statistical air quality mapping

Promotor:	Prof. dr. ir. A. Stein Hoogleraar in de Wiskundige en Statistische Methoden, Wageningen Universiteit, Wageningen
Copromotor:	Dr. A.L.M. Dekkers Senior onderzoeker bij het Centrum voor Informatisering en Methodologische Advisering, Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Bilthoven
Promotiecommissie:	Dr. H. Wackernagel Centre de Geostatistique de l'Ecole des Mines de Paris, Fontainebleau, France
	Prof. dr. ir. J. Grasman Wageningen Universiteit, Wageningen
	Prof. dr. A.A.M. Holtslag Wageningen Universiteit, Wageningen
	Prof. ir. N.D. van Egmond Milieu- en Natuurplanbureau (MNP), Bilthoven

Dit onderzoek is uitgevoerd binnen de C.T. de Wit Graduate School for Production Ecology and Resource Conservation.

Jan van de Kassteele

statistical air quality mapping

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M.J. Kropff,
in het openbaar te verdedigen
op vrijdag 7 april 2006
des namiddags te vier uur in de Aula

Kassteele, Jan van de
Statistical Air Quality Mapping
Doctoral Thesis Wageningen University. With ref. and summary in Dutch
ISBN 90-8504-380-8

Voorwoord

Een voorwoord is altijd een geliefd stuk om te lezen, en zeker die van een proefschrift! Hoe is dit proefschrift tot stand gekomen? In de zomer van 2001 werd ik door Bert Holtslag, professor bij de leerstoelgroep Meteorologie en Luchtkwaliteit, op de hoogte gesteld van een AIO-baan bij de leerstoelgroep Wiskundige en Statistische Methoden. Ik besloot een afspraak te maken met Alfred Stein, professor bij de betreffende leerstoelgroep, en werd aangenomen om per 1 november aan een vierjarig AIO-project te beginnen. Ik kwam terecht in het Wiskundegebouw op de Dreijen, bij de mensen waarvan je een paar jaar daarvoor nog onderwijs kreeg. Daarnaast zat ik ook een groot deel van de week op het RIVM bij het laboratorium voor Luchtonderzoek (LLO). De mensen daar hielden zich bezig met het meten, berekenen, rapporten en adviseren van en over luchtkwaliteit.

Ik begon met het vergelijken van verschillende interpolatiemethoden voor luchtkwaliteitmetingen. De eerste resultaten waren dat kriging, een statistische interpolatiemethode, als beste naar voren kwam, behalve voor het interpoleren van het aantal smogdagen per jaar. Dat zou het onderwerp voor het eerste artikel worden.

Hoewel ik in het eerste half jaar flink wat literatuur over interpolatiemethoden tot me had genomen, was het statistisch correct interpoleren van tellingen iets wat, wetenschappelijk gezien, nog in de kinderschoenen stond. Het inmiddels klassieke artikel van Diggle, Tawn en Moyeed diende als basis, maar was moeilijk te begrijpen. Als afgestudeerd meteoroloog was het wel even slikken als je belandt in de wereld van de statistiek. Op een congres in Genua hoorde ik van het bestaan van een library voor het interpoleren dit soort data in R, een statistisch softwarepakket. Aangezien ik deze library wilde gebruiken, moest ik ook in leren programmeren in R en tot de dag van vandaag heb ik daar nooit spijt van gehad.

In Genua pikte ik ook mee dat “external drift kriging” een methode zou kunnen zijn om metingen en modeluitvoer te combineren. In het tweede artikel behandelde ik deze methode, in relatie met het aantal metingen. Dit heb ik gepresenteerd op een congres in het Spaanse Santiago de Compostela. Daarna was de vraag hoe om te gaan met onderzekerheden in zowel metingen als modeluitvoer. De theorie der “measurement error modellen” kwam aan de orde. Die had ik vervolgens met de external drift kriging verweven. Dit had ik succesvol gepresenteerd in München. Het vierde artikel werd een uitbreiding van het derde artikel en ging over het berekenen van toekomstige luchtkwaliteit op locale schaal. Dit resulteerde zelfs in een klus voor het Milieu- en Natuurplanbureau (MNP). Tenslotte werd in het vijfde artikel de ontwikkelde theorie toegepast op fijn stof metingen, modeluitvoer en remote sensing data boven Europa.

Hoe clichématig het ook mag klinken, maar in aflopen vier jaar ik veel bijgeleerd. In de zomer van 2001 was ik in staat iemand alles te vertellen over het weer, maar nu kan ik ook nog iedereen, gevraagd of ongevraagd, enige Wiskunde of Statistiek bijbrengen. Van het lesgeven aan de studenten heb ik dan ook erg genoten. Het artikel van Diggle, Tawn en Moyeed leest nu makkelijk weg. Als persoon ben ik ook wijzer geworden en ik ben er van overtuigd dat deze nieuwe persoonlijk bagage me in de toekomst nog zeker van pas gaat komen.

Promoveren is niet alleen maar hard werken. Uiteraard mocht ik de afgelopen vier jaar ook genieten van de dubbele uitjes, sportmiddagen en borrels bij zowel Biometris als het RIVM/MNP. Wat verder voor wat afwisseling zorgde waren de verhuizingen. Ten eerste in Wageningen van het Wiskundegebouw naar het Wisselgebouw, waar we omgedoopt werden tot Biometris. Uiteindelijk toch in de AIO kamer terechtgekomen. En binnen het RIVM van het LLO naar het MNP bij het team Luchtkwaliteit Europese Duurzaamheid (LED).

Tenslotte wil ik nog een aantal mensen bedanken. Dit proefschrift werd mede mogelijk gemaakt door ten eerste de Heren begeleiders: Alfred Stein, Arnold Dekkers en Guus Velders. Alfred, je opbouwende commentaren op de artikelen hebben de kwaliteit ervan verder verbeterd en je efficiënte manier van werken heeft ertoe bijgedragen dat alles ruim op tijd klaar was. Arnold, je enthousiasme, betrokkenheid en behulpzaamheid heb ik als erg prettig ervaren en ik hoop dat we in de toekomst samen nog veel inspirerende discussies mogen hebben. Guus, jij hebt er voor gezorgd dat dit onderzoek niet alleen theoretisch wetenschappelijk van aard was, maar je hebt er ook voor gezorgd dat het praktische toepassingen wist te vinden in de huidige luchtkwaliteitsproblematiek.

Natuurlijk was er ook ondersteuning vanuit Biometris van Bas Engel (Bayesiaanse methoden), Jacques Withagen (statistiek algemeen), en alle mensen aan de koffietafel. Dan zijn er natuurlijk van het RIVM/MNP: Patrick van Hooydonk (meetdata), Jan de Ruiter en Jan Aben (OPS model), Robert Koelemeijer (MODIS) en alle mensen in de koffieplaza. Ook alle kamergenoten die indirect hun bijdrage hebben geleverd: binnen Biometris Johan van Ooijen, Jacques Withagen en Marjolein Lof, bij het RIVM/MNP Remus Hanea en alle stagiaires die zich daar in de afgelopen vier jaar in onze kamer hebben aangediend: Dave, Serge, Andy, Willem, Karin, Sandra, Carine, Femke en Maartje. Ook mijn vrienden Erik Holtslag, Arjan van ’t Zelfde, Harm de Coninck, Sandra Snel, Gert-Jan Steeneveld, de BWA-clan en mijn (ex)huisgenoten: de mensen van Dijkgraaf 13B en Hoogstraat 20a voor de gezellige en soms kansloze avon(d)(t)uren. Ja, en dan natuurlijk iedereen wiens naam hier ook had moeten staan, tot en met de vriendelijke caissière bij de EDAH aan toe. Pa, ma en Leo, jullie ook! Het zit erop.

Jan van de Kassteele
Wageningen, januari 2006

Contents

List of symbols

The following list contains symbols that are frequently used in this thesis. Scalars are shown in *italic*, vectors and matrices in **bold**.

$\boldsymbol{\beta}$	vector with trend or drift parameter
$\boldsymbol{\varepsilon}_q$	nugget or equation error term
$\boldsymbol{\varepsilon}_s$	spatial correlated error term
$\boldsymbol{\varepsilon}_x$	explanatory variable error term
$\boldsymbol{\varepsilon}_y$	measurement or observation error term
$\boldsymbol{\eta}$	latent variable for **y**
λ	Box-Cox transformation parameter
$\boldsymbol{\mu}$	large scale external trend or drift
μ_ξ	expectation of $\boldsymbol{\xi}$
$\boldsymbol{\xi}$	latent variable for **x**
ρ	correlation function
σ	uncertainty on data
σ_ξ^2	variance of $\boldsymbol{\xi}$
σ_q^2	nugget or equation error variance
$\sigma_{q,rel}^2$	relative nugget
σ_s^2	partial sill or spatially correlated error variance
σ_x^2	variance of covariate or explanatory variable
σ_y^2	variance of measurement or observation
ϕ	range parameter
g	Box-Cox transformation function
H	distance matrix
I	identity matrix
i	index number, usually $i = 1 \ldots n$
j	index number, usually $j = 1 \ldots m$
m	number of prediction locations
n	number of observation locations
p	number of covariates or explanatory variables, or distribution function
q	quantile function
R	correlation matrix
s	spatial locations
$\mathbf{s}_1$	spatial locations in WE direction

$\mathbf{s}_2$	spatial locations in SN direction
$\mathbf{t}$	time steps
$\mathbf{V}$	covariance matrix
$\mathbf{X}$	matrix with covariates or explanatory variables
$\mathbf{y}$	observations
^	estimate
~	prediction or sign indicating a distribution

1. General introduction

1.1 Air quality, health and policy making

It is estimated that 3400-5700 people in the Netherlands have died prematurely through short term exposure to air pollution in 2003, of which 1/3 is caused by ozone (O_3) and 2/3 by particulate matter (PM) (Fischer et al., 2004). Ground level (tropospheric) ozone is a major air pollutant in Western Europe. It causes inflammatory responses and reduction in lung function caused when humans are exposed to periods of several days with high ozone concentration. It can also affect ecosystems, mainly through damage to leaves and other parts of plants (WHO, 1996; UNECE, 1996). Effects of long-term exposure to PM is uncertain, but is believed to have a much greater effect (Pope et al., 1995). Besides ozone and PM, it is well known that nitrogen dioxide (NO_2) in high concentrations also causes respiratory problems for humans (EPA, 1998; WHO, 2003).

As a protection instrument for human health, the European Commission has set several targets and objectives for pollutants. For ozone this is for example the number of ozone exceedance days (EC, 2002), and for PM and NO_2 there are standards for ambient yearly and daily averaged concentrations (EC, 1999). Although emissions have substantially decreased in the Netherlands and Europe over the past decade, still at many places these air quality limit values are not met. This has major societal consequences in, for example, the Netherlands, where the highest court of justice has recently rejected many projects for construction of new houses, business parks, and highways (Backes *et al.*, 2005; Folkert *et al.*, 2002; Van Velze *et al.*, 2000).

For all these reasons, policy makers require accurate and spatially highly resolved maps showing concentrations of pollutants. These maps form a basis for assessing individual human exposures and serve to decide on infrastructural projects near residential areas. It is the task of the Netherlands Environmental Assessment Agency to measure and model air quality. This must be done better than ever because of the big social-economical impacts and small margins that are left concerning air quality standards. Since policy makers tend to focus more and more on uncertainties as well, the question is then how accurate these concentration maps are and how should uncertainties be handled? For example, one might question what the probability of exceedance of an air quality standard is and how this should be communicated. This thesis focuses on statistical techniques for detailed mapping of

major pollutants ozone, NO_2 and PM and its uncertainties. From these maps, several statistics can be derived which can be used for policy making (Van Soest *et al.*, 2001).

Ozone results from photochemical reactions with ozone precursors under the influence of solar radiation. Ozone typically forms on large scales. NO_2 is completely formed by combustion. Sources are for example traffic, industry, power plants, and inland waterway shipping. NO_2 is a typical local scale pollutant. PM originates from natural processes, like soil dust and sea salt, and anthropogenic processes, like combustion in car engines. It can also be formed from PM precursors (NO_x, SO_2, NH_3). PM is therefore also a large scale pollutant, but higher concentrations may be found near highways and large cities.

In this thesis the relation to air quality standards will be handled differently for these three major pollutants. For ozone, we will look at the annual number of exceedance days above the standard in the present over the Netherlands, Belgium and Germany. For NO_2, we will look at yearly averaged concentrations in the present and future at national and local scales in the Netherlands. For PM, we will look at yearly averaged concentrations in the present over Western Europe.

1.2 Air quality mapping

1.2.1 Monitoring

Concentration maps can in principle be based on measurements of air quality only. In the Netherlands the Dutch air quality monitoring network (Landelijk Meetnet Luchtkwaliteit, LML) exists. Its purpose is to monitor air quality on a continual basis. The observations provide a general description of national, regional and local air quality, along with information on smog episodes.

Different types of stations are distinguished, depending on the surroundings. Background stations are found at rural and (sub)urban sites and provide information about concentrations on a regional scale, i.e. concentrations far from streets. Street stations are usually found in cities, in and near streets. They provide information about traffic related air quality on a local scale (Fig 1.1). The network has undergone some changes in the past decades. The largest took place in the mid-nineteen eighties (Van Elzakker, 2001).

Other European countries have similar networks. Information from those is stored in databases and is freely available to the public. One of these databases is the AirBase database. AirBase is the public air quality database system of the European Environmental Assessment Agency (EEA). It contains information submitted by the

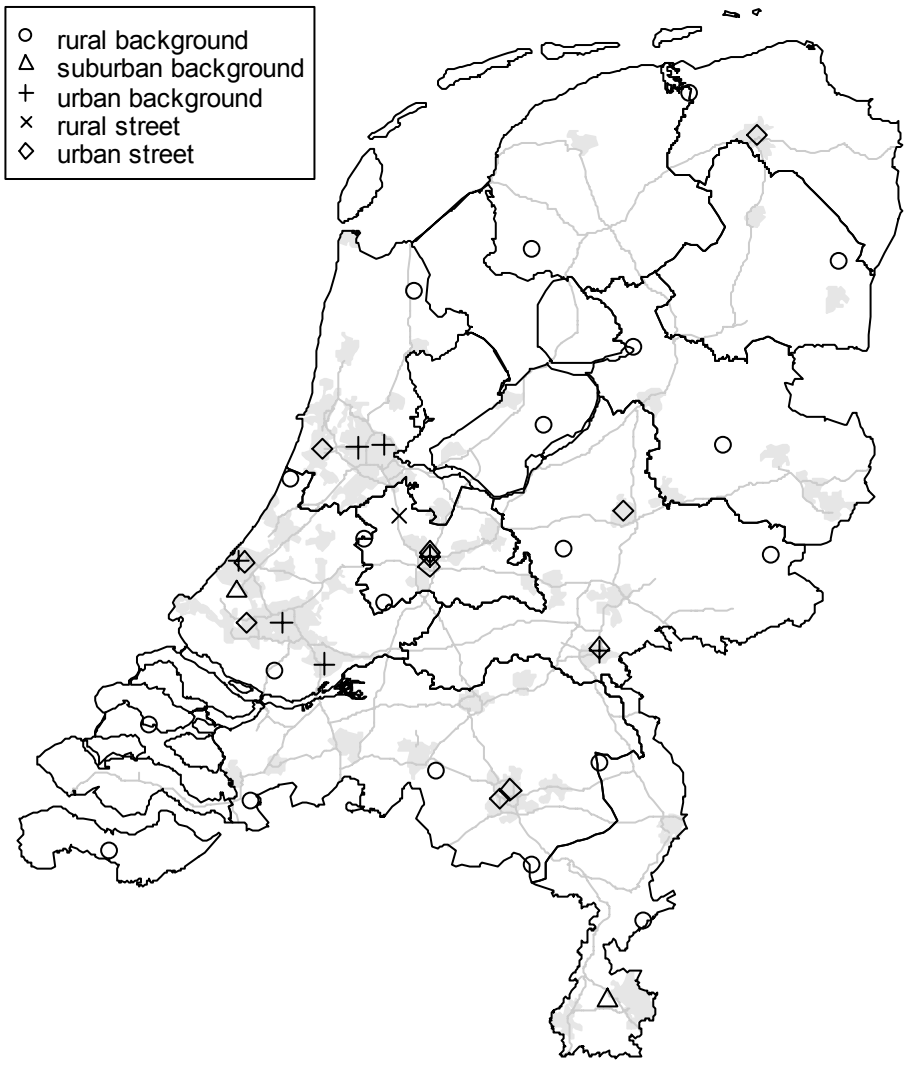

Fig. 1.1. The Dutch air quality monitoring network for ozone, NO_2 and PM10 on 1 January 2004. The different symbols indicate different types of stations and surroundings.

participating countries throughout Europe. The air quality database consists of multi-annual time series and statistics for a number of pollutants. Besides, it provides meta-information on the involved monitoring networks, their stations, and measurements (ETC-ACC, 2005). The measurements are quality checked. Monitoring data used in this thesis are extracted from this database.

1.2.2 Modeling

As an alternative to measuring, concentration maps can also be obtained using atmospheric dispersion models. In such models, chemical and transport processes are described partly by physical laws and partly by empirical relations. Examples of

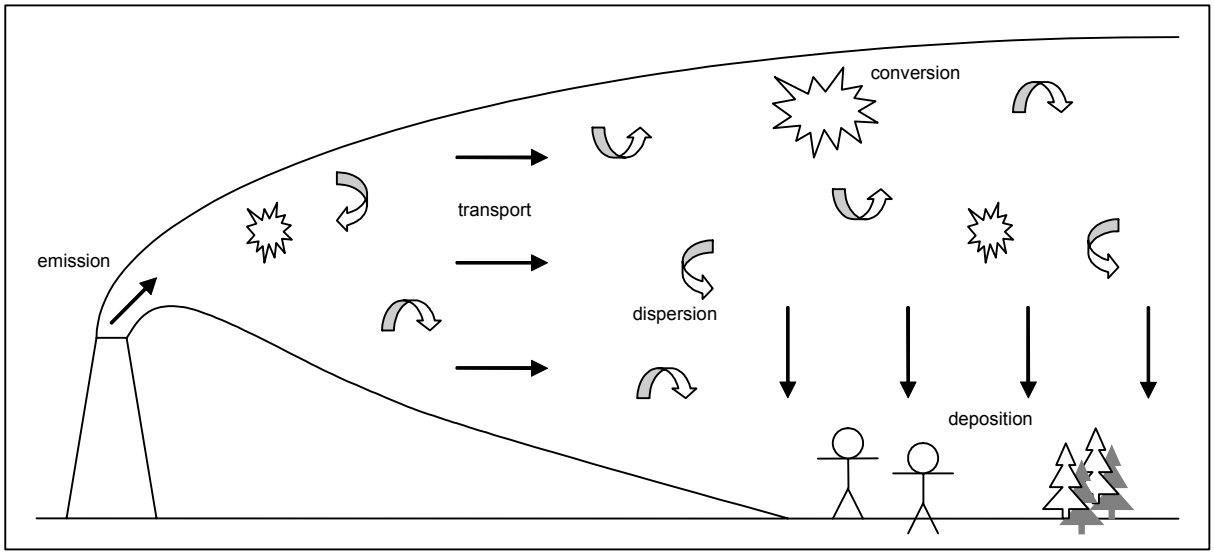

Fig. 1.2. Schematic illustration of air pollution modeling. From the left to the right: emission, dispersion, transport and conversion, deposition.

processes are emission, dispersion, transport, conversion and deposition. Input usually consists of emissions from sources into the atmosphere. Source properties like emission height determine the dispersion. Meteorological forcing and land surface characteristics determine the dispersion of pollutants in the atmosphere. Output is represented by concentration fields on an hourly, daily, monthly or yearly basis, usually on a domain covering regular grid. The processes are schematically shown in Fig. 1.2.

In this thesis, two major dispersion models are used. The Operational Priority Substances (OPS) dispersion model provides information for NO_2 background concentrations in the Netherlands on national, regional and local scales (Van Jaarsveld and De Leeuw, 1993; Van Jaarsveld, 2004). It calculates average atmospheric concentrations and deposition from the atmosphere on the basis of emissions within the Netherlands and Europe, using the Gaussian plume model to describe transport and dispersion. The model is suitable for a series of substances of which the behavior can be described by first-order linear chemical reactions; it cannot be used, for example, for describing ozone concentrations (Van Jaarsveld, 1995). Its output is usually in yearly averages.

The LOTOS-EUROS model is a 3D chemistry transport model that is used to simulate air pollutants over Europe. Based on emission estimates, meteorological data and process knowledge the concentration of, only in this thesis, PM is simulated on an hourly basis (Schaap *et al.*, 2005a,b).

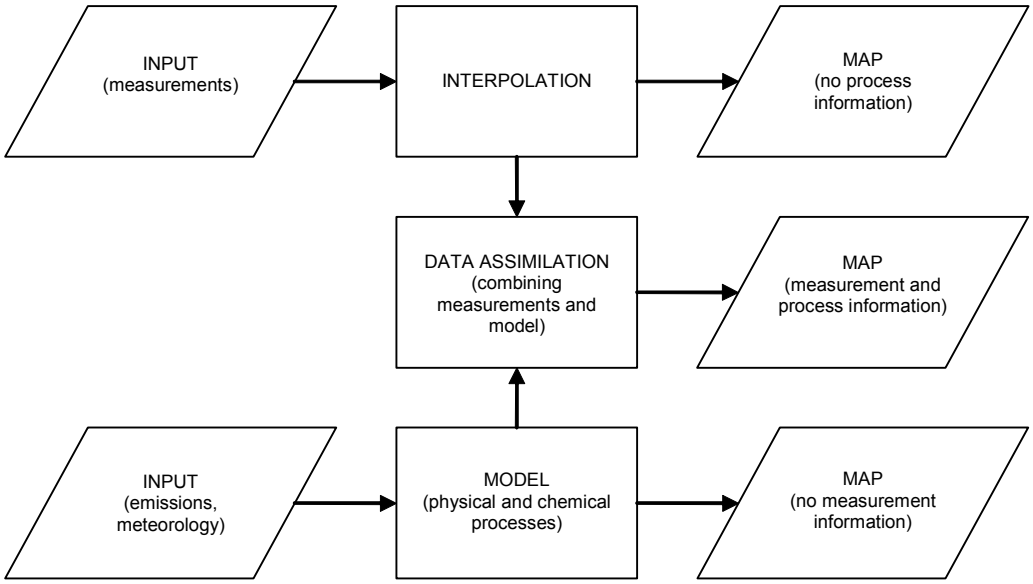

Fig. 1.3. Illustration of data assimilation.

1.2.3 Data assimilation

To base concentration maps on measurements of air quality only, every km² should be monitored continuously in time and internally consistent. Measurements however are only taken at a limited number of locations, and interpolation to a regular grid followed by a display as a map is necessary. Between the monitoring locations relevant information will be missing or can only be predicted, i.e. interpolated, leading to uncertainty in the map. Besides, the measurements have errors arising from various sources, such as instrumental noise, environmental noise, sampling, and the interpretation of sensor measurements. No information about the physical and chemical processes about the concerned component is taken into account.

To base concentration maps on modeling of components only, these models must contain a complete description of physical and chemical processes of the concerning component, including emissions and meteorology. All models are imperfect however, with errors arising from approximate physics (transport or chemistry) and different spatiotemporal scales. Such models need to be calibrated en validated by measurements and the models remain dependent on the quality of its input, which may lead to biased output and uncertainties.

This thesis combines the two approaches by means of data assimilation (Fig. 1.3). Data assimilation is a collective term for methods where biased and/or uncertain measurements are being combined with biased and/or uncertain model output. The combination always results into a more detailed and more accurate map than maps based on measurements or model output alone.

Data assimilation consists of three components: measurements, a physical process model, and a data assimilation algorithm. A central concept is the concept of errors, error estimation and error modeling, usually with a statistical approach. In this thesis focus is on static model output, like yearly averages. In that case we can apply geostatistics. A geostatistical model can be written as a linear mixed model with a correlated error structure (Ribeiro and Diggle, 1999; Pinheiro and Bates, 2000) and the dispersion model output can be considered as an explanatory variable. However, explanatory variables in this context are usually considered to be deterministic, so the geostatistical model must be extended.

1.2.4 The geostatistical approach

Geostatistical interpolation, or kriging, is being applied for since the nineteen seventies in many spatial applications (Cressie, 1993; Chilès and Delfiner, 1999). During the last ten years, we observe also a strong increase in applying kriging to air quality studies, often to interpolate measurements only. Emphasis usually lies on spatiotemporal interpolation (e.g. Wikle *et al*., 1998; Kyriakidis and Journel, 1999; Huerta *et al*., 2004). More recently, additional information is being used in air quality mapping studies, as found in Pauly and Drueke (1996), where ozone is mapped using a digital elevation model, in Bertino and Wackernagel (2002), where dispersion model output is used to map ozone concentrations around Paris, and in Genikhovich *et al*. (2002), where dispersion model output is combined with measurements to describe urban air quality.

All studies in the previous paragraph use kriging with external drift (KED) to combine measurements with additional information. KED merges two sources of information: a primary variable that is accurate and precise but only available at a limited number of locations, and a secondary variable that covers the full domain on a fine-mazed grid but is less accurate. Collocated co-kriging can be used as an alternative, but KED requires a less demanding variogram analysis. Furthermore, comparison studies (Pardo-Igúzquiza, 1998; Goovaerts, 2000) show KED interpolation to perform better than collocated co-kriging. KED is a form of data assimilation based on a regression-based interpolation method. The secondary information is treated as a covariate or explanatory variable and, as such, partly explains the variation in the spatially correlated observations. This allows the model to have a bias, whereas other external information may be included as well. KED has been applied in many environmental mapping of sparsely sampled data using dense external information.

During the last decade, a shift has occurred towards a more model-based geostatistics approach, starting with Diggle *et al*. (1998). The word "model" in this context refers to statistical models for inference. A model-based approach allows us, besides interpolation of traditional continual data, interpolation of count data or fractions for example. In general, it is a link between geostatistics and generalized linear models (GLM) (McCullagh and Nelder, 1989; Breslow and Clayton, 1993). The use of secondary information by means of KED, as described in the previous paragraph, fits easily in this model-based approach because of the GLM setup.

1.2.5 Measurement Error models

KED it its current form does not account for uncertain explanatory variables. This is a limitation. It does however take into account dependence between spatial observations. On the other hand, there are the so-called measurement error models or error-in-variable models. These models allow regression of dependent variables with uncertain explanatory variables. They do however not take into account spatial dependence. According to the theory of measurement error models, we have two or more related quantities that are not observable and therefore unknown. These quantities are called latent variables and can only be observed with additive errors. The error variances are assumed to be known at each location and can differ from location to location. Part of the variance in the observations can be explained by the variance of a linear function of the explanatory variables. The residuals, or equation errors, are considered independent and identically distributed (Cheng and Van Ness, 1999).

It would be nice to have the best of both: KED that allows for explanatory variables and spatial dependence between observations, and error-in-variable models, that allow uncertainties in both the observations and explanatory variables. In this thesis, we put those two together. The residuals are considered a sum of a spatially correlated part, as in KED, and the so-called equation error part, due to an imperfect relation between our variables. For this reason we call this new concept "error-in-variable KED".

Estimation of parameters in these statistical models requires modern statistical techniques, because analytical solutions are not possible and the classical approach using the variogram, as in Chilès and Delfiner (1999), is not sufficient anymore. An alternative is Bayesian inference (Handcock and Stein, 1993; Diggle *et al*., 1998; Gelman *et al*., 2004). Its advantage is that geostatistical models can be written down explicitly, and can be evaluated by numerical techniques, like Markov Chain Monte Carlo (Lunn *et al*., 2000; Gilks *et al*., 1994). Furthermore, we can use prior information.

In this thesis, we shall often apply the model-based geostatistical approach for combining measurements and chemistry transport model output.

1.3 Objectives and scope

Air quality measurements are in general accurate and precise at a certain location, but interpolation of a limited number of observations causes imprecise maps. Dispersion model output on the other hand is inaccurate (biased) and imprecise compared to measurements, but it has a much higher spatial resolution and provides more detailed information. A combination of both leads to better maps.

The main objective of this thesis is to develop and use geostatistical methods to combine a limited number of air quality measurements and inaccurate and imprecise, but domain covering dispersion model output to acquire detailed air quality maps on local, national and international scales, and to show that this improves spatial predictions, i.e. smaller bias and smaller uncertainties. Uncertainties will be quantified better than ever.

First ideas were an extension of universal kriging with Poisson regression, the same extension as linear regression to generalized linear regression for interpolating exceedance days. Second, a combination of spatially sparse measurements with spatially dense model output using KED. During the study, the power of KED emerged and also the urging question of assessing uncertainties in the final predictions, based on the uncertainties in measurements and dispersion model output. These ideas led to the following research questions:

Q1. How to interpolate spatial count data such as the number of ozone exceedance days?
Q2. Can additional information from dispersion models improve interpolation of measurements?
Q3. Which role does prior knowledge play in this context?
Q4. Should parameters kept fixed or should they be re-estimated every time?
Q5. How many measurements are actually needed for a detailed map of sufficient quality?
Q6. How do different uncertainties in the measurements and dispersion models affect the accurateness and precision of the final predictions?
Q7. Can these methods be applied in scenario studies for determining future air quality?
Q8. How do the results relate to the European air quality standards?
Q9. How do we communicate uncertainty in air quality maps to policy makers?

Q10. How do we deal with different measurement techniques between countries?
Q11. Can we use other sources of additional information, such as satellite images, besides dispersion model output?

1.4 Thesis outline

The objectives are dealt with in five chapters of this thesis. Chapter two starts with a model-based geostatistical interpolation of the annual number of ozone smog days in the Netherlands, Belgium and Germany (Q1). Two statistical models are being compared to describe these count data.

Chapter three illustrates the use of dispersion model output on the interpolation of NO_x measurements in the Netherlands (Q2). In this chapter we describe the basic techniques of universal kriging and kriging with external drift for such mapping procedures. We further discuss the impact of different parameter estimation techniques and the use of prior knowledge (Q3 and Q4). The methods are applied to a reduced air quality monitoring network in the Netherlands to see what the impact is when the number of stations decreases (Q5).

Chapter four, a theoretical chapter, is the core of this thesis. The KED method is extended and a new method is being developed were uncertain measurements and uncertain explanatory data can be combined in a solid geostatistical setting. We call this method error-in-variable KED. The impact of different measurement and model uncertainties is illustrated by a simulated data example and applied to NO_2 concentration mappings at an urban and rural site in the Netherlands (Q6).

Chapter five extends the newly developed error-in-variable KED to make future predictions. It is applied to a NO_2 scenario for 2010. The emphasis is on the assessment of local air quality near Rotterdam, and an attempt is made to simply communicate spatial uncertainties to end users, like policy makers (Q7, Q8 and Q9).

Chapter six considers air quality at the European scale, focusing on particulate matter. Measurements from different countries and surroundings are standardized first by means of statistical methods and are then combined with uncertain additional information form a chemistry transport model and remote sensing data to result in maps that completely cover Western-Europe (Q10 and Q11).

Finally, chapter seven formulates the concluding remarks of this thesis, based on the questions and statistical methods described in the first six chapters.

2. Interpolation of ozone exceedance days

This chapter discusses two model-based geostatistical methods for spatial interpolation of the number of days that ground level ozone exceeds a threshold level. The first method assumes counts to approximately follow a Poisson distribution, while the second method assumes a log-Normal distribution. First, these methods were compared using an extensive data set covering the Netherlands, Belgium and Germany. Secondly, the focus was placed on only the Netherlands, where only a small data set was used. Bayesian techniques were used for parameter estimation and interpolation. Parameter estimates are comparable due to the log-link in both models. Incorporating data from adjacent countries improves parameter estimation. The Poisson model predicts more accurately (maximum kriging standard deviation of 2.16 compared to 2.69) but shows smoother surfaces than the log-Normal model. The log-Normal approach ensures a better representation of the observations and gives more realistic patterns (an RMSE of 2.26 compared to 2.44). Model-based geostatistical procedures are useful to interpolate limited data sets of counts of ozone exceedance days. Spatial risk estimates using existing prior information can be made relating health effects to environmental thresholds.

This chapter is based on J. van de Kassteele, A.L.M. Dekkers, A. Stein and G.J.M. Velders (2005). Model-based geostatistical interpolation of the annual number of ozone exceedance days in the Netherlands. *Stochastic Environmental Research and Risk Assessment*, **19**(3), 173-183.

2.1 Introduction

Ground level (tropospheric) ozone is a major air pollutant in Western Europe. Tropospheric ozone results from photochemical reactions with ozone precursors, volatile organic compounds, nitrogen oxides, carbon monoxide and methane in the atmosphere. Environmental focus on ozone concentrations has increased as a result of the possible inflammatory responses and reduction in lung function caused when humans are exposed to periods of several days with high ozone concentrations. Ozone can also affect ecosystems, mainly through damage to leaves and other parts of plants (WHO, 1996; UNECE, 1996).

As a protection instrument for human health, the European Commission has set several targets and objectives for ozone levels in the atmosphere. The indicator applied in this study is the number of days per year in which an eight-hour moving average ozone concentration exceeds 120 $\mu g/m^3$ (EC, 2002).

Currently, rural ozone concentrations in the Netherlands are measured hourly within the Netherlands Air Quality Monitoring Network at 23 stations, spread across the country (van Elzakker, 2001). Each station registers the annual number of exceedance days. EU regulations (EC, 2002) require the number of exceedance days to be reported at the measuring sites. Interpolation of the exceedance days to produce maps for the Netherlands are a basis for assessment studies related to public health and environmental effects (for example, see EEA, 1998).

Here we analyze the use of geostatistical interpolation of annual ozone count data. So far, no attention has been paid in the literature to geostatistical interpolation of counts for ozone exceedance days. In geostatistics, spatial data are assumed to be a realization of a random field, and often without the assumption of any stochastic model being declared. Usually, normality is implicitly assumed (Christakos, 1992). The data analyzed in this chapter, however, are positively valued count data, without constant variance and normally distributed errors. In this case, the normality assumption may no longer be appropriate. Count data require a different approach. The question addressed in this chapter is then which interpolation procedure will be most appropriate and practically applicable for environmental scientists.

The aim of this study is to investigate the applicability of either a Poisson procedure or a log-Normal model-based geostatistical procedure (Diggle *et al*., 1998 and Ribeiro and Diggle, 1999) to interpolate the number of exceedances of the 120 $\mu g/m^3$ threshold. A complication is sparseness of the data. Therefore, data from 2000 measured at 120 rural ozone monitoring stations in the Netherlands, Belgium and Germany were analyzed first. Then there is a focus on the small subset of 23 stations in the Netherlands.

2.2 Material and methods

2.2.1 Data

Verified hourly data for 2000 were collected from the Airbase database (ETC-ACC, 2005) at 120 rural background ozone stations for the Netherlands, Belgium and Germany. These data were then aggregated, first, by calculating eight-hour moving averages, and, secondly, by taking the daily maxima. Finally, the days on which these maxima exceeded the threshold of 120 $\mu g/m^3$ were summed to obtain the annual number of exceedance days. The data were aggregated according to the guidelines of the European Commission for missing data (EC, 2002). Nine stations had therefore to be excluded. It was assumed that small differences between measurement techniques in the monitoring networks had not influenced the annual number of exceedances, since all observations of ozone had to satisfy the same quality control specifications (EC, 2002).

Fig. 2.1 shows the 111 observations over the whole study region. The coordinates were obtained by transforming the geographical coordinates with an azimuthal equidistant projection centered on 51° north and 9° east. This projection preserves a correct absolute distance between the stations in the region considered. The number of exceedances was lower near the North Sea coast and higher in the south-east of the region. High ozone concentrations are caused by photochemical reactions during warm and sunny days, while the strong dependence on these meteorological conditions caused the number of exceedance days to fluctuate sharply from year to year (Feister and Balzer, 1991). In and near large cities, ozone concentrations are usually lower than in rural areas (see, for example, Gregg *et al.*, 2003). Since this effect introduces local non-stationarities, only stations in rural areas were considered in this study.

Next to showing the observations in the data set of 111 stations, the study focused on analysing data from the national air quality monitoring network of the Netherlands only. This network consists of 23 rural ozone stations (van Elzakker, 2001), but one station had to be excluded due to missing data. The reason for analyzing this small subset only is practical: i.e. the Netherlands Environmental Assessment Agency needs to report the number of exceedance days to the European Commission as soon as the data has become available.

Fig. 2.1 shows the number of exceedance days to increase from the north-west to the south-east. The spatial variability increases along with this trend, making the random field a spatial non-stationary process, typical behavior of a Poisson-like process. To gain more insight into the distribution, the count data were analyzed and

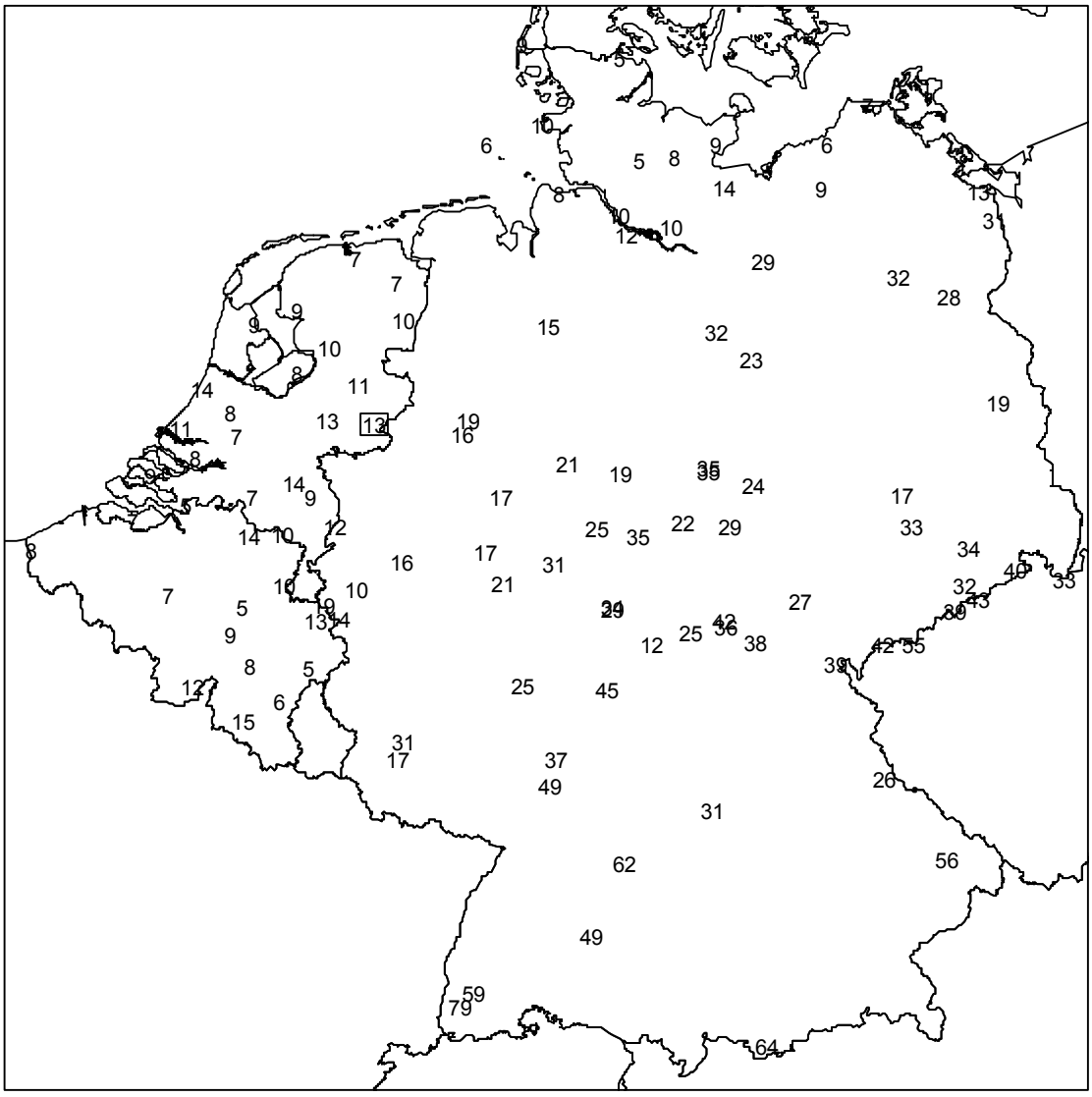

Fig. 2.1. The annual number of days where the daily maximum eight-hour moving average ozone concentration exceeds the 120 μg/m³ threshold value in 2000. The rectangle indicates the location of Eibergen station.

simulated at one single point. At other locations a similar process may occur. Observations of daily maximum eight-hour moving average ozone concentration from 1991 to 2000 are shown in Fig. 2.2, at one particular station, Eibergen, a village in the eastern part of the Netherlands. Meteorological conditions and human activities contribute to fluctuations in the concentration. The graph shows extreme concentrations during the spring/summer season. Circles indicate the days that the concentration exceeds the threshold of 120 μg/m³. One particular difficulty in assigning any statistical distribution to these data is the clustering of the exceedances. Such behavior is typical for extreme events. This can be described with an extreme value model (Smith, 1989). Shively (1991) models the sequence of exceedances as a non-homogeneous Poisson process. A simpler and more straightforward approach

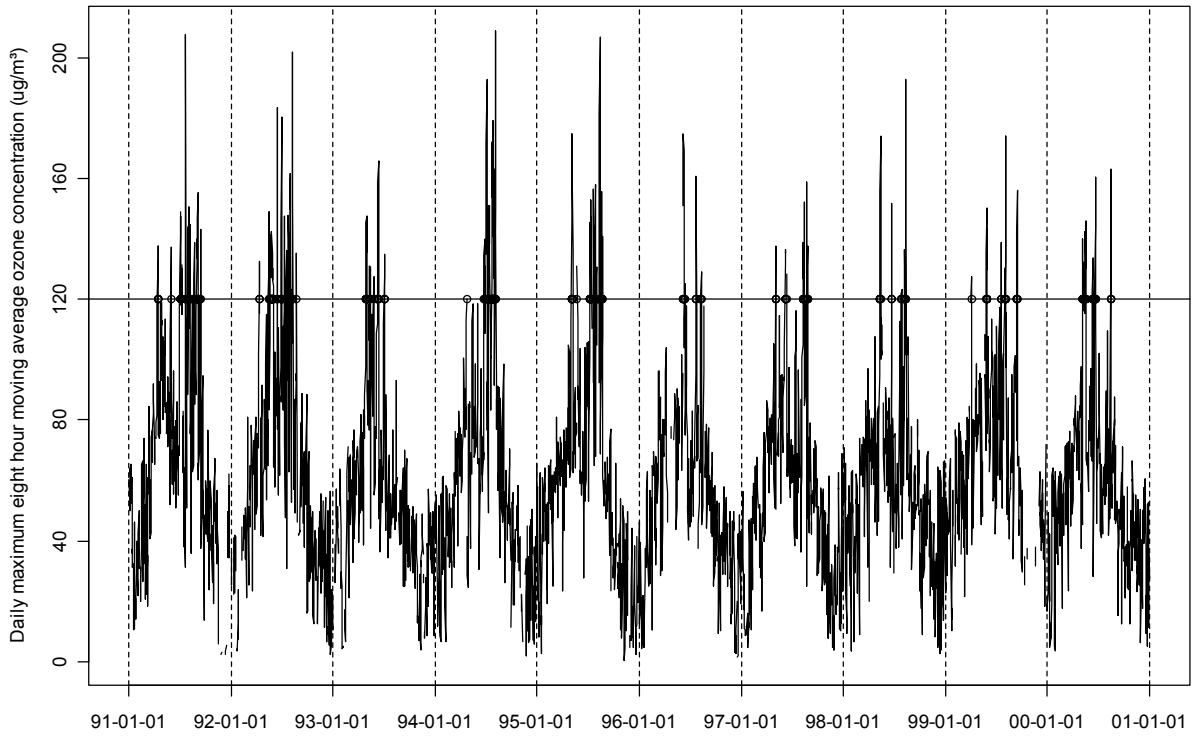

Fig. 2.2. Daily maximum eight-hour moving average ozone concentration (µg/m³) at Eibergen station for 1991 to 2000. Days that this concentration exceeds the threshold of 120 µg/m³ are indicated by circles.

assumes the exceedance days to follow a Poisson process in the limit, i.e. that the dependence between daily maxima separated by a given number of days decreases sufficiently fast as the separation increases.

As an experiment, we simulated the annual number of exceedance days. We modeled the occurrence of exceedance days over one ozone season at the Eibergen station by sampling from the Bernoulli distribution. To account for temporal dependence, the probabilities were conditional on the outcome of the previous day. These conditional probabilities were estimated from the 10 years of observations shown in Fig. 2.1. Summing the resulting sequences of zeros and ones yielded the annual numbers of exceedance days. The distribution of simulated exceedance days is presented as a histogram in Fig. 2.3, to which a Poisson distribution was fitted. It describes the average well (17.9), and is suitable for handling count data. It shows a smaller variance (17.9 for the Poisson distribution) than the simulated data (64.4), however. It overestimates the top and underestimates the tails. Also the log-Normal density function was fitted, which better accounts for the tails and the mean (18.2) of the simulations. The function however handles the data as being continuous and overestimates (105.6) the variance of the simulated data (64.4).

This exploring analysis showed that neither a Poisson distribution nor a log-Normal distribution can describe ozone exceedance count data very well, but both

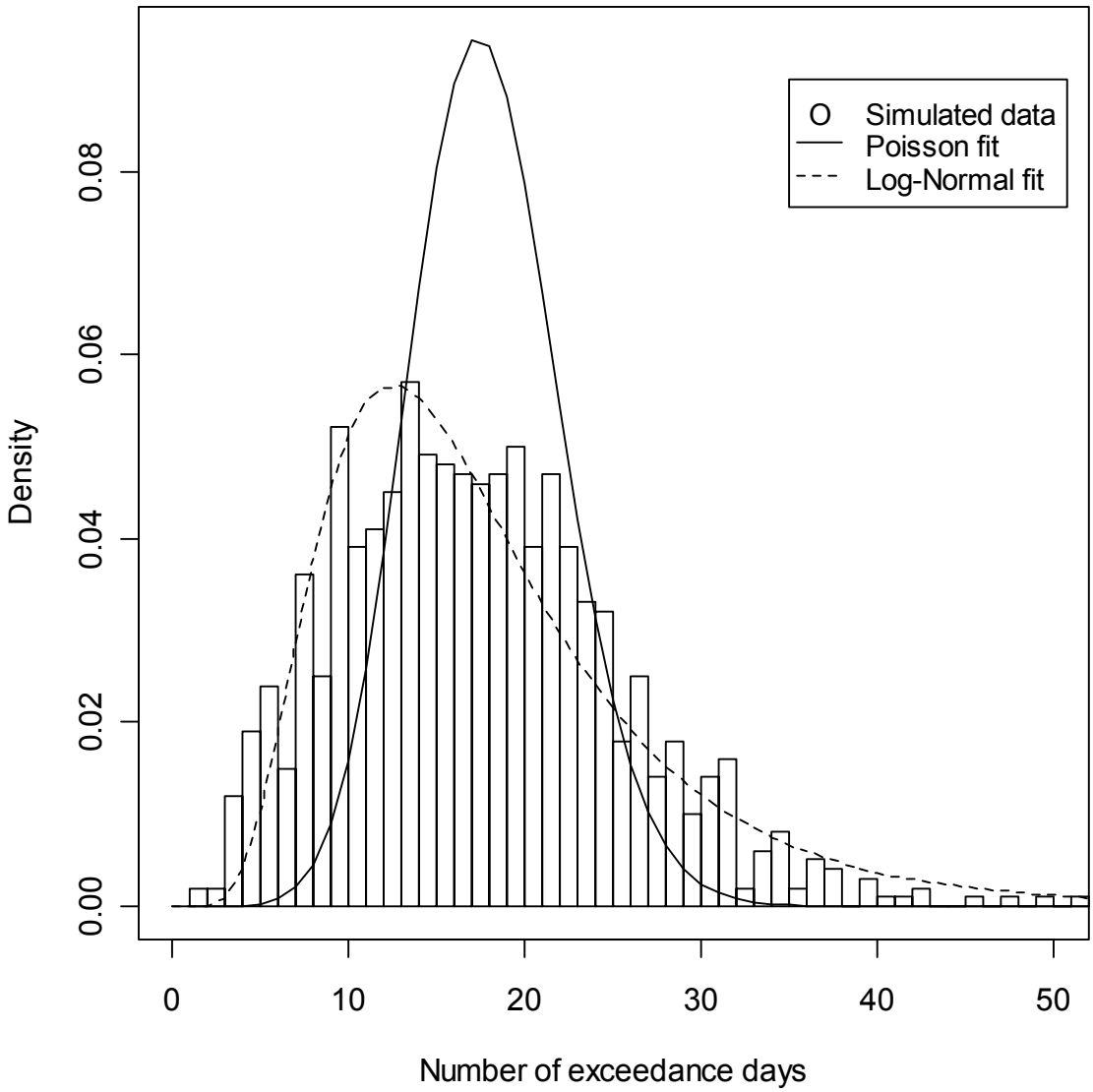

Fig. 2.3. Simulations of the annual number of exceedance days (bars) at Eibergen station with a fitted Poisson distribution (solid line) and log-Normal distribution (dashed line).

distributions have properties that do fit the data. On the other hand, the Bernoulli simulation of the occurrence of exceedance days might not have been correct, since it was only a very simple model for the real situation (see the discussion section). It should also be realized that comparison with only one observed datum could be performed, since only one realization was available.

2.2.2 The Poisson Model

Assuming counts to be spatially independent Poisson distributed, they could be analyzed with a generalized linear model (McCullagh and Nelder, 1989). Generalized linear models allow data to follow any distribution of the exponential

family, accommodating both continuous and discrete non-Normal distributions. Generalized linear-mixed models (Breslow and Clayton, 1993) allow for correlated data. Diggle *et al*., (1998) embedded kriging into the framework of generalized linear models, providing a way to analyze spatially correlated Poisson data. This model was applied in this study.

Considering n spatial observations y_i, with $i = 1 \ldots n$, as realizations of a random field process Y_i at locations $\mathbf{s}_i$, each observation for spatial correlated Poisson can written as

$$y_i \mid \varepsilon_{s_i} \sim \text{Poisson}(m_i) \tag{2.1}$$

The interpretation is that conditional on an underlying surface $\boldsymbol{\varepsilon}_s$, $\mathbf{y}$ comes from an independent Poisson distributed variable with the conditional expectation $\mathbf{m}$. $\mathbf{m}$ is a variable containing the deterministic large-scale trend $\boldsymbol{\mu}$, and the underlying surface $\boldsymbol{\varepsilon}_s$. For Poisson data the relation between $\mathbf{m}$, $\boldsymbol{\mu}$, and $\boldsymbol{\varepsilon}_s$ is attenuated by the log link function

$$\log(m_i) = \mu_i + \varepsilon_{s_i} \tag{2.2}$$

The trend $\boldsymbol{\mu}$ is a linear function $\mathbf{X}\boldsymbol{\beta}$ of known functions of covariates $\mathbf{X}$, which, in our study, only depends on the location variable $\mathbf{s}$, and unknown regression or trend parameters $\boldsymbol{\beta}$. The underlying surface $\boldsymbol{\varepsilon}_s$ is modeled by a zero mean second-order stationary Gaussian process with covariance matrix $\mathbf{V} = \sigma_s^2 \mathbf{R}$

$$\boldsymbol{\varepsilon}_s \sim \text{N}(0, \mathbf{V}) \tag{2.3}$$

Here σ_s^2 is the variance of the underlying random field. Elements of $\mathbf{R}$ depend upon the distance between two locations, $\mathbf{s}_i$ and $\mathbf{s}_j$, by means of a permissible correlation function ρ with a range parameter ϕ. The Poisson model predicts the intensity in space $\mathbf{m}$. For the Poisson model, intensity and variance are equal. Hence, the predicted intensity field will always be smoothed because it can explain deviations from the intensity value by its corresponding Poisson variance.

2.2.3 The log-Normal model

We considered the log-Normal model (Cressie, 1993) as an alternative method. We now assume the logarithm of the random field followed a Normal distribution. The

log-Normal model can be written in an equivalent model-based formulation of a linear mixed model (Ribeiro and Diggle, 1999; Pinheiro and Bates, 2000). Conditional on the underlying surface $\boldsymbol{\varepsilon}_S$ the log(**y**) are independently normally distributed, with conditional expectation **m** and variance σ_q^2

$$\log(y_i) \mid \varepsilon_{s_i} \sim \mathrm{N}(m_i, \sigma_q^2) \tag{2.4}$$

Note that here log(**y**) is in fact a noisy version of **m**, with residual variance σ_q^2. The relationship between **m** $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}_S$ is the identity link for a Gaussian model so

$$m_i = \mu_i + \varepsilon_{s_i} \tag{2.5}$$

Interpretation of $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}_S$ remains unchanged in comparison to the Poisson model. In conventional geostatistics, the variance σ_q^2 is called the nugget, $\sigma_s^2 + \sigma_q^2$, the sill, and σ_s^2, the partial sill. The parameter σ_q^2 can be considered to resemble variations that cannot be attributed to spatial correlation and thus introduces smoothing. Finally, the log-Normal model makes spatial predictions of the expected number of exceedance days. This is a major difference with the Poisson model.

2.2.4 Parameter estimation and spatial prediction

Parameters were estimated using Bayesian inference (Gelman *et al*., 2004), in particular using Markov Chain Monte Carlo (MCMC) methods (Gilks *et al*., 1996) based upon the Langevin-Hastings algorithm (Besag, 1994; Papaspilliopoulus *et al*., 2003). This is a Metropolis-Hastings algorithm in which the proposal distribution uses gradient information from the log-posterior distribution. The algorithm iteratively generates a chain, where in each step a proposal is generated for an update of the current state of the chain. The update is then accepted or rejected according to a certain acceptance probability. Proposal variances for σ_s^2 and ϕ have to be found manually in such a way that approximately 60 percent of the proposals is accepted (Christensen and Ribeiro, 2002). The predictive distribution is obtained by first sampling from the posterior distributions, and then taking, for each, samples from the multivariate Gaussian distribution of $\boldsymbol{\varepsilon}_S$. This procedure automatically incorporates parameter uncertainty in the predictions. For the mathematical formulation of the above process we refer to Diggle *et al*., 1998 and Gelman *et al*., 2004.

A re-parameterization of the nugget σ_q^2 as a relative nugget $\sigma_{q,rel}^2 = \sigma_q^2 / \sigma_s^2$ was carried out to still be able to write the covariance matrix **V** as a product between σ_s^2 and the correlation matrix **R**. Discrete intervals for ϕ and $\sigma_{q,rel}^2$ had to be taken, because their posteriors cannot be written as a standard statistical distribution. (Ribeiro and Diggle, 1999; Christensen and Waagepetersen, 2002).

2.2.5 Prior specification and setup of the MCMC algorithm

Bayesian inference needs a specification of prior distributions of the parameters. Prior knowledge was available (see Fig. 2.1). To allow modeling of the trend towards the south-east, we included covariates $\mathbf{X} = [\mathbf{1}\ \mathbf{s}_1\ \mathbf{s}_2]$, where $\mathbf{s}_1$ and $\mathbf{s}_2$ are the coordinates in the east-west and north-south directions, respectively, and associated regression parameter $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]$. The variance σ_s^2 of the log-data is positive, approximately equal to 0.1, and the correlation distance ϕ a few hundred kilometers, which is typical for ozone concentrations found in previous years. An exponential variogram model $\rho(\mathbf{h}) = \exp(-\mathbf{h})$ was chosen for the covariance structure. The resulting priors for both the Poisson and log-Normal models are:

$$\beta_0 \sim \mathrm{N}(3, 0.5) \tag{2.6a}$$
$$\beta_1 \sim \mathrm{N}(0.001, 0.001) \tag{2.6b}$$
$$\beta_2 \sim \mathrm{N}(-0.002, 0.001) \tag{2.6c}$$
$$\sigma_s^2 \sim \text{Inv-}\chi^2\,(1, 0.1) \tag{2.6d}$$
$$\phi \sim \mathrm{Exp}(1/100) \tag{2.6e}$$
$$\sigma_{q,rel}^2 \propto 1 \tag{2.6f}$$

We chose Gaussian priors for the trend parameters, an inverse-χ^2 distribution with one degree of freedom and a scale parameter 0.1 for σ_s^2 and an exponential prior with an expectation of 100 km for the range parameter. The relative nugget $\sigma_{q,rel}^2 = \sigma_q^2 / \sigma_s^2$, only used in the log-Normal model, was given a uniform prior. The Gaussian distributions and inverse-χ^2 distribution are conjugate priors for the trend and sill parameters, respectively. The exponential distribution for the range parameter leads to more equally spaced correlations at a fixed distance (Ribeiro and Diggle, 1999).

The variances of the trend parameters seem rather strict. They are not however, because these variances are scaled by the partial sill parameter. In

combination with the fact that coordinates are given in kilometers, these priors are relatively flat.

Proposal variances for σ_s^2 and ϕ were found to be 0.002 and 100. To check on convergence and mixing, we considered trace plots of the individual samples and their corresponding auto-correlation functions. The samples preferably show stationarity with low auto-correlation. The chain's burn-in time was set at 10000 iterations and it was sampled every 200th iteration to reduce the auto correlation.

The prior specification for the subset of 22 observations for the Netherlands only was based on information on the full set. Only the intercept parameter was given a lower value, and no prior trend was specified. The resulting priors for the subset are:

$$\beta_0 \sim N(2, 0.5) \quad (2.7a)$$
$$\beta_1 \sim N(0, 0.001) \quad (2.7b)$$
$$\beta_2 \sim N(0, 0.001) \quad (2.7c)$$
$$\sigma_s^2 \sim \text{Inv-}\chi^2\,(0.1, 1) \quad (2.7d)$$
$$\phi \sim \text{Exp}(1/100) \quad (2.7e)$$
$$\sigma_{q,rel}^2 \propto 1 \quad (2.7f)$$

The proposal variances for σ_s^2 and ϕ were found to be 0.01 and 300, respectively. The chain's burn-in time and thinning remained unchanged.

2.2.6 Validation

A cross validation by "leaving one out" was carried out to see which interpolation method performs better. The root-mean-squared error RMSE was chosen as the error measure. The two models do not predict the same quantity. Therefore, results have to be interpreted with care.

2.3 Results

The data sets were analyzed with the software packages *geoR* (Ribeiro and Diggle, 2001) and its extension, *geoRglm* (Christensen and Ribeiro, 2002). Both packages run under the programming environment of *R* (Ihaka and Gentleman,1996). The *geoR* package contains several functions for handling (log-)Normal spatial data; *geoRglm*

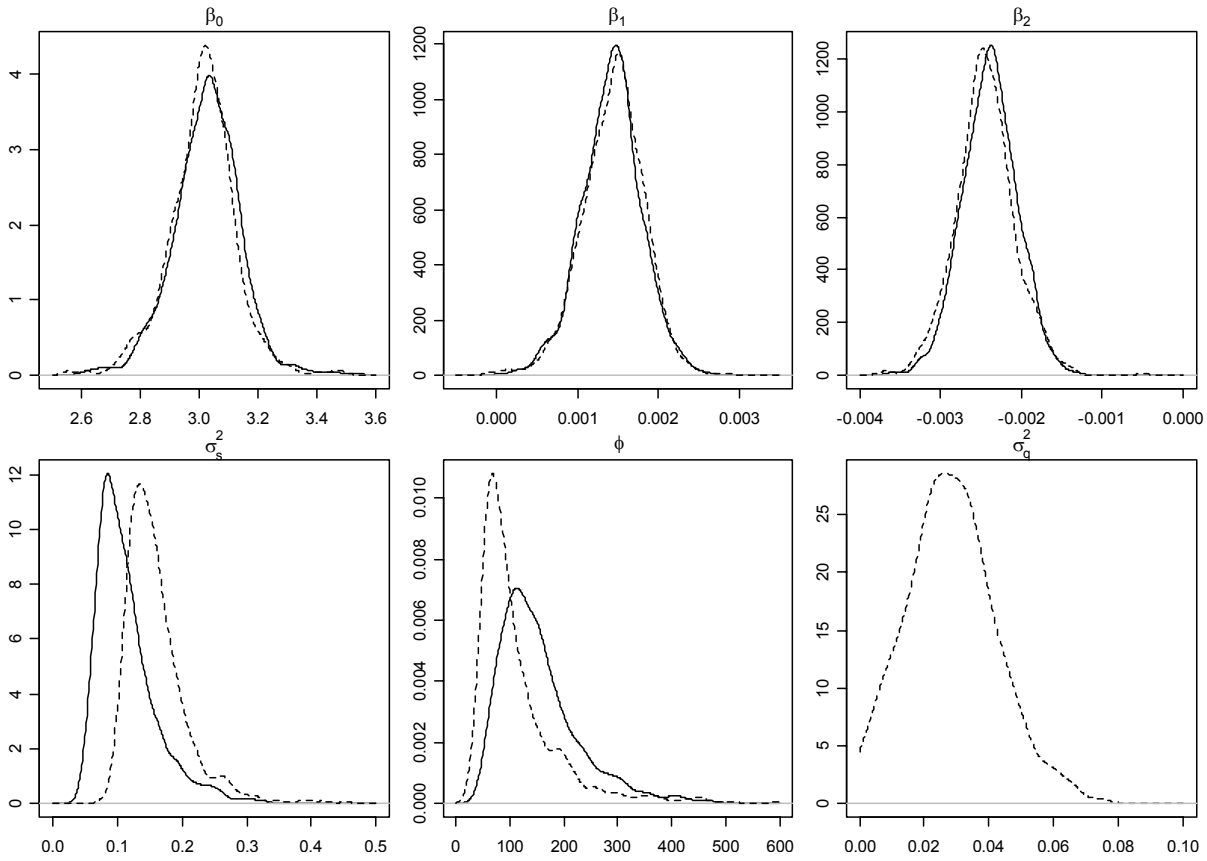

Fig. 2.4. Posterior distributions of the Poisson model parameters (solid line) and log-Normal model (dashed line) using the full data set.

can deal with spatial Poisson data. *R* and both packages are available free of charge on the Internet.

The results are presented under three headings: 1) parameter estimation and interpolation using the full data set; 2) interpolation results of 1), focusing on the Netherlands, and 3) parameter estimation and interpolation using the subset of the Netherlands only. Interpolation was done on a 15 × 15 km grid for the Netherlands, Belgium and Germany, while for the Netherlands a 5 × 5 km grid has been taken.

2.3.1 Part 1: analysis and interpolation using the full set

Posterior densities of the six model parameters are shown in Fig. 2.4. Since the Poisson model does not contain a nugget effect, no posterior is shown. Values of the modes and standard deviations are given in Table 2.1. Since we work with a number of days, sill and nugget have no units.

The posterior densities of the trend parameter vector $\boldsymbol{\beta}$ of both models are practically identical. This is not surprising since both models estimate the trend on a

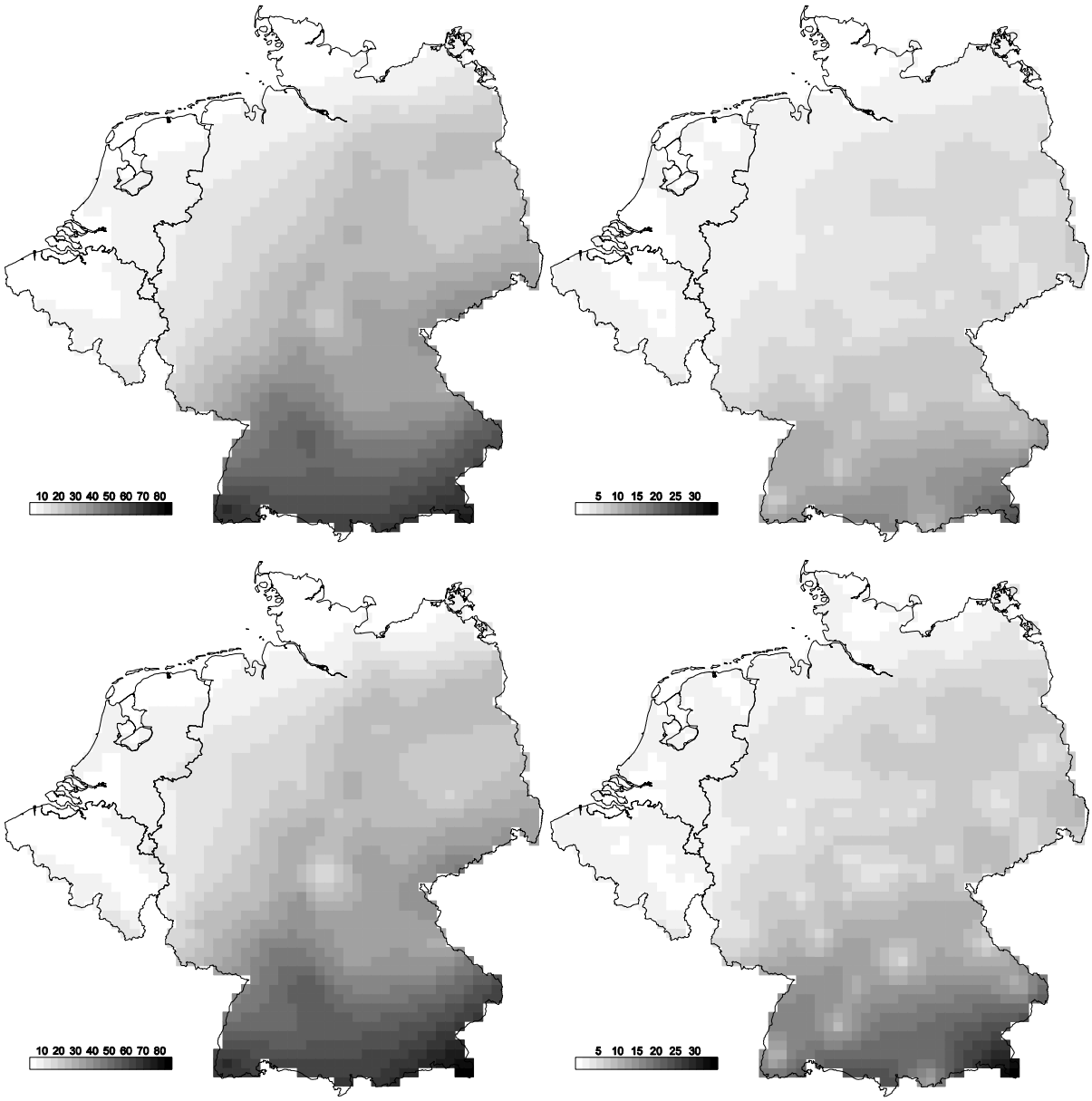

Fig. 2.5. Predicted number of ozone exceedance days (left) and corresponding kriging standard deviations (right) with the Poisson model (top) and log-Normal model (bottom) for 2000 using the full data set.

log-scale. The modes are β = (3.01, 0.0015, -0.0024) and β = (3.00, 0.0013, -0.0025) for the Poisson model and log-Normal model, respectively.

The partial sill σ_s^2 of the Poisson model is smaller than that of the log-Normal model. Their modes are 0.093 and 0.12, respectively. Although they have the same order of magnitude due to the log-scale, we can understand the difference from distributional assumptions of both models. The Poisson model predicts the intensity field. The Poisson model can describe the variation in the original data by its corresponding Poisson variance. For this reason, σ_s^2 may be smaller than for the log-

Table 2.1. Modes and standard deviations (calculated as ¼ of the 95% credible interval) of the posterior distributions (Figs 4 and 7).

		β_0 (-)	β_1 (km^{-1})	β_2 (km^{-1})	σ_s^2 (-)	ϕ (km)	σ_q^2 (-)
Full set	Poisson	3.01	0.0015	-0.0024	0.093	102	
		(0.11)	(0.00034)	(0.00033)	(0.037)	(68)	
	log-Normal	3.00	0.0013	-0.0025	0.12	68	0.028
		(0.12)	(0.00040)	(0.00037)	(0.044)	(74)	(0.014)
Subset	Poisson	2.07	-0.00010	0.000076	0.029	63	
		(0.14)	(0.00036)	(0.00036)	(0.039)	(113)	
	log-Normal	2.13	-0.000041	0.00011	0.073	56	0.013
		(0.16)	(0.00047)	(0.00048)	(0.052)	(76)	(0.011)

Normal model. If simulations of equal probable fields were made, these fields would be close to the original data.

For the posterior range distribution for the Poisson model we have a mode ϕ = 102 km and for the log-Normal model ϕ = 68 km. The effective correlation distance is three times larger, because of the exponential correlation function. The Poisson posterior is more uncertain, shown by its smaller peak and wider tail (see also Table 2.1). Further, we found the range to be positively correlated with the partial sill. The nugget effect is estimated only in the log-Normal model. Its mode is relatively small, σ_q^2 = 0.028, but it will introduce smoothing in the interpolation.

Fig. 2.5 shows the predicted spatial fields (left panels) and their corresponding standard deviations (right panels). Minimum and maximum values are given in Table 2.2. From the original number of exceedance days, the Poisson model (upper panels) predicts the Poisson intensity of the number of exceedance days, while the log-Normal model (lower panels) predicts the expected number of exceedance days. The patterns and values of both models are rather similar. Both models smoothed the original data.

The standard deviations clearly show the difference between what both models predict. Related to its lower sill and higher range, the Poisson model shows considerable lower values. This indicates the predicted intensity to be more certain than the expected number of exceedance days. In both models, the log-link yields larger standard deviations at those locations where predictions are larger. Cross validation (Table 2.3) shows that the RMSE for the Poisson model was smaller (7.09) than that for log-Normal model (7.15).

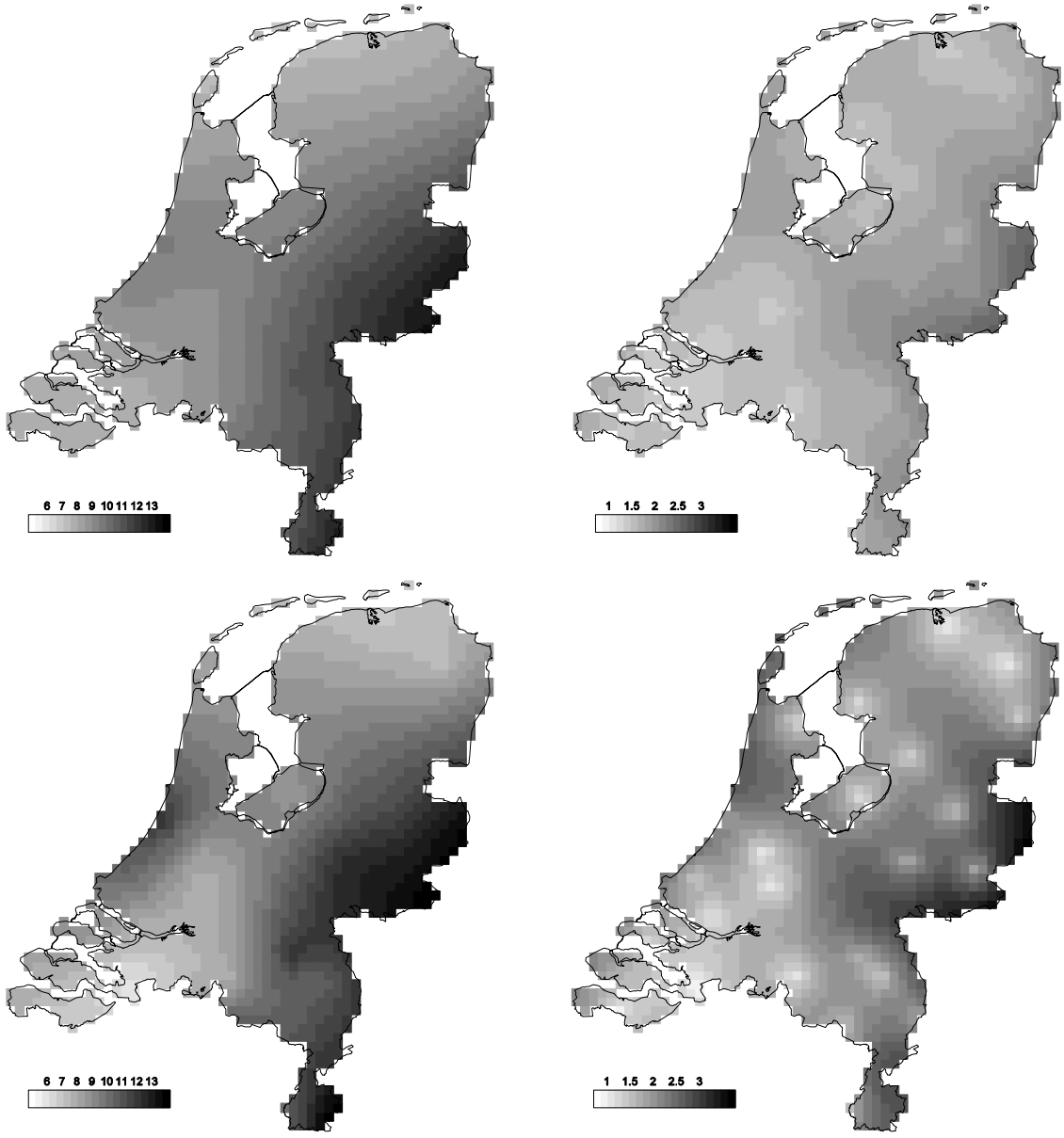

Fig. 2.6. Predicted number of ozone exceedance days (left) and corresponding kriging standard deviations (right) with the Poisson model (top) and log-Normal model (bottom) for the Netherlands in 2000 using the full data set.

2.3.2 Part 2: interpolation in the Netherlands using the full set

In the previous section, we showed the most important properties of both models. Because the effective range was approximately a few hundred kilometers, we also incorporated data from the surrounding countries Belgium and Germany. In this section we zoom in on the interpolation results for the Netherlands only, while using the parameter estimates from the full set.

Table 2.2. Minimum and maximum values of the data, model predictions and model standard deviations.

		Full set NL-B-D		Full set NL		Subset NL	
		min	max	min	max	min	max
	Observations	5.00	79.00	5.00	14.00	5.00	14.00
Poisson	Predictions	6.60	75.84	7.23	13.28	8.11	10.03
	Standard deviations	1.40	23.48	1.39	2.60	1.26	2.16
log-Normal	Predictions	5.78	83.71	6.05	13.97	5.85	12.42
	Standard deviations	1.20	34.47	1.00	3.56	0.85	2.69

Fig. 2.6 shows the predicted fields in the Netherlands (left panels) and their corresponding standard deviations (right panels). The presence of observations from Germany leads to higher values near the Netherlands-German border. The Poisson model (upper panels) shows more smoothing than the log-Normal model (lower panels), as is verified from the minimum and maximum predicted values (Table 2.2).

Standard deviations of the Poisson model are smaller and show less variation. This indicates that the predicted intensities are more certain than the expected number of exceedance days predicted by the log-Normal model. The RMSE of both models are practically equal, 2.28 and 2.27, respectively. The log-Normal model has shown less smoothing, but this has only a little effect on the RMSE.

2.3.3 Part 3: analysis and interpolation in the Netherlands using the subset

This section focuses on the Netherlands only. Parameter estimation and interpolation has been done using the subset of 22 observations. Posterior densities are shown in Fig. 2.7, with values of the modes and standard deviations given in Table 2.1. The posterior trend parameter vector of the Poisson model is again similar to that of the log-Normal model. Posterior modes are β = (2.07, -0.00010, 0.000076) and β = (2.13, -0.000041, 0.00011), respectively, indicating no significant trend in the data. The estimates for the Poisson model are more certain (Table 2.1).

The partial sill σ_s^2 in both models diminished in comparison to the values found using the full set. Posterior modes are 0.029 and 0.073, respectively. In particular for the Poisson model, the sill value has become very small, indicating that the Poisson model will only show little variation around its mean. As for the trend parameters, differences in uncertainty of σ_s^2 between the two models have grown.

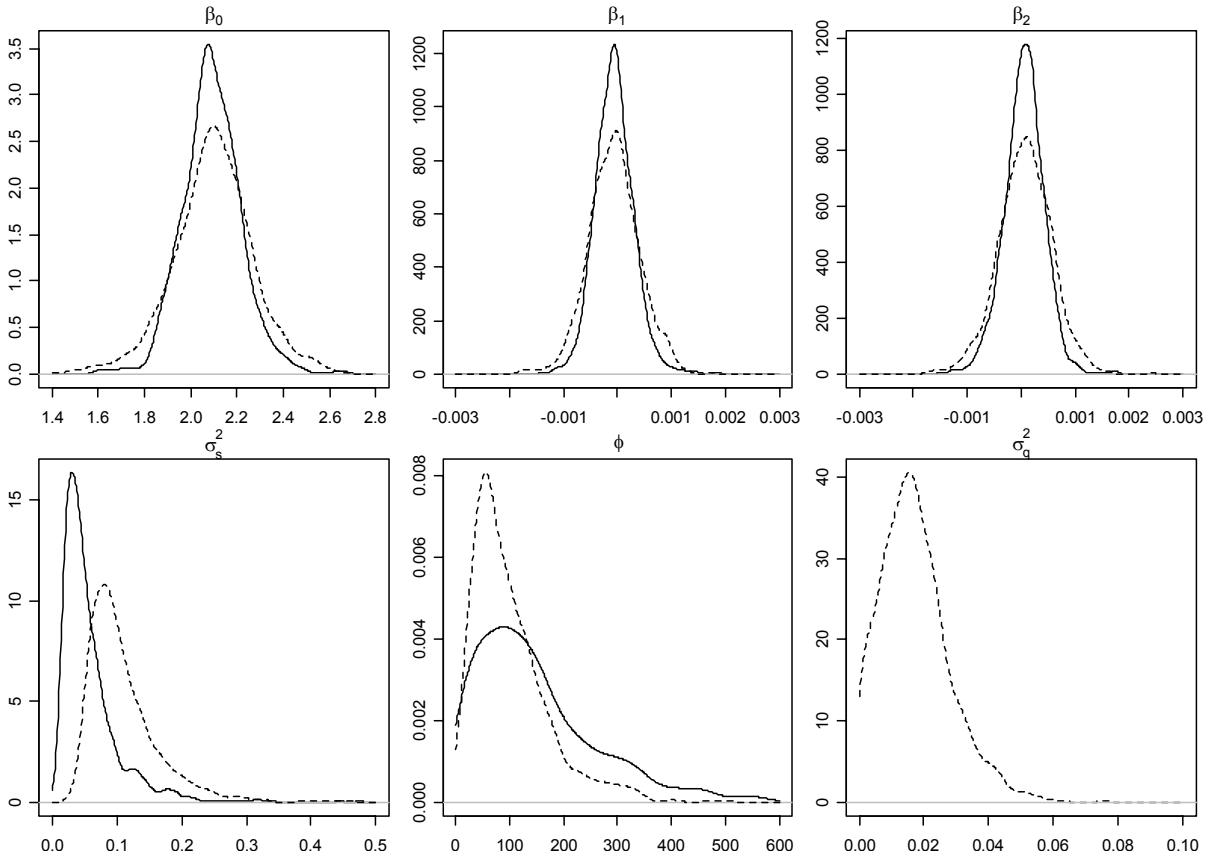

Fig. 2.7. Posterior distributions of the Poisson model parameters (solid line) and log-Normal model (dashed line) using the subset.

The posterior range parameter ϕ has also become smaller, with modes of 63 km and 56 km for the Poisson model and log-Normal model, respectively. The posterior range of the Poisson model (Fig. 2.7) has become more uncertain than the range in Fig. 2.4 (Table 2.1). It appeared to strongly depend on its prior. The nugget of the log-Normal model has also reduced (σ_q^2 = 0.013). Compared to the estimates using the full set, the standard deviations of all parameters, except the nugget, have increased (Table 2.1).

Fig. 2.8 shows the predictions (left panels) and corresponding standard deviations (right panels) of both models. Contrary to Fig. 2.6, three aspects can be clearly seen. First, the influence of the observations from the surrounding countries has disappeared, especially near the Netherlands-German border. Second, the Poisson model has larger smoothing, and third, the log-Normal model has less smoothed. Standard deviations for both models decreased. Minimum and maximum values are given in Table 2.2. Cross validation shows a lower RMSE for the log-Normal model (2.26) than for the Poisson model (2.44) (Table 2.3).

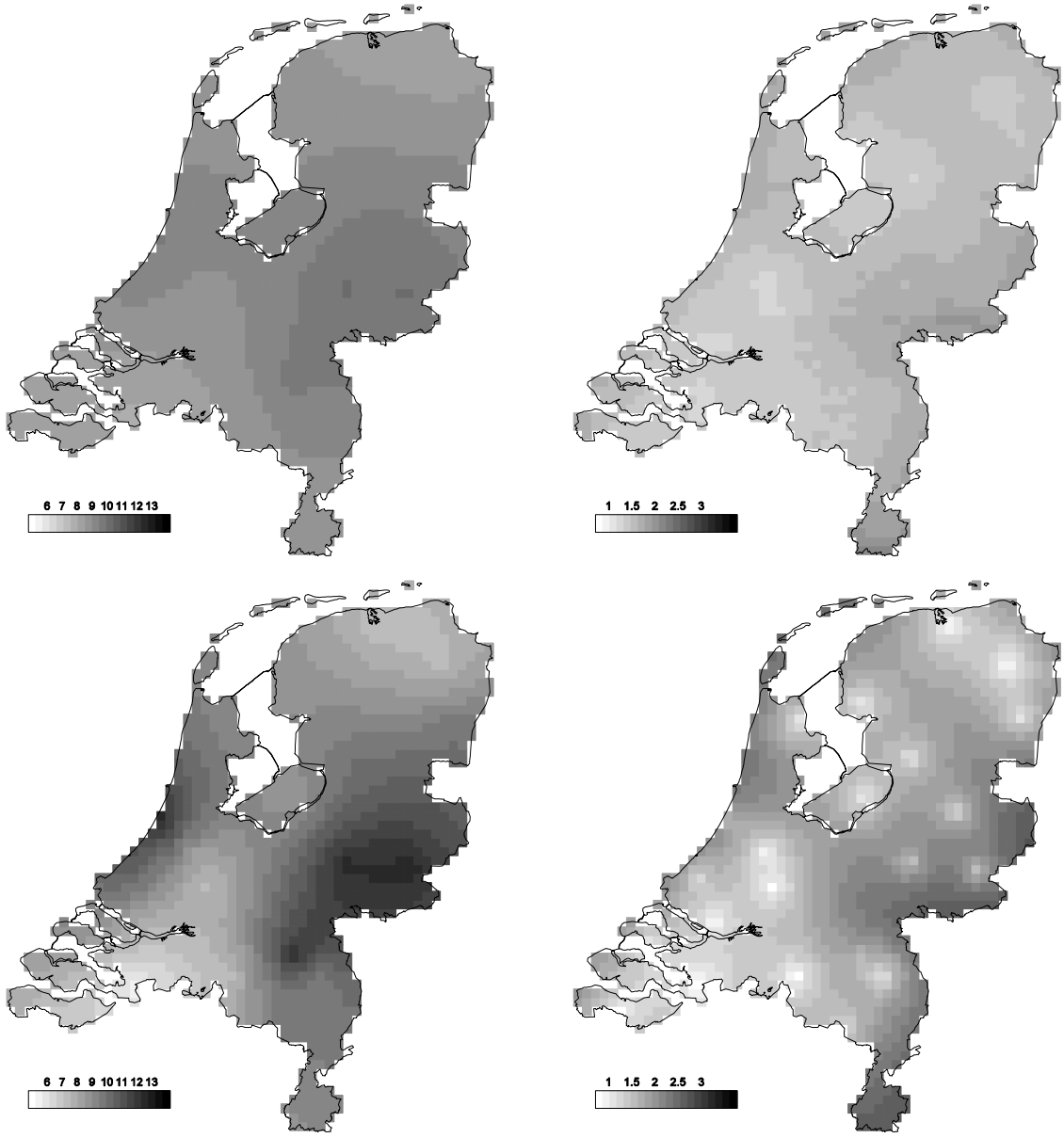

Fig. 2.8. Predicted number of ozone exceedance days (left) and corresponding kriging standard deviations (right) with the Poisson model (top) and log-Normal model (bottom) for the Netherlands in 2000 using the subset.

2.4 Discussion

The data used in this study represent the annual number of days in which ozone exceeds a threshold level. Observations were used from the Netherlands, Belgium and Germany. One may question the possibility of pooling data, since different countries may use different ozone measurement techniques. According to European quality control specifications (EC, 2002) however, measuring was done in a standardized way with calibrated equipment, but an intercomparison study

Table 2.3. Root mean-squared error values of the cross validation.

	Full set NL-B-D	Full set NL	Subset NL
Poisson	7.09	2.28	2.44
log-Normal	7.15	2.27	2.26

(Borowiak *et al.*, 2000) showed that the Netherlands ozone concentrations were measured 4% lower than its surrounding countries. We performed a recalculation of the number of exceedance days in the Netherlands, and the number of exceedance days increased from 0 to 5 days, depending on the station, with an average of 2.05 days. We could have corrected the data in advance, but on the other hand, in our research we analyzed data that were actually reported by the Netherlands Environmental Assessment Agency, without correcting them first. The correction should be done by the Agency before releasing the data.

The study showed the effective correlation distance to be approximately a few hundred kilometers. This satisfies analysis of the extensive data set covering the three countries. It further implies that when interpolating for the Netherlands only, data from surrounding countries have to be taken into account. One practical issue remains important as well: the Netherlands Environmental Assessment Agency needs to report the number of exceedance days as soon as the data has become available. Since data from other agencies can arrive late, analyzing only data from the Netherlands is then the ultimate possibility, but on the other hand, information from previous years can be used as prior information.

In the study we chose an explicit model-based geostatistical approach to interpolate the annual number of exceedance days. First, we assumed an approximation by a Poisson distribution, and second, a log-Normal distribution. The log-link in both models made model and parameter comparison easier. The advantage of using a Poisson model was that data could be analyzed as count data, with corresponding properties. This was indicated by increasing variance with increasing mean (Fig. 2.1) as well as by the simulation study (Fig. 2.3). Its disadvantage was that it did not properly fit the simulation study. The variance of the Poisson distribution was too small as compared to the variance of the simulation. Occurrences of exceedance days cluster in time, which weakens the assumption of a Poisson process, and it predicted the intensity of the annual number of exceedance days, as such complicating direct comparison with observations. The log-Normal model better fitted (Fig. 2.3). It also showed increasing variance with an increasing mean and it predicted the expected number of exceedance days. The disadvantage is that it handled data as continuous, which was not the case.

The most important difference between the models was that the Poisson model predicts an intensity field, whereas the log-Normal model predicted expected values. For this reason, σ_s^2 was smaller for the Poisson model (Table 2.2). As kriging standard deviations were smaller, the spatial predictions by the Poisson model seemed more accurate. Since expectation and variance are equal for a Poisson distribution, the Poisson model could describe more variation in the original data. Therefore the Poisson model described the original data by a smoothed intensity field that seemed more accurate.

The nugget of the log-Normal model can be considered similar to the Poisson variance and has a comparable effect to the smoothing properties of the Poisson model. This became clear for the full data set. For the subset, the nugget was lower, resulting in a less smoothed surface. The log-Normal model described most variation in the data by the underlying surface. The Poisson model, on the contrary, described this variation with its Poisson variance and showed a smoothed surface.

The choice for the prior of the range parameter was important. Earlier estimates using non-informative priors resulted in poor convergence in the MCMC algorithm. The choice of an exponential prior was an appropriate choice because it has the property that the correlation at a fixed distance was more uniformly distributed (Ribeiro and Diggle, 1999). The other parameters seemed less sensitive and the chains always converged to reasonable values given our priors. The priors could be more specified if data from past years were analyzed.

The log-Normal model described the variation in the original data well and incorporates it into the estimates of the spatial correlation parameters σ_s^2 and ϕ. Consequently, the predictions passed practically through the observations at the monitoring stations. The Poisson model on the other hand avoided this. Apparently, the original data could be described by the predicted Poisson intensity parameter. When predicting exceedances near a critical level, e.g. the maximum allowed exceedance days per year, the log-Normal model approach was more appropriate.

As a further extension, the number of exceedance days may in fact follow a Negative Binomial distribution. This distribution can account for overdispersion and may fit the number of exceedance days better than the Poisson distribution or log-Normal distribution. In this case, the intensity $\mathbf{m}$ come from a Gamma distribution where the parameters vary in space.

The question remains how to interpolate this kind of count data exactly. The real situation is complex. The conceptual process is as follows: during smog days, the concentration in one area (a range of about 100 km) increases, while in another area it does not. In the first area, an exceedance may occur, while in the other is does not. On another day, in the other area an exceedance may occur, while in the first area it

does not. On average, there will be more exceedances in a certain area, in this case the southern part of Germany. The data are in fact a summation of different spatially correlated data over time. This may introduce large variability in space on small scales. To avoid interpolation of count data directly and using the bulk of information in hourly observations, one could imagine spatial-temporal interpolation. This can by done by interpolating hourly observed ozone concentrations (for example, Guttorp *et al*., 1994) or daily maxima. In a second step, one can determine the number of days on every grid cell. Not only the inappropriate data assumptions or laborious MCMC parameter estimates can be avoided, but a more detailed map may also result. The primary interest is still the creation of an accurate national map showing the *actual* number of exceedance days at a certain location.

2.5. Conclusions

Two methods were discussed here for a model-based geostatistical interpolation of the annual number of exceedance days. The Poisson model was found to give a better representation of the random field process of the number of exceedance days. For environmental assessment applications, however, we concluded the log-Normal model to be the preferred method for interpolation, considering its capacity to predict the expected number of exceedance days instead of an intensity field.

When making interpolations of a component with a large spatial range, such as ozone exceedance days, for a small area such as the Netherlands, incorporating observations from surrounding countries in the analysis was beneficial since the effective correlation distance of the data was approximately 300 km. Predictions near the Dutch border still depend on observations in Germany. Furthermore, inclusion of more observations improved parameter estimation and resulted into more precise predictions.

Use of prior information in the Bayesian inference procedures avoids problems with convergence of the MCMC algorithm, which kept on fluctuating if flat priors were used in the subset. Also, even use of a limited data set allowed us to map the number of exceedance days. These maps, including their uncertainties, might be used in the future to study environmental relations between ozone and risks for public health.

3. External drift kriging with dispersion model output

In the mid-eighties the Dutch NO_x air quality monitoring network was reduced from 73 to 32 rural and city background stations, leading to higher spatial uncertainties. This necessitated the use of another source of information to help reduce uncertainties. Here, we focus on the use of secondary information from a dispersion model for obtaining high-resolution maps to compensate for the loss of spatial precision caused by a reduction in the number of stations. Our objective was to map atmospheric NO_x concentrations on rural and urban scales using kriging with external drift (KED) to merge the two sources of information. We also compared KED with universal kriging (UK). Because less reliable parameter estimates are obtained due to the reduced number of stations, Bayesian inference was applied and compared with parameter estimation by restricted maximum likelihood. The impact of several parameter estimation and spatial interpolation methods, the number of observations and configuration of the network on uncertainty were quantified by cross-validation. With KED, more accurate predictions and lower prediction standard deviations were obtained at the nodes of a fine-maze prediction grid. Where observations were sparse, the predictions were substantially improved by including dispersion model output and available prior information.

This chapter is based on J. van de Kassteele, A. Stein, A.L.M. Dekkers, and G.J.M. Velders (2005). External drift kriging of NO_x concentrations with dispersion model output in a reduced air quality monitoring network. *Environmental and Ecological Statistics* (under review).

3.1 Introduction

Accurate and spatially highly resolved maps of NO_x levels are essential to assessing individual human exposures to NO_x. It is well known that NO_x in high concentrations causes respiratory problems for humans (EPA, 1998; WHO, 2003). In the Netherlands, NO_x maps are based on a limited number of monitoring stations. In the mid-eighties the Dutch NO_x air quality monitoring network (Van Elzakker, 2001) was reduced from 73 to 32 stations for budgetary reasons. The increased distance between stations has caused a substantial loss of information and resulted in higher uncertainties in the maps. The combination of measurements with related external information from a dispersion model is likely to result in more accurate maps.

Kriging with external drift (KED) merges two sources of information: a primary variable that is precise but only available at a small number of locations, and a secondary variable that covers the full domain on a fine-mazed grid. KED has been applied in the past in environmental mapping of sparsely sampled data using dense external information. Examples include combining rainfall data with a digital elevation model (DEM) as a covariate (Pardo-Igúzquiza, 1998; Goovaerts, 2000; Drogue *et al*., 2002), combining rainfall data with satellite imagery (Grimes *et al*., 1999) or radar data (Cassiraga *et al*., 1997), combining temperature with a DEM (Hudson and Wackernagel, 1994), and combining temperature and land use for application in a crop growth model (Monestiez *et al*., 2001). In soil science and hydrology KED is applied in mapping soil horizons using a DEM (Bourennane *et al*., 1996; Bourennane *et al*., 2000), soil variables (Hengl *et al*., 2004), erosion mapping (Goovaerts, 1999), water table depths (Desbarats *et al*., 2002), and soil moisture content with precipitation (Snepvangers *et al*., 2003). Applications in air quality mapping are found in Pauly and Drueke (1996), where ozone is mapped using a DEM, in Bertino and Wackernagel (2002), where dispersion model output is used to map ozone concentrations around Paris, and in Genikhovich *et al*. (2002), where dispersion model output is combined with measurements to describe urban air quality.

This study focused on mapping yearly average atmospheric NO_x concentrations on rural and urban scales in the Netherlands. Secondary information was provided by the Operational Priority Substances (OPS) dispersion model (Van Jaarsveld, 1995), which calculates yearly average concentrations and deposition on the basis of emissions, dispersion, transport, chemical conversion, and wet and dry deposition. OPS also accounts for transport from adjacent countries. The model output consists of a national map with a spatial resolution of 5×5 km.

We explored the use of external drift kriging with the OPS model output in a reduced monitoring network A comparison was made with universal kriging (UK) by comparing UK and KED before (1983) and after network reduction (1987). Parameter estimation was carried out by means of restricted maximum likelihood and Bayesian inference. Our hypothesis was that Bayesian inference would show lower prediction uncertainties. By cross-validation, we quantified the relationship between the number of stations and occurring errors. A range of errors resulted from selecting several random configurations of different station densities from the 1983 configuration, describing explicitly the effect of the number of stations, and implicitly the effect of the station configuration.

3.2 Material and methods

3.2.1 Observations

The purpose of the Dutch air quality monitoring network is to monitor air quality in the Netherlands on a continual basis (Van Elzakker, 2001). The measurements obtained provide a general description of national, regional and local air quality, along with information on smog episodes; measurements are also tested against international air quality standards. The size of the area considered is about 260 × 310 km.

The monitoring network has undergone several changes in the past 25 years. The first reorganization took place in 1985/1986, resulting in a considerable reduction in the number of monitoring sites for SO_2 and NO_x measurements. A second reduction took place in 1994. In this study we used observations from 1983, two years before the first reduction in 1985. A total of 85 yearly average NO_x concentrations were available. Twelve street stations were omitted since they were not representative on the scale considered in this study. Three regional stations had to be omitted because of non-representative values due to missing data, leaving 64 rural background and 6 city background stations (Fig. 3.1a). High concentrations (in ppb) occurred in the western part of the Netherlands, around the major cities of Rotterdam, The Hague, Amsterdam and Utrecht, and near roadways. High concentrations were also found in the south-east, under influence of the German industrial Ruhr area, 50 km east of the Dutch-German border.

After the reduction, the total number of rural, city and street stations in 1987 came to 22, 5 and 7, respectively. A few regional stations had moved. For our analysis of the station configuration in 1987, we matched the existing locations of the

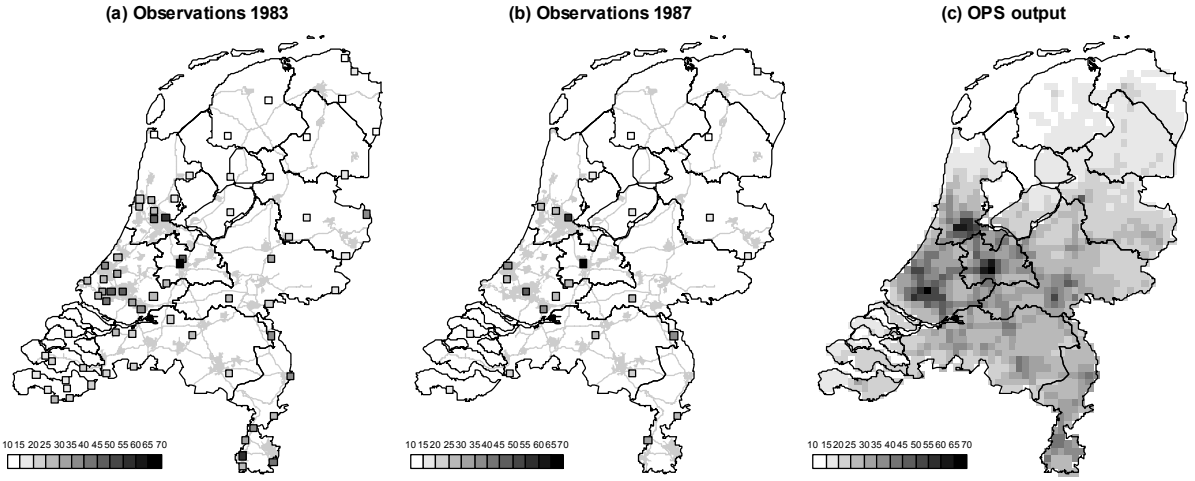

Fig. 3.1. Measured yearly average NO_x concentrations (ppb) in 1983 (a), in 1987 (b) and OPS model output for 1983 (c). The black lines indicate provincial boundaries, and gray lines and patches major roads and cities.

stations with those of 1983 and, where necessary, by selecting the nearest station (Fig. 3.1b). One regional station was excluded, since there was no possible match with a 1983 regional station. We maintained the concentrations of 1983.

3.2.2 The OPS dispersion model

The OPS model calculates average atmospheric concentrations and deposition from the atmosphere on the basis of emissions within the Netherlands and Europe. The model is suitable for a series of substances of which the behavior can be described by first-order linear chemical reactions; it cannot be used, for example, for describing ozone concentrations (Van Jaarsveld, 1995).

The processes described by the OPS model are emission, dispersion, transport, conversion, and wet and dry deposition. It is an analytical model using the Gaussian plume model for dispersion at local scales and operates as a Lagrangian trajectory model for long-distance transport. The OPS model is driven by actual meteorological observations and is statistical in the sense that dispersion is distributed over specific classes according to transport direction, atmospheric stability and scale of transport. Accompanying dispersion parameters are determined according to properties of all trajectories within that class. Yearly average concentration and deposition fields are found by weighting all classes according to the frequency of occurrence. Computationally speaking, this procedure is relatively rapid.

Input consists of emissions from sources into the atmosphere. Source properties like emission height are determinative for the dispersion. Since a detailed emission inventory for 1983 was not available for this study, emissions from 1995

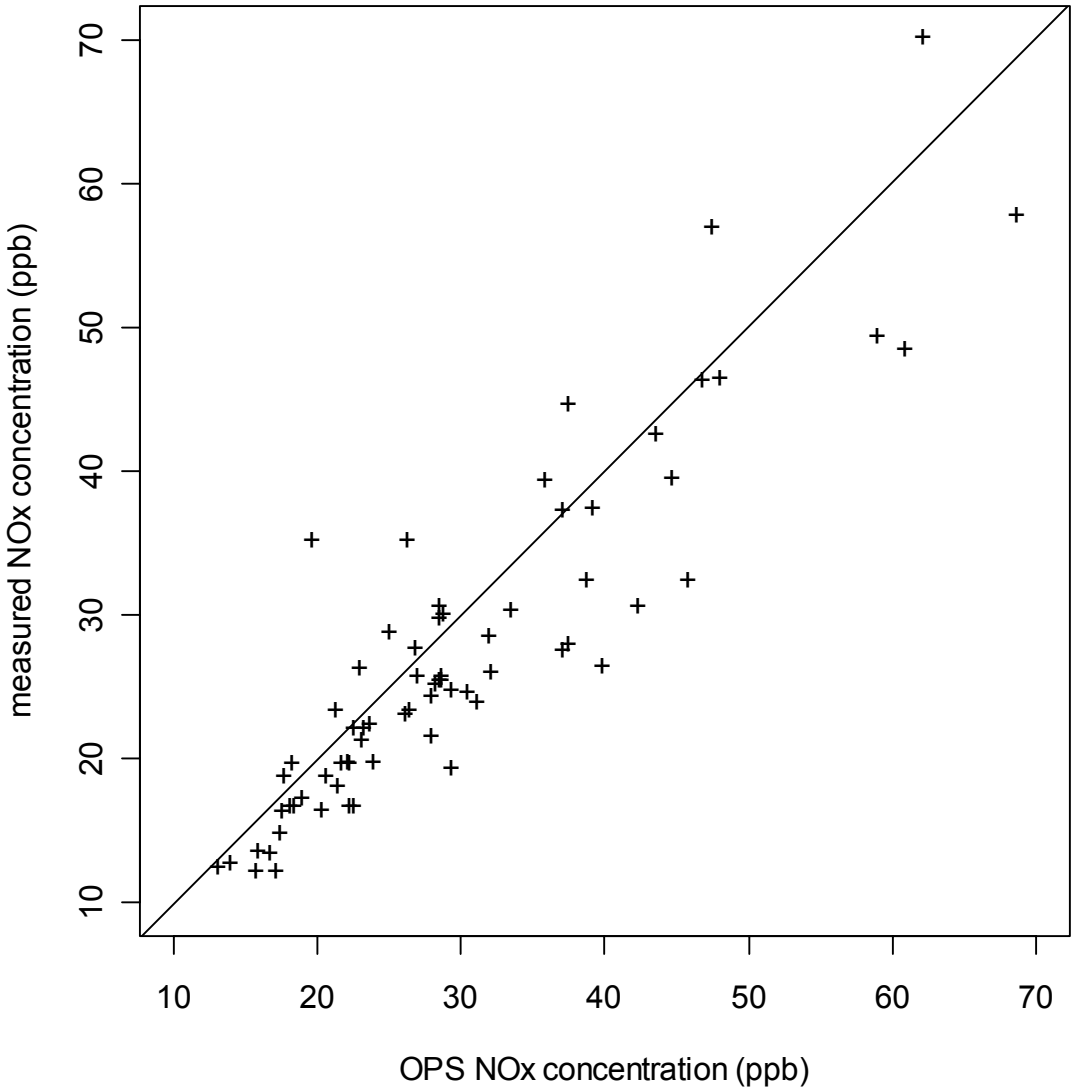

Fig. 3.2. Measured yearly average NO_x concentration (ppb) vs. modeled OPS NO_x concentration (ppb) at the 70 monitoring stations in 1983.

were taken and scaled proportionally to known total emissions per source group for 1983. Output is represented by a concentration field on a 5 × 5 km grid (Fig. 3.1c). The results are valid for the center of grid boxes only and do not represent averages for the whole grid box area.

The OPS model is described in Van Jaarsveld (1991), Van Jaarsveld and De Leeuw (1993), and Van Jaarsveld (1995). In Van Jaarsveld (1995), the model is compared with measurements on different levels, e.g. process descriptions such as mixing height and descriptions of vertical dispersion. The model has played a role in international comparison studies (Derwent *et al.*, 1989). It also generates data at the monitoring station locations (Fig. 3.2) and does well at predicting yearly average NO_x concentrations, although predictions are systematically higher than observations.

3.2.3 Universal kriging vs. external drift kriging

For kriging measured NO_x concentrations, we applied the model-based approach of Ribeiro and Diggle (1999) and Diggle and Ribeiro (2002). After a Box-Cox transformation, the n observations y_i, with $i = 1 \ldots n$, were interpreted as a realization of a Gaussian random variable Y_i at locations $\mathbf{s}_i$. Each observation can be decomposed into

$$y_i = \mu_i + \varepsilon_{s_i} + \varepsilon_{q_i} \tag{3.1}$$

where $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ is the deterministic trend component, $\mathbf{X}$ an $n \times p$ matrix consisting of p known trend components at each location and $\boldsymbol{\beta}$, a vector with p unknown trend parameters. $\boldsymbol{\varepsilon}_s$ is a realization of a zero-mean stationary Gaussian random process with a partial sill variance σ_s^2. The vector $\boldsymbol{\varepsilon}_s$ accounts for spatial correlation between observations by means of a spatial correlation function $\rho(.)$, with range parameter ϕ. Finally, $\boldsymbol{\varepsilon}_q$ is an error term with variance σ_q^2 (nugget).

Kriging with external drift (KED) is a particular case of universal kriging (UK) (Bourennane *et al.*, 2000). The difference between UK and KED lies in the trend component. For UK, $\mathbf{X}$ is a function of the coordinates $\mathbf{s}_1$ and $\mathbf{s}_2$ in two orthogonal directions, whereas for KED, $\mathbf{X}$ is a function of the OPS model output at locations $\mathbf{s}$, i.e. $\mathbf{X} = [1\ \mathbf{s}_1\ \mathbf{s}_2]$ and $\mathbf{X} = [1\ OPS(\mathbf{s})]$, for UK and KED, respectively. Note that $\boldsymbol{\varepsilon}_s$ is different for UK and KED.

3.2.4 Restricted maximum likelihood vs. Bayesian inference

A common method to estimate parameters is fitting the parametric correlation function $\rho(.)$ to the empirical variogram, obtained by binning and averaging the squared differences between point pairs after de-trending. Here, we applied restricted maximum likelihood (RML) and Bayesian inference, methods working directly from the sample data.

Maximum likelihood estimation is widely used in statistics. It has the advantage of providing joint estimation of trend and covariance parameters. However, it introduces a bias if the number of observations, n, is small compared to the number of covariates p. With RML, the trend parameters are integrated out of the likelihood function, leaving unbiased estimators for the covariance parameters (Kitanidis and Shen, 1996). RML estimation requires the data to be a realization of a multivariate Gaussian distribution (see section 3.2.5). If the observations $\mathbf{y}$ and

covariates **X** are regarded as fixed, the restricted likelihood function of the above model will depend only on the unknown parameters σ_s^2, ϕ, and σ_q^2. When calculating the likelihood, it is useful to re-parameterize, using the relative nugget, $\sigma_{q,rel}^2 = \sigma_q^2 / \sigma_s^2$ (Ribeiro and Diggle, 1999). This re-parameterization allows us to state the covariance matrix $\mathbf{V}_y$ as a product of σ_s^2 and matrix $\mathbf{R}_y = \rho(\mathbf{H}_y/\phi) + \mathbf{I}_n \sigma_{q,rel}^2$, where $\mathbf{H}_y$ is the $n \times n$ distance matrix containing the distances between stations. The restricted log likelihood equals:

$$RLL(\mathbf{y} \mid \sigma_s^2, \phi, \sigma_{q,rel}^2) = -\frac{n-p}{2}\log(2\pi\sigma_s^2) - \frac{1}{2}\log\left(\frac{\left|\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X}\right|\left|\mathbf{R}_y\right|}{\left|\mathbf{X}^{\mathrm{T}}\mathbf{X}\right|}\right) - \frac{\mathbf{y}^{\mathrm{T}}(\mathbf{R}_y^{-1} - \mathbf{R}_y^{-1}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1})\mathbf{y}}{2\sigma_s^2} \quad (3.2)$$

where the superscripts $^{\mathrm{T}}$ and $^{-1}$ denote the transpose and inverse of a matrix, respectively. For any fixed value of ϕ and $\sigma_{q,rel}^2$, RML estimates for σ_s^2 and $\boldsymbol{\beta}$ are equal to

$$\hat{\sigma}_s^2 = \frac{1}{n-p}\mathbf{y}^{\mathrm{T}}(\mathbf{R}_y^{-1} - \mathbf{R}_y^{-1}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1})\mathbf{y} \quad (3.3a)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{y} \quad (3.3b)$$

Estimates for ϕ and $\sigma_{q,rel}^2$ cannot be written in closed form. They are found by substituting $\hat{\sigma}_s^2$ in the restricted log likelihood function and minimizing the negative log likelihood over the ϕ and $\sigma_{q,rel}^2$ parameter space.

In the case of parameter estimation with (restricted) maximum likelihood, the predictive distribution is multivariate Normal, with mean and variance equal to

$$E[\tilde{\mathbf{y}} \mid \mathbf{y}] = \tilde{\mathbf{X}}\hat{\boldsymbol{\beta}} + \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (3.4a)$$

$$\mathrm{Var}[\tilde{\mathbf{y}} \mid \mathbf{y}] = \sigma_s^2[\mathbf{R}_{\tilde{y}} - \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{R}_{y,\tilde{y}} + (\tilde{\mathbf{X}} - \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X})^{-1}(\tilde{\mathbf{X}} - \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{X})^{\mathrm{T}}] \quad (3.4b)$$

where $\tilde{\mathbf{X}}$ is an $m \times p$ matrix containing the p covariates at m prediction locations $\tilde{\mathbf{s}}$, $\mathbf{R}_y = \rho(\mathbf{H}_y/\phi) + \mathbf{I}_n \sigma_{q,rel}^2$ the $n \times n$ correlation matrix between observation locations, $\mathbf{R}_{y,\tilde{y}} = \rho(\mathbf{H}_{y,\tilde{y}}/\phi)$ the $n \times m$ correlation matrix between the observation and prediction locations, and $\mathbf{R}_{\tilde{y}} = \rho(\mathbf{H}_{\tilde{y}}/\phi) + \mathbf{I}_m \sigma_{q,rel}^2$ the $m \times m$ correlation matrix between the prediction locations. The expression for the variance has three components: the first

is the marginal variance for $\tilde{\mathbf{y}}$, the second is the variance reduction after observing $\mathbf{y}$ and the third component accounts for the uncertainty in the trend parameter $\boldsymbol{\beta}$.

Bayesian inference treats the parameters as unknown stochastic variables and, as such, incorporates the uncertainty of all parameters (Gelman *et al.*, 2004; Berger *et al.*, 2001). Prior distributions, jointly written as $p(\boldsymbol{\beta},\ \sigma_s^2,\ \phi,\ \sigma_{q,rel}^2)$, are assigned to the parameters. This is attractive in the case of sparse data, since additional prior information can be used, thereby improving the estimation accuracy (see, for example, Cui *et al.*, 1995). The priors are updated by observations with the use of the likelihood function

$$p(y \mid \beta, \sigma_s^2, \phi, \sigma_{q,rel}^2) = (2\pi\sigma_s^2)^{-n/2} \mid \mathbf{R}_y \mid^{-1/2} \exp[-(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}} \mathbf{R}_y^{-1} (\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) / 2\sigma_s^2] \tag{3.5}$$

resulting in the posterior parameter distribution

$$p(\boldsymbol{\beta},\ \sigma_s^2,\ \phi,\ \sigma_{q,rel}^2 \mid \mathbf{y}) \propto p(\boldsymbol{\beta},\ \sigma_s^2,\ \phi,\ \sigma_{q,rel}^2)\, p(\mathbf{y} \mid \boldsymbol{\beta},\ \sigma_s^2,\ \phi,\ \sigma_{q,rel}^2) \tag{3.6}$$

The above posterior distribution can be written in analytical form if we use conjugate priors or flat priors for $\boldsymbol{\beta}$ and σ_s^2, and if ϕ and $\sigma_{q,rel}^2$ are known (Ribeiro and Diggle, 1999). However, ϕ and $\sigma_{q,rel}^2$ are unknown. If we factorize the above distribution and rearrange it, we obtain an expression for the joint posterior for ϕ and $\sigma_{q,rel}^2$

$$p(\phi,\ \sigma_{q,rel}^2 \mid \mathbf{y}) \propto \frac{p(\boldsymbol{\beta}, \sigma_s^2 \mid \phi, \sigma_{q,rel}^2)\, p(\phi, \sigma_{q,rel}^2)\, p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma_s^2, \phi, \sigma_{q,rel}^2)}{p(\boldsymbol{\beta}, \sigma_s^2 \mid \mathbf{y}, \phi, \sigma_{q,rel}^2)} \tag{3.7}$$

where the distributions conditional on ϕ and $\sigma_{q,rel}^2$ can still be written in analytical form. If the prior $p(\phi,\ \sigma_{q,rel}^2)$ is specified, the posterior $p(\phi,\ \sigma_{q,rel}^2 \mid \mathbf{y})$ can be evaluated. Ribeiro and Diggle (2001) do this by evaluating the above posterior on a grid of points that covers the parameter space of interest. Once the grid of density values is computed, random draws from $p(\phi,\ \sigma_{q,rel}^2 \mid \mathbf{y})$ are taken to approximate the posterior densities for ϕ and $\sigma_{q,rel}^2$ (Gelman *et al.*, 2004).

The posterior densities for $\boldsymbol{\beta}$ and σ_s^2 are then easily evaluated by plugging the obtained draws for ϕ and $\sigma_{q,rel}^2$ in the analytical expression for the posterior density $p(\sigma_s^2 \mid \mathbf{y},\ \phi,\ \sigma_{q,rel}^2)$, a scaled inverse-$\chi^2$ distribution, and the posterior density $p(\boldsymbol{\beta} \mid \mathbf{y},$

σ_s^2, ϕ, $\sigma_{q,rel}^2$), a multivariate normal distribution. After drawing from these distributions, the joint posterior parameter distribution is fully evaluated.

The predictive distribution $p(\tilde{\mathbf{y}} \mid \mathbf{y}, \boldsymbol{\beta}, \sigma_s^2, \phi, \sigma_{q,rel}^2)$ is a multivariate normal distribution, with mean and variance

$$E[\tilde{\mathbf{y}} \mid \mathbf{y}] = \tilde{\mathbf{X}}\boldsymbol{\beta} + \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \tag{3.8a}$$

$$\mathrm{Var}[\tilde{\mathbf{y}} \mid \mathbf{y}] = \sigma_s^2[\mathbf{R}_{\tilde{y}} - \mathbf{R}_{y,\tilde{y}}^{\mathrm{T}}\mathbf{R}_y^{-1}\mathbf{R}_{y,\tilde{y}}] \tag{3.8b}$$

and is evaluated analogously by plugging in the obtained parameter samples in the distribution and, subsequently, drawing samples from this distribution. In conventional geostatistics, the above expression is that for simple kriging, since all parameters are considered known. However, parameter uncertainty is incorporated directly by the draws from the posterior parameter distribution. Prediction using RML only accounts for uncertain trend parameters, but RML and Bayesian inference are similar if a uniform prior for $\boldsymbol{\beta}$ is chosen and the other parameters are kept fixed (i.e. are known).

Bayesian geostatistics are described by Handcock and Stein (1993). Ribeiro and Diggle (1999) and Diggle and Ribeiro (2002) further expand on this description. More details on UK and KED can be found in Chilès and Delfiner (1999) and Wackernagel (1995).

3.2.5 Set-up of the kriging models

The Gaussian model-based approach requires residuals after de-trending to be stationary and normally distributed. We applied the Box-Cox transformation (Box and Cox, 1964) of the original data $\mathbf{y}_{org}$ to obtain such residuals:

$$\mathbf{y} = g(\mathbf{y}_{org}) = \begin{cases} (\mathbf{y}_{org}^{\lambda} - 1)/\lambda & (\lambda \neq 0) \\ \log(\mathbf{y}_{org}) & (\lambda = 0) \end{cases} \tag{3.9}$$

Here $g(.)$ is the transformation function and λ is the Box-Cox parameter. Maximum likelihood estimates, without taking into account spatial effects, yielded Box-Cox parameters of λ = -0.63 for UK and λ = 0.34 for KED. We could have applied a Bayesian estimation of λ, but as in Ribeiro and Diggle (2001), we kept them known throughout. If not, each λ would have changed the location and scale of the transformed data, as well as the correlation structure (De Oliveira *et al.*, 1997). Other

transformations, for example, the Normal Score transformation (Lehmann, 1975), could also have been applied.

The assumption of normality of the residuals after transformation is strong, but it cannot be tested in practice because there is only one realization of the random field. However, to get more insight into this, we employed a Shapiro-Wilk test to check for evidence against normality. Spatial correlation in the residuals is not taken into account. The test yielded a p-value = 5.038E-7 and a p-value = 1.470E-4 for UK and KED, respectively, before transformation, and a p-value = 0.2371 and a p-value = 0.01461 for UK and KED, respectively, after transformation. This provided strong evidence (p-value < 0.01) that the residuals of both UK and KED before transformation were not normally distributed. It also showed that there was little or no real evidence (p-value > 0.1) that the UK residuals after transformation were not normally distributed and, finally, that there was moderate evidence (0.01 < p-value < 0.05) that the KED residuals after transformation were not normally distributed.

The methods that we described in the previous section produce spatial predictions on a transformed scale. One cannot apply the inverse Box-Cox function directly to the expectations and variances because it would introduce biased predictions. This can be avoided by back-transforming the conditional simulations of the predictive distribution $p(\tilde{\mathbf{y}} \mid \mathbf{y})$. The most common predictor is the expectation of the back-transformed conditional simulations, but for many back-transformations of a Gaussian random field, the mean may not exist. Therefore, just as De Oliveira *et al*. (1997), we used the median as predictor and a fourth of the 95% credible interval as standard deviation:

$$\text{med}[\,\tilde{\mathbf{y}}_{org} \mid \mathbf{y}_{org}\,] = q_{50}[g^{-1}(\tilde{\mathbf{y}} \mid \mathbf{y})] \tag{3.10a}$$

$$\text{sd}[\tilde{\mathbf{y}}_{org} \mid \mathbf{y}_{org}] = \frac{q_{97.5}[g^{-1}(\tilde{\mathbf{y}} \mid \mathbf{y})] - q_{2.5}[g^{-1}(\tilde{\mathbf{y}} \mid \mathbf{y})]}{4} \tag{3.10b}$$

where $q(.)$ produces sample quantiles corresponding to the given probabilities. A probability of exceeding a threshold level can be obtained analogously (Ribeiro and Diggle, 2001).

An exponential correlation function suited both UK and KED for the spatial covariance structure. The effective range or correlation distance equals 3ϕ.

Prior parameter distributions have to be specified for Bayesian inference and should be independent of the data. For UK, independent prior information was available in the form of the OPS model output, where the OPS model output had first been transformed with the UK Box-Cox parameter. After that, we started computations with non-informative flat priors for $\boldsymbol{\beta}$, σ_s^{-2}, ϕ^1, and $\sigma^2_{q,rel}$. For

computational stability we had to limit the number of transformed OPS values to 200, at randomly chosen locations. The posterior distributions obtained were used as priors for the UK analysis. We fitted a multivariate Normal distribution to (β | **OPS**) and a scaled inverse-χ^2 distribution to (σ_s^2 | **OPS**). No standard statistical distributions were fitted to (ϕ | **OPS**) and ($\sigma_{q,rel}^2$ | **OPS**), because these samples were used directly in the UK analysis. The priors are

$$\beta \sim N\left[\begin{pmatrix}1.53\text{E}+0\\ 1.80\text{E-}4\\ -5.18\text{E-}4\end{pmatrix}, \sigma_s^2\begin{pmatrix}4.48\text{E}+0 & -2.01\text{E-}3 & -7.72\text{E-}3\\ -2.01\text{E-}3 & 2.21\text{E-}5 & -3.43\text{E-}6\\ -7.72\text{E-}3 & -3.43\text{E-}6 & 1.79\text{E-}5\end{pmatrix}\right] \quad (3.11a)$$

$$\sigma_s^2 \sim \text{Inv-}\chi^2\,(11,\ 2.27\text{E-}3) \quad (3.11b)$$

$$\phi \sim p(\phi \mid \textbf{OPS}) \quad (3.11c)$$

$$\sigma_{q,rel}^2 \sim p(\sigma_{q,rel}^2 \mid \textbf{OPS}) \quad (3.11d)$$

The small (co)variances for β and σ_s^2 are due to the Box-Cox transformation. The values of ϕ and $\sigma_{q,rel}^2$ were evaluated over a grid defined by (0, 300] km and [0, 0.3] respectively.

Since no prior information was available for the KED parameters, we started with non-informative priors, similar to those we used for analyzing the OPS output. The priors for the KED parameters are

$$\beta \sim \text{N}(0,\ 1000) \quad (3.12a)$$

$$\sigma_s^{-2} \sim \text{Gamma}(0.001,\ 0.001) \quad (3.12b)$$

$$\phi^{-1} \sim \text{Uniform}(0.00005,\ 1000) \quad (3.12c)$$

$$\sigma_{q,rel}^2 \sim \text{Uniform}(0,\ 0.5) \quad (3.12d)$$

Here ϕ and $\sigma_{q,rel}^2$ were evaluated over a grid defined by (0, 20000] km and [0, 0.5], respectively. The range parameter should be large, since the range for KED was expected to grow considerably (see section 3.3.1). The bounds of ϕ and $\sigma_{q,rel}^2$ were also applied to the RML parameter estimation. The eventual posterior parameter densities are discussed in section 3.3.1.

3.2.6 Validation procedure

A cross-validation by 'leaving one out' was carried out to study the performance of the models, given a number of stations n_s and network configurations. As in Cui *et al*. (1995) we used three error measures for validation: mean error (ME or bias), unbiased root mean squared error (URMSE) and mean squared standardized error (MSSE):

$$ME = \frac{1}{n_s}\sum_{i=1}^{n_s}(\tilde{y}_{org,i} - y_{org,i}) \tag{3.13a}$$

$$URMSE = \sqrt{\frac{1}{n_s}\sum_{i=1}^{n_s}(\tilde{y}_{org,i} - y_{org,i})^2 - ME^2} \tag{3.13b}$$

$$MSSE = \frac{1}{n_s}\sum_{i=1}^{n_s}\frac{(\tilde{y}_{org,i} - y_{org,i})^2}{\mathrm{Var}(\tilde{y}_{org,i})} \tag{3.13c}$$

The ME indicates the bias of the predictions $\tilde{\mathbf{y}}_{org}$ to the original observations $\mathbf{y}_{org}$ and should be close to zero. The URMSE indicates the bias corrected standard deviation of the model and should be close to zero. We used the URMSE because it can be close to zero even in presence of a bias. The MSSE compares the squared differences with the kriging model variance, and yields a value that should be close to one.

Unlike other studies, for example, in Cui *et al*., 1995 and Bourennane *et al*., 2000, where only one random set of different sizes is selected from the full set, we studied the effect of different configurations by selecting several random configurations of n_s stations (n_s = 10, 20, …, 70) from the complete set of 70. Cross-validation was performed for each selection. The rationale for selecting several configurations is that every configuration leads to a different value for the ME, URMSE and MSSE. Since many combinations were possible, new configurations of n_s stations were selected until the 2.5, 50 and 97.5-quantiles of the outcomes of ME, URMS and MSSE reached stability. These values indicate the sensitivity of the models to the network configurations considered. If we had just selected one random configuration for every station sample density, we could have ended with accidental low or high values for some station sample density. Note that different configurations cannot be independent of each other if they share one or more stations within the correlation distance.

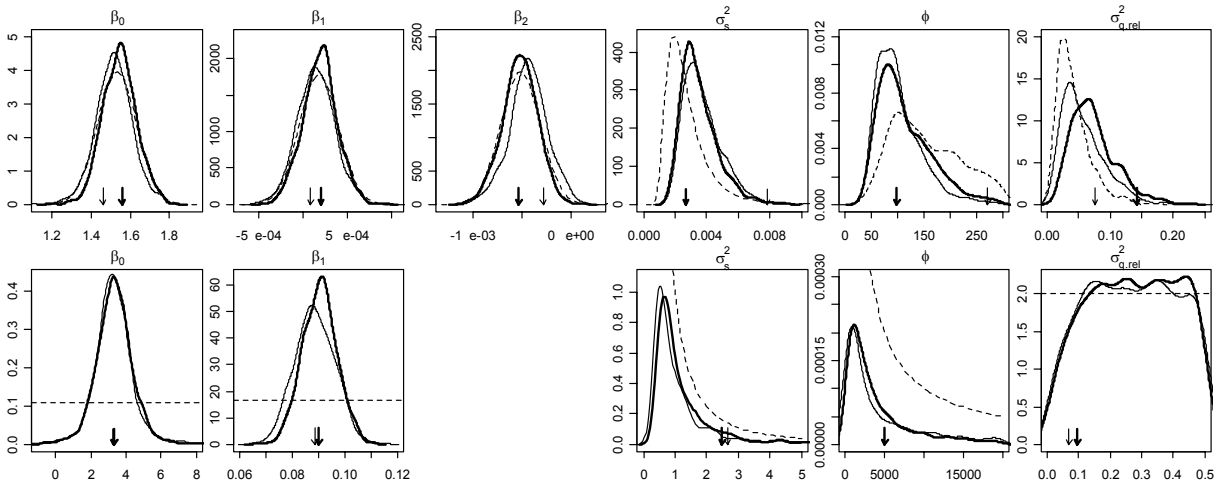

Fig. 3.3. Bayesian posterior parameter densities (curves) and the RML parameter estimates (arrows) for both UK (upper 6 panels) and KED (lower 5 panels). The bold solid lines represent 1983 posteriors and the fine solid lines 1987 posteriors. The dashed lines represent the prior distributions. The priors for UK are the posteriors based on OPS output, while priors for KED are non-informative.

3.3 Results

3.3.1 Parameters before and after network reduction

Fig. 3.3 shows the Bayesian posterior parameter densities (curves) and the RML parameter estimates (arrows). The 1983 posterior trend parameters β are more precise for both UK and KED; i.e. they show a density curve that is steeper and higher than curves for 1987 because of the larger number of stations. The trend parameters appear to be very precise, but this is due to the Box-Cox transformation. The RML estimates correspond well with the posterior modes.

The covariance parameters show different results. The posterior for the partial sill σ_s^2 for UK shows larger values than the prior based on OPS output. For KED, the posterior σ_s^2 is totally determined by the observations because of the non-informative prior. Posterior σ_s^2 obtained from the 1983 or the 1987 data are largely similar.

For UK, the posterior for the range parameter ϕ also differs from the prior. Based on OPS output, the mode is approximately 100 km, whereas the posterior mode for both 1983 and 1987 is approximately 80 km. The RML estimates differ considerably. For KED, both 1983 and 1987 posterior ϕ look similar, but have high values. The RML estimates for ϕ are the same. For UK, the posterior relative nugget

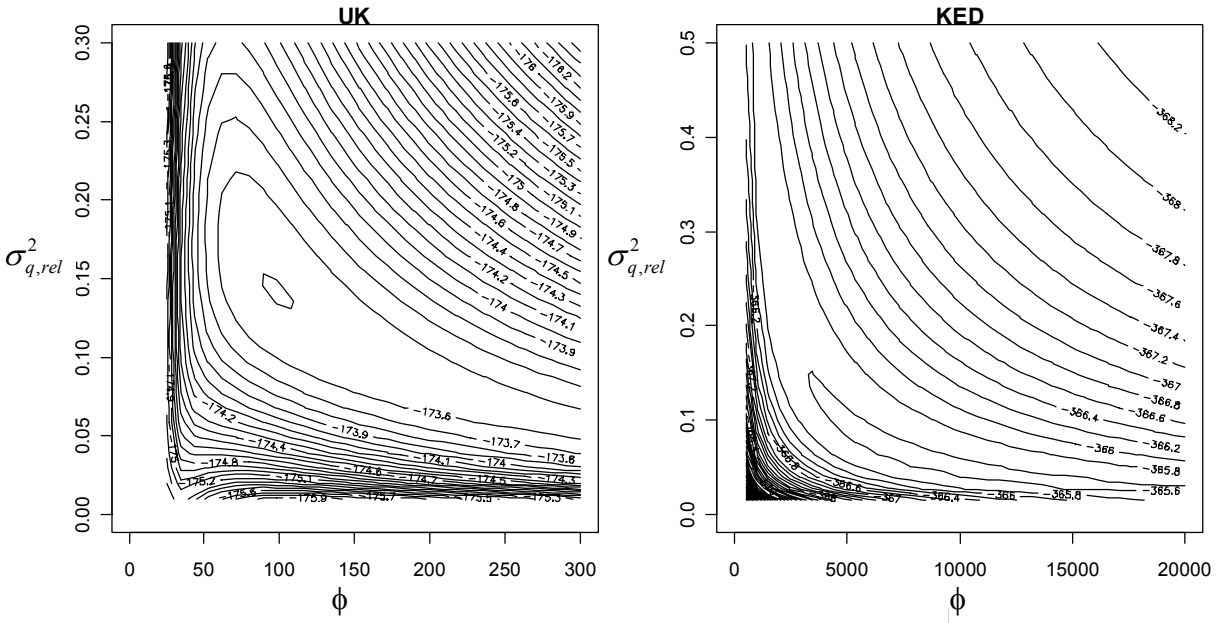

Fig. 3.4. Profile likelihood as a function of ϕ and $\sigma^2_{q,rel}$ for UK and KED for the 1983 data.

$\sigma^2_{q,rel}$, is higher than its prior, which is more pronounced due to more available stations. For KED, the posterior $\sigma^2_{q,rel}$ is very flat, just like its prior. The RML estimates are approximately equal to 0.1 for both years.

The ϕ and $\sigma^2_{q,rel}$ parameters are related (Fig. 3.4). For UK, a maximum occurs at approximately ϕ = 100 km and $\sigma^2_{q,rel}$ = 0.15. For KED, this maximum is very flat, stretching out over a whole range of possible ϕ and $\sigma^2_{q,rel}$ values. For KED, it is therefore difficult to estimate these parameters using RML, of which the maximum is found at approximately ϕ = 5000 km and $\sigma^2_{q,rel}$ = 0.1. For 1987, both profile likelihood surfaces were found to be flatter due to fewer observations. However, the figure is not shown.

Fig. 3.4 explains the shape of the posterior density curves of ϕ and $\sigma^2_{q,rel}$ in Fig. 3.3. High values for both UK and KED are shown in Fig. 3.4 in the banana-shaped area between low values for ϕ and corresponding high values for $\sigma^2_{q,rel}$, and high values for ϕ and corresponding low values for $\sigma^2_{q,rel}$. If ϕ is low, the corresponding values for $\sigma^2_{q,rel}$ will be high. For UK this results in a density curve similar to the posterior for $\sigma^2_{q,rel}$ in Fig. 3.3. For KED this results in a flat posterior for $\sigma^2_{q,rel}$. The banana-shaped area also explains the differences between the RML estimates and the posterior modes. Note prior information is included for UK in Fig. 3.3.

For KED, ϕ extends to values far beyond the largest area of the Netherlands (approximately 400 km). Furthermore, ϕ has to be multiplied by three to compute the

effective range of the exponential model. For such large values, the exponential model for the variogram approaches a linear model for distances inside the Netherlands. The correlation between stations is no longer a function of their distance apart. Apparently, the information from the OPS model is so good, that it takes most of the spatial correlation in the residuals away. In this case, KED almost becomes similar to a special case of KED: linear regression. However, it does not change the results, because KED still applies to the data.

3.3.2 Spatial predictions and standard deviations before and after network reduction

We made predictions for 1) UK and KED, 2) based on the 1983 and 1987 data, 3) parameter estimates by RML and Bayesian inference, and 4), for the 1987 data, keeping the 1983 parameters fixed and re-estimating the parameters for the 1987 data. We call the respective predictions: UK 83 RML, UK 83 Bayes, UK 87 RML fix, UK 87 Bayes fix, UK 87 RML re, UK 87 Bayes re, KED 83 RML, KED 83 Bayes, KED 87 RML fix, KED 87 Bayes fix, KED 87 RML re and KED 87 Bayes re.

Fig. 3.5 illustrates the spatial predictions and corresponding standard deviations of NO_x concentrations with UK and KED on the basis of 1983 data and parameter estimation with RML. For UK the predicted concentrations vary gradually in space, whereas for KED the predicted NO_x concentrations show more spatial variation in terms of alternating higher and lower values. Individual cities and highways are clearly visible as a result of using the OPS output as external trend. For UK, the standard deviations look very similar to the predicted expectation and show more variation in space than the KED standard deviations. Because of the back-transformation, the standard deviations are correlated with the predicted expectations. In fact, the predictive distribution at each location is skewed.

The spatial patterns of the other predictions look similar and the corresponding figures are therefore not shown. Instead, Fig. 3.6 shows box plots based on the 1405 individual locations with the minimum, 25-quantile, median, 75-quantile and maximum values for the twelve predictions and standard deviations.

Based on Figs. 3.5 and 3.6, UK and KED can be said to differ considerably. We will therefore discuss UK and KED separately. Of interest are the differences before and after network reduction (1983 and 1987), the method of parameter estimation and, for 1987, keeping parameters fixed or not.

The median for UK 87 is systematically 9% lower than for UK 83. Re-estimating the parameters gives the same result as keeping them fixed. Bayesian inference shows higher extremes than RML. The standard deviations for UK 87 are

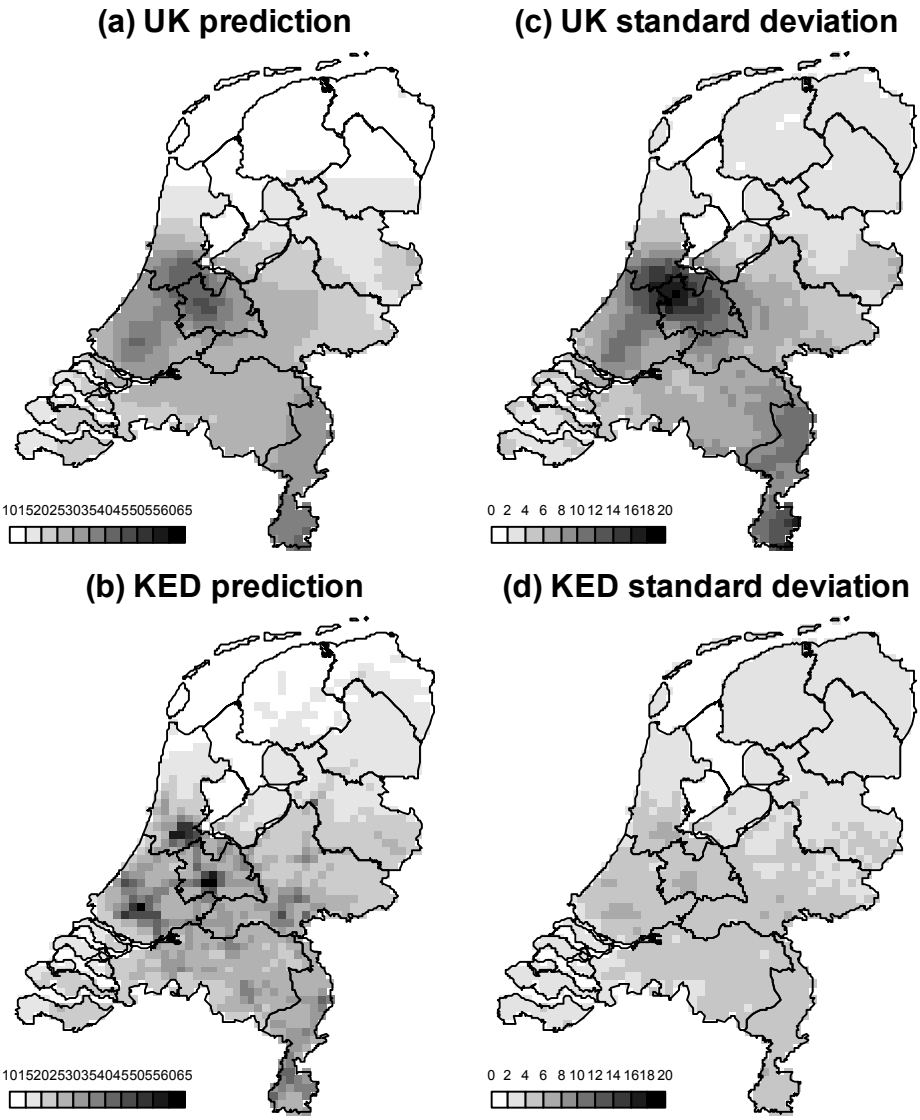

Fig. 3.5. Spatial predictions (a,c) and their standard deviations (b,d) for NO_x concentrations with UK (a,b) and KED (c,d), based on 1983 data and parameter estimation by RML.

higher than for UK 83 due to fewer observations. Re-estimating the parameters yields higher standard deviations, since parameter uncertainty has emerged.

The median for KED 87 is systematically 4% lower than the median for KED 83. Extremes for KED predictions are similar for all approaches. Apparently, the external trend determines the predicted values. The standard deviations look very similar, except for KED 87 re. The range of the KED standard deviations is almost equal to the 50% box of the UK standard deviations. The OPS output explains the spatial variation in the observations.

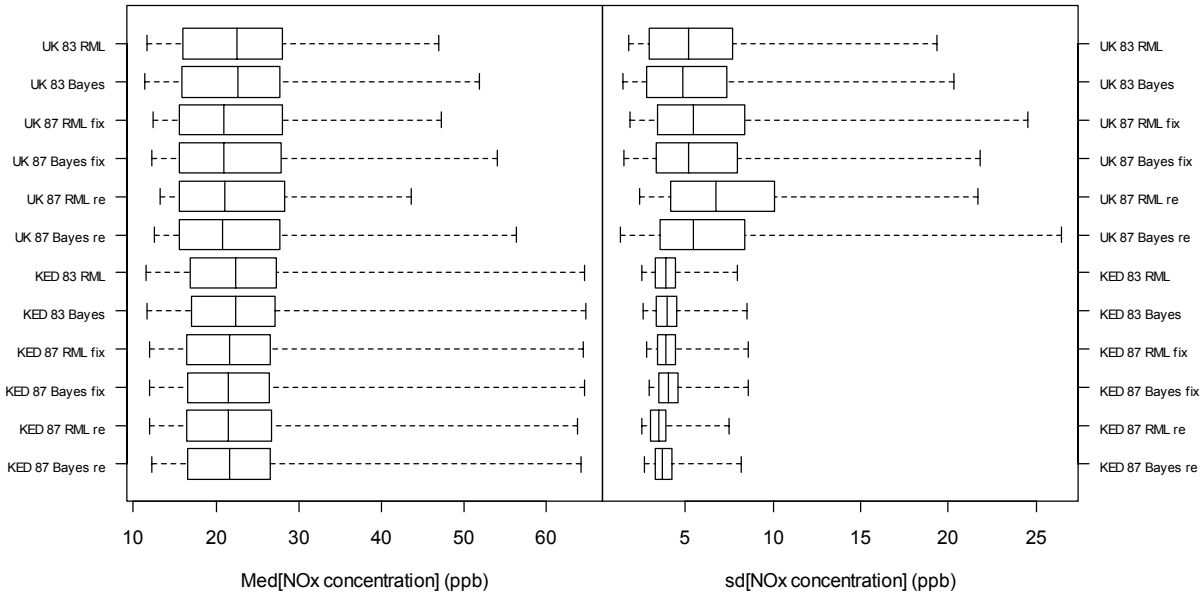

Fig. 3.6. Box plots based on the 1405 individual locations with the minimum, 25-quantile, median, 75-quantile and maximum values for the 12 predictions and standard deviations.

3.3.3 Validation

The core study results are reflected in Fig. 3.7. Validations were made for: 1) UK and KED, 2) parameter estimates by RML and Bayesian inference, and 3) keeping the 1983 parameters fixed and re-estimating the parameters for each station configuration. We called the respective validations: UK RML fix, UK Bayes fix, UK RML re, UK Bayes re, KED RML fix, KED Bayes fix, KED RML re, and KED Bayes re. The last step was to add a validation of the OPS model and we present, in succession, the validation results of OPS, UK, and KED.

OPS is positively biased, with an average of 2 ppb (Fig. 3.7a). The median is independent of the number of observations, because it represents a model validation without using observations. The ME deviance increases with fewer observations. The URMSE (Fig. 3.7b) of OPS is on average 5 ppb. The deviance, as seen in Fig. 3.7a, is almost symmetric around the median and increases as well for fewer observations. The MSSE of OPS is not available since the OPS model does not provide a standard deviation.

For UK, the median ME for 70 to 40 observations shows almost no bias as we were working here with observations only. Note the advantage of keeping parameters fixed (for RML) for sparse observations (20 to 10 observations), or the advantage of prior information over RML. This can be seen in Fig. 3.7b as well. For both UK RML and UK Bayes, the MSSE deviances for fixed parameters are equal (Fig. 3.7c). If parameters are re-estimated, the values are closer to one. This indicates

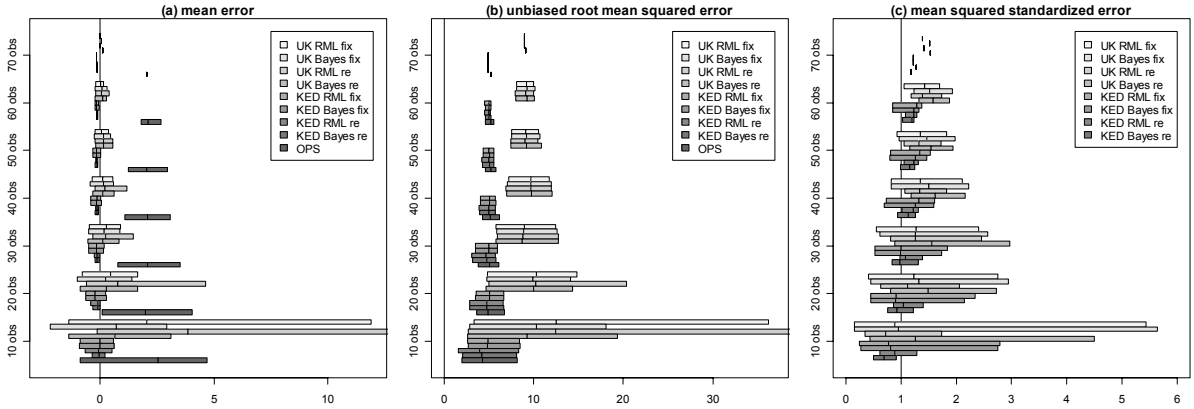

Fig. 3.7. The 95% intervals of outcomes of (a) mean errors (ppb), (b) unbiased root mean-squared errors (ppb), and (c) mean-squared standardized errors (-) for the eight spatial prediction methods and the OPS model output (ME and URMSE only), taken over random subsets of 10, 20, …, 60 points from the total test set of 70 stations.

that the absolute prediction error is almost equal to the predicted kriging standard deviation. The 97.5-quantiles that lie outside Figs. 7a and 7b are 22.9 ppb and 64.3 ppb, respectively.

KED outperforms OPS and UK since observations are combined with OPS output. The ME deviance (Fig. 3.7a) is smaller and closer to zero for KED, even for sparse observations. KED RML and KED Bayes are very similar because of the non-informative priors. Re-estimating the parameters causes a smaller deviance in the ME. For 40 or more observations, the URMSEs are systematically lower than the URMSEs of UK (Fig. 3.7b). Somewhat lower URMSE values are obtained by re-estimating the parameters. Compared to UK, the MSSE deviances (Fig. 3.7c) remain closer to one. Re-estimating the parameters causes smaller deviances than keeping them fixed.

3.4 Discussion

In our comparison of several estimation and interpolation methods, we observed KED to be superior to both UK and OPS. With reference to UK, KED emphasizes the details of OPS, as shown in Fig. 3.5. KED is less sensitive to the network configuration, resulting in smaller intervals for ME values, URMSE values and MSSE values (Fig. 3.7). With reference to OPS, KED is almost unbiased due to the use of observations, which is not the case for OPS alone. Even if OPS is biased, the combination of both the observations and OPS is superior to UK or OPS alone. One may note, however, that the KED results are somewhat optimistic since uncertainty in calculations by OPS are only treated as part of the nugget effect. For UK we used a

linear trend. We could have chosen a more complex model, like a second order trend, which probably would have resulted in lower prediction variances. Our goal, however, was to focus on parsimonious models. Also, the trend in the observations could be initially approximated with a linear trend in a reasonable way.

The OPS output explains much of the variation in the observations, resulting in a flat covariance structure of the KED model. In most cases, the effective range becomes large and in combination with a large relative nugget, the KED covariance structure is nearly constant within the horizontal scales of the Netherlands. Nevertheless, spatial prediction using the OPS model is still beneficial, as it reduces the uncertainties commonly occurring in UK models when predicting data beyond the correlation distance.

Should we use RML or Bayesian inference? In the case of many observations (>20), it does not matter. If prior information is available, Bayesian inference leads to lower prediction standard deviations. In case of sparse observations (10 to 20 observations), Bayesian inference with prior information is more robust than RML estimation.

Should we keep parameters fixed or should we re-estimate them? The lower the number of observations, the larger the parameter uncertainties become. Re-estimated parameters are then to be preferred, because specific uncertainties are then incorporated, thereby approaching reality the best. The MSSE shows that re-estimating parameters brings the values close to one. On the other hand, when parameters are kept fixed, there is no problem with sparse observations. The effect of parameter uncertainty on predictions is small relative to the effect of network configuration. The network configuration (number and location of stations) can have large effects on the accuracy of the kriging models and might need to be further optimized (Boer *et al*., 2002).

How many observations are actually needed? This will depend on what interpolation method is used. For UK RML re, this will mean 30 observations. For UK RML fix or UK Bayes, 20 observations might be needed. For KED, the results were mainly determined by the external trend. So for only 10 observations, KED still produced sensible results.

KED allows spatial prediction of a primary variable, accounting for the dense secondary variable. Collocated co-kriging could have been used as an alternative. We preferred the use of KED, requiring a less demanding variogram analysis. Furthermore, comparison studies (Pardo-Igúzquiza, 1998; Goovaerts, 2000) show KED interpolation to perform better than collocated co-kriging. KED is a form of data assimilation, but differs from Kalman filtering, for example, which results in an optimal estimation by combining observations and model output weighted by their inverse variances (Jazwinski, 1970; Meinhold and Singpurwalla, 1983). KED is a

regression-based interpolation method, where the model output is regarded as a covariate and, as such, partly explains the variation in the observations. This allows the model to have a systematic error (bias) and, more generally speaking, other external information may be included as well. It is though limited because the covariates are considered to be deterministic.

A further thing that needs to be addressed is the change-of-support problem (Wackernagel, 1995; Lajaunie and Wackernagel, 2000; Bertino and Wackernagel, 2002). We applied point kriging. We also ran the OPS model as a test on a 1 × 1 km resolution and aggregated this to a 5 × 5 km resolution. Small differences occurred in the rural areas (-1% to +1%), whereas larger differences occurred in and near urban areas (+2% to +5%) because of the presence of many point emissions.

Note that the OPS dispersion model output contains uncertainties as well. These were not considered as first approach. If all uncertainties are incorporated in the analysis, uncertain observations should be accounted for, along with uncertain dispersion model output and uncertain kriging parameters (Van de Kassteele and Stein, 2005). In this study, we only considered uncertain kriging parameters by applying Bayesian techniques. Yearly average observations are expected to carry small uncertainties. Dispersion model uncertainties, however, can be large, since the model input is uncertain, along with uncertain model parameters and uncertainty introduced by model simplification.

3.5 Conclusions

We conclude from this study that a combination of observations and a deterministic dispersion model by a model-based geostatistical interpolation procedure is successful in reducing uncertainties. The combination led to more accurate and precise spatial interpolation results, in particular, outside the sampling area. If applied as an external drift, the dispersion model output provided more detail in spatial maps than universal kriging. Standard deviations for KED are smaller than those for UK. Furthermore, KED also allows handling of biased additional information. This can be beneficial if the pollution sources are missing or unknown. KED accounts for systematic errors by use of regression parameters.

Cross-validating the different interpolation procedures was done by replacing common ME, URMSE and MSSE values by intervals obtained by repeatedly selecting different subsets from a set of test data. Reliability intervals of the interpolation results could be compared. Where small subsets were applied (10 observations), KED led to a much lower spread in mean error values than UK.

Bayesian interpolation methods have advantages over RML methods. In the case of >20 observations, this has little effect. If prior information is available however, Bayesian inference leads to lower prediction standard deviations. In case of 10-20 observations, Bayesian inference with prior information is more robust than RML estimation. Re-estimated parameters are to be preferred, because specific process changes and parameter uncertainties are then incorporated, thereby approaching reality the best.

The number of observations that is needed to make an accurate map of sufficient quality depends on the interpolation method, but also on the spatial variation and the assumption that processes described by the OPS model does not change. For UK with re-estimated parameters using RML estimation, this will be 30 observations. For UK with fixed parameters using RML estimation or Bayesian inference, 20 observations might be needed. For KED, the results were mainly determined by the external trend. So for only 10 observations, KED still produced sensible results. These numbers may apply to NO_x in this chapter, but in general certainly not.

Further use of improved prior distributions chosen and more detailed modeling needs to be further investigated in the future, especially for KED. A further optimization of the location of network stations is as yet to be carried out.

3.6 Acknowledgments

The authors would like to thank Jan de Ruiter of the Netherlands Environmental Assessment Agency – RIVM for providing the OPS model output data, and the referees for their comments and suggestions, which were all very helpful to increase the quality of this chapter.

4. Error-in-variable external drift kriging

We present a method that combines uncertain air quality measurements with uncertain secondary information from an atmospheric dispersion model. The method combines external drift kriging and a measurement error model, and uses Bayesian techniques for inference. An illustration with simulated data shows what can theoretically be expected. The method is flexible for assigning different error variances to both the primary information and secondary information at each location. Next, we address actual NO_2 data collected at an urban and a rural site in the Netherlands. Uncertainty assessments in terms of exceeding air quality standards are given. The study shows that biased uncertain secondary information can be used successfully in a spatial interpolation study at the national scale.

This chapter is based on J. van de Kassteele and A. Stein (2005). A model for external drift kriging with uncertain covariates applied to air quality measurements and dispersion model output. *Environmetrics* (in press, published online DOI 10.1002/env.771).

4.1 Introduction

Accurate and spatially highly resolved maps of NO_2 levels are essential for assessing individual human exposures to NO_2. Such maps have to be based on data from monitoring stations. In the Netherlands, 45 NO_2 monitoring stations, to which secondary information is added from a high-resolution dispersion model, are used for this purpose. Kriging with external drift (KED) has been shown to be useful in merging these two sources of information (Van de Kassteele *et al*., 2005). The primary variable, given by the monitoring data, is combined with a secondary variable, the dispersion model output, as the external drift covering the whole domain. Additionally, the secondary variable should be known at locations of the primary variable. KED has been successfully applied in air quality mapping problems as seen, for example, in Pauly and Drueke (1996), Bertino and Wackernagel (2002), Genikhovich *et al*. (2002).

Recently, more focus has been put on uncertainty assessment of these maps (Van Aalst *et al*., 1999; RIVM, 1999; Van Asselt *et al*., 2001). The air-quality mapping studies above have only considered deterministic secondary information. One may question what the effect of uncertain secondary information is on the interpolation of the primary information. The idea of including uncertain secondary information is not new, having been used, for example, in disease mapping (Bernardinelli *et al*., 1997, and Held *et al*., 2005).

This chapter describes a method to perform a detailed uncertainty assessment of mapping NO_2 concentrations in the Netherlands at rural and urban scales using uncertain secondary information from an atmospheric dispersion model. Four different sources of uncertainty were investigated. The first source are the NO_2 measurements. The second is the dispersion model output. The third source is an imperfect relationship between the observed concentrations and those modeled by the dispersion model. The fourth uncertainty is the spatial prediction procedure itself. The method presented here assimilates these uncertainties, being based on KED (Wackernagel, 2003) with features of a measurement error model (Cheng and Van Ness, 1999).

We used a simulated data experiment to illustrate the method. Next, the method was applied to real NO_2 data at two sites in the Netherlands. The final products consisted of maps with: 1) optimal spatial predictions of NO_2 rural and urban background concentrations, 2) corresponding standard deviations and 3) probabilities of exceeding the NO_2 air quality standard of 40 $\mu g/m^3$ (EC, 1999). Our investigation here was to determine how different values of uncertainty in the

observations and in the dispersion model output affected the interpolation parameters and maps and to compare and discuss the results.

4.2 Material and Methods

4.2.1 Primary information: observed NO_2 concentrations

We illustrate the use our method by data from 2002 in which 45 annual averaged NO_2 concentrations were available. Data consisted of 26 rural background and 6 city background stations, and 13 street stations (Van Elzakker, 2001). The street stations were omitted because these are specific for local street conditions and not representative of the scale considered in this study. Another 6 stations were omitted because they did not suit the standards for missing data, leaving us with 26 stations.

The observations are shown graphically in Fig. 4.1a. The size of the area is approximately 260 × 310 km. NO_2 is formed by combustion. Sources are traffic, industry, power plants, and inland waterway shipping. High concentrations are therefore found in the large cities (Rotterdam, The Hague, Amsterdam and Utrecht) and near highways. For measured annual averaged NO_2 concentrations, the uncertainty (in this case twice the standard deviation) was assumed to be known and being approximately equal to 10% of the annual averaged concentration (Van Aalst *et al.*, 1999). We also considered different percentages in section 4.3.2.

4.2.2 Secondary information: dispersion model output

The Operational Priority Substances (OPS) dispersion model provided the secondary information. The OPS model calculates annual averaged concentrations and depositions, based on emissions and meteorology. It also accounts for transport from adjacent countries. The model is an analytical model using the Gaussian plume model for dispersion on local scale. A Lagrangian trajectory model is used for long-distance transport.

Input consists of emissions from sources into the atmosphere. A detailed emission inventory for 2002 was not available, therefore emissions from 1995 were scaled proportional to known total emissions per source group for 2002. Output consists of concentration fields on a desired spatial resolution (5 × 5 km in Fig. 4.1b). The model generated data at the monitoring station locations as well. The OPS model is described in Van Jaarsveld and De Leeuw (1993) and in Van Jaarsveld (2004).

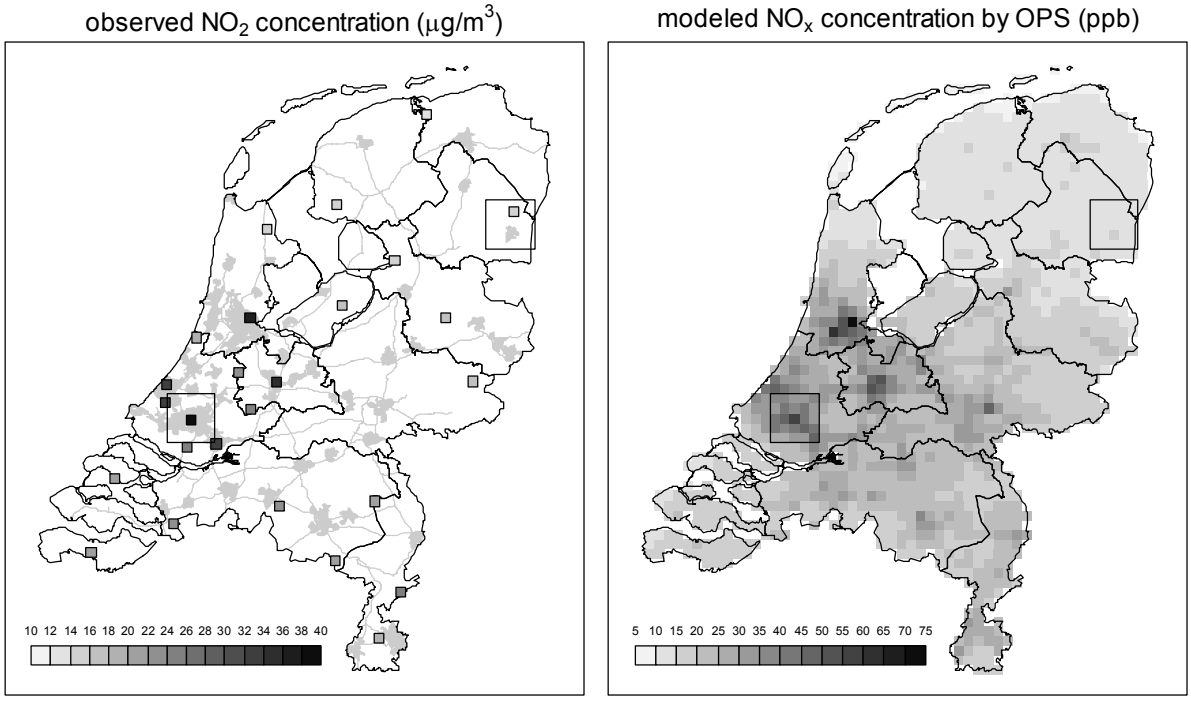

Fig. 4.1. Annual averaged NO_2 concentration ($\mu g/m^3$) in the Netherlands, monitored at 26 rural and urban background stations and NO_x concentrations (ppb) modeled by the OPS model on a 5 × 5 km grid. Regional (provincial) boundaries are indicated by black lines. The squares indicate the locations of the two study sites.

Uncertainties of the OPS model, the sum of uncertain model input, uncertain model parameters, and uncertainty caused by a simplified model structure can become large (Janssen at al., 1994; Acharya, 1994). On the basis of results from Van Aalst *et al*. (1999) and Van Jaarsveld (2004), we assumed that twice the standard deviation was equal to 20% of the predicted value, but, like the observations, we looked at different percentages (see section 4.3.2).

4.2.3 KED with error-in-variable features

The method used here is based on kriging with external drift and has the features of a measurement error model (ME model or error-in-variable model); this is the reason we call it "error-in-variable KED". According to the theory of ME models, we have two related quantities, η and ξ, that are not observable and therefore unknown. These quantities are called latent variables and can only be observed with additive errors:

$$y_i = \eta_i + \varepsilon_{y_i} \tag{4.1a}$$

$$x_i = \xi_i + \varepsilon_{x_i} \tag{4.1b}$$

where y_i are observations and x_i are covariates at location $\mathbf{s}_i$, where $i = 1 \ldots n$, where n is the number of observations. The errors are assumed to be normally distributed and independent:

$$\varepsilon_{y_i} \sim \mathrm{N}(0, \sigma^2_{y_i}) \tag{4.2a}$$

$$\varepsilon_{x_i} \sim \mathrm{N}(0, \sigma^2_{x_i}) \tag{4.2b}$$

$$\mathrm{cov}(\varepsilon_{y_i}, \varepsilon_{x_k}) = 0 \tag{4.2c}$$

The error variances $\sigma^2_{y_i}$ and $\sigma^2_{x_i}$ are assumed to be known at each location and can differ from location to location. Part of the variance in the observations can be explained by the variance of a linear function of the covariates. In the context of a ME model, this relation should be written between the latent variables (Cheng and Van Ness, 1999). The residuals represent a sum of a spatially correlated part, as in KED, and a so-called equation error part, due to an imperfect relation between $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$:

$$\eta_i = \beta_0 + \beta_1 \xi_i + \varepsilon_{s_i} + \varepsilon_{q_i} \tag{4.3}$$

In this equation, β_0 and β_1 are unknown trend parameters. We assume that the ε_{s_i} are multivariate normally distributed with variance $\sigma^2_s \mathbf{R}_y$, and that the ε_{q_i} are univariate normally distributed with common variance σ^2_q for each i

$$\boldsymbol{\varepsilon}_s \sim \mathrm{MVN}(0, \sigma^2_s \mathbf{R}_y) \tag{4.4a}$$

$$\varepsilon_{q_i} \sim \mathrm{N}(0, \sigma^2_q) \tag{4.4b}$$

Here $\mathbf{R}_y = \exp(-\mathbf{H}_y/\phi)$ is an $n \times n$ correlation matrix and $\mathbf{H}_y$ an $n \times n$ distance matrix between the monitoring station locations. Elements of $\mathbf{R}_y$ represent a exponential function of minus the distance divided by an unknown range parameter, ϕ, representing the decay rate of a spatial correlation function. Contrary to σ^2_y and σ^2_x, both variances σ^2_s and σ^2_q are unknown.

Spatial prediction is a simple extension of the data modeling. Variable of interest is $\tilde{\mathbf{y}}$. Latent variables $\tilde{\boldsymbol{\eta}}$ and $\tilde{\boldsymbol{\xi}}$ at prediction locations $\tilde{\mathbf{s}}$ can only be observed with additive errors:

$$\tilde{y}_j = \tilde{\eta}_j + \varepsilon_{\tilde{y}_j} \tag{4.5a}$$

$$\tilde{x}_j = \tilde{\xi}_j + \varepsilon_{\tilde{x}_j} \tag{4.5b}$$

The tildes indicate that we are dealing with a prediction at j = 1 … m prediction locations contained in the vector $\tilde{\mathbf{s}}$. Again, errors are assumed to be normally distributed and independent:

$$\varepsilon_{\tilde{y}_j} \sim \mathrm{N}(0, \sigma^2_{\tilde{y}_j}) \tag{4.6a}$$

$$\varepsilon_{\tilde{x}_j} \sim \mathrm{N}(0, \sigma^2_{\tilde{x}_j}) \tag{4.6b}$$

$$\mathrm{cov}(\varepsilon_{\tilde{y}_j}, \varepsilon_{\tilde{x}_k}) = 0 \tag{4.6c}$$

The error variances $\sigma^2_{\tilde{y}_j}$ and $\sigma^2_{\tilde{x}_j}$ are assumed to be known at each prediction location and can differ from location to location. Variance $\sigma^2_{\tilde{x}}$ follows directly from the data, whereas for $\sigma^2_{\tilde{y}}$ it is assumed to be the average of all $\sigma^2_{y_i}$. The relation between $\tilde{\boldsymbol{\eta}}$ and $\tilde{\boldsymbol{\xi}}$ is given by

$$\tilde{\eta}_j = \beta_0 + \beta_1 \tilde{\xi}_j + \varepsilon_{\tilde{s}_j} + \varepsilon_{q_j} \tag{4.7}$$

In equation (4.7) the variables and parameters have the same meaning as equation (4.3), except for $\varepsilon_{\tilde{s}_j}$, which are multivariate normally distributed according to

$$\boldsymbol{\varepsilon}_{\tilde{s}} \sim \mathrm{MVN}(\mathbf{R}^{\mathrm{T}}_{y,\tilde{y}} \mathbf{R}^{-1}_{y} \boldsymbol{\varepsilon}_s, \sigma^2_s (\mathbf{R}_{\tilde{y}} - \mathbf{R}^{\mathrm{T}}_{y,\tilde{y}} \mathbf{R}^{-1}_{y} \mathbf{R}_{y,\tilde{y}})) \tag{4.8}$$

In equation (4.8), $\mathbf{R}_{\tilde{y}} = \exp(-\mathbf{H}_{\tilde{y}} / \phi)$ and $\mathbf{R}_{y,\tilde{y}} = \exp(-\mathbf{H}_{y,\tilde{y}} / \phi)$ are m × m and n × m correlation matrices, and $\mathbf{H}_{\tilde{y}}$ and $\mathbf{H}_{y,\tilde{y}}$ are distance matrices between the mutual prediction locations, and between the observation locations and prediction locations, respectively. Elements of $\mathbf{R}_{\tilde{y}}$ and $\mathbf{R}_{y,\tilde{y}}$ have the same exponential function of minus the distance, both divided by the range parameter ϕ. Predictions further away from monitoring sites are more uncertain.

4.2.4 Hierarchical structure of the model

The model can be written as a Bayesian hierarchical model (Fig. 4.2). In the Bayesian framework, all parameters are treated as random variables. The hierarchical model

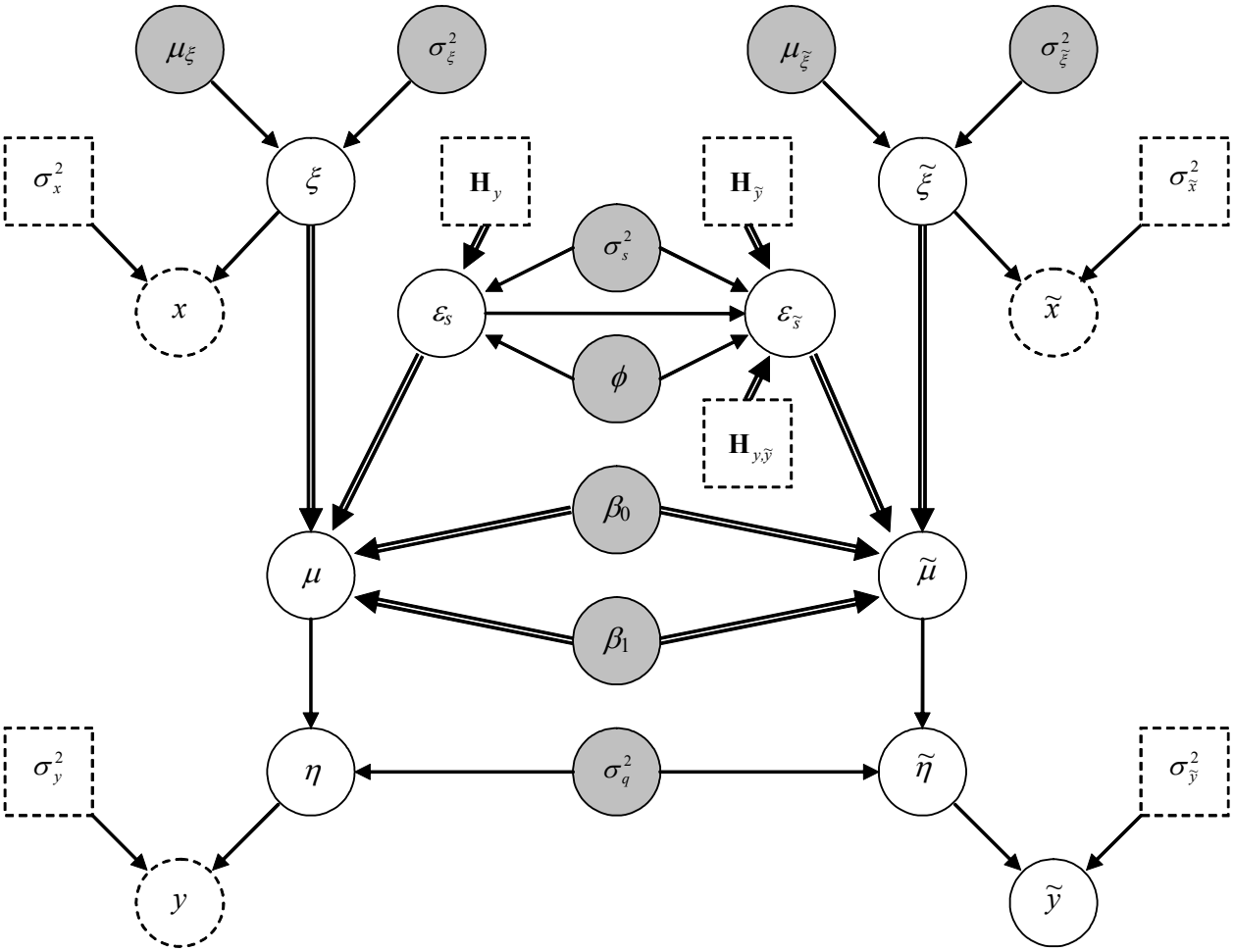

Fig. 4.2. Hierarchical structure of the error-in-variable KED model. Dashed circles represent data, dashed squares constants, solid circles parameters, gray circles parameters with a prior distribution, single arrows stochastic dependence, and double arrows deterministic dependence.

consists of a joint distribution over all unknown parameters and unobserved data, represented in the figure as solid circles. The joint posterior distribution is obtained by conditioning the observed data, represented in the figure as dashed circles. Constants, like the known variances and distance matrices are represented by dashed squares.

On the left-hand side of Fig. 4.2, the data are observations **y** and covariates **x**, both known at the monitoring station locations. According to equation (4.1a), both are realizations of a univariate normal distribution with expectation $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$, and variance σ_y^2 and σ_x^2, respectively. For convenience, we split up equation (4.3) and introduce an additional parameter $\boldsymbol{\mu}$, which is the expectation of $\boldsymbol{\eta}$ conditional on $\boldsymbol{\varepsilon}_s$. Consequently, $\boldsymbol{\eta}$ is a realization of a univariate normal distribution with expectation $\boldsymbol{\mu}$, and common variance σ_q^2. Variable $\boldsymbol{\mu}$ represents the link between $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ by the linear relation with trend parameters β_0 and β_1, plus spatial error $\boldsymbol{\varepsilon}_s$, which is a realization of multivariate normal distribution (4.4a). We assume $\boldsymbol{\xi}$ and $\tilde{\boldsymbol{\xi}}$ as being

independent identically distributed random variables with expectations μ_ξ and $\mu_{\tilde{\xi}}$ and variances σ_ξ^2 and $\sigma_{\tilde{\xi}}^2$ respectively, making it a structural measurement error model. The right-hand side of Fig. 4.2 is the prediction model, which is similar to the data model. As represented in Fig. 4.2, the model can be easily evaluated with Bayesian techniques.

Because β_0, β_1, σ_s^2, ϕ, σ_q^2, μ_ξ, $\mu_{\tilde{\xi}}$, σ_ξ^2 and $\sigma_{\tilde{\xi}}^2$ do not have parental nodes, they are given prior distributions. It is also possible to set priors on $\boldsymbol{\xi}$ and $\tilde{\boldsymbol{\xi}}$ directly, but writing it as a structural measurement error model makes the procedure more stable. We chose the following non-informative priors:

$$\beta_0 \sim \mathrm{N}(0, 1000) \tag{4.9a}$$
$$\beta_1 \sim \mathrm{N}(0, 1000) \tag{4.9b}$$
$$\sigma_s^{-2} \sim \mathrm{Gamma}(0.001, 0.001) \tag{4.9c}$$
$$\phi^{-1} \sim \mathrm{Uniform}(0.001, 1000) \tag{4.9d}$$
$$\sigma_q^{-2} \sim \mathrm{Gamma}(0.001, 0.001) \tag{4.9e}$$
$$\mu_\xi \sim \mathrm{N}(0, 1000) \tag{4.9f}$$
$$\sigma_\xi^{-2} \sim \mathrm{Gamma}(0.001, 0.001) \tag{4.9g}$$
$$\mu_{\tilde{\xi}} \sim \mathrm{N}(0, 1000) \tag{4.9h}$$
$$\sigma_{\tilde{\xi}}^2 \sim \mathrm{Gamma}(0.001, 0.001) \tag{4.9i}$$

We then applied Markov chain Monte Carlo (MCMC) integration (Gelman *et al*., 2004), using the Gibbs sampler. Typically, the chains are run for a number of iterations until the outputs are stable (burn-in) and then a large number of additional iterations are run. The stationary distributions of the chains are the posterior distributions of the nodes. We performed the Gibbs sampler using the Bayesian analysis Using Gibbs Sampler (BUGS) software (Gilks *et al*., 1994).

After sampling from the posterior predictive distribution and application of a possible back transformation (section 4.3.2) different quantities, such as the median as predictor, standard deviation as a fourth of the 95% credible interval, and a probability of exceeding the air quality standard of 40 μg/m^3, can be estimated empirically (De Oliveira *et al*., 1997). Note the difference between the known variance $\sigma_{\tilde{y}}^2$ and the prediction variance (standard deviation squared) of $\tilde{\mathbf{y}}$. The first one results only from uncertain observations $\mathbf{y}$, equations (4.2a) and (4.6a), while the second one includes all uncertainties.

4.3 Results

4.3.1 Simulated data example

We illustrate the model in Fig. 4.3 using four cases of a one-dimensional simulated data example. The circles represent 20 observations **y** at fixed randomly chosen spatial locations, while the triangles represent a covariate **x** at the same locations. We took $\beta_0 = 5$, $\beta_1 = 1$, $\sigma_s^2 = 1$, $\phi = 10$, $\sigma_q^2 = 0$, and an exponential correlation function. The goal was to estimate these parameters, make spatial predictions of the primary data (resulting in the solid line through the circles), and quantify the uncertainties (resulting in the dashed lines and 95% interval around the interpolations) based on the 20 observations. We used additional information from the covariate (solid line though the triangles) and included uncertainties in both the observations (error bars on the circles in Fig. 4.3b and Fig. 4.3d) and covariate (dashed lines around the covariate in Fig. 4.3c and Fig. 4.3d)

If **y** and **x** are measured without error ($\sigma_{y_i}^2 = 0$, $\sigma_{x_i}^2 = 0$, $\sigma_{\tilde{x}_j}^2 = 0$), the prediction is similar to kriging with external drift (Fig. 4.3a). The prediction passes through each point, following the covariate. The prediction error is small near the points and increases between two points. Although it is zero at the observation locations, it appears to be larger than zero. This is because prediction locations do not exactly coincide with observation locations.

If only **y** is measured with error ($\sigma_{y_i}^2 \in [0, 0.2]$, $\sigma_{x_i}^2 = 0$, $\sigma_{\tilde{x}_j}^2 = 0$), the prediction is smoother (i.e. towards the linear function of covariate) at locations with a larger measurement error (Fig. 4.3b). For example, this appears if we compare locations $s = 46$, $s = 83$, and $s = 96$ with Fig. 4.3a. Overall, it is somewhat difficult to distinguish from Fig. 4.3a. The contribution of errors in the observations to the total prediction error is only small, but the prediction error has increased and the spatial correlated error (Fig. 4.4) has decreased in comparison to Fig. 4.3a.

If only **x** is measured with error ($\sigma_{y_i}^2 = 0$, $\sigma_{x_i}^2 \in [0.01, 0.30]$, $\sigma_{\tilde{x}_j}^2 \in [0.01, 0.30]$), the prediction has smoothed everywhere, especially at locations where σ^2_{x0} is larger, because uncertainty is now present over the whole domain (Fig. 4.3c). Uncertain covariates have the same effect as uncertain observations. For example, compare locations $s = 46$, $s = 83$, and $s = 96$ again with Fig. 4.3b. Overall, the prediction error has increased more, while the spatial correlated error (Fig. 4.4) has decreased more in comparison to Fig. 4.3b.

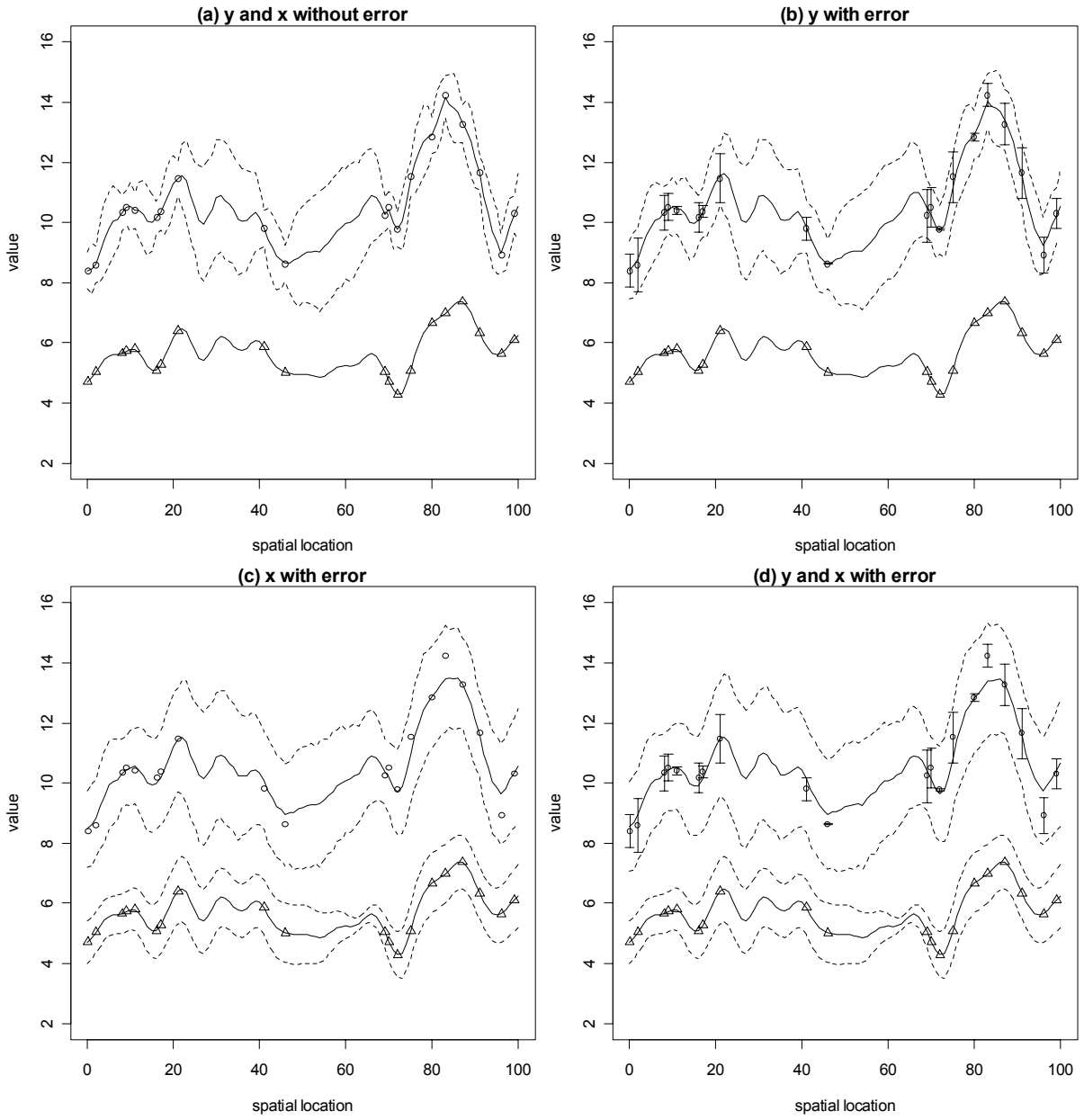

Fig. 4.3. Illustration of error-in-variable KED.

Finally, if both **y** and **x** are measured with error ($\sigma^2_{y_i} \in [0, 0.20]$, $\sigma^2_{x_i} \in [0.01, 0.30]$, $\sigma^2_{\tilde{x}_j} \in [0.01, 0.30]$), the prediction is smoother and the prediction error has further increased (Fig. 4.3d). For the same reason as in Fig. 4.3b, it is somewhat difficult to distinguish it from Fig. 4.3c.

This simulated data example did not contain an equation error. The effect of such an error is similar to that of additional measurement errors. If the error variances of the observations and covariate further increase, both the equation error and the spatial correlated error decrease, eventually becoming equal to zero. Predicted values approximate the linear function of the covariate and the prediction error increases. We can assign different error variances to each location, both to the

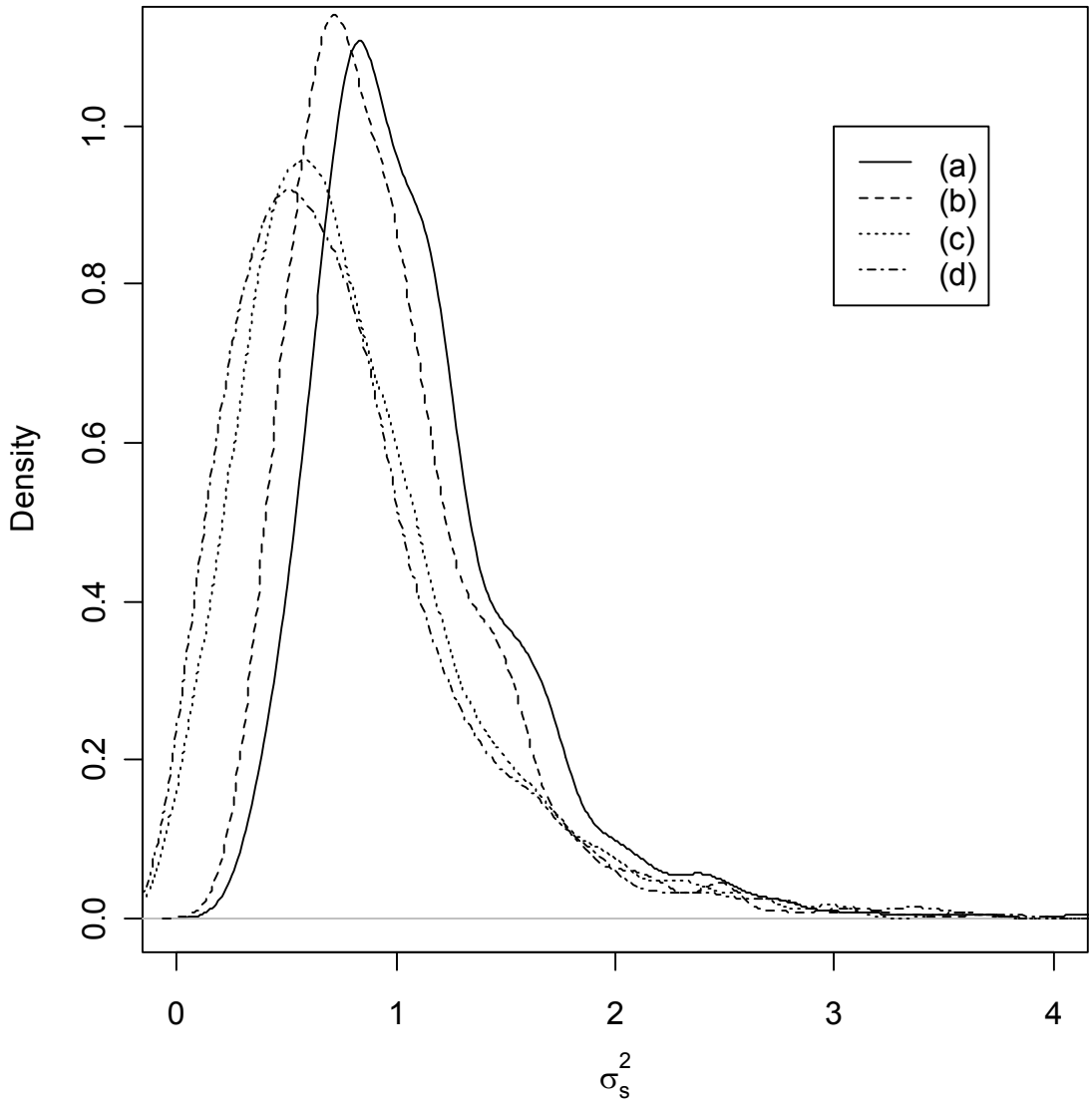

Fig. 4.4. Posterior densities of σ_s^2 for the four cases of the simulated data experiment in Fig. 4.3.

monitoring stations and the OPS output. This makes the error-in-variable KED model flexible for application. Uncertain observations (or outliers) may have a large variance and thus be confronted with less weight during parameter estimation and spatial prediction. This also applies to uncertain covariates at observation locations.

4.3.2 NO_2 concentrations in the Netherlands

We considered two sites of 25 km² at a 1 × 1 km spatial resolution, each in the Netherlands (see Fig. 4.1). The first site was the urban area of Rotterdam, with buildings, petrochemical industrial plants, highways, rivers, and harbor activities. The second site was a rural area near the monitoring station, Valthermond, located in

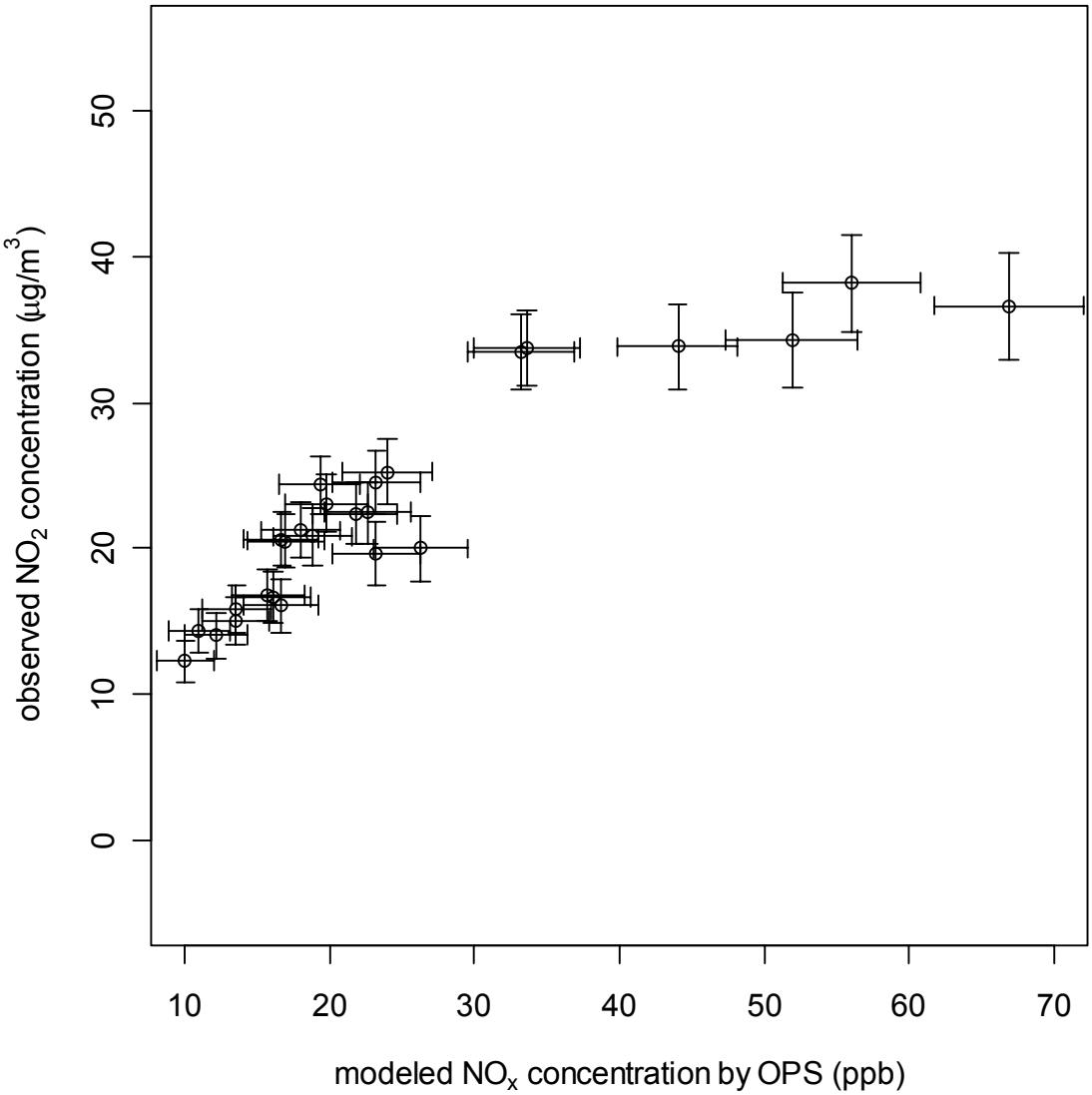

Fig. 4.5. Observed NO_2 concentrations vs. NO_x concentrations modeled by OPS. Vertical and horizontal bars between -2σ and $+2\sigma$ indicate twice the standard deviation (10% in the observations and 20% in OPS output, respectively).

the northeast of the Netherlands. This area, located in the former Dutch peat area, is used for agriculture, nowadays mainly consisting of potato fields. A small city nearby is Emmen, 10 km south of Valthermond. We expected the urban area of Rotterdam to show more spatial variation and higher uncertainties than the agricultural area around Valthermond. These sites were chosen to investigate differences between these two types of land use.

Fig. 4.5 shows the relation at the station locations between observations of annual averaged NO_2 ($\mu g/m^3$) and OPS output of annual averaged NO_x concentration (ppb). The errors increased with increasing concentrations. After a log-log-transformation, the error structure became additive and variances stabilized.

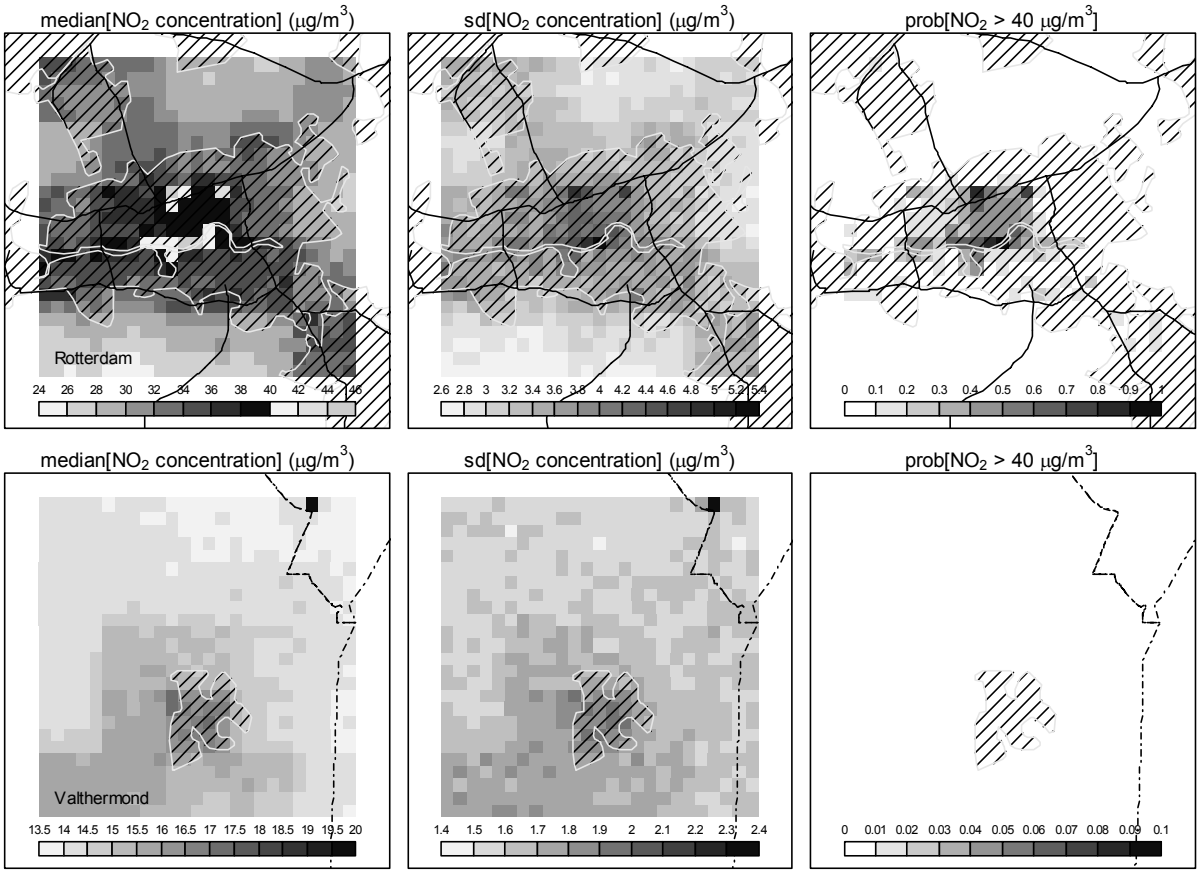

Fig. 4.6. Median for the posterior distribution of the annual NO_2 concentration, the standard deviation, and the probability of exceeding the 40 $\mu g/m^3$ air quality standard around Rotterdam and Valthermond for $2\sigma_y$ = 10% and $2\sigma_x$ = 20%. The shaded areas indicate cities and/or industrial zones, the solid black lines highways, and the dot-dashed lines to country and provincial boundaries.

Each sample is back-transformed after sampling from the posterior predictive distribution.

An earlier study using a comparable data set of NO_x concentrations (Van de Kassteele *et al.*, 2005) showed that a large part of the variation in the observations could be explained by the OPS output. This resulted in a high value for the range parameter (~ 1000 km), leading to an almost linear shape of the exponential correlation function on the Dutch scale. We kept the range parameter fixed at 1000 km, because its posterior distribution showed a high skewness and a very long tail, and therefore not a clear mode.

Fig. 4.6 consists of maps showing the median of the posterior distribution of NO_2 concentration, the standard deviation, and the probability of exceeding the air quality standard of 40 $\mu g/m^3$. These maps are constructed with values of $2\sigma_y$ = 10% and $2\sigma_x$ = 20% of the annual averaged concentration and apply to background concentrations. Concentrations are shown to become higher when they are close to

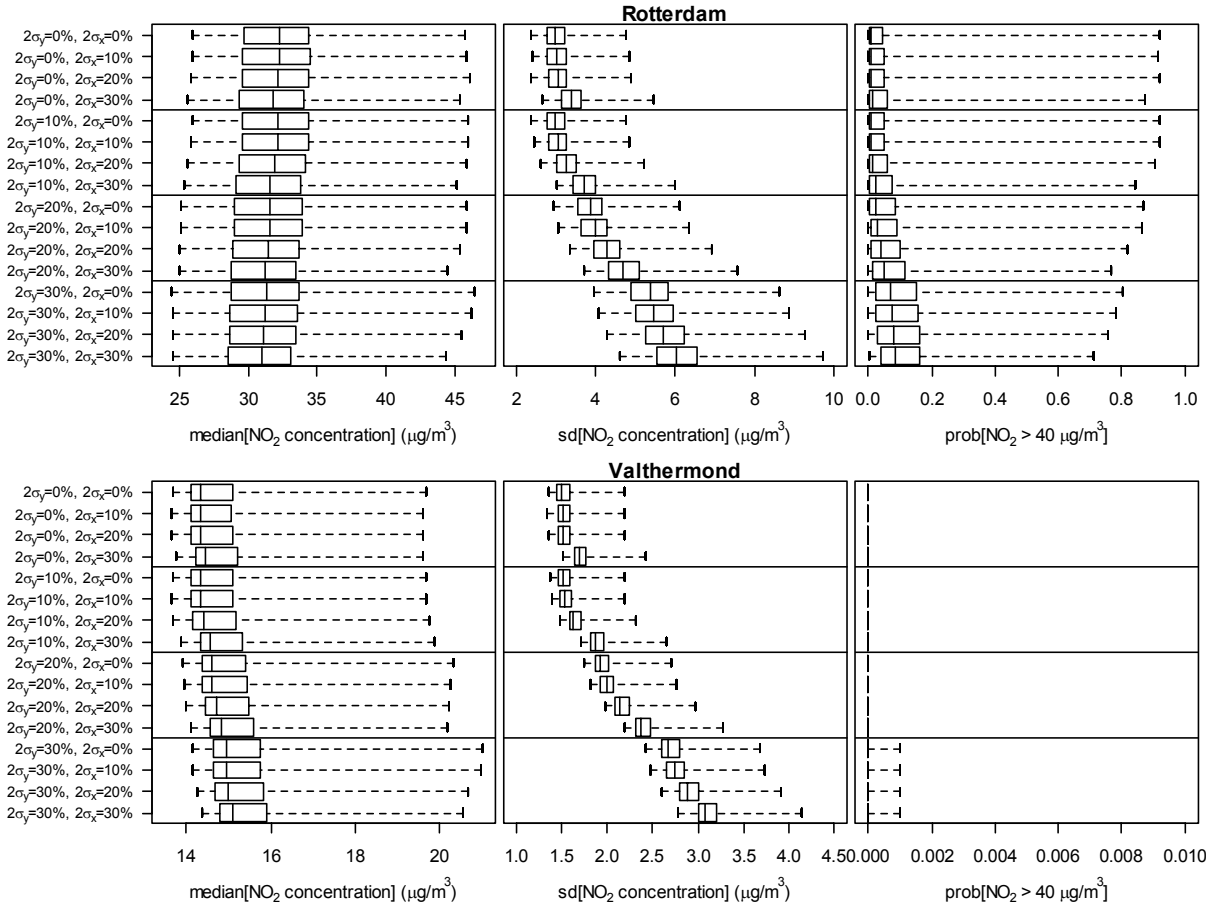

Fig. 4.7. Boxplots of the median, standard deviation, and probability of exceeding 40 μg/m³, based on 625 individual locations for different values for $2\sigma_y$ and $2\sigma_x$. The box represents the inter-quartile range, and the whiskers the minimum and maximum values.

highways. The highest concentrations occur in the center of Rotterdam, close to the River Maas, and near the highway north of Rotterdam. The air quality standard at these locations was exceeded. Standard deviation increased with increasing concentrations because of the log-transform, resulting in a higher probability of exceeding 40 μg/m³. However, it did remain under 1. The highest concentrations around Valthermond occur in the city of Emmen. A potato flour factory was responsible for the single peak in the northeast. The probability of exceeding 40 μg/m³ is zero everywhere.

Next, we studied different percentages, 2σ = 0%, 10%, 20%, and 30%, of the annual averaged concentrations. The reason for this was unknown actual uncertainties in observations and OPS output. We performed error-in-variable KED for each combination, resulting in 16 possible combinations. Fig. 4.7 shows boxplots of the median and standard deviation, and the probability of exceeding 40 μg/m³ at the 2 × 625 prediction locations. The median concentration decreases around Rotterdam and increases evenly over space for increasing values of 2σ around

Table 4.1. Change in the probability of exceeding 40 $\mu g/m^3$ in case 2σ increases for different values of the median, where P_1 represents the probability for a low 2σ and P_2 the probability for a high 2σ.

	case 1: 2σ low	case 2: 2σ high	case 2 – case 1
median < 40	$P_1 < 0.5$	$P_1 < P_2 < 0.5$	$P_2 - P_1 > 0$
median = 40	$P_1 = 0.5$	$P_1 = P_2 = 0.5$	$P_2 - P_1 = 0$
median > 40	$P_2 > 0.5$	$P_1 > P_2 > 0.5$	$P_2 - P_1 < 0$

Valthermond. Changes are small, however. Similar to the simulated data example, the concentrations are smoothed towards a linear function of the OPS output.

The standard deviation increases rapidly with increasing values for 2σ. The increase is highest at locations with a high standard deviation. The increase in the prediction standard deviation is also higher if $2\sigma_y$ increases from 0% to 30% for a fixed value of $2\sigma_x$, than if $2\sigma_x$ increases from 0% to 30% for a fixed value of $2\sigma_y$. Therefore, with respect to prediction uncertainty, having more accurate observations can be preferred to having more accurate OPS output.

If 2σ increases, the probability of exceeding 40 $\mu g/m^3$ changes, as seen in Table 4.1. The value of the median determines what happens if 2σ becomes higher. In general, the probability for high 2σ approaches 0.5. This also explains the narrowing of the boxplots for Rotterdam for the probability of exceeding 40 $\mu g/m^3$. If 2σ increases, the probability evens out over space to become 0.5 if it goes to infinity. Around Valthermond, the probability remains zero everywhere, except near the potato flour factory.

4.4 Discussion

We have presented a method to interpolate uncertain spatial air quality observations using uncertain secondary information from an atmospheric dispersion model. This method was based on KED using error-in-variable features. Error-in-variable KED allows performance of a detailed uncertainty assessment of air quality mapping on rural and urban scales. It takes into account uncertainty in the observations, dispersion model output, the imperfect relation between both, and uncertainty due to spatial prediction. In all, a more flexible method is developed as compared to earlier KED-based methods.

Inference was done using a Bayesian approach. It has as an advantage that there was no need to write the model in a complicated analytical formulation, which

is the case if uncertain covariates are introduced. Furthermore, the Bayesian approach automatically takes into account parameter uncertainty.

The simulated data example showed that uncertainty in the covariates had a similar effect on the predictions as uncertainty in the observations. In fact, the prediction tended towards a linear function of the covariate, whereas prediction errors increased. This was also the case if the equation error differed from zero. A contradiction may be that if the observations are without error and the covariates are with error smoothing emergences. An explanation is that uncertain covariates lead to uncertain predictions, with subsequent smoothing.

For the annual averaged NO_2 concentrations we assumed the uncertainty ($2\sigma_y$) to be approximately 10% of the averaged concentration. Observation uncertainty should be taken into account into the prediction uncertainty. If this is ignored, inference is being be made on $\tilde{\boldsymbol{\eta}}$ instead of on $\tilde{\mathbf{y}}$, then prediction uncertainty actually decreases as σ_y^2 increases. However, no information about $\sigma_{\tilde{y}}^2$ was available at the prediction locations. As a solution, the average of σ_y^2 has been used.

A detailed quantification of OPS output uncertainty was outside the scope of this chapter. We did not look at model input, parameters, and uncertainty caused by a simplified model structure. We assumed it to be 20% of the annual averaged concentration. Furthermore, we assumed the error in the OPS output to be spatially independent. As an extension, we could have replaced $\sigma_{\tilde{x}}^2$ by a covariance matrix and with its elements depending on distance. To avoid further complications, we restricted ourselves to independence.

We studied different percentages of the annual averaged NO_2 concentrations, because actual uncertainties were unknown. As concerns prediction uncertainty, it appeared that accuracy of observations was more important than accuracy of OPS output. As uncertainty in the observations is already low (10%), however, less improvement occurs than might be expected. On the other hand, as OPS output may be much more uncertain (up to 30%), it may be preferred to improve upon this.

Fuentes and Raftery (2005) modeled both the observations and model output as a function of a hidden truth. Our interest was to assess the primary variable. Therefore, instead of considering OPS output as a function of a truth, we turned it around and considered the observations and the OPS output as being related in an orthogonal regression. This is an improvement over the traditional KED. The intention of this study has been to provide a method to combine different sources of uncertain information, and not so much a statistical evaluation of the performance of the OPS model as such. As in Fuentes and Raftery (2005), the OPS output is corrected for bias. Additive bias is included in the intercept parameter β_0, and a multiplicative bias in the slope parameter β_1. We considered the change of support problem as

neglectable on our scale. In a future study, change of support issues can be addressed as an extension of the current study.

Further extensions and applications of the model may be the inclusion of more covariates. Also, it could possibly be successfully used to assess uncertainties in long-term scenario studies for spatial decision support. The OPS model also addresses other pollutants and can be applied to different countries and regions.

4.5 Conclusions

This study showed a successful creation of concentration maps based on uncertain measurements and uncertain dispersion model output. Error-in-variable KED proved sufficiently flexible for assigning different error variances to each location, both to the monitoring stations and the OPS output, and handling biased secondary information. This can be beneficial if the pollution sources are missing or unknown. Examples are the amount of sea salt or wind blown dust in the atmosphere, e.g. for PM10. Our method accounts for the systematic error by use of regression parameters.

The Bayesian approach for spatial modeling was extremely useful in this context, because it had the advantage that there was no need to write the model in a complicated analytical formulation. Furthermore, the Bayesian approach automatically took into account parameter uncertainty.

Different uncertainties in the measurements and dispersion models affect the eventual results. Use of error-in-variable KED approach showed that high accuracy in the measurements is to be preferred above high accuracy in the OPS dispersion model output. Since the measurement precision was already high (2σ = 10%), it is more beneficial to increase precision of the OPS output. Evaluating OPS output uncertainty, however, requires additional research.

4.6 Acknowledgments

The authors would like to thank Bas Engel of Biometris and Arnold Dekkers and Guus Velders of the Netherlands Environmental Assessment Agency – RIVM for their useful comments and discussions.

5. Uncertainty assessment of future local NO_2 concentrations

Local NO_2 concentrations near Rotterdam (Netherlands) were assessed for the year 2010, focusing on the uncertainties and the changes in exceedance of European air quality standards. In the first step of the 2-step assessment method, the background contribution was determined by error-in-variable external drift kriging, where measurements and dispersion model output in the 1987–2003 period were combined. The result was subsequently extrapolated using dispersion model output and an emission scenario for 2010. In the second step, the local traffic contribution was added on the basis of a local generic dispersion model with use of an emission scenario for 2010. This resulted in maps showing local NO_2 concentrations, upper and lower limits, and probabilities of exceeding the 40 $\mu g/m^3$ air quality standard. The probabilistic measures were calculated in numbers and translated into words for easier communication. Using this method and scenario we found that within about 100 m from the highways near Rotterdam the mean NO_2 concentrations are likely to exceed the standard in 2010. The chance of exceeding the standard is unlikely up to 1 km from the highways, where the mean is expected to be below the standard in 2010.

This chapter is based on J. van de Kassteele and G.J.M. Velders (2006). Uncertainty assessment of local NO_2 concentrations derived from error-in-variable external drift kriging and its relationship to the 2010 air quality standard. *Atmospheric Environment* (in press).

5.1 Introduction

Policy makers require detailed local information about future air quality to be able to decide on infrastructural projects near residential areas. These areas may be exposed to too high concentrations of air pollutants or plans for new infrastructural projects may be obstructed if European air quality standards (EC, 1999) are exceeded at a certain location (Van Velze *et al*., 2000; Folkert *et al*., 2002). This is the reason we need maps showing concentrations of pollutants. Policy makers use results from deterministic or statistical models to assess if future air quality standards will be met. Such models inevitably have assumptions and simplifications which result in uncertainties in the results. Since these models are, in general, used quite precisely (for example, to indicate if the air quality standard is exceeded or not), the inclusion of a measure of uncertainty in such model outputs yields valuable additional information (ApSimon *et al*., 2002). Since policy makers tend to focus more and more on uncertainties as well, the question is then, how certain are future concentration maps and how should uncertainties be handled? For example, one might question what the probability of exceedance of an air quality standard is and how this should be communicated (Van Asselt *et al*., 2001).

Usually, local air quality is determined in two steps: 1) via a regional background contribution and 2) via a local contribution, which usually are line or point sources. On a regional scale, the background concentrations may be based on a combination of air quality measurements and regional dispersion model output by kriging with external drift (KED). This method is already successfully applied in air quality mapping studies, such as in Genikhovich (2002), Wackernagel (2003) and Van de Kassteele and Stein (2005). Measurements are only available at a limited number of locations, while dispersion model output over the whole region is known. We can use this information to improve interpolation of the measurement data. Adding errors to the measurements and dispersion model output makes a detailed uncertainty assessment possible (Van de Kassteele and Stein, 2005). On a local scale, measurements are usually not available, so the local contribution can be determined by local dispersion model output only. We have only considered traffic, i.e. line sources, for contributing to the local concentrations.

Currently, concentration maps used for assessing future air quality are based on information on future emissions and long-term meteorology only; see, for example, Jonson *et al*. (2001); MNP (2005). However, measurements from the past can be used too. In this study, we first explored the possibility of predicting future NO_2 background concentrations based on a future emission scenario and information from the past. The process is schematically shown in Fig. 5.1. Since dispersion model

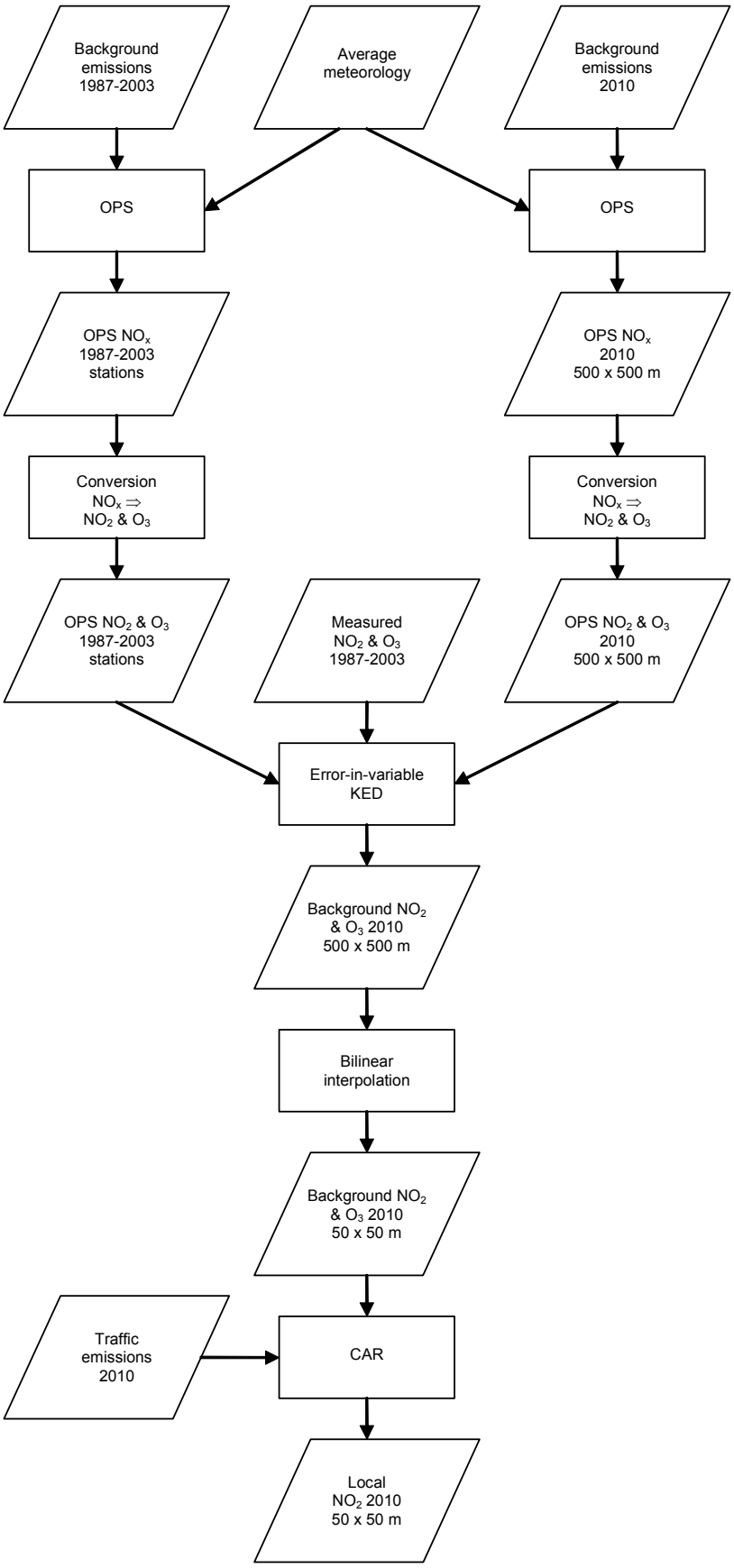

Fig. 5.1. Flow chart of the procedure to determine local NO_2 concentrations for 2010, based on a combination of measurements, dispersion model output and the local traffic contribution.

output is determined through emissions and meteorological aspects, these input data are available for both recent and past years. Future emissions can be estimated from current emissions, and technological and policy developments. Average meteorological data (climatology) can be applied to get around the lack of future meteorological factors. We applied error-in-variable KED to combine measured background concentrations and regional dispersion model output. After resampling the concentrations to fit a local scale, the next step was to add the traffic contribution by means of local dispersion model output. Uncertainties and the relationship between with the air quality standard for NO_2 (40 $\mu g/m^3$, EC, 1999) were being assessed. Finally, we will discuss communication of the results to policy makers. Since uncertainties are currently not in use in the policy process, a translation of scientific results into policy would be desirable.

5.2 Available data

5.2.1 NO_2 measurements

We used rural and urban background measurements from the Dutch national air quality monitoring network (Van Elzakker, 2001). Measurements from 1978 and beyond are available; however, the number of stations in the monitoring network diminished in the mid-1980s, so we only took data for the 1987-2003 period. After adjusting the standards for missing data for annual average concentrations, 45 stations remained, of which 38 were characterized as rural stations and 7 as city stations. Stations in street canyons were not used here. Concentrations were measured by chemiluminescence, which is the same for all stations in the Netherlands.

For measured annual average NO_2 concentrations, the uncertainty (in this case twice the standard deviation) could be found in the literature; this was approximately 10% of the annual average concentration (Blank, 2001; Van Aalst *et al.*, 1999).

5.2.2 Atmospheric dispersion model output: OPS and CAR

The OPS dispersion model (Operational Priority Substances) is used to provide information for determining NO_2 background concentrations (Van Jaarsveld and De Leeuw, 1993; Van Jaarsveld, 2004) in the Netherlands. OPS calculates annual average

concentrations based on emissions, dispersion, transport, conversion and deposition, and takes transport from adjacent countries into account. Substances dealt with show behavior that can be described by first-order chemical reactions. The model uses the Gaussian plume model for dispersion on local scale and a Lagrangian trajectory model for long-distance transport.

Input consisted of emissions from sources into the atmosphere. Source properties, like emission height, are determinative for the dispersion. A detailed emission inventory was not available for the 1987–2003 period, so known emissions from 1995 were scaled proportionally to known total emissions per source group for each year. Output consisted of concentration fields on a 500 × 500 m spatial resolution. This resolution was chosen as a compromise between the usual resolution of OPS (> 1 km) and the required resolution for the local traffic contribution. A higher resolution did not add new information. The model also generated data at the monitoring station locations for error-in-variable KED (section 5.3.2).

Quantification of OPS output uncertainty was difficult, since emissions and their effect on the output were not well known. Many model parameters were empirically obtained values. The Gaussian plume model also has some shortcomings, being a simple conceptualization of reality. On the basis of Van Aalst *et al*. (1999) and Van Jaarsveld (2004), we assumed that twice the standard deviation was approximately 20% of the predicted value. Van de Kassteele and Stein (2005) discussed the effect of different values of uncertainty and concluded that it would be more beneficial to increase the accuracy of the OPS output than go for measurement uncertainty. However, evaluating OPS output uncertainty will require additional research.

Uncertainties such as historical emissions, future economic, social and technological developments, current negotiations and modeling instruments were also taken into account in the reference projection for 2010 (Van Dril and Elzenga, 2005). This led to an additional 15% uncertainty for 2010 (Gijsen and Seebregts, 2005).

The CAR model (Calculation of Air Pollution from Road Traffic) is a generic model for determining air quality near roads in cities (Eerens *et al*., 1993). The traffic contribution is calculated by multiplying the traffic emissions by a dilution factor. The traffic emission depends on the traffic properties (composition, intensity and speed). The dilution factor depends on street properties (buildings, trees and distance to road). As in the OPS model, CAR output consists of NO_x concentrations, but on a local scale. CAR uses an empirical relationship to convert the NO_x into NO_2 through O_3 background concentrations. Uncertainty of the CAR output was estimated through Monte Carlo simulations (section 5.3.3).

OPS output consists of NO_x concentrations (in ppb). Since we are interested in NO_2 concentrations (in $\mu g/m^3$) and the fact that the CAR model requires O_3 as input

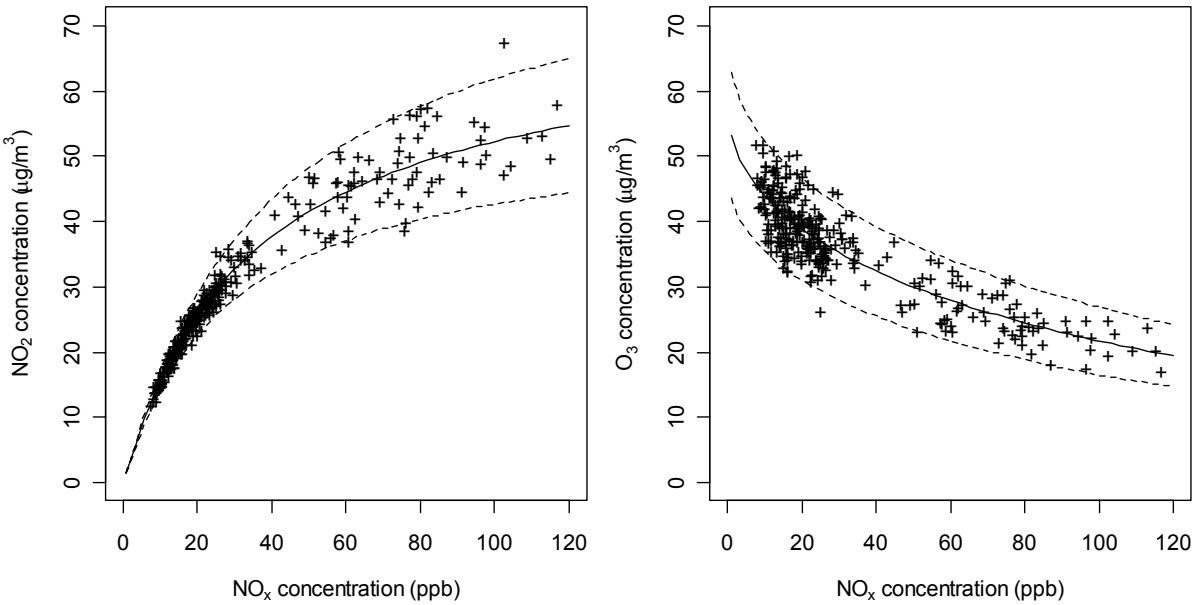

Fig. 5.2. Relationship between measured NO_2 ($\mu g/m^3$) and measured NO_x (ppb) (left), and the relationship between measured O_3 ($\mu g/m^3$) and the same measured NO_x (ppb) (right).

variable, the NO_x were first converted to NO_2 and O_3. Usually this requires complicated chemical reactions (Seinfeld and Pandis, 1998), but we used two empirical relationships (Fig. 5.2) based on measurements at rural and urban background monitoring stations for 1990-1999. Cleary, two clusters can be identified; however, differences between these two types of stations are captured well by the empirical relationships. The dashed lines in Fig. 5.2 represent 95% prediction intervals. The relationships for the predicted value (E) and prediction variance (Var) are given by:

$$E[NO_2] = (0.061 + 0.728\ NO_x^{-0.829})^{-1.544} \quad (5.1a)$$

$$Var[NO_2] = 9.34 \cdot 10^{-5}\ E(NO_2)^{3.135} \quad (5.1b)$$

$$E[O_3] = (3.419 - 0.085\ NO_x^{0.508})^{3.301} \quad (5.1c)$$

$$Var[O_3] = 0.104\ E(O_3)^{1.358} \quad (5.1d)$$

The uncertainty introduced by (5.1b) and (5.1d) are taken into account in further calculations.

Fig. 5.3. Location of the study area in the Netherlands; the dashed line shows where a cross section is made.

5.3 Methodology

5.3.1 A case study area near Rotterdam

We considered an area of 5 × 5 km on the north side of Rotterdam (Netherlands). Its location is shown in Fig. 5.3. The area is traversed by two highways: one going from east to west (the A20) and the other from north to south (the A13). North of the A20, the urban area is scattered, while in the south, towards the city centre, it is more compact. Highways A13 and A20 cause severe air pollution in this area (Van Velze *et al.*, 2000). We did not consider local roads because their contribution is not very high, mainly because of their much smaller traffic intensities compared to highways.

5.3.2 Background concentration

It was not appropriate to run OPS with the emission reference projection for 2010 (Van Dril and Elzenga, 2005) and use average meteorology as a forcing to obtain NO_2 background concentrations for 2010. This is because OPS model output is biased (Van de Kassteele and Stein, 2005). To correct the biased modeled NO_2 concentrations for 2010, we related modeled and measured NO_2 concentrations from 1987-2003 and extended the measured NO_2 time series to 2010, using the modeled NO_2 time series as an explanatory variable. Linear regression is then a simple option, but requires independent residuals in time. The regression parameters may also change in time.

We used the emissions known for the years 1987-2003 and the same average meteorological data. This resulted in modeled concentrations that depend only on changing emissions, slowly decreasing in time. If we assume meteorological factors over the years that are independent and fluctuations in the measured NO_2 concentrations that are driven mainly by actual meteorological factors, the measured concentrations (after detrending with the modeled concentrations) will be independent in time. For stations with at least 15 years of data, we checked the assumption using the test for randomness, i.e. no time dependence in the data. There is no time dependence if there are autocorrelations between the confidence bands $\pm z_{1-\alpha/2}/\sqrt{n}$,where n is the sample size, z the percent point function of the standard normal distribution and α = 0.05 the significance level. After detrending with the modeled NO_2 concentrations, all series showed randomness, so this assumption is correct. This leaves us with only a possible spatial correlation in the residuals. Similar results were obtained for O_3.

To check if the regression parameters changed with time, we estimated the parameters for each year using all stations (ignoring any spatial correlation and measurement errors). This showed that the regression parameters changed linearly with time. For NO_2, the intercept increased from -1.5 to -0.5 and the slope decreased from 1.4 to 1.1. In other words, OPS performed better for NO_2 in recent years. For O_3, the intercept increased from 0.0 to 1.0, and the slope decreased from 1.0 to 0.7. In other words, OPS performed poorer for O_3 in recent years. The associated 95% confidence intervals of the estimated parameters for 1987 and 2003 did not overlap, so we concluded the change in time to be significant. Residual variance remained constant over the years.

To model spatial dependence we applied error-in-variable KED (Van de Kassteele and Stein, 2005), a method interpolating uncertain spatial observations using uncertain secondary information. Error-in-variable KED allows a detailed uncertainty assessment of mapping rural and urban background concentrations. It takes into account uncertainty in the observations, dispersion model output, the imperfect relationship between the two and uncertainty due to spatial prediction.

Error-in-variable KED consists of two steps: a parameter estimation step and a spatial prediction step. In the estimation step we have n_s monitoring stations within n_t years. Here, we applied a log transformation to the concentrations to stabilize the error variances. We denoted measurements and OPS output as **y** and **x,** respectively. These quantities represent observations of the underlying unknown latent variables, η and ξ, which can only be observed with additive errors:

$$y_{i,j} = \eta_{i,j} + \varepsilon_{y_{i,j}} \tag{5.2a}$$

$$x_{i,j} = \xi_{i,j} + \varepsilon_{x_{i,j}} \tag{5.2b}$$

The errors are assumed to be normally distributed and independent:

$$\varepsilon_{y_{i,j}} \sim \mathrm{N}(0, \sigma^2_{y_{i,j}}) \tag{5.3a}$$

$$\varepsilon_{x_{i,j}} \sim \mathrm{N}(0, \sigma^2_{x_{i,j}}) \tag{5.3b}$$

The indices $i = 1 \ldots n_s$ and $j = 1 \ldots n_t$ represent locations in space and moments in time. The error variances $\sigma^2_{y_{i,j}}$ and $\sigma^2_{x_{i,j}}$ are known at each location and time, and can differ for each location and time. In the context of a measurement error model (Cheng and Van Ness, 1999), the latent variables are related. In error-in-variable KED, the residuals represent the sum of a spatially correlated error, as in classical KED, and a so-called equation error:

$$\eta_{i,j} = \beta_0 + \beta_1 \xi_{i,j} + \beta_2 (t_j - 1987) + \varepsilon_{s_i} + \varepsilon_{q_{i,j}} \tag{5.4}$$

To account for the linearly changing regression coefficients in time, the original error-in-variable KED model was extended with an extra term including year t_j. In the above equation, β_0, β_1, and β_2 are unknown trend parameters. The spatially correlated error is the same for each year but changes over space, meaning that each year is a realization of the same random field. Stationarity of the data was assumed because of the large-scale background concentrations for the relatively small domain of the Netherlands. We further assumed ε_{s_i} to be multivariate and normally distributed, with variance $\sigma^2_s \mathbf{R}_y$ for each i, and $\varepsilon_{q_{i,j}}$ as univariate normally distributed, with a common variance σ^2_q for each of i and j:

$$\boldsymbol{\varepsilon}_s \sim \mathrm{MVN}(0, \sigma^2_s \mathbf{R}_y) \tag{5.5a}$$

$$\varepsilon_{q_{i,j}} \sim \mathrm{N}(0, \sigma^2_q) \tag{5.5b}$$

Here $\mathbf{R}_y = \exp(-\mathbf{H}_y/\phi)$ is an $n_s \times n_s$ correlation matrix and $\mathbf{H}_y$ an $n_s \times n_s$ distance matrix between the monitoring station locations. Elements of $\mathbf{R}_y$ represent a exponential function of minus the distance divided by an unknown range parameter, ϕ. This represents the decay rate of a spatial correlation function. Contrary to σ^2_y and σ^2_x, both variances σ^2_s and σ^2_q are unknown.

In the prediction step there are $i = 1 \ldots m_s$ prediction locations and $j = 1 \ldots m_t$ prediction times. In our study, $m_t = 1$ and $\tilde{t}$ is the year 2010. Spatial prediction can be phrased as a simple extension of the estimation step. The eventual variable of interest is $\tilde{\mathbf{y}}$. Latent variables $\tilde{\boldsymbol{\eta}}$ and $\tilde{\boldsymbol{\xi}}$ at prediction locations $\tilde{\mathbf{s}}$, can only be observed with additive errors. The tildes indicate that we are dealing with predictions. Again, errors are assumed to be normally distributed and independent. The error variances, $\sigma^2_{\tilde{y}_{i,j}}$ and $\sigma^2_{\tilde{x}_{i,j}}$, are assumed to be known at each prediction location and can differ from location to location. Variances $\sigma^2_{\tilde{x}_{i,j}}$ follow directly from the data, whereas $\sigma^2_{\tilde{y}_{i,j}}$ is assumed to be the average of all $\sigma^2_{y_{i,j}}$. The relationship between $\tilde{\boldsymbol{\eta}}$ and $\tilde{\boldsymbol{\xi}}$ is given by

$$\tilde{\eta}_{i,j} = \beta_0 + \beta_1 \tilde{\xi}_{i,j} + \beta_2 (\tilde{t}_j - 1987) + \varepsilon_{\tilde{s}_i} + \varepsilon_{q_{i,j}} \tag{5.6}$$

In equation (5.6) the variables and parameters have the same meaning as in equation (5.4), except for $\varepsilon_{\tilde{s}_i}$, which are multivariate normally distributed according to

$$\boldsymbol{\varepsilon}_{\tilde{s}_i} \sim \text{MVN}(\mathbf{R}^{\text{T}}_{y,\tilde{y}} \mathbf{R}^{-1}_{y} \boldsymbol{\varepsilon}_s, \sigma^2_s (\mathbf{R}_{\tilde{y}} - \mathbf{R}^{\text{T}}_{y,\tilde{y}} \mathbf{R}^{-1}_{y} \mathbf{R}_{y,\tilde{y}})) \tag{5.7}$$

In equation (5.7), $\mathbf{R}_{\tilde{y}} = \exp(-\mathbf{H}_{\tilde{y}} / \phi)$ and $\mathbf{R}_{y,\tilde{y}} = \exp(-\mathbf{H}_{y,\tilde{y}} / \phi)$ are $m_s \times m_s$ and $n_s \times m_s$ correlation matrices, and $\mathbf{H}_{\tilde{y}}$ and $\mathbf{H}_{y,\tilde{y}}$ are distance matrices between the mutual prediction locations, and between the observation and prediction locations, respectively. Elements of $\mathbf{R}_{\tilde{y}}$ and $\mathbf{R}_{y,\tilde{y}}$ have the same exponential function of minus the distance, and both are divided by the range parameter ϕ. Predictions further away from monitoring sites are more uncertain.

The statistical model for error-in-variable KED can be written as a Bayesian hierarchical model (Van de Kassteele and Stein, 2005). We furthermore assumed that $\boldsymbol{\xi}$ and $\tilde{\boldsymbol{\xi}}$ are independent, identically distributed, random variables with expectations μ_ξ and $\mu_{\tilde{\xi}}$ and variances σ^2_ξ and $\sigma^2_{\tilde{\xi}}$. This makes the model a structural measurement error model, which makes the estimation procedure more stable. Parameters β_0, β_1, β_2, σ^2_s, ϕ, σ^2_q, μ_ξ, $\mu_{\tilde{\xi}}$, σ^2_ξ and $\sigma^2_{\tilde{\xi}}$ were given prior distributions. We chose the following non-informative priors: β_0, β_1, β_2, μ_ξ, $\mu_{\tilde{\xi}} \sim \text{N}(0, 0.001)$, σ^{-2}_s, σ^{-2}_q, σ^{-2}_ξ, $\sigma^2_{\tilde{\xi}} \sim \text{Gamma}(0.001, 0.001)$, $\phi^{-1} \sim \text{Uniform}(0.001, 0.1)$ and then applied Markov chain Monte Carlo (MCMC) integration (Gelman *et al.*, 2004), a technique making an analysis of hierarchical models more feasible. A common MCMC integration is obtained with the Gibbs sampler using WinBUGS (Lunn *et al.*, 2000; Gilks *et al.*, 1994).

Typically, the chains are run for a number of iterations until the outputs are stable (burn-in); after this, a large number of additional iterations are run. Having thinned the chain, we obtained 1000 samples of $\tilde{\boldsymbol{\eta}}$ for each prediction location. The individual samples were then back-transformed and the measurement error for NO_2 and O_3 included, to get the variable of interest $\tilde{\mathbf{y}}$.

The background NO_2 and O_3 concentrations are known on a 500 × 500 m grid (MNP, 2005). The CAR model, however, requires a much higher resolution because traffic contributions can decay rapidly with distance from the road axis. We should therefore apply a resampling of the background concentrations to a 50 × 50 m grid, in our case, by bilinear interpolation. Since the individual samples between the 500 × 500 m grid points were independent in space, a bilinear interpolation would have led to an underestimation of the 97.5-quantile, and an overestimation of the 2.5-quantile on the 50 × 50 m grid. This restriction was forced by WinBUGS (Lunn *et al*., 2000; Gilks *et al*., 1994). Therefore, the individual samples of both components at each location had to be re-ordered first, which means high concentrations with high neighboring concentrations on the grid.

Because of the chemical relationship between NO_2 and O_3 (Seinfeld and Pandis, 1998), we could not arrange both the NO_2 and O3 samples in ascending order. From the observations from 1990-1999 used earlier in section 5.2.2, we found a linear relationship between both components, with a correlation coefficient of -0.86. This correlation should also be present in the individual samples of both components on the prediction grid. We therefore arranged the 1000 individual samples of NO_2 and O_3 separately in ascending order and then drew 1000 pairs from a multivariate normal distribution with correlation -0.86. We sorted the first column and kept the second column coupled to the first, and finally assigned the rank numbers of the second column to the O_3 samples. This resulted in a correlation of approximately -0.86 between the NO_2 and O_3 samples at each location.

5.3.3 Traffic contribution

The CAR model for local traffic contributions is a relatively simple and generally accepted and applied model in the Netherlands (Eerens *et al*., 1993). The roads were each divided into parts, each with its own properties, e.g. traffic intensity and composition. Each part subsequently consisted of several smaller linear segments that described the location and curvature of the part. CAR calculated concentrations for each part symmetrically and perpendicular to the smaller segments.

Because the CAR model is relatively simple, a Monte Carlo uncertainty analysis could be performed. A random draw from a Gaussian distribution was

Table 5.1. Mean, and upper and lower boundaries (95% confidence interval) for trend parameters, spatial covariance parameters and equation error variance.

Parameter	NO_2		O_3	
β_0	1.63	(1.29, 2.00)	-2.29	(-10.9, 14.0)
β_1	0.54	(0.42, 0.62)	1.66	(-2.84, 4.06)
β_2	-0.015	(-0.017, -0.013)	7.1E-4	(-1.65E-3, 3.13E-3)
σ_s^2	0.058	(0.025, 0.14)	0.023	(9.5E-3, 0.071)
ϕ	50.0	(15.06, 155.0)	41.8	(10.6, 171.7)
σ_q^2	6.0E-4	(1.9E-4, 1.4E-3)	2.7E-3	(5.4E-4, 5.2E-3)

made for all parameters and input of the CAR model with $2\sigma = 30\%$ of the parameter and input value (Eerens *et al*., 1993; Van Oorschot *et al*., 2003). For each individual resampled background concentration and random draw, the local traffic contribution was calculated in a 7-step procedure. In step 1 the dilution factor was calculated as a function of distance to the road and step 2 emission as a function of traffic intensity and composition. Step 3 produced traffic NO_x concentrations as a result of steps 1 and 2 and in step 4 the NO_x contributions from each road were summed. In step 5 the NO_x concentrations were converted to NO_2 concentrations by applying an empirical relationship using O_3 background concentrations; in step 6 a correction was made for traffic contributions in the background concentrations by OPS, and in step 7 we determined the total NO_2 concentration by adding the corrected NO_2 traffic contribution to the NO_2 background concentration. We eventually got an empirical probability density function of the total NO_2 concentration at each location.

5.4 Results

Table 5.1 shows parameter estimates for the error-in-variable KED in determining the background NO_2 and O_3 concentrations for 2010. Trend parameter estimates for NO_2 have smaller credible intervals than those for O_3 because of: 1) a higher accuracy of the NO_x to NO_2 conversion and 2) a better relationship between the converted modeled OPS NO_x (modeled NO_2) and measured NO_2. Although the trend parameters are not significant for O_3, we followed the same approach for both components, based on the results of the time varying coefficients as described in section 5.3.2. The variance of the spatial correlated error σ_s^2 is higher than the equation error variance σ_q^2 for both components. This indicates that the error-in-variable KED found a substantial spatially correlated effect in the residuals. Since we

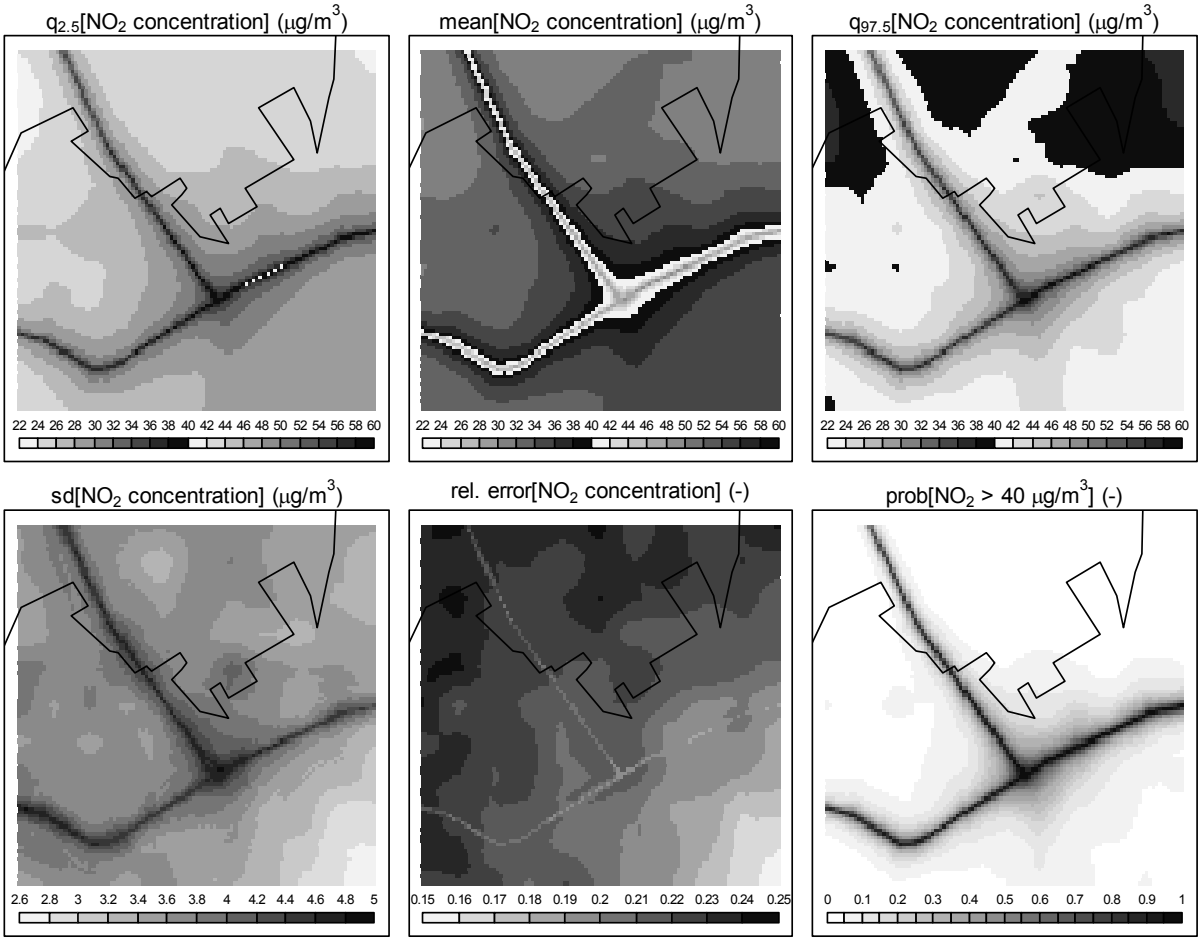

Fig. 5.4. Total NO_2 concentrations near Rotterdam for 2010: 2.5-quantile, mean, 97.5-quantile, standard deviation (sd), relative error and probability of exceeding 40 μg/m³.

used average meteorological data to feed the OPS model, the spatial effects are more pronounced than if OPS output with actual meteorological factors had been used.

Fig. 5.4 shows the total NO_2 concentration (background + traffic) near Rotterdam for 2010. The upper panels show the expectation, and 2.5 and 97.5-quantiles (95% confidence interval of the posterior of the NO_2 concentration). The lower panels show the standard deviation, relative error (i.e. twice the standard deviation divided by the mean) and the probability of exceeding the 40 μg/m³ air quality standard. The standard deviation can also be approximated by $(q_{97.5} - q_{2.5})/4$. The standard deviation and relative error are smaller towards the city centre (southeast) because a monitoring site is located there.

The influence of the highways decreases rapidly with increasing distance from the road axis, reaching approximately 500 m. Based on the expectation for this emission scenario, the NO_2 air quality standard in 2010 is only exceeded within a distance of 100 m to the road axis, except near the highway junction, where both highways contribute to the traffic-related concentrations. Based on the 2.5-quantile, the concentrations in the whole area remain below the air quality standard, whereas

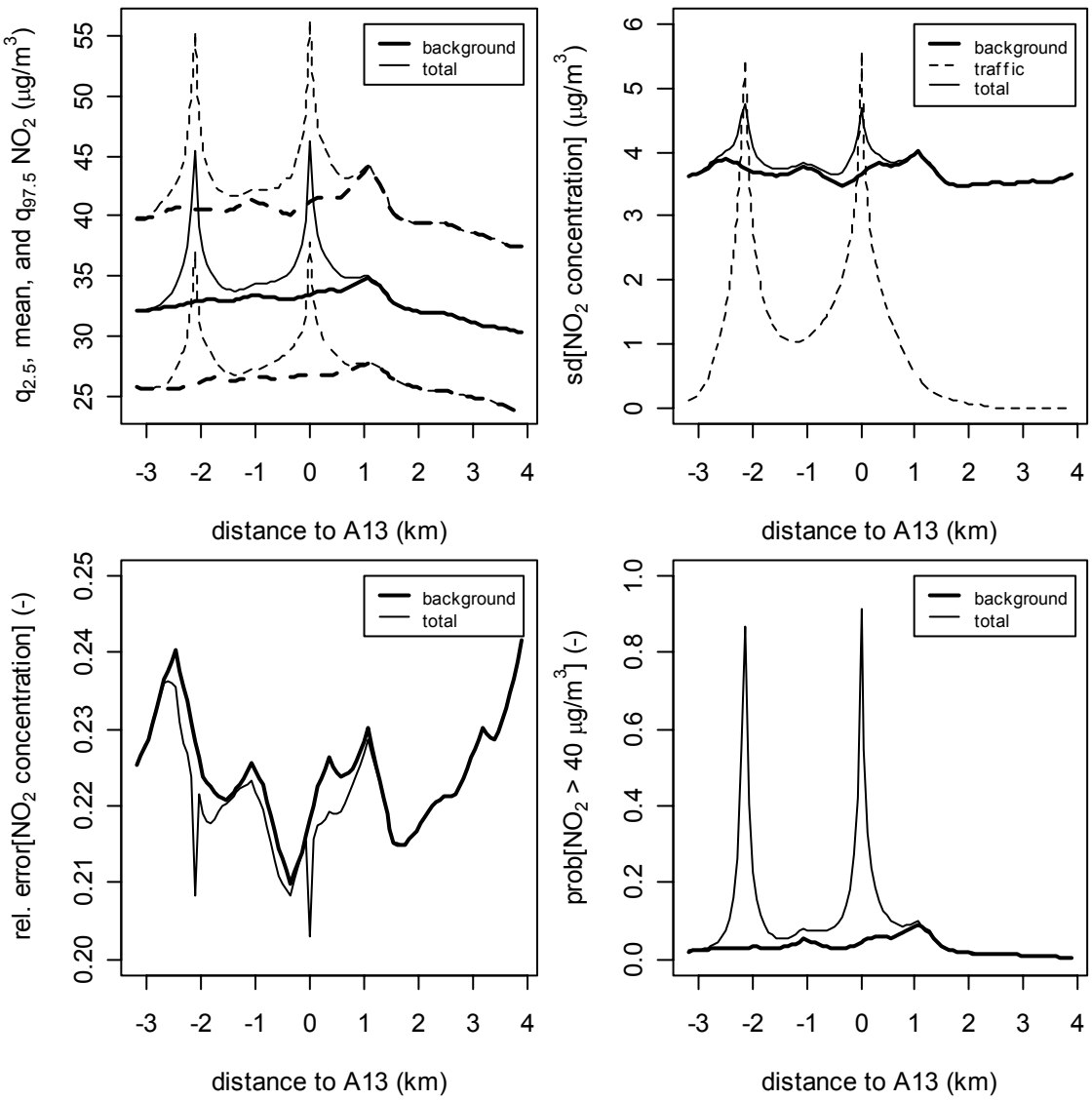

Fig. 5.5. NO_2 profiles of the background and total concentration, with the cross-section of the area from SW to NE as shown in Fig. 5.3. The fine lines in all panels represent background concentrations, the heavy lines total NO_2 concentrations. Additionally, the dashed lines in the upper left panel represent the 2.5 and 97.5 quantiles, and the dashed line in the upper right panel represents the standard deviation of the CAR model output.

for the 97.5-quantile, the standard is exceeded in a large area (also due the high background concentrations).

Standard deviations near highways do not seem as large as expected on the basis of standard deviations of the CAR model only. This is caused by the addition of background concentration and the traffic contribution. Low NO_2 background concentrations are connected with high O_3 background concentrations. As a consequence, more NO_x contributed by traffic can be converted to NO_2 traffic contributions. Furthermore, high NO_2 background concentrations are connected with low O_3 background concentrations. Therefore less NO_x can be converted to NO_2 from

traffic. This chemical mechanism dims the NO_2 variability near highways. This effect is more pronounced in the figure of the relative error.

The exceedance probability of the air quality standard rapidly decreases with increasing distance to the road axis. The 0.50-contour is related to the mean concentration. Exactly the same contour would emerge for the median concentration. The figure indicates that on the 0.10 probability contour, for example, the standard is exceeded with a probability of 10%. Interesting in this example is the exposure of the residential area in terms of an exceedance distance (Folkert *et al.*, 2002). From that distance to the road, the 40 $\mu g/m^3$-contour, it is very unlikely to exceptionally unlikely that the standard will be exceeded.

To illustrate the traffic contribution in more detail, cross-sections of NO_2 background and the total NO_2 concentration profiles are shown in Fig. 5.5 [from the lower left corner (SW) to the upper right corner (NE)]. The two peaks correspond to the presence of highways A20 and A13. The upper left panel shows the two highways to cause a sharp increase in concentration. In the upper right panel showing the standard deviations, we note that the standard deviation of the CAR model exceeds that of the background concentration near the highways; however, the total standard deviation for NO_2 concentration at that location is lower than expected. Furthermore, near highways, the relative error (lower left panel) of the total concentration is smaller than that of the background concentration. The exceedance probability only rises near highways.

5.5 Discussion

We assessed future local NO_2 concentrations by first determining background concentrations using error-in-variable KED, and subsequently adding a traffic contribution. Results were from empirical probability density functions at each location, from which several statistics could be derived.

First, one remark about the methodology we used to account for the time-varying coefficients in the error-in-variable KED model. In equation (4) we assumed that the observations were a linear function of the covariate and time. This is a strong assumption, considering we enforce a relationship that resulted in lower standard errors for the trend parameter estimates. Furthermore, the relationship may be different for components other than NO_2 and O_3. A more elegant and flexible approach would have been to estimate the trend dynamically, as in a structural time series approach (Harvey, 1989). This approach, however, was outside the scope of this chapter.

Second, there is the relationship between NO_2 and O_3. The error-in-variable KED model did not allow simultaneous estimation and interpolation of multiple components. This could result in a distorted image of reality. We therefore accounted for the negative correlation between NO_2 and O_3 pragmatically. The background concentrations could perhaps have been determined simultaneously by applying co-kriging (Wackernagel, 2003), but this is outside the scope of this chapter.

The third remark concerns the CAR model. Its actual purpose is to determine air quality in city streets (canyon) surrounded by buildings (Eerens *et al*., 1993). CAR did not account for traffic stagnation, noise barriers, or other objects near the road, except for trees. Although more advanced models are available, for example, the ADMS-Urban model (McHugh *et al*., 1997), we used CAR here as an initial approach to test our statistical model. Application of a more advanced model may take place in future research.

A fourth remark concerns the uncertainty of the input data. Percentages used here were based on those found in the literature. We note here, however, that these percentages may themselves be uncertain, but do not expect this factor to have more than marginal effects. Undertaking more research on data uncertainty is one possibility for studying the effect of different percentages, such as that demonstrated in Van de Kassteele and Stein (2005). Error-in-variable KED is sufficiently flexible for assigning different error variances to each location.

Finally, it was impossible to validate our model, since no measurements were available on the local level. We refer to Van de Kassteele and Stein (2005) for more details on the error-in-variable KED and to Eerens *et al*. (1993) for the CAR model.

The method developed here is very general and could, if measurements are available, be applied to other components, like PM10, in other years, at other locations (Van Velze *et al*., 2000), in other countries and for other emission scenarios (Beck *et al*., 2001). More information, e.g. a climate scenario, may also be included.

An interesting question now arises as to how these spatial uncertainties and probabilities of exceedances can be communicated to policy makers and other end users. By applying the above methodology, an environmental assessment agency can assess and review uncertainties. Bear in mind though that the Council Directive (EC, 1999) provides only targets and limit values. As a consequence, legislation cannot include uncertainties because uncertain information is not expressed in laws. However, if some sort of information on the uncertainties of results presented is available, it would be desirable if this could be communicated to the end users. Special interest groups might look at the 2.5- and 97.5-quantiles. They might make an appeal for these maps as offering an option, with a certain probability, depending on their own interests, e.g. economic or environmental.

Table 5.2. IPCC terminology for communicating probabilities (IPCC, 2001).

from	to	IPCC terminology
0%	1%	Exceptionally unlikely
1%	10%	Very unlikely
10%	33%	Unlikely
33%	66%	Medium likelihood
66%	90%	Likely
90%	99%	Very likely
99%	100%	Virtually certain

Fisher (2003) illustrates applications of fuzzy set theory for decision-making in air quality. Like ApSimon *et al*. (2002), he noticed that all models introduce uncertainty. The use of fuzzy numbers was avoided in our study because we have uncertainty expressed automatically by the Bayesian approach, resulting in an empirical probability density function at each location. This makes hypothesis testing in relation to the NO_2 standard possible, for example. The standard is not significantly exceeded if the 97.5-quanitle is below it, but is significantly exceeded if the 2.5-quantile is above the standard. The cases in between (2.5-quantile below the standard and 97.5-quantile above it) should be considered as fuzzy, with implications for subsequent action plans, by allowing some exceedances of the standard. Especially these areas require a more detailed assessment, because here it is difficult to exactly diagnose if the standard is exceeded or not.

The exceedance probability maps presented here provide additional information. IPCC (2001) has developed terminology for communicating the uncertainties associated with climate change (Table 5.2), with the idea that instead of giving probabilities, terms would help to communicate uncertainties. The results as presented in section 5.4 could consequently also be formulated in a new map, showing how terms could be applied to each location (Fig. 5.6). This figure is similar to the lower right panels of Fig. 5.4 and Fig. 5.5, except that the probabilities have been replaced by words. These figures are used to provide information to a policy maker on the acceptability of a risk.

5.6 Conclusions

This chapter has shown how to predict future NO_2 concentrations maps based on an emission scenario and to assess uncertainties inside and outside Rotterdam. The error-in-variable KED, which combines past measurements, dispersion model output for the past and future. With a generic dispersion model for local traffic-related emissions, we produced results for local traffic-related air quality that seemed to be

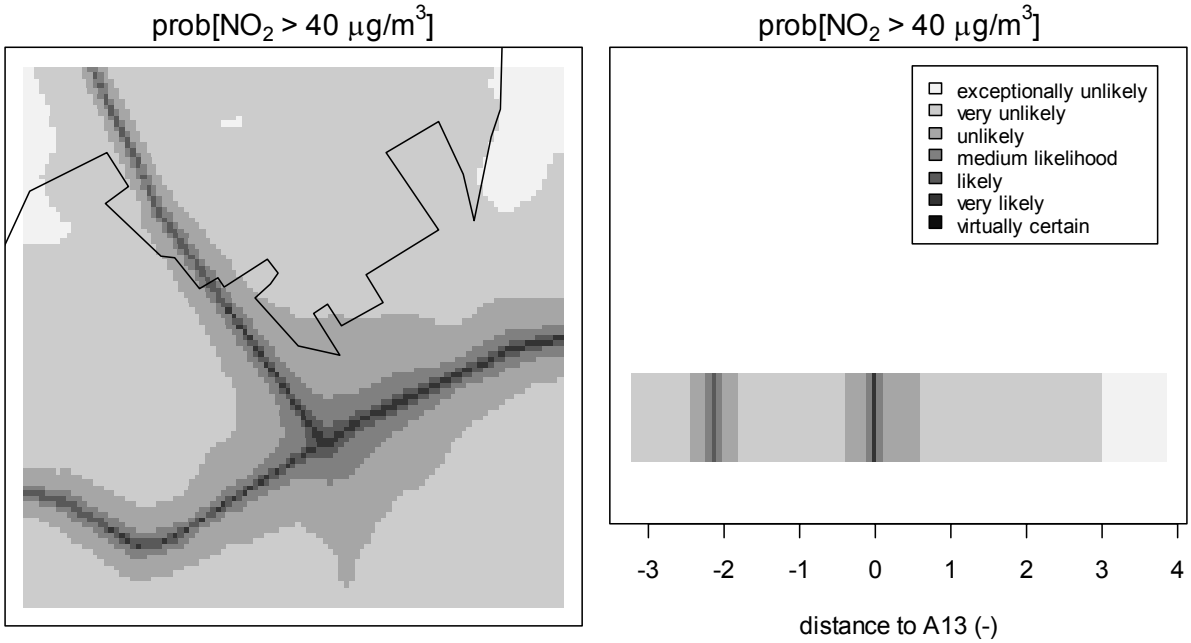

Fig. 5.6. Probability of exceeding the NO_2 air quality standard according to IPCC terminology. The left panel shows the map as in Fig. 5.4; the right panel shows the cross-section as in Fig. 5.5.

valuable in relation to the European air quality standards. Several statistics could be derived from the empirical probability density functions at each location and results were related to the 2010 NO_2 air quality standard.

The mean NO_2 concentrations were shown to exceed the air quality standard at a distance of up to 100 m from highways; however, the areas where the standard was exceeded were fuzzy. For distances of 100 m to 1 km, exceedance probabilities occurred in a range from 0.10 to 0.50. Further away, at 1 km from the highway, where the mean of 35 $\mu g/m^3$ was below the standard according to regulations, the probability of exceedance was still 0.10.

Our methods resulted in probabilistic measures, allowing hypothesis testing. To allow simple communication with users, this study showed how these measures could be translated verbally into probabilistic statements with the use of IPCC terminology.

5.7 Acknowledgements

We would like to thank Wim Blom, Karel van Velze and Hub Diederen of the Netherlands Environmental Assessment Agency for the useful discussions contributing to this chapter.

6. Statistical PM10 mapping using dispersion modeling and satellite observations

This chapter illustrates the use of statistical techniques to standardize ground based measurements of particulate matter (PM10). Concentrations are interpolated over Western Europe using uncertain secondary information from a chemical transport model and of aerosol optical thickness from MODIS satellite observations. A consistent overview of PM10 concentrations over Europe based solely on ground based measurements is complicated by differences between countries. Different monitoring methods are used and calibrations are applied. There also is an inherent limitation to the spatial representativeness of ground based measurements. Validation showed that adding secondary information from either the chemical transport model or the satellite observations improved the PM10 mapping. The URMSE decreased from 5.14 to 4.26 and 4.58 respectively. A combination of both sources of secondary information gave the most accurate and precise predictions, with an URMSE of 3.62. This means that both external sources contain additional information on the spatial distribution of PM10 concentrations and should therefore be preferred.

This chapter is based on J. van de Kassteele, R.B.A. Koelemeijer, A.L.M. Dekkers, M. Schaap, C.D. Homan and A. Stein (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* (submitted).

6.1 Introduction

Exposure to particulate matter is the largest contributor to morbidity and mortality from air pollution in Europe and the US (e.g., Dockery *et al*., 1993; Brunekreef and Holgate, 2002; Brunekreef and Forsberg, 2005). It is estimated that in 2000, about 350,000 people have died prematurely in Europe through long-term exposure to PM2.5, which by far dominates health effects of other air pollutants (EEA, 2005). Recent studies in Europe indicate that some fractions of PM10, especially particles that are emitted from combustion, are of more importance than other fractions (e.g. sea salt or secondary inorganic components) (Hoek *et al*., 2002).

The European Commission has set standards for ambient yearly and daily averaged PM10 concentrations, which are 40 $\mu g/m^3$ and 50 $\mu g/m^3$ respectively (EC, 1999). The latter is not to be exceeded more than 35 days a year. Although emissions of primary PM, and PM precursors (NO_x, SO_2, NH_3), have substantially decreased in Europe over the past decade, still most EU countries do not meet these air quality limit values.

To check compliance with the limit values, assessment of yearly and daily PM10 concentrations is needed for the whole of Europe. However, quantification of PM concentrations is complicated because of large spatial and temporal heterogeneity of PM sources, differences in chemical composition and the relatively short atmospheric lifetime of particulate matter that is typically up to several days in the atmosphere in combination with a large variety of sources and formation pathways. Several ground based measurement networks are currently in operation to monitor PM, but these are limited in space and hence inadequate to provide full European coverage of PM10 concentrations.

In addition, measurement techniques for PM10 other than gravimetry, i.e. the reference method (CEN, 1998), are prone to artifacts. These techniques systematically underestimate PM mass (e.g. Hitzenberger *et al*., 2004; Charron *et al*., 2004). Correction factors are therefore applied (Buijsman and de Leeuw, 2004), but differences between correction factors and the application itself hinder integration of PM10 data on a European scale (CAFE-WGPM, 2004).

In this study we attempt to statistically standardize the measurements and to map the annual average PM10 concentration over Western Europe. This is done by exploiting ground-based measurements, obtained from the AirBase database of the European Environment Agency (EEA), in combination with two explanatory variables: PM2.5 fields as modeled by the chemical transport model (CTM) LOTOS-EUROS and measurements of aerosol optical thickness (AOT), as measured by the Moderate-Resolution Imaging Spectroradiometer (MODIS) instruments on board the

EOS-Terra and EOS-Aqua polar satellites. The benefits of a combination of ground based measurements and modeled fields and satellite data have been illustrated by Van de Kassteele and Stein (2005) and Hutchison *et al*. (2003, 2004, 2005) for the mapping of NO_2 levels in the Netherlands and air quality prediction in Texas, respectively.

Atmospheric modeling provides concentrations consistent for the whole European region. In addition, modeling provides data for all situations; continuous in time. On the other hand, modeled distributions are subject to significant model uncertainties and have a relatively coarse resolution (in our case approximately 35 × 25 km). The satellite observations of AOT are available at higher resolution (approximately 7 × 10 km). The AOT is a measure for the attenuation of light by aerosols, and is, among others, proportional to the number of aerosol particles in the atmosphere. A disadvantage is that the AOT values are only obtained under cloud free conditions and only provide a single snapshot around noon. Previous studies indicate that the spatial gradients in the yearly average AOT and modeled fields are correlated to those of surface PM concentrations (Koelemeijer *et al*., 2005; van Loon *et al*., 2004). Hence, these data may provide useful secondary information to map the PM10 concentrations over Europe. Therefore, a secondary objective of this study is to evaluate the added value of AOT satellite retrievals and modeled fields to the mapping of annual average PM10 concentrations.

As a benchmark we interpolate the standardized PM10 data with universal kriging. In addition, we use error-in-variable external drift kriging (Van de Kassteele and Stein, 2005), which allows interpolation of ground based measurements using domain covering secondary information from multiple uncertain sources and allows for quantification of uncertainties.

6.2 Input data

We have limited our study on the region between 42N and 55N latitude and 6W and 16E longitude. This area covers a large part of Western Europe (e.g. England, The Netherlands, Germany, France, Switzerland and Northern Italy). The rationale for choosing this area is that it provides a relatively dense network of monitoring stations and captures two areas with considerable pollution: the Benelux and the Po Valley (Northern Italy). Furthermore, the aerosol (PM) type is similar in this region. We neglect areas over open sea for three reasons. Firstly, there are no PM observations. Secondly, satellite retrieved AOT is rather insensitive to coarse sea salt particles. Third, models are known to be inaccurate concerning sea salt concentrations over the open ocean. The PM10 measurements, model results and

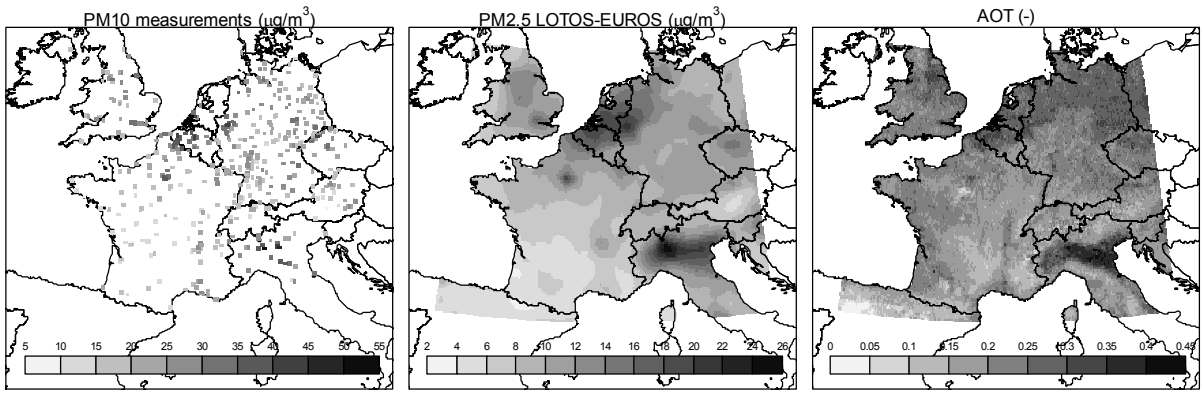

Fig. 6.1. Yearly averages of standardized PM10 measurements, LOTOS-EUROS dispersion model output and AOT satellite data for 2003.

AOT data for this area were selected. The data are shown in Fig. 6.1 and are presented below.

6.2.1 Ground-based PM10 measurements

The PM10 data for 2003 were obtained from the AirBase data base of the ETC-ACC (ETC-ACC, 2005). Annual averages were calculated for all stations, which have at least 90% data availability. Besides hourly or daily concentration data, meta-data such as location, station type and measurement technique were obtained from AirBase. We have used stations, which are representative for background conditions in rural, suburban and urban areas. The concentrations at these stations are typically representative for areas of several km² or larger which allows comparison with modeled distributions and satellite data. We omitted the traffic stations, because they are only representative for more limited areas (scale of several meters). The number of stations per type and country is summarized in Table 6.1. As urban stations are expected to have systematically higher concentrations then rural sites due to urban sources (Lenshow *et al*., 2001), we have used station type as an explanatory variable in the statistical procedure to be able to differentiate between station types.

As mentioned above, PM10 is measured gravimetrically or with several techniques equivalent to the reference gravimetrical method. Most networks use a beta-absorption technique or the tapered element oscillating microbalance (TEOM). These techniques systematically underestimate PM mass because of loss of semi-volatile particles during the measurement process (e.g. Hitzenberger *et al*., 2004; Charron *et al*., 2004) and have to be corrected by a correction factor which may vary substantially in space and even seasonally. As a consequence of the different techniques, procedures and correction factors used in European countries the data in AirBase may not be consistent between countries (Buijsman and de Leeuw, 2004),

Table 6.1. Number of stations per type of surroundings per country in the selected region. Differentiation to measurement technique has been omitted.

	rural	suburban	urban	unknown
Austria	10	8	11	0
Belgium	3	12	6	0
Czech Republic	7	6	11	0
France	10	66	64	49
Germany	57	60	89	0
Great Britain	2	5	35	0
Italy	1	9	14	0
Netherlands	1	0	4	0
Switzerland	6	4	4	0

causing concentration jumps at borders. Hence, we have also used countries and measurement technique as an explanatory variable in the statistical procedure.

Further, an uncertainty estimate is needed in the mapping procedure. For the Netherlands, the measurement uncertainty or precision (2σ) for annual average PM10 concentrations in the Netherlands is 9% (Blank, 2001). We assume that this value is representative for all PM10 stations.

6.2.2 LOTOS-EUROS model data

In this study we used yearly average PM2.5 distribution for 2003 as modeled with the LOTOS-EUROS model as the first explanatory variable. LOTOS-EUROS is a 3D chemistry transport model that is used to simulate the fate of air pollutants over Europe. Based on emission estimates of PM and PM precursors, meteorological data and process knowledge the concentrations of polluting gases and aerosols are simulated on an hour by hour basis. The model has recently been developed based on the models LOTOS and EUROS (Schaap *et al.*, 2005a,b). The horizontal resolution of the model is 0.5° × 0.25° (approximately 35 × 25 km in Europe). LOTOS-EUROS includes primary particles from combustion processes, fine mode sea salt and the formation of secondary inorganic components (SO_4, NO_3 and NH_4). These components make up the largest part of PM2.5 mass concentrations in Europe and the modeled fields of these components were combined to assess the PM2.5 distribution over Europe. The modeled data were interpolated to match the MODIS data by bilinear interpolation.

The results and performance in comparison with observations of the LOTOS-EUROS model are comparable to those of other regional models over Europe (Van

Loon *et al*., 2004; Schaap *et al*., 2005a). Overall, the model underestimates PM2.5 levels systematically, which is mostly caused by the underestimation of primary particulate matter by about a factor 2. The reasons for this feature are discussed in detail for LOTOS by Schaap *et al*. (2004). A systematic (absolute) bias does not strongly affect our methods as we only used the spatial gradients of the distributions modeled by LOTOS-EUROS. We estimated the uncertainty (2σ) of modeled PM2.5 levels to approximately 20%.

6.2.3 MODIS satellite data

The Moderate-Resolution Imaging Spectroradiometer (MODIS) instruments, onboard both the EOS-Terra and EOS-Aqua polar satellites, observe Europe twice per day, around 10:30 and 13:30 local solar time. The retrieval algorithm over land is described in Kaufman and Tanré (1998) and Remer *et al*. (2005). In this study we used the aerosol optical thickness originating from fine aerosol particles. One AOT retrieval is made up of 20 × 20 pixels, corresponding to 10 × 10 km for each AOT value in an image. All individual MODIS image were remapped to a regular grid of 0.1° × 0.1° resolution (approximately 7 × 10 km in Europe), and subsequently, a yearly average AOT map was produced.

The retrieved AOT has been validated against AOT measurements from the ground-based AERONET network. The standard deviation of the retrieved AOT is within 0.05 ± 0.20AOT over land (e.g., Kaufman *et al*., 1997), except in situations with possible cloud contamination, over surfaces with surface water such as coastal areas, and over surfaces with snow or ice cover (Chu *et al*., 2002). The main source of errors in the AOT is uncertainty in the surface reflectance and in the aerosol model. Since our focus was on mapping yearly average PM10 concentrations over land, we estimated the uncertainty in the annual mean by deviding twice the AOT standard deviation for land pixels (assumed to be random errors) by the square root of the number of retrievals. This value (2σ) is approximately 8% of the yearly average AOT. Any systematic error is automatically taken into account by the mapping procedure and can be ignored.

6.3 Methodology

To meet the objective of this study we followed a two step approach: the first step is the standardization of the PM10 measurements (section 6.3.1). The second step is the actual mapping procedure. We followed four variants to map PM10 concentrations.

The first variant is based on the standardized PM10 measurements only (section 6.3.2), the second is based on combining the standardized PM10 measurements with the PM2.5 distributions modeled by LOTOS-EUROS, the third variant is based on combining standardized PM10 measurements with the MODIS observations of AOT, and the fourth variant is based on combining standardized PM10 measurements with both the modeled distributions and the MODIS observations of AOT (all three described in section 6.3.3). Finally, results were validated using a training and an independent validation set (section 6.3.4).

6.3.1 Standardization of the PM10 measurements

In order to account for differences in measured PM10 levels stemming from differences between countries, surroundings and measurement techniques, the measurements have to be standardized using these three factors. We call these factors internal explanatory variables. In addition, the procedure needs external information on pollution levels. Otherwise in one specific country with high pollution levels, these high levels would be regarded as a country effect. This in fact is not true, because it is external due to, for example, higher emissions. Therefore, we have built a linear model using the measured PM10 concentrations as response vector, the LOTOS-EUROS results and AOT as external explanatory variables and assume the spatial distribution of both external explanatory variables represent that of real PM distributions.

It is expected that the residuals of this linear model have different properties in different countries, because the countries in our study area (Fig. 6.1 and section 6.4.1) all have their own measurement techniques and correction factors. Furthermore, it is expected that urban stations give higher residuals than rural stations, since the dispersion model and satellite do not capture concentration variations on scales of (smaller) cities. These properties are accounted for by the internal explanatory variables.

First we applied a log-transformation to the PM10 measurements and LOTOS-EUROS data to ensure that these data follow a Gaussian distribution. This was not necessary for the AOT data. Then, we performed an analysis of variance (ANOVA), ignoring the uncertainties in measurement, model, and satellite data in this stage. Our response vector was measured PM10 ($\mathbf{y}$), external explanatory variables were the LOTOS-EUROS results ($\mathbf{x}_1$) and AOT ($\mathbf{x}_2$), and three factors, or internal explanatory variables, were "country" (C), "surroundings" (S) and "measurement technique" (M). Allowing interactions between the external and internal explanatory variables, the complete linear model looks like:

$$\begin{aligned}\mathbf{y} = {} & \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_{3,i}C_i + \beta_{4,j}S_j + \beta_{5,k}M_k + \\ & \beta_{6,i}C_i\mathbf{x}_1 + \beta_{7,j}S_j\mathbf{x}_1 + \beta_{8,k}M_k\mathbf{x}_1 + \beta_{9,i}C_i\mathbf{x}_2 + \beta_{10,j}S_j\mathbf{x}_2 + \beta_{11,k}M_k\mathbf{x}_2 + \boldsymbol{\varepsilon}\end{aligned} \tag{6.1}$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients. The indices *i*, *j* and *k* represent the corresponding levels of the three factors respectively and $\boldsymbol{\varepsilon}$ is a vector of residuals. We had nine countries, four types of surroundings and seven types of measurement technique (Table 6.1). The linear model (6.1) can be analyzed using any advanced statistical software package. Next we standardized the PM10 measurements for "country", "surroundings" and "measurement technique" and interactions with the two explanatory variables by subtracting the levels and interactions per factor, resulting in

$$\begin{aligned}&\mathbf{y} - \beta_{3,i}C_i - \beta_{4,j}S_j - \beta_{5,k}M_k - \beta_{6,i}C_i\mathbf{x}_1 - \beta_{7,j}S_j\mathbf{x}_1 - \beta_{8,k}M_k\mathbf{x}_1 - \\ &\beta_{9,i}C_i\mathbf{x}_2 - \beta_{10,j}S_j\mathbf{x}_2 - \beta_{11,k}M_k\mathbf{x}_2 = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \boldsymbol{\varepsilon}\end{aligned} \tag{6.2}$$

A reference country, surroundings and measurement technique needs to be chosen to which the others were standardized (section 6.4.1). We call the left hand side of equation (6.2) $\mathbf{y}^*$, the standardized PM10. This $\mathbf{y}^*$ is used to obtain PM10 distributions. By multivariate simulation, using the expectation and covariance of the concerning regression coefficients, we determined the 95% prediction intervals for the standardized PM10 concentrations. These uncertainties were taken along in further calculations.

6.3.2 Universal kriging with measurement error

To map concentrations based on measurements only, the best linear unbiased predictor is kriging (Chiles and Delfiner, 1999). Kriging consists of two steps: a parameter estimation step and a spatial prediction step. We denoted the log of the standardized PM10 measurements with $\mathbf{y}^*$. This quantity however is subject to errors: measurement error and the error due to the standardization. The true value is an unknown latent variable, which we call $\boldsymbol{\eta}^*$, that can only be observed with an additive error, which are assumed to be normally distributed and independent:

$$y_i^* = \eta_i^* + \varepsilon_{y_i}, \text{ with } \varepsilon_{y_i^*} \sim \mathrm{N}(0, \sigma^2_{y_i^*}) \tag{6.3}$$

The indices $i = 1 \ldots n$ represent the monitoring stations with spatial locations $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2]$. The error variances $\sigma^2_{y_i^*}$ are known at each location, and can differ for each

location (sections 6.2.1 and 6.3.1). In case of universal kriging, at each location the latent variable is modeled by the sum of the large scale trend, which is, in this study, a linear function of the location, a spatially correlated error and a so-called equation error:

$$\eta_i^* = \beta_0 + \beta_1 s_{1_i} + \beta_2 s_{2_i} + \varepsilon_{s_i} + \varepsilon_{q_i} \tag{6.4}$$

In the above equation, β_0, β_1, and β_2 are unknown trend parameters. β_0 applies to an offset level, β_1 applies to a linear trend in the WE direction, while β_2 does for the SN direction. We assume that the spatially correlated errors ε_{s_i} are spatially stationary and multivariate normally distributed with variance $\sigma_s^2 \mathbf{R}_{y^*}$ for each *i*, and that ε_{q_i} are univariate normally distributed with a common variance σ_q^2 for each *i*:

$$\boldsymbol{\varepsilon}_s \sim \text{MVN}(0, \sigma_s^2 \mathbf{R}_{y^*}) \tag{6.5a}$$

$$\varepsilon_{q_i} \sim \text{N}(0, \sigma_q^2) \tag{6.5b}$$

Here $\mathbf{R}_{y^*} = \exp(-\mathbf{H}_{y^*} / \phi)$ is an $n \times n$ correlation matrix and $\mathbf{H}_{y^*}$ an $n \times n$ distance matrix between the monitoring station locations. Elements of $\mathbf{R}_{y^*}$ represent an exponential function of minus the distance divided by an unknown range parameter, ϕ, representing the decay rate of spatial correlation with increasing distance. Contrary to $\sigma_{y^*}^2$, both variances σ_s^2 and σ_q^2 are unknown. In geostatistics, these variances are usually called the partial sill and nugget variance.

We estimated the unknown parameters by maximum likelihood parameter estimation. This procedure optimizes the likelihood function over the whole parameter space. For more details we refer to Van de Kassteele *et al*. (2005) or Kitanidis and Shen (1996).

In the prediction step there are $j = 1 \ldots m$ prediction locations. The MODIS grid (section 6.2.3) is used as prediction grid. Spatial prediction can be phrased as a simple extension of the estimation step. Variable of interest is $\tilde{\mathbf{y}}^*$. Latent variable $\tilde{\boldsymbol{\eta}}^*$ at prediction locations $\tilde{\mathbf{s}} = [\tilde{\mathbf{s}}_1, \tilde{\mathbf{s}}_2]$ can only be observed with an additive error. The tildes indicate that we are dealing with predictions. Again, errors are assumed to be normally distributed and independent. The error variance $\sigma_{\tilde{y}_j^*}^2$ is assumed to be the average of all $\sigma_{y_i^*}^2$. The relation for spatial prediction is given by:

$$\tilde{\eta}_j^* = \beta_0 + \beta_1 \tilde{s}_{1_j} + \beta_2 \tilde{s}_{2_j} + \varepsilon_{\tilde{s}_j} + \varepsilon_{q_j} \tag{6.6}$$

In equation (6.6) the variables and parameters have the same meaning as equation (6.4), except for $\varepsilon_{\tilde{s}_j}$, which are multivariate normally distributed according to:

$$\boldsymbol{\varepsilon}_{\tilde{s}} \sim \text{MVN}(\mathbf{R}^{\text{T}}_{y^*,\tilde{y}^*}\mathbf{R}^{-1}_{y^*}\boldsymbol{\varepsilon}_s, \sigma^2_s(\mathbf{R}_{\tilde{y}^*} - \mathbf{R}^{\text{T}}_{y^*,\tilde{y}^*}\mathbf{R}^{-1}_{y^*}\mathbf{R}_{y^*,\tilde{y}^*})) \tag{6.7}$$

In equation (6.7), $\mathbf{R}_{\tilde{y}^*} = \exp(-\mathbf{H}_{\tilde{y}^*}/\phi)$ and $\mathbf{R}_{y^*,\tilde{y}^*} = \exp(-\mathbf{H}_{y^*,\tilde{y}^*}/\phi)$ are $m \times m$ and $n \times m$ correlation matrices, and $\mathbf{H}_{\tilde{y}^*}$ and $\mathbf{H}_{y^*,\tilde{y}^*}$ are distance matrices between the mutual prediction locations, and between the observation locations and prediction locations, respectively. Elements of $\mathbf{R}_{\tilde{y}^*}$ and $\mathbf{R}_{y^*,\tilde{y}^*}$ have the same exponential function of minus the distance, both divided by the range parameter ϕ. Predictions further away from monitoring sites are more uncertain.

We simulated 1000 realizations of our $\tilde{\mathbf{y}}^*$ field, which is still on a log-scale. These conditional simulations allowed us to directly back-transform to the original scale without introducing a bias.

6.3.3 Linear regression with measurement error

For the combination of the standardized PM10 measurement with the secondary information, we can apply error-in-variable external drift kriging (KED) (Van de Kassteele and Stein, 2005). It takes into account uncertainty in the PM10 measurements, the LOTOS-EUROS results, the AOT, the imperfect relationship between the response and external explanatory variables and uncertainty due to spatial prediction.

In the estimation step we have again n monitoring stations. Besides the latent variable $\boldsymbol{\eta}^*$, we introduce two latent variables for the LOTOS-EUROS results and AOT, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ that can only be observed with normally distributed and independent additive errors:

$$y^*_i = \eta^*_i + \varepsilon_{y_i}, \text{ with } \varepsilon_{y^*_i} \sim \text{N}(0, \sigma^2_{y^*_i}) \tag{6.8a}$$

$$x_{1_i} = \xi_{1_i} + \varepsilon_{x1_i}, \text{ with } \varepsilon_{x1_i} \sim \text{N}(0, \sigma^2_{x1_i}) \tag{6.8b}$$

$$x_{2_i} = \xi_{2_i} + \varepsilon_{x2_i}, \text{ with } \varepsilon_{x2_i} \sim \text{N}(0, \sigma^2_{x2_i}) \tag{6.8c}$$

The variable error variances $\sigma^2_{y^*_i}$, $\sigma^2_{x1_i}$ and $\sigma^2_{x2_i}$ are known at each location (see sections 6.2.1 - 6.2.3). In the context of a measurement error model (Cheng and Van Ness, 1999), the latent variables are related. In error-in-variable KED the residuals

represent the sum of a spatially correlated error, as in classical KED, and an equation error. However, a preliminary analysis of our data showed that, because of the explanatory variables, the spatial correlated residuals became zero, meaning that the error-in-variable KED method reduced to a measurement error model. Therefore, to avoid a nonsensical model, we removed the spatial term, so eventually we have:

$$\eta_i^* = \beta_0 + \beta_1 \xi_{1_i} + \beta_2 \xi_{2_i} + \varepsilon_{q_i} \tag{6.9}$$

In the above equation, the parameters have the same meaning as in section 6.3.2, except that the parameters do now apply to the two external explanatory variables. They are comparable with the regression parameters that we saw earlier in section 6.3.1, except that we had to re-estimate them again, since we are now dealing with latent variables.

Spatial prediction is similar to universal kriging. Latent variables $\tilde{\boldsymbol{\eta}}^*$, $\tilde{\boldsymbol{\xi}}_1$ and $\tilde{\boldsymbol{\xi}}_2$ can only be observed with additive errors. Again, errors are assumed to be normally distributed and independent. The error variances $\sigma^2_{\tilde{y}_j^*}$, $\sigma^2_{\tilde{x}1_j}$ and $\sigma^2_{\tilde{x}2_j}$ are assumed to be known at each prediction location and can differ from location to location. Variances $\sigma^2_{\tilde{x}1_j}$ and $\sigma^2_{\tilde{x}2_j}$ follow directly from the LOTOS-EUROS results and AOT data. The relation between $\tilde{\boldsymbol{\eta}}^*$, $\tilde{\boldsymbol{\xi}}_1$ and $\tilde{\boldsymbol{\xi}}_2$ is given by

$$\tilde{\eta}_j^* = \beta_0 + \beta_1 \tilde{\xi}_{1_j} + \beta_2 \tilde{\xi}_{2_j} + \varepsilon_{q_j} \tag{6.10}$$

In equation (6.10) the variables and parameters have the same meaning as equation (6.9). Since we could drop the spatial correlated error term, spatial prediction became simpler as in the original error-in-variable KED model.

The statistical model linear regression with measurement error can be written as a Bayesian hierarchical model. Parameters β_0, β_1, β_2, σ_q^2, ξ_1, ξ_2, $\tilde{\xi}_1$ and $\tilde{\xi}_2$ were given non-informative prior distributions. This, in fact, is the same as maximum likelihood estimation, except that estimation is now done by Markov chain Monte Carlo (MCMC) integration (Gelman *et al.*, 2004). We obtained 1000 samples of $\tilde{\mathbf{y}}^*$ for each prediction location. As in section 6.3.2, these samples were back-transformed to the original scale. We refer to Van de Kassteele and Stein (2005) for further details about these procedures.

6.3.4 Validation

We validated the four mappings by means of a training set and a validation set. From the total set of 554 monitoring stations, we randomly chose 100 stations for a training set, while the validation set consisted of the remaining 454 stations. As in Van de Kassteele *et al*. (2005), we calculated the following three error measures: the mean error (ME or bias), the unbiased root mean squared error (URMSE) and the mean squared standardized error (MSSE):

$$ME = \frac{1}{n_s}\sum_{i=1}^{n_s}(\tilde{y}_i^* - y_i^*) \tag{6.11a}$$

$$URMSE = \sqrt{\frac{1}{n_s}\sum_{i=1}^{n_s}(\tilde{y}_i^* - y_i^*)^2 - ME^2} \tag{6.11b}$$

$$MSSE = \frac{1}{n_s}\sum_{i=1}^{n_s}\frac{(\tilde{y}_i^* - y_i^*)^2}{\mathrm{var}(\tilde{y}_i^*)} \tag{6.11c}$$

The ME indicates the bias of the predictions $\tilde{\mathbf{y}}^*$ to the original observations $\mathbf{y}^*$ and should be close to zero. The URMSE indicates the bias corrected standard deviation of the model and should be close to zero. We used the URMSE because it can be close to zero even in presence of a bias. The MSSE compares the squared differences with the prediction variance, and yields a value that should be close to 1. If this value is greater than 1, then there is an overestimation of variability of the predictions.

We repeated the above procedure 30 times by randomly selecting other training sets. This allowed us to determine the sensitivity to different training sets and to compare the means by significance tests.

6.4 Results

6.4.1 Standardized PM10 measurements

We explored the relation between the measured PM10 and the selected variables by means of an ANOVA. The ANOVA (Table 6.2) shows that both the external explanatory variables LOTOS-EUROS results and AOT were significant (p-value = 0.00 for both), with α = 0.05. Also the variables "country" and "surroundings" were significant (p-value = 0.00 for both). "Measurement technique" was not significant (p-value = 0.52), however, placing "measurement technique" over "country" made

Table 6.2. ANOVA table. Response vector is measured PM10. The other variables are LOTOS-EUROS results (x_1), AOT data (x_2), country (C), surroundings (S) and measurement technique (M). Interactions are denoted by a colon.

	Df	Sum Sq	Mean Sq	F-value	p-value
x_1	1	7.670	7.6700	255.742	0.00
x_2	1	2.803	2.803	93.461	0.00
C	8	7.979	0.997	33.257	0.00
S	3	3.406	1.135	37.854	0.00
M	5	0.126	0.025	0.843	0.52
$x_1 : C$	8	0.420	0.053	1.751	0.08
$x_1 : S$	3	0.423	0.141	4.698	0.00
$x_1 : M$	5	0.494	0.099	3.292	0.01
$x_2 : C$	8	0.175	0.022	0.730	0.67
$x_2 : S$	3	0.251	0.084	2.791	0.04
$x_2 : M$	5	0.085	0.017	0.566	0.73
residuals	503	15.085	0.030		

this factor significant. This can be explained since countries incline to use a preferred technique.

For the interactions, the LOTOS-EUROS-country interaction was almost significant (p-value = 0.08), the LOTOS-EUROS-surroundings and LOTOS-EUROS-measurement technique interactions were significant (p-values = 0.00 and 0.01 respectively). Only the AOT-surroundings interaction was significant (p-value = 0.04). Placing AOT before LOTOS-EUROS gave no different results, except that then AOT explained most of the variance and the significance of the interactions switched with that of LOTOS-EUROS. This can be explained by the fact that LOTOS-EUROS and AOT are related (correlation coefficient = 0.59).

A reference country, surroundings and measurement technique had to be chosen to which the others were standardized. We chose the German rural monitoring stations using gravimetry as our reference. The reason for this choice will be elaborated in the discussion section. Fig. 6.2 shows the standardized PM10 versus the original PM10, both on log-scale. In Fig. 6.3 we summarize the correction factors per country and surroundings, defined as the ratio between the standardized and measured concentrations.

By definition, in Fig. 6.2 the reference measurements are on the 1:1 line. In general, most concentrations are corrected downwards, because most stations are urban and suburban stations, which are not captured by the explanatory variables. It is expected that a country that does not apply correction factors to their measurements shows a ratio larger than one between the standardized PM10 and

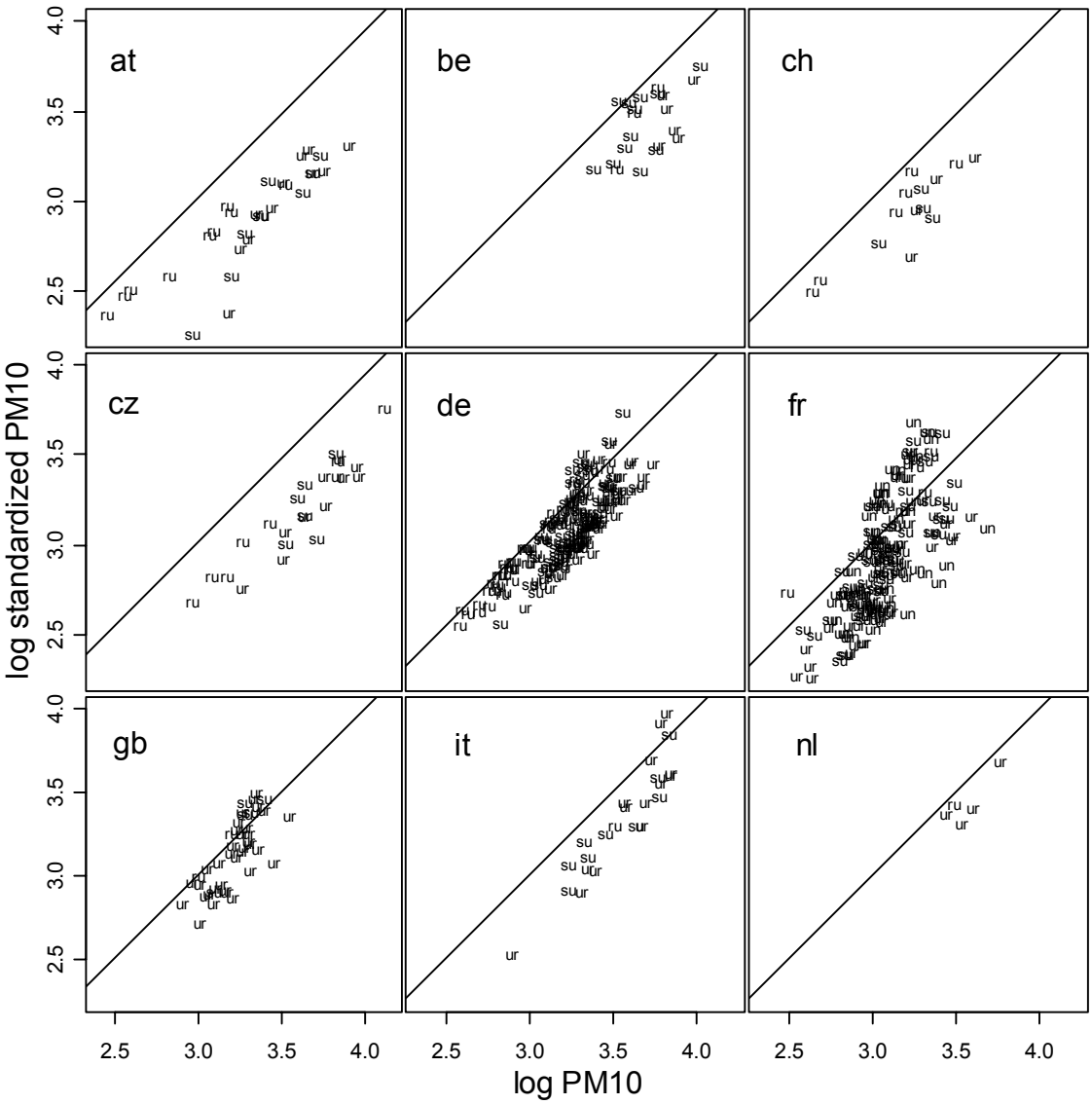

Fig. 6.2. Standardized PM10 versus the original PM10 (both on log-scale) for the nine countries in the study area, grouped by surroundings. The diagonal line is the 1:1 line.

measured PM10 in Fig. 6.3. This is indeed observed for France. The rural stations in the alpine countries show large downward corrections. This may be an artificial phenomenon as regional models are not able to represent the pollution in the valleys. Similarly, AOT values also represent larger areas that also incorporate the cleaner high altitude areas. Besides, AOT over mountains is expected to be underestimated because of shading of the mountains and is uncertain due to glint for example.

Because of the standardization, the correlation coefficient between the PM10 measurements and LOTOS-EUROS results increased from 0.44 to 0.77, and between the PM10 measurements and AOT it increased from 0.48 to 0.72.

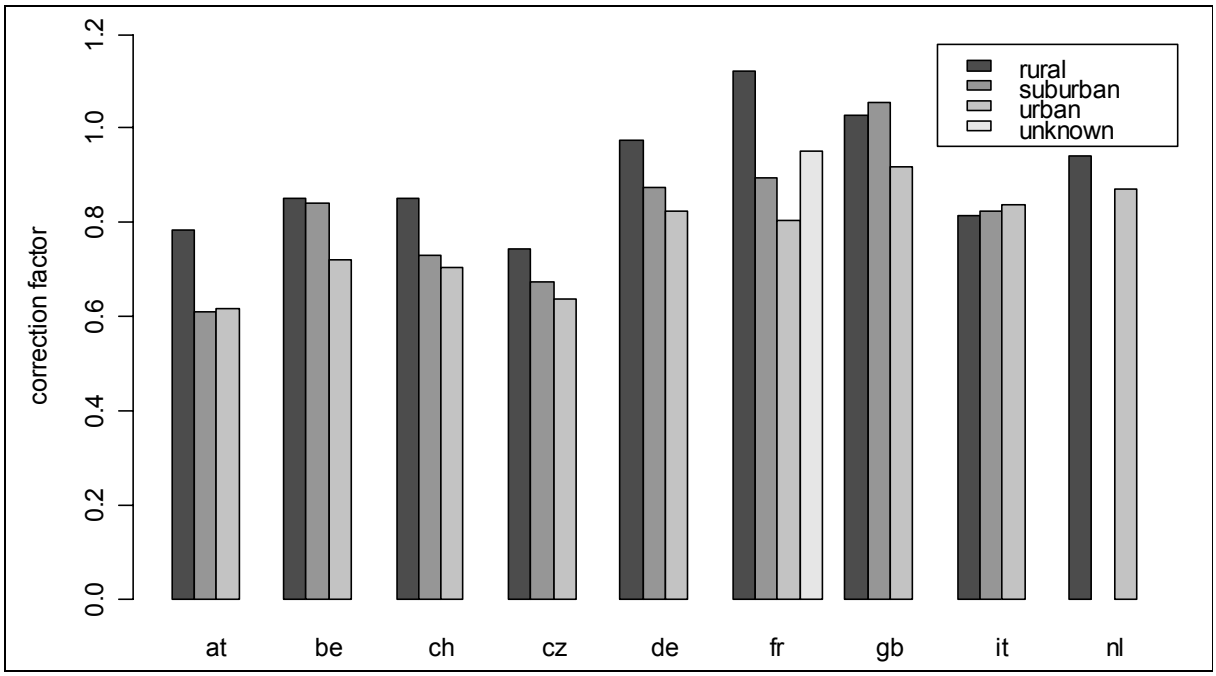

Fig. 6.3. Averaged correction factors for the nine countries.

6.4.2 PM10 concentration maps

Fig. 6.4 shows the expectation of the yearly average PM10 concentrations for 2003, obtained by universal kriging of the measurements only, by linear regression using the LOTOS-EUROS results, the AOT, and both the explanatory variables. High concentrations are found in the Netherlands/Belgium, the Ruhr Area in Western Germany and the Po Valley Area. Smaller areas with high concentrations are found around London, Paris, Lyon and several areas in the Czech Republic. The maps of the measurements and the LOTOS-EUROS results look similar and show smooth surfaces. However, the map based on measurements only clearly show the influence of (groups of) data points evident as circular shaped areas with similar concentrations, e.g. around Milan. The LOTOS-EUROS based map contains process knowledge and therefore appears to contain more spatially realistic information than the map based on the measurements only. The map based on the AOT data is broadly similar to that based on LOTOS-EUROS. For example, the concentrations around the Rhone river valley in southern France are very similar and significantly different from the map based on measurements alone.

In some areas differences can be observed. The mapping based on AOT data provides higher concentrations in the Po Valley. On the other hand, the maps based on the model and the measurements show a more pronounced maximum in the PM10 distribution over the Netherlands and the Ruhr area. Hence, the spatial information provided by the satellite data and the model differs from that of LOTOS-

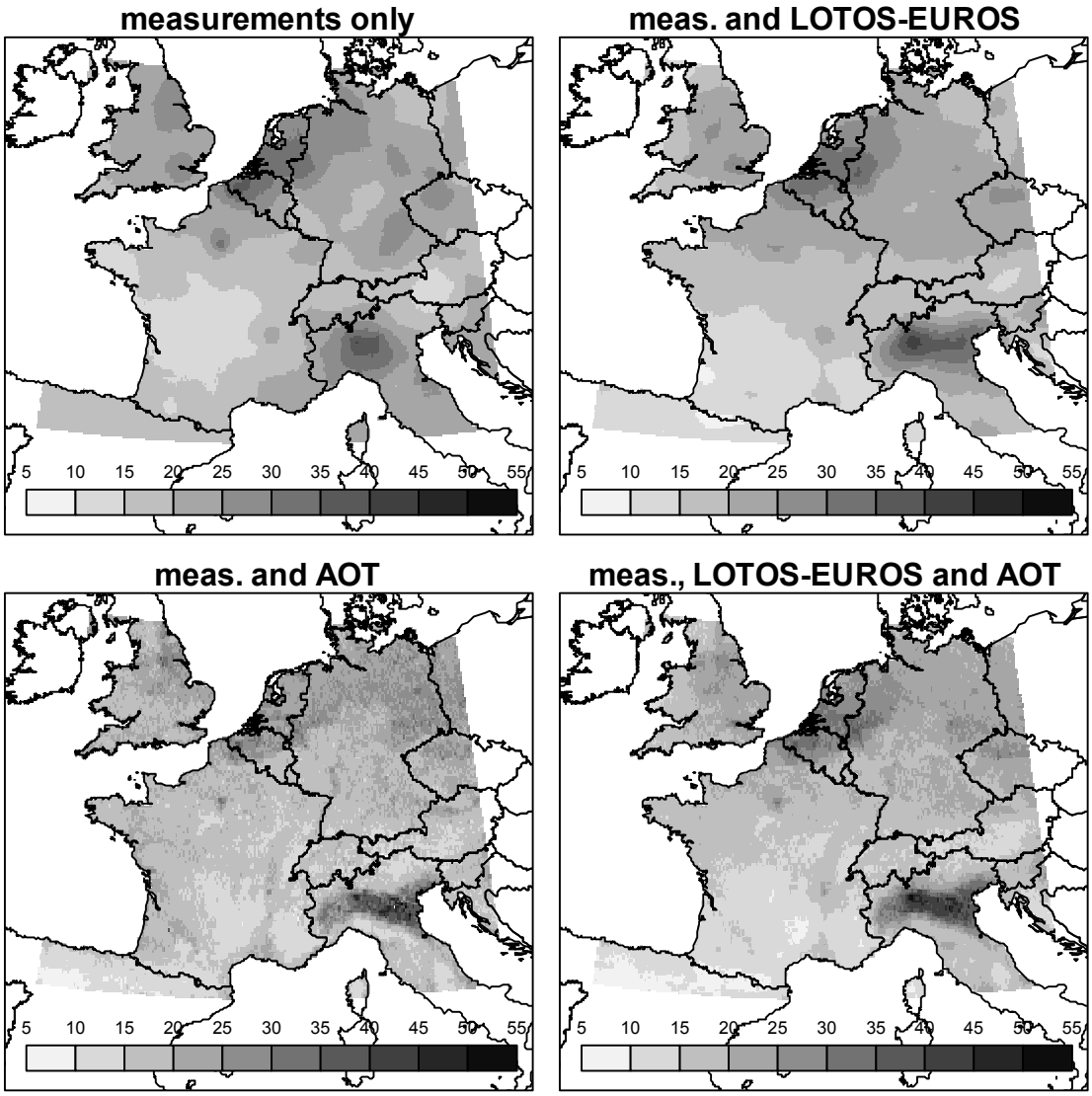

Fig. 6.4. Expectation of the yearly average PM10 concentrations (μg/m³) over Western Europe for 2003.

EUROS. The combination of both the explanatory variables and the PM10 measurements shows a smoother surface than the map based on the AOT data alone and is clearly influenced by the LOTOS-EUROS results, e.g. Germany and the Netherlands/Ruhr area.

Fig. 6.5 shows the corresponding prediction standard deviations. Because of the back-transformation from log-scale, standard deviations increase with higher concentrations, so these maps look similar to the concentrations maps. There are however some differences. For the map based on the measurements only, higher standard deviations are found far away from monitoring sites, like in Northern Spain, Center of Italy and Southern France. Furthermore, standard deviations are slightly higher for the AOT based map, because the correlation between the standardized PM10 measurements and AOT observations is not as high as that with

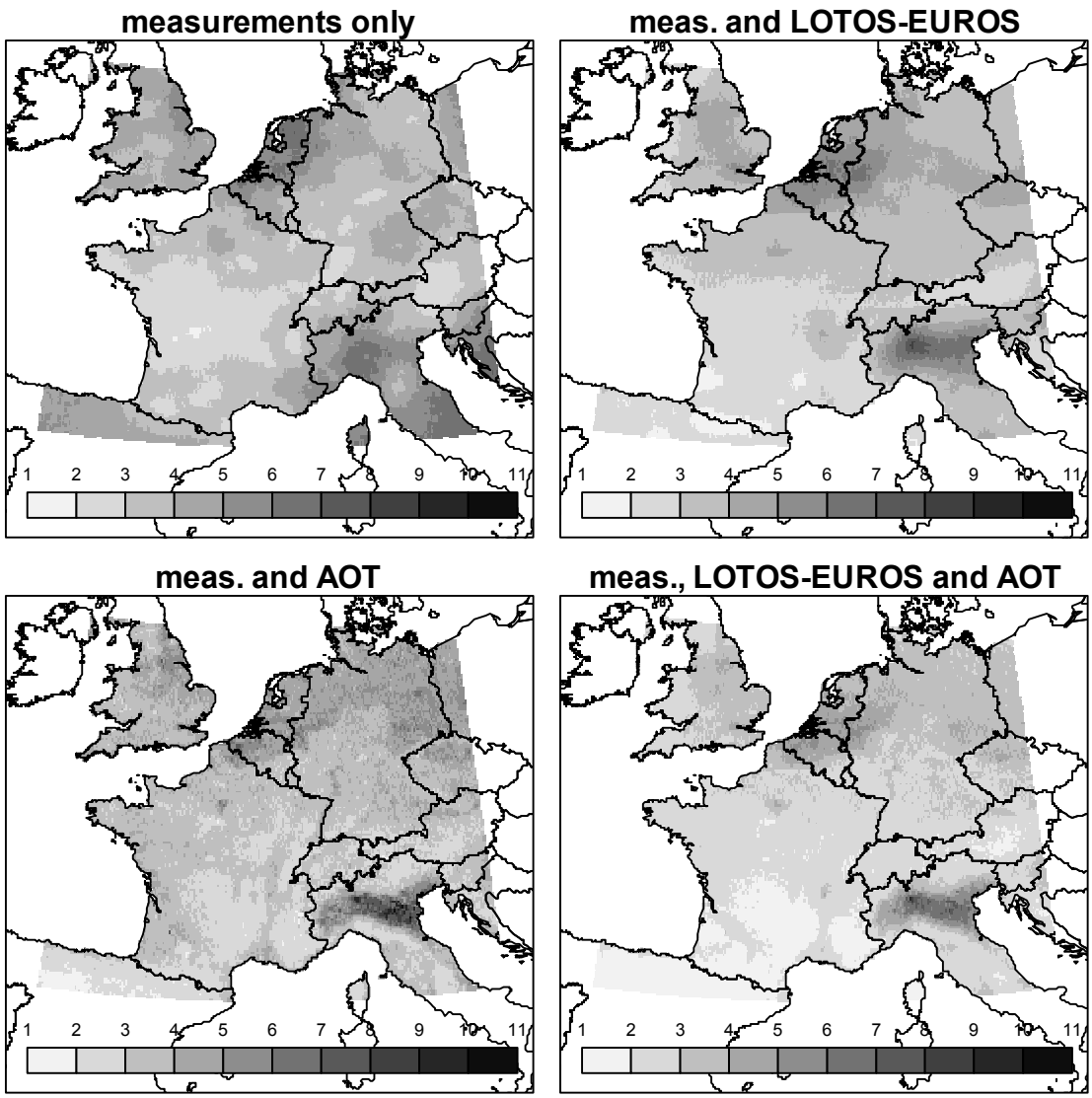

Fig. 6.5. Standard deviation of the yearly average PM10 concentrations ($\mu g/m^3$) over Western Europe for 2003.

the LOTOS-EUROS results. The lowest standard deviations are found for the LOTOS-EUROS based map and the combination of LOTOS-EUROS and AOT based maps.

Fig. 6.6 summarizes all values of these eight figures in box plots. Based on all the individual prediction locations, the minimum, 25-quantile, median, 75-quantile and maximum values for the four expectations and standard deviations are shown. The left panel confirms that the maps based on the measurements only and on based the LOTOS-EUROS results show smooth surfaces. The AOT based map and AOT and LOTOS-EUROS based map show more spatial variation, resulting in a larger range of concentrations. In the right panel we summarize the standard deviations. The AOT based map shows standard deviations that compare to those obtained by kriging the measurements. The LOTOS-EUROS based map yields somewhat lower

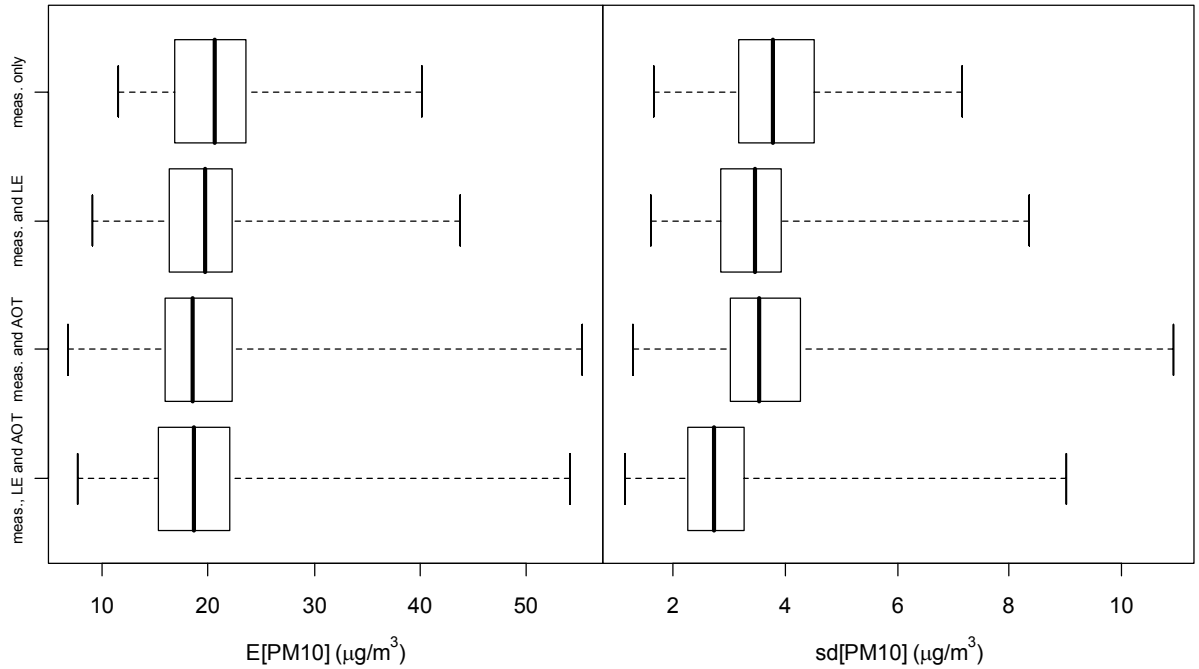

Fig. 6.6. Box plots of the expectations and standard deviations for the four mapping procedures, based on all the individual prediction locations.

standard deviations. Using both the modeled field and AOT data results in the lowest standard deviations and is therefore the most precise. The median standard deviation is reduced by about 25 % compared to the other maps.

6.4.3 Validation

Table 6.3 shows the means and standard errors of the three errors measurers for the four mapping procedures. From these values 95% confidences intervals can be calculated. These are given by $mean \pm t_{29}(0.025)\ se(mean)$, where $t_{29}(0.025) = 2.045$ is the critical value at the 2.5-point of the Student's t-Distribution with 29 degrees of freedom. The mean errors (ME) for the four maps are very small, and all around zero (0.00, 0.06, -0.19, 0.00 $\mu g/m^3$ respectively) and do not significantly differ from zero, except for the map based on AOT data (p-value = 0.97, 0.51, 0.05, 0.99 respectively). The four maps do not differ significantly from each other regarding the ME.

The unbiased root mean squared error (URMSE) is the highest for the maps based on the measurements only (5.14), and the lowest for the combination based on the LOTOS-EUROS results and AOT data (3.62). The maps based on LOTOS-EUROS (4.26) and the AOT data (4.58) show URMSE value in between these. The URMSEs all differ significantly from zero (p-value = 0.00 for all). The four maps all differ significantly from each other regarding the URMSE.

Table 6.3. Mean error (ME), unbiased root mean squared error (URMSE) and mean squared standardized error (MSSE) for validation of the four mapping procedures. The values are averages over 30 samples. Standard error of the average between parentheses.

	ME	URMSE	MSSE
measurements only	0.00 (0.121)	5.14 (0.059)	1.19 (0.073)
meas. and LOTOS-EUROS	0.06 (0.094)	4.26 (0.014)	1.18 (0.037)
meas. and AOT	-0.19 (0.091)	4.58 (0.012)	1.20 (0.041)
meas., LOTOS-EUROS and AOT	0.00 (0.077)	3.62 (0.012)	1.31 (0.057)

Regarding the mean squared standardized error (MSSE), al maps are not significantly different from each other, but all are significantly different from, but close to, 1 (p-value = 0.01, 0.00, 0.00, 0.00 respectively). The MSSE is closest to 1 for the map based on the LOTOS-EUROS results (1.18), however, the maps based on the measurements only and the AOT data are very similar to the LOTOS-EUROS based map regarding the MSSE (1.19 and 1.20 respectively). The map based on the combination of the LOTOS-EUROS results and AOT data shows the highest MSSE (1.31).

The validation exercise shows that the most accurate predictions are based on a combination of the LOTOS-EUROS results and AOT data (lowest ME and URMSE), although prediction variances are underestimated.

6.5 Discussion

This chapter showed the successful application of (geo)statistical techniques to give a consistent PM10 field over Western Europe. First we presented a statistical method to standardize PM10 measurements across Europe. Next we introduced two external explanatory variables to increase the spatial information and to reduce uncertainties in the interpolated PM10 field. The advantage of statistical mapping of PM10 concentrations using explanatory variables is that these explanatory variables do not need to be the same geophysical quantity as the primary variable. The only requirement is that an explanatory variable is linearly related to the primary variable. Any bias or multiplication can be taken into account. Furthermore, if any spatial correlated residuals were present, these effects can be modeled by applying external drift kriging (Van de Kassteele and Stein, 2005). Below, we discuss the issues relating to the standardization of the PM10 data and the explanatory variables.

The concentration maps derived in this study are sensitive to the choice of the reference in the standardization procedure. Another reference would yield systematically higher or lower depending on the reference chosen. The German rural

data were chosen because they are obtained with the prescribed method in Europe. Equally important is that the number of reference data points influences the quality of the results. If there are too little reference points to fit he data to, the influence of single points may become very important. Only one outlier with respect to the predictors, a leverage point, affects the regression model considerably. Therefore, a reference should be chosen with care. In case of a large number of reference points (as applicable for Germany), leverage points have less effect.

We used all PM10 monitoring stations in Western Europe that measured concentrations at rural and (sub)urban scales. Street stations were not used because these local scales can not be resolved by LOTOS-EUROS and the satellite measurements. However, LOTOS-EUROS and the satellite also do not show enhanced PM or AOT over small cities. This resulted in higher residual values for stations located in these cities, and hence, the corresponding PM data were all corrected downwards. We could have omitted all urban stations. However, this would have caused large data reduction (97 versus 554 stations) causing larger estimation errors. Due to the standardization our maps should be regarded as a regional background concentrations map of PM10, without showing the contribution of concentration variations by smaller cities.

We used the LOTOS-EUROS results for PM2.5 as an explanatory variable. In all procedures we have implicitly assumed that the ratio PM10:PM2.5 is constant over the whole region. A number of studies indicate that this ratio is quite stable: about 70 +/- 10% on average in the area under study, e.g. see compilation by Putaud *et al*. (2004). Hence, this assumption appears to be valid under most circumstances. We feel that the uncertainty in this assumption is part of the overall uncertainty in the modeled gradients, which is directly related to the ability of models to simulate the distribution of PM and its components. Important uncertainties are associated with the emissions of carbonaceous and natural particles, vertical mixing under stabile conditions and the possible importance of secondary organic aerosol. For a detailed discussion on model uncertainties we refer Schaap *et al*. (2004).

The AOT data have other characteristics than the measurements and the modeled fields. The retrieval of AOT data is only possible when no clouds are present. Hence, the AOT distribution represents a composite of cloud free situations at noon. Although the MODIS data were found to represent daily average aerosol loading (Kaufman *et al*., 2000), the AOT distribution may be biased to good weather situations. Furthermore, some areas may be influenced by systematic biases in observed AOT due to undetected clouds and sun glint. In general, when uncertainties in the explanatory variable increase the prediction variances increase as well. The effect of different uncertainties of explanatory variables to mapping procedures is described in detail in Van de Kassteele and Stein (2005).

Despite the uncertainties mentioned above we show that the use of secondary information for mapping PM10 concentrations over Europe is advantageous. Although the interpolation of the measurements is not completely independent from the modeled and AOT data (through the standardization), adding the secondary information yields lower prediction errors and shows a better agreement with the validation data sets. Using either of the secondary variables alone gave similar improvements of the predicted fields. Using both variables, however, yields significantly better results. Hence, both the satellite retrievals as the model provide unique information to the procedure. The uncertainties in the interpolated PM10 fields will decrease with an improved representation of particulate matter in models, improved retrieval algorithms and the development and use of artifact free and consistent PM10 measurement methods. This study also illustrates the benefits of combining different sources of information: in situ measurements, remote sensing and modeling.

6.6 Conclusions

It is impossible to get a comprehensive overview of PM10 concentrations in Europe based on ground based measurements only. This is because of the limited spatial representativeness of ground-based measurements, and systematic differences stemming from different measurement methods used in different countries. To improve mapping of PM10 concentrations over Europe, the use of secondary information -from the LOTOS-EUROS chemical transport model and MODIS satellite observations of aerosol optical thickness (AOT)- has been explored.

The analysis showed that, accounting for the pollution levels provided by the secondary information, there were significant differences between countries and surroundings. Measurement technique was not significant. Most measured concentrations were corrected downwards, because most stations were urban and suburban stations, which were not captured by the secondary information. The standardization caused the correlation coefficients between the PM10 measurements and LOTOS-EUROS and AOT data to increase.

This study showed to which degree external information improves the spatial prediction of PM10. Validation showed that adding external information from either the LOTOS-EUROS model or the AOT satellite observations improved the spatial prediction of PM10 levels in Europe. The URMSE decreased from 5.14 to 4.26 and 4.58 respectively. The most precise predictions are based on a combination of the LOTOS-EUROS results and AOT data, with an URMSE of 3.62. The study illustrates

the benefits of combining different sources of information: in situ measurements, remote sensing and modeling.

6.7 Acknowledgements

We would like to thank the NASA staff for providing the MODIS data and Jacques Withagen of Biometris for his useful suggestions. Part of this work was financed by the Netherlands Agency for Aerospace Programmes - NIVR through the PARMA project.

7. General conclusions

Air quality measurements are in general accurate and precise at a certain location, but interpolation of a limited number of observations causes imprecise maps. Dispersion model output on the other hand is inaccurate (biased) and imprecise compared to measurements, but it has a much higher spatial resolution and provides more detailed information. A combination of both leads to better maps.

This thesis showed how to use geostatistics to combine a limited number of air quality measurements and inaccurate and imprecise, but domain covering dispersion model output to acquire detailed air quality maps on local, national and international scales, and to show that this improves spatial predictions, i.e. smaller bias and smaller uncertainties. Uncertainties were quantified. The objectives were dealt with in five chapters, which Table 7.1 summarizes. This chapter considers the research questions Q1-Q11 from chapter 1.

To interpolate spatial count data, such as the number of ozone exceedance days, we chose an explicit model-based geostatistical approach. First, we assumed an approximation by a Poisson distribution, and second, a log-Normal distribution. The log-link in both models made model and parameter comparison easier. The Poisson model gave a better representation of the random field process of the number of exceedance days. For environmental assessment applications, however, we concluded the log-Normal model to be preferred, considering its capacity to predict the expected number of exceedance days instead of an intensity field (Q1).

When making interpolations of a component with a large spatial range, such as ozone exceedance days, for a small area such as the Netherlands, incorporating observations from surrounding countries in the analysis was beneficial. Predictions near the Dutch border still depend on observations in Germany. Furthermore, inclusion of more observations improved parameter estimation and resulted into more precise predictions (Q5). For a small dataset, e.g. mapping of ozone exceedance days in de Netherlands, use of prior information avoids problems with convergence of the MCMC algorithm, which kept on fluctuating if flat priors were used. Also, even use of a limited data set allowed us to map the number of exceedance days (Q3). These maps, including their uncertainties, might be used in the future to study environmental relations between ozone and risks for public health.

Additional information from dispersion models improved interpolation of measurements. A model-based geostatistical interpolation procedure using

Table 7.1. Summary of all chapters showing the component, quantity of interest, secondary information source, mapping technique, parameter estimation procedure, spatial scale and time.

compo-nent	variable	secondary information	mapping method	parameter estimation	scale	time	chapter
O_3	exceedance days	none	UK	Bayesian	inter-national	present	2
NO_x	yearly averages	OPS	UK/ KED	RML/ Bayesian	national	present	3
NO_2	yearly averages	OPS	eiv-KED	Bayesian	local	present	4
NO_2	yearly averages	OPS	eiv-KED	Bayesian	local	2010	5
PM10	yearly averages	LOTOS -EUROS/ MODIS	UK/ ME	RML/ Bayesian	inter-national	present	6

additional information led to more accurate and precise spatial interpolation results. If applied as an external drift (KED), the dispersion model output provided more detail in the concentration maps than maps based the measurements only (Q2). Standard deviations for kriging with external drift were much smaller than those for universal kriging. Furthermore, KED also allows handling of biased additional information. This can be beneficial if the pollution sources are missing or unknown. KED accounts for systematic errors by use of regression parameters.

Cross-validating of the different interpolation procedures was done by replacing common error values, i.e. mean error, unbiased root mean squared error and mean squared standardized error, by intervals obtained by repeatedly selecting different subsets from a set of test data. Reliability intervals of the interpolation results could be compared, showing the inclusion of the deterministic OPS model to lead to a substantial improvement in the predictions. Where small subsets are applied (10 observations), KED leads to a much lower spread in mean error values than universal kriging.

Bayesian interpolation methods had advantages over restricted maximum likelihood (RML) methods. In the case of >20 observations, this has little effect. If prior information is available however, Bayesian inference leads to lower prediction standard deviations. In case of 10-20 observations, Bayesian inference with prior information is more robust than RML estimation (Q3).

Should parameters kept fixed or should they be re-estimated every time? The lower the number of observations, the larger uncertainties become. Re-estimated parameters are to be preferred, because specific process changes and parameter uncertainties are then incorporated, thereby approaching reality the best. On the

other hand, when parameters are kept fixed, there is no problem with sparse observations (Q4). The number of observations that is needed to make an accurate map of sufficient quality depends on the interpolation method, but also on the spatial variation and the assumption that processes described by the OPS model does not change. For universal kriging with re-estimated parameters using RML estimation, this will be 30 observations. For universal kriging with fixed parameters using RML estimation or Bayesian inference, 20 observations might be needed. For KED, the results were mainly determined by the external trend. So for only 10 observations, KED still produced sensible results (Q5). These numbers may apply to NO_x in this thesis, but in general certainly not.

In this thesis the model-based approach of KED was extended by allowing uncertain secondary information. This new interpolation approach was called error-in-variable KED. It showed a successful creation of concentration maps based on uncertain measurements and uncertain dispersion model output. Error-in-variable KED proved sufficiently flexible for assigning different error variances to each location, both to the measurement and the OPS dispersion model output. The Bayesian approach for spatial modeling was extremely useful in this context, because it had the advantage that there was no need to write the model in a complicated analytical formulation, which should be the case if uncertain covariates are introduced and parameters are to be estimated by RML. Furthermore, the Bayesian approach automatically took into account parameter uncertainty.

Different uncertainties in the measurements and dispersion models affect the eventual results. Use of error-in-variable KED approach showed that high accuracy in the measurements is to be preferred above high accuracy in the OPS dispersion model output. Since the measurement precision was already high (2σ = 10%), it is more beneficial to increase precision of the OPS output (Q6). Evaluating OPS output uncertainty, however, was outside the scope of this thesis.

These methods can be applied in scenario studies for determining future air quality. We have shown how to predict future NO_2 concentrations maps based on an emission scenario and to assess uncertainties inside and outside Rotterdam by applying error-in-variable KED (Q7). With a simple extension allowing for time, it combined past measurements with dispersion model output for the past and future. With a generic dispersion model for local traffic-related emissions, we produced results for local traffic-related air quality that seemed to be valuable in relation to the European air quality standards. Several statistics could be derived from the empirical probability density functions at each location and results were related to the 2010 NO_2 air quality standard (Q8). It showed that the mean NO_2 concentrations exceeded the air quality standard at a distance of up to 100 m from highways; however, the areas where the standard was exceeded were fuzzy. For distances of 100 m to 1 km,

exceedance probabilities occurred in a range from 0.10 to 0.50. Further away, at 1 km from the highway, where the mean of 35 $\mu g/m^3$ was below the standard according to regulations, the probability of exceedance was still 0.10.

Our methods resulted in probabilistic measures, allowing hypothesis testing. To allow simple communication with users, this thesis showed how these measures could be translated verbally into probabilistic statements with the use of IPCC terminology (Q9).

We showed how to deal with different measurement techniques between countries. It is impossible to get a comprehensive overview of PM10 concentrations in Europe based on ground based measurements only. This is because of the limited spatial representativeness of ground-based measurements, and systematic differences stemming from different measurement methods used in different countries. To improve mapping of PM10 concentrations over Europe, the use of secondary information -from the LOTOS-EUROS chemical transport model and MODIS satellite observations of aerosol optical thickness (AOT)- has been explored (Q10).

It was shown that, accounting for the pollution levels provided by the secondary information, there were significant differences between countries and surroundings. Most measured concentrations were corrected downwards, because most stations were urban and suburban stations, which were not captured by the secondary information. The standardization caused the correlation coefficients between the PM10 measurements and LOTOS-EUROS and AOT data to increase.

Validation showed that adding external information from either the LOTOS-EUROS model or the AOT satellite observations improved the spatial prediction of PM10 levels in Europe. The URMSE decreased from 5.14 to 4.26 and 4.58 respectively. The most precise predictions are based on a combination of the LOTOS-EUROS results and AOT data, with an URMSE of 3.62. We illustrated the benefits of combining different sources of information: in situ measurements, remote sensing and modeling (Q11).

Concluding, KED appeared to be flexible method to combine measurements of air quality with secondary information. The KED method was extended (error-in-variable KED) to allow for multiple sources of uncertain secondary information, which has many applications in accurate and precise air quality mapping.

References

Acharya, R.C. (1994). From emission to concentration. Analysis of ammonia predictions by the OPS Model. *M.Sc. Thesis H.H. 203.* IHE Delft, Delft, the Netherlands.

ApSimon, H.M., Warren, R.F., Kayin, S. (2002). Addressing uncertainty in environmental modeling: a case study of integrated assessment of strategies to combat long-range transboundary air pollution. *Atmospheric Environment*, **36**, 5417-5426.

Backes, Ch.W., Van Nieuwerburgh, T., and Koelemeijer, R.B.A. (2005). Transformation of the first Daughter Directive on air quality in several EU member states and its application in practice. *European Environmental Law Review, 14(6)*, 157-164.

Beck, J.P., Van Bree, L., Van Esbroek, M.L.P., Freijer, J.I., Van Hinsberg, A., Marra, M., Van Velze, K., Vissenberg, H.A., and Van Pul, W.A.J. (2001). Evaluation of the Acidification targets: the emission alternatives. *RIVM report 725501002*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Berger, J.O., De Oliveira, V., and Sanso, B. (2001). Objective Bayesian Analysis of Spatially Correlated Data. *Journal of the American Statistical Association*, **96**, 361-1374.

Bernardinelli, L., Pascutto, C., Best, N.G. and Gilks, W.R. (1997). Disease mapping with errors in covariates. *Statistics in Medicine*, **16**, 741–752.

Bertino, L. and Wackernagel, H. (2002). Case studies of change-of-support problems. *Technical report N–21/02/G.* ENSMP - ARMINES, Centre de Géostatistique, Fontainebleau, France.

Besag, J.E. (1994). Discussion on the paper by Grenander and Miller. *Journal of the Royal Statistical Society B*, **56**, 591-592.

Blank, F.T. (2001). Measurement uncertainty of the Dutch National Air Quality Monitoring Network. *KEMA Report 50050870-KPS/TCM 01-3063*, KEMA, Arnhem, the Netherlands.

Boer, E.P.J., Dekkers, A.L.M., and Stein, A. (2002). Optimization of a monitoring network for sulfar dioxide. *Journal of Environmental Qualtiy*, **31**, 121-128.

Borowiak, A., Lagler, F., Gerboles, M., and DeSaeger, E. (2000). EC harmonization programme for air quality measurements. Intercomparison exercises 1999/2000 for SO_2, CO, NO_2 and O_3. *EC report EUR 19629*, Joint Research Centre, Ispra, Italy.

Bourennane, H., King, D., and Couturier, A. (2000). Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. *Geoderma*, **97**(3-4), 255-271.

Bourennanne, H., King, D., Chery, P., and Bruand, A. (1996). Improving the kriging of a soil variable using slope gradient as external drift. *European Journal of Soil Science*. **47**(4), 473-483.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B*. **26**, 211–246.

Breslow, N.E. and Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**: 9-25.

Brunekreef, B. and Forsberg, B. (2005). Epidemiological evidence of effects of coarse airborne particles on health. *European Respiratory Journal,* **26**, 309-318.

Brunekreef, B. and Holgate, S.T. (2002). Air pollution and health. *The Lancet,* **360**, 1233-1242.

Buijsman, E. and De Leeuw, F.A.A.M. (2004). PM10 measurement results and correction factors in Airbase. *ETC/ACC Technical Paper 2004/4*, November 2004.

Buringh, E. and Opperhuizen, A. (2002). On health risks of ambient PM in the Netherlands. *RIVM report 650010032*, Bilthoven, The Netherlands.

CAFE-WGPM (2004). Second position on particulate matter. CAFE Working Group on Particulate Matter.

Cassiraga, E.F., Gomez-Hernandez, J.J., Soares, A., and Froidevaux, R. (1997). Improved rainfall estimation by integration of radar data: A geostatistical approach. *Proceedings GeoENV I - Geostatistics for environmental applications*. Lisbon, Portugal, 18-19 November 1996. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 363-374.

CEN (1998) Air Quality. Determination of the PM10 fraction of suspended particulate matter. Reference method and field test procedure to demonstrate. Reference equivalence of measurement methods. *Report EN 12341*, Brussels.

Charron, A., Harrison, R.M., Moorcroft, S., and Booker, J. (2004). Quantitative interpretation of divergence between PM10 and PM2.5 mass measurement by TEOM and gravimetric (Partisol) instruments. *Atmospheric Environment,* **38**, 415-423.

Cheng, C. and Van Ness, J.W. (1999). Statistical regression with measurement error. *Kendall's Library of Statistics 6,* Arnold Publishers, London.

Chilès, J.P. and Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. Wiley, New York.

Christakos, G. (1992). *Random field models in earth sciences*. Academic Press.

Christensen, O.F. and Ribeiro Jr., P.J. (2002). GeoRglm - a package for generalised linear spatial models. *R News,* **2**(2), 26-28 .

Christensen, O.F. and Waagepetersen, R.P. (2002). Bayesian prediction of spatial count data using generalised linear mixed models. *Biometrics*, **58**, 280-286.

Chu, D.A., Kaufman, Y.J., Ichoku, C., Remer, L.A., Tanré, D., and Holben, B.N. (2002). Validation of MODIS aerosol optical depth retrieval over land. *Geophysical. Research Letters,* **29**(12*)*, doi:10.1029/2001GL013205.

Cressie, N.A.C. (1993). *Statistics for Spatial Data, revised edition*. John Wiley & Sons, New York.

Cui H., Stein, A. and Meyers, D.E. (1995). Extension of spatial information, Bayesian kriging and updating of prior variogram parameters. *Environmetrics*, **4**, 373-384.

De Oliveira, V., Kedem, B., and Short, D.A. (1997). Bayesian prediction of transformed gaussian random fields. *Journal of the American Statistical Association*, **92**, 1422-1433.

Derwent R.G., Hov, O., Asman, W.A.H., Van Jaarsveld, J.A., and De Leeuw, F.A.A.M. (1989). An intercomparison of long-term atmospheric transport models; the budgets of acidifying species for the Netherlands. *Atmospheric Environment*, **23**, 1893-1909.

Desbarats, A.J., Logan, C.E., Hinton, M.J., and Sharpe, D.R. (2002). On the kriging of water table elevations using collateral information from a digital elevation model. *Journal of Hydrology*, **255**(1-4), 25-38.

Diggle, P.J., Tawn, J.A., and Moyeed, R.A. (1998). Model-based geostatistics (with discussion). *Journal of the Royal Statistical Society C*, **47**, 299-350.

Diggle, P.J. and Ribeiro Jr., P.J. (2002). Bayesian Inference in Gaussian Model-based Geostatistics. *Geographical and Environmental Modelling,* **6**(2), 129-146.

Dockery, D.W., Pope III, C.A., Xu, X., Spengler, J.D., Ware, J. H., Fay, M.E., Ferris, B.G., and Speizer, F.E. (1993). An association between air pollution and mortality in six U.S. cities. *The New England Journal of Medicine,* **329**, 1753-1759.

Drogue, G., Humbert J., Deraisme J., Mahr, N., and Freslon, N. (2002). A statistical-topographic model using an omnidirectional parameterization of the relief for mapping orographic rainfall. *International Journal of Climatology*, **22**(5), 599-613.

EC (1999). Council Directive 99/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen in the air. *EU Publication nr. L163/41.*

EC (2002). Directive 2002/3/EC of the European Parliament and of the Council of 12 February 2002 relating to ozone in ambient air. *Official Journal L 067, 09/03/2002.*

EEA (1998). Europe's environment, The second assessment. European Environment Agency. ISBN 92-828-3351-8, Office for Official Publications of the European Communities, Luxembourg. Elsevier Science Ltd. Kidlington, UK, 94-108.

EEA (2005). Environment and health, EEA report 10/2005, Copenhagen.

Eerens, H.C., Sliggers, C.J., Van den Hout, K.D., 1993. The CAR model: the Dutch method to determine city street air quality. *Atmospheric Environment,* **27**B, 389-399.

EPA (1998). National Air Quality and Emissions Trends. *Report 1997*, U.S. Environmental Protection Agency, Triangle Park N.C., America.

ETC-ACC (2005). AirBase air quality information system. European Topic Centre on Air and Climate Change, Bilthoven, The Netherlands. *http://air-climate.eionet.eu.int/databases/airbase/index_html*

Feister, U. and Balzer, K. (1991). Surface ozone and Meteorological Predictors on a sub-regional scale. *Atmospheric Environment,* **25**A, 1781-1790.

Fisher, B. (2003). Fuzzy environmental decision-making: applications to air pollution. *Atmospheric Environment,* **37**, 1865-1877.

Folkert, R.J.M., Eerens, H.C., Odijk, M., Van Breugel, P.B., and Van Bree, L. (2002). Attainment of EU NO_2 standards in the Netherlands. Implementation of the 1st EU daughter directive. *RIVM report 725601009*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Fuentes, M. and Raftery, A.E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics,* **61**, 36-45.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004). *Bayesian data analysis - 2nd ed*. Chapman and Hall/CRC, Boca Raton FL.

Genikhovich, E., Filatova, E. and Ziv, A. (2002). A method for mapping the air pollution in cities with the combined use of measured and calculated concentrations. *International Journal of Environment and Pollution*, **18**(1), 56-63.

Gijsen, A., and Seebregts, A.J. (2005). Uncertainties in the reference projections, supplementary on report 'Reference projections energy and emissions 2005 – 2020'. *RIVM report 773001032*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996). *Markov chain Monte Carlo in practice*. Chapman & Hall, London.

Gilks, W.R., Thomas, A., and Spiegelhalter, D.J. (1994). A language and program for complex Bayesian modelling. *The Statistician* **43**, 169-78.

Goovaerts, P. (1999). Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena*, **34**(3-4), 227-242.

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, **228**(1-2), 113-129.

Gregg, J.W., Jones, C.G., and Dawson, E. (2003). Urbanization effects on tree growth in the vicinity of New York City. *Nature*, **424**, 83-187.

Grimes, D.I.F., Pardo-Igúzquiza, E., and Bonifacio, R. (1999). Optimal areal rainfall estimation using raingauges and satellite data. *Journal of Hydrology*, **222**(1-4), 93-108.

Guttorp, P., Meiring, W., and Sampson, P. (1994). A space-time analysis of ground-level ozone data. *Environmetrics*, **5**, 241-254.

Handcock, M.S. and Stein, M.L. (1993). A Bayesian Analysis of Kriging. *Technometrics*, **35**(4), 403-410.

Harvey, A.C. (1989). *Forecasting Structural Times Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Held, L., Nata, I., Fenton, S.E., Rue, H., and Becker, N. (2005). Towards joint disease mapping. *Statistical Methods in Medical Research,* **14**, 1-22.

Hengl, T., Heuvelink, G.B.M., and Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, **120**, 75-93.

Hitzenberger, R., Berner, A., Galambos, Z., Maenhaut, W., Cafmeyer, J., Schwarz, J., Muller, K., Spindler, G., Wieprecht, W., Acker, K., Hillamo, R., and Makela, T. (2004). Intercomparison of methods to measure the mass concentration of the atmospheric aerosol during INTERCOMP2000 - influence of instrumentation and size cuts. *Atmospheric Enviroment,* **38**, 6467-6476.

Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., and Van den Brandt, P.A. (2002). Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet*, **360**(9341), 1203-1209.

Hudson, G. and Wackernagel, H. (1994) Mapping temperature using kriging with external drift: theory and an example from Scotland. *International Journal of Climatology*, **14**(1), 77-91.

Huerta, G., Sansó, B., and Stroud, J.R. (2004). A spatiotemporal model for Mexico City ozone levels. *Applied Statistics*, **53**(2), 231-248.

Hutchison, K.D. (2003). Applications of MODIS satellite data and products for monitoring air quality in the state of Texas. *Atmospheric Environment*, **37**, 2403-2412.

Hutchison, K.D., Smith, S. and Faruqui, S. (2004). The use of MODIS data and aerosol products for air quality prediction. *Atmospheric Environment*, **38**, 5057–5070.

Hutchison, K.D., Smith, S. and Faruqui, S.J. (2005). Correlating MODIS aerosol optical thickness data with ground-based PM2.5 observations across Texas for use in a real-time air quality prediction system. *Atmospheric Environment*, **39**(37), 7190-7203.

Ihaka, R. and Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, **5**, 299-314.

IPCC (2001). Climate Change 2001: The scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge.

Janssen, A.J., Heuberger, P.S.C., and Sanders, R. (1994). UNCSAM: a tool for automating sensitivity and uncertainty analysis. *Environmental Software*, **9**, 1-11.

Jazwinski, A.H. (1970). Stochastic Processes and Filtering Theory. *Mathematics in Science and Engineering* 64, Academic Press, New York.

Jonson, J.E., Sundet, J.K., and Tarrason, L. (2001). Model calculations of present and future levels of ozone and ozone precursors with a global and a regional model. *Atmospheric Environment*, **35**(3), 525-537.

Kaufman, Y.J. and Tanré, D. (1998). Algorithm for remote sensing of tropospheric aerosol from MODIS. *Product ID MOD04*.

Kaufman, Y.J., Tanré, D., Remer, L.A., Vermote, E., Chu, A., and Holben, B.N. (1997). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *Journal of Geophysical Research,* **102**, 17051-17068.

Kaufman, Y.J., Holben, B.N., Tanré, D., Slutsker, I., Smirnov, A., and Eck, T.F. (2000). Will aerosol measurements from Terra and Aqua polar orbiting satellites represent the daily aerosol abundance and properties? *Geophysical Research Letters*, **27**, 3861– 3864.

Kitanidis, P.K. and Shen, K.F. (1996). Geostatistical estimation of chemical concentration. *Advances in Water Resources*, **19**(6), 369-378.

Koelemeijer, R.B.A., Homan, C.D. and Matthijsen, J. (2005) Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. *Atmospheric Environment* (submitted).

Kyriakidis, P.C. and Journel, A.G. (1999). Geostatistical Space–time models: a review. *Mathematical Geology*, **31**(6), 651-684.

Lajaunie, C. and Wackernagel, H. (2000). Geostatistical Approaches to Change of Support Problems, Theoretical Framework. *Technical Report N–30/01/G*, ENSMP - ARMINES, Centre de Géostatistique, Fontainebleau, France.

Lehmann, E.L. (1975). *Nonparametrics: statistical methods based on ranks*. Holden-Day, San Francisco.

Lenschow, P., Abraham, H. J., Kutzner, K., Lutz, M., Preuß, J.D., and Reichenbächer, W. (2001). Some ideas about the sources of PM10. *Atmospheric Environment*, **35**, 23-33.

Lunn, D.J., Thomas, A., Best, N.G., and Spiegelhalter, D.J. (2000). WinBUGS - a Bayesian modelling framework: concepts, structure and extensibility. *Statistics and Computing*, **10**(4), 321-333.

McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models, 2nd edition*. Chapman and Hall, London.

McHugh, C.A., Carruthers, D.J., and Edmunds, H.A. (1997). ADMS-Urban: a model of traffic, domestic and industrial pollution. *International Journal of Environment and Pollution,* **8**(3-6), 666-674.

Meinhold, R.J. and Singpurwalla, N.D. (1983). Understanding the Kalman Filter. *The American Statistician*, **37**(2), 123-127.

MNP (2005). Background concentrations maps of air pollutants in the Netherlands, http://www.mnp.nl/mnc/i-nl-0506.html.

Monestiez, P., Courault D., Allard D., and Ruget F. (2001). Spatial interpolation of air temperature using environmental context: Application to a crop model. *Environmental and Ecological Statistics*, **8**(4), 297-309.

Papaspilliopoulus, O., Roberts, G.O., and Skold, M. (2003). Non-centered parameterizations for hierarchical models and data augmentation. In: Bernardo JM, Bayarri S, Dawid JO, Heckerman D, Smith AFM, West M (eds) *Bayesian statistics* 7. Oxford University Press.

Pardo-Igúzquiza, E. (1998). Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography. *International Journal of Climatology*, **18**(9), 1031-1047.

Pauly M. and Drueke, M. (1996). Mesoscale spatial modelling of ozone immissions. An application of geostatistical methods using a digital elevation model. *Gefahrstoffe reinhaltung der luft*, **56**(6), 225-230.

Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-effects models in S and S-PLUS*. Springer, New York.

Pope III, C.A., Dockery, D.W., and Schwartz, J. (1995). Review of epidemiological evidence of health effects of particulate air pollution. *Inhalation Toxicology*, **7**, 1–18.

Remer, L.A., Kaufman, Y. J. , Tanré, D., Mattoo, S., Chu, D.A., Martins, J.V., Li, R-R., Ichoku, C., Levy, R. C., Kleidman, R.G., Eck, T. F., Vermote, E., and Holben, B.N. (2005). The MODIS Aerosol Algorithm, Products and Validation. *Journal of the Atmospheric Sciences*, **62**, 947-973.

Ribeiro Jr., P.J. and Diggle, P.J. (1999). Bayesian inference in Gaussian model-based geostatistics. *Technical report ST-99-08*, Dept. of Mathematics and Statistics, Lancaster University, Lancaster, United Kingdom.

Ribeiro Jr., P.J. and Diggle, P.J. (2001). GeoR: A package for geostatistical analysis. *R-NEWS*, **1**(2), 15-18.

Putaud, J.P, Raes, F., *et al.* (2004). A European aerosol phenomenology - 2: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe, *Atmospheric Environment*, **38**, 2579-2595.

RIVM (1999). Measuring, modelling and uncertainty. *RIVM report 408129005*. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

Schaap, M., Denier Van der Gon, H.A.C., Dentener, F.J., Visschedijk, A.J.H., Van Loon, M., ten Brink, H.M., Putaud, J.P., Guillaume, B., Ciousse, C. and Builtjes, P.J.H. (2004). Anthropogenic black carbon and fine aerosol distribution over Europe. *Journal of Geophysical Research*, **109**, D18207.

Schaap, M., Roemer, M., Sauter, F.J., Boersen, G.A.C., Timmermans, R.M.A., and Builtjes, P.J.H. (2005). LOTOS-EUROS Documentation, *TNO report B&O 2005/297*, TNO, Apeldoorn, the Netherlands.

Schaap, M., Timmermans, R.M.A., Sauter, F.J., Roemer, M., Velders, G.J.M., Boersen, G.A.C., Beck, J.P., and Builtjes, P.J.H. (2005). The LOTOS-EUROS model: description, validation and latest developments. *International Journal of Environment and Pollution* (accepted).

Seinfeld, J.H., and Pandis, S.N. (1998). *Atmospheric Chemistry and Physics: from air pollution to climate change*. Wiley, New York.

Shively, T.S. (1991). An analysis of the trend in ground-level ozone using non-homogeneous Poisson processes. *Atmospheric Environment*, **25**B, 387-395.

Smith, R.L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, **4**, 367-393.

Snepvangers, J.J.J.C., Heuvelink, G.B.M., and Huisman, J.A. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*, **112**(3-4), 253-271.

UNECE (1996). Critical levels for ozone in Europe: test and finalising the concept. In: Kärenlampi L, Skärby L (eds) UN-ECE workshop report. University of Kuopio, Finland.

Van Aalst, R., Edwards, L., Pulles, T., De Saeger, E., Tombrou, M., and Tønnesen, D. (1999). Guidance report on preliminary assessment under EC air quality directives. *Technical report No 11*. European Environment Agency, Copenhagen, Denmark.

Van Asselt, M,B,A,. Langendonck, R., Van Asten, F., Van der Giessen, A., Janssen P.H.M., Heuberger, P.S.C., and Geuskens, I. (2001). Uncertainty and RIVM's Environmental Outlooks: Documenting a learning process. *RIVM report 550002001*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Van de Kassteele, J. and Stein A. (2005). A model for external drift kriging with uncertain covariates applied to air quality measurements and dispersion model output. *Environmetrics* (in press).

Van de Kassteele, J., Stein, A., Dekkers., A.L.M., and Velders, G.J.M. (2005). External drift kriging of NO_x concentrations with dispersion model output in a reduced air quality monitoring network. *Environmental and Ecological Statistics* (under review).

Van Dril, A.W.N. and Elzenga, H.E. (2005). Reference projections energy and emissions 2005 – 2020. *RIVM report 773001031*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Van Elzakker, B.G. (2001). Monitoring activities in the Dutch National Air Quality Monitoring Network in 2000 and 2001. *RIVM report 723101055*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands, *http://www.lml.rivm.nl/*

Van Jaarsveld, J.A. (1991). An Operational atmospheric transport model for Priority Substances; specification and instructions for use. *RIVM report 222501002*, National Institute of Public Health and the Environment, Bilthoven, Netherlands.

Van Jaarsveld, J.A. (1995). Modelling the long-term atmospheric behaviour of pollutants on various spatial scales. Ph.D. thesis, Utrecht University, the Netherlands.

Van Jaarsveld, J.A. and De Leeuw, F.A.A.M. (1993). OPS: An operational atmospheric transport model for priority substances. *Environmental Software*, **8**, 93-100.

Van Jaarsveld, J.A., 2004. The Operational Priority Substances model. *RIVM report 500045001*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands, *http://www.mnp.nl/ops/*

Van Loon, M., Roemer, M. and Builtjes, P. (2004). Model intercomparison in the framework of the review of the Unified EMEP model. *TNO report R2004/282*, Apeldoorn, The Netherlands.

Van Oorschot, M.M.P., Kraan, B.C.P., Van den Brink, R.M.M., Janssen, P.H.M., and Cooke, R.M. (2003). Uncertainty analysis for NO_x emissions from Dutch passenger cars in 1998. Applying structured expert elicitation and distinguishing different types of uncertainty. *RIVM report 550002004*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Van Soest, F., Stein, A., Dekkers, A.L.M., and Van Duijvenboden, W. (2001). Development of a method for a quantitative evaluation of region oriented policy. *Journal of Environmental Management*, **61**, 215-225.

Van Velze, K., Diederen, H.S.M.A., Ter Maat, H.W., and Van Bree, L. (2000). An outline of air quality problems in the Netherlands. *RIVM report 725601003*. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Wackernagel, H. (2003). *Multivariate Geostatistics*. Springer, Berlin.

WHO (1996). Update and revision of the WHO air quality guidelines for Europe. Classical air pollutants; ozone and other photochemical oxidants. European Centre for Environment and Health, Bilthoven, The Netherlands.

WHO (2003). Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide. *Report on a WHO Working Group EUR/03/5042688*, Bonn, Germany.

Wikle, C.K., Berliner, L.M., and Cressie, N. (1998). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, **5**, 117-154.

Summary

This thesis handles statistical mapping of air quality data. Policy makers require more and more detailed air quality information to take measures to improve air quality. Besides, researchers need detailed air quality information to assess health effects. Accurate and spatially highly resolved maps of air pollution levels form a basis. Since policy makers and researchers tend to focus more and more on uncertainties as well, the question is how precise these concentration maps are.

To base concentration maps on measurements of air quality only, every km^2 should be monitored. Measurements, however, are only taken at a limited number of locations, so between the monitoring locations relevant information will be missing or can only be predicted, i.e. interpolated, leading to uncertainty in the map. Furthermore, no information about the physical and chemical processes about the concerned component is taken into account. On the other hand, concentration maps can also be based on physical and chemical processes modeling of components only. This model output covers the full domain on a fine-mazed grid. All dispersion models are imperfect however, which may lead to biased output and uncertainties.

A combination of the two approaches always results into more detailed and more accurate maps. In this thesis this is done by means of a geostatistical approach: kriging with external drift (KED). KED allows mapping of a primary variable that is accurate and precise but only available at a limited number of locations, and a secondary variable that covers the full domain on a fine-mazed grid but is less accurate.

First, we focus on the use of atmospheric dispersion model output as secondary information source to compensate for the loss of spatial precision caused by a reduction in the Dutch air quality monitoring network in the mid-nineteen eighties. We compare KED with universal kriging. The impact of several parameter estimation and spatial interpolation methods, the number of observations and configuration of the network on uncertainty are quantified by cross-validation. With KED, more accurate and precise predictions are obtained where observations were sparse. However, the dispersion model output in this context was considered to be deterministic, i.e. without uncertainties, so the geostatistical model must be extended.

We present a method, error-in-variable KED, which combines uncertain air quality measurements with uncertain secondary information from the atmospheric dispersion model. The new method combines KED and a measurement error model, and uses Bayesian techniques for inference. The method is flexible for assigning different error variances to both the primary information and secondary information

at each location. We address actual NO_2 data collected at an urban and a rural site in the Netherlands. Uncertainty assessments in terms of exceeding air quality standards are given.

The error-in-variable KED procedure is further extended with a time component to assess future local NO_2 concentrations near Rotterdam for the year 2010, focusing on uncertainties and exceedances of European air quality standards. The background concentration is determined by the extended error-in-variable KED. A local traffic contribution is added based on a local generic dispersion model with use of an emission scenario for 2010. This results in maps showing local NO_2 concentrations, upper and lower limits, and probabilities of exceeding the air quality standard. The probabilistic measures are calculated in numbers and translated into words for easier communication to policy makers.

Finally, the use of two secondary information sources is explored to map particulate matter (PM10) over Western Europe. It is almost impossible to get a consistent overview of PM10 concentrations based solely on ground based measurements because of differences between countries regarding monitoring methods used and monitoring station surroundings. We illustrate the use of statistical techniques to standardize the ground based measurements of PM10 and interpolate these standardized concentrations by combining them uncertain secondary information from a chemical transport model and from MODIS satellite observations of aerosol optical thickness. The secondary variables contain different information and a combination of both gives the most accurate and precise predictions and should therefore be preferred.

Samenvatting

Dit proefschrift gaat over statistische methoden om luchtkwaliteitsdata te karteren. Beleidsmakers hebben steeds meer gedetailleerde luchtinformatie nodig om maatregelen te kunnen treffen voor de verbetering van de luchtkwaliteit. Daarnaast hebben onderzoekers gedetailleerde luchtkwaliteitsinformatie nodig om gezondheidseffecten van luchtvervuiling te kunnen berekenen. Nauwkeurige kaarten met een hoge ruimtelijke resolutie vormen daarvoor de basis. Omdat beleidsmakers en onderzoekers zich ook steeds meer op onzekerheden richten, is de vraag hoe nauwkeurig deze concentratiekaarten zijn.

Als de concentratiekaarten alleen maar gebaseerd zouden worden op luchtkwaliteitsmetingen, dan zou er op elke vierkante kilometer gemeten moeten worden. Omdat de metingen maar op een beperkt aantal locaties gedaan worden, ontbreekt relevante informatie tussen de meetstations of kan alleen maar voorspeld of geïnterpoleerd worden, wat tot onzekerheden in de kaart leidt. Bovendien wordt er geen informatie meegenomen over de fysische en chemische processen van de betreffende component. Aan de andere kant kunnen concentratiekaarten gebaseerd worden op alleen maar fysische en chemische modellering van componenten. Deze modeluitvoer bedekt het hele gebied op een fijnmazig grid. Alle verspreidingsmodellen zijn echter imperfect, wat tot afwijkende resultaten en onzekerheden kan leiden.

Een combinatie van beide aanpakken leidt altijd tot gedetailleerdere en nauwkeurigere kaarten. Dit wordt in dit proefschrift gedaan door middel van een geostatistische aanpak: kriging met externe drift (KED). KED laat kartering toe van een primaire variabele die exact en nauwkeurig is, maar alleen bekend is op een beperkt aantal locaties en een secondaire variabele die minder nauwkeurig is, maar wel op een fijnmazig grid het hele gebied bedekt.

We kijken eerst naar uitvoer van een atmosferisch verspreidingsmodel als secondaire informatiebron om het verlies van ruimtelijke nauwkeurigheid te compenseren als gevolg van een uitdunning van het Landelijk Meetnet Luchtkwaliteit halverwege de jaren tachtig. We vergelijken KED met univeral kriging. Door middel van kruisvalidatie wordt de invloed van verschillende parameterschatting- en ruimtelijke interpolatiemethoden, het aantal waarnemingen en de configuratie van het meetnet op de onzekerheid gekwantificeerd. Door KED te gebruiken worden exactere en nauwkeurigere predicties verkregen wanneer er weinig waarnemingen zijn. Echter, de uitvoer van het verspreidingsmodel werd als deterministisch gezien, zonder onzekerheden, dus moet dit eerste geostatistische model uitgebreid worden.

We presenteren een methode, error-in-variable KED, die onzekere luchtkwaliteitsmetingen combineert met onzekere secondaire informatie van het verspreidingsmodel. De nieuwe methode combineert KED met een meetfoutmodel en gebruikt Bayesiaanse inferentietechnieken. De methode is flexibel genoeg om verschillende varianties aan zowel de primaire als secondaire informatie toe te kennen op elke locatie. We richten ons op actuele NO_2 data op een stedelijke- en plattelandslocatie in Nederland. Onzekerheidsanalyses in termen van overschrijding van luchtkwaliteitsnormen wordt gegeven.

De error-in-variable KED procedure wordt verder uitgebreid met een tijdscomponent om toekomstige locale NO_2 concentraties nabij Rotterdam te bepalen voor het jaar 2010, kijkend naar onzekerheden and overschrijdingen van Europese luchtkwaliteitsstandaarden. De achtergrondconcentratie wordt bepaald aan de hand van de uitgebreide error-in-variable KED. Aan de hand van een lokaal generiek verspreidingsmodel met het gebruik van een emissiescenario voor 2010 wordt er een locale verkeersbijdrage bij opgeteld. Dit resulteert in kaarten van locale NO_2 concentraties, onder- en bovengrenzen en overschrijdingskansen van de luchtkwaliteitsnorm. De probabilistische maten worden berekend en vervolgens vertaald in woorden om communicatie naar beleidsmakers te versoepelen.

Ten slotte wordt het gebruik van twee secondaire informatiebronnen verkend om fijn stof (PM10) boven West Europa in kaart te brengen. Het is vrijwel onmogelijk om een consistent beeld te krijgen van PM10 dat alleen maar gebaseerd is op metingen van een meetnet, omdat meetmethoden tussen landen verschillen, evenals de omgeving van de meetstations. We illustreren het gebruik van statistische methoden om meetnetgegevens van PM10 te standaardiseren en om deze gestandaardiseerde concentraties te interpoleren door ze te combineren met onzekere secondaire informatie van een chemisch transportmodel en MODIS satellietobservaties van aërosol optische dikte. De secondaire variabelen bevatten verschillende soorten informatie en een combinatie van beide geeft de meest exacte en nauwkeurige voorspellingen en dient daarom geprefereerd te worden.

Curriculum Vitae

Jan van de Kassteele werd op 13 december 1976 geboren te Zierikzee. In 1994 behaalde hij zijn HAVO diploma en in 1996 zijn VWO diploma aan de Scholengemeenschap Prof. Zeeman te Zierikzee. Zijn "hobby" om te rekenen aan de levenloze natuur was toen al duidelijk. Tot zijn 19e woonde hij in Noordgouwe, maar hij ruilde dit dorp in voor Wageningen door Bodem, Water en Atmosfeer te gaan studeren aan de Wageningen Universiteit. Daar specialiseerde hij zich in de Meteorologie. Afstudeervakken en stage deed hij bij de leerstoelgroep Meteorologie en Luchtkwaliteit en het KNMI, met als hoofdthema land-atmosfeer wisselwerkingen. In november 2001 rondde hij zijn studie af en begon hij aan zijn promotie-onderzoek bij de Leerstoelgroep Wiskundige en Statistische methoden - Biometris. Het onderwerp was het ontwikkelen en toepassen van statistische methoden om luchtkwaliteit boven Nederland in kaart te brengen. Hij deed dit in samenwerking met het Rijksinstituut voor de Volksgezondheid en Milieu (RIVM) en Milieu- en Natuurplanbureau (MNP) in Bilthoven. Resultaten waren onder andere een vijftal artikelen en presentaties op verschillende internationale congressen. Het werk wordt operationeel gemaakt en is al voor een deel toegepast in luchtkwaliteitberekeningen van het MNP. Per 1 maart 2006 werkt hij als statisticus bij het RIVM bij het Centrum voor Informatisering en Methodologische Advisering in Bilthoven.

List of publications

Peer-reviewed publications:

J. van de Kassteele, A.L.M. Dekkers, A. Stein and G.J.M. Velders (2005). Model-based geostatistical interpolation of the annual number of ozone exceedance days in the Netherlands. *Stochastic Environmental Research and Risk Assessment*, **19**(3), 173-183.

J. van de Kassteele and A. Stein (2005). A model for external drift kriging with uncertain covariates applied to air quality measurements and dispersion model output. *Environmetrics* (in press, published online DOI 10.1002/env.771).

Peer-reviewed publications in press:

J. van de Kassteele, A. Stein, A.L.M. Dekkers, and G.J.M. Velders (2005). External drift kriging of NO_x concentrations with dispersion model output in a reduced air quality monitoring network. *Environmental and Ecological Statistics* (under review).

J. van de Kassteele and G.J.M. Velders (2006). Uncertainty assessment of local NO_2 concentrations derived from error-in-variable external drift kriging and its relationship to the 2010 air quality standard. *Atmospheric Environment* (in press).

J. van de Kassteele, R.B.A. Koelemeijer, A.L.M. Dekkers, M. Schaap, C.D. Homan and A. Stein (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* (submitted).

PE&RC PhD Education Statement Form

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 22 credits (= 32 ECTS = 22 weeks of activities).

Review of Literature (4 credits)

- Spatial interpolation of air quality data (2002)

Post-Graduate Courses (3.75 credits)

- Uncertainty modelling and analysis (2002)
- Estimation of human impact on the environment (2002)
- Bayesian inference in Markov random fields (2002)
- Bayesian statistics (2003)
- Environmental spatial statistics and their applications (2003)
- Statistics of spatiotemporal systems (2004)

Deficiency, Refresh, Brush-up and General Courses (2.25 credits)

- Career orientation and planning (2004)
- Ethics of scientific publishing (2004)
- Time planning and project management (2005)

PhD Discussion Groups (4 credits)

- Statistics, mathematics and modelling in production ecology and resource conservation (2002-2005)

PE&RC Annual Meetings, Seminars and Introduction Days (1.5 credits)

- PE&RC annual meeting "Ethics in Science" (2002)
- PE&RC annual meeting "Global Climate Change and Biodiversity" (2003)
- PE&RC annual meeting "Biological Disasters" (2004)
- PE&RC weekend (2005)
- PE&RC annual meeting "The Truth of Science" (2005)

International Symposia, Workshops and Conferences (6.5 credits)

- The International Environmetrics Society (TIES) Annual Conference, Genoa, Italy (2002)
- Wageningen Biometry Symposium, Wageningen, the Netherlands (2002)

- ISI International Conference on Environmental Statistics and Health, Santiago de Compostela, Spain (2003)
- Séminaire Européen de Statistique – Statistics (SemStat 2004) for Spatiotemporal Systems, Bernried, Germany (2004)
- Statistics and Models, Wageningen Biometry Symposium, Wageningen, the Netherlands (2004)

Printed by PrintPartners Ipskamp, Nijmegen