

**SPATIO-TEMPORAL MODELLING
OF URBAN SENSOR NETWORK
DATA:
MAPPING AIR QUALITY RISKS IN
EINDHOVEN, THE NETHERLANDS**

Veronella Maria van Zoest

**SPATIO-TEMPORAL MODELLING OF URBAN
SENSOR NETWORK DATA:
MAPPING AIR QUALITY RISKS IN EINDHOVEN,
THE NETHERLANDS**

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. T. T. M. Palstra,
on account of the decision of the Doctorate Board,
to be publicly defended
on Wednesday, January 15, 2020 at 14.45

by

Veronella Maria van Zoest
born on August 5, 1992
in Voorburg, the Netherlands

This dissertation is approved by:

prof. dr. ir. A. Stein (supervisor)
dr. F.B. Osei (co-supervisor)
dr. ir. G. Hoek (co-supervisor)

ITC dissertation number 372
ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN: 978-90-365-4929-5
DOI: <http://dx.doi.org/10.3990/1.9789036549295>
Printed by: ITC Printing Department, Enschede, The Netherlands

© Veronella Maria van Zoest, Enschede, The Netherlands
All rights reserved. No part of this publication may be reproduced without the
prior written permission of the author.



Graduation committee**Chair**

prof. dr. ir. A. Veldkamp University of Twente

Supervisor

prof. dr. ir. A. Stein University of Twente

Co-supervisors

dr. F.B. Osei University of Twente

dr. ir. G. Hoek Utrecht University

Members

prof. dr. R.V. Sliuzas University of Twente

prof. dr. R. Zurita Milla University of Twente

prof. dr. D.P. Lee University of Glasgow

prof. dr. ir. A.K. Bregt Wageningen University

This research was conducted under the auspices of the Graduate School
for Socio-Economic and Natural Sciences of the Environment (SENSE).

Summary

Low-cost urban air quality sensor networks are increasingly used to study the spatio-temporal variability in air pollutant concentrations. In Eindhoven, the Netherlands, a low-cost air quality sensor network was set up in 2013 as part of the civil initiative AiREAS. The aim of this thesis is to evaluate the data quality of the collected data and its usability in spatio-temporal modelling and health effect estimation.

The first research objective addresses outliers. Those could reflect measurement errors or unusually high or low air pollution events. In this first chapter I present a novel outlier detection method based upon a spatio-temporal classification. The focus is on hourly nitrogen dioxide (NO_2) concentrations, as NO_2 has a large spatio-temporal variability and strong association with health effects. Different spatio-temporal classes are defined, reflecting urban background vs. urban traffic stations, weekdays vs. weekends and four periods per day. Truncated normal distributions are used to set thresholds for the definitions of outliers in each spatio-temporal class. Based on this study, I conclude that this method is able to detect outliers while maintaining the spatio-temporal variability of air pollutant concentrations in urban areas.

The second research objective addresses the calibration of low-cost sensor networks. Field calibration is typically performed at one location, while little is known about the spatial transferability of correction factors. This chapter evaluates three calibration methods: (1) an iterative Bayesian approach for daily estimation of the parameters in a multiple linear regression model, (2) a daily updated correction factor and (3) a correction factor updated only when concentrations are uniformly low. Performance of the calibration methods is compared in terms of temporal stability, spatial transferability, and sensor specificity. A poor spatial transferability of the calibration parameters was found for all methods. This is consistent with different responses of individual sensors to environmental factors such as temperature and relative humidity. Due to their spatial and temporal variability, calibration parameters require regular updates and sensor-specific recalibrations.

Summary

The third research objective addresses prediction of air pollutant concentrations at unobserved locations. Spatio-temporal regression kriging was applied to map NO₂ at a 25 m spatial resolution and hourly temporal resolution. The trend is modelled separately from autocorrelation in the residuals. The trend part of the model consists of a set of spatial and temporal covariates including population density, road type and meteorological variables. Spatio-temporal autocorrelation in the residuals is modelled by fitting a sum-metric spatio-temporal variogram model. The method provides local estimates of the strength and association of air pollution sources and sinks, and allows for near real-time prediction of air pollutant concentrations. The resulting maps visualize these in space and time and can be used to assess exposure for the evaluation of short-term health effects.

The fourth research objective addresses health effect estimates related to air pollution, focusing on daily respiratory symptoms in children with asthma. Bayesian estimates of the exposure-response function were obtained by updating *a priori* information from a meta-analysis with data from a panel study. Positive associations between NO₂ and lower respiratory symptoms and medication use were observed. Credible intervals substantially narrowed when adding prior information from the meta-analysis. Burden of disease maps showed a strong spatial variability in the number of asthmatic symptoms associated with ambient NO₂. Bayesian methods provided accurate local air pollution effect estimates and subsequent local burden of disease calculations.

To summarize, this thesis evaluates the use of low-cost air quality sensor network data from data collection to application. After careful evaluation of the data quality and removal of outliers, it shows that the data can be used to map air pollutant concentrations at a fine spatial and temporal resolution. These maps can be used to estimate burden of disease at the within-city level. Future research may address a wide range of applications, including sensor network development, policy making, and further health risk assessment.

Samenvatting

Relatief goedkope stedelijke netwerken voor het meten van luchtkwaliteit worden steeds vaker gebruikt om de ruimtelijk-temporele variatie in concentraties luchtvervuilende stoffen te bestuderen. In Eindhoven is in 2013 een dergelijk netwerk opgezet geïnitieerd door het burgerinitiatief AiREAS. Het doel van dit proefschrift is het evalueren van de kwaliteit van de sensordata die met dit netwerk verzameld zijn alsmede de bruikbaarheid van deze gegevens, zowel voor het ontwikkelen van ruimtelijk-temporele modellen, als voor het maken van schattingen voor effecten van luchtkwaliteit op de gezondheid.

Het eerste onderzoeksdoel behandelt uitschieters. Deze kunnen worden veroorzaakt door meetfouten of door gebeurtenissen die leiden tot ongebruikelijk hoge of lage concentraties van luchtvervuilende stoffen. In dit eerste hoofdstuk presenteert ik een nieuwe detectiemethode voor uitschieters. Deze is gebaseerd op een ruimtelijk-temporele classificatie. Ik heb ervoor gekozen om de methode toe te passen op stikstofdioxide (NO_2) concentraties, omdat die een hoge ruimtelijk-temporele variatie laten zien en een sterke associatie hebben met gezondheidseffecten. Ruimtelijk-temporele klassen zijn gedefinieerd als reflectie van de dagelijkse variatie in verkeersdrukte, weekdagen vs. weekenddagen, en achtergrond vs. verkeerslocaties in de stad. De grenswaarden voor de definities van uitschieters in iedere klasse zijn gebaseerd op de afgeknotte normale verdeling. De studie laat zien dat de methode in staat is om uitschieters te detecteren met behoud van de ruimtelijk-temporele variatie in luchtvervuilende stoffen in een stedelijk gebied.

Het tweede onderzoeksdoel behandelt de kalibratie van relatief goedkope netwerken voor het meten van luchtkwaliteit. Kalibratie wordt vaak op één locatie gedaan, terwijl weinig bekend is over de ruimtelijke verplaatsbaarheid van de correctiefactoren. Dit hoofdstuk evalueert drie kalibratiemethoden: (1) een iteratieve Bayesiaanse benadering voor de dagelijkse inschatting van de parameters in een multivariaat lineair regressiemodel, (2) een dagelijks aangepaste correctiefactor en (3) een correctiefactor die slechts aangepast wordt wanneer de concentraties uniform en laag zijn. De kalibratiemethoden zijn met elkaar vergeleken op basis van temporele stabiliteit, ruimtelijke verplaatsbaarheid, en spe-

Samenvatting

cificiteit van de sensoren. Alle methoden laten een beperkte ruimtelijke verplaatsbaarheid zien, die in overeenstemming is met de verschillende gevoeligheden van de individuele sensoren voor omgevingsfactoren zoals temperatuur en relatieve luchtvochtigheid. Op basis van de ruimtelijke en temporele variabiliteit in de kalibratieparameters adviseert dit hoofdstuk om reguliere aanpassingen en sensor-specifieke herkalibraties toe te passen.

Het derde onderzoeksdoel behandelt de voorspellingen van concentraties luchtvervuilende stoffen op locaties waar geen metingen zijn gedaan. Om NO₂ in kaart te brengen met een ruimtelijke resolutie van 25 m en een temporele resolutie van 1 uur, is ruimtelijk-temporele regressiekriging toegepast. Deze methode modelleert de trend apart van de autocorrelatie in de residuen. Het trenddeel bestaat uit ruimtelijke en temporele variabelen, zoals bevolkingsdichtheid, type van de weg en meteorologische variabelen. Met behulp van een ruimtelijk-temporeel variogram is de autocorrelatie in de residuen gemodelleerd. De methode verbetert lokale schattingen van de sterkte en associatie van factoren die van invloed zijn op de luchtvervuiling, en maakt *near real-time* voorspellingen van luchtvervuilende stoffen mogelijk. De resulterende kaarten kunnen worden gebruikt bij de schatting van korte termijn gezondheidseffecten.

Het vierde onderzoeksdoel behandelt het schatten van gezondheidseffecten ten gevolge van luchtvervuiling, met een focus op de dagelijkse variatie in luchtwegsymptomen bij kinderen met astma. Bayesiaanse schattingen van de blootstelling-responsfunctie zijn verkregen door *a priori* informatie van een meta-analyse te verrijken met gegevens uit een panelstudie. De resultaten suggereren positieve associaties tussen NO₂ en lagere luchtwegklachten en medicijngebruik. De betrouwbaarheidsintervallen zijn sterk verkleind door het gebruik van *a priori* informatie uit de meta-analyse. Kaarten van de gezondheidsbelasting tonen een sterke ruimtelijke variabiliteit in het aantal astmasymptomen gerelateerd aan NO₂ in de buitenlucht. Bayesiaanse methoden geven accurate schattingen van de lokale luchtvervuilingseffecten en daarmee nauwkeuriger berekeningen van de gezondheidslast.

Samenvattend evaluateert dit proefschrift het gebruik van gegevens die met relatief goedkope netwerken voor het meten van luchtkwaliteit zijn verkregen. Na een zorgvuldige evaluatie van de gegevenskwaliteit en het verwijderen van uitschieters, laat het proefschrift zien dat de gegevens gebruikt kunnen worden om met een hoge ruimtelijk-temporele resolutie de concentraties luchtvervuilende stoffen in kaart te brengen.

Acknowledgements

You will have to stand on the shoulders of giants to become a giant yourself. In the past four years I grew a lot to become an independent researcher. This would not have been possible without the people around me. My special thanks go to professor Alfred Stein, who has always supported me through the mountainous terrain of my PhD. You pushed me forward and gave me confidence during the hard times in the valleys, but also pulled me back when I ran too fast up the hill. I cannot wish for a promotor more skilled and devoted to his PhD students. When I had to spend a year without a daily supervisor, you showed your flexibility to take this role as well, efficiently and always there when needed.

Many thanks go to Frank Osei, who took over the role of daily supervisor during the last year of my PhD. You encouraged me to work quickly and efficiently and boosted my self-confidence to climb the highest hills. We never planned a meeting, because your door was always open. Many thanks also go to Gerard Hoek, who always seemed to remember exactly what I was working on although our meetings were less frequent. You taught me many lessons in environmental epidemiology, but also in clear and transparent writing. Thanks for hosting me when I was working on the last research objective.

I thank NWO for funding this research project and putting emphasis on the usability of the results. This encouraged me to focus on the practical applicability of the research outcomes and on communication of the results with a wider audience than just the scientific community. My gratitude goes to the members of the user group, who have been actively participating in user group meetings and always accessible when I needed more information or data.

My stay at ITC has been very pleasant. For this I thank my colleagues and friends, with whom I shared many valuable experiences. I enjoyed the scientific and non-scientific discussions a lot and thank you for the memories. Because of you I have made friends all over the world, who will make me feel at home wherever I travel. Thanks especially to my officemates for creating a pleasant climate to work in. Peaceful and quiet

Acknowledgements

while working, but always there to discuss problems or to have a friendly chat. The supporting staff of ITC has also contributed to a pleasant working environment. Besides the many people that assisted me in one way or the other with administrative matters or technical issues, I would especially like to mention Roelof Schoppers for the welcoming smile every day, and Teresa Brefeld and Marga Koelen for all their support. My thanks also go to Marieke Oldenwening at Utrecht University for her support during the panel study.

Doing a PhD keeps your mind working for much longer than eight hours a day. I thank my friends and family for their understanding. I especially thank my parents for their support, contributions and valuable discussions from a local perspective. My sincerest gratitude goes to Hugo for being patient and supporting, and following me wherever my path leads me. You kept loving me in stressful times and always encouraged me to get the best out of myself.

Who knows, one day I might be a giant myself. Thanks to the support of the giants around me, I am confident that I will keep growing for many years to come.

Contents

Summary	i
Samenvatting	iii
Acknowledgements	v
Contents	vii
List of Figures	ix
List of Tables	xi
List of Symbols	xiii
1 Introduction	1
1.1 Motivation	1
1.2 Spatio-temporal big data analysis	2
1.3 Spatial data quality	3
1.4 Modelling and mapping	4
1.5 Uncertainty	6
1.6 Air pollution: sources and sinks	6
1.7 Health effects	7
1.8 Limit values and guidelines	8
1.9 Problem statement	9
1.10 Research objectives	10
1.11 Outline	10
1.12 Author contributions	11
2 Case study area	13
2.1 Study area: the city of Eindhoven	13
2.2 The ILM air quality network	13
2.3 Spatial sampling scheme	14
2.4 Airbox sensors	15
2.5 Reference measurements	16
3 Outlier detection in urban air quality sensor networks	19
3.1 Introduction	20

Contents

3.2 Data preprocessing	22
3.3 Methods	22
3.4 Results	27
3.5 Discussion	32
3.6 Conclusions	36
4 Calibration of NO₂ sensors in an urban air quality network	37
4.1 Introduction	38
4.2 Methods	39
4.3 Results	47
4.4 Discussion and conclusions	55
5 Spatio-temporal regression kriging for modelling urban NO₂ concentrations	65
5.1 Introduction	66
5.2 Methods	67
5.3 Application	69
5.4 Results and discussion	71
5.5 Conclusions	78
6 Short-term impact of NO₂ exposure on local burden of asthmatic symptoms	81
6.1 Introduction	83
6.2 Methods	84
6.3 Results	91
6.4 Discussion	97
6.5 Conclusions	98
7 Synthesis	103
7.1 Main findings	103
7.2 Significance	104
7.3 Limitations	106
7.4 Prospects	108
Bibliography	111

List of Figures

2.1	Location of Eindhoven within the Netherlands	14
2.2	Locations of the airboxes in the city of Eindhoven, the Netherlands	15
2.3	Airbox attached to light pole	16
3.1	Locations of the airboxes measuring NO ₂ at urban background locations and urban traffic locations	24
3.2	Distribution of NO ₂ concentrations, before square root transformation and after square root transformation	25
3.3	The truncated normal distribution of square root transformed NO ₂ concentrations and its underlying normal distribution . .	25
3.4	Boxplots of the outliers in each spatio-temporal class	29
3.5	NO ₂ concentrations measured at an urban background location.	30
3.6	NO ₂ concentrations measured at an urban traffic location. . .	30
3.7	NO ₂ concentrations measured by a conventional monitor at an urban traffic location.	32
3.8	Scatterplot of traffic airbox outliers vs. the maximum NO ₂ concentration measured by the two conventional monitors. .	32
3.9	Comparison of NO ₂ concentrations measured by an airbox and a conventional monitor on a weekday at urban traffic locations	35
4.1	Locations of the airboxes and conventional monitors	40
4.2	Scatterplots of hourly NO ₂ values in 2016: airbox vs. conventional monitor	47
4.3	Difference between mean airboxes and mean conventional monitors over time and fitted smooth curves.	49
4.4	Time series of the coefficients of the daily INLA models, using model 8.	51
4.5	Posterior mean estimates of different airboxes, for airbox NO ₂ vs. covariates 'relative humidity' and 'temperature'	54
4.6	Posterior distributions of slopes for reference monitor NO ₂ per covariate	55
A4.1	Histogram and uniform(0,1) Q-Q plot of the PIT values on June 15, 2016 at location z_1	59

List of Figures

A4.2 Time series of the relative correction factors $\gamma_{rel,d,z}$ and absolute correction factors $\gamma_{abs,d,z}$	60
A4.3 Time series of the correction factor γ_{uni}	60
A4.4 Time series of the correction factor $\gamma_{night,d,s}$	61
A4.5 Scatterplots before calibration, after calibration without random effects, and after calibration with random effects	62
A4.6 Residual plots for INLA models without random effects and with random effects	63
5.1 Locations of the airboxes used for modelling NO ₂ in November 2016	70
5.2 Spatio-temporal sample variogram and sum-metric fitted variogram model	74
5.3 Prediction maps of NO ₂ concentrations at four time stamps on Monday November 7, 2016	75
5.4 Prediction maps of NO ₂ concentrations at four Sundays in November 2016, between 5 and 6 p.m.	76
5.5 Kriging variance map	77
6.1 Flowchart of daily asthma symptom calculation	84
6.2 Results of the meta-analysis on NO ₂ and lower respiratory symptoms (LRS)	92
6.3 Results of the meta-analysis on NO ₂ and cough	93
6.4 (a) Number of children per neighborhood. (b) Mean NO ₂ exposure in 2016. (c) NO ₂ attributable cases of LRS per day	95
6.5 Posterior densities of (a) odds ratio of lower respiratory symptoms (LRS), and (b) the attributable cases of LRS in neighborhood 't Hofke.	96
7.1 Framework of this thesis	105

List of Tables

1.1 WHO guidelines in comparison with EU directives and Dutch national law on the concentration limit for different pollutants	8
2.1 Variables measured and instruments used in the airboxes	17
3.1 Upper thresholds for hourly average NO ₂ concentrations above which considered outliers, per spatio-temporal class	28
3.2 Percentage outliers per spatio-temporal NO ₂ concentration class for hourly values in 2016	28
A3.1 Mean (\pm standard deviation) of the distribution underlying the truncated normal distribution of each spatio-temporal class . .	36
4.1 Overview of potential covariates for the calibration model	41
4.2 DIC performance statistics for different models	50
4.3 RMSE before and after temporal and spatiotemporal calibration using different models	52
4.4 RMSE values before and after applying a daily correction factor	52
4.5 RMSE values before and after updating the correction factor y_{uni}	53
4.6 RMSE before and after night-time calibration	53
A4.1 RMSE before and after temporal calibration, for different lengths of the calibration dataset	59
A4.2 RMSE for the models with random effects vs. without random effects	61
5.1 $\hat{\beta}$ and p-values for the fixed effects part of the regression model	73
5.2 Spatio-temporal variogram parameter estimates for the fitted sum-metric variogram.	74
6.1 Uninformative priors for Bayesian estimation of the parameters in the model	88
6.2 Frequency of reported daily symptoms in the panel study	91
6.3 Association between NO ₂ and daily symptoms, expressed as odds ratios (95% C.I.) based on panel study without informative prior information.	94

List of Tables

6.4 Comparison between prior OR based on meta-analysis, local OR based on uninformative prior, and OR based on informative prior	94
A6.1 Odds ratios (95% C.I.) related to a $10 \mu\text{g m}^{-3}$ increase in NO ₂ ambient air pollution, based on REML estimation of the model parameters	99
A6.2 Odds ratios (95% C.I.) related to a $10 \mu\text{g m}^{-3}$ increase in PM ₁ ambient air pollution, based on REML estimation of the model parameters	99
A6.3 Odds ratios (95% C.I.) related to a $10 \mu\text{g m}^{-3}$ increase in PM _{2.5} ambient air pollution, based on REML estimation of the model parameters	100
A6.4 Odds ratios (95% C.I.) related to a $10 \mu\text{g m}^{-3}$ increase in PM ₁₀ ambient air pollution, based on REML estimation of the model parameters	100
A6.5 Odds ratios (95% C.I.) related to a 10000 particle # increase in UFP ambient air pollution, based on REML estimation of the model parameters	100
A6.6 Comparison of June, November, mean June and November, and annual mean concentrations measured at the two RIVM reference monitors in Eindhoven, 2016	100
A6.7 Association between NO ₂ and daily symptoms, expressed as odds ratios (95% C.I.) based on panel study and prior information from the meta-analysis on LRS and cough	101

List of Symbols

$\#$	Number
α	Significance level
β_0	Intercept
$\hat{\beta}_0$	Estimated intercept
$\beta_{0,d,z}$	Intercept at d, z
$\hat{\beta}_{0,d,z}$	Estimated intercept at d, z
$\beta_{c,d,z}$	Regression coefficient of covariate c at d, z
$\beta_{NO2,y}$	NO_2 regression coefficient for symptom y
$\hat{\beta}_{c,d,z}$	Estimated regression coefficient of covariate c at d, z
γ	Semivariance
Γ	Semivariance matrix of all possible combinations of space-time observations
γ_0	Vector of semivariances between observation locations and prediction location
$\gamma_{abs,d,z}$	Absolute correction factor at d, z
$\gamma_{night,d,s}$	Nightly established correction factor at d, s
$\gamma_{rel,d,z}$	Relative correction factor at d, z
γ_{uni}	Correction factor at uniformly low concentrations
γ_v	Individual intercept per participant v
δ	Threshold for standard deviation
$\Delta_{d,t}$	NO_2 difference between mean of airboxes and mean of conventional monitors, at day d and hour t
$\Delta x_{q=0}^{NO2}$	NO_2 exposure in q above baseline of zero
$\epsilon_{d,t,z}$	Error at d, t, z , $\epsilon_{d,t,z} \sim N(0, \sigma^2)$
ζ	Confidence interval size indicator
η	Residual
$\hat{\eta}$	Predicted residual
$\bar{\eta}$	Vector of observed space-time residuals
θ_1	Set of parameters $\{m_K^{-i}, s_K^{-i}, a, b\}$
θ_2	Set of parameters $\{n_K^{-i}, t_K^{-i}\}$
$\theta_{c,d,z}$	Set of parameters $\{\mu_{\beta_{c,d,z}}, \tau_{\beta_{c,d,z}}\}$
κ	Space-time anisotropy parameter
λ_0	Vector of kriging weights
μ	Mean / trend
$\hat{\mu}$	Predicted mean / trend

List of Symbols

$\mu_{\beta_{c,d,z}}$	Mean of the posterior distribution of $\beta_{c,d,z}$
σ	Standard deviation
σ^2	Variance
σ_0^2	Kriging variance
σ_y	Standard deviation of γ_v
σ_c	Standard deviation of covariate c
σ_{joint}^2	Joint partial sill
σ_s^2	Spatial partial sill
σ_t^2	Temporal partial sill
σ_y	Standard deviation of $\beta_{NO2,y}$
$\tau_{\beta_{c,d,z}}^2$	Precision of the posterior distribution of $\beta_{c,d,z}$
τ_{joint}^2	Joint nugget
τ_s^2	Spatial nugget
τ_t^2	Temporal nugget
$\phi(\cdot)$	Probability density function
$\Phi(\cdot)$	Cumulative distribution function
ϕ_{joint}	Joint range
ϕ_s	Spatial range
ϕ_t	Temporal range
χ	Threshold for mean
ψ	Fixed intercept
a	Lower truncation limit
AC_{city}	Total number of attributable cases in the city
AC_q	Number of attributable cases in q
AF_q	Attributable fraction in q
ap	Air pollutant in set $\{NO2, PM10, PM2.5, PM1, UFP\}$
b	Upper truncation limit
c	Covariate
C	Set of covariates
d	Day
\bar{D}	Posterior mean of deviance
d_{+1}	Day after day d
E	Approximation
\mathbb{E}	Expected value
$east$	Easting coordinate
flu	Reported flu
$g_c(\cdot)$	Covariate transformation function
h	Spatial separation distance
i	Index
j	Index
k	Observation index in $\{1 \dots N_K\}$
K	Spatio-temporal class
$L(\cdot)$	Log likelihood function
m_K^{-i}	Mean of all observations in K except the i^{th}
n_K	Mean of underlying normal distribution of observations in K
n_K^{-i}	Mean of underlying normal distribution of observations

N	in K except the i^{th}
$N_{+\infty}$	Normal distribution
N_K	Half-normal distribution truncated at zero
N_K^{child}	Total number of observations in spatio-temporal class K
N_q^{child}	Number of children in q
N_q^{pop}	Number of inhabitants in q
N_s	Total number of airbox locations
N_t	Total number of timestamps
N_z	Total number of conventional monitor locations
$NO2$	Nitrogen dioxide
$\widehat{NO2}$	Predicted NO ₂ concentration
$NO2_{ab}$	Observed airbox NO ₂ concentration
$NO2_k$	NO ₂ observation with index k
$NO2_{ref}$	Observed reference NO ₂ concentration
$O3$	Ozone
OR	Odds ratio
p	P-value
p_{asthma}	Proportion of children with asthma
p_D	Effective number of parameters
P_q	Population at risk in q
$p_{v,d}$	Probability of occurrence of symptom $y_{v,d}$
$PM1$	Particulate matter <1 μm
$PM2.5$	Particulate matter <2.5 μm
$PM10$	Particulate matter <10 μm
pop	Population density
PR	Prevalence rate
q	Neighborhood index
Q	Environmental variable
r	Pearson's correlation coefficient
\hat{R}	Gelman-Rubin diagnostic
RH	Relative humidity
$road$	Road type
$RMSE_{post}$	RMSE after calibration
$RMSE_{pre}$	RMSE before calibration
RR_q	Relative risk in q
rw_1	First-order random walk
s	Spatial location / airbox location
S	Set of airbox locations $\{s_1 \dots s_{25}\}$
s_0	Spatial prediction location
s_K^{-i}	Standard deviation of all observations in K except the i^{th}
$sday$	Day of study participation
t	Hour / timestamp
T	Temperature
t_0	Temporal prediction location
T_d	Total number of non-missing hours in day d
t_K	Standard deviation of underlying normal distribution of observations in K

List of Symbols

t_K^{-i}	Standard deviation of underlying normal distribution of observations in K except the i^{th}
u	Temporal separation distance
UFP	Ultrafine particles
$Unif$	Uniform distribution
v	Participant index
WD	Wind direction
$wday$	Week/weekend day factor
WS	Wind speed
x_c	Value of covariate c
$x_{c,d,t,z}$	Value of covariate c at d, t, z
$\bar{x}_{NO2,d,t}$	Mean NO_2 concentration of all airboxes at d, t
$x_{NO2,d,t,s}$	NO_2 concentration value at d, t, s
$x_{NO2,d,t,z}$	NO_2 concentration value at d, t, z
x_k	Square root transformed $NO2_k$
$x_{k,i}$	i^{th} observation of x_k
\boldsymbol{y}	Vector of all square root transformed NO_2 observations
$\boldsymbol{y}_{-d,t}$	Vector of all square root transformed NO_2 observations except d, t
$y_{d,t}^*$	Replicate square root transformed NO_2 observations at d, t
$y_{d,t}$	Square root transformed NO_2 observation at d, t
$y_{d,t,z}$	Square root transformed reference NO_2 concentration at d, t, z
$\hat{y}_{d,t,z}$	Estimate of $y_{d,t,z}$
$y_{NO2,d,t,z}$	Reference NO_2 concentration at d, t, z
$\hat{y}_{NO2,d,t,z}$	Predicted and backtransformed NO_2 concentration at d, t, z
$y_{v,d}$	Symptom prevalence of participant v on day d
z	Location of conventional monitor
Z	Set of conventional monitor locations $\{z_1, z_2\}$
z_i	Location of conventional monitor i
z_j	Location of another conventional monitor j

Introduction

1

1.1 Motivation

Air pollution has major effects on human health (Cohen et al., 2017) and globally causes about 7 million premature deaths each year (World Health Organization (WHO), 2015). Whilst the air pollution levels increase in developing countries due to rising industry, developed countries are taking measures to reduce emissions. Even in developed countries, target limit values are not as strict as those suggested by the WHO to minimize health impacts (WHO, 2006). Meanwhile, the WHO states that the suggested guidelines cannot fully protect human health, as there are no lower limits known below which no health effects occur. A better quantification of health effects at lower air pollution levels is therefore required.

To quantify health risks related to ambient air pollution levels, a good estimate of personal exposure is required. Since personal exposure monitoring is expensive and time-consuming (Brandt et al., 2015), it is rarely used and limited to short study periods (e.g. Linn et al., 1996; Spira-Cohen et al., 2011). Typically, health risk assessments are based on central monitors of national ambient air quality monitoring networks (Roemer et al., 1993; Van der Zee et al., 1999, 2000; Dales et al., 2009). In Europe, these monitoring networks are operated by national environmental agencies and comply with high quality standards (EC Working Group on GDE, 2010), with the aim to evaluate exceedance of limit values determined by European guidelines (European Parliament and Council of the European Union, 2008).

Although national ambient air quality monitoring networks can be used to obtain high quality measurements, their spatial coverage is limited. Due to the high costs of the instruments, maintenance and calibration, typically only one or two monitors are located in each city. However, air pollution levels typically strongly vary within short distances. This spatial variation is strongest in urban areas where there is a strong variety of road types, traffic intensities and land uses (Hoek et al., 2008). To increase the spatial coverage of air pollution measurements within a

1. Introduction

city, low-cost urban air quality sensor networks have recently been set up by civil initiatives (Snyder et al., 2013).

1.2 Spatio-temporal big data analysis

Smart city sensor networks can capture multivariate data at multiple spatial locations and at a high temporal frequency (e.g. every 10 minutes). Over the span of multiple years, these data add up to big data sets. These data have a number of challenges when treating them statistically. Depending on the intended use, data may first have to be selected within a spatial and temporal window of interest. The resolution should fit the intended use and data quality, for which aggregation may be needed. To detect anomalies in the data that could point at errors or observations of interest, automatic filtering techniques may be useful. Next, one may be interested in the detection of spatio-temporal patterns and relations to build prediction models. To do so, data could be analyzed in a spatio-temporal statistical framework. Finally, the results of the spatio-temporal data analysis should be visualized to present and communicate results with users and stakeholders.

We consider the continuous multivariate spatio-temporal field $Q(s, t)$, where measurements of environmental variables Q are taken at any spatial location s and time stamp t (Caselton and Zidek, 1984; Sølna and Switzer, 1996). Here, a spatial location is a three-dimensional set of spatial coordinates. In practice, however, the vertical height of the measurement locations in sensor networks is typically kept constant and ignored in subsequent analyses. A spatio-temporal statistical framework allows to assess the data quality of a measurement at (s, t) , e.g. by comparing to reference measurements or the expected value at (s, t) given spatio-temporal patterns in the data. Next, it allows prediction of Q at any unobserved spatio-temporal location (s_0, t_0) based on spatio-temporal autocorrelation in Q and relations between Q and other variables measured at (s, t) (Cressie and Wikle, 2011; Sherman, 2011; Bivand et al., 2013).

The advantage of using low-cost sensors is that a relatively high number of sensors N_s can be used in a relatively small area. In this way the measurements are better able to reflect the spatial variability in Q , and thus more useful to model its spatial autocorrelation and predict its values at unobserved locations. There is, however, a trade-off between the relative costs and data quality (Snyder et al., 2013). Reis et al. (2015) state that the number of local sensor networks is small due to the expectation that all low-cost sensors need to function at the same quality level as the reference instruments used for legislative purposes. This leads to high costs for instrumentation, calibration and maintenance. When combining data of multiple sensors in a model, information content becomes more

important than data quality, as long as the data quality is known (Reis et al., 2015). Another advantage of a smart city sensor network is the ability for sensors to connect with each other – observations of one sensor can for example be used to calibrate other sensors.

In the city of Eindhoven, the Netherlands, a low-cost air quality sensor network has been set up by the AiREAS civil initiative (Close, 2016). This sensor network is used throughout this thesis. More information about this network can be found in Chapter 2.

1.3 Spatial data quality

As the development of low-cost sensor networks recently started, their data quality is still unknown (Snyder et al., 2013). It is important to know the spatial data quality of the sensor data, however, as it will influence the quality of the output models, maps and exposure estimates. In literature different lists are named of elements which could be included in spatial data quality assessment (Guptill and Morrison, 1995; van Oort, 2006). In order to assess spatial data quality in a transparent way, an international standard is needed. Such standard is provided by the International Organization for Standardization (ISO) in ISO 19157 (ISO/TC 211 Secretariat, 2013). Six elements of spatial data quality are defined: completeness, logical consistency, positional accuracy, thematic accuracy, temporal quality, and usability element (ISO/TC 211 Secretariat, 2013).

In terms of completeness, an important issue for air quality sensor networks is missing data. Several methods exist to impute missing data in air pollution time series (e.g. Basu and Meckesheimer, 2007; Nguyen and Hoogerbrugge, 2014; Harrell, 2018), which can deal with longer periods of missing data. Logical consistency deals with the validity of attribute values and the adherence of relationships and compositions between objects to logical rules of structure and compatibility (Kainz, 1995). Negative air pollution values should for example be removed, as they are impossible to occur. Positional accuracy defines the accuracy of positions of features and is always related to some kind of spatial reference system (ISO/TC 211 Secretariat, 2013). It deals with the nearness of the true values in comparison to the observed values in this reference system (Drummond, 1995). In sensor networks where the sensors all have fixed and known locations, this is less of an issue. The reported positions, however, can be used to assess whether a sensor is at its usual location or moved for maintenance or calibration. Thematic accuracy refers to classification correctness, non-quantitative attribute correctness, and quantitative attribute accuracy. Quantitative attribute accuracy refers to the closeness of the value of a quantitative attribute to the true value (ISO/TC 211 Secretariat, 2013). This ‘true value’ often

1. Introduction

refers to a value of a reference dataset which is accepted to be true. There are different calibration methods available to calibrate low-cost sensors to reference stations, to improve quantitative attribute accuracy in the concentration values (Spinelle et al., 2015). An important step in assessing the non-quantitative attribute correctness of the raw sensor observations is detecting outliers in the data. Outliers, sometimes referred to as anomalies (Chandola et al., 2009), are those observations that differ from the expected observations (Basu and Meckesheimer, 2007; Zhang et al., 2012). Erroneous outliers are those observations that deviate from the true values. These are different from events, which are observations that can be detected as outliers, but do not deviate from the true values (Zhang et al., 2012). Events rather reflect a real change in the measured phenomenon and can therefore be of interest depending on the user perspective. Temporal quality explains the quality of temporal attributes and relationships. It consists of accuracy of a time measurement, temporal consistency, and temporal validity. Van Oort (2006) adds three elements which are not in ISO 19157 but which were present in the European pre-standard ENV 12656 (CEN/TC 287, 1998): last update, rate of change, and temporal lapse. The temporal lapse represents the average time between change in the real world and change representations in the data, and is thus related with the temporal resolution used when averaging the air pollutant concentrations over a period of time (e.g. ten minutes, hourly, daily). Usability element refers to the suitability of the data for a specific application. All previously mentioned elements can be used to describe and assess the usability of the data (ISO/TC 211 Secretariat, 2013).

1.4 Modelling and mapping

As air quality can only be monitored at point locations while the true air quality changes over the continuous spatio-temporal field Q , modelling is required to map air pollutant concentrations at unobserved spatio-temporal locations (s_0, t_0) . Many models and their classifications exist. Dispersion models have been used for a long time as they are relatively easy to use. The Gaussian plume model, as an example, is typically used for modelling the air pollutant dispersion. Historically, the model was mainly used for point sources such as factory stacks (Weil et al., 1992), but it has been improved over time to be used for line sources and to be applicable even under calm and changing wind conditions (Shorshani et al., 2015). Empirical models are based on measurements which are typically interpolated to create air quality maps. Beelen et al. (2009) compared different methods to map the background air pollution in the European Union, including kriging and a land use regression model. Interpolation with kriging of point observations at locations s is based on a stochastic process that is split into a trend, a spatially dependent error term and spatially independent noise. In ordinary kriging the trend is constant but unknown; in simple kriging the trend is constant and

1.4. Modelling and mapping

known. Land use regression (LUR) is based on a regression equation between predictor variables and measured concentrations. Air pollutant concentration mapping with LUR predicts values at unsampled locations using measured concentrations at a number of locations. Those are combined with a stochastic model using predictor variables such as land use, altitude and meteorology (Hoek et al., 2008). Regression kriging combines estimation of the trend using a regression model with simple kriging on the error component, which is assumed to have a known mean of zero. Van de Kassteele et al. (2005) evaluated methods for predicting the annual number of days that the ozone (O_3) limit value is exceeded, using model-based spatial interpolation. In several other studies the methods for spatial interpolation of particulate matter (PM) in Europe are evaluated (Van de Kassteele et al., 2006; Hamm et al., 2015). Van de Kassteele and Stein (2006) developed a geostatistical model for mapping PM at the European scale, using error-in-variable external drift kriging (KED). In KED, secondary information is added to the statistical interpolation model (Van de Kassteele and Stein, 2006). Generally, the primary variable is expected to be the most precise while the number of locations is low, whereas the secondary variable can be less precise but is sampled more densely (Van de Kassteele et al., 2009). In epidemiological studies, urban scale maps are often used, accounting for short-distance spatial variability. Klompmaker et al. (2015) studied the spatial variation in ultrafine particles (UFPs) and black carbon (BC) in Amsterdam and Rotterdam, the Netherlands. LUR is often applied in urban air quality mapping, for example for mapping nitrogen dioxide (NO_2) (Sahsuvaroglu et al., 2006; Jerrett et al., 2007; Hoek et al., 2008; Basagaña et al., 2012), particulate matter (PM) (Hoek et al., 2008; Saraswat et al., 2013), and more recently also for mapping UFP (Saraswat et al., 2013; Montagne et al., 2015).

Studies of health effects of long term exposure typically take into account the spatial component only, ignoring the temporal variability or adjusting for it (Gulliver and Briggs, 2004; Gehring et al., 2010; Beelen et al., 2014). Other studies measure exposure in different transportation modes and combine it with findings from health studies to assess the health effects of using specific transportation modes, without addressing the spatial and temporal variation (Knibbs and de Dear, 2010; Knibbs et al., 2011). Sensor networks measuring at a high spatio-temporal resolution provide opportunities for assessing short-term health effects. Exposure can be estimated near the school or work address, instead of at one central location in the city. The best spatial and temporal resolution can be achieved by modelling towards the maximum scalable unit, being the maximum unit in space and time where the air pollutant concentrations are considered to be homogeneous.

1. Introduction

1.5 Uncertainty

When maps are used to assess health effects, it is important to communicate its related uncertainty. Incorrect or inaccurate dose-response characterizations may lead to overestimation or underestimation of health effects (Burns et al., 2014). The confidence and trust of the user in map products depends on user's awareness of the uncertainties that they bring along (Sacha et al., 2016).

The U.S. National Research Council (NRC) defines uncertainty as a lack of information, incorrect information, or incomplete information (U.S. NRC, 2009). Not surprisingly, it is recognized as an important element in geo-information science (Foody, 2003), as all geo-information contains uncertainty to some degree (Hwang et al., 1998). Uncertainty leads to imperfection and is a result of vague, ambiguous, imprecise, inaccurate, or incomplete information (Tavana et al., 2016). Low spatial data quality leads to an increase in uncertainty of the output models and maps. Uncertainty depends upon the density of observations and the mapping procedure; increasing the number of observations reduces the uncertainty about the spatial variability of an attribute rather than it reduces the spatial variability itself (Heuvelink, 1998).

Tavana et al. (2016) make a distinction between statistical and non-statistical methods for assessing uncertainty. Uncertainty is a result of vagueness, ambiguity, imprecision, inaccuracy, or incompleteness. Vagueness and ambiguity can be assessed using non-statistical methods such as fuzzy set and possibility theory. Imprecision, inaccuracy and incompleteness can be assessed using statistical methods, such as probability theory or Dempster-Shafer theory (Wang et al., 2005b; Tavana et al., 2016). Other methods include Monte Carlo, Taylor series expansion and Relative Variance Contribution (RVC) (Wang et al., 2005a). Fuzzy membership approaches have been used to assess uncertainty in air pollutant models (Guo et al., 2007; Shad et al., 2009), as well as probabilistic methods (Colvile et al., 2002; Yegnan et al., 2002). The latter refers to uncertainty as the variance in the input data compared to the variance in the output results.

1.6 Air pollution: sources and sinks

In order to model air pollution levels, it is important to understand the processes that lead to an increase in air pollutant concentrations (sources) and those that lead to a decrease in air pollutant concentrations (sinks). Sources can be of natural or anthropogenic origin, or the result of chemical processes in the atmosphere. Sinks can be related to meteorological and chemical processes. We differentiate between PM and gases. PM can be divided into different categories based upon their aerodynamic

diameter: PM₁₀ (particles <10 µm), PM_{2.5} (<2.5 µm), PM₁ (<1 µm) and UFPs (<0.1 µm). Fine particles (PM_{2.5}, PM₁ and UFPs) penetrate deepest into the gas-exchange part of the lung (Brunekreef and Holgate, 2002). Sources of fine particles include vehicular traffic and exhaust, construction activities, factories and power generation plants, wood burning, and agricultural activities (Graham, 2004). The concentration of particulate matter consists of primary and secondary pollutants. Primary pollutants include for example motor vehicle emissions. Secondary pollutants are formed by collision of smaller particles and gases (Gulliver and Briggs, 2004).

Nitric oxide (NO) is brought into the atmosphere by combustion of fossil fuels in power generators and motor vehicles (Brunekreef and Holgate, 2002; Graham, 2004) The toxic pollutant NO₂ is formed when the non-toxic NO is oxidized in the atmosphere in a chemical reaction using the O₃ present in the atmosphere (Brunekreef and Holgate, 2002; Fenger, 2009).

1.7 Health effects

Long-term exposure to traffic-related pollutants may have large health effects and shorten life expectancy (Hoek et al., 2002). An association has been suggested between long-term exposure to particulate matter air pollution and increased mortality from lung cancer, respiratory diseases and cardiovascular diseases (Dockery et al., 1993; Pope et al., 1995; Abbey et al., 1999; Hoek et al., 2002; Brook et al., 2010; Beelen et al., 2014). In different studies, different exposure-response associations were found. Those are related to differences in methods used for exposure assessment, differences in infiltration of particles indoors, particle composition and population composition (Hoek et al., 2013). NO₂ exposure has been associated with all-cause mortality in adults (Hoek et al., 2013) and respiratory infections, lung function growth, and asthma exacerbation in children (Goldizen et al., 2016). NO₂ is also a tracer for other traffic-related air pollutants such as black carbon and UFPs (Health Effects Institute, 2010). Pollutants have an indirect effect on asthma, interacting with pollen grains and enhancing the release of antigen, causing inflammation in the airways (Graham, 2004). The number of prospective cohort studies on the relationship between traffic-related air quality and asthma is limited (Gehring et al., 2010). It is difficult to distinguish between effects of specific pollutants due to a large overlap between the symptoms of different pollutants and a high correlation between different air pollutants in space and time (WHO, 2013a,b).

Asthma exacerbation in children specifically is a highly relevant endpoint, which needs better quantification to be used in future health impact assessment of outdoor air pollution (WHO, 2013a). Young children with asthma are very sensitive to the effects of air pollution. Increased

1. Introduction

sensitivity compared to adults is due to a combination of increased time spent exercising outside, high ventilation rates per body weight, developing lungs and immature metabolic pathways (Guarnieri and Balmes, 2014). Few epidemiological studies have directly compared effects in children and adults in the same study. In two studies conducted in the Netherlands in parallel, of which one was focused on adults and one on children, Van der Zee et al. found significant effects of increased PM₁₀ concentration levels on lung function in children, but not in adults. Similar results were found for acute symptoms in the lower respiratory tract, which were significantly related to PM₁₀ in symptomatic children but not in symptomatic adults (Van der Zee et al., 1999, 2000). School age children have more predictable time activity patterns, allowing more precise exposure assessment based on outdoor monitors.

1.8 Limit values and guidelines

Based on the scientific literature available on the health effects of air pollution, the WHO has provided a set of guidelines to reduce the health impacts of air pollution. The European Commission (EC) also formulated a directive in which limit values are given that should not be exceeded by the member states of the European Union (European Parliament and Council of the European Union, 2008). Table 1.1 shows an overview of the different maximum concentration values as advised by the WHO and implemented in the European directive and in the Dutch national law.

Table 1.1 WHO guidelines in comparison with EU directives and Dutch national law on the concentration limit for different pollutants (in $\mu\text{g m}^{-3}$) (Ministry of Infrastructure and the Environment, 1979; WHO, 2006; European Parliament and Council of the European Union, 2008).

	PM _{2.5} (annual mean)	PM _{2.5} (24 hr mean)	PM ₁₀ (annual mean)	PM ₁₀ (24 hr mean)	O ₃ (8 hr mean)	NO ₂ (annual mean)	NO ₂ (1 hr mean)
WHO	10	25	20	50	100	40	200
EC	25	-	40	50*	120**	40	200***
Dutch law	25	-	40	50*	120**	40	200***

* value may be exceeded maximum 35 times a year.

** value may be exceeded maximum 25 times a year based on a 3-year average.

*** value may be exceeded maximum 18 times a year.

The WHO guidelines concern the concentrations of PM₁₀, PM_{2.5}, O₃, and NO₂. There has been too little research on the health effects of UFPs to set guidelines for those pollutants (WHO, 2006). Unregulated pollutants, such as PM₁ and UFPs, are often not measured by ambient

1.9. Problem statement

air quality monitoring stations (Goldizen et al., 2016). This creates a vicious circle: without monitoring sites, little research is being done, which in turn makes it difficult to come up with guidelines – and without guidelines, no monitoring is performed.

As can be derived from Table 1.1, the values from the European directive are implemented in Dutch national law without changes. The EC limit values are up to twice as high as the WHO guidelines for PM_{2.5} and PM₁₀. The WHO has also set up a set of interim targets, including a description of the difference in health effects between the interim targets (IT-x) and the air quality guidelines (AQG). The aim of the interim targets is to provide policy makers all over the world with various options for air quality management (WHO, 2006).

Even well below the EC limit values, long-term exposure to fine particulate matter causes health effects and mortality (Beelen et al., 2014). Brunekreef and Holgate (2002) mention that a limit value for concentrations to cause health effects is absent or at very low value, because exposure to low concentrations of air pollutants already causes damage. Acknowledging this, the WHO states that the guideline values mentioned cannot fully protect human health (WHO, 2006). As lower limits are absent, there is an interest in health effects at lower levels of air pollution. To study these, air pollutant models are often used to estimate exposure. A distinction is made between short-term exposure and long-term exposure. Short-term exposure refers to the exposure for hours up to days (WHO, 2013b), whereas long-term exposure refers to exposures of a year or longer (Hoek et al., 2013). A distinction is made as well between short-term health effects and long-term health effects. Short-term health effects are acute effects of exposure on daily symptoms or lung function (Weinmayr et al., 2010). Long-term health effects include the development of diseases over a longer period of time. Long-term health effects are often studied in relation to long-term exposure (Gehring et al., 2010; Hoek et al., 2013; Beelen et al., 2014), whereas short-term health effects are studied in relation to short-term exposure (Weinmayr et al., 2010; Goldizen et al., 2016). This is also reflected in the WHO guidelines (WHO, 2006) that associate daily averages with short-term mortality risks and annual averages with long-term mortality risks.

1.9 Problem statement

Recently developed low-cost urban air quality sensor networks offer the possibility for monitoring air pollution at a fine spatio-temporal resolution. However, low-cost sensors may be more prone to report outliers, and their data quality is often unknown (Snyder et al., 2013). Compared to conventional monitors, measurements of low-cost air quality sensors are more sensitive to interference effects of humidity and other pollutants, as well as a loss of sensitivity to the target pollutant over time,

1. Introduction

referred to as sensor drift. An evaluation of the data quality of low-cost sensors is of major importance, as it determines the usefulness of the measurements in different applications. The communication of data quality issues is of importance to avoid misinterpretation of the data when open to the public, media, politicians and researchers. Next, the possibilities of using low-cost urban sensor networks for modelling and mapping air quality should be evaluated. Using those fine resolution air quality maps in combination with health data, the spatially explicit health risks related to air pollution can be visualized. Here, a neglected topic is the propagation of uncertainty from the input data to the output maps.

1.10 Research objectives

The key objectives of this thesis are:

1. To develop an outlier detection method suitable to detect outliers in space and time while accounting for the large spatio-temporal variability of air pollutant concentrations in an urban area.
2. To develop and evaluate automatic calibration methods for low-cost sensors in an urban air quality sensor network, accounting for drift and interference effects.
3. To develop a spatio-temporal kriging framework for modelling air pollutant concentrations using a low-cost sensor network.
4. To create burden of disease maps, expressing the spatial variability in health risks related to ambient air pollution, using a low-cost sensor network to allow spatially refined human exposure assessment.

1.11 Outline

Chapter 1 introduces the challenges related to spatio-temporal big data analysis and data quality issues associated with low-cost sensor network data. It provides context and background information related to statistical modelling and mapping, as well as air pollution sources and sinks, health effects and limit values. It presents research gaps in this field and the related research objectives of this thesis.

Chapter 2 gives a detailed overview of the case study area and the low-cost air quality sensor network used in this study.

Chapter 3 presents a new outlier detection method, in which observations are classified in spatio-temporal classes to determine outlier threshold levels based on the location and time at which an observation was taken. Transformations are applied to account for non-normality of air pollutant concentrations.

Chapter 4 presents a novel iterative Bayesian calibration method, and compares the method to several existing calibration methods. The methods are compared in terms of temporal stability, spatial transferability, and sensor-specificity.

Chapter 5 presents a spatio-temporal regression kriging framework to model air pollution in an urban area. The trend part of the model consists of a set of spatial and temporal covariates, and the residuals are interpolated using simple kriging.

Chapter 6 presents a burden of disease assessment based on a new panel study on asthmatic children. The panel study data is combined with a priori effect estimates from literature in a Bayesian framework. The updated effect estimates are combined with modelled air pollution concentrations to obtain a burden of disease map. The propagation of uncertainty from the input data to the burden of disease map is evaluated.

Chapter 7 provides a synthesis of this thesis. The main results are summarized, as well as the implications, limitations and suggestions for further research.

1.12 Author contributions

Chapters 3, 4 and 5 are based on published papers and Chapter 6 is currently under review. For the purpose of consistency throughout this thesis, small changes have been made compared to the published versions. Case study area descriptions have been removed from the papers and merged in Chapter 2. Variable names and symbols have been changed to avoid confusion due to multiple meanings and definitions. Chapters 3–6 contain references to ‘we’, referring to the authors of the publication. In these publications, VZ carried out all the scientific analyses and wrote the manuscript. FO advised on the spatio-temporal statistical framework, GH advised on air quality, sensor calibration and health effects, and AS advised on data interpretation and spatio-temporal data analysis. All authors were involved in textual editing of the final manuscripts.

Case study area

2

2.1 Study area: the city of Eindhoven

The study area is the city of Eindhoven, located in the southern part of the Netherlands (Figure 2.1). Its background concentrations of air pollutants are relatively high, mainly due to surrounding agricultural activities and to the city's position with respect to industrial areas such as Antwerp and Ghent in Belgium, the Ruhr area in Germany and Rotterdam area in the Netherlands. According to the Royal Netherlands Meteorological Institute (Koninklijk Nederlands Meteorologisch Instituut, KNMI), the prevailing wind direction in Eindhoven is from the south-west (KNMI, 2011), causing long range transport of pollutants. The city has a high population density and traffic intensity, elevating levels of traffic-related pollutants such as NO₂.

Together with the high background concentrations, there is a large short-distance spatial variation. This short-distance variation is not only found for PM, but most evidently in gases such as NO₂. Traffic-related air pollutants are a major source of ambient air pollution in urban areas (Goldizen et al., 2016). Therefore, the spatial variability of PM_{2.5} and NO₂ is highest in the city, because of the relation of these pollutants with traffic (Beelen et al., 2014). In recent years, inhabitants of the city have become more aware of the health effects of traffic-related air pollution. Because of the low density of existing monitoring networks and a general mistrust in routine dispersion models used for the evaluation of exceedance of limit values, the AiREAS civil initiative has been set up to monitor the air quality at a fine spatio-temporal resolution.

2.2 The ILM air quality network

AiREAS is a civil initiative in Eindhoven in which inhabitants cooperate with companies, universities and governmental organizations (Close, 2016). As part of this initiative, an air quality sensor network has been set up in Eindhoven, referred to as Innovatief Luchtmeetnet (ILM). It is the first fine resolution urban air quality sensor network in the Netherlands

2. Case study area



Figure 2.1 Location of Eindhoven (black dot) within the Netherlands

(Figure 2.2). It was installed in November 2013 and has been operated continuously since. The network consists of 35 weatherproof 'airboxes'. Since the total area of the municipality of Eindhoven is approximately 90 km², the sensor network is relatively dense. The airboxes are of size 43 × 33 × 20 cm (Figure 2.3) and contain an array of sensors. They are manufactured by the former Energy Research Centre of the Netherlands (ECN), now part of the Netherlands Organisation for applied scientific research (TNO). Each airbox measures PM₁₀, PM_{2.5}, PM₁, O₃, temperature and humidity as the air flows through. A large part, 25 airboxes, also measures NO₂ since 2015 based on available budget. Because of the high sensor costs, UFPs are measured at six locations. The UFP sensors are installed in separate boxes which are attached to the airboxes for power supply and GPRS connection. From November 2016 to February 2017 the UFP sensors were attached to different airboxes every three weeks to cover multiple locations. All AiREAS data is publicly available (AiREAS, 2016).

2.3 Spatial sampling scheme

The spatial locations of the airboxes were chosen based on several criteria (Close, 2016), following the philosophy of the ESCAPE study (Eeftens et al., 2012). Most importantly, sampling sites represent locations where humans are exposed. The airboxes are located in the build-up area of the city, near residential areas and schools. The set of locations covers urban background locations in quiet neighborhoods as well as urban traffic locations near busy roads. One airbox is located outside of the city for regional background monitoring. All airboxes are installed in fixed positions at lamp posts to supply electricity. They are located at 2.5–3 m height, representing human exposure as closely as possible, while minimizing the risk of accessibility by third persons. At two locations in the city, an airbox is collocated with a reference monitor (Section 2.5).

2.4. Airbox sensors

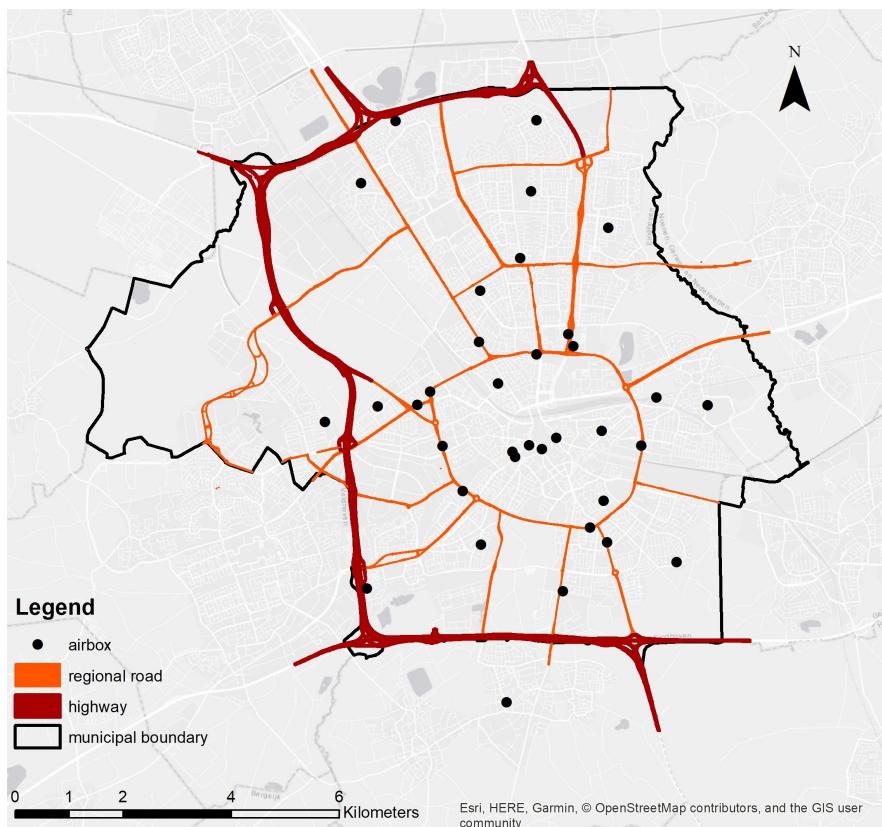


Figure 2.2 Locations of the airboxes in the city of Eindhoven, the Netherlands

2.4 Airbox sensors

Ventilation strips on the sides of the airbox allow air to flow through. A gauze protects the airbox from insects and the air is dried to minimize interference from relative humidity. An overview of the installed sensors is shown in Table 2.1. An optical PM sensor is used to count the size and number of particles that flow through with the help of an resistive heater. The light of an infrared LED is scattered by PM and then measured by a photo-diode detector. The raw output of the sensor consists of digital pulses proportional to particle count concentrations (Austin et al., 2015). A combination of particle size and count is used to convert the particle counts to concentrations in $\mu\text{g m}^{-3}$. The particles entering the UFP sensor are charged and enter a Faraday cage in which the deposited charge is measured using a very sensitive current meter and converted to particle number concentrations (Marra et al., 2010). O_3 is measured using a metal oxide sensor after heating and ambient temperature correction. Sensor resistance is converted to O_3 concentrations (Hamm et al., 2016). NO_2 is measured using the electrochemical cell Citytech Sensoric NO2

2. Case study area



Figure 2.3 Airbox attached to light pole.

3E50 in a differential measurement setup. A switching valve and reagent cartridges are used in front of the electrochemical cell to dry the air. Observations are discarded when temperature and humidity fall outside acceptable ranges. The airboxes are attached to light poles for power supply. The data of all sensors is sent to a server every 10 minutes via a GPRS connection.

After initial lab and field calibration of the sensors, data have been collected since November 2013. There are some gaps in the data for moments in time in which the instruments were removed for testing, adjusting or calibration. The sensors were recalibrated at the end of 2015, together with the installation and calibration of the NO₂ sensors. Throughout this thesis, NO₂ sensor data of 2016 are used.

2.5 Reference measurements

The national air quality sensor network (LML) is maintained and operated by the National Institute for Public Health and the Environment (RIVM, 2019a). The LML sensor network consists of around 60 measurement stations throughout the Netherlands, of which two are situated in Eind-

2.5. Reference measurements

Table 2.1 Variables measured and instruments used in the airboxes

Variable	Units	Instruments
PM ₁₀	$\mu\text{g m}^{-3}$	Shinyei PPD42 ECN revised
PM _{2.5}	$\mu\text{g m}^{-3}$	Shinyei PPD42 ECN revised
PM ₁	$\mu\text{g m}^{-3}$	Shinyei PPD42 ECN revised
UFP	# cm ⁻³	Aerasense NanoMonitor PNMT 1000
O ₃	$\mu\text{g m}^{-3}$	E2V MICS 2610
NO ₂	$\mu\text{g m}^{-3}$	Citytech Sensoric NO ₂ 3E50 ECN revised
Temperature	°C	Sensirion SHT75
Relative humidity	%	Sensirion SHT75

hoven. Although the LML has a lower spatial and temporal resolution, the uncertainty of the measurements is expected to be lower than the uncertainty of the ILM measurements. The measurement uncertainty of the LML sensor network is about 15-20% for PM_{2.5} and PM₁₀ (RIVM, 2014). The required maximum uncertainty to follow the European directives equals 25% for PM₁₀ (Nguyen and Hoogerbrugge, 2014). Although the uncertainty of the LML sensor network is below the threshold set by the European directives, it may create problems for specific applications. For example, when the concentrations of PM_{2.5} and PM₁₀ are close to each other, the measured values of PM_{2.5} could be higher than the measured values of PM₁₀ (RIVM, 2014). This does of course not represent a valid situation, with PM_{2.5} being part of PM₁₀. This uncertainty level also allows for negative values of PM_{2.5} and PM₁₀ concentrations to occur. For calibration and validation purposes, airboxes are located within a few meters from the LML measurement stations.

Outlier detection in urban air quality sensor networks

3

Abstract

Low-cost urban air quality sensor networks are increasingly used to study the spatio-temporal variability in air pollutant concentrations. Recently installed low-cost urban sensors, however, are more prone to result in erroneous data than conventional monitors, e.g. leading to outliers. Commonly applied outlier detection methods are unsuitable for air pollutant measurements that have large spatial and temporal variations as occur in urban areas. We present a novel outlier detection method based upon a spatio-temporal classification, focusing on hourly NO₂ concentrations. We divide a full year's observations into 16 spatio-temporal classes, reflecting urban background vs. urban traffic stations, weekdays vs. weekends and four periods per day. For each spatio-temporal class, we detect outliers using the mean and standard deviation of the normal distribution underlying the truncated normal distribution of the NO₂ observations. Applying this method to a low-cost air quality sensor network in the city of Eindhoven, the Netherlands, we found 0.1-0.5% of outliers. Outliers could reflect measurement errors or unusual high air pollution events. Additional evaluation using expert knowledge is needed to decide on treatment of the identified outliers. We conclude that our method is able to detect outliers while maintaining the spatio-temporal variability of air pollutant concentrations in urban areas.

This chapter is published as: Van Zoest, V.M., Stein, A., Hoek, G., 2018. Outlier Detection in Urban Air Quality Sensor Networks. *Water, Air, & Soil Pollution* 229, 111. doi:10.1007/s11270-018-3756-7.

3. Outlier detection in urban air quality sensor networks

3.1 Introduction

Air quality is monitored globally, with national monitoring networks being used to assess air pollution in relation to environmental limit values. In Europe, national, regional and local environmental agencies operate these monitoring networks according to EU guidelines (European Parliament and Council of the European Union, 2008), complying to high standards of equivalency (EC Working Group on GDE, 2010). Each European country has a network of air quality monitoring stations that are located in urban, suburban and rural areas.

Health effects of air pollution have attracted public and scientific attention globally as the global burden of disease of outdoor air pollution is significant (Cohen et al., 2017). The health risks are typically highest in urban areas because of their high population density, a high density of schools and hospitals and higher air pollution concentrations. In recent local networks, urban air quality is measured using a larger number of sensors than in national air quality networks, allowing detection of more local sources. In response to the increasing civil interest in the air they breathe, more local initiatives have resulted in extended low-cost monitoring networks. These provide more detailed spatio-temporal data on air quality. Data from such sensor networks however are more prone to result in errors and their spatio-temporal data quality is often unknown (Snyder et al., 2013). This leads to an increased need for data evaluation. Data evaluation of low-cost air quality networks typically includes outlier detection, comparison with classical monitors, comparison of inter-sensor measurements and evaluation of the stability of sensors. In this paper we focus on outlier detection.

Outlier detection is an important part of data cleaning and particularly relevant for low-cost air quality sensor networks. Outlier detection is defined as the detection of values that are statistically significantly different from the expected value at a given time and location. Outlier detection is important for detecting air pollution events, but also for removing errors that might otherwise affect data analysis and comparison, including unnecessary unrest among the population if data are publicly available online. Errors in this context refer to inaccuracies due to air quality sensor faults, mistakes in the human handling of the sensors, or positioning of the sensors under conditions for which they are not designed. Events are valid observations of very high or low air pollutant concentrations compared to the concentrations expected at a given time in a given location (Zhang et al., 2012). True events can be related to very local sources (e.g. a small fire, truck idling within meters of a monitor) or to very unusual weather circumstances such as low mixing height and high atmospheric stability resulting in poor dispersion of emitted pollutants.

Functional outlier detection, as a common type of temporal outlier detection, compares various function curves of fixed time periods. In the past, this method was applied to PM₁₀, SO₂, NO, NO₂, CO and O₃ to detect months with unusually high air pollutant concentrations (Martínez Torres et al., 2011), or to detect working days and non-working days with outlying NO_x levels (Febrero et al., 2007, 2008; Sguera et al., 2016). Functional outlier detection is used to compare entire vectors of measurements (e.g. all observations in a month) and is therefore less suitable for the detection of individual outliers. Comparing an observation only to its temporal neighborhood may also lead to the neglect of a systematic bias in the sensor.

In spatial outlier detection, an observation is compared to the observations in its spatial neighborhood. Bobbia et al. (2015) used kriging to detect outliers in PM₁₀ concentrations on a provincial scale. Spatio-temporal outlier detection combines the spatial neighborhood with a temporal neighborhood. It has been applied to PM₁₀ measurements at the European scale (Kracht et al., 2014). At this scale level however only rural and urban background stations can be used, as the methods are not suitable for dealing with the wide spatial variation of air pollutants in an urban area.

For an urban air quality sensor network, both spatial and spatio-temporal outlier detection have only been applied to air pollutants that show a low spatial variation. Hamm (2016) and Shamsipour et al. (2014) applied spatial and spatio-temporal outlier detection methods on PM₁₀, which in cities is mostly dominated by regional background concentrations from sources outside the city (Eeftens et al., 2012). Distance-weighting techniques such as kriging were successfully applied to urban PM₁₀ for filling missing values and for outlier detection. There was no need for space varying covariates because PM₁₀ concentration was not related to the type of location or street (Hamm, 2016). For NO₂, however, the concentrations can vary over short distances, e.g. governed by the traffic density of a street (Briggs et al., 1997; Cyrys et al., 2012). As the distances over which NO₂ concentrations vary (10-s of meters) are commonly shorter than the distances between sensor locations (km's), spatial outlier detection methods based on distance-weighting cannot be applied to NO₂ measurements in cities.

The objective of this study was to develop an adequate outlier detection method for an urban air quality sensor network. Such a network is characterized by a fine-scale spatial and temporal variation in air quality. For this study, we use NO₂ data from the ILM air quality sensor network located in the city of Eindhoven, the Netherlands (Chapter 2).

3. Outlier detection in urban air quality sensor networks

3.2 Data preprocessing

We focus on NO₂, as an air pollutant with a high spatial variability in urban areas (Cyrys et al., 2012). The hourly concentrations measured by the conventional monitors in Eindhoven ranged from 2.5 to 123.8 µg m⁻³ in 2016, with a mean of 28.6 µg m⁻³ and a standard deviation of 16.5 µg m⁻³. The distribution of NO₂ concentrations is skewed with a long right-tail ($P_{95} = 61.0 \mu\text{g m}^{-3}$, $P_{99} = 78.8 \mu\text{g m}^{-3}$). We used airbox NO₂ data for this study, as described in Chapter 2. To reduce the noise, the 10-minute NO₂ measurements were averaged to hourly values for the current analysis. Data for the full year of 2016 were used for this study.

The data were cleansed before being used. Negative concentration values occurred when the concentrations were below the limit of detection and were removed from the dataset (1.5%). Zeroes in the data indicated a sensor failure and were removed from the dataset (1%). High peaks in NO₂ concentrations can occur in 10-minute data if the sensor is exposed to a high concentration peak for a short period of time. Similar peaks in hourly concentration data however are more likely to be caused by sensor failure and influence the outlier detection. To carefully remove extreme peaks in hourly concentrations we turned to the two conventional NO₂ monitors in Eindhoven, which are part of the national air quality monitoring network. We set a threshold equal to three times the maximum hourly concentration measured in 2016. In doing so, concentration values $x_i > 372 \mu\text{g m}^{-3}$ were removed (0.02%). Such extreme peaks are impossible to occur under natural conditions in this city and are most probably caused by sensor failures. Such failures also caused frozen concentration values for several hours or days. Those values were removed from the dataset as well (1.5%). One airbox showed a consistent positive bias. Including it in the analysis showed the many outliers of the airbox, but also strongly influenced the percentage of outliers that could be detected in other airboxes, which almost dropped to zero. Therefore, data of this airbox was removed prior to the final outlier detection shown here.

3.3 Methods

Outlier detection is based upon checking whether an observed concentration value falls within a given confidence interval, set by:

$$\mu \pm \zeta \times \sigma \quad (3.1)$$

where μ is the mean NO₂ concentration level in µg m⁻³, σ the standard deviation, and ζ is an indicator of the size of the confidence interval. We consider Eq. 3.1 for grouped NO₂ concentration observations within temporal, spatial and spatio-temporal neighborhoods. Assuming

3.3. Methods

independence and normality, the value of ζ is set at 1.96 for a 95% confidence level (Kracht et al., 2014) or at 2.97 for a 99.7% confidence interval, depending on the required strictness of the outlier detection. We used $\zeta = 2.97$ which in related studies has been rounded to $\zeta = 3$ (Martínez Torres et al., 2011; Shamsipour et al., 2014).

NO_2 concentrations in an urban setting, however, highly depend on the proximity of busy roads and therefore too much noise in concentrations is found within the neighborhood to detect values that are abnormally high given their location. Similarly, temporal neighborhoods have a highly temporally dependent variation in air pollutant concentrations over the day.

We propose to overcome this by classifying the locations and time periods into 16 spatio-temporal categories distinguished by different levels of air pollution. To do so, we divided the measurement locations into two categories: urban traffic and urban background locations (Figure 3.1). These take into account the positions of the airboxes near specific land-use types, the presence of traffic and distance from the center. We take four intervals: ‘traffic hours’ (6:01-9:00 and 16:01-20:00 UTC time), ‘off-peak hours’ (9:01-16:00 and 20:01-22:00 UTC time), ‘transition periods’ (22:01-1:00 and 5:01-6:00 UTC time), and ‘night hours’ (1:01-5:00 UTC time).

Days of the week were divided into two classes: weekdays (Monday to Friday) and weekend days (Saturday and Sunday). This all resulted into sixteen classes: eight temporal classes and two spatial classes. For each spatio-temporal class K , the three steps described below are taken to detect outliers.

1. We transformed the NO_2 concentrations using the square root transformation to obtain approximately normally distributed values (Figure 3.2), i.e. to justify the use of Eq. 3.1.

Before transforming the NO_2 concentration values, in line with Kracht et al. (2013) we added a value of (1 - minimum value of all observations) to all observations, to prevent values $< 1 \mu\text{g m}^{-3}$ from increasing during square root transformation while values $> 1 \mu\text{g m}^{-3}$ decrease:

$$x_k = \sqrt{\text{NO}_2k + (1 - \min(\text{NO}_2))} \quad (3.2)$$

where NO_2k is an observation and x_k is the transformed observation in spatio-temporal class K where k is an observation index in $\{1 \dots N_K\}$ for N_K total number of observations in class K . Note that x_k has coordinates in space and time.

2. As a result of the transformation in Eq. 3.2, the distribution of NO_2 concentrations is truncated at the left at $1 \mu\text{g m}^{-3}$. The resulting

3. Outlier detection in urban air quality sensor networks

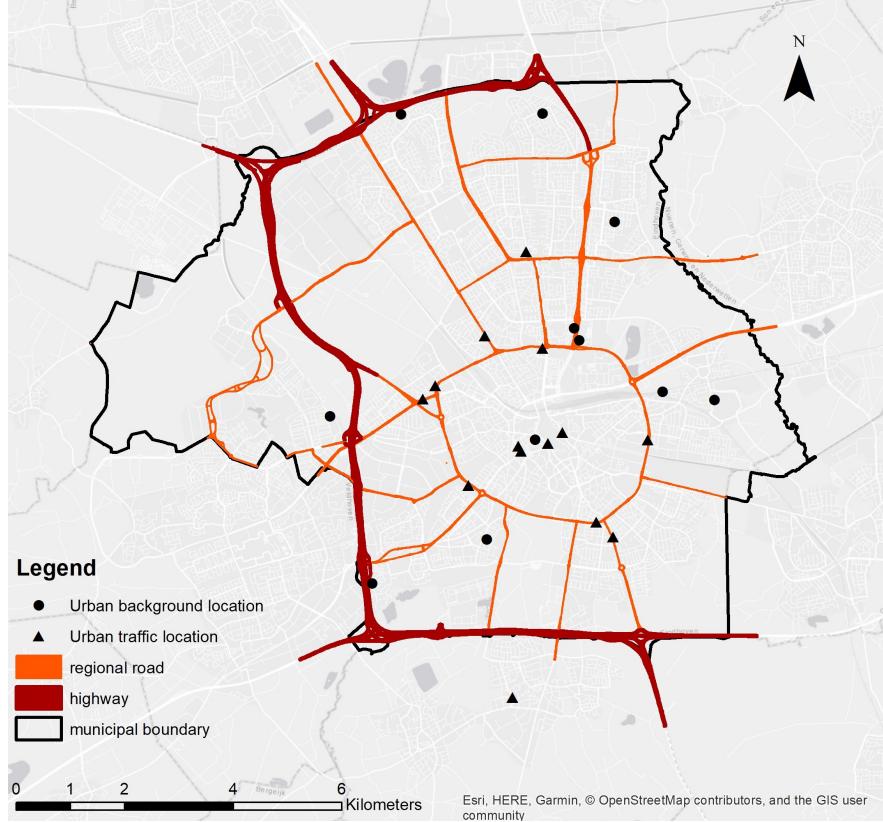


Figure 3.1 Locations of the airboxes measuring NO₂ at urban background locations (circles) and urban traffic locations (triangles).

distribution thus showed a truncated normal distribution (Figure 3.3). For each square root transformed NO₂ observation $x_{k,i}$, we temporarily excluded the i th observation from the NO₂ concentration dataset, in order to avoid impact of the observation, a potential outlier, on the standard deviation and mean. We then obtained the mean and standard deviation of the remainder of the dataset as:

$$m_K^{-i} = \frac{\sum_k (x_k) - x_{k,i}}{N_K - 1} \quad (3.3)$$

$$s_K^{-i} = \sqrt{\frac{\sum_k (x_k - m_K^{-i})^2 - (x_{k,i} - m_K^{-i})^2}{(N_K - 2)}} \quad (3.4)$$

where summation extends over all hourly NO₂ observations x_k in one spatio-temporal class K , and m_K^{-i} and s_K^{-i} are the mean and the standard deviation of all hourly NO₂ observations excluding the i th observation $x_{k,i}$, respectively.

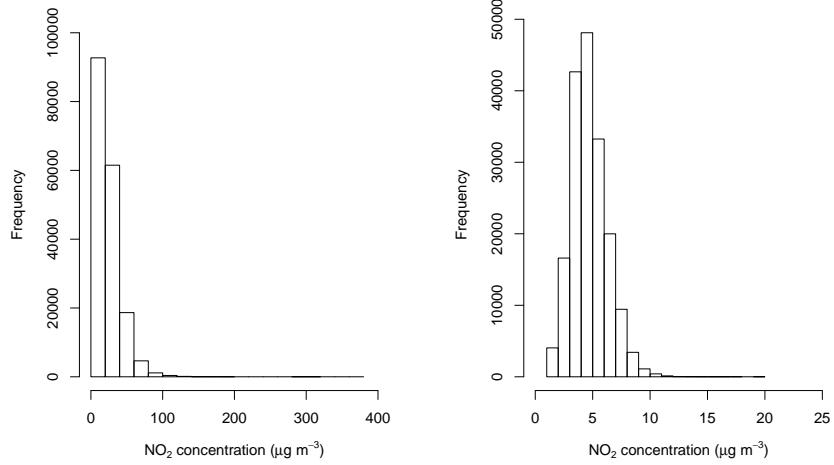


Figure 3.2 Distribution of NO_2 concentrations, before square root transformation (left) and after square root transformation (right)

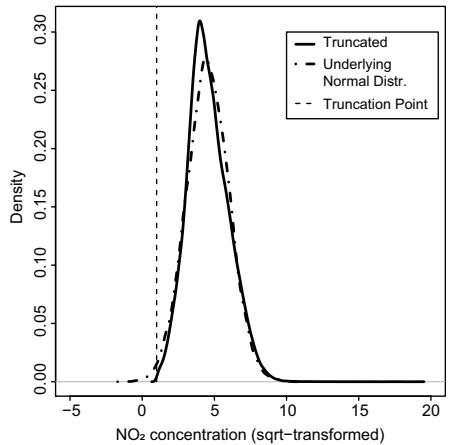


Figure 3.3 The truncated normal distribution of square root transformed NO_2 concentrations (solid line) and its underlying normal distribution (dot dashed line). The truncation point is set at 1 (dotted line).

Eqs. 3.3 and 3.4 provided both the mean and the standard deviation of the truncated normal distribution of NO_2 concentrations, referred to as m_K^{-i} and s_K^{-i} . Equation 3.1 requires a normal distribution, and therefore we are more interested in the mean and standard deviation of the underlying normal distribution, referred to n_K^{-i} and t_K^{-i} , respectively, rather than the mean and standard deviation of the truncated normal distribution. We use a maximum likelihood estimator to obtain estimated values n_K^{-i} and t_K^{-i} . The

3. Outlier detection in urban air quality sensor networks

log likelihood function is given as:

$$\sum_k \ln(f(x_k | \theta_1)) \quad (3.5)$$

where $f(x_k | \theta_1)$ is the probability density function of the truncated normal distribution of NO₂ concentrations, returning the probability of observing x_k given a set of parameters $\theta_1 = (m_K^{-i}, s_K^{-i}, a, b)$, for $a \leq x \leq b$. In our case of left-truncation we have $a = 1$ and $b = \infty$. Then the probability density function is given as:

$$f(x_k | \theta_1) = \frac{\phi\left(\frac{x_k - n_K^{-i}}{t_K^{-i}}\right)}{t_K^{-i} \left(1 - \Phi\left(\frac{a - n_K^{-i}}{t_K^{-i}}\right)\right)} \quad (3.6)$$

Imputing Eq. 3.6 into the log likelihood function and taking $\theta_2 = (n_K^{-i}, t_K^{-i})$ gives:

$$L(\theta_2) = \sum_k \left(\ln \left(\phi \left(\frac{x_k - n_K^{-i}}{t_K^{-i}} \right) \right) - \ln \left(t_K^{-i} \left(1 - \Phi \left(\frac{a - n_K^{-i}}{t_K^{-i}} \right) \right) \right) \right) \quad (3.7)$$

where $\phi(\cdot)$ is the probability density function of the normal distribution and $\Phi(\cdot)$ is the corresponding cumulative distribution function. Optimization of the log likelihood function Eq. 3.7 using Nelder and Mead (1965) gives maximum likelihood values for n_K^{-i} and t_K^{-i} . We used the parameters m_K^{-i} and s_K^{-i} as starting values.

For each observation $x_{k,i}$ removed from the dataset, n_K^{-i} and t_K^{-i} are computed on the remainder of the spatio-temporal class dataset as described above.

3. Next, Eq. 3.1 is adapted to find the lower and upper thresholds of values considered outliers:

$$n_K^{-i} \pm \zeta \times t_K^{-i} \quad (3.8)$$

which is computed for each individual observation. If the i th observation $x_{k,i}$ falls outside this interval, it is considered to be an outlier. The observations of spatio-temporal class K are backtransformed after the outlier detection:

$$NO2_k = (x_k)^2 - (1 - min(x_k)) \quad (3.9)$$

returning the NO₂ concentrations in $\mu\text{g m}^{-3}$. Depending upon the purpose of the outlier detection, the outlying observations can then be removed or further investigated.

We further computed the thresholds for the entire dataset, without removal of observation $x_{k,i}$ in Eqs. 3.3 and 3.4. The mean and standard deviation of the underlying normal distribution are then expressed by n_K and t_K , respectively, which results in the following thresholds:

$$n_K \pm \zeta \times t_K \quad (3.10)$$

which are also back-transformed using Eq. 3.9. These thresholds are not used for actual outlier detection, but as an approximation of the thresholds for each spatio-temporal class. This allowed us to compare the thresholds of the sixteen spatio-temporal classes. Given the large number of observations in each class, the thresholds are not highly affected by removing one of the observations.

For comparison with conventional monitors, the same analysis was repeated with data from the two NO₂ monitors in Eindhoven which are part of the national air quality monitoring network. Both conventional monitors are located in an urban traffic location and therefore considered as the same spatial class. We used the temporal classification similar to the one used in the analysis of the airbox data.

3.4 Results

Of the 25 airboxes measuring NO₂ that were used for this analysis, 11 were classified as urban background locations, and 14 were classified as urban traffic locations (Figure 3.1). Table 3.1 shows the approximated upper thresholds for outliers in each spatio-temporal class (Eq. 3.10). All lower thresholds were equal to zero. For the values of n_K and t_K of each spatio-temporal class we refer to Table A3.1 in the Appendix. Table 3.2 shows the percentage of outliers detected per spatio-temporal NO₂ concentration class using a full year of hourly NO₂ data. Note that our method defines unusual observations, which are not necessarily errors, but which could also be very unusual air pollution events related to local sources, or extreme weather conditions of low wind speed and high atmospheric stability.

Table 3.2 shows that the period of night hours during the weekend has an increase in the number of outliers, both for urban traffic locations and urban background locations. Both n_K and t_K are relatively small in these spatio-temporal classes compared to other spatio-temporal classes. The combination of a short right-tail and the relatively small n_K and t_K cause the upper threshold to be low while detecting a relatively high number of outliers in the thicker tail. All categories have an approximately similar percentage of outliers and there are no large deviations.

The boxplots in Figure 3.4 show the range in concentrations that were considered outliers for each spatio-temporal class. The lower whiskers

3. Outlier detection in urban air quality sensor networks

Table 3.1 Upper thresholds for hourly average NO₂ concentrations ($\mu\text{g m}^{-3}$) above which considered outliers, per spatio-temporal class, using $\zeta = 2.97$. Between brackets: N_K shows the number of hourly concentration values in this class.

	Urban traffic		Urban background	
	Week	Weekend	Week	Weekend
Rush hours	96.6 ($N_K=17,761$)	78.4 ($N_K=7,127$)	81.0 ($N_K=17,660$)	62.3 ($N_K=6,983$)
Off-peak hours	87.3 ($N_K=22,768$)	76.7 ($N_K=9,153$)	72.9 ($N_K=22,554$)	61.3 ($N_K=8,961$)
Night hours	63.2 ($N_K=10,161$)	63.6 ($N_K=4,123$)	58.6 ($N_K=9,983$)	57.3 ($N_K=3,995$)
Transition hours	76.5 ($N_K=10,195$)	67.1 ($N_K=4,129$)	67.9 ($N_K=10,031$)	56.4 ($N_K=3,983$)

Table 3.2 Percentage outliers per spatio-temporal NO₂ concentration class for hourly values in 2016, using $\zeta = 2.97$.

	Urban traffic		Urban background	
	Week	Weekend	Week	Weekend
Rush hours	0.2%	0.2%	0.2%	0.2%
Off-peak hours	0.2%	0.2%	0.2%	0.2%
Night hours	0.2%	0.5%	0.1%	0.5%
Transition hours	0.3%	0.3%	0.3%	0.3%

are short and close to the threshold values shown in Table 3.1. Especially during off-peak hours in the weekend, the range in concentrations of the outliers is large. Extreme outliers, denoted by the dots, representing observations outside $1.5 \times IQR$ (interquartile range) of the outliers, occur in many spatio-temporal classes. Note that these boxplots are only based on the outliers, which is a small number of observations.

Figures 5 and 6 show NO₂ measurements during two weeks in 2016 containing outliers. Figure 3.5 shows the week from April 25 until May 1, of an urban background location, whereas Figure 3.6 shows the week from February 8 until February 14 of an urban traffic location. The concentrations at the urban traffic location were higher than those at the urban background location. Due to the spatial classification, some concentration values are considered outliers at the urban background location, while they are non-outliers at the urban traffic location. The temporal classification is also visible in Figure 3.6: concentration values that are considered outliers at one point in time can be considered non-outliers at other points in time, e.g. during rush hours in which higher concentrations are expected. This is a major difference as compared to applying the outlier threshold on the entire dataset without classification

3.4. Results

(Eq. 3.1), yielding an expected 0.3% of outliers as cut off peaks without taking spatio-temporal variability in the NO_2 concentrations into account.

Figure 3.5 shows two outliers, labelled (a) and (b), occurring during the night, in the early morning (1:00-3:00) of April 28. During weekday night hours at an urban background location, the transformed (Eq. 3.2) parameter estimations are $n_K=3.965$ and $t_K=1.265$. Entered in Eq. 3.8 with $\zeta = 2.97$, and back-transformed using Eq. 3.9, this gives an upper threshold of $58.6 \mu\text{g m}^{-3}$. The concentrations measured at outliers (a) and (b) were 75 and $70.8 \mu\text{g m}^{-3}$, respectively, both exceeding the upper threshold. Given that these are consecutive observations and within the range of thresholds of other periods, it is not clear whether these observations reflect instrument error.

From Figure 3.6 we identify four outliers, labelled (a)-(d). Three outliers, specifically (a), (c) and (d), are clearly higher than expected concentration values in any of the spatio-temporal categories. They are furthermore single observations. Outlier (b) occurred on February 9 from 23:00-0:00 in the temporal class ‘transition period’. In this spatio-temporal class, with (transformed) $n_K=4.76$ and $t_K=1.36$, the upper threshold is approximately $(4.76 + 2.97 \times 1.36)^2 - (1 - 0.0244) = 76.5 \mu\text{g m}^{-3}$. The concentration measured at (b) is $81.8 \mu\text{g m}^{-3}$, exceeding the upper

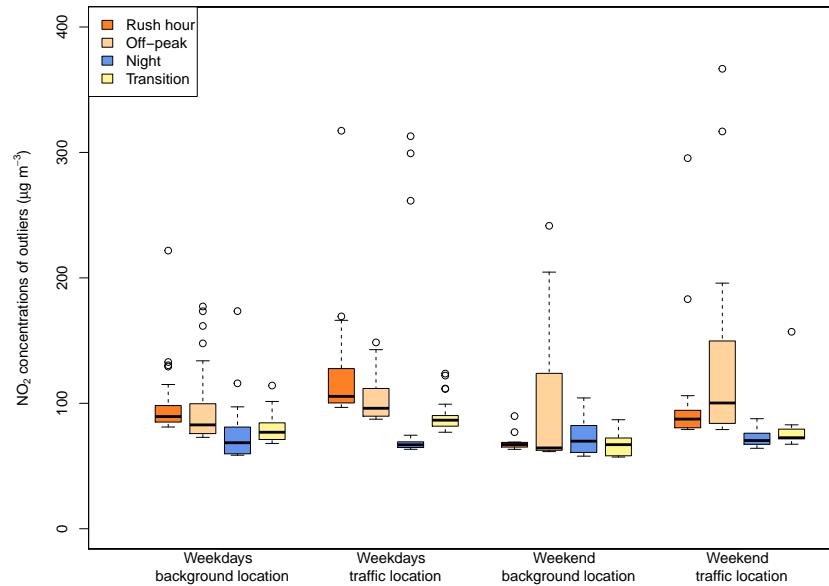


Figure 3.4 Boxplots of the outliers in each spatio-temporal class

3. Outlier detection in urban air quality sensor networks

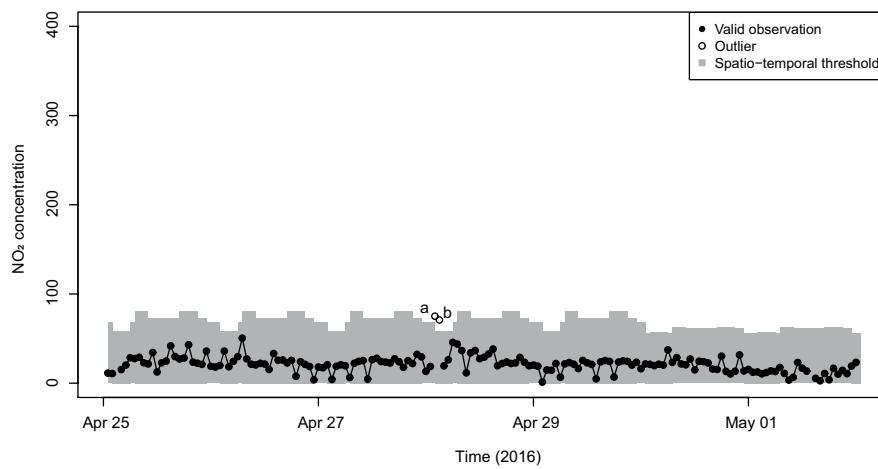


Figure 3.5 NO₂ concentrations ($\mu\text{g m}^{-3}$) measured by airbox 6, at an urban background location. Filled circles indicate non-outlying observations; unfilled circles indicate outliers using $\zeta = 2.97$. The gray bars indicate the threshold values for each temporal class, for urban background airboxes.

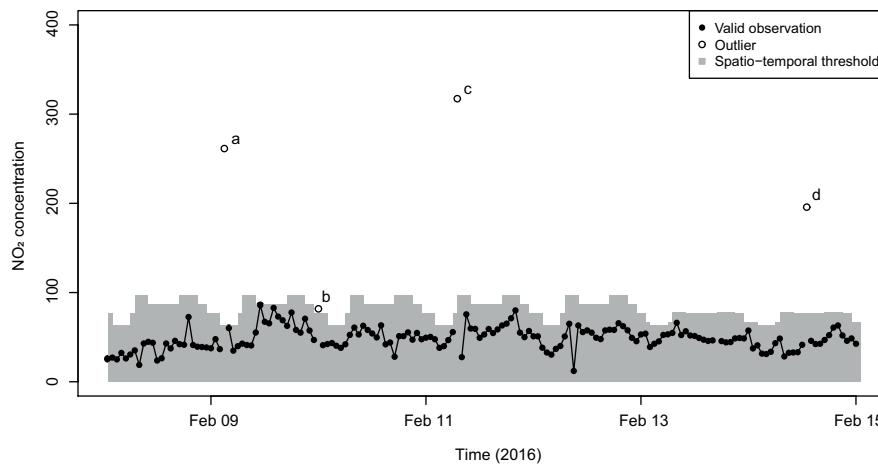


Figure 3.6 NO₂ concentrations ($\mu\text{g m}^{-3}$) measured by airbox 26, at an urban traffic location. Filled circles indicate non-outlying observations; unfilled circles indicate outliers using $\zeta = 2.97$. The gray bars indicate the threshold values for each temporal class, for urban traffic airboxes.

3.4. Results

threshold. However, during the daytime such a concentration value would have been within expected concentration values.

There was seasonal deviation in the number of outliers: a higher number of outliers was detected in spring (0.37%) compared to the mean percentage of outliers of the entire year (0.22%). In summer, the number of outliers was relatively low (0.09%).

Table 3.2 shows no difference in the percentage of outliers between urban traffic locations and urban background locations. Some individual airboxes however show more outliers than others. Most airboxes have 0–0.1% outliers for a year of data, whereas a few airboxes have a larger percentage of outliers for some spatio-temporal classes, up to a maximum of 2.5% for one airbox for one spatio-temporal class. The highest percentages of outliers are found in airboxes with the highest mean concentration values. The percentage of outliers of an airbox varies between spatio-temporal classes.

Similar results were found using hourly NO₂ observations of 2016 from the two conventional monitors. The total number of outliers detected was 0.3% of the dataset, which varied from 0–0.7% depending on the temporal class. In Figure 3.7 we observe a different pattern in the spatio-temporal thresholds compared to the threshold pattern of the airboxes (Figures 3.5 and 3.6). Note that for the conventional monitors we also observe positive lower threshold values, though close to zero. In Figure 3.7 we identify one outlier, which occurred in the off-peak hours period after the evening rush hour. This period after the evening rush hour is the period in which most outliers occurred for the conventional monitors.

We compared the outliers in the traffic airboxes with the NO₂ concentrations measured with the conventional monitors at the same time. A scatterplot is shown in Figure 3.8. The plot shows many observations down-right in the plot that have similarly high concentrations measured by the airbox and the conventional monitor, though at different locations. Some outliers occurred in multiple airboxes at the same time. This may be an indication of a pollution event that has an effect on the entire city. Down-left in the plot we find observations that are considered outliers by the airboxes, but are within normal range of concentrations according to the conventional monitors. These could be errors or very local air pollution events. In the upper part of the plot we find very high concentrations measured by the airbox which are higher than any value measured by the conventional monitor in the entire year. These are most likely errors.

3. Outlier detection in urban air quality sensor networks

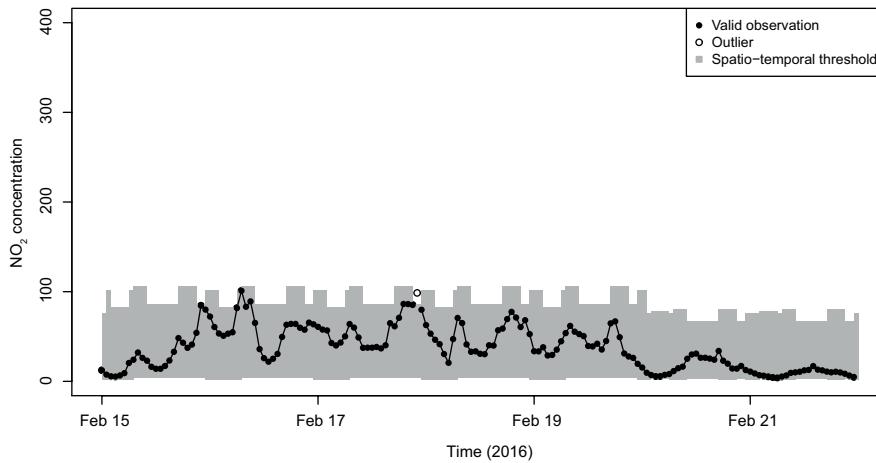


Figure 3.7 NO₂ concentrations ($\mu\text{g m}^{-3}$) measured by a conventional monitor at an urban traffic location. Filled circles indicate non-outlying observations; unfilled circles indicate outliers using $\zeta = 2.97$. The gray bars indicate the threshold values for each temporal class, for urban traffic conventional monitors.

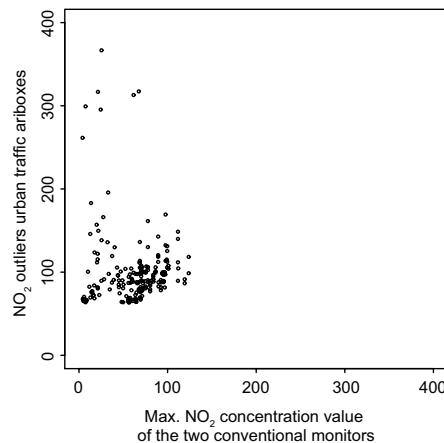


Figure 3.8 Scatterplot of traffic airbox outliers vs. the maximum NO₂ concentration ($\mu\text{g m}^{-3}$) measured at the same moment in time by the two conventional monitors located in traffic sites.

3.5 Discussion

The results show that the spatio-temporal classification of NO₂ concentration values in an urban sensor network is a simple outlier detection method in an area with high spatial and temporal variability of air pollutant concentrations. The number of outliers detected using the classification (0.1-0.5% for the airboxes and 0-0.7% for the conventional monitors) matches expectation when using $\zeta = 2.97$ as a threshold for

the number of standard deviations, including 99.7% of the observations under the assumption of a normal distribution. The value of ζ can be tuned depending on the application. A lower value of ζ will result in more concentration values to be considered outliers. Brown and Brown (2012) suggest that the choice of the threshold value should be a trade-off between the extra work associated with investigating false positives, i.e. observations falsely detected as outliers, and the likelihood of false negatives, i.e. true outliers that are not detected.

We aimed to compare the above procedure with kriging-based outlier detection (Zhang et al., 2012). We found that the NO₂ concentrations vary over shorter distances than the distances between measurement locations, resulting in a pure noise variogram. Sampling NO₂ over shorter distances, e.g. within a few meters, might make it possible to apply kriging-based outlier detection methods, especially when including covariates such as road distance and wind direction into the model.

Air pollutant concentrations are generally considered lognormally distributed (Ott, 1990). Applying the proposed outlier detection method on log-transformed NO₂ concentrations would however result in an implausible number of outliers detected on the left side of the distribution (99.5%) compared to the right side of the distribution (0.5%). Instead, we are mostly interested in high peaks in the data, which can be used to detect air pollution events and errors. Therefore, we used a square root transformation of the NO₂ concentration data.

The temporal classification used in this analysis is mostly based on expected traffic during certain hours of the day. Other factors that may influence the temporal variability in NO₂ concentrations are meteorological factors such as wind speed, wind direction, air pressure, temperature and solar radiation. An analysis of seasonal and diurnal variation at a UK city is presented by Bigi and Harrison (2010). NO₂ concentrations in Europe tend to be higher in the winter than in the summer season. Hence observations in the summer season had a lower chance to be detected as outliers by our method. Our method can be expanded by defining more classes, for example taking into account season and meteorological factors, or by taking into account temporal autocorrelation. For simplicity reasons we used full year data for the current study.

Public holidays occurring on a weekday are classified as weekdays, although the concentrations are likely lower, and therefore more similar to weekend concentrations. A visual analysis of the data showed that there was no increase in low-peak outliers during such holidays. High-peak outliers occurred and were also detected during the weekday holidays.

3. Outlier detection in urban air quality sensor networks

In this study, we aggregated the NO₂ concentrations to hourly values. Using 10-minute data, the outlier detection method would give more detailed instances of outliers compared to using hourly data. The results of 10-minute outlier detection should be interpreted differently from the results of hourly outlier detection. In hourly outlier detection, peaks occurring as a result of a strongly emitting vehicle passing by are more likely to be averaged out as they may occur every hour. In 10-minute data, such peaks are more likely to be considered outliers. Hourly outliers give a better overview of hours in which there is an abnormal number of peaks rather than showing individual peaks, as in the case of 10-minute outlier detection.

For the conventional monitors, the largest number of outliers was found during the off-peak period after the evening rush hours. Comparing the daily threshold pattern of the airbox to that of the conventional monitor on a weekday (Figure 3.9), both at an urban traffic location, we see that the upper threshold of the airbox in off-peak periods ($87.3 \mu\text{g m}^{-3}$) lays between the upper threshold of rush hours ($96.6 \mu\text{g m}^{-3}$) and the upper threshold of transition periods ($76.5 \mu\text{g m}^{-3}$). For the conventional monitor, the upper threshold for off-peak periods ($86.4 \mu\text{g m}^{-3}$) is below the threshold for both rush hours ($106 \mu\text{g m}^{-3}$) and transition periods ($101.6 \mu\text{g m}^{-3}$). The threshold for off-peak periods is calculated using the observations between morning rush hour and evening rush hour ((9:01-16:00 UTC time)) combined with the observations after evening rush hour (20:01-22:00 UTC time). For the airboxes this is alright because the concentrations are within a similar range. The conventional monitors, however, still measure high concentrations for two hours after the evening rush hour. This leads to underestimation of the threshold after evening rush hour. The cause of this difference is unclear, but most likely it is caused by differences between the sensor system of the airbox and the conventional monitor, and could be solved by defining different temporal classes depending upon the measurement instrument used.

The spatial classification method has been applied to the city of Eindhoven, the Netherlands. The spatio-temporal variability of NO₂ concentrations in this city is determined mainly by road traffic, like in many European cities (Cyrys et al., 2012). The spatial classification used in this analysis, distinguishing between urban background locations and urban traffic locations, is based upon this spatial variability. In Asian cities where, for example, industry plays a major role in the spatio-temporal variability of NO₂ concentrations (Cui et al., 2016), other classifications may be more relevant.

The proposed method for outlier detection using a spatio-temporal classification of the NO₂ variability was found useful for distinguishing outliers in an area with high spatial and temporal variability of air pollutant concentrations. This provides a basis for future work on distin-

guishing between types of outliers, e.g. errors and events. Air pollution events are often characterized by lasting for a period of time, which would lead to a number of outliers in a row for the same sensor. Such events can also be characterized by covering a large area in space. The occurrence of outliers at multiple locations at the same moment may indicate such an event.

The method provides a useful outlier detection method for those involved in urban air quality sensor networks. Its use in other fields of environmental variables with a high spatial and temporal variability is to be further investigated and will largely depend on the ability to classify the observations in various spatial and temporal categories.

Future research is needed in order to deal with the application of this method for (near) real-time outlier detection, in which each new observation can be compared to previous observations in the same spatio-temporal class. By using a moving average over the last hour, applied every 10 minutes, the method can be applied to (near) real-time data. Its applicability is currently mostly limited by the computation time, which is too long for real-time outlier detection. This may in the future be improved by using higher computation power or smaller datasets, or a combination of these two.

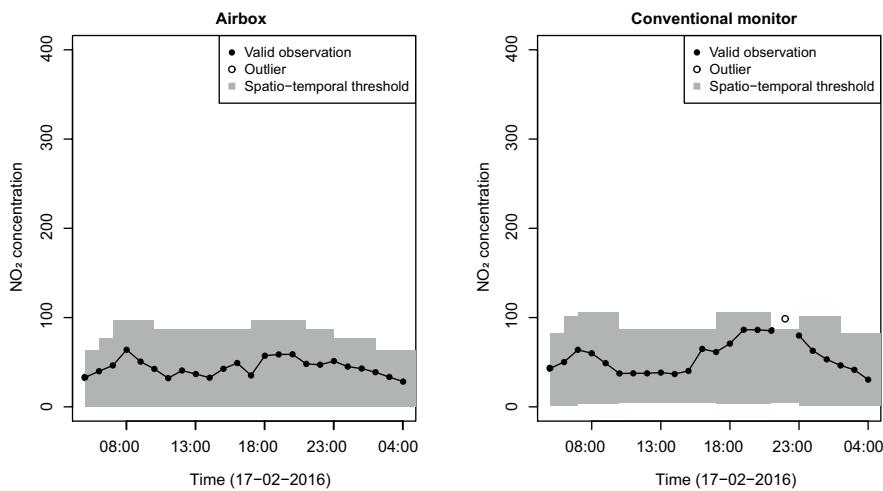


Figure 3.9 Comparison of NO_2 concentrations ($\mu\text{g m}^{-3}$) measured by an airbox (left) and a conventional monitor (right) on a weekday at urban traffic locations. Filled circles indicate non-outlying observations; unfilled circles indicate outliers using $\zeta = 2.97$. The gray bars indicate threshold values for each temporal class and are specific for each dataset, characterized by a spatial class and measurement instrument.

3. Outlier detection in urban air quality sensor networks

3.6 Conclusions

We presented a novel method for outlier detection in urban air quality sensor networks, based on dividing the observations in two spatial and eight temporal classes. Each of the sixteen resulting spatio-temporal classes represents a range of typical air pollutant concentrations for this class. By finding outliers in each class separately, the spatio-temporal variability in concentrations is maintained. In doing so, this work addressed an important challenge in outlier detection in urban areas.

In our analysis using hourly NO₂ data from an air quality sensor network in Eindhoven, the Netherlands, we detected 0.1-0.5% of outliers using a 99.7% confidence interval. The size of the confidence interval can be changed depending on the application. The non-normality of air pollutant concentrations is taken into account by using a truncated normal distribution of square-root transformed concentrations. The method is easy to implement and simple to adjust to other cities and pollutants by choosing spatio-temporal classes based on the sources of the air pollutants.

This research is a first step in outlier detection of NO₂ concentrations in urban areas. The detected outliers are unusually high concentrations, which can be either errors or events. Expert knowledge is however required to evaluate each outlier and decide on its treatment. Further research is needed with a focus on automatically distinguishing errors from events and (near) real-time outlier detection.

Appendix

Table A3.1 Mean n_K (\pm standard deviation t_K) of the distribution underlying the truncated normal distribution of each spatio-temporal class. Note that these are not actual concentration values, but transformed concentration values. Threshold values are obtained by $n_K \pm \zeta \times t_K$ and back-transformation of the resulting values using Eq. 3.9.

	Urban traffic		Urban background	
	Week	Weekend	Week	Weekend
Rush hours	5.26 (± 1.55)	4.54 (± 1.47)	4.67 (± 1.48)	4.05 (± 1.31)
Off-peak hours	5.02 (± 1.47)	4.50 (± 1.45)	4.46 (± 1.39)	4.01 (± 1.31)
Night hours	4.30 (± 1.25)	4.22 (± 1.28)	3.97 (± 1.27)	3.83 (± 1.28)
Transition hours	4.76 (± 1.36)	4.39 (± 1.30)	4.33 (± 1.34)	3.99 (± 1.21)

Calibration of NO₂ sensors in an urban air quality network

4

Abstract

Low-cost air quality sensors measuring air quality at fine spatio-temporal resolutions, typically suffer from sensor drift and interference. Field calibration is typically performed at one location, while little is known about the spatial transferability of correction factors. We evaluated three calibration methods using a year of hourly NO₂ observations from low-cost sensors, collocated at two sites with a conventional monitor as reference: (1) an iterative Bayesian approach for daily estimation of the parameters in a multiple linear regression model, (2) a daily updated correction factor and (3) a correction factor updated only when concentrations are uniformly low. We compared the performance of the calibration methods in terms of temporal stability, spatial transferability, and sensor specificity. We documented drift within the 1-year period. The correction factor updated under uniformly low concentrations performed poorly. The iterative Bayesian approach and daily correction factor reduced the root mean squared error (RMSE) by 21-46% at the calibration locations, but did not reduce RMSE at the other location. By examining the posterior distributions of the regression coefficients, we found that the poor spatial transferability is consistent with different responses of individual sensors to environmental factors. We conclude that the spatial and temporal variability in the calibration parameters requires them to be updated regularly, including sensor-specific recalibrations.

This chapter is published as: Van Zoest, V., Osei, F.B., Stein, A., Hoek, G., 2019. Calibration of low-cost NO₂ sensors in an urban air quality network. *Atmospheric Environment* 210, 66-75. doi:10.1016/j.atmosenv.2019.04.048.

4. Calibration of NO₂ sensors in an urban air quality network

4.1 Introduction

The interest in the use of low-cost air quality sensors at the city level to extend and densify conventional air quality monitoring is increasing (Snyder et al., 2013; Jerrett et al., 2017). A dense air quality monitoring network can be used for modeling and mapping air quality on a finer resolution in space and time than conventional monitoring networks that typically include only one or two monitors per city (Schneider et al., 2017). Such models and maps may be of added value e.g., for health studies and policy making. The data quality of low-cost air quality sensor networks however is often poor or unknown, potentially leading to wrong policy decisions or bias when applied in epidemiological studies (Snyder et al., 2013). In order to adjust, correct or improve sensor observations, calibration is important in the development and maintenance of such networks (Lewis and Edwards, 2016).

Calibration starts in the lab with finding an optimal function to convert absorbance or conductivity to pollutant concentrations (Neri et al., 2002). Performance indicators include sensitivity, selectivity, stability, response time, saturation, sensitivity to humidity, and the limit of detection (LoD) (Santos et al., 1997; Colin et al., 1998; Penza et al., 1998; Morales et al., 2002). Lab calibration, however, is not sufficient for field deployment of air quality sensors. The sensors perform less accurately under changing weather conditions and when exposed to different mixtures of gases as compared to calibration in the lab (Kamionka et al., 2006; De Vito et al., 2009). For long-term performance in the field, there are two main challenges: drift and interference effects.

The electrochemical cells typically used in low-cost gas sensors are more prone to lose sensitivity as compared to conventional monitors. This leads to sensor drift: an increasing bias in the sensor response. Xiang et al. (2016) showed that the measurement error due to drift in NO₂ sensors increased by a factor three within two months. Low-cost air quality sensors are also sensitive to relative concentration distribution changes, for example caused by seasonality or pollution events (De Vito et al., 2009; Moltchanov et al., 2015). The gas sensor can show an unwanted response to other pollutants, gases, temperature and relative humidity (Kamionka et al., 2006; Mead et al., 2013). We refer to these unwanted responses as interference effects. Field calibration is needed to estimate the parameters which correct the sensor response such that it represents the true concentrations as good as possible, accounting for drift and interference.

Much research has already been done to find the best methods for field calibration of air quality sensors. Multiple linear regression (MLR) has been used in the past few years to build calibration functions in which covariates account for environmental and meteorological vari-

ability (Piedrahita et al., 2014; Spinelle et al., 2015; Kizel et al., 2018). Feed-forward artificial neural networks (ANNs) have also been used to adjust for interference effects (De Vito et al., 2009). Xiang et al. (2016) also accounted for multiple sensors drifting simultaneously.

In air quality sensor networks, one commonly used approach is that only one sensor is collocated with a reference monitor during regular operation. Before operation and in regular intervals during operation, sensors may be collocated to be calibrated or recalibrated. In addition, some studies report automatic calibration procedures, aiming at reducing calibration costs. When recalibrating the sensors in a sensor network to account for drift and varying meteorological conditions in this way, it is assumed that all sensors behave similarly, and are influenced by the same conditions and similar drift. Although there is evidence that this assumption is not valid (Barakeh et al., 2016), calibration parameters established at a location with a reference instrument are often assumed to be transferable to other locations in the network.

The aim of this study is to evaluate the spatio-temporal variability and sensor specificity of calibration parameters. To do so, we evaluate three different calibration methods that account for interference effects and drift. The methodology is applied to the city of Eindhoven, where a low-cost sensor network was set up with NO₂ measurements at 25 locations (Hamm et al., 2016). At two locations the low-cost sensors were co-located with reference monitors, allowing us to assess spatial and temporal variability of correction factors. The focus of this study is on NO₂, because earlier passive sampler measurements have documented large spatial variation in European cities related especially to road traffic (Cyrys et al., 2012).

4.2 Methods

We evaluated the performance of (1) an iterative Bayesian approach for daily estimation of the parameters in a multiple linear regression model, (2) a daily updated simple correction factor and (3) a uniform concentration correction similar to the method suggested in Tsujita et al. (2005). For Bayesian inference we used Integrated Nested Laplace Approximations (INLA). We evaluated the methods in terms of the stability of the correction factors or calibration parameters in time, the usability of the correction factors or calibration parameters at other locations within the same urban area, and the transferability of the calibration parameters to different sensors of the same type.

4.2.1 Data

The measurement sites of the ILM sensor network were purposely selected reflecting background and traffic sites (Figure 4.1). Background sites

4. Calibration of NO₂ sensors in an urban air quality network

were located in minor residential roads; traffic sites in roads with typically more than 10000 vehicles/day. The two conventional monitors sites in Eindhoven are also designated as traffic sites in the Dutch national air quality measurement network. The focus of this study is on NO₂ which is measured in 25 airboxes of the ILM network since 2016, after an initial calibration at the end of 2015. More details about the sensors can be found in Chapter 2.

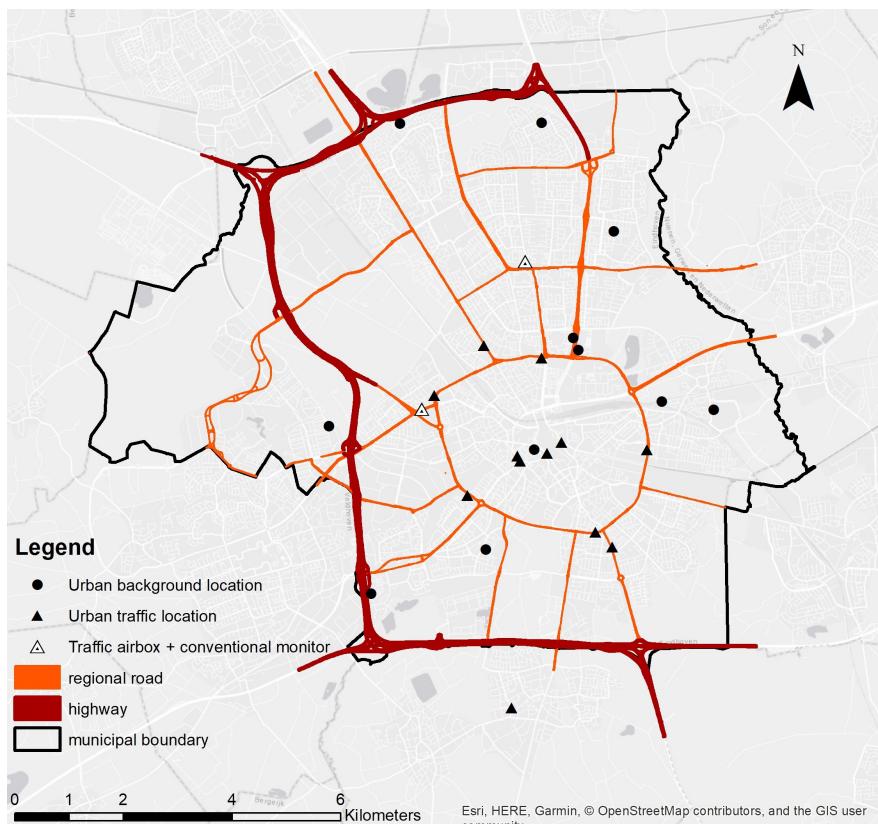


Figure 4.1 Locations of the airboxes and conventional monitors in Eindhoven

Two conventional chemiluminescence monitors of the Dutch national air quality measurement network (Buijsman, 2013), operated by the RIVM, are both located in similar urban traffic locations in Eindhoven and are used as reference instruments. With each conventional monitor, an airbox is collocated (Figure 4.1). Airbox NO₂ measurements are averaged to hourly values for analysis of the calibration methods, similar to the temporal resolution of the conventional monitors. Data cleaning and outlier detection were performed as described in Chapter 3. We refer to $S = \{s_1, \dots, s_{25}\}$ as the collection of airboxes measuring NO₂ and $Z = \{z_1, z_2\}$ as the collection of conventional monitors, where $Z \subset S$.

In this study we use covariates $c \in C = \{NO_2, O_3, RH, T, WS, WD\}$. Nitrogen dioxide (NO_2), relative humidity (RH) and temperature (T) are measured within the airbox. Ozone (O_3) data in the airboxes at z_1 and z_2 were missing for most of the year due to sensor failure. Therefore we used the O_3 data available at the conventional monitor at location z_1 (O_3 not measured at z_2). Wind speed (WS) and wind direction (WD) were obtained from the Royal Netherlands Meteorological Institute weather station in Eindhoven (KNMI, 2016). For the multiple linear regression model, NO_2 concentrations were square root transformed to approximate a normal distribution. Where needed, the covariates were also transformed to obtain distributions closer to the normal distribution and to obtain a more linear relationship between the covariate and square root transformed NO_2 concentrations. An overview of the potential covariates, their sources and transformations is given in Table 4.1.

Table 4.1 Overview of potential covariates for the calibration model

Covariate	Units	Source	Instrument/method	Transformation
NO_2	$\mu g m^{-3}$	Airbox	Citytech Sensoric NO_2 3E50 ECN revised	square root
O_3	$\mu g m^{-3}$	RIVM	Ultraviolet photometry	log
Temperature (T)	$^{\circ}C$	Airbox	Sensirion SHT75	none
Relative humidity (RH)	%	Airbox	Sensirion SHT75	squared
Wind speed (WS)	$m s^{-1}$	KNMI	Automatic Weather Station (AWS)	square root
Wind direction (WD)	degrees	KNMI	Automatic Weather Station (AWS)	categorized into nine categories: 8 wind rose directions and 1 class calm/variable

4.2.2 Sensor drift

Sensor drift is caused by the loss of sensitivity of the electrochemical cell measuring NO_2 . We examine the average drift of the sensor network by plotting a time series of the difference $\Delta_{d,t}$ between the mean NO_2 concentration of the two conventional monitors and the mean NO_2 concentration observed by all airboxes deployed in the city, for all hourly observations in 2016:

$$\Delta_{d,t} = \frac{\sum_s (x_{NO2,d,t,s})}{N_s} - \frac{\sum_z (y_{NO2,d,t,z})}{N_z} \quad (4.1)$$

where $x_{NO2,d,t,s}$ is the hourly NO_2 concentration measured at day d and hour $t = 1, \dots, 24$ at airbox location s for N_s number of airboxes, $y_{NO2,d,t,z}$ is the hourly NO_2 concentration measured at day d and hour $t = 1, \dots, 24$ at conventional monitor location z for N_z number of conventional monitors. For calculation of $\Delta_{d,t}$ we use the actual NO_2 observations before square root transformation. We take a smoothed line

4. Calibration of NO₂ sensors in an urban air quality network

through the time series of $\Delta_{d,t}$ to largely separate the effect of drift from temporally varying spatial variability.

4.2.3 Multiple linear regression model

Calibration

The term calibration refers to two processes. First, it establishes a relationship between indicative measurements and standard (reference) measurements, i.e., estimating the parameters of the calibration function; second, it uses an established relationship, i.e., the calibration function, for obtaining a measurement result from an indicative measurement (Rasch et al., 1994). In this paper, our focus is on the first process, and we will use the term correction for the second process.

MLR functions have been widely used to build calibration functions accounting for environmental and meteorological variables (Piedrahita et al., 2014; Spinelle et al., 2015). We adapted the method, (1) to allow for transformations of the response variable and covariates to obtain better linear relationships, (2) to estimate the calibration parameters including their uncertainty using Bayesian inference, and (3) to iteratively update the calibration parameters on a daily basis using observations of the previous 30 days.

The calibration function resembles a generalized additive model (GAM):

$$y_{d,t,z} = \beta_{0,d,z} + \sum_c \beta_{c,d,z} g_c(x_{c,d,t,z}) + \varepsilon_{d,t,z} \quad (4.2)$$

where $y_{d,t,z}$ is the square root transformed reference NO₂ level at day d , hour $t = 1, \dots, 24$ and location z , $g_c(x_{c,d,t,z})$ are covariate-dependent known functions or transformations applied to covariate $x_{c,d,t,z}$, $\beta_{0,d,z}$ is the intercept and $\beta_{c,d,z}$ are the unknown coefficients for covariates c for day d at location z and the error is assumed $\varepsilon_{d,t,z} \sim N(0, \sigma^2)$. The coefficients $\beta_{c,d,z}$ have a posterior distribution with mean $\mu_{\beta_{c,d,z}}$ and precision $\tau_{\beta_{c,d,z}}$. The covariate-dependent transformations $g_c(x_{c,d,t,z})$ are chosen such that $x_{c,d,t,z}$ approximates a normal distribution, and $g_c(x_{c,d,t,z})$ has an approximately linear relation with $y_{d,t,z}$.

We used hierarchical Bayesian estimation and inference. Bayesian inference provides a posterior distribution for each $\beta_{c,d,z}$ rather than a single estimate, and therefore allows for the comparison of estimates of different airboxes, including their uncertainty. The parameters of the posterior distributions were estimated using Integrated Nested Laplace Approximation (INLA). INLA provides fast and accurate Bayesian parameter estimates through Laplace approximations. The advantage of INLA

over Markov Chain Monte Carlo (MCMC) simulations is that the computation time is significantly shorter, while INLA gives an approximation which is as good or better (Rue et al., 2009).

We built a set of multiple linear regression models including combinations of covariates which are often used in calibration of NO₂ sensors, as they are known for causing interference effects or sensor bias. All covariates are scaled and centered to zero before running INLA. The calibration equation is updated every day d to account for gradual drift and meteorological variability using all non-missing observations in the previous 30 days, adding up to maximum 720 hourly observations. Moltchanov et al. (2015) found variability in calibration parameters over shorter periods of time, e.g. 4 days, using 30-minute averages. We chose a period of 30 days based on prior analyses (Table A4.1), aiming to include enough variation in air pollutant concentrations while minimizing the longer term effects of drift and seasonality. We further note that our calibration parameters change daily in a smooth manner. Our method does not distinguish between day and night periods in calibration parameters. Using R-INLA (Martins et al., 2013), we built a model on the calibration set to estimate the parameters of the posterior distribution $\theta_{c,d,z} = (\mu_{\beta_{c,d,z}}, \tau_{\beta_{c,d,z}})$ of the coefficients $\beta_{c,d,z}$. The models are built at the two locations z_1 and z_2 where a conventional monitor is collocated with an airbox. The model is rebuilt for every day in 2016, such that there is an overlap of 29 days between the data used for calibration on day d and for calibration on the next day $d+1$.

Calibration performance measures

The fit of a Bayesian model is commonly evaluated using posterior predictive checks or leave-one-out cross-validation. For the first check the posterior predictive p-values,

$$p(\gamma_{d,t}^* \leq \gamma_{d,t} | \boldsymbol{\gamma}) \quad (4.3)$$

for replicate observations $\gamma_{d,t}^*$, are evaluated to be uniformly distributed. For none of the models explored in this study the distribution of the posterior predictive p-values was uniformly distributed. Wang et al. (2018) however argue that in some cases the posterior predictive p-values can be affected by the nature of the data in such a way that they would never be uniformly distributed even in the case of a perfect model. Therefore they suggest using the probability integral transform (PIT) instead:

$$PIT_{d,t} = p(\gamma_{d,t}^* \leq \gamma_{d,t} | \boldsymbol{\gamma}_{-d,t}) \quad (4.4)$$

where $\boldsymbol{\gamma}_{-d,t}$ are all observations except for the observation at time stamp t on day d . The performance of the different calibration models is evaluated based on the Deviance Information Criterion (DIC), a generalization

4. Calibration of NO₂ sensors in an urban air quality network

of the Akaike Information Criterion (AIC), accounting for both model complexity and fit in a Bayesian model (Spiegelhalter et al., 2002):

$$DIC = \bar{D} + p_D \quad (4.5)$$

where \bar{D} is the posterior mean of the deviance and p_D is the effective number of parameters. A smaller DIC denotes a better fit.

Validation

For temporal validation, we use the calibration function from Eq. 4.2 and replace the unknown β by $\hat{\beta}$ to predict $y_{d,t,z}$:

$$\hat{y}_{d,t,z} = \hat{\beta}_{0,d,z} + \sum_c \hat{\beta}_{c,d,z} g_c(x_{c,d,t,z}) + \varepsilon_{d,t,z} \quad (4.6)$$

for which we now know the posterior distributions of $\hat{\beta}_{0,d,z}$ and $\hat{\beta}_{c,d,z}$. On every day d , Eq. 4.6 is applied on $t = 1 \dots 24$ using the parameters of $\hat{\beta}_{0,d,z}$ and $\hat{\beta}_{c,d,z}$ estimated during the calibration phase. The calibration and validation are repeated daily, so each 24 hour period is validated with a new set of calibrated parameters based on the hourly data available in the previous 30 days.

The number of locations with a collocated reference monitor is always sparse in low-cost air quality sensor networks. If the drift and the influence of external variables are similar for each airbox and location, a calibration model built at one location can be transferred to the other locations. To test this, we apply spatiotemporal validation by adjusting Eq. 4.6 to predict $y_{d,t,z}$ at a different location (z_i) from where the model is built (z_j):

$$\hat{y}_{d,t,z_i} = \hat{\beta}_{0,d,z_j} + \sum_c \hat{\beta}_{c,d,z_j} g_c(x_{c,d,t,z_j}) + \varepsilon_{d,t,z_i} \quad (4.7)$$

for the two locations where a conventional monitor is located. Note the different subscripts of z to denote the different locations used in spatiotemporal validation. Similar to the temporal validation, Eq. 4.7 is applied on $t = 1 \dots 24$ on the current day d using the parameters of $\hat{\beta}_{0,d,z}$ and $\hat{\beta}_{c,d,z}$ estimated during the calibration phase, but now at a different location z_j .

Validation performance measures

Prediction performance is based on the Root Mean Squared Error (RMSE). We consider two RMSE values: the RMSE before calibration ($RMSE_{pre}$) and the RMSE after calibration ($RMSE_{post}$). $RMSE_{pre}$ is obtained as:

$$RMSE_{pre} = \sqrt{\frac{\sum_{t=1}^{T_d} (y_{NO2,d,t,z} - x_{NO2,d,t,z})^2}{T_d}} \quad (4.8)$$

where $\gamma_{NO2,d,t}$ is the observed reference NO₂ concentration ($NO2_{ref}$) and $x_{NO2,d,t}$ is the observed airbox NO₂ concentration ($NO2_{ab}$) for time stamp t in $1, \dots, T_d$ where T_d is the total number of non-missing hours in validation day d . $RMSE_{post}$ is obtained as:

$$RMSE_{post} = \sqrt{\frac{\sum_{t=1}^{T_d} (\gamma_{NO2,d,t,z} - \hat{\gamma}_{NO2,d,t,z})^2}{T_d}} \quad (4.9)$$

where $\hat{\gamma}_{NO2,d,t}$ is the predicted and back-transformed NO₂ concentration. A smaller RMSE denotes better prediction.

The calibration approach using INLA is compared to two other techniques that are often used for correction of low-cost air quality sensor networks: (1) a simple daily updated correction factor and (2) a correction factor which is updated when the concentrations are low and uniform across the sensor network.

4.2.4 Daily correction factor

Miskell et al. (2018) calibrated low-cost O₃ sensors using a conventional monitor in the vicinity with similar land use type. This method assumes that the drift between the airboxes is similar and that high peaks missed by the airboxes are due to meteorological factors which are the same across the sensor network. We apply a similar method to NO₂, distinguishing between an absolute correction factor and a relative correction factor. On each day d we find the relative difference correction factor $\gamma_{rel,d,z}$:

$$\gamma_{rel,d,z} = \sum_{t=1}^{T_d} \left(\frac{\gamma_{NO2,d,t,z}}{x_{NO2,d,t,z}} \right) \times \frac{1}{T_d} \quad (4.10)$$

and the absolute difference correction factor $\gamma_{abs,d,z}$:

$$\gamma_{abs,d,z} = \frac{\sum_{t=1}^{T_d} (\gamma_{NO2,d,t,z} - x_{NO2,d,t,z})}{T_d} \quad (4.11)$$

The correction factors are computed on a location z where an airbox is collocated with a conventional monitor. On a daily basis, $\gamma_{rel,d,z}$ or $\gamma_{abs,d,z}$ corrects all hourly airbox measurements of that day, for all airboxes located at a similar site type (urban traffic or urban background). The spatial transferability of the correction factor is evaluated by applying it at the other airbox location in Z and comparing the corrected airbox NO₂ concentrations with the observations of the conventional monitor at that location. Since both conventional monitors are located at urban traffic locations, we could only evaluate the method for this site type. The RMSE is calculated before and after correction.

4. Calibration of NO₂ sensors in an urban air quality network

4.2.5 Uniform concentration correction

Tsujita et al. (2005) proposed a method for automatic calibration of low-cost air quality sensor networks. The method differs from the previously mentioned correction factor, in the sense that the correction factor is only updated under conditions of uniform low NO₂ concentrations. We tested a similar method. When NO₂ concentrations are uniform and low for any hourly timestamp, the baseline of the low-cost sensors is adjusted to the mean of the conventional monitors. All hourly NO₂ observations are corrected using a fixed correction factor γ_{uni} which is the same for all airboxes. This correction factor is based on the ratio between the mean of the two conventional monitors and the mean of all low-cost sensors. The correction factor is applied to all future observations of the low-cost sensors, until it is updated at time stamp d, t when standard deviations of the NO₂ concentrations drop below a threshold δ and the mean NO₂ concentrations drops below a threshold χ :

$$\sqrt{\frac{\sum_s (x_{NO2,d,t,s} - \bar{x}_{NO2,d,t})}{N_s - 1}} < \delta \wedge \frac{\sum_s (x_{NO2,d,t,s})}{N_s} < \chi \quad (4.12)$$

We evaluated different values of $\delta = \{4, 5, 6\} \mu\text{g m}^{-3}$ and values $\chi = \{10, 12, 15\} \mu\text{g m}^{-3}$.

Following Moltchanov et al. (2015) we also applied night-time calibrations. Every night the concentrations are assumed to be uniform, and a new airbox-specific correction factor $\gamma_{night,d,s}$ is retrieved from the ratio between each individual airbox and the average between the two conventional monitors between 1:00-4:00 a.m. The correction factor is used to correct the airbox values during the next day.

4.2.6 Sensitivity of individual airboxes to environmental factors

We evaluated whether individual airboxes reacted similarly to interfering gases, temperature and humidity. First, using INLA we examined the relationship between $NO2_{ab}$ and the other variables RH and T measured in the airbox. For the 25 airboxes the posterior mean estimates were compared based on slope direction and strength. The differences in posterior mean estimates between the 25 non-collocated airboxes reflect both spatial variability in the calibration parameters, as well as inter-sensor variability (Broday and The Citi-Sense Project Collaborators, 2017). Second, for a comparison independent of the airbox location, we compared the posterior mean estimates of ten airboxes which were simultaneously collocated with a conventional monitor for ten days (240 hourly observations per airbox). A separate model, with $NO2_{ref}$ as the response variable, was built using INLA for each covariate measured in the airbox: $NO2_{ab}$, RH , and T . The posterior mean estimates in this case solely reflect inter-sensor variability.

4.2.7 Temporal autocorrelation

A first-order random walk model is added to Eq. 4.2 to account for possible temporal autocorrelation, as suggested in Blangiardo and Cameletti (2015):

$$y_{d,t,z} = \beta_{0,d,z} + \sum_c \beta_{c,d,z} g_c(x_{c,d,t,z}) + rw_1(d, t) + \varepsilon_{d,t,z} \quad (4.13)$$

where $rw_1(d, t)$ is a first-order random walk function on the time series of d, t . For both locations z_1 and z_2 , and for both the model with and without random walk component, the $RMSE_{pre}$ and $RMSE_{post}$ are obtained. A lower $RMSE_{post}$ for the model with random walk component compared to the model without random walk component suggests the presence of temporal autocorrelation. The models in Eqs. 4.2 and 4.13 are also compared using a full year of data, for which 80% of all hourly observations in 2016 are randomly selected for calibration and the remaining 20% of the observations is used for validation.

4.3 Results

4.3.1 Descriptive statistics

Based on the initial calibration, the airbox showed good agreement with the conventional monitors. Pearson's correlation coefficient for the full year 2016 for hourly NO_2 concentrations is 0.75 at z_1 and 0.83 at z_2 . Scatterplots for both locations nevertheless showed substantial differences of individual observations (Figure 4.2).

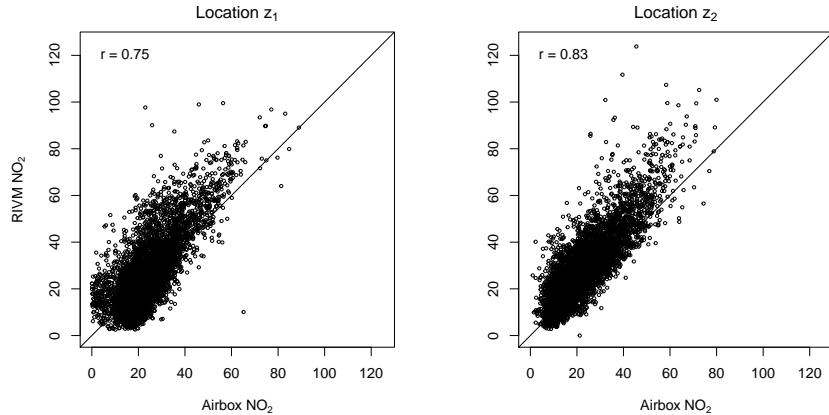


Figure 4.2 Scatterplots of hourly NO_2 values in 2016: airbox vs. conventional monitor at location z_1 and z_2 . Concentrations in $\mu\text{g m}^{-3}$.

The percentage of missing values in the dataset of 2016 was 23.1% for the airboxes, of which 4.4% was removed during data cleaning and outlier

4. Calibration of NO₂ sensors in an urban air quality network

detection (Chapter 3). The remainder is caused by sensor malfunctioning (6.6%) or displacement during maintenance periods (12.1%). To reduce the service costs of the sensor network, maintenance periods were long. The airbox located at z_1 was removed for maintenance from March 22 until May 3, 2016. The airbox at z_2 was removed from February 22 until May 3, 2016. In future development of the network, maintenance time clearly needs to be reduced. For the two conventional monitors, 2.8% of the hourly NO₂ observations was missing in 2016.

4.3.2 Sensor drift

Figure 4.3 shows the time series of $\Delta_{d,t}$, the difference between the mean NO₂ concentration of the two conventional monitors and the mean NO₂ concentration observed by all airboxes. Since all airboxes are at different locations and their mean concentrations are not necessarily equal to the mean of the conventional monitors which are both located at a traffic location, we do not require $\Delta_{d,t}$ to be zero. Neither do we require $\Delta_{d,t}$ to be stable throughout the year, as it could possibly vary with the seasonality of NO₂, meteorological conditions or interfering gases. However, the downward trend in $\Delta_{d,t}$ in Figure 4.3 suggests sensor drift. After four months of deployment, there is a sudden decrease in $\Delta_{d,t}$, which leads to a systematic decrease of $\sim 10 \mu\text{g m}^{-3}$. This is related to a change in the initial calibration factor around May 10. The bias further increases with time. Compared to the average drift of all airboxes in the sensor network, the two collocated airboxes do not show the same decreasing trend. At z_1 the fitted smooth curve of the difference is around zero after May 10, showing no signs of drift. At z_2 , the fitted smooth curve follows that of the average drift after May 10 until the end of August. Due to malfunctioning of the sensor there are no data available for the last month of the year. The high values at the end of November would therefore strongly influence the fitted curve to increase from September onwards. For a better comparison with z_1 and the mean of all airboxes, the line of z_2 is drawn until September.

4.3.3 Multiple linear regression model

Calibration

The histogram and Uniform Q-Q plot of the PIT values (Eq. 4.4) are created for every daily iteration of the models. A sample is visually inspected to check for uniformity; an example is shown in Figure A4.1 in the Appendix. The PIT values show uniformity for all models, which means that the models suitably fit the data. The model performance, based on the DIC, can thus be evaluated for the models. In Table 4.2 we present the DIC values for different models. Since the INLA model is iteratively rebuilt, giving a new DIC value every day, we report the mean DIC and median DIC for 2016. The lowest DIC, indicating the best model fit, is found for model 9 including all covariates. Model 8 (excluding wind

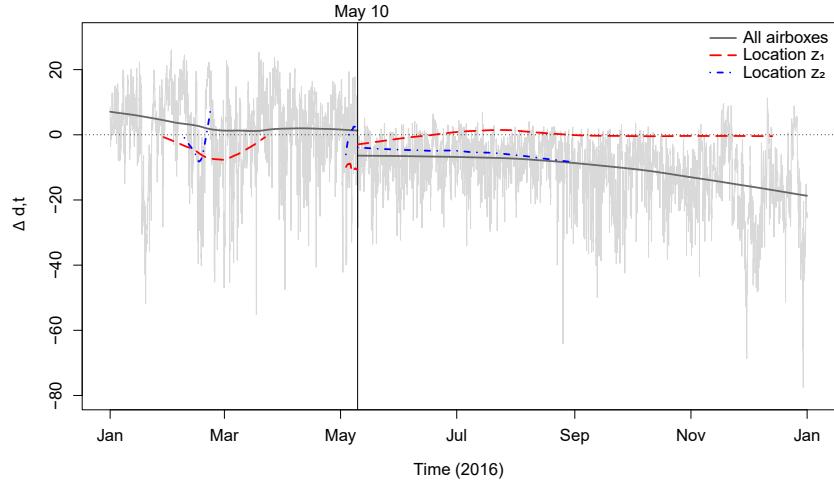


Figure 4.3 Difference between mean airboxes and mean conventional monitors over time (light gray line) and fitted smooth curves before and after May 10 (dark gray solid lines). Vertical solid black line: May 10, after which we observe a sudden decrease in $\Delta_{d,t}$. Red dashed and blue dot-dashed lines: fitted smooth curves of difference between conventional monitor and airbox at location z_1 and location z_2 , respectively. The line at z_2 has been drawn until September. Missing values in December would otherwise influence the line up towards the high levels in November, making comparison with z_1 and the mean of all airboxes impossible.

direction) has the next lowest DIC, only slightly higher than model 9. All models show a better fit at z_2 compared to z_1 .

Figure 4.4 shows the change of the coefficients of covariates over time when the model parameters are recalibrated on a daily basis for model 8. We show model 8 (without wind direction) because DIC values are similar to model 9 and wind direction is represented by 9 slopes, increasing the complexity of the figure. The intercept β_0 for the daily INLA models is positive between 3 and 7 $\mu\text{g m}^{-3}$. At z_1 , $\beta_{NO2} < 1$ and at z_2 , $\beta_{NO2} > 1$. Transferring the coefficients to another location where the bias is in a different direction will lead to an increase in bias rather than a decrease. Coefficient β_{O3} is negative throughout the year, and β_{RH} is close to zero. Coefficient β_T is mostly positive for both locations. Coefficient β_{WS} shows a pattern close to zero but mostly negative. The month of hourly data used in each calibration iteration should contain enough temporal variability in the covariates to avoid overfitting. However, both locations show a large temporal variability in the coefficients. This is probably due to seasonal variation, as the temporal variability at the two locations is very similar over time. The patterns are smoothed by the overlap of the calibration datasets. When the direction of the slopes would be the

4. Calibration of NO₂ sensors in an urban air quality network

Table 4.2 DIC performance statistics for different models. A lower DIC denotes better model fit. In each model, the dependent variable is the square root of hourly average reference monitor concentrations of NO₂, $\sqrt{NO_2}_{ref}$. NO₂_{ab} is NO₂ measured by the low-cost airbox sensor; O₃ is ozone measured by one reference monitor (z_1). RH is relative humidity and T is temperature, both measured by the airbox. WS is wind speed and WD is wind direction, both measured by the Royal Netherlands Meteorological Institute (KNMI)

#	Covariates	Location z_1		Location z_2	
		DIC mean	DIC median	DIC mean	DIC median
1	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab}$	1459	1630	1352	1464
2	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{O3} \log(O3)$	1291	1319	1328	1446
3	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{RH}RH^2$	1444	1619	1312	1439
4	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_T T$	1406	1506	1136	1242
5	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{WS} \sqrt{WS}$	1306	1348	1320	1421
6	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{WD} factor(WD)$	1403	1589	1292	1410
7	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{WS} \sqrt{WS} + \beta_{O3} \log(O3)$	1221	1243	1297	1409
8	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{O3} \log(O3) + \beta_{RH}RH^2 + \beta_T T + \beta_{WS} \sqrt{WS}$	1134	1208	828	877
9	$\beta_0 + \beta_{NO2}\sqrt{NO_2}_{ab} + \beta_{O3} \log(O3) + \beta_{RH}RH^2 + \beta_T T + \beta_{WS} \sqrt{WS} + \beta_{WD} factor(WD)$	1104	1161	778	815

same for each location, this would be beneficial for the transferability of the model from one location to the other. However, when coefficients tend to have a different direction at different locations at any point in time, correction may lead to a deterioration.

Validation

A model with only NO₂ improved the RMSE modestly at both locations (Table 4.3). Adding additional covariates substantially further reduced RMSE. Model 8, with all covariates except wind direction, performed best at location z_1 . At z_2 , model 9 with all covariates performed best. At both locations, no improvement in RMSE was achieved by correcting the observations using the calibration models built at the other location.

4.3.4 Daily correction factor

The relative correction factor $\gamma_{rel,d,z}$ shows a higher $RMSE_{post}$ than the absolute correction factor $\gamma_{abs,d,z}$, even exceeding $RMSE_{pre}$ at z_1 (Table 4.4). Since this method could be applied to all non-missing NO₂ observations at each location, while the Bayesian models could only be applied to observations non-missing for all covariates at each location, the $RMSE_{pre}$ and $RMSE_{post}$ are not directly comparable to those retrieved using INLA (Table 4.3).

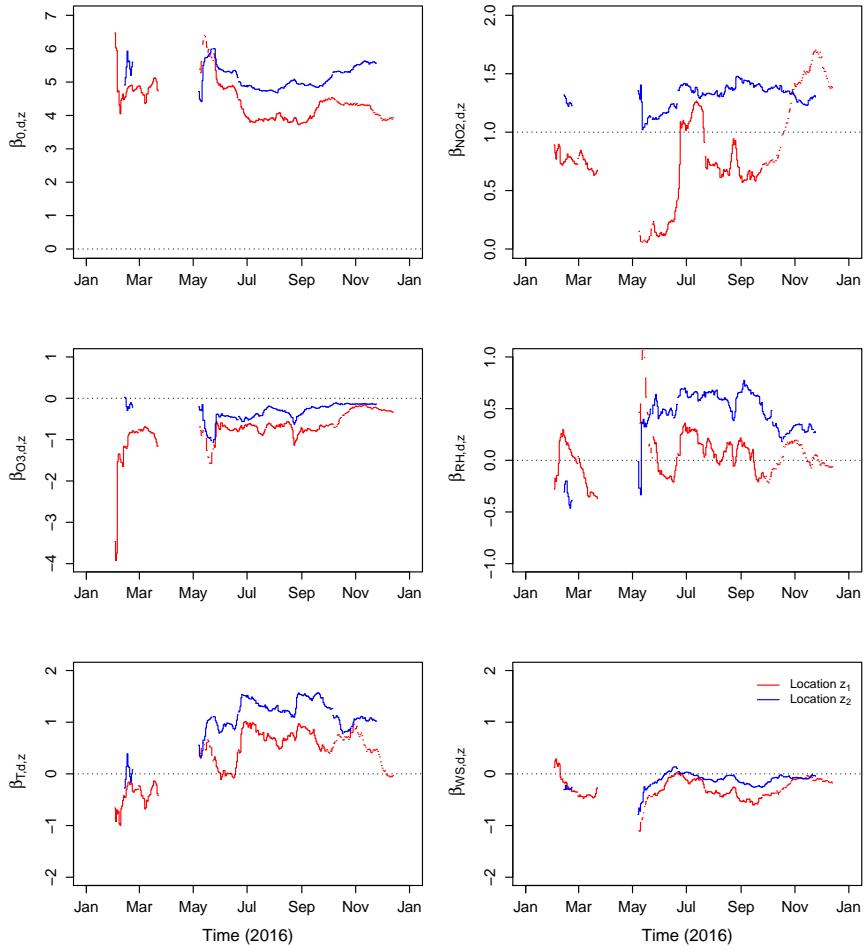


Figure 4.4 Time series of the coefficients of the daily INLA models, using model 8. Red: location z_1 , blue: location z_2 .

Table 4.4 also shows the $RMSE_{pre}$ and $RMSE_{post}$ when we tested the correction factor determined at the other airbox location collocated with a conventional monitor. At z_1 , RMSE is higher after correction; at z_2 a modest decrease was found using γ_{abs,d,z_1} . The RMSE values can be influenced by a few extremes in the corrected values, especially when an extreme correction factor is established at one location and is transferred to another location. This led to a high $RMSE_{post}$ of $120.86 \mu\text{g m}^{-3}$ at z_2 . Removing extreme correction factors led to a decrease of this value, however not decreasing below $RMSE_{pre}$. A time series plot of the correction factors illustrates its variability and the extremes (Figure A4.2).

4. Calibration of NO₂ sensors in an urban air quality network

Table 4.3 RMSE before and after temporal and spatiotemporal calibration using different models. Covariates in each model are similar to those in Table 4.2.

#	Temporal calibration				Spatiotemporal calibration			
	Location z_1 (# obs.=5684)		Location z_2 (# obs.=4816)		Model location z_2 validated at z_1 (# obs.=4751)		Model location z_1 validated at z_2 (# obs.=4885)	
	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
1	9.91	9.44	10.67	8.54	10.00	11.35	10.64	11.39
2	9.91	8.43	10.67	8.43	10.00	10.81	10.64	11.68
3	9.91	9.52	10.67	8.22	10.00	11.86	10.64	12.70
4	9.91	8.83	10.67	7.26	10.00	12.70	10.64	11.62
5	9.91	8.05	10.67	8.28	10.00	10.29	10.64	11.10
6	9.91	9.47	10.67	8.62	10.00	11.55	10.64	11.56
7	9.91	7.97	10.67	8.23	10.00	10.27	10.64	11.42
8	9.91	7.62	10.67	5.80	10.00	11.28	10.64	11.55
9	9.91	7.80	10.67	5.74	10.00	11.66	10.64	11.58

Table 4.4 RMSE values before and after applying a daily correction factor on hourly values at the same location (temporal calibration) and at the other location (spatiotemporal calibration)

Correction factor	Temporal calibration				Spatiotemporal calibration			
	Location z_1		Location z_2		Model location z_2 validated at z_1		Model location z_1 validated at z_2	
	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
$\gamma_{rel,d,z}$	9.94	19.57	10.55	7.34	9.46	13.53	10.68	120.86
$\gamma_{abs,d,z}$	9.94	6.54	10.55	5.78	9.46	10.17	10.68	9.75

4.3.5 Uniform concentration correction

The results of correction factor γ_{uni} updated under conditions of uniform and low concentrations are shown in Table 4.5. Depending on the threshold values of standard deviation δ and mean χ , the number of updates of γ_{uni} in the year ranged between 1 and 39 for the chosen thresholds. For none of the threshold combinations, the correction method improved the RMSE value. Instead, the $RMSE_{post}$ was 27-145% higher than the $RMSE_{pre}$, making the method not suitable for NO₂ in this sensor network. Figure A4.3 shows the time series of the correction factor for $\delta = 5 \mu\text{g m}^{-3}$ and $\chi = 12 \mu\text{g m}^{-3}$. As Figure 4.3 suggested, a change occurred in May. This is reflected in an update of the correction

factor in Figure A4.3. Before May the correction factor remains at 1.

Table 4.5 RMSE values before and after updating the correction factor γ_{uni} under conditions of uniform and low NO₂ concentrations

δ	χ	# updates	Location z_1		Location z_2	
			$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
4	10	1	9.94	24.33	10.55	19.73
4	12	1	9.94	24.33	10.55	19.73
4	15	2	9.94	17.67	10.55	13.44
5	10	1	9.94	15.97	10.55	12.33
5	12	16	9.94	18.42	10.55	14.60
5	15	36	9.94	23.01	10.55	19.58
6	10	8	9.94	20.00	10.55	15.91
6	12	39	9.94	22.12	10.55	18.33
6	15	39	9.94	22.96	10.55	19.58

Table 4.6 shows the results of the night-time calibration. At z_1 the RMSE is almost doubled after night-time calibration, while at z_2 there is a slight improvement in RMSE. The increase in RMSE at z_1 is mainly due to some extreme values for $\gamma_{night,d,s}$ in May (Figure A4.4).

Table 4.6 RMSE before and after night-time calibration

Location z_1		Location z_2	
$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
9.94	19.51	10.55	9.86

4.3.6 Sensitivity of individual airboxes to environmental factors

Figure 4.5 shows a boxplot of the posterior mean estimates for relative humidity and temperature for a model with airbox NO₂ as response variable. When using a full year of data at once, the posterior mean estimate for temperature is negative for each airbox, varying between -0.1 and -0.7 for different airboxes. The posterior mean of relative humidity differed between the different airboxes from -0.5 to +0.3. A difference in slope direction can have large influence on the transferability of calibration models to other airboxes.

4. Calibration of NO₂ sensors in an urban air quality network

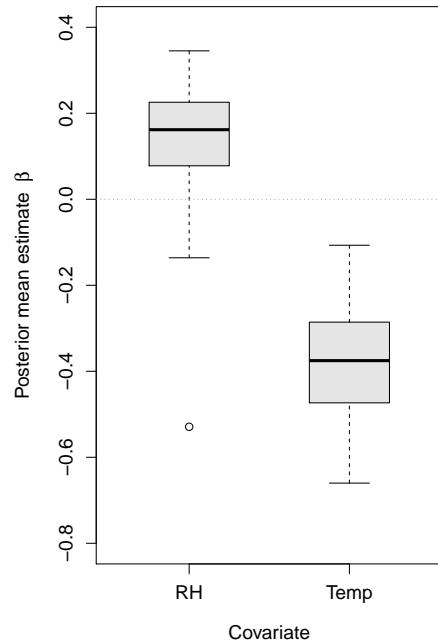


Figure 4.5 Posterior mean estimates of different airboxes, for airbox NO₂ vs. covariates ‘relative humidity’ and ‘temperature’ measured in the same airbox, full year 2016.

Figure 4.6 shows the posterior distributions of β_{NO_2} , β_{RH} and β_T for ten collocated airboxes for a shorter period of time with NO₂ from the conventional monitor as the response (240 hours). The posterior distributions of β_{NO_2} are around 1. For some of the NO₂ sensors, however, the β coefficient was below 1 while for other airboxes the β coefficient was above 1. The NO₂ concentrations would thus be corrected in the wrong direction when using the estimated coefficients of another airbox. β_{RH} and β_T have posterior distributions around zero, indicating different slope directions for different airboxes. We note that the short duration may have contributed to some uncertainty in the estimates.

4.3.7 Temporal autocorrelation

In the iterative calibration procedure, the dataset for calibration was not large enough to model the temporal autocorrelation in the NO₂ data. Inclusion of the random walk component (Eq. 4.13) did not lead to improvements in $RMSE_{post}$. When applying the model in Eq. 4.2 to the full year dataset, however, there were clear signs of temporal autocorrelation in the residuals. Including random effects in the model using Eq. 4.13 led to a significant decrease in $RMSE_{post}$ from 8.30 to 3.12 at z_1 , and from 6.76 to 3.71 at z_2 . A complete overview is given in Table A4.2. Inclusion of random effects narrowed the scatterplot closer to the

1:1 line (Figure A4.5). A substantial decrease in residuals and removal of the temporal pattern is visible in the residual plot (Figure A4.6).

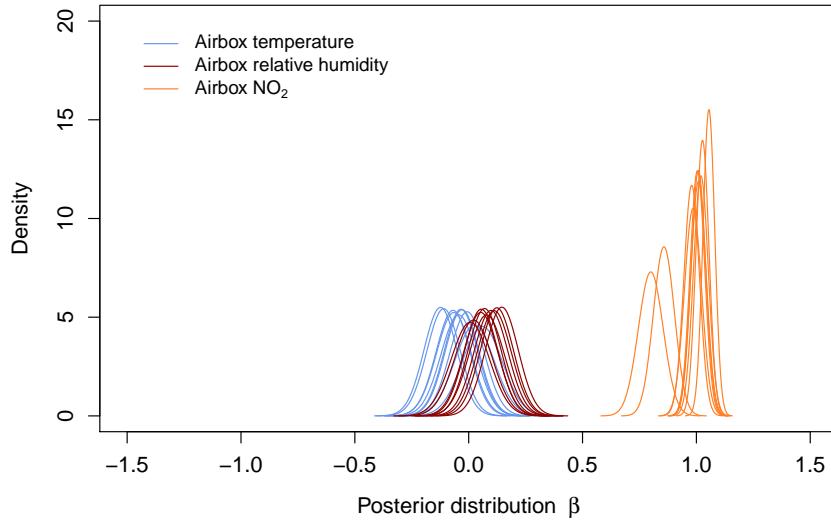


Figure 4.6 Posterior distributions of slopes for reference monitor NO₂ per covariate, for 10 airboxes collocated for ten days.

4.4 Discussion and conclusions

After approximately two to six months after the initial calibration, the airbox NO₂ sensors showed signs of drift. We have evaluated three different methods for regular calibration: daily updated correction factors, corrections based on uniform low concentrations, and a Bayesian regression model. The Bayesian regression model and the daily correction factors both worked very well on the airbox for which they were created, accounting for both systematic bias due to drift and non-systematic errors due to interference effects. However, we found that the transferability of the correction parameters and coefficients to another airbox was limited, though the other airbox was within a short distance and in a similar traffic situation. The poor spatial transferability is consistent with the different sensitivity of individual airboxes to environmental factors including temperature and relative humidity, in agreement with Broday and The Citi-Sense Project Collaborators (2017).

The sensitivity of electrochemical cell NO₂ sensors to temperature and relative humidity has already been evaluated for different types of

4. Calibration of NO₂ sensors in an urban air quality network

sensors (Neri et al., 2002; Mead et al., 2013; Phala et al., 2016). With a set of collocated sensors, we found that the interference effects might be different for individual sensors of the same type when deployed in an outdoor environment. We also showed, in line with Kizel et al. (2018), that for each airbox the coefficients strongly vary over time, emphasizing the need for regular recalibration. These results may be different for other pollutants or sensors. Zalel et al. (2015), for example, found a good temporal stability and spatial transferability for benzo(a)pyrene.

Wind speed data was only available at a single location in Eindhoven. In the model we could therefore only include the temporal variation in wind speed but not its spatial variation. Lerner et al. (2015) found that local wind speed affects the NO₂ concentrations. The NO₂ sensors in our study are covered by the airbox to minimize the influence of direct wind. The regression model may be improved by adding measurements of wind speed and wind direction on each airbox location. Ozone data are measured by the airboxes, but were only available for two months in 2016 for the two airboxes collocated with a conventional monitor, due to sensor failure. Therefore we used O₃ data available at one of the conventional monitor sites. This limited the O₃ data to only one location, but the relation between NO₂ and O₃ is strong and the temporal variability in O₃ is typically higher than the spatial variability. This solution was therefore preferred over using no O₃ data or using only two months of data for analysis. Due to a change in initial calibration factor, there is a sudden change in NO₂ concentrations around May 10. This has affected the models based on 30-day periods including May 10, but not other periods. The change thus does not affect our general conclusions.

We evaluated the performance of the calibration models, their temporal stability and spatial transferability by comparing the RMSE values before and after calibration. This measure is widely used to evaluate sensor performance, but can be influenced by extreme values (Fishbain et al., 2017). We cleaned the data from outliers before the analysis to minimize the influence of extreme values on calibration parameters and their performance (Chapter 3). Besides the RMSE, other performance measures can be used to assess the quality and usability of low-cost air quality sensors. For example, Fishbain et al. (2017) developed a tool kit to evaluate the performance of air quality micro-sensing units. Here, our focus is on calibration performance only.

Miskell et al. (2018) suggested to calibrate low-cost sensors using a conventional monitor in the vicinity with similar land use type. They successfully applied the method on O₃ with an averaging time of 72 hours. We applied a similar method on NO₂ using daily absolute and relative correction factors and an averaging time of 24 hours. This method accounts for drift and daily variability in interference effects. Despite similar traffic conditions at both locations, and traffic being the major

4.4. Discussion and conclusions

contributor to NO₂ levels in Eindhoven, the correction factors could not successfully be transferred from one location to the other. This is probably due to variability between individual airboxes in the strength of drift or interference.

The uniform concentration correction method was proposed by Tsujita et al. (2005) and only adjusts the correction factor when concentrations of pollutants are uniform over the city. This method accounts mostly for drift but also for long-term variability in interference effects. Moltchanov et al. (2015) applied this method using nighttime calibrations (1:00-4:00 a.m.) when O₃ concentrations are uniformly negligible. For NO₂ they could not apply the method because they did not find periods of sufficiently long duration with negligible spatial variation in NO₂ concentrations. We faced the same issue for NO₂.

We evaluated the presence of temporal autocorrelation in the residuals by adding a first-order random walk component to the model, as described in Blangiardo and Cameletti (2015). In our case the addition of this random effect only led to an improvement in RMSE when applied on 80% of the full dataset (2279 hourly non-missing observations) rather than iteratively using the hourly observations of the previous 30 days (maximum 720 hourly non-missing observations).

The need for regular recalibration of parameters is clear. The time series plots of the daily correction factors, the correction factors based on uniform low concentrations, and β coefficients of the INLA model show that independent of the method used, there is significant variability in the correction factors and parameters over time. Also, the parameters are dependent on the individual sensor. Hasenfratz et al. (2012) proposed an on-the-fly calibration procedure for gas sensors mounted on public transport vehicles, calibrating the sensors when in each other's vicinity or when in the vicinity of a conventional monitor. In static sensor networks, a moving reference sensor could be used for regular calibration and data quality evaluation of the sensors in the network as suggested by Kizel et al. (2018). A moving reference sensor takes the different response of individual sensors into account, and would be a suitable solution to account for the spatio-temporal variability in the calibration parameters. A disadvantage is the added workload. Besides a moving reference sensor, it would still be of added value to collocate the sensors once a year to compare performance differences.

Low-cost air quality networks provide data of a fine spatial and temporal resolution. They provide valuable opportunities for spatiotemporal modelling and health risk mapping. It can be debated whether one should use modelled values, as derived from the calibration model, as an input for spatiotemporal modelling purposes and health studies, in which the same covariates are likely to be used again as potential confounders in

4. Calibration of NO₂ sensors in an urban air quality network

for example time series studies. The purpose of this model is however different, the calibration parameters are optimized for the calibration of NO₂, and the covariates have been transformed and scaled. Therefore we do not expect major issues regarding the use of calibrated values in future modelling.

Calibration procedures are important to correct air pollution data before online publishing, to avoid misinterpretation of the results. This study has highlighted the need for such calibration procedure to not just account for drift and interference effects, but also for the variability in drift and interference effects in space, time and between sensors. The strength of this variability may differ between locations, pollutants and sensors used. Transferability of calibration parameters from one sensor to the other and similarities in drift are often assumed, but this assumption is not always justified. Regular calibration should therefore be performed at the location of the low-cost sensor, for example using a moving reference sensor.

Low-cost air quality sensors are valuable instruments to increase the spatial and temporal resolution of air quality sensor networks. When aware of their limitations, sensor-specific differences and when communicating the uncertainties related to their measurements, they could prove useful in various settings.

Appendix

Table A4.1 RMSE before and after temporal calibration, for different lengths of the calibration dataset. In each case, the full model including all covariates (model 9) is used. The length of the validation dataset is 1 day (max. 24 hourly observations) in each case. The minimum number of observations in a calibration dataset is 72 to continue calibration. Slight differences in $RMSE_{pre}$ occur because this minimum is more easily met in a larger calibration dataset.

# cal. days (max. # hourly observations)	Location z_1		Location z_2	
	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
	7 days (168 hours)	9.91	8.45	10.67
14 days (336 hours)	9.91	7.79	10.67	5.68
30 days (720 hours)	9.91	7.80	10.67	5.74
45 days (1080 hours)	9.91	7.56	10.70	5.80

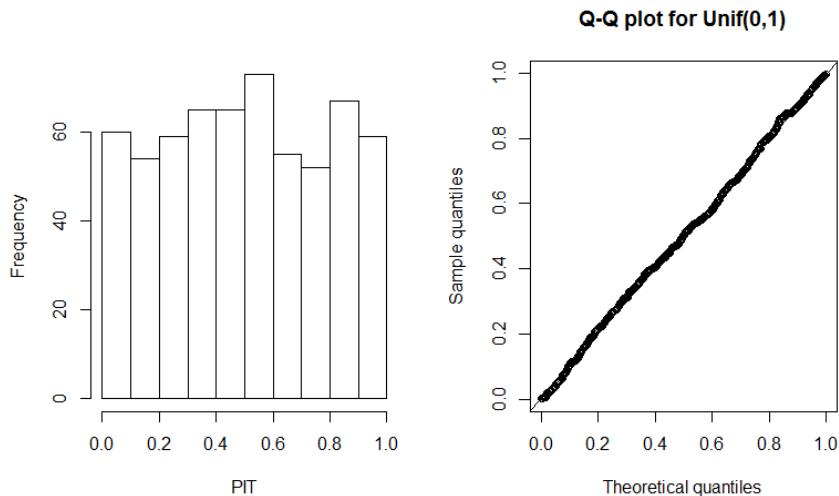


Figure A4.1 Histogram and uniform(0,1) Q-Q plot of the PIT values on June 15, 2016 at location z_1 .

4. Calibration of NO₂ sensors in an urban air quality network

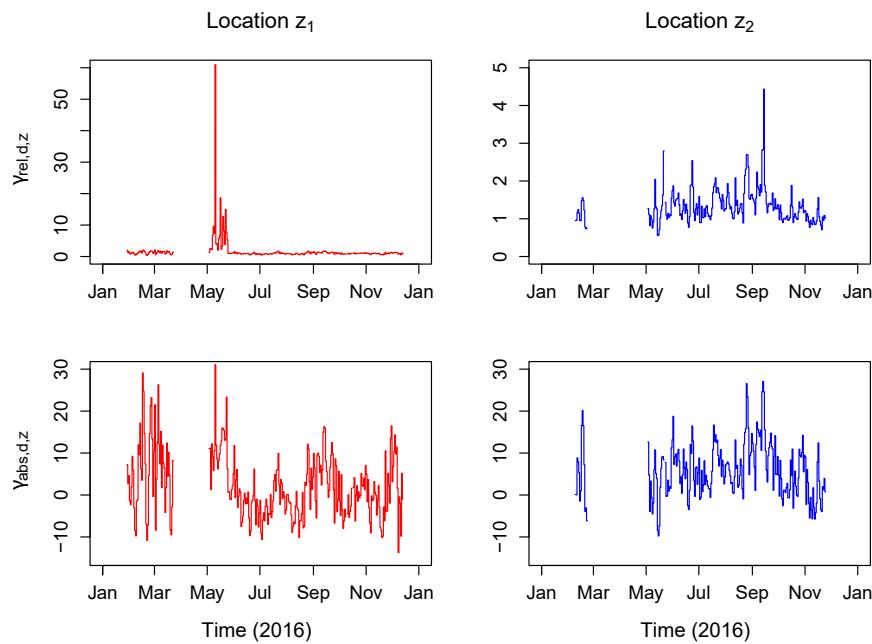


Figure A4.2 Time series of the relative correction factors $\gamma_{rel,d,z}$ and absolute correction factors $\gamma_{abs,d,z}$

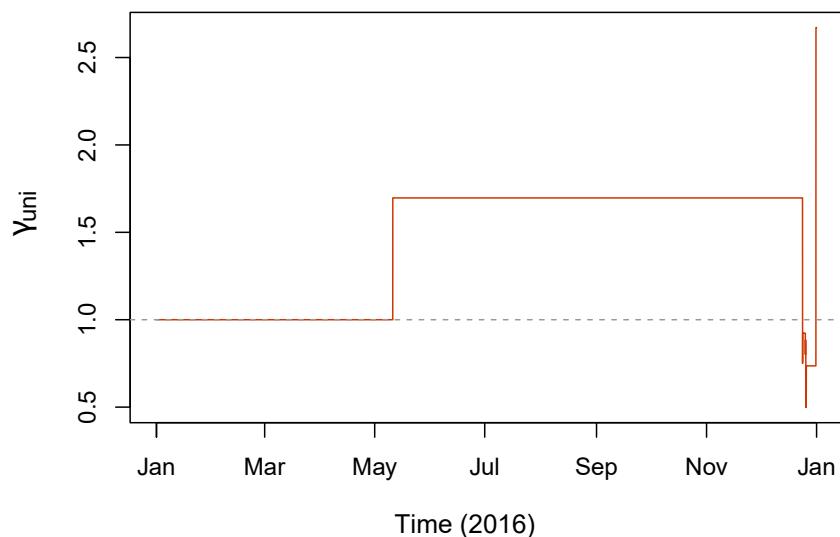


Figure A4.3 Time series of the correction factor γ_{uni} using airbox NO₂ standard deviation threshold $\delta = 5 \mu\text{g m}^{-3}$ and airbox mean NO₂ threshold $\chi = 12 \mu\text{g m}^{-3}$

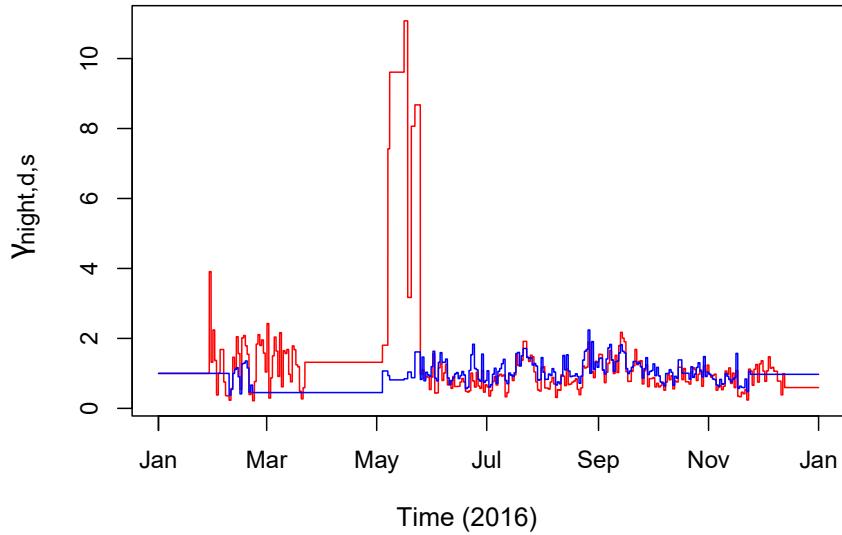


Figure A4.4 Time series of the correction factor $\gamma_{night,d,s}$ using night-time calibrations. Red: location z_1 , blue: location z_2

Table A4.2 RMSE for the models with random effects vs. without random effects. In each model, the dependent variable is $\sqrt{NO2_{ref}}$

Model	Location z_1		Location z_2	
	$RMSE_{pre}$	$RMSE_{post}$	$RMSE_{pre}$	$RMSE_{post}$
Model excluding random effects: $\beta_0 + \beta_{NO2}\sqrt{NO2_{ab}}$ $+ \beta_T T + \beta_{RH} RH^2 +$ $\beta_{O3} \log(O3) + \beta_{WS} \sqrt{WS}$	9.32	8.30	10.77	6.76
Model including random effects: $\beta_0 + \beta_{NO2}\sqrt{NO2_{ab}}$ $+ \beta_T T + \beta_{RH} RH^2 +$ $\beta_{O3} \log(O3) + \beta_{WS} \sqrt{WS}$ $+ rw_1(d, t)$	9.32	3.12	10.77	3.71

4. Calibration of NO₂ sensors in an urban air quality network

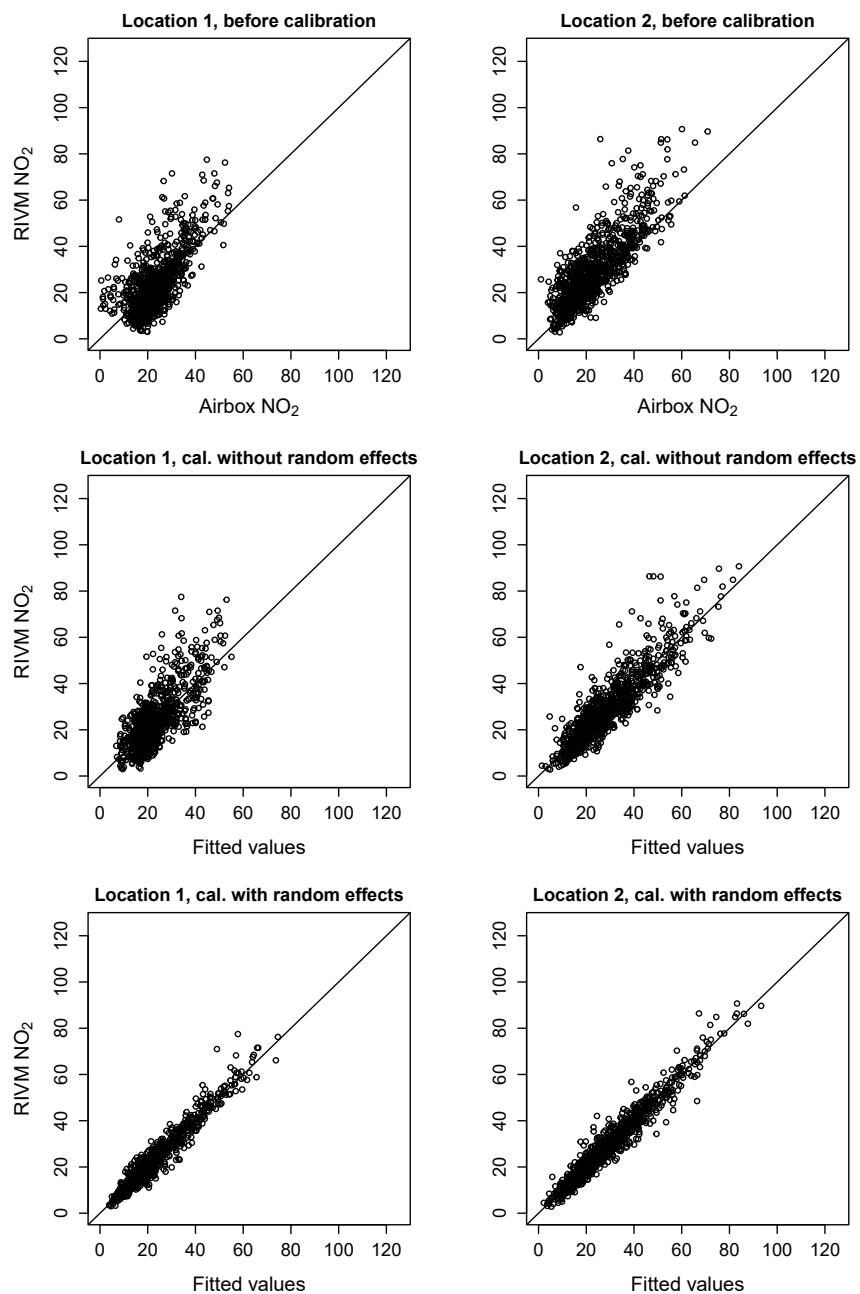


Figure A4.5 Scatterplots before calibration, after calibration without random effects, and after calibration with random effects, for locations z_1 and z_2 .

4.4. Discussion and conclusions

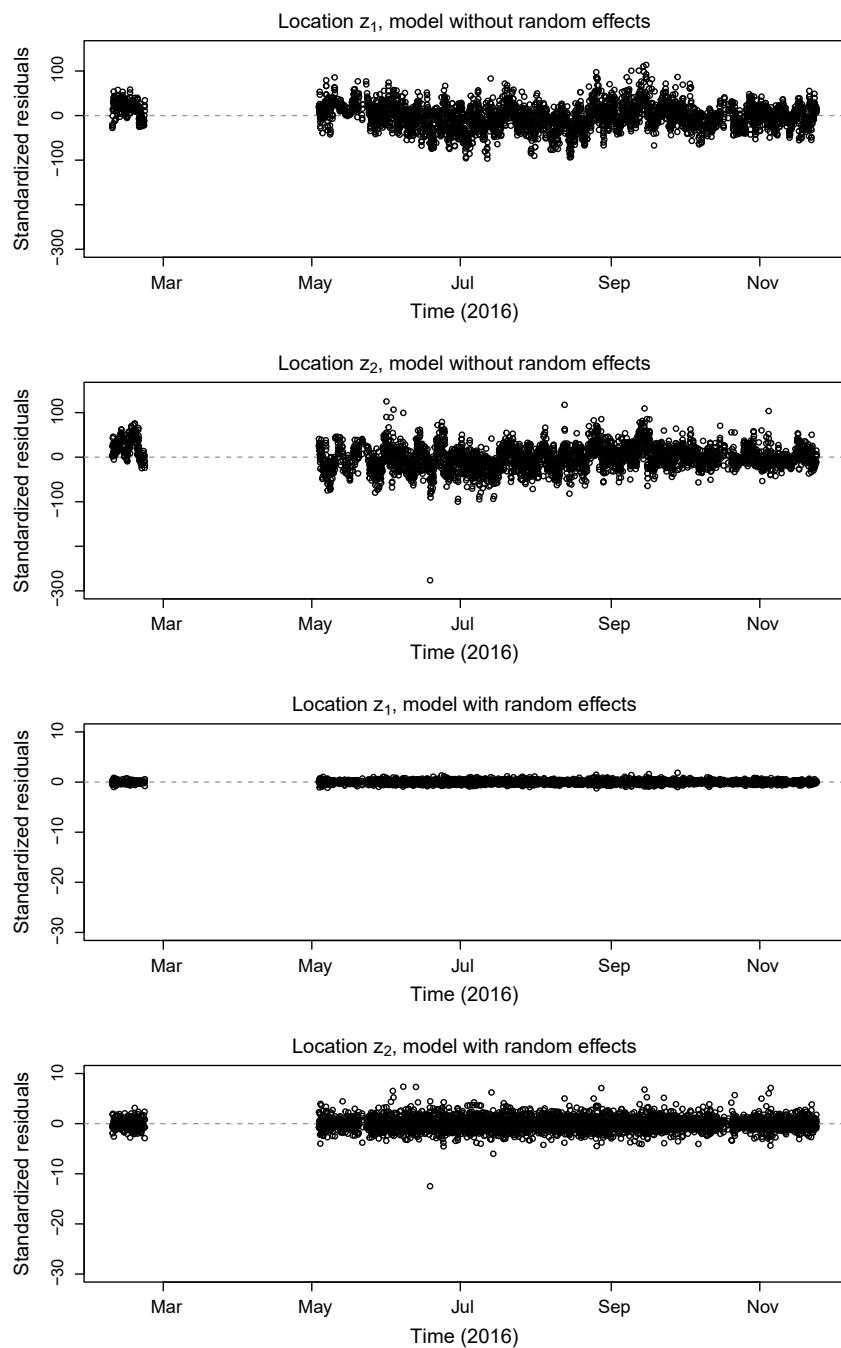


Figure A4.6 Residual plots for INLA models without random effects and with random effects. Note the different scaling used on the y-axes to improve readability.

Spatio-temporal regression kriging for modelling urban NO₂ concentrations

5

Abstract

Recently developed urban air quality sensor networks are used to monitor air pollutant concentrations at a fine spatial and temporal resolution. The measurements are however limited to point support. To obtain areal coverage in space and time, interpolation is required. A spatio-temporal regression kriging approach was applied to predict NO₂ concentrations at unobserved space-time locations in the city of Eindhoven, the Netherlands. Prediction maps were created at 25 m spatial resolution and hourly temporal resolution. In regression kriging, the trend is separately modelled from autocorrelation in the residuals. The trend part of the model, consisting of a set of spatial and temporal covariates, was able to explain 49.2% of the spatio-temporal variability in NO₂ concentrations in Eindhoven in November 2016. Spatio-temporal autocorrelation in the residuals was modelled by fitting a sum-metric spatio-temporal variogram model, adding smoothness to the prediction maps. The accuracy of the predictions was assessed using leave-one-out cross-validation, resulting in an RMSE of 9.91 $\mu\text{g m}^{-3}$, a Mean Error (ME) of -0.03 $\mu\text{g m}^{-3}$ and a Mean Absolute Error (MAE) of 7.29 $\mu\text{g m}^{-3}$. The method allows for easy prediction and visualization of air pollutant concentrations, and can be extended to a near real-time procedure.

This chapter is published as: Van Zoest, V.M., Osei, F.B., Hoek, G., Stein, A., 2019. Spatio-temporal regression kriging for modelling urban NO₂ concentrations. International Journal of Geographical Information Science. doi:10.1080/13658816.2019.1667501

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

5.1 Introduction

Geo-information science supports the computation and visualization of large amounts of spatio-temporal data. In combination with spatio-temporal statistics, prediction maps can be made, for example to visualize air pollution. Air pollution is worldwide a major cause of morbidity and mortality (Cohen et al., 2017) and national air quality monitoring networks have been set up to monitor whether it exceeds legal limit levels. Due to the high costs of the instruments, their maintenance and validation, the number of measurement locations is typically limited to one or two in each city. Low-cost air quality sensor networks, measuring pollutant concentrations at a fine spatio-temporal resolution at urban scale level, are gaining interest (Snyder et al., 2013). These networks can be operated for a longer period of time (>1 year) and can be used to model air pollution at a fine spatial and temporal resolution.

Modelling air pollutant concentrations is done with the aim to predict air pollutant concentrations at unmeasured locations. Land use regression (LUR) models, referring to regression models including land use covariates from a geo-information system (Hoek et al., 2008), are used for modelling fine scale variation in air quality at an urban scale level. In most LUR studies, the models are used to obtain seasonal or annual average predictions (Lee et al., 2017; Kashima et al., 2018; Weissert et al., 2018). These are suited for applications where mostly spatial variability is of importance, such as policy decisions regarding polluted areas of a city or assessing long-term health effects of air pollution. All temporal variation will be neglected when an LUR model is used to predict annual mean concentrations, leading to a loss of precision and power (Klompmaker et al., 2015). The spatial covariates in an LUR model will also not be able to account for all spatial variability, nor for spatial autocorrelation in the residuals.

When a regression model is combined with spatial kriging, the spatial autocorrelation structure can be accounted for. Spatial kriging is then used to interpolate the concentration values between different measurement locations (Beelen et al., 2009; Van de Kassteele et al., 2009). To account for temporal variability and temporal autocorrelation, the regression model can be extended to include temporal covariates, and the residuals of the model can be interpolated using spatio-temporal kriging (Kilibarda et al., 2014; Hu et al., 2015). A spatio-temporal variogram function is then used to describe the spatio-temporal autocorrelation structure (Gräler et al., 2016).

The main objective of this study is to model spatio-temporal variability in urban air pollutant concentrations using a spatio-temporal regression kriging model. We applied the method on a low-cost urban air quality

sensor network in the city of Eindhoven, the Netherlands (Chapter 2), focusing on NO₂.

5.2 Methods

5.2.1 Model formulation

We consider a sensor network with N_s number of sensor locations s_i , $i = \{1, \dots, N_s\}$. The sensor measurements are taken at point support in space, represented by a two-dimensional set of spatial coordinates for each sensor location. At sensor location s_i , the air pollutant concentration $NO_2(s_i, t_j)$ is stored for each of the N_t number of time stamps t_j , $j = \{1, \dots, N_t\}$. Each time stamp denotes the end of an hourly averaging period.

We modelled the air pollutant concentration $NO_2(s_i, t_j)$ with mean $\mu(s_i, t_j)$ and residual $\eta(s_i, t_j)$:

$$NO_2(s_i, t_j) = \mu(s_i, t_j) + \eta(s_i, t_j) \quad (5.1)$$

The mean $\mu(s_i, t_j)$ incorporates the trend component of the model, consisting of a linear combination of covariate values x_c from a set of spatial and temporal covariates C . The trend part of the model is estimated as:

$$\hat{\mu}(s_i, t_j) = \hat{\beta}_0 + \sum_c \hat{\beta}_c x_c(s_i, t_j). \quad (5.2)$$

Here $\hat{\beta}_0$ is the estimated intercept and $\hat{\beta}_c$ are the estimated regression coefficients for covariates $c \in C$, based on ordinary least squares.

After estimating the parameters and predicting $\mu(s_i, t_j)$ at each observed space-time location (s_i, t_j) , we model the spatio-temporal autocorrelation in the residuals. We take the residuals as the differences between the observations and the estimated trend, and in this approach we ignore any uncertainties in the estimated trend. Their inclusion could be part of future research. The distribution of the residuals is visually checked for normality. To explore the spatio-temporal dependency in the residuals, we use the spatio-temporal variogram (Cressie and Wikle, 2011; Sherman, 2011). The spatio-temporal variogram represents the semivariance between any pair of residuals which are separated by spatial lag h and/or temporal lag u :

$$\gamma(h, u) = \frac{1}{2} \mathbb{E} (\eta(s_i, t_j) - \eta(s_i + h, t_j + u))^2. \quad (5.3)$$

Here, \mathbb{E} denotes the expected value. A spatio-temporal sample variogram is formed by averaging the semivariance in regularly spaced spatio-temporal bins, similar to standard spatial variograms. A space-time

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

variogram model is fitted to the spatio-temporal sample variogram. Based on the smallest mean square error (MSE) between the sample variogram and fitted variogram we fitted a sum-metric space-time variogram model. The sum-metric model combines a spatial, temporal and metric model accounting for space-time anisotropy (Gräler et al., 2016):

$$\gamma(h, u) = \gamma(h) + \gamma(u) + \gamma\left(\sqrt{h^2 + (\kappa \times u)^2}\right) \quad (5.4)$$

where $\gamma(h, u)$ is the space-time variogram, $\gamma(h)$ is the spatial variogram, $\gamma(u)$ is the temporal variogram, $\gamma\left(\sqrt{h^2 + (\kappa \times u)^2}\right)$ is the joint variogram, and κ is a spatio-temporal anisotropy scaling parameter. This requires estimation of κ , as well as a set of spatial variogram model parameters $\theta_s = \{\tau_s^2, \sigma_s^2, \phi_s\}$, a set of temporal variogram model parameters $\theta_t = \{\tau_t^2, \sigma_t^2, \phi_t\}$ and a set of joint variogram model parameters $\theta_{joint} = \{\tau_{joint}^2, \sigma_{joint}^2, \phi_{joint}\}$. Each set of parameters contains, respectively, the nugget, partial sill and range of – in our case – a spherical variogram model (Zimmerman and Stein, 2010). The nugget consists of two components: microscale variance and variance induced by inaccuracies in the measurement device (Cressie and Wikle, 2011). The partial sill and range affect the shape of the variogram model. Depending on the spatio-temporal autocorrelation of the pollutant, different pollutants have different variogram models. Variogram model parameter estimation is done through optimization using the bound constrained BFGS method (Byrd et al., 1995). We assume isotropy and stationarity in space and time, which allows for the same variogram to be used in all directions and at all spatio-temporal locations.

5.2.2 Spatio-temporal predictions

We combine the estimated parameters of the regression model and the estimated semivariance parameters to predict the NO₂ concentration $\widehat{NO_2}(s_0, t_0)$ at any unobserved spatio-temporal location (s_0, t_0) located on a spatio-temporal prediction grid:

$$\widehat{NO_2}(s_0, t_0) = \hat{\mu}(s_0, t_0) + \hat{\eta}(s_0, t_0) \quad (5.5)$$

where the predicted trend component $\hat{\mu}(s_0, t_0) = \hat{\beta}_0 + \sum_c \hat{\beta}_c x_c(s_0, t_0)$ is based on the covariate values $x_c(s_0, t_0)$ at spatio-temporal prediction location (s_0, t_0) . Kriging gives us the Best Linear Unbiased Predictor (BLUP) of the residual component, $\hat{\eta}(s_0, t_0) = \lambda_0 \bar{\eta}$, where $\bar{\eta}$ is a vector of observed space-time residuals and λ_0 is a vector of kriging weights (Diggle and Ribeiro, 2007). The kriging weights express the strength of the association between observation locations and the prediction location, estimated as $\lambda_0 = y_0^T \Gamma^{-1}$. Here, y_0 is a vector containing the semivariances between the observation locations and the prediction location (s_0, t_0) , and Γ is a matrix containing the semivariances between all possible combinations of space-time observations. We apply simple kriging, as we assume the residuals to have a known mean of zero.

The prediction maps on the full space-time grid were accompanied with kriging variance maps to evaluate the uncertainty of the kriging predictions. The kriging variance σ_0^2 at a prediction location (s_0, t_0) is defined as (Webster and Oliver, 2001):

$$\sigma_0^2 = \gamma_0^T \Gamma^{-1} \gamma_0. \quad (5.6)$$

Variogram parameter estimation and spatio-temporal kriging were done in R using the ‘gstat’ package (Gräler et al., 2016). The used code is available from the authors upon request. Spatio-temporal kriging is computationally demanding, as it requires computation of the inverse of the spatio-temporal semivariance matrix at every location on the spatio-temporal prediction grid. To improve efficiency and to reduce computation time, we limit the temporal observation locations used for predictions, i.e. perform local kriging on the temporal part. While using all spatial locations, we limit the number of temporal neighbors to those within a temporal distance of $\max(\phi_t, \phi_{joint}/\kappa)$, rounded up to the next whole number. This should not meaningfully influence the predictions, as the kriging weights approach zero when $u > \max(\phi_t, \phi_{joint}/\kappa)$.

5.2.3 Validation

To evaluate the accuracy of the kriging predictions, we performed leave-one-out cross-validation (LOOCV) at all observed space-time locations (s_i, t_j) . For one space-time location (s_i, t_j) at a time, the value $\eta(s_i, t_j)$ is removed from the dataset. The remainder of the dataset, temporally limited to $t_j \pm \max(\phi_t, \phi_{joint}/\kappa)$, is used to predict $\eta(s_i, t_j)$. This process is repeated for each observed space-time location (s_i, t_j) . The Root Mean Square Error (RMSE) is then used to assess the accuracy of the kriging predictions:

$$RMSE = \sqrt{\frac{1}{N_s N_t} \sum_{i=1, j=1}^{N_s, N_t} (\widehat{NO2}(s_i, t_j) - NO2(s_i, t_j))^2} \quad (5.7)$$

as well as the Mean Error (ME):

$$ME = \frac{1}{N_s N_t} \sum_{i=1, j=1}^{N_s, N_t} (\widehat{NO2}(s_i, t_j) - NO2(s_i, t_j)) \quad (5.8)$$

and Mean Absolute Error (MAE):

$$MAE = \frac{1}{N_s N_t} \sum_{i=1, j=1}^{N_s, N_t} (|\widehat{NO2}(s_i, t_j) - NO2(s_i, t_j)|). \quad (5.9)$$

5.3 Application

For this study we used the AiREAS NO₂ data as described in Chapter 2. We estimated the model parameters for one month of data at a time, to

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

account for sensor drift and seasonal variability in the regression coefficients and semivariance parameters. As an illustration of the method, we used in this study hourly data from November 2016, when NO₂ data were available for 20 airbox locations (Figure 5.1). The airboxes measure NO₂ every 10 minutes, and the data were averaged to hourly concentration values to reduce noise and to match the data with the temporal support of the meteorological covariates. Data cleaning and outlier removal were performed as described in Chapter 3 and missing values (9.8%) in each airbox were imputed using regression on the NO₂ values of the remaining airboxes, following Harrell (2018). The observed hourly average NO₂ concentrations varied between 0 and 96 $\mu\text{g m}^{-3}$ in November 2016.

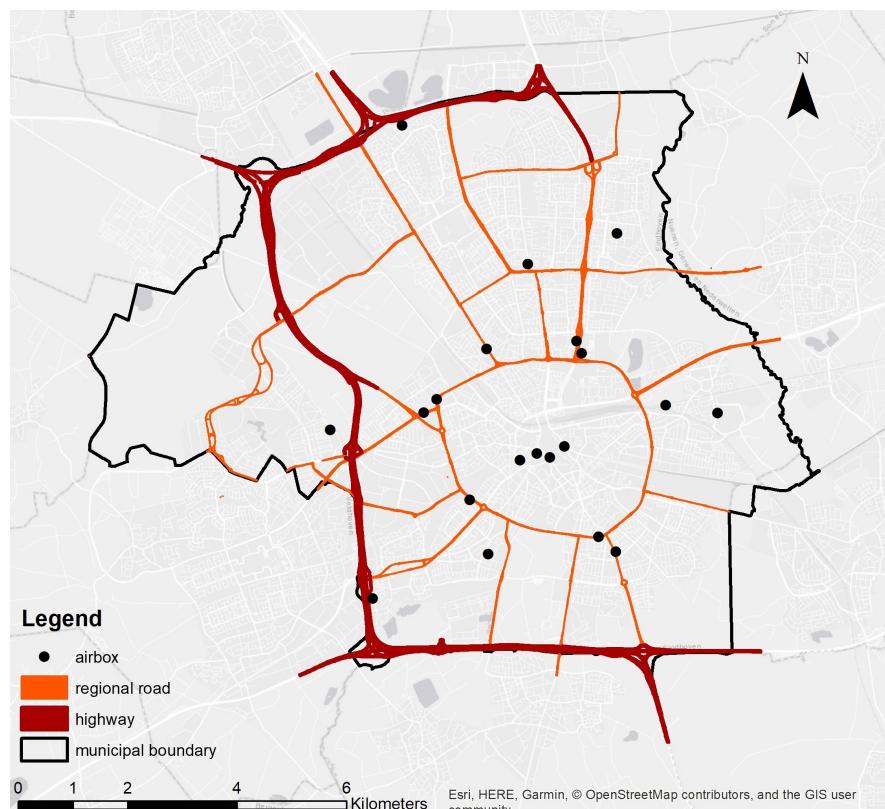


Figure 5.1 Locations of the airboxes in Eindhoven used for modelling NO₂ in November 2016. The coloured lines represent major roads.

The main source of NO₂ is traffic inside the city and it may be trapped in street canyons between high buildings, especially in the areas with a high population density. In the trend part of the model, three spatial covariates and five temporal covariates were included, which significantly affected the NO₂ concentrations at significance level $\alpha = 0.05$ as shown

in the results section. This set of covariates

$$C = \{pop, road, east, RH, WS, WD, hour, wday\}$$

contains population density (*pop*), road type (*road*), easting coordinates (*east*), relative humidity (*RH*), wind speed (*WS*), wind direction (*WD*), hour of day (*hour*) and weekday/weekend (*w*), respectively. The final prediction model then becomes:

$$\begin{aligned} \widehat{NO_2}(s_0, t_0) = & \hat{\beta}_0 + \hat{\beta}_{pop}x_{s_0}^{pop} + \hat{\beta}_{road}x_{s_0}^{road} + \hat{\beta}_{east}x_{s_0}^{east} \\ & + \hat{\beta}_{RH}x_{t_0}^{RH} + \hat{\beta}_{WS}x_{t_0}^{WS} + \hat{\beta}_{WD}x_{t_0}^{WD} \\ & + \hat{\beta}_{hour}x_{t_0}^{hour} + \hat{\beta}_{wday}x_{t_0}^{wday} + \gamma_0^T \Gamma^{-1} \bar{\eta}. \quad (5.10) \end{aligned}$$

Population density was obtained as the number of inhabitants km⁻², from Statistics Netherlands (CBS, 2018) at neighborhood level. The lattice data were converted to a raster with a grid cell size of 25 m, similar to the spatial resolution of the prediction grid. Road type data was obtained from the topographic base dataset TOP10NL (Kadaster, 2018). We reclassified the road types to distinguish between small roads (width 2–7 m) and main roads (width >7 m). In rasterizing the vector data to fit the prediction grid, any raster cell containing a piece of road was classified as a road cell. The distinction between small roads and main roads was based upon the maximum combined area of each road type overlapping with the raster cell. Easting coordinates were included as the coordinates of the prediction grid. Relative humidity, wind speed and wind direction are centrally monitored at the Royal Netherlands Meteorological Institute weather station in Eindhoven (KNMI, 2016), and are therefore considered as temporal covariates only. We distinguish between weekdays and weekends, since the traffic patterns are highly different during weekdays as compared to weekends, thus causing different air pollutant concentrations. Similarly, we include hour of the day in the model to account for diurnal variability in traffic intensity and weather. The prediction grid has a temporal resolution of 1 hour.

5.4 Results and discussion

5.4.1 Model parameter estimation

The regression model, representing the spatio-temporal trend part of the model, explained 49.2% of the variability in NO₂ concentrations in Eindhoven in November 2016. The estimated coefficients and their p-values are shown in Table 5.1. Population density was positively related to NO₂ concentrations, as areas with higher population density tend to have a higher traffic intensity and more high-rise buildings. Areas between high-rise buildings form street canyons in which the pollutants are easily trapped. Road type is related to the amount of traffic and

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

therefore an important predictor of air pollution. Especially the presence of main roads had a large influence on the NO₂ concentrations, due to the higher traffic intensity. Easting is a case study area specific covariate, which is likely related to the prevailing west/south-west wind direction and accumulation of air pollution in the east. Relative humidity and wind speed were negatively related to NO₂ concentrations. A higher wind speed dilutes the pollutant concentrations in the air and therefore naturally leads to lower NO₂ concentrations. Wind direction was related to the temporal variability in NO₂ concentrations, with winds from the South, South-East, East, North-East and North leading to lower NO₂ concentrations than winds from the South-West, West and North-West. The β coefficients of the latter three wind directions were not significantly different from zero, the baseline β coefficient for calm or variable winds.

Relative humidity, wind speed and wind direction cannot be controlled to reduce air pollution in the city. However, policy makers can consider these in spatial planning. Based on prevailing winds from the west/south-west and their impact on the transportation of air pollutants, spatial planners would be advised to locate main sources of air pollution, such as highways and the airport, on the east side of the city. We also observe a strong relationship between NO₂ concentrations and population density. The exposure to air pollutants would be more equally divided amongst inhabitants when spatial planners would step away from the traditional city plan, in which high-rise buildings are clustered in the center and low-rise buildings are clustered in the suburbs. Making main roads smaller might decrease local air pollution, but will likely create congestion and increase air pollution elsewhere.

The residuals follow an approximately normal distribution. The left panel in Figure 5.2 shows the spatio-temporal sample variogram of the residuals of the fitted regression model, using a temporal bin size of 1 hour and a spatial bin size of 500 m. The sample variogram shows some periodicity along the spatial axis, likely due to the limited number of spatial locations on which the variogram is based. The fitted sum-metric variogram model is shown in the right panel of Figure 5.2. Its MSE of 288 was lowest compared to metric and separable variogram models, as also found in the example shown in Gräler et al. (2016). Visually, the fitted variogram model well represents the overall shape of the sample variogram both in terms of spatial, temporal and joint spatio-temporal dependencies. The estimated parameters of the sum-metric variogram model are shown in Table 5.2. We observe that the spatial parameters indicate a pure nugget variogram. The spatial dependencies are therefore only considered in the joint variogram model. The temporal dependencies are considered both in the temporal variogram model and in the joint variogram model.

5.4. Results and discussion

Table 5.1 $\hat{\beta}$ and p-values for the fixed effects part of the regression model. The baseline for road type is ‘no road’. The baseline wind direction is ‘Calm/Variable’, the baseline weekday/weekends is ‘weekday’, and the baseline for hour is ‘0’ (23:00-0:00).

Covariate	Covariate unit	$\hat{\beta}$	p-value
Intercept	N/A	-113.38	<0.001
Population density	inhabitants km ⁻² (s)	1.47	<0.001
Road type: Small roads	factor	0.59	0.02
Road type: Main roads	factor	16.16	<0.001
Easting	km	0.82	<0.001
Relative humidity	%	-0.07	<0.001
Wind speed	m s ⁻¹	-1.03	<0.001
Wind dir.: North	factor	-2.11	<0.001
Wind dir.: North-East	factor	-2.00	<0.001
Wind dir.: East	factor	-4.23	<0.001
Wind dir.: South-East	factor	-2.67	<0.001
Wind dir.: South	factor	-1.76	<0.001
Wind dir.: South-West	factor	0.19	0.71
Wind dir.: West	factor	-0.92	0.10
Wind dir.: North-West	factor	-1.28	0.07
Weekday: Weekend	factor	-3.79	<0.001
Hour: 1	factor	-1.61	0.004
Hour: 2	factor	-2.25	<0.001
Hour: 3	factor	-2.97	<0.001
Hour: 4	factor	-3.20	<0.001
Hour: 5	factor	-2.77	<0.001
Hour: 6	factor	-0.50	0.37
Hour: 7	factor	2.59	<0.001
Hour: 8	factor	3.91	<0.001
Hour: 9	factor	3.29	<0.001
Hour: 10	factor	2.59	<0.001
Hour: 11	factor	2.04	<0.001
Hour: 12	factor	2.14	<0.001
Hour: 13	factor	2.78	<0.001
Hour: 14	factor	4.32	<0.001
Hour: 15	factor	6.03	<0.001
Hour: 16	factor	8.05	<0.001
Hour: 17	factor	9.70	<0.001
Hour: 18	factor	8.99	<0.001
Hour: 19	factor	6.78	<0.001
Hour: 20	factor	5.46	<0.001
Hour: 21	factor	4.68	<0.001
Hour: 22	factor	4.36	<0.001
Hour: 23	factor	2.70	<0.001

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

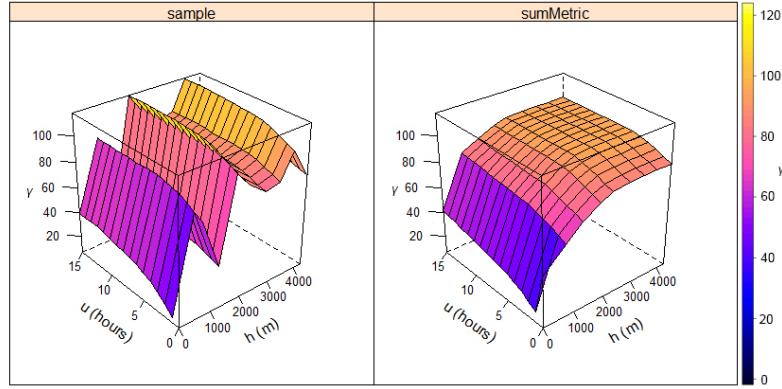


Figure 5.2 Spatio-temporal sample variogram (left) and sum-metric fitted variogram model (right)

Table 5.2 Spatio-temporal variogram parameter estimates for the fitted sum-metric variogram.

Parameter	Estimate
τ_s^2	35.2 ($\mu\text{g m}^{-3}$) ²
σ_s^2	0.00 ($\mu\text{g m}^{-3}$) ²
ϕ_s	2692 m
τ_t^2	0.00 ($\mu\text{g m}^{-3}$) ²
σ_t^2	10.8 ($\mu\text{g m}^{-3}$) ²
ϕ_t	6.1 hrs
τ_{joint}^2	10.6 ($\mu\text{g m}^{-3}$) ²
σ_{joint}^2	40.3 ($\mu\text{g m}^{-3}$) ²
ϕ_{joint}	2670 m
κ	61.1 m hr ⁻¹

5.4.2 Prediction maps

Figure 5.3 shows the prediction maps of four time stamps on Monday November 7, 2016. The maps represent the spatial variability as well as the diurnal variability in NO₂ concentrations. The dark blue colors between 2 and 3 a.m. suggest that the concentrations are low. This can be expected during night hours when traffic intensity is low as well. Main roads have a substantially higher NO₂ concentration than small roads within the neighborhoods. The neighborhoods can be clearly distinguished due to the effect of population density. Some smoothing is visible thanks to kriging of the residuals. During rush hours, e.g. between 7 and 8 a.m. and between 5 and 6 p.m., the concentrations are

5.4. Results and discussion

higher than during the night, both at the main roads and at background locations. Especially the air close to the roads south of the center is most polluted with concentrations $>40 \mu\text{g m}^{-3}$. At this location, we find one of the main roads connecting the highway to the city center. At noon, background levels slightly drop, but still a hotspot exists around the southern main entrance road (red ellipse).

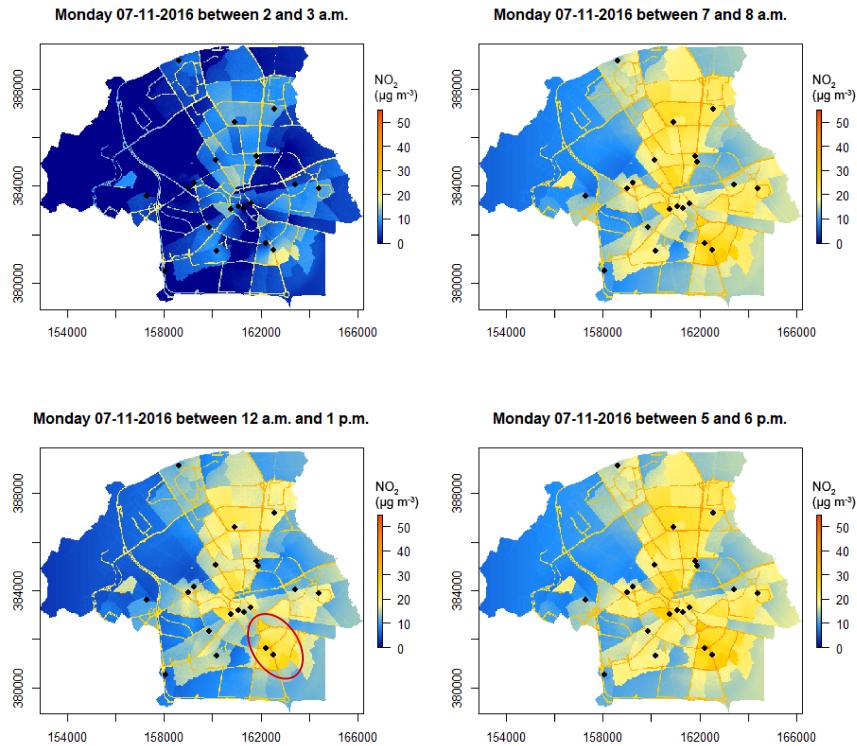


Figure 5.3 Prediction maps of NO₂ concentrations at four time stamps on Monday November 7, 2016 (UTC time; local time is one hour later). The covariate ‘population density’ was included as lattice data, creating clearly distinguished features for the neighborhoods. The red ellipse indicates a hotspot, with locally elevated NO₂ concentrations around the southern main city entrance road.

The prediction maps also allow for visual inspection of extreme values. Spatial extremes could be identified as local hotspots on the map. Temporal extremes can be identified by comparing predictions at different time stamps. Figure 5.4, for example, shows the prediction maps of four different Sundays in November 2016 between 5 and 6 p.m. Clearly, the NO₂ concentrations on November 27 were extremely high throughout the city. The high concentrations on November 27 could not be explained by the meteorological covariates in the trend part of the model, nor

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

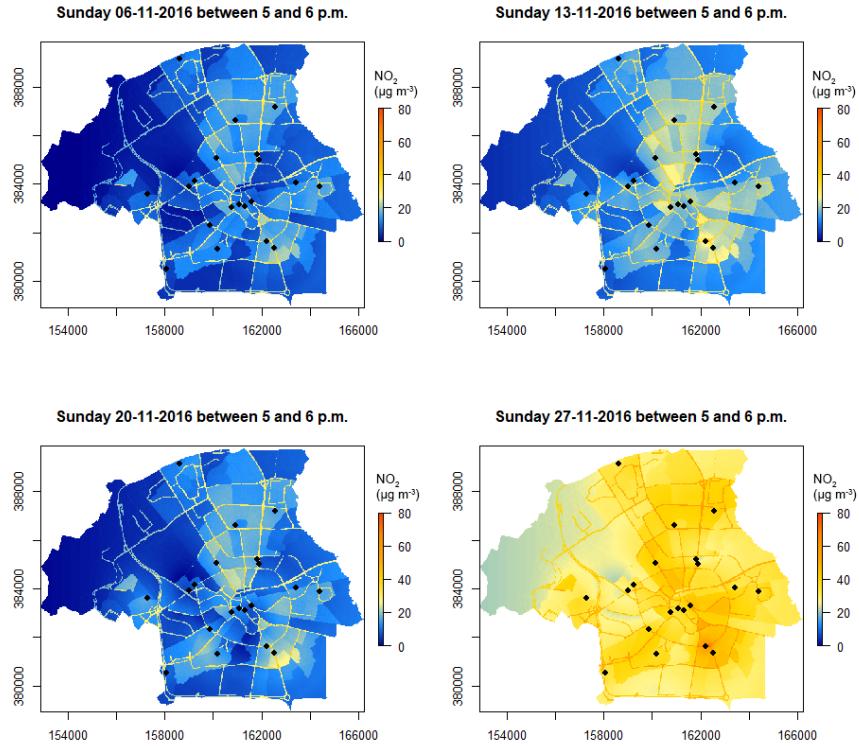


Figure 5.4 Prediction maps of NO₂ concentrations at four Sundays in November 2016, between 5 and 6 p.m. (UTC time; local time is one hour later). Note that different concentration limits were used as compared to Figure 5.3, to visualize the high concentrations on November 27.

by other extreme weather conditions, public events or traffic intensity. However, it should be noted that air pollution levels are based on a very complex combination of sources and sinks, which can be anthropogenic, natural or chemical (Fenger, 2009; Brook et al., 2010). Since all airboxes measured high values on November 27, the extreme values are likely due to a real air pollution event rather than measurement error.

Prior to the analysis, covariates were selected to be included in the trend part of the model. Some covariates were not included due to a lack of significant association with NO₂ concentrations ($\alpha = 0.05$) or lack of improvement in the amount of explained variability. Distance to the nearest road had no significant impact on NO₂ concentrations, because all airboxes were attached to light poles near a road and the variability in distance was only minor. As an alternative, distance to the nearest main road was explored as a covariate. For most airboxes at background air pollution locations, however, these distances were

too large to find significant effects. When systematically sampling at different distances from the road smaller than ϕ_s , it is more likely to find significant effects for this covariate. Instead, we included road type as a factor covariate, distinguishing between no road, small roads and main roads. The difference between the predictions for ‘no road’ and ‘small road’ is small, as can be seen in Table 5.1 and in the prediction maps. This is no surprise due to the low traffic intensity in smaller streets, practically diluting to background concentrations. We expected distance to highway to be negatively related to NO₂ but found opposing results, likely because of an inverse relationship between distance to highway and population density. The final model included three spatial covariates and five temporal covariates, a number small enough to avoid overfitting.

5.4.3 Model performance

The RMSE obtained using LOOCV was 9.91 $\mu\text{g m}^{-3}$, the ME was -0.03 $\mu\text{g m}^{-3}$ and the MAE was 7.29 $\mu\text{g m}^{-3}$. Due to the use of lattice data for the covariate ‘population density’, the boundaries between neighborhoods are clearly visible on the prediction map. Although this may partly be caused by differences in building patterns, some smoothness is expected. The covariate ‘Easting’ should be interpreted as one specific to the study area, and $\hat{\beta}_{east}$ should therefore not be used outside the study area. A combination of this covariate with a low population density on the west side of the study area, creates low NO₂ predictions in the western part of the city. No airboxes are located in this area, therefore the kriging variances are higher here (Figure 5.5). Due to the airport and highways located in this area, true concentrations could be higher as well.

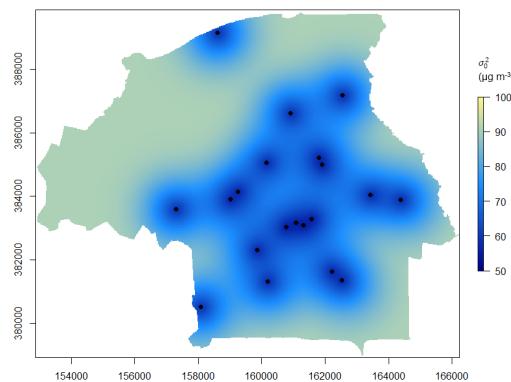


Figure 5.5 Kriging variance map (Monday 07-11-2016 between 7 and 8 a.m.)

For the spatial variogram, we found a pure nugget effect. Optimization of the sampling scheme, e.g. by using shorter distances between some of the sensors, may further improve the estimation of the spatial variogram

5. Spatio-temporal regression kriging for modelling urban NO₂ concentrations

parameters. The sampling scheme in Eindhoven is mostly based on variability in air pollutant concentrations and on the locations of people at risk. Sampling near sources of air pollutants, e.g. the airport, and sampling at different distances from the road may lead to additional covariates of interest and improved model predictions. Further research is needed on sampling scheme optimization, for which the rise of low-cost sensor networks provides valuable opportunities.

5.5 Conclusions

In this paper we predicted urban NO₂ concentrations in space and time using a spatio-temporal regression kriging approach. We applied the model on a low-cost urban air quality sensor network in the city of Eindhoven, the Netherlands. A set of spatial covariates, including road type, population density and easting coordinate, and a set of temporal covariates accounting for meteorological variability and periodicity, were included in the trend part of the model. Kriging of the residuals led to more smoothness in the prediction maps compared to a trend model only. Due to the strong temporal variability in the data, spatio-temporal kriging was more useful than spatial kriging. It also allowed for more accurate variogram estimation using all 14400 space-time locations rather than the limited 20 spatial locations. Using the sum-metric variogram model, the spatial and temporal dependencies were not only modelled independently, but also their joint dependencies. In our case of a pure nugget spatial variogram, these joint dependencies were stronger than the purely spatial dependencies.

The method was useful for spatio-temporal prediction of NO₂ in an urban area, where the resulting maps can assist policy makers in infrastructural decision making and epidemiologists in health risk mapping. They can also improve the development of healthy cyclist route planning (Sharker and Karimi, 2014) and they can be of use in outlier detection to distinguish between errors and events. After selection of relevant site-specific covariates, the method can be applied in other urban areas where fine resolution urban air quality sensor networks are emerging. While traffic-related covariates are of importance in Eindhoven, other covariates such as distance to factories may be of relevance in highly industrial cities. As the emissions of factories are, like traffic, also dependent on hour of the day and weekday/weekends, including these covariates will likely also be of added value in industrial cities. Interactions between covariates can also be included in the trend part of the model, when enough spatial and temporal locations are available to avoid overfitting.

The estimates of the β coefficients and spatio-temporal variogram parameters should be regularly updated, e.g. every month, to account for drift and seasonal variability in the estimates. Prediction could be extended to a near real-time procedure in a straightforward way, for

5.5. Conclusions

example by creating prediction maps of the air pollutant concentrations each hour. In this way, air pollutant concentrations can be efficiently visualized, allowing for communication with citizens and creating awareness about the quality of the air they breathe.

6

Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

Abstract

Short-term exposure to air pollution has been associated with exacerbation of respiratory diseases such as asthma. Substantial heterogeneity in effect estimates has been observed between previous studies. This study aims to quantify the local burden of daily asthma symptoms in asthmatic children in a medium-sized city. Air pollution exposure was estimated using the nearest sensor in a fine resolution urban air quality sensor network in the city of Eindhoven, the Netherlands. Bayesian estimates of the exposure-response function were obtained by updating *a priori* information from a meta-analysis with data from a panel study using a daily diary. Five children participated in the panel study, resulting in a total of 400 daily diary records. Positive associations between NO₂ and lower respiratory symptoms (LRS) and medication use were observed. The odds ratio (OR) for any lower respiratory symptoms was 1.07 (95% C.I. 0.92, 1.28) expressed per 10 µg m⁻³ for current day NO₂ concentration, using data from the panel study only (uninformative prior). ORs for dry cough and phlegm were close to unity. The pattern of associations agreed well with the updated meta-analysis. The meta-analytic random effects summary estimate was 1.05 (1.02, 1.07) for LRS. Credible intervals substantially narrowed when adding prior information from the meta-analysis. The OR for LRS with an informative prior was 1.06 (0.99, 1.14). Burden of disease maps showed a strong spatial variability in the number of asthmatic symptoms associated with ambient NO₂ derived from a regression kriging model. In total, 70 cases of asthmatic symptoms per day can be attributed to NO₂ exposure in the city of Eind-

This chapter is submitted as: Van Zoest, V., Hoek, G., Osei, F.B., Stein, A. Bayesian analysis of the short-term impact of NO₂ exposure on local burden of asthmatic symptoms in children. Environmental Research.

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

hoven. We conclude that Bayesian estimates are useful in estimation of specific local air pollution effect estimates and subsequent local burden of disease calculations. With the fine resolution air quality network, neighborhood-specific burden of asthmatic symptoms was assessed.

6.1 Introduction

Air pollution has major effects on human health including respiratory and cardiovascular diseases (Brunekreef and Holgate, 2002; Brook et al., 2010; Guarneri and Balmes, 2014; Goldizen et al., 2016). In the recent Health Risks of Air Pollution In Europe (HRAPIE) report, the WHO presents the key health endpoints which require further quantification for a full health impact assessment of outdoor air pollution (WHO, 2013a). One of the endpoints is the association between short-term exposure to air pollutants and incidence of daily asthma symptoms in children (WHO, 2013a). Asthma symptoms cause a significant burden of disease (Martinez and Vercelli, 2013). A number of panel studies have been conducted with the aim to quantify the short-term effects of air pollutants on asthmatic symptoms, particularly in Europe and North-America (Weinmayr et al., 2010).

To assess the burden of disease related to asthmatic symptoms in a specific city, generic concentration-response functions can be used, based on previous studies conducted in other areas. Alternatively, data from typically a single epidemiological study in the specific city can be used. Both approaches have pros and cons, including more robust evidence from multiple studies versus differences in the magnitude of effect between different study areas. An interesting option is to combine both approaches in a Bayesian analysis. Previous studies have suggested that accurate estimates and narrow credible intervals can be acquired with Bayesian estimates that include prior information in addition to local effect estimates on the exposure-effect relationship (Post et al., 2001; Le Tertre et al., 2005; Liu et al., 2009; Beach et al., 2012).

Daily exposure is estimated in various ways in panel studies, including measurements at a central monitoring site, personal monitoring and satellite imagery (Chambers et al., 2018). Personal monitoring of air quality is expensive and time-consuming (Brandt et al., 2015), so its rare use is mostly limited to short measurement campaigns (e.g. Linn et al., 1996; Spira-Cohen et al., 2011). In the majority of panel studies, one or two central monitors in a city are used to estimate exposure (Roemer et al., 1993; Van der Zee et al., 1999, 2000; Dales et al., 2009). However, air pollution levels typically show a strong short-distance variability within urban areas, because of the variety of road types and land uses within the city (Hoek et al., 2008). A more accurate estimate of exposure can be acquired by estimating exposure at or near the individual residence and work or school address.

The objective of this study is to estimate the spatially explicit burden of air pollution exposure on daily symptom prevalence in asthmatic children in a medium-sized urban area. We used a Bayesian analysis for effect estimation, with the aim to combine prior information from

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

a meta-analysis with data from a panel study conducted in the city of Eindhoven, the Netherlands. The ILM fine resolution urban air quality sensor network in this city was used for exposure estimation (Chapter 2).

6.2 Methods

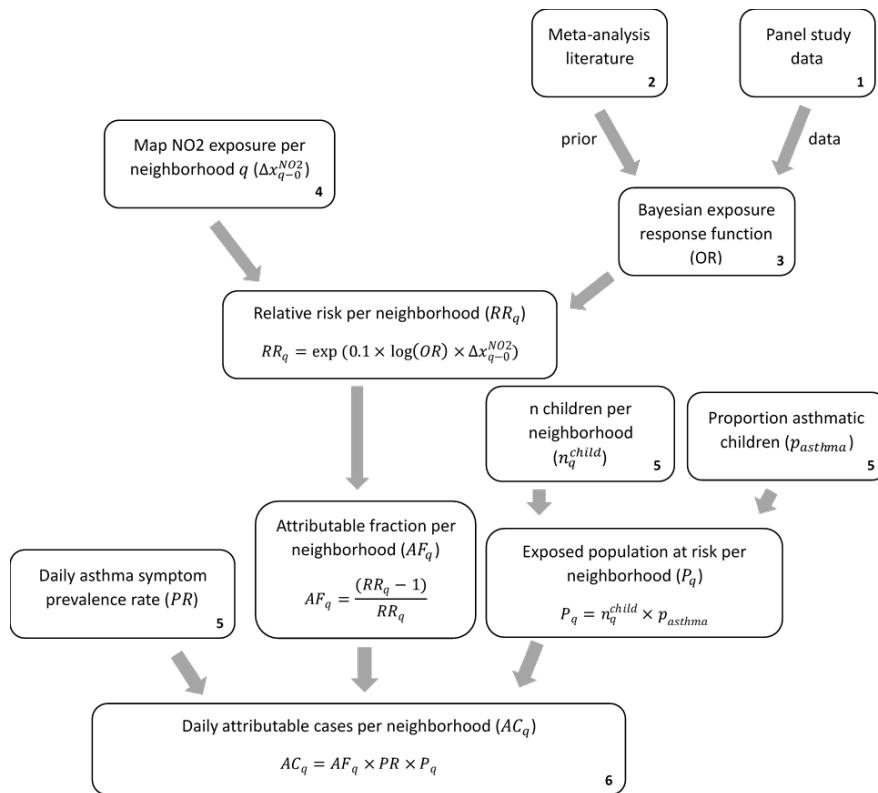


Figure 6.1 Flowchart of daily asthma symptom calculation

Our study included the following components to assess the local burden of disease (Figure 6.1):

1. Panel study among asthmatic children in Eindhoven, combining a daily symptom diary and a fine resolution sensor network to obtain local effect estimates
2. Update of the meta-analysis by Weinmayr et al. (2010) to obtain prior information on the association between air pollution and daily asthma symptoms in children
3. Bayesian analysis with and without informative prior to obtain the concentration-response function to be used in the burden of asthma symptom calculation

4. Spatial modelling of outdoor air pollution across the city of Eindhoven, averaged to annual mean concentrations per neighborhood
5. Collection of data on number of children per neighborhood, proportion of children with asthma, and daily asthma symptom prevalence rate
6. Calculation of neighborhood-specific and overall city burden of daily asthma symptoms, combining prevalence of daily asthma symptoms with the attributable fraction related to the average neighborhood air pollution

6.2.1 Panel study

The design of the panel study is based upon previous panel studies in the Netherlands (Roemer et al., 1993; Van der Zee et al., 1999, 2000) and elsewhere (Weinmayr et al., 2010). An important lesson from previous studies is that short observation periods (<2 months) are prone to confounding by infection episodes which are difficult to characterize (Weinmayr et al., 2010). Therefore, in our panel study, each participant was asked to participate for 4 months. To increase variability in air pollutant concentrations, participants started at different moments in time. The total study period was 13 months (March 29, 2017 - April 22, 2018).

Participants were recruited in different ways. Articles advertising about the panel study were published in local newspapers, on social media and on websites. Flyers were spread door-to-door to ~5000 households, and flyers were displayed in schools, at general practitioners and at pulmonologists in hospitals in Eindhoven. In order to be eligible to participate in the panel study, a subject had to meet all of the following criteria: (1) the child has current asthma, (2) the child is 7 to 11 years old at the start of his or her participation, and (3) the child lives and attends school in the municipality of Eindhoven. A child was considered to have current asthma when at least two of the following three criteria were met: (1) asthma diagnosis by a physician ever, (2) wheeze in the past 12 months, and (3) use of asthma medication in the past 12 months. Children living in a house in which people smoke indoors were excluded from participation, to avoid noise related to this strong risk factor. Children aged 7-11 years were selected because a minimum age of 7 is required for lung function measurements.

The parents of the children were asked to fill out an electronic diary with the child every evening during the 4 months of participation. Electronic diaries are considered more accurate than paper diaries, since the electronic diaries could only be filled in retrospectively for a limited amount of days up to a week (Velická et al., 2015). The diary questions were based on previous panel studies in children (Roemer et al., 1993; Van der Zee et al., 1999, 2000). The symptoms included wheeze,

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

shortness of breath in rest, shortness of breath after exercise, waking up during the night with breathing problems, dry cough, phlegm, and nose complaints. In the diary, the participants reported on-demand medication use and the presence of asthma symptoms as absent, mild, or moderate / severe. The participants were visited at home three times: at the start of the 4-month study period, after 2 months, and at the end of the study period. During the first visit, the study was explained to the child and parent(s), there was an opportunity to ask questions, informed consent forms were signed, and instructions for filling out the daily diary were given. Furthermore, the parent(s) filled out a baseline questionnaire to characterize the child's health status, medication use, daily activities and indoor sources of air pollutants. After 2 months the research assistant visited the participants to keep motivation up, answer potential questions and monitor general progress. All parents signed an informed consent form. The study was approved by the Medical Ethical Committee of University Medical Center Utrecht.

6.2.2 Statistical analysis

Multiple logistic regression models are often used to model the relation between air pollutant concentrations and occurrence of asthma-related symptoms (Roemer et al., 1993; Forsberg et al., 1998; Van der Zee et al., 1999, 2000; Migliore et al., 2009; Schinasi et al., 2011; Ranzi et al., 2015). The binary outcome $y_{v,d}$ denotes the presence or absence of an outcome reported by participant v on day d . Separate models were built for each of the symptoms and for any lower respiratory symptoms (LRS), including wheeze, shortness of breath in rest, shortness of breath after exercise, and waking up during the night with breathing problems. Daily exposure to NO₂, PM₁₀, PM_{2.5} and PM₁ was estimated by the ILM air quality monitoring network (Chapter 2), using the airbox closest to the house and school of the child. All airboxes measure particulate matter, whereas 25 airboxes measure NO₂. Only six airboxes measure UFPs, so the distance from the house of the child to the nearest UFP sensor could be >2.5 km in some cases, while UFP varies strongly over shorter distances. Therefore, daily exposure to UFP was estimated by the mean of the six airboxes. We evaluated the associations of air pollution concentrations of the same day (lag 0), day before (lag 1), and average of lag 0-2 to allow for delayed associations on the various symptoms. The effect estimates were calculated per 10 $\mu\text{g m}^{-3}$ increase in air pollutant concentrations for NO₂, PM₁₀, PM_{2.5} and PM₁, and per 10000 particles increase for UFP. Separate models were built for each lag. The presence of a symptom $y_{v,d}$ follows a Bernoulli distribution with probability of occurrence $p_{v,d}$:

$$y_{v,d} \sim \text{bernoulli}(p_{v,d}) \quad (6.1)$$

where

$$\begin{aligned} \text{logit}(p_{v,d}) = & \beta_{0,v} + \beta_{ap}x_{v,d}^{ap} + \beta_{sday}x_{v,d}^{sday} + \beta_{RH}x_d^{RH} + \beta_Tx_d^T \\ & + \beta_{flu}x_{v,d}^{flu} + \beta_{wday}x_d^{wday} \end{aligned} \quad (6.2)$$

Here, β_c are the coefficients and x^c are the covariate values for the different covariates c , consisting of the air pollutant ap in $\{NO_2, PM1, PM2.5, PM10, UFP\}$ and confounders. Confounders were included a priori based upon previous studies: day of follow-up ($sday$), daily mean relative humidity (RH), daily mean temperature (T), daily reported flu (flu), and day of the week categorized as weekdays/weekend days ($wday$). Daily temperature and relative humidity were obtained from the Royal Netherlands Meteorological Institute weather station in Eindhoven (KNMI, 2019). Characteristics of the child that are expected to remain constant over the full study period, such as gender, socio-economic status, and presence of pets in the household, were not included in the analysis because the form of the model considers each subject as its own control. To adjust for differences in baseline symptom reporting, individual-specific intercepts $\beta_{0,v} = \psi + \gamma_v$ were specified for each individual, consisting of an overall intercept ψ and a random intercept γ_v for each participant v . Our interest is in the odds ratio (OR): $OR = \exp(\beta_{ap})$.

We first estimated the parameters using restricted maximum likelihood (REML). Based on this pre-liminary analysis (Tables A6.1–A6.5 in the Appendix), we found no informative results (i.e. wide confidence intervals) for PM_{10} , $PM_{2.5}$, PM_1 and UFP. We therefore continued our Bayesian analysis and burden of disease study on NO_2 only.

Bayesian estimation of the parameters was performed using JAGS (Plummer, 2003), through the ‘R2jags’ package in R (Su and Yajima, 2015). We used the Gelman-Rubin diagnostic \hat{R} to evaluate convergence (Gelman and Rubin, 1992), by comparing variances of different Markov Chain Monte Carlo (MCMC) simulation chains. Values close to 1 indicate convergence. Two chains of MCMC simulations were ran until convergence was achieved for all parameters in the model ($\hat{R} < 1.1$).

6.2.3 Meta-analysis and prior selection

First, we used an uninformative prior for all parameters in the model (Table 6.1). The individual-specific intercepts were considered as exchangeable random effects $\gamma_v \sim N(\mu = 0, \sigma = \sigma_\gamma)$ with an uninformative prior on σ_γ .

We then evaluated the effect of choosing an informative prior for β_{NO_2} . To obtain these informative priors we performed a meta-analysis of the literature, updating the work by Weinmayr et al. (2010). They performed

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

Table 6.1 Uninformative priors for Bayesian estimation of the parameters in the model (Eq. 6.2)

Variable	Prior distribution ^a
$\beta_{NO2,y}$ for each symptom	$\beta_{NO2,y} \sim N(\mu = 0, \sigma = \sigma_y)$ $\sigma_y \sim N_{+\infty}(\mu = 0, \sigma = \sqrt{0.1})$
β_c for each confounder	$\beta_c \sim N(\mu = 0, \sigma = \sigma_c)$ $\sigma_c \sim N_{+\infty}(\mu = 0, \sigma = \sqrt{10})$
ψ (fixed intercept)	$\psi \sim Unif(-100, 100)$
γ_v (random intercept)	$\gamma_v \sim N(\mu = 0, \sigma = \sigma_y)$ $\sigma_y \sim N_{+\infty}(\mu = 0, \sigma = \sqrt{10})$

^a $N_{+\infty}$ denotes a half-normal distribution truncated at zero, such that the standard deviation can only take positive values.

a systematic review and meta-analysis of literature published between 1990 and July 2008 on short-term health effects of NO₂ on respiratory health among children with asthma or asthma-like symptoms, providing estimates for asthmatic symptoms and cough. To get a more up-to-date estimate, we updated the meta-analysis to include publications from August 2008 to March 2019. We followed the same procedure applied by Weinmayr et al. (2010). The MEDLINE database was searched through the PubMed search engine, using the search string [("asthma" OR "wheeze" OR "cough" OR "bronchitis" OR "lung function") AND ("air" AND pollut*) AND ("NO2" OR "NO(2)" OR "nitrogen dioxide")] and limits were set to retrieve only publications about children (0-18 years). Based on the abstracts, we excluded indoor air pollution studies, laboratory studies and studies on infants. The meta-analysis included only panel studies on asthmatic or symptomatic children which reported a quantitative effect estimate and which controlled for temperature and day of the week or temporal autocorrelation.

The effect estimates of all panel studies (1990 to March 2019) were combined in a random effects meta-analysis model (DerSimonian and Laird, 1986) using the 'metafor' package in R (Viechtbauer, 2010). The effect estimates of all studies were standardized to β_{NO2} coefficients per 10 $\mu\text{g m}^{-3}$. Where needed, concentrations in ppb were converted to $\mu\text{g m}^{-3}$ using the standard conversion at 20 °C: 1 ppb = 1.91 $\mu\text{g m}^{-3}$. ORs were converted to β_{NO2} coefficients before standardization, using the natural logarithm $\beta_{NO2} = \log(OR)$. The combined effects estimate of β_{NO2} was used as a normally distributed prior in the Bayesian estimation.

6.2.4 Burden of disease mapping

To obtain the burden of disease we calculated the potential health risk reductions when NO₂ exposure levels would be reduced from its actual

6.2. Methods

levels per neighborhood to zero. The analysis was conducted per neighborhood because this was the smallest unit with data on number of children available. We determined the number of attributable cases AC_q for each neighborhood q :

$$AC_q = AF_q \times PR \times P_q \quad (6.3)$$

where the attributable fraction $AF_q = \frac{RR_q - 1}{RR_q}$ (Shaddick et al., 2018) is based on the neighborhood-specific relative risk RR_q and PR is the prevalence rate. We use the OR as a proxy for RR, as it typically represents RR well (Liu et al., 2009). Since we obtain the OR per $10 \mu\text{g m}^{-3}$, multiplication by 0.1 is required for standardization. RR_q is then obtained as:

$$RR_q = \exp \left(0.1 \times \log (OR) \times \Delta x_{q-0}^{NO_2} \right) \quad (6.4)$$

where $x_{q-0}^{NO_2} = x_q^{NO_2} - x_0^{NO_2}$ is the annual average NO_2 concentration in neighborhood q minus the baseline NO_2 concentration. We set $x_0^{NO_2} = 0$. The NO_2 concentrations were estimated on a $25 \times 25 \text{ m}$ grid using a regression kriging model for each hour of the day and weekdays/weekends separately. Details are described in Chapter 5. The sensor network data was not available for the entire year 2016 due to a few months of maintenance. The lack of complete daily data for 2016 also precluded linking the exposure response function with daily spatial exposure data. Therefore, an average of the modelled values for June and November was used. This is a reasonable proxy for the annual average when consulting the measurements of the central monitoring stations in Eindhoven (Table A6.6), which are part of the national ambient air quality monitoring network (RIVM, 2019b). All raster cells within neighborhood q were used to obtain the average NO_2 concentration in neighborhood q .

In Eq. 6.3, P_q is the exposed population at risk in neighborhood q . Assuming that all people are exposed to ambient NO_2 air pollutant concentrations, P_q represents the population at risk $P_q = n_q^{\text{child}} \times p_{\text{asthma}}$, where n_q^{child} is the number of children in neighborhood q , and p_{asthma} is the average proportion of asthmatic children, estimated at 0.126 based on the PIAMA birth cohort study in the Netherlands (Scholtens et al., 2009). In the absence of local data on the daily prevalence of asthmatic symptoms in asthmatic children, we estimated the prevalence rate PR at 0.17 asthmatic symptoms per person-day in the exposed population at risk, based upon the HRAPIE report (WHO, 2013a).

To obtain a measure of uncertainty propagated from the input data into the number of attributable cases, we ran MCMC simulations drawing samples from prior distributions of the input variables. For $\log (OR)$, this prior distribution equals the posterior distribution of β_{NO_2} , which was obtained through Bayesian estimation of the parameters in Eq. 6.2.

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

The number of children per neighborhood, approximated by $E(n_q^{child}) = p_q^{child} \times E(n_q^{pop})$, is retrieved from the Statistics Netherlands 2016 population data (CBS, 2018). Here, p_q^{child} is the proportion of children between 0-14 years as part of the total population n_q^{pop} . The percentage of children is considered to be measured without uncertainty. However, to avoid privacy concerns, CBS rounds the population data to a multiple of five in a random manner. Therefore, CBS provides an approximation of the number of inhabitants $E(n_q^{pop})$ for each neighborhood, and we consider the uncertainty in the true number of inhabitants per neighborhood by posing $n_q^{pop} \sim Unif(E(n_q^{pop}) - 5, E(n_q^{pop}) + 5)$. We account for uncertainty in the proportion of asthmatic children by sampling from a truncated positive normal distribution $p_{asthma} \sim N_{+\infty}(\mu = 0.126, \sigma = 0.05)$. For the prevalence rate we sample values from $PR \sim N_{+\infty}(\mu = 0.17, \sigma = 0.05)$. The modelled NO₂ exposure at 25 m raster resolution is within the 30% uncertainty required for modelling annual average values (European Parliament and Council of the European Union, 2008). Averaging to neighborhood averages however leads to added uncertainty at measurement locations, since concentrations near main roads can be substantially higher than the background concentrations in the neighborhoods. As most people live in areas with background concentrations, we consider the average uncertainty at background locations only (Figure 3.1). We sample values from the distribution $x_q^{NO2} \sim N_{+\infty}(\mu = E(x_q^{NO2}), \sigma = \sigma_q^{NO2})$, where $E(x_q^{NO2})$ is the average modelled NO₂ concentration of the raster cells in neighborhood q , and $\sigma_q^{NO2} = 2.64$ is the standard deviation of the absolute differences between modelled and observed annual average NO₂ concentration values at the background airbox locations.

We obtained the total number of LRS related to ambient NO₂ concentrations in the city, AC_{city} , by summing the number of attributable cases from all neighborhoods:

$$AC_{city} = \sum_q AC_q. \quad (6.5)$$

We then compared AC_{city} with the number of LRS which would have been obtained in a non-spatial analysis, in which only one central monitoring station is used to estimate exposure. Typically, when no fine resolution sensor network would be available, a central monitor of the national ambient air quality network would have been used for this purpose. However, the two central monitors in Eindhoven are both in traffic locations, which are not representative for the background concentrations in which most people live. We therefore use the airboxes from the sensor network in background locations to obtain a burden of disease estimate from one airbox at a time. We report the minimum and maximum values to show the variability in estimates depending on the location of the central monitoring site.

6.3 Results

6.3.1 Panel study descriptives

Despite extensive recruitment efforts, only seven children could be recruited to participate in the panel study. Two participants stopped filling in the diary within the first three days of the study and were excluded from the analysis. The remaining five children were two boys and three girls with a mean age of 9.4 yr (range 7-11 yrs). The participants were living in different areas of the city with highly different NO₂ concentrations: the mean concentration during the study period varied between 19.7 and 49.3 µg m⁻³ for the different participants, based on the closest airbox. Differences in mean NO₂ between children were accounted for using random intercepts in the model. Daily variability is therefore considered more important than variability between individuals and locations.

Days with missing NO₂ values were removed from the analysis. In total, 394 diary entries were included. Table 6.2 shows the frequency of symptoms reported in these diary entries. There were signs of clustering of symptoms within participants, but each individual symptom was reported by at least three of the five participants. One of the participants only reported one symptom throughout the entire study period, which was removed due to missing NO₂ data. The remaining non-symptomatic days of this participant were included in the analysis.

Table 6.2 Frequency of reported daily symptoms in the panel study

Symptom	Frequency (%)
Any lower respiratory symptoms (LRS)	111 (28.3 %)
Wheezing	7 (1.8 %)
Shortness of breath in rest	65 (16.5 %)
Shortness of breath after exercise	38 (9.7 %)
Waking up with breathing problems	17 (4.3 %)
Dry cough	98 (24.9 %)
Phlegm	44 (11.2 %)
Nose complaints	138 (35.0 %)
Medication use	37 (9.4 %)

6.3.2 Meta-analysis and prior selection

Weinmayr et al. (2010) included 20 studies on NO₂ and symptoms in their meta-analysis. Based on our extended literature search, we added 5 more publications with one study population each. All studies included in the meta-analysis reported effect estimates for asthmatic symptoms, with varying definitions. From the 39 study populations, effect estimates for cough were reported in 32 study populations.

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

Based on the combined meta-analysis, the combined ORs (95% confidence intervals) were 1.048 (1.023, 1.074) for asthmatic symptoms (Figure 6.2) and 0.995 (0.973, 1.018) for cough (Figure 6.3). We used prediction intervals instead of confidence intervals to obtain prior distributions. A prediction interval represents possible outcomes of single studies rather than the overall OR, and therefore allows for heterogeneity in individual studies. The prediction intervals of the meta-analysis resulted in a prior $\beta_{NO_2} \sim N(\mu = 0.047, \sigma = 0.040)$ for any lower respiratory symptoms and $\beta_{NO_2} \sim N(\mu = -0.005, \sigma = 0.039)$ for cough. We used the prior distribution for cough for the effect estimates of both dry cough and phlegm, as both were combined in the original meta-analysis (Weinmayr et al., 2010).

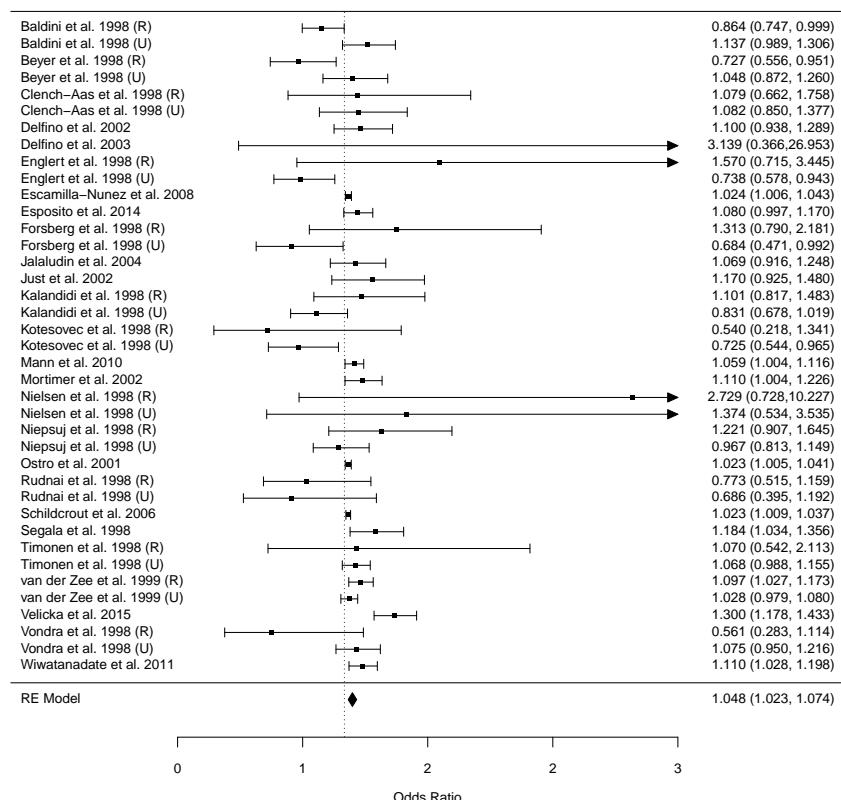


Figure 6.2 Results of the meta-analysis on NO₂ and lower respiratory symptoms (LRS). Some studies distinguished between study populations in rural (R) and urban (U) areas. OR per 10 µg m⁻³

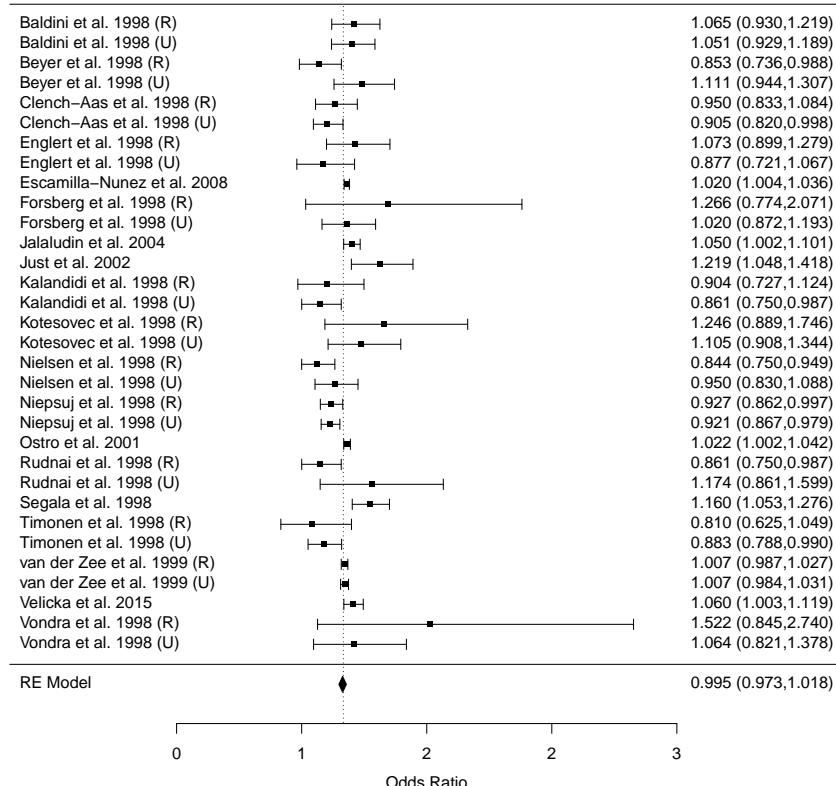


Figure 6.3 Results of the meta-analysis on NO₂ and cough. Some studies distinguished between study populations in rural (R) and urban (U) areas. OR per $10 \mu\text{g m}^{-3}$

6.3.3 Effect estimates

Table 6.3 shows the odds ratios (95% C.I.) for the different symptoms related to NO₂, obtained with uninformative priors. Despite the small number of children, we observed some associations between NO₂ and lower respiratory symptoms and medication use. ORs for cough were close to unity. This pattern of associations agrees well with the meta-analysis of previous studies. Table 6.4 shows the ORs for lag 0 obtained using informative priors based on the meta-analysis for LRS, dry cough and phlegm, in comparison to the ORs obtained using an uninformative prior. The ORs for lag 1 and mean lag 0-2 using the same informative priors are shown in Table A6.7. Convergence was achieved for all parameters in the model ($\hat{R} < 1.1$) and convergence was strong for the ORs

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

of all symptoms ($\hat{R} < 1.02$). The fastest and strongest convergence was found for the models using informative priors. The credible intervals were also much smaller when using informative priors, indicating less uncertainty about the estimates. ORs with informative prior for lag 1 and mean of lag 0-2 were similar to ORs for lag 0 (Table A6.7), reflecting the large impact of the prior.

Table 6.3 Association between NO₂ and daily symptoms, expressed as odds ratios (95% C.I.) based on panel study without informative prior information.

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	1.07 (0.92, 1.28)	1.16 (0.96, 1.40)	1.20 (0.95, 1.52)
Wheezing ^a	1.02 (0.80, 1.32)	1.00 (0.66, 1.43)	1.00 (0.63, 1.50)
Shortness of breath in rest	1.12 (0.95, 1.44)	1.23 (0.97, 1.58)	1.24 (0.93, 1.66)
Shortness of breath after exercise	1.15 (0.96, 1.55)	1.30 (0.99, 1.69)	1.50 (1.06, 2.06)
Waking up with breathing problems ^a	0.98 (0.73, 1.23)	0.79 (0.46, 1.10)	0.81 (0.46, 1.18)
Dry cough	1.03 (0.87, 1.23)	1.02 (0.83, 1.22)	1.01 (0.80, 1.27)
Phlegm	1.07 (0.87, 1.40)	1.17 (0.88, 1.59)	1.16 (0.80, 1.65)
Nose complaints	1.04 (0.88, 1.26)	1.02 (0.81, 1.25)	1.02 (0.78, 1.31)
Medication use	1.03 (0.84, 1.29)	1.35 (1.00, 1.95)	1.35 (0.97, 1.94)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5% (Table 6.1); results should be interpreted with care.

Table 6.4 Association between NO₂ at lag 0 and daily symptoms, expressed as odds ratios (95% C.I.) for LRS and cough. Comparison between local OR based on uninformative prior and OR based on informative prior

Symptom	OR based on uninformative prior	OR based on informative prior
Any lower respiratory symptoms (LRS)	1.07 (0.92-1.28)	1.06 (0.99-1.14)
Dry cough	1.03 (0.87, 1.23)	1.00 (0.93, 1.08)
Phlegm	1.07 (0.87, 1.40)	1.00 (0.93, 1.08)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

6.3.4 Burden of disease calculations

Figure 6.4 shows maps of the number of children per neighborhood, the mean NO₂ exposure per neighborhood, and the number of attributable cases per neighborhood per day, based on the association between NO₂ and LRS at lag 0 using an informative prior. In most neighborhoods, between 0 and 1 LRS per day are attributable to ambient NO₂ concentrations. In some more populated neighborhoods, this can increase up to 3 LRS per day. Some neighborhoods show 'No Data', where the percentage of children is unavailable due to privacy issues. This only occurs if the number of inhabitants <50. Since the number of children is also expected to be small here (<8), the attributable number of cases is expected to be close to zero.

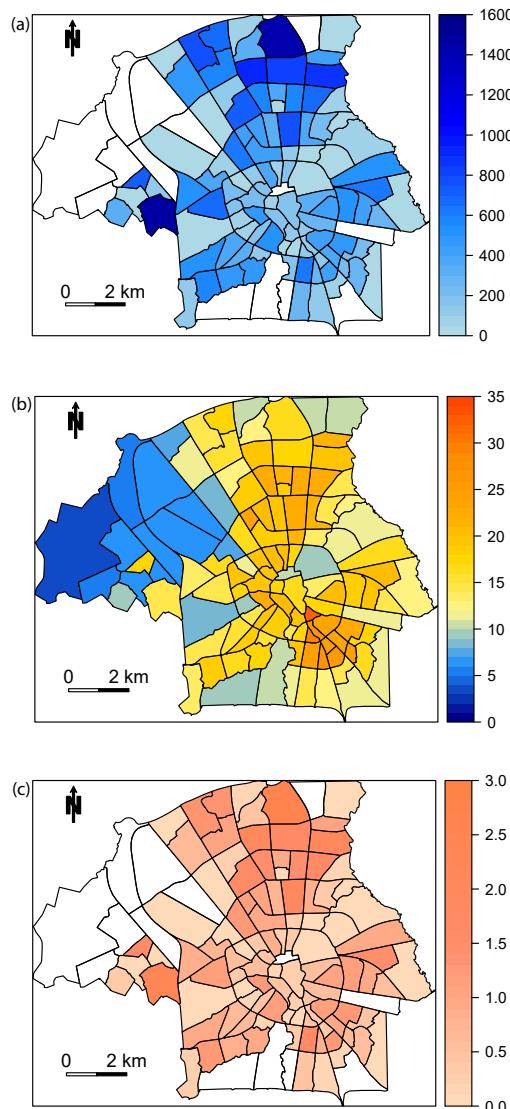


Figure 6.4 (a) Number of children per neighborhood. (b) Mean NO₂ exposure in 2016. (c) Number of lower respiratory symptoms per day attributable to ambient NO₂ concentrations in Eindhoven, per neighborhood. White areas represent ‘No Data’, which applies to neighborhoods where the percentage of children is unavailable due to privacy issues, related to a very small number of inhabitants.

6. Short-term impact of NO_2 exposure on local burden of asthmatic symptoms

The number of attributable cases of LRS per neighborhood strongly reflects the number of children per neighborhood, as expected (Pearson's correlation coefficient $r = 0.96$). NO_2 exposure is modelled using population density as one of the covariates and is therefore also, though less strongly, related to the number of children ($r = 0.27$). NO_2 exposure has a stronger relation with the number of LRS ($r = 0.45$).

Summing the AC_q for each neighborhood, $AC_{city} = 70$ cases of LRS attributable to ambient NO_2 exposure on a daily basis. When only one background concentration monitor would have been consulted to estimate the burden of disease, the number of attributable cases would vary between 45 and 73, depending on the location of the background monitor. The spatial variability in exposure and the related burden is large, however, and AC_{city} therefore represents a much better estimate of the true burden of disease.

We chose a normal distribution to represent the uncertainty in PR and a uniform distribution to represent the uncertainty in P_q . Even though no parametric distribution is posed on the OR, the posterior density function of the OR is close to a normal distribution (Figure 6.5a). This is also reflected in AF_q , which has a symmetrical distribution with long tails on both sides. Since Eq. 6.3 is a multiplicative function in which the distributions of PR , P_q and AF_q are multiplied, the uncertainty propagated in AC_q is skewed (Figure 6.5b) – a characteristic of multiply-

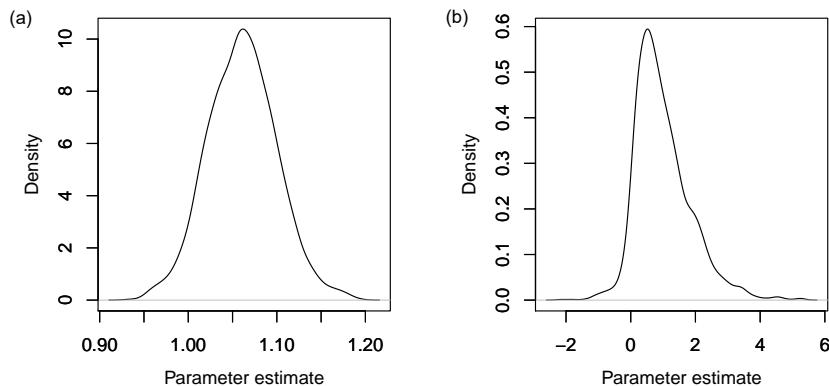


Figure 6.5 Posterior densities of (a) odds ratio of lower respiratory symptoms (LRS), and (b) the attributable cases of LRS in neighborhood 't Hofke. Neighborhood 't Hofke here serves as an example; similarly shaped posterior densities were found for the other neighborhoods.

ing normal distributions. The 95% credible interval of AC_q is therefore highly unsymmetrical around the mean.

6.4 Discussion

In this study we estimated the burden of NO_2 exposure on daily symptom prevalence in asthmatic children in the city of Eindhoven, the Netherlands. We used a Bayesian method for estimation of the effect estimates, combining prior information from an updated meta-analysis with data acquired in a local panel study. Odds ratios for lower respiratory symptoms for an increment of $10 \mu\text{g m}^{-3}$ NO_2 were 1.07 (0.92, 1.28) and 1.06 (0.99, 1.14) using an uninformative and informative prior respectively. The OR based on the meta-analysis was 1.05 (1.02, 1.07).

Although the number of participants in the panel study was low, we saw that the data affected the posterior distributions for all symptoms, by modestly pushing up the estimates of the prior distributions. The modest difference between posterior and prior distributions was due to the small size of the panel study but also because of the similarity in effect estimates between the local panel study and the meta-analytic combined estimate. The use of prior information narrowed the credible interval compared to the use of the panel study alone, in line with previous studies (Post et al., 2001; Le Tertre et al., 2005; Liu et al., 2009). Where feasible, this combination of local and generic exposure response data seems preferable to using only one of the two sources. The amount of work to recruit subjects for a panel study may however be problematic, unless strong cooperation with medical specialists is obtained. In the current setting, medical doctors were only willing to passively inform their patients by allowing leaflets in their facilities.

The priors used in this study were obtained by updating the meta-analysis by Weinmayr et al. (2010) with more recent panel studies on associations of NO_2 with daily asthmatic symptoms. . The inclusion of 5 more recent studies modestly increased the summary OR: from OR = 1.031 (95% C.I. 1.001, 1.062) in the published meta-analysis to 1.048 (1.023, 1.074) in the current review.

There were signs of heterogeneity between studies, for example due to differences in study design, exposure estimation and study area. The effect estimates from the meta-analysis were based on the most significant lag in each study, which varied between different studies. To obtain a general overall estimate to be used as prior information, we combined all studies despite their heterogeneity and different lags, and used the same combined prior for all lags. For the priors we used the prediction interval rather than the confidence interval of the random effects estimates from the meta-analysis. The prediction interval is

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

wider, representing the uncertainty of one single study rather than the uncertainty of the mean of all studies. The first is more appropriate, as we use the prior to predict the effect estimates of a single study.

We accounted for uncertainty in the number of attributable cases of LRS per neighborhood by sampling from prior distributions representing the uncertainty in the population at risk, attributable fraction and prevalence rate. The resulting posterior distribution was highly skewed. However, there may still be remaining sources of uncertainty, and the burden of disease may likely be an underestimation of the actual number of attributable cases in children with asthmatic symptoms. For example, we here only used the number of children between 0-14 years old, based on available CBS data. Also, the NO₂ exposure estimates used here are an average over each neighborhood, while concentrations near traffic roads may be substantially higher than background concentrations. Compared to the use of one central monitor to represent exposure, however, we believe that our method shows a more realistic picture of the burden of disease per neighborhood, dependent on the neighborhood-specific population at risk and relative risk.

This study has a number of strengths and limitations. Strengths include the Bayesian analysis combining prior and local information of exposure response functions and the use of a spatially refined low-cost sensor network to assess exposure more refined than in typical epidemiological panel studies (Weinmayr et al., 2010). Limitations include the small size of the local panel study, the inability to derive daily spatial maps that could have been linked with exposure response functions, and the lack of local data on number of children with asthma and the frequency of daily asthma symptoms in children with asthma. Although linking the annual average concentration with the exposure response function does not result in identical AF values compared to averaging daily AF values in a full year, the difference is likely small as β_{NO_2} is small and thus $\exp(\beta_{NO_2})$ is close to $1 + \beta_{NO_2}$.

With 70 cases of LRS per day attributable to ambient NO₂ exposure only in the city of Eindhoven, the air pollution problem is one that should not be neglected even in cities where the air pollutant concentrations are generally below European limit values. The absolute number of cases is affected by the choice to calculate the burden compared to a zero NO₂ concentration. We do not know whether associations with NO₂ extend to zero, but so far there is little evidence of a threshold in the relationship between outdoor air pollution and respiratory symptoms.

6.5 Conclusions

We conclude that a Bayesian analysis is useful to estimate location-specific air pollution effects and subsequent local burden of disease.

6.5. Conclusions

Despite the small number of participants in the panel study, we were able to derive narrow credible intervals around the effect estimates, by incorporating prior information from updating an existing meta-analysis using a Bayesian framework. With the help of a fine resolution urban air quality sensor network we were able to obtain air pollution exposure estimates close to the houses of the participants, rather than at a central location in the city. This allowed for spatial variability in the exposure of different participants. We created burden of disease maps, showing the spatial variability in the number of LRS attributable to ambient NO₂ exposure. The uncertainty propagation analysis showed that the uncertainty in the number of LRS is positively skewed. Imposing a normal distribution would have led to a biased mean and an unrealistically symmetrical 95% credible interval. By means of the Bayesian analysis we obtained more realistic estimates of the number of LRS attributable to ambient NO₂ exposure.

Appendix

Table A6.1 Odds ratios (95% C.I.) related to a 10 µg m⁻³ increase in NO₂ ambient air pollution, based on REML estimation of the model parameters

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	1.15 (0.92, 1.43)	1.19 (0.97, 1.48)	1.25 (0.97, 1.61)
Wheezing ^a	1.12 (0.67, 1.88)	1.00 (0.59, 1.68)	1.01 (0.58, 1.78)
Shortness of breath in rest	1.32 (0.99, 1.75)	1.34 (1.01, 1.76)	1.34 (0.96, 1.86)
Shortness of breath after exercise	1.43 (1.06, 1.94)	1.44 (1.08, 1.91)	1.76 (1.25, 2.49)
Waking up with breathing problems ^a	0.78 (0.44, 1.36)	0.51 (0.28, 0.94)	0.51 (0.25, 1.04)
Dry cough	1.06 (0.85, 1.34)	1.02 (0.82, 1.26)	1.01 (0.78, 1.32)
Phlegm	1.35 (0.94, 1.95)	1.35 (0.97, 1.89)	1.38 (0.94, 2.03)
Nose complaints	1.08 (0.83, 1.41)	1.00 (0.78, 1.28)	1.00 (0.74, 1.35)
Medication use	1.10 (0.79, 1.54)	1.58 (1.13, 2.19)	1.58 (1.05, 2.36)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5%; results should be interpreted with care.

Table A6.2 Odds ratios (95% C.I.) related to a 10 µg m⁻³ increase in PM₁ ambient air pollution, based on REML estimation of the model parameters

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	2.16 (0.38, 12.46)	2.41 (0.41, 14.24)	5.03 (0.54, 46.39)
Wheezing ^a	0.08 (0.00, 31.60)	0.10 (0.00, 33.18)	0.26 (0.00, 37.85)
Shortness of breath in rest	2.66 (0.16, 44.69)	3.63 (0.20, 67.40)	5.67 (0.12, 263.67)
Shortness of breath after exercise	2.91 (0.34, 25.06)	3.04 (0.34, 26.85)	8.09 (0.53, 122.69)
Waking up with breathing problems ^a	0.03 (0.00, 28.01)	1.00 (0.00, 231.55)	0.02 (0.00, 91.86)
Dry cough	0.75 (0.16, 3.48)	1.01 (0.22, 4.62)	1.24 (0.19, 8.14)
Phlegm	2.45 (0.13, 45.21)	1.14 (0.04, 30.96)	1.35 (0.03, 69.87)
Nose complaints	1.05 (0.14, 7.83)	0.14 (0.01, 1.53)	0.24 (0.02, 3.46)
Medication use	0.00 (0.00, 0.67)	0.03 (0.00, 8.45)	0.00 (0.00, 0.00)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5%; results should be interpreted with care.

6. Short-term impact of NO₂ exposure on local burden of asthmatic symptoms

Table A6.3 Odds ratios (95% C.I.) related to a 10 µg m⁻³ increase in PM_{2.5} ambient air pollution, based on REML estimation of the model parameters

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	1.14 (0.34, 3.80)	1.53 (0.44, 5.30)	2.44 (0.55, 10.89)
Wheezing ^a	0.54 (0.03, 8.85)	0.33 (0.01, 9.24)	0.55 (0.03, 11.47)
Shortness of breath in rest	1.39 (0.17, 11.72)	2.19 (0.25, 19.44)	3.06 (0.21, 44.29)
Shortness of breath after exercise	1.21 (0.28, 5.20)	1.71 (0.40, 7.40)	2.94 (0.51, 16.91)
Waking up with breathing problems ^a	0.37 (0.01, 17.69)	2.29 (0.08, 64.71)	0.96 (0.01, 66.62)
Dry cough	0.95 (0.33, 2.72)	0.87 (0.29, 2.60)	1.12 (0.30, 4.11)
Phlegm	3.36 (0.78, 14.55)	1.53 (0.21, 10.86)	2.30 (0.26, 20.35)
Nose complaints	1.26 (0.36, 4.43)	0.39 (0.10, 1.61)	0.45 (0.08, 2.40)
Medication use	0.05 (0.00, 2.16)	0.10 (0.00, 3.55)	0.00 (0.00, 0.07)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5%; results should be interpreted with care.

Table A6.4 Odds ratios (95% C.I.) related to a 10 µg m⁻³ increase in PM₁₀ ambient air pollution, based on REML estimation of the model parameters

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	0.84 (0.31, 2.28)	1.24 (0.46, 3.35)	1.67 (0.47, 5.95)
Wheezing ^a	0.93 (0.15, 5.64)	0.57 (0.07, 4.43)	0.82 (0.10, 6.44)
Shortness of breath in rest	1.28 (0.23, 7.06)	3.00 (0.57, 15.79)	5.12 (0.57, 46.05)
Shortness of breath after exercise	0.75 (0.21, 2.70)	1.23 (0.36, 4.26)	1.88 (0.38, 9.24)
Waking up with breathing problems ^a	0.79 (0.10, 6.55)	1.70 (0.20, 14.79)	1.21 (0.09, 17.04)
Dry cough	1.29 (0.57, 2.92)	0.88 (0.37, 2.06)	1.25 (0.43, 3.59)
Phlegm	2.73 (0.77, 9.63)	1.06 (0.23, 5.02)	1.67 (0.28, 9.94)
Nose complaints	1.13 (0.41, 3.12)	0.46 (0.15, 1.39)	0.45 (0.11, 1.86)
Medication use	0.45 (0.05, 3.76)	0.99 (0.14, 7.14)	0.16 (0.01, 1.94)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5%; results should be interpreted with care.

Table A6.5 Odds ratios (95% C.I.) related to a 10000 particle # increase in UFP ambient air pollution, based on REML estimation of the model parameters

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	1.00 (0.72, 1.37)	1.12 (0.81, 1.54)	1.03 (0.71, 1.48)
Wheezing ^a	0.65 (0.13, 3.17)	0.56 (0.09, 3.32)	0.67 (0.12, 3.85)
Shortness of breath in rest	1.28 (0.86, 1.93)	1.34 (0.88, 2.03)	1.24 (0.77, 1.99)
Shortness of breath after exercise	0.99 (0.57, 1.73)	1.19 (0.72, 1.98)	1.23 (0.70, 2.15)
Waking up with breathing problems ^a	0.14 (0.01, 2.55)	0.07 (0.00, 2.66)	0.10 (0.00, 2.99)
Dry cough	0.55 (0.34, 0.89)	0.32 (0.16, 0.63)	0.36 (0.18, 0.72)
Phlegm	0.76 (0.40, 1.44)	0.55 (0.23, 1.30)	0.51 (0.19, 1.35)
Nose complaints	1.13 (0.80, 1.60)	0.83 (0.56, 1.23)	1.02 (0.67, 1.54)
Medication use	0.83 (0.39, 1.79)	1.11 (0.59, 2.07)	0.73 (0.27, 1.99)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

^a Symptom prevalence < 5%; results should be interpreted with care.

Table A6.6 Comparison of June, November, mean June and November, and annual mean concentrations measured at the two RIVM reference monitors in Eindhoven, 2016. Concentrations in µg m⁻³

Monitoring station	Jun. mean	Nov. mean	Jun.+Nov. mean	Annual mean
Genovevalaan	19.0	30.4	24.7	25.7
Noord-Brabantlaan	25.8	36.5	31.2	31.5

Table A6.7 Association between NO₂ and daily symptoms, expressed as odds ratios (95% C.I.) based on panel study and prior information from the meta-analysis on LRS and cough. The same prior is used for all lags.

Symptom	Lag 0	Lag 1	Mean lag 0-2
Any lower respiratory symptoms (LRS)	1.06 (0.99, 1.14)	1.07 (0.99, 1.15)	1.07 (0.99, 1.14)
Dry cough	1.00 (0.93, 1.07)	1.00 (0.93, 1.08)	1.00 (0.93, 1.08)
Phlegm	1.00 (0.93, 1.08)	1.01 (0.93, 1.08)	1.00 (0.93, 1.08)

All OR are adjusted for daily temperature, relative humidity, day of follow-up, daily flu, weekday/weekend day and participant ID.

Synthesis

7

7.1 Main findings

In this thesis I have investigated the usability of a low-cost air quality sensor network. The research comprised key steps from data collection to application. It focused on outlier detection, calibration, spatio-temporal modelling and health effect assessment. In this section, my aim is to summarize the research objectives and major findings, and to guide further discussion of their implications in the next section.

Research objective 1: To develop an outlier detection method suitable to detect outliers in space and time while accounting for the large spatio-temporal variability of air pollutant concentrations in an urban area.

Air pollutant concentrations show a large spatial and temporal variability within urban areas. This spatio-temporal variability argues against the use of a single threshold level for the definition of outliers. This study proposes a spatio-temporal classification of observations, where each class of observations has its own truncated normal distribution after transformation. The underlying normal distributions are used to obtain threshold levels for each spatio-temporal class individually. The proposed method provided spatio-temporally varying threshold levels for outlier detection, well reflecting the spatio-temporal variability in NO₂ concentrations. Spatio-temporal correlations could be used to further distinguish between errors and events.

Research objective 2: To develop and evaluate automatic calibration methods for low-cost sensors in an urban air quality sensor network, accounting for drift and interference effects.

Low-cost electrochemical gas sensors are prone to loose sensitivity over time (drift) and show signs of interference effects from other gases. Regular calibration can help to minimize these effects. Calibration can, however, only take place at locations where reference measurements are available. Calibration parameters need to be transferred from those locations to locations without available reference data. In this study I developed an automatic calibration method and assessed two existing calibration methods. The methods were evaluated in terms of temporal

7. Synthesis

stability, spatial transferability, and sensor specificity. All methods showed a poor spatial transferability, consistent with different responses of individual sensors to environmental factors such as temperature and relative humidity. These results led to the advice to regularly recalibrate sensors using a moving reference sensor.

Research objective 3: To develop a spatio-temporal kriging framework for modelling air pollutant concentrations using a low-cost sensor network. Sensor observations are limited to point support. This study proposes spatio-temporal regression kriging to interpolate air pollutant concentration values between observed space-time locations. First, a set of spatial and temporal covariates was selected. Population density, road type, easting coordinate, wind speed, wind direction, relative humidity, hour of the day and weekday/weekend were useful predictor variables. Second, the residuals of the linear regression model were checked for spatio-temporal autocorrelation, and the resulting space-time variogram was used for interpolation. The method provided local estimates of the strength and association of air pollution sources and sinks, and allowed for near real-time prediction and spatio-temporal mapping of air pollutant concentrations. This research also led to advices for sampling scheme optimization in future sensor networks.

Research objective 4: To create burden of disease maps, expressing the spatial variability in health risks related to ambient air pollution.

Burden of disease maps can be used to express the spatial variability in health risks related to environmental exposure. In this study, asthma symptom exacerbation was related to air pollutant exposure. Population data, symptom prevalence data, exposure data and exposure-response functions were combined to create burden of disease maps. A panel study was set up to obtain estimates of the exposure-response functions. Despite the low number of participants recruited in the panel study, accurate Bayesian estimates of the exposure-response functions and local burden of disease could be obtained by combining the panel study data with information from an updated meta-analysis of the literature. Burden of disease maps showed a strong spatial variability in the number of asthmatic symptoms associated with ambient NO₂.

7.2 Significance

The number of low-cost sensor networks is growing and new insights support further development progress. This thesis contributes to the further development of low-cost sensor networks. It firstly does so by providing methods for data quality evaluation, outlier detection, calibration, modelling and health risk assessment. Secondly, it provides guidance to improve calibration frameworks and to optimize sampling schemes. These results are useful to sensor developers and sensor

network developers, as well as researchers in this field. The development of sensor networks can lead to new applications in the future, such as healthy cyclist route planning and dynamic smart cities.

This thesis provided key steps to assess the data quality of low-cost sensors and to visualize air quality in space and time (Figure 7.1). Sensor observations have spatial and temporal locations, and so do health effects of air pollution. These can be linked based on their space-time coordinates, for example to create exposure-effect estimates and health risk maps. The quality of the sensor observations is reflected in the results. Data quality evaluation, through assessment of drift, interference effects, accuracy and outliers, as well as calibration, is therefore an important aspect of low-cost sensor data analysis.

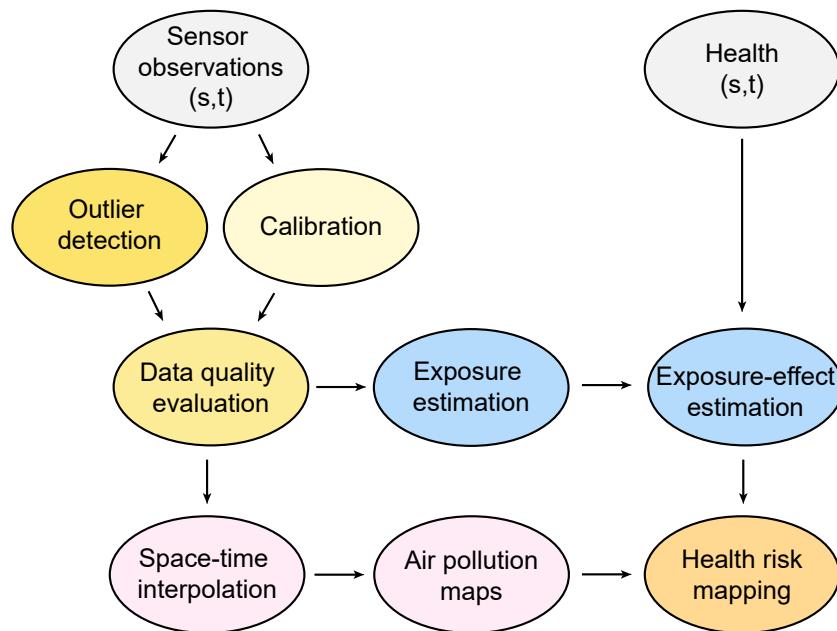


Figure 7.1 Framework of this thesis. The grey ellipses refer to inputs with a spatio-temporal location (s,t). The yellow ellipses refer to work related to data quality assessment, including outlier detection in darker yellow (Chapter 3) and calibration in lighter yellow (Chapter 4). The pink ellipses refer to modelling and prediction work (Chapter 5). The blue ellipses refer to health effect assessment (Chapter 6) and the orange ellipse to refers to a combination of the work in Chapters 5 and 6.

After data quality evaluation and outlier removal, air quality maps are able to provide policy makers with valuable information on air pollution events in space and time. Since the trend part of the model is based

7. Synthesis

on a monthly average, the residuals account for short-term variability. In case of polluting events, the contribution of the kriged residuals to the predictions will be higher compared to the trend part of the model. The fine resolution air pollution maps can be used to create awareness about highly polluting events such as fireworks, and may help to start a discussion on the trade-off between traditions and health effects. In the long run, spatial locations with consistently high concentrations can be marked. This information can lead to infrastructural decisions, for example reducing speed limits, converting traffic lights to roundabouts, and adding or removing road lanes. Combined with temporal air pollution trends and real-time pollution data, spatio-temporal models can lead to smart city interventions such as dynamic speed limits and dynamic directions based on current traffic and air pollution patterns.

The temporal covariates used in the air pollution model relate to meteorological and diurnal variability. Spatial covariates, such as road type and population density, are related to city planning. Information about important spatial predictor variables of air pollution can therefore help city planners to develop cleaner and healthier cities.

For people with respiratory diseases such as asthma, real-time maps and forecasts of the air pollution can be used to advice on the optimal time and location to move through the city. This research provided critical steps towards real-time mapping and forecasting, which can be further developed in e.g. smartphone applications. The burden of disease calculations can be used to create awareness on the severity of the health effects attributable to air pollution. The corresponding maps show the spatial variability of the burden of disease.

7.3 Limitations

Throughout the research conducted as part of this thesis, useful lessons were learnt on sensor data quality, calibration and sampling scheme optimization. Part of the developed methods could directly be applied – for example, the removal of outliers in the development of calibration methods. Some, however, could not directly be applied. For example, we found that calibration of the low-cost sensors is not possible in the current measurement setup, due to sensor specificity and a lack of spatial transferability of the calibration parameters. The advice to use moving reference sensors would be useful in the future development of the network and other sensor networks worldwide, but could not be applied in the current network before a further data analysis. In general, however, the sensors performed well and the data have been valuable for modelling and health risk analysis. In the future, more stable and consistent sensors may be developed and implemented.

The advantage of low-cost sensor networks is the large number of sensors that can be deployed at relatively low cost, resulting in data at a fine spatial and temporal resolution (Snyder et al., 2013). One may argue, however, that the costs of calibrating and maintaining a large number of sensors quickly rises to a level that the network can no longer be considered ‘low-cost’. Moving reference sensors would for example be beneficial in low-cost sensor networks, but expensive to implement. In the future, this may be solved by, for example, mounting reference sensors on buses or other public transport vehicles (Hasenfratz et al., 2012). In the sensor network used in this study, transportation of the sensors for maintenance was relatively expensive. Therefore, sensors were only repaired after a few were broken – or fixed – so they could be transported at the same time. This sometimes led to long periods of missing data. This is undesirable when interpolating and modelling health effects. The method used to impute missing values, however, was able to fill large data gaps.

Throughout this thesis, a spatial location is assumed to consist of a two-dimensional set of coordinates. With this assumption, the vertical position of the sensors is ignored. As long as those are similar across the sensor network, and the height of the sensor closely represents exposure of the population at risk, this issue can be ignored. In practice, however, concentrations higher in apartment buildings differ from those at ground level (Amato et al., 2019). Further research could focus on three-dimensional modelling and mapping, for example using sensors at different levels of tall buildings (Azimi et al., 2018) or using unmanned aerial vehicles (Li et al., 2018b,a).

A spatio-temporal regression kriging framework was used for modeling and prediction of air pollutant concentrations in this thesis. The trend part of the model is very similar to a LUR model. In the LUR literature, typically tens of variables are evaluated at different buffer sizes to select the best covariates in the model (Montagne et al., 2015; Lee et al., 2017; Kashima et al., 2018). Acknowledging that covariate selection was not the main purpose of this study, this thesis only evaluated a small number of variables and buffer sizes based on the ones most often related to NO₂. It is possible that a larger part of the spatio-temporal variability can be explained after a full assessment of possible LUR covariates. This could also create a smoother spatio-temporal variogram of the residuals.

One of the major challenges faced in this thesis was the recruitment of participants for the panel study. Despite extensive efforts and contacts with local newspapers, health care practitioners and schools, only five children participated in the panel study while the proposal aimed at 100. Given the attention given to the study via newspapers, social media, websites, posters and leaflets, the study should have reached the target

7. Synthesis

group of potential participants. The parents of many asthmatic children apparently did not see the advantages of participation. A possible reason could be the burden of participation for four months compared to the perceived benefits. Further research on this aspect is required to avoid this happening in future studies, as much money and time is spent on obtaining medical ethical permission and recruitment of participants. In this thesis, narrow credible intervals were found for the odds ratio's relating NO₂ exposure to asthmatic symptoms, when combining the data with a priori information from literature in a Bayesian analysis. When more participants would have been recruited, more accurate estimates could have been made for other air pollutants as well.

The main focus of this thesis has been on NO₂, while data on other pollutants is also collected in the sensor network. This choice was made largely because of the spatio-temporal variability in NO₂, which within the city is much larger than spatio-temporal variability in particulate matter (PM₁₀, PM_{2.5}, PM₁). Low-cost sensors were not available for particle metrics that do show large spatio-temporal variation, such as UFP and BC. UFP was measured at only six locations precluding detailed analyses, while BC was not measured at all. For modelling and prediction, it is important that the spatio-temporal variability is higher than the measurement error of the sensors. This is most prevalent in the NO₂ measurements, as is also reflected in the results of the health effects estimation. Since the effects of single pollutants are hard to distinguish (WHO, 2013a,b), one may also argue to evaluate the mixture of pollutants rather than specific pollutants.

7.4 Prospects

This thesis provides methods and recommendations to improve the development and use of low-cost air quality sensor networks. There are, however, remaining challenges for future research. In terms of methodological development, future work may extend the current work on outlier detection, by distinguishing between erroneous values and true air pollution events. The first steps have been made to accomplish this. The spatio-temporal air pollution maps allow for visual inspection of outliers in space and time. When extreme values occur for longer periods of time or at multiple locations in space, they are more likely to represent actual air pollution events (Zhang et al., 2012). When a single observation is an outlier, it is more difficult to distinguish errors from events. Correlation with other air pollutants may be useful to obtain, for example, further conclusions on the reliability of individual measurements (Shahid et al., 2015).

Some suggestions on sampling scheme optimization are already given in this thesis. Further research is needed to optimize those for different

pollutants, for example to determine air pollution level degradation as a local function of distance to the road (Amato et al., 2019). For different sources of air pollution, depending upon the location and pollutant, other covariates may be of interest. Sampling scheme optimization should target multiple aspects. Sampling near the population at risk is beneficial for epidemiological studies, as measurements are used to approximate exposure (Chambers et al., 2018). Meanwhile, for optimal geostatistical modelling, a varying distance between sensors locations is important (Webster and Oliver, 2001). For accurate regression models, e.g. as trend part in a regression kriging model, covariates should represent the sources and sinks of air pollution. Sampling at different distances of the sources leads to better regression models. All levels of a covariate should be covered by multiple sensors to avoid overfitting. The latter is important to support accurate predictions.

The methods developed in this thesis are applicable to other low-cost air quality sensor networks and may as well be more generally applicable to other environmental phenomena. This will require a good understanding of the other variables (e.g. sources and sinks) that are related to spatio-temporal variability of the location and phenomenon of interest. Recalibration of the parameters and reconsideration of covariates and classes is always required. Moreover, when working with low-cost sensors, the data quality of the sensors should be carefully evaluated for the purpose at hand (Snyder et al., 2013). The precision and accuracy of the low-cost sensors can be further evaluated with reference monitors (Thomson et al., 2005). In case the absolute measurement values are close to the reference values, they can be used to evaluate the exceedance of limit values. When the accuracy is too low for this purpose, the relative values of the sensors could still be useful to assess spatio-temporal variability, under the condition that the uncertainty of the sensor measurements is below the spatio-temporal variability of the phenomenon of interest. This can be an air pollutant or another variable measured using low-cost sensors.

The maps created throughout this research project can be used to visualize many air pollution problems. Visualizations can help to create awareness about their health effects. When anthropogenic sources of air pollutants are clearly visible, maps also create awareness about the human influence on these air pollution risks. This awareness is often not enough, however, for people to change their behavior. Behavioral changes require more than a sense of awareness – they require a sense of urgency. This urgency may be lacking in developed countries, as the life expectancy is already high, and quality of life is considered more important than quantity of life. More research, for example in the field of behavioral science, is needed to create the sense of urgency that is needed to tackle the air pollution problem.

7. Synthesis

In the advancements of smart cities and sensor networks, more and more data will become available. This will increase the importance of big data analysis methods, data quality analysis, pattern analysis and filtering (Lau et al., 2019). With increasing computational power, it is possible to process larger datasets in shorter amounts of time, allowing for near real-time modelling and mapping at ever finer spatio-temporal resolutions. Moreover, it will allow for the integration of data from multiple sources and scale levels (Xu et al., 2019). When more applications emerge, the demand for low-cost sensors will increase. This will contribute to the development of higher quality sensors at lower cost. In turn, the higher quality of the sensor data will propagate into better products, such as more accurate models and exposure estimates.

Bibliography

- Abbey, D.E., Nishino, N., McDonnell, W.F., Burchette, R.J., Knutson, S.F., Lawrence Beeson, W., Yang, J.X., 1999. Long-term inhalable particles and other air pollutants related to mortality in non-smokers. *American Journal of Respiratory and Critical Care Medicine* 159, 373–382. doi:10.1164/ajrccm.159.2.9806020.
- AiREAS, 2016. Aireas api v2 documentation. <http://data.aireas.com/docs/>. [Accessed 23rd of July 2019].
- Amato, F., Pérez, N., López, M., Ripoll, A., Alastuey, A., Pandolfi, M., Karanasiou, A., Salmatidis, A., Padoan, E., Frasca, D., Marcoccia, M., Viana, M., Moreno, T., Reche, C., Martins, V., Brines, M., Minguillón, M.C., Ealo, M., Rivas, I., van Drooge, B., Benavides, J., Craviotto, J.M., Querol, X., 2019. Vertical and horizontal fall-off of black carbon and NO₂ within urban blocks. *Science of The Total Environment* 686, 236–245. doi:10.1016/j.scitotenv.2019.05.434.
- Austin, E., Novoselov, I., Seto, E., Yost, M.G., 2015. Laboratory evaluation of the Shinyei PPD42NS low-cost particulate matter sensor. *PLOS one* 10. doi:10.1371/journal.pone.0137789.
- Azimi, P., Zhao, H., Fazli, T., Zhao, D., Faramarzi, A., Leung, L., Stephens, B., 2018. Pilot study of the vertical variations in outdoor pollutant concentrations and environmental conditions along the height of a tall building. *Building and Environment* 138, 124–134. doi:10.1016/j.buildenv.2018.04.031.
- Barakeh, Z.A., Breuil, P., Redon, N., Pijolat, C., Locoge, N., Viricelle, J.P., 2016. Development of a normalized multi-sensors system for low cost on-line atmospheric pollution detection. *Sensors and Actuators B: Chemical* 241, 1235–1243. doi:10.1016/j.snb.2016.10.006.
- Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment* 54, 634–642. doi:10.1016/j.atmosenv.2012.01.064.
- Basu, S., Meckesheimer, M., 2007. Automatic outlier detection for time series: an application to sensor data. *Knowledge and Information Systems* 11, 137–154. doi:10.1007/s10115-006-0026-6.

Bibliography

- Beach, J., Burstyn, I., Cherry, N., 2012. Estimating the extent and distribution of new-onset adult asthma in British Columbia using frequentist and Bayesian approaches. *Annals of Work Exposures and Health* 56, 719–727. doi:10.1093/annhyg/mes004.
- Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Science of the Total Environment* 407, 1852–1867. doi:10.1016/j.scitotenv.2008.11.048.
- Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z.J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet* 383, 785–795. doi:10.1016/S0140-6736(13)62158-3.
- Bigi, A., Harrison, R.M., 2010. Analysis of the air pollution climate at a central urban background site. *Atmospheric Environment* 44, 2004–2012. doi:10.1016/j.atmosenv.2010.02.028.
- Bivand, R.S., Pebesma, E.J., Gomez-Rubio, V., 2013. Applied Spatial Data Analysis with R. volume 747248717. Springer Science+Business Media, New York.
- Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-Temporal Bayesian Models with R-INLA. John Wiley & Sons, Chichester.
- Bobbia, M., Misiti, M., Misiti, Y., Poggi, J.M., Portier, B., 2015. Spatial outlier detection in the PM10 monitoring network of Normandy (France). *Atmospheric Pollution Research* 6, 476–483. doi:10.5094/apr.2015.053.
- Brandt, E.B., Myers, J.M.B., Ryan, P.H., Hershey, G.K.K., 2015. Air pollution and allergic diseases. *Current Opinion in Pediatrics* 27, 724–735. doi:10.1097/mop.0000000000000286.
- Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science* 11, 699–718. doi:10.1080/136588197242158.
- Broday, D.M., The Citi-Sense Project Collaborators, 2017. Wireless distributed environmental sensor networks for air pollution measurement – the promise and the current reality. *Sensors* 17. doi:10.3390/s17102263.
- Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation* 121, 2331–2378. doi:10.1161/CIR.0b013e3181dbece1.
- Brown, R.J.C., Brown, A.S., 2012. Principal component analysis as an outlier detection tool for polycyclic aromatic hydrocarbon concentrations in ambient air. *Water, Air, & Soil Pollution* 223, 3807–3816. doi:10.1007/s11270-012-1149-x.

Bibliography

- Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. *The Lancet* 360, 1233-1242.
- Buijsman, E., 2013. De luchtkwaliteitsmeetnetten van het RIV en het RIVM. LUVO reeks 16, Tinsentiep, Houten.
- Burns, C.J., Wright, J.M., Pierson, J.B., Bateson, T.F., Burstyn, I., Goldstein, D.A., Klaunig, J.E., Luben, T.J., Mihlan, G., Ritter, L., 2014. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. *Environmental Health Perspectives* 122, 1160-1165. doi:10.1289/ehp.1308062.
- Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 1190-1208. doi:10.1137/0916069.
- Caselton, W.F., Zidek, J.V., 1984. Optimal monitoring network designs. *Statistics & Probability Letters* 2, 223-227. doi:10.1016/0167-7152(84)90020-8.
- CBS, 2018. Wijk- en buurtkaart 2016 versie 3. <http://nationaalgeoregister.nl/geonetwork/srv/dut/catalogsearch#/metadata/6cde2ce7-3ecd-4785-b7b8-d62c3381efe9>. [Accessed 6th of September 2018].
- CEN/TC 287, 1998. ENV 12656:1998 Geographic Information - data description - quality.
- Chambers, L., Finch, J., Edwards, K., Jeanjean, A., Leigh, R., Gonem, S., 2018. Effects of personal air pollution exposure on asthma symptoms, lung function and airway inflammation. *Clinical and Experimental Allergy* 48, 798-805. doi:10.1111/cea.13130.
- Chandola, V., Banerjee, A., Kumar, V., 2009. Anomaly detection: A survey. *ACM computing surveys (CSUR)* 41. doi:10.1145/1541880.1541882.
- Close, J.P. (Ed.), 2016. AiREAS: Sustainocracy for a Healthy City. The Invisible made Visible Phase 1. SpringerBriefs on Case Studies of Sustainable Development, Springer. doi:10.1007/978-3-319-26940-5.
- Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 389, 1907-1918. doi:10.1016/S0140-6736(17)30505-6.
- Colin, F., Shepherd, P.D., Carter, T.J.N., Wright, J.D., 1998. Development of a piezo-optical chemical monitoring system for nitrogen dioxide. *Sensors and Actuators B: Chemical* 51, 244-248. doi:10.1016/s0925-4005(98)00201-9.

Bibliography

- Colvile, R., Woodfield, N., Carruthers, D., Fisher, B., Rickard, A., Neville, S., Hughes, A., 2002. Uncertainty in dispersion modelling and urban air quality mapping. *Environmental Science & Policy* 5, 207-220. doi:10.1016/S1462-9011(02)00039-4.
- Cressie, N., Wikle, C.K., 2011. Statistics for spatio-temporal data. John Wiley & Sons, Hoboken (New Jersey).
- Cui, Y.Z., Lin, J.T., Song, C.Q., Liu, M.Y., Yan, Y.Y., Xu, Y., Huang, B., 2016. Rapid growth in nitrogen dioxide pollution over Western China, 2005-2013. *Atmospheric Chemistry and Physics* 16, 6207-6221. doi:10.5194/acp-16-6207-2016.
- Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dédelé, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Graulevičienė, R., Grivas, G., Gruzieva, O., Gustafsson, A.H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Möller, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E.G., Taimisto, P., Tsai, M.Y., Vaskövi, É., Villani, S., Wang, M., Brunekreef, B., Hoek, G., 2012. Variation of NO₂ and NO_x concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmospheric Environment* 62, 374-390. doi:10.1016/j.atmosenv.2012.07.080.
- Dales, R., Chen, L., Frescura, A.M., Liu, L., Villeneuve, P.J., 2009. Acute effects of outdoor air pollution on forced expiratory volume in 1 s: a panel study of schoolchildren with asthma. *European Respiratory Journal* 34, 316-323. doi:10.1183/09031936.00138908.
- De Vito, S., Piga, M., Martinotto, L., Di Francia, G., 2009. CO, NO₂ and NO_x urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization. *Sensors and Actuators B: Chemical* 143, 182-191. doi:10.1016/j.snb.2009.08.041.
- DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials* 7, 177-188. doi:10.1016/0197-2456(86)90046-2.
- Diggle, P., Ribeiro, P., 2007. Model-based geostatistics. Springer Series in Statistics, Springer, New York.
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine* 329, 1753-1759. doi:10.1056/NEJM199312093292401.
- Drummond, J., 1995. Positional accuracy. Elsevier Science Ltd, Oxford. pp. 31-58.
- EC Working Group on GDE, 2010. Guide to the demonstration of equivalence of ambient air monitoring methods.

Bibliography

- Eeftens, M., Tsai, M.Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dédelé, A., Eriksen, K., Galassi, C., Gražulevičienė, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Kat-souyanni, K., Korek, M., Krämer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Möller, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskövi, É., Wein-mayr, G., Brunekreef, B., Hoek, G., 2012. Spatial variation of PM_{2.5}, PM₁₀, PM_{2.5} absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO₂ — results of the ESCAPE project. *Atmospheric Environment* 62, 303–317. doi:10.1016/j.atmosenv.2012.08.038.
- European Parliament and Council of the European Union, 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Union L152 51, 1–44.
- Frerero, M., Galeano, P., Gonzalez-Manteiga, W., 2007. A functional analysis of NO_x levels: Location and scale estimation and outlier detection. *Computational Statistics* 22, 411–427. doi:10.1007/s00180-007-0048-x.
- Frerero, M., Galeano, P., Gonzalez-Manteiga, W., 2008. Outlier detection in functional data by depth measures, with application to identify abnormal NO_x levels. *Environmetrics* 19, 331–345. doi:10.1002/env.878.
- Fenger, J., 2009. Urban Air Pollution. Wiley & Sons Ltd., Chichester. pp. 243–267.
- Fishbain, B., Lerner, U., Castell, N., Cole-Hunter, T., Popoola, O., Broday, D.M., Iñiguez, T.M., Nieuwenhuijsen, M., Jovasevic-Stojanovic, M., Topalovic, D., Jones, R.L., Galea, K.S., Etzion, Y., Kizel, F., Columbic, Y.N., Baram-Tsabari, A., Yacobi, T., Drahler, D., Robinson, J.A., Kocman, D., Horvat, M., Svecova, V., Arpacı, A., Bartonova, A., 2017. An evaluation tool kit of air quality micro-sensing units. *Science of The Total Environment* 575, 639–648. doi:10.1016/j.scitotenv.2016.09.061.
- Foody, G.M., 2003. Uncertainty, knowledge discovery and data mining in GIS. *Progress in Physical Geography* 27, 113–121. doi:10.1091/0309133303pp345pr.
- Forsberg, B., Stjernberg, N., Linne, R., Segerstedt, B., Wall, S., 1998. Daily air pollution levels and acute asthma in southern Sweden. *European Respiratory Journal* 12, 900–905. doi:10.1183/09031936.98.12040900.
- Gehring, U., Wijga, A.H., Brauer, M., Fischer, P., de Jongste, J.C., Kerkhof, M., Oldenwening, M., Smit, H.A., Brunekreef, B., 2010. Traffic-related air pollution and the development of asthma and allergies during the first 8 years of life. *American Journal of Respiratory and Critical Care Medicine* 181, 596–603. doi:10.1164/rccm.200906-0858OC.

Bibliography

- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–472. doi:10.1214/ss/1177011136.
- Goldizen, F.C., Sly, P.D., Knibbs, L.D., 2016. Respiratory effects of air pollution on children. *Pediatric Pulmonology* 51, 94–108. URL: 10.1002/ppul.23262, doi:10.1002/ppul.23262.
- Graham, L.M., 2004. All I need is the air that I breath: Outdoor air quality and asthma. *Paediatric Respiratory Reviews* 5, Supplement 1, S59–S64. doi:10.1016/S1526-0542(04)90012-7.
- Gräler, B., Pebesma, E., Heuvelink, G., 2016. Spatio-temporal interpolation using gstat. *The R Journal* 8, 204–218.
- Guarnieri, M., Balmes, J.R., 2014. Outdoor air pollution and asthma. *The Lancet* 383, 1581–1592. doi:10.1016/S0140-6736(14)60617-6.
- Gulliver, J., Briggs, D., 2004. Personal exposure to particulate air pollution in transport microenvironments. *Atmospheric Environment* 38, 1–8. doi:10.1016/j.atmosenv.2003.09.036.
- Guo, D., Guo, R., Thiart, C., 2007. Predicting air pollution using fuzzy membership grade Kriging. *Computers, Environment and Urban Systems* 31, 33–51. doi:10.1016/j.compenvurbsys.2005.07.006.
- Guptill, S.C., Morrison, J.L., 1995. Elements of Spatial Data Quality. Elsevier Science Ltd, Oxford.
- Hamm, N.A.S., 2016. Spatial Temporal Modelling of Particulate Matter for Health Effects Studies. volume 41 of *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*. pp. 1403–1406. doi:10.5194/isprs-archives-XLI-B8-1403-2016.
- Hamm, N.A.S., Finley, A.O., Schaap, M., Stein, A., 2015. A spatially varying coefficient model for mapping PM10 air quality at the European scale. *Atmospheric Environment* 102, 393–405. doi:10.1016/j.atmosenv.2014.11.043.
- Hamm, N.A.S., Van Lochem, M., Hoek, G., Otjes, R., Van der Sterren, S., Verhoeven, H., 2016. “The Invisible Made Visible”: Science and Technology. Springer. SpringerBriefs on Case Studies of Sustainable Development, pp. 51–78. doi:10.1007/978-3-319-26940-5.
- Harrell, F., 2018. Function aregImpute. <https://cran.r-project.org/web/packages/Hmisc/Hmisc.pdf>. [Accessed 11th of November 2018].
- Hasenfratz, D., Saukh, O., Thiele, L., 2012. On-the-fly calibration of low-cost gas sensors, in: Picco, G., Heinzelman, W. (Eds.), European Conference on Wireless Sensor Networks, Springer, Berlin Heidelberg. pp. 228–244.
- Health Effects Institute, 2010. Traffic-related Air Pollution: a Critical Review of the Literature on Emissions, Exposure and Health Effects. Report. Health Effects Institute. Boston.

Bibliography

- Heuvelink, G.B., 1998. Error Propagation in Environmental Modelling with GIS. Taylor & Francis, London.
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric environment* 42, 7561–7578. doi:10.1016/j.atmosenv.2008.05.057.
- Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., van den Brandt, P.A., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *The Lancet* 360, 1203–1209. doi:10.1016/S0140-6736(02)11280-3.
- Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environmental Health* 12, 43. doi:10.1186/1476-069X-12-43.
- Hu, Y., Li, R., Bergquist, R., Lynn, H., Gao, F., Wang, Q., Zhang, S., Sun, L., Zhang, Z., Jiang, Q., 2015. Spatio-temporal transmission and environmental determinants of Schistosomiasis Japonica in Anhui province, China. *PLOS Neglected Tropical Diseases* 9, e0003470. doi:10.1371/journal.pntd.0003470.
- Hwang, D.M., Karimi, H.A., Byun, D.W., 1998. Uncertainty analysis of environmental models within GIS environments. *Computers & Geosciences* 24, 119–130. doi:10.1016/s0098-3004(97)00133-7.
- ISO/TC 211 Secretariat, 2013. ISO/TC 211 geographic information / geomatics.
- Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N., Finkelstein, M.M., 2007. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *Journal of Toxicology and Environmental Health, Part A* 70, 200–212. doi:10.1080/15287390600883018.
- Jerrett, M., Donaire-Gonzalez, D., Popoola, O., Jones, R., Cohen, R.C., Almanza, E., de Nazelle, A., Mead, I., Carrasco-Turigas, G., Cole-Hunter, T., Triguero-Mas, M., Seto, E., Nieuwenhuijsen, M., 2017. Validating novel air pollution sensors to improve exposure estimates for epidemiological analyses and citizen science. *Environmental Research* 158, 286–294. doi:10.1016/j.envres.2017.04.023.
- Kadaster, 2018. TOP10NL. <http://nationaalgeoregister.nl/geonetwork/srv/dut/catalog.search#/metadata/29d5310f-dd0d-45ba-abad-b4ffc6b8785f>. Accessed 4th of June 2018.
- Kainz, W., 1995. Logical consistency. Elsevier Science Ltd, Oxford. pp. 109–137.
- Kamionka, M., Breuil, P., Pijolat, C., 2006. Calibration of a multivariate gas sensing device for atmospheric pollution measurement. *Sensors and Actuators B: Chemical* 118, 323–327. doi:10.1016/j.snb.2006.04.058.

Bibliography

- Kashima, S., Yorifuji, T., Sawada, N., Nakaya, T., Eboshida, A., 2018. Comparison of land use regression models for NO₂ based on routine and campaign monitoring data from an urban area of Japan. *Science of the Total Environment* 631-632, 1029-1037. doi:10.1016/j.scitotenv.2018.02.334.
- van de Kassteele, J., Dekkers, A., Stein, A., Velders, G., 2005. Model-based geostatistical interpolation of the annual number of ozone exceedance days in the Netherlands. *Stochastic Environmental Research and Risk Assessment* 19, 173-183. doi:10.1007/s00477-004-0215-3.
- van de Kassteele, J., Koelemeijer, R.B.A., Dekkers, A.L.M., Schaap, M., Homan, C.D., Stein, A., 2006. Statistical mapping of PM10 concentrations over Western Europe using secondary information from dispersion modeling and MODIS satellite observations. *Stochastic Environmental Research and Risk Assessment* 21, 183-194. doi:10.1007/s00477-006-0055-4.
- van de Kassteele, J., Stein, A., 2006. A model for external drift kriging with uncertain covariates applied to air quality measurements and dispersion model output. *Environmetrics* 17, 309-322. doi:10.1002/env.771.
- van de Kassteele, J., Stein, A., Dekkers, A.L.M., Velders, G.J.M., 2009. External drift kriging of NO_x concentrations with dispersion model output in a reduced air quality monitoring network. *Environmental and Ecological Statistics* 16, 321-339. doi:10.1007/s10651-007-0052-x.
- Kilibarda, M., Hengl, T., Heuvelink, G.B.M., Gräler, B., Pebesma, E., Perćec Tadić, M., Bajat, B., 2014. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres* 119, 2294-2313. doi:10.1002/2013JD020803.
- Kizel, F., Etzion, Y., Shafran-Nathan, R., Levy, I., Fishbain, B., Bartonova, A., Broday, D.M., 2018. Node-to-node field calibration of wireless distributed air pollution sensor network. *Environmental Pollution* 233, 900-909. doi:10.1016/j.envpol.2017.09.042.
- Klompmaker, J.O., Montagne, D.R., Meliefste, K., Hoek, G., Brunekreef, B., 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. *Science of the Total Environment* 508, 266-275. doi:10.1016/j.scitotenv.2014.11.088.
- Knibbs, L.D., Cole-Hunter, T., Morawska, L., 2011. A review of commuter exposure to ultrafine particles and its health effects. *Atmospheric Environment* 45, 2611-2622. doi:10.1016/j.atmosenv.2011.02.065.
- Knibbs, L.D., de Dear, R.J., 2010. Exposure to ultrafine particles and PM_{2.5} in four Sydney transport modes. *Atmospheric Environment* 44, 3224-3227. doi:10.1016/j.atmosenv.2010.05.026.
- KNMI, 2011. Windrozen. <http://knmi.nl/nederland-nu/klimatologie/grafieken/maand/windrozen>. [Accessed 14th of March 2016].

Bibliography

- KNMI, 2016. Uurgegevens van het weer in Nederland - download. <http://projects.knmi.nl/klimatologie/uurgegevens/selectie.cgi>. [Accessed 16th of January 2017].
- KNMI, 2019. Daggegevens van het weer in Nederland. <https://www.knmi.nl/nederland-nu/klimatologie/daggegevens>. [Accessed 22nd of January 2019].
- Kracht, O., Gerboles, M., Reuter, H.I., 2014. First evaluation of a novel screening tool for outlier detection in large scale ambient air quality datasets. *International Journal of Environment and Pollution* 55, 120-128. doi:10.1504/ijep.2014.065912.
- Lau, B.P.L., Marakkalage, S.H., Zhou, Y., Hassan, N.U., Yuen, C., Zhang, M., Tan, U.X., 2019. A survey of data fusion in smart city applications. *Information Fusion* 52, 357-374. doi:10.1016/j.inffus.2019.05.004.
- Lee, M., Brauer, M., Wong, P., Tang, R., Tsui, T.H., Choi, C., Cheng, W., Lai, P.C., Tian, L., Thach, T.Q., Allen, R., Barratt, B., 2017. Land use regression modelling of air pollution in high density high rise cities: A case study in Hong Kong. *Science of The Total Environment* 592, 306-315. doi:10.1016/j.scitotenv.2017.03.094.
- Lerner, U., Yacobi, T., Levy, I., Moltchanov, S.A., Cole-Hunter, T., Fishbain, B., 2015. The effect of ego-motion on environmental monitoring. *Science of the Total Environment* 533, 8-16. doi:10.1016/j.scitotenv.2015.06.066.
- Lewis, A., Edwards, P., 2016. Validate personal air-pollution sensors. *Nature Comments* 535, 29-31. doi:10.1038/535029a.
- Li, X.B., Wang, D., Lu, Q.C., Peng, Z.R., Fu, Q., Hu, X.M., Huo, J., Xiu, G., Li, B., Li, C., Wang, D.S., Wang, H., 2018a. Three-dimensional analysis of ozone and PM2.5 distributions obtained by observations of tethered balloon and unmanned aerial vehicle in Shanghai, China. *Stochastic Environmental Research and Risk Assessment* 32, 1189-1203. doi:10.1007/s00477-018-1524-2.
- Li, X.B., Wang, D.S., Lu, Q.C., Peng, Z.R., Wang, Z.Y., 2018b. Investigating vertical distribution patterns of lower tropospheric PM2.5 using unmanned aerial vehicle measurements. *Atmospheric Environment* 173, 62-71. doi:10.1016/j.atmosenv.2017.11.009.
- Linn, W.S., Shamoo, D.A., Anderson, K.R., Peng, R.C., Avol, E.L., Hackney, J.D., Gong, H., 1996. Short-term air pollution exposures and responses in Los Angeles area schoolchildren. *Journal of Exposure Analysis and Environmental Epidemiology* 6, 449-472.
- Liu, J., Gustafson, P., Cherry, N., Burstyn, I., 2009. Bayesian analysis of a matched case-control study with expert prior information on both the misclassification of exposure and the exposure-disease association. *Statistics in Medicine* 28, 3411-3423. doi:10.1002/sim.3694.
- Marra, J., Voetz, M., Kiesling, H.J., 2010. Monitor for detecting and assessing exposure to airborne nanoparticles. *Journal of Nanoparticle Research* 12, 21-37. doi:10.1007/s11051-009-9695-x.

Bibliography

- Martinez, F.D., Vercelli, D., 2013. Asthma. *The Lancet* 382, 1360–1372. doi:10.1016/S0140-6736(13)61536-6.
- Martínez Torres, J., García Nieto, P.J., Alejano, L., Reyes, A.N., 2011. Detection of outliers in gas emissions from urban areas using functional data analysis. *Journal of Hazardous Materials* 186, 144–149. doi:10.1016/j.jhazmat.2010.10.091.
- Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: New features. *Computational Statistics & Data Analysis* 67, 68–83. doi:10.1016/j.csda.2013.04.014.
- Mead, M.I., Popoola, O.A.M., Stewart, G.B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J.J., McLeod, M.W., Hodgson, T.F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J.R., Jones, R.L., 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment* 70, 186–203. doi:10.1016/j.atmosenv.2012.11.060.
- Migliore, E., Berti, G., Galassi, C., Pearce, N., Forastiere, F., Calabrese, R., Armenio, L., Biggeri, A., Bisanti, L., Bugiani, M., Cadum, E., Chellini, E., Dell'Orco, V., Giannella, G., Sestini, P., Corbo, G., Pistelli, R., Viegi, G., Ciccone, G., 2009. Respiratory symptoms in children living near busy roads and their relationship to vehicular traffic: Results of an Italian multicenter study (SIDRIA 2). *Environmental Health* 8. doi:10.1186/1476-069x-8-27.
- Ministry of Infrastructure and the Environment, 1979. Bijlage 2 bij de wet milieubeheer. Wet milieubeheer. wetten.overheid.nl/BWBR0003245/Bijlage2/. [Last edited 9th of December 2015].
- Miskell, G., Salmond, J.A., Williams, D.E., 2018. Solution to the problem of calibration of low-cost air quality measurement sensors in networks. *ACS Sensors* 3, 832–843. doi:10.1021/acssensors.8b00074.
- Moltchanov, S., Levy, I., Etzion, Y., Lerner, U., Broday, D.M., Fishbain, B., 2015. On the feasibility of measuring urban air pollution by wireless distributed sensor networks. *Science of the Total Environment* 502, 537–547. doi:10.1016/j.scitotenv.2014.09.059.
- Montagne, D.R., Hoek, G., Klompmaker, J.O., Wang, M., Meliefste, K., Brunekreef, B., 2015. Land use regression models for ultrafine particles and black carbon based on short-term monitoring predict past spatial variation. *Environmental Science & Technology* 49, 8712–8720. doi:10.1021/es505791g.
- Morales, J.A., Walsh, J.E., Treacy, J., Garland, W.E., 2002. Miniaturized differential optical absorption spectroscopy (DOAS) system for the analysis of NO₂, in: Blau, W.J., Donegan, J.F., Duke, A.F., MacCraith, J.A., McMillan, N.D., OConnor, G.M., Omongain, E., Toal, V., McLaughlin, J.A. (Eds.), Opto-Ireland 2002: Optics and Photonics Technologies and Applications, International Society for Optical Engineering (SPIE). pp. 1229–1235. doi:10.1117/12.463920.

Bibliography

- Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *The Computer Journal* 7, 308–313. doi:10.1093/comjnl/7.4.308.
- Neri, G., Bonavita, A., Galvagno, S., Siciliano, P., Capone, S., 2002. CO and NO₂ sensing properties of doped-Fe₂O₃ thin films prepared by LPD. *Sensors and Actuators B: Chemical* 82, 40–47. doi:10.1016/s0925-4005(01)00987-x.
- Nguyen, P., Hoogerbrugge, R., 2014. Methods used to Compensate for the Effect of Missing Data in Air Quality Measurements. Report RIVM Letter report 2014-0079. National Institute for Public Health and the Environment.
- van Oort, P., 2006. Spatial Data Quality: from Description to Application. Phd thesis. Wageningen University, Wageningen, the Netherlands. URL: <http://edepot.wur.nl/38987>.
- Ott, W.R., 1990. A physical explanation of the lognormality of pollutant concentrations. *Journal of the Air & Waste Management Association* 40, 1378–1383. doi:10.1080/10473289.1990.10466789.
- Penza, M., Martucci, C., Cassano, G., 1998. NO_x gas sensing characteristics of WO₃ thin films activated by noble metals (Pd, Pt, Au) layers. *Sensors and Actuators B: Chemical* 50, 52–59. doi:10.1016/s0925-4005(98)00156-7.
- Phala, K.S.E., Kumar, A., Hancke, G.P., 2016. Air quality monitoring system based on iso/iec/ieee 21451 standards. *IEEE Sensors Journal* 16, 5037–5045. doi:10.1109/JSEN.2016.2555935.
- Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., Li, K., Dick, R.P., Lv, Q., Hannigan, M., Shang, L., 2014. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques* 7, 3325–3336. doi:10.5194/amt-7-3325-2014.
- Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, in: Hornik, K., Leisch, F., Zeileis, A. (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vienna, Austria.
- Pope, C.A., Thun, M.J., Namboodiri, M.M., Dockery, D.W., Evans, J.S., Speizer, F.E., Heath, C.W., 1995. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *American Journal of Respiratory and Critical Care Medicine* 151, 669–674. doi:10.1164/ajrccm/151.3_Pt_1.669.
- Post, E., Hoaglin, D., Deck, L., Larntz, K., 2001. An empirical Bayes approach to estimating the relation of mortality to exposure to particulate matter. *Risk Analysis* 21, 837–842. doi:10.1111/0272-4332.215155.
- Ranzi, A., Freni Sterrantino, A., Forastiere, F., Sartini, C., Casale, G., Cavallini, R., De Togni, A., Gallo, L., Lauriola, P., 2015. Asthmatic symptoms and air pollution: a panel study on children living in the Italian Po Valley. *Geospatial Health* 10, 248–254. doi:10.4081/gh.2015.366.

Bibliography

- Rasch, D., Tiku, M.L., Sumpf, D., 1994. Elsevier's Dictionary of Biometry. Elsevier, Amsterdam.
- Reis, S., Seto, E., Northcross, A., Quinn, N.W.T., Convertino, M., Jones, R.L., Maier, H.R., Schlink, U., Steinle, S., Vieno, M., Wimberly, M.C., 2015. Integrating modelling and smart sensors for environmental and human health. *Environmental Modelling & Software* 74, 238-246. doi:10.1016/j.envsoft.2015.06.003.
- RIVM, 2014. Kunnen PM2.5 metingen hoger zijn dan PM10 metingen? http://www.rivm.nl/Documenten_en_publicaties/Algemeen_Actueel/Veelgestelde_vragen/Milieu_Leefomgeving/Kunnen_PM2_5_metingen_hoger_zijn_dan_PM10_metingen? [Accessed 11th of January 2016].
- RIVM, 2019a. Air quality network. <https://www.luchtmeetnet.nl/uitleg>. [Accessed 23rd of August 2019].
- RIVM, 2019b. Luchtmeetnet rapportage. <https://www.luchtmeetnet.nl/rapportage>. [Accessed 9th of April 2019].
- Roemer, W., Hoek, G., Brunekreef, B., 1993. Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *American Review of Respiratory Disease* 147, 118-124. doi:10.1164/ajrccm/147.1.118.
- Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71, 319-392. doi:10.1111/j.1467-9868.2008.00700.x.
- Sacha, D., Senaratne, H., Bum Chul, K., Ellis, G., Keim, D.A., 2016. The role of uncertainty, awareness, and trust in visual analytics. *IEEE Transactions on Visualization and Computer Graphics* 22, 240-249. doi:10.1109/TVCG.2015.2467591.
- Sahsuvaroglu, T., Arain, A., Kanaroglou, P., Finkelstein, N., Newbold, B., Jerrett, M., Beckerman, B., Brook, J., Finkelstein, M., Gilbert, N.L., 2006. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. *Journal of the Air & Waste Management Association* 56, 1059-1069. doi:10.1080/10473289.2006.10464542.
- Santos, J., Serrini, P., O'Beirn, B., Manes, L., 1997. A thin film SnO₂ gas sensor selective to ultra-low NO₂ concentrations in air. *Sensors and Actuators B: Chemical* 43, 154-160. doi:10.1016/s0925-4005(97)00115-9.
- Saraswat, A., Apte, J.S., Kandlikar, M., Brauer, M., Henderson, S.B., Marshall, J.D., 2013. Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. *Environmental Science & Technology* 47, 12903-12911. doi:10.1021/es401489h.

Bibliography

- Schinasi, L., Horton, R.A., Guidry, V.T., Wing, S., Marshall, S.W., Morland, K.B., 2011. Air pollution, lung function, and physical symptoms in communities near concentrated swine feeding operations. *Epidemiology* 22, 208–215. doi:10.1097/EDE.0b013e3182093c8b.
- Schneider, P., Castell, N., Vogt, M., Dauge, F.R., Lahoz, W.A., Bartonova, A., 2017. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environment International* 106, 234–247. doi:10.1016/j.envint.2017.05.005.
- Scholtens, S., Wijga, A.H., Brunekreef, B., Kerkhof, M., Hoekstra, M.O., Gerritsen, J., Aalberse, R., de Jongste, J.C., Smit, H.A., 2009. Breast feeding, parental allergy and asthma in children followed for 8 years. the PIAMA birth cohort study. *Thorax* 64, 604–609. doi:10.1136/thx.2007.094938.
- Sguera, C., Galeano, P., Lillo, R.E., 2016. Functional outlier detection by a local depth with application to NO(x) levels. *Stochastic Environmental Research and Risk Assessment* 30, 1115–1130. doi:10.1007/s00477-015-1096-3.
- Shad, R., Mesgari, M.S., Shad, A., 2009. Predicting air pollution using fuzzy genetic linear membership kriging in GIS. *Computers, Environment and Urban Systems* 33, 472–481. doi:10.1016/j.compenvurbsys.2009.10.004.
- Shaddick, G., Ranzi, A., Thomas, M.L., Aguirre-Perez, R., Dunbar, M.B.N., Parmagnani, F., Martuzzi, M., 2018. Towards an assessment of the health impact of industrially contaminated sites: Waste landfills in Europe. *Epidemiologia & Prevenzione* 42, 69–75. doi:10.19191/EP18.5-6.S1.P069.089.
- Shahid, N., Naqvi, I.H., Qaisar, S.B., 2015. Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey. *Artificial Intelligence Review* 43, 193–228. doi:10.1007/s10462-012-9370-y.
- Shamsipour, M., Farzadfar, F., Gohari, K., Parsaeian, M., Amini, H., Rabiei, K., Hassanvand, M.S., Navidi, I., Fotouhi, A., Naddafi, K., Sarrafzadegan, N., Mansouri, A., Mesdaghinia, A., Larijani, B., Yunesian, M., 2014. A framework for exploration and cleaning of environmental data — Tehran air quality data experience. *Archives of Iranian Medicine* 17, 821–829.
- Sharker, M.H., Karimi, H.A., 2014. Computing least air pollution exposure routes. *International Journal of Geographical Information Science* 28, 343–362. doi:10.1080/13658816.2013.841317.
- Sherman, M., 2011. *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. John Wiley & Sons, Ltd, Chichester.
- Shorshani, M.F., Seigneur, C., Rehn, L.P., Chanut, H., Pellan, Y., Jaffrezo, J.l., Charron, A., André, M., 2015. Atmospheric dispersion modeling near a roadway under calm meteorological conditions. *Transportation*

Bibliography

- Research Part D: Transport and Environment 34, 137–154. doi:10.1016/j.trd.2014.10.013.
- Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. The changing paradigm of air pollution monitoring. Environmental Science & Technology 47, 11369–11377. doi:10.1021/es4022602.
- Sølna, K., Switzer, P., 1996. Time trend estimation for a geographic region. Journal of the American Statistical Association 91, 577–589. doi:10.1080/01621459.1996.10476927.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.v.d., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64, 583–639. doi:10.1111/1467-9868.00353.
- Spinelle, L., Gerboles, M., Villani, M.G., Aleixandre, M., Bonavitacola, F., 2015. Field calibration of a cluster of low-cost available sensors for air quality monitoring. part A: Ozone and nitrogen dioxide. Sensors and Actuators B-Chemical 215, 249–257. doi:10.1016/j.snb.2015.03.031.
- Spira-Cohen, A., Chen, L.C., Kendall, M., Lall, R., Thurston, G.D., 2011. Personal exposures to traffic-related air pollution and acute respiratory health among Bronx schoolchildren with asthma. Environmental Health Perspectives 119, 559–565. doi:10.1289/ehp.1002653.
- Su, Y.S., Yajima, M., 2015. R2jags: A package for running jags from R. <http://CRAN.R-project.org/package=R2jags>. R package version 0.5-7.
- Tavana, M., Liu, W.R., Elmore, P., Petry, F.E., Bourgeois, B.S., 2016. A practical taxonomy of methods and literature for managing uncertain spatial data in geographic information systems. Measurement 81, 123–162. doi:10.1016/j.measurement.2015.12.007.
- Le Tertre, A., Schwartz, J., Touloumi, G., 2005. Empirical Bayes and adjusted estimates approach to estimating the relation of mortality to exposure of PM10. Risk Analysis 25, 711–718. doi:10.1111/j.1539-6924.2005.00606.x.
- Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M., Pavel, M., 2005. A typology for visualizing uncertainty, in: Conference on Visualization and Data Analysis 2005, International Society for Optics and Photonics. pp. 146–157.
- Tsujita, W., Yoshino, A., Ishida, H., Moriizumi, T., 2005. Gas sensor network for air-pollution monitoring. Sensors and Actuators B: Chemical 110, 304–311. doi:10.1016/j.snb.2005.02.008.
- U.S. NRC, 2009. Science and Decisions: Advancing Risk Assessment. National Academy of Sciences, Washington.
- Velická, H., Puklová, V., Keder, J., Brabec, M., Malý, M., Bobák, M., Kotlík, B., Jiřík, V., Janout, V., Kazmarová, H., 2015. Asthma exacerbations and

Bibliography

- symptom variability in children due to short-term ambient air pollution changes in Ostrava, Czech Republic. Central European Journal of Public Health 23, 292–298. doi:10.21101/cejph.a4548.
- Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software 36, 1–48.
- Wang, G.X., Gertner, G.Z., Fang, S.F., Anderson, A.B., 2005a. A methodology for spatial uncertainty analysis of remote sensing and GIS products. Photogrammetric Engineering and Remote Sensing 71, 1423–1432. doi:10.14358/PERS.71.12.1423.
- Wang, S.L., Shi, W.Z., Yuan, H.N., Chen, G.Q., 2005b. Attribute Uncertainty in GIS Data. volume 3614 of *Lecture Notes in Artificial Intelligence*. pp. 614–623.
- Wang, X., Yue, Y., Faraway, J., 2018. Bayesian Regression Modeling with INLA. Chapman and Hall/CRC, New York.
- Webster, R., Oliver, M.A., 2001. Geostatistics for Environmental Scientists. Statistics in Practice. second edition ed., John Wiley & Sons Ltd., Chichester.
- Weil, J.C., Sykes, R.I., Venkatram, A., 1992. Evaluating air-quality models — review and outlook. Journal of Applied Meteorology 31, 1121–1145. doi:10.1175/1520-0450(1992)031<1121:eaqmra>2.0.co;2.
- Weinmayr, G., Romeo, E., De Sario, M., Weiland, S.K., Forastiere, F., 2010. Short-term effects of PM10 and NO₂ on respiratory health among children with asthma or asthma-like symptoms: A systematic review and meta-analysis. Environmental Health Perspectives 118, 449–457. doi:10.1289/ehp.0900844.
- Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., Williams, D.E., 2018. Development of a microscale land use regression model for predicting NO₂ concentrations at a heavy trafficked suburban area in Auckland, NZ. Science of the Total Environment 619, 112–119. doi:10.1016/j.scitotenv.2017.11.028.
- WHO, 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005 - Summary of risk assessment. Report. World Health Organization.
- WHO, 2013a. Health Risks of Air Pollution in Europe — HRAPIE Project. Report. WHO Regional Office for Europe.
- WHO, 2013b. Review of Evidence on Health Aspects of Air Pollution — REVIHAAP Project. Report. WHO Regional Office for Europe.
- WHO, 2015. Health in 2015: from MDGs, Millennium Development Goals, to SDGs, Sustainable Development Goals. Report. World Health Organization.
- Xiang, Y., Tang, Y., Zhu, W., 2016. Mobile sensor network noise reduction and recalibration using a Bayesian network. Atmospheric Measurement Techniques 9, 347–357. doi:10.5194/amt-9-347-2016.

Bibliography

- Xu, Z., Cao, Y., Kang, Y., 2019. Deep spatiotemporal residual early-late fusion network for city region vehicle emission pollution prediction. *Neurocomputing* 355, 183–199. doi:10.1016/j.neucom.2019.04.040.
- Yegnan, A., Williamson, D., Graettinger, A., 2002. Uncertainty analysis in air dispersion modeling. *Environmental Modelling & Software* 17, 639–649. doi:10.1016/S1364-8152(02)00026-9.
- Zalel, A., Yuval, Svecova, V., Sram, R.J., Bartonova, A., Broday, D.M., 2015. Modeling airborne benzo(a)pyrene concentrations in the Czech Republic. *Atmospheric Environment* 101, 166–176. doi:10.1016/j.atmosenv.2014.11.031.
- van der Zee, S.C., Hoek, G., Boezen, H.M., Schouten, J.P., van Wijnen, J.H., Brunekreef, B., 1999. Acute effects of urban air pollution on respiratory health of children with and without chronic respiratory symptoms. *Occupational and Environmental Medicine* 56, 802–812. doi:10.1136/oem.56.12.802.
- van der Zee, S.C., Hoek, G., Boezen, M., Schouten, J., Van Wijnen, J., Brunekreef, B., 2000. Acute effects of air pollution on respiratory health of 50–70 yr old adults. *European Respiratory Journal* 15, 700–709.
- Zhang, Y., Hamm, N.A., Meratnia, N., Stein, A., van de Voort, M., Havinga, P.J., 2012. Statistics-based outlier detection for wireless sensor networks. *International Journal of Geographical Information Science* 26, 1373–1392. doi:10.1080/13658816.2012.654493.
- Zimmerman, D., Stein, M., 2010. Classical Geostatistical Methods. CRC Press, Boca Raton, FL. Chapter 3. pp. 29–44.



*Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment*

D I P L O M A

for specialised PhD training

The Netherlands research school for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

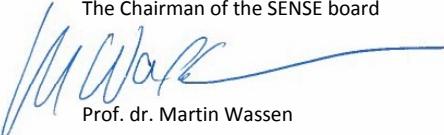
Veronella Maria van Zoest

born on 5 August 1992, in Voorburg, The Netherlands

has successfully fulfilled all requirements of the
educational PhD programme of SENSE.

Enschede, 15 January 2020

The Chairman of the SENSE board



Prof. dr. Martin Wassen

the SENSE Director of Education



Dr. Ad van Dommelen

The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)



K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N



The SENSE Research School declares that **Veronella Maria van Zoest** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 53.2 EC, including the following activities:

SENSE PhD Courses

- o Environmental research in context (2017)
- o Research in context activity: 'Preparing and presenting Geo-health workshop on GIS and Spatial Statistics in Geo-Health at HCM Pasteur Institute (Ho Chi Minh city, Viet Nam, 20-24 August 2018)'

Other PhD and Advanced MSc Courses

- o TGS Introductory workshop, University of Twente (2016)
- o Introduction to Epidemiology, Utrecht University (2016)
- o Environmental Epidemiology, Utrecht University (2016)
- o UT/CBS Data Camp, University of Twente (2016)
- o Geostat Summer School, GEOSTAT, Croatia (2017)
- o Technical Writing & Editing, University of Twente (2017)
- o Taste of Teaching, University of Twente (2018)
- o Bayesian Statistics, Wageningen Graduate Schools (2018)
- o Uncertainty propagation in Spatial Environmental Modelling, Wageningen Graduate Schools (2018)
- o Thesis Writing with LaTeX, ITC, University of Twente (2019)

Management and Didactic Skills Training

- o Supervising MSc student with thesis entitled 'Modelling and mapping of ultrafine particles in space and time in the city of Eindhoven, the Netherlands' (2018-2019)
- o Assisting practicals of the MSc courses 'Principles of Spatial Data Quality' and 'Acquisition and exploration of geospatial data'(2018-2019)
- o Organizing user group and stakeholder meetings (2015-2019)

Oral Presentations

- o *Data quality evaluation of low-cost PM10 air quality sensors.* 52N Geospatial Sensor Webs Conference, 29-31 August 2016, Münster, Germany
- o *Spatio-temporal modelling of NO₂ in urban air quality sensor networks.* Spatial Statistics, 10-13 July 2019, Sitges, Spain

SENSE coordinator PhD education



Dr. ir. Peter Vermeulen