarily affected by the meteorological conditions [Monks et al., 2009]. Studies focusing on finding trends in ambient pollutant concentrations following an intervention emphasise the importance of controlling for the meteorology so trend analysis can be assessed robustly [Grange and Carslaw, 2019]. This can be achieved by applying meteorological normalisation, a technique to estimate the variability in concentrations induced by meteorology in air quality time series. It consists of training, on historical data, a statistical model that is able to predict variation in pollutant concentrations as a function of parameters representing both the meteorological conditions and the regular emissions patterns. Such a model can then be used to remove the influence of weather conditions on the variation of the pollutant concentrations [Grange and Carslaw, 2019]. Meteorological normalisation was successfully used to detect and quantify the impact of large scale interventions ranging from the decrease in the sulphur content of shipping fuel, to the introduction of the congestion charge in London or to the COVID19 pandemic [Grange and Carslaw, 2019, Wyche et al., 2021]. The main drawback of meteorological normalisation is the requirement for long-term historical time series that might not be available, especially when the scale of interest is very local (i.e. street).

Most studies aiming to quantify the impact of an intervention on air quality are either concerned by a change in emissions or by a change in air quality. However, it appears crucial to jointly consider emissions, meteorological conditions and air quality when evaluating the impact of any intervention aiming to improve air quality. To the best of our knowledge, no studies have combined traffic simulations, emission modelling and real-world pollutant concentration measurements from various sensors to this aim. One particular challenge consists in evaluating the impact of the traffic interventions on air quality independently from the variability in pollutant concentrations induced by the meteorological conditions. Finally, the data collected during the field study, jointly with simulations, also represent a unique opportunity to evaluate on a real case study the performance of the NLR and the REVS models presented in earlier chapters. We aim to address these research gaps in the next Sections of this Chapter.

5.2 Methods

This section firstly gives an overview of the field study area and provides a description of the traffic interventions carried out in Section 5.2.1. The methodology to process the raw sensor data is described in Section 5.2.2, while Section 5.2.3 describes the meteorological conditions encountered during the two weeks of the field study. Finally, Section 5.2.4 provides an overview of the methodology and assumptions used to set up the complete simulation of the two traffic interventions. The inputs of the traffic microsimulations used to describe the traffic flow are described in Section 5.2.4.1, Section 5.2.4.2 then describes and discusses how emissions were modelled and finally Section 5.2.4.3 provides a detailed description of the parameters used to calibrate the operational air dispersion simulation.

5.2.1 Field study description

The field study was conducted on London Road, a busy street connecting St Georges Circus to Elephant and Castle in Southwark, London, United Kingdom. The area of interest was specifically situated at the junction between London Road and Garden Row, a one-way street accessible from London Road. London Road is a three-lanes street, with one northbound lane going from Elephant and Castle towards St Georges Circus and two southbound lanes going from St Georges Circus towards Elephant and Castle. The northbound lane is exclusively reserved to buses and taxis, while the two southbound lanes do not have any vehicle type restrictions. Vehicles on the southbound lanes can access Garden Row by turning to their right at the junction. Two pedestrian crossings are situated at the studied junction, one to cross London Road and the other to cross Garden Row. The junction is controlled by a complex traffic light management system that regulates the conflicting usage of the junction by the different users. Measurements were conducted Monday to Friday, from 12 am to 4 pm, between the 16th and the 27th of September 2019. A map depicting the area and pointing out the locations of all the instruments is presented in Figure 5.1.

Two traffic interventions were conducted, during which the traffic light cycle of the junction

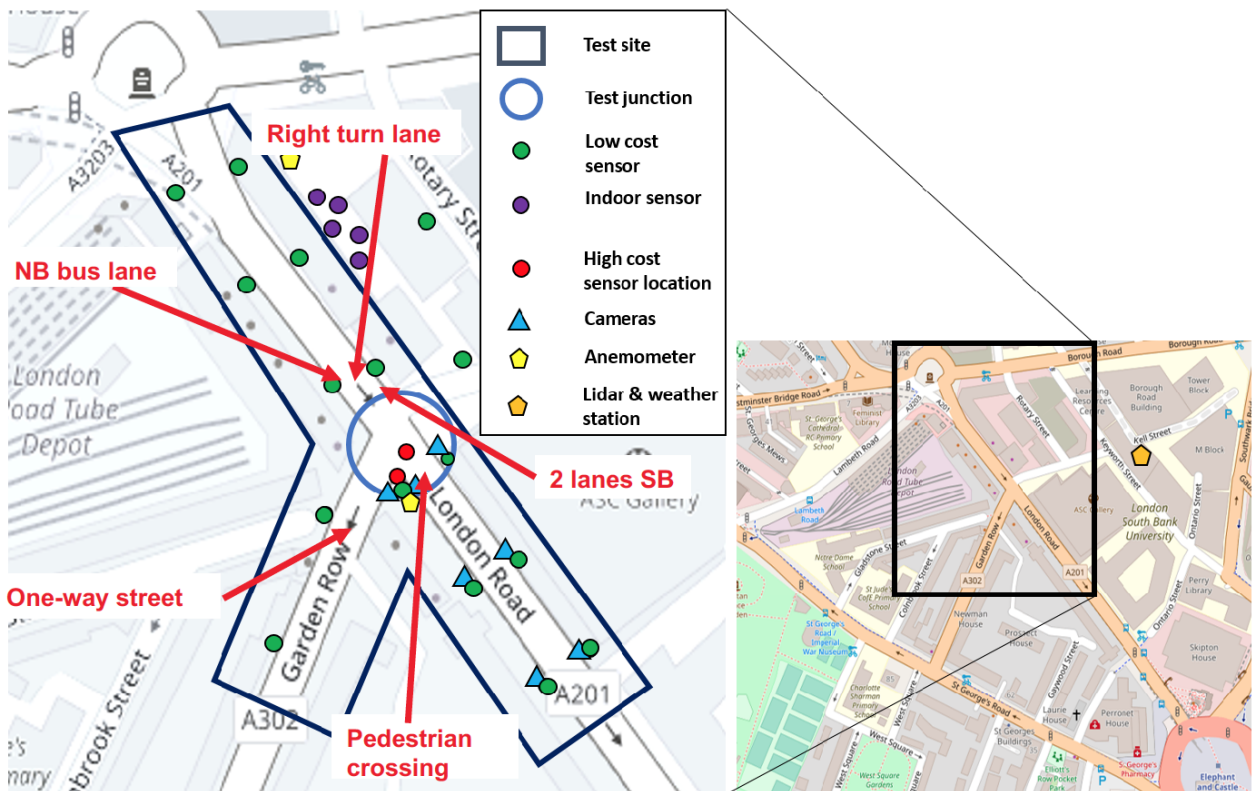


Figure 5.1: Map of the field study area. The emplacement of sensors is depicted with different markers. Courtesy: Dr. Anna Schroeder, University of Cambridge.

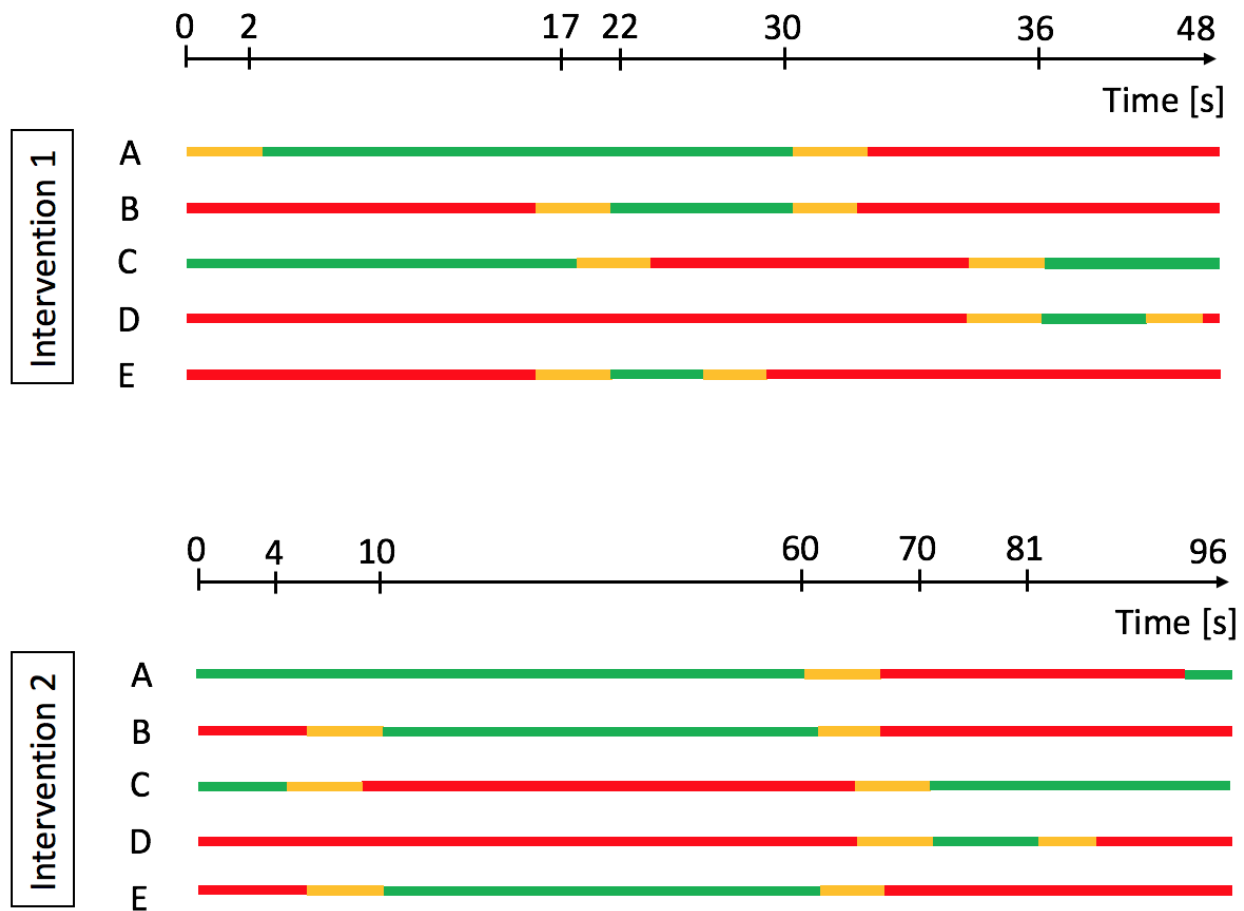


Figure 5.2: Description of the traffic light timings for the two interventions. A is the phase allocated to the Southbound lane, B is the phase allocated to the Northbound lane, C is the phase allocated to the right turning lane, D is the phase allocated to pedestrian to cross London Road and E is the phase allocated to pedestrians to cross Garden Row.

was modified. The first traffic intervention consisted in a cycle of 48 seconds (hereafter denoted as the A-48 intervention), while the second intervention consisted in a cycle of 96 seconds (hereafter denoted as the B-96 intervention) that aimed to create a green way (high proportion of green time) for the buses on the Northbound lane. The first traffic intervention (A-48) was implemented on even days while the second intervention (B-96) was implemented on odd days. Both interventions are compared in Figure 5.2.

5.2.2 Instrumentation

A network of low cost sensors and traffic monitoring cameras were installed on the street lampposts, where they stayed for the entire two weeks of the field study, their batteries being regularly replaced. The network of low cost sensors was not exploited in this study and is therefore not mentioned further. The other sensors were set up every day, and were placed in two different locations: on a central island situated in the middle of the pedestrian crossing on London Road, the sensors inlets facing the southbound lane and next to the northbound lane facing the bus lane. In this study, we focus primarily on NO_x emissions and ambient NO_2 concentrations, and data recorded by sensors other than the CAPS (located on the central island, hereafter denoted CI) and the M200E (located next to the bus lane hereafter denoted BL) is not exploited here. A summary of the main technical specifications for both sensors is displayed in Section C.1 in Appendix C.

Sensors were covered with a tarpaulin during rain episodes to avoid damaging them. Any data recorded during these rain episodes was consequently not reliable and dismissed for the analysis. Regular checks of the drift of the M200E (BL) requiring to disable for a few minutes the instrument were conducted during the measurement campaign, and we consequently dismissed these checks from the analysis. Finally, the M200E (BL) data was completely lost on the 17th of September.

At the time of the writing, analysis of the camera data was still on-going and only traffic counts could be used. A summary of the traffic total vehicle counts, split per day and per vehicle type, is shown in Table 5.1 (courtesy of Dr. Anna Schroeder). At the time of the writing, the accuracy of the traffic count algorithm is estimated to be $\pm 10\%$ on the northbound lane and the southbound lane, and $\pm 25\%$ on the right-turn lane (Camera A, personal communication from Dr. Anna Schroeder). Since the maximum difference in traffic between the days is about 7%, it can therefore be concluded from Table 5.1 that the traffic is globally invariant during the 10 days of field study.

High resolution weather stations measuring the background concentration was not available

Table 5.1: Cameras traffic counts for the two weeks of field study between 12 am and 4 pm. Data for the 17th and the 25th of September have been largely lost and are therefore not exploitable to get reliable traffic counts.

	NBL - bus	SBL - cars	SBL - HGVs	SBL - bus	Total
16	306.5	1235.5	402.0	550.5	2537.0
17	NaN				
18	282.5	1147.0	447.0	555.5	2432.0
19	294.0	1230.5	438.0	574.5	2537.0
20	285.5	1210.5	396.5	550.0	2442.5
23	291.0	1132.0	377.5	553.5	2354.0
24	287.5	1248.0	363.5	475.0	2374.0
25	NaN				
26	288.0	1255.5	335.5	499.5	2378.5
27	277.5	1256.5	365	538.5	2437.5

in the immediate area. The closest available London Air Quality Network station measuring background concentration and is situated in Elephant and Castle. Unfortunately, such measurements are available only at a resolution of 15 minutes. We therefore estimated the background concentration using the methodology described in [Popoola et al., 2018]: the background concentration was taken as the 5th percentile of the measured NO₂ concentration on rolling window of 5 minutes for the CAPS (CI) or as the minimum measured NO₂ concentration on rolling window of 5 minutes for the M200E (BL).

Finally, to facilitate the comparison between the data recorded by the M200E (BL) (1 minute resolution) and the CAPS (CI) (1 second resolution), the resolution of the CAPS (CI) was degraded to 1 minute. The wind speed and wind direction data available from the anemometer was available at an initial resolution of 4 Hz, but due to concerns about the accuracy of such measurements (courtesy Dr. Shiwei Fang), it was decided to aggregate them to a 1 minute time step.

5.2.3 Meteorological conditions

Analysis of the LIDAR data was conducted by Matthew Williams and Prof. Janet Barlow from the University of Reading (personal communication). Their results are used to support

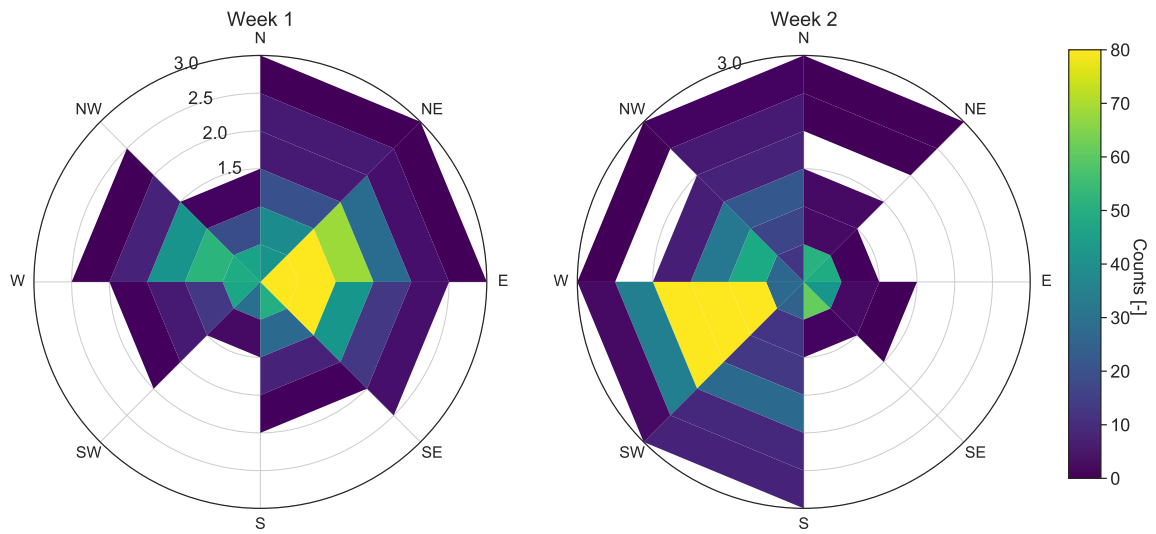


Figure 5.3: Comparison of the distribution in wind speed and wind direction for the two weeks of field study measured by the anemometer located on London Road.

our analysis.

Analysis of the LIDAR data revealed that the meteorological conditions during the two weeks of the field study varied tremendously: in the first week (16th to 20th of September 2019) conditions were typically anticyclonic, with Northerly and Easterly winds, while in the second week (23rd to 27th of September 2019) a low-pressure coming from the Atlantic resulted in weather fronts, with Westerly and Southerly winds. Over the two weeks, the background wind flow never aligned with London Road. Winds were stronger during the second week, averaging at 2.62 m/s at the top of the LSBU building, compared to average winds of 1.95 m/s during the first week. The LIDAR anemometer was situated at 30.33 meters above the street level. Comparison conducted versus the street anemometer showed good agreement in both the vertical and horizontal components. The average wind speed and wind direction from 12 am to 4 pm on the days of the field study are available in Table C.2 in Appendix C.

During both weeks, the angle of incidence of the wind resulted in an helical flow in the direction of the wind. On the first week, it lead to an advection of the pollutants from North to South and from West to East, increasing the concentrations measured at the CAPS (CI) compared

to the concentrations measured at the M200E (BL) and advecting any surplus of emissions occurring on the northern section of London Road towards the sensors. On the second week, the helical flow leads to an advection of the pollutants from South to North and from West to East, increasing the concentrations measured at the CAPS (CI) compared to the concentrations measured at the M200E (BL) and advecting the surplus of emissions that is occurring on the southern section of London Road towards the sensors.

5.2.4 Modelling of the field study

The bottom up analysis of the impact of a traffic intervention from the estimation of the local emissions to the analysis of the measured real-world pollutant concentrations is not commonly performed in the literature and we aim to fill this gap. Moreover, it provides the unique opportunity to compare the performance of the two emissions introduced and developed in Chapters 3 and 4 on a real-world case study. These observations provided the motivation to: (i) conduct an analysis of the change in traffic patterns induced by both interventions, (ii) compare the performance of the two emission models introduced in Chapter 3 and 4 and (iii) use an operational air dispersion model to quantify the theoretical impact of both traffic interventions on the local air quality and compare this result to the concentrations effectively measured. We described the methodology adopted to do so in the next Sections.

5.2.4.1 Traffic microsimulation with VISSIM

A representation of the traffic flow on London Road is first necessary to estimate vehicular emissions. The traffic flow characteristics were obtained using a VISSIM traffic microsimulation. VISSIM (version 9.00-12 is used in this study) is a traffic, timestep and behaviour microsimulation model [Bloomberg and Dale, 2000], based on both traffic flow and light signal control models [AG, 2014]. The traffic flow model is derived from a psycho-physical approach which states that the behaviour of each driver can be classified into four different modes (free driving, approaching, following and braking), responding to the perceived relative speed and distance

to the preceding vehicle [Gomes et al., 2004]. The software also accounts for driving styles via a distribution of speed and distance behaviour [AG, 2014]. The road layout for London Road was given by Transport for London (TfL), the fixed traffic lights were implemented as described in Figure 5.2 and manual count records, more precise than the camera counts, conducted over 3 days between 12 am and 2 pm in June 2019 were used to calibrate the model. The simulation was run for 4,500 seconds, and the first 900 seconds were discarded for the analysis. Instantaneous data of location, speed and acceleration was given for every vehicle in the network on a second-by-second basis.

5.2.4.2 Emissions modelling

The two emission models developed in Chapter 3 and 4 were used to estimate the NO_x emissions. Major adjustments of the models were required, including their extension to other vehicles types such as buses, vans and trucks. The NLR model, presented in Chapter 4, was limited to diesel passenger vehicles and an extension to petrol vehicles was therefore needed. For the sake of comparability, the same underlying datasets should be used to train both models, and the DfT, the DVSA and the EA2019 datasets were included in the analysis alongside with the EA2016 dataset.

Description of the buses, vans and trucks datasets

The dataset used to calibrate both models for vans and HGVs was provided by Transport for London (TfL) and was already described in the study from [Moody and Tate, 2017]. It is comprised of 12 Euro 5/V and Euro 6/VI vehicles that were tested at loads comprised between 0 to 100%, for 11 of them on a chassis-dynamometer using the TfL cycle and for 1 of them in real-world driving conditions with PEMS. The TfL cycle aims to mimic London traffic situation, at peak and off peak hours, and can be considered as representative of real-world driving conditions. All the 12 vehicles were tested on both typical peak and off peak traffic cycles. For more details on the data acquisition process, we refer the interested reader to [Moody and Tate, 2017].

The dataset used to calibrate both models for buses was also provided by Transport for London and was already described in the study from [Phantawesak et al., 2021]. It is comprised of 10 Euro V retrofitted TfL buses tested at unknown loads in real-world driving conditions with PEMS. Nine buses were double decked buses and only one was single decked.

Fleet characteristics

London Road is situated within the London Ultra Low Emission Zone. The London ULEZ was introduced on the 8th of April 2019 in Central London (Covent Garden, Bloomsbury, City of London and Southwark) and operates 24 hours a day, every day of the year [GLA, 2020]. Vehicles must meet strict emission standards to drive in the ULEZ area: Euro 4 and above for petrol passenger vehicles and vans, Euro 6 and above for diesel passenger vehicles and vans, Euro VI for lorries, buses and coaches. In January 2020, after the first 10 months of operation, the compliance rate with the ULEZ standard was of 82% for cars, 63.6% for vans, 90.3% for HGVs, 100% for TfL buses and 76.8% for non-TfL buses in a 24 hour period [GLA, 2020]. We applied this respective share to our fleet.

NLR model

The NLR model presented in Chapter 3 was based on clusters defined after a clustering analysis carried out before hand. One main limitation of the approach was the fundamental assumption that all the vehicles were tested on the same route, in similar traffic conditions. However, the datasets collected later on in time (EA2019 dataset) or by other testing facilities (DVSA and DfT datasets) used different testing routes, making the approach unusable. It was therefore decided to train one NLR model specifically for each vehicle. One specific model was randomly assigned to each simulated vehicle in the network with the following characteristics: 1) the final distribution of the vehicles' fuel type and engine size followed the licensed statistics for Great Britain for fuel type and engine size [Department for Transport, 2020] and 2) the final distribution of the vehicles' Euro standards followed the ULEZ characteristics described in the previous Section. The distribution of fuel type and engine size used to attribute a model to a vehicle is displayed in Table 5.2.

Table 5.2: Distribution of fuel type and engine size used to estimate the instantaneous NO_x emission using the NLR model.

	<1000 cc	1001 - 1550 cc	1551 - 2000 cc	2001 - 2500 cc	2501 - 3000 cc	>3001 cc	Total
Petrol	8.8%	28.3%	19.7%	1.2%	1.5%	1.7%	61.4%
Diesel	0%	5.9%	25.9%	5.6%	1.2%	0%	38.6%

Due to a limited size of the buses, trucks and vans datasets, it was not possible to use a realistic distribution of engine size when attributing randomly an individual NLR model to a simulated vehicle. For these categories of vehicles, an individual model was instead randomly selected only based on the ULEZ fleet characteristics. To evaluate the uncertainty due to the random attribution of a model type to a vehicle in the network, the emission calculation procedure was conducted 1,000 times. The resulting distribution in instantaneous NO_x average emissions per street is available in Figure C.1 in Appendix C.

REVS model

The REVS model presented in Chapter 3 is, at the time of the writing, up to date for passenger vehicles and was used as such. Following the methodology to extract trips from real-world driving tests and build the joint probability distribution described in Chapter 3, aggregated models for Euro V and Euro VI HGVs, Euro 5 and Euro 6 vans and retrofitted Euro V buses were built. For HGVs, the load was not considered separately, and the model obtained is thus valid at all load.

Similarly to the NLR model, a model was randomly assigned for each vehicle, based on the ULEZ fleet characteristics. The same proportion of petrol and diesel passenger vehicles as for the NLR model was assumed. The average NO_x emission factor and its standard deviation was estimated for each vehicle technology category using the average speed and average VSP over the network of the vehicles belonging to this specific category. The instantaneous emissions for each individual vehicle can then be simply obtained by multiplying by the instantaneous speed at each time step. In order to get the average instantaneous NO_x emission per street, the total hourly NO_x emissions were calculated and further divided by 3,600. To evaluate the uncertainty due to the random attribution of a model type to a vehicle in the network,

the emission calculation procedure was conducted 1,000 times. The resulting distribution in instantaneous NO_x average emissions and its standard deviation per street is available in Figure C.1 in Appendix C.

5.2.4.3 Operational air dispersion simulation

SIRANE is the operational air dispersion model developed by the Ecole Centrale de Lyon. We refer the interested reader to Section 2.3.2 for a detailed description of the model properties. In this Section, the inputs to the SIRANE models are discussed, and the pedestrian concentrations are retrieved using the one parameter-parametrisation presented in [Grylls et al., 2019].

Description of SIRANE inputs

The representation of the urban morphology was obtained using one metre resolution, vector-based LIDAR data with an applied building mask [Environmental Agency, 2017]. SIRANE does not model explicitly the buildings within the domain but represents the streets as straight boxes, with their beginning and end points at the centre point of the connecting intersections. The average heights and widths on either side of the box centre line were calculated from the height of and respective distance to the surrounding buildings. The network consists only of three boxes. Vehicles emissions are specified as mass flow (in g/s) for every box in the network (see Table C.1). In order to estimate the uncertainty in pollutant concentration resulting from the uncertainty in emissions, three simulations were conducted per traffic intervention and emission model. The first simulation consisted in setting the emissions to their average estimate, the second to their average estimate minus one standard deviation and the third to their average estimate plus one standard deviation.

Background concentrations were obtained directly from the sensors measurements, as described in Section 5.2.2. The averaged NO_2 background concentration recorded by the M200E (BL) and both traffic interventions was selected, with its corresponding NO background concentration. Similarly to SIRANE internal assumptions for the final pollutant concentrations, a

photostationary steady-state was assumed to derive the ozone background concentration:

$$\frac{J_{\text{NO}_2}}{k_3} = \frac{[\text{NO}][\text{O}_3]}{[\text{NO}_2]} \quad (5.1)$$

where J_{NO_2} is the photolysis rate for NO_2 and k_3 is the $\text{NO}+\text{O}_3$ reaction rate. J_{NO_2} (0.0064 s^{-1}) and k_3 ($10041.8 \text{ m}^3/(\text{mole} \cdot \text{s})$) were calculated through the parametrisation used within SIRANE [Soulhac et al., 2011], as a function of the solar elevation and cloud cover and air temperature respectively. The final background concentrations are available in Table C.3.

Typical meteorological conditions of Southwark, London, were applied within SIRANE: an end of summer's day (September, 1 pm), with neutral stability (enforced by setting an albedo of 1), a partial cloud cover, no precipitation. One simulation was run per day of field study, using the average wind speed and wind direction recorded by the LIDAR (see Table C.2).

Pedestrian concentrations

SIRANE is not able to catch the vertical gradients observed in the daily NO_x concentration profile within a typical street canyon [Xie et al., 2003], and a direct comparison of the predicted canyon averaged concentrations with the sensors measurements is therefore not relevant. Nevertheless, the parametrisation developed by [Grylls et al., 2019] can be applied to derive the pedestrian average concentrations (at 1.80 m) from the background and the canyon average concentrations:

$$\phi_{\text{SIRANE-ped}} = \phi_{\text{b}} + m(\phi_{\text{SIRANE-canyon}} - \phi_{\text{b}}) \quad (5.2)$$

where ϕ_{b} is the background concentration, $\phi_{\text{SIRANE-canyon}}$ is the SIRANE canyon average prediction, $\phi_{\text{SIRANE-ped}}$ is the predicted concentration at pedestrian level and m is a correction factor that depends on the specie and the wind direction (1.39 for NO_2).

5.3 Results

The present study was conducted with two main point of focus: the real-world versus the modelled pollutants concentrations. The analysis of the measured NO_2 concentrations includes: (i)

the analysis of the joint impact of the traffic interventions and meteorological conditions on the reported concentrations and (ii) the analysis of the daily variations in NO₂ concentrations. The analysis of the simulated traffic conditions and NO₂ concentrations include: (i) the analysis of the change in traffic patterns induced by both interventions, (ii) a comparison of the performance of the two emission models introduced in Chapter 3 and 4 and (iii) a comparison of the simulated NO₂ concentrations to the NO₂ concentrations effectively measured by the two sensors. We attract the reader attention that all the concentrations are given in ppb, as measured by the sensors, rather than $\mu\text{g}/\text{m}^3$. At 20°C, the conversion factor from ppb to $\mu\text{g}/\text{m}^3$ is of 1.88 for NO₂, 1.25 for NO and 2.00 for O₃.

This Section firstly presents and discusses the NO₂ concentrations measured during the field study with respect to both the traffic interventions and the meteorological conditions. In a second step, it analyses the impact of both traffic interventions on the traffic flow. Section 5.3.3 then describes how the REVS and the NLR model compare to each other in terms of total emissions over the entire network while Section 5.3.4 discusses the spatial variability in emissions. Section 5.3.5 presents the results of the operational air dispersion simulation, and particularly discusses how the predicted pedestrian average NO₂ concentrations relate to the sensors measurements and how the variability in emissions impacts the variability in the predicted pollutant concentrations. Special focus is given to quantifying the impact of both traffic interventions independently from the effects of meteorology.

5.3.1 NO₂ concentrations

This Section examines and discusses the impact of both traffic interventions on the measured NO₂ concentrations. First, the overall impact of both traffic intervention over the two weeks of field study is discussed and a comparison between the two interventions is made. In a second step, daily variation in the recorded NO₂ concentrations are analysed, with a special focus on the effects of meteorological conditions.

5.3.1.1 Average difference between the two interventions

Figure 5.4 displays the distribution of the background NO₂ concentration and the above-background NO₂ concentrations for the CAPS (CI) and the M200E (BL) over the two weeks of field study. The red and the blue dashed lines depict to the average above-background NO₂ concentrations for the A-48 and the B-96 intervention respectively. It can be observed from Figure 5.4 that the B-96 intervention results in concentrations on average lower of 4.99 ppb (respectively 3.67 ppb) for the CAPS (CI) (respectively the M200E (BL)) compared to the A-48 intervention. A two-sided Kolmogorov-Smirnov test conducted on the two-weeks aggregated data rejected the null hypothesis that both distributions were drawn from the same underlying distribution (p-value = $7.9 \cdot 10^{-9}$ for the CAPS (CI) and $1.2 \cdot 10^{-5}$ for the M200E (BL)). It can therefore be concluded that the combination of meteorological conditions and traffic intervention result in significantly different NO₂ concentration distribution. We attract the reader attention to the inherent bias of the two weeks average concentration introduced by the loss of one entire day of data for the M200E (BL). Moreover, the meteorological conditions were significantly different from one day to the another, but we believe that this drawback can be addressed by looking independently at the 10 days of field study. Looking independently at the 10 days of the field study will also allow to quantify, at least roughly, the impact of both traffic interventions and the impact of the changing meteorological conditions on the local air quality.

5.3.1.2 Daily variation in NO₂ concentration

A close inspection of Figure 5.4 reveals that, for both instruments and traffic interventions, the NO₂ concentrations were higher during the first week compared to the second week. This is clearly reflected by the data presented in Table 5.3 that displays the average background NO₂ concentration, the average NO₂ concentration and the average above-background NO₂ concentration split by day.

Firstly, it can be noticed that the average above-background NO₂ concentration presents large variations from one day to another for the same traffic intervention. Since traffic can be

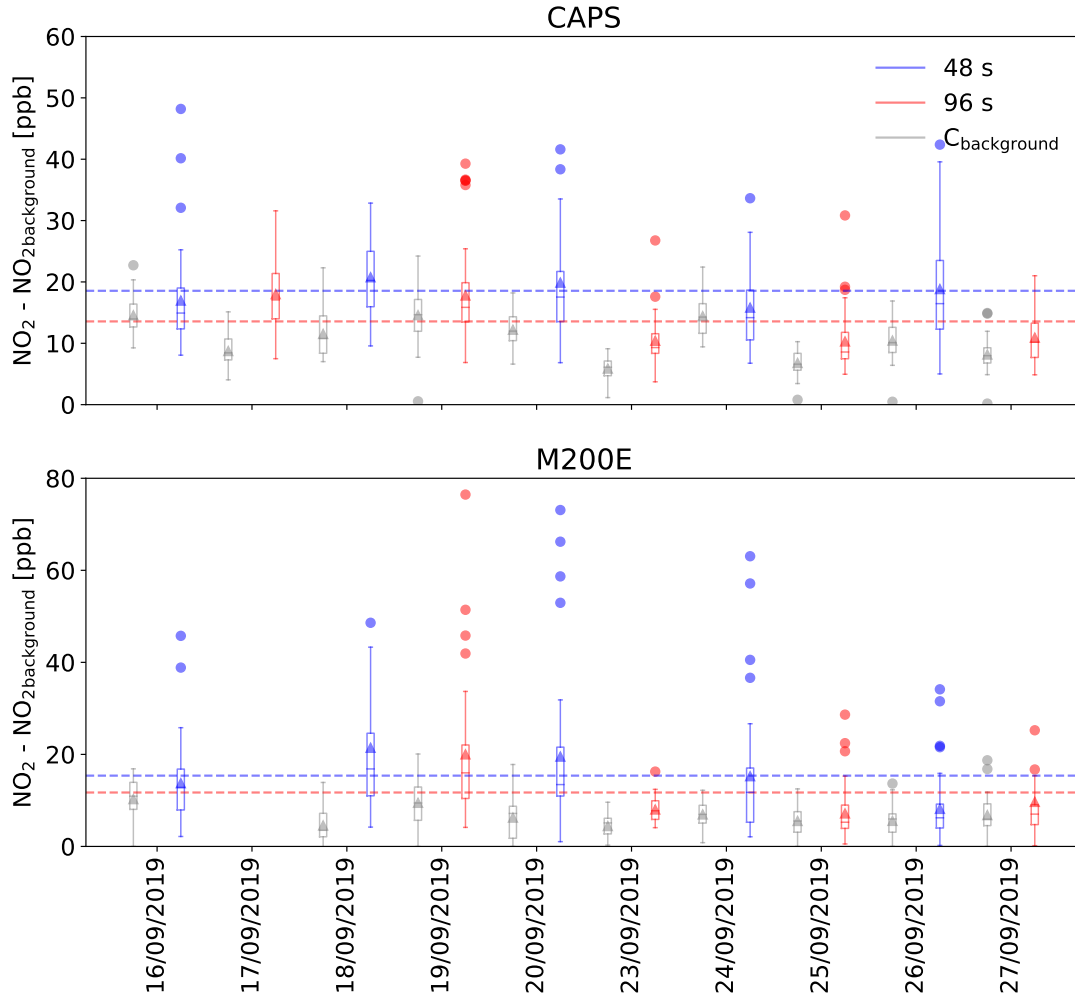


Figure 5.4: Distribution of the background NO₂ concentration and the above-background NO₂ concentrations for the CAPS (CI) and the M200E (BL) over the two weeks of field study. The background concentration distributions are depicted in grey while the above-background concentrations are depicted in blue during the A-48 traffic intervention and in red during the B-96 traffic intervention. The blue and red dashed lines represent the average above-background concentrations for the A-48 and the B-96 intervention respectively. For the CAPS (CI), the data resolution was degraded to 1 minute to match the resolution of the M200E (BL).

Table 5.3: Average background NO₂ concentrations, NO₂ concentrations and average measured above-background NO₂ concentrations for the CAPS (CI) and the M200E (BL) over the two weeks of field study. Days where the B-96 intervention was implemented are depicted in bold font, while days where the A-48 intervention was implemented are depicted in normal font.

	CAPS (CI)			M200E (BL)		
	NO _{2,bc} [ppb]	NO ₂ [ppb]	Diff. [ppb]	NO _{2,bc} [ppb]	NO ₂ [ppb]	Diff. [ppb]
16.09.2020	14.6	31.5	17.0	10.3	24.0	13.7
17.09.2020	8.7	26.8	18.0	-	-	-
18.09.2020	11.5	32.3	20.8	4.6	26.0	21.4
19.09.2020	14.6	32.4	17.8	9.5	29.5	20.0
20.09.2020	12.2	32.1	19.9	6.3	25.8	19.5
23.09.2020	5.9	16.3	10.4	4.4	12.5	8.0
24.09.2020	14.4	30.2	15.8	7.0	22.2	15.3
25.09.2020	6.8	17.1	10.3	5.5	12.7	7.2
26.09.2020	10.4	29.3	18.8	5.6	13.7	8.2
27.09.2020	8.2	19.1	10.9	6.8	16.5	9.7

considered as approximately invariant over the 10 days of field study, it indicates that local meteorological conditions are responsible for variations in concentration by up to 12.8 ppb, which corresponds to a variation by about 50%. Moreover, it can also be observed from Table 5.3 that the difference in concentrations between the two interventions is larger on the second week than on the first week, particularly for the CAPS (CI). For the M200E (BL), the difference in concentration between the two interventions is on average of 1.8 ppb for the first week and of 3.4 ppb for the second week. For the CAPS (CI) however, the difference in concentration between the two interventions is on average of 1.3 ppb for the first week and of 6.8 ppb for the second week. Since the traffic was approximately invariant over the 10 days of the field study, it can be inferred that the relative impact of the traffic interventions with respect to each other is in a large proportion determined by the meteorological conditions.

A close inspection of Tables 5.3 and C.2 reveals that the 26th and the 27th of September presented extremely similar meteorological conditions in terms of wind speed and wind direction. Since traffic was approximately invariant and meteorological conditions were similar, it can be inferred that any difference in concentrations can be directly linked to the traffic interventions, allowing us to quantify the direct impact of the traffic interventions on the NO₂ concentrations based on the sensors measurements. The CAPS (CI) measurements indicate that the A-48

intervention results in an increase in NO₂ concentration above the background of 18.8 ppb, compared to only 10.9 ppb for the B-96 intervention, or a difference of 7.9 ppb. The M200E (BL) measurements indicate that the difference in increase above NO₂ background concentration is only of 1.5 ppb. The large difference between the two sensors is probably attributable to the difference in location of the two sensors: the CAPS (CI) was located very close from the main source of traffic and concentrations recorded by the CAPS (CI) are therefore extremely influenced by the individual vehicle emissions.

Finally, it can be observed from Table 5.3 that for the CAPS (CI), the average above-background NO₂ concentration is 18.5 (s.d. 1.8) ppb for the A-48 intervention and 13.5 (s.d. 3.6) ppb for the B-96 intervention. For the M200E (BL), the average above-background NO₂ concentration is 15.6 (s.d. 4.7) ppb for the A-48 intervention and 11.2 (s.d. 5.2) ppb for the B-96 intervention. Nevertheless, the proportionally high standard deviations (comprised between 10% and 46%) indicate that the extend of this impact is extremely variable. For the A-48 intervention, absolute concentrations on average higher by 2.8 ppb (respectively 2.3 ppb) were recorded by the CAPS (CI) (respectively the M200E (BL)) compared to the B-96 intervention. The difference in concentrations recorded by the two instruments is most likely due to two main factors: 1) the proximity of the CAPS (CI) to the main source of traffic and 2) the helical recirculation induced by the wind within the canyon.

5.3.2 Traffic conditions

Figure 5.5 presents a comparison of the difference in traffic conditions, in terms of average speed, number of stops and percentage of time spent idling, induced by the two traffic interventions as a function of the distance to the studied junction. It is clearly visible from Figure 5.5 that the traffic interventions impact the most the Northbound lane (bus lane). The 96s cycle results in a significant increase of the vehicles average speed, and a significant decrease in the number of stops and in the time spent idling upstream from the junction. It indicates undoubtedly that the 96s cycle creates smoother driving conditions on the Northbound lane.

It can be observed from Figure 5.5 that the impact of the traffic interventions on the Southbound

lane is more subtle. The B-96 intervention seem to lead to an increase in the number of stops and time spent idling 30 to 50 meters upstream from the junction. It can therefore be inferred that the B-96 intervention leads to an increase in the length of the queue of vehicles along the Southbound lane upstream from the junction. It can therefore be deduced that the B-96 intervention causes a displacement of the emissions induced by acceleration events consequential to a stop. The emissions are more dispersed, avoiding the concentration of emissions (and consequently pollution) near the junction, therefore decreasing the concentrations measured at the CAPS (CI).

5.3.3 Comparison of the NLR and REVS models

The predictive skills of REVS in modelling aggregated emissions is analysed by spatially averaging the NLR model's predictions. An analysis of the spatial variability in emissions induced by each intervention is performed in a second step, using the predictions of the NLR model only.

Figure 5.6 displays a comparison of the probability distribution of the predicted average NO_x emission factors (left panel) and the predicted total NO_x emissions (right panel), split by vehicle types. The distributions were obtained by conducting the emission calculation procedure 1,000 times, as described in Section 5.2.4.2. The absence of distribution for the buses emissions predicted by the REVS model is due to the underlying assumptions: buses are exclusively retrofitted Euro V and therefore a unique model is applied. We make the reader attentive to the fact that the trucks, vans and buses emissions models are based on a very limited dataset, and that the predicted emissions should therefore not be interpreted as a reliable absolute representation of reality. Nevertheless, we believe that the relative differences in emissions are preserved and that it does not jeopardise the relative comparison of the differences in emissions induced by the two traffic interventions.

It can be first observed from Figure 5.6 that the REVS model predicts NO_x emissions factors lower by 6.8% (0.02 g/km) for the A-48 (respectively 13.2% (0.03 g/km) for the B-96 intervention) for passenger vehicles, by 6.1% (0.08 g/km) for the A-48 (respectively 10.9% (0.14 g/km)

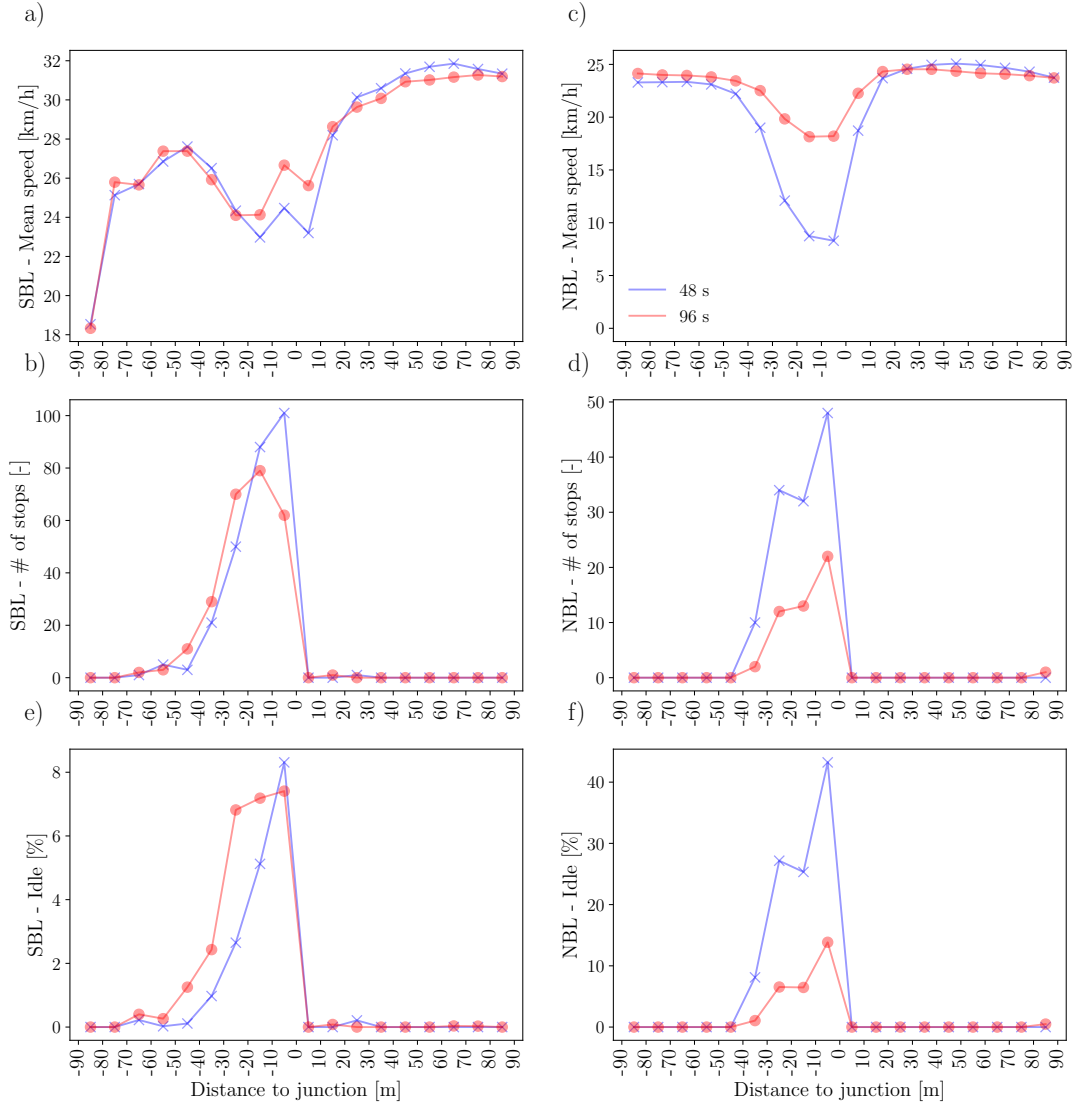


Figure 5.5: Comparison of the impact of the two traffic interventions on the traffic conditions. The left panel (a, b, c) corresponds to the Southbound lane, while the right panel (d, e, f) corresponds to the Northbound lane. Each datapoint corresponds to the average metric for all the vehicles at the range of distance to the junction considered. Negative distances (respectively positive distances) are situated upstream (respectively downstream) from the junction in the direction of the flow. The A-48 intervention is depicted in blue, while the B-96 intervention is depicted in red.

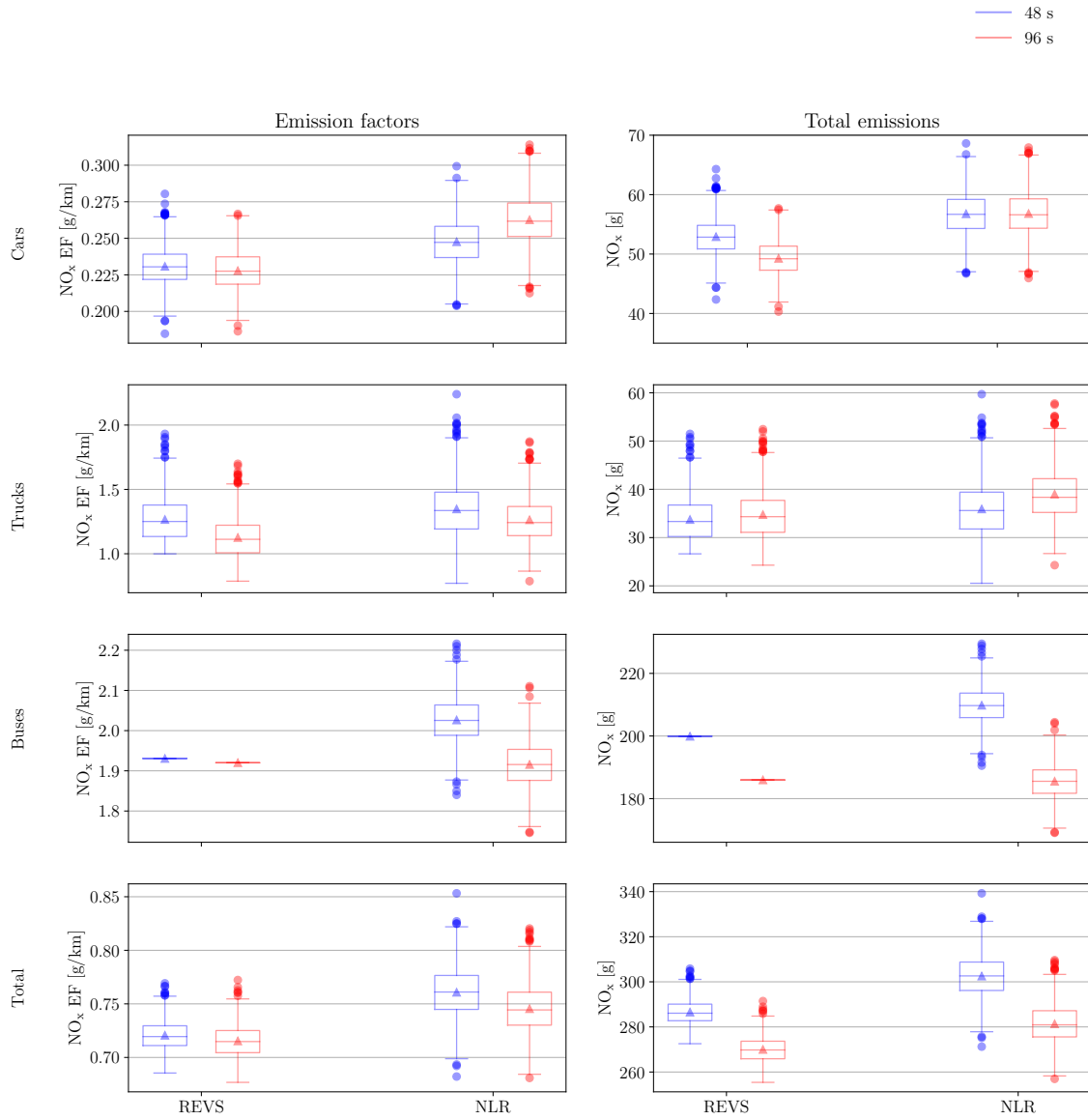


Figure 5.6: Comparison of the predicted emissions from the REVS and the NLR models, split by vehicle category. Distribution of the emission factors is given to the left, while the total emissions are given to the right. Distribution of the emissions resulting from the A-48 intervention is depicted in blue, while the distribution of the emissions resulting from the B-96 intervention is depicted in red.

for the B-96 intervention) for trucks and vans, by 4.7% (0.1 g/km) for the A-48 (respectively 0.3% (0.01 g/km) for the B-96 intervention) for buses and by 5.3% (0.04 g/km) for the A-48 (respectively 4.0% (0.03 g/km) for the B-96 intervention) for all the vehicles compared to the NLR model. This can be explained by the fundamental differences between the two models: the NLR is an instantaneous emission model that it able to catch local peaks in emissions whereas the REVS model is an aggregated emission model that predicts the average emission factor and its standard deviation over a trip, based on the average speed and VSP over this trip. Contrary to the NLR model, it therefore does not provide any indication on the spatial distribution of emissions over the network but gives an estimate of the variability via the standard deviation.

It can be also noticed from Figure 5.6 that overall the B-96 intervention results in a decrease in emissions by 7% (21.17 g) for NLR and by 6% (16.57 g) for REVS compared to the A-48 intervention. This decrease in emissions can be primarily attributed to the decrease in buses emissions. Finally, it can be noticed from Figure C.1 in Appendix C that the standard deviation of the average instantaneous NO_x emissions per street for the REVS model is about 10%, which is in agreement with the NLR model.

5.3.4 Spatial distribution in emissions

Figure 5.7 displays a comparison of the spatial distribution of the average predicted NLR NO_x emissions induced by the two traffic interventions. The average NO_x emissions distributions were obtained by averaging the predicted instantaneous NO_x emissions over the 1,000 runs of the emission calculation procedure. The panel the furthest right displays the difference of the A-48 intervention spatial distribution to the B-96 intervention spatial distribution.

It can be observed from Figure 5.7 that non surprisingly the maximum emissions occur around junctions. It confirms that the B-96 intervention leads to a decrease in emissions on the bus lane around the studied junction. A closer inspection of Figure 5.7.c) also reveals that the B-96 intervention results in a increase in emissions on the SBL upstream from the junction (for both the SBL and the RT direction), by 22.55 g or 13.5%, while the A-48 intervention present an increase in emissions on the SBL at the junction and downstream (for both the SBL and the

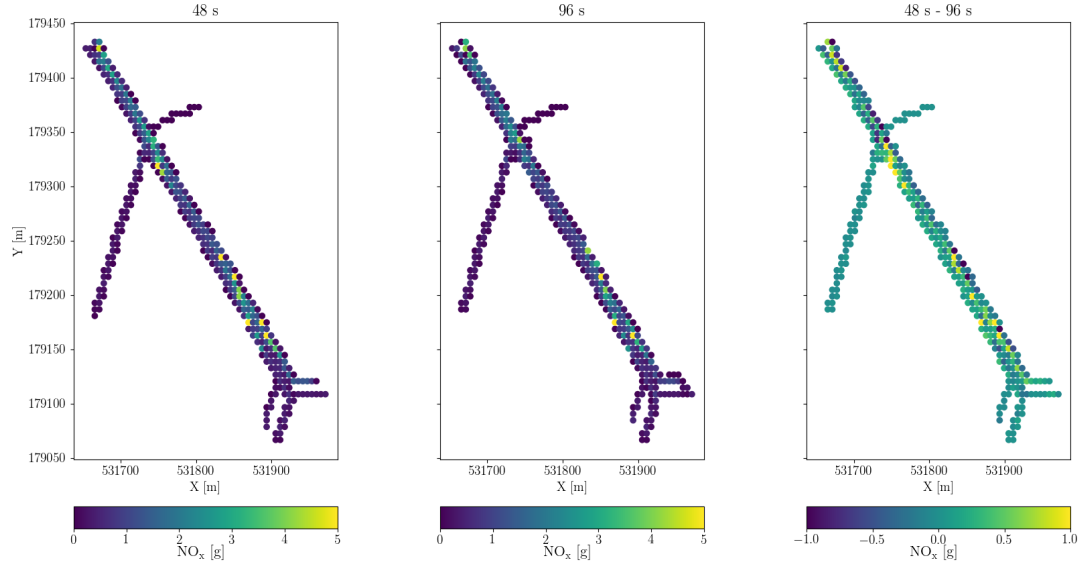


Figure 5.7: Comparison of the spatial distribution of the NLR predicted NO_x emissions induced by the two traffic interventions. The furthest right panel displays the difference of the 48s cycle spatial distribution to the 96s cycle spatial distribution. Datapoints colored in blue indicate areas where the B-96 intervention resulted in higher emissions than the A-48 intervention, while the datapoints colored in light green to yellow indicate areas where the A-48 intervention resulted in lower emissions than the B-96 intervention.

RT direction), by 2.59 g or 3.6%. This result supports the idea that, at the network level, the implemented interventions primarily influence the extent of buses emissions on the NBL. On the SBL and RT however, the two interventions do not present a fundamental difference in absolute emissions, however, the spatial distribution of them depends on the implemented cycle.

5.3.5 Operational air dispersion simulation

Figure 5.8 provides a comparison of the predicted canyon-averaged NO , NO_2 and O_3 concentrations in the three boxes of the network, for both emission models and traffic interventions, over the two weeks of field study, split per day. For every day and specific meteorological conditions, the impact of both traffic interventions on the local air quality were simulated, allowing to precisely quantify the impact of both traffic interventions disregarding of the meteorological conditions. The entire data is available in Section C.5. It can be observed that, according to the

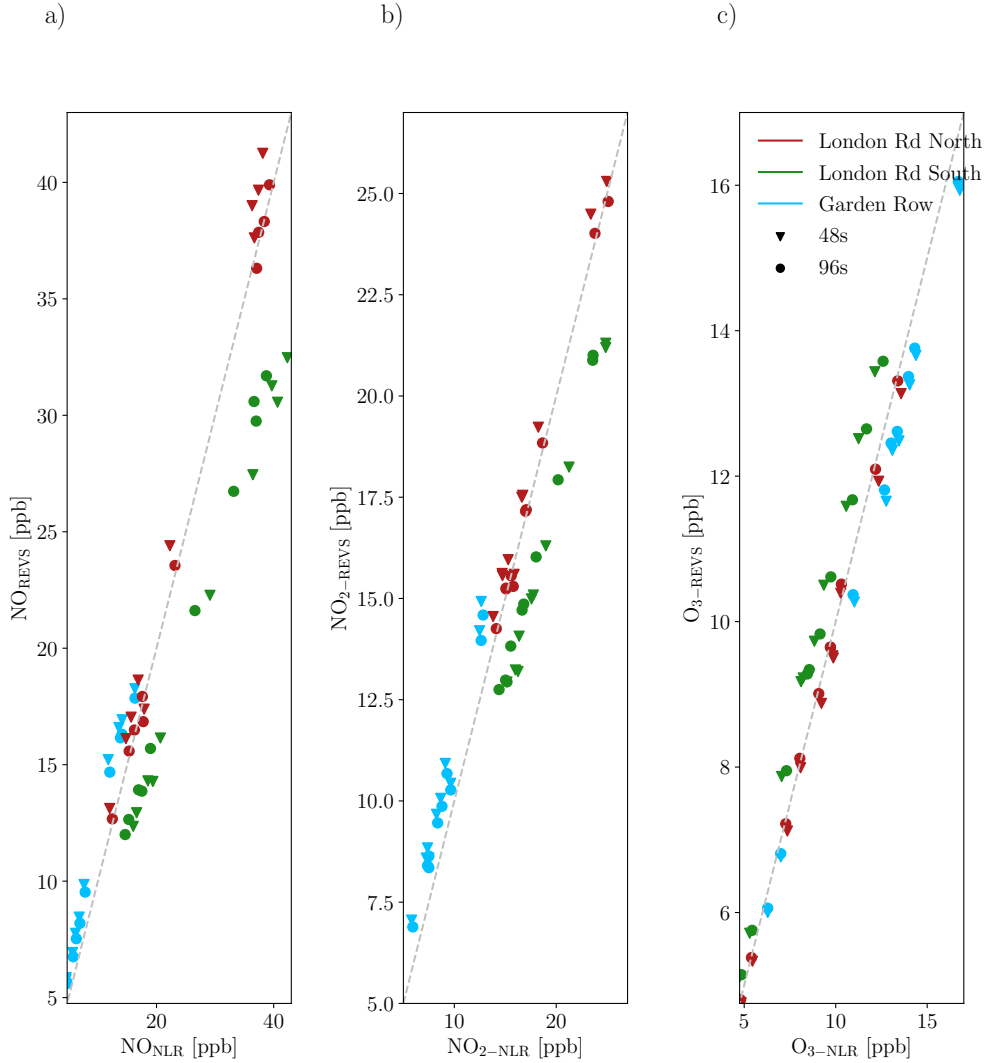


Figure 5.8: Comparison of the predicted canyon averaged (a) NO concentrations, (b) NO_2 concentrations and (c) O_3 concentrations in the three streets composing the network for both emission models. The concentrations resulting from the A-48 intervention are depicted with a cross, while the concentrations resulting from the B-96 intervention are depicted with a circle. The colour of the datapoints indicates the location of the SIRANE boxes. We attract the reader attention on the fact that the impact of both interventions were simulated for every day.

NLR model, over the two weeks of interventions, the B-96 intervention results on an absolute increase in NO_2 concentrations of on average 0.3 (s.d. 0.2) ppb on London Road North, of 0.1 (s.d. 0.06) ppb on Garden Row and a decrease of on average 1.1 (s.d. 0.2) ppb on London Road South. This result is in agreement with the observed difference in spatial distribution of emissions induced by both traffic interventions (increase in emissions on London Road North by 3.6% and decrease in emissions by 13.5% on London Road South for the B-96 intervention compared to the A-48 intervention). On the other hand, according to simulations conducted based on the NLR model predictions, it can be noticed that the average above-background NO_2 concentrations presents large variations from one day to another for the same traffic intervention, with a maximum difference of 7.1 ppb for the A-48 intervention on London Road North, of 6.9 ppb for the A-96 intervention on London Road North, of 6.3 ppb for the A-48 intervention on London Road South, of 5.8 ppb for the B-96 intervention on London Road South, of 1.6 ppb for the A-48 intervention on Garden Row and of 1.7 ppb for the B-96 intervention on Garden Row. Since the traffic is globally invariant over the two weeks of field study, it can be inferred that the meteorological conditions are responsible for variations in concentrations by up to 40%. At the canyon average resolution, it can therefore be concluded that the variation in concentration attributed to the change in meteorological conditions is up to one order of magnitude larger than the variation in concentration that can be directly attributed to the traffic interventions.

The predicted NO_2 concentrations at pedestrian level for each day of the field study are given in Table 5.4, and will be compared to the sensors measurements in Section 5.3.6. It can be observed from Table 5.4 that the change in emissions induced by the B-96 intervention compared to the A-48 intervention, as calculated by the NLR model, is responsible for a difference in pedestrian level average NO_2 concentration comprised between -0.6 and 0.2 ppb on London Road North, 1.1 and 1.8 ppb on London Road South and -0.3 and 0.01 ppb on Garden Row. Based on Table 5.4 and Table C.3, it can be noticed that the average above-background pedestrian NO_2 concentrations also presents large variations from one day to another for the same traffic intervention, with a maximum difference of 13.9 ppb for the A-48 intervention on London Road North, of 13.7 ppb for the A-96 intervention on London Road North, of 12.2 ppb for the

Table 5.4: Predicted NO₂ concentrations at pedestrian level. The first number corresponds to the NLR model, while the second number corresponds to the REVS model. Numbers in bold correspond to the B-96 intervention while numbers in normal font correspond to the A-48 intervention.

	NO ₂ [ppb]		
	London Road North	London Road South	Garden Row
16.09.2020	28.5 / 30.0	30.5 / 25.5	13.3 / 15.7
17.09.2020	20.9 / 21.1	20.5 / 17.8	10.6 / 11.4
18.09.2020	19.5 / 20.4	20.8 / 16.6	8.7 / 10.2
19.09.2020	31.2 / 30.7	29.0 / 25.3	14.1 / 16.6
20.09.2020	19.6 / 19.3	18.7 / 15.6	9.1 / 11.1
23.09.2020	19.3 / 19.5	19.2 / 16.4	6.5 / 7.9
24.09.2020	22.7 / 24.0	26.9 / 22.7	10.0 / 12.5
25.09.2020	17.5 / 17.7	19.4 / 17.1	8.1 / 9.5
26.09.2020	18.3 / 19.4	22.2 / 18.6	8.1 / 10.1
27.09.2020	21.0 / 21.2	22.4 / 19.6	9.6 / 11.1

A-48 intervention on London Road South, of 11.5 ppb for the B-96 intervention on London Road South, of 7.5 ppb for the A-48 intervention on Garden Row and of 7.6 ppb for the B-96 intervention on Garden Row. Similarly to the average SIRANE box concentration, it can be concluded that the variability in the above-background NO₂ concentrations directly caused by the variations of the meteorological conditions is up to 54%, which is higher than the variability induced by the two traffic interventions. In reality, SIRANE does not capture well the spatial variability in emissions and the concentrations in the immediate vicinity from the junction are thus likely to be higher than predicted. This effect is expected to be more pronounced for the CAPS (CI), situated in next to the largest source of traffic, than for the M200E (BL), located a few meters away from the main source of traffic. Finally, it can be observed that the fundamental differences between the REVS and the NLR emissions models are also reflected in Table 5.4: the systematic lower NO_x emissions predicted by the REVS model (by up to 6%) compared to the NLR model is reflected in the lower predicted pedestrian averaged NO₂ concentrations (by about 9%).

We make the reader attentive to the fact that the parametrisation used to derive pedestrian averaged concentrations does not account for the role of canyon asymmetry on vertical concentration profiles within the street canyon [Xie et al., 2003]. As a result of canyon asymmetry,

concentrations on the leeward side of the canyon are typically higher than on the windward side due to the dispersion induced by the canyon vortex. On both weeks, the angle of incidence of the wind leads to an increase in concentrations on the side opposite to the M200E (BL) (leeward). SIRANE does not account for this aspect, but other operational dispersion models, such as the Danish Operational Street Pollution Model (OSPM), provide parametrisations to account for differences in concentrations on the leeward and the windward side of the street induced by the canyon asymmetry. The OSPM model splits the canyon into a leeward side and a windward side and assumes that on the leeward side, the measured concentrations are the sum of the direct and recirculated contributions while on the windward side only the direct contribution of the emissions generated outside the recirculation zone are taken into account [Berkowicz et al., 1997]. As the authors underline [Berkowicz et al., 1997], this approach is very simple, but it allows to quantify the order of magnitude of the increase in concentration on the leeward side of the canyon due to the canyon asymmetry. Considering the simple case where the vortex is totally immersed in the canyon ($W/H < 1$), the recirculation contribution is parametrised as:

$$C_{\text{rec}} = \frac{Q}{\sigma_{\text{wt}} \cdot W} \quad (5.3)$$

where Q is the emission rate of the street in $\text{g}/(\text{m} \cdot \text{s})$, W is the width of the street in m, $\sigma_{\text{wt}} \approx 0.1 \cdot u_t$ is the canyon ventilation velocity, determined by the turbulence at the top of the canyon and u_t is the wind speed at roof level. Recirculation contribution on London Road is comprised between 2.85 and 3.84 ppb. We make the reader attentive to the fact that the aspect ratio of London Road is larger than 1, and therefore the calculated recirculated contributions give an upper limit rather than a precise estimate of the impact of canyon asymmetry on concentrations.

5.3.6 Comparison of the SIRANE predicted pedestrian average NO_2 concentrations with the sensors measurements

Figure 5.9 provides a comparison of the SIRANE predicted pedestrian NO_2 concentrations with the CAPS (CI) and the M200E (BL) measured NO_2 concentrations. The variability in

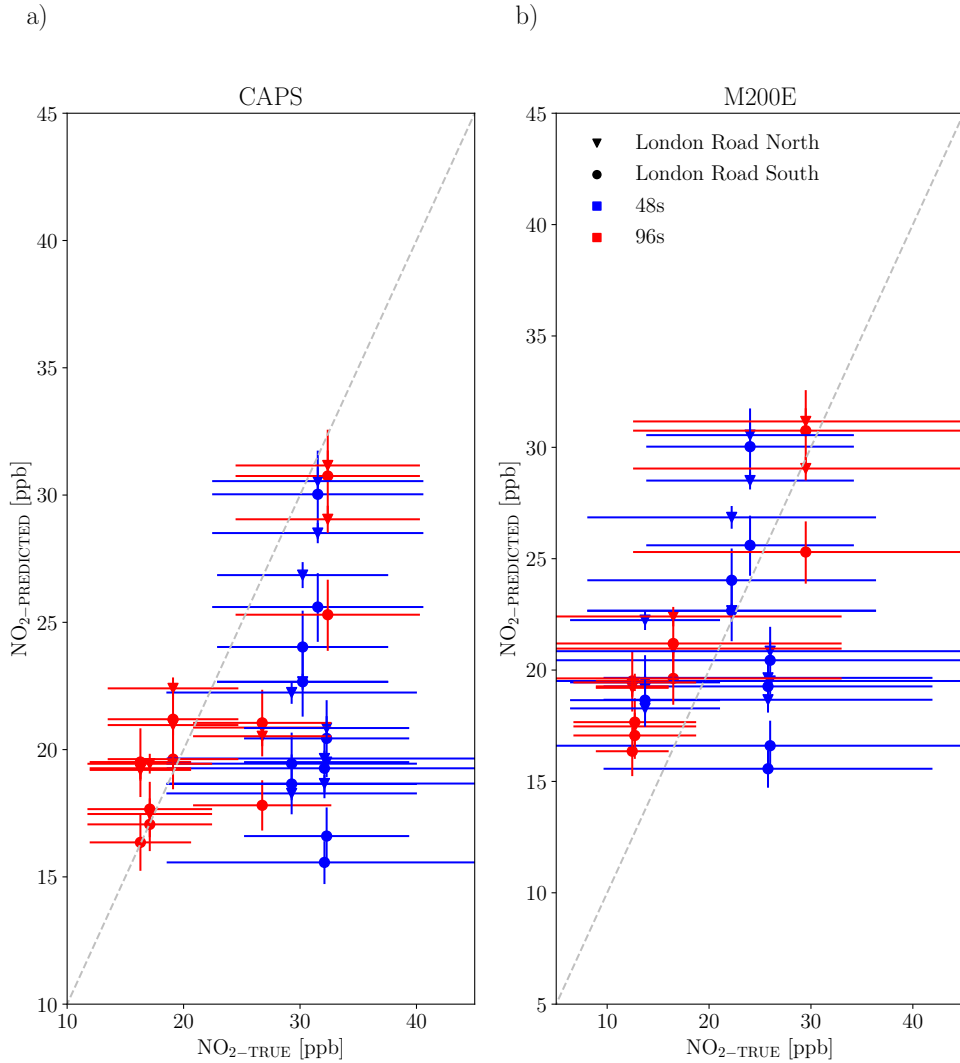


Figure 5.9: Comparison of the SIRANE predicted pedestrian NO_2 concentrations with the CAPS (CI) and the M200E (BL) measured NO_2 concentrations. The concentrations resulting from the A-48 intervention are depicted in blue, while the concentrations resulting from the B-96 intervention are depicted in red. The predicted concentrations on London Road North are depicted with a triangle while the predicted concentrations on London Road South are depicted with a circle. The errorbars depict one standard deviation. The silver dashed line represents the identity line.

the measured NO₂ concentrations are depicted with horizontal errorbars while the variability in the predicted NO₂ concentrations (induced by the real-world variability in emissions) are depicted with vertical errorbars. It can be observed from Figure 5.9 that the predicted NO₂ concentrations are globally in agreement with the NO₂ concentrations recorded by both sensors, although it is clearly noticeable that the predicted variability (of the order of $\pm 10\%$) is systematically underestimated. Most of the real-world variability in pollutant concentration can therefore be attributed to a sum of complex factors including variations in meteorological conditions, variations in the traffic flow, the spatial variability in emissions or the respective location of the sensors with respect to the main source of traffic. These factors are impossible to capture with an operational box dispersion model such as SIRANE.

As shown in Tables 5.3 and 5.4 and in Figure 5.9, SIRANE tends to underestimate the NO₂ pedestrian concentration by about 7.1 (s.d. 4.8) ppb for the A-48 intervention and by about 7.1 (s.d. 4.7) ppb for the B-96 intervention during the first week at the CAPS (CI) location. SIRANE also underestimates NO₂ pedestrian concentration by about 2.7 (s.d. 5.6) ppb for the A-48 intervention and by about 0.4 (s.d. 4.1) ppb for the B-96 intervention on the second week at the CAPS (CI) location. Similar conclusions can be drawn for the M200E (BL): SIRANE underestimates the NO₂ pedestrian concentration by about 1.6 (s.d. 4.8) ppb for the A-48 intervention and by about 1.5 (s.d. 5.3) ppb for the B-96 intervention during the first week, while it overestimates by about 6.5 (s.d. 1.3) ppb for the A-48 intervention and by about 4.2 (s.d. 2.1) ppb for the B-96 intervention during the second week. In addition, it can be noticed that SIRANE does not capture well the increase in NO₂ concentration above-background recorded by both sensors, especially during the first week. SIRANE tends to underestimate the increase in above-background NO₂ concentration by 9.0 (s.d. 2.7) ppb during the first week and by about 3.4 (s.d. 3.8) ppb during the second week. These observations can be attributed to the inability of SIRANE to account for the spatial variability in emissions and to resolve the flow field within the street canyon.

In Section 5.3.1, it was inferred that the traffic interventions resulted in difference in above-background NO₂ concentration of 7.92 ppb for the CAPS (CI) and 1.52 ppb for the M200E (BL). While the difference recorded by the M200E (BL) is well within the SIRANE's prediction,

the difference recorded by the CAPS (CI) is between 5 and 6 times higher than what SIRANE predicts. It can be undoubtedly concluded that the proximity of the CAPS (CI) to the traffic causes the sensor to be primarily affected by the individual level of emissions of the vehicles passing nearby.

The discrepancies between the sensors measurements and SIRANE can therefore be explained by 3 main factors: 1) the inability of SIRANE to account for the spatial variability in emissions, 2) its inherent resolution (i.e. the detailed flow field within a canyon is not solved and recirculation within the canyon are not accounted for) and 3) the location of the sensors with respect to the main source of traffic. Not surprisingly, while SIRANE comes in handy to estimate the large scale impact of an intervention, its inherent assumptions limit its capacity in analysing very local interventions.

5.4 Conclusion

In this chapter, a comparison of the real-world impact of two traffic interventions on the local emissions and air quality is conducted, using data collected during a two-weeks field study at a busy junction on London Road (Southwark, London, United Kingdom). The first traffic intervention consisted in a 48s cycle (A-48), while the second traffic intervention consisted in a 96s cycle (B-96) designed specifically to create a green way for buses on the Northbound lane. The data collected and the simulations conducted allowed us to: 1) analyse how the traffic interventions impacted the traffic flow and 2) quantify the variability in NO_2 concentration attributable to the traffic interventions, the meteorological conditions and the location of the sensors with respect to the main source of emissions (Southbound lane).

Analysis of the data collected during the two weeks of field study revealed that the meteorological conditions were responsible for large variations in NO_2 concentration: by up to 50% (more than 12 ppb) from one day to another. The exact extent of the impact of the traffic interventions independently from the meteorological conditions is difficult to quantify from the sensors measurements only, since the measured concentrations had the other intervention been

implemented cannot be known. Nevertheless, comparison of the two-weeks average NO_2 concentrations show that the B-96 intervention resulted in concentrations on average lower of 4.99 ppb (for the CAPS (CI)) and 3.67 ppb (for the M200E (BL)) compared to the A-48 intervention, which is significant. The significantly lower NO_2 concentrations recorded by the M200E (BL) compared to the CAPS (CI) (on average 2.5 ppb) can be linked to the distance of the both instruments to the main source of traffic.

Simulations were therefore used to study the impact of both interventions on: 1) the traffic flow, 2) the NO_x emissions and 3) the NO_2 concentrations. Analysis of the predicted real-world emissions data indicated that the B-96 intervention lead to a decrease in NO_x emissions on the bus lane that can be directly attributed to the decrease in driving dynamicity caused by the traffic intervention. On the southbound lane however, the B-96 intervention increases queuing 30 to 50 meters away from the junction, displacing the emissions upstream from the junction. Analysis of the outcomes of the SIRANE simulation allow to quantify the impact of both traffic interventions on NO_2 concentrations independently from the meteorological conditions. Traffic interventions were able to explain a difference in NO_2 concentrations comprised between -0.6 and 1.8 ppb at pedestrian level on London Road while the meteorological conditions were able to explain a variation in NO_2 concentrations by up to 13.9 ppb at pedestrian level. These numbers are globally in agreement with the measured NO_2 concentrations, despite the clear underestimation of the final NO_2 concentrations from SIRANE. Predictions from SIRANE confirmed that the higher NO_2 concentrations recorded by the CAPS (CI) were most likely due to the proximity of the sensor with the main source of traffic and to the recirculation vortex induced by the wind within the street canyon, two parameters that SIRANE does not account for.

This study is the first to conduct both an empirical and a theoretical analysis of the impact of a traffic intervention on the local emissions and air quality. It could be concluded that, while the impact of the traffic interventions on traffic conditions, emissions and local air quality were significant, the difference in measured concentrations that can be attributed to the meteorological conditions is at least one order of magnitude higher than the difference in measured concentrations that can be attributed to the traffic interventions. It therefore provided promis-

ing insights on the effectiveness of traffic light optimisation to reduce emissions and improve air quality, and we believe that a future step would be to evaluate the impact of these local improvements on pedestrian exposure.

Chapter 6

Conclusions

6.1 Main findings and contributions

This PhD thesis presents new insights on the factors affecting vehicles real-world driving NO_x emissions, develops new emission models able to account for the aforementioned factors and finally applies the acquired knowledge to quantify the impact of two traffic interventions on emissions and air quality. For a detailed list of key contributions, the reader is referred to Section 1.4. Let us recall here the main keywords from the list of novel results.

Factors affecting real-world NO_x emissions

- I. Chapter 3 demonstrates that driving dynamicity (characterised by the VSP) impacts directly real-world driving NO_x emissions, however, the extend of this impact is more pronounced for diesel than petrol vehicles and tend to decrease with the successive regulations (Euro standards).
- II. Chapter 3 also establishes by a bootstrap procedure that the uncertainty due to the existing sampling of the fleet compared to a full sampling of the fleet is around 20%.
- III. Chapter 4 shows that physical factors such as engine size or vehicle weight have little if any impact on real-world driving NO_x emissions of diesel vehicles. Only the type of

aftertreatment system fitted on the vehicle (and consequently the operating regulation at the time of type-approval) presents a relatively homogeneous influence on NO_x emissions.

- IV. Chapter 4 also indicates that while the inter-vehicles real-world variability in emissions is important and should not be neglected, the intra-vehicle real-world variability in emissions is also of major importance. In particular, it is shown in Chapter 4 that a repeated test of the same physical vehicle, on the same test road, with the same driver, can lead to differences in emission factors by up to a factor 7, with differences by a factor of 2 not being unusual.

Modelling of real-world driving NO_x emissions

- V. Chapter 3 introduces the REVS model, an aggregated emission model that uses a simple parametrisation of speed and VSP to derive the average NO_x emission factor and its standard deviation per vehicle technology category. In addition of providing an accurate estimate the standard deviation of the NO_x emission factor of a fleet, the REVS model is demonstrated to perform better than COPERT5 in predicting the average NO_x emission factor of a fleet.
- VI. Chapter 4 introduces a refitting of a non-linear regression model (NLR) developed by [Panis et al., 2006] to predict instantaneous NO_x emissions from instantaneous speed and acceleration. The novelty of the approach consists in developing one set of coefficient per set of vehicles presenting similar emitting behaviour, as opposed to one model per vehicle technology category. Although the NLR model misses some peaks in emissions, it does not present systematic errors in prediction and has a better accuracy than the COPERT5 model.

Impact of a traffic interventions on local emissions and air quality

- VII. Closely linked to both previous chapters, Chapter 5 uses the framework of a field campaign aiming to study the impact of a traffic intervention on emissions and air quality to compare

the REVS model against the NLR model. It is demonstrated that the REVS model predicts total NO_x emission factor over the entire network lower by about 6% compared to the NLR model, which can be considered as satisfactory. The predicted standard deviation by the REVS model is about 10%, which is in good agreement with the NLR model. However, the REVS model does not provide any indication of the spatial variability in emissions, which represents a major drawback to conduct the high resolution analysis required by the study.

- VIII. In particular, close inspection of the spatial variability in emissions induced by the traffic interventions reveals that not only traffic interventions modify the intensity of emissions but also spatial distribution of them on the network.
- IX. Chapter 5 then quantifies using an operational air dispersion model, SIRANE, coupled to the predicted emissions for both interventions, the expected difference in NO_2 concentrations caused, on one hand, by the traffic interventions and, on the other hand, by the change in meteorological conditions. It is demonstrated that most variability could be attributed to the meteorological conditions rather than to the traffic interventions, although the impact of these can also be quantified.
- X. The NO_2 concentrations measured by the sensors located next to the southbound lane and to the bus lane respectively confirmed the results obtained with SIRANE: the impact of both traffic interventions on the NO_2 concentrations is measurable, but this impact is about one order of magnitude smaller than the variability in concentrations that can be attributed to changes in meteorological conditions.

6.2 Limitations and future work

There are several limitations to this PhD thesis and some future research efforts should focus on further improving the results presented in this thesis:

- I. A major drawback of the emissions models developed in Chapter 3 and 4 is their limitation

- to Euro 5 and current Euro 6 passenger vehicles. In the future, data collected during the RDE type approval procedure could be used to: 1) update the passenger vehicles models and 2) develop the emissions models for other vehicles types (buses, HGVs).
- II. A major limitation of the current NLR model is the impossibility to determine a priori to which cluster a new, untested vehicle belongs to. Any newly tested vehicle would therefore require its own independent model, multiplying the total number of models to an unmanageable number. The data collected during the RDE type approval procedure could help solving this issue. Despite not defining the RDE type approval like a driving cycle, the driving dynamics experienced by a vehicle during type approval is well-defined and segments with similar driving conditions could be extracted and used to perform the clustering procedure before fitting the model.
- III. Hybrids vehicles are not considered in this thesis. While they still represent, at the time of writing, a minor fraction of the fleet, this fraction will further increase in the future. It is therefore crucial to understand: 1) how real-world driving conditions affect the emissions from the different types of hybrids, 2) what is the impact of ageing on the real-world emissions of hybrids vehicles.
- IV. This thesis focuses on exhaust NO_x emissions. However, it appears very likely that non-exhaust emissions (tyres, brakes and road dust) will also be regulated in a near future. It is therefore crucial to better assess and quantify the real-world factors affecting their emissions (e.g. type of brakes and tyres, driving dynamicity). Such understanding could be use to develop non-exhaust emissions model as well as type approval procedures.
- V. This thesis is exclusively based on PEMS data. However, remote sensing typically provides a snapshot of real-world emissions on a large fraction of the fleet. Quantifying how representative of real-world emissions remote sensing data is would allow to regularly perform large scale evaluation of the fleet performance in terms of emissions. It could ultimately be used as a way to monitor real-world driving emissions and ensure that another Dieselgate scandal does not happen.
- VI. This thesis focuses on conditions that are typical of Western European urban areas, but

more interest has to be dedicated to developing countries urban areas that often suffer a greater burden in terms of air quality than Western European cities and where typical driving conditions and fleet differ tremendously from Europe. Conclusions drawn in this thesis might not apply to them.

- VII. The analysis presented in Chapter 5 relies only upon 2 weeks of data, each of them with specific weather conditions. Definitive quantification of the impact of both the change in emissions induced by the traffic interventions and the meteorological conditions would require the collection of data under other weather conditions.
- VIII. Moreover, the field campaign was conducted only during off peak hours, with similar traffic conditions over the two weeks of study. Collecting data at different times during the day would allow to quantify the relative impact of one traffic intervention with respect to the other as a function of traffic flow.
- IX. Finally, Chapter 5 does not address the question of the impact of both traffic interventions on pedestrian exposure. The use of high resolution dispersion models such as Fluidity, uDALES or ADMS that resolve the flow field within the street canyon and that can directly account for the spatial variability in emissions will certainly help answering this question.

6.3 Final thoughts

To conclude, this thesis contributes substantially to the literature on real-world driving NO_x emissions and modelling, and also provides useful insights for the assessment of the impact of a traffic intervention on both the emissions and air quality. However, future research is needed to update and extend the developed emission models and to understand more in depth the mechanisms governing the real-world emissions of the new types of vehicles arriving on the market (such as hybrids). More investigation is also needed to better ascertain the potential of traffic interventions such as optimisation of traffic management systems to improve air quality. This thesis is definitely the start of a challenging and exciting journey to improve air quality.

Appendix A

REVS: Road transport Emissions and Variance Simulator

A.1 Datasets description

A.1.1 Description of the EA2019 dataset

Table A.1: Vehicles characteristics EA2019 dataset

ID	Make	Model	Engine size [cm ³]	Power	Fuel	Year	Mileage [km]	NO _x af- tertreatment technology
27	Land Rover	Discovery Sport	1999	240 [ps]	Diesel	2017	18093	EGR+SCR
2	BWM	X3	2993	265 [ps]	Diesel	2017	10230	EGR + SCR
7	BMW	3 Series 320D	1995	163 [ps]	Diesel	2017	10316	EGR + LNT
18	Honda	Civic	1597	120 [ps]	Diesel	2018	5436	EGR + LNT
60	Volkswagen	Touareg	2967	285 [ps]	Diesel	2018	8143	EGR + SCR
30	Mercedes-Benz	GLE-Class	2987	258 [ps]	Diesel	2017	9886	EGR + SCR
47	Opel/Vauxhall	Insignia	1956	209 [ps]	Diesel	2017	10139	EGR + SCR
2	Audi	A6	2967	286 [ps]	Diesel	2018	4615	EGR + SCR

51	Porsche	Macan	2967	258 [ps]	Diesel	2016	103610	EGR + SCR
64	Volvo	V60	1969	190 [ps]	Diesel	2018	3300	EGR + SCR
11	BMW	3 Series 318d	1995	148 [bhp]	Diesel	2018	3384	EGR + SCR
10	Citroen	Relay	1997	130 [bhp]	Diesel	2018	7533	EGR + SCR
40	Nissan	Qashqai	1598	130 [bhp]	Diesel	2018	1387	EGR + LNT
19	Hyundai	Santa Fe	2199	200 [ps]	Diesel	2018	6163	EGR + LNT + SCR?
48	Opel/Vauxhall	Combo	1499	130 [ps]	Diesel	2018	5683	EGR + SCR
25	Kia	Sportage	1995	185 [ps]	Diesel	2018	8451	EGR + SCR + LNT
35	Mini	Countryman	1995	150 [ps]	Diesel	2018	3883	EGR + SCR
18	BMW	3 Series 318d	1995	150 [ps]	Diesel	2018	3877	EGR + LNT + SCR
49	Opel/Vauxhall	Astra	1595	136 [ps]	Diesel	2018	4103	EGR + SCR
61	Volkswagen	Golf	1598	105 [ps]	Diesel	2011	149353	EGR
11	Citroen	Berlingo	1499	130 [ps]	Diesel	2018	8210	EGR + SCR
31	Mercedes-Benz	C-Class	1950	194 [bhp]	Diesel	2019	1320	EGR + SCR
3	Audi	A7	2967	286 [bhp]	Diesel Hybrid	2018	2076	EGR + SCR
44	Opel/Vauxhall	Crossland X	1199	130 [ps]	Gasoline	2018	5245	-
52	Renault	Megane	1197	132 [ps]	Gasoline	2017	7236	-
39	Nissan	Micra	1000	70 [ps]	Gasoline	2017	9757	-
58	Volkswagen	Polo	998	75 [ps]	Gasoline	2017	8982	-
20	Kia	Stinger	1998	255 [ps]	Gasoline	2017	5681	-
36	Mitsubishi	Eclipse Cross	1499	163 [ps]	Gasoline	2017	6863	-
62	Volvo	XC40	1477	156 [ps]	Gasoline	2018	3099	-
34	Mini	Hatch	1499	136 [ps]	Gasoline	2018	3648	-
12	Fiat	500	1242	69 [bhp]	Gasoline	2017	3976	-
16	Honda	Civic	998	129 [bhp]	Gasoline	2017	2547	-
23	Kia	Picanto	998	66 [bhp]	Gasoline	2018	6507	-

38	Mitsubishi	Space Star	999	70 [ps]	Gasoline	2018	2793	-
59	Volkswagen	Polo	1984	200 [ps]	Gasoline	2018	9111	-
63	Volvo	V60	1969	310 [ps]	Gasoline	2018	2236	-
13	Ford	Fiesta	998	100 [ps]	Gasoline	2018	10296	-
26	Land Rover	Range Rover Velar	1997	300 [ps]	Gasoline	2018	4422	-
55	Ssangyong	Tivoli	1597	128 [ps]	Gasoline	2018	6614	-
53	Skoda	Kodiaq	1984	180 [ps]	Gasoline	2018	13853	-
45	Opel/Vauxhall	Adam	1229	69 [bhp]	Gasoline	2018	896	-
28	Mazda	Mazda2	1496	90 [ps]	Gasoline	2018	4981	-
37	Mitsubishi	Eclipse Cross	1499	161 [bhp]	Gasoline	2018	461	-
4	BMW	3 Series 340i	2998	322 [bhp]	Gasoline	2018	2930	-
21	Kia	Stonic	998	120 [ps]	Gasoline	2017	20627	-
14	Ford	Fiesta	998	140 [ps]	Gasoline	2018	4079	-
50	Peugeot	508	1598	224 [ps]	Gasoline	2018	4309	-
0	Audi	Q3	1395	150 [bhp]	Gasoline	2018	5775	-
15	Ford	Fiesta	1496	200 [ps]	Gasoline	2018	3996	-
1	Audi	A3	1984	187 [bhp]	Gasoline	2018	101	-
29	Mazda	Mazda2	1496	115 [ps]	Gasoline	2018	5372	-
41	Opel/Vauxhall	Astra	1598	200 [ps]	Gasoline	2018	6575	-
17	Honda	CR-V	1498	193 [ps]	Gasoline	2018	8326	-
24	Kia	Picanto	1248	83 [bhp]	Gasoline	2018	3315	-
42	Opel/Vauxhall	Zafira Tourer	1598	136 [ps]	Gasoline	2018	6284	-
46	Opel/Vauxhall	Corsa	1364	150 [ps]	Gasoline	2018	6833	-
56	Suzuki	Jimny	1462	101 [ps]	Gasoline	2018	2996	-
54	Skoda	Octavia	1498	150 [ps]	Gasoline	2018	4270	-
43	Opel/Vauxhall	Astra	1598	200 [bhp]	Gasoline	2019	1278	-
22	Kia	Ceed	1353	140 [ps]	Gasoline	2018	11597	-
5	BMW	8 Series M850i	4395	489 [ps]	Gasoline	2018	5723	-
57	Suzuki	Ignis	1242	88 [bhp]	Gasoline Hybrid	2017	8079	-

A.1.2 Description of the EA-2016 dataset

A Portable Emissions Measurements System (PEMS) was used to measure real world driving CO, CO₂, NO, NO₂ and NO_x emissions of 149 Euro 5 and Euro 6 petrol and diesel passenger vehicles in the Greater London area between 2012 and 2016. The test fleet contained 37 Petrol Euro 5, 35 Petrol Euro 6, 36 Diesel Euro 5, 39 Diesel Euro 6, 1 Euro 5 Petrol Hybrid and 1 Euro 6 Petrol Hybrid vehicles. The vehicles were made by 28 different manufacturers (Ford, Nissan, Mercedes, Citroen, Kia, Peugeot, Skoda, Volkswagen, Toyota, Audi, Land Rover, Hyundai, Mitsubitshi, Ssangyong, BMW, Fiat, Volvo, Mazda, Vauxhall, Infiniti, Mini, Seat, Suzuki, Chevrolet, Subaru, Porsche, Renault, Alfa Romeo). All the diesel vehicles were fitted with a Diesel Oxidation Catalyst (DOC), a Diesel Particulate Filter (DPF) and an Exhaust Gas Recirculation (EGR) system. The diesel Euro 6 vehicles had a mixture of NO_x aftertreatment systems, 7 used EGR only, 16 EGR + Lean NO_x Trap (LNT) and 11 used EGR + Selective Catalytic Reduction (SCR) while the diesel Euro 5 vehicles were only equipped with an EGR. All petrol vehicles in the test fleet were fitted with three-way catalysts (TWCs) which control NO_x, carbon monoxide (CO) and hydrocarbon emissions. We refer the interested reader to [O’Driscoll et al., 2018] for a more detailed description of the data collection process.

A.1.3 Description of the DfT dataset

A Portable Emissions Measurement System (PEMS) was used to measure real-world driving CO, CO₂, NO, NO₂ and NO_x emissions of 19 Euro 5 and 19 Euro 6 diesel passenger vehicles across the UK. The work was commissioned in 2016 by the UK Department for Transport (DfT) following the Volkswagen scandal in 2015 and vehicles were selected to represent at least 50% of the British diesel passenger fleet. The vehicles were from 16 different manufacturers (Audi, BMW, Ford, Honda, Hyundai, Jaguar, Kia, Mazda, Mercedes, Mini, Peugeot, Renault, Skoda, Toyota, Vauxhall, Volkswagen). The vehicles had a mixture of NO_x aftertreatment systems, 1 used EGR only, 14 used EGR + LNT and 4 used EGR + SCR. We refer the interested reader to [Department for Transport, 2016] for a more detailed description of the data collection process.

A.1.4 Description of the DVSA dataset

A Portable Emissions Measurement System (PEMS) was used to measure real-world driving CO, CO₂, NO, NO₂ and NO_x emissions of 2 Euro 5 and 31 Euro 6 petrol and diesel passenger vehicles across the UK. The work was commissioned in 2017 and 2018 by the UK Department for Transport (DfT) via the Vehicle Market Surveillance Unit program. The vehicles were from 21 different manufacturers (Abarth, BMW, Fiat, Ford, Honda, Hyundai, Jaguar, Jeep, Kia, Mazda, Mercedes, Mini, Mitsubishi, Nissan, Peugeot, Renault, Seat, Toyota, Vauxhall, Volkswagen, Volvo). The diesel vehicles had a mixture of NO_x aftertreatment technologies: 8 used EGR + LNT, 1 used EGR + SCR and 2 are fitted with EGR + LNT + SCR. All the petrol engines were fitted with a TWC. Ten of the petrol engines were GDI while the other 12 were PFI. We refer the interested reader to [Driver and Vehicle Standards Agency, 2018, Driver and Vehicle Standards Agency, 2019] for a more detailed description of the data collection process. It must be noticed that to keep the testing route secrete, neither the GPS coordinates nor the altitude is available and this latest factor was neglected in the analysis.

A.2 Processing of the raw data

This section gives an overview of the raw processing of the raw data. For each test, the resolution of the acceleration (a_{res}) is given. If $a_{\text{res}} > 0.01 \text{ m/s}^2$, the T4253 smoothing function was applied. Once this smoothing procedure was performed, any negative speed was set equal to 0. The topography correction consisted in replacing the GPS altitude measurement by the real topographic altitude if the discrepancies between these two variables was higher 40 meters. Finally, an altitude correction was run.

A.2.1 EA2019

Table A.2: Raw processing of the tests of EA2019 dataset

Test ID	vehicle ID	Place	a_{res} [m/s ²]	Topography correction?	Altitude correction?
0	15	Ge	0.014	Y	Y
1	16	UK	0.014	Y	Y

2	41	Ge	0.014	Y	Y
3	30	Ge	0.014	Y	Y
4	49	Ge	0.014	Y	Y
5	18	Ge	0.014	Y	Y
6	34	Ge	0.014	Y	Y
7	29	Ge	0.014	Y	Y
8	28	UK	0.014	Y	Y
9	1	UK	0.014	Y	Y
10	18	Ge	0.014	Y	Y
11	36	Ge	0.014	Y	Y
12	19	Ge	0.014	Y	Y
13	45	UK	0.014	Y	Y
14	62	Ge	0.028	Y	Y
15	59	Ge	0.014	Y	Y
16	16	UK	0.014	Y	Y
17	50	Ge	0.014	Y	Y
18	25	Ge	0.014	Y	Y
19	27	Ge	0.014	Y	Y
20	2	Ge	0.014	Y	Y
21	4	UK	NaN	N	N
22	23	UK	0.014	Y	Y
23	9	Ge	0.014	Y	Y
24	2	Ge	0.014	Y	Y
25	57	UK	0.014	Y	Y
26	60	Ge	0.014	Y	Y
27	40	UK	0.014	Y	Y
28	46	Ge	0.014	Y	Y
29	38	Ge	0.014	Y	Y
30	31	UK	0.014	Y	Y
31	14	Ge	0.014	Y	Y

32	3	UK	0.014	Y	Y
33	48	Ge	0.014	Y	Y
34	21	Ge	0.014	Y	Y
35	10	UK	NaN	N	N
36	54	Ge	0.014	Y	Y
37	6	Ge	0.014	Y	Y
38	22	Ge	0.014	Y	Y
39	26	Ge	0.014	Y	Y
40	5	Ge	0.014	Y	Y
41	7	Ge	0.014	Y	Y
42	63	Ge	0.014	Y	Y
43	47	Ge	0.014	Y	Y
44	58	Ge	0.014	Y	Y
45	28	Ge	0.014	Y	Y
46	56	Ge	0.014	Y	Y
47	8	UK	0.014	Y	Y
48	40	UK	0.014	Y	Y
49	13	Ge	0.014	Y	Y
50	36	Ge	0.014	Y	Y
51	61	Ge	0.014	Y	Y
52	52	Ge	0.014	Y	Y
53	10	UK	NaN	N	N
54	42	Ge	0.014	Y	Y
55	9	Ge	0.014	Y	Y
56	20	Ge	0.014	Y	Y
57	43	UK	0.014	Y	Y
58	58	Ge	0.014	Y	Y
59	42	Ge	0.014	Y	Y
60	0	UK	NaN	N	N
61	55	Ge	0.014	Y	Y

62	5	Ge	0.014	Y	Y
63	35	Ge	0.014	Y	Y
64	13	Ge	0.014	Y	Y
65	62	Ge	0.014	Y	Y
66	39	Ge	0.014	Y	Y
67	60	Ge	0.014	Y	Y
68	28	Ge	0.014	Y	Y
69	30	Ge	0.014	Y	Y
70	51	Ge	0.014	Y	Y
71	21	Ge	0.014	Y	Y
72	62	Ge	0.014	Y	Y
73	38	Ge	0.014	Y	Y
74	64	Ge	0.014	Y	Y
75	12	UK	0.014	Y	Y
76	53	Ge	0.014	Y	Y
77	37	UK	0.014	Y	Y
78	46	Ge	0.014	Y	Y

A.2.2 EA2016

None of the tests required a topographic correction. The altitude correction was not needed neither.

A.2.3 DfT

None of the tests required a topographic correction. The altitude correction was not needed neither.

A.2.4 Extraction of the trips

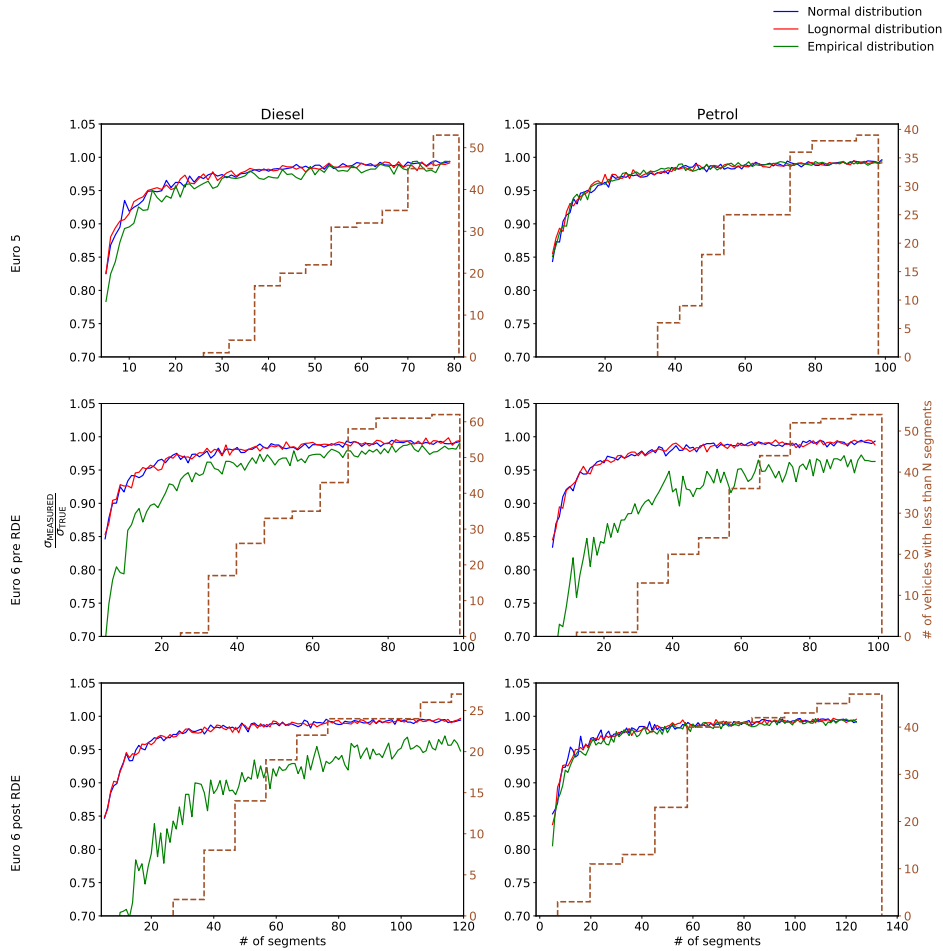


Figure A.1: Comparison of the minimum number of trips needed per vehicle to capture the true standard deviation with an error lower than 20%. The blue dashed line represents the case of a normal distribution, the red dashed line represents the case of a lognormal distribution and the green dashed line represents the case of the empirical distribution of the vehicle presenting the maximum number of trips per vehicle technology category. The brown dashed line displays the distribution of the number of trips per vehicle and per vehicle technology category.

To eliminate any potential bias due to an individual vehicle being under-sampled (i.e. not having enough trips), we empirically computed the minimum number of trips necessary to reconstruct an individual vehicle pdf. Three distributions were considered: 1) a normal distribution, 2) a lognormal distribution and 3) an empirical distribution. The mean and standard deviation of the three types of distribution were taken as the mean and standard deviation of the vehicle having the highest number of trips within a specific vehicle technology category. One thousand iterations were performed and the average measured standard deviation was compared to the true standard deviation. Since the average

is by definition an unbiased estimator, the minimum number of trips necessary to recover the standard deviation will be enough to recover the mean.

Figure A.1 displays the ratio of the average measured standard deviation to the real standard deviation as a function of the number of trips considered, for the three distributions. Not surprisingly, it can be observed that recovering the standard deviation is the most difficult in the case of an empirical distribution. However, it can also be noticed that if more than 10 trips are considered, the error on the standard deviation is lower than 20%, which can be considered as satisfactory.

A.2.5 Bootstrapping

Bootstrapping is a Monte Carlo technique belonging to the wide family of resampling methods. It was firstly introduced by [Efron, 1992] in 1979 and is based on the idea that inference about a population from sample data can be made by resampling with replacement the original sample data. Bootstrapping methods are particularly useful to draw inferences about populations, and are widely used to derive confidence intervals, variances or standard errors from an approximating distribution.

The basic bootstrapping algorithm consists in building the distribution of a statistic of interest (e.g. standard error) for a population in a recursive manner. At each step, named bootstrap replication, a random sample is drawn with replacement from the available data and a so-called bootstrap copy of the statistic is calculated. After B replications (generally $B \geq 1000$) the collected estimators constitute a virtual copy of the real population behind the sample data. It is important to bear in mind that the fundamental assumption behind bootstrapping is that the original sample accurately represents the real population. In particular, if small perturbations to the data generating process result in large variations of the sampling distribution, the bootstrapping is extremely likely to miserably fail [Shalizi, 2015].

Bootstrapping methods can be used in multiple contexts to estimate parameters of a population or probability distribution. The non-parametric bootstrap makes no assumption on the population distribution. In this case, the empirical distribution of the sample data is considered as the best guess of the population distribution and the distribution of the statistic of interest is obtained as previously described. Often however, some knowledge on the population distribution is available. Parametric bootstrap is designed to integrate this extra information about the population distribution. In this

case, the Maximum Likelihood Estimation parameters for the distribution are estimated from the sample data. The distribution of the statistic of interest is obtained as previously described.

A complete description of all the existing bootstrap methods is beyond the scope of this report and the interested reader is referred to [Shalizi, 2015] for a discussion on the original bootstrap procedure and to [Lee, 1994] for a discussion on parametric and nonparametric bootstrap.

A.2.6 VSP operating modes

Table A.3: VSP operating modes

Speed [km/h]			Speed [km/h]		Speed [km/h]	
VSP [kW/t]	[0; 1.6)	[1.6; 55)	VSP [kW/t]	[55; 90)	VSP [kW/t]	>90
<-4	Bin 1	Bin 11	<2	Bin 21	<6	Bin 31
[-4; -2)	x	Bin 12	[2; 4)	Bin 22	[6; 8)	Bin 32
[-2; 0)	x	Bin 13	[4; 6)	Bin 23	[8; 10)	Bin 33
[0; 2)	x	Bin 14	[6; 8)	Bin 24	[10; 12)	Bin 34
[2; 4)	x	Bin 15	[8; 10)	Bin 25	[12; 14)	Bin 35
[4; 6)	x	Bin 16	[10; 12)	Bin 26	[14; 18)	Bin 36
[6; 8)	x	Bin 17	[12; 14)	Bin 27	[18; 20)	Bin 37
>8	x	Bin 18	>14	Bin 28	>20	Bin 38

A.2.7 Comparison of the typical VSP distribution in London to the VSP distribution of the dataset

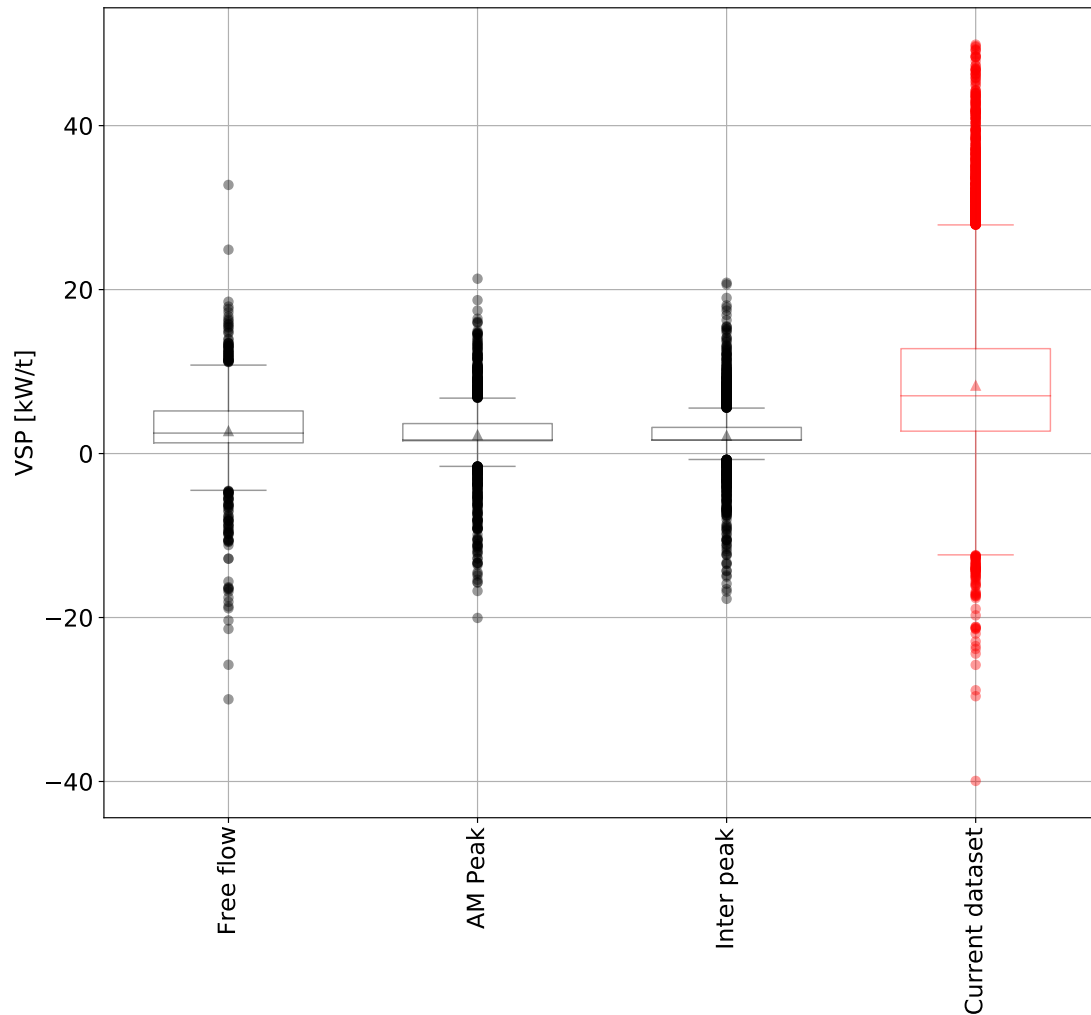


Figure A.2: Comparison of the typical VSP distribution in London, based on the TfL car cycles, in different traffic flow conditions, with the measured VSP over the entire dataset. Typical London VSP are depicted in black, while the current dataset distribution is depicted in red.

Figure A.2 displays a comparison of the typical distribution of VSP in London, based on the TfL car cycles, with the measured distribution of VSP over the entire dataset utilised in this study. It can be observed that there is globally a good agreement between the TfL typical VSP and the measured

VSP. The distribution of the measured VSP presents a larger spread and a slightly higher mean than the TfL typical VSP.

A.3 Computing the average and standard deviation of a sum of different distributions

Calculating the mean μ and the standard deviation σ of an ensemble of vehicles belonging to different vehicle technology categories corresponds to calculate the mean and the standard deviation of an union of sets containing N_S samples of respective mean μ_S and standard deviation σ_S . Let's denote the average, standard deviation and number of samples of the entire set as μ , σ and N . If these sets are independent, the entire set average is simply a weighted average of the samples average:

$$\mu = \frac{\sum_S \mu_S N_S}{N} \quad (\text{A.1})$$

The entire set standard deviation σ is given by:

$$\sigma = \sqrt{\frac{\sum_{N_S} (x_j - \mu)^2}{N - 1}} \quad (\text{A.2})$$

However, the value of $(x_j - \mu)^2$ is unknown.

The standard deviation of one set σ_S is given by:

$$\sigma_S = \sqrt{\frac{\sum_{N_S} (x_j - \mu_S)^2}{N_S - 1}} \quad (\text{A.3})$$

Therefore we have:

$$\sigma_S^2 (N_S - 1) = \sum_{N_S} (x_j - \mu_S)^2 \quad (\text{A.4})$$

If we sum over all the samples, we get:

$$\sum_S \sum_{N_S} (x_j - \mu_S)^2 = \sum_N (x_j - \mu)^2 \quad (\text{A.5})$$

Looking at the left hand side of the equation and introducing some terms, we have:

$$\sum_{N_S} (x_j - \mu_S) = \sum_{N_S} (x_j^2 - 2x_j\mu_S + \mu_S^2) + \left(\sum_{N_S} \mu^2 - \sum_{N_S} \mu^2 \right) + \left(2 \sum_{N_S} x_j(\mu - \mu_S) - 2 \sum_{N_S} x_j(\mu - \mu_S) \right) \quad (\text{A.6})$$

By summing over N_S , the above equation can be rewritten as:

$$= \sum_{N_S} (x_j^2 - 2x_j(\mu_S + \mu - \mu_S) + \mu^2) + \sum_{N_S} \mu_d^2 - \sum_{N_S} \mu^2 + 2 \sum_{N_S} x_j(\mu - \mu_S) \quad (\text{A.7})$$

Which can be further simplified in:

$$= \sum_{N_S} (x_j^2 - 2x_j\mu + \mu^2) + N_S\mu_S^2 - N_S\mu^2 + 2 \sum_{N_S} x_j(\mu - \mu_S) \quad (\text{A.8})$$

and:

$$= \sum_{N_S} (x_j - \mu)^2 + N_S\mu_S^2 - N_S\mu^2 + 2 \sum_{N_S} x_j(\mu - \mu_S) \quad (\text{A.9})$$

As we know the sample average, the equation can be rewritten further as:

$$= \sum_{N_S} (x_j - \mu)^2 + N_S\mu_S^2 - N_S\mu^2 + 2(\mu - \mu_S)N_S \left(\frac{1}{N_S} \sum_{N_S} x_j \right) = \sum_{N_S} (x_j - \mu)^2 + N_S\mu_S^2 - N_S\mu^2 + 2(\mu - \mu_S)N_S\mu_S \quad (\text{A.10})$$

If we plug into the first equation, we get:

$$\sigma_S^2(N_S - 1) = \sum_{N_S} (x_j - \mu)^2 + N_S\mu_d^2 - N_S\mu^2 + 2(\mu - \mu_S)N_S\mu_S \quad (\text{A.11})$$

$$\Leftrightarrow \sigma_S^2(N_S - 1) - N_S\mu_S^2 + N_S\mu^2 - 2N_S\mu_S(\mu - \mu_S) = \sum_{N_S} (x_j - \mu)^2 \quad (\text{A.12})$$

$$\Leftrightarrow \sigma_S^2(N_S - 1) + N_S(\mu - \mu_S)^2 = \sum_{N_S} (x_j - \mu)^2 \quad (\text{A.13})$$

By replacing into Equation (5), The final standard deviation sigma is therefore given by:

$$\sigma = \sqrt{\frac{\sum_S (\sigma_S^2(N_S - 1) + N_S(\mu - \mu_S)^2)}{N - 1}} \quad (\text{A.14})$$

A.4 REVS model

A.4.1 Model coefficients

The REVS model, described in details in the main text, is given by:

$$\bar{e}_{\text{REVS}}(\bar{v}, \overline{VSP}) = \exp(\alpha(\bar{v}) \cdot \overline{VSP} + \beta(\bar{v})) \quad (\text{A.15})$$

$$\sigma_{\text{REVS}}(\bar{v}, \overline{VSP}) = \exp(\gamma(\bar{v}) \cdot \overline{VSP} + \delta(\bar{v})) \quad (\text{A.16})$$

where $\bar{e}_{\text{REVS}}$ is the average NO_x emission factor in g/km, σ_{REVS} is the standard deviation around the average NO_x emission factor in g/km, α , β , γ and δ are the model parameters, $\bar{v}$ is the average speed in km/h and $\overline{VSP}$ is the average vehicle specific power in kW/t.

Table A.4 and A.5 give the coefficients of the model, for all the vehicle technology categories. Please note that for the sake of clarity, the coefficients are given only up to the 4th decimal. The full coefficients are also available in as a numpy array for download at [Le Cornec, 2021].

Table A.4: Coefficients of the REVS model for diesel vehicles, per Euro standard

	D.E.5				D.E.6 pre RDE				D.E.6 post RDE			
	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ
[10;20) [km/h]	0.1386	-0.0311	0.0932	-0.3644	0.2789	-1.2216	0.214384	-1.1311	-0.1793	-1.3530	0.0305	-1.4743
[20;30) [km/h]	0.1038	-0.2411	0.0583	-0.5242	0.0823	-0.8703	-0.085194	-0.3640	0.2057	-2.4610	0.1764	-2.0766
[30;40) [km/h]	0.1293	-0.5821	0.1234	-0.8036	0.1316	-1.3681	0.132893	-1.1869	0.1393	-2.6891	0.1637	-2.3268
[40;50) [km/h]	0.1366	-0.9405	0.0859	-0.8278	0.2154	-2.1642	0.129413	-1.6269	0.0608	-2.7695	-0.0264	-1.9486
[50;60) [km/h]	-0.0008	0.0166	-0.0447	-0.3904	0.1745	-1.5467	0.124073	-1.1518	0.1294	-3.1825	0.2315	-3.4727
[60;70) [km/h]	0.0874	-0.5300	0.0999	-1.0814	0.1884	-2.0255	0.111803	-1.2724	0.1910	-3.5942	0.1395	-2.7981
[70;80) [km/h]	0.1010	-0.9289	0.0729	-0.8812	0.1516	-1.7825	0.209930	-1.7611	0.0469	-3.3539	0.0487	-2.4960
[80;90) [km/h]	0.0907	-1.0618	0.0730	-0.9661	0.1340	-1.8015	0.154925	-1.2857	0.0577	-3.6070	0.1091	-3.0189
[90;100) [km/h]	0.0583	-0.7759	0.0088	-0.4911	0.1494	-2.5044	0.058279	-1.1847	0.1929	-5.1686	0.1957	-4.1885
[100;110) [km/h]	0.0822	-1.0276	0.0575	-0.9907	0.1149	-2.1910	0.114251	-1.9256	0.0545	-3.7763	0.0336	-2.2309
[110;120) [km/h]	0.0829	-1.4962	0.0754	-1.5387	0.0819	-2.3097	0.040689	-1.3889	0.0099	-2.8571	0.0110	-2.4262
[120;130) [km/h]	0.0891	0.0882	0.1081	0.1007	0.0942	0.0935	0.1189	0.0977	0.2639	0.8984	-0.1713	0.6891

Table A.5: Coefficients of the REVS model for petrol vehicles, per Euro standard

	P.E.5				P.E.6 pre RDE				P.E.6 post RDE			
	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ
[10;20) [km/h]	-0.2014	-1.7569	-0.2723	-1.3735	0.2538	-3.4088	0.4923	-3.6955	-0.2070	-3.0591	-0.2144	-2.7803
[20;30) [km/h]	0.1803	-2.6958	0.2982	-2.5401	0.0674	-3.1848	0.0320	-2.4438	0.0347	-3.7651	0.0835	-3.4998
[30;40) [km/h]	0.0191	-2.5812	0.0657	-2.1464	0.1527	-3.7141	0.2557	-3.4838	-0.0220	-3.7451	0.0230	-3.2249
[40;50) [km/h]	0.1263	-3.4926	0.1810	-3.1597	-0.1115	-3.5824	-0.2185	-2.6611	0.0351	-4.5179	0.1823	-4.2783
[50;60) [km/h]	-	-	-	-	0.6384	-7.1395	0.2802	-3.9548	0.1286	-5.0148	0.2962	-5.0534
[60;70) [km/h]	-	-	-	-	0.0859	-3.2808	0.0757	-2.3833	0.0080	-4.2218	-0.0037	-3.4846
[70;80) [km/h]	-	-	-	-	0.1591	-3.9410	0.1438	-2.6301	0.0223	-4.7879	0.1260	-4.5738
[80;90) [km/h]	-	-	-	-	0.0966	-4.6462	0.2154	-4.8667	0.0879	-5.1273	0.1274	-4.6126
[90;100) [km/h]	0.1208	-5.3673	0.1842	-5.9511	0.0154	-3.7752	0.0111	-3.1557	0.1640	-6.6127	0.1867	-6.4019
[100;110) [km/h]	0.0356	-4.0340	0.0296	-3.6846	0.0568	-3.7617	0.0279	-2.3602	0.0560	-5.0983	0.0913	-4.8308
[110;120) [km/h]	0.1102	-5.1091	0.1005	-4.8484	0.0800	-4.6724	0.0994	-4.4526	0.0238	-4.8315	0.0558	-4.3735
[120;130) [km/h]	0.0996	0.0996	0.1006	0.0993	0.0993	0.1000	0.0942	0.0995	-0.4323	-2.6067	-0.4213	-2.7452

A.4.2 Performance of the REVS model per driving environment and comparison with COPERT5

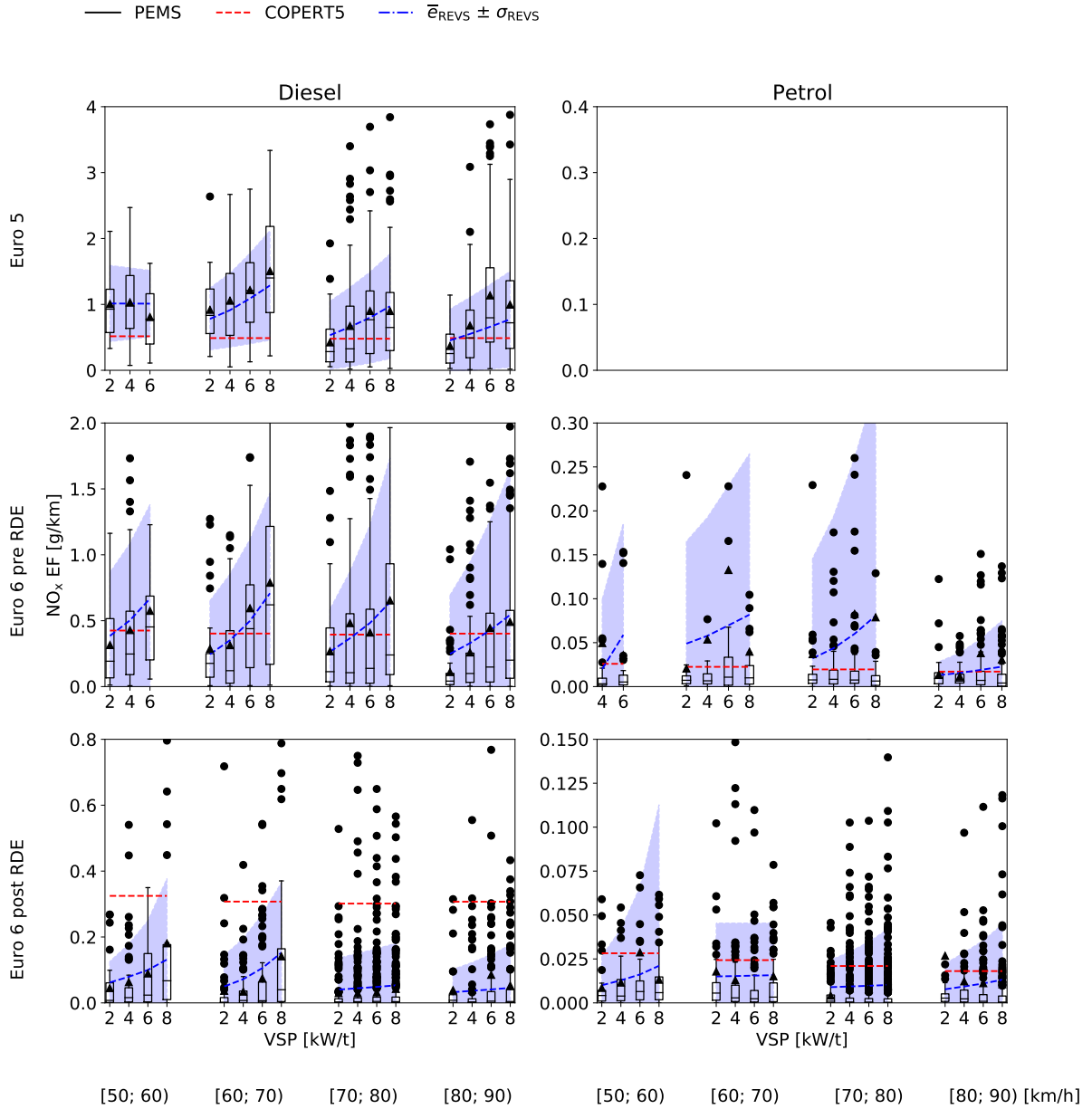


Figure A.3: Comparison of the measured and predicted NO_x emission factors in rural driving conditions split by vehicle technology category. The average NO_x emission factor predicted by COPERT5 (respectively the REVS model) is depicted by a dashed red (respectively blue) line. The predicted standard deviation by the REVS model is depicted by blue shaded areas around the average. The average speed of the trip is comprised between 50 and 90 km/h aggregated per bin of 10 km/h. The average VSP of the trip is comprised between -2 and 6 kW/t aggregated per bin of 2 kW/t (only the lower bin is indicated on the Figure). The distributions of the measured NO_x emission factor are constituted of all the extracted trips with the corresponding binning in speed and VSP. Note: Some outliers have been cropped and distributions are displayed only then more than 10 trips per bin are available.

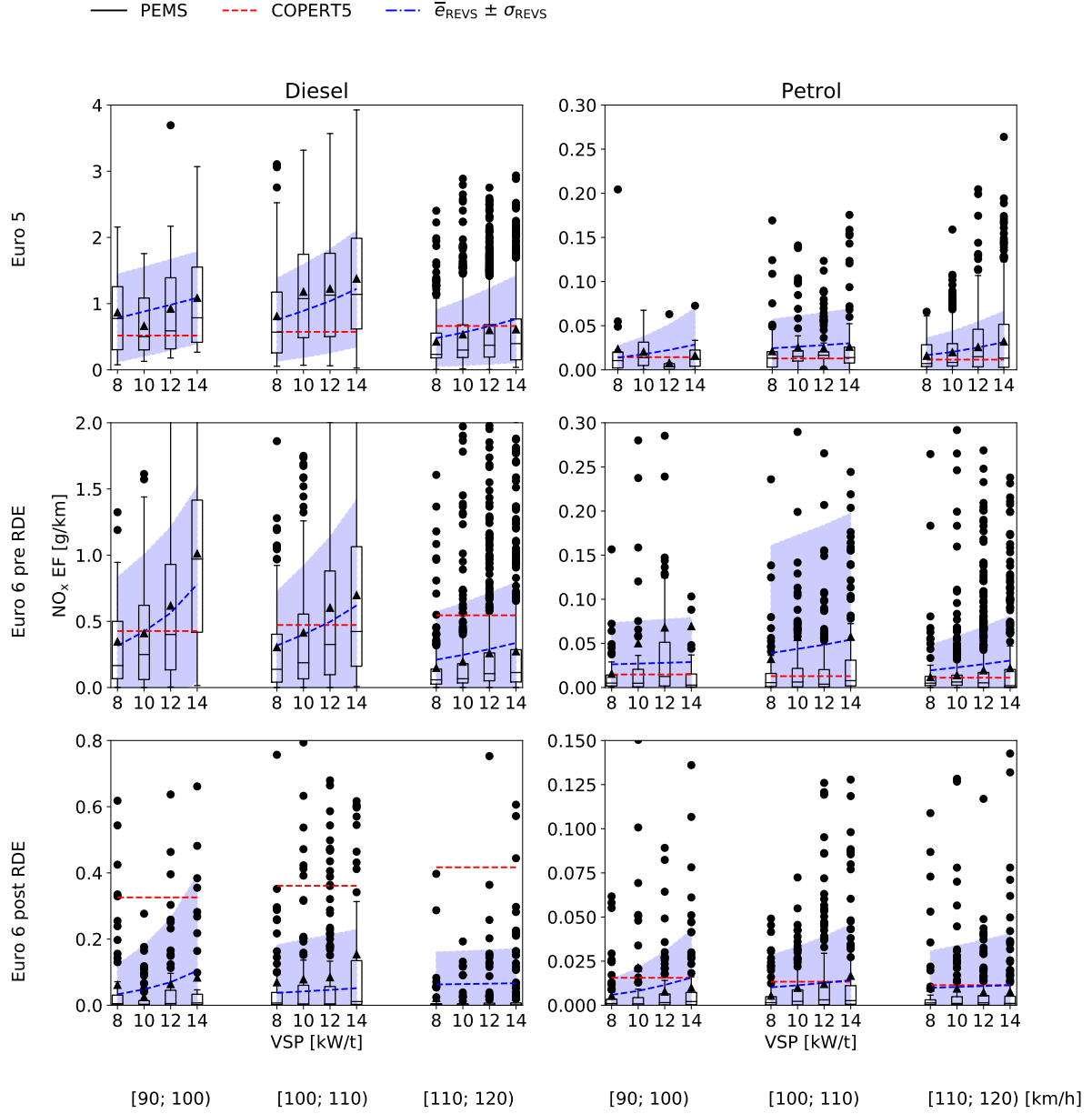


Figure A.4: Comparison of the measured and predicted NO_x emission factors in motorway driving conditions split by vehicle technology category. The average NO_x emission factor predicted by COPERT5 (respectively the REVS model) is depicted by a dashed red (respectively blue) line. The predicted standard deviation by the REVS model is depicted by blue shaded areas around the average. The average speed of the trip is comprised between 90 and 130 km/h aggregated per bin of 10 km/h. The average VSP of the trip is comprised between -2 and 6 kW/t aggregated per bin of 2 kW/t (only the lower bin is indicated on the Figure). The distributions of the measured NO_x emission factor are constituted of all the extracted trips with the corresponding binning in speed and VSP. Note: Some outliers have been cropped and distributions are displayed only then more than 10 trips per bin are available.

A.5 Procedure to apply the REVS model

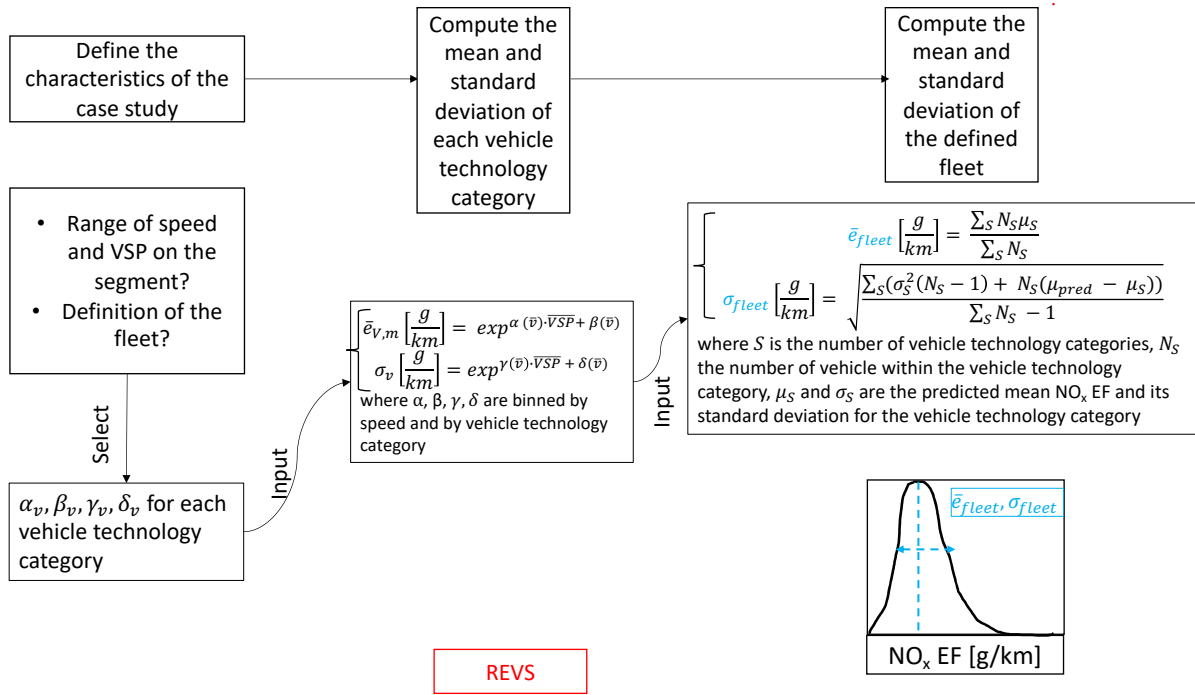


Figure A.5: Illustration of the procedure to follow to apply the model to a case study.

Appendix B

Modelling of instantaneous emissions from diesel vehicles with a special focus on NO_x : Insights from machine learning techniques

B.1 Dynamic Time Warping

Dynamic Time Warping (DTW) is a class of algorithms that allow comparing two (time-dependent) sequences with respect to their shape under certain restrictions [Muller, 2007]. Intuitively, the two sequences are matched by distorting the time axis, stretching some segments while shrinking others. Similar dynamic patterns occurring at different timesteps of the sequences are matched together and consequently the algorithm is able to cope with time deformations and different speeds associated with time-dependent data [Izakian et al., 2015]. The optimal alignment of these two sequences should satisfy the following conditions: the first and the last indices of both sequences are matched together, the mapping of indices from the first to the second sequence must be monotonically increasing and continuity should be insured [Muller, 2007]. The monotonicity condition reflects to requirement for faithful timing (e.g. if

an element in the first sequence precede a second, it should be the same in the second sequence) while the continuity condition forces the algorithm to consider all the data points.

DTW algorithms were successfully applied to a wide range of problems, including speech recognition [Myers et al., 1980], gesture recognition [Sempena et al., 2011, Corradini, 2001], clustering [Petitjean et al., 2014], etc.

B.2 K-means++

k-means is a widely used clustering algorithm aiming at partitioning N observations into k clusters by minimizing the within-cluster sum-of-squares [Arthur and Vassilvitskii, 2007]. The original algorithm begins with k centres, typically chosen randomly from the data points. Each point is then assigned to the nearest cluster and the cluster center is recomputed as the centre of mass of all the points assigned to it. These two stages are repeated until the clusters stabilize. Although simple and quick, accuracy is not insured and the results can converge to a suboptimal solution [Jain, 2010]. [Arthur and Vassilvitskii, 2007] proposed an approximation to this NP-hard problem, by choosing the initial values (seeds) randomly from the datapoints similarly as for the original algorithm but additionally weighting the datapoints according to their squared distance from the closest center already chosen. This so called k-means++ is $O(\log k)$ -competitive and achieves performances 20% better and almost twice as fast than the original k-means algorithm for a real-world dataset [Arthur and Vassilvitskii, 2007].

B.3 Calinski-Harabatz and Davies-Bouldin indices

Evaluation of the optimal number of clusters, in the context of unsupervised clustering where little or none previous knowledge about the data is available, is a complicated task. Internal validation measures, based on separation and compactness, are a commonly used methodology. The Calinski-Harabatz (CH) index is based on the average sum-of-squared between and within each cluster while the Davies-Bouldin (DB) index is a function of the intra-cluster similarity

and inter-cluster separation [Maulik and Bandyopadhyay, 2002, Czink et al., 2006]. Both the CH and the DB indices allow to determine the partition with the most compact and separated clusters.

B.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) are a wide family of machine learning algorithms and a computational approach to problem solving. They are used in a large variety of applications from classification, pattern recognition, image processing to data-processing tasks. ANNs are intended for modelling the organization principles of a biological central nervous system. The important feature of these networks is their adaptive nature and therefore their ability to learn and generalize from examples. This flexibility makes ANNs very appealing for solving complex non-linear problems especially when the underlying physic is poorly understood. Various ANNs architectures aiming at tackling different problems have been developed throughout the years. Among those, it is worth citing multilayer perceptrons, convolutional and recurrent networks.

An ANN consists of one or more nodes (or neurons) each receiving inputs $x(x_1, \dots, x_n)$. Each neuron performs a linear combination of input signal and adds a bias. The single neuron is a feedforward element with the connections directed from the inputs to the output of the neuron. The result is then passed to an activation or transfer function and then to the next layer until the output layer is reached. Activation functions are used to map the weighted sum of inputs to outputs in a non-linear way inside a node. The neuron is activated or not depending on the result. The final output of the network is finally compared to the real value. The error is then back propagated to the network and the weights are adjusted to minimize the final error (starting from the final neurons of the output layer up to the input layer). A graphic description of a neuron structure is given in Figure B.1.

A complete description of neural network architectures is beyond the scope of this paper and the interested reader is referred to [Leonard and Kramer, 1990] for a description of the back-

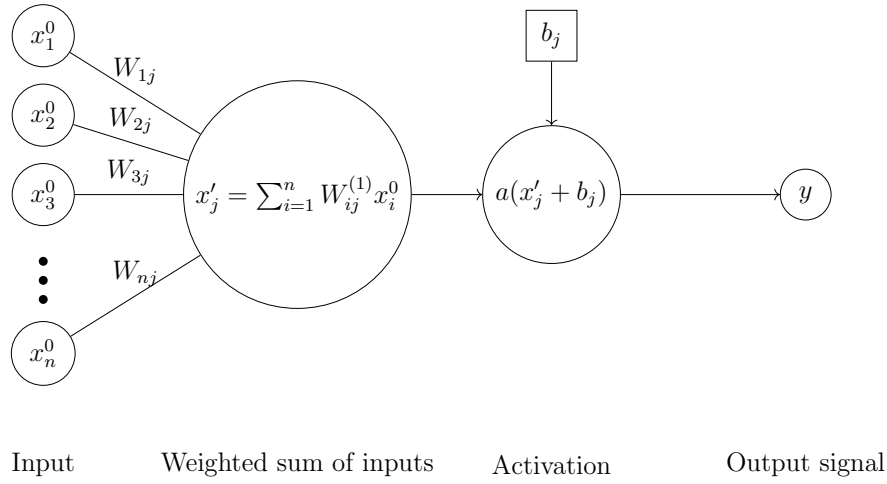


Figure B.1: Illustration of a single artificial neuron

propagation algorithm, to [Ruder, 2016], [Duchi et al., 2011], [Zeiler, 2012], [Kingma and Ba, 2014], [Sutskever et al., 2013], [Dozat, 2016] for descriptions of the optimization algorithms and to [Kalman and Kwasny, 1992], [Zeiler et al., 2013], [Xu et al., 2015], [Clevert et al., 2015] for detailed descriptions of the activation functions used in the field. In this paper, we use Exponential Linear Units (ELUs) that are known to prevent dead units (neurons systematically outputting zero), speeds up the learning process and perform with a better accuracy in classification according to [Clevert et al., 2015]. In this paper, only a description of multilayer perceptron is given.

B.5 Multilayers perceptron

A multilayer perceptron (MLP) simply consists of a stack of single layer perceptron units. They are theoretically able to solve any non-linear classification problem. The MLP can be of any complexity (number of inputs, size of layers and number of hidden layers). A MLP network consists of an input layer, one or several hidden layers (computational nodes) and an output layer. The input data flows through the network on a layer by layer basis (also called feedforward direction). They are usually trained using the popular backpropagation algorithm.

MLPs are used in the literature to perform pattern classification [Gardnera and Dorlinga, 1998].

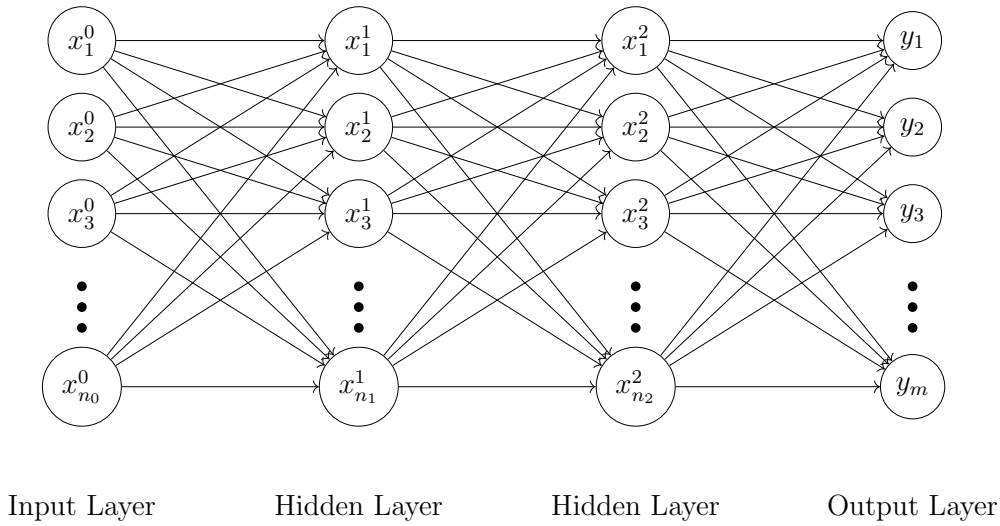


Figure B.2: Illustration of a fully-connected multilayer perceptron with 2-hidden layers. Note that the layers do not have necessarily the same dimension, i.e. $n_0 \neq n_1 \neq n_2 \neq m$

These kinds of networks have been used for face classification [Zhang et al., 1998] (although more efficient techniques are used today) or land-cover classification [Pacifi et al., 2009].

They are also used to perform air pollutant forecasting [Gardnera and Dorlinga, 1998]. MLPs have been shown to perform better than regression. [Navone and Ceccatto, 1994] evaluate the performance of a MLP to predict Indian monsoon rainfall whereas [Moseholm et al., 1996] uses MLP to predict the carbon monoxide concentration due to vehicle emissions at a urban intersection. It should be noted that in the studies mentioned previously the input data used are mainly meteorological data (wind, precipitation).

B.6 Results

B.6.1 Intra-vehicle real-world driving variability

Table B.1 presents the dynamic and emitting characteristics of the tests from the same vehicle falling in different clusters. This data was presented as Figure 3 in the main paper and discussed in Section 3.1.3. Exact numbers are given here for the sake of completeness. We remind the interested reader that vehicles 1, 3, 4, 5, 8, 9, 10, 12, 16 and 17 present variation in NO_x

emission factor by more than a factor of 1.4. On the other hand, vehicles 2, 6, 7, 11, 13, 14, 15 present variation in the NO_x emission factor by less than 1.2, which can be attributed to normal real-world variability.

Table B.2 shows the mean, the standard deviation and the coefficient of variation of the instantaneous NO_x emission time series of the tests from the same vehicle falling in different clusters. It can be observed that the mean, the standard deviation and the coefficient of variation of the instantaneous NO_x time series varies among the tests. In particular, tests of the vehicle classified in the same cluster present more similar mean, standard deviation and coefficient of variation than tests classified in different clusters. One example is the Nissan Qashqai: all the three tests present similar mean NO_x emissions but the test classified in cluster 8 presents a much higher standard deviation than the two tests classified in cluster 11. A similar comment can be made for the BMW 218D LUXURY 2014: while the three tests present a similar mean NO_x emission, the test associated to cluster 4 presents a significantly higher standard deviation than the two tests associated to cluster 10. Apart from the Nissan X-Trail, similar observations can be made for the other vehicles.

It can be observed that for all vehicles presenting differences in NO_x emission factor higher than 1.4 also presented significant differences in their NO_x time-series. Vehicles 1, 5, 8, 10, 12, 16 and 17 present only local variations in the emission time-series and local peaks in emissions can be noticed from one test to the other. On the other hand, vehicles 3, 4, 9 present very distinctive NO_x emission time series. For vehicles 1, 3, 8 and 12, local differences in traffic patterns, which are reflected in the instantaneous speed profile, constitute the most likely explanation for the discrepancies noticed between the tests. It was not possible to explain the discrepancies in emissions of the other vehicles based on driving dynamic or change in local traffic patterns. Vehicles 2, 6, 7, 11, 13, 14 and 15 present variation in emission factor smaller than 20% and striking differences in time-series are not noticeable at first glance. However, tests classified in different clusters present, according to Table B.2, variations in the mean and standard deviation of the instantaneous NO_x emissions, which allows to explain the results of the clustering. More

work is currently on-going to fully understand these discrepancies will be published in a future paper.

Table B.1: Dynamic and emitting characteristics of tests from the same vehicle falling in different clusters

	#	Test	Cluster	Avg. speed [km/h]	EF [g/km]	% of idle	KI [km ⁻¹]	RPA [m/s ²]	v*a _{pos} [95] [m ² /s ³]	Mean temp. [°C]
BMW - 220D M SPORT 2015 - LNT	1	1	4	19.4	0.07	29.0	1.77	0.17	8.8	21.7
		2	10	25.7	0.13	16.7	1.71	0.18	9.9	22.8
Citroen - Grand C4 Picasso - EGR	2	1	1	25.7	0.60	14.5	2.39	0.20	16.8	23.8
		2	11	25.8	0.53	13.8	2.16	0.17	15.5	23.9
Citroen - DS5 DSTYLE BLUEHDI 2014 - SCR	3	1	5	27.3	0.31	20.5	1.68	0.17	10.7	28.4
		2	11	21.9	0.58	19.0	2.26	0.21	10.5	28.8
BMW - 116D ED PLUS 2015 - LNT	4	1	4	20.0	0.11	28.3	2.71	0.23	21.0	17.5
		2	5	23.0	0.46	20.0	2.29	0.23	20.9	17.3
BMW - 220d SE - LNT	5	1	5	23.3	0.45	18.8	1.89	0.16	16.1	15.0
		2	10	22.9	0.22	25.2	2.15	0.17	17.7	14.6
Volkswagen - Passat - EGR	6	1	5	23.4	0.42	23.6	2.29	0.18	17.9	25.4
		2	11	21.1	0.55	24.8	2.55	0.22	25.0	25.4
Ford - S-Max TDCi Duratorq Titanium X - LNT	7	1	1	27.9	0.62	13.3	1.96	0.20	16.9	24.3
		2	11	26.4	0.54	18.2	1.78	0.18	10.7	24.2
Volkswagen - Golf SV SE TDi - LNT	8	1	4	26.3	0.09	16.2	1.69	0.16	8.3	22.3
		2	10	23.9	0.15	15.5	2.03	0.18	8.7	23.7
Volvo - S60 R-DESIGN LUX NAV D3 2015 - LNT	9	1	4	19.8	0.08	33.1	1.99	0.16	7.9	22.0
		2	10	22.0	0.16	26.8	2.07	0.18	9.1	21.2
BMW - X3 XDRIVE20D M SPORT 2014 - LNT	10	1	1	22.1	0.62	26.5	2.08	0.23	10.4	24.0
		2	6	23.8	0.97	16.2	2.57	0.23	11.0	25.0
Mercedes - C200 AMG Line BlueTec - SCR	11	1	5	20.9	0.44	23.2	3.11	0.21	9.7	22.4
		2	11	20.0	0.52	23.6	2.71	0.20	10.3	22.9
		3	11	22.0	0.53	22.6	2.75	0.21	11.2	19.3
Peugeot - 308 - EGR	12	1	13	23.5	1.43	24.3	2.06	0.18	17.4	27.3
		2	14	25.8	1.17	17.3	1.95	0.17	19.0	27.7
		3	15	25.7	1.66	20.9	1.99	0.21	18.5	27.6
Nissan - X-Trail - EGR	13	1	1	27.2	0.62	16.1	2.00	0.20	10.2	27.6
		2	8	25.0	0.80	18.2	1.90	0.21	11.3	27.3
		3	8	25.6	0.80	18.6	1.83	0.19	11.6	28.2
Nissan - Qashqai - EGR	14	1	8	23.4	0.76	24.0	1.62	0.18	10.7	19.7
		2	11	16.2	0.70	36.3	2.40	0.15	8.2	19.0
		3	11	22.0	0.69	17.0	2.30	0.19	9.1	20.8
Audi - Q3 - EGR	15	1	8	25.4	0.80	19.5	1.72	0.18	8.4	19.0
		2	11	27.1	0.62	10.2	1.47	0.11	5.9	19.8
		3	11	21.5	0.68	24.5	1.72	0.14	7.6	20.3
Volvo - V40 CROSS COUNTRY SE NAV 2014 - LNT	16	1	4	18.5	0.13	30.1	2.61	0.18	9.6	21.3
		2	9	23.9	0.90	14.3	2.15	0.15	8.5	22.5
		3	10	23.2	0.22	16.9	2.36	0.17	9.5	21.6
BMW - 218D LUXURY 2014 - LNT	17	1	4	23.9	0.20	24.3	1.90	0.21	13.2	26.9
		2	10	21.5	0.13	25.9	2.00	0.19	11.3	26.9
		3	10	23.4	0.20	22.9	1.84	0.18	11.0	26.6

B.6.2 NLR_1 coefficients

The NLR regression model, defined in Equation 4 in [Panis et al., 2006] is given by:

$$E_n(t) = f_1 + f_2 \cdot v_n(t) + f_3 \cdot v_n(t)^2 + f_4 \cdot a_n(t) + f_5 \cdot a_n(t)^2 + f_6 \cdot v_n(t) \cdot a_n(t) \quad (\text{B.1})$$

where $E_n(t)$ is the instantaneous emission at time t , $f_1, f_2, f_3, f_4, f_5, f_6$ are the emission constants specific for each vehicle and pollutant type, $v_n(t)$ is the instantaneous speed at time t , $a_n(t)$ is the instantaneous acceleration at time t . For NO_x emissions, two functional forms are derived, for acceleration (defined as $a_n(t) \geq -0.5\text{m/s}^2$) and deceleration (defined as $a_n(t) < -0.5\text{m/s}^2$).

Figure B.37 depicts the distribution of the 12 coefficients of the NLR_1 regression for 20 successive training runs. For each training run, 20% of the datafiles were selected randomly as a validation set, while the other 80% were used to fit the coefficients. It can be observed that the spread of the 12 coefficients is generally small. The biggest spreads can be observed for the coefficients of cluster 8, which can be attributed to the limited number of tests belonging to cluster 8: since cluster 8 only contains 5 tests (compared to 9, 19, 21, 18 and 24 for clusters 1, 4, 5, 10 and 11 respectively), randomly selecting out 20% means that only 4 tests are used to refit the coefficients. Results are consequently more sensitive to specific patterns in the data than for other clusters.

B.6.3 Clustering

A clustering analysis was also performed on CO and CO_2 . For comparison purposes with the NO_x clustering analysis, 17 clusters were defined in both cases, comprising six main clusters for CO_2 (regrouping 81% of the data) and eight main clusters for CO (regrouping 84% of the data).

Table B.2: Mean, standard deviation and coefficient of variation (CV) of the NO_x emission time serie for each vehicles whose tests are classified in different clusters

Vehicle	Test	Cluster	Mean NO _x	STD NO _x	CV
BMW - 220D M SPORT 2015 - LNT	1	4	0.00036	0.0019	5.13
	2	10	0.00094	0.0017	1.79
Citroen - Grand C4 Picasso - EGR	1	1	0.0043	0.0068	1.59
	2	11	0.0038	0.0051	1.35
Citroen - DS5 DSTYLE BLUEDEHI 2014 - SCR	1	5	0.0023	0.0032	1.40
	2	11	0.0035	0.0048	1.34
BMW - 116D ED PLUS 2015 - LNT	1	4	0.00059	0.0019	3.26
	2	5	0.0029	0.0051	1.77
BMW - 220d SE - LNT	1	10	0.0014	0.0018	1.31
	2	5	0.0029	0.0035	1.19
Volkswagen - Passat - EGR	1	5	0.0027	0.0034	1.25
	2	11	0.0032	0.0046	1.42
Ford - S-Max TDCi Duratorcq Titanium X - LNT	1	11	0.0040	0.0048	1.24
	2	1	0.0048	0.0053	1.10
Volkswagen - Golf SV SE TDi - LNT	1	4	0.00067	0.00088	1.31
	2	10	0.0010	0.0012	1.22
Volvo - S60 R-DESIGN LUX NAV D3 2015 - LNT	1	10	0.00099	0.0022	2.25
	2	4	0.00043	0.0012	2.88
BMW -X3 XDRIVE20D M SPORT AUTO - LNT	1	1	0.0038	0.0074	1.95
	2	6	0.0064	0.010	1.57
Mercedes - C200 AMG Line BlueTec - SCR	1	5	0.0025	0.0032	1.28
	2	11	0.0029	0.0044	1.49
	3	11	0.0032	0.0046	1.42
Peugeot - 308 - EGR	1	15	0.012	0.019	1.59
	2	14	0.0084	0.012	1.46
	3	13	0.0093	0.013	1.43
Nissan X-Trail - EGR	1	1	0.0047	0.0062	1.33
	2	8	0.0055	0.0069	1.24
	2	8	0.0057	0.0075	1.31
Nissan - Qashqai - EGR	1	11	0.0032	0.0047	1.48
	2	8	0.0049	0.0070	1.40
	3	11	0.0042	0.0052	1.25
Audi - Q3 - EGR	1	8	0.0057	0.0069	1.21
	2	11	0.0047	0.0053	1.13
	2	11	0.0041	0.0043	1.07
Volvo - V40 CROSSCOUNTRY SE NAV 2014 - LNT	1	4	0.00067	0.0011	1.58
	2	10	0.0014	0.0033	2.33
	3	9	0.0059	0.025	4.17
BMW - 218D LUXURY 2014 - LNT	1	10	0.00081	0.0036	4.51
	2	4	0.0013	0.0095	7.21
	3	10	0.0013	0.0030	2.35

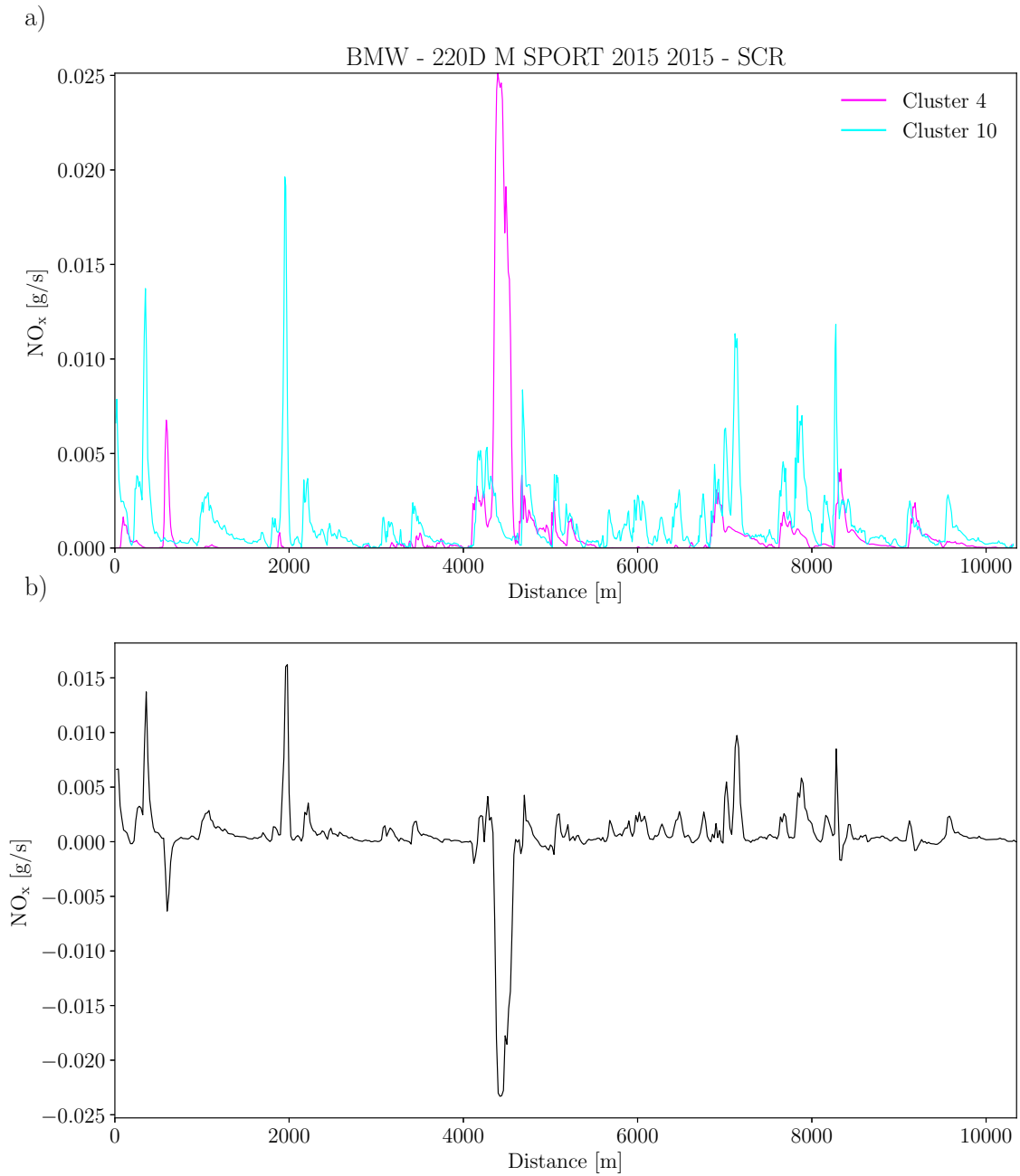


Figure B.3: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 4 and the test scenario is the test classified in cluster 10.

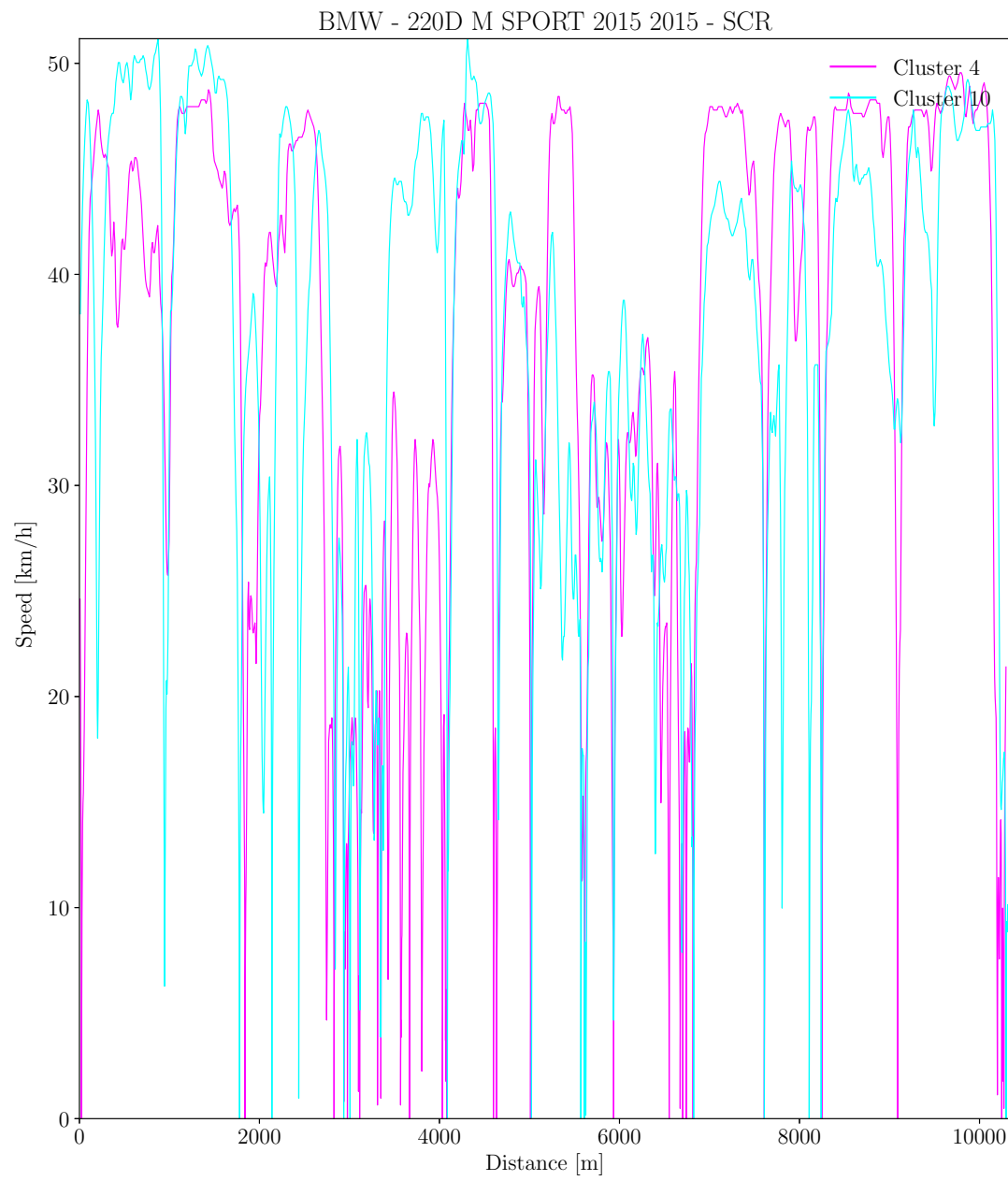


Figure B.4: Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.

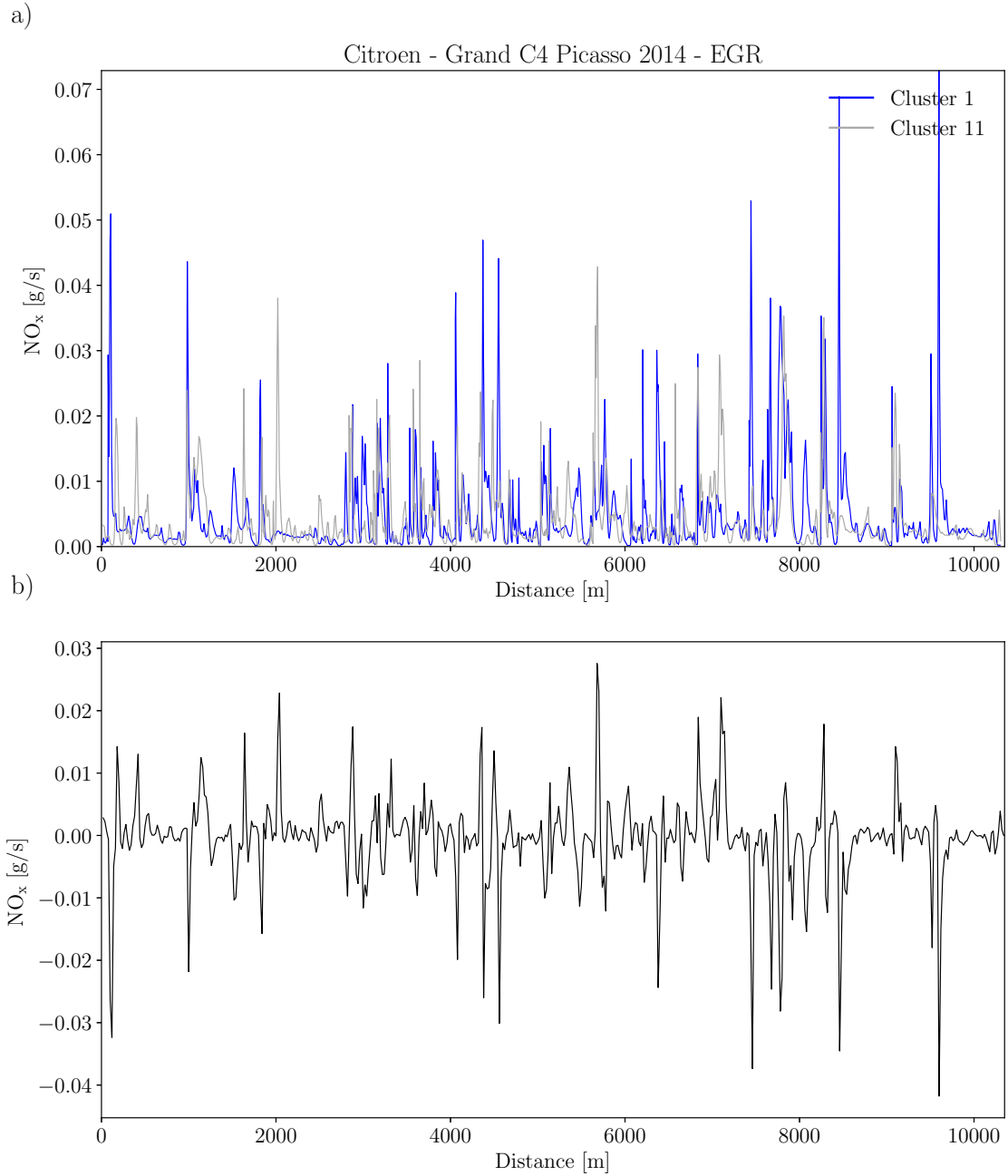


Figure B.5: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Citroen Grand C4 Picasso. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 1 and the test scenario is the test classified in cluster 11.

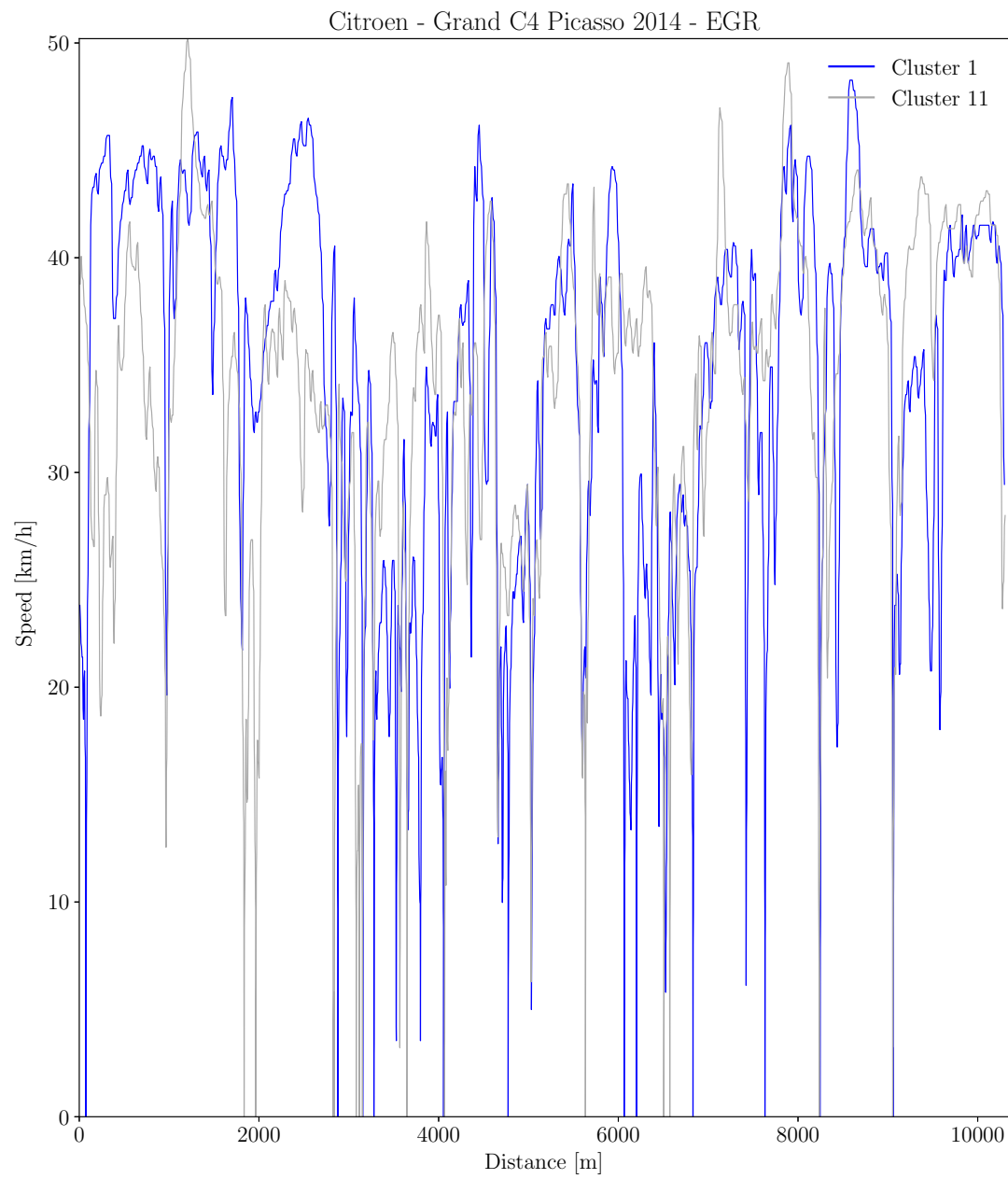


Figure B.6: Instantaneous speed as a function of the distance travelled as a function of the distance travelled for the Citroen Grand C4 Picasso.

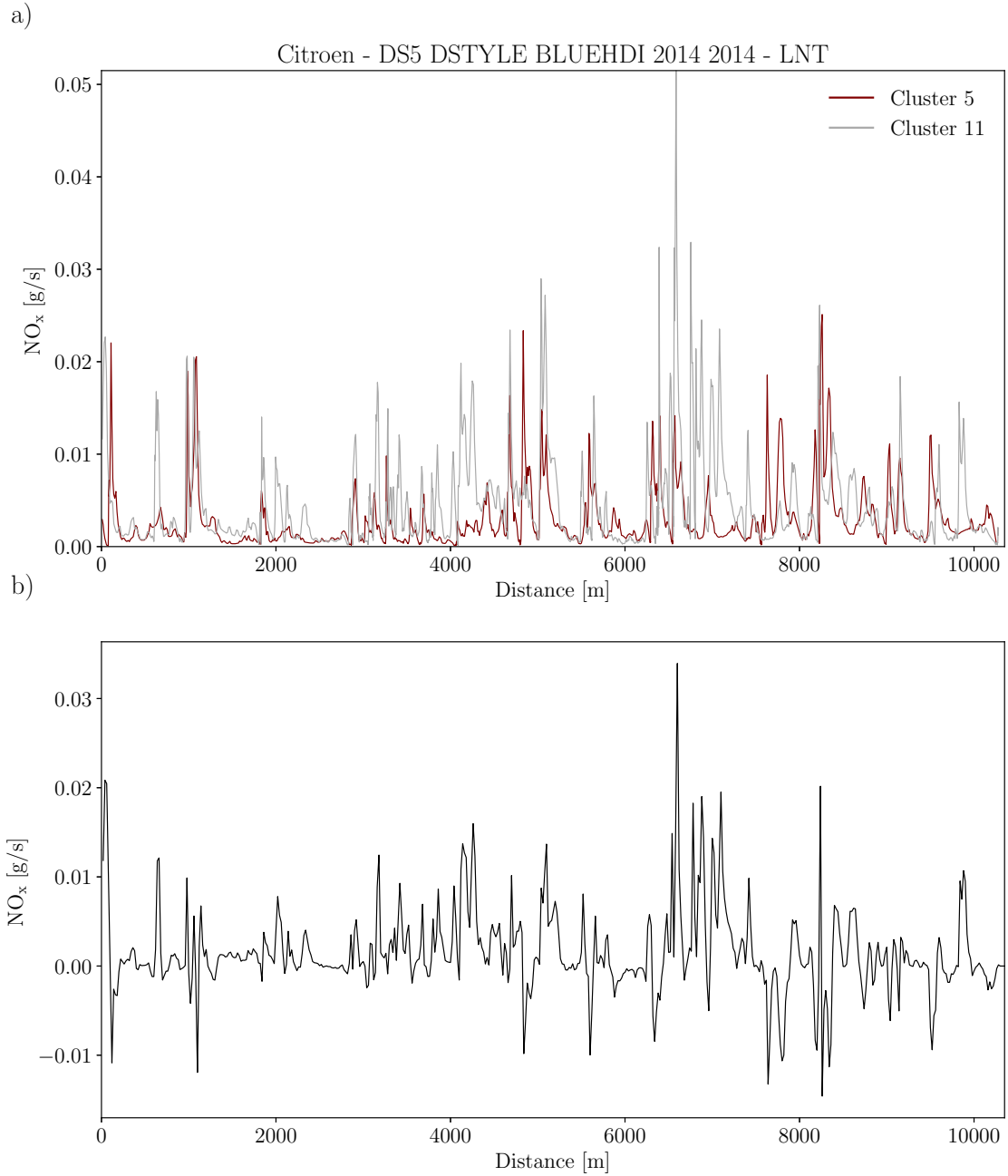


Figure B.7: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Citroen DS5 DSTYLE BLUEHDI. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 5 and the test scenario is the test classified in cluster 11.

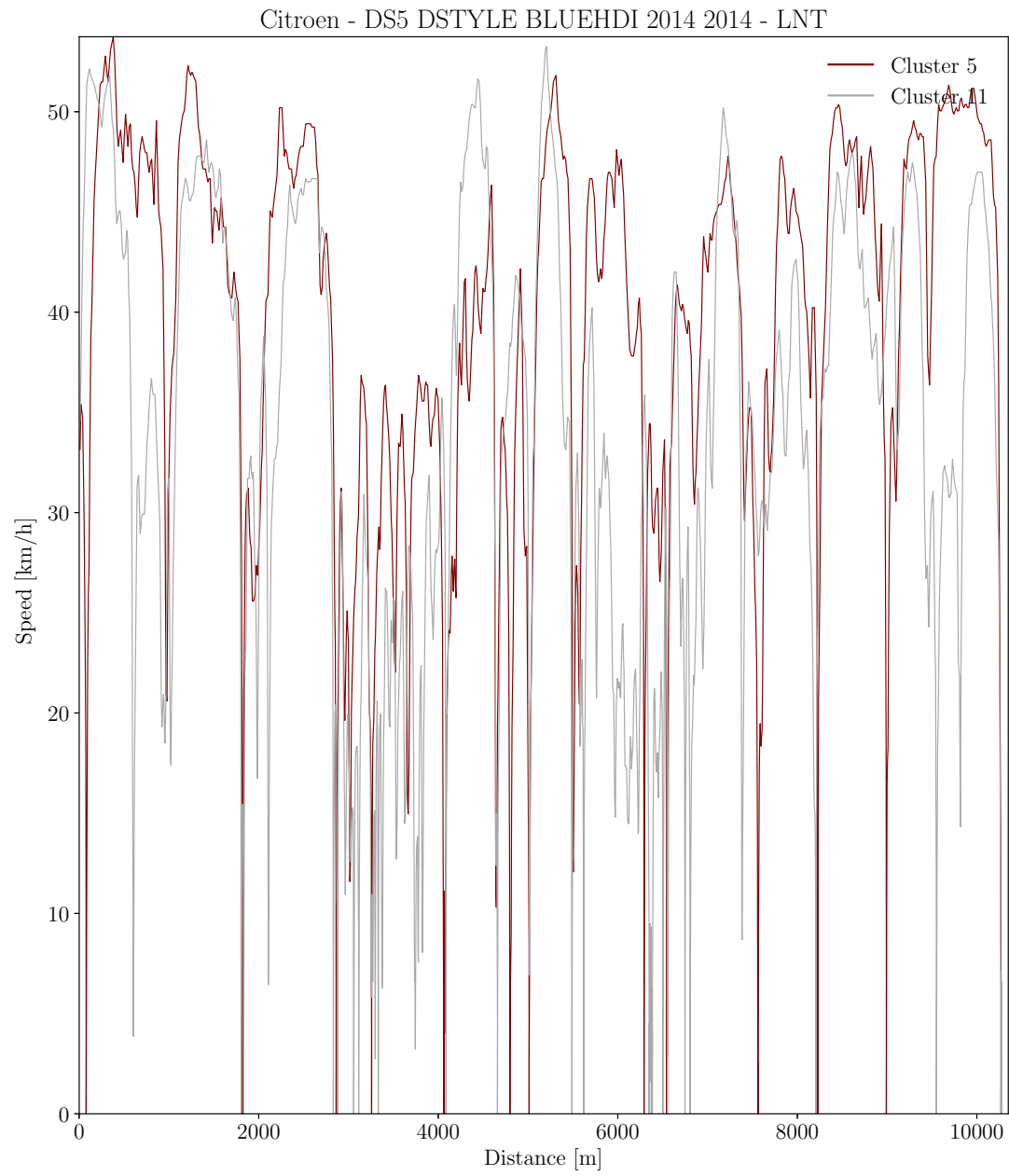


Figure B.8: Instantaneous speed as a function of the distance travelled for the Citroen DS5 DSTYLE BLUEHDI.

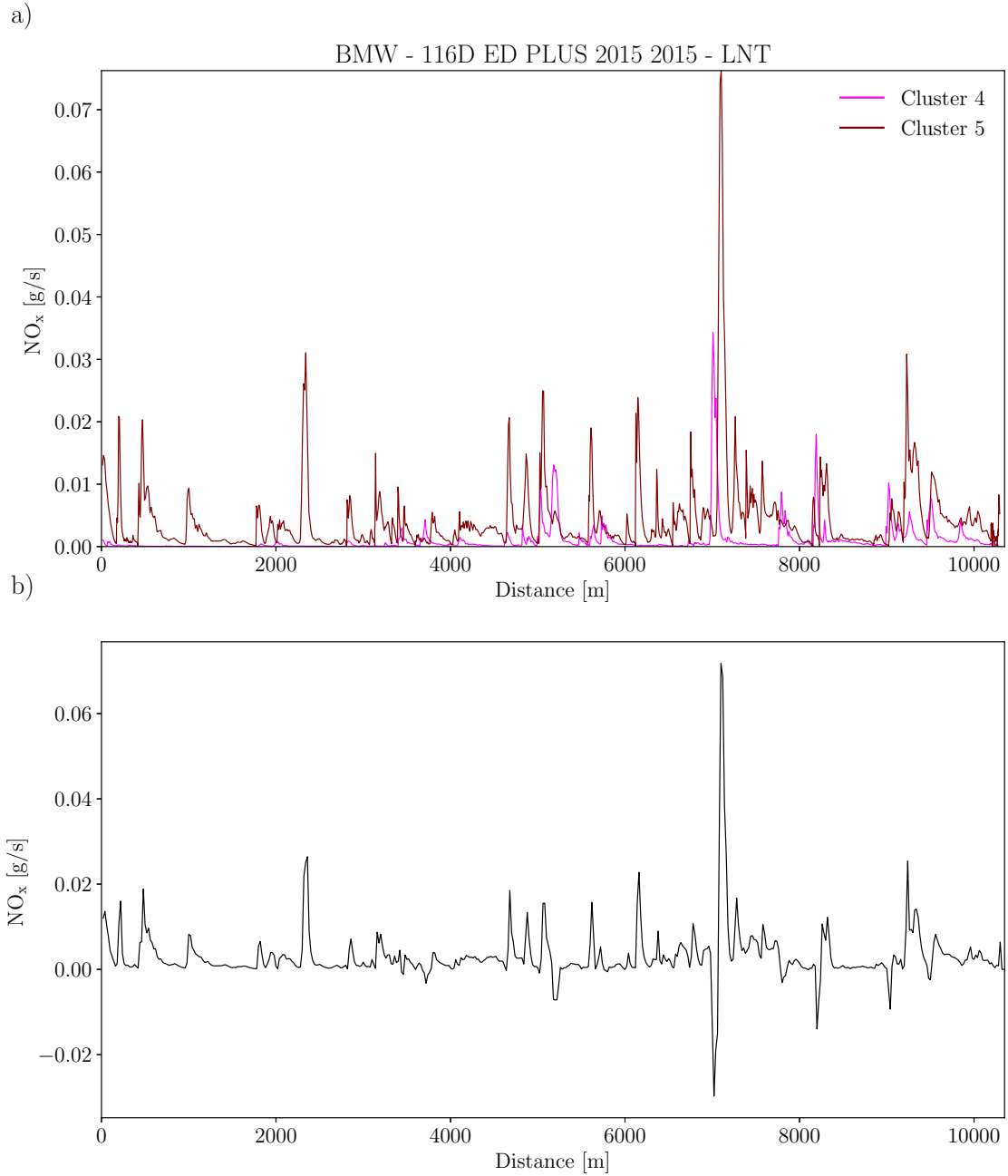


Figure B.9: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 116D ED PLUS. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 4 and the test scenario is the test classified in cluster 5.

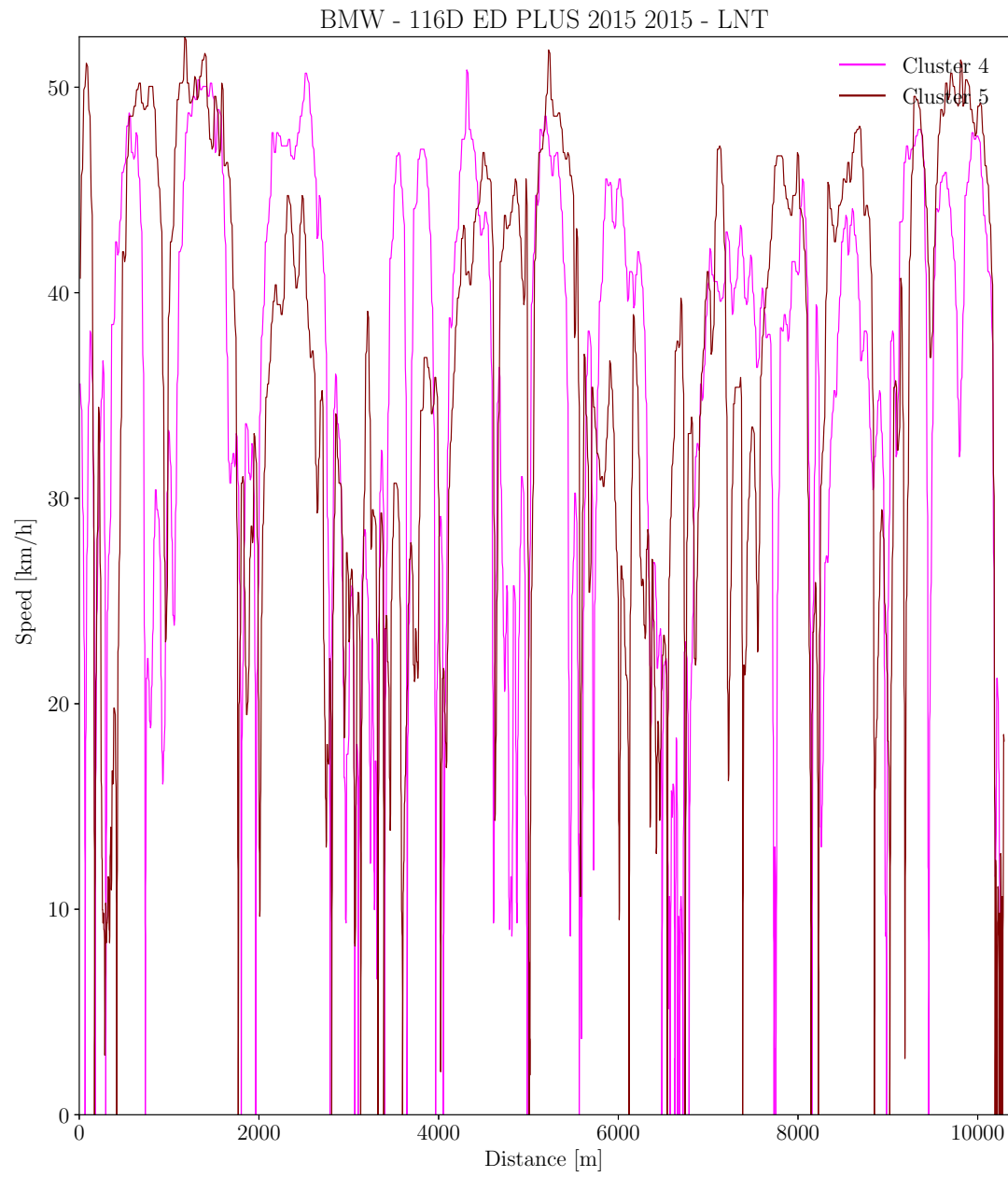


Figure B.10: Instantaneous speed as a function of the distance travelled for the BMW 116D ED PLUS.

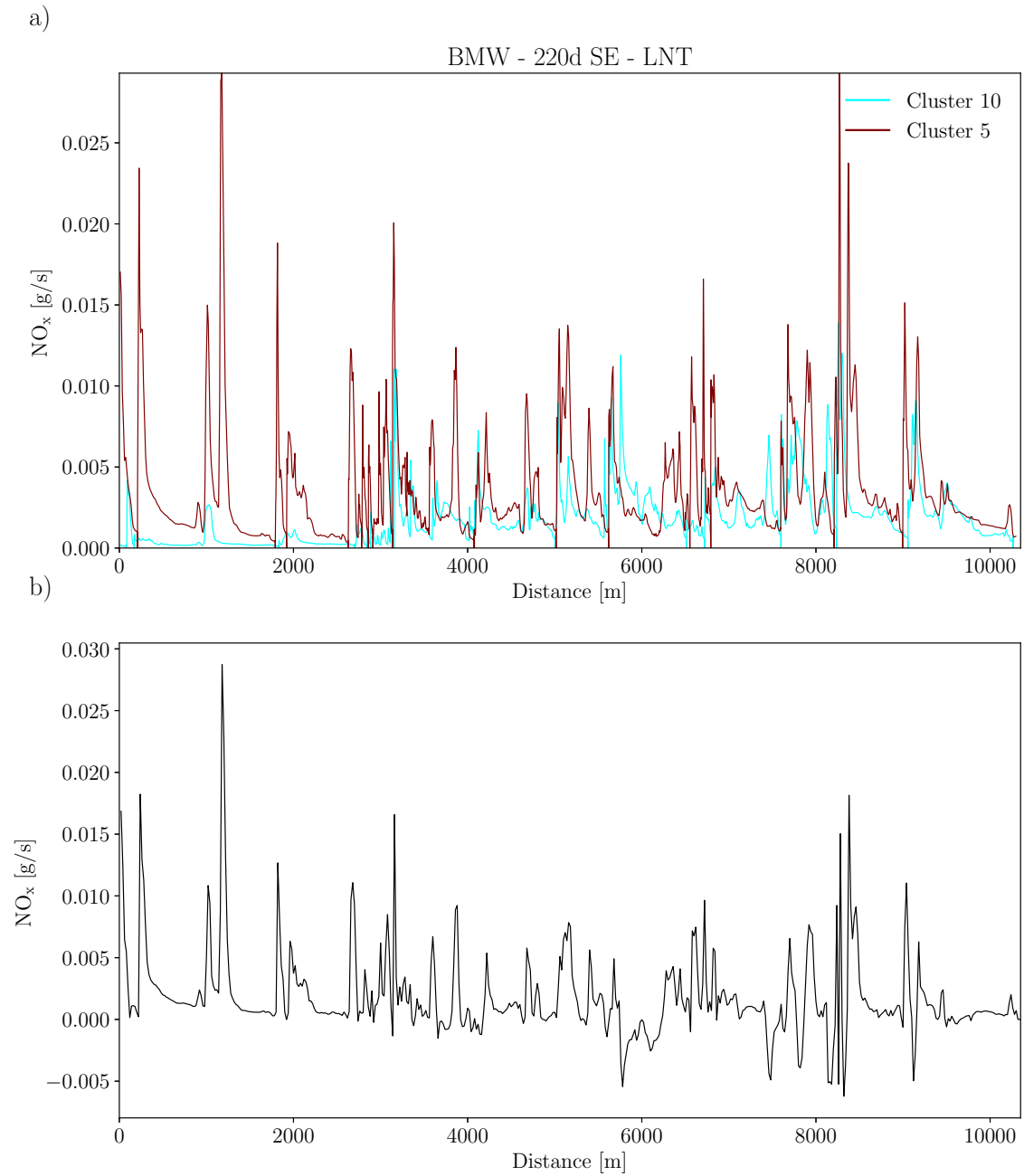


Figure B.11: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 220d SE. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 10 and the test scenario is the test classified in cluster 5.

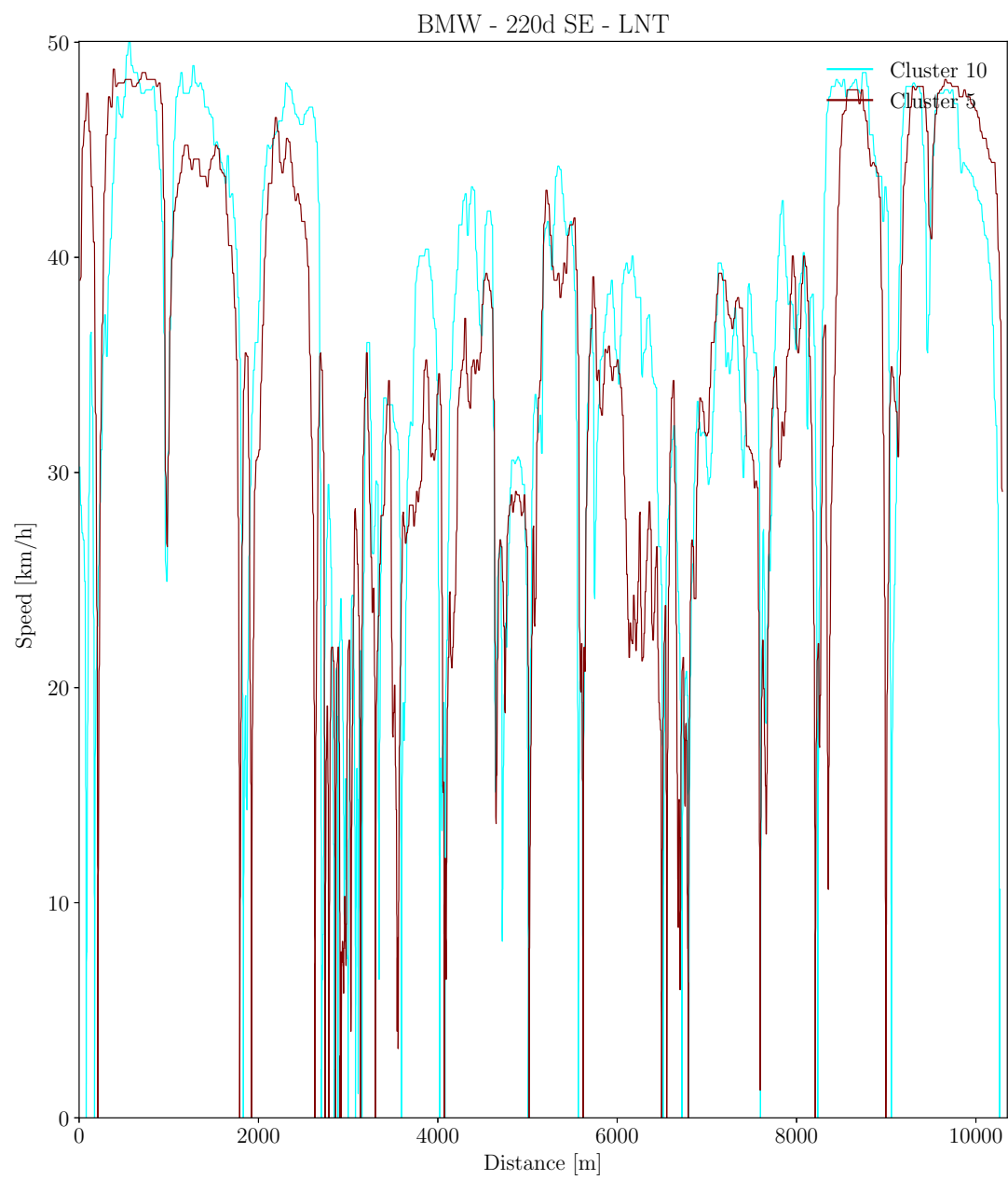


Figure B.12: Instantaneous speed as a function of the distance travelled for the BMW 220d SE.

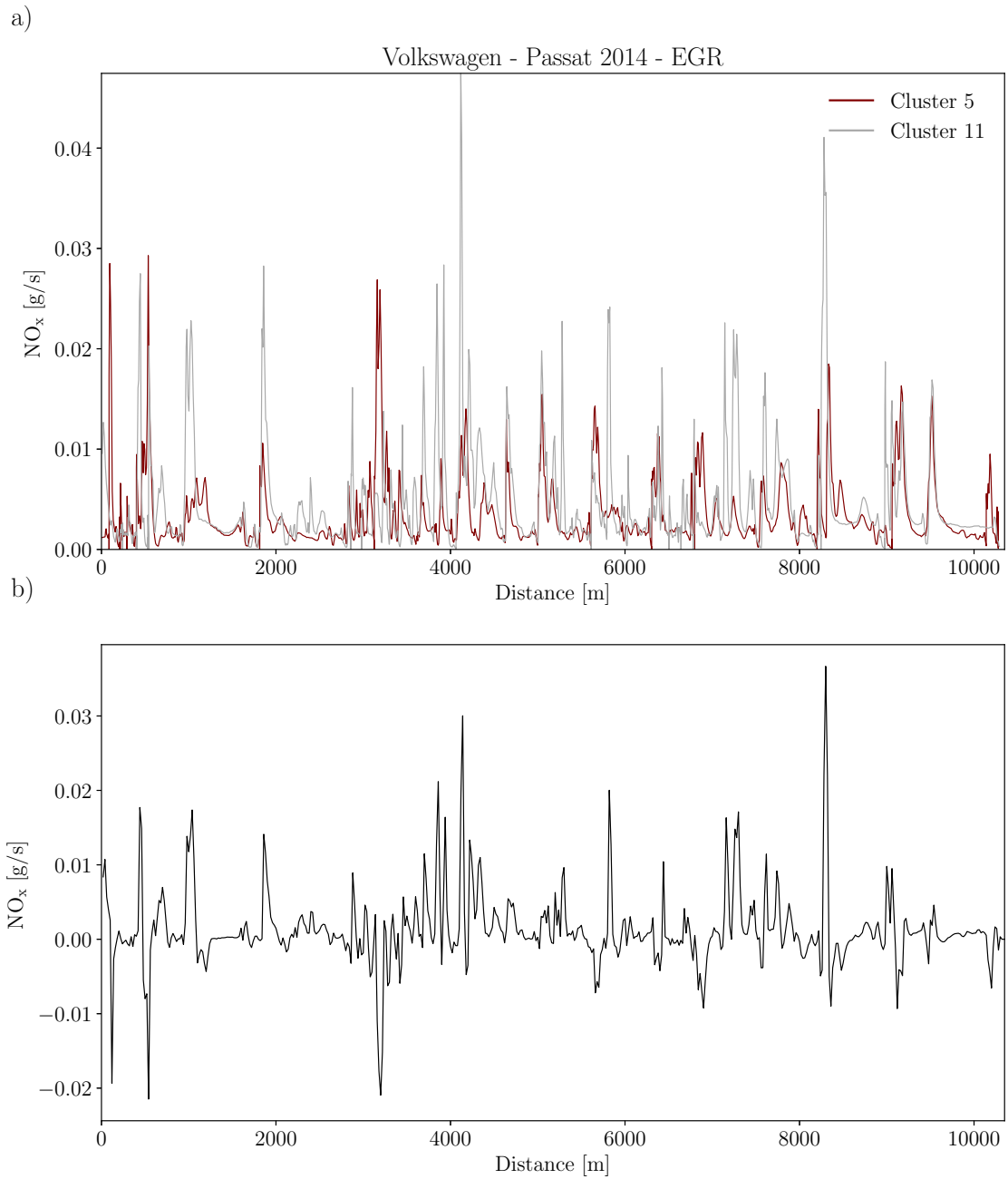


Figure B.13: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Volkswagen Passat. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 5 and the test scenario is the test classified in cluster 11.

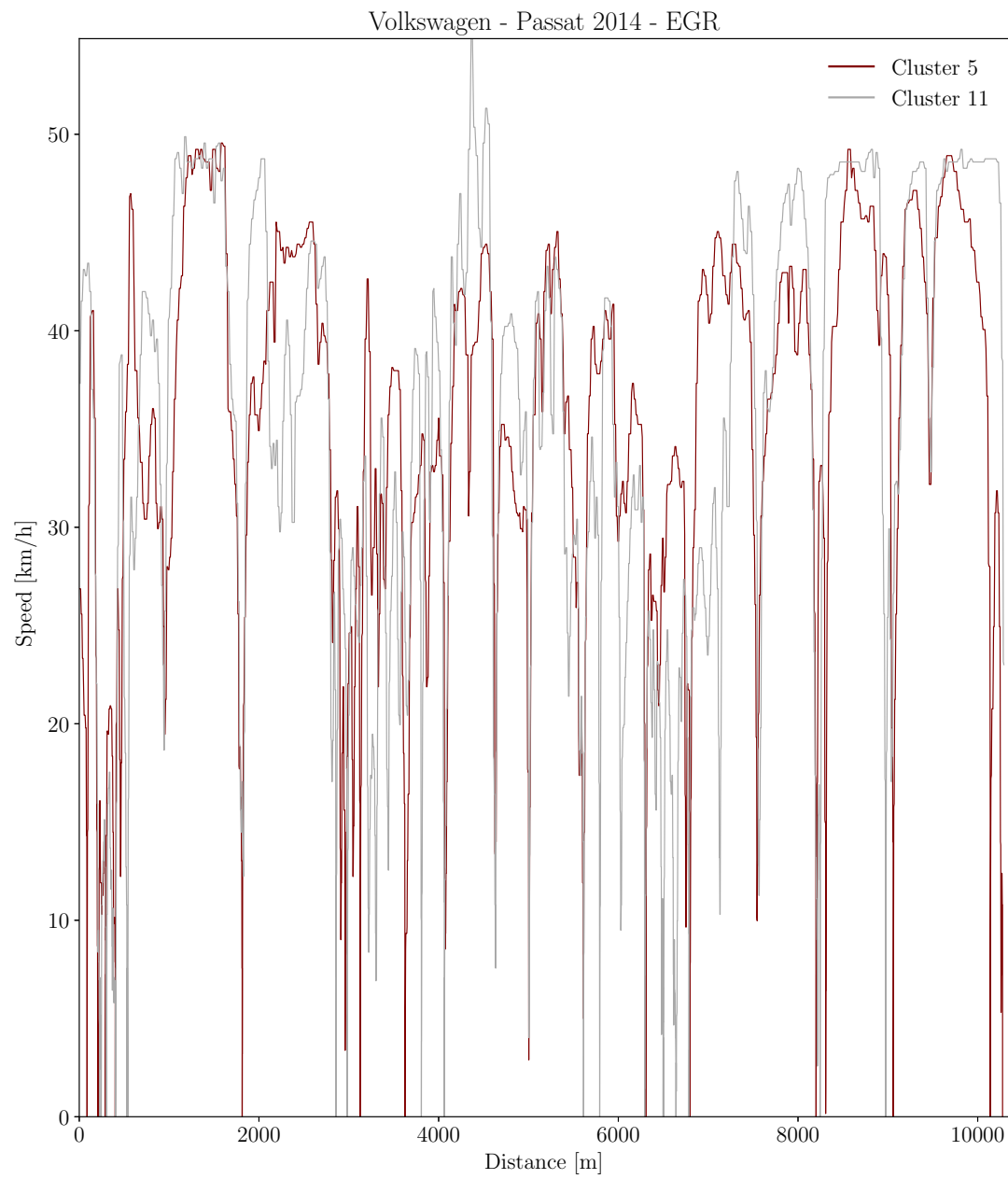


Figure B.14: Instantaneous speed as a function of the distance travelled for the Volkswagen Passat.

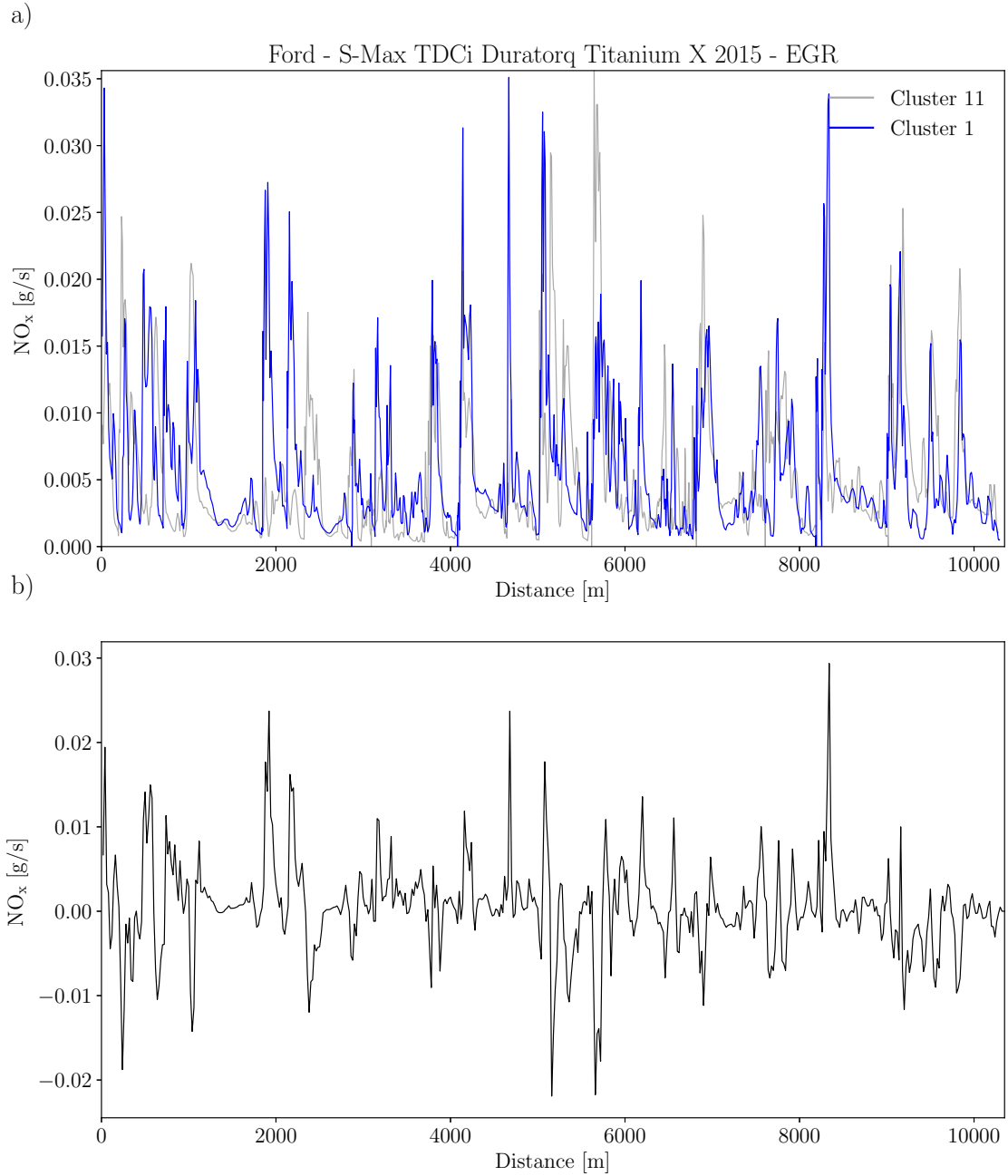


Figure B.15: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Ford S-Max TDCi Duratorq Titanium X. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 11 and the test scenario is the test classified in cluster 1.

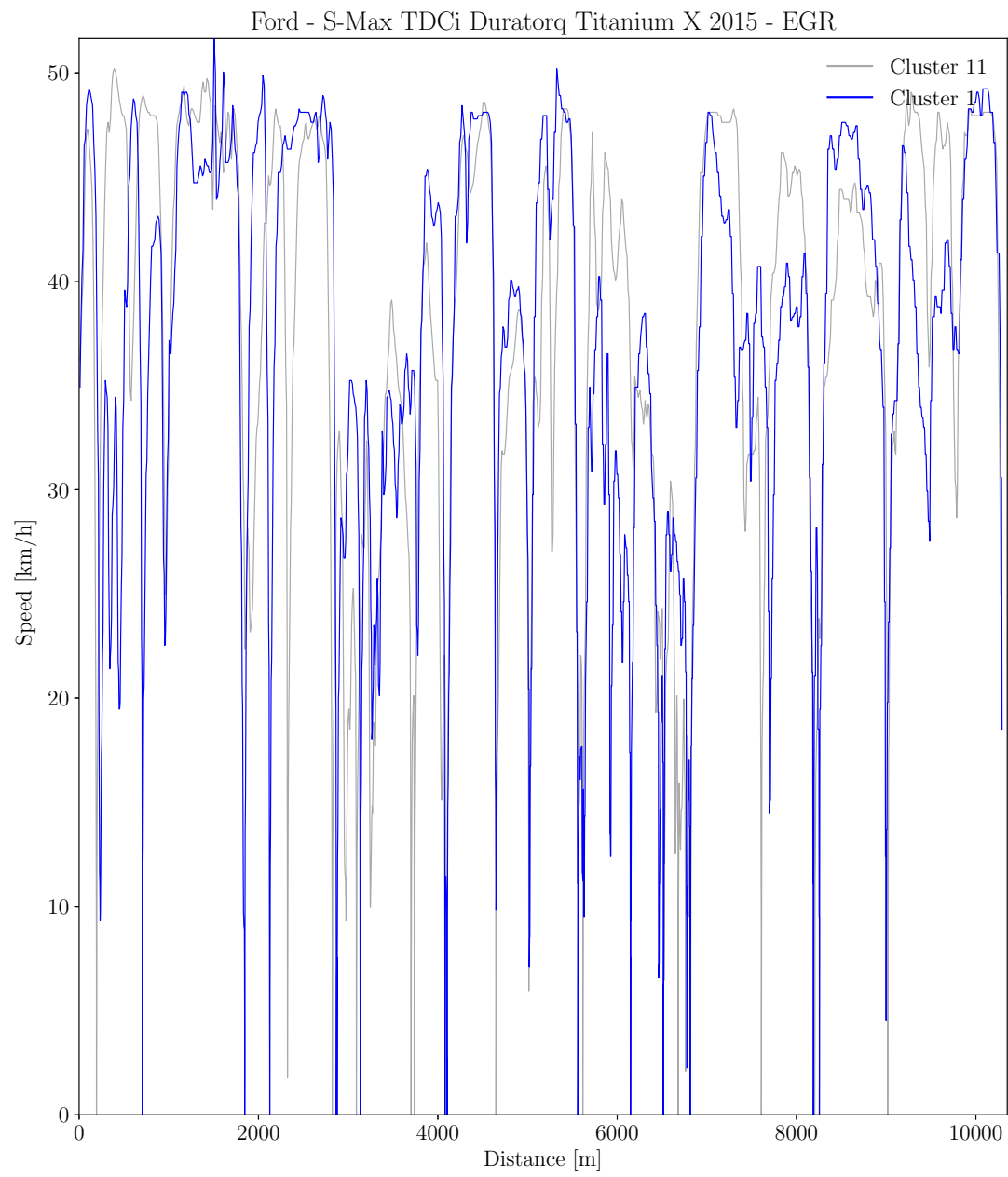


Figure B.16: Instantaneous speed as a function of the distance travelled for the Ford S-Max TDCi Duratorq Titanium X.

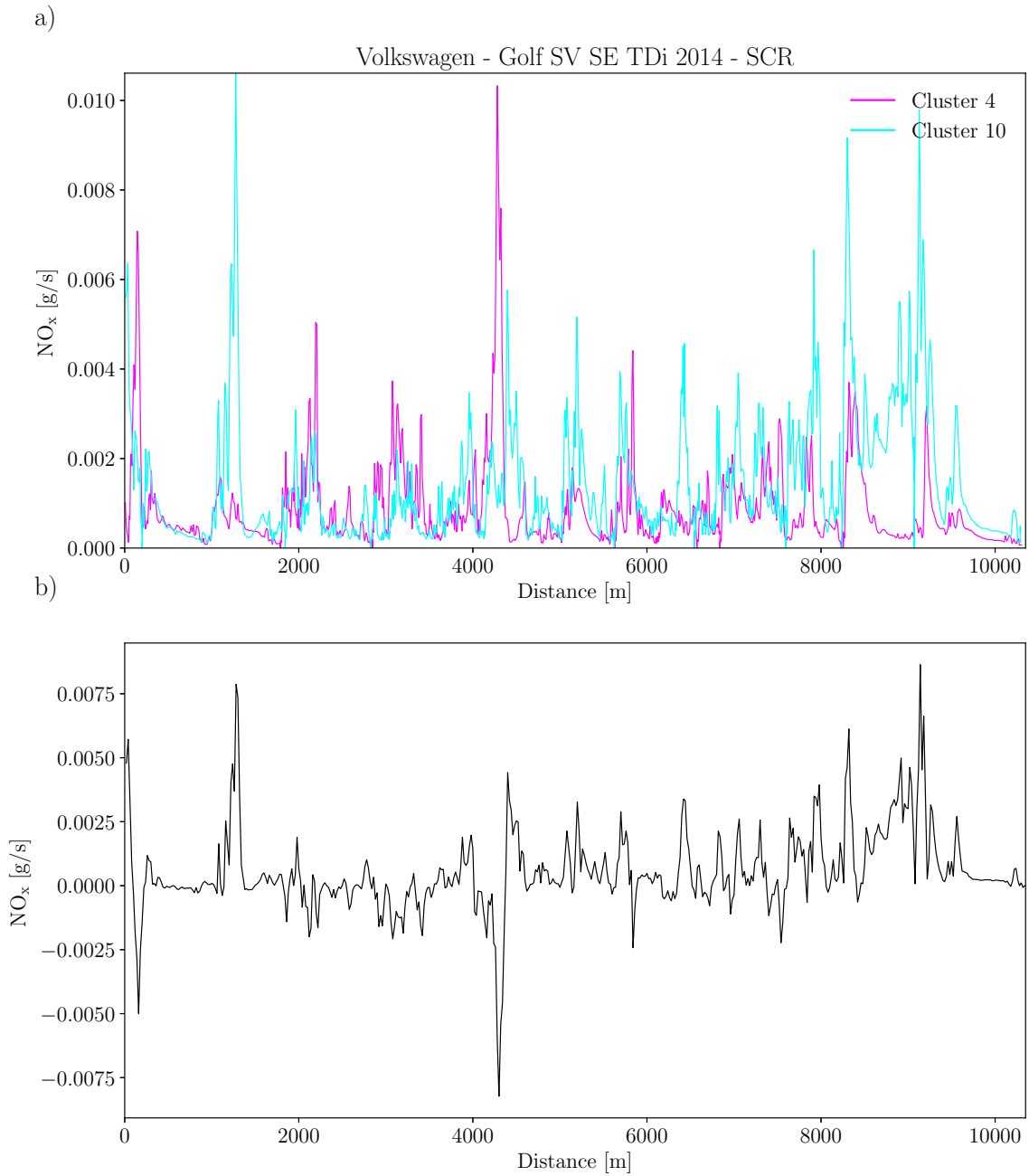


Figure B.17: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Golf SV SE TDi. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 4 and the test scenario is the test classified in cluster 10.

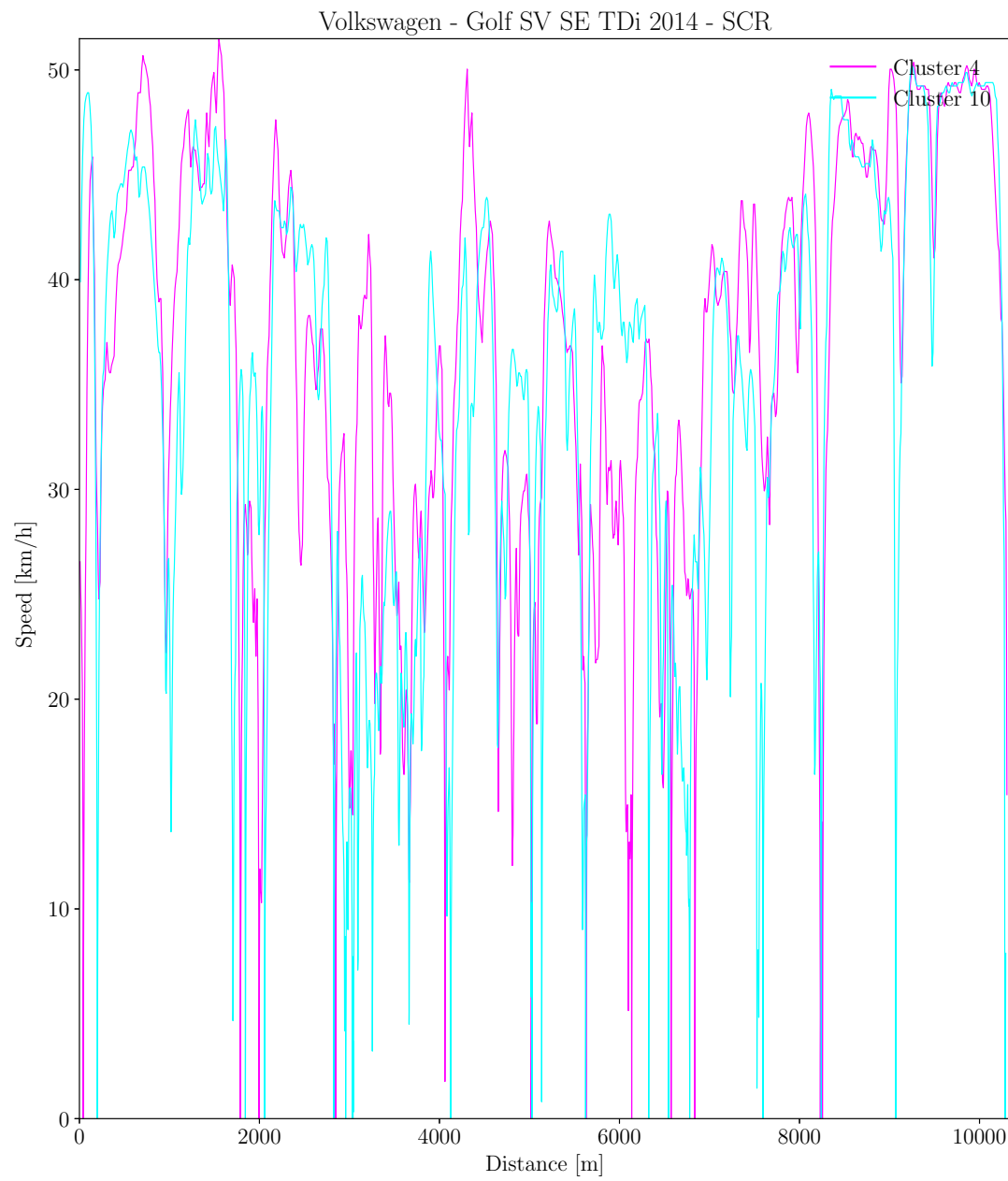


Figure B.18: Instantaneous speed as a function of the distance travelled for the Golf SV SE TDi.

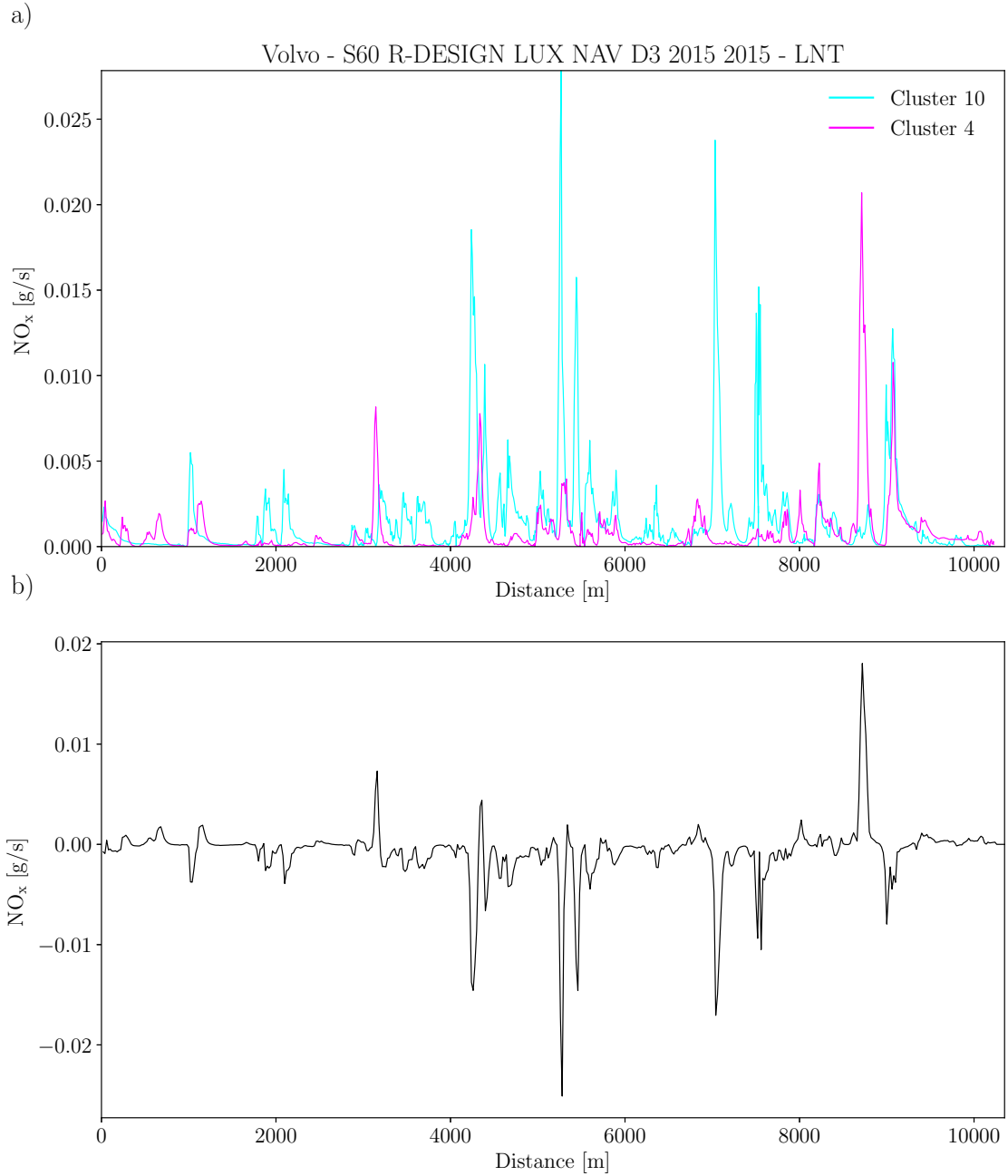


Figure B.19: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Volvo S60 R-DESIGN LUX NAV D3. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 10 and the test scenario is the test classified in cluster 4.

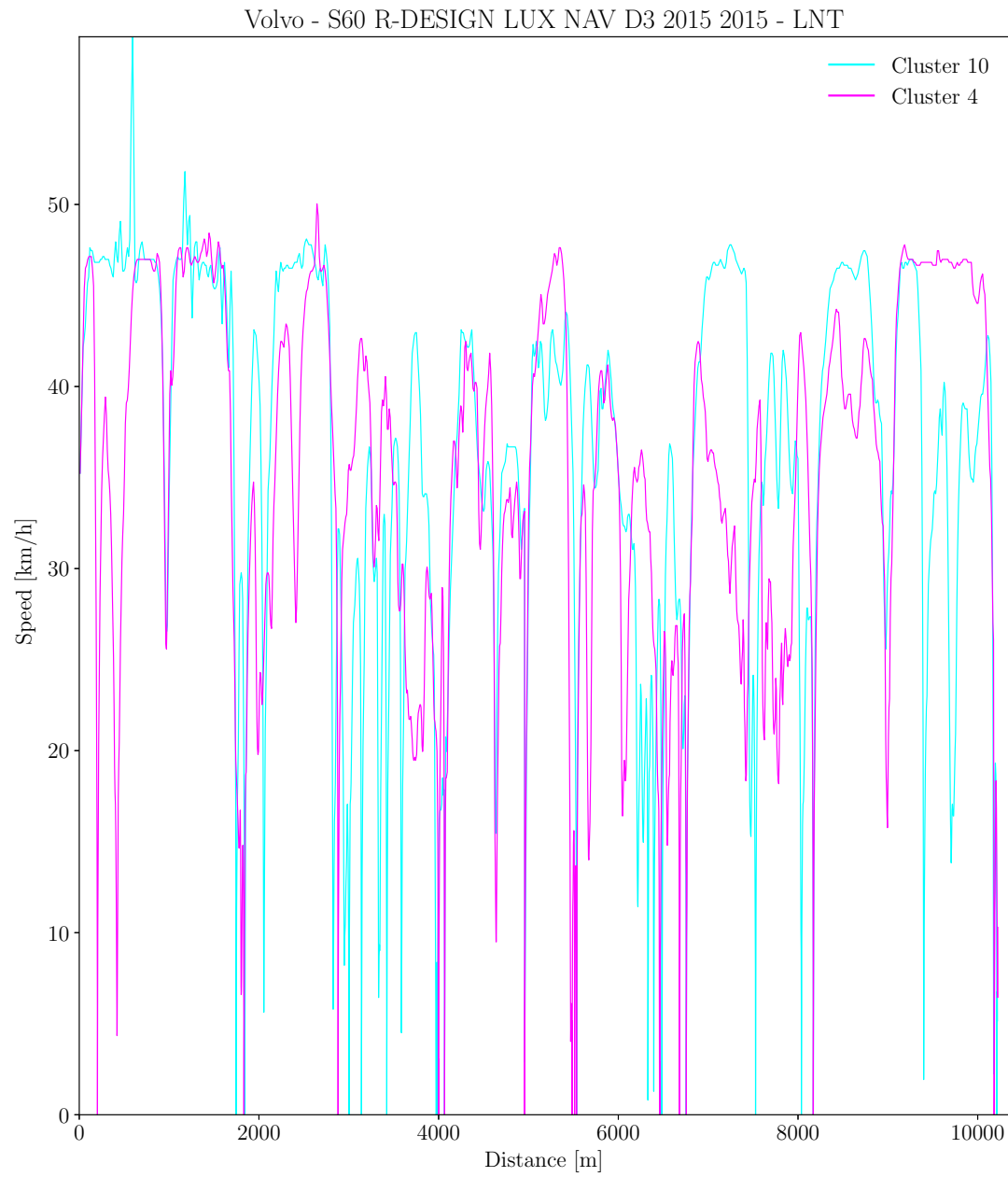


Figure B.20: Instantaneous speed as a function of the distance travelled for the Volvo S60 R-DESIGN LUX NAV D3.

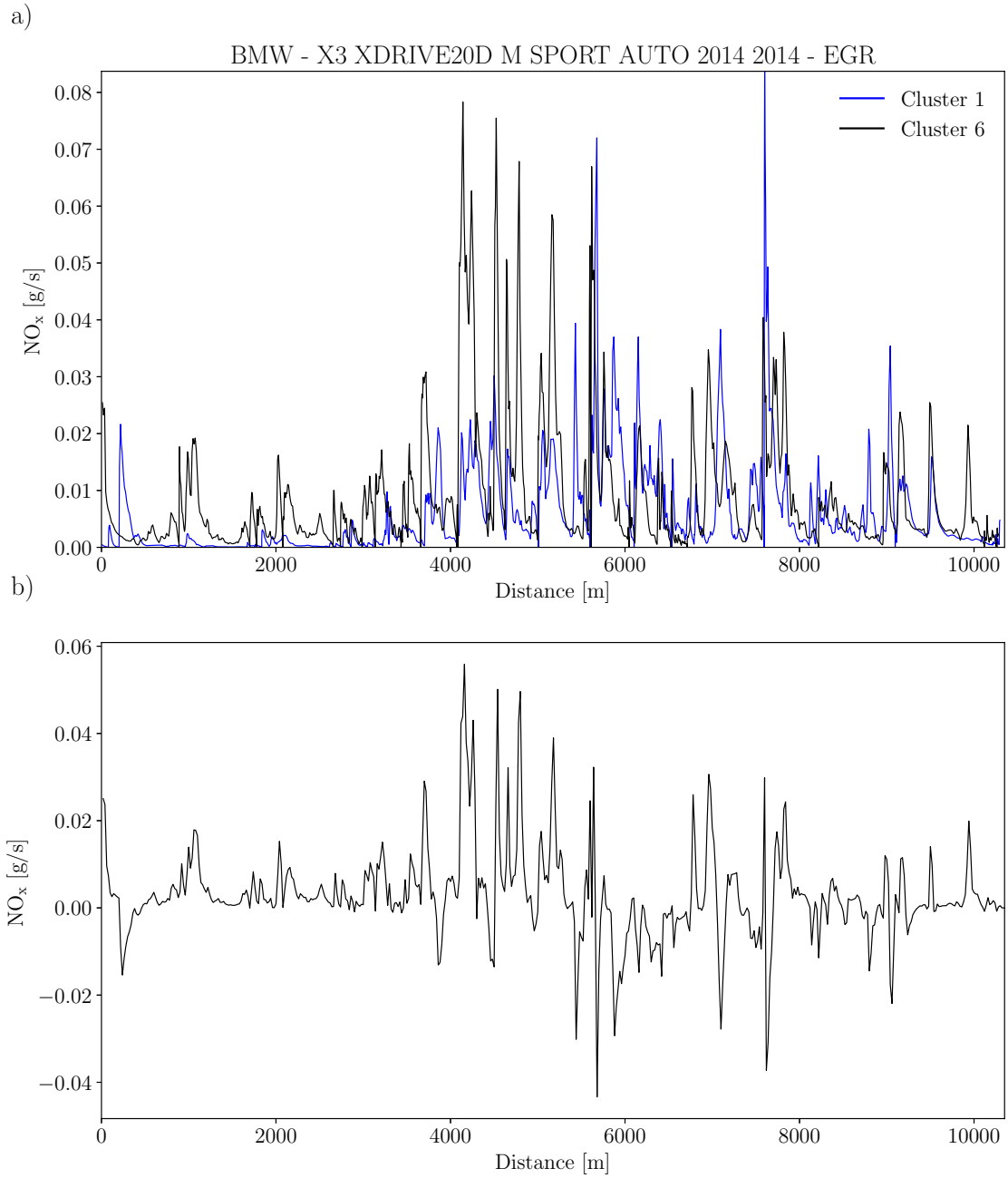


Figure B.21: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW X3 XDRIVE20D M SPORT AUTO. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 1 and the test scenario is the test classified in cluster 6.

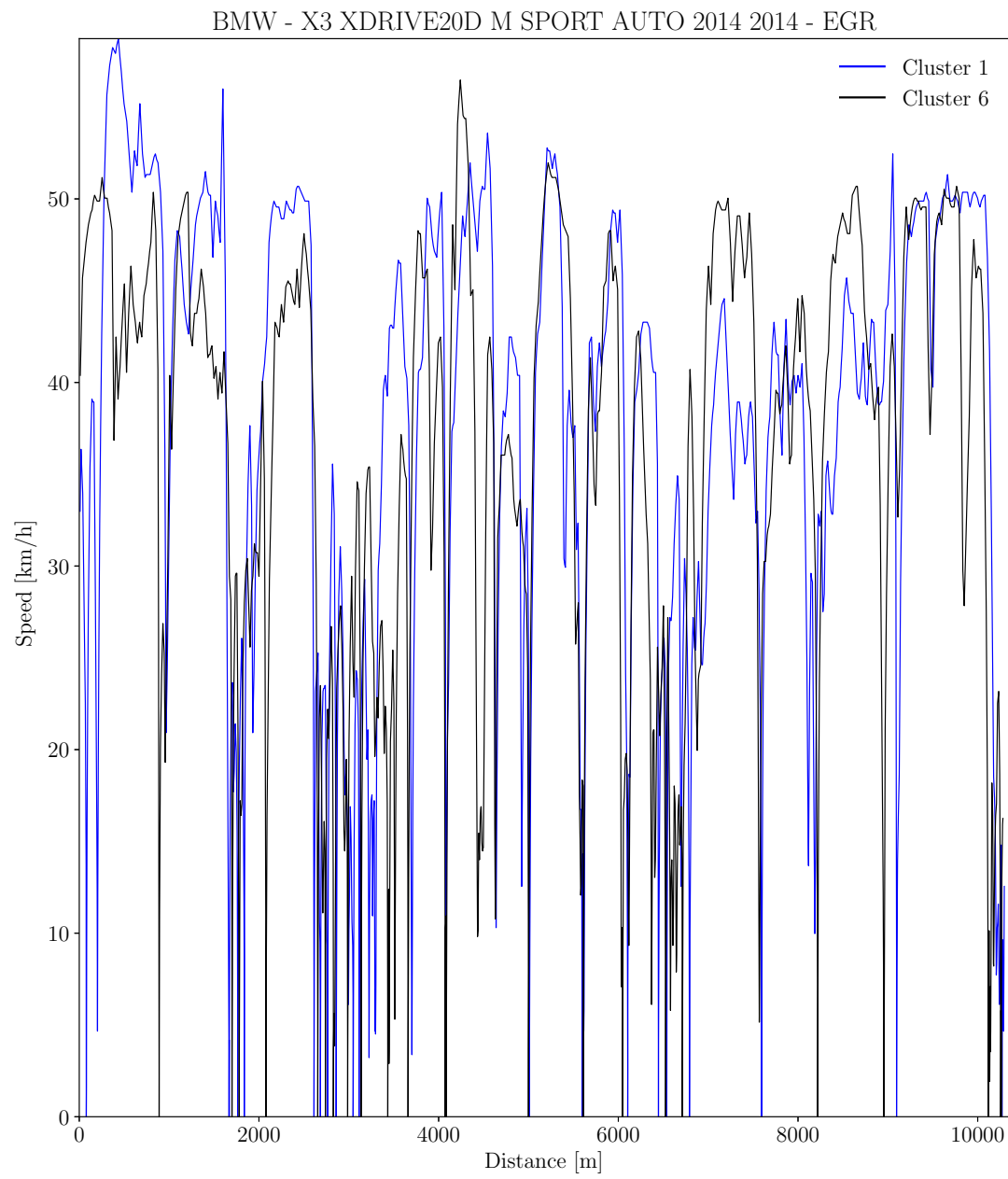


Figure B.22: Instantaneous speed as a function of the distance travelled for the BMW X3 XDRIVE20D M SPORT AUTO.

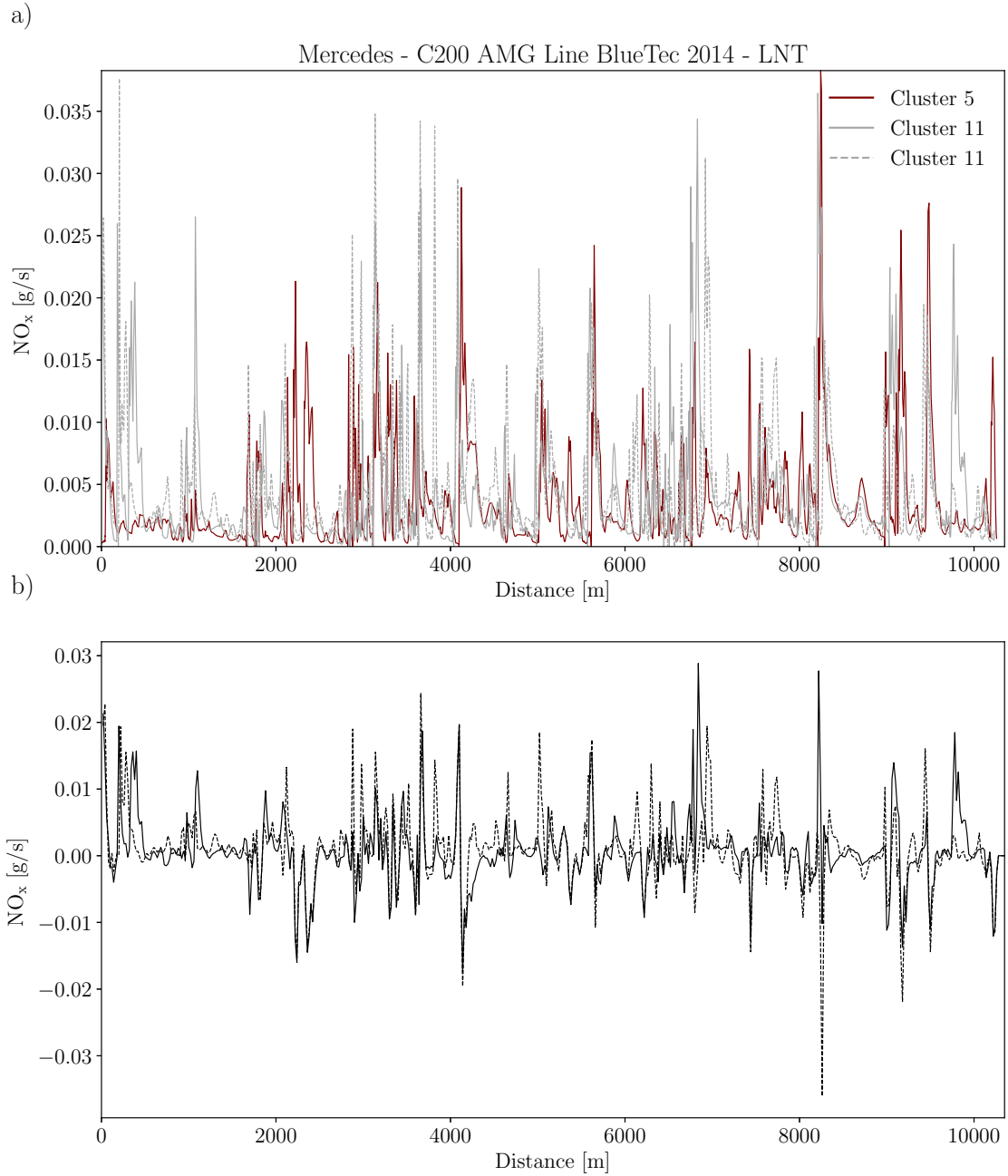


Figure B.23: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Mercedes C200 AMG Line BlueTec. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 4 and the tests scenario are the tests classified in cluster 11.

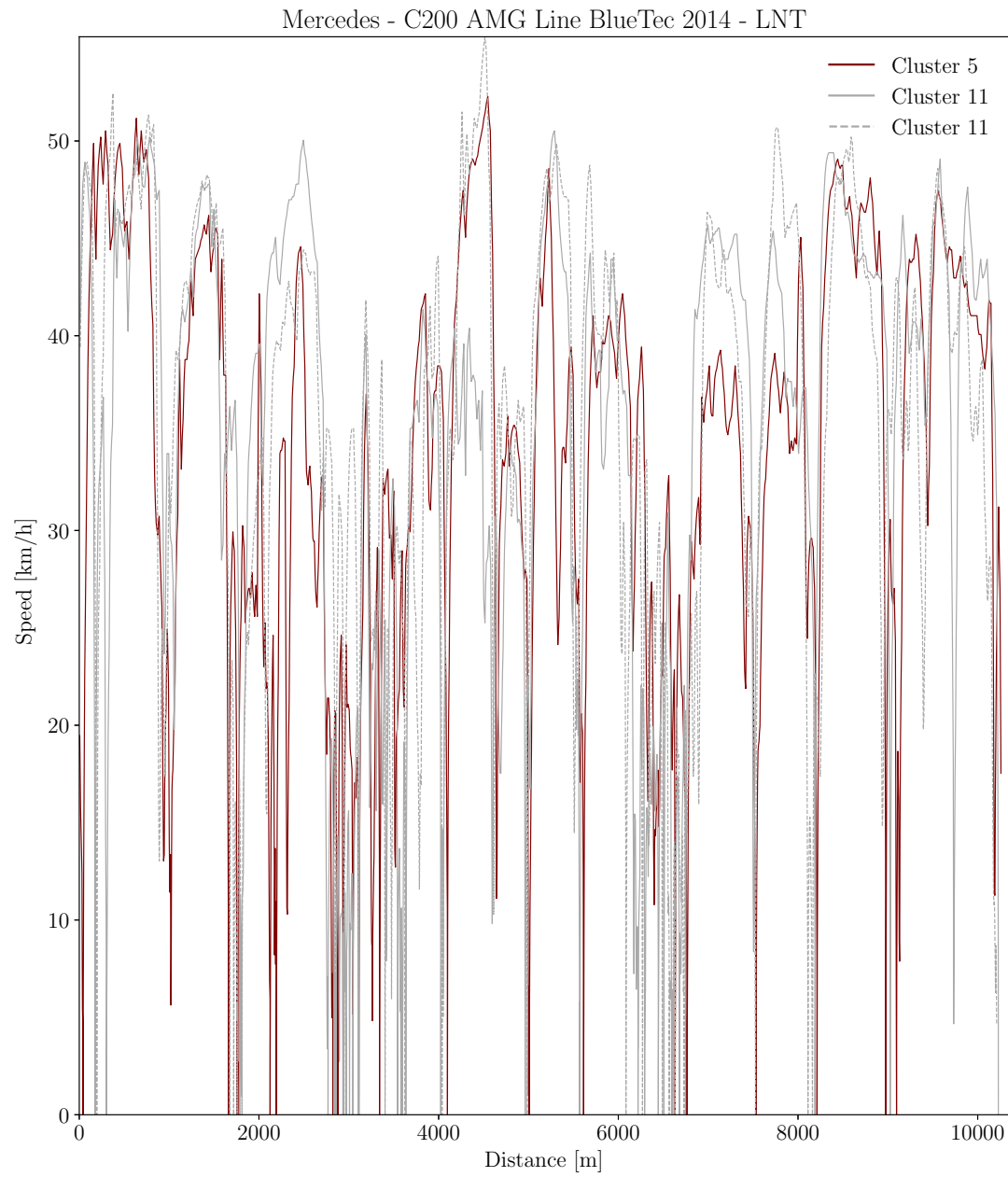


Figure B.24: Instantaneous speed as a function of the distance travelled for the Mercedes C200 AMG Line BlueTec.

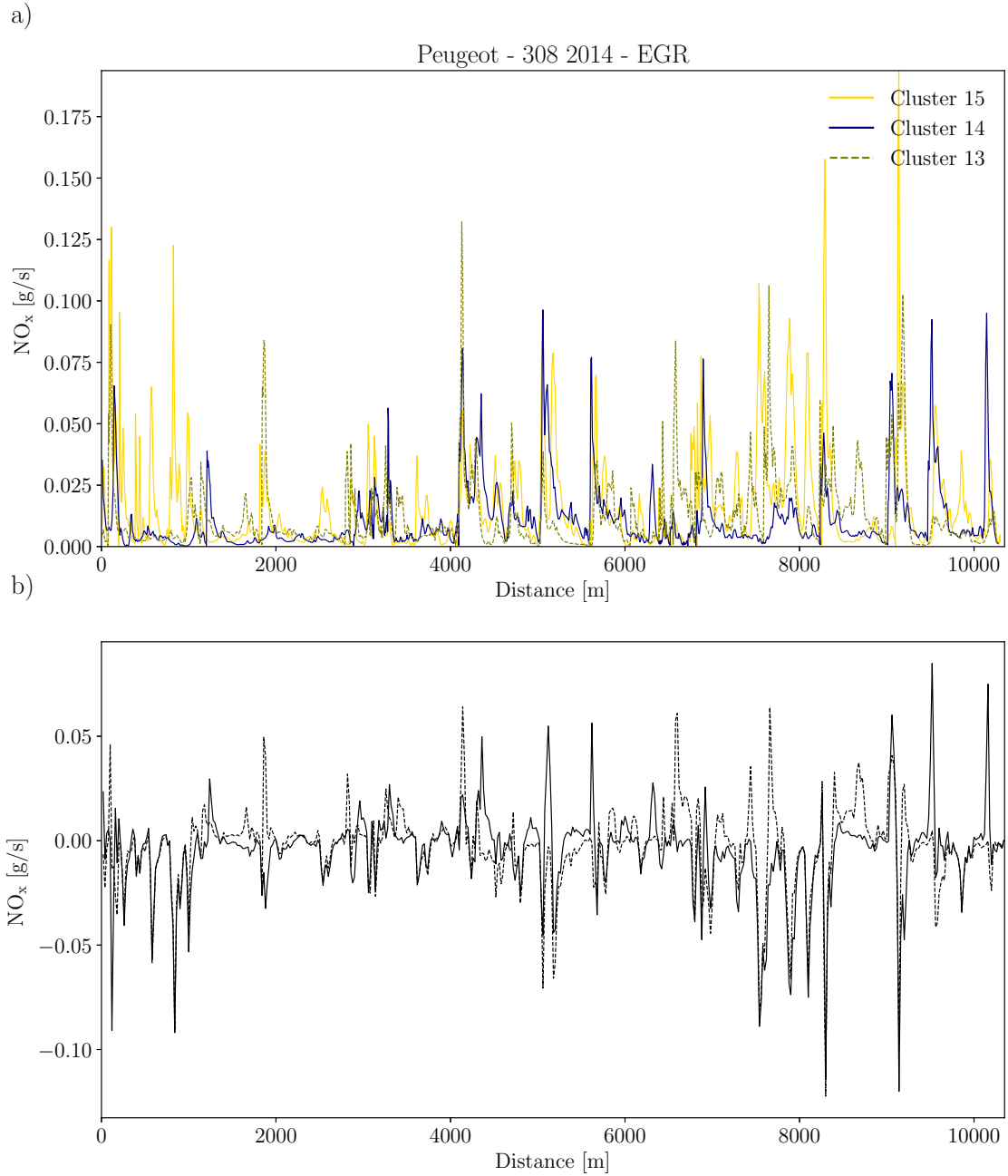


Figure B.25: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Peugeot 308. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 15 and the tests scenario are the tests classified in cluster 13 and 14.

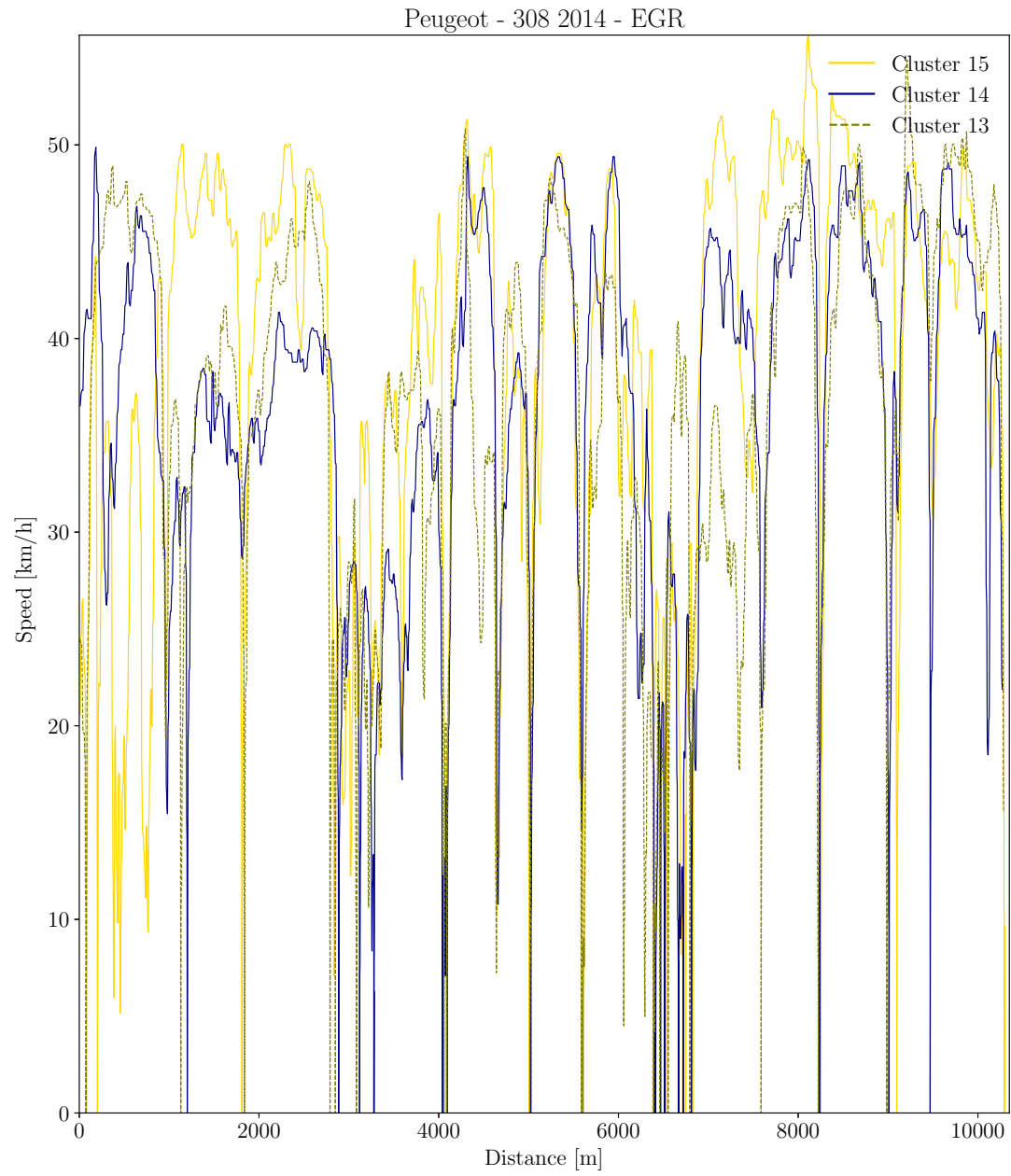


Figure B.26: Instantaneous speed as a function of the distance travelled for the Peugeot 308.

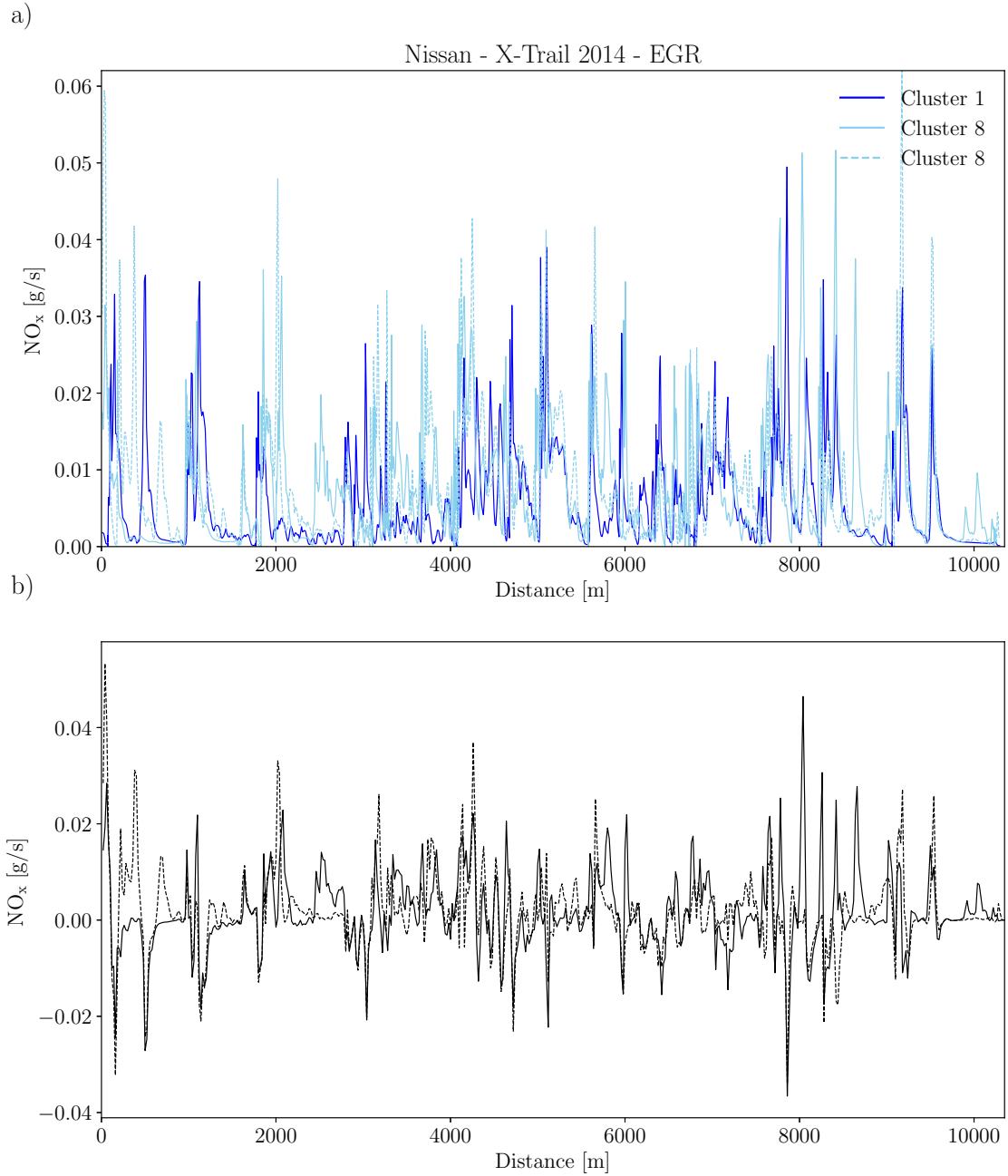


Figure B.27: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 1 and the tests scenario are the tests classified in cluster 8.

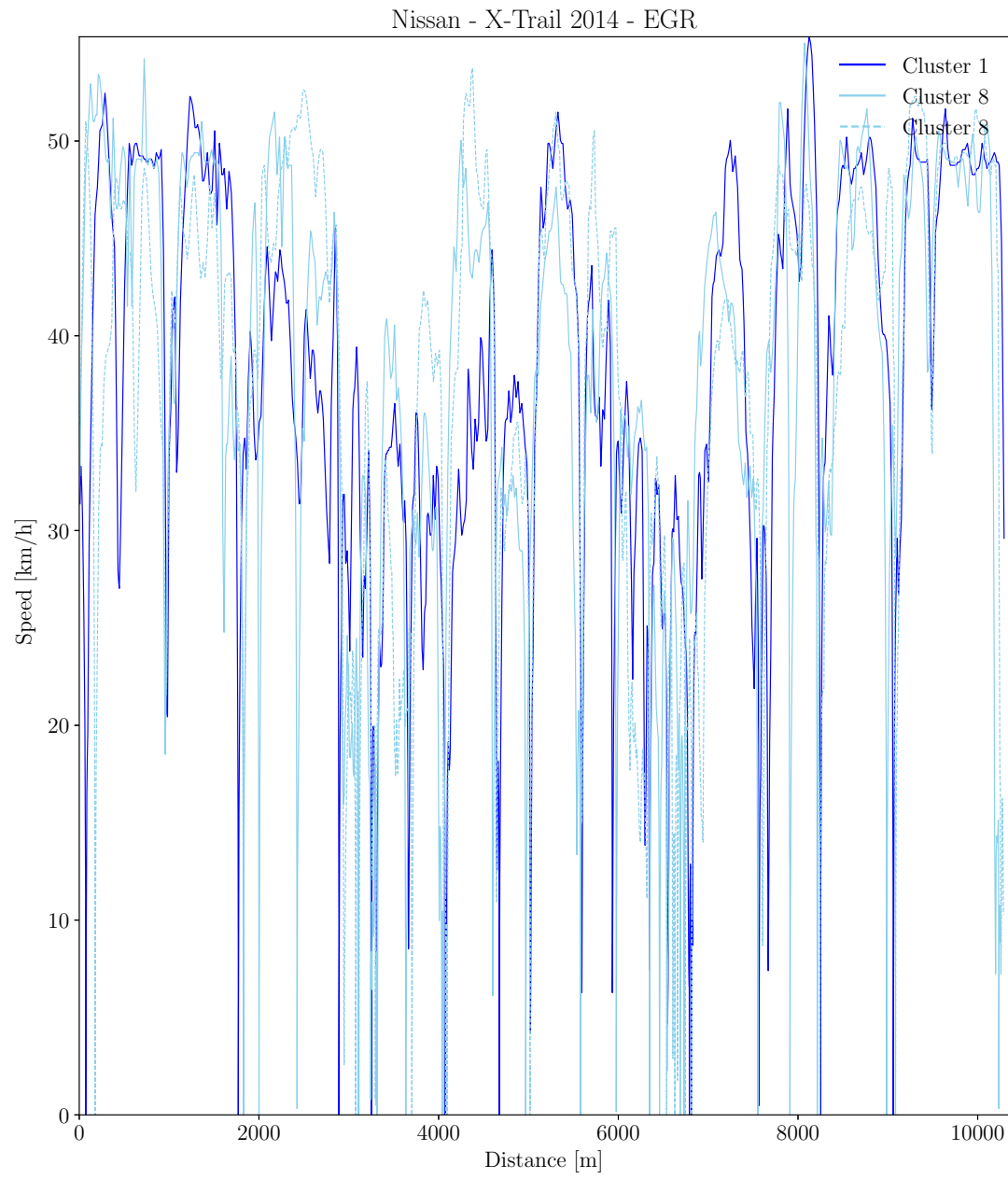


Figure B.28: Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.

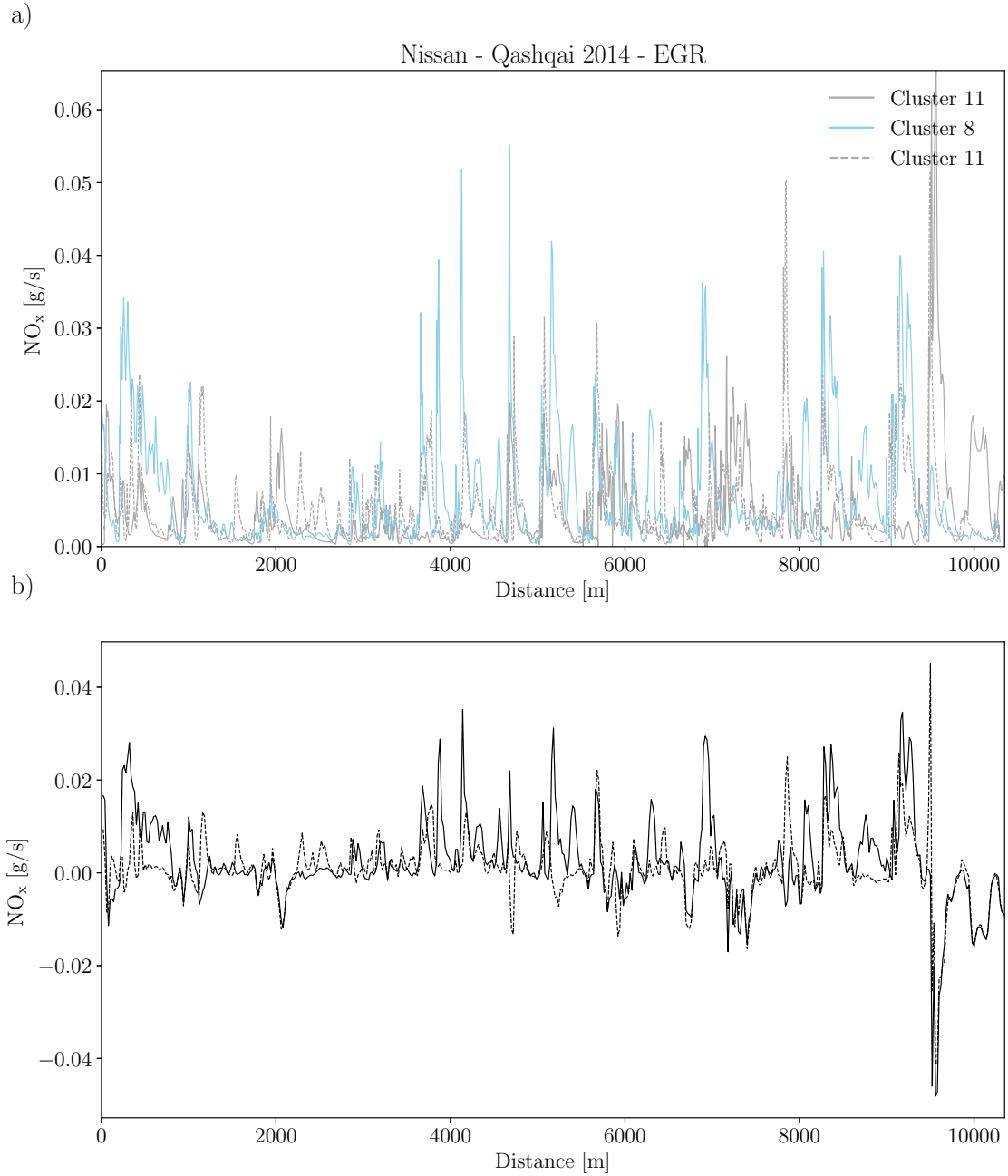


Figure B.29: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Nissan Qashqai. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 11 and the tests scenario are the tests classified in clusters 8 and 11.

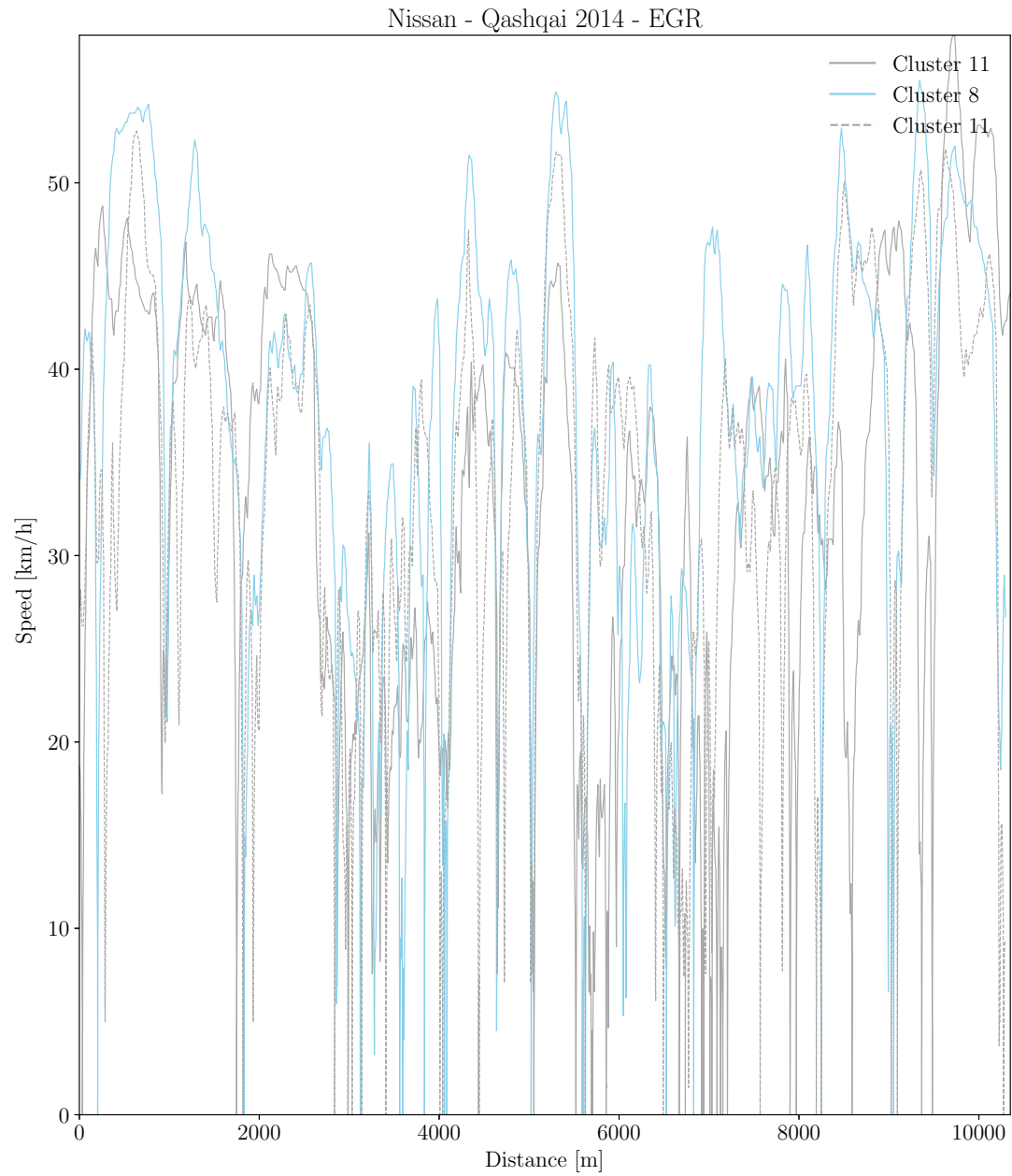


Figure B.30: Instantaneous speed as a function of the distance travelled for the Nissan Qashqai.

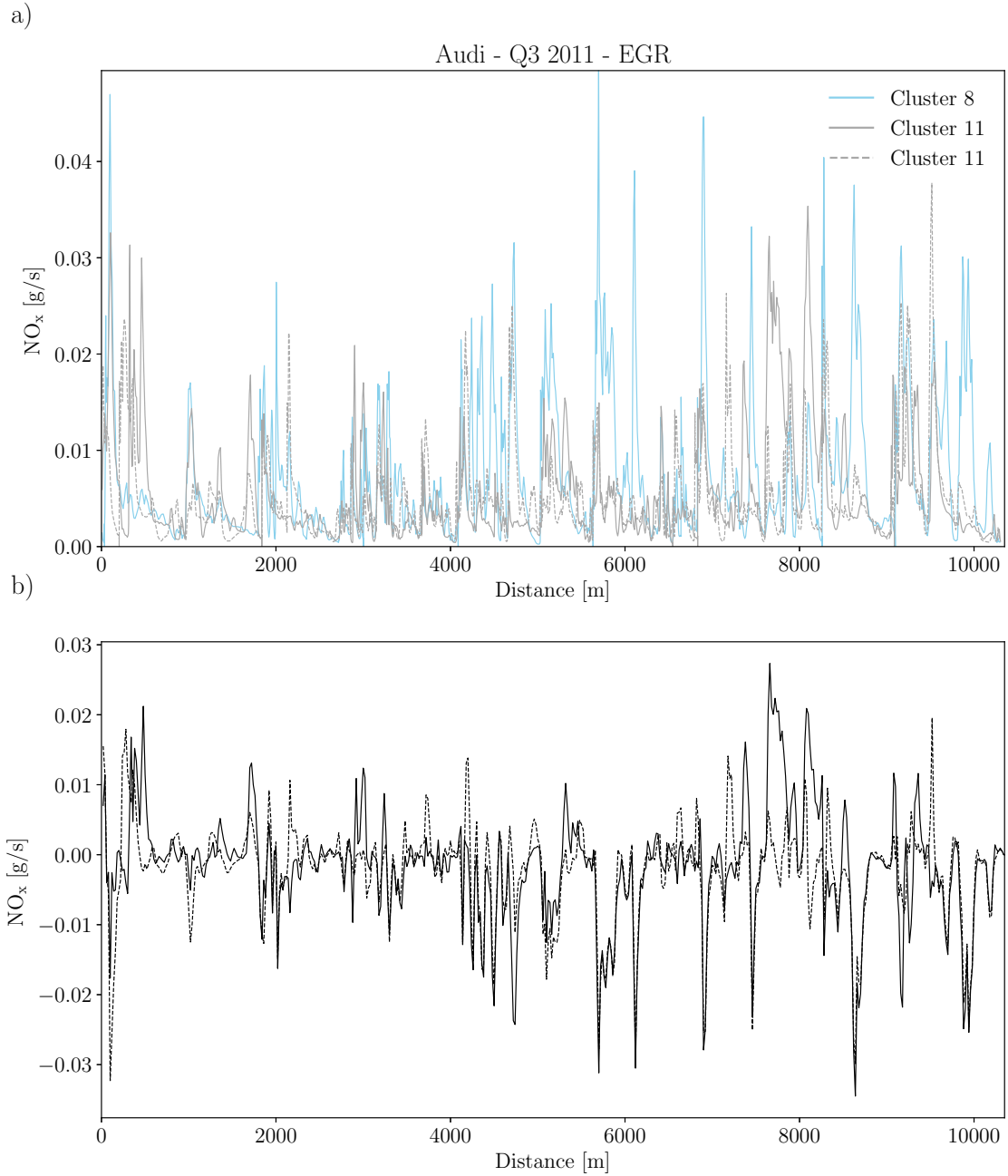


Figure B.31: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the Audi Q3. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 8 and the tests scenario are the tests classified in cluster 11.

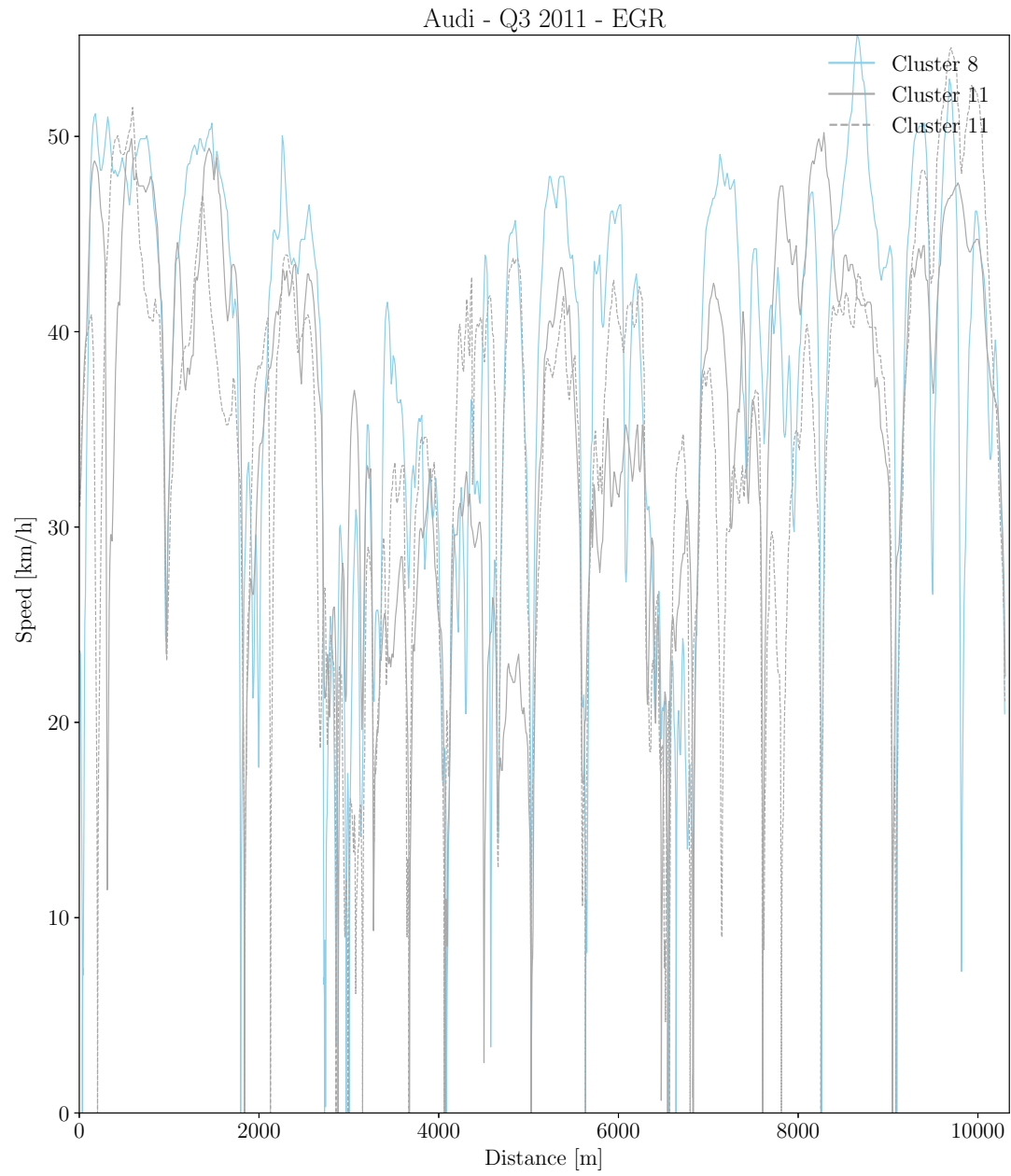


Figure B.32: Instantaneous speed as a function of the distance travelled for the Audi Q3.

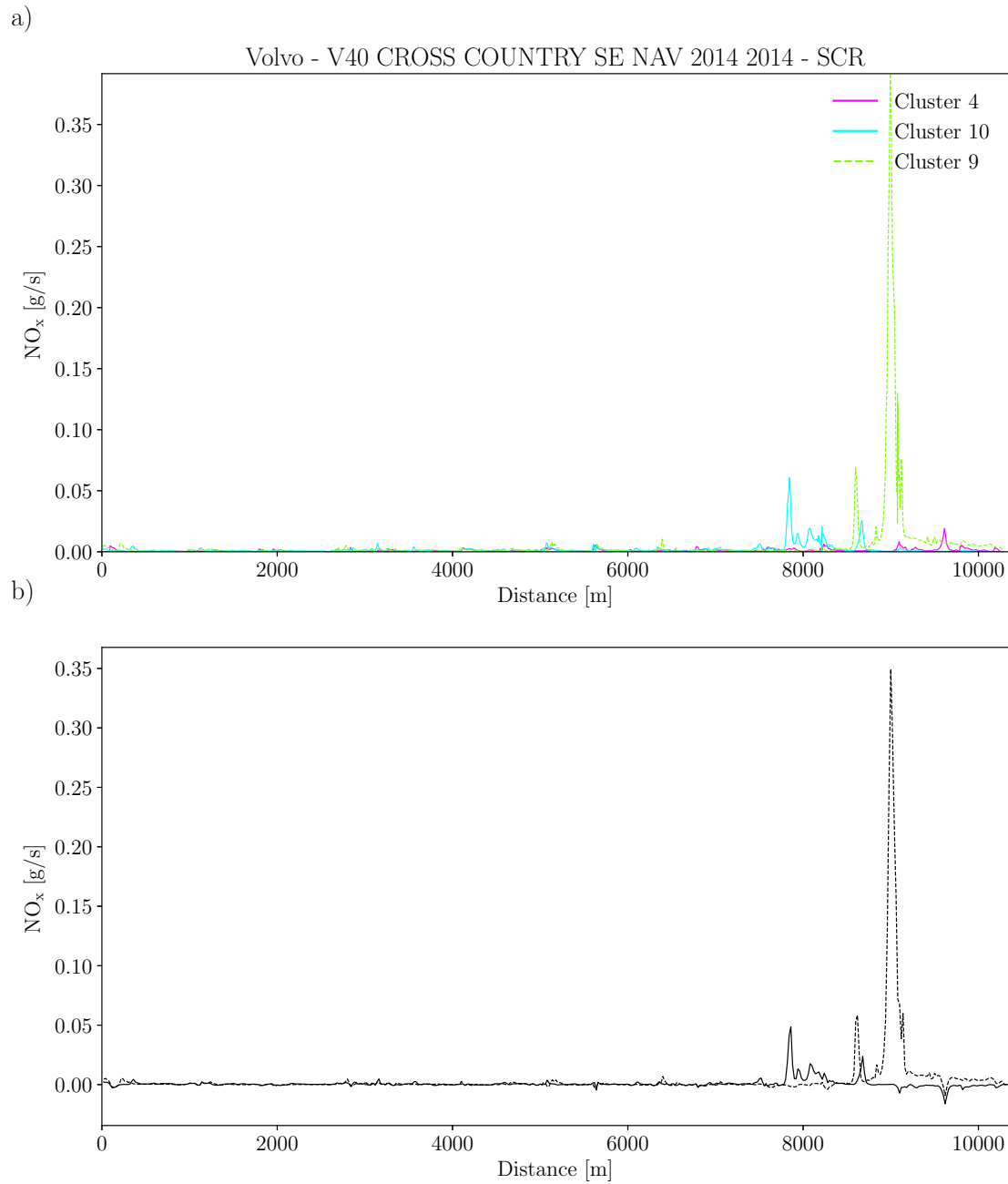


Figure B.33: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 220D M SPORT. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 4 and the tests scenario are the tests classified in clusters 9 and 10.

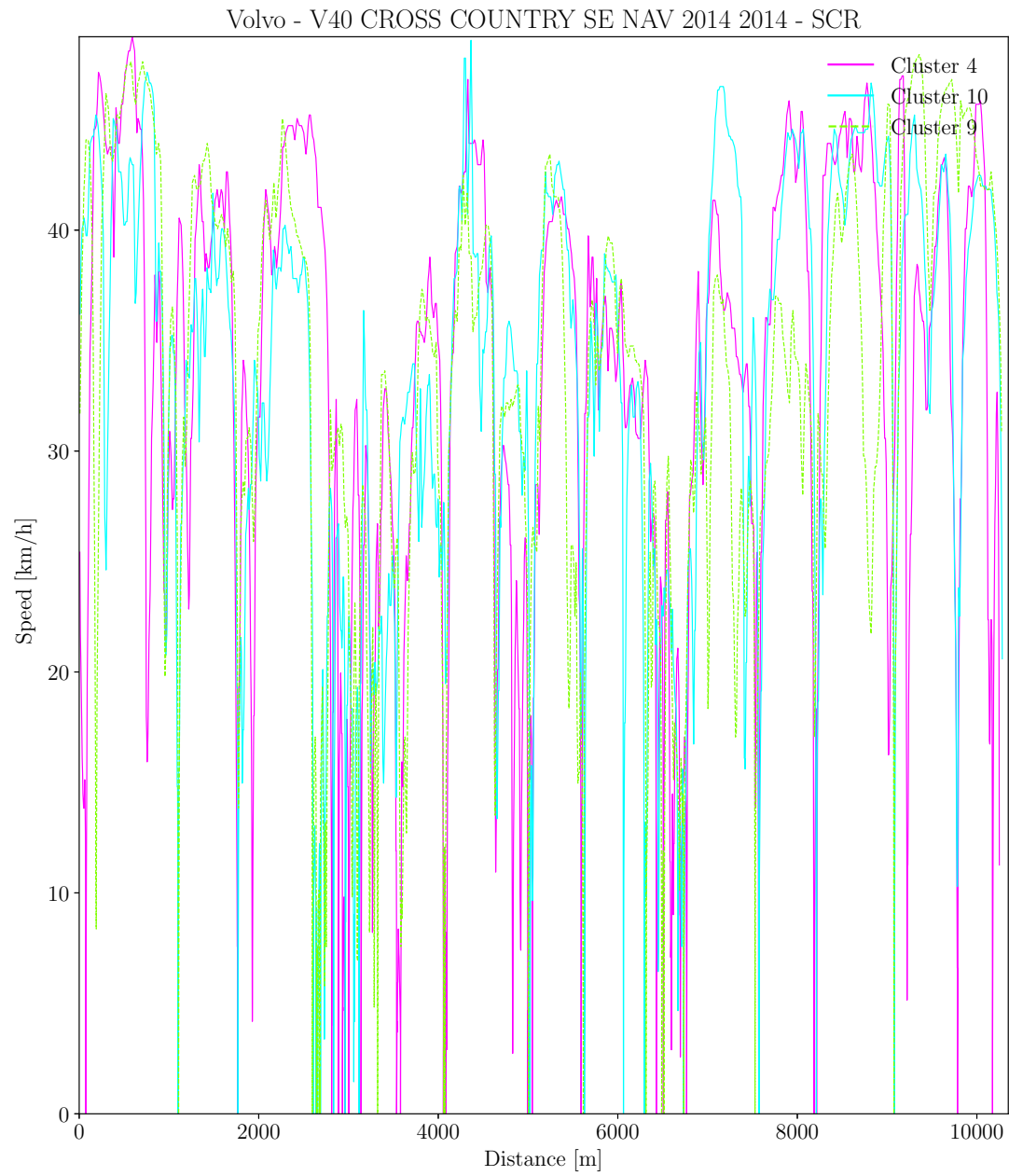


Figure B.34: Instantaneous speed as a function of the distance travelled for the BMW 220D M SPORT.

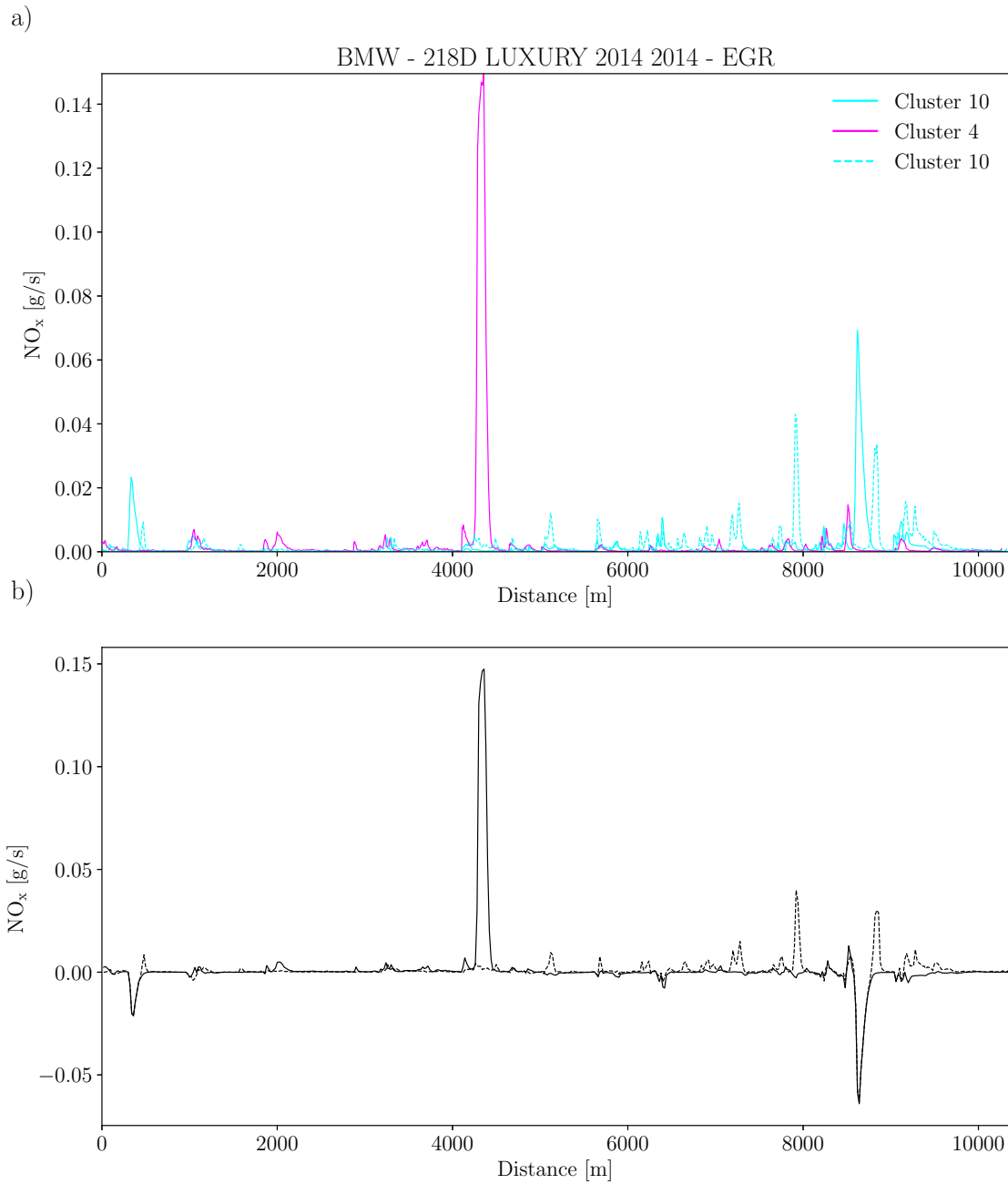


Figure B.35: (a) Instantaneous NO_x emissions as a function of the distance travelled and (b) Residuals of the two time series as a function of the distance travelled for the BMW 218D LUXURY. The residuals were computed as test scenario minus baseline, where the baseline is test classified in cluster 10 and the tests scenario are the tests classified in clusters 4 and 10.

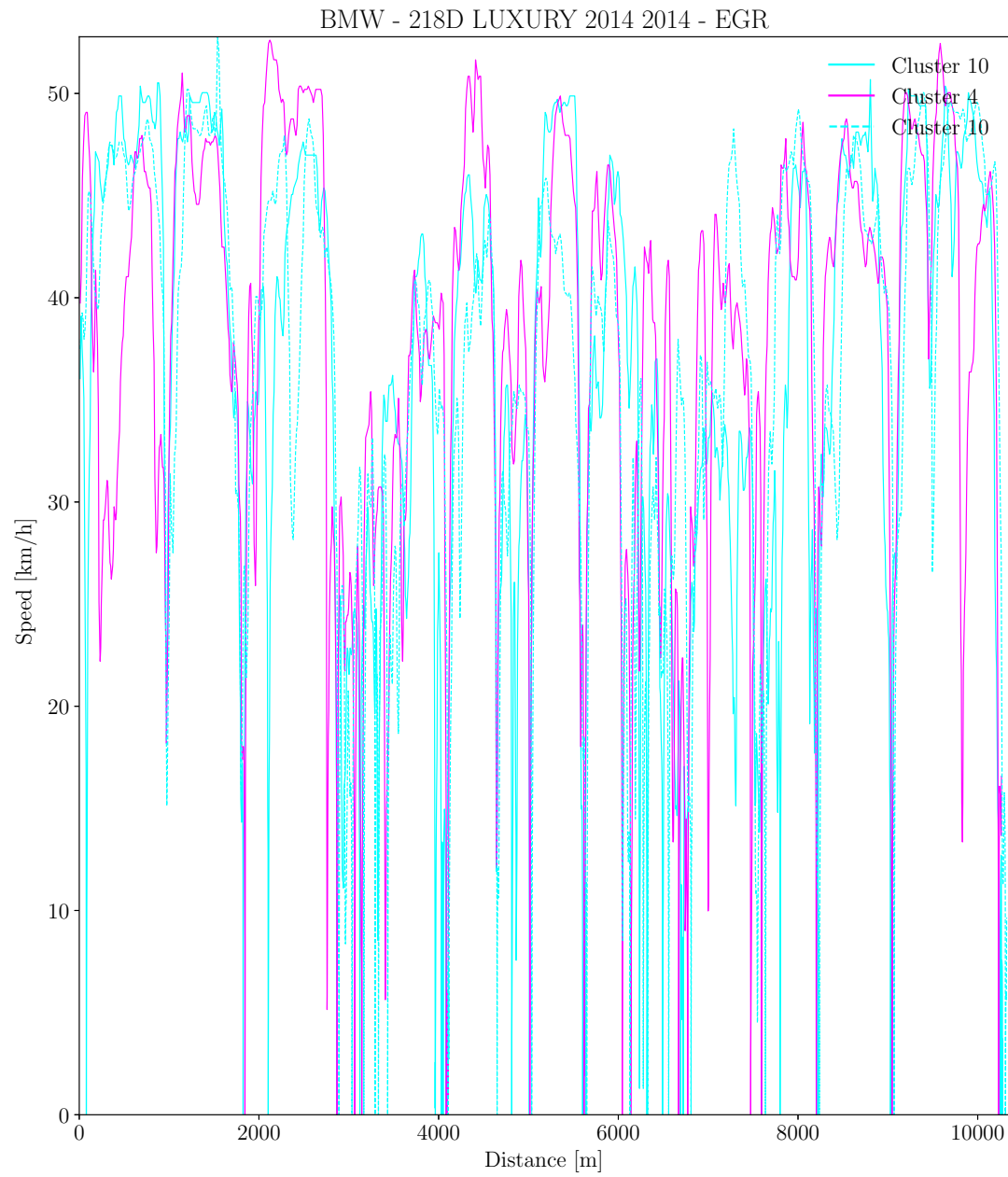


Figure B.36: Instantaneous speed as a function of the distance travelled for the BMW 218D LUXURY.

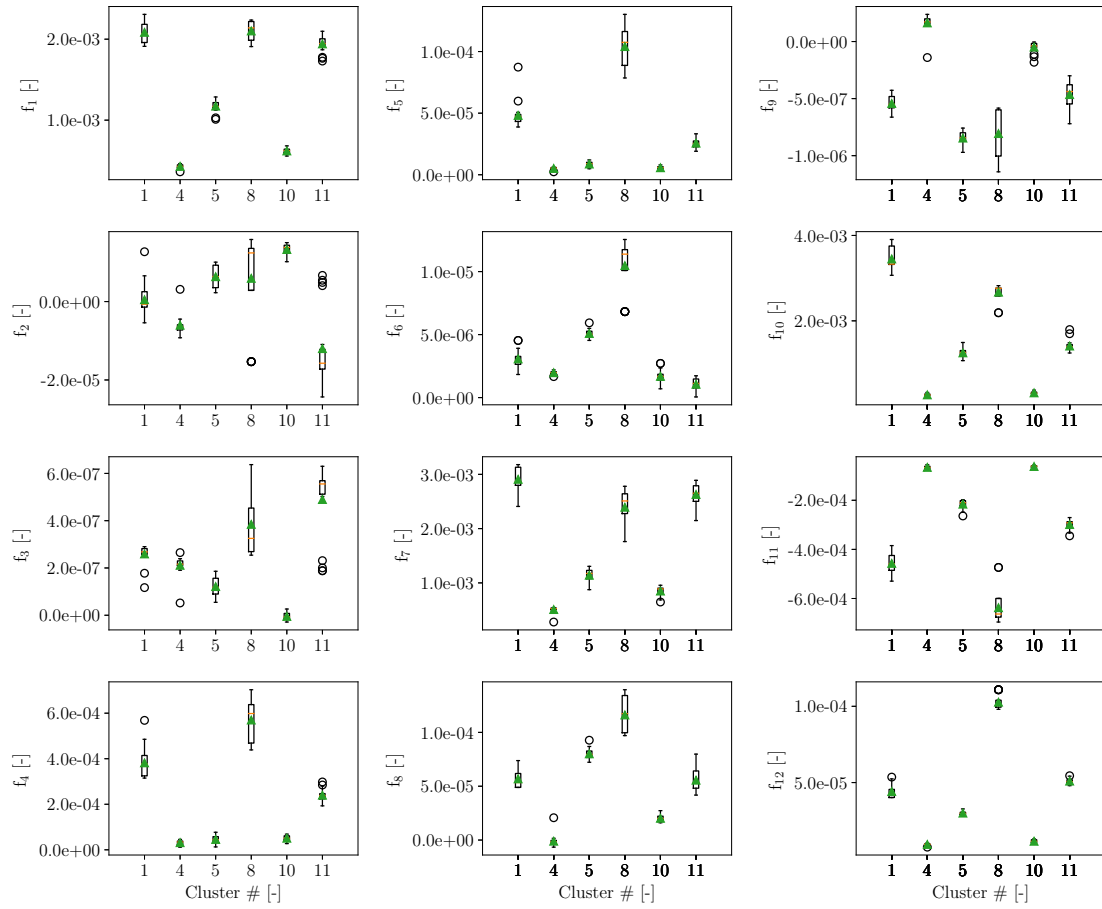


Figure B.37: Distribution of the 12 coefficients of the NLR_1 model for each main cluster for 20 runs of training, with 10% of the datafiles randomly taken out as validation set.

Table B.3: NLR₁ coefficients during deceleration

	$a < -0.5 \text{ m/s}^2$					
	f_1	f_2	f_3	f_4	f_5	f_6
Cluster 1	2.0845e-03	4.7957e-07	2.5933e-07	3.8190e-04	4.8169e-05	3.0714e-06
Cluster 4	4.2900e-04	-6.0311e-06	2.1132e-07	3.1453e-05	5.1147e-06	1.9760e-06
Cluster 5	1.1731e-03	6.3499e-06	1.2127e-07	4.5498e-05	8.7184e-06	5.0873e-06
Cluster 8	2.1017e-03	5.9321e-06	3.8447e-07	5.6994e-04	1.0405e-04	1.0479e-05
Cluster 10	6.1977e-04	1.3288e-05	-4.8776e-09	5.0933e-05	5.6941e-06	1.6844e-06
Cluster 11	1.9417e-03	-1.2003e-05	4.8961e-07	2.3975e-04	2.5670e-05	1.0582e-06

Table B.4: NLR₁ coefficients during acceleration

	$a \geq -0.5 \text{ m/s}^2$					
	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}
Cluster 1	2.9082e-03	5.6845e-05	-5.4230e-07	3.4495e-03	-4.5701e-04	4.4116e-05
Cluster 4	5.0885e-04	-1.0681e-06	1.6350e-07	2.6640e-04	-6.5976e-05	9.4263e-06
Cluster 5	1.1410e-03	8.0077e-05	-8.4596e-07	1.2503e-03	-2.1577e-04	2.9978e-05
Cluster 8	2.3939e-03	1.1618e-04	-8.0644e-06	2.6718e-03	-6.3748e-04	1.0270e-04
Cluster 10	8.4963e-04	2.0206e-05	-4.8663e-08	3.1357e-04	-6.2477e-05	1.1526e-05
Cluster 11	2.6294e-03	5.5594e-05	-4.6483e-07	1.4101e-03	-2.9889e-04	5.0963e-05

Figure B.38 presents the main clusters with respect to the NO_x emission factor (a), the CO₂ emission factor (b), the CO emission factor (c), and the fraction of NO_x primarily emitted as NO₂. It can be observed that all the clusters present a similar distribution of the CO and CO₂ emission factors. Vehicular NO_x emissions seem therefore uncorrelated to CO and CO₂ emissions and specific knowledge on CO and CO₂ emissions is unlikely to give any insights on NO_x emissions. The fraction of NO_x emitted primarily as NO₂ is high (average $f_{\text{NO}_2} = 44 \pm 20 \%$), which can be attributed to the use of DPF and oxidation catalysts [Mavroidis and Chaloulakou, 2011]. Overall, f_{NO_2} presents significant variations within each cluster, which is consistent with other studies [Mavroidis and Chaloulakou, 2011], and further suggests that primary NO_x emissions from diesel vehicles can have a significant impact on local NO₂ concentrations.

Figure B.39 presents the CO₂ main clusters with respect to the CO₂ emission factor (a), the engine size (b), the vehicle weight (c). Each manufacturer is depicted with a different marker, while the aftertreatment technology is displayed in red when the vehicle is equipped with an EGR system only, in blue when the vehicle is equipped with an LNT system in addition to the

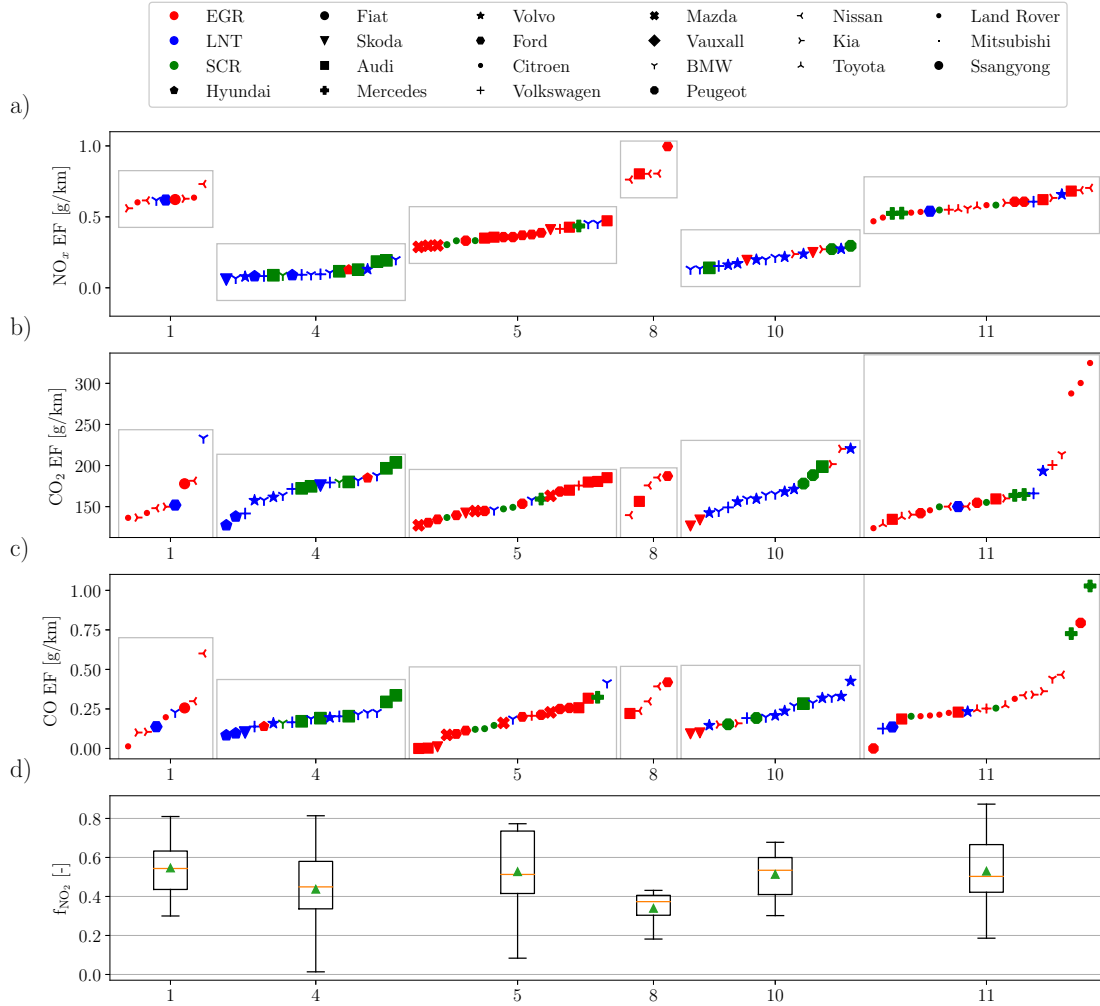


Figure B.38: (a) Vehicle manufacturer, NO_x aftertreatment technology and NO_x emission factor of each vehicle in the cluster, (b) Vehicle manufacturer, NO_x aftertreatment technology and CO_2 emission factor of each vehicle in the cluster, (c) Vehicle manufacturer, NO_x aftertreatment technology and CO emission factor of each vehicle in the cluster and (d) Boxplots of f_{NO_2} per cluster. Clusters containing less than three tests are not depicted here.

EGR system and in green when the vehicle is equipped with an SCR system in addition to the EGR system. Finally, the distribution of the CO₂ emission factors within each cluster is shown in (d). Interestingly, it can be observed that all the six main clusters contain vehicles with very similar CO₂ emission factors.

Figure B.40, which presents the mean instantaneous CO₂ emission as a function of the distance travelled, shows that even though small differences in emitting behaviour can be noticed among the clusters, these are considerably less pronounced than for NO_x emissions. It demonstrates therefore that among a fleet of Euro 5 and 6 diesel vehicles, the variability in CO₂ emissions is significantly less marked than the variability in NO_x emissions. From a modelling perspective, performing a clustering analysis in order to improve the CO₂ prediction accuracy is likely to be unnecessary.

Figure B.41 presents the CO main clusters with respect to the CO emission factor (a), the engine size (b), the vehicle weight (c). Each manufacturer is depicted with a different marker, while the aftertreatment technology is displayed in red when the vehicle is equipped with an EGR system only, in blue when the vehicle is equipped with an LNT system in addition to the EGR system and in green when the vehicle is equipped with an SCR system in addition to the EGR system. Finally, the distribution of the CO emission factors within each cluster is shown in (d). It can be observed that none of the vehicle exceeded the European emission standard. Similarly to the clustering performed on NO_x, it can be seen that clusters consist of tests presenting a similar emission factor, even though some clusters are also overlapping in terms of CO emission factor. Figure S7 in SI confirms that conclusions similar to NO_x can be drawn.

Figure B.42 presents the mean instantaneous CO emission as a function of the distance travelled. Results are discussed in Section 3.1.3 of the main text.

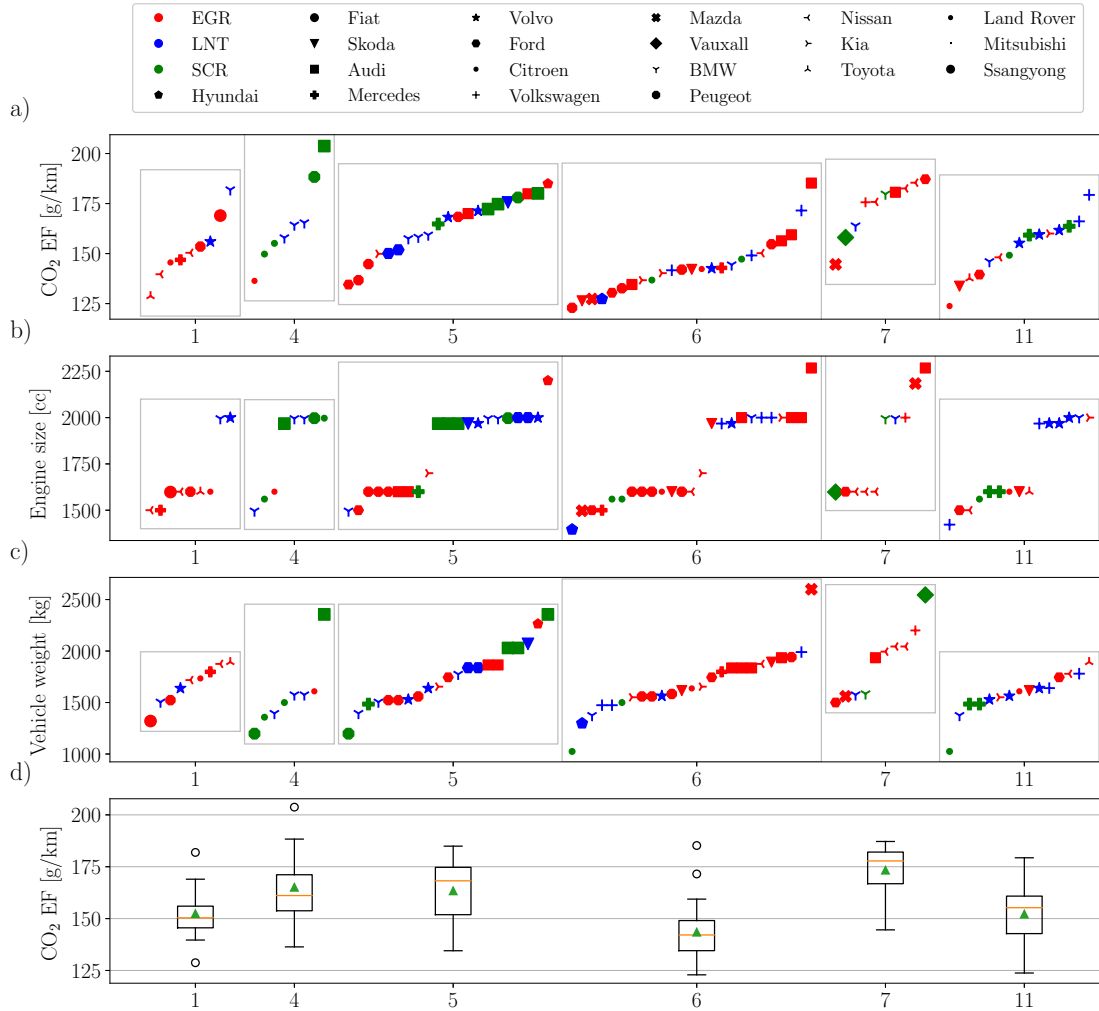


Figure B.39: (a) Vehicle manufacturer, NO_x aftertreatment technology and CO_2 emission factor of each vehicle in the cluster, (b) Vehicle manufacturer, NO_x aftertreatment technology and engine size of each vehicle in the cluster, (c) Vehicle manufacturer, NO_x aftertreatment technology and vehicle weight of each vehicle in the cluster and (d) Boxplots of CO_2 emission factors per cluster. Clusters containing less than three tests are not depicted here.

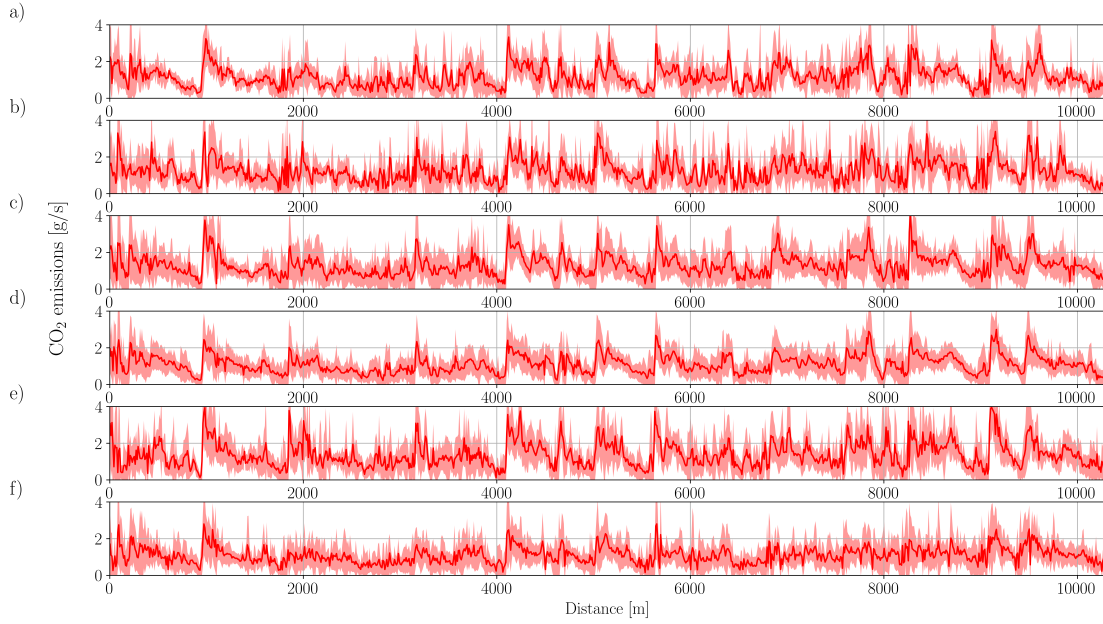


Figure B.40: Aggregated instantaneous CO₂ emissions for the main clusters (a) Cluster 1, (b) Cluster 4, (c) Cluster 5, (d) Cluster 8, (e) Cluster 10 and (f) Cluster 11. The standard deviation is depicted by the light red envelop.

B.6.4 Robustness

Table B.5 presents the evolution of the mean and standard deviation of the predicted NO_x emission factor, the fractional bias (FB) and the normalised mean squared error (NMSE) through the three performed runs. We remind the reader that each run comprised 20 unseen tests, which are used to evaluate the robustness of the results. It can be observed that these metrics are virtually the same over the three runs for the three studied models, which illustrates that the results are robust.

B.6.5 Multilayer perceptron

Figure B.43 depicts the map of the standard deviation for the MLP model for the 400 runs (a), and the residuals between the mean MLP emission map and the mean NLR emission map. The white dashed lines represent the 5th and 95th percentiles of speed and acceleration. Inside this interval, the amount of data is considered to be sufficient to create reliable models. The

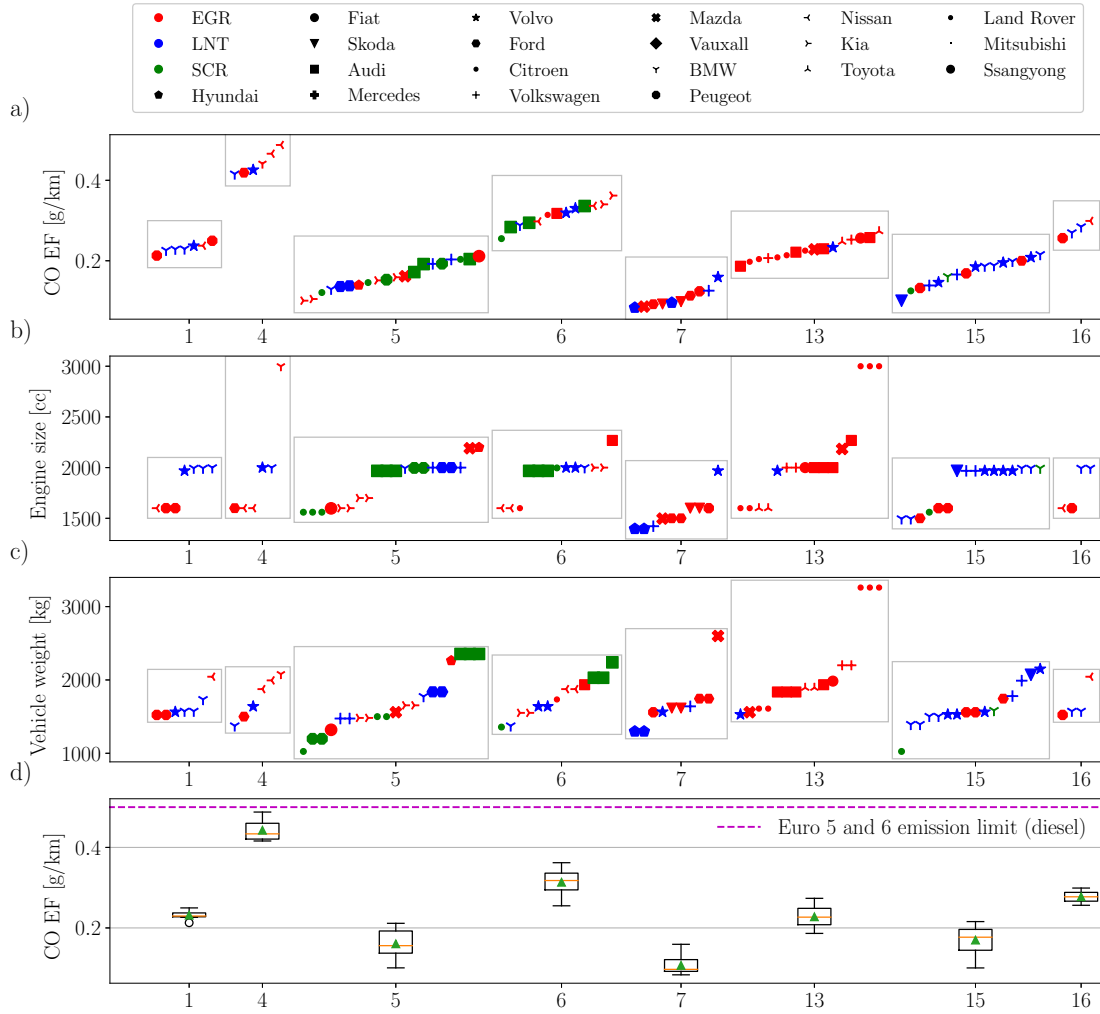


Figure B.41: (a) Vehicle manufacturer, NO_x aftertreatment technology and CO emission factor of each vehicle in the cluster, (b) Vehicle manufacturer, NO_x aftertreatment technology and engine size of each vehicle in the cluster, (c) Vehicle manufacturer, NO_x aftertreatment technology and vehicle weight of each vehicle in the cluster and (d) Boxplots of CO emission factors per cluster. Clusters containing less than three tests are not depicted here.

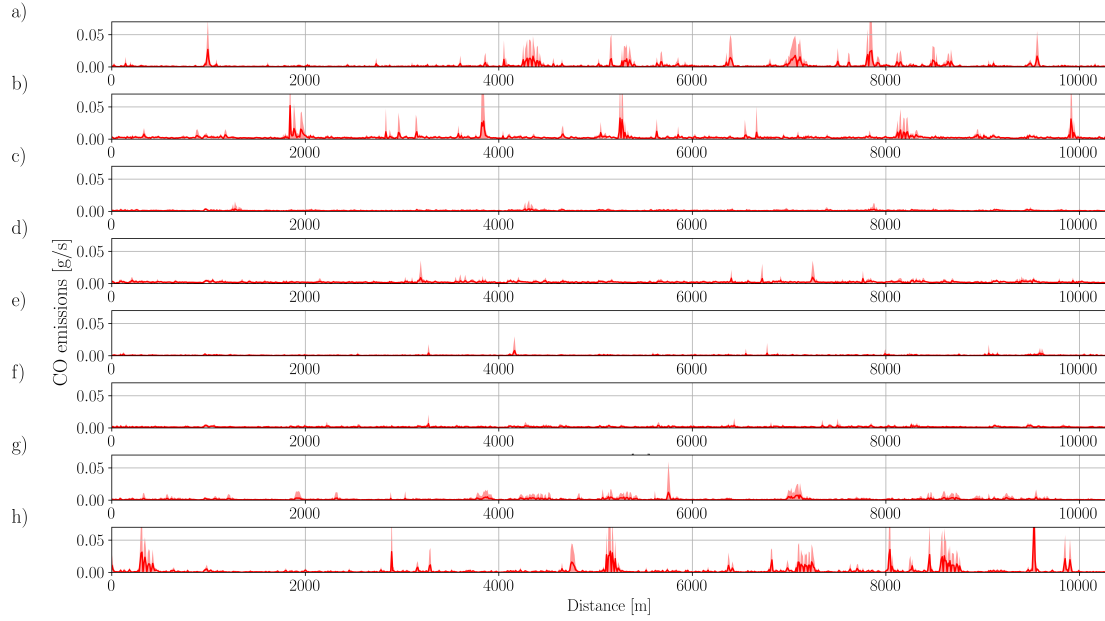


Figure B.42: Aggregated instantaneous CO emissions for the main clusters (a) Cluster 1, (b) Cluster 4, (c) Cluster 5, (d) Cluster 6, (e) Cluster 7, (f) Cluster 13, (g) Cluster 15 and (h) Cluster 16. The standard deviation is depicted by the light red envelop.

Table B.5: Evolution of the mean and the standard deviation of the predicted NO_x emission factor, the fractional bias and the normalised mean squared error for the three performed runs.

	LT			NLR			MLP		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
Predicted EF [g/km]	0.22 (s.d. 0.12)	0.20 (s.d. 0.12)	0.20 (s.d. 0.10)	0.27 (s.d. 0.14)	0.26 (s.d. 0.14)	0.26 (s.d. 0.14)	0.25 (s.d. 0.15)	0.24 (s.d. 0.15)	0.23 (s.d. 0.13)
FB [-]	0.03 (s.d. 0.04)	0.04 (s.d. 0.05)	0.04 (s.d. 0.04)	-0.02 (s.d. 0.02)	-0.01 (s.d. 0.03)	-0.02 (s.d. 0.02)	-0.003 (s.d. 0.03)	0.009 (s.d. 0.02)	0.01 (s.d. 0.02)
NMSE [-]	0.33 (s.d. 0.16)	0.33 (s.d. 0.14)	0.36 (s.d. 0.16)	0.19 (s.d. 0.08)	0.21 (s.d. 0.08)	0.23 (s.d. 0.09)	0.24 (s.d. 0.10)	0.24 (s.d. 0.10)	0.26 (s.d. 0.11)

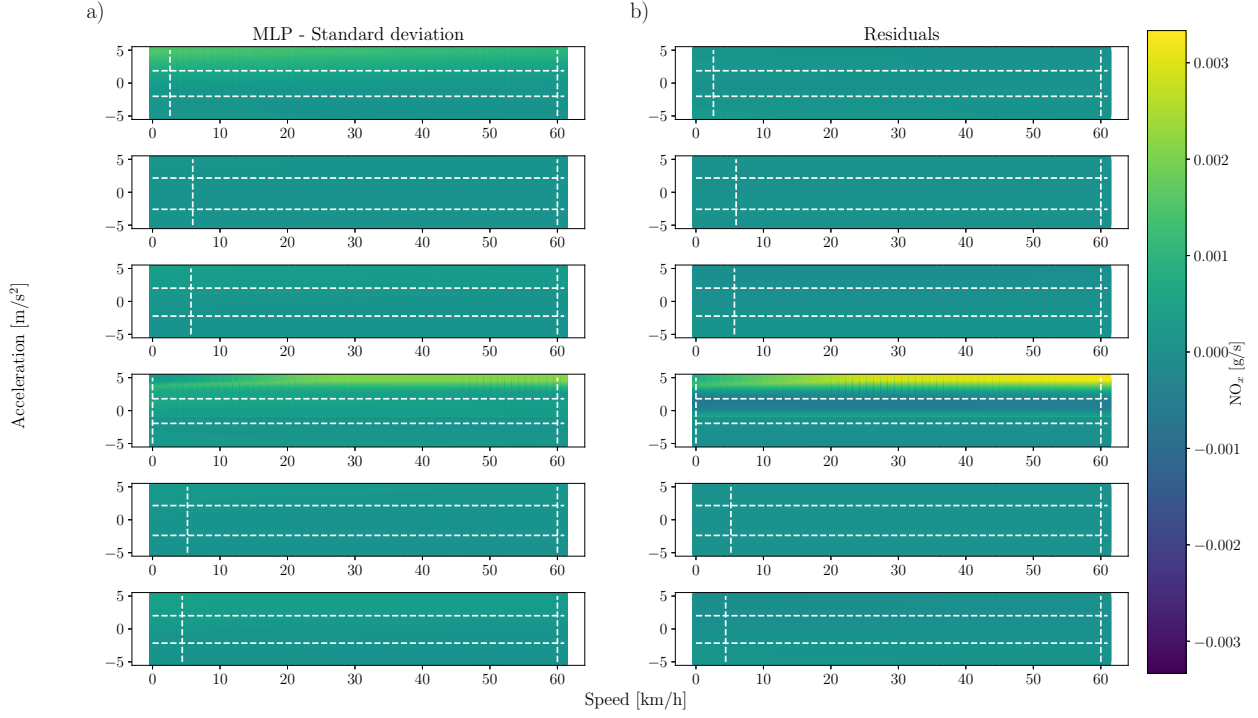


Figure B.43: (a) Map of the standard deviation for the 400 runs and (b) Residuals between the mean MLP emission map and the mean NLR model emission map. White dash lines represent the 5th and 95th percentiles. Clusters are presented in numerical order (1, 4, 5, 8, 10, 11) from the top to the bottom of the graph.

standard deviation over the 400 runs of the MLP model tend to 0.004 inside the interval of speed and acceleration containing 90% of the data, which can be considered as small since a typical instantaneous NO_x emission for a Euro 6 vehicle is in between 0 and 0.08 g/s [Weiss et al., 2012]. Inside this interval, the results can therefore be considered as robust, whereas extrapolating the model outside this range lead to wrong estimates. The residuals between the mean MLP emission map and the mean NLR emission map tend to 0.002 inside the interval of speed and acceleration containing 90% of the data, which can be considered as small since a typical instantaneous NO_x emission for a Euro 6 vehicle is in between 0 and 0.08 g/s [Weiss et al., 2012]. Inside this range, the two models present therefore a good agreement with each other.

B.6.6 Influence of time history

Figure B.44 depicts the evolution of the relative error distribution for the MLP model as a function of the time history considered for the three training runs performed. The grey dashed lines represents the 20% relative error. It can be observed that the distribution of the relative error does not vary with the length of the time interval considered, and it can therefore be concluded that accounting for time history does not improve the accuracy of the MLP prediction.

B.6.7 Case study

Figure B.45 presents the aggregated speed as a function of the distance travelled (procedure described in Section 2.5 of the main text) for the measured speed and the 1-km long mean speed as computed for the COPERT5 model. It can be observed that the measured speed presents higher variations than the 1-km long COPERT5 mean speed. Nevertheless, the variations presented by the 1-km long COPERT5 mean speed are significant, which explains the peaks in emissions observed in Figure 7, Section 3.4 of the main text.

Figure B.46 presents the aggregated instantaneous NO_x emission factor as a function of the distance travelled (procedure described in Section 2.5 of the main text). It can be observed that the agreement between the aggregated measured instantaneous NO_x emission factor and the aggregated NLR_1 , MLP_1 and COPERT5 NO_x emission factors are globally good. However concerning COPERT5, as mentioned in section 3.4 of the main text, even though the aggregated results present a satisfactory accuracy, the individual tests present large errors.

Figure B.47 presents the residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions by the NLR_1 model (a), the residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions by the MLP_1 model (b), and the residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions by the COPERT5 model (c). It can be observed that the residuals between the measurements and the prediction from the COPERT5 model are higher than between the measurements and the

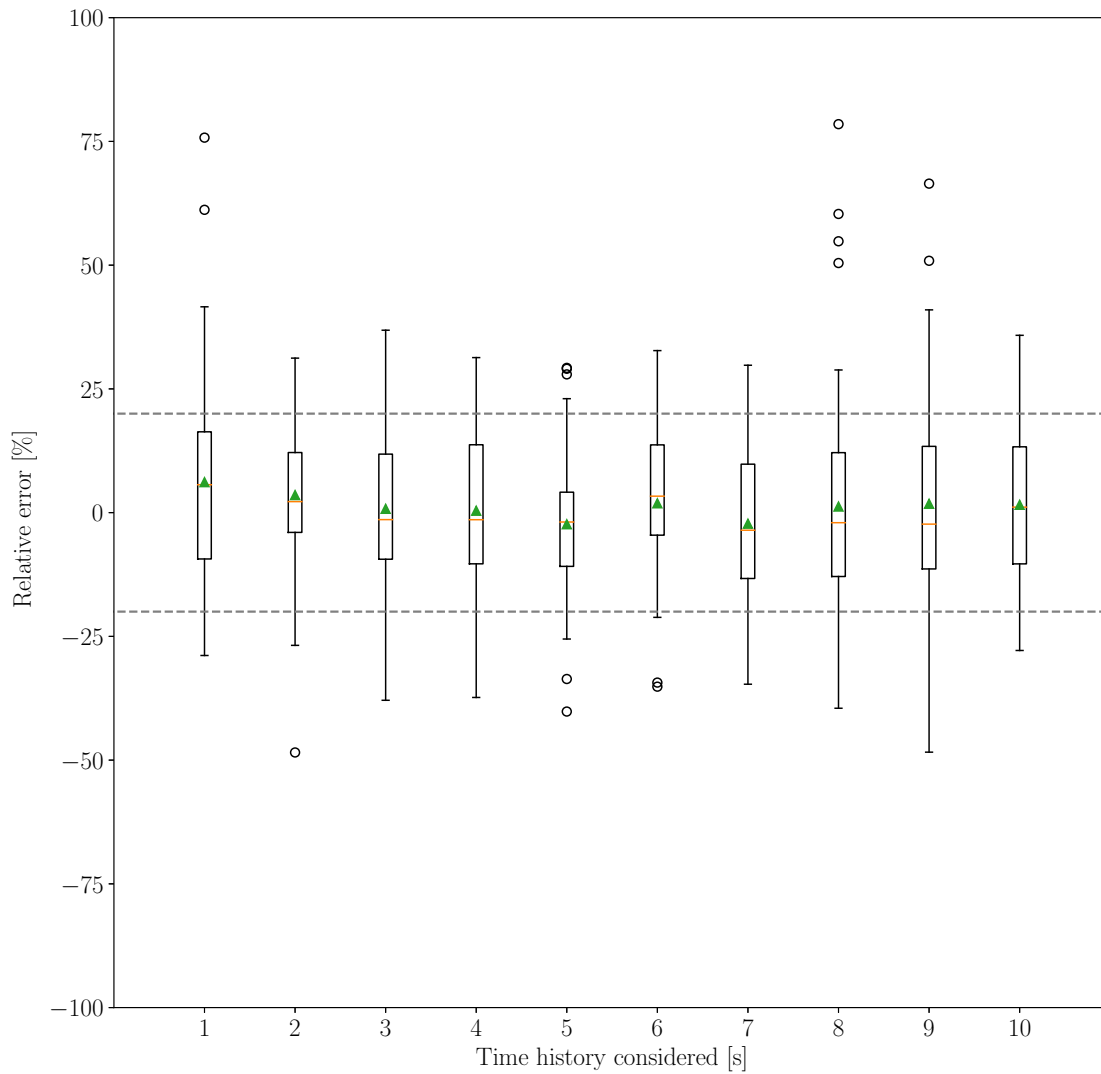


Figure B.44: Evolution of the relative error distribution for the MLP model as a function of the time history considered for the three training runs. The grey dashed lines represent the 20% error.

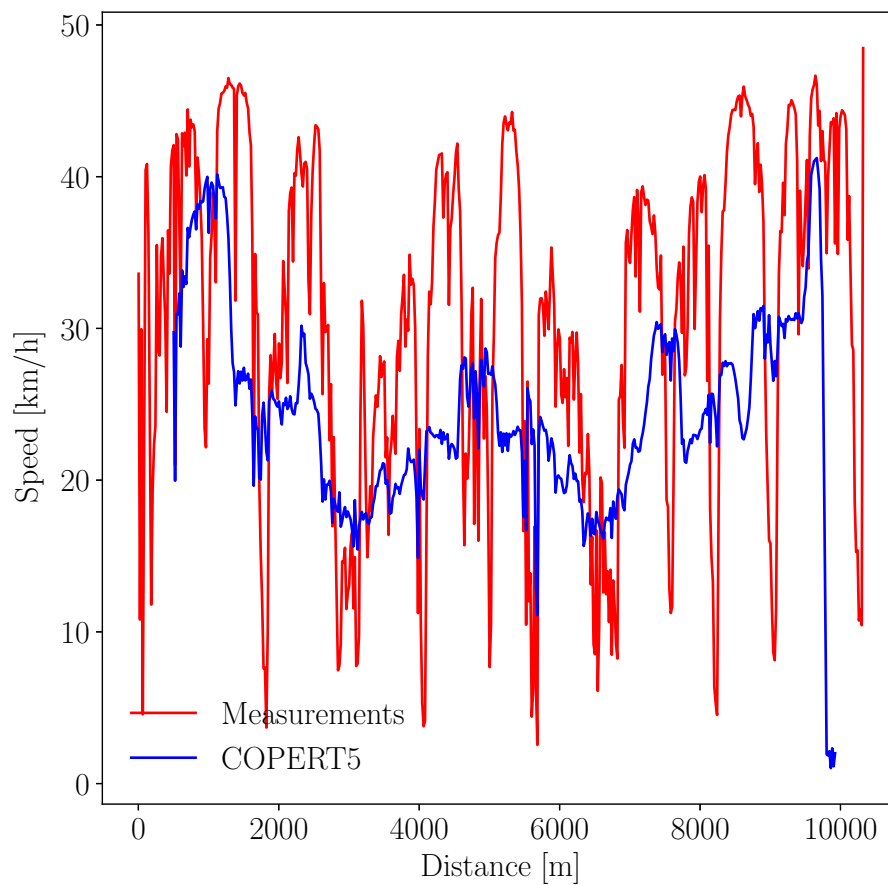


Figure B.45: Aggregated speed as a function of the distance travelled. The red line represents the aggregated measured mean speed. The blue line depicts the aggregated 1-km long mean speed computed at each data point used to compute the COPERT5 emission prediction.

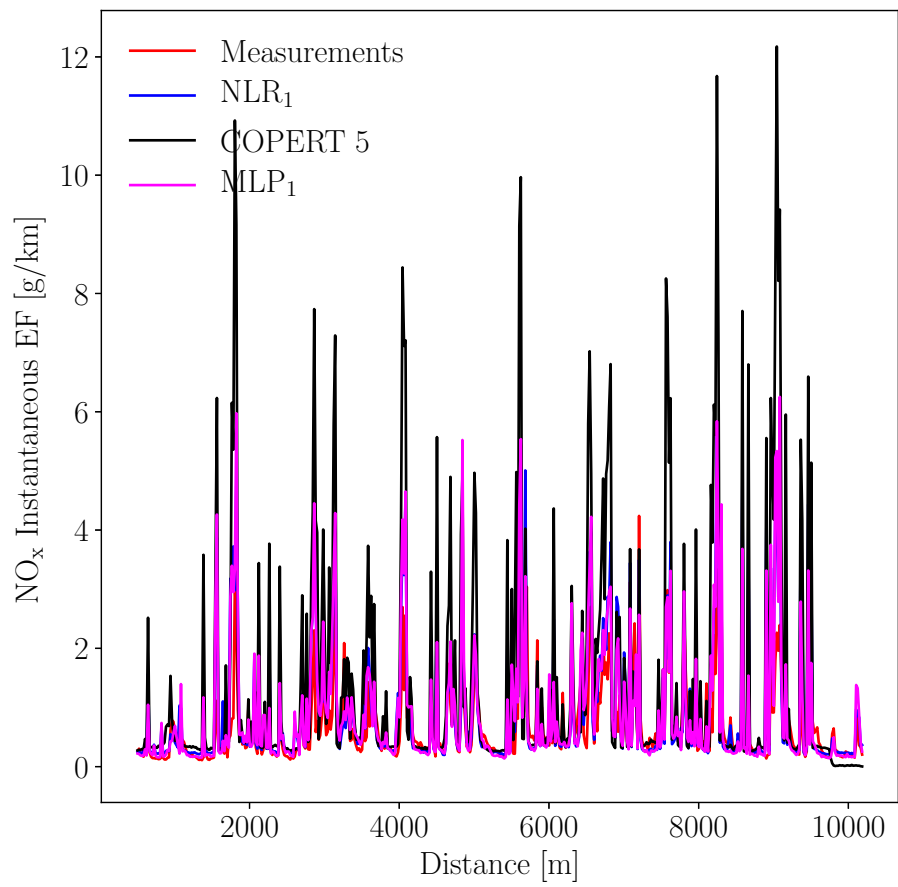


Figure B.46: Aggregated instantaneous NO_x emission factor as a function of the distance travelled.

prediction from the NLR_1 or the MLP_1 model. Moreover, it can be observed that the residuals of the MLP_1 and the NLR_1 models present similarities: peaks are found at the same places and are of a similar order of magnitude.

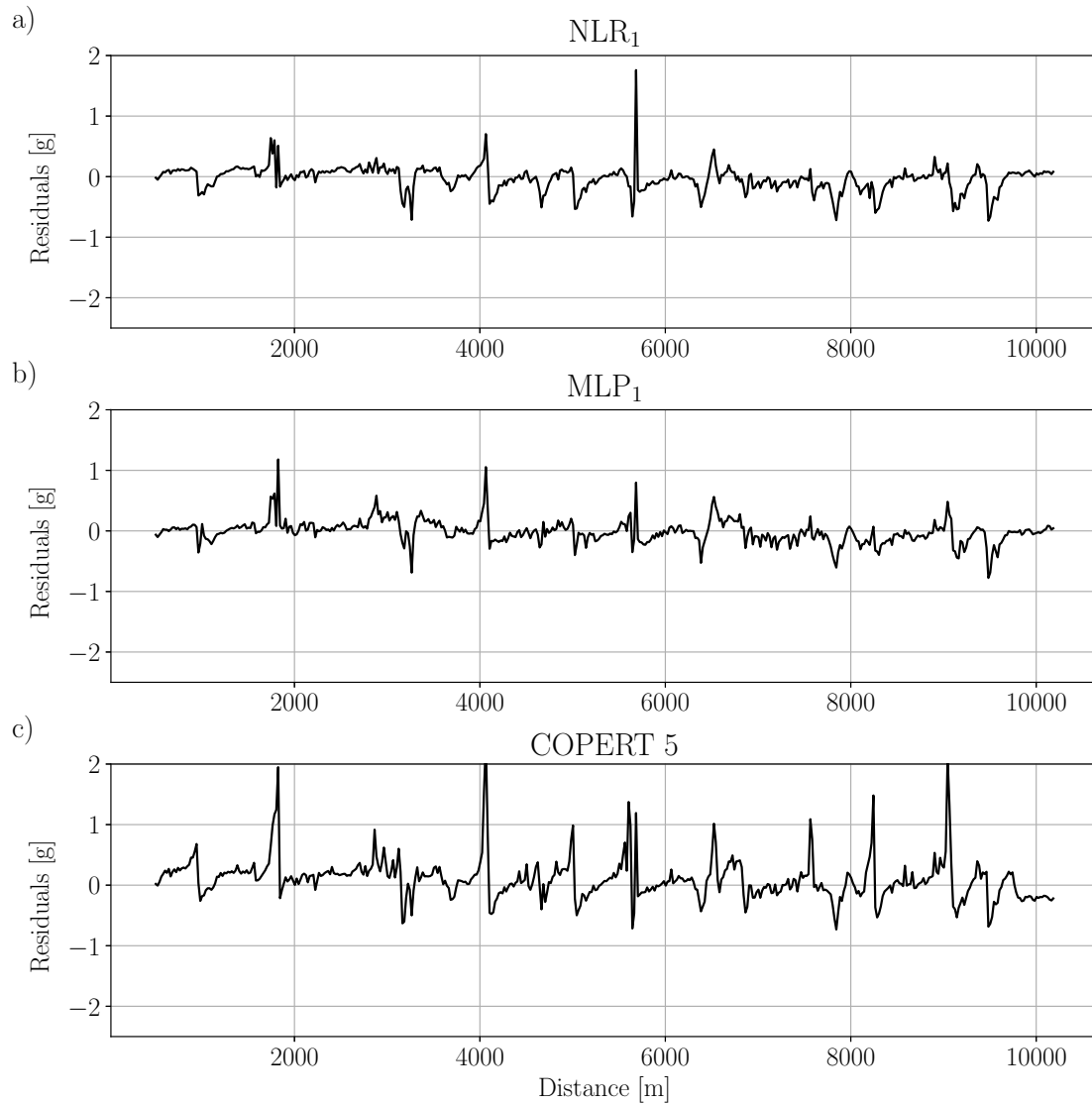


Figure B.47: (a) Residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions from the NLR_1 model, (b) Residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions from the MLP_1 model and (c) Residuals between the aggregated measured NO_x emissions and the predicted NO_x emissions from the COPERT5 model.

Appendix C

Impact of a traffic intervention on the local air quality

C.1 Instrumentation specification

The CAPS, or Serinus 60 NO₂ analyser, commercialised by ACOEM Ecotech, uses a Cavity Attenuated Phase Shift (CAPS) spectroscopy technology to measure directly the NO₂ concentration of the sample. The main specifications of the CAPS are available in Table C.1. We refer the interested reader to [ACOEM Ecotech, 2016] for a more detailed description of the CAPS specifications.

The M200E, or Model 200E Chemiluminescence NO/NO₂/NO_x analyser, commercialised by Teledyne Instruments, uses chemiluminescence detection to measure directly the NO and NO_x concentration of the sample. The NO₂ concentration is obtained as the difference between NO_x and NO. The main specifications of the M200E are available in Table C.1. We refer the interested reader to [Teledyne Instruments, 2015] for a more detailed description of the M200E specifications.

Table C.1: Comparison of the technical specifications of the CAPS and the M200E

	Range	Noise	Precision	Response time	Temperature range
CAPS	0-1 ppm	<0.02 ppb	0.04 ppb or 0.5%	< 30 s	0-45 °C
M200E	0-20,000 ppb	0.2 ppb	0.5%	20 s	5-40 °C

Table C.2: LIDAR average wind speed and wind direction from 12am to 4pm during the two weeks of field study. Days where the 96s cycle was implemented are displayed in bold font while days where the 48s cycle was implemented are displayed in normal font.

	Wind speed [m/s]	Wind direction [°]
16.09.2019	1.94	211.53
17.09.2019	2.85	58.00
18.09.2019	1.98	78.14
19.09.2019	1.30	128.33
20.09.2019	4.07	120.23
23.09.2019	3.83	239.04
24.09.2019	2.76	258.92
25.09.2019	3.74	263.00
26.09.2019	4.07	262.49
27.09.2019	3.90	254.25

C.2 LIDAR average wind speed and wind direction

Table C.2 displays the average wind speed and wind direction recorded from 12am to 4pm during the two weeks of field study.

C.3 Background concentration per day

Table C.3 depicts the average NO, NO₂ and O₃ background concentration used for the SIRANE simulation. These background concentrations were derived from the M200E that measures not only NO₂ but also NO, allowing the ozone concentration to be derived using the photostationary steady-state assumption.

Table C.3: Background concentrations per day

Day	NO_2 [$\mu\text{g}/\text{m}^3$]	NO [$\mu\text{g}/\text{m}^3$]	O_3 [$\mu\text{g}/\text{m}^3$]
16	19.33	15.43	24.99
17	13.72	16.63	16.47
18	8.39	11.00	15.22
19	17.97	12.14	29.53
20	11.63	7.48	30.99
23	8.16	5.26	30.93
24	13.00	8.62	30.08
25	10.38	7.21	28.73
26	10.54	5.48	38.41
27	12.78	6.99	36.45

C.4 Distribution of the mean NO_x emissions and its standard deviation per street

Figure C.1 depicts the distribution of the instantaneous average NO_x emission factor for the NLR and the REVS model and its associated predicted standard deviation for the REVS model per box of the SIRANE network. These instantaneous mass flow are calculated as the total NO_x emissions happening within the box over one hour, divided by 3,600.

C.5 Operational air dispersion modelling

Tables C.4 to C.9 displays the predicted canyon average NO_2 , NO and O_3 concentrations for both interventions, over the two weeks of field study. Emissions were estimated using the NLR or the TOFFEE model.

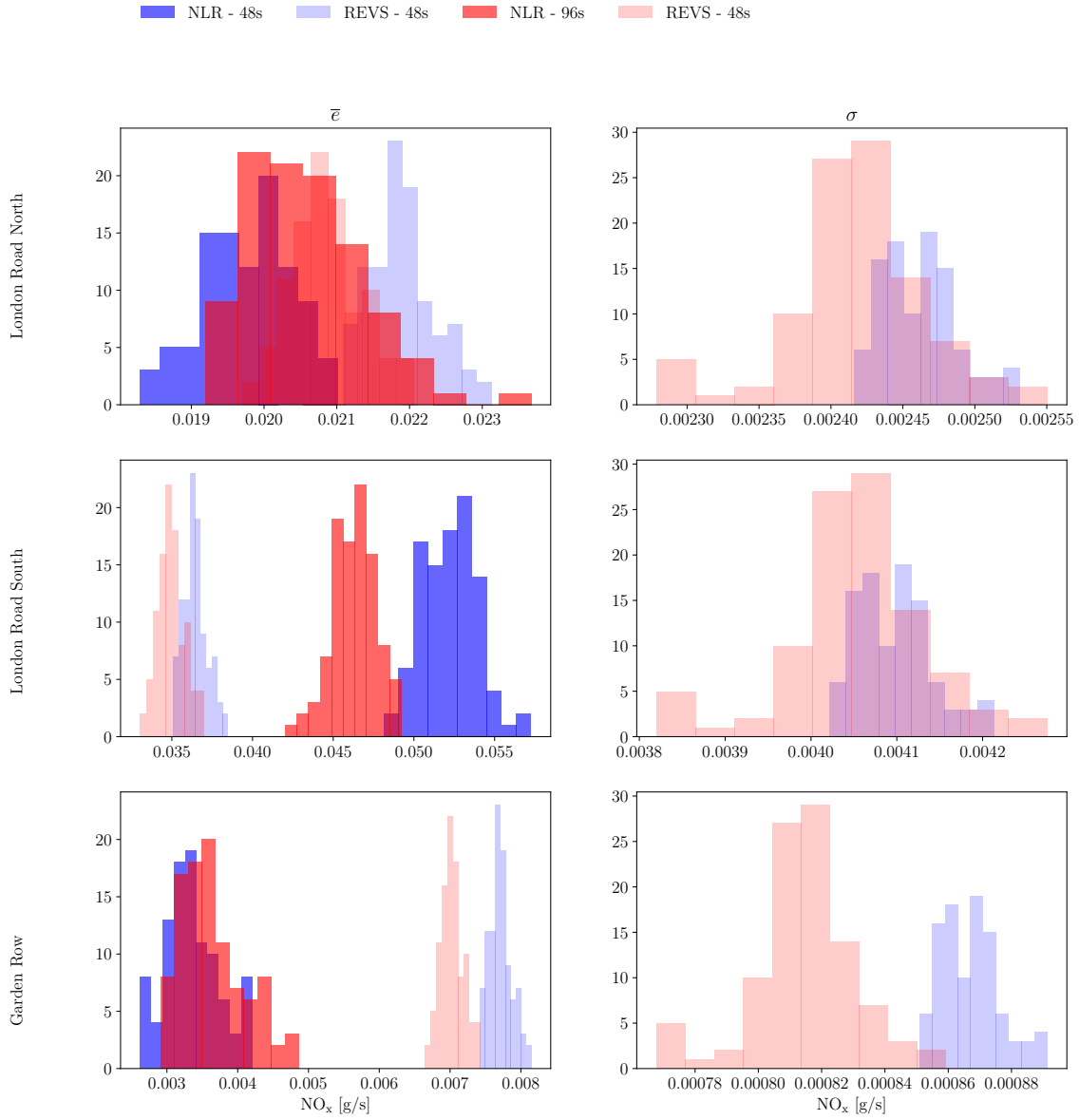


Figure C.1: Distribution of the mean NO_x emissions and its standard deviation per street for both the NLR and the REVS model for the 1,000 runs. Predictions for the NLR model are depicted in dark red and dark blue, while the predictions for the REVS model are depicted in light blue and light red. The predictions for the A-48 intervention are depicted in blue while the predictions for the B-96 intervention are depicted in red.

Table C.4: Predicted canyon average NO₂ concentration in ppb. Emissions were calculated using the NLR model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	23.39	23.80	24.86	23.62	12.48	12.63
17.09.2020	16.71	17.06	17.76	16.81	9.64	9.65
18.09.2019	15.29	15.58	16.25	15.17	7.51	7.50
19.09.2019	24.95	25.10	24.86	23.58	12.64	12.82
20.09.2019	15.87	15.78	15.16	14.39	8.25	8.36
23.09.2019	14.69	15.08	15.99	15.03	5.81	5.93
24.09.2019	18.24	18.66	21.26	20.19	9.12	9.28
25.09.2019	13.81	14.12	16.36	15.54	7.24	7.37
26.09.2019	14.72	15.07	17.57	16.65	7.37	7.51
27.09.2019	16.62	16.99	18.98	18.03	8.66	8.79

Table C.5: Predicted canyon average NO concentration in ppb. Emissions were calculated using the NLR model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	38.13	39.29	42.33	38.75	13.58	13.85
17.09.2020	36.32	37.43	39.68	36.64	16.30	16.34
18.09.2019	37.40	38.38	40.65	37.01	14.14	14.11
19.09.2019	36.70	37.09	36.44	33.17	11.77	12.03
20.09.2019	17.89	17.73	16.61	15.26	6.81	6.93
23.09.2019	16.87	17.59	19.38	17.50	4.58	4.69
24.09.2019	22.28	23.15	29.12	26.56	7.63	7.82
25.09.2019	15.66	16.23	20.66	18.97	6.18	6.31
26.09.2019	12.01	12.46	16.03	14.65	4.57	4.67
27.09.2019	14.78	15.33	18.53	16.94	5.67	5.78

Table C.6: Predicted canyon average O₃ concentration in ppb. Emissions were calculated using the NLR model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	7.36	7.27	7.04	7.31	11.02	10.95
17.09.2020	5.45	5.39	5.30	5.43	7.00	7.00
18.09.2019	4.83	4.80	4.73	4.85	6.29	6.29
19.09.2019	8.08	8.04	8.11	8.45	12.76	12.66
20.09.2019	10.28	10.31	10.57	10.91	14.04	13.98
23.09.2019	9.87	9.71	9.35	9.73	14.38	14.32
24.09.2019	9.22	9.07	8.22	8.55	13.46	13.37
25.09.2019	9.84	9.71	8.84	9.14	13.10	13.03
26.09.2019	13.57	13.39	12.14	12.59	17.86	17.77
27.09.2019	12.35	12.18	11.25	11.69	16.78	16.70

Table C.7: Predicted canyon average NO₂ concentration in ppb. Emissions were calculated using the REVS model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	24.49	24.02	21.30	21.01	14.20	13.96
17.09.2020	17.55	17.19	15.09	14.86	10.44	10.27
18.09.2019	15.96	15.56	13.20	12.94	8.59	8.35
19.09.2019	25.30	24.80	21.20	20.88	14.93	14.59
20.09.2019	15.60	15.30	12.94	12.75	9.67	9.46
23.09.2019	15.62	15.26	13.23	12.98	7.06	6.89
24.09.2019	19.23	18.84	18.24	17.93	10.93	10.68
25.09.2019	14.55	14.26	14.07	13.82	8.59	8.40
26.09.2019	15.56	15.23	14.99	14.71	8.84	8.64
27.09.2019	17.50	17.15	16.30	16.03	10.06	9.87

Table C.8: Predicted canyon average NO concentration in ppb. Emissions were calculated using the REVS model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	41.25	39.90	32.47	31.70	16.59	16.15
17.09.2020	39.00	37.86	31.26	30.59	18.26	17.86
18.09.2019	39.66	38.32	30.56	29.75	16.93	16.30
19.09.2019	37.62	36.31	27.45	26.74	15.22	14.68
hline 20.09.2019	17.38	16.85	12.94	12.65	8.46	8.19
23.09.2019	18.63	17.93	14.28	13.86	5.86	5.67
24.09.2019	24.40	23.55	22.27	21.62	9.86	9.53
25.09.2019	17.04	16.49	16.14	15.70	7.76	7.53
26.09.2019	13.12	12.67	12.35	12.00	5.78	5.60
27.09.2019	16.12	15.58	14.30	13.92	6.94	6.75

Table C.9: Predicted canyon average O₃ concentration in ppb. Emissions were calculated using the REVS model.

	London Road North		London Road South		Garden Row	
	48s	96s	48s	96s	48s	96s
16.09.2019	7.12	7.22	7.87	7.95	10.27	10.37
17.09.2020	5.33	5.38	5.71	5.75	6.76	6.81
18.09.2019	4.76	4.80	5.11	5.14	6.00	6.06
19.09.2019	7.99	8.12	9.18	9.28	11.65	11.81
20.09.2019	10.39	10.52	11.59	11.68	13.26	13.37
23.09.2019	9.50	9.64	10.50	10.62	13.66	13.76
24.09.2019	8.88	9.01	9.22	9.34	12.48	12.62
25.09.2019	9.53	9.65	9.73	9.83	12.36	12.46
26.09.2019	13.13	13.31	13.44	13.58	16.95	17.07
27.09.2019	11.93	12.10	12.52	12.65	15.93	16.05

Bibliography

- [Abadi et al., 2015] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org.
- [Abd-Alla, 2002] Abd-Alla, G. (2002). Using exhaust gas recirculation in internal combustion engines: a review. *Energy Conversion and Management*, 43(8):1027 – 1042.
- [ACOEM Ecotech, 2016] ACOEM Ecotech (2016). Serinus 60 NO2 Analyser. <https://www.ecotech.com/wp-content/uploads/2016/03/ACOEM-Ecotech-Serinus-60-CAPS-NO2-Analyser-20200731.pdf>. Accessed: 2021-01-20.
- [AG, 2014] AG, P. (2014). PTV Vissim 7 User Manual. *PTV Group, Haid-und-Neu-Str*, 15.
- [Alfieri et al., 2015] Alfieri, L., Burek, P., Feyen, L., and Forzieri, G. (2015). Global warming increases the frequency of river floods in europe. *Hydrology and Earth System Sciences*, 19(5):2247–2260.
- [Alonso et al., 2007] Alonso, J. M., Alvarruiz, F., Desantes, J. M., Hernandez, L., Hernandez, V., and Molto, G. (2007). Combining neural networks and genetic algorithms to predict and

- reduce diesel engine emissions. *IEEE Transactions on Evolutionary Computation*, 11(1):46–55.
- [Andersen et al., 2012] Andersen, Z. J., Kristiansen, L. C., Andersen, K. K., Olsen, T. S., Hvidberg, M., Jensen, S. S., Ketzel, M., Loft, S., Sorensen, M., Tjonneland, A., Overvad, K., and Raaschou-Nielsen, O. (2012). Stoke and Long-Term Exposure to Outdoor Air Pollution From Nitrogen Dioxide - A cohort study. *Stroke*, 43:320–325.
- [André and Joumard, 2005] André, J.-M. and Joumard, R. (2005). Modelling of cold start excess emissions for passenger cars. Technical report. 239 p.
- [Andre et al., 2014] Andre, M., Roche, A.-L., and Bourcier, L. (2014). Statistiques de parcs et trafic pour le calcul des émissions de polluants des transports routiers en france.
- [Arthur and Vassilvitskii, 2007] Arthur, D. and Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics.
- [Atkinson et al., 1998] Atkinson, C. M., Long, T. W., and Hanzevack, E. L. (1998). Virtual sensing: a neural network-based intelligent performance and emissions prediction system for on-board diagnostics and engine control. Technical report, SAE Technical Paper.
- [Bachman et al., 2000] Bachman, W., Sarasua, W., Hallmark, S., and Guensler, R. (2000). Modeling regional mobile source emissions in a geographic information system framework. *Transportation Research Part C: Emerging Technologies*, 8(1):205 – 229.
- [Baik and Kim, 1999] Baik, J.-J. and Kim, J.-J. (1999). A numerical study of flow and pollutant dispersion characteristics in urban street canyons. *Journal of Applied Meteorology*, 38(11):1576–1589.
- [Baldino et al., 2017] Baldino, C., Tietge, U., Muncrief, R., Bernard, Y., and Mock, P. (2017). Road tested: Comparative overview of real-world versus type-approval nox and co2 emissions from diesel cars in europe. *The international council on clean transportation*.

- [Barratt and Carslaw, 2014] Barratt, B. and Carslaw, D. (2014). Impacts of the bus retrofit programme on no₂ concentrations along putney high street. *Environ. Res. Group, Rep. Prep, London brgh. Wandsworth*.
- [Beattie et al., 2001] Beattie, C., Longhurst, J., and Woodfield, N. (2001). Air quality management: evolution of policy and practice in the uk as exemplified by the experience of english local government. *Atmospheric Environment*, 35(8):1479 – 1490.
- [Beevers and Carslaw, 2005] Beevers, S. D. and Carslaw, D. C. (2005). The impact of congestion charging on vehicle emissions in london. *Atmospheric Environment*, 39(1):1 – 5.
- [Beevers et al., 2016] Beevers, S. D., Carslaw, D. C., Dajnak, D., Stewart, G. B., Williams, M. L., Fussell, J. C., and Kelly, F. J. (2016). Traffic management strategies for emissions reduction: recent experience in london. *Energy Emission Control Technol*, 4(5):27–39.
- [Bell and Davis, 2001] Bell, M. L. and Davis, D. L. (2001). Reassessment of the lethal london fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environmental Health Perspectives*, 109(suppl 3):389–394.
- [Bergmann et al., 2009] Bergmann, M., Kirchner, U., Vogt, R., and Benter, T. (2009). On-road and laboratory investigation of low-level pm emissions of a modern diesel particulate filter equipped diesel passenger car. *Atmospheric Environment*, 43(11):1908 – 1916.
- [Berkowicz et al., 1997] Berkowicz, R., Hertel, O., Larsen, S. E., Sørensen, N. N., and Nielsen, M. (1997). Modelling traffic pollution in streets.
- [Bickel and Friedrich, 2005] Bickel, P. and Friedrich, R. (2005). Externe: Externalities of energy. methodology 2005 update. Technical report, European Commission.
- [Bloomberg and Dale, 2000] Bloomberg, L. and Dale, J. (2000). A comparison of the vissim and corsim traffic simulation models. In *Institute of Transportation Engineers Annual Meeting*, volume 1727, pages 52–60.

- [Borge et al., 2012] Borge, R., [de Miguel], I., [de la Paz], D., Lumbreras, J., Pérez, J., and Rodríguez, E. (2012). Comparison of road traffic emission models in madrid (spain). *Atmospheric Environment*, 62:461 – 471.
- [Boriboonsomsin and Barth, 2009] Boriboonsomsin, K. and Barth, M. (2009). Impacts of road grade on fuel consumption and carbon dioxide emissions evidenced by use of advanced navigation systems. *Transportation Research Record*, 2139(1):21–30.
- [Borken-Kleefeld and Chen, 2015] Borken-Kleefeld, J. and Chen, Y. (2015). New emission deterioration rates for gasoline cars – results from long-term measurements. *Atmospheric Environment*, 101:58 – 64.
- [Boulter and McCrae, 2007] Boulter, P. and McCrae, I. (2007). ARTEMIS: Assessment and Reliability of Transport Emission Models and Inventory Systems-Final Report. *TRL Published Project Report*.
- [Brimblecombe, 2006] Brimblecombe, P. (2006). The clean air act after 50 years. *Weather*, 61(11):311–314.
- [Brimblecombe and Makra, 2005] Brimblecombe, P. and Makra, L. (2005). Selections from the history of environmental pollution, with special attention to air pollution. part 2*: From medieval times to the 19th century. *International journal of environment and pollution*, 23(4):351–367.
- [Carslaw et al., 2019] Carslaw, D. C., Farren, N. J., Vaughan, A. R., Drysdale, W. S., Young, S., and Lee, J. D. (2019). The diminishing importance of nitrogen dioxide emissions from road vehicle exhaust. *Atmospheric Environment: X*, 1:100002.
- [Carslaw et al., 2015] Carslaw, D. C., Priestman, M., Williams, M. L., Stewart, G. B., and Beevers, S. D. (2015). Performance of optimised scr retrofit buses under urban driving and controlled conditions. *Atmospheric Environment*, 105:70 – 77.
- [Carslaw and Rhys-Tyler, 2013] Carslaw, D. C. and Rhys-Tyler, G. (2013). New insights from comprehensive on-road measurements of nox, no2 and nh3 from vehicle emission remote sensing in london, uk. *Atmospheric Environment*, 81:339 – 347.

- [Caudet and Pietila, 2015a] Caudet, L. and Pietila, H. (2015a). Commission welcomes member states' agreement on robust testing of air pollution emissions by cars. http://europa.eu/rapid/press-release_IP-15-5945_en.htm. Accessed: 2018-06-18.
- [Caudet and Pietila, 2015b] Caudet, L. and Pietila, H. (2015b). Faq - air pollutant emissions standards. https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_5705. Accessed: 2020-01-24.
- [Cecil, 2016] Cecil, N. (2016). London parents see toxic air as 'the biggest health threat to their children'. <https://www.standard.co.uk/news/london/london-parents-see-toxic-air-as-the-biggest-health-threat-to-their-children-a3208221.html>. Accessed: 2020-11-18.
- [Çelik and Ozdalyan, 2010] Çelik, M. B. and Ozdalyan, B. (2010). *Gasoline direct injection*. London: InTech Open.
- [CERAM, 2016] CERAM, U. (2016). Dossier de presse: résultats des contrôles des émissions de polluants atmosphériques et de co2 menés sur les 52 premiers véhicules 2016. *Ministère de l'environnement, de l'énergie et de la mer*.
- [Cha et al., 2019] Cha, J., Lee, J., and Chon, M. S. (2019). Evaluation of real driving emissions for euro 6 light-duty diesel vehicles equipped with lnt and scr on domestic sales in korea. *Atmospheric Environment*, 196:133 – 142.
- [Chan et al., 2012] Chan, T. W., Meloche, E., Rosenblatt, D., Kubsh, J., Brezny, R., and Rideout, G. (2012). Reducing particulate emissions for future gasoline direct injection vehicles with a gasoline particulate filter. In *ETH Conference on Combustion Generated Nanoparticles*.
- [Chang and Hanna, 2004] Chang, J. C. and Hanna, S. R. (2004). Air quality model performance evaluation. *Meteorology and Atmospheric Physics*, 87(1):167–196.
- [Choi et al., 2007] Choi, J.-S., Partridge, W. P., and Daw, C. S. (2007). Sulfur impact on nox storage, oxygen storage, and ammonia breakthrough during cyclic lean/rich operation of a commercial lean nox trap. *Applied Catalysis B: Environmental*, 77(1):145 – 156.

- [Chu et al., 2019] Chu, T., Wang, J., Codecà, L., and Li, Z. (2019). Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):1086–1095.
- [Ciardelli et al., 2007] Ciardelli, C., Nova, I., Tronconi, E., Chatterjee, D., Bandl-Konrad, B., Weibel, M., and Krutzsch, B. (2007). Reactivity of no/no₂-nh₃ scr system for diesel exhaust aftertreatment: Identification of the reaction network as a function of temperature and no₂ feed content. *Applied Catalysis B: Environmental*, 70(1):80 – 90. Papers presented at the 4th International Conference on Environmental Catalysis (4th ICEC)Heidelberg, Germany, June 05 – 08, 2005.
- [Clevert et al., 2015] Clevert, D., Unterthiner, T., and Hochreiter, S. (2015). Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *CoRR*, abs/1511.07289.
- [COMEAP, 2017] COMEAP (2017). The effects of long-term exposure to ambient air pollution on cardiovascular morbidity: Mechanistic evidence. Technical report, Committee on the Medical Effects of Air Pollutants.
- [Contag et al., 2017] Contag, M., Li, G., Pawlowski, A., Domke, F., Levchenko, K., Holz, T., and Savage, S. (2017). How they did it: An analysis of emission defeat devices in modern automobiles. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 231–250.
- [Copat et al., 2020] Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., Conti, G. O., and Ferrante, M. (2020). The role of air pollution (pm and no₂) in covid-19 spread and lethality: A systematic review. *Environmental Research*, 191:110129.
- [Corradini, 2001] Corradini, A. (2001). Dynamic time warping for off-line recognition of a small gesture vocabulary. In *Proceedings IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, pages 82–89.
- [Czink et al., 2006] Czink, N., Cera, P., Salo, J., Bonek, E., Nuutinen, J.-P., and Ylitalo, J. (2006). A framework for automatic clustering of parametric MIMO channel data including path powers. In *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*, pages 1–5. IEEE.

- [Davison et al., 2020] Davison, J., Bernard, Y., Borken-Kleefeld, J., Farren, N. J., Hausberger, S., Åke Sjödin, Tate, J. E., Vaughan, A. R., and Carslaw, D. C. (2020). Distance-based emission factors from vehicle emission remote sensing measurements. *Science of The Total Environment*, 739:139688.
- [de Lucas et al., 2001] de Lucas, A., Durán, A., Carmona, M., and Lapuerta, M. (2001). Modeling diesel particulate emissions with neural networks. *Fuel*, 80(4):539 – 548.
- [de Nazelle et al., 2010] de Nazelle, A., Morton, B. J., Jerrett, M., and Crawford-Brown, D. (2010). Short trips: An opportunity for reducing mobile-source emissions? *Transportation Research Part D: Transport and Environment*, 15(8):451 – 457.
- [Degraeuwe et al., 2016] Degraeuwe, B., Thunis, P., Clappier, A., Weiss, M., Lefebvre, W., Janssen, S., and Vranckx, S. (2016). Impact of passenger car nox emissions and no2 fractions on urban no2 pollution – scenario analysis for the city of antwerp, belgium. *Atmospheric Environment*, 126:218 – 224.
- [Degraeuwe and Weiss, 2017] Degraeuwe, B. and Weiss, M. (2017). Does the new european driving cycle (nedc) really fail to capture the nox emissions of diesel cars in europe? *Environmental Pollution*, 222:234 – 241.
- [Demuynck et al., 2019] Demuynck, J., Favre, C., Bosteels, D., Bunar, F., Spitta, J., and Kuhrt, A. (2019). Diesel vehicle with ultra-low nox emissions on the road. Technical report, SAE Technical Paper.
- [Demuynck et al., 2017] Demuynck, J., Favre, C., Bosteels, D., Hamje, H., and Andersson, J. (2017). Real-world emissions measurements of a gasoline direct injection vehicle without and with a gasoline particulate filter. Technical report, SAE Technical Paper.
- [Department for Transport, 2016] Department for Transport (2016). Vehicle emissions testing programme: data and conclusions. <https://www.gov.uk/government/publications/vehicle-emissions-testing-programme-conclusions>. Accessed: 2020-11-30.

- [Department for Transport, 2019] Department for Transport (2019). Road traffic statistics: Site number 161841. <https://roadtraffic.dft.gov.uk/manualcountpoints/16841>. Accessed: 2020-12-16.
- [Department for Transport, 2020] Department for Transport (2020). Data on all licensed and registered cars, produced by department for transport. <https://www.gov.uk/government/statistical-data-sets/veh02-licensed-cars>. Accessed: 2020-11-30.
- [Department of the Environment Great Britain (Transport and the Regions.), 2000] Department of the Environment Great Britain (Transport and the Regions.) (2000). *The Air Quality Strategy for England, Scotland, Wales and Northern Ireland: Working Together for Clean Air*. Department of the Environment, transport and the regions.
- [Department of the Environment Great Britain (Transport and the Regions.), 2007] Department of the Environment Great Britain (Transport and the Regions.) (2007). *The Air Quality Strategy for England, Scotland, Wales and Northern Ireland: Working Together for Clean Air*. Department of the Environment, transport and the regions.
- [Dozat, 2016] Dozat, T. (2016). Incorporating Nesterov Momentum into Adam. Technical report, Stanford University.
- [Driver and Vehicle Standards Agency, 2018] Driver and Vehicle Standards Agency (2018). Vehicle market surveillance unit programme results, 2017. <https://www.gov.uk/government/publications/vehicle-market-surveillance-unit-programme-results-2017>. Accessed: 2020-11-30.
- [Driver and Vehicle Standards Agency, 2019] Driver and Vehicle Standards Agency (2019). Vehicle market surveillance unit programme results, 2018. <https://www.gov.uk/government/publications/vehicle-market-surveillance-unit-programme-results-2018>. Accessed: 2020-11-30.
- [Duchi et al., 2011] Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12:2121–2159.

- [EEA, 2019] EEA (2019). Air quality in Europe — 2019 report. Technical report, European Environment Agency.
- [Efron, 1992] Efron, B. (1992). *Bootstrap Methods: Another Look at the Jackknife*, pages 569–593. Springer New York, New York, NY.
- [Environmental Agency, 2017] Environmental Agency (2017). OS MasterMap Building Heights Layer. Last accessed 15 December 2017.
- [Ericsson, 2001] Ericsson, E. (2001). Independent driving pattern factors and their influence on fuel-use and exhaust emission factors. *Transportation Research Part D: Transport and Environment*, 6(5):325 – 345.
- [European Commission, 2017] European Commission (2017). Reducing co2 emissions from passenger cars. https://ec.europa.eu/clima/policies/transport/vehicles/cars_en. Accessed: 2020-02-17.
- [European Commission, 2018a] European Commission (2018a). Testing of emissions from cars. https://europa.eu/rapid/press-release_MEMO-18-3646_en.htm. Accessed: 2019-08-01.
- [European Commission, 2018b] European Commission (2018b). Testing of emissions from cars. https://europa.eu/rapid/press-release_MEMO-18-3646_en.htm. Accessed: 2019-08-01.
- [Favez et al., 2009] Favez, J.-Y., Weilenmann, M., and Stilli, J. (2009). Cold start extra emissions as a function of engine stop time: Evolution over the last 10 years. *Atmospheric Environment*, 43(5):996 – 1007. Air Pollution Related to Transport.
- [Ferguson and Kirkpatrick, 2015] Ferguson, C. R. and Kirkpatrick, A. T. (2015). *Internal combustion engines: applied thermosciences*. John Wiley & Sons.
- [Fomunung et al., 2001] Fomunung, I., Washington, S., Guensler, R., and Bachman, W. (2001). Validation of the measure automobile emissions model: a statistical analysis. *Journal of Transportation and Statistics (September 2000)*, pages 65–84.

- [Fontaras et al., 2017a] Fontaras, G., Ciuffo, B., Zacharof, N., Tsiakmakis, S., Marotta, A., Pavlovic, J., and Anagnostopoulos, K. (2017a). The difference between reported and real-world co2 emissions: How much improvement can be expected by wltip introduction. *Transportation Research Procedia*, 25:3933–3943.
- [Fontaras et al., 2014] Fontaras, G., Franco, V., Dilara, P., Martini, G., and Manfredi, U. (2014). Development and review of euro 5 passenger car emission factors based on experimental results over various driving cycles. *Science of The Total Environment*, 468-469:1034 – 1042.
- [Fontaras and Samaras, 2010] Fontaras, G. and Samaras, Z. (2010). On the way to 130gco2/km—estimating the future characteristics of the average european passenger car. *Energy Policy*, 38(4):1826 – 1833. Energy Security - Concepts and Indicators with regular papers.
- [Fontaras et al., 2017b] Fontaras, G., Zacharof, N.-G., and Ciuffo, B. (2017b). Fuel consumption and co2 emissions from passenger cars in europe – laboratory versus real-world emissions. *Progress in Energy and Combustion Science*, 60:97 – 131.
- [Franco et al., 2014] Franco, V., Sánchez, F. P., German, J., and Mock, P. (2014). Real-world exhaust emissions from modern diesel cars. *communications*, 49(30):847129–102.
- [Frey et al., 2008] Frey, H. C., Zhang, K., and Rouphail, N. M. (2008). Fuel use and emissions comparisons for alternative routes, time of day, road grade, and vehicles based on in-use measurements. *Environmental Science & Technology*, 42(7):2483–2489.
- [Gallus et al., 2017] Gallus, J., Kirchner, U., Vogt, R., and Benter, T. (2017). Impact of driving style and road grade on gaseous exhaust emissions of passenger vehicles measured by a portable emission measurement system (pems). *Transportation Research Part D: Transport and Environment*, 52:215 – 226.
- [Gallus et al., 2016] Gallus, J., Kirchner, U., Vogt, R., Börensen, C., and Benter, T. (2016). On-road particle number measurements using a portable emission measurement system (pems). *Atmospheric Environment*, 124:37 – 45.

- [Gao et al., 2017] Gao, J., Shen, Y., Liu, J., Ito, M., and Shiratori, N. (2017). Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network. *CoRR*, abs/1705.02755.
- [Gardnera and Dorlinga, 1998] Gardnera, M. W. and Dorlinga, S. R. (1998). Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14-15):2627–2636.
- [Genders and Razavi, 2016] Genders, W. and Razavi, S. N. (2016). Using a deep reinforcement learning agent for traffic signal control. *CoRR*, abs/1611.01142.
- [Giechaskiel et al., 2014] Giechaskiel, B., Manfredi, U., and Martini, G. (2014). Engine exhaust solid sub-23 nm particles: I. literature survey. *SAE International Journal of Fuels and Lubricants*, 7(3):950–964.
- [GLA, 2020] GLA (2020). Central london ultra low emission zone - 10 months report. Technical report, Greater London Authority.
- [Gomes et al., 2004] Gomes, G., May, A., and Horowitz, R. (2004). Congested freeway microsimulation model using vissim. *Transportation Research Record*, 1876(1):71–81.
- [Gordon et al., 2004] Gordon, N. D., McMahon, T. A., Finlayson, B. L., Gippel, C. J., and Nathan, R. J. (2004). *Stream hydrology: an introduction for ecologists*. John Wiley and Sons.
- [Grange and Carslaw, 2019] Grange, S. K. and Carslaw, D. C. (2019). Using meteorological normalisation to detect interventions in air quality time series. *Science of The Total Environment*, 653:578 – 588.
- [Grylls et al., 2019] Grylls, T., Cornec, C. M. L., Salizzoni, P., Soulhac, L., Stettler, M. E., and [van Reeuwijk], M. (2019). Evaluation of an operational air quality model using large-eddy simulation. *Atmospheric Environment: X*, 3:100041.

- [Hagena et al., 2006] Hagena, J. R., Filipi, Z., and Assanis, D. N. (2006). Transient Diesel Emissions: Analysis of Engine Operation During a Tip-In. In *SAE 2006 World Congress and Exhibition*. SAE International.
- [Harold, 2012] Harold, M. P. (2012). Nox storage and reduction in lean burn vehicle emission control: a catalytic engineer’s playground. *Current Opinion in Chemical Engineering*, 1(3):303 – 311. Energy and environmental engineering / Reaction engineering and catalysis.
- [Harrison et al., 2012] Harrison, R. M., Laxen, D., Moorcroft, S., and Laxen, K. (2012). Processes affecting concentrations of fine particulate matter (pm2.5) in the uk atmosphere. *Atmospheric Environment*, 46:115 – 124.
- [Hatzopoulou and Miller, 2010] Hatzopoulou, M. and Miller, E. (2010). Linking an activity-based travel demand model with traffic emission and dispersion models: Transport’s contribution to air pollution in Toronto. *Transportation Research Part D: Transport and Environment*, 15(6):315 – 325.
- [Hausberger, 2003] Hausberger, S. (2003). *Simulation of Real World Vehicle Exhaust Emission*. PhD thesis.
- [Hausberger et al., 2003] Hausberger, S., Rodler, J., Sturm, P., and Rexeis, M. (2003). Emission factors for heavy-duty vehicles and validation by tunnel measurements. *Atmospheric Environment*, 37(37):5237 – 5245. 11th International Symposium, Transport and Air Pollution.
- [He and Hou, 2012] He, J. and Hou, Z. (2012). Ant colony algorithm for traffic signal timing optimization. *Advances in Engineering Software*, 43(1):14 – 18.
- [Higuera et al., 2019] Higuera, C., Lozano, F., Camacho, E. C., and Higuera, C. H. (2019). Multiagent reinforcement learning applied to traffic light signal control. In Demazeau, Y., Matson, E., Corchado, J. M., and De la Prieta, F., editors, *Advances in Practical Applications of Survivable Agents and Multi-Agent Systems: The PAAMS Collection*, pages 115–126, Cham. Springer International Publishing.

- [Hinton, 2016] Hinton, S. R. (2016). ChainConsumer. *The Journal of Open Source Software*, 1:00045.
- [Holmes and Morawska, 2006] Holmes, N. and Morawska, L. (2006). A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available. *Atmospheric Environment*, 40(30):5902 – 5928.
- [Holmén and Niemeier, 1998] Holmén, B. A. and Niemeier, D. A. (1998). Characterizing the effects of driver variability on real-world vehicle emissions. *Transportation Research Part D: Transport and Environment*, 3(2):117 – 128.
- [Hu et al., 2012] Hu, J., Wu, Y., Wang, Z., Li, Z., Zhou, Y., Wang, H., Bao, X., and Hao, J. (2012). Real-world fuel efficiency and exhaust emissions of light-duty diesel vehicles and their correlation with road conditions. *Journal of Environmental Sciences*, 24(5):865 – 874.
- [Huang et al., 2017] Huang, J., Yu, H., Dai, A., Wei, Y., and Kang, L. (2017). Drylands face potential threat under 2 c global warming target. *Nature Climate Change*, 7(6):417.
- [Izakian et al., 2015] Izakian, H., Pedrycz, W., and Jamal, I. (2015). Fuzzy clustering of time series data using dynamic time warping distance. *Engineering Applications of Artificial Intelligence*, 39:235 – 244.
- [Jackson et al., 2006] Jackson, E., Qu, Y., Holmén, B. A., and Aultman-Hall, L. (2006). Driver and road type effects on light-duty gas and particulate emissions. *Transportation Research Record*, 1987(1):118–127.
- [Jaikumar et al., 2017] Jaikumar, R., Nagendra, S. S., and Sivanandan, R. (2017). Modeling of real time exhaust emissions of passenger cars under heterogeneous traffic conditions. *Atmospheric Pollution Research*, 8(1):80 – 88.
- [Jain, 2010] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8):651 – 666. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR).

- [Johnson, 2008] Johnson, T. (2008). Diesel engine emissions and their control. *Platinum Metals Review*, 52(1):23–37.
- [Johri and Filipi, 2014] Johri, R. and Filipi, Z. (2014). Neuro-fuzzy model tree approach to virtual sensing of transient diesel soot and nox emissions. *International Journal of Engine Research*, 15(8):918–927.
- [Jonson et al., 2017] Jonson, J. E., Borken-Kleefeld, J., Simpson, D., Nyíri, A., Posch, M., and Heyes, C. (2017). Impact of excess NO_x emissions from diesel cars on air quality, public health and eutrophication in europe. *Environmental Research Letters*, 12(9):094017.
- [Kalman and Kwasny, 1992] Kalman, B. L. and Kwasny, S. C. (1992). Why tanh: choosing a sigmoidal function. In *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, volume 4, pages 578–581 vol.4.
- [Kesgin, 2004] Kesgin, U. (2004). Genetic algorithm and artificial neural network for engine optimisation of efficiency and nox emission. *Fuel*, 83(7):885 – 895.
- [Kingma and Ba, 2014] Kingma, D. P. and Ba, J. (2014). Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980.
- [Klutse et al., 2018] Klutse, N. A. B., Ajayi, V. O., Gbobaniyi, E. O., Egbebiyi, T. S., Kouadio, K., Nkrumah, F., Quagraine, K. A., Olusegun, C., Diasso, U., Abiodun, B. J., et al. (2018). Potential impact of 1.5 c and 2 c global warming on consecutive dry and wet days over west africa. *Environmental Research Letters*, 13(5):055013.
- [Kousoulidou et al., 2010] Kousoulidou, M., Ntziachristos, L., Gkeivanidis, S., Samaras, Z., Franco, V., and Dilara, P. (2010). Validation of the copert road emission inventory model with real-use data. In *US EPA 19th Annual International Emission Inventory Conference. Emissions Inventories—Informing Emerging Issues*, volume 40.
- [Kühlwein and Friedrich, 2000] Kühlwein, J. and Friedrich, R. (2000). Uncertainties of modelling emissions from road transport. *Atmospheric Environment*, 34(27):4603 – 4610.

- [Lairenlakpam et al., 2018] Lairenlakpam, R., Jain, A., Gupta, P., Kamei, W., Badola, R., and Singh, Y. (2018). Effect of real world driving and different drive modes on vehicle emissions and fuel consumption. *SAE Technical Paper*.
- [Lapuerta et al., 2012] Lapuerta, M., Rodríguez-Fernández, J., and Oliva, F. (2012). Effect of soot accumulation in a diesel particle filter on the combustion process and gaseous emissions. *Energy*, 47(1):543 – 552. Asia-Pacific Forum on Renewable Energy 2011.
- [Le Cornec et al., 2020] Le Cornec, C. M., Molden, N., van Reeuwijk, M., and Stettler, M. E. (2020). Modelling of instantaneous emissions from diesel vehicles with a special focus on nox: Insights from machine learning techniques. *Science of The Total Environment*, 737:139625.
- [Le Cornec, 2021] Le Cornec, C. M. A. (2021). Revs. <https://github.com/clemlecornec/REVS>.
- [Le Cornec, Clemence M. A, 2019] Le Cornec, Clemence M. A (2019). Ann.emissions. <https://github.com/clemlecornec/ANN.Emissions>, [Online; accessed 13-January-2020].
- [Lee, 1994] Lee, S. M. S. (1994). Optimal choice between parametric and non-parametric bootstrap estimates. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 115, pages 335–363. Cambridge University Press.
- [Lejri et al., 2018] Lejri, D., Can, A., Schiper, N., and Leclercq, L. (2018). Accounting for traffic speed dynamics when calculating copert and phem pollutant emissions at the urban scale. *Transportation Research Part D: Transport and Environment*, 63:588 – 603.
- [Lejri and Leclercq, 2020] Lejri, D. and Leclercq, L. (2020). Are average speed emission functions scale-free? *Atmospheric Environment*, 224:117324.
- [Leonard and Kramer, 1990] Leonard, J. and Kramer, M. (1990). Improvement of the back-propagation algorithm for training neural networks. *Computers and Chemical Engineering*, 14(3):337 – 341.
- [Liang et al., 2013] Liang, B., Ge, Y., Tan, J., Han, X., Gao, L., Hao, L., Ye, W., and Dai, P. (2013). Comparison of pm emissions from a gasoline direct injected (gdi) vehicle and a port

fuel injected (pfi) vehicle measured by electrical low pressure impactor (elpi) with two fuels: Gasoline and m15 methanol gasoline. *Journal of Aerosol Science*, 57:22 – 31.

[Liang et al., 2019] Liang, X., Du, X., Wang, G., and Han, Z. (2019). A deep reinforcement learning network for traffic light cycle control. *IEEE Transactions on Vehicular Technology*, 68(2):1243–1253.

[Ligterink et al., 2013] Ligterink, N., Kadijk, G., van Mensch, P., Hausberger, S., and Rexeis, M. (2013). *Investigations and real world emission performance of Euro 6 light-duty vehicles*. Delft: TNO.

[Linares et al., 2018] Linares, C., Falcón, I., Ortiz, C., and Díaz, J. (2018). An approach estimating the short-term effect of no2 on daily mortality in spanish cities. *Environment International*, 116:18 – 28.

[Luján et al., 2018] Luján, J. M., Bermúdez, V., Dolz, V., and Monsalve-Serrano, J. (2018). An assessment of the real-world driving gaseous emissions from a euro 6 light-duty diesel vehicle using a portable emissions measurement system (pems). *Atmospheric Environment*, 174:112 – 121.

[Lv et al., 2013] Lv, J., Zhang, Y., and Zietsman, J. (2013). Investigating emission reduction benefit from intersection signal optimization. *Journal of Intelligent Transportation Systems*, 17(3):200–209.

[Madireddy et al., 2011] Madireddy, M., De Coensel, B., Can, A., Degraeuwe, B., Beusen, B., De Vlieger, I., and Botteldooren, D. (2011). Assessment of the impact of speed limit reduction and traffic signal coordination on vehicle emissions using an integrated approach. *Transportation Research Part D: Transport and Environment*, 16(7):504 – 508.

[Mahesh et al., 2018] Mahesh, S., Ramadurai, G., and Nagendra, S. S. (2018). Real-world emissions of gaseous pollutants from diesel passenger cars using portable emission measurement systems. *Sustainable Cities and Society*, 41:104 – 113.

- [Makra and Brimblecombe, 2004] Makra, L. and Brimblecombe, P. (2004). Selections from the history of environmental pollution, with special attention to air pollution. part 1. *International journal of environment and pollution*, 22(6):641–656.
- [Manzie et al., 2007] Manzie, C., Watson, H., and Halgamuge, S. (2007). Fuel economy improvements for urban driving: Hybrid vs. intelligent vehicles. *Transportation Research Part C: Emerging Technologies*, 15(1):1 – 16.
- [Martin et al., 2017] Martin, R. S., Woods, C., and Thomas, J. (2017). Assessment of automobile start and idling emissions under utah specific conditions: Cold start, hot start, and idle emissions as measured on northern utah vehicles.
- [Matzer et al., 2017] Matzer, C., Hausberger, S., Lipp, S., and Rexeis, M. (2017). A new approach for systematic use of pems data in emission simulation. *Journal of Earth Sciences and Geotechnical Engineering*, 7(1):33–50.
- [Maulik and Bandyopadhyay, 2002] Maulik, U. and Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12):1650–1654.
- [Mavroidis and Chaloulakou, 2011] Mavroidis, I. and Chaloulakou, A. (2011). Long-term trends of primary and secondary NO₂ production in the Athens area. Variation of the NO₂/NO_x ratio. *Atmospheric Environment*, 45(38):6872 – 6879.
- [May et al., 2014] May, J., Bosteels, D., and Favre, C. (2014). An assessment of emissions from light-duty vehicles using pems and chassis dynamometer testing. *SAE International Journal of Engines*, 7(3):1326–1335.
- [Mollenhauer et al., 2010] Mollenhauer, K., Tschöke, H., and Johnson, K. G. (2010). *Handbook of diesel engines*, volume 1. Springer.
- [Monks et al., 2012] Monks, P., ApSimon, H., Carruthers, D., Carslaw, D., Derwent, D., Harrison, R., Laxen, D., and Stedman, J. (2012). Fine particulate matter (pm_{2.5}) in the united kingdom. Technical report, Air Quality Expert Group.

- [Monks et al., 2009] Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R., Carslaw, K., Cooper, O., Dentener, F., Fowler, D., Fragkou, E., Frost, G., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I., Jenkin, M., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M., Lee, J., Liousse, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J., O'Dowd, C., Palmer, P., Parrish, D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A., Reeves, C., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., ten Brink, H., Theloke, J., van der Werf, G., Vautard, R., Vestreng, V., Vlachokostas, C., and von Glasow, R. (2009). Atmospheric composition change – global and regional air quality. *Atmospheric Environment*, 43(33):5268 – 5350. ACCENT Synthesis.
- [Moody and Tate, 2017] Moody, A. and Tate, J. (2017). In service co2 and nox emissions of euro 6/vi cars, light-and heavy-dutygoods vehicles in real london driving: Taking the road into the laboratory. *Journal of Earth Sciences and Geotechnical Engineering*, 7(1):51–62.
- [Moseholm et al., 1996] Moseholm, L., Silva, J., and Larson, T. (1996). Forecasting carbon monoxide concentrations near a sheltered intersection using video traffic surveillance and neural network. *Transportation research*.
- [Mousavi et al., 2017] Mousavi, S. S., Schukat, M., and Howley, E. (2017). Traffic light control using deep policy-gradient and value-function-based reinforcement learning. *IET Intelligent Transport Systems*, 11(7):417–423.
- [Muller, 2007] Muller, M. (2007). *Dynamic Time Warping*, pages 69–84. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Myers et al., 1980] Myers, C., Rabiner, L., and Rosenberg, A. (1980). Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(6):623–635.

- [Myung et al., 2017] Myung, C.-L., Jang, W., Kwon, S., Ko, J., Jin, D., and Park, S. (2017). Evaluation of the real-time de-nox performance characteristics of a lnt-equipped euro-6 diesel passenger car with various vehicle emissions certification cycles. *Energy*, 132:356 – 369.
- [Myung et al., 2013] Myung, C.-L., Ko, A., Kim, J., Choi, K., Kwon, S., and Park, S. (2013). Specific engine performance and gaseous emissions characteristics of european test cycle and worldwide harmonized driving cycle for a heavy-duty diesel engine. *Journal of Mechanical Science and Technology*, 27(12):3893–3902.
- [National Atmospheric Emissions Inventory, 2020] National Atmospheric Emissions Inventory (2020). Emission factors for transport. <https://naei.beis.gov.uk/data/ef-transport>. Accessed: 2020-12-16.
- [Navone and Ceccatto, 1994] Navone, H. D. and Ceccatto, H. A. (1994). Predicting Indian monsoon rainfall: a neural network approach. *Climate Dynamics*, 10:305–312.
- [Neeft et al., 1996] Neeft, J. P., Makkee, M., and Moulijn, J. A. (1996). Diesel particulate emission control. *Fuel Processing Technology*, 47(1):1 – 69.
- [Negrenti, 1999] Negrenti, E. (1999). The ‘corrected average speed’ approach in enea’s tee model: an innovative solution for the evaluation of the energetic and environmental impacts of urban transport policies. *Science of The Total Environment*, 235(1):411 – 413.
- [Nowak and Pielecha, 2017] Nowak, M. and Pielecha, J. (2017). Comparison of exhaust emission on the basis of real driving emissions measurements and simulations. In *MATEC Web of Conferences*, volume 118, page 00026. EDP Sciences.
- [Ntziachristos et al., 2009] Ntziachristos, L., Gkatzoflias, D., Kouridis, C., and Samaras, Z. (2009). Copert: a european road transport emission inventory model. In *Information technologies in environmental engineering*, pages 491–504. Springer.
- [Ntziachristos et al., 2016] Ntziachristos, L., Papadimitriou, G., Ligterink, N., and Hausberger, S. (2016). Implications of diesel emissions control failures to emission factors and road transport NOx evolution. *Atmospheric Environment*, 141:542 – 551.

- [Obodeh and Ajuwa, 2009] Obodeh, O. and Ajuwa, C. (2009). Evaluation of artificial neural network performance in predicting diesel engine nox emissions. *European Journal of Scientific Research*, 33(4):642–653.
- [ODriscoll, 2017] ODriscoll, R. (2017). *Uncertainty in exhaust emissions from passenger cars; a policy perspective*. PhD thesis, Imperial College London.
- [O’Driscoll et al., 2016] O’Driscoll, R., ApSimon, H. M., Oxley, T., Molden, N., Stettler, M. E., and Thiyagarajah, A. (2016). A portable emissions measurement system (pems) study of nox and primary no2 emissions from euro 6 diesel passenger cars and comparison with copert emission factors. *Atmospheric Environment*, 145:81 – 91.
- [O’Driscoll et al., 2018] O’Driscoll, R., Stettler, M. E., Molden, N., Oxley, T., and ApSimon, H. M. (2018). Real world co2 and nox emissions from 149 euro 5 and 6 diesel, gasoline and hybrid passenger cars. *Science of The Total Environment*, 621:282 – 290.
- [Official Journal of the European Union, 2017] Official Journal of the European Union (2017). COMMISSION REGULATION (EU) 2017/1151 of 1 June 2017 supplementing Regulation (EC) No 715/2007 of the European Parliament and of the Council on type-approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information, amending Directive 2007/46/EC of the European Parliament and of the Council, Commission Regulation (EC) No 692/2008 and Commission Regulation (EU) No 1230/2012 and repealing Commission Regulation (EC) No 692/2008. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R1151&from=EN>. Accessed: 2020-05-21.
- [Oke, 1978] Oke, T. R. (1978). *Boundary layer climates*. Methuen, London.
- [O’Keefe et al., 2007] O’Keefe, M. P., Simpson, A., Kelly, K. J., and Pedersen, D. S. (2007). Duty cycle characterization and evaluation towards heavy hybrid vehicle applications. Technical report, SAE Technical Paper.

- [Olivera et al., 2015] Olivera, A. C., García-Nieto, J., and Alba, E. (2015). Reducing vehicle emissions and fuel consumption in the city by using particle swarm optimization. *Applied Intelligence*, 42(3):389–405.
- [Oxley et al., 2009] Oxley, T., Valiantis, M., Elshkaki, A., and ApSimon, H. (2009). Background, road and urban transport modelling of air quality limit values (the brutal model). *Environmental Modelling and Software*, 24(9):1036 – 1050.
- [Pacifici et al., 2009] Pacifici, F., Chini, M., and Emery, W. J. (2009). A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113(6):1276–1292.
- [Panis et al., 2006] Panis, L. I., Broekx, S., and Liu, R. (2006). Modelling instantaneous traffic emission and the influence of traffic speed limits. *Science of the total environment*, 371(1-3):270–285.
- [Papadopoulos et al., 2018] Papadopoulos, G., Keramydas, C., Ntziachristos, L., Lo, T.-S., Ng, K.-L., Wong, H.-L. A., and Wong, C. K.-L. (2018). Emission factors for a taxi fleet operating on liquefied petroleum gas (lpg) as a function of speed and road slope. *Frontiers in Mechanical Engineering*, 4:19.
- [Park and Jun, 2009] Park, H.-S. and Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications*, 36(2, Part 2):3336 – 3341.
- [Parlak et al., 2006] Parlak, A., Islamoglu, Y., Yasar, H., and Egrisogut, A. (2006). Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a diesel engine. *Applied Thermal Engineering*, 26(8):824 – 828.
- [Pavlovic et al., 2018] Pavlovic, J., Ciuffo, B., Fontaras, G., Valverde, V., and Marotta, A. (2018). How much difference in type-approval co2 emissions from passenger cars in europe can be expected from changing to the new test procedure (nedc vs. wltp)? *Transportation Research Part A: Policy and Practice*, 111:136 – 147.

- [Pavlovic et al., 2017] Pavlovic, J., Clairotte, M., Anagnostopoulos, K., Arcidiacono, V., Fontaras, G., and Ciuffo, B. (2017). Characterisation of real-world CO₂ variability and implications for future policy instruments. Technical report, European Commission.
- [Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- [Pelkmans and Debal, 2006] Pelkmans, L. and Debal, P. (2006). Comparison of on-road emissions with emissions measured on chassis dynamometer test cycles. *Transportation Research Part D: Transport and Environment*, 11(4):233 – 241.
- [Peters et al., 2009] Peters, J., McCourt, R., and Hurtado, R. (2009). Reducing carbon emissions and congestion by coordinating traffic signals. *Institute of Transportation Engineers. ITE Journal*, 79(4):25.
- [Petitjean et al., 2014] Petitjean, F., Forestier, G., Webb, G. I., Nicholson, A. E., Chen, Y., and Keogh, E. (2014). Dynamic Time Warping Averaging of Time Series Allows Faster and More Accurate Classification. In *2014 IEEE International Conference on Data Mining*, pages 470–479.
- [Phantawesak et al., 2021] Phantawesak, N., Coyle, F., and Stettler, M. E. J. (2021). In-use NOx emissions from London buses with retrofitted NOx aftertreatment. *in prep.*
- [Popoola et al., 2018] Popoola, O. A., Carruthers, D., Lad, C., Bright, V. B., Mead, M. I., Stettler, M. E., Saffell, J. R., and Jones, R. L. (2018). Use of networks of low cost air quality sensors to quantify air quality in urban settings. *Atmospheric Environment*, 194:58 – 70.
- [Prakash and Bodisco, 2019] Prakash, S. and Bodisco, T. A. (2019). An investigation into the effect of road gradient and driving style on nox emissions from a diesel vehicle driven on urban roads. *Transportation Research Part D: Transport and Environment*, 72:220 – 231.
- [Putha et al., 2012] Putha, R., Quadrifoglio, L., and Zechman, E. (2012). Comparing ant colony optimization and genetic algorithm approaches for solving traffic signal coordina-

- tion under oversaturation conditions. *Computer-Aided Civil and Infrastructure Engineering*, 27(1):14–28.
- [Péres et al., 2018] Péres, M., Ruiz, G., Nesmachnow, S., and Olivera, A. C. (2018). Multiobjective evolutionary optimization of traffic flow and pollution in montevideo, uruguay. *Applied Soft Computing*, 70:472 – 485.
- [Reşitoğlu et al., 2015] Reşitoğlu, İ. A., Altinişik, K., and Keskin, A. (2015). The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. *Clean Technologies and Environmental Policy*, 17(1):15–27.
- [Royal College of Physician, 2016] Royal College of Physician (2016). Every breath we take: the lifelong impact of air pollution. <https://www.rcplondon.ac.uk/projects/outputs/every-breath-we-take-lifelong-impact-air-pollution>. Accessed: 2020-11-17.
- [Ruder, 2016] Ruder, S. (2016). An overview of gradient descent optimization algorithms. *CoRR*, abs/1609.04747.
- [Saharidis and Konstantzos, 2018] Saharidis, G. K. and Konstantzos, G. E. (2018). Critical overview of emission calculation models in order to evaluate their potential use in estimation of greenhouse gas emissions from in port truck operations. *Journal of Cleaner Production*, 185:1024 – 1031.
- [Saliba et al., 2017] Saliba, G., Saleh, R., Zhao, Y., Presto, A. A., Lambe, A. T., Frodin, B., Sardar, S., Maldonado, H., Maddox, C., May, A. A., Drozd, G. T., Goldstein, A. H., Russell, L. M., Hagen, F., and Robinson, A. L. (2017). Comparison of gasoline direct-injection (gdi) and port fuel injection (pfi) vehicle emissions: Emission certification standards, cold-start, secondary organic aerosol formation potential, and potential climate impacts. *Environmental Science & Technology*, 51(11):6542–6552.
- [Salizzoni et al., 2009] Salizzoni, P., Soulhac, L., and Mejean, P. (2009). Street canyon ventilation and atmospheric turbulence. *Atmospheric Environment*, 43(32):5056 – 5067.
- [Salvador and Chan, 2007] Salvador, S. and Chan, P. (2007). FastDTW: Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11.5:561 – 580.

- [Salvator, Stan and Chan, Philip, 2019] Salvator, Stan and Chan, Philip (2019). fastdtw 0.3.4. <https://pypi.org/project/fastdtw/>, [Online; accessed 13-January-2020].
- [Salvi and Brickman, 2015] Salvi, C. and Brickman, S. (2015). Air pollution costs european economies us\$ 1.6 trillion a year in diseases and deaths, new who study says. [http://www.euro.who.int/en/media-centre/sections/press-releases/2015/04/air-pollution-costs-european-economies-us\\\$-1.6-trillion-a-year-in-diseases-and-deaths,-new-who-study-says](http://www.euro.who.int/en/media-centre/sections/press-releases/2015/04/air-pollution-costs-european-economies-us\$-1.6-trillion-a-year-in-diseases-and-deaths,-new-who-study-says). Accessed: 2020-04-21.
- [Samaras et al., 2014] Samaras, C., Tsokolis, D., Toffolo, S., Garcia-Castro, A., Vock, C., Ntziachristos, L., and Samaras, Z. (2014). Limits of applicability of COPERT model to short links and congested conditions. In *20th International Transport and Air Pollution Conference*.
- [Santos and Fraser, 2006] Santos, G. and Fraser, G. (2006). Road pricing: lessons from london. *Economic Policy*, 21(46):264–310.
- [Schejbal et al., 2010] Schejbal, M., Štěpánek, J., Marek, M., Kočí, P., and Kubíček, M. (2010). Modelling of soot oxidation by no2 in various types of diesel particulate filters. *Fuel*, 89(9):2365 – 2375.
- [Scora and Barth, 2006] Scora, G. and Barth, M. (2006). Comprehensive modal emissions model (CMEM), version 3.01. *User guide. Centre for environmental research and technology. University of California, Riverside*, 1070.
- [Sempena et al., 2011] Sempena, S., Maulidevi, N. U., and Aryan, P. R. (2011). Human action recognition using Dynamic Time Warping. In *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, pages 1–5.
- [Sentoff et al., 2015] Sentoff, K. M., Aultman-Hall, L., and Holmén, B. A. (2015). Implications of driving style and road grade for accurate vehicle activity data and emissions estimates. *Transportation Research Part D: Transport and Environment*, 35:175 – 188.
- [Shalizi, 2015] Shalizi, C. (2015). The bootstrap. <https://www.stat.cmu.edu/~cshalizi/uADA/15/lectures/06.pdf>. Accessed: 2020-11-12.

- [Sileghem et al., 2014] Sileghem, L., Bosteels, D., May, J., Favre, C., and Verhelst, S. (2014). Analysis of vehicle emission measurements on the new wltp, the nedc and the cadc. *Transportation Research Part D: Transport and Environment*, 32:70 – 85.
- [Sjödin and Jerksjö, 2008] Sjödin, A. and Jerksjö, M. (2008). *Evaluation of European road transport emission models against on-road emission data as measured by optical remote sensing*. Citeseer.
- [Smit et al., 2008] Smit, R., Brown, A., and Chan, Y. (2008). Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow? *Environmental Modelling & Software*, 23(10):1262 – 1270.
- [Smit et al., 2010] Smit, R., Ntziachristos, L., and Boulter, P. (2010). Validation of road vehicle and traffic emission models – a review and meta-analysis. *Atmospheric Environment*, 44(25):2943 – 2953.
- [Smit et al., 2007] Smit, R., Smokers, R., and Rabé, E. (2007). A new modelling approach for road traffic emissions: Versit+. *Transportation Research Part D: Transport and Environment*, 12(6):414 – 422.
- [Soulhac et al., 2017] Soulhac, L., Nguyen, C., Volta, P., and Salizzoni, P. (2017). The model sirane for atmospheric urban pollutant dispersion. part iii: Validation against no2 yearly concentration measurements in a large urban agglomeration. *Atmospheric Environment*, 167:377 – 388.
- [Soulhac et al., 2011] Soulhac, L., Salizzoni, P., Cierco, F.-X., and Perkins, R. (2011). The model sirane for atmospheric urban pollutant dispersion; part i, presentation of the model. *Atmospheric Environment*, 45(39):7379 – 7395.
- [Soulhac et al., 2012] Soulhac, L., Salizzoni, P., Mejean, P., Didier, D., and Rios, I. (2012). The model sirane for atmospheric urban pollutant dispersion; part ii, validation of the model on a real case study. *Atmospheric Environment*, 49:320 – 337.

- [Sun et al., 2020] Sun, Z., Wang, C., Ye, Z., and Bi, H. (2020). Long short-term memory network-based emission models for conventional and new energy buses. *International Journal of Sustainable Transportation*, 0(0):1–10.
- [Sutskever et al., 2013] Sutskever, I., Martens, J., Dahl, G., and Hinton, G. (2013). On the importance of initialization and momentum in deep learning. In Dasgupta, S. and McAllester, D., editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1139–1147, Atlanta, Georgia, USA. PMLR.
- [Teledyne Instruments, 2015] Teledyne Instruments (2015). Model 200e chemiluminescence no/no₂/no_x analyser. <https://pdf.directindustry.com/pdf/teledyne-api/analyzers-nitrogen-compounds-200e/22283-371488.html>. Accessed: 2021-01-20.
- [Thiel et al., 2014] Thiel, C., Schmidt, J., Zyl, A. V., and Schmid, E. (2014). Cost and well-to-wheel implications of the vehicle fleet co2 emission regulation in the european union. *Transportation Research Part A: Policy and Practice*, 63:25 – 42.
- [Thompson et al., 2000] Thompson, G., Atkinson, C., Clark, N., Long, T., and Hanzevack, E. (2000). Neural network modelling of the emissions and performance of a heavy-duty diesel engine. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 214(2):111–126.
- [Traver et al., 1999] Traver, M. L., Atkinson, R. J., and Atkinson, C. M. (1999). Neural network-based diesel engine emissions prediction using in-cylinder combustion pressure. *SAE Transactions*, 108:1166–1180.
- [Triantafyllopoulos et al., 2019] Triantafyllopoulos, G., Dimaratos, A., Ntziachristos, L., Bernard, Y., Dornoff, J., and Samaras, Z. (2019). A study on the co2 and nox emissions performance of euro 6 diesel vehicles under various chassis dynamometer and on-road conditions including latest regulatory provisions. *Science of The Total Environment*, 666:337 – 346.

- [Tsanakas et al., 2017] Tsanakas, N., Ekström, J., and Olstam, J. (2017). Reduction of errors when estimating emissions based on static traffic model outputs. *Transportation Research Procedia*, 22:440 – 449.
- [Tsiakmakis et al., 2017] Tsiakmakis, S., Fontaras, G., Cubito, C., Pavlovic, J., Anagnostopoulos, K., and Ciuffo, B. (2017). From nedc to wltp: effect on the type-approval co2 emissions of light-duty vehicles. *Publications Office of the European Union: Luxembourg*, page 50.
- [Tsokolis et al., 2016] Tsokolis, D., Tsiakmakis, S., Dimaratos, A., Fontaras, G., Pistikopoulos, P., Ciuffo, B., and Samaras, Z. (2016). Fuel consumption and co2 emissions of passenger cars over the new worldwide harmonized test protocol. *Applied Energy*, 179:1152 – 1165.
- [Turns, 1996] Turns, S. R. (1996). *An introduction to combustion*, volume 499. McGraw-hill New York.
- [Unal et al., 2003] Unal, A., Rouphail, N. M., and Frey, H. C. (2003). Effect of arterial signalization and level of service on measured vehicle emissions. *Transportation Research Record*, 1842(1):47–56.
- [United Nations, 2019] United Nations (2019). World Urbanization Perspective - The 2018 Revision. <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>. Accessed: 2020-11-17.
- [United States Environmental Protection Agency, 2016a] United States Environmental Protection Agency (2016a). Description and history of the mobile highway vehicle emission factor model: What was mobile? <https://www.epa.gov/moves/description-and-history-mobile-highway-vehicle-emission-factor-model>. Accessed: 2020-03-27.
- [United States Environmental Protection Agency, 2016b] United States Environmental Protection Agency (2016b). Final rule for control of air pollution from motor vehicles: Tier 3 motor vehicle emission and fuel standards. <https://www.epa.gov/regulations-emissions-vehicles-and-engines/>

final-rule-control-air-pollution-motor-vehicles-tier-3. Accessed: 2020-01-20.

[United States Environmental Protection Agency, 2016c] United States Environmental Protection Agency (2016c). Vehicle weight classifications for the emission standards reference guide. <https://www.epa.gov/emission-standards-reference-guide/vehicle-weight-classifications-emission-standards-reference-guide>. Accessed: 2020-01-20.

[United States Environmental Protection Agency, 2019] United States Environmental Protection Agency (2019). Can i use moves to model emissions from vehicles in other countries? <https://www.epa.gov/moves/can-i-use-moves-model-emissions-vehicles-other-countries>. Accessed: 2020-04-01.

[US EPA, 2005] US EPA (2005). Exhaust Emission Rates for Light-Duty On-road Vehicles in MOVES2014 - Final Report. Technical report, United States Environmental Protection Agency.

[Vallamsundar and Lin, 2011] Vallamsundar, S. and Lin, J. (2011). Overview of us epa new generation emission model: Moves. *International Journal on Transportation and Urban Development*, 1(1):39.

[Varella et al., 2017] Varella, R. A., Duarte, G., Baptista, P., Villafuerte, P. M., and Sousa, L. (2017). Analysis of the influence of outdoor temperature in vehicle cold-start operation following eu real driving emission test procedure. *SAE International Journal of Commercial Vehicles*, 10(2017-24-0140):596–697.

[Varella et al., 2019] Varella, R. A., Faria, M. V., Mendoza-Villafuerte, P., Baptista, P. C., Sousa, L., and Duarte, G. O. (2019). Assessing the influence of boundary conditions, driving behavior and data analysis methods on real driving co2 and nox emissions. *Science of The Total Environment*, 658:879 – 894.

- [Velleman, 1980] Velleman, P. F. (1980). Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association*, 75(371):609–615.
- [Velmurugan, 2012] Velmurugan, T. (2012). Efficiency of k-means and k-medoids algorithms for clustering arbitrary data points. *Int. J. Computer Technology & Applications*, 3(5):1758–1764.
- [Vineis et al., 2006] Vineis, P., Gerard, H., Michal, K., Federica, V.-T., Fabrizio, V., Luisa, A., Herman, A., Alison, D., Seymour, G., Pierre, H., Christian, M., Giuseppe, M., Kim, O., Ole, R.-N., Francoise, C.-C., Jacob, L., Heiner, B., Antonia, T., Domenico, P., Marco, P., Vittorio, K., Rosario, T., Salvatore, P., Bas, B.-D.-M. H., H., P. P., Eylin, L. E., A., G. C., Carmen, M., Miren, D., Aurelio, B., Lluís, C., Ramon, Q. J., Goran, B., Bertil, F., E., D. N., J., K. T., Rodolfo, S., Rudolf, K., and Elio, R. (2006). Air pollution and risk of lung cancer in a prospective study in Europe. *International Journal of Cancer*, 119(1):169–174.
- [Virtanen et al., 2020] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*.
- [Wang et al., 2014] Wang, C., Xu, H., Herreros, J. M., Wang, J., and Cracknell, R. (2014). Impact of fuel and injection system on particle emissions from a gdi engine. *Applied Energy*, 132:178 – 191.
- [Wei et al., 2018] Wei, H., Zheng, G., Yao, H., and Li, Z. (2018). Intellilight: A reinforcement learning approach for intelligent traffic light control. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2496–2505.

- [Weilenmann et al., 2009] Weilenmann, M., Favez, J.-Y., and Alvarez, R. (2009). Cold-start emissions of modern passenger cars at different low ambient temperatures and their evolution over vehicle legislation categories. *Atmospheric Environment*, 43(15):2419 – 2429.
- [Weilenmann et al., 2005] Weilenmann, M., Soltic, P., Saxer, C., Forss, A.-M., and Heeb, N. (2005). Regulated and nonregulated diesel and gasoline cold start emissions at different temperatures. *Atmospheric Environment*, 39(13):2433 – 2441. 12th International Symposium, Transport and Air Pollution.
- [Weiss et al., 2011] Weiss, M., Bonnel, P., Hummel, R., Provenza, A., and Manfredi, U. (2011). On-road emissions of light-duty vehicles in europe. *Environmental Science & Technology*, 45(19):8575–8581. PMID: 21815612.
- [Weiss et al., 2012] Weiss, M., Bonnel, P., Kühlwein, J., Provenza, A., Lambrecht, U., Alessandrini, S., Carriero, M., Colombo, R., Forni, F., Lanappe, G., Le Lijour, P., Manfredi, U., Montigny, F., and Sculati, M. (2012). Will euro 6 reduce the nox emissions of new diesel cars? – insights from on-road tests with portable emissions measurement systems (pems). *Atmospheric Environment*, 62:657 – 665.
- [Williams and Carslaw, 2011] Williams, M. L. and Carslaw, D. C. (2011). New directions: Science and policy – out of step on nox and no2? *Atmospheric Environment*, 45(23):3911 – 3912.
- [World Health Organisation, 2016] World Health Organisation (2016). Ambient air pollution: A global assessment of exposure and burden of disease. Technical report, World Health Organisation.
- [Wurzel and Wurzel, 2002] Wurzel, R. K. and Wurzel, R. (2002). *Environmental policy-making in Britain, Germany and the European Union: The Europeanisation of air and water pollution control*. Manchester University Press.
- [Wyatt, 2017] Wyatt, D. W. (2017). *Assessing micro-scale carbon dioxide (CO₂) emission on UK road networks using a coupled traffic simulation and vehicle emission model*. PhD thesis, University of Leeds.

- [Wyche et al., 2021] Wyche, K., Nichols, M., Parfitt, H., Beckett, P., Gregg, D., Smallbone, K., and Monks, P. (2021). Changes in ambient air quality and atmospheric composition and reactivity in the south east of the uk as a result of the covid-19 lockdown. *Science of The Total Environment*, 755:142526.
- [Xie et al., 2003] Xie, S., Zhang, Y., Qi, L., and Tang, X. (2003). Spatial distribution of traffic-related pollutant concentrations in street canyons. *Atmospheric Environment*, 37(23):3213 – 3224.
- [Xu et al., 2015] Xu, B., Wang, N., Chen, T., and Li, M. (2015). Empirical Evaluation of Rectified Activations in Convolutional Network. *CoRR*, abs/1505.00853.
- [Zeiler, 2012] Zeiler, M. D. (2012). ADADELTA: An Adaptive Learning Rate Method. *CoRR*, abs/1212.5701.
- [Zeiler et al., 2013] Zeiler, M. D., Ranzato, M., Monga, R., Mao, M., Yang, K., Le, Q. V., Nguyen, P., Senior, A., Vanhoucke, V., Dean, J., and Hinton, G. E. (2013). On rectified linear units for speech processing. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3517–3521.
- [Zhai et al., 2008] Zhai, H., Frey, H. C., and Rouphail, N. M. (2008). A vehicle-specific power approach to speed- and facility-specific emissions estimates for diesel transit buses. *Environmental Science & Technology*, 42(21):7985–7991. PMID: 19031891.
- [Zhang et al., 2013] Zhang, L., Yin, Y., and Chen, S. (2013). Robust signal timing optimization with environmental concerns. *Transportation Research Part C: Emerging Technologies*, 29:55 – 71.
- [Zhang et al., 2016] Zhang, Q., Wu, L., Yang, Z., Zou, C., Liu, X., Zhang, K., and Mao, H. (2016). Characteristics of gaseous and particulate pollutants exhaust from logistics transportation vehicle on real-world conditions. *Transportation Research Part D: Transport and Environment*, 43:40 – 48.

- [Zhang et al., 2014] Zhang, S., Wu, Y., Liu, H., Huang, R., Un, P., Zhou, Y., Fu, L., and Hao, J. (2014). Real-world fuel consumption and co₂ (carbon dioxide) emissions by driving conditions for light-duty passenger vehicles in china. *Energy*, 69:247 – 257.
- [Zhang et al., 1998] Zhang, Z., Lyons, M., Schuster, M., and Akamatsu, S. (1998). Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. In *Third IEEE International Conference on Automatic Face and Gesture Recognition*.
- [Zhou et al., 2010] Zhou, Y., Wu, Y., Yang, L., Fu, L., He, K., Wang, S., Hao, J., Chen, J., and Li, C. (2010). The impact of transportation control measures on emission reductions during the 2008 olympic games in beijing, china. *Atmospheric Environment*, 44(3):285 – 293.