

# Quantification and attribution of urban fossil fuel emissions through atmospheric measurements



Ingrid Super

# Quantification and attribution of urban fossil fuel emissions through atmospheric measurements

Ingrid Super

## **Thesis committee**

### **Promotor**

Prof. Dr W. Peters

Personal chair at the Meteorology and Air Quality Group

Wageningen University & Research

Professor Atmospheric Composition Modeling

University of Groningen

### **Co-promotors**

Dr M.K. van der Molen

Associate professor, Meteorology and Air Quality Group

Wageningen University & Research

Dr H.A.C. Denier van der Gon

Senior scientist, Built Environment and Geosciences

Netherlands Organisation for applied scientific research (TNO), Utrecht

### **Other members**

Prof. Dr A.H. Weerts, Wageningen University & Research

Dr G. Broquet, Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France

Dr J. Marshall, Max Planck Institute for Biogeochemistry, Jena, Germany

Prof. Dr H.A.J. Meijer, University of Groningen

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

# Quantification and attribution of urban fossil fuel emissions through atmospheric measurements

Ingrid Super

## **Thesis**

submitted in fulfilment of the requirements for the degree of doctor  
at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Thursday 11 October 2018

at 11 a.m. in the Aula.

Ingrid Super

Quantification and attribution of urban fossil fuel emissions through atmospheric measurements,  
192 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)

With references, with summaries in English and Dutch


ISBN: 978-94-6343-498-0

DOI: <https://doi.org/10.18174/457839>

## Table of contents

Chapter 1	Introduction: The carbon cycle and methods to monitor urban fossil fuel emissions	7
Chapter 2	Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam	25
Chapter 3	A multi-model approach to monitor emissions of CO <sub>2</sub> and CO from an urban–industrial complex	51
Chapter 4	Development of a dynamic emission model to predict high-resolution fossil fuel CO <sub>2</sub> emissions	79
Chapter 5	Optimizing a dynamic fossil fuel CO <sub>2</sub> emission model for an urban area	105
Chapter 6	General discussion on monitoring urban fossil fuel emissions per source sector	137
	Summary	153
	Samenvatting	157
	Appendix A	161
	References	167
	Dankwoord	185
	List of publications	189
	Education certificate	190



A thick, dark blue vertical bar runs along the left edge of the page. To its right, several thin, curved lines in shades of blue and grey sweep upwards and outwards from the bottom left corner, creating an abstract, organic shape.

## Chapter 1

### **Introduction: The carbon cycle and methods to monitor urban fossil fuel emissions**





## 1.1 Background

Carbon dioxide is an important greenhouse gas affecting our climate, making our planet inhabitable for flora and fauna. Carbon is exchanged between the biosphere, lithosphere, atmosphere and hydrosphere in a natural biogeochemical cycle (Berner, 2003). The carbon reservoirs have a turnover time of several years (e.g. the atmosphere) to thousands of years (e.g. the deep ocean). The feedbacks in the carbon cycle ensure that the temperature on our planet remains within certain limits (Riebeek, 2011). In that respect the carbon cycle can be considered to be in equilibrium. However, large and persistent perturbations can distort the equilibrium for a long period of time while the feedbacks strengthen or weaken to restore the equilibrium.

The largest perturbation of the atmospheric carbon content in recent history is the addition of fossil  $\text{CO}_2$  by anthropogenic combustion (Ballantyne et al., 2012). Since the fossil fuel emission is only partly compensated by an increase in terrestrial and oceanic uptake, the amount of  $\text{CO}_2$  in the atmosphere has increased rapidly from about 280 ppm before the industrial revolution to 400 ppm nowadays (IPCC, 2013). The time to restore the equilibrium is estimated to be tens of thousands of years (Hausfather, 2010) under the restriction that no more fossil fuel  $\text{CO}_2$  is added to the atmosphere. The effects of the rise in atmospheric  $\text{CO}_2$  levels will be plentiful, from sea level rise to extreme precipitation events and heat waves (IPCC, 2013).

Therefore, efforts are made to reduce the future impacts of human-induced climate change by limiting anthropogenic emissions from fossil fuel combustion. In 2015, 195 countries agreed on an extensive climate action plan known as the Paris Agreement, including climate mitigation and adaptation efforts, with the aim to limit global warming to 1.5-2.0°C (UNFCCC, 2015). Besides setting climate targets and implementing climate actions all parties involved are obliged to monitor and report their progress. This requires methods to objectively monitor (trends in) emissions and the impact of specific policies. Moreover, parties are encouraged to collaborate and exchange information about best practices and failures related to emission reduction policies. As such, emission monitoring is a key factor for successfully reducing emissions worldwide.



Figure 1.1: Schematic illustration of the dispersion of an air pollutant emitted from a stack.

An emission is simply the amount of CO<sub>2</sub> that is released to the atmosphere. A simple method to estimate fossil fuel emissions is to calculate it from fossil fuel consumption, based on the amount of fuel that is consumed and the amount of carbon each fuel type contains (emission inventory). Emissions can also be measured directly with CO<sub>2</sub> flux measurements. There are several instruments capable of measuring fluxes, although they usually only cover a small area. Therefore, often a more indirect method is used, namely observing changes in the atmospheric CO<sub>2</sub> levels that are the result of CO<sub>2</sub> fluxes. The relationship between a stack emission and the atmospheric concentration is depicted in Fig. 1.1. The CO<sub>2</sub> is released to the atmosphere where a plume forms, which becomes wider and higher when it is transported further away from the stack by the wind. This process is called dispersion, which causes the amount of CO<sub>2</sub> per volume of air (concentration) to decrease (lighter grey in Fig. 1.1). To connect measurements in the atmosphere to fluxes that occurred elsewhere at the surface atmospheric transport models can be used. A transport model takes prescribed CO<sub>2</sub> fluxes and calculates the transport and dilution of the emitted CO<sub>2</sub>, resulting in an atmospheric CO<sub>2</sub> concentration at a measurement site. The prescribed fossil fuel fluxes often come from an emission inventory. In case the transport is calculated correctly, the difference between the modelled and observed atmospheric concentrations can be attributed to an error in the prescribed fluxes and thus the emissions can be improved. Such a combined method is called inverse modelling or data assimilation (DA) (these terms will be used interchangeably in this thesis). It has been applied successfully to constrain biogenic fluxes of CO<sub>2</sub> by combining continental monitoring networks (like the Integrated Carbon Observation System) with regional to global transport models, leading to flux estimates at large spatiotemporal scales (0.5 to 10° and weeks to seasons) (e.g. (Basu et al., 2016; Broquet et al., 2013; Liu and Bowman, 2016; Meesters et al., 2012; Peters et al., 2007; Peters et al., 2010; Ray et al., 2014; Rödenbeck et al., 2009; Tolk et al., 2011; Van der Laan-Luijkx et al., 2015; Van der Laan-Luijkx et al., 2017)). However, most of the anthropogenic CO<sub>2</sub> emissions come from cities and therefore monitoring should be done at much smaller scales. Therefore, inverse modelling has recently been applied to urban areas as well (Bréon et al., 2015; Brioude et al., 2012; Brioude et al., 2013; Lauvaux et al., 2013; Lauvaux et al., 2016). This introduces some additional challenges as much smaller scales (1 km and sub-hour) need to be resolved.

An additional challenge related to monitoring the effect of emission reduction policies is that most policies focus on specific source sectors, such as the industry or road traffic. A trend in overall fossil fuel CO<sub>2</sub> fluxes can be the sum of different processes, making the interpretation of trends in observed concentrations complicated. Whereas emission inventories often estimate emissions per source sector, all sectors emit the same CO<sub>2</sub> and no distinction can be made based on observed CO<sub>2</sub> concentrations. However, all source sectors also emit other trace gasses, like CO or NO<sub>x</sub>, in a ratio that is specific for that sector

and measuring atmospheric concentrations of co-emitted gases could thus help to attribute atmospheric CO<sub>2</sub> to specific source sectors.

In this thesis I explore the possibility to monitor (trends in) emissions at the (sub-)urban scale. A combination of (semi-)urban observation networks and regional atmospheric transport models is applied to investigate the complex relationship between urban anthropogenic emissions and observed CO<sub>2</sub> mixing ratios at (sub)urban measurement sites. This requires resolving small scale processes that affect the observed mixing ratios, which is challenging with current state-of-the-art transport models. Moreover, additional trace gases are monitored and modelled to support source attribution. Besides quantifying the CO<sub>2</sub> emissions themselves, I also try to attribute atmospheric CO<sub>2</sub> signals to specific source sectors. This step is especially useful when operationalising this method for actual emission reporting and verification.

The aim is to increase our understanding of how observed CO<sub>2</sub> mixing ratios can provide information on multiple source sector emissions and of the monitoring requirements to obtain this source sector specific information. With this knowledge, an inverse modelling framework can be developed that is suited to urban scales and to attribute emissions to specific source sectors. In the next sections, we shortly describe the carbon cycle and more specifically the role of (urban) fossil fuel emissions therein, how fossil fuel emissions are currently estimated and the limitations of these methods, and finally a range of techniques that can be used to monitor urban fossil fuel emissions.

## 1.2 The carbon cycle

The carbon cycle is very important for our planet. Although the Earth's climate can vary over time due to perturbations, moving from ice ages to interglacial periods, the Earth's temperature always remains within certain limits in which life is preserved. The temperature is mainly determined by the radiative balance, in which CO<sub>2</sub> also plays a role. Due to its characteristics, atmospheric CO<sub>2</sub> traps longwave radiation and increases the temperature at the Earth's surface (Lacis et al., 2010). This is the natural greenhouse effect. Without the greenhouse effect Earth would be too cold for human life to exist.

The exchange of carbon in the carbon cycle is depicted in Fig. 1.2. Carbon is exchanged between reservoirs (arrows in Fig. 1.2). Some exchange processes are faster than others, depending on the gradient between the reservoirs and the ease with which the exchange can take place. For example, the transport of atmospheric CO<sub>2</sub> to the deep ocean is slow, because the ocean is layered and transport between layers is relatively difficult. The pace of this exchange determines how long carbon is stored in each reservoir. In its pre-industrial state, the carbon cycle is in quasi steady-state with small changes in the reservoirs over time-scales of a century. The exchange processes ensure that the carbon is distributed over the reservoirs and that a disturbance in one of the exchanges or reservoirs is counteracted by another exchange mechanism.

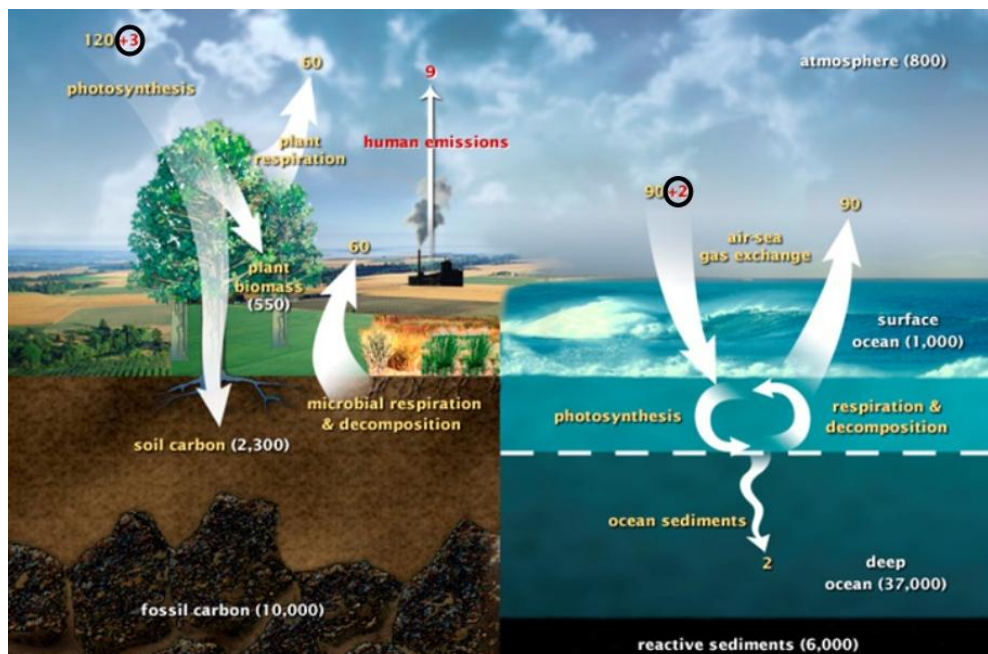
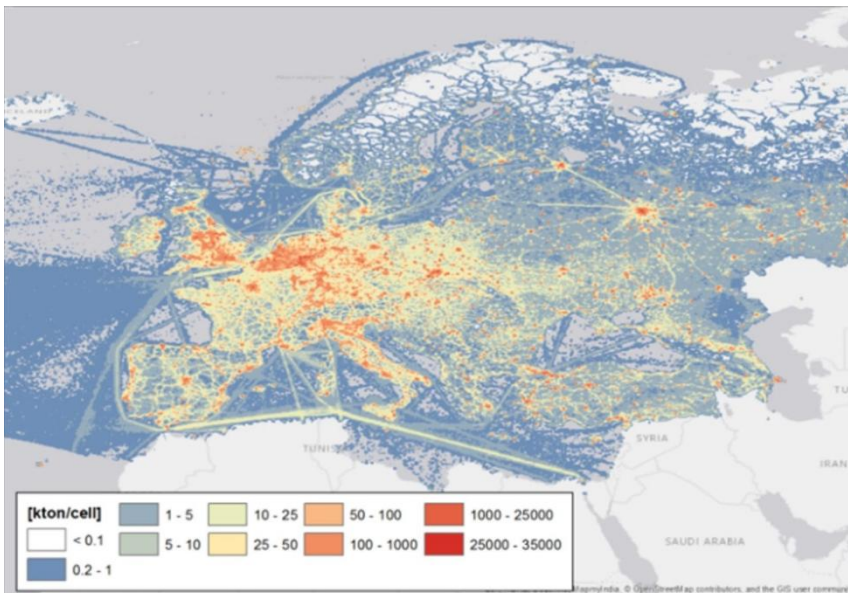


Figure 1.2: Schematic representation of the carbon cycle, where the numbers between brackets indicate the storage size of the carbon reservoirs and the numbers with the arrows indicate carbon fluxes (in Petagrams (Pg) of carbon per year). The circled +3 and +2 indicate changes in the fluxes due to human emissions. Figure adapted from <https://earthobservatory.nasa.gov/Features/CarbonCycle>.

Naturally, fossil carbon is transported to the atmosphere by volcanic activity and rock weathering over millions of years (Lacis et al., 2010). But due to human activity, fossil carbon is now released to the atmosphere at a much higher pace, perturbing the carbon cycle (Berner, 2003). Currently, about 9 Petagrams ( $\text{Pg} = 10^{12} \text{ kg}$ ) of carbon is emitted to the atmosphere in one year by human activities (Fig. 1.2), of which 3 Pg is taken up additionally by the biosphere and 2 Pg by surface waters (Le Quéré et al., 2018). The remainder is accumulated in the atmosphere where it affects the radiation balance. The deep ocean and the soil could potentially store a lot of this carbon, but given the slow transfer of atmospheric carbon to these large carbon reservoirs this process will take a long time. Anthropogenic emissions are therefore an important factor in forcing our climate at decadal to centennial timescales. Due to the increased level of  $\text{CO}_2$  in our atmosphere, the temperature increases. But the higher  $\text{CO}_2$  level will also affect many other aspects of our climate and (indirectly) other physical and biological processes. One example is that rising temperatures cause melting of ice caps and glaciers, that affects oceanic transport (IPCC, 2013). Another example is the acidification of the ocean waters due to the increased uptake of  $\text{CO}_2$  that poses a threat to marine life that in turn can impact the exchange of  $\text{CO}_2$  between the ocean and atmosphere. Hence, knowledge on the carbon cycle and its perturbations is crucial to understand our future climate.

### 1.3 Fossil fuel emissions

In order to understand the impact of human activities on our climate, fossil fuel emissions need to be quantified. Fossil fuels are used for many human activities, such as industrial processes, transportation, and heating of buildings. Since fossil fuel emissions are so strongly linked with human activities, urbanized areas contribute approximately 70% to the global fossil fuel CO<sub>2</sub> emissions according to the International Energy Agency (IEA, 2008). Therefore, recent monitoring efforts focus on urban areas. The importance of urbanized areas is also clearly visible on a European emission map, where major cities like Paris, Madrid and Moscow stand out (Fig. 1.3). The variability in combustion processes results in a large spatiotemporal variability in emissions at the urban scale, which makes it difficult to quantify the impact of mitigation measures.



**Figure 1.3: Gridded CO<sub>2</sub> fossil fuel emissions from the TNO-CAMS emission inventory. Source: Denier van der Gon et al. (2017).**

Emissions are often estimated from activity data (e.g. total fuel consumption) and emission factors (e.g. the amount of CO<sub>2</sub> emitted per amount of fuel consumed). For developed countries the annual national emissions of CO<sub>2</sub> are relatively well-known with an uncertainty of less than 8% (Nassar et al., 2013). Larger uncertainties of up to 50% exist for countries with a poorly developed infrastructure (Andres et al., 2012). These uncertainties are related to the emissions that are considered in the accounting method, while assumptions in the accounting method (e.g. about which sectors should be included) might result in much larger uncertainties (Ciais et al., 2010).

More detailed knowledge on fossil fuel emissions is often gathered in so-called emission inventories. A wide range of inventories is available, varying in the domain covered,

resolution, amount of detail and approach. Inventory builders can use the annual emissions at the national level and disaggregate them using statistical data (e.g. population density) to reach higher spatiotemporal resolution, a so-called “scale-down” approach. Also temporal downscaling, e.g. from annual to hourly emissions, can be done using statistical data such as traffic counts. These are averaged over several years to make the time profiles applicable for multiple years. As such, these emission inventories contain a larger uncertainty than the annual total emissions due to errors introduced during downscaling (Ciais et al., 2010; Nassar et al., 2013). Examples are the limited representativeness of a proxy used for disaggregation and the allocation of point sources to a gridded emission map (Hogue et al., 2016). These uncertainties increase further when going to higher resolutions.

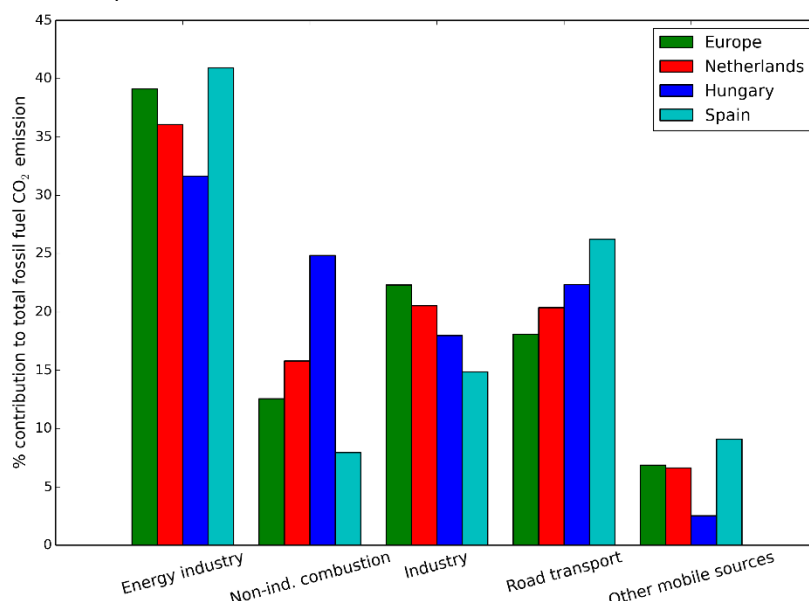
Another method to build an emission inventory is the “scale-up” approach, which makes use of local data such as local activity data and emission factors or reported values from specific industrial/power plants. Although both the “scale-down” and “scale-up” approach can be used on their own, often emission inventories are built with a mixture of both methods. The reason is that not all data are available at the local level and consistency with annual, national emissions is often desired. An advantage of combining both methods is that the “scale-up” data can partly offset the errors in the downscaling. It is therefore mostly relevant for high-resolution inventories that often cover a specific region (a country or even a single city). For example, instead of using large-scale and long-term average traffic counts we can use the specific traffic counts for Rotterdam. This adds more detailed information, although these activity data can also be highly uncertain (Olivier et al., 2009). The uncertainty in high-resolution, local emission inventories is not well-known, as there are no direct measurements to compare the estimates to. However, some studies have used inverse modelling at urban scales and have estimated the total fossil fuel CO<sub>2</sub> emissions for the urban areas to have an uncertainty of 20-50% for a period of several hours to several days (Bréon et al., 2015; Brioude et al., 2012; Lauvaux et al., 2016).

**Table 1.1: Overview of SNAP (Selected Nomenclature for sources of Air Pollution) categories.**

<b>SNAP categories</b>	
<b>1</b>	Combustion in energy and transformation industries
<b>2</b>	Non-industrial combustion plants
<b>3+4</b>	Industrial processes (Combustion in manufacturing industry; Production processes)
<b>5</b>	Extraction and distribution of fossil fuels and geothermal energy
<b>6</b>	Solvent and other product use
<b>7</b>	Road transport
<b>8</b>	Other mobile source and machinery
<b>9</b>	Waste treatment and disposal
<b>10</b>	Agriculture
<b>11</b>	Other sources and sinks

Most emission inventories divide the emissions over different source sectors, for example following the SNAP (Selected Nomenclature for source of Air Pollution) categories listed in Table 1.1. Sometimes these categories are further subdivided, for example by fuel type or vehicle type for road transport. Moreover, some inventories distinguish between area sources and point sources. Point sources are mostly related to industrial processes and emissions come from elevated stacks, like shown in Fig. 1.1, although their height is often not included in the inventories. Area sources have a larger horizontal extent and often emit at ground level.

Figure 1.4 shows the distribution of fossil fuel CO<sub>2</sub> emissions over the different source sectors in 2011 for Europe and some European countries based on an emission inventory from the Monitoring Atmospheric Composition & Climate (MACC) project (Kuenen et al., 2014). Not all sectors distinguished in Table 1.1 are relevant for CO<sub>2</sub> and only the largest contributors are shown in Fig. 1.4. The European fossil fuel emissions are strongly dominated by (energy) industry. Also road transport and non-industrial combustion are significant sources of CO<sub>2</sub>. However, economic activities, climate conditions, and the implementation of new clean technologies can differ per country, which leads to large differences between countries. For example, Fig. 1.4 suggests a relatively small fraction of centralized energy production in Hungary and less heating of buildings (non-industrial combustion) in Spain. Also within a country the spatial variability can be large, because industrial plants are often clustered in industrial areas.



**Figure 1.4:** Relative contribution of several source sectors to the total fossil fuel CO<sub>2</sub> emission in 2011 for Europe, Netherlands, Hungary and Spain. Source: MACC III emission inventory.

While existing emission inventories can provide valuable insight in the emission landscape of a country or urban area, there are also some downsides. Large



inconsistencies can exist between inventories due to large uncertainties in different attribution methods (Denier van der Gon et al., 2012). Moreover, emission reporting takes some time and therefore emission inventories which make use of such reports appear with a lag of several years. In other words, knowledge on current emissions will be available only several years from now. This complicates monitoring of emission trends and the effectiveness of emission reduction policies.

## **1.4 Monitoring of urban fossil fuel emissions**

In the previous sections we have established that there is an urgent need for techniques that can monitor emission trends in detail and in near real-time. Current emission inventories are not suitable for this task as they are too coarse (based on yearly data) and have a lag of several years. Therefore, other techniques are required that make use of higher resolution and more up-to-date information. In this thesis we explore the use of atmospheric measurements in urban fossil fuel monitoring, we identify the advantages and limitations of different network and model configurations, and we explore new opportunities by combining several existing and new techniques. Therefore, the main research question is:

**Main RQ: What are the monitoring requirements to constrain fossil fuel CO<sub>2</sub> emissions from an urban area at source sector level?**

This section discusses relevant monitoring techniques and the related research questions addressed in this thesis.

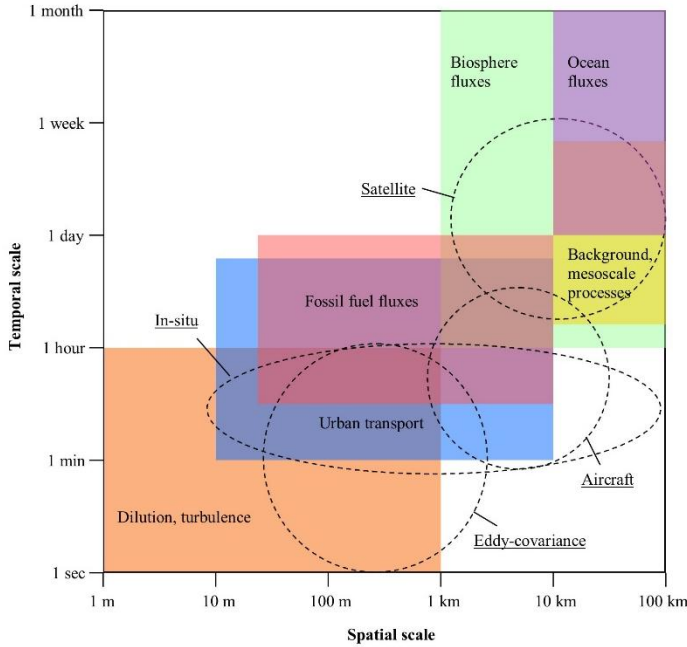
### **1.4.1 Observations**

Observations play an important role in monitoring emissions, as they provide an objective top-down constraint on emissions and they can be available in near real-time. There are several important considerations when choosing an instrument or measurement technique:

- 1) The temporal resolution and scale: how often does the instrument do a measurement and what time frame does it cover.
- 2) The spatial resolution and scale: how large is the area that affects the measurement (its footprint) and how much detail can be separated within that footprint.
- 3) The accuracy/precision of the measurement.
- 4) The measured quantity: what exactly is measured, for example fluxes versus concentrations.

For urban applications the spatial and temporal resolution are very important, because there is a lot of small-scale variability that needs to be resolved. Figure 1.5 gives an overview of processes that are relevant for understanding the urban carbon budget. Their location in the diagram indicates at which spatial and temporal scale they cause variations in atmospheric CO<sub>2</sub> levels. We see that time scales of less than an hour and spatial scales

of 10-100m are of importance for the urban carbon cycle and therefore instruments are needed that can capture this small-scale variability.



**Figure 1.5: The spatial and temporal scale at which different processes affect atmospheric CO<sub>2</sub> concentrations (coloured rectangles). The circles represent the resolution of different types of instruments.**

Different types of instruments are also displayed in the diagram with the scales that they can resolve. Note that the temporal resolution and the covered time frame are not necessarily the same, because high-resolution measurements that are done continuously for several years can also resolve seasonal and yearly variations. Similarly, an instrument with a high spatial resolution often covers a smaller domain. The domain can be extended by using multiple instruments. As such, each method has its pros and cons and therefore all these techniques have been used to monitor urban CO<sub>2</sub> fields and emissions. For example, aircraft measurements can cover an entire urban area or transects up- and downwind of a city (Mays et al., 2009). This gives information on the spatial variability of CO<sub>2</sub> concentrations. However, flights only cover a limited time window. In contrast, eddy-covariance measurements (or flux measurements) can be used continuously (Kleingeld et al., 2017), but such measurements often have a limited footprint (i.e. they are affected by a small area and are thus not representative for an entire urban area). A promising method for the future might also be the use of satellite data (Kort et al., 2012; Silva et al., 2013), although currently there are still some limitations in the resolution and accuracy.

In this thesis we used an alternative method based on in-situ measurements of atmospheric concentrations. Our in-situ measurements are continuous, providing a long time series of concentration measurements at high temporal resolution. Therefore, a wide

range of temporal scales can be resolved. The covered domain is highly dependent on the location of the instrument. In the scientific literature there is an ongoing debate about the best location for in-situ measurements. On the one hand, in-city sites are extremely influenced by local sources (100-1000 m) and therefore provide a lot of information, but on the other hand they have a limited footprint. Therefore, a lot of instruments are needed to cover the entire urban area. Moreover, in-city sites are prone to large model-data mismatches due to the complex transport processes in the city that are not easily resolved with regional transport models (Bréon et al., 2015). On the other hand, measurement sites outside the city are less often impacted by the city, depending on the wind direction (Kort et al., 2013). Yet they potentially have the entire city in their footprint, providing a constraint on the entire area of interest albeit with a lower spatial resolution. Moreover, elevated measurements (i.e. towers) also have a larger footprint and are less affected by local sources.

Although in this thesis the focus is on CO<sub>2</sub> emissions, this work can also be useful to monitor other co-emitted species. These are trace gasses that are emitted during fossil fuel combustion, similarly to CO<sub>2</sub>. Well-known co-emitted species are CO and NO<sub>x</sub> (NO+NO<sub>2</sub>) which are both emitted in relatively high quantities by road traffic and shipping, but also SO<sub>2</sub> is considered in this thesis. Since CO<sub>2</sub> and co-emitted air pollutants result from the same combustion processes the emission reduction policies will also affect air pollution levels. Therefore, it is beneficial to monitor CO<sub>2</sub> and co-emitted species simultaneously to provide information for climate projections, but also for policies aiming to reduce air pollution. Moreover, the presence of co-emitted tracers indicates the presence of fossil fuel combustion so that these tracers can help separate between biogenic and fossil fuel signals of CO<sub>2</sub> (Djuricin et al., 2010; Lopez et al., 2013; Turnbull et al., 2006).

Another advantage of combining information on multiple species is that the co-emitted species can help attribute observed CO<sub>2</sub> to specific source sectors. CO<sub>2</sub> and co-emitted species are emitted in a specific ratio that is dependent on the conditions under which combustion takes place, but also on the technology (e.g. sometimes specific species are filtered from the exhaust). For example, modern power plant combustion is relatively clean and their emission only contains very small levels of co-emitted species. In contrast, road traffic emits a relatively large amount of CO and NO<sub>x</sub>. Monitoring these species simultaneously and looking at the concentration ratios can thus provide information on the dominant combustion processes that took place in the footprint of a measurement. This tracer method has only seen limited application for source attribution (Lindenmaier et al., 2014). The reason is that the emission ratios need to be accurately known, while these can be variable in space and time. We will explore this method in more detail in this thesis.

Besides co-emitted species isotopes can provide valuable information. Isotopes are different variants of an atom that only differ in the number of neutrons. Specific isotopes

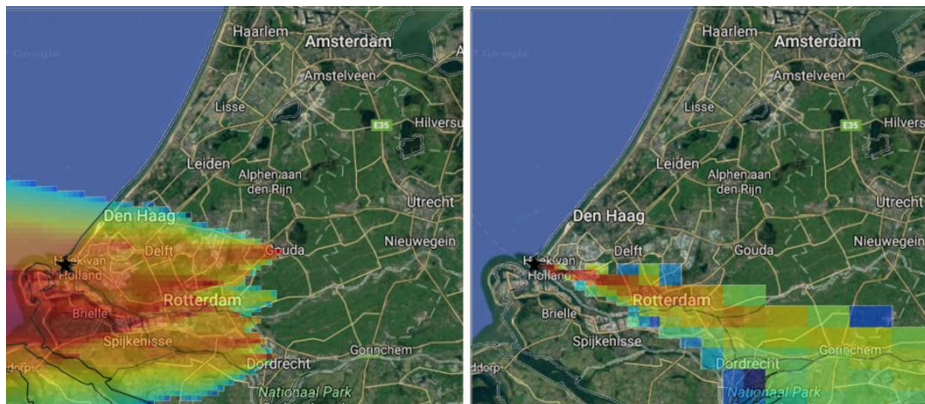
are more or less abundant in specific fuels and can therefore give information about dominant fuel types (Djuricin et al., 2010; Lopez et al., 2013). Moreover, carbon isotopes can be used to separate the fossil fuel signal from the influence of other CO<sub>2</sub> sources and sinks, shown in Fig. 1.2. The carbon in atmospheric CO<sub>2</sub> can contain either 6, 7 or 8 neutrons and the most abundant isotope is the one with 6 neutrons (called <sup>12</sup>C because of the number of neutrons and number of protons (6) adds up to twelve). The <sup>14</sup>C isotope with 8 neutrons is much less abundant in the atmosphere but is nevertheless present in organic materials. Interestingly, <sup>14</sup>C is radioactive, meaning it decays over time, and therefore no <sup>14</sup>C is left in fossil fuels that have been buried for millions of years. Therefore, combustion of fossil fuel causes a detectable decrease in <sup>14</sup>C in the atmosphere that can help to constrain the total fossil fuel signal of CO<sub>2</sub>.

These considerations related to the optimal configuration of monitoring sites and tracers lead to the first research question that will be addressed in Chapters 2, 3 and 5:

**RQ 1: What are the (dis)advantages of different observation networks and combinations of tracers in monitoring urban fossil fuel CO<sub>2</sub> emissions per source sector?**

#### 1.4.2 Atmospheric transport modelling

Atmospheric concentrations are the result of a wide range of processes, such as emission, transport, dilution, and uptake. In order to relate observed concentrations to emissions these processes need to be quantified, using simple parameterisations (Chapter 2) and/or transport models (Chapter 3). Basically, a transport model is a simulation of the reality and tries to explain how air pollutants are transported through the atmosphere. Because the exact description of each process that occurs in the atmosphere is often complex or unknown, a model consists of simplified descriptions of these processes called parameterisations.



**Figure 1.6: Examples of model output from a Lagrangian model (the OPS model). The left panel shows results from a simulation in forward mode, giving CO<sub>2</sub> concentrations resulting from point source emissions. The right panel shows results from a simulation in backward mode, giving the contribution of each discrete (Eulerian) grid cell to the concentration at a specific measurement site (black star).**

Transport models can be divided roughly into two categories: Eulerian and Lagrangian models. Eulerian models provide grid-box average concentrations (usually around  $10 \times 10$  km<sup>2</sup> to  $1 \times 1$  km<sup>2</sup> for (sub)national applications) which makes them efficient when large domains need to be covered. However, comparing a grid-box average with a point observation results in a significant model-observation mismatch (Houweling et al., 2000; Karamchandani et al., 2011; Peters et al., 2004). Moreover, a tracer from a point source which is emitted into a grid box becomes instantly mixed throughout the full volume of the grid boxes of the model. This causes an underestimation of the local and downwind simulated concentration.

Lagrangian models make use of trajectories to predict a concentration at a certain measurement location: they release a pollutant from a source and follow the plume as it is transported with the wind. This can be done in forward mode to calculate the dispersion of all emissions over the domain (Fig. 1.6, left panel) or in backward mode to determine the footprint of a specific measurement site (Fig. 1.6, right panel). Note that the prevailing wind on this day was easterly. In the latter case the transport is reversed and the contribution of each grid cell to the total concentration is calculated. Lagrangian models have no horizontal or vertical discretization, which makes them very suitable for transporting pollutants from sources with a small spatial extent (i.e. point sources). However, using them over a large area is computationally expensive, and the accuracy of transport still depends on the accuracy and resolution of the available meteorological information. Finally, Lagrangian models contain relatively simple parameterisations compared to most Eulerian models, resulting in less realistic estimates of dispersion and mixing.

In a complex urban environment, transport plays an important role in determining CO<sub>2</sub> concentrations. However, most regional transport models have no parameterisations to correctly represent small-scale urban transport, such as the effect of street canyons on turbulent mixing. This makes it difficult to correctly quantify such transport processes, causing potentially large discrepancies between observed and modelled concentrations. Therefore, it is important to identify the most suitable modelling framework to reduce transport errors as much as possible. This leads to the second research question:

**RQ 2: How well can different types of modelling frameworks represent observed atmospheric CO<sub>2</sub> concentrations at multiple monitoring sites?**

### 1.4.3 Emission modelling

As mentioned before, a downside of emission inventories is that they may have large uncertainties and often lag several years. Such inventories often downscale emissions using large-scale and long-term averaged activity data. The uncertainty of these high-resolution inventories is so far unknown. However, an increasing amount of high-resolution data is becoming available that can be used to downscale emissions more accurately. Therefore, we aimed at developing a dynamic emission model (Chapter 4). The

model calculates total yearly emissions from emission factors (the amount of CO<sub>2</sub> emitted per amount of fuel consumed), total yearly activity and the relation between activity and the amount of fuel consumed (energy efficiency). For example, the emissions from residential heating strongly correlate with the outside temperature, such that the average temperature over a year can be used as a proxy for the total emissions from residential heating (activity). The emission factor of natural gas used for heating is relatively well-known and the energy efficiency is dependent on for example the amount of insulation. So we try to estimate emissions, which we do not know, from physical drivers of these emissions (like temperature) which can be measured easily. Similarly, we calculate the emissions for each source sector. We then downscale the yearly emissions to hourly and 1x1 km<sup>2</sup> resolution using proxy data. The temporal variability of residential heating can be estimated from daily variations in the outside temperature, while the spatial distribution can be estimated using population density. So in principle the idea behind an emission model is quite similar to that of an emission inventory, i.e. using activity data and proxies to calculate emissions. Yet, the dynamic emission model makes use of more localized data with a high temporal resolution. Moreover, the emission model offers an alternative computation of the national annual emissions.

The dynamic emission model has several advantages over an emission inventory. First, the model is very flexible and can be adapted to the area under study. It can also be used for different species, which is useful when attributing CO<sub>2</sub> emissions to source sectors using tracer ratios. Second, the model can be more easily transported to regions where currently no information is available about emissions. An example is that a model developed for the coastal regions of the Netherlands might be used in coastal cities in Asia as well, where emission data is simply not reported and emission inventories do not yet exist. Third, instead of having emission fields the model consists of a wide range of parameters with a physical meaning. It therefore gives more insight in the underlying processes that determine the emissions. Fourth, the model makes use of typical parameter values that can have a wide range due to differences in technology. This information allows us to estimate the uncertainties in high-resolution emissions caused by uncertainties in these model parameters. With the dynamic emission model we can thus address the following research question:

**RQ 3: Can we predict hourly emissions per source sector in an urban environment using proxy data and what are the uncertainties?**

#### **1.4.4 Inverse modelling**

All of the previous methods (observations, transport model, emission model) can be combined in an inverse modelling system. Inverse modelling is a way to optimize a parameter by minimizing the difference between observations and model output. In the case of CO<sub>2</sub> monitoring, inverse models try to minimize the difference between observed and simulated concentrations by scaling a first estimate of the emissions (the prior), taking

into account uncertainties in the emissions, transport models and observations. For this purpose the following cost function is optimized:

$$J(x) = (y^0 - \mathcal{H}(x))^T R^{-1} (y^0 - \mathcal{H}(x)) + (x - x^b)^T (P^b)^{-1} (x - x^b) \quad (1)$$

Here,  $x$  is a vector containing all emissions that are optimized (the state vector, where  $x^b$  is the prior state vector).  $\mathcal{H}$  is the model that returns simulated concentrations given a certain emission field (observation operator), which in this thesis is a combination of Eulerian and Lagrangian (in backward mode, see Fig. 1.6) transport models.  $y^0$  is a vector containing all observations.  $R$  and  $P^b$  contain information about the uncertainties in the transport model/observations and the prior state vector, respectively. Together they determine the weight given to the observations and the prior estimate. The uncertainty of the optimized state vector can also be estimated ( $P^a$ ).

Often, the state vector consists of gridded emissions from an emission inventory that are optimized to be in better agreement with observations. In this thesis we use the dynamic emission model as a first estimate of the emissions, similar to the set-up of the Fossil Fuel Data Assimilation System (FFDAS) with global coverage (Raupach et al., 2010; Rayner et al., 2010). Our dynamic emission model allows us to estimate parameters inside the dynamical emission model instead of the total fluxes. This means that these parameters are in the state vector  $x$  and that the observation operator  $\mathcal{H}$  also includes the dynamic emission model to return  $\text{CO}_2$  mixing ratios as a function of the parameters to be estimated. The methodology of optimising model parameters has been applied before to, for example, biosphere models (Tolk et al., 2011). However, the application to a fossil fuel emission model is new. Our inverse modelling system also makes use of multiple tracers in order to attribute changes in  $\text{CO}_2$  emissions to specific source sectors, namely  $\text{CO}$ ,  $\text{NO}_x$  and  $\text{SO}_2$ .


The challenge with inverse modelling is that the total mismatch between observed and modelled concentrations can have different causes. The error in the emissions is what we are interested in. But the model parameterisations would cause a model-data mismatch even if the exact emissions are known and used, because the model transport is not perfect. This is especially relevant at higher resolutions as the urban atmospheric transport is very complex to model. As such, the transport errors are identified as a major challenge for urban inversions and often lead to the selection of well-mixed daytime conditions (Boon et al., 2016; Bréon et al., 2015; Brioude et al., 2013; Lauvaux et al., 2013; McKain et al., 2012). Moreover, in our system we only optimize fossil fuel fluxes, while other sources and sinks of  $\text{CO}_2$  (inflow from outside the model domain and biospheric fluxes) are also uncertain and cause part of the model-data mismatch. Therefore, a good representation of the errors is essential to get reliable emission estimates and the dynamic emission model can help with that.

This leads to the following research question:

**RQ 4: How well can we constrain urban emissions per source sector given (systematic) errors in the emission and transport model?**





A thick dark blue vertical bar runs down the left side of the page. A light teal arrow points to the right from this bar, containing the chapter title. Below the bar, several thin, curved lines in dark blue and light grey sweep upwards from the bottom left towards the center of the page.

## Chapter 2

# **Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam**

*This chapter is published as: Super, I., Denier van der Gon, H. A. C., Visschedijk, A. J. H., Moerman, M. M., Chen, H., Van der Molen, M. K., and Peters, W.: Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam, Atmos. Pollut. Res., 8, 174–187, 10.1016/j.apr.2016.08.008, 2017.*



**Abstract**

Large networks of expensive instruments are often used to independently quantify and monitor urban CO<sub>2</sub> emissions with sufficient level of detail. However, many developing regions cannot afford such a monitoring effort. We explore the use of a simple, less costly method to constrain urban emissions using only two measurement sites, one upwind and one downwind of the city of Rotterdam in the Netherlands. This provides an interesting dataset of concentration gradients of multiple combustion tracers over an urban-industrial complex. We find clear emission signals from three source sectors, mainly related to industrial activities in the port and from residential areas. We estimate the anthropogenic CO<sub>2</sub> emissions for three footprints from our observations and find them in reasonable agreement with the Dutch National Emission Registration (NER) database after accounting for biogenic fluxes. The large confidence interval for one of the footprints illustrates that the presence of point sources complicates the flux estimates. Additionally, we were able to pinpoint a limitation in the emission database using observed fossil fuel CO:CO<sub>2</sub> ratios, although the applicability of this method is limited for the footprint with a large influence from point source emissions. There is also a large variability in the observed ratios per footprint, which indicates that the dominant source type varies over time. Finally, we show that the fossil fuel CO concentration can be used to calculate fossil fuel CO<sub>2</sub> if their emission ratio is well-known.

## 2.1 Introduction

Urban areas are densely populated and have a wide range of anthropogenic activities – such as transportation, industry and residence – that result in emissions of many different pollutants. Although urban areas only cover a small land area, they contribute about 70% to the global greenhouse gas (GHG) emissions due to human activities (Font et al., 2014; Järvi et al., 2012). By monitoring a small hotspot of anthropogenic activity it is possible to get insight in the spatiotemporal variability of emissions, identify emission sources, and detect emission trends. Therefore, monitoring urban anthropogenic emissions of CO<sub>2</sub> and CH<sub>4</sub>, two of the most important GHGs, is receiving increasing attention (Bréon et al., 2015; Font et al., 2014; Peischl et al., 2015; Turnbull et al., 2015).

Ciais et al. (2014) stress the importance of different types of GHG observations and the deployment of GHG observation networks around megacities worldwide to identify the magnitude and spatiotemporal variability of emissions. Effective GHG observation networks could assist in establishing the effectiveness of emission-reduction policies (Turnbull et al., 2015). Over the past years different types of monitoring strategies have been used to constrain urban emissions, each with their own (dis)advantages, and several lessons that can be learned from these studies. First, if the footprint of an instrument is small – like with eddy-covariance flux towers (Bergeron and Strachan, 2011; Buckley et al., 2014) – many measurements are needed to constrain an entire urban area due to the spatial heterogeneity (Grimmond et al., 2002). In contrast, a large footprint makes it difficult to identify the emissions of a particular source sector, for example when using one observational site and a large wind sector (Levin et al., 2011). The footprint is also related to the location of the measurement site. Whereas in-city observations provide valuable information on local emissions (Gratani and Varone, 2005; Zimnoch et al., 2010), they are only sensitive to part of the urban area. In contrast, remote sites are only occasionally affected by the urban emissions, but do see a larger part of the urban area. So the location and footprint of the measurement need to be considered carefully. Second, the spatial and temporal resolution of the measurements is important, especially in a heterogeneous urban-industrial environment. While several studies have used satellite retrievals of CO<sub>2</sub> columns to constrain urban emissions (Kort et al., 2012; Silva et al., 2013), this method provides too little spatial detail for source attribution. Similarly, using <sup>14</sup>C flask or plant samples to identify the fossil fuel contribution to the total CO<sub>2</sub> signal (e.g. (Bozhinova et al., 2014; Djuricin et al., 2010; Lopez et al., 2013; Turnbull et al., 2015)) is expensive and time-consuming (Djuricin et al., 2010). Therefore, the temporal resolution of this method is generally limited. Methods in which ratios of multiple tracers are measured continuously (Levin and Karstens, 2007; Lopez et al., 2013; Vogel et al., 2010) seem promising, with the main challenge to delineate the relations between different emission sources from the observed mole fraction ratios.

In addition, most of these studies have used long-term observations and/or an extensive observational network. Although the required number of monitoring sites depends on the size, population density, and complexity (heterogeneity) of the area under study, constraining urban emissions asks for sufficient observations of atmospheric concentrations to account for the large spatiotemporal variability. Kort et al. (2013) compared different network configurations to identify the optimal monitoring strategy for Los Angeles' ( $\sim 1300 \text{ km}^2$ )  $\text{CO}_2$  emissions. They conclude that at least eight measurement sites are needed to have sufficient sensitivity to emissions from the entire basin. Additionally, McKain et al. (2012) argue that 5 monitoring sites are needed to constrain the emissions from Salt Lake City, while Turnbull et al. (2015) have access to 12 measurement sites in Indianapolis. Such monitoring networks necessitate the availability of numerous resources, including sufficient funds and trained technicians to maintain the instruments. However, such resources are evidently not available in all regions. Ciais et al. (2014) note that, despite all efforts to monitor (urban) GHG emissions, many regions are still systematically undersampled. These regions include some developing economies that experience a high urbanization rate, thus becoming important contributors to anthropogenic GHG emissions. However, such regions may have neither the resources nor the priority to build and preserve an extensive monitoring network. An interesting question is thus whether a simpler approach can provide sufficient detail to monitor and constrain the emissions of a developing urbanized and/or industrialized centre.

In this paper we present measurements from a relatively simple and cheap monitoring framework that has a high temporal resolution and a footprint that covers an urban area. It consists of only two observational sites, which are located upwind and downwind of a medium-sized urban area with respect to the prevailing wind direction of  $190^\circ$ . We selected the city of Rotterdam, the Netherlands (dark blue shape in Fig. 2.1) as our case study. Rotterdam is a relatively small city (about  $320 \text{ km}^2$  and 625.000 inhabitants) compared to megacities used in previous studies. However, Rotterdam is part of a larger urbanized area containing several urban centres and industrial areas, surrounded by pasture and agricultural land. This area, which is called the Rijnmond area ( $\sim 860 \text{ km}^2$ , light blue outline in Fig. 2.1), has 1.2 million inhabitants. The Rijnmond area includes the largest sea port of Europe ( $\sim 100 \text{ km}^2$ , yellow shape in Fig. 2.1) with activities in storage and transshipment of dry and wet bulk and containers, refineries, chemistry, and energy industries. Along the coast (North Sea) between the port and The Hague, there is a large area with gas-heated and  $\text{CO}_2$ -enriched glasshouses (green shape in Fig. 2.1). Additionally, other major cities, such as The Hague and Utrecht, are only at 20 and 45 km distance from the city centre of Rotterdam, respectively. Thus the Rijnmond area is characterised by a wide variety of scattered anthropogenic activities and influence of other nearby urban centres. As such, it is an interesting case for other European coastal urban areas that often display a more heterogeneous emission landscape than most megacities in the US or Asia.

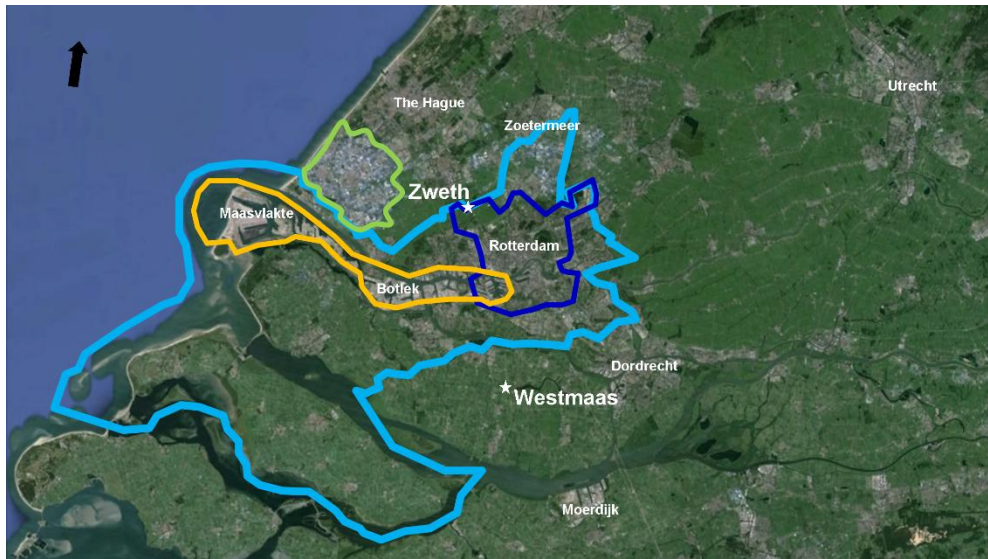


Figure 2.1: Map of the Rijnmond area (light blue outline), including the city of Rotterdam (dark blue outline), the port (yellow outline) and the glasshouse agriculture source sector (green outline); the observation sites are indicated with white stars; the prevailing wind direction during the observations is given by the black arrow. Source: Google Earth.

We measure (semi)continuous mole fractions of  $\text{CO}_2$ ,  $\text{CH}_4$ , and  $\text{CO}$ . The latter has the potential to function as a tracer for fossil fuel  $\text{CO}_2$  emissions (Djuricin et al., 2010; Turnbull et al., 2015). With these observations we have created a dataset of long-term continuous concentration gradients of multiple combustion tracers over an urban-industrial complex. We perform several analyses to examine how much two continuous observations sites (Westmaas and Zweth) can explain about the  $\text{CO}_2$  emissions from important source sectors within this highly urbanized area. Moreover, we study the use of  $\text{CO}$  as tracer to identify the fossil fuel contribution to the total  $\text{CO}_2$  signal found at our measurement sites. In this paper we discuss the use of our observations to identify signals from Rotterdam and the port, following four research questions:

- Can we explain the temporal variations in the  $\text{CO}_2$  concentration time series?
- Under which conditions is the upwind site (Westmaas) suitable as background station to estimate  $\text{CO}_2$  fluxes from the Zweth-Westmaas gradient?
- Can we get a first estimate of the fossil fuel  $\text{CO}_2$  fluxes of important source sectors with the two measurement sites?
- Can  $\text{CO}$  be used to quantify the fossil fuel contribution to the total  $\text{CO}_2$  signal from our case study?

We start with a description of our observations and method to estimate fossil fuel fluxes, which is based on a mass balance approach. Then we demonstrate the spatiotemporal variability in the  $\text{CO}_2$  background concentration, estimate  $\text{CO}_2$  fluxes for

several footprints, and reconstruct fossil fuel CO<sub>2</sub> concentrations using CO. Finally, we present recommendations for effectively monitoring and estimating fossil fuel fluxes.

## 2.2 Methods

### 2.2.1 Sampling sites

We selected two measurement locations marked by a star in Fig. 2.1, 15 km south (location Westmaas) and 7 km northwest (location Zweth) of the city centre of Rotterdam (see Table 2.1 for details). The Westmaas site is in an open area with mostly low agricultural crops and a few trees along a subsidiary road (at 150 m distance). There are highways at 1.3 km to the north and at 1.7 km to the west. Westmaas is also a station in the Dutch National Air Quality network (RIVM). This site is considered the upwind site due to the prevailing southwesterly wind and because of its relatively rural to suburban location. The Zweth site is not part of an existing network and was selected specifically for this study. It is located between a highway (at 650 m), a railway (at 1 km), and an airport (at 2.6 km) and at the edge of a small village (De Zweth) with a medium-large canal used for occasional inland shipping. Rotterdam The Hague Airport is the third largest airport of the Netherlands, which serves many European metropolitan areas (>1.6 million passengers in 2014) and is also regularly used for military aircrafts. Therefore, this site is more affected by anthropogenic emissions and is used as downwind site. The field in which the instrument is installed is surrounded by trees and is about 85 by 30 m large. With a sampling height of 10 m (almost similar to the height of the trees) the influence of the trees on advection may be an important factor. A sampling height of 10 m is sub-optimal, but installing higher masts was not allowed in the area.

**Table 2.1: Specifications of the two measurement sites Westmaas and Zweth.**

	Coordinates	Start time series	Sampling height [m above surface]	Elevation [m above sea level]
<b>Westmaas</b>	51.786666° N, 4.450536° E	2014-01-01	10	-0.5
<b>Zweth</b>	51.964381° N, 4.394650° E	2014-05-01	3 (until 2014-07-03 11:00) 10 (from 2014-07-03 11:00)	-3

At both locations we installed a low-drift analyser based on cavity ring-down spectroscopy (CRDS) (Picarro Inc., CA, USA, type G2401 (Picarro, 2015)) to measure atmospheric concentrations of CO<sub>2</sub>, CH<sub>4</sub> and CO (more information on the observations and their availability can be found in the Supplementary Information). It requires little maintenance and can be operated at relatively low cost without the need for highly trained technicians. Moreover, the long-term stability reduces the consumption of expensive calibration gases (Andrews et al., 2014; Richardson et al., 2012; Schmidt et al., 2014; Welp et al., 2013; Winderlich et al., 2010). The low cost and ease with which the



analysers can be operated are essential for possible expansion of the monitoring network to other cities in developing countries and remote areas.

### 2.2.2 Methods for flux estimates

When an air parcel is transported from Westmaas to Zweth, the mole fraction of CO<sub>2</sub> in that air parcel is affected by processes that occur during transport. Therefore, comparing CO<sub>2</sub> mole fractions at Westmaas and Zweth can provide information on the magnitude of those processes in their footprint. Previous research has shown that observations made during flights upwind and downwind of a source sector of interest can be used to estimate the flux for that sector (Caulton et al., 2014; Karion et al., 2013; Mays et al., 2009; Peischl et al., 2015). The integrated flux of X ( $F_X$ ) over the total sector between the two flight tracks can be calculated using a mass balance equation:

$$F_X = v \cdot \cos(\alpha) \int_0^{z_l} \int_{-y}^y \Delta X_{obs} dy dz \quad (1)$$

where  $v \cdot \cos(\alpha)$  is the proportion of the wind velocity parallel to the observed concentration gradient, and  $\Delta X_{obs}$  is the observed concentration enhancement over a background value which is integrated over a horizontal (y) and vertical (z) plane where  $z_l$  is boundary-layer height adjusted for vertical transport and boundary-layer growth (Peischl et al., 2015). In contrast to these studies we lack a y-dimension as the location of our observational sites is stationary in time, but we do have a longer time series that allows us to apply data selection.

To get a flux estimate from continuous in-situ observations we adapt the previous mass balance equation to the one-box model (Fig. 2.2). The change in concentration [ $\text{kg m}^{-3}$ ] of species X in a box is a function of the inflow  $F_{in}$  and outflow  $F_{out}$  (advection), chemical reactions ( $C$  and  $L$ ), emission ( $E$ ) and deposition ( $D$ , all fluxes in [ $\text{kg km}^{-2} \text{ hr}^{-1}$ ]):

$$\Delta X_{obs} = (F_{in} - F_{out} + E - D + C - L) \cdot \Delta t / h \quad (2)$$

where  $\Delta t$  is the time difference between two measurements of X in hours and  $h$  the boundary layer height in km. We adopt a Lagrangian approach in which this box moves with the wind from an upwind to a downwind site and advection becomes zero. Since CO<sub>2</sub> is chemically inert only emission (both fossil fuel emissions  $E_{ff}$  and biogenic respiration  $E_{bio}$ ) and deposition (photosynthesis  $A_n$ ) remain, resulting in the following equation:

$$F_X = E_{ff} + E_{bio} - A_n = \Delta X_{obs} \cdot h / \Delta t \quad (3)$$

where  $\Delta X_{obs}$  is equal to the concentration at Zweth measured at time  $t + \Delta t$  minus the concentration at Westmaas measured at time  $t$ . With a distance of 20.15 km between the sites the time difference  $\Delta t$  for wind parallel to the gradient is typically between 0.5 and

1.5 hours. Note that we round off  $\Delta t$  to full hours since we binned our observations (approx. 1 per second) into hourly concentrations.

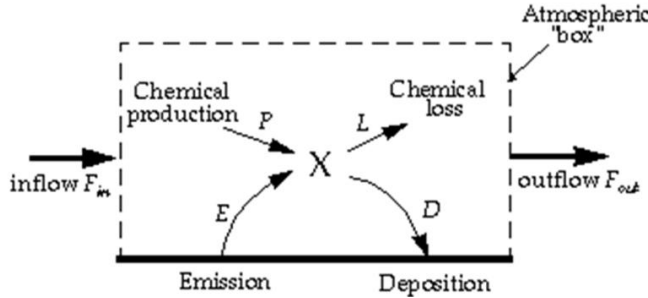


Figure 2.2: Representation of a one-box model. Source: Jacob (1999)

We define a footprint from which the fluxes are observed at Zweth when the wind is coming from Westmaas. We need to identify this area of influence in order to compare the flux estimates made with Eq. 3 with those given by emission databases for the same area. We define the footprint as the area covered by a Gaussian plume at neutral stability (to represent dominant wind speed and mixing conditions) from the downwind to the upwind site (i.e. it has a triangular shape). The total area of the footprint between Westmaas and Zweth is about  $44 \text{ km}^2$  (see Table 2.2) following:

$$A = 3\sigma_y \cdot x = 3F \cdot x^f \cdot x \quad (4)$$

where  $F$  (0.32) and  $f$  (0.78) are empirical stability parameters,  $x$  is the travel distance in m and  $\sigma_y$  is the standard deviation of the Gaussian plume in the horizontal direction in m. We multiply by 3 to get the  $3\sigma$  width of the normal distribution.

To estimate the fossil fuel flux we need to approximate the biogenic fluxes in the footprints and account for their contribution to the  $\text{CO}_2$  concentration gradient. This is done based on observations in the Netherlands of annual cycles of the net ecosystem exchange ( $\text{NEE} = E_{\text{bio}} - A_n$ ) (Hendriks et al., 2007), from which we determine monthly mean NEE. An average daily cycle for May is determined based on the work of Jacobs et al. (2003) and scaled according to the monthly mean NEE. This gives us an average daily cycle per month (Appendix A), which will be used to calculate the total biogenic flux. Both the Hendriks et al. and Jacobs et al. studies used NEE observations over well-watered/wet grassland, which match with the agricultural land use in our study area.

Note that the mass balance approach assumes the emissions to be well-mixed throughout the boundary layer by the time it reaches the downwind site and that  $h$  is taken as a constant during transport. We use a monthly value of  $h$  based on measurements at the nearby Cabauw site (Royal Netherlands Meteorological Institute (KNMI), 2007) and assume this value to be maximal and relatively constant during the afternoon (see Appendix A). These assumptions on constant  $h$  and neutral stability are

only valid during well-mixed afternoon conditions. Therefore we select only 12-16h UTC data from our time series that correspond to a wind speed of at least  $3 \text{ m s}^{-1}$ .

Although this process of using an up- and downwind observational site to constrain urban emissions seems rather straightforward, Turnbull et al. (2015) have identified several difficulties related to this strategy for Indianapolis. First, the mole fraction is heterogeneous in space and time. Spatial heterogeneity implies that the upwind observation is not necessarily representative for the mole fraction of the entire ‘background’ air parcel moving over the city. This is especially relevant if the wind direction is not exactly parallel to the concentration gradient. The issue of temporal heterogeneity is addressed by accounting for the travel time between the up- and downwind site. Note, however, that the gradient then becomes:

$$\Delta X = X_{Zweth}(t + \tau) - X_{Westmaas}(t) \quad (5)$$

where time lag  $\tau$  between the observations at the upwind and downwind site is:

$$\tau = (x/v) * \cos(\alpha) \quad (6)$$

and  $x$  is the distance between the sites,  $v$  the wind speed and  $\alpha$  the wind direction relative to the gradient. The time lag is zero when the wind blows perpendicular to the gradient (so the gradient consists of measurements taken simultaneously at both sites). The value of  $\Delta t$  in Eq. 3 is independent of the wind direction, but merely a function of wind speed and travel distance.

Second, other sources and sinks affect the mole fraction during transport, which makes it difficult to extract the impact of fossil fuel emissions from the total signal. We have previously discussed the biogenic sources and sinks, but Turnbull et al. (2015) also mention the importance of entrainment when the free-tropospheric concentration significantly differs from the boundary-layer background concentration. However, estimating entrainment requires an estimate of the free-tropospheric mole fractions. Additionally, mixing and entrainment of  $\text{CO}_2$  can be highly variable as those processes depend on atmospheric conditions and the land surface (Vilà-Guerau de Arellano et al., 2004). Therefore, similar to Turnbull et al. (2015) we assumed entrainment did not affect the mole fraction during transport. During the afternoon hours we selected here, this assumption is supported by a typically quite constant boundary layer height  $h$  (Vilà-Guerau de Arellano et al., 2004).

Finally, the distinction has to be made between in-city and remote observational sites. Kort et al. (2013) argue that in-city observational sites are continuously affected by local emissions. Therefore, they are better suitable for constraining urban fluxes than remote sites, even though they are only sensitive to a small part of the entire urban area. In contrast, remote sites only detect urban emissions at certain wind directions and even

then the signal is diluted by other fluxes, i.e. the urban air is mixed with other air masses. Turnbull et al. (2015) also found an occasional and diluted signal at a 130m-tower 24 km downwind of the centre of Indianapolis. In contrast, our downwind site is at only 7 km from the city centre and within an urbanized area, so we consider it to be semi-urban rather than remote. However, the sampling height is only 10 m. Therefore, we will explore the use of such sites for monitoring urban fluxes. We expect that the effect of dilution is limited while at the same time we are able to detect urban signals for most wind directions. We will go more in-depth in the discussion.

### 2.2.3 Flux estimates for non-aligned cases

The methodology presented in the previous section requires an air parcel that travels straight from Westmaas to Zweth, as the gradient between the sites is only determined by the biogenic and fossil fuel fluxes in the footprint  $P_1$  (Fig. 2.3). However, this limits the amount of useful observations. For example, including westerly winds allows us to get an impression of the emissions from the glasshouse source sector.

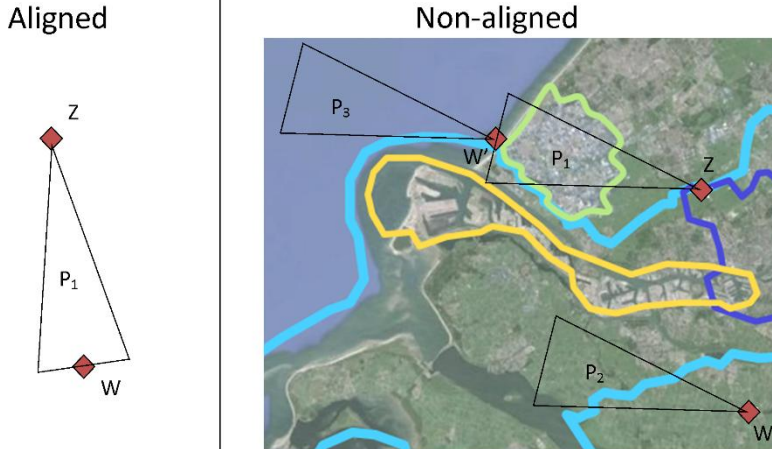


Figure 2.3: Illustration of the footprints ( $P_1$ ,  $P_2$ ,  $P_3$ ) in the aligned and non-aligned case, where Z and W represent the Zweth and Westmaas measurement sites, W' is the actual upwind location and triangles illustrate the footprints.

Therefore, we adapt the method to apply to other, 'non-aligned' source sectors by rotating the footprint  $P_1$  of Zweth into the wind direction such that it overlaps the glasshouse source sector (Fig. 2.3). If we want to estimate the flux in  $P_1$  we need to account for the difference between the mole fraction measured at W and the actual upwind mole fraction at W'. This is done by estimating the difference between the fluxes in the footprints of W ( $P_2$ ) and W' ( $P_3$ ) and assuming that Westmaas is not affected by large, local fluxes. We will show that this is a valid assumption for several wind sectors. Eq. 3 then becomes:

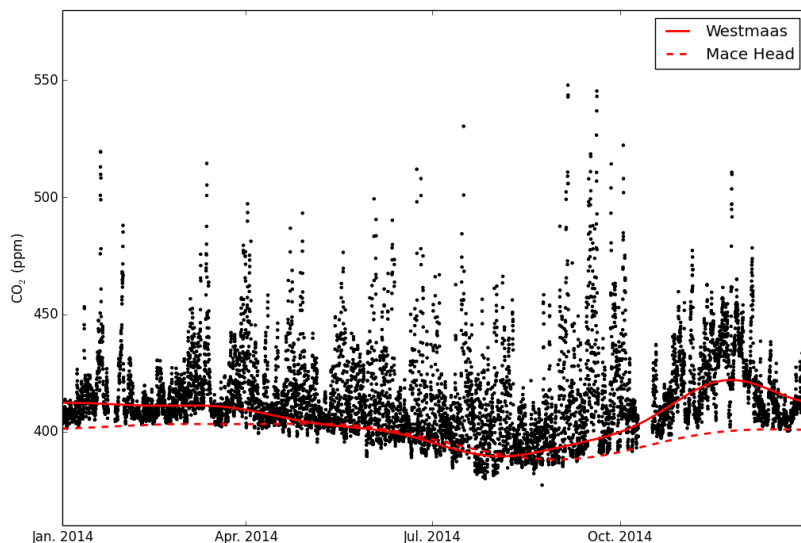
$$E_{ff} = \Delta X_{obs} \cdot h / \Delta t + A_n - E_{bio} + (F_{P_2} - F_{P_3}) \quad (7)$$

where  $F_{P2}$  and  $F_{P3}$  are the total fluxes in footprints  $P_2$  and  $P_3$ . Since the footprints  $P_2$  and  $P_3$  are not over urbanized or industrialized areas, the difference between  $F_{P2}$  and  $F_{P3}$  can be considered to be caused by biogenic fluxes.

## 2.3. Results

### 2.3.1 Temporal variability in CO<sub>2</sub> concentrations

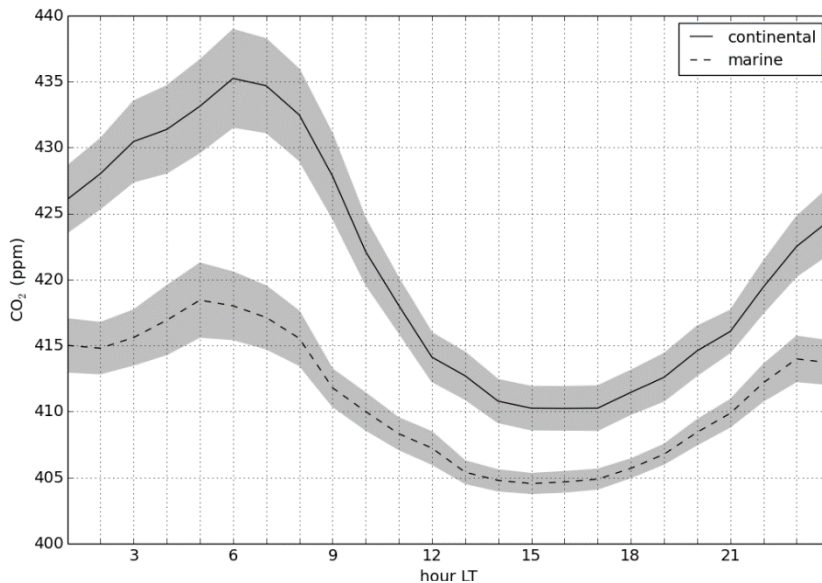
Our CO<sub>2</sub> concentration time series show a large temporal variability (Fig. 2.4). If we want to use these observations to constrain urban fossil fuel fluxes, we need to ensure that the observations are indeed highly affected by those anthropogenic sources. In order to extract the anthropogenic influence from the total observed CO<sub>2</sub> concentrations we need to consider several processes that explain the full range of variability in the time series of the CO<sub>2</sub> concentration.



**Figure 2.4:** Smoothed annual cycle of daytime CO<sub>2</sub> mole fractions at Westmaas (solid red curve) and Mace Head, Ireland (dashed red curve). The dots indicate the observed hourly CO<sub>2</sub> mole fractions at Westmaas.

The CO<sub>2</sub> concentration has a distinct annual cycle, mainly driven by the biosphere. The seasonal cycle in the background afternoon CO<sub>2</sub> concentration is estimated with the curve fitting method described by Press et al. (1992). We fit a function to the observations to describe the annual cycle. Next, we apply a low-pass filter with a default cut-off value of 80 days to the hourly residuals to account for synoptic variations. Figure 2.4 shows the seasonal cycle of Westmaas and Mace Head (Bousquet et al., 1996; Derwent et al., 2002) (from the GLOBALVIEWplus product (Cooperative Global Atmospheric Data Integration Project, 2015)). Mace Head is located at the west coast of Ireland and is representative for marine background air without the influence of local (anthropogenic and biogenic) sources and sinks. Therefore, the short-term variability in its CO<sub>2</sub> concentrations is limited

to a few ppm. The two curves show a similar cycle, with lower than average concentrations from July to September. This is related to the large-scale biospheric uptake that exceeds the ecosystem respiration during the Northern Hemisphere summer. During the winter, ecosystem and soil respiration cause an increase in the CO<sub>2</sub> concentration. Interestingly, the smoothed seasonal cycle at Westmaas has a larger amplitude, showing higher concentrations in the winter and slightly lower concentrations in the summer compared to Mace Head. This suggests that there is an additional local influence of biogenic fluxes compared to the marine background air sampled at Mace Head. This additional influence would be expected given the vegetated surface in the footprint of the Westmaas site.



**Figure 2.5:** Average daily cycle of the CO<sub>2</sub> mole fraction at Westmaas for continental (NE-S wind direction, full line) and marine (SSW-N wind direction, dashed line) regime and 95% confidence interval of the mean. LT is the local wintertime.

In addition, the CO<sub>2</sub> concentration at Westmaas has a more distinct peak towards the end of the year, whereas Mace Head shows the highest CO<sub>2</sub> concentrations at the end of the winter. This peak can be explained by a change in wind regime. During the period of high concentrations the prevailing wind direction is easterly and continental air is advected. During the rest of the year mainly marine air with a lower CO<sub>2</sub> concentration is advected by the predominant (south-)westerly wind. Figure 2.5 shows the average daily cycles for the marine (wind from SSW-N) and continental (wind from NNE-S) regimes. These average cycles are created by subtracting the smoothed annual cycle in Fig. 2.4 and then adding the yearly average value of this smoothed cycle to each data point. This provides a time series of seasonally detrended CO<sub>2</sub> concentrations to prevent a bias due to the change in dominant wind regime between winter and summer. The continental

regime shows a significantly higher concentration, because continental air contains emissions from the European mainland. In contrast, marine air is only (partly) affected by emissions from the UK. Moreover, the CO<sub>2</sub> concentration under the marine regime shows less diurnal variability. Marine air is relatively well-mixed and only boundary-layer development has some impact on the concentrations. In contrast, the amplitude during the continental regime is much larger. During daytime the difference between the two regimes is small due to intensive mixing. Yet at night the continental emissions are trapped in a much smaller boundary-layer, causing the concentration to increase.

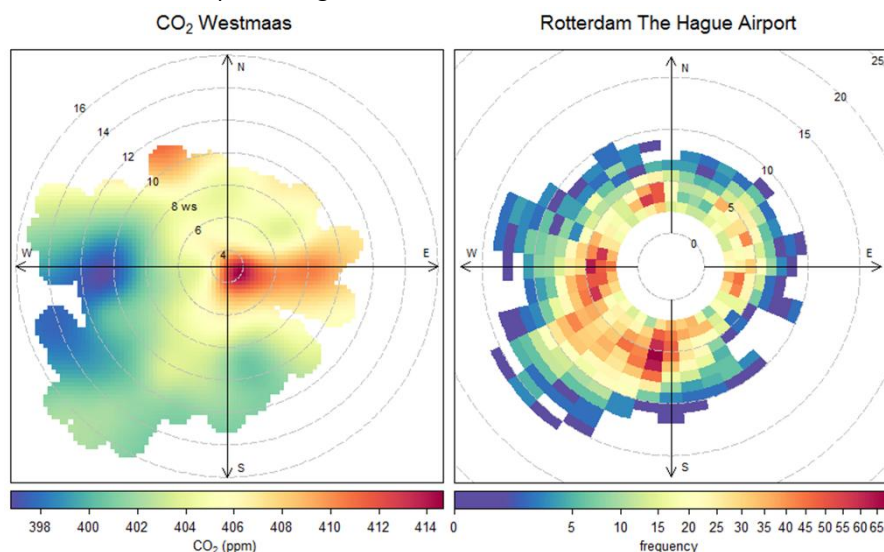
Even though a large part of the variability in the CO<sub>2</sub> concentration can thus be explained by the seasonal cycle and changes in wind regime, a part of the variability remains unexplained. The smoothed function in Fig. 2.4 can be considered a background concentration and the increases above this background can be largely explained by the impact of regional sources and sinks. So indeed our observations are affected by anthropogenic sources. Therefore, we subtract the seasonal cycle from the observations for all further analyses to ensure that the signals we see are coming from those regional influences rather than large-scale atmospheric changes.

### **2.3.2 Spatial variability in CO<sub>2</sub> concentrations**

Now we know that our observations contain anthropogenic signals, we can examine the observations in more detail. The case study contains several regions, such as the port and the glasshouse area, with large anthropogenic emissions. Therefore, if we want to get a full spectrum of urban signals we need to use Westmaas-Zweth gradients from different wind directions. However, it is important to ensure that the upwind site is suitable as background station. This means that the concentration measured at Westmaas should be relatively homogeneous and similar to the observed concentration at Zweth in absence of the source sector.

Figure 2.6 (left panel) shows a bivariate polar plot of the observed CO<sub>2</sub> mole fractions at Westmaas. The polar plot shows the median mole fraction for the period 2014-2015 (which is used for all further analyses) as a function of wind direction and wind speed. Based on the position of Zweth relative to (the port of) Rotterdam we are mainly interested in north-westerly to south-easterly winds. For Westmaas we clearly see a marine regime with uniformly lower concentrations with a south-westerly to north-westerly wind. Even the impact of emissions from the UK hardly appears in this plot, making Westmaas a suitable background site for Zweth at these wind directions. Winds from the southwest to southeast advect continental air coming from Belgium, causing a small increase in the CO<sub>2</sub> concentration. For a wind direction of 170° this is no issue, because the wind blows parallel to the Westmaas-Zweth transect and this is thus included correctly in the Westmaas as background. However, for south to south-westerly winds the observed gradients might be less useful. Also, at lower wind speeds we seem to detect traffic emissions from the highway west of Westmaas. The (south)east winds, which are

expected to advect emissions from Rotterdam The Hague Airport and part of the city of Rotterdam to Zweth, are not suitable for our analysis due to large increments in the concentration at Westmaas. These are likely to be caused by emissions from Moerdijk (southeast of Westmaas), the city of Dordrecht and its industrial area (east of Westmaas), and at high wind speed the Ruhr area in Germany. To summarize, we expect that winds from south-southeast and north-northwest can be used to constrain emission from a part of the city of Rotterdam and south-westerly to north-westerly winds are suitable to detect emissions from the port and glasshouse area.



**Figure 2.6:** Median  $\text{CO}_2$  mole fraction observed at Westmaas (left) as a function of wind direction (angular coordinate) and wind speed (radial coordinate); Frequency of occurrence of each combination of wind speed and wind direction at Rotterdam The Hague Airport (right).

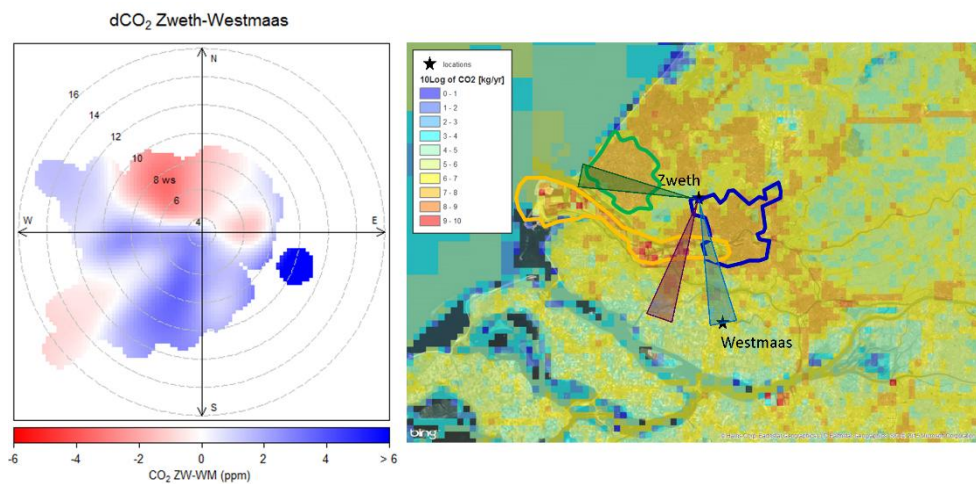
### 2.3.3 $\text{CO}_2$ flux estimates for footprints

We have seen in Sect. 2.3.2 that there are several wind sectors that contain important source sectors and have a suitable background concentration estimate. Combining these results with the Zweth-Westmaas gradients (Fig. 2.7, left panel) allows us to identify interesting footprints. For Fig. 2.7a we used median values to reduce the influence of single events since we are interested in sustained concentration gradients associated with large source sectors. For Rotterdam, which also includes a small fraction of the port, we find a gradient of +2 and -2 ppm (positive gradients mean Zweth has higher mole fractions than Westmaas) for southeasterly and northwesterly winds, respectively. Because this wind direction is parallel to the gradient, the negative gradients can also be used to estimate the flux from this source sector, reversing the role of Zweth to background.

In addition, we can identify two more wind direction regimes with large positive  $\text{CO}_2$  gradients; one with a gradient of +3-4 ppm with south-westerly to southerly winds and a smaller one with a gradient of up to +2 ppm with westerly winds. The first one is the result



of a wide range of industrial activities in the port. The largest gradient is probably caused by emissions from a section of the port mainly occupied by refineries, energy production, and chemical industries (hereafter referred to as port-Botlek). We have seen in Sect. 2.3.2 that Westmaas is not an appropriate background site for this wind direction. That we find a very clear signal even though the background concentration is already elevated at Westmaas indicates that this air travels over an important source sector. Therefore, we will try to estimate the flux of this source sector too to see how important the background concentration is. The smaller signal of +2 ppm is likely to be caused by emissions from the Maasvlakte, a section of the port where we mainly find activities related to transshipment of containers, dry and wet bulk storage and transshipment, a large coal-fired power plant, and the glasshouse area between the port and The Hague (hereafter referred to as Maasvlakte/glasshouse).



**Figure 2.7:** Difference between Zweth and Westmaas CO<sub>2</sub> mole fractions (left) as a function of wind direction (angular coordinate) and wind speed (radial coordinate) for a minimum bin size of 2; Yearly CO<sub>2</sub> emissions for 2012 (right) (source: Netherlands PRTR (2014)) with source sectors similar to Fig. 2.1 and the selected footprints are given by triangles (blue: metropolitan Rotterdam; red: port-Botlek; green: Maasvlakte/glasshouse).

So we have identified three interesting wind sectors that represent part of the Port-Botlek source sector (200°) with large point source emissions, Maasvlakte/glasshouse source sector (280°), and metropolitan Rotterdam (170°). The corresponding footprints are shown in Fig. 2.7b. In reality the footprint that represents Rotterdam also contains part of the port source sector, but this part contains no large point sources. We estimate the fluxes in these footprints using Eq. 3 and Eq. 7.

We now estimate the fossil fuel fluxes in the footprints based on the observed gradients ( $F_{\text{obs}}$ ) and compare them to estimates based on the emissions inventory ( $F_{\text{NER}}$ ) for Rijnmond (2012) from the Dutch National Emission Registration (NER) (Netherlands PRTR, 2014). The average wind speed is about  $7 \text{ m s}^{-1}$  for the Port-Botlek footprint and  $6 \text{ m s}^{-1}$  for the other two footprints. The results are given in Table 2.2. Since the individual flux

estimates from the observations are not normally distributed the median fluxes are given and their 95% confidence interval. Moreover, we should note that there is a significant uncertainty in the emission database, especially at high spatiotemporal resolution. For example, Ciais et al. (2010) found an uncertainty of 19% in emission databases at the European level and spatiotemporal downscaling causes uncertainties to increase up to 50% at 100 km horizontal resolution. At the Dutch national level the uncertainty of the annual CO<sub>2</sub> budget is estimated to be 3% (Coenen et al., 2012), but the uncertainty at 1km hourly resolution is unknown. In Table 2.2 also the vegetative fraction in the footprints is given, which is used to calculate the biogenic fluxes. As the port-Botlek and Maasvlakte/glasshouse footprints are non-aligned, an additional correction is required. For the port-Botlek footprint the  $F_{p2}$  and  $F_{p3}$  (Eq. 7) are very similar and  $F_{p2}-F_{p3}$  is nearly zero, but for the Maasvlakte/ glasshouse footprint we account for the difference in  $f_{veg}$  in  $F_{p2}$  (0.7) and  $F_{p3}$  (0, as it is located over sea).

**Table 2.2:** CO<sub>2</sub> fluxes in kg km<sup>-2</sup> hr<sup>-1</sup> estimated from the observed Zweth-Westmaas gradient ( $F_{obs}$ ) and as reported in the NER emission database ( $F_{NER}$ ) for three footprints and the corresponding wind direction, distance from upwind to downwind site (x), area of the footprint (A), the fraction of vegetation in the footprint ( $f_{veg}$ ) and number of observations used for the flux estimates (N).

Footprint	$F_{obs}$ [kg km <sup>-2</sup> hr <sup>-1</sup> ]	95% Confidence interval $F_{obs}$ [kg km <sup>-2</sup> hr <sup>-1</sup> ]	$F_{NER}$ [kg km <sup>-2</sup> hr <sup>-1</sup> ]	Wind direction (deg.)	x [km]	A [km <sup>2</sup> ]	$f_{veg}$	N
<b>Metropolitan Rotterdam</b>	2670	(2211-4139)	3396	170	20.15	44	0.2	10
<b>Port-Botlek</b>	6752	(5198-9170)	8750	200	20.15	44	0.4	43
<b>Maasvlakte/ glasshouse</b>	2645	(2193-3888)	2295	280	19.24	40	0.3	116

For the footprint metropolitan Rotterdam, sampled when the wind blows parallel to the gradient, there are only 10 cases (N=10 hours) which match our criteria. However, every individual hour represents the average of ~300 measurements within that hour; hence the hourly concentrations are very robust with a low uncertainty. Variations in  $F_{obs}$  are the result of short-term variations in the emissions, while the same 10 fluxes estimated from the emission database contain little temporal variability. Nevertheless,  $F_{obs}$  is in reasonable agreement with  $F_{NER}$ . Also the Maasvlakte/glasshouse and port-Botlek  $F_{NER}$  fall within the  $F_{obs}$  confidence interval, although this range is much larger for port-Botlek than for metropolitan Rotterdam and Maasvlakte/glasshouse. We have shown in Sect. 2.3.2 that the Westmaas concentration at this wind direction is elevated by local emissions, which can result in an underestimation of the flux. Moreover, in this part of the port a fraction of the CO<sub>2</sub> is emitted from stacks, which may not always reach ground level at the measurement site. Therefore, we conclude that making a reliable flux estimate of the port-Botlek footprint with just these two measurement sites is not possible without additional information on the transport to constrain the estimate.

### 2.3.4 $\Delta\text{CO}:\Delta\text{CO}_2$ ratios

Now we have seen that the two sites can be used to get a first idea of the urban fluxes, another interesting question is whether CO can be used as tracer for fossil fuel CO<sub>2</sub> in this area. Many source types have a distinct emission ratio between co-emitted species like CO and CO<sub>2</sub> (Djuricin et al., 2010; Levin and Karstens, 2007; Lopez et al., 2013) and see Table 2.3). Therefore, if we know the CO:CO<sub>2</sub> emission ratio in a footprint we could use CO to reconstruct the fossil fuel CO<sub>2</sub> time series. Moreover, the tracer ratios ( $\Delta\text{CO}:\Delta\text{CO}_2$ ) in the observations can provide information on dominant source types.  $\Delta\text{CO}_2$  and  $\Delta\text{CO}$  represent an elevation in concentration above the background concentration, which we assume is completely caused by fossil fuel emissions. We can calculate  $\Delta\text{CO}_2$  and  $\Delta\text{CO}$  by subtracting the smoothed curve from the original concentration time series (e.g. Fig. 2.4 for CO<sub>2</sub>), similar to Tohjima et al. (2014). We correct the  $\Delta\text{CO}_2$  for the biogenic contribution based on our previous findings.

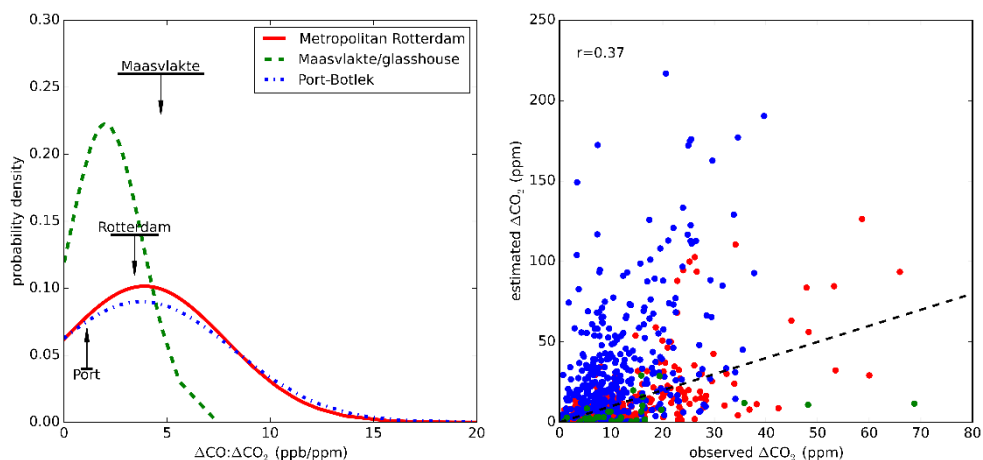
**Table 2.3: Average CO:CO<sub>2</sub> emission ratios per emission category for the entire Rijnmond area, based on the National Emission Registration 2012.**

Source type	CO:CO <sub>2</sub> (ppb/ppm)
Power generation	0.16
Non-industrial combustion	2.77
Industrial combustion and processes	1.14
Road transport	16.72
Non-road transport	5.81

Since every footprint has a characteristic mixture of source types, they also have a distinct emission ratio. The CO<sub>2</sub> footprint of metropolitan Rotterdam is dominated by power generation (55%), non-industrial combustion (19%) and road traffic (13%). In contrast, the majority of the CO<sub>2</sub> emissions in the Maasvlakte/glasshouse footprint comes from non-industrial combustion (94%) and only 5% of the emissions is from road traffic. The port-Botlek footprint is dominated by industrial production (62%) and energy production (31%) emissions, which have a low CO:CO<sub>2</sub> emission ratio. The corresponding ratios vary over time due to temporal changes in the dominant source types. This makes interpretation of  $\Delta\text{CO}:\Delta\text{CO}_2$  in the observations a challenging task. Nevertheless, the fractions of CO and CO<sub>2</sub> from fossil fuel emissions are strongly correlated ( $R^2 > 0.5$  for October 2014-March 2015), which suggests that CO can provide additional information about CO<sub>2</sub>.

To test this method for our case study we first examine how well the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratios correspond to the ratios suggested by the NER. Figure 2.8 (left) shows the smooth Gaussian fit of the probability density functions of observed  $\Delta\text{CO}:\Delta\text{CO}_2$  for the three footprints. The corresponding NER ratios are given by arrows and their standard deviation by black horizontal lines. For the metropolitan Rotterdam footprint the mean observed ratio is 3.9 ppb/ppm which corresponds well to the NER ratio of 3.4 ppb/ppm. The Maasvlakte/glasshouse footprint has a mean observed ratio of 2.0 ppb/ppm, while

the NER suggests a ratio of 4.7 ppb/ppm. The selected hours with correct wind speed and direction for this footprint are mainly summertime data, when the non-industrial combustion that dominates the CO<sub>2</sub> emissions is limited and road traffic becomes relatively more important. This results in a higher emission ratio. However, the observations suggest that the impact of non-industrial combustion is underestimated. Probably, these emissions are mainly from the glasshouse area rather than residential heating and the heating of glasshouses has less seasonal variation. The NER database has no distinct emission category for glasshouse heating and therefore the same seasonal cycle is applied as for residential heating, causing this discrepancy with the observed ratio. Moreover, for port-Botlek the observed mean ratio is 3.8 ppb/ppm, which is much higher than the 1.1 ppb/ppm given by the NER database. A possible explanation could be that the low ratio in the NER is dominated by point sources (industrial production and power generation) that are relatively devoid of CO, while the observations are more affected by the (surface) area sources that have a higher emission ratio. These ratios are lower than what was found in previous studies due to relatively complete (clean) combustion (Djuricin et al., 2010; Lopez et al., 2013; Turnbull et al., 2015).



**Figure 2.8: Smooth Gaussian fit of probability density functions of  $\Delta\text{CO}:\Delta\text{CO}_2$  for the three footprints based on the observations. Black arrows indicate the emission ratio for co-sampled times according to the NER emission database while black horizontal lines illustrate their standard deviation (left); Scatter plot of the observed  $\Delta\text{CO}_2$  and those estimated from the NER reported ratios multiplied with observed  $\Delta\text{CO}$  and the 1:1 line in black (right, colours correspond to the left graph).**

Additionally, the Maasvlakte/glasshouse footprint shows a much narrower distribution in the observed ratios than the other two footprints. It is likely that the emissions from the Maasvlakte/glasshouse footprint are better mixed by the time they reach Zweth due to the larger travel distance and the lack of large sources along the transect after leaving the port. In contrast, the other two footprints include several sources close to the Zweth measurement site. Due to temporal variations in the local atmospheric conditions and in the emissions Zweth can be affected by different source types at different times. For

example, during the morning rush hour road traffic could be the dominant source type, while at another moment point source emissions affect the measurements. This results in a wider range of the  $\Delta\text{CO}:\Delta\text{CO}_2$  PDF that is not fully captured by the NER database, with the port-Botlek footprint most clearly illustrating this difference.

Finally, we calculate  $\Delta\text{CO}_2$  values for each footprint based on the  $\Delta\text{CO}$  observations and the NER emission ratios (Fig. 2.8, right). For the metropolitan Rotterdam footprint the  $\Delta\text{CO}_2$  values are scattered around the 1:1 line, as was expected based on the good match with the observed emission ratio. However, the  $\Delta\text{CO}_2$  is generally overestimated for the port-Botlek footprint using this approach, while the Maasvlakte/glasshouse footprint shows an underestimation. The correlation coefficient  $r$  for all data combined is 0.37, while for only the metropolitan Rotterdam data  $r$  is 0.61.

## 2.4 Discussion and conclusions

Over the past years, several studies have pointed out the importance of in-situ observations to verify urban emissions of greenhouse gases (Ciais et al., 2014; McKain et al., 2012; Zimnoch et al., 2010) and they attempted to reach this goal using a wide range of methods; from ground-based in-situ observations to aircraft campaigns and from flux measurements to remotely sensed atmospheric concentrations (e.g. (Brioude et al., 2012; Font et al., 2014; Järvi et al., 2012; Lauvaux et al., 2013; Mays et al., 2009; Silva et al., 2013)). These studies are often complemented with an extensive and expensive urban  $\text{CO}_2$  monitoring network.

To test a more cost-efficient method we used only 2 continuous measurements, one on each side of an urban area. This method is relatively simple and therefore suitable to extend to other regions in the world. Using a mass balance approach and assuming a neutral to well-mixed boundary layer with constant height the gradient between the two sites was used to estimate the emissions along this transect. This method has been used before by several studies, but only with aircraft measurements rather than ground-based observations (Caulton et al., 2014; Karion et al., 2013; Mays et al., 2009; Peischl et al., 2015). The advantage of aircraft measurements is that there is more information on the vertical profile of concentrations and the horizontal extent of the plume, including the wind field along the transect. But our continuous observations give us a much larger dataset, which allows us to find variations in the gradient over time due to emission variations and select subsets from the database based on atmospheric conditions.

The objective of this study was to determine whether continuous observations of CO and  $\text{CO}_2$  at two sites (Zweth and Westmaas) can provide interesting information about the different source sectors and whether they can be used for an initial estimate of the  $\text{CO}_2$  fluxes. The results indicate that our observations are indeed affected by urban emissions and that Westmaas is suitable as background station for different wind sectors. The mass balance approach gives good results for the aligned cases. For non-aligned cases

additional estimates and assumptions are needed, increasing the uncertainty of the flux estimates. For our case study the non-aligned flux estimates show reasonable agreement with the emission database because for the selected wind sectors Westmaas seems to be relatively unaffected by local fluxes. Previous flux estimates for large cities range from 175-625 kg km<sup>-2</sup> hr<sup>-1</sup> for Paris (Bréon et al., 2015) and 745 kg km<sup>-2</sup> hr<sup>-1</sup> for Helsinki (Järvi et al., 2012) to 3600 kg km<sup>-2</sup> hr<sup>-1</sup> for Houston, 4000 kg km<sup>-2</sup> hr<sup>-1</sup> for Los Angeles (Brioude et al., 2013), and 7300-16500 kg km<sup>-2</sup> hr<sup>-1</sup> for Greater London (Font et al., 2014). Our  $F_{\text{obs}}$  ranges from about 2600-6800 kg km<sup>-2</sup> hr<sup>-1</sup>, whereas  $F_{\text{NER}}$  takes values of 2250-8800 kg km<sup>-2</sup> hr<sup>-1</sup>. The Rotterdam footprint fluxes are thus well in line with those of other large cities.

Since the method used in this study gives a good first impression of the emissions, it can be used to explore an area where little is known about the emissions. Based on the findings an efficient measurement strategy can be developed to estimate fluxes from important source sectors. Based on the results in this paper we recommend using aligned flux estimates to reduce the uncertainty. If sufficient resources are available all gradients can be monitored continuously by installing multiple instruments. Another, more cost-efficient method is to assign one fixed measurement site and use a mobile instrument that can be relocated depending on the wind direction. Moreover, our footprints cover only part of the source sectors and the metropolitan Rotterdam footprints also covers part of the port. This illustrates the importance of finding good locations for the observations when flux estimates for particular source sectors are made. Both the distance from the source sector (which determines the size of the footprint) and the location with respect to the dominant wind direction and upwind fluxes should be taken into account. However, in densely build areas often practical limitations exist that reduce the number of possible observational sites. Nevertheless, our results show that one upwind site that is relatively unaffected by large local emissions and an upwind site that is in between the important source sectors could already provide valuable information.

The main difficulty related to the mass balance approach is that it is sensitive to the choice of boundary layer height ( $h$ ). We used monthly values for  $h$  based on observations, but in reality there would be a lot of temporal variability depending on the synoptic conditions. A change in  $h$  affects both the  $F_{\text{obs}}$  and the range of its confidence interval. Based on the variability of  $h$  found in previous research (Lee and De Wekker, 2016), we estimate the uncertainty to be about 30%. In addition, our biogenic flux estimates are quite uncertain, yet a 50% decrease of these fluxes only results in maximum 2% decrease of the fossil fuel flux estimates. Moreover, the mass balance approach assumes that emissions are well-mixed by the time they reach the upwind site. To favour such well-mixed conditions we used the minimum wind speed and afternoon data criteria. However, we think that well-mixed conditions are not always reached, especially for stack emissions and sources that emit close to the measurement site. Sources with a smaller spatial extent and a smaller distance to the measurement site could be mixed throughout a much thinner layer than the actual boundary layer height and, as illustrated before, this could

have an impact of several percent on the flux estimates. Also, we rounded-off the travel time to full hours in order to correspond to hourly averaged measurements. We estimated the error introduced by this approximation by using the concentration gradient an hour after the selected hour. We found that the error is limited (1% for the port-Botlek footprint) because of the relative constancy of the concentration gradient for the selected moments. In addition, we assumed neutral stability to establish the footprint dimensions, while in reality the footprint could have a different width. This introduces an error especially in heterogeneous source sectors. For example, an increase of the footprint width of 10% will result in an increase of the  $F_{\text{NER}}$  of 180%. In contrast, for the other two footprints the flux changes only with a few percent (-8 and +4%). Finally, we assumed that the air mass moving from Westmaas to Zweth is not affected by other air masses (dilution) or entrainment. However, entrainment of air with a lower  $\text{CO}_2$  concentration causes the downwind observations to be lower than expected based on the fossil fuel emissions and this dilution effect has shown to be significant compared to biogenic uptake (Vilà-Guerau de Arellano et al., 2004).

All these assumptions increase the uncertainty in our flux estimates, but are necessary due to a lack of information about these processes. Although the observations provide useful information on important source sectors and their anthropogenic emissions, a chemical transport model can quantify the effect of dilution, entrainment, biogenic fluxes, et cetera and improve the flux estimates by providing a value for  $h$ . Model simulations for the same cases including explicit biospheric contributions, as well as plume dispersion will be done in a follow-up study. In addition, we aim to extend our dataset by continuing this monitoring framework, which will further contribute to the uncertainty reduction.

In addition, our data inclusion criteria, although underpinned by our observations, are somewhat arbitrary. Nevertheless, careful examination of the Zweth-Westmaas gradients for varying selection criteria and even with the full dataset shows that the signals identified in this paper are consistent. Only the magnitude of the signals is sensitive to our choice of selection criteria (Appendix B shows 95% confidence interval).

Finally, we illustrated that the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio agrees well with the  $\text{CO}:\text{CO}_2$  emission ratio in the emission database for metropolitan Rotterdam, albeit with a large observed range. The observed  $\Delta\text{CO}$  can therefore be used to reconstruct fossil fuel  $\text{CO}_2$  concentrations. However, the presence of point sources had a large impact on the estimated ratio for the port-Botlek footprint. Whereas the NER ratio is dominated by the low point source emission ratios, the observations are mostly affected by area source emissions at the surface with a much higher  $\text{CO}:\text{CO}_2$  ratio. If we exclude the point sources in the port-Botlek footprint the NER indeed gives a  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio of 8.8 ppb/ppm compared to 1.1 ppb/ppm with point sources. This large difference presents a good opportunity to use our monitoring sites to specifically target emissions from this area. Although the metropolitan Rotterdam footprint also includes several point sources, these are less dominant over the area sources than for the port-Botlek footprint and the

agreement between observations and the NER is better. The NER also seems to underestimate the emissions from glasshouse heating, while overestimating their seasonality due to the way the database was constructed. Moreover, the observations show a wide range in ratios, which is related to the temporal variability in the emissions and meteorological conditions. For example, point sources will only affect the measurements from time to time, resulting in a low observed ratio. In contrast, during rush hour traffic emissions will be dominant and the observed ratio increases drastically.

The direct calculation of  $\Delta\text{CO}_2$  from the observations is complicated by the presence of biogenic  $\text{CO}_2$  fluxes. We have included a correction for this by estimating the biogenic  $\text{CO}_2$  fraction based on the flux estimates with and without biogenic fluxes. Without this correction the observation-based ratios would be 1, 7 and 11% higher for the metropolitan Rotterdam, port-Botlek and Maasvlakte/glasshouse footprint, respectively. Miller et al. (2012) underpin our finding that exact quantification of  $\Delta\text{CO}$  and  $\Delta\text{CO}_2$  is challenging due to the impact of chemistry and biogenic contributions, although this is especially relevant during summer months. Indeed, it has previously been concluded that the use of CO to constrain fossil fuel  $\text{CO}_2$  emissions is limited by the uncertainty and variability in CO: $\text{CO}_2$  emission ratios (Turnbull et al., 2006; Vardag et al., 2015). Nevertheless, several studies have also shown that regular  $^{14}\text{C}$  observations can be used to quantify fossil fuel  $\text{CO}_2$  emissions and calibrate  $\Delta\text{CO}$  (Levin and Karstens, 2007; Turnbull et al., 2006; Vardag et al., 2015; Vogel et al., 2010). Using plume modelling could also be useful to get more spatial detail in CO: $\text{CO}_2$  emission ratios, especially when it comes to stack plumes with lower emission ratios and their impact on the observations.

The method presented in this study has proven useful and it is recommended for further exploitation in other areas to identify important source sectors and test its general applicability. The final derived fluxes are within -23% - + 15% of the estimates from a state of the art bottom-up inventory and for every footprint within the 95% confidence interval (Table 2.2). For many urban regions or cities this would be an enormous improvement in their emission estimate. However, to verify emission reductions, the uncertainty in the flux estimates will need to be reduced significantly. Therefore, we recommend future studies to explore the use of a high-resolution transport model and an inverse modelling approach. Such method could improve the flux estimates by including additional information on processes that remain unquantified with our method, such as entrainment, dilution and additional (biogenic) sources and sinks of  $\text{CO}_2$ . Such a full modelling framework presents a more expensive and time-consuming methodology, while the easy-to-use mass balance approach gives a reasonable first, rough estimate. Moreover, a transport and/or plume model can be useful for source attribution using tracer ratios and footprint analysis. Furthermore, observations of additional tracers, such as  $^{14}\text{C}$  and  $\text{O}_2/\text{N}_2$ , can help to separate  $\text{CO}_2$  fossil fuel fluxes from biogenic sources and sinks (Turnbull et al., 2006). In combination with high quality datasets, as presented here, we expect this would lead to improved flux estimates and monitoring of  $\text{CO}_2$  emissions for



heterogeneous urban-industrial landscapes, which are responsible for the majority of the global anthropogenic CO<sub>2</sub> emissions.

Appendix 2A: Methodology

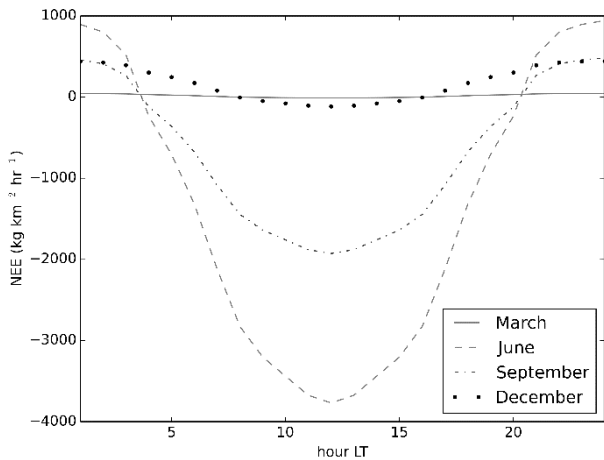


Figure A1: Mean daily cycle of the net ecosystem exchange for several months.

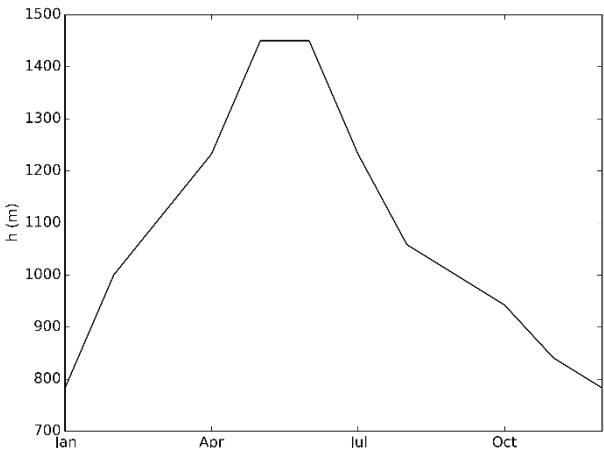


Figure A2: Annual cycle of the maximum boundary layer height.

## Appendix 2B: Polar plots

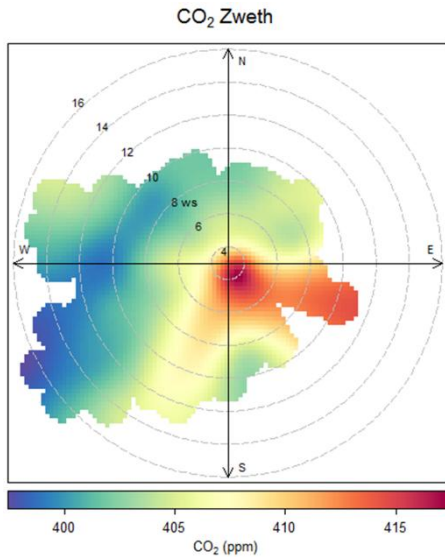


Figure B1: Median CO<sub>2</sub> mole fraction observed at Zweth as a function of wind direction (angular coordinate) and wind speed (radial coordinate).

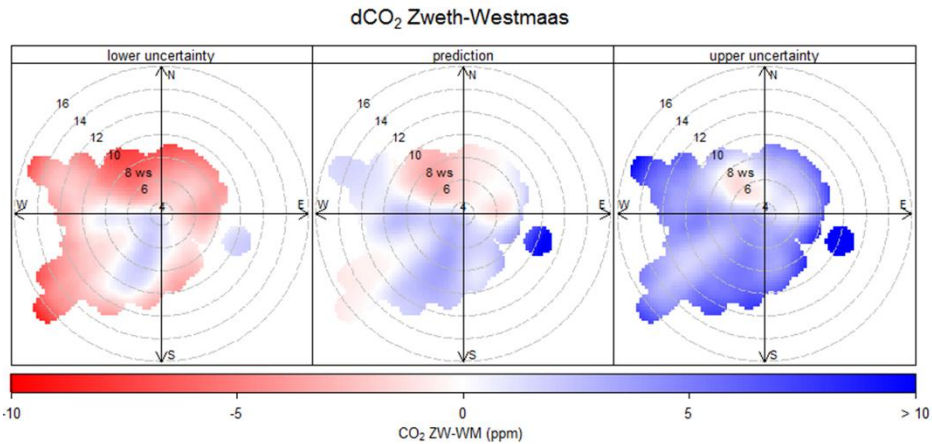


Figure B2: Difference between Zweth and Westmaas CO<sub>2</sub> mole fractions as a function of wind direction (angular coordinate) and wind speed (radial coordinate) for a minimum bin size of 2 and the lower and upper limit of the 95% confidence interval.



A thick dark blue vertical bar runs along the left edge of the page. A light teal arrow points to the right from this bar, containing the chapter title. In the bottom left corner, several thin, curved lines in dark blue and light grey sweep upwards and to the right.

## Chapter 3

# **A multi-model approach to monitor emissions of CO<sub>2</sub> and CO from an urban–industrial complex**

*This chapter is published as: Super, I., Denier van der Gon, H. A. C., Van der Molen, M. K., Sterk, H. A. M., Hensen, A., and Peters, W.: A multi-model approach to monitor emissions of CO<sub>2</sub> and CO from an urban–industrial complex, Atmos. Chem. Phys., 17, 13297–13316, 10.5194/acp-17-13297-2017, 2017.*



## Abstract

Monitoring urban-industrial emissions is often challenging, because observations are scarce and regional atmospheric transport models are too coarse to represent the high spatiotemporal variability in the resulting concentrations. In this paper we apply a new combination of a Eulerian model (WRF with chemistry) and a Gaussian plume model (OPS). The modelled mixing ratios are compared to observed CO<sub>2</sub> and CO mole fractions at four sites along a transect from an urban-industrial complex (Rotterdam, Netherlands) towards rural conditions for October–December 2014. Urban plumes are well-mixed at our semi-urban location, making this location suited for an integrated emission estimate over the whole study area. The signals at our urban measurement site (with average enhancements of 11 ppm CO<sub>2</sub> and 40 ppb CO over the baseline) are highly variable due to the presence of distinct source areas dominated by road traffic/residential heating emissions or industrial activities. This causes different emission signatures that are translated into a large variability in observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratios, which can be used to identify dominant source types. We find that WRF-Chem is able to represent synoptic variability in CO<sub>2</sub> and CO (e.g. the median CO<sub>2</sub> mixing ratio is 9.7 ppm (observed) against 8.8 ppm (modelled)), but it fails to reproduce the hourly variability of daytime urban plumes at the urban site ( $R^2$  up to 0.05). For the urban site, adding a plume model to the model framework is beneficial to adequately represent plume transport especially from stack emissions. The explained variance in hourly, daytime CO<sub>2</sub> enhancements from point source emissions increases from 30 % with WRF-Chem to 52 % with WRF-Chem in combination with the most detailed OPS simulation. The simulated variability in  $\Delta\text{CO}:\Delta\text{CO}_2$  ratios decreases drastically from 1.5 to 0.6 ppb ppm<sup>-1</sup> which agrees better with the observed standard deviation of 0.4 ppb ppm<sup>-1</sup>. This is partly due to improved wind fields (increase in  $R^2$  of 0.10), but also due to improved point source representation (increase in  $R^2$  of 0.05) and dilution (increase in  $R^2$  of 0.07). Based on our analysis we conclude that a plume model with detailed and accurate dispersion parameters adds substantially to top-down monitoring of greenhouse gas emissions in urban environments with large point source contributions within a ~10 km radius from the observation sites.

### 3.1 Introduction

Cities are major contributors to anthropogenic CO<sub>2</sub> and air pollutant emissions (Brioude et al., 2013; Turnbull et al., 2015; Velasco et al., 2014). Both monitoring and modelling of urban/regional concentrations of CO<sub>2</sub> and co-emitted air pollutants, such as CO and NO<sub>x</sub>, has therefore received a lot of attention (Brioude et al., 2013; Font et al., 2014; Huszar et al., 2016; Lac et al., 2013; Mays et al., 2009; McKain et al., 2012; Rayner et al., 2014; Ribeiro et al., 2016; Silva et al., 2013; Tolk et al., 2009; Wunch et al., 2009; Zhang et al., 2015). Since current emission inventories at small scales contain substantial uncertainties (Pouliot et al., 2012; Vogel et al., 2013), data assimilation has been applied to urban environments in order to better quantify fossil fuel fluxes. However, modelling urban atmospheric composition remains challenging as the urban environment is complex in both the emission landscape and atmospheric transport. This means that to independently estimate urban emissions from atmospheric observations, urban inversions require a detailed and accurate transport model that allows the mismatch between model and observations to be attributed to errors in the emission inventory, rather than to transport errors (Boon et al., 2016). Previous inversion studies relied heavily on a strict data selection to favour well-mixed conditions with more reliable model output, which results in very small data sets and therefore increased uncertainty on the estimated emissions (Bréon et al., 2015; Brioude et al., 2013). This could be overcome by improving the model representation of urban transport, taking into account that the model requirements are strongly dependent on the type of observation site used in the inversion. In this paper we aim to construct a promising observation and modelling framework to quantify the CO<sub>2</sub> budget of an urban area by addressing two important questions in the context of inverse modelling at the urban scale.

The first question is what type of measurement location (urban vs. rural) can best be used to monitor urban fluxes. Generally, urban sites are most strongly exposed to nearby (<1 km) fluxes and therefore show a large variability (Bréon et al., 2015; Lac et al., 2013). In contrast, rural sites show a much smaller response to urban emissions due to the small range of wind directions at which the site is affected by the urban area. Moreover, the dilution of urban plumes increases with distance (Calabrese, 1990; Finn et al., 2007) and the observed signal at the rural site can be small. Another consideration is that near-ground measurements, as commonly found in cities, are highly influenced by local sources (<100 m) that mask the overall urban signal. Boon et al. (2016) suggested that, even if strict data selection is applied, the usefulness of such sites in inversions with high-resolution Eulerian models (1–10 km) might be limited. Together, these papers suggest that a useful measurement location should be just downwind of an urban area relative to the dominant wind direction at a distance that ensures enough exposure to the urban plume and limits model errors due to large heterogeneity and local emissions. We will examine a transect of measurement sites to see which site best matches this criterion.

The second question we address is what type of modelling framework is best capable of explaining urban transport and the resulting mole fractions at the measurement sites. Since the measurement location determines the level of spatiotemporal variation that can be observed in the concentrations, it also determines the requirements imposed on the modelling framework. In atmospheric composition modelling both Eulerian and Lagrangian (plume, puff or Gaussian) models are used, or a combination of both (Kim et al., 2014; Korsakissok and Mallet, 2010b). Eulerian models use a grid that can be adapted to cover either small or large areas at different resolutions and are therefore widely used. However, Eulerian models assume that trace gasses are instantly mixed within individual grid boxes, which may enhance dispersion in the horizontal and vertical. The resulting errors in transport and mixing are reflected in unrealistic concentrations (Karamchandani et al., 2011; Tolk et al., 2009). The magnitude of the concentration error depends on the heterogeneity of the emissions and the grid resolution (Tolk et al., 2008). A plume model improves the description of horizontal and vertical mixing and can account for higher spatial heterogeneity of emissions and concentrations. The use of such models has proven useful for both inert and reactive species, and point and line sources at local/urban scales (Briant and Seigneur, 2013; Korsakissok and Mallet, 2010a, b; Rissman et al., 2013; Vinken et al., 2011). However, a plume model is usually only applied to local sources to reduce computational expenses. It therefore does not resolve the impact of remote emissions and synoptic transport. So, when assessing the carbon balance of a whole city or larger areas, a combination of both models might be needed.

Oney et al. (2015) examined an extensive CO<sub>2</sub>, CH<sub>4</sub> and CO measurement network in combination with the FLEXPART-COSMO model. However, their framework focused on regional (~100–500 km), terrestrial fluxes. Several other studies focussed on urban scales (Boon et al., 2016; Bréon et al., 2015; Turnbull et al., 2015), but only few incorporated a Lagrangian model. For example, McKain et al. (2012) and Lauvaux et al. (2016) used a Lagrangian model to optimise urban fluxes of CO<sub>2</sub>, while Brioude et al. (2013) compared simulated FLEXPART CO<sub>2</sub>, CO and NO<sub>x</sub> concentrations to small observational datasets from seven flights over Los Angeles. Here, we compare and combine simulations with two different models: the Eulerian WRF-Chem model and the segmented Gaussian plume model OPS. The Gaussian plume model is used here specifically to transport point source emissions. The model output is compared to continuous observations of CO<sub>2</sub> and CO at several measurement sites along an urban-to-rural transect. We included CO, because this species can act as a useful tracer for source attribution. We use the Rijnmond area (The Netherlands) including the city of Rotterdam as our case study, which is surrounded by scattered urban, agricultural, and rural areas. We chose this area because of the availability of a 1x1 km<sup>2</sup> emission inventory and its complex combination of residential, transport (including shipping), greenhouse and industrial activities. This makes Rijnmond an interesting test case, albeit not a simple one.



This paper starts with a description of the case study (Sect. 3.2.1), the modelling framework (Sect. 3.2.2–3.2.5), and a summary of data selection criteria and methods (Sect. 3.2.6). Subsequently, we examine the ability of our measurement sites to detect urban signals, and demonstrate the added value of both urban and semi-urban sites (Sect. 3.3.1). Sect. 3.3.2 examines the ability of WRF-Chem to represent the urban signals at the measurement sites. Finally, we discuss the advances made by implementing the Gaussian OPS plume model (Sect. 3.3.3) and we examine the relative importance of improved meteorological conditions and source representation in Sect. 3.3.4. Our results lead to recommendations for monitoring and modelling of urban atmospheric composition in Sect. 3.4.

## 3.2 Methods

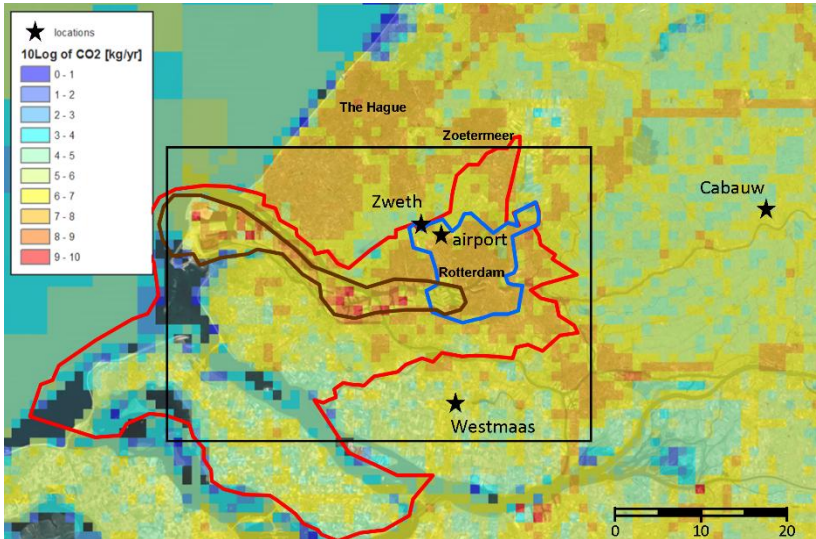
### 3.2.1 Study area and measurements

We take the Rijnmond area (Fig. 3.1) in the Netherlands for our case study in which Rotterdam is the major urban area (625.000 inhabitants). The area is situated in flat terrain near the west coast of The Netherlands and includes a large harbour and industrial area. The bottom-up estimated emissions in this area are about 35 Mt CO<sub>2</sub> and 48 kt CO in 2012 (Netherlands PRTR, 2014). In the port area, over three times more CO<sub>2</sub> is emitted than in the city of Rotterdam. In contrast, more than 60 % of all CO is emitted in the city of Rotterdam. The reason for this difference is that emissions within the city are dominated by road traffic, which emits relatively much CO (CO:CO<sub>2</sub> emission ratio of almost 17 ppb ppm<sup>-1</sup>). The principal source of CO<sub>2</sub>, namely energy production and industrial processes, is mainly found in the port area and barely emits any CO (CO:CO<sub>2</sub> emission ratio of less than 1 ppb ppm<sup>-1</sup>). The CO<sub>2</sub> emissions are therefore dominated by point sources (~80 %).

We have installed two measurement sites to monitor CO<sub>2</sub> and CO mixing ratios 15 km south (Westmaas, 51.79° N, 4.45° E) and 7 km northwest (Zweth, 51.96° N, 4.39° E) of the city centre with an inlet at 10 m a.g.l. We consider Zweth to be an urban site which is highly affected by urban emissions. Westmaas functions as a background site close to – but not within – the city and it is usually located upwind of the major source areas. Therefore, Westmaas provides information on the air mass entering the Rijnmond area and we only use this site to validate the large-scale patterns in WRF-Chem. These measurements have been described in more detail by Super et al. (2017b). At Rotterdam-The Hague airport (Fig. 3.1) meteorological observations are made, which we also use for transport model validation purposes.

We include two additional, more remote, sites in our framework. The Cabauw site (51.97° N, 4.93° E) is situated 32 km east of the centre of Rotterdam and is considered a semi-urban site (Van der Laan et al., 2016; Vermeulen et al., 2011). This means the sampled air masses are influenced by urban emissions, but less often than a truly urban location. CO<sub>2</sub> is measured at several heights (20, 60, 120 and 200 m a.g.l.) along a 200 m

tall tower by the Energy research Centre of the Netherlands (ECN). CO is measured at ground level (2.5–4 m a.g.l.) by the National Institute for Public Health and the Environment (RIVM). Another observation site is located at Lutjewad (53.40° N, 6.35° E), close to the coast in the north of the Netherlands. At this rural site, CO and CO<sub>2</sub> mixing ratios are observed at 60 m a.g.l. (Van der Laan et al., 2009b; Van der Laan et al., 2016). These four stations together describe a transect from the city towards rural areas.



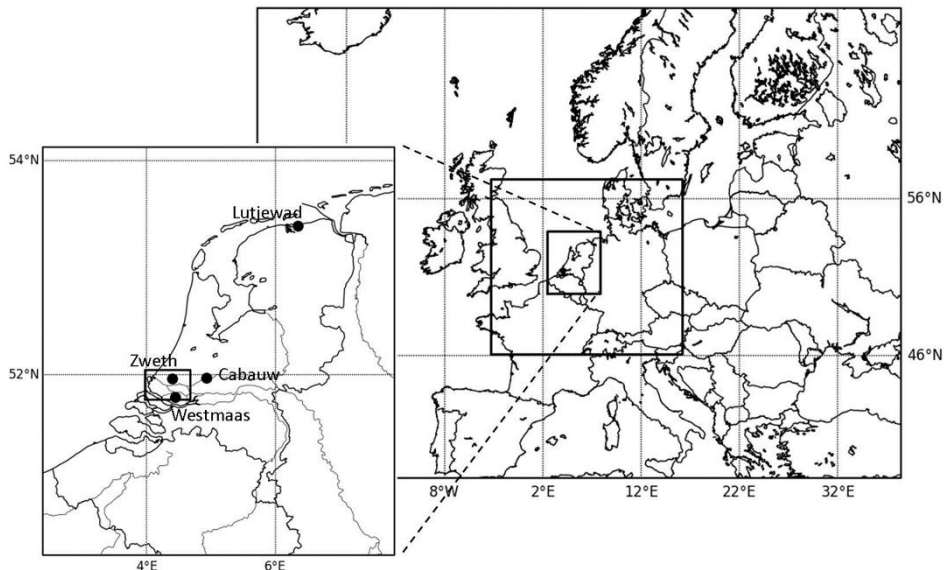
**Figure 3.1:** CO<sub>2</sub> emission map of the Rijnmond area (red outline), including the city of Rotterdam (blue outline) and the port area (brown outline); the observation sites are indicated with black stars (Lutjewad is shown in Fig. 3.2). The boundaries of domain 4 in WRF-Chem are indicated by the black square. Source: Netherlands PRTR (2014).

For the Cabauw CO<sub>2</sub> measurements we selected the 60 m level. On average the CO<sub>2</sub> mixing ratios are similar at all levels during well-mixed daytime conditions (Vermeulen et al., 2011), but a large gradient is observed for stable conditions when the 20 m level is highly affected by surface fluxes surrounding the tower. Similarly, Turnbull et al. (2015) suggested that measurements closer to the surface are more sensitive to local fluxes and therefore a higher level than 20 m is more suitable to obtain information on more remote fluxes. We choose the 60 m level observations to be able to compare easily to the Lutjewad site. However, a higher level could have been used without affecting our conclusions.

### 3.2.2 Eulerian model

The Eulerian model used in this study is WRF-Chem V3.2.1 (Skamarock et al., 2008). For its initial and boundary conditions we use meteorological fields from the National Centers for Environmental Prediction (NCEP) Final (FNL) Operational Global Analysis (National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000) at 1x1° horizontal resolution and a temporal resolution of 6 hours. We

define four 2-way nested domains (Fig. 3.2) which have a horizontal resolution of 48x48, 12x12, 4x4 and 1x1 km respectively, and a vertical resolution of 29 eta levels with the lowest model layer 40 m deep and a total of 8 levels in the lowest 1 km. The outer domain is situated over Europe. Domains 2–4 zoom in on the Rijnmond area in the southwest of the Netherlands. Based on previous studies over the Netherlands (Bozhinova et al., 2014; Daniels et al., 2016; Steeneveld et al., 2014), we have used the Yonsei University (YSU) boundary layer scheme (Hong et al., 2006), the Dudhia scheme for shortwave radiation (Dudhia, 1989), the Rapid Radiation Transfer Model (RRTM) as longwave radiation scheme (Mlawer et al., 1997), and the Unified Noah Land-Surface Model as the surface physics scheme (Ek et al., 2003). We also used the single-layer urban canopy model (UCM) to account for changes in roughness length and heat fluxes in the urban environment (Chen et al., 2011), although the impact of the UCM model on simulated mixing ratios is very small in our domain.



**Figure 3.2:** Location of the domains is indicated with squares. The horizontal resolutions of the domains are (from outer to inner domain): 48x48 km, 12x12 km, 4x4 km and 1x1 km. Black circles represent the observation sites.

The CO<sub>2</sub> initial and boundary conditions are taken from the 3D mole fractions from CarbonTracker Europe (Peters et al., 2010). The CarbonTracker 3D fields have a horizontal resolution of 1x1° and 34 vertical levels. Therefore, they are both horizontally and vertically interpolated onto the WRF-Chem grid. The CO initial and boundary conditions are calculated with IFS-MOZART (Flemming et al., 2009) and obtained from the Monitoring Atmospheric Composition and Climate (MACC) project. The boundary conditions are updated every 6 hours (only for the outer domain).

We have implemented a CO<sub>2</sub> budget based on the methodology used by Bozhinova et al. (2014), described in Eq. (1).

$$X_{CO_2,obs} = X_{CO_2,lsbg} + X_{CO_2,ff} + X_{CO_2,bf} + X_{CO_2,p} + X_{CO_2,r} \quad (1)$$

where the indices express the origin of CO<sub>2</sub>: obs – total observed concentration at a particular location, lsbg – large-scale background mole fraction, ff – fossil fuels, bf – biofuels, p – photosynthetic uptake, r – ecosystem respiration. Similar to the original study of Bozhinova et al. (2014), we omitted the stratosphere-troposphere exchange and ocean fluxes and assume they are accounted for in the large-scale background. With Eq. (1) we thus only consider regional contributions to the carbon budget in addition to the large-scale background. In the model, any change in the large-scale background CO<sub>2</sub> mole fraction ( $X_{CO_2,lsbg}$ ) is only caused by advection and exchange at the domain boundaries.

In addition, we added the CO budget to WRF-Chem following Eq. (2). The main sources of CO are fossil fuel combustion and oxidation of hydrocarbons (US EPA, 1991). Several scholars have argued that the hydrocarbon oxidation term is important for the large-scale background CO concentration (Gerbig et al., 2003; Griffin et al., 2007; Hudman et al., 2008), contributing a significant percentage to the total CO burden. Yet, these studies were all based on summer time measurements and under conditions favourable for photochemistry. Photochemical oxidation is likely to be less important in the winter months considered here. Moreover, Griffin et al. (2007) found the CO fraction from local anthropogenic emissions to dominate at measurement sites. We assume this is also valid in the urban-industrial environment of our case study. We nevertheless consider that this introduces an uncertainty in the modelled CO mixing ratios. For summer time studies the oxidation term might be significant.

The main sink of CO is the reaction with the hydroxyl radical (chemical loss term  $L$ ), which we account for with a simple first order loss term. We assume steady-state, i.e. the OH concentration is taken as a constant ( $10^6$  molecules cm<sup>-3</sup>). This results in a lifetime for CO of about 2 months at mid-latitudes (Jacob, 1999) during the winter months used in our study:

$$X_{CO,obs} = X_{CO,lsbg} + X_{CO,ff} + X_{CO,bf} + X_{CO,L} \quad (2)$$

The different contributions in Eq. (1) and Eq. (2) are separated as different additive tracers (i.e. labelled) in the WRF-Chem simulations.

### 3.2.3 Gaussian plume model

The plume dispersion model OPS (Operational Priority Substances) is a segmented Gaussian plume model that calculates the transport, dispersion, chemical conversion and deposition of pollutants (Sauter et al., 2016; Van Jaarsveld, 2004). It is used to calculate large-scale, yearly averaged concentration and deposition maps for the Netherlands at 1x1 km<sup>2</sup> resolution. It was initially developed to model dispersion of pollutants like particulate

matter and ammonia, but has also been used to study the dispersion of pathogens (Van Leuken et al., 2015).

In this paper we use the so-called "short-term" version of this model (version 10.3.5), which contains mostly the same parameterisations as the "long-term" model described by Sauter et al. (2016). The short-term model provides hourly concentrations at receptors that can be individual sites, or across a gridded domain. The model keeps track of a trajectory forward in time, for which plumes consist of so-called segments, taking into account time-varying transport over longer distances (e.g. changes in wind direction and dispersion). If for a time step a specific plume affects the receptor, a Gaussian plume formulation is used to calculate the concentration caused by that source based on the true travel distance along the trajectory.

The OPS model uses primary meteorological variables which are measured by the Royal Dutch Meteorological Institute, and calculates secondary variables such as boundary layer height and friction velocity, but also the turning of the wind with height and a vertical wind profile. Primary meteorological variables are spatially interpolated over the Netherlands to  $10 \times 10 \text{ km}^2$  using 19 observation sites with a weighing factor depending on the distance to the grid point. The variables are subsequently averaged over a pre-defined area (for more information see Sauter et al. (2016)). The use of observed meteorology in OPS versus model-calculated meteorology in WRF-Chem could result in an unfair comparison of the models, and we therefore replaced the primary parameters (temperature, humidity, wind speed, and wind direction) and the boundary layer height with those calculated by WRF-Chem. The secondary (dispersion) parameters are automatically also updated, since they are calculated from the primary parameters. Note that the meteorological conditions in OPS remain constant during each simulated hour and over a large region.

Although potentially the OPS model can be used for both area and point source emissions, we believe that point sources will benefit most from a more detailed description of dispersion as they are affected most by the instant dilution in a Eulerian model. When using OPS, we assume wet deposition plays no role due to the relative insolubility of  $\text{CO}_2$ , while dry deposition of  $\text{CO}_2$ , i.e. photosynthetic uptake, is accounted for by WRF-Chem (Eq. (1)). We do not simulate CO with the OPS model. The point source contribution to the total CO concentrations is very small and therefore the impact of OPS is limited.

#### **3.2.4 Emissions**

The fossil fuel and biofuel emissions for domains 1–3 in the WRF-Chem simulation are taken from the TNO-MACC III inventory for 2011 (Kuenen et al., 2014) and have a horizontal resolution of  $0.125 \times 0.0625^\circ$ . Fossil fuel and biofuel emissions for domain 4 in WRF-Chem are collected from the Dutch Emission Registration (Netherlands PRTR, 2014) and compiled by TNO (Netherlands Organization for Applied Scientific Research) to a  $1 \times 1$

km<sup>2</sup> emission map for the year 2012. In the OPS simulations we only include the point source emissions from domain 4 in WRF-Chem (hereafter referred to simply as the Rijnmond area).

**Table 3.1: Overview of SNAP categories and the vertical distribution of point source emissions in WRF-Chem.**

SNAP	Description	% of point source emissions per model layer [m above surface]				
		0–55	55–130	130–235	235–360	>360
1	Combustion in energy and transformation industries			18.5 %	42 %	39.5 %
2	Non-industrial combustion plants					
3	Combustion in manufacturing industry	12.2 %	37.3 %	46.2 %	4.3 %	
4	Production processes	12.2 %	37.3 %	46.2 %	4.3 %	
5	Extraction and distribution of fossil fuels					
6	Solvents and other product use					
7	Road transport					
8	Other mobile sources and machinery	100 %				
9	Waste treatment and disposal		16.5 %	44.5 %	39 %	
10	Agriculture					

The emissions are divided over ten SNAP emission categories, summarised in Table 3.1, which may include both area and point sources. We apply a temporal profile to the emissions by assigning hourly, daily and monthly fractions to the emissions per emission category (Denier van der Gon et al., 2011). In WRF-Chem, area source emissions are added to the lowest surface model level every hour. Point source emissions (only SNAP 1, 3, 4, 8 and 9) are given a simplified, fixed vertical distribution based on previous research with plume rise calculations (Bieser et al., 2011). These emissions are emitted at the heights shown in Table 3.1. OPS allows for more detailed point source characteristics and accounts for stack height and plume rise (based on heat content) per individual point source.

The biogenic (non-biofuel) CO<sub>2</sub> fluxes in WRF-Chem are generated as described by Bozhinova et al. (2014). The SiBCASA model (Schaefer et al., 2008) calculates monthly averaged 1x1° photosynthetic uptake ( $A_n$ ) and ecosystem respiration ( $R$ ) for nine different land use types. Combining the high-resolution land-use map of WRF-Chem with the SiBCASA fluxes gives us biogenic fluxes on the resolution of the WRF-Chem grid. The temporal resolution is enhanced by scaling the  $A_n$  and  $R$  at each WRF-Chem time step with modelled shortwave solar radiation ( $SW_{in}$  in W m<sup>-2</sup>) and 2m temperature ( $T_{2m}$  in K):

$$A_n = A_{n,f} \cdot SW_{in} \quad (3)$$

$$R = R_f \cdot 1.5^{(T_{2m}-273.15)/10} \quad (4)$$

where  $A_{n,f}$  is the monthly average photosynthetic flux divided by the monthly total incoming shortwave radiation ( $\text{mole CO}_2 \text{ km}^{-2} \text{ h}^{-1} (\text{W m}^{-2})^{-1}$ ), and  $R_f$  the monthly average respiration flux ( $\text{mole CO}_2 \text{ km}^{-2} \text{ h}^{-1}$ ) divided by the monthly total of the empirical function  $1.5^{(T_{2m}-273.15)/10}$  (unitless). This procedure was first described in Olsen and Randerson (2004). It neglects the impact of water stress, temperature and  $\text{CO}_2$  concentration on the photosynthetic uptake. Given that we consider only winter months in which photosynthesis is limited, we assume the error resulting from this simplification to be small.

### 3.2.5 Overview of simulations

We simulated a period of 3 months, October–December 2014. We choose this period because of the high data coverage at all measurement sites and to limit the impact of biogenic fluxes and hydrocarbon oxidation. We considered four simulations for  $\text{CO}_2$ , using two different model systems as described in Table 3.2. All simulations include the WRF-Chem contributions of  $X_{\text{CO}_2,\text{lsbg}}$ ,  $X_{\text{CO}_2,\text{p}}$ ,  $X_{\text{CO}_2,\text{r}}$  and  $X_{\text{CO}_2,\text{bf}}$  and  $X_{\text{CO}_2,\text{ff}}$  from area sources. Also, the first three simulations make use of meteorological conditions as simulated by WRF-Chem. Therefore, the simulations only differ in the representation of point source emissions in the Rijnmond area. To identify the importance of a correct representation of meteorological conditions we do an additional OPS simulation with interpolated meteorological observations (see Sect. 3.2.3). The simulations are designed to gradually increase the complexity of the point source representation towards more realistic point source contributions:

- In simulation 1 (WRF-Chem) the point sources are represented as area sources in WRF-Chem;
- In simulation 2 (WRF+OPS-area) the point sources are treated as area sources in OPS;
- In simulation 3 (WRF+OPS-point) the point sources are represented as true point sources with detailed source characteristics in OPS;
- In simulation 4 (WRF+OPS-point-obsmet) the point sources are represented as true point sources with detailed source characteristics in OPS and the meteorology in OPS is replaced by interpolated observations and OPS calculated boundary layer height.

In the WRF-Chem run we labelled the point source emissions from the Rijnmond area separately, so we can replace them by the OPS counterparts. The OPS model simulates concentrations directly at the measurement sites, whereas from WRF-Chem we extract the grid box average mixing ratio of the boxes in which the measurement sites are located.

**Table 3.2: Overview of the simulations, which model is used to calculate the urban plume mixing ratio from point sources in the Rijnmond area, how point sources are represented and the source of meteorological conditions.**

Simulation name	Point source contribution	Point source representation	Meteorological input
WRF-Chem	WRF-Chem	area	WRF-Chem
WRF+OPS-area	OPS	area	WRF-Chem
WRF+OPS-point	OPS	point	WRF-Chem
WRF+OPS-point-obsmet	OPS	point	observations

### 3.2.6 Baseline determination and data selection criteria

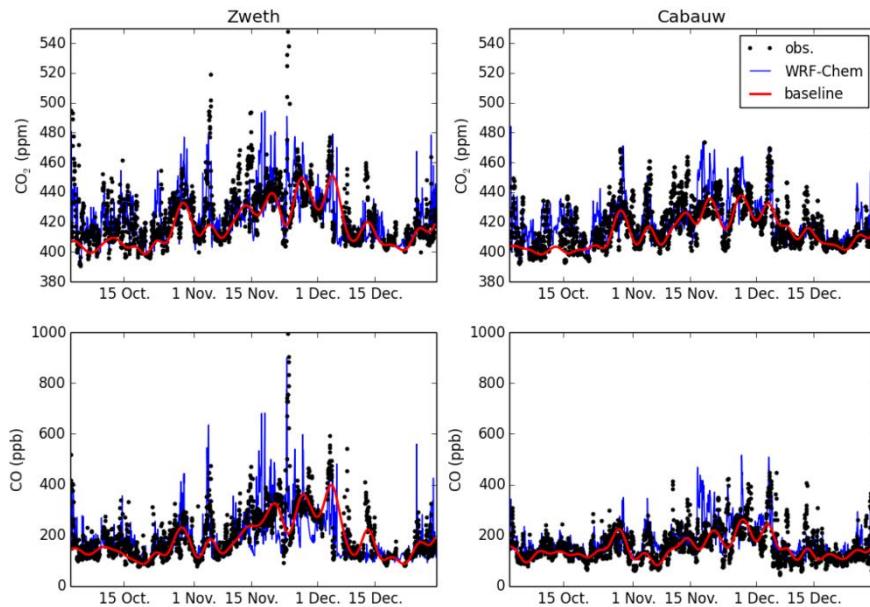
In this study we are especially interested in the contribution of urban emissions and the ability of the models to represent the transport of those emissions to the observation sites. However, the observed  $\text{CO}_2$  and CO mixing ratios are also affected by background signals and other fluxes. Therefore, in order to purely compare the transport of urban emissions, we need to separate the fossil fuel contribution from all other contributions. In the models we can separate the fossil fuel contribution  $X_{\text{CO}_2,ff}$  coming from the Rijnmond area (hereafter referred to as “urban plume”) from all other contributions (i.e.  $X_{\text{CO}_2,ff}$  from outside the Rijnmond area,  $X_{\text{CO}_2,lsbg}$ ,  $X_{\text{CO}_2,p}$ ,  $X_{\text{CO}_2,r}$ , and  $X_{\text{CO}_2,bf}$ , hereafter referred to as “baseline”) by using labelled tracers. To quantify the urban plume contribution to the total observed mixing ratio, we also need to subtract a baseline.

Previous studies have suggested various methods to calculate the baseline from observations, for example using a remote/upwind measurement site or statistical methods (e.g. (Djuricin et al., 2010; Lopez et al., 2013; Turnbull et al., 2015; Van der Laan et al., 2010)). An in-model comparison with WRF-Chem shows that Westmaas is a suitable background site for Zweth (Super et al., 2017b), but Westmaas gives a biased baseline estimate for the more remote sites (Cabauw and Lutfjewad) because of the interference of other sources and sinks along the transect from Rijnmond to the measurement site. Another suggested method is to subtract a smoothed representation of the original time series (Press et al., 1992; Super et al., 2017b; Thoning et al., 1989) which filters out variations below a certain cut-off time scale. For seasonal cycle smoothing for example, a typical cut-off value is 80 days. In our study however, the baseline needs to filter out synoptic variations across the domain and we therefore chose a cut-off time of 5 days. We tested this baseline definition by applying it to the WRF-Chem time series and comparing the resulting concentrations to the true WRF-Chem baseline based on the labelled tracers. We found satisfactory agreement ( $R^2$  is between 0.65 and 0.81 for both species at all 3 locations). Note that this method does not account for short pollution events bringing polluted air into the domain as only synoptic variations are captured.

To prevent any differences between model and observations resulting from the baseline selection, we choose to apply this subtraction of a smooth cycle method with a 5-day cut-off to both observations and our model time series at all measurement sites (see Fig. 3.3 for an example). The concentrations above the baseline are considered to be the urban



plume concentrations and are denoted  $\Delta\text{CO}_2$  and  $\Delta\text{CO}$ . Note that data points can also be below the baseline if clean air is advected and only a small fossil fuel contribution is calculated. We discard these data points, because we cannot accurately estimate the fossil fuel concentrations in those urban plumes.



**Figure 3.3:** Time series of modelled (WRF-Chem) and observed  $\text{CO}_2$  and CO mixing ratios at Zweth (left) and Cabauw (right). The observation-based baseline used in this study is also shown.

In all the analyses, we applied a wind sector selection to ensure that the observations are affected by emissions in the Rijnmond area rather than from other urban areas nearby. For Zweth we selected wind directions of 90–220 degrees, for Cabauw 230–270 degrees, and for Lutjewad 210–230 degrees. For Zweth we can also separate between signals from the residential area (90–150 degrees, Zweth-city) and industrial area (160–220 degrees, Zweth- port). Wind direction observations at Rotterdam airport are used for this purpose. Additionally, a daytime selection criterion (8:00–17:00 LT) is applied to favour well-mixed conditions.

### 3.3 Results

#### 3.3.1 Comparison of measurement sites

The urban-to-rural transect of observation sites provides an opportunity to evaluate the ability of different types of sites to detect urban plumes. We find that a semi-urban site can provide a constraint on the total emissions in the Rijnmond area, whereas an urban site is able to separate between different source areas. This is illustrated in Fig. 3.4 (left panel), where we display the probability density functions of the urban plume  $\text{CO}:\text{CO}_2$  concentration ratio (i.e.  $\Delta\text{CO}:\Delta\text{CO}_2$ ) at the three sites. A probability density function

illustrates the likelihood that an observed urban plume concentration ratio takes a certain value. The narrower the distribution, the less variable the ratios are and the more likely a ratio is to take the mean value (largest probability). Figure 3.4 also displays the mean bottom-up derived emission ratio of the Rijnmond area (vertical solid line, 2.5 ppb ppm<sup>-1</sup>) and its range, which is taken from the emission inventory taking into account the temporal profiles of the separate emission categories.

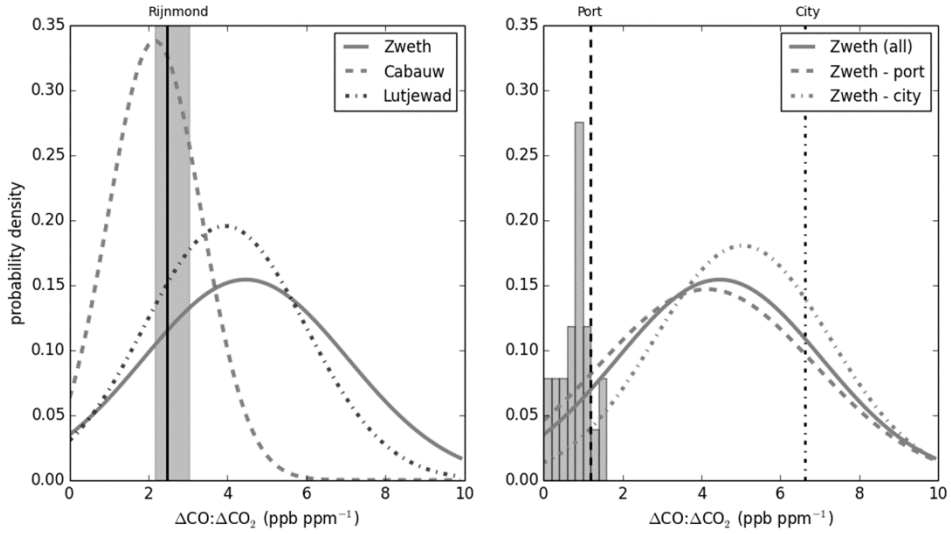


Figure 3.4: Left: Smooth Gaussian fit of probability density functions of observed  $\Delta\text{CO}:\Delta\text{CO}_2$  at the Zweth, Cabauw and Lutjewad measurement sites. The solid vertical line (shaded area) shows the mean emission ratio (Q1–Q3 range) for all emissions integrated over the Rijnmond area (see Fig. 3.1). Right: The Zweth observations separated into two distinct source areas based on the observed wind direction. The dash-dotted and dashed vertical lines represent the mean emission ratios from the residential area and the port, respectively. Generally, there is a reasonable match between the bottom-up emission ratio and the concentration-derived ratio, but observed ratios from the Zweth-port wind sector are much higher than expected because of the intermittency of plume transport from the many stacks in this area. The grey bars in the right panel show the point source events selected in Sect. 3.3.3.

We see that the  $\Delta\text{CO}:\Delta\text{CO}_2$  distribution at Cabauw is relatively narrow. Also, the mean  $\Delta\text{CO}:\Delta\text{CO}_2$  at Cabauw (2.2 ppb ppm<sup>-1</sup>) is very close to the bottom-up Rijnmond emission ratio. This indicates that Cabauw observes an integrated, well-mixed signal from the Rijnmond area and therefore contains information on the entire urban area. Interestingly, Lutjewad shows a much wider distribution with a mean of 3.9 ppb ppm<sup>-1</sup>. The urban plume from Rijnmond is mixed with signals from other industrial and urban areas (such as Amsterdam) before it reaches Lutjewad, causing more variability. This suggests that a site too far away from the urban sources is unable to uniquely identify the urban plume coming from a specific region. Also, the wind direction is heterogeneous between Rijnmond and Lutjewad. So, despite that the wind in Rijnmond is blowing towards Lutjewad according to our wind sector selection, the urban plume might never reach the site if the wind direction is changing during transport. This makes it difficult to filter out the Rijnmond urban plume and Lutjewad will be disregarded for the remainder of this

study. The Zweth site has an even wider distribution than Lutjewad and a mean ratio of  $4.5 \text{ ppb ppm}^{-1}$ . This site is affected by different source areas with distinct emission ratios depending on the wind direction, resulting in a large variability in observed concentration ratios. This variability contains a lot of information about the Rotterdam emissions and their spatiotemporal variations. Therefore, we examine the Zweth distribution in more detail by selecting wind sectors that sample different source areas with distinct emission characteristics (Fig. 3.4, right panel). Zweth-city is illustrative for the signal from the urban residential area dominated by road traffic and the Zweth-port signal contains mostly industrial and power plant emissions.

We find a large difference in bottom-up emission ratios for the residential ( $6.6 \text{ ppb ppm}^{-1}$ , vertical dash-dotted line) and port area ( $1.2 \text{ ppb ppm}^{-1}$ , vertical dashed line), which is not fully reproduced by the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratios. Whereas the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio for Zweth-city ( $5.0 \text{ ppb ppm}^{-1}$ ) is in reasonable agreement with the emission ratio, Zweth-port has a mean observed ratio that is much higher than expected ( $4.1 \text{ ppb ppm}^{-1}$ ). This discrepancy is related to the presence of high stack emissions in this area, which make up almost 75 % of the total Rijnmond  $\text{CO}_2$  emissions. The stack emissions from industrial processes and energy production have a small emission ratio of  $\sim 1 \text{ ppb ppm}^{-1}$  and dominate the total emission ratio. However, stack emissions have small plume dimensions that can easily be missed at the Zweth site and not be visible in the observations, especially for stacks in the vicinity of Zweth. Therefore, the observed concentration ratio can turn out much higher than what is expected based on the emission inventory including stack emissions. Indeed, the emission ratio of the Zweth-port area without point sources would be  $3.9 \text{ ppb ppm}^{-1}$ , which is very close to the observed  $4.1 \text{ ppb ppm}^{-1}$ . This finding indicates that stack emissions only occasionally affect the Zweth observations and it is very important to represent those events well with a model in order to constrain this large fraction of  $\text{CO}_2$  emissions. Although there might be an uncertainty in the emission inventory, reported emissions from industrial stacks are relatively accurate. Thus, it is unlikely that this explains the full discrepancy found for Zweth-port. Another potential cause of the discrepancy could be that the emission ratio is variable in time – for example due to a change in fuels used for energy production, while this is not accounted for in the inventory. However, this would likely have a smaller impact than the discrepancy found here. The impact of stack emissions on the Zweth observations is discussed in more detail in Sect. 3.3.3.

### 3.3.2 WRF-Chem urban plume transport

We have now seen that the observations at Zweth and Cabauw contain valuable information about the emissions in the Rijnmond area. In order to use that information to estimate the emissions, we explore the ability of WRF-Chem to represent observed time series, and especially their urban plume components.

Table 3.3: Statistics for WRF-Chem daytime (8:00–17:00 LT) average meteorological variables and total CO<sub>2</sub> and CO mixing ratios as compared to observed daytime averages (full simulation period).  $\bar{X}_{obs}$  is the average observed mixing ratio and N gives the number of days included. This table shows that WRF-Chem is able to represent day-to-day variations in meteorological conditions and mixing ratios, except for the wind direction.

Variable	Site	R <sup>2</sup>	RMSE	bias	$\bar{X}_{obs}$	N
Temperature	Rotterdam airport	0.77	2.5 °C	+ 0.9 °C		90
Specific humidity	Rotterdam airport	0.81	1.0 g kg <sup>-1</sup>	+ 0.5 g kg <sup>-1</sup>		90
Wind speed	Rotterdam airport	0.72	1.2 m s <sup>-1</sup>	<0.1 m s <sup>-1</sup>		90
Wind direction	Rotterdam airport	0.20	53 degrees	- 13 degrees		90
CO <sub>2</sub> mixing ratio	Westmaas	0.65	8.8 ppm	+ 1.1 ppm	418 ppm	83
	Zweth	0.45	13.0 ppm	+ 2.5 ppm	423 ppm	85
	Cabauw (60 m)	0.48	10.6 ppm	+ 3.6 ppm	417 ppm	86
CO mixing ratio	Westmaas	0.53	55 ppb	- 23 ppb	187 ppb	83
	Zweth	0.41	69 ppb	- 1 ppb	198 ppb	85
	Cabauw (60 m)	0.35	53 ppb	+ 18 ppb	156 ppb	89

First, we analyse the model performance on a day-to-day basis by looking at daytime averages and find that WRF-Chem is able to resolve day-to-day variations reasonably well. Table 3.3 shows that, respectively, 65 % and 53 % of the variability in the CO<sub>2</sub> and CO mixing ratios is captured at the Westmaas background site. Although the explained variances are slightly smaller at the urban (Zweth) and semi-urban (Cabauw) site, the performance at Cabauw for CO<sub>2</sub> is comparable to previous modelling studies (Bozhinova et al., 2014; Tolk et al., 2009). Yet, the RMSE (Root Mean Square Error) is relatively large for CO and CO<sub>2</sub> at all sites. Since Westmaas is nearly unaffected by urban emissions, the cause of the large RMSE is related to larger scale transport. Looking at meteorological variables, there is a good agreement for temperature, humidity and wind speed. However, the model has difficulties simulating the correct wind direction, which is especially expressed in the large RMSE. Similar errors have been observed before (Deng et al., 2017; Srinivas et al., 2016). The largest error is found in the second half of November, causing a large model-data discrepancy (also visible in Fig. 3.3). Table 3.3 also shows that the RMSE in the mixing ratios further increases for sites that are more influenced by the urban area. This finding indicates that WRF-Chem has difficulties representing the full variability caused by urban-industrial emissions.

Second, looking closer at the urban plumes we find that WRF-Chem represents the typical characteristics of urban plumes reasonably well, but it simulates the peaks at the wrong time at the wrong location compared to the measurements (Table 3.4). We tried to isolate the impact of errors in urban transport by looking statistically at the urban plume concentrations ( $\Delta\text{CO}_2$  and  $\Delta\text{CO}$ ) at Zweth and Cabauw. We select all data points that satisfy our criteria, separately for the observed and modelled time series such that both data sets can have a different size. We disregard data points associated with wind speeds of less than 3 m s<sup>-1</sup> to favour well-mixed conditions that are easier to interpret. However, we find that the inclusion of low wind speed data has limited impact on the average statistics. Table 3.4 shows that, on average, there is a good agreement between WRF-

Chem and the observations in the median and the 80<sup>th</sup> percentile. The median values of CO<sub>2</sub> are somewhat lower in WRF-Chem, indicating there are more small values and less high peak values in the model. Because the frequency distribution of the wind direction is similar between the observations and WRF-Chem, we expect no bias is introduced by the wind direction error. However, if we now co-sample WRF-Chem and the observations in time (i.e. we select observations that match our criteria and then take the same time from the WRF-Chem time series, which creates two data sets of equal size) we find a very small explained variance ( $R^2$ ) for both species at both sites based on hourly data. An inversion using these hourly data would thus be subject to a large model-data mismatch that increases the uncertainty in the optimised fluxes. Therefore, we next look more specifically at the data points responsible for the highest mismatch in observed and simulated  $\Delta\text{CO}_2$ .

**Table 3.4: Statistics for the distribution of the observed and modelled (WRF-Chem) urban plume mixing ratios ( $\Delta\text{CO}_2$  and  $\Delta\text{CO}$ ) at the Zweth and Cabauw site. N is number of hours included for either the observed or simulated time series. The  $R^2$  in the final column is based on co-sampling of WRF-Chem with the observations. The agreement between WRF-Chem and the observations is satisfactory when considering the distribution of the plume mixing ratios, but the low explained variance when co-sampling suggests a large impact of transport errors on individual plumes.**

Species	Site	Obs/model	Median	80 <sup>th</sup> percentile	N	$R^2$
CO <sub>2</sub>	Zweth	Observed	9.7 ppm	17.3 ppm	284	0.05
		WRF-Chem	8.8 ppm	16.9 ppm	249	
	Cabauw (60 m)	Observed	6.0 ppm	9.1 ppm	32	<0.01
		WRF-Chem	5.6 ppm	6.4 ppm	37	
CO	Zweth	Observed	29 ppb	57 ppb	274	0.01
		WRF-Chem	33 ppb	50 ppb	207	
	Cabauw (60 m)	Observed	13 ppb	28 ppb	58	<0.01
		WRF-Chem	18 ppb	31 ppb	51	

We find that the largest differences between WRF-Chem and the observations at Zweth when co-sampling urban plumes results from errors in simulated wind direction, as well as from an inability of WRF-Chem to simulate the impact of point source emissions. This is illustrated in Fig. 3.5, where we binned the absolute errors in hourly  $\Delta\text{CO}_2$  into four magnitude classes of 10 ppm each and correlate them with the error in simulated wind direction (as binned into three classes of 20 degrees, scatter plots) and with the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio (whisker plots). We find that the smallest  $\Delta\text{CO}_2$  model error class (0–10 ppm) is dominated by the smallest wind direction error (0–20 degrees, 68 %), while in the largest  $\Delta\text{CO}_2$  model error class (30–40 ppm) 70 % of the data points have a wind direction error of more than 20 degrees. With such large wind direction errors, the trajectory of urban plumes is misrepresented and the modelled mixing ratios are affected by the wrong source area, or plumes may even entirely miss the sites in the model. In addition, we find that in the largest  $\Delta\text{CO}_2$  model error class (30–40 ppm) the observed  $\Delta\text{CO}:\Delta\text{CO}_2$  is lower (2.5 ppb ppm<sup>-1</sup>) and less variable than in the other classes, suggesting a larger influence of

industrial (stack) emissions. Although the number of data points in the largest  $\Delta\text{CO}_2$  model error class is small ( $N=14$ ), these tendencies give a good indication of what might cause these errors. At Cabauw, the impact of stack emissions is not visible, because the point source emissions are already well-mixed when the air mass arrives at Cabauw. Hence, we will next examine the added value of the OPS plume model only at Zweth to better represent the dispersion of  $\text{CO}_2$  emitted from stacks and the impact of wind direction in OPS.

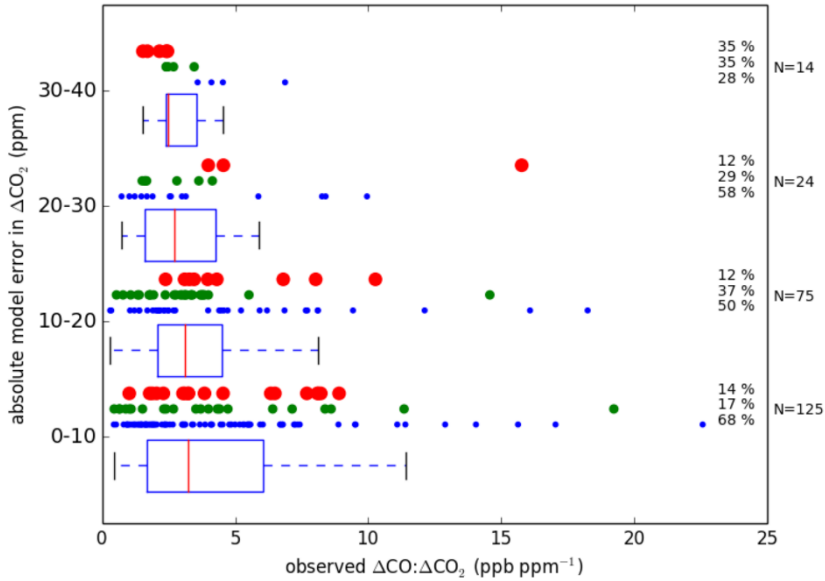


Figure 3.5: This figure shows four classes of the absolute model error in  $\Delta\text{CO}_2$  compared with the Zweth measurement site. For each class two quantities are displayed. 1) A whisker plot of observed  $\Delta\text{CO}:\Delta\text{CO}_2$ , which shows that the largest absolute  $\Delta\text{CO}_2$  model error (y-axis) is related to small observed concentration ratios (x-axis). This indicates an important role for low-ratio stack emissions (industrial and power plant sources) in the large model error class. 2) A coloured scatter plot for which data points are divided into three classes based on the absolute error in simulated wind direction (<20 degrees in small blue dots on bottom row, 20–40 degrees in larger green dots on middle row, and >40 degrees in large red dots on top row). Each dot represents one hour. The percentage contribution of each wind direction error class to the total number of data points (N) is shown on the right. These numbers show that the model error in wind direction also plays an important role in the  $\Delta\text{CO}_2$  model error.

### 3.3.3 WRF-Chem and OPS point source representation

When we focus exclusively on point source emissions, we find that all simulations that include the OPS plume model are in better agreement with the observations than the WRF-Chem simulation (based on the  $R^2$  and regression slope). This is illustrated in Table 3.5, where we compare co-sampled simulated and observed events with a high point source contribution (see also Appendix A for more details). These events are selected based on a low observed  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio (the threshold is  $1.5 \text{ ppb ppm}^{-1}$ , events illustrated as grey bars in Fig 3.4). In the models, these events are highly correlated with a high point source contribution (of at least 90 %) in the simulated  $\Delta\text{CO}_2$  mixing ratio ( $r$  is -

0.76 (WRF+OPS-point-obsmet) and -0.61 (WRF-Chem)). Including low wind speed data deteriorates most of the statistics for all simulations (not shown), meaning that the models have difficulties representing stagnant conditions.

For WRF-Chem the explained variance in the co-sampled observations is limited ( $R^2=0.30$ ) and the regression slope of  $\Delta\text{CO}_2$  is significantly lower than one (i.e. the 1:1 line of modelled vs. observed  $\Delta\text{CO}_2$ ). Both the mean  $\Delta\text{CO}:\Delta\text{CO}_2$  and the standard deviation are larger than the observed mean and standard deviation. This suggests that the lack of agreement is partly caused by an error in the WRF-Chem wind direction, causing the model to sample air from a wrong source area.

**Table 3.5: Statistics for  $\text{CO}_2$  point source peaks at Zweth in four different model simulations as compared to observations. N is number of hours included and the slope is based on a linear regression.  $\Delta\text{CO}:\Delta\text{CO}_2$  denotes the mean ( $\pm 1\sigma$  standard deviation) of the urban plume concentration ratio in  $\text{ppb ppm}^{-1}$ .**

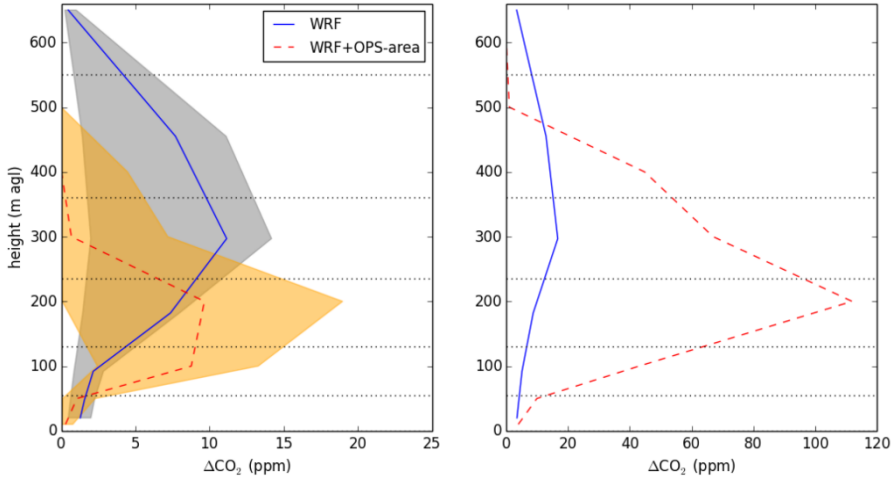
Model run	$R^2$	$\Delta\text{CO}:\Delta\text{CO}_2$	$\Delta\text{CO}_2$ slope	N
WRF-Chem	0.30	0.9 ( $\pm 1.5$ )	0.82	42
WRF+OPS-area	0.37	1.2 ( $\pm 1.1$ )	0.87	42
WRF+OPS-point	0.42	1.2 ( $\pm 1.6$ )	0.86	42
WRF+OPS-point-obsmet	0.52	0.7 ( $\pm 0.6$ )	0.99	40
Observed		0.7 ( $\pm 0.4$ )		

In contrast to WRF-Chem, WRF+OPS-point-obsmet shows a larger explained variance ( $R^2=0.52$ ), a regression slope that is nearly one, and a  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio that agrees with observations both in mean and in standard deviation. Since only about 10 % of the Zweth-port observations are affected by stack emissions due to the small dimension of the plumes ( $N=42$ ), a better representation of atmospheric conditions has a large impact. An advantage of the OPS model is the ability to estimate the model uncertainty by providing a plume cross-section. Receptor points can be positioned anywhere and by adding several receptor points around the true measurement location we can account for transport errors (e.g. in the wind direction). If we allow for a maximum wind direction error of 5 degrees, this has no significant impact on the  $R^2$  or slope (results not shown), suggesting that the results from the WRF+OPS-point-obsmet simulation are robust against small random errors in wind direction. However, systematic errors in the wind direction or the treatment of point source emissions such as present in WRF-Chem will have an impact on its performance, as we will explore next.

### **Dispersion**

When comparing WRF-Chem and WRF+OPS-area we find that the OPS model reduces the dispersion of point source emissions, which causes emissions from high stacks to barely reach ground level. Vertical profiles of  $\Delta\text{CO}_2$  near an energy production stack for both model simulations are shown in Fig. 3.6. Energy production sources often have the highest stacks and the lowest  $\Delta\text{CO}:\Delta\text{CO}_2$  ratios. Near an energy production stack the vertical

dimension of the plume in WRF+OPS-area is smaller than in WRF-Chem. The plume remains more concentrated in WRF+OPS-area, leading on average to lower mixing ratios at ground level (left panel) and to higher maximum values at around 200 m (right panel). This effect is also clearly visible at Zweth (not shown) and results in a higher mean  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio in Table 3.5 for WRF+OPS-area (i.e. less influence of the low-ratio stack emissions) and a higher explained variance (37 %).



**Figure 3.6:** Vertical profiles of the median (Q1–Q3) (left panel) and maximum (right panel)  $\Delta\text{CO}_2$  mixing ratio at 14 h UTC at about 500 m from an energy production point source in WRF-Chem and WRF+OPS-area. The horizontal lines represent the boundaries of the vertical levels in WRF-Chem. Emissions are taking place in levels 3, 4 and 5 in WRF-Chem or at 130, 235 and 360 m in WRF+OPS-area. The figure shows on average lower mixing ratios at ground level in WRF+OPS-area than in WRF-Chem, despite an identical treatment of the vertical emission structure. WRF+OPS-area also shows higher maximum values, reflecting a reduction in vertical dispersion compared to the Eulerian box representation in WRF-Chem.

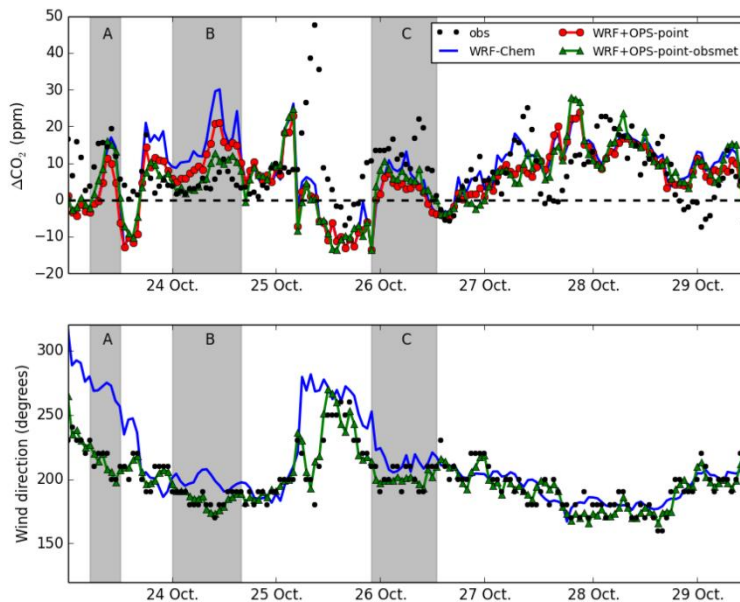
### *Point source representation*

From a comparison of WRF-Chem, WRF+OPS-point and WRF+OPS-point-obsmet it follows that having a plume model with full point source characteristics can improve the agreement with the observed mixing ratios, even if the meteorological conditions are biased. Implementing detailed source characteristics (WRF+OPS-point) not only increases the explained variance to 42 %, it also increases the  $\Delta\text{CO}:\Delta\text{CO}_2$  standard deviation. This is the result of larger spatial (both horizontal and vertical) variability in the emission landscape.

These effects are also visible in Fig. 3.7, which shows a time series of six days of observations and model output. When differences between the simulations are small, this indicates the absence of point source signals. On October 23 (event A) an improvement is made by using observed meteorological conditions due to the large wind direction error, while the difference between WRF-Chem and WRF+OPS-point is small. However, on other occasions the use of the OPS model, irrespective of the meteorology used, already improves the simulated urban plume mixing ratio. For example, on October 24 (event B)



both OPS runs reduce the urban plume mixing ratios and are in better agreement with the observations. On October 26 (event C) the opposite is happening. Whereas WRF-Chem is only above the background for four hours, the observations show a longer and more severe pollution event, despite a relatively small wind direction error. Although an additional improvement can be made using the observed wind fields, using WRF+OPS-point already improves the length and strength of the pollution event. Note that, although WRF-Chem sometimes performs better than the simulations including OPS, the overall statistics suggest that it is recommended to use WRF+OPS-point-obsmet.



**Figure 3.7:** Time series of  $\Delta\text{CO}_2$  at Zweth from observations and three model simulations (top panel) and of the wind direction at Rotterdam airport from WRF-Chem, WRF+OPS-point-obsmet, and observations (bottom panel). Shaded areas indicate specific events discussed in more detail in the text.

### 3.4 Discussion

In this study we focused on two major questions in urban greenhouse gas modelling studies: what type of measurement locations can provide the best information on urban fluxes of  $\text{CO}_2$  and  $\text{CO}$ , and what type of modelling framework can best represent urban plume mixing ratios at these measurement sites. In a previous study, Lauvaux et al. (2016) have used nine observation towers to estimate  $\text{CO}_2$  fluxes from Indianapolis. They have argued that the optimum number of towers is dependent on the spatial heterogeneity of the emissions within the city. They also state that it is impossible to attribute changes in the total  $\text{CO}_2$  concentration to specific source sectors when only  $\text{CO}_2$  observations are available. Based on our current findings, we believe that with the use of other co-emitted species, like  $\text{CO}$ , information can be gained about source sector contributions, as was also shown by Turnbull et al. (2015). Additionally, Brioude et al. (2013) have shown that with

only a few flights a reasonably robust flux estimate can be made for CO and NO<sub>y</sub>. These studies thus show that with additional species and strategically placed measurements the need for a large observation network can be reduced. However, an important pre-condition is that atmospheric transport is correctly represented. Lauvaux et al. (2016) discussed that the atmospheric transport in high-resolution Eulerian models might suffer from errors due to assumptions about turbulence and other fine-scale processes, which causes urban plumes to violate the well-mixed assumptions of the model. This is especially relevant for emission sources with dimensions that are significantly smaller than the model resolution, i.e. point sources. Indeed, in this study we find that a plume model is a useful addition to the Eulerian model to correctly represent the transport of emissions from large point sources.

### 3.4.1 Comparison of observation sites

We first examined the use of the measurement sites to detect urban plumes, since the measurement sites in an inversion determine the demand put on the model performance. At the rural site (Lutjewad), the urban plume has become mixed with other signals and the urban plume is difficult to distinguish. This site (at ~200 km from the Rijnmond area) is therefore too far removed to specifically constrain the Rijnmond emissions, although it was shown to constrain emissions from the larger urban conglomerate of the Randstad quite well (Van der Laan et al., 2009a; Van der Laan et al., 2010). The semi-urban site (Cabauw) detects urban plumes from Rijnmond which have already become well-mixed during transport. Moreover, the mean concentration ratio matches well with the emission ratio for the Rijnmond area. We therefore argue that the Cabauw site could constrain the overall emissions of the Rijnmond area due to its integrating power without the need for a multi-model approach. In contrast, the urban location (Zweth) is highly exposed to the urban fluxes and is able to detect spatial variations in emissions inside the urban area. We find distinct concentration ratios for different source areas that can provide valuable information about dominant source types and areas. These findings are similar to a previous study concluding that a network of in-city sites provides good constraints due to their high exposure and ability to separate between different parts of the source area (Kort et al., 2013). However, the difference between the emission ratio and observed concentration ratio for the Zweth-port area indicates that stack emissions might frequently be missed at the Zweth measurement site due to the limited plume dimensions. Therefore, a correct representation of the transport becomes increasingly important. Thus, we conclude that the Cabauw and Zweth site have their own particular (dis)advantages and a combination of an urban and semi-urban site could be most beneficial to constrain urban fluxes in detail. Note that this conclusion is specifically valid for the Rijnmond area with the presence of major point sources and the requirements might be different for other urban topologies.

### 3.4.2 Model skill

Next, we evaluated the skill of the Eulerian WRF-Chem set-up. The ability of our WRF-Chem framework to represent daytime average mixing ratios is comparable with other model frameworks in the urban environment (Bozhinova et al., 2014; Bréon et al., 2015; Lac et al., 2013; Tolk et al., 2009). However, WRF-Chem has a large wind direction bias that makes it difficult to compare modelled and observed mole fractions. The monthly average WRF-Chem wind direction shows an absolute bias of 1 (October), 51 (November) and 10 (December) degrees compared to the observed wind direction at Rotterdam airport. The error in November is large compared to previous findings (Jiménez et al., 2016) and this results in a large model-observation mismatch in tracer mixing ratios (Fig. 3.3). Also at the Cabauw site, which is less influenced by build-up areas, the model-data agreement for the 10 m wind direction in November is limited. Previous research has also shown an uncertainty of 30–40 % in the tracer mixing ratio due to the uncertainty in meteorological conditions (Angevine et al., 2014; Srinivas et al., 2016). Additionally, Angevine et al. (2014) have shown that using an ensemble mean of model simulations with different meteorology does not necessarily lead to a better representation of plume transport and dispersion in a Lagrangian model for area sources. We therefore speculate that assimilating observed wind fields in WRF-Chem, as was done by Lauvaux et al. (2013), could be more beneficial to improve the modelled wind fields and as such improve the plume transport. Furthermore, the model performance under stratified and low wind speed conditions need to be addressed, since removing these data can lead to biased emission estimates.

Some studies argued that the main limitations of a Eulerian model are the enhanced dispersion due to instant mixing of species throughout the grid box and, related to that, the absence of a good point source representation (Karamchandani et al., 2011; Tolk et al., 2009). Our results show evidence for both limitations in the WRF-Chem set-up. First, WRF-Chem underestimates the median urban plume mixing ratios of both CO<sub>2</sub> and CO which should mainly be attributed to errors in transport and mixing. Whereas CO mixing ratios at the Zweth site are dominated by area sources, CO<sub>2</sub> mixing ratios are also highly affected by point source emissions. Therefore, their consistent underestimation cannot be caused solely by errors in point source emissions. Second, looking more specifically at the point source contribution, WRF-Chem can only explain 30 % of the variance and the spread in the  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio is too large compared to the observations. Thus the resolution appears to be too low to fully represent the transport of the urban plumes from point sources, similar to previous findings related to power plant plumes (Lindenmaier et al., 2014) and megacities (Boon et al., 2016).

In order to overcome the limitations of WRF-Chem related to point source representation and wind field errors, we evaluated the use of the OPS plume model with full point source characteristics and observed meteorological conditions. As discussed before, the OPS plume has limited impact on the CO mixing ratios as point sources only

contribute a small fraction to the total CO emissions. Therefore, the focus in the remainder is on CO<sub>2</sub>. The OPS plume model requires limited effort to be run in addition to WRF-Chem (it requires 2 input files and takes only a few seconds to run) and is therefore a relatively easy solution to improve the point source representation.

Several previous plume modelling studies with different species showed improvements compared to the gridded approach (Briant and Seigneur, 2013; Ganshin et al., 2012; Karamchandani et al., 2006; Karamchandani et al., 2012; Korsakissok and Mallet, 2010b; Rissman et al., 2013). In this study we find a significant improvement with WRF+OPS-point-obsmet at Zweth, both in the explained variance and the  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio. Also the observed-vs-simulated regression slope of the point source  $\Delta\text{CO}_2$  mixing ratio becomes nearly one. In this analysis the number of selected data points is relatively small, because stack emissions can easily be missed by an observation site due to the small plume dimensions. Therefore, only a few events can be used to constrain point source emissions and a good representation of the plume transport is essential. Although there are only ~100 individual point sources in the Rijnmond area, they make up about 75 % of the total CO<sub>2</sub> emissions. Thus we argue that in an urban-industrial area with a significant point source contribution the use of a plume model is critical to get a reliable emission estimate. If detailed point source characteristics are unknown, these would have to be estimated and this adds an uncertainty to the modelled mixing ratios. Nevertheless, we have shown that even with the WRF-Chem point source representation (i.e.  $1\times 1\text{km}^2$  in size and fixed vertical distribution) the plume model can already improve the agreement with the observations. Further improvements can possibly be made by representing traffic emissions as line source emissions in a plume model (Briant and Seigneur, 2013) rather than considering them as gridded area sources in the Eulerian model.

Although part of the OPS-driven improvement can be attributed to the use of observed meteorological conditions, we have shown with the WRF+OPS-point simulation that there is also an improvement in point source representation. We found that a higher spatial variability in the emissions causes more variability in the concentration ratios. Representing point sources as area sources, as is done in WRF-Chem, results in lower correlations and less variability in concentration ratios, which is consistent with previous studies that demonstrated the importance of a good source representation (Kim et al., 2014; Korsakissok and Mallet, 2010a; Touma et al., 2006). Besides the ability to include detailed source characteristics and to use observed meteorology, the OPS model has some additional advantages. We have shown that looking at individual stacks can provide valuable information about the underlying transport and dispersion processes and how they are affected by source characteristics. Additionally, receptor sites can be positioned anywhere, which allows us to study the spatial variations at much higher resolution than currently possible with WRF-Chem.

At Cabauw, the difference between WRF-Chem and the WRF+OPS-point simulation is small, although the model-data mismatch at Cabauw is further reduced when observed

meteorology is used. This leads to the question for which spatial extent a plume model is beneficial. In previous plume-in-grid models at high resolution (<25 km) plumes or puffs are often injected in the Eulerian parent model when the width of the plume is similar to the grid size (Karamchandani et al., 2006; Kim et al., 2014; Korsakissok and Mallet, 2010a). According to the definition of the lateral dispersion factor in OPS this would mean that a plume will have reached a horizontal width of 4 km (the resolution of the domain in which Cabauw is located) after about 8 km travel distance under well-mixed conditions. To test this, we compared a monthly average WRF-Chem CO<sub>2</sub> mixing ratio field in and around Rijnmond with a monthly averaged gridded OPS mixing ratio field. The OPS model was only applied for emissions within the Rijnmond area and therefore the distance outside the WRF-Chem domain 4 at which the mixing ratio fields become similar gives an indication of the spatial extent for which the OPS model is still beneficial. We find that the difference between the mixing ratio fields disappears quickly outside the Rijnmond area and WRF-Chem and OPS become similar at about 10–14 km outside the boundary of domain 4.

### 3.5 Conclusions

Our ultimate ambition is to quantify the total urban CO<sub>2</sub> budget using multiple observation sites and an inverse modelling system. Such information could be used to monitor the impact of implemented policies and progress towards objectives. Based on the work reported here, we state that the modelling framework should ideally consist of a Eulerian model in combination with a plume model for point source emissions within the city, preferably driven by locally observed meteorology. The use of a plume model is of great added value to correctly represent the transport of point source emissions in a diameter closer than ~10 km to the site. Although the additional computational demand with the OPS plume model is limited, detailed model input is required given that the results are very sensitive to source characteristics and wind fields. Given the importance of observed local meteorology for the model performance, we strongly recommend inclusion of a (simple) meteorological station in any similar monitoring set-up. Also, Lagrangian particle dispersion models driven by WRF meteorological fields have proven useful in describing the transport of point source emissions and in inverse modelling (Brioude et al., 2013; Pan et al., 2014; Srinivas et al., 2016), but such set-up would suffer from wind field errors. The optimal set-up for an urban monitoring network requires a semi-urban measurement site (here ~30 km from the urban area with no other urban areas in between) and at least one additional urban measurement site (here at the edge of the urban area, at ~7 km from the city centre). The semi-urban site provides a robust and integral constraint on the urban fluxes and can be used in combination with a high-resolution Eulerian model framework. The urban measurement site can provide useful information about local differences, such as the dominance of road traffic in a certain source area or local changes due to

implemented measures. Observing additional species besides CO, like  $^{14}\text{CO}_2$ ,  $^{13}\text{CO}_2$ ,  $\text{O}_2/\text{N}_2$ ,  $\text{NO}_2$ ,  $\text{SO}_2$  or black carbon, could be a useful extension of our framework for identifying source sector contributions. Such a set-up is a promising step towards independent verification of urban  $\text{CO}_2$  budgets.

### Appendix 3A: Scatter plots

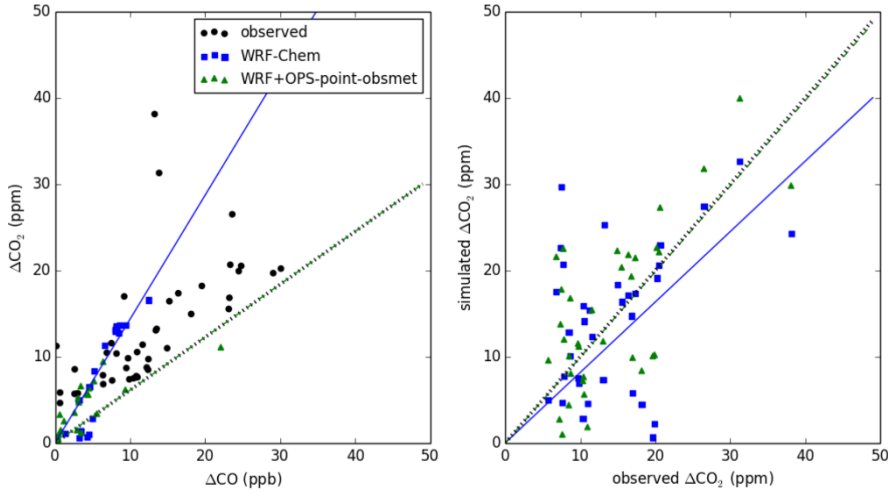


Figure A1: Left: A scatter plot of  $\Delta\text{CO}$  and  $\Delta\text{CO}_2$ , where the slopes (represented by lines) represent the  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio for the observed and modelled values. The slope of WRF+OPS-point-obsmet coincides with the slope of the observations, suggesting a good agreement. Right: A scatter plot of simulated  $\Delta\text{CO}_2$  to observed  $\Delta\text{CO}_2$ . The slope of WRF+OPS-point-obsmet coincides with the 1:1 line (dotted line), suggesting a good agreement with the observations.



A thick dark blue vertical bar runs down the left side of the page. A light teal arrow points to the right from this bar, containing the text 'Chapter 4'. Below the bar, several thin, curved lines in dark blue and light grey sweep upwards from the bottom left towards the center of the page.

## Chapter 4

# **Development of a dynamic emission model to predict high-resolution fossil fuel CO<sub>2</sub> emissions**

*Authors: Ingrid Super, Wouter Peters, Michiel van der Molen, Hugo Denier van der Gon*





## 4.1 Introduction

Recently, much effort is put into climate agreements to limit the effect of global warming. An important part of these agreements is the reduction of greenhouse gas emissions from fossil fuel combustion. An example is the Nationally Determined Contributions related to the Paris Agreement. These emission reduction efforts need to be monitored (UNFCCC, 2015), preferably independently using atmospheric observations. A potentially interesting tool in an independent emission verification system is data assimilation (DA), in which atmospheric observations provide a constraint on the emissions and emission trends.

Data assimilation has been used for a wide range of applications, also related to the carbon cycle (Brioude et al., 2013; Broquet et al., 2013; Lauvaux et al., 2016; Peters et al., 2005; Van der Laan-Luijkx et al., 2017). One important precondition is the presence of enough observations in the area of interest. Therefore, DA of atmospheric CO<sub>2</sub> observations has mostly been applied on regional scales and mainly related to biogenic fluxes. Focusing on smaller (sub-national/urban) scales requires a denser network of CO<sub>2</sub> observations than currently available in most regions. However, in several megacities (e.g. Indianapolis, Paris, Los Angeles) an effort has been made to install a dense network of instruments with a high temporal resolution in order to monitor CO<sub>2</sub> emissions (Bréon et al., 2015; Turnbull et al., 2015). This is important, because urban areas are responsible for about 70% of the global fossil fuel CO<sub>2</sub> emissions and the largest potential of emission reduction can be found in those areas. Therefore, applying DA at the urban scale is receiving increasing attention.

An important factor in DA is the definition of the prior emission map (i.e. a first estimate) and its uncertainties (Wang et al., 2018). Currently, fossil fuel CO<sub>2</sub> emissions are often estimated on a yearly basis per country with an uncertainty of a few percent for developed countries and up to 50% for data-sparse regions (Andres et al., 2012). Mostly, these emissions are disaggregated to higher spatiotemporal resolution to create state-of-the-art emission maps. However, the relationship between the data used for disaggregation and the emissions is not always straightforward and can be dynamic in space and time. Moreover, most inventories use fixed time profiles to introduce a diurnal or seasonal cycle, which is based on long-term, regional average activity data at best, or else expert judgement. Therefore, the uncertainty at higher spatiotemporal resolution is much larger than at the annual time scales. This places a higher demand on the DA system at the urban scale as less confidence can be put in the prior and the system is allowed to move further away from the starting point. Moreover, the uncertainties in high-resolution inventories have not been quantified accurately, which hampers the creation of a realistic error covariance matrix. Such a matrix defines the error structure of the prior emissions and is used, together with the error structure of the observations, to constrain the emissions. As such, incorrect definition of the errors can steer the outcome too much towards the observations, or towards the prior.

We can identify two main challenges related to emission monitoring at high resolution. First, the quality of the emission data need to be improved by including new or more specific data streams, as previously suggested by Ciais et al. (2014). For example, current time profiles for road traffic are based on traffic counts averaged over several years. Instead, real-time traffic counts for the area of interest can be used to better mimic the dynamic nature of this emission sector. An additional advantage of using these data streams is that they are available in near real-time, whereas current emission inventories often have a lag of several years. Second, we need to quantify the uncertainty in the emissions at the spatiotemporal resolution of the observations to be able to use them in the DA system. Since it is difficult to assess the uncertainty of an emission product due to the lack of verification methods a more promising method is to estimate the uncertainty in underlying parameters, such as emission factors (see Chapter 1).

In this chapter we develop a dynamic fossil fuel emission model that calculates high-resolution emissions from a wide range of data streams. With "dynamic" we refer to the use of hourly activity data for the specific study region to create hourly variability in the emissions instead of using fixed hourly profiles based on long-term averaged activity data or socio-economic characteristics. A dynamic emission modelling system allows us to study the uncertainty of underlying data streams (e.g. traffic counts, energy consumption, household heating intensity) and to asses critical assumptions that are needed to come to the spatiotemporal resolutions that allow independent fossil fuel emission monitoring. Similar to most state-of-the-art inventories the emission model gives emissions per source sector. In the DA system we can thus optimize model parameters per source sector. To ease the attribution of CO<sub>2</sub> enhancements to specific source sectors we also include additional tracers (CO, NO<sub>x</sub> and SO<sub>2</sub>) in the emission model and DA system. These tracers are co-emitted with CO<sub>2</sub> in a ratio that is specific for each source sector and can therefore potentially be used to attribute CO<sub>2</sub> signals to specific source sectors. We will demonstrate the optimization of emission factors, emission ratios and time profiles in the next chapter.

The purpose of this chapter is twofold. First, to examine the use of different data streams to calculate emissions of CO<sub>2</sub> and other co-emitted species for a specific area of interest. Second, to gain insight in the uncertainty of emissions and the parameters that have the largest impact on the overall uncertainty. We start with a description of the dynamic fossil fuel emission model in Sect. 4.2, including an inventory of the different data streams. Sect. 4.3 examines the uncertainty in model parameters, which are used to estimate the uncertainty in the emissions in Sect. 4.4. Data assimilation within the dynamical emission model will be demonstrated in Chapter 5 of this thesis.

## 4.2 Dynamic fossil fuel emission model

### 4.2.1 Case study and general set-up

Our dynamic emission model aims to calculate hourly emissions per source sector based on different data streams. We developed the model for the Netherlands with a specific focus on the Rijnmond area, including the city of Rotterdam (about 625,000 inhabitants) and its large industrial-port area. The Rijnmond area is located at the west coast of the Netherlands in a larger urbanized area (Randstad, about 7 million inhabitants). The area of interest is thus characterized by a complex mixture of residential and industrial activities. We focus on the year 2014.

**Table 4.1: Overview of sectors and subsectors distinguished in the dynamic emission model, including their short name used in the figures.**

Sector	Subsector	Short name
<b>Power plants</b>	Coal-fired power plants	1A
	Gas-fired power plants	1B
<b>Non-industrial combustion</b>	Households	2A
	Glasshouses	2B
<b>Industry</b>		3
<b>Road traffic</b>	Cars	7A
	- Highway	7A1
	- Main road	7A2
	- Urban road	7A3
	HDV	7B
	- Highway	7B1
	- Main road	7B2
	- Urban road	7B3
<b>Shipping</b>	Ocean shipping	8A
	Inland shipping	8B
	Recreational shipping	8C

We identified five source sectors (based on SNAP categories, Selected Nomenclature for sources of Air Pollution) that are relevant for the CO<sub>2</sub> emissions in the area of interest. These are power plants, industry, non-industrial combustion, road transport, and shipping. These sectors contribute 37%, 39%, 15%, 6% and 3% to the total CO<sub>2</sub> emissions in the Rijnmond area, respectively. The non-industrial combustion sector is subdivided into households and glasshouses. In the port of Rotterdam power plants and industry are the major sources of CO<sub>2</sub>, but in the urbanized area residential combustion and road transport are dominant. All emission categories and subcategories are listed in Table 4.1.

The dynamic emission model consists of three parts: calculation of the total yearly emissions for the entire country, temporal disaggregation to hourly scale, and spatial disaggregation to 1x1 km<sup>2</sup> for the area of interest. The following sections will explain each of these parts in more detail. A summary of all data used in the emission model is given in Appendix A.

#### 4.2.2 Calculation of total emissions

Total emissions ( $F$  in  $\text{kg yr}^{-1}$ ) are calculated as a function of the economic activity and an emission factor per unit of activity (adapted from Raupach et al. (2007)):

$$F = A \left( \frac{E}{A} \right) \left( \frac{F}{E} \right) \quad (1)$$

where  $A$  is the amount of activity, such as vehicle kilometres driven or generated power, and  $E$  is the primary energy consumption (petajoule (PJ)). In this equation the term  $F/E$  is the emission factor (EF), i.e. the amount of  $\text{CO}_2$  emitted per amount of energy consumed. The term  $E/A$  can be seen as a measure of energy efficiency, in which technological development plays an important role (Nakicenovic et al., 2000).

For each source sector we describe the total  $\text{CO}_2$  emission with Eq. 1. For the activity data we use either the degree day sum calculated from the outside temperature (for non-industrial combustion) or the GDP (Gross Domestic Product; for power plants, industry, road transport and shipping), which are simple and widely available proxies for activity. The degree day function is explained in more detail in Sect. 4.2.3. The temperature can be measured and GDP is provided by several institutions, like the International Monetary Fund. Therefore, this method is applicable in other countries as well. The second term in Eq. 1 ( $E/A$ ) can be estimated from energy consumption statistics, such as available from the International Energy Agency. Note that this term can show a large trend in case of technological development. The last term in Eq. 1 ( $F/E$ ) is the most difficult one, because the emission factor is dependent on the fuel mix and the energy efficiency, which itself can vary with environmental conditions (e.g. a cold engine on a winter day burns less efficiently). It can therefore differ significantly between countries.

The industrial sector consists of a wide range of activities, from refineries and chemical industry to food industry, that have their own characteristics. Although the industrial energy consumption shows a correlation with GDP, this is not consistent with trends in the reported emissions. The current development in the Netherlands is that energy intensive industry is exported and that the industrial production is decreasing per unit GDP. This causes a shift in dominant industrial processes that affects the average characteristics of the industrial sector. Therefore, predicting industrial emissions requires additional information on dominant processes. For now we estimate the industrial emissions using the energy consumption per unit GDP given the absence of a good alternative.

From the total  $\text{CO}_2$  emissions we can also calculate emissions of co-emitted tracers using emission ratios of each tracer to  $\text{CO}_2$ , which is an additional multiplier at the right-hand side of Eq. 1. The ratios can be highly variable in space and time and depend on the technology and combustion characteristics. We apply this method to calculate emissions of  $\text{CO}$ ,  $\text{NO}_x$  and  $\text{SO}_2$ .

### 4.2.3 Temporal disaggregation

The next step is to disaggregate the yearly emissions to hourly emissions by calculating time profiles, such that Eq. 1 becomes "dynamic":

$$F = A \left( \frac{E}{A} \right) \left( \frac{F}{E} \right) T \quad (2)$$

where  $T$  is the hourly time factor. Averaged over a year the value of  $T$  is 1.0, so that it only alters the temporal evolution and not the total emissions.  $T$  can be calculated in two ways: 1) by directly using activity data or 2) by parameterizing temporal variations from environmental and/or economic conditions. When activity data is available the first option is preferable. However, in data-sparse regions the second option might be necessary to implement. Nevertheless, the second option is still an improvement compared to long-term average profiles as commonly used. Therefore we explore both methods. In this section we explore the possibilities for temporal disaggregation for each source sector.

#### *Non-industrial combustion*

Non-industrial combustion is considered to be dominated by households' natural gas consumption to heat houses, for cooking, and for warm water supply. A Dutch energy provider has a small dataset publicly available from about 80 smart meters for the year 2013 with hourly gas consumption (Liander, 2018). It clearly shows a seasonal cycle, but also more small-term variations (daily data are shown in Fig. 4.1). We also see higher gas consumption in the beginning of the year, where the first three months of 2013 had some long, cold spells.

Previous studies have used the concept of heating degree days to describe the temporal variability in emissions from households (Mues et al., 2014; Terrenoire et al., 2015). This concept assumes that heating only takes place below a certain temperature threshold (here 18°C) and the hourly time factor can be defined as:

$$T_{2A} = H / \bar{D} \quad (3)$$

where  $H$  is the heating degree day factor ( $H = \max(291.15 - \bar{T}, 0)$ ) based on the daily mean outside temperature at 2m.  $\bar{D}$  is the yearly average heating degree day ( $\bar{D} = \frac{1}{N} \sum_{j=1}^N H$ ).

However, gas consumption related to warm water supply and cooking is largely independent of the outside temperature and therefore a constant offset is included in the heating degree day factor:

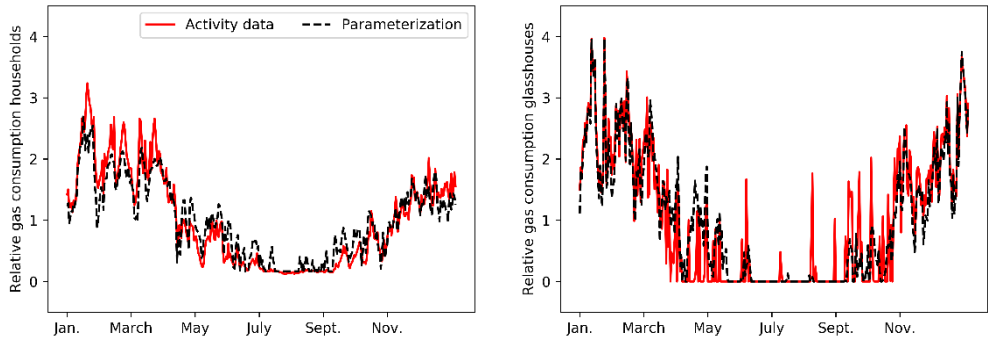
$$H_f = H + f \cdot \bar{D} \quad (4)$$

where  $f$  is the constant offset. We assumed an offset of 20%, similar to Mues et al. (2014). The time factor can now be defined as:

$$T_{2A} = H_f / \overline{D_f} \quad (5)$$

where the average heating degree day accounted for the constant offset  $\overline{D_f} = (1 + f)\overline{D}$ .

We compared the heating degree day method with gas consumption data on a daily basis (Fig. 4.1). The degree day function follows the gas consumption data very well, including the higher consumption at the start of the year, reaching an  $R^2$  of 0.90 (N=365). The gas consumption of consumers also has a diurnal pattern with peaks in the early morning and late afternoon. Therefore, a diurnal profile can be estimated based on typical working hours. For hourly data  $R^2$  is 0.80 (N=8760, not shown).



**Figure 4.1:** Daily time profiles for households (left) and glasshouses (right). Full lines are based on true activity data, whereas dashed lines are parameterizations based on the degree day function.

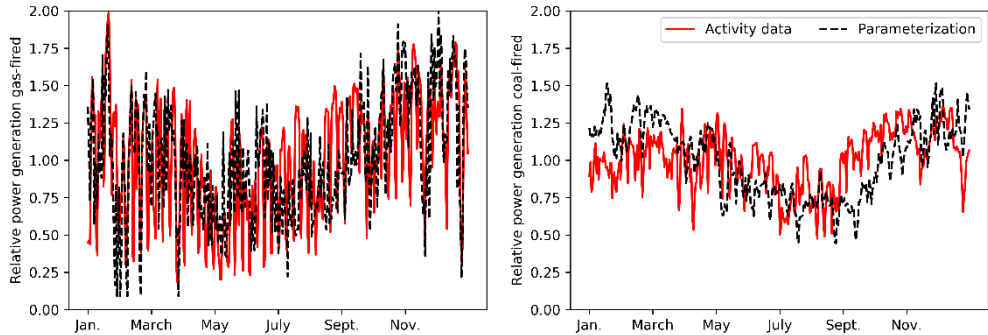
Interestingly, in the Rijnmond area 45% of the non-industrial combustion emissions of  $\text{CO}_2$  are caused by glasshouses. Similar to residential heating, the glasshouse emission factors are assumed to be a function of temperature. Yet, the energy consumption of glasshouses is dependent on many other variables as well, such as the humidity, light intensity, and the crop that is cultivated (Van der Velden and Smit, 2015). Moreover, most glasshouses have cogeneration plants that produce electricity and heat. Both can be used in the glasshouses, but sometimes electricity is supplied to the power grid depending on energy prices.

For the energy consumption of glasshouses there is no true activity data available. Instead, we use modelled daily energy consumption for a typical Dutch glasshouse cultivating tomatoes (courtesy of Bas Knoll, TNO) as the ‘truth’ (activity data) in Fig. 4.1. This time profile is calculated for typical meteorological conditions, such that the order of magnitude and the peaks are representative for an average year. There is almost no energy consumption during the summer, which indicates that there is no constant offset. So, we use Eq. 3 to determine the emission factor. Moreover, we use a lower temperature threshold of  $15^\circ\text{C}$  to get a better fit with the observed energy consumption. The estimated

function compares well with the activity data (Fig. 4.1) with an  $R^2$  of 0.85 ( $N=365$ ). The diurnal cycle of glasshouse emissions is likely to be different from that of household emissions. Yet we lack data to establish a diurnal cycle. We therefore use the same diurnal profile as for households, although this is likely to be incorrect.

### Power plants

Power plants can use different fuels such as hard coal, natural gas or biomass. In the Netherlands coal-fired and gas-fired power plants account for 80-85% of the total energy production. The remainder comes mainly from wind energy and biomass burning. Coal-fired power plants are currently the main source of energy and their generation is relatively stable compared to other sources. It does show a seasonal cycle with less energy production during the summer months. Gas-fired power plants have a larger temporal variability as they are mainly used as back-up for peak hours, depending also on the amount of renewable energy that is available. This is shown in Fig. 4.2 based on actual power generation in the Netherlands in 2016 reported by the European Network of Transmission System Operators for Electricity (ENTSO-E), which has detailed data available for the whole of Europe. The data are averaged over the day.



**Figure 4.2:** Daily time profiles for gas-fired (left) and coal-fired (right) power plants. Full lines are based on true activity data, whereas dashed lines are parameterizations based on observed temperature (coal) and wind speed/radiation (gas).

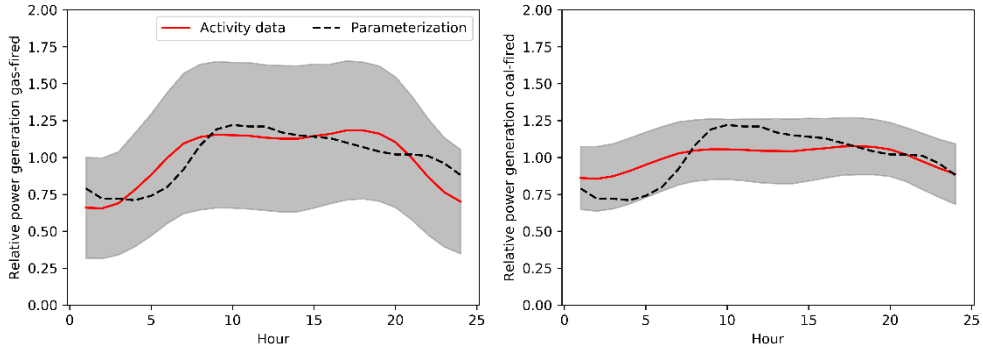
We use Eq. 5 to estimate the time profiles of power plants. Linear regression analysis shows that the variability in coal-fired power generation is (negatively) correlated with degree days. In this case we use a large constant offset of 80% and a threshold of 25°C which were chosen to best match the actual power generation data. The offset is much larger than for households because there is always a basic energy demand from the industry. The estimated time profile resembles the one based on activity data, although there is a clear mismatch in January-February and September-October. In contrast, the gas-fired power plants are (negatively) correlated with the wind speed and incoming solar radiation, indicating the need for gas-fired power generation in the absence of renewable sources. Therefore, we replace the temperature in Eq. 5 with the multiplication of wind speed and incoming solar radiation:



$$H = \max(10 - \overline{FF}, 0) \cdot \max(150 - \overline{R}, 0) \quad (6)$$

where  $FF$  is the wind speed ( $\text{m s}^{-1}$ ) and  $R$  the incoming solar radiation ( $\text{J cm}^{-2}$ ). Here we use a constant offset of 10% and a threshold of  $10 \text{ m s}^{-1}$  and  $150 \text{ J cm}^{-2}$ .

The diurnal cycles for power plants can be based on socio-economic factors. For example, the energy demand peaks early in the morning when people get ready to go to work and at the end of the afternoon when they get home. We find this pattern in the actual power generation data, with coal-fired power plants being less variable during the day than gas-fired power plants. The fixed profile from the European MACC-III emission inventory (Denier van der Gon et al., 2011; Kuenen et al., 2014) matches reasonably well with gas-fired power plant profiles, but it is less applicable for coal-fired power plants (Fig. 4.3). Overall, the estimated profiles for gas-fired power plants (hourly data) have an  $R^2$  of 0.32 ( $N=8784$ ) when compared to the activity data. For coal-fired power plants this is 0.21 ( $N=8784$ ).



**Figure 4.3:** Average diurnal cycle for gas-fired (left) and coal-fired (right) power plants. Full lines are based on true activity data, whereas dashed lines are fixed profiles from the MACC inventory (Denier van der Gon et al., 2011; Kuenen et al., 2014). Shading gives the  $1\sigma$  variability of the diurnal cycle based on activity data.

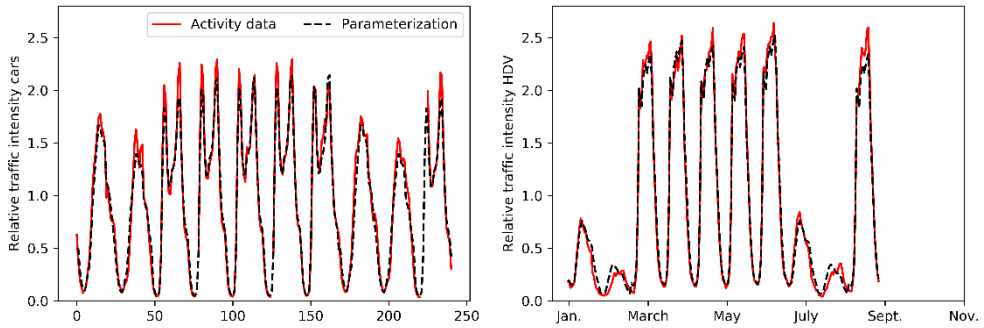
### Industry

This source sector consists of a wide range of activities, of which some are semi-continuous and only interrupted by maintenance stops while others follow working hours. This makes it very difficult to predict the temporal variability, especially for the overall sector. Since the largest  $\text{CO}_2$  emissions are related to refineries and heavy industry we will focus on these activities. We find a seasonal cycle in the reported industrial activity, with a small decline during the summer and Christmas holidays. However, the variations are very small (max. 1%). Therefore, we assume constant emissions.

### Road transport

Road transport emissions can vary between different road and vehicle types (Mues et al., 2014), but are also strongly dependent on environmental, socio-economic and driving conditions (such as the amount of stops, free-flow versus stagnant conditions, and engine

temperature). For example, during holidays or weekends there can be a lot of traffic heading for the coast or other attractions, which relocates emissions compared to regular weekdays.



**Figure 4.4:** Time profiles of road transport on highways for ten randomly chosen days in March. Full lines are based on true activity data, whereas dashed lines are parameterizations based on averaged traffic counts for Rotterdam.

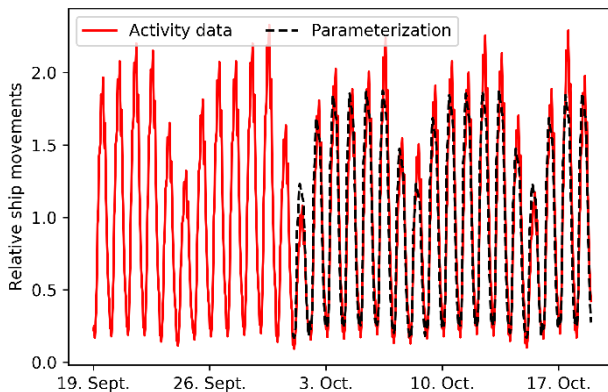
Traffic count data are often used to create average time profiles for road traffic emissions. However, with traffic counts we are unable to account for environmental and driving conditions. Traffic counts for the Netherlands are made available by the Nationale Databank Wegverkeersgegevens (NDW) and similar data is available in many developed countries. We differentiate between two vehicle types (passenger cars + motorcycles (hereafter referred to as cars) and light duty + heavy duty vehicles (LDV+HDV) (hereafter referred to as HDV) and three road types (highway, main road, urban road). We selected all available locations for 2014 within or close to Rotterdam that separate between 3-5 vehicle lengths and filtered for a minimum data coverage of 75%. This leaves us with 25 highway, 6 main road and 13 urban road locations. From this data we make average time profiles (daily, weekly and monthly) per road and vehicle type, as is often done to disaggregate road traffic emissions. Note that this method excludes any spatial variations (e.g. highways leading towards the city vs. the beach), except for differentiating between road types.

Generally, HDV show a larger spread due to the low counts during the weekend (Fig. 4.4). Car counts show a morning and evening rush hour and they go down in between. In contrast, HDV counts peak throughout the day and only go down after the evening rush hour. Moreover, the diurnal cycles are different during the weekend than on weekdays. These patterns can be explained from socio-economic factors. Current time profiles are often based on cars and are unable to correctly represent the temporal variability of HDV. This also affects the spatial distribution of emissions and therefore we create average diurnal, weekly and seasonal profiles separately for cars and HDV, for different road types and taking into account the day of the week. The comparison of true traffic counts and averaged traffic counts results in  $R^2$  values between 0.83 and 0.95 for hourly data for the whole year (N between 2665 and 6471).

### Shipping

Shipping emissions are dependent on the type of fuel they use and whether ships apply slow-steaming. Additionally, during loading and unloading ships still emit CO<sub>2</sub> and other pollutants, even though they are not moving. Such information is currently not available, so instead we used information about the arrival and departure of ships in the port of Rotterdam to make a time series of ship movements. Note that this only applies to large vessels that transport goods and passengers and that the time profile will look quite different for recreational shipping. However, large ships account for approximately 80% of the total shipping emissions in the area of interest. Since we lack information about other type of shipping movements we will only account for large ships in the time profiles.

We collected ship movements for one month (daily data) and an average diurnal profile (Fig. 4.5). The diurnal cycle shows a peak throughout the day, which corresponds well with the HDV road transport emission patterns on highways. The reason for this is that HDV road transport is related to shipping movements, as HDV takes care of good transport further inland after the goods have arrived by ship. We also find a clear weekly pattern with less ship movements during the weekend, although the decrease is less than for HDV road transport. This is likely because large ships, such as entering the port of Rotterdam, continue travelling during the weekend. Therefore, the weekly pattern resembles more that of car road transport on highways. Thus we can estimate ship movements by using the temporal profiles of HDV and cars on highways. This method is specifically tested for Rotterdam, but different patterns might be visible elsewhere. We also use HDV patterns for the seasonal variability. With this method the  $R^2$  reaches a value of 0.89 for a period of 18 days with hourly data ( $N=432$ ).



**Figure 4.5:** Daily time profiles for shipping. Full line is based on true activity data, whereas dashed line is a parameterization based on traffic counts of heavy-duty vehicles (diurnal cycle) and cars (day-to-day variations) on highways.

#### 4.2.4 Spatial disaggregation

The spatial disaggregation of emissions has received much more attention by inventory builders than temporal disaggregation. Therefore, the spatial patterns in inventories are relatively good and we will spend no further effort on improving them, neither will their

uncertainties be included in the following analyses. Moreover, our DA system will be set up to estimate only domain-average parameters. However, we do need to determine which fraction of the national total emissions are taking place in the Rijnmond area. For this purpose we disaggregate the total emissions calculated with our emission model to a  $1 \times 1 \text{ km}^2$  grid using simple proxies described below and we compare this to the Rijnmond emissions in the Dutch national inventory (Netherlands PRTR, 2014). We find that the yearly emissions for the Rijnmond region based on the dynamic emission model and the simplified spatial disaggregation based on these proxies are within a few percent from the total reported emissions for the same area.

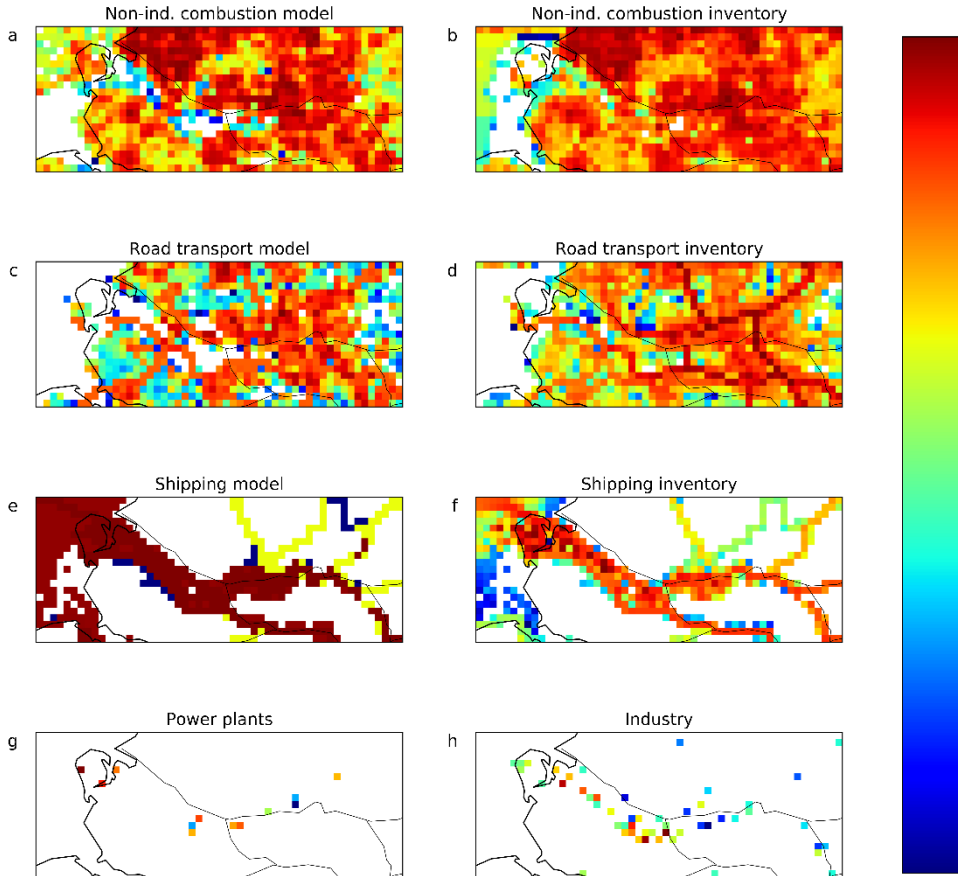


Figure 4.6: Spatial distribution of non-industrial combustion, road transport and shipping as in the dynamic emission model (left) and in the Dutch national emission inventory (Netherlands PRTR, 2014) (right). Spatial patterns of power plants and industry are taken from the Dutch national emission inventory. The colours indicate the relative emission strength in each pixel, each panel summing up to one for the entire Rijnmond area displayed here.

Household emissions are most easily located in space by using population density. However, for glasshouses this is not applicable, but they do appear very clearly on satellite images. This implies that the spatial distribution can be scaled with the area of a pixel

covered by glasshouses. We use these data to create an overall emission map of the non-industrial combustion emissions (Fig. 4.6a). The colours indicate the relative emission strength. We can observe two large centres of emissions. The first is the city of Rotterdam with abundant residential heating, the other is the glasshouse area near the coast. The spatial pattern in the emission model compares well to that of the emission inventory (Fig. 4.6b).

Road transport and shipping takes place only on roads and waterways. Therefore, tools like OpenStreetMap or Google Earth can be helpful to identify pixels in which to place road transport or shipping emissions. Main roads and urban roads are busiest in densely populated areas and we distribute the total road traffic emissions over these pixels using population density. Highways are used for transport between cities and therefore emissions take place outside densely populated areas as well. Nevertheless, highway transport is usually to and from densely populated areas, such that most emissions will take place close to cities. We therefore relate these emissions with the population density in the area of interest (here Rijnmond) relative to the rest of the country. We find that the overall emission pattern looks quite realistic, but too much of the emissions seem to be clustered in the city of Rotterdam compared to the inventory (Fig. 4.6c-d).

Shipping emissions are dependent on the type of ships and waterways. Large ships are the biggest emitters and use larger waterways. Rivers are often used for recreational purposes, resulting in less emissions on rivers (therefore recreational shipping is excluded in the remainder of this study). Therefore, the waterway characteristics are important to consider. We therefore separate between three categories: 1) (heavy) ocean shipping of bulk goods, oil, etc. which are dominant in the port and at sea on the main shipping routes, 2) inland shipping, mainly for transporting goods further inland via large rivers and canals, and 3) recreational shipping on (smaller) waterways away from the busy port area. With these categories we can attribute emissions to specific waterways, but we lack information about the spatial distribution within waterways. For example, at sea, just outside the main fairway, ships are at anchor while waiting before they are allowed to enter the port, which causes less emissions than a moving vessel. This is not taken into account in the simplified spatial distribution and the result is that too much emissions are located at sea outside the main fairway (Fig. 4.6e-f).

Power plants and industrial stacks are relatively easy to locate in space, as they can be identified through platforms like Google Earth. Even though this emission category also includes some diffuse sources, the stacks emit the vast majority of CO<sub>2</sub>. Locating these would thus be sufficient to represent the spatial distribution of these source sectors. The disaggregation of emissions over these stacks can be done using the capacity of each industrial plant. In Europe these plants are required to report their yearly emissions and this information can be used. However, a difficulty with stack emissions is that the stack height is of great importance for the dilution and transport of pollutants. In Rijnmond the stack heights are known and in addition the heat content of the plume is reported, which

is needed to calculate the plume rise. When such data is not available an estimate must be made per source sector, as was done by Bieser et al. (2011).

### 4.3 Parameter uncertainties

Using Eq. 1 and the temporal and spatial proxies for the activity within each emission category, we created dynamic hourly emissions of CO<sub>2</sub> and co-emitted tracers for the Rijnmond area. We can estimate the uncertainty in these high-resolution emissions from the uncertainties in the emission model parameters. All parameter values and their error shapes (mostly lognormal for emission ratios and normal for other variables) are summarized in Appendix A.

The second term in Eq. 1, the amount of energy consumed per amount of activity, is estimated from activity data (degree day sum or GDP) and reported energy consumption data. These data have a relatively small uncertainty. However, it is important to use recent numbers, because the trend in this term can be large. For example, the gas consumption by Dutch households per degree day has decreased by almost 15% between 1990 and 2014 (CBS, 2017).

Emission factors for CO<sub>2</sub> are determined by the fuel type that is burned and the carbon content of that fuel. Emissions of SO<sub>2</sub> are mostly dependent on the sulphur content of the fuel. Emission factors of CO and NO<sub>x</sub> are dependent on the conditions under which combustion takes place, such as the temperature. Moreover, sometimes catalysts or other (end-of-pipe) techniques are in place to reduce the emissions of air pollutants. Therefore, emission ratios are extremely variable in space and time and often their errors are correlated (see Appendix B for the full covariance matrix). In most cases the estimated emission factor for the Netherlands is within the range given in Appendix A. The only exception is the SO<sub>2</sub> emission factor from power plants. In the Netherlands SO<sub>2</sub> is effectively removed from the power plant flue gas, e.g. by desulphurization, resulting in much lower emission factors than derived from the general databases. We therefore scale the SO<sub>2</sub> emission factor for power plants down with a factor of 100 to reach agreement with the Dutch emission factor. We use the emission factors to convert CO<sub>2</sub> emissions to emissions of the other tracers (CO, NO<sub>x</sub>, SO<sub>2</sub>).

To estimate the uncertainty related to the time profiles we compare the time profiles based on activity data with those estimated from environmental and socio-economic conditions. We use the hourly profiles, as these will be used in the DA system for the best comparison with observations. Moreover, we assume the errors in some time profiles are correlated (i.e. covariance exists). We use the same time profiles for all tracers, although in reality differences exist as expressed in the temporal variability of emission ratios.

Most sectors have been subdivided in subsectors to better capture the wide range of emission factors and ratios and their specific spatiotemporal distribution. Although more detailed information is needed, this method has the advantage that the overall

uncertainty of each sector is reduced. For power plants, road traffic and shipping we estimate a fraction of the total energy consumption for each subsector and assume these are correct. A similar strategy is needed for the industrial sector, which contains a wide range of activities and processes with their own characteristics. Including one full range of emission factors and ratios for the industry in the emission model means that the industry fully dominates the uncertainty in the emissions. However, the industrial sector is so diverse that many subsectors need to be specified and a lot of information is required to represent their spatiotemporal variability. Given the complexity of this sector we assume, for now, a range in emission factors and ratios that only captures the most dominant subsectors. In the future we may have to revisit this assumption.

#### **4.4 Emission uncertainties**

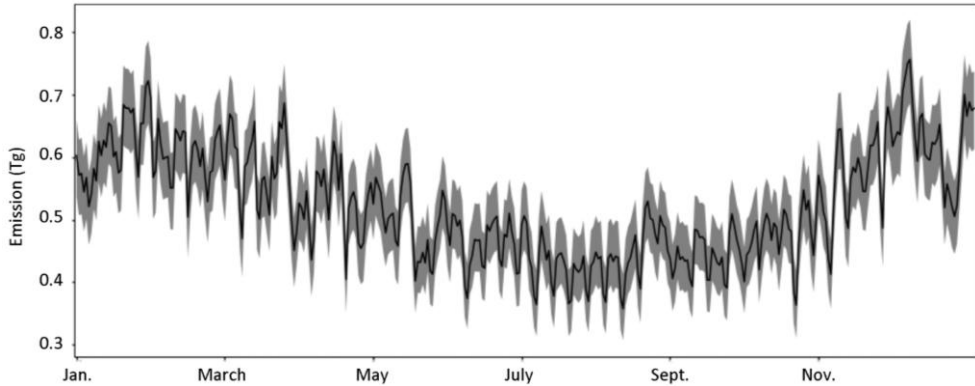
With the covariance matrix of all parameters in our dynamic emission model we create an ensemble ( $N=500$ ) that represents the variance of the parameter distributions, and then use these to create an ensemble of emission calculations in a Monte Carlo simulation. This provides us with a range in the total emissions, which represents the uncertainty in the emissions due to the uncertain input parameters. We report uncertainties in % ( $1\sigma$ ) for normal distributions ( $\text{CO}_2$ ) or as a 90% confidence interval (CI) for lognormal distribution (co-emitted species).

##### **4.4.1 Uncertainty in total emissions**

The total yearly emission of  $\text{CO}_2$  for the Netherlands calculated by the emission model is 180 Tg with an uncertainty of about 15%. In the Dutch national emission inventory the uncertainty for 2004 is estimated with a Monte Carlo simulation to be only 1% for  $\text{CO}_2$  (Ramírez et al., 2006), because national specific emission factors are used with a smaller uncertainty range (see Sect. 4.5 for more details). For CO,  $\text{NO}_x$  and  $\text{SO}_2$  the uncertainties in the emission model are much larger, with medians (CI's) of  $6.5 \times 10^8$  ( $1.3 \times 10^8 - 6.8 \times 10^9$ ) kg CO,  $5.0 \times 10^8$  ( $1.2 \times 10^8 - 5.1 \times 10^9$ ) kg  $\text{NO}_x$ , and  $1.3 \times 10^8$  ( $5.1 \times 10^6 - 2.2 \times 10^{10}$ ) kg  $\text{SO}_2$  per year. These ranges are the result of uncertainties in emission factors and the assumed ratios of release per unit of  $\text{CO}_2$  emitted. The time profiles play no role here, because the average time profile over a year is always one. In other words, the time profile has no impact on the total yearly emissions. Additionally, no spatial disaggregation is included in the estimated uncertainty of the national emission inventory.

For shorter time periods the time profiles have an impact on the error as well. The daily emissions of the Netherlands depend on the day and the season (Fig. 4.7) and range from 0.36 to 0.76 Tg  $\text{CO}_2$ . The time series shows a seasonal cycle with lower emissions during the summer. There is a clear weekly cycle with reduced emissions during the weekend. The uncertainty in the total daily emission varies between 8 and 15 %, which is similar to or lower than the uncertainty in the yearly total emissions. This is surprising, because

additional uncertainties are introduced. Indeed, the time profiles themselves cause an uncertainty in daily CO<sub>2</sub> emissions of about 7% if the other uncertainties are excluded from the analyses. A likely explanation for the relatively low uncertainties in the daily CO<sub>2</sub> emissions is that deviations are cancelled out. Each parameter is perturbed for each ensemble member. Possibly, positive deviations in some parameters are balanced by negative deviations in other parameters, causing a lower uncertainty in the total emissions than expected based on the uncertainties of individual parameters.



**Figure 4.7:** Time series of daily CO<sub>2</sub> emissions for the Netherlands (in Tg day<sup>-1</sup>). Given is the interquartile range (shaded area) and the median (line) from the ensemble.

Therefore, we further examine the impact of individual parameters on the uncertainty in the total emissions of CO<sub>2</sub> (Fig. 4.8, upper panel). In other words, we calculate the uncertainty in the total CO<sub>2</sub> emission caused by one individual parameter by setting the uncertainty of all other parameters to zero. For this purpose we performed a Monte Carlo simulation per parameter, setting all other parameters to the expected value. To ease the interpretation we normalized all the emissions by dividing by the median value of each simulation. For this analysis (and all following ones) we consider daily emissions for a random day in January as we plan to update the state vector in our DA system (see Chapter 5 of this thesis) on a daily basis. We find that the uncertainty in the daily CO<sub>2</sub> emission is mainly caused by the emission factors of the dominant source sectors, namely coal-fired power plants (4%, EF 1A), industry (13%, EF 3) and households (3%, EF 2A). Also the time profiles of the coal-fired power plants (T 1A) and households (T 2A) add some uncertainty (1% and 2%, respectively).



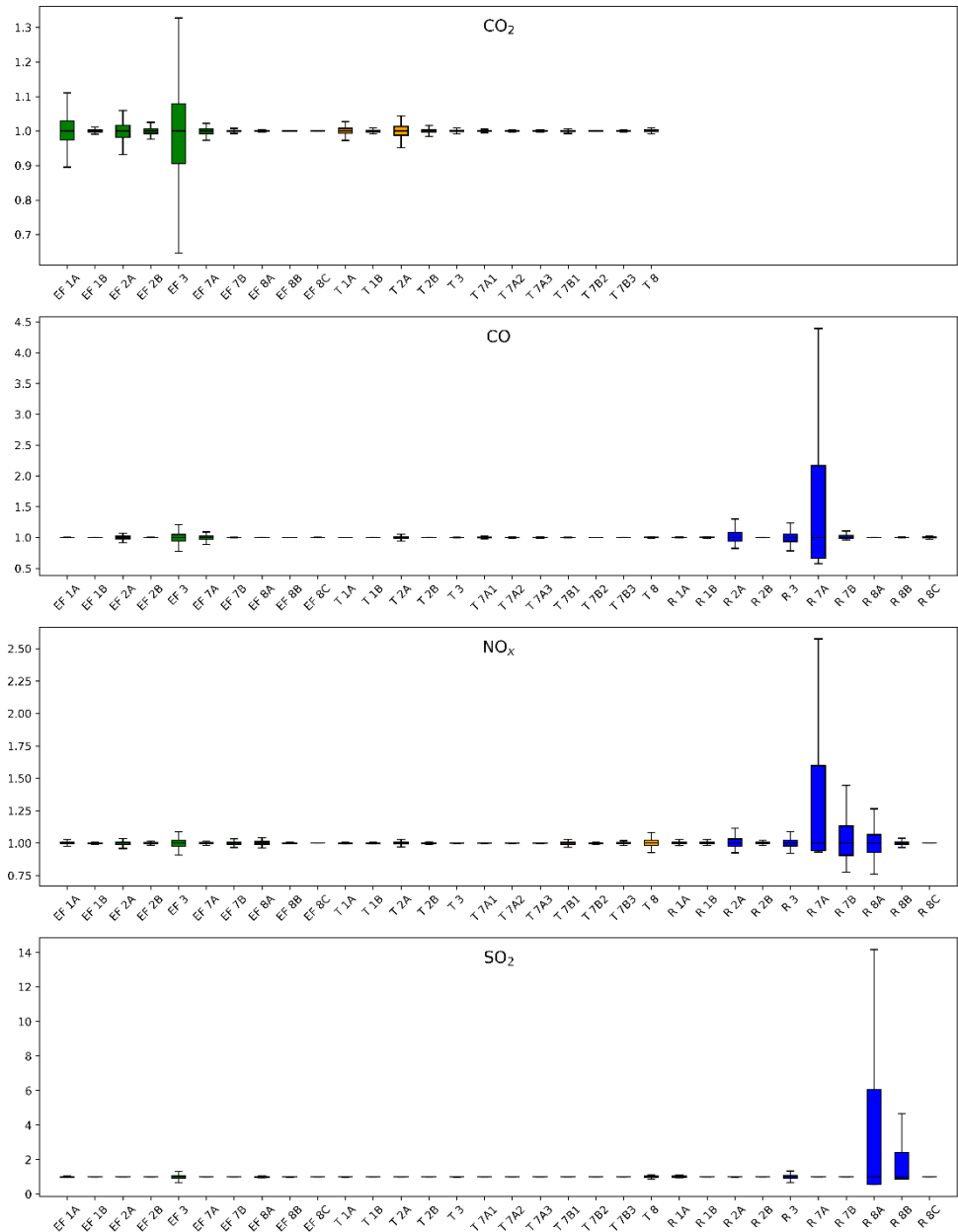


Figure 4.8: Box plots showing the uncertainty in the overall emission of CO<sub>2</sub>, CO, NO<sub>x</sub> and SO<sub>2</sub> caused by individual parameters. Emissions are normalized. EF refers to an emission factor (green bars), T to a time profile (orange bars) and R to an emission ratio (blue bars). (Sub)sectors are indicated with their short names as summarized in Table 4.1. In this analysis, uncertainties in industry (category 3), road transport (category 7) and shipping (category 8) dominate the national total emissions of the tracers.

#### 4.4.2 Uncertainty in emissions per sector

The main contributors to the overall uncertainty are not only the most uncertain parameters, but also those related to the dominant source sectors. If we want to study individual source sectors we need to know which parameters cause the largest uncertainties in each source sector. For this purpose we perform similar Monte Carlo simulations per parameter as before, but this time we calculate the uncertainty caused by a parameter in the CO<sub>2</sub> emission from the source sector related to that specific parameter. For example, we calculate the uncertainty caused by the emission factor of road traffic on the total road traffic emissions, and the effect of the industrial time profile on the total industrial emissions. Figure 4.9 shows that the road traffic and shipping sectors contain the smallest uncertainties, although the time profile for shipping causes an uncertainty of about 7% in the total shipping emissions. The industrial emissions are most uncertain and this is almost exclusively due to the emission factor, which causes an uncertainty of 41% in the total industrial emissions. Similarly, the power plant emissions have a large uncertainty due to the uncertain emission factor of coal-fired power plants (19%). Also for households and glasshouses the emission factor is uncertain (14% and 26%, respectively), but here the time profiles also have a large impact (10% and 16%, respectively).

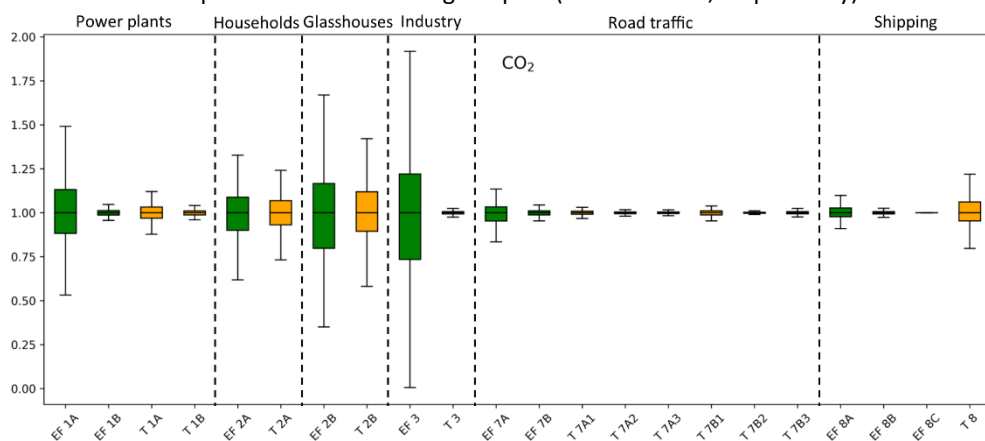


Figure 4.9: Box plots showing the uncertainty in the emission of power plants (1A+1B), households (2A), glasshouses (2B), industry (3), road traffic (7A+7B) and shipping (8A+8B+8C) caused by individual parameters affecting that sector. Emissions are normalized. EF refers to an emission factor (green bars) and T to a time profile (orange bars). (Sub)sectors are indicated with their short names as summarized in Table 4.1.

Moreover, we intent to use additional tracers in the DA system for source attribution. In order to do so we also need to consider the parameters that cause the largest uncertainty in those tracers. For example, if a species is dominated by road traffic emissions (like NO<sub>x</sub>) it can be used to attribute CO<sub>2</sub> emissions to that sector. However, an incorrect estimate of an emission ratio can attribute too much or too little CO<sub>2</sub> emissions to road traffic. Figure 4.8 shows the uncertainty caused by the individual parameters on the total emissions of the co-emitted species. Note that the overall uncertainties are much larger for these species, especially for SO<sub>2</sub>. We find that the largest uncertainty in CO emissions is caused

by the emission ratio of road traffic (cars). Additional parameters that are of importance for CO are the emission factors for households, industry and road traffic, and emission ratios of households and industry. Similar to CO, the  $\text{NO}_x\text{:CO}_2$  emission ratio for road traffic dominates the uncertainty in the overall  $\text{NO}_x$  emissions, but also the emission ratios for ocean shipping, industry and households have a significant impact. Finally, the emission factor of industry and the time profile of the shipping sector are of importance. For  $\text{SO}_2$  the ratios of shipping and industry cause the largest uncertainties. For industry the emission factor is also relevant.

#### **4.5 Discussion and conclusion**

The aim of this study is to gain insight in how we can calculate emissions from different data streams, and to quantify the uncertainties in the resulting emissions. For this purpose we developed a dynamic fossil fuel emission model. We have used relatively simple and easily accessible data in our model, such that it can be applied successfully to other cities and in other countries as well. However, more detailed data could significantly improve the emission estimates and generate emissions in near real-time. For example, with the emergence of smart meters in residential buildings a new source of data becomes available that can be used to predict emissions from households. Moreover, new vehicles are equipped with an on-board computer system including GPS. The on-board computer can monitor, amongst others, engine temperature, exhaust temperature and exhaust composition (Pucher, 2017). With such data the road transport emissions can be well established, also with more spatial detail. Similarly, shipping movements can be tracked with GPS data to improve the timing and spatial distribution of shipping emissions. However, such detailed data are not (yet) easily accessible and processing so much data would involve a large computational effort.

Each fuel type has its specific emission factor for  $\text{CO}_2$ , which is relatively well known and can be used to accurately determine the  $\text{CO}_2$  emission. However, the fuel mix can differ significantly between countries. For example, in the Netherlands most of the household heating is based on natural gas, whereas in eastern European countries more wood and coal is used with a different emission factor. Similarly, the composition of the vehicle fleet can vary strongly. Therefore, the emission model needs to be updated for each application, changing the assumptions on fuel mixes. Moreover, a different set of source sectors might be needed. For example, the importance of glasshouse emissions is very specific to Rijnmond and can be excluded in most other cases. Given the flexible set-up of the model this can be done easily.

We find an uncertainty of about 15% in the total yearly  $\text{CO}_2$  emission for the Netherlands. The Dutch Pollution Release and Transfer Register (PRTR) reports an uncertainty of only 1%. The main reason for this difference is that the PRTR uses emission factors established specifically for the Netherlands which have a smaller reported

uncertainty than the general ones used in this work. The reason we nevertheless start from these general emission factors is to test whether our DA system would be able to independently estimate the true emission factors in an area where local data is absent. Moreover, we doubt whether the 1% uncertainty reported by the PRTR is realistic. For example, reported emissions from diesel cars were found to be underestimated by a factor 3 (for CO<sub>2</sub>) up to a factor 40 (for NO<sub>x</sub>) (Brand, 2016). This scandal, also known as Dieselgate, illustrates that our assumptions about the real world might be less certain than we expect. Therefore, we want to include larger prior uncertainties to allow our DA system to drift away from what we currently assume is true. In this way we can independently estimate the emission factors from atmospheric observations. Note that the emission factors used by the PRTR are within the uncertainty range of our estimates.

Looking at the uncertainty in the total emissions we find that the emission factors are the most important parameters for CO<sub>2</sub>. Clearly the parameters related to the most dominant source sectors have the largest influence on the overall uncertainty. The impact of uncertainties in the time profiles are limited, except for households. We conclude that to reduce uncertainty in dynamically calculated emissions (the aim of Chapter 5), the emission factors of coal-fired power plants, households and industry and the time profile of households need to be optimized to better constrain the total CO<sub>2</sub> emissions.

We can further zoom in on specific source sectors and how co-emitted species can help to attribute CO<sub>2</sub> emissions to specific sectors. The power plants emit very little air pollutants, such that they cause large enhancements in CO<sub>2</sub> and very small enhancements in the other tracers. These signals are thus relatively easy to identify, which presents a nice opportunity to constrain CO<sub>2</sub> emissions from this sector. Moreover, the emission ratios from this sector are close to zero and thus not interesting to optimize. Non-industrial combustion contributes significantly to the CO emissions (14%). Both the emission factor, emission ratio and time profile of households contribute to the uncertainty in the CO emissions. Parameters related to glasshouses have no significant impact on total emissions of any of our tracers. The industry causes about 17% of the total CO emissions, 10% of the total NO<sub>x</sub> emissions and 35% of the total SO<sub>2</sub> emissions. It is therefore the only sector causing a significant signal in all the tracers included in the system. It is also a sector with large uncertainties due to the wide range of activities. As such, all parameters related to industrial emissions would benefit from optimization, leading to a better constraint on the CO<sub>2</sub> emissions from the industry. The industrial sector could potentially also benefit from a further division into refineries, chemical industry, etc. Road traffic dominates the CO emissions (60%) and also has a large contribution to NO<sub>x</sub> emissions (28%). The emission ratios related to cars are very uncertain and therefore have a huge impact on the uncertainty in CO and NO<sub>x</sub> emissions. The uncertainty in time profiles has little impact and therefore there is little need to optimize the road traffic time profiles. Finally, shipping is the dominant sector for both NO<sub>x</sub> (46%) and SO<sub>2</sub> (47%) emissions. The uncertainties related to this sector are specifically large for SO<sub>2</sub>, caused by

the emission ratios. Also the shipping time profile contributes significantly to the uncertainty of both species. Previous studies have shown that satellite data could put a constraint on shipping emissions of  $\text{NO}_x$  (Boersma et al., 2015), so they could prove useful to inform our emission model on temporal variations as well.

A side note to the usage of the co-emitted species is that we currently assume the same time profiles apply to all tracers. In reality, the emission of CO from road traffic is relatively large when the engine is cold (cold start), causing the emission ratio of CO to  $\text{CO}_2$  to be variable (Andrews et al., 2004). Similar processes occur for other source sectors and species as well, thus making the uncertainties in the emission ratios even larger than suggested here. This can be overcome by implementing species-specific time profiles, taking into account processes that affect the tracer ratios.

Based on this work we have thus identified the most uncertain and most important model parameters to improve the estimate of  $\text{CO}_2$  emissions per source sector. However, the industrial sector complicates doing this based on atmospheric measurements as used in an inversion, since it lacks a clear tracer ratio signature. This poses the risk that all unexplained variability in measured concentrations could be attributed to this sector, leaving other model parameters at their initial estimated values. Thus with the limited amount of information that is currently available, estimating hourly industrial emissions could benefit from continuous monitoring of the largest stack emissions, as was previously suggested by Ciais et al. (2014). This would also provide a constraint on the total yearly emissions and emission factors for all species.

## Appendix 4A: Dynamic emission model parameters

**Table A1: Overview of all parameters in the dynamic emission model, their unit, function type (log = lognormal), expected value and uncertainty (range). For normal distributions we quote the 1 $\sigma$  standard deviation.**

Parameter	(Sub)sector	Unit	Type	Expected value	Uncertainty
<b>Emission factor <sup>(A)</sup></b>	Coal-fired power plants <sup>(1)</sup>	kg/PJ	normal	1.01x10 <sup>8</sup>	23%
	Gas-fired power plants <sup>(1)</sup>	kg/PJ	normal	5.61x10 <sup>7</sup>	10%
	Households <sup>(1)</sup>	kg/PJ	normal	5.89x10 <sup>7</sup>	14%
	Glasshouses <sup>(1)</sup>	kg/PJ	normal	5.61x10 <sup>7</sup>	25%
	Industry <sup>(2)</sup>	kg/PJ	normal	7.66x10 <sup>7</sup>	40%
	Road traffic cars <sup>(3)</sup>	kg/PJ	normal	7.24x10 <sup>7</sup>	10%
	Road traffic HDV <sup>(3)</sup>	kg/PJ	normal	7.33x10 <sup>7</sup>	5%
	Ocean shipping <sup>(4)</sup>	kg/PJ	normal	7.76x10 <sup>7</sup>	5%
	Inland shipping <sup>(4)</sup>	kg/PJ	normal	7.30x10 <sup>7</sup>	5%
	Recreational shipping <sup>(4)</sup>	kg/PJ	normal	7.10x10 <sup>7</sup>	5%
<b>Emission ratio CO:CO<sub>2</sub></b>	Coal-fired power plants <sup>(3)</sup>	kg/kg	log	1.29x10 <sup>-4</sup>	8.7x10 <sup>-7</sup> –2.9x10 <sup>-4</sup>
	Gas-fired power plants <sup>(3)</sup>	kg/kg	log	8.47x10 <sup>-4</sup>	3.4x10 <sup>-4</sup> –2.5x10 <sup>-3</sup>
	Households <sup>(3)</sup>	kg/kg	log	3.88x10 <sup>-3</sup>	8.3x10 <sup>-4</sup> –9.6x10 <sup>-3</sup>
	Glasshouses <sup>(3)</sup>	kg/kg	log	5.40x10 <sup>-4</sup>	3.1x10 <sup>-5</sup> –7.7x10 <sup>-4</sup>
	Industry <sup>(2)</sup>	kg/kg	normal	2.06x10 <sup>-3</sup>	40%
	Road traffic cars <sup>(3)</sup>	kg/kg	log	1.32x10 <sup>-2</sup>	8.0x10 <sup>-5</sup> –6.5x10 <sup>-2</sup>
	Road traffic HDV <sup>(3)</sup>	kg/kg	log	2.22x10 <sup>-3</sup>	9.3x10 <sup>-5</sup> –1.3x10 <sup>-2</sup>
	Ocean shipping <sup>(4)</sup>	kg/kg	normal	2.32x10 <sup>-3</sup>	30%
	Inland shipping <sup>(4)</sup>	kg/kg	normal	3.42x10 <sup>-3</sup>	30%
	Recreational shipping <sup>(4)</sup>	kg/kg	normal	2.96x10 <sup>-1</sup>	30%
<b>Emission ratio NO<sub>x</sub>:CO<sub>2</sub></b>	Coal-fired power plants <sup>(3)</sup>	kg/kg	log	5.94x10 <sup>-4</sup>	3.0x10 <sup>-4</sup> –9.4x10 <sup>-4</sup>
	Gas-fired power plants <sup>(3)</sup>	kg/kg	log	2.00x10 <sup>-3</sup>	2.6x10 <sup>-4</sup> –3.7x10 <sup>-3</sup>
	Households <sup>(3)</sup>	kg/kg	log	1.50x10 <sup>-3</sup>	4.8x10 <sup>-4</sup> –3.3x10 <sup>-3</sup>
	Glasshouses <sup>(3)</sup>	kg/kg	log	1.63x10 <sup>-3</sup>	5.0x10 <sup>-4</sup> –3.5x10 <sup>-3</sup>
	Industry <sup>(2)</sup>	kg/kg	normal	6.56x10 <sup>-4</sup>	40%
	Road traffic cars <sup>(3)</sup>	kg/kg	log	1.76x10 <sup>-3</sup>	9.0x10 <sup>-5</sup> –7.5x10 <sup>-3</sup>
	Road traffic HDV <sup>(3)</sup>	kg/kg	log	1.11x10 <sup>-2</sup>	3.3x10 <sup>-4</sup> –3.7x10 <sup>-2</sup>
	Ocean shipping <sup>(4)</sup>	kg/kg	normal	2.32x10 <sup>-2</sup>	30%
	Inland shipping <sup>(4)</sup>	kg/kg	normal	1.37x10 <sup>-2</sup>	30%
	Recreational shipping <sup>(4)</sup>	kg/kg	normal	1.97x10 <sup>-3</sup>	30%

Table A1 (continued)

Parameter	(Sub)sector	Unit	Type	Expected value	Uncertainty
<b>Emission ratio SO<sub>2</sub>:CO<sub>2</sub></b>	Coal-fired power plants <sup>(3)</sup>	kg/kg	log	1.66x10 <sup>-4</sup>	2.9x10 <sup>-5</sup> –4.4x10 <sup>-4</sup>
	Gas-fired power plants <sup>(3)</sup>	kg/kg	log	5.01x10 <sup>-6</sup>	2.9x10 <sup>-6</sup> –7.2x10 <sup>-6</sup>
	Households <sup>(3)</sup>	kg/kg	log	2.21x10 <sup>-5</sup>	1.4x10 <sup>-5</sup> –6.7x10 <sup>-5</sup>
	Glasshouses <sup>(3)</sup>	kg/kg	log	8.91x10 <sup>-6</sup>	5.2x10 <sup>-6</sup> –1.3x10 <sup>-5</sup>
	Industry <sup>(2)</sup>	kg/kg	normal	4.28x10 <sup>-4</sup>	40%
	Road traffic cars <sup>(5)</sup>	kg/kg	normal	1.01x10 <sup>-6</sup>	100%
	Road traffic HDV <sup>(5)</sup>	kg/kg	normal	8.16x10 <sup>-7</sup>	100%
	Ocean shipping <sup>(4)</sup>	kg/kg	log	6.18x10 <sup>-3</sup>	3.3x10 <sup>-4</sup> –2.0x10 <sup>-2</sup>
	Inland shipping <sup>(4)</sup>	kg/kg	log	6.57x10 <sup>-3</sup>	3.5x10 <sup>-4</sup> –3.0x10 <sup>-2</sup>
	Recreational shipping <sup>(4)</sup>	kg/kg	log	3.14x10 <sup>-4</sup>	1.1x10 <sup>-4</sup> –7.0x10 <sup>-4</sup>
<b>Hourly time factor <sup>(6)</sup></b>	Coal-fired power plants	-	normal	1	28%
	Gas-fired power plants	-	normal	1	43%
	Industry	-	normal	1	5%
	Households	-	normal	1	43%
	Glasshouses	-	normal	1	74%
	Road traffic cars highw.	-	normal	1	18%
	Road traffic cars main r.	-	normal	1	18%
	Road traffic cars urban r.	-	normal	1	18%
	Road traffic HDV highw.	-	normal	1	41%
	Road traffic HDV main r.	-	normal	1	18%
	Road traffic HDV urban r.	-	normal	1	48%
	Total shipping	-	normal	1	31%
<b>Energy consumption per activity data <sup>(7)</sup></b>	Total power plants	PJ/mIn €	-	8.22x10 <sup>-4</sup>	-
	Households	PJ/dd <sup>(*)</sup>	-	0.199	-
	Glasshouses	PJ/dd <sup>(*)</sup>	-	0.061	-
	Industry	PJ/mIn €	-	7.05x10 <sup>-4</sup>	-
	Road traffic cars	PJ/mIn €	-	3.98x10 <sup>-4</sup>	-
	Road traffic HDV	PJ/mIn €	-	2.01x10 <sup>-4</sup>	-
	Total shipping	PJ/mIn €	-	1.51x10 <sup>-4</sup>	-

<sup>(a)</sup> Emission factor for coal-fired and gas-fired power plants include uncertainty due to variations in fuel type, including burning of biomass (5% uncertainty). For households we assumed 8% wood combustion ([http://www.emissieregistratie.nl/erpubliek/documenten/Lucht%20\(Air\)/Consument,%20Kleinbedrijf%20en%20HDO%20\(Consumers\)/2016%20R10318%20Wood-burning%20stove.pdf](http://www.emissieregistratie.nl/erpubliek/documenten/Lucht%20(Air)/Consument,%20Kleinbedrijf%20en%20HDO%20(Consumers)/2016%20R10318%20Wood-burning%20stove.pdf)), the remainder is natural gas (with 10% uncertainty). For glasshouses we assumed only natural gas combustion, including 20% additional uncertainty due to use of cogeneration plants. For road traffic cars we assumed 69% gasoline, 29% diesel and 2% LPG (with 5% uncertainty); for road traffic HDV we assumed 100% diesel.

Data sources:

(2) Expected value based on Emissieregistratie (emission) and CBS (energy consumption); uncertainty based on expert judgement

[http://efdb.apps.eea.europa.eu/?source=%7B%22query%22%3A%7B%22match\\_all%22%3A%7B%7D%7D%2C%22display\\_type%22%3A%22tabular%22%7D](http://efdb.apps.eea.europa.eu/?source=%7B%22query%22%3A%7B%22match_all%22%3A%7B%7D%7D%2C%22display_type%22%3A%22tabular%22%7D)

(4) Expected value and uncertainty based on  $CO_2$ ,  $CH_4$ , and  $N_2O$  emissions from transportation-water-borne navigation, by Paul Jun, Michael Gillenwater, and Wiley Barbour ([http://www.ipcc-nggip.iges.or.jp/public/gp/bgp/2\\_4/Water-borne\\_Navigation.pdf](http://www.ipcc-nggip.iges.or.jp/public/gp/bgp/2_4/Water-borne_Navigation.pdf))

(5) Expected value based on Air Pollutant Emission Factor Library (<http://www.apec-library.fi/>); uncertainty based on expert judgement

(6) Uncertainties based on comparison activity data-based time profiles and estimated time profiles from environmental/socio-economic factors

<sup>(7)</sup> Expected value based on CBS (energy consumption, GDP) and KNMI (degree day sum)

(8) Expected value based on Emissieregistratie



Appendix 4B: Covariance matrix

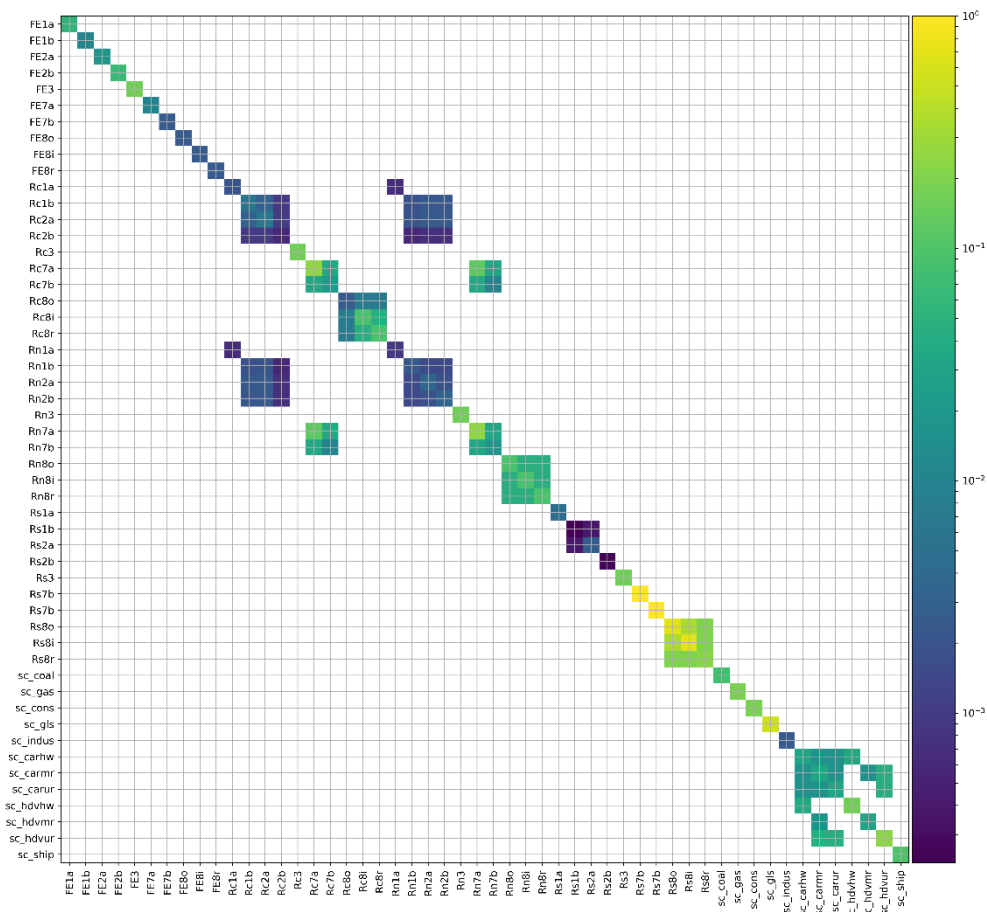



Figure B1: Covariance matrix for all parameters in the dynamic emission model. For all covariances we assume a correlation coefficient of 0.5.

A thick, dark vertical bar runs down the left side of the page. A light teal arrow points to the right from this bar, containing the text 'Chapter 5'. Below the bar, several thin, curved lines in dark and light grey sweep upwards from the bottom left towards the center of the page.

## Chapter 5

# **Optimizing a dynamic fossil fuel CO<sub>2</sub> emission model for an urban area**

*Authors: Ingrid Super, Wouter Peters, Michiel van der Molen, Hugo Denier van der Gon*



**Abstract**

Emission monitoring in urban areas is receiving increasing attention as a way to support decision making in the context of CO<sub>2</sub> emission reductions, but so far quite some challenges remain when applying inverse modelling at such small scale. In this study we address some of these challenges in order to gain more insight in the limitations of urban inversions, the (dis)advantages of several network configurations and the importance of having a prior estimate at high resolution. We apply an inverse modelling framework to an urban-industrial complex and make use of pseudo-observations to be able to perform sensitivity tests. As prior we employ a dynamic fossil fuel emission model and we optimize the model parameters. This method shows promising results in improving our understanding of the emission landscape of the city, as the model parameters can be translated into physically relevant variables and processes. Moreover, we show that the optimization can be successful in data-scarce regions where, for example, the temporal variability of emissions is unknown. In addition, we extend our inversions with co-emitted tracers (CO, NO<sub>x</sub> and SO<sub>2</sub>) that help attribute atmospheric CO<sub>2</sub> signals to specific source sectors. In this way we are able to distinguish the effect of emission reduction policies targeting specific source sectors, especially with urban measurement sites. However, when using erroneous atmospheric transport the results deteriorate and may even become physically unfeasible. This illustrates the importance of modelling correct atmospheric transport in inversions using true observations, which remains the number one challenge in urban inversions.

## 5.1 Introduction

During recent years, international debate has resulted in important climate agreements that support climate mitigation and adaptation efforts. Within the 2015 Paris Agreement 195 nations agreed with a climate action plan, in which each nation sets its own climate targets, also related to carbon emission reductions (UNFCCC, 2015). An important role in reaching emission reduction targets is laid out for cities, which emit a large portion of the global fossil fuel CO<sub>2</sub> emissions (about 70% according to the International Energy Agency (IEA, 2008)). All efforts related to the Paris agreement should be reported regularly to the *United Nations Framework Convention on Climate Change* (UNFCCC). The Paris Agreement also states that parties should strengthen their cooperation, also with respect to the sharing of information and good practices. Within this context it becomes increasingly important to map fossil fuel emissions and to quantify emission trends. Monitoring emission trends can help understand the effectiveness of mitigation strategies, which is valuable information to share with other member states. Besides, monitoring can reveal unexpected growth in emissions that reduce the effect of emission reduction policies.

Most greenhouse gas emission estimates are currently based on yearly, national emissions reported to the UNFCCC, which are often disaggregated using activity data or proxies for activities like population density (Hutchins et al., 2017; Kuenen et al., 2014). Although the yearly estimates are reasonably accurate at national scale (estimated uncertainty for developed countries is less than 8% for CO<sub>2</sub>), their uncertainty increases when disaggregating them towards finer spatiotemporal resolutions, such as hourly/city scale (Ciais et al., 2010; Nassar et al., 2013). A method to improve emission estimates is using transport models in combination with independent observations of atmospheric mixing ratios, generally called data assimilation (DA). Recently, efforts have been made to apply DA techniques to the urban environment (Boon et al., 2016; Bréon et al., 2015; Brioude et al., 2013; Brophy et al., 2018; Fischer et al., 2017; Graven et al., 2018; Lauvaux et al., 2013; Lauvaux et al., 2016; McKain et al., 2012). Yet, for urban DA the fine scales (less than 1km and less than 1 hour) need to be resolved, therefore putting a higher demand on the DA system. For example, Boon et al. (2016) mentioned that sources with a small spatial extent (point sources) are not correctly represented on a 2x2 km<sup>2</sup> grid, while these sources have a significant impact on the locally observed mixing ratios. Similarly, we have previously shown that a plume model is necessary to correctly represent sources with a limited spatial extent (Super et al., 2017a). These model limitations result in strict data selection criteria to favour conditions that can be represented correctly by atmospheric transport models. This in turn can lead to biased emission estimates by the inversion framework as specific sources are underrepresented, or even missed, in the observations.

Additionally, there is an ongoing debate about the type of in-situ measurement sites that could best be used in urban inversions. On the one hand, Kort et al. (2013) argue that

in-city surface sites are most exposed to urban fluxes and therefore provide the best constraints. On the other hand, Bréon et al. (2015) found that their in-city observation location was too prone to transport errors and they used horizontal concentration gradients from measurements on the edge of the city. We have shown similar results in Chapter 3 where we show that observations at urban sites contain the most information, but the models have trouble simulating the correct transport, especially in the presence of point source emissions (Super et al., 2017a). Moreover, Turner et al. (2016) argued that a dense network with moderate precision is preferred over a smaller network with high precision. Other types of observations, such as flux measurements, also have their own (dis)advantages. Nevertheless, most studies show that DA is a promising tool to objectively estimate CO<sub>2</sub> fluxes from urban areas, especially when transport errors can be reduced or better quantified (Lauvaux et al., 2016).

So far, urban DA studies have tried to constrain the total fossil fuel flux to validate bottom-up CO<sub>2</sub> inventories, often without considering the underlying emission process that caused the mismatch between observed and modelled concentrations. As one of very few exceptions, Lauvaux et al. (2013) used the CO:CO<sub>2</sub> ratio to conclude that the emission reduction in Davos during the World Economic Forum 2012 was likely related to reduced traffic, but without a quantification. However, emission reduction policies usually target specific source sectors. Therefore, an increase in fossil fuel emissions from one source sector can cause the total CO<sub>2</sub> emissions to appear stable, although a policy targeting another source sector can be effective in itself. In order to monitor the effect of each measure independently it becomes essential to attribute changes in the total CO<sub>2</sub> emissions to these policies and thus to specific source sectors. It is, therefore, not sufficient to constrain the total CO<sub>2</sub> flux, but we need to differentiate the CO<sub>2</sub> signal into signals from the different source sectors. One way to accomplish this is by using additional measurements of co-emitted species and isotopes. Such methods have previously been used to differentiate between biogenic and anthropogenic emissions or between fuel types (Djuricin et al., 2010; LaFranchi et al., 2013; Lopez et al., 2013; Super et al., 2017b; Turnbull et al., 2015), but also to separate between different fossil fuel sources (Lindenmaier et al., 2014; Super et al., 2017a).

The aim of this study is to explore how well we are able to constrain urban CO<sub>2</sub> emissions per source sector given the errors/uncertainties that exist in the emissions and transport models. The research questions are:

- Which type of monitoring framework is most suitable to constrain urban fossil fuel emissions?
- Is the DA system able to attribute residuals to the correct parameters when the prior contains large errors?
- How sensitive is the DA system to atmospheric transport errors?
- What are the (dis)advantages of propagating gained knowledge to the next DA cycle?

In this study we use observing system simulation experiments (OSSEs, experiments using pseudo-observations), applied to the urban-industrial complex of Rotterdam (Netherlands), to answer these questions. For this purpose, we developed a data assimilation framework that uses atmospheric observations to optimize CO<sub>2</sub> emissions and attributes them to specific source sectors. Underlying the inverse method is a dynamic emission model that estimates hourly, 1x1 km<sup>2</sup> resolution emissions from different source sectors using a wide range of (statistical) data. The DA system optimizes parameters in the emission model rather than the total CO<sub>2</sub> emissions and by including four co-emitted trace gases (CO<sub>2</sub>, CO, NO<sub>x</sub> and SO<sub>2</sub>) we enable source attribution. The model is dynamic in the sense that it responds to environmental conditions and can be used to calculate emissions in near real-time. We use data that is widely available so that the emission model can be applied in regions where currently little is known about the emission landscape. The emission model is also flexible and can be adapted easily to different areas. The description of the dynamic emission model is found in Chapter 4 of this thesis.

This idea of a parameter inversion is not new, but has been applied mainly to biogenic fluxes of CO<sub>2</sub> (Koffi et al., 2013; Rayner et al., 2014; Tolk et al., 2011). Previous research has shown that a parameter inversion infers a good estimate of net ecosystem exchange, although non-linear model parameters can be more difficult to solve (Tolk et al., 2011). An advantage of a parameter inversion is that the result has more physical meaning than an updated emission. For example, emissions from consumer heating can decrease from one year to another, but that could either be the result of higher temperatures or a lower emission factor. With the emission model we can predict the emissions based on the observed temperature and the resulting discrepancy should thus be attributed to a change in emission factors.

We start with an overview of the DA system components and the model settings. Next, we discuss the results following the research questions. We start with the comparison of different observation networks, including the ability to use co-emitted tracers to attribute CO<sub>2</sub> emissions to specific source sectors. Next, we quantify the uncertainty caused by model errors, both in the dynamic emission model (prior) and the atmospheric transport model. Finally, we discuss the effect of propagating optimized parameter values and/or posterior uncertainties.

## **5.2 Methods**

### **5.2.1 Data assimilation principles**

The goal of data assimilation is to find a state at which the system is in optimal agreement with observations. In this study we employ the CarbonTracker Data Assimilation Shell (CTDAS) described in detail in Van der Laan-Luijkx et al. (2017). Briefly, the CTDAS system is a flexible implementation of a square-root Ensemble Kalman Filter (Whitaker and Hamill, 2002), which also allows lagged windows (i.e. smoothing instead of filtering).

Although originally based on the DA system CarbonTracker (Peters et al., 2007) and the TM5 atmospheric transport model (Krol et al., 2005), CTDAS's implementation in python offers a flexible DA system than can be easily adapted to other transport models and new data sets. In our study we have for example replaced the observation operator of CTDAS (the TM5 model) with a combination of WRF-STILT footprints, the OPS plume model, and a fossil fuel emission model (discussed in detail in the following sections).

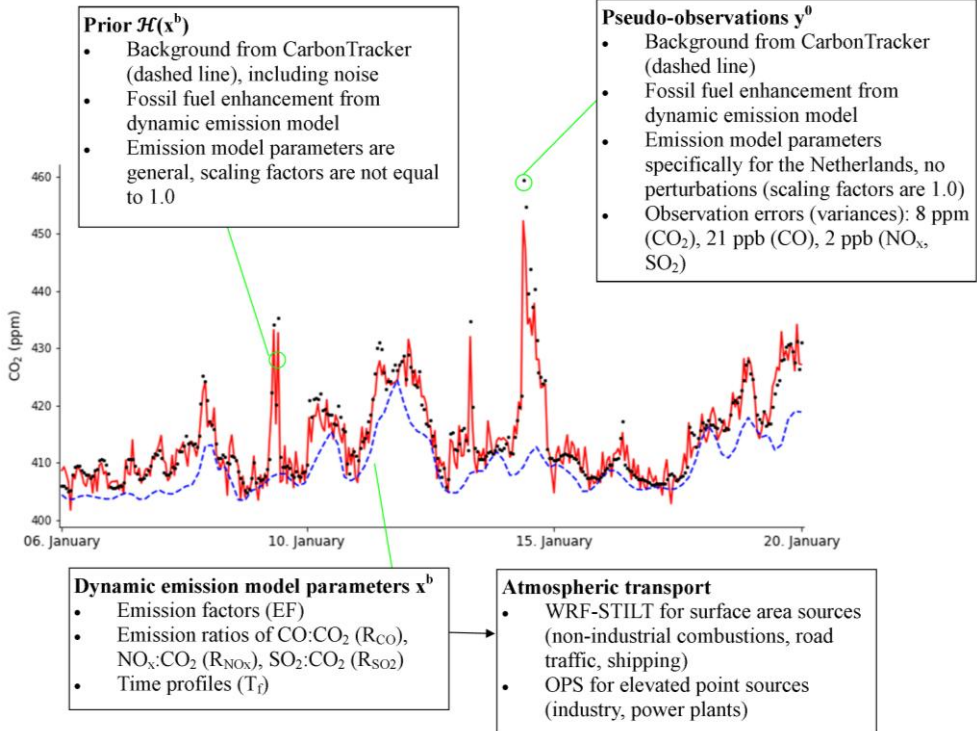


Figure 5.1: Time series of pseudo-observations and prior  $\text{CO}_2$  mixing ratios and a summary of how these time series were created. The CarbonTracker background (used for the pseudo-observation) is given by the dashed line.

We use the ensemble Kalman smoother as optimization method (Cohn et al., 1994; Peters et al., 2005). This method optimizes the cost function for variables in the state vector  $x$  using information from observations ( $y^0$  with covariance  $R$ ) and a prior estimate of the state vector ( $x^b$  with covariance  $P^b$ ).

$$J(x) = (y^0 - \mathcal{H}(x))^T R^{-1} (y^0 - \mathcal{H}(x)) + (x - x^b)^T (P^b)^{-1} (x - x^b) \quad (1)$$

In this function,  $\mathcal{H}$  is the observation operator that returns simulated mole fractions given the state vector.  $R$  and  $P$  determine how much weight is given to the observations and prior estimate, respectively.



The optimized state vector (indicated with superscript  $a$ , whereas  $b$  refers to the prior estimates) which minimizes the cost function is

$$x^a = x_t^b + K(y_t^0 - \mathcal{H}(x_t^b)) \quad (2)$$

and its covariance is

$$P_t^a = (I - KH)P_t^b \quad (3)$$

Here,  $H$  is the linearized observation operator and  $K$  is the Kalman gain matrix:

$$K = (P_t^b H^T)(H P_t^b H^T + R)^{-1} \quad (4)$$

The solutions of Eq. 2 and Eq. 3 are calculated as in Peters et al. (2005) using an ensemble of 40 members. More details in the individual parts of this system are provided below and are summarized in Fig. 5.1.

### 5.2.2 Modelling framework

#### ***Dynamic emission model***

In most inversions the observation operator uses  $\text{CO}_2$  flux fields from a biosphere model or emission inventory as input. In this study we use a dynamic fossil fuel emission model to drive our transport models such that we can optimize the parameters in this model. The emission model is developed for the Netherlands and is focused over Rotterdam (Fig. 5.2). This is one of the major cities in the Netherlands (about 625,000 inhabitants) with the largest sea port of Europe to the west. It is located in a larger urbanized area (Randstad, about 7 million inhabitants) with The Hague, Amsterdam and Utrecht being other major cities. A large area to the southwest of The Hague is covered with glasshouses. It is characterized by a complex mixture of anthropogenic and industrial activities and therefore we distinguish five source sectors and a total of ten sub-sectors to construct its emissions (see Table 5.1).

The emissions are calculated in three steps. Firstly, the yearly, national emission is calculated per sector using activity data and  $\text{CO}_2$  emission factors. Secondly, we apply temporal disaggregation to hourly emissions using local activity data or environmental conditions. Finally, we downscale the national totals to  $1 \times 1 \text{ km}^2$  resolution using statistical data, such as population density. The emission model thus contains several parameters per source sector: activity, emission factor, and time profile.

Table 5.1: Overview of sectors and subsectors distinguished in the dynamic emission model, including their short name used in the figures and their approximate contribution to the total CO<sub>2</sub> emission in Rotterdam. Crosses indicate which emission factors (EF), tracer ratios of CO, NO<sub>x</sub> or SO<sub>2</sub> ( $R_{CO}$ ,  $R_{NOx}$ ,  $R_{SO2}$ ) and time profiles ( $T_f$ ) are part of the state vector.

Sector	Subsector	Short name	Contribution	EF	$R_{CO}$	$R_{NOx}$	$R_{SO2}$	$T_f$
Power plants	Gas-fired	1A	37%	X				
	Coal-fired	1B		X				
Non-industrial combustion	Households	2A	15%	X	X			X
	Glasshouses	2B		X				X
Industry		3	39%	X	X	X	X	X
Road traffic	Cars	7A	6%	X	X	X		
	Heavy duty vehicles	7B		X	X	X		
Shipping	Ocean shipping	8A	3%	X		X	X	X
	Inland shipping	8B		X		X	X	
	Recreational shipping	8C						

The model parameters are based, as much as possible, on local/national data. For example, we use traffic counts for the Rotterdam area to perform temporal disaggregation of road traffic emissions. These data are considered to be the true state of the emission landscape in Rotterdam. When such data are not available more general data can be used. For example the emission factors can be taken from large-scale databases from the IPCC and other organisations (see Chapter 4 of this thesis). However, these values then have a large uncertainty, which propagates into the calculated emissions of the dynamic emission model. The advantage of this model is that it uses data that are widely available in near real-time, so that it can be applied around the world as prior for DA systems. More detailed information about the dynamic emission model can be found in Chapter 4 of this thesis.

One major challenge in this study is to attribute the mismatch between observations and model to a specific sector, as a CO<sub>2</sub> observation alone provides no details on the origin of the CO<sub>2</sub>. To retrieve more information about a single CO<sub>2</sub> observation we included three additional tracers in the DA system: CO, NO<sub>x</sub> and SO<sub>2</sub>. These species are co-emitted with CO<sub>2</sub> during fossil fuel combustion, but often with a distinct ratio (referred to as  $R_{CO}$ ,  $R_{NOx}$  and  $R_{SO2}$ ) (Fig. 5.3). These ratios act as a signature of each source sector, and the co-emitted tracers thereby allow the DA system to attribute observed CO<sub>2</sub> changes to a specific sector. The emissions of these trace gases are calculated from the CO<sub>2</sub> emissions using typical emission ratios per source sector, which in reality are highly variable and uncertain and can thus affect the optimization results, as we will investigate and discuss further below. The emission ratios are also parameters in the dynamic emission model.

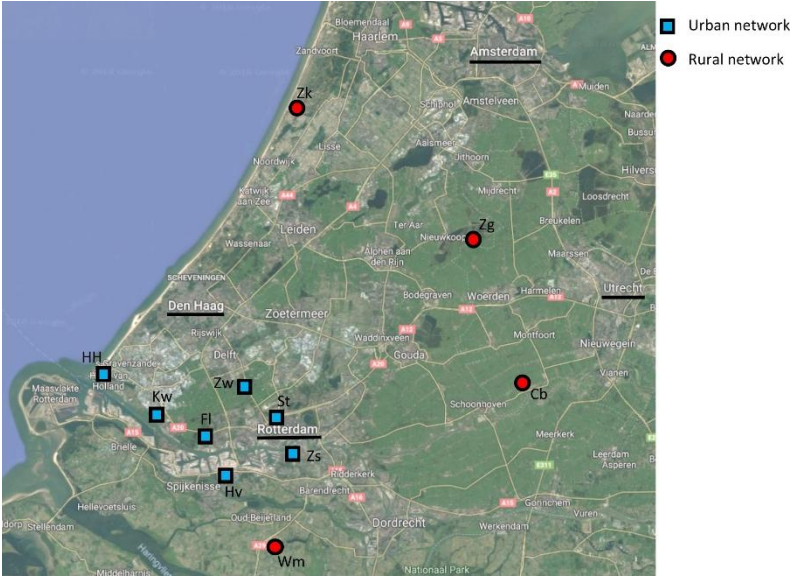


Figure 5.2: Map of the domain covered within the inversions, including several major cities (underlined). The squares and circles show the locations of the measurement sites within the urban (7 sites) and rural (4 sites) network configurations.

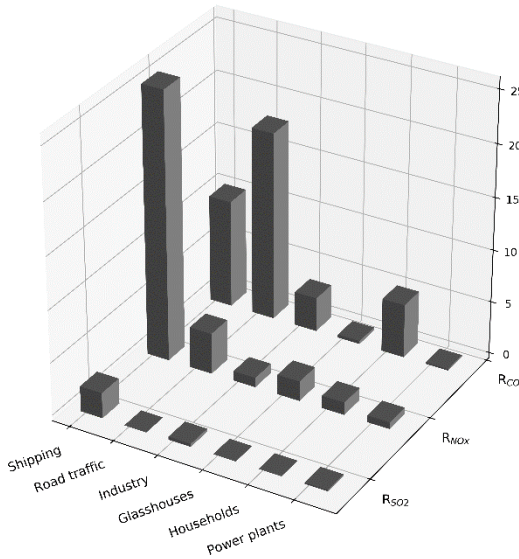


Figure 5.3: Emission ratios of CO<sub>2</sub>:CO<sub>2</sub> ( $R_{CO}$ ), NO<sub>x</sub>:CO<sub>2</sub> ( $R_{NOx}$ ) and SO<sub>2</sub>:CO<sub>2</sub> ( $R_{SO2}$ ) for specific source sectors based on the Dutch Pollution Release and Transfer Register (Netherlands PRTR, 2014). Units are in ppb ppm<sup>-1</sup>. A value of 10 on the y-axis thus implies that for each 1000 moles of CO<sub>2</sub>, 10 moles of the auxiliary tracer is emitted.

### Observation operator

The observation operator takes care of the translation from the state vector (in our case parameters of the emission model) first into emissions (through the emission model) and then into atmospheric mixing ratios that can be compared to observations (through

transport modelling). In our study the transport modelling consists of two parts. The first part, the Weather Research and Forecasting-Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model, is used for surface area emissions only. STILT is a Lagrangian particle dispersion model that describes the footprint of a single measurement by dispersing particles back in time (Gerbig et al., 2003; Lin et al., 2003). With this footprint the surface influence of emissions on a single observation can be described. An advantage of this method is that it allows the pre-calculation of linear atmospheric transport, which makes this part of the observation operator less computationally demanding than running an ensemble of a full atmospheric transport model (like WRF with chemistry). The total domain covered with WRF-STILT is 77 x 88 km (Fig. 5.2) and includes most of the Randstad.

To generate a footprint, 75 particles are released at the observation site at the start of the back-trajectory and followed back in time. Given that the variability in hourly observations at an urban location is dominated by local signals, we choose to construct back-trajectories of 6 hours. This is based on the domain size, which could be covered within 6 hours for typical wind speeds. Within this time frame emissions can become well-mixed throughout the boundary layer under typical mixing conditions, such that emissions outside this range can be represented by a boundary inflow. Footprints are generated for each hour within the back-trajectory to account for hourly variations in the emissions. We drive STILT with meteorology from the WRF model (v3.5.1) (Nehrkorn et al., 2010). The WRF model was set up with two nested domains (15x15 and 3x3 km<sup>2</sup> horizontal resolution) and the footprints have a 1x1 km<sup>2</sup> resolution over the entire domain.

The second part of the transport modelling is a plume model. In a previous study we have shown that point source (stack) emissions should be modelled with a plume model to better represent the limited dimensions of the stack plume (Super et al., 2017a). Similarly, Vogel et al. (2013) have shown that the surface influence calculated by STILT can lead to large model errors for stack emissions. Therefore, we include the OPS (Operational Priority Substances) plume model in our framework to model the transport and dispersion of stack emissions (Sauter et al., 2016; Van Jaarsveld, 2004). OPS provides hourly concentrations at pre-defined receptor points, which represent our measurement sites. The model keeps track of a plume trajectory, taking into account time-varying transport over longer distances (e.g. changes in wind direction and dispersion). If for a time step a specific plume affects the receptor, a Gaussian plume formulation is used to calculate the mixing ratio caused by that source based on the true travel distance along the trajectory. We drive the model with the same WRF meteorology as STILT.

Similar to the WRF-STILT model, we assume an influence time of 6 hours on our observations. However, in this case we run the OPS model forward from -6 hours to the time of observation. We apply the OPS model only to point source emissions within the Rijnmond area, as we found in Chapter 4 that a plume model only has an added value less than 10 km downwind from the stack (Super et al., 2017a). Point sources at more than 10

km from the observation site can be sufficiently represented with a Eulerian model. The OPS model input includes detailed information about the exact stack height and heat content of the plume.

In addition to the fossil fuel contribution we also include background mixing ratios for CO<sub>2</sub> and CO. NO<sub>x</sub> and SO<sub>2</sub> are short-lived and therefore the variations in the background are relatively small compared to the fossil fuel signals. The CO<sub>2</sub> background is taken from the 3-D mole fractions from CarbonTracker Europe (Peters et al., 2010) and also accounts for biogenic fluxes. The resolution of these CO<sub>2</sub> fields is 1x1° and we select the pixel that is situated over Rotterdam. The 3-hourly data are linearly interpolated to get hourly background mixing ratios and are added to the fossil fuel signals calculated by the transport models. We use the strong correlation between CO<sub>2</sub> and CO mixing ratios ( $r = 0.73$ ) to calculate CO background conditions from the CO<sub>2</sub> background. This is not very accurate, but for the purpose of this OSSE it provides us with a decent estimate of the variability in background mixing ratios. For the pseudo-observations (see Sect. 5.2.4) the original time series is used, whereas in the inversion random noise is added to the background mixing ratios with a standard deviation of 2 ppm for CO<sub>2</sub>.

### 5.2.3 State vector

As mentioned before, the emission model uses as much as possible local or national data that we consider to be the truth and are used to create the pseudo-observations (see Sect. 5.2.4). However, we start from a different state where we lack local data and typical values from large-scale data bases are used (as listed in Appendix A of Chapter 4). The prior state vector ( $x^b$ ) consists of 9, 21 or 25 scaling factors that each relate to a parameter from the dynamic emission model, depending on the experiment (see Sect. 5.2.5). The scaling factors quantify the difference between the truth and the prior, where true scaling factors are always 1.0. In Chapter 4 we examined the importance of each model parameter and we selected the most important ones for the state vector. Parameters are important when they are either very uncertain or when they have a large contribution to the overall emissions. For example, power plants have a large contribution to the total CO<sub>2</sub> emissions in Rotterdam and therefore the CO<sub>2</sub> emission factor related to this source sector needs to be well-known in order to obtain an accurate emission estimate. Some additional emission factors and tracer ratios are not in the state vector, but they do have an error (i.e. different values are used for the pseudo-observations than in the inversion). This causes a residual mixing ratio that might be attributed incorrectly to other model parameters.

The covariance matrix  $P$ , which describes the error structure of the state vector, is based on the same data as the prior parameter values that define the state vector. Most databases provide a range of values for emission factors from a particular source or fuel type and this is used to estimate the uncertainty. These uncertainty values can also be found in Appendix A of Chapter 4. The true values are always within the uncertainty range.

For this study we selected an arbitrary two-week period in January 2014 (6-20 January). Note that during the summer the importance of source sectors might be different, e.g. there will be less heating from households. Nevertheless, this period is sufficient to test the applicability of our DA system. We update our state vector every day starting from the same prior values and prior uncertainties such that we obtain 14 independent estimates of the state vector. Therefore, the state vector only contains one time step of one day (recall that the footprints also only extend backwards for a short period). In this way we can test the robustness of the parameter estimates and examine whether a short period of observations would be sufficient to constrain the most important emission model parameters. Another option is to create one optimized state vector for the entire period. In that case the state vector is only updated if the newly added observations contain relevant information that reduces the uncertainty. We tested this option too and the conclusions were not different from is presented. Note that by estimating time-average parameters we would be susceptible to temporal aggregation errors (Kaminski et al., 2001). Moreover, we would be unable to constrain the day-to-day variability, which is relevant for optimizing the time profiles.

Optimizing time profiles requires a different method than optimizing the other parameters. The reason is that the average value of the yearly time profile should equal 1.0, as it only affects the variability and not the total yearly emission. For example, a colder year results in more emissions from households (heating), which is accounted for in the calculation of the total emission. The time profile should not change the total emissions, but only attribute most emissions to the coldest periods and less emissions to warmer periods. That means that if the DA system recognizes an increase in heating during a cold period, it lowers the time profile for the remainder of the year (only data that have not yet been optimized) such that the average value stays 1.0 and total emissions are not affected.

#### 5.2.4 Pseudo-observations

The observations used to optimize the emission model parameters are created using the same modelling system as described above. The dynamic emission model is used to create realistic emissions with a high spatiotemporal resolution. Yet, in contrast to the prior, the model parameters are not scaled (scaling factors are 1.0) and the specific Dutch parameter values are used in the emission model. This is also true for parameters that are not included in the state vector. This means that some part of the prior is wrong and is not adjusted in the inversion. These parameters, however, have a small impact on the emissions (Chapter 4).

The resulting emissions are used in combination with the transport calculated by WRF-STILT and OPS model to create pseudo-observations at the locations shown in Fig. 5.2. The simulated time series are illustrated in Fig. 5.1. The urban network consists of seven sites that are scattered over the city of Rotterdam and the port. The rural network consists of

four sites that form an arch around the Rotterdam area. All sites are existing sites in the CO<sub>2</sub> or air quality measurement networks, although not all species used in the inversion are observed at all locations. We only use the daytime (12-16h LT) observations to constrain our emissions. This is normally done to favour well-mixed conditions with better simulated transport and we want to mimic these limitations. We assume all instruments have an inlet at 10m above ground level. In reality this is lower for several sites, but during the well-mixed daytime conditions the difference should be minimal.

The covariance matrix  $R$  describes the observation error. It accounts for errors related to instrumentation, but also representativeness errors due to model transport and interpolation. Since the pseudo-observations are created with the same modelling framework as used for the inversion the model error is non-existent. Also the measurement error is irrelevant. Nevertheless, realistic estimates of these errors need to be included in the OSSE. Moreover, the boundary inflow is subject to errors and some parameters in the emission model are not optimized even though they contain an error. Therefore, the  $R$  matrix is based on the estimated errors and variability caused by these differences and we end up with variances of 8 ppm (CO<sub>2</sub>), 21 ppb (CO), and 2 ppb (NO<sub>x</sub> and SO<sub>2</sub>).

### 5.2.5 Experiments

We perform various experiments to examine the sensitivity of the system to different sources of error. The experiments are discussed here and the detailed set-up of the inversions is summarized in Table 5.2.

1) Network configurations: We start with a comparison of two monitoring networks which differ in the amount of measurement sites and their location (Fig. 5.2). We perform inversions with CO<sub>2</sub> as the only tracer and with the full range of tracers to assess the added value of including co-emitted species for source attribution (inversions Rural-CO<sub>2</sub>, Rural-all, Urban-CO<sub>2</sub> and Urban-all). These tests address the question which type of monitoring sites (urban/semi-urban/rural) are most useful in urban inversions. The results are discussed in Sect. 5.3.1.

2) Time profiles: The second experiment uses different time profiles to describe hourly variations (inversion Temporal). Whereas the pseudo-observations are based on the dynamic profiles developed in Chapter 4, the inversion uses non-dynamic profiles (Denier van der Gon et al., 2011). With this experiment we examine how the DA system deals with large errors in the prior and whether it is able to attribute the resulting residuals to the correct model parameters. The results are discussed in Sect. 5.3.2.

3) Transport model errors: The third experiment is used to examine the impact of erroneous transport (inversion Transport). The pseudo-observations for this experiments are created with OPS as the only transport model, whereas the inversion used STILT as the only observation operator (both models are driven by the same WRF meteorology). The OPS and STILT models differ strongly in their underlying basis and assumptions, but also in

model parameterizations. We have shown previously that the dispersion in WRF-STILT is much faster than in OPS and the model results show large differences even if the same meteorology is used (Super et al., 2017a). The results are discussed in Sect. 5.3.3.

4) Propagation of gained knowledge: The fourth experiment is similar to the Urban-all inversion, except that we propagate knowledge from the previous day (inversions Urban-all\_parameter and Urban-all\_covariance). It addresses the question whether the DA system benefits from the propagation of knowledge in an environment where the observations are complex mixtures of signals from different source sectors. In the Urban-all\_parameter inversion we use the optimized value of a previous day as prior for the next day. In the Urban-all\_covariance inversion we also use the posterior covariance matrix of the previous day as the prior covariance matrix (error propagation) (Sect. 5.3.5).

**Table 5.2: Overview of the inversions: which network and tracers are included, which parameters make up the state vector, and which transport models and time profiles are used. Some settings are different for pseudo-observations (obs) and the inversion (inv).**

Run name	Network	Tracers	State vector	Observation operator	Time profiles
Rural-CO <sub>2</sub>	Rural	CO <sub>2</sub>	EF	STILT-OPS	Dynamic
Rural-all	Rural	all	EF, R <sub>CO</sub> , R <sub>NOx</sub> , R <sub>SO2</sub>	STILT-OPS	Dynamic
Urban-CO <sub>2</sub>	Urban	CO <sub>2</sub>	EF	STILT-OPS	Dynamic
Urban-all (parameter, covariance)	Urban	all	EF, R <sub>CO</sub> , R <sub>NOx</sub> , R <sub>SO2</sub>	STILT-OPS	Dynamic
Temporal	Urban	all	EF, R <sub>CO</sub> , R <sub>NOx</sub> , R <sub>SO2</sub> , T <sub>f</sub>	STILT-OPS	Dynamic (obs), fixed (inv)
Transport	urban	all	EF, R <sub>CO</sub> , R <sub>NOx</sub> , R <sub>SO2</sub>	OPS (obs), STILT (inv)	Dynamic

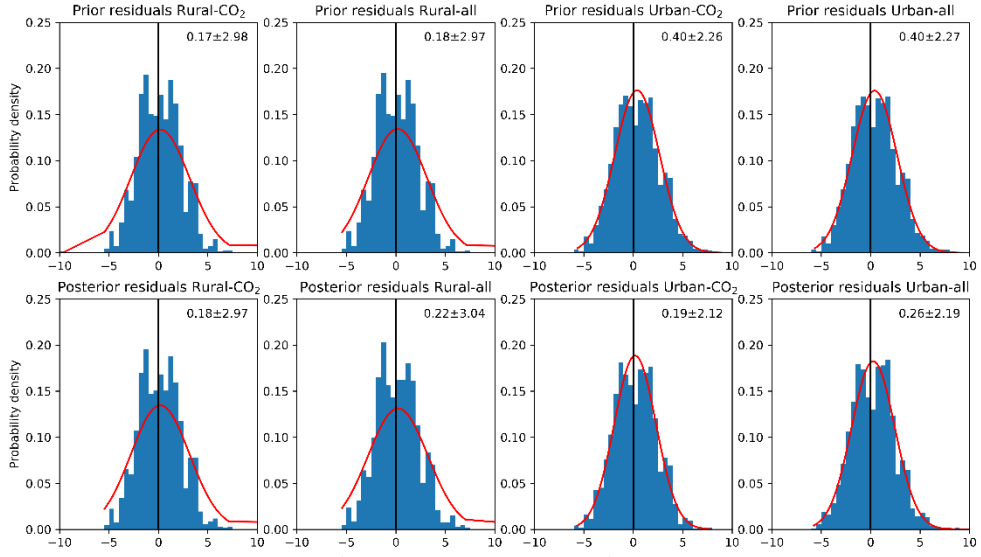
## 5.3 Results

### 5.3.1 Monitoring requirements

#### *Network configurations*

The urban network is more sensitive to urban emissions than the rural network. This follows from the comparison of the prior and posterior residual CO<sub>2</sub> mixing ratios (Fig. 5.4). The prior residuals indicate the difference between the pseudo-observations and the simulated mixing ratios (in ppm) with the prior estimate of the emission factors. The posterior residuals represent the differences between the same pseudo-observations and the simulated mixing ratios with the optimized emission factors. The prior and posterior residuals for Rural-CO<sub>2</sub> and Rural-all are almost similar, indicating that little information is gained during optimization due to the insensitivity of the rural network to urban plumes. In contrast, both simulations with the urban network show a smaller bias in CO<sub>2</sub> mixing ratios after optimization. Nevertheless, the spread remains large.



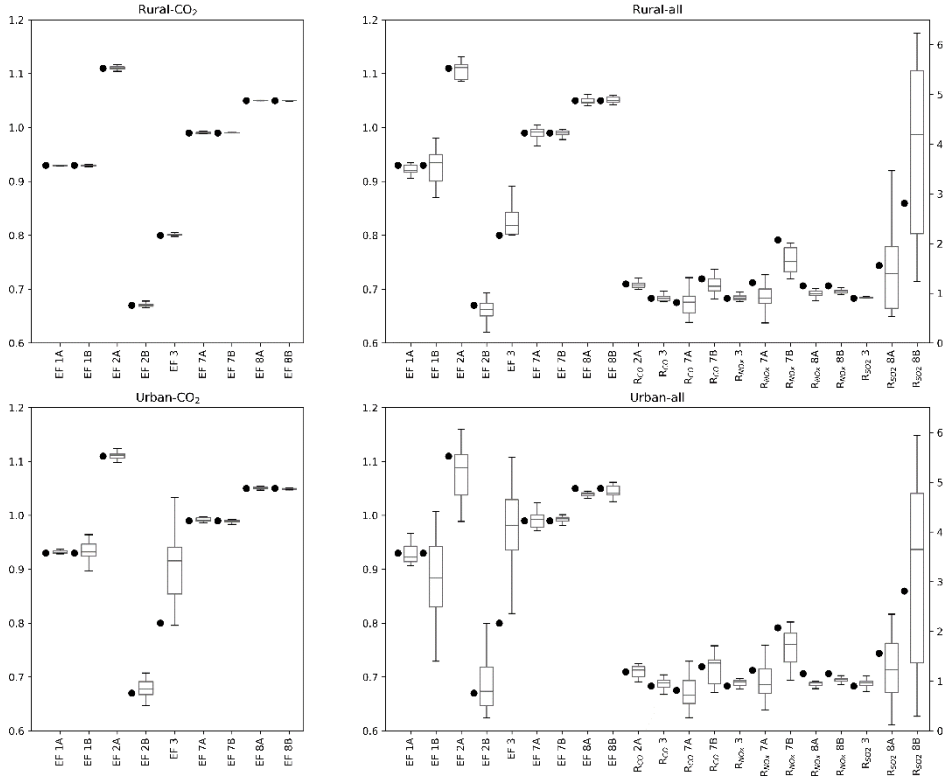


**Figure 5.4: Normalized distributions (probability density functions) of the pseudo-observations – modelled mixing ratios (in ppm) for four different simulations. Given are the bias (deviation from zero) and standard deviation of each distribution. The prior residuals are calculated based on the prior state vector; the posterior residuals are calculated based on the optimized state vector. The runs with the urban network show reduced biases. Note that residuals are included for all observations and not just those assimilated, and thus the agreement can also become worse after optimization.**

A possible explanation for the small differences when using the rural network is the larger distance between the measurement sites and the emission hotspots. This causes the plumes with enhanced tracer mixing ratios to often miss the measurement sites. Or the plumes are so diluted that the enhancement (and therefore also the prior residual) is very small compared to the uncertainty caused by transport and background mixing ratio errors. In both cases, limited information can be gained from the observations (note that the adjustment to the state vector in Eq. 2 scales with the residual term  $y^0 - \mathcal{H}(x^b)$  squared). Indeed, the pseudo-observations show CO<sub>2</sub> signals of more than 8 ppm (the estimated observation error) only about 9% of the selected hours at one of the rural observation sites. There are no cases when two or more rural sites measure significant signals. We therefore conclude that rural sites do not provide useful information on time scales of less than two weeks. In contrast, the urban network picks up strong signals for about 75% of the selected hours, of which 60% affects two or more sites.

The finding that the urban network is more sensitive than the rural network is supported by Fig. 5.5, which shows the variability (spread) in the 14 independently optimized emission factors and tracer ratios. The simulations with the rural network show limited day-to-day variability in the estimated parameter values, which remain close to the prior estimates (dots). The optimized emission factors for the simulations with the urban network show more variability, suggesting that these sites are sensitive to the urban signals. Several emission factors (for example for industry (EF 3), coal-fired power plants

(EF 2B), and households (EF 2A)) differ more from the prior estimates (dots) and display a wide range in optimized values.

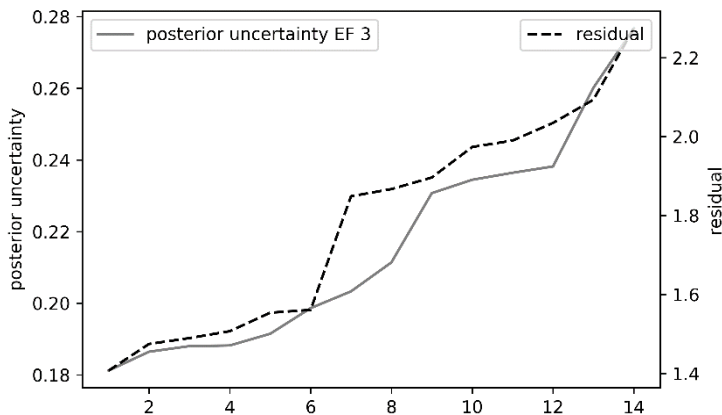


**Figure 5.5: Spread (Q1-Q3) and median values of the parameter scaling factors for the fourteen days included in the inversions. The prior values are indicated by the dots to the left of the box plots. The left y-axis is for the emission factors, the right y-axis for the tracer ratios. The simulations with the urban network show more variability and larger deviations from the prior values (black dots).**

### Co-emitted tracers

The additional co-emitted species also have an impact on the optimization. The Urban-CO<sub>2</sub> inversion attributes the existing residuals mainly to the industrial emission factor (EF 3), because this parameter is highly uncertain (Fig. 5.5). The industrial sector has large CO<sub>2</sub> emissions, so that an update of the emission factor can easily improve the simulated CO<sub>2</sub> mixing ratios. In contrast, in the Urban-all inversion the emission factors of households (2A), glasshouses (2B) and coal-fired power plants (1B) also show large changes in optimized values (Fig. 5.5), because the co-emitted tracers show that part of the residual is unlikely to be caused by the industrial sector. For example, power plants and industrial plants are often located in the same areas and their signals are often mixed. In the absence of co-emitted species the residual is most likely attributed to the industry, which is most uncertain. However, when including the co-emitted tracer CO the DA system can more easily identify the source of the residual. If the industry would be the cause of the

residual there should also be a residual in CO. However, if the power plants are actually the cause of the residual the residual in CO would be limited, as power plants co-emit almost no CO. As such, in the absence of a CO residual the DA system knows that the residual is not caused by the industry, but by power plants. Therefore, we conclude that the co-emitted species contain valuable information about the emission model parameters, despite the uncertainties in the emission ratios.



**Figure 5.6:** Posterior uncertainty ( $1\sigma$  of unitless scaling factor) in the industrial emission factor (EF 3) and absolute mean residual of CO<sub>2</sub> (in ppm) from highest- to lowest-ranked days. The prior uncertainty in EF 3 is 0.4.

Next we take a closer look at the optimized parameter values. For this purpose, we focus on the Urban-all inversion. We exclude the emission factors for road traffic and shipping, because these are relatively certain and they barely affect the total CO<sub>2</sub> emissions. The 14 daily independent parameter estimates are ranked based on their posterior uncertainty and the remaining residual to find the most trustworthy parameter estimates. This ranking is done per parameter, so the best estimate of different parameters can be related to different days. The increase in residual (same for all parameters) and posterior uncertainty (of the industrial emission factor) is shown in Fig. 5.6. Both show that the 3-5 highest ranked days have similar characteristics. After that the reliability decreases, suggesting that only the best few days are fit to make emission estimates. The reason is that on other days the atmospheric signals from that particular source sector are too small, or even absent, and are not informative to update the parameters related to that source sector. A similar pattern is found for the other parameters (not shown), with the exception of the glasshouse emission factor (EF 2B) which shows a large increase in posterior uncertainty as of the third ranked day. Therefore, we focus on the average of the three (two for EF 2B) most likely values.

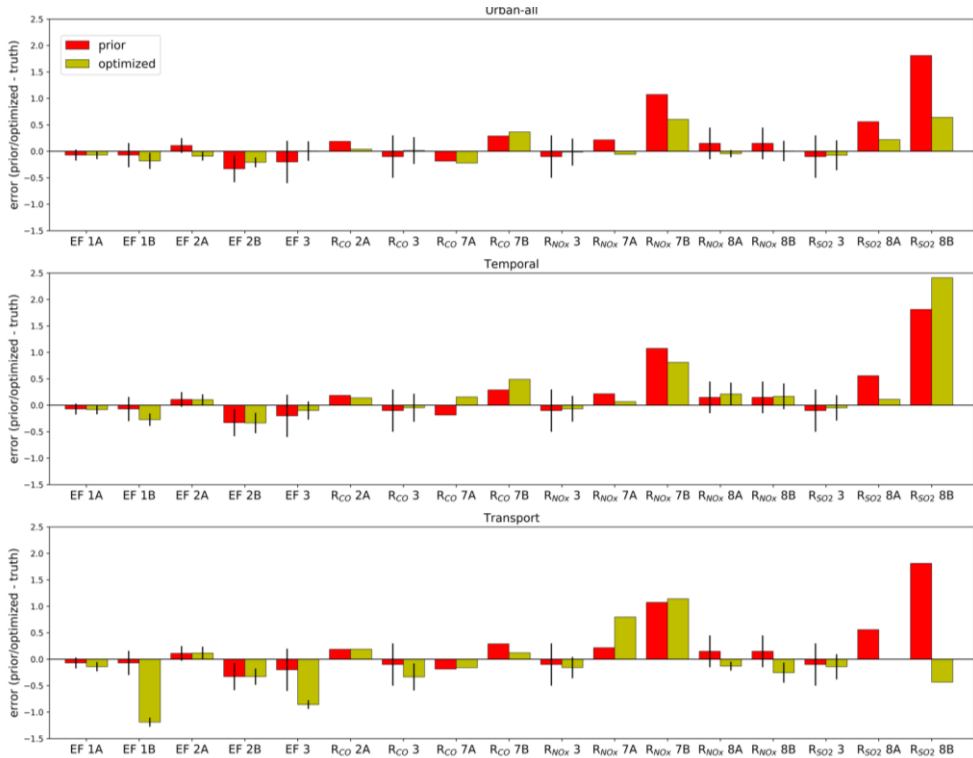
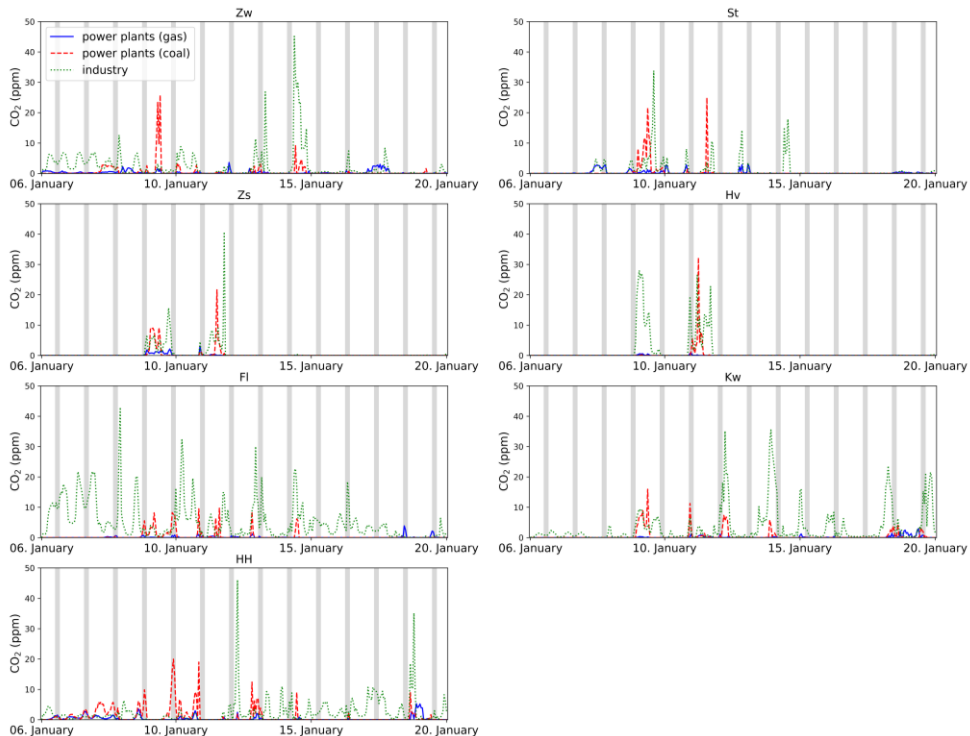


Figure 5.7: Error in the prior or optimized scaling factors (top-3 average) compared to the truth (which is always 1.0) for a range of model parameters for several runs. Positive (negative) values indicate an overestimation (underestimation) of the true values as the truth has been subtracted. The ranking of the optimized values is done based on the posterior uncertainty and posterior residuals. The uncertainties of parameters with a lognormal distribution are not shown due to their large range. These uncertainties are only slightly reduced.

For most parameters the top-3 average optimized value from the Urban-all simulation is closer to the true value compared to the prior estimate (Fig. 5.7). This is true for all tracer ratios, except for  $R_{CO}$  of road traffic (cars and HDV). The optimized emission factors for households, glasshouses and industry are also closer to the truth. The other emission factors, related to power plants, are not well-constrained. If we check the pseudo-observation time series of  $CO_2$  mixing ratios resulting from power plants and industry we find that there are no distinct peaks from gas-fired power plants (Fig. 5.8). There are only few gas-fired power plants in the area with high stacks and their emissions are likely to miss the observation sites. Therefore, the emission factor for this subsector is barely updated. In contrast, coal-fired power plants do generate large signals, although these mostly occur during night time and these data are excluded from the inversion. Nevertheless, some daytime peaks exist and the corresponding optimized emission factor is quite different from the prior. However, coal-fired power plant emissions are mostly mixed with industrial signals, posing a major challenge to the inversion system in

attributing the residuals to the correct sector. This is even made more difficult by the lack of a distinct emission ratio for either of these sectors.



**Figure 5.8:** Pseudo-observation time series of CO<sub>2</sub> mixing ratios at the urban observation sites resulting from three source sectors: gas-fired power plants, coal-fired power plants and industry. Shaded areas indicate daytime hours used in the inversion.

We can also use the top-3 averaged parameter values to calculate the total CO<sub>2</sub> emissions and the emissions of the largest source sectors (Table 5.3). Road traffic and shipping are not included (similar for Fig. 5.7), since their contribution to the total CO<sub>2</sub> emissions is minor. The prior underestimates the emissions with 11% ( $-3.97 \text{ Tg yr}^{-1}$ ), indicating that a reasonable first estimate of the total emissions can be made based on general emission factors that are applied globally, rather than on the Rijnmond-specific numbers used in the truth. This sets quite a challenging benchmark for our observation-based assimilation method to improve on. The Rural-CO<sub>2</sub> run estimates the emission to be more or less similar to the prior, which is low relative to the truth. However, the other runs have higher (less negative deviations) estimates that are closer to the truth. The best estimate for total CO<sub>2</sub> is obtained with the Urban-CO<sub>2</sub> simulation with an underestimation of only 5% ( $-1.71 \text{ Tg yr}^{-1}$ ). However, when we look at specific source sectors we find that the Urban-all simulation improves the estimated emissions of glasshouses and industry. In contrast, the coal-fired power plant emissions deteriorate while being the second largest source of CO<sub>2</sub>, resulting in a stronger underestimation of the total emissions. The gas-fired

power plant emissions are barely affected by the optimizations. This comparison shows that the inclusion of co-emitted species significantly affects the estimated emissions per source sector, but that this does not necessarily lead to a better total CO<sub>2</sub> emission estimate.

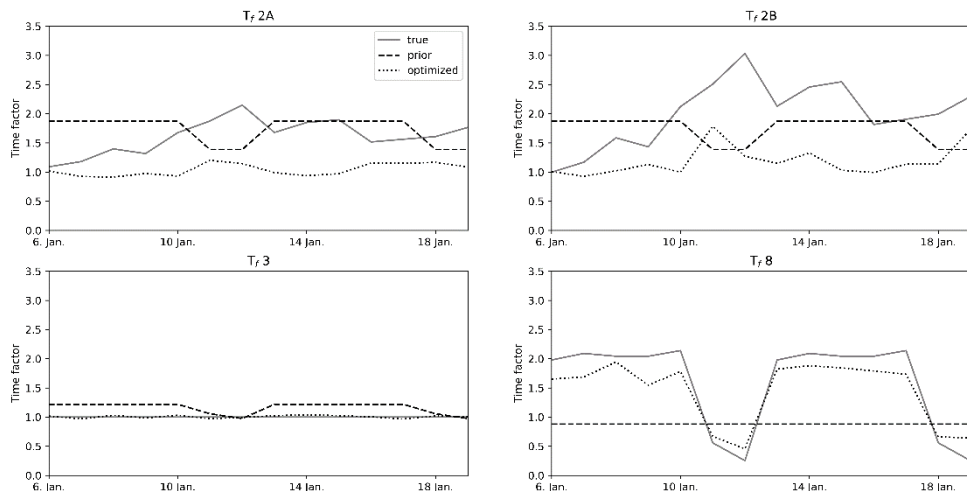
**Table 5.3: Reported yearly CO<sub>2</sub> emission for Rijnmond (in Tg yr<sup>-1</sup>) and deviations from the truth based on the prior and top-3 optimized parameter values for all runs. Bold (underlined) values signify the best (worst) emission estimates per category.**

Run	Total	Power plants (gas)	Power plants (coal)	Households	Glasshouses	Industry
True	34.4	3.60	10.6	2.29	1.31	14.2
Prior	-3.97	-0.25	-0.74	+0.25	-0.43	-2.84
Rural-CO <sub>2</sub>	-3.96	-0.25	-0.74	+0.26	-0.44	-2.82
Rural-all	-3.26	-0.29	<b>-0.53</b>	+0.25	<u>-0.45</u>	-2.29
Urban-CO <sub>2</sub>	<b>-1.71</b>	-0.24	-5.19	<b>+0.20</b>	-0.44	-0.74
Urban-all	-2.54	-0.25	-1.91	-0.21	-0.27	<b>+0.08</b>
Temporal	-4.71	-0.29	-2.88	+0.24	-0.44	-1.39
Transport	<u>-25.4</u>	<u>-0.51</u>	<u>-12.6</u>	+0.27	-0.43	<u>-12.2</u>
Urban-all parameter	-3.27	-0.12	-2.52	<u>-0.34</u>	-0.25	<b>-0.08</b>
Urban-all covariance	-6.19	<b>-0.05</b>	-4.07	-0.29	<b>-0.05</b>	-1.91

### 5.3.2 Time profiles

In the previous sections we showed that our system is able to constrain a range of model parameters. In this section we take it a step further by introducing an additional error that is typically present in these type of inversions, namely uncertain time profiles (Temporal simulation). The time profiles determine which fraction of emissions for a sector occur at which hour of the day, and on which day of the year. Like the emission factors and emission ratios, the time profiles act linearly on the emissions per sector. Yet, unlike the emission factors and emission ratios the time profiles are constrained to be on average 1.0 over a year and thus only modify the temporal structure of the solution.

The optimized time profiles are in better agreement with the true profiles than the priors (Fig. 5.9). We show here daily values, because the time profiles are optimized on a daily basis similar to the other parameters. Especially the day-to-day variations are captured well, but the time profiles of households (T<sub>f</sub> 2A) and glasshouses (T<sub>f</sub> 2B) have a strong negative bias, so that the emissions are not correctly spread over the year. This has an effect on the optimized emission factors as well. For example, it seems that the overestimation of household emissions is attributed to the time profiles instead of the emission factor, causing the emission factor to remain too large (Fig. 5.7). This hypothesis is supported by the presence of strong negative correlations between the time profiles and emission factors of both sectors, which indicates that the system had difficulties separating the two from atmospheric observations (see Sect. 5.3.4 for more details on correlations).



**Figure 5.9: Different daily time profiles for the simulated period based on the selected hours (12–16h LT). The true profiles are dynamic, the prior profiles are fixed, and the optimized profiles are the result of the inversion. The optimized profiles follow the day-to-day variations of the dynamic profiles much better than the priors.**

Additionally, the uncertainty in the hourly variations is much larger than the uncertainty in the yearly average emission factors. Therefore, a system that seeks to minimize uncertainties for this sector gains most from updating the time profiles. This is especially true for the shipping emissions, which have a relatively certain emission factor. The shipping time profile is therefore very well-constrained ( $T_r$  8). The only exception is the industry, which has a very uncertain emission factor. We find that this emission factor is better constrained than the ones for households and glasshouses as residuals are more easily attributed to the emission factor. The underestimation of the emission factors leads to a too low estimate of the total  $\text{CO}_2$  emissions of  $29.7 \text{ Tg yr}^{-1}$ .

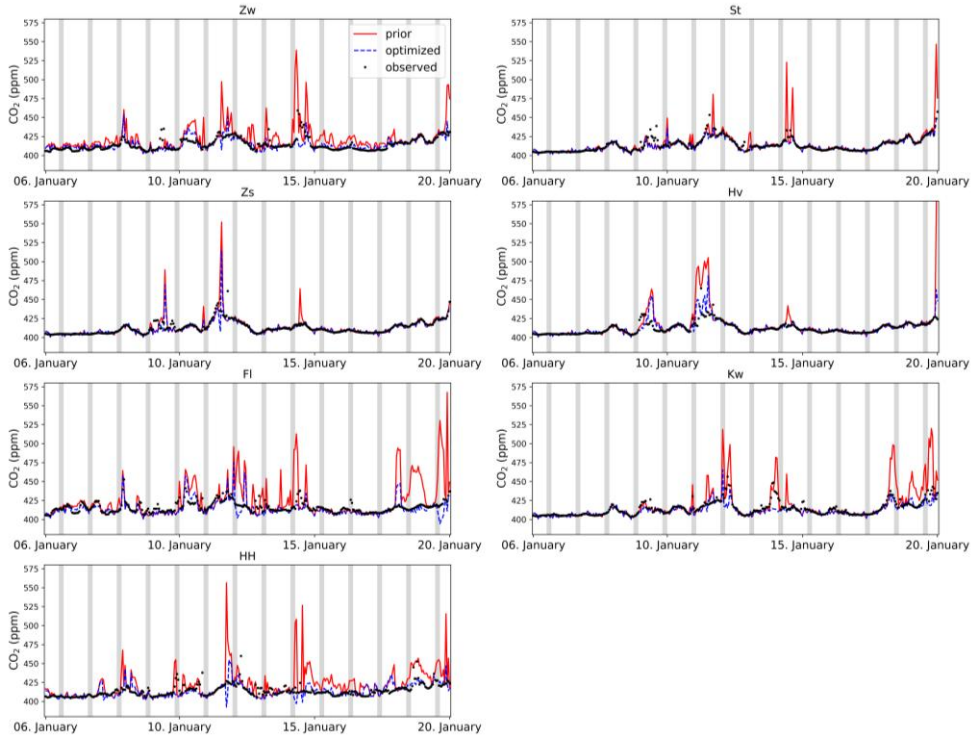
These results show that for some sectors it is possible to obtain time profiles from an optimization in the absence of activity data to correctly represent day-to-day variations. Such time profiles are needed to correctly optimize the emission factors and ratios, while without time profiles the inversion suffers from a structural model error that gets aliased into the other model parameters.

### 5.3.3 Transport model errors

In the previous simulations the model transport was assumed to be exactly known, as the inversion used the same models that were used to create the pseudo-observations. In the case of real observations the transport adds an additional uncertainty to the inversion. We examine the impact of erroneous transport by using different models for the inversion and the generation of the pseudo-observations. The pseudo-observations are the same as in previous simulations, but the inversion only uses STILT. Therefore, there is only a difference in the transport and dispersion of point source emissions that were performed with OPS to create pseudo-observations. Treating these point sources with STILT causes

high peak mixing ratios (Fig. 5.10), because with STILT we assume all emissions take place at ground level whereas OPS accounts for stack heights and plume rise. We thus consider the results in the context of erroneous representation of point source emissions.

We find that the difference in transport leads to much larger prior residuals of  $-4.78 \pm 16.51$  ppm (not shown) compared to the Urban-all simulation. The optimized mixing ratios are generally lower than the prior mixing ratios and are in much better agreement with the pseudo-observations (posterior residuals are  $0.67 \pm 7.06$  ppm). However, sometimes the posterior mixing ratios suddenly show a sharp decline whereas the observations and prior show a large peak (such as on 11-12 January). This indicates the presence of negative emissions in the posterior estimate (i.e. uptake of  $\text{CO}_2$ ). Indeed, the emission factor of coal-fired power plants becomes negative (smaller than  $-1.0$ , as  $1.0$  is the truth) (Fig. 5.7). So we find that an overestimation of atmospheric mixing ratios due to erroneous representation of stack emissions in the transport model results in a strong decrease in emission factors. Moreover, the source sectors that are not simulated with the plume model (households and glasshouses) also show deteriorating results, which is a compensating effect for the large changes made to the other sectors. On average, the erroneous transport leads to a large underestimation of the total emissions of  $25.4 \text{ Tg yr}^{-1}$ .



**Figure 5.10:** Time series of prior (full) and posterior (dashed) simulated  $\text{CO}_2$  mixing ratios and the pseudo-observations (dots) for the Transport simulation at all seven urban sites. Shaded areas indicate daytime hours used in the inversion.



The implications of these results are important, as transport model errors are currently the main challenge for urban inversions with real observations. Not only does an erroneous description of model transport have a huge effect on the optimized emissions, it also results in large posterior uncertainties. The comparison shown here is an extreme case, resulting in obviously incorrect and untrustworthy emissions. But in reality it might be difficult to assess the effect of transport errors on the results.

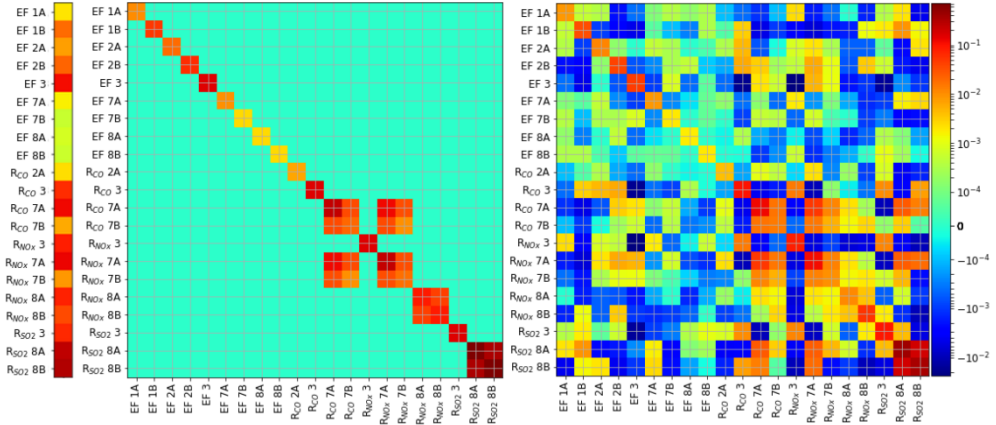
### 5.3.4 Error covariance structure

In order to better understand previous findings we have a closer look at the covariance structure of the Urban-all simulation (Fig. 5.11). The prior covariance matrix (left panel) mainly shows diagonal values, which indicate the uncertainty in a parameter. Some parameter errors are also correlated (we will call this parameter correlation), indicated by off-diagonal values. A positive (negative) parameter correlation means that a positive deviation in one parameter has to be compensated by a positive (negative) deviation in the other parameter. This affects the cost of tuning a parameter and thus provides additional information to the inversion. Additionally, the presence of a parameter correlation between two variables means that they are difficult to differentiate from an atmospheric signal. In other words, if A and B are correlated a  $\text{CO}_2$  residual can be reduced by tuning A or B, or both. The inversion system is unable to identify with which magnitude to correct either of the parameters. Unless the variables also affect another observed quantity in a different way, for example the mixing ratio of a co-emitted tracer.

The prior covariance matrix (Fig. 5.11, middle panel) shows that we assumed positive parameter correlations between tracer ratios from the same sector (road traffic: cars (7A) and HDV (7B), shipping: ocean (8A) and inland (8B)), but also between the  $R_{\text{CO}}$  and  $R_{\text{NO}_x}$  of road traffic. The reason is that the emission ratios are mainly determined by technological implementations that we think affect the subsectors equally. Moreover, we also assume the technology to impact the emissions of CO and  $\text{NO}_x$  equally. In contrast,  $\text{SO}_2$  emissions are usually reduced by desulphurization methods which is different from the techniques to reduce CO and  $\text{NO}_x$  emissions. Therefore, we assume  $R_{\text{SO}_2}$  not to be correlated with  $R_{\text{CO}}$  and  $R_{\text{NO}_x}$ . We indeed see that the prior errors show the same sign for the correlated parameters and the optimized values of correlated parameters are both reduced or increased. The only exception is the  $R_{\text{CO}}$  for road traffic, which is underestimated for cars and overestimated for HDV. Our inversion is able to recognize the discrepancy between the  $R_{\text{CO}}$  from cars and HDV as it decreases the one while increasing the other, unlike suggested by the positive correlation. However, the corrections are in the wrong direction.

Equally interesting is the posterior covariance matrix. The diagonal values in the prior covariance matrix are larger than in the posterior covariance matrix (Fig. 5.11, left panel), indicating that the uncertainty of the parameters is reduced. In contrast to the prior covariance matrix, the posterior covariance matrix (Fig. 5.11, right panel) displays a significant number of off-diagonal values. Note that this is an average matrix for all

fourteen days and that variations exist between the days. Therefore, the parameter correlations in this graph are not averaged out and are consistently present. Now the question is whether the presence of parameter correlations is beneficial (extra information) or harmful (difficult to discern variables) for the inversion.



**Figure 5.11: Prior (middle panel) and 14-day average posterior (right panel) covariance matrix  $P$  from the run Urban-all. On the diagonal the variances (uncertainties in a variable) are shown, whereas off-diagonal values indicate covariances or correlations between different variables. The left panel shows the difference between the prior and posterior diagonal values. All differences are positive, indicating an uncertainty reduction.**

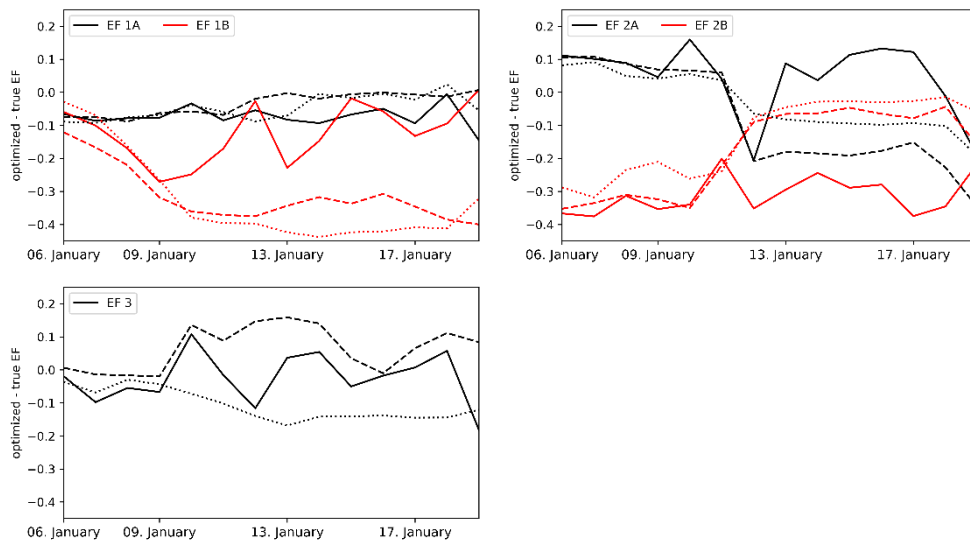
If we compare the posterior uncertainties of a parameter for the individual days against the absolute sum of the correlations this parameter has with the other parameters for the same days, then we find a positive relationship for all parameters with correlation coefficients between 0.23 and 0.89 (on average 0.59). So days with a high posterior uncertainty of a parameter correspond to days where the error of that parameter also correlates strongly with errors of other parameters. A weak correlation means that a parameter is less sensitive to the presence of error correlations with other variables, therefore we will call this the sensitivity correlation. The weakest sensitivity correlations are related to the  $R_{NOx}$ ,  $R_{SO2}$  and emission factor of the industrial sector. These parameters show little variability during the fourteen days and an average over the entire period gives a robust estimate of the true parameter values. The parameters with the strongest sensitivity correlations are  $R_{CO}$  of households and road traffic (HDV). These parameters show large fluctuations during the fourteen days and the 2-week average values tend to be dominated by a few outliers. If a few days can be selected on which the parameter correlations are weak (i.e. the atmospheric signal clearly contains information about this specific parameter), the top-3 estimate displayed in Fig. 5.7 can still give a good estimate of the parameter value, as is true for the household emission ratio. However, for road traffic (HDV) the parameter correlations always seem to have an effect and this parameter is less well-constrained.

A few parameters with a particularly large amount of strong parameter correlations are  $R_{CO}$  from industry (3),  $R_{NOx}$  from cars (7A) and  $R_{SO2}$  from ocean shipping (8A). Although

these three emission ratios are improved during optimization, the remaining uncertainties are large. The optimized  $R_{CO}$  of industry is 3.3 ppb/ppm with an uncertainty of about 25%, the optimized  $R_{NOx}$  of road traffic (cars) is 1.9 ppb/ppm with an uncertainty of about 28% and the optimized  $R_{SO2}$  of shipping (ocean) is 3.3 ppb/ppm with an uncertainty of about 65%. These results suggest that the presence of parameter correlations in the posterior covariance matrix makes the optimization procedure more difficult for that specific parameter, due to the inability to isolate the effect of this parameter on the atmospheric observations. This results in poorly constrained parameters and/or large posterior uncertainties. However, this is only true when parameters are sensitive to parameter correlations (high sensitivity correlation). Why some parameters are more sensitive to the presence of parameter correlations than others needs to be investigated. One hypothesis is that the presence of parameter correlations is less problematic when the parameter can be discerned by another signal, for example from a co-emitted tracer.

### 5.3.5 Propagation of gained knowledge

In Rijnmond some source sectors are strongly clustered. For example, glasshouse emissions take place mainly to the northwest of Rotterdam. Therefore, this sector can only be constrained by the urban network when the wind blows from the northwest and affects the measurements sites. As a consequence, each day of measurements can only constrain a limited number of source sectors, while improved knowledge on one of the source sectors would also affect the other source sectors (as seen in the previous section).



**Figure 5.12:** Time series of the optimized – true scaling factors for the simulations Urban-all (full line), Urban-all\_parameter (dashed line) and Urban-all\_covariance (dotted line).

In previous inversions we have made fourteen independent estimates of the model parameters, starting each day with the same prior estimates and uncertainties. A

downside of this method is that knowledge that is gained during previous days is not propagated to the next day. Including such propagation scheme means that each day starts with an updated prior estimate, such that for example a new estimate of the industrial emissions can be constrained by better knowledge about power plant emissions. Normally, propagation schemes also use the covariance matrix of the previous time step. This would imply that after some time the model parameters become more or less fixed, while in reality some parameters might change over longer periods of time. We therefore perform two new simulations which propagate information from the previous day and compare them. The first one only propagates the optimized parameters, but starts each day with the same covariance matrix  $P^b$  as shown in Fig. 5.11 (Urban-all\_parameter). The second one propagates both the optimized parameters and the covariance matrix (Urban-all\_covariance). Otherwise, the simulations are similar to the Urban-all run.

In general, the more information is propagated the less variability we find from one day to the next (Fig. 5.12). The regular Urban-all inversion (full lines) shows large variability for those parameters that are regularly updated with new information, such as the industrial emission factor (EF 3). The industrial emission factor is likely to be different for specific sub-sectors, so that it matters whether a plume comes from a refinery or from the food industry, resulting in large variations. In the Urban-all\_parameter inversion the variability is slightly reduced by starting each day with a different prior. For example, the large reduction in the household emission factor (EF 2A) on day 7, which is visible in all inversions, is maintained during the following days. The Urban-all\_covariance inversion shows even smoother time series due to the reduced freedom of the system to change the parameter values after information has been added (reduced uncertainties).

For parameters that occasionally affect the observations both the propagation schemes have a clear advantage. Fig. 5.12, which displays the difference between the optimized and true emission factors, shows that the emission factor of gas-fired power plants (EF 1A) grows steadily towards the correct value with both propagation schemes. The emission factor of glasshouses (EF 2B) is well-constrained on day 6-7 (11-12 January) and this value remains relatively stable during the rest of the period. As such, the estimated emissions of these subsectors are the best for the inversions with propagation. Also the industrial emission factor (EF 3) is doing reasonably well when propagating the posterior parameter values, fluctuating around the truth during the two weeks. Yet when the covariance matrix is also propagated this emission factor estimate becomes too low. The Urban-all and Urban-all\_parameter simulations have a similar error in the industrial emission, although reversed in sign. However, the Urban-all\_covariance simulation strongly underestimates the industrial emission.

For both inversions the emission factor of households (EF 2A) and coal-fired power plants (EF 1B) deteriorate during the two-week period. EF 1B decreases during the first few days, moving further from the truth. As mentioned before, the signals of coal-fired power plants are often mixed with those of the industry (Fig. 5.8). The posterior

covariance matrix in Fig. 5.11 also shows a relatively strong correlation between the emission factors of these two source sectors. The decrease in the coal-fired power plant emission factor is thus likely the result of the inability of the inversion system to discern this parameter from the industrial emission factor.

The household emission factor shows a sudden decline on day 7 (12 January), related to a shift in the wind direction from west to south-east. The atmosphere apparently contains information about household emissions during a south-easterly flow. Indeed, the city of Rotterdam is within the footprint of several observation sites on this day. We see the same happening on days 13 and 14 (18-19 January) when the wind again shifts to the south-east and the emission factor of households is constrained. Unfortunately, the emission factor is overcorrected, possibly due to the overestimation of the road traffic and shipping emissions that are also in the footprint. This results in an underestimation of the household emissions.

## 5.4 Discussion and conclusion

The aim of this study was to examine how well our DA system is able to quantify urban CO<sub>2</sub> emissions per source sector. Since the prior consists of a dynamic fossil fuel emission model the model parameters are optimized rather than the emissions themselves. We find that most of our inversions improve the total yearly CO<sub>2</sub> flux for the Netherlands compared to the prior. Note that, given the improvement in emission ratios, also the emissions of CO, NO<sub>x</sub> and SO<sub>2</sub> are improved in these inversions. However, the disaggregation between source sectors remains challenging. This is partially due to the fact that some parameters have a similar effect on the observations (their errors are correlated) and are therefore inseparable. We also identified some additional challenges for urban scale inversions, such as the definition of the temporal distribution and simulating urban scale transport. Introducing these uncertainties challenges the DA system to find a reasonable emission estimate and to separate between source sectors. Nevertheless, we have shown that optimizing time profiles seems viable for some categories in case dynamic profiles are not available.

Based on previous work we expected the rural network to be the best option to constrain the total emissions (Super et al., 2017a). However, here we find that the urban network gives the best emission estimate, which could be due to the relatively short simulation period in which the rural sites are barely affected by the urban emissions. We do expect the rural network to be of larger value once we consider emissions to come from multiple cities in a region. The urban network shows the largest sensitivity to the different source sectors and is therefore the preferred network for source attribution. The coal-fired power plant emissions are nevertheless better constrained by the rural network. Yet, we find no configuration with which we can optimize all the source sectors and the total CO<sub>2</sub> emission correctly. For example, the inversion with propagation of both

optimized parameter values and covariances (Urban-all\_covariance) gives the best results for gas-fired power plants and glasshouses, but performs poorly for coal-fired power plants. Therefore, more research is needed to establish the best DA configuration, possibly with an extended simulation period.

The signals we work with in our DA system are driven mainly by differences in fossil fuel emissions. For this purpose we subtract a background mixing ratio from our observations. Fossil fuel signals during daytime are typically on the order of 1-10 ppm. Residuals ( $y^0 - \mathcal{H}(x^b)$ ) are only a fraction of this, and are often close to, or smaller than, the prescribed model-data mismatch ( $R$ ). Given the linear relationship between the residuals and the parameters, illustrated in Eq. 2, a small residual means that the leverage on the parameters is relatively small. The strength of this relationship depends on  $K$  (given in Eq. 4) and could be increased by decreasing the observation error ( $R$ ), increasing the prior parameter error ( $p^b$ ), or increasing the footprint sensitivity ( $H$ ). However, larger signals, and therefore larger residuals, exist during night time and the early morning (Fig. 5.8 and Fig. 5.10). These data could provide more information about the parameters, but they are currently not included in the inversions. We only selected afternoon observations, as is often done in high-resolution atmospheric inversions, to exclude the impact of atmospheric transport errors related to stable boundary layers. Therefore, the inversion could benefit strongly from an improved description of stable boundary layers so that the large night time enhancements can be used to constrain the fossil fuel fluxes.

The largest discrepancies between the optimized and true emission model parameters are found when we use different transport models for the pseudo-observations and for the inversion. Transport errors can either be random or systematic. Random errors, such as errors in the wind direction, are unlikely to affect the optimized emissions much when averaged over a longer time period and domain. This was shown by Deng et al. (2017), who found little variation in the average  $\text{CO}_2$  emission for Indianapolis using different configurations of WRF to calculate the transport. They do find an impact on the spatial distribution of the emissions. This becomes important when optimizing a specific source sector that is clustered in one place, such as the glasshouses. We found that the glasshouse sector is only correctly optimized with a specific wind direction. If the modelled wind direction is wrong the residuals would thus not be attributed to the glasshouse sector as it is not in the modelled footprint of the measurement site.

Systematic errors, whether in the modelled transport or in the observations, are more difficult to solve as they do not cancel out when simulating a longer period and this can lead to biased emission estimates (Meirink et al., 2008; Su et al., 2011). An example of a bias in modelled transport is the systematic error in pollutant dispersion under stable conditions, which has often lead previous inversion studies to only select well-mixed afternoon hours (Boon et al., 2016; Bréon et al., 2015; Lauvaux et al., 2013). Such data selection also leads to a bias in the estimated emissions when the diurnal cycle is not accounted for. In this study the systematic error is introduced by emitting all point source

emissions at surface level instead of accounting for the stack height. This even resulted in a negative emission factor for coal-fired power plants.

A process that becomes important when using real observations is chemistry, since  $\text{SO}_2$  and  $\text{NO}_x$  are chemically reactive with a lifetime of several hours to days. CO is also chemically reactive (lifetime of several months in mid-latitude winter), but at the scale considered here this is relatively unimportant. Chemical removal of these tracers affects the observed tracer ratios with  $\text{CO}_2$  and thus complicates source attribution. This is especially true for the rural network since the transport time is longer and the chemical removal has advanced more. Therefore, the modelling framework should consider chemical removal, either by implementing chemistry or by estimating the removal based on travel time and a decay function. Nevertheless, an additional uncertainty is introduced.

Additionally, a complicating factor in interpreting observed  $\text{CO}_2$  mixing ratios is the presence of a background and a biogenic signal, while we are only interested in that part of the mixing ratio that is caused by fossil fuel emissions. An incorrect definition of the background causes a large bias in the optimized emissions (Göckede et al., 2010). Here, we ignore the uncertainty in the background, except in the definition of the covariance matrix  $R$ , and attribute all tracer residuals to the fossil fuel emissions. Yet, there are several other methods to deal with the non-fossil fuel related  $\text{CO}_2$  signals. First, the uncertain background can be added to the state vector and be optimized in the inversion. For example, He et al. (2017) have shown that high-altitude aircraft observations are suitable to improve regional biosphere flux estimates by correcting the bias in boundary conditions. Second, a mixing ratio gradient over the area of interest can be calculated using an upwind and downwind site such that the boundary inflow plays no role anymore (Turnbull et al., 2015). This method was shown to reduce the impact of boundary inflow, but only when the wind direction is more or less perpendicular to the gradient (Bréon et al., 2015; Staufer et al., 2016). Therefore, this method limits the amount of useful measurements. Finally, the fossil fuel signal can be separated from other  $\text{CO}_2$  sources using the radiocarbon isotope ( $^{14}\text{CO}_2$ ). Although observations of carbon isotopes are expensive and currently not widely available, previous studies have shown promising results using  $\Delta^{14}\text{C}$  (the ratio of  $^{14}\text{CO}_2$  to  $^{12}\text{CO}_2$ ) to simulate fossil fuel  $\text{CO}_2$  records and in inverse modelling studies (Basu et al., 2016; Bozhinova et al., 2014; Levin and Karstens, 2007; Miller et al., 2012; Turnbull et al., 2006; Turnbull et al., 2015; Wang et al., 2018). Our urban network detects average fossil fuel  $\text{CO}_2$  signals of about 5 ppm with peaks up to 50 ppm. This would result in  $\Delta^{14}\text{C}$  signals of around 13 up to 130 per mille, which are certainly detectable with current techniques. Besides  $\Delta^{14}\text{C}$  other isotope signatures and tracers can also provide additional information. For example,  $^{13}\text{CO}_2$  and  $\text{O}_2/\text{N}_2$  can give insight in the dominant sources and sinks or fuel types (Lopez et al., 2013; Van der Laan et al., 2014) and as such also be an indicator for the transition from fossil fuels to biofuels.

Finally, we compared two optimization methods with a different propagation scheme. We have shown that this affects the results significantly and therefore the choice of

propagation scheme should be made carefully. As mentioned before, the propagation of error statistics has the advantage that observations with little information about a specific parameter will have little impact when the uncertainty is already reduced by previous days. However, this method does not allow us to track changes in the parameter values over a longer period of time while the uncertainties are becoming too small. Developing methods to properly inflate these errors, or even the creation of an error model to aid this propagation, would be an interesting follow-up to our work. Yet, we expect changes in emission factors and ratios not to occur at short time scales and therefore applying error propagation to two-weekly inversions could be a possibility, starting the following two-weekly inversion with the original covariance matrix. The alternative method, which consists of only propagating the optimized parameter value and not the uncertainty, does give the DA system the freedom to identify temporal variability. However, the individual estimates are not very robust for some parameters, thus leaving us with a large uncertainty.

To conclude, inverse modelling at the urban scale is feasible when the observations contain a lot of information about the different source sectors. However, additional tracers are an important addition to the inversion framework in order to discern the information belonging to specific source sectors. Moreover, we argue that a dynamic emission model has some major advantages over regular emission maps, allowing us to constrain physically relevant parameters even in the absence of good prior information. Nevertheless, quite some challenges remain. Transport modelling at this small scale needs to be improved to be able to use real urban observations, as under current conditions the transport error strongly dominates the results. Moreover, adding additional tracers and isotopes can add more information to the system to constrain the emissions (for example for coal-fired power plants) and to improve the definition of the fossil fuel signal. For the future, additional advances need to be made to include satellite observations in the inverse modelling framework. The advantage of satellite data is that it covers data-sparse regions and with a larger view it can differentiate between the urban dome with high pollution levels and the cleaner rural areas. Nevertheless, currently the uncertainties in satellite retrievals are too large to significantly improve emission estimates for short time windows in addition to in-situ measurements.





A thick, dark vertical bar runs down the left side of the page. To its right, several thin, curved lines in dark and light grey originate from the bottom left and sweep upwards and to the right, creating an abstract, organic shape.

## Chapter 6

### **General discussion on monitoring urban fossil fuel emissions per source sector**



## 6.1 Introduction

There are significant efforts needed to reduce the amount of CO<sub>2</sub> emitted to our atmosphere from fossil fuel combustion at a wide range of scales. There are continental and global agreements (e.g. Paris Agreement, EU 2020 Climate & Energy Package), local efforts (e.g. emission reduction by better waste separation in Arnhem, Netherlands (Gemeente Arnhem, 2014)) and industry-wide efforts (e.g. emission reductions in the Dutch energy-intensive industry (VEMW, 2017)). International agreements, such as the Paris agreement, play an important role in creating a solid base for coordinated action. In such agreements emission reduction targets are set, but often there is no strict guideline on how to reach those targets. Each member state thus implements policies targeting specific source sectors, but often there is a lack of knowledge on the efficiency of such policies. As such emission monitoring, reporting and verification (MRV) is an important aspect of emission reduction policies in order to monitor the progress made and to exchange best practices.

In this thesis several methods for emission monitoring are explored. Emissions are reported at a national level and often emission reductions are guided by policies at the (inter)national level. However, most emissions take place in urban areas and therefore the effect of emission reductions will certainly become visible here. As such, there are two main research strategies that we could have followed. The first option is to focus on large spatial scales and make emission estimates at regional to national scale using rural measurement sites or even satellite observations. A supporting argument for this option is that this would support implementation and verification of the Paris agreement. An example of such an approach is illustrated by Van der Laan et al. (2010), who used long-term observations of CO<sub>2</sub>, CO, radon and radiocarbon at one urban site to constrain fossil fuel emissions of the Netherlands. The second option is to focus on emission hotspots, i.e. urban areas, and use (semi-)urban measurement sites. An important reason to choose this option is that urban areas contribute a large fraction to the global fossil fuel CO<sub>2</sub> emissions (Duren and Miller, 2012) and therefore the largest emission reductions can be achieved in urban areas. Emission reductions in high-emitting areas cause large changes in atmospheric CO<sub>2</sub> signals in the surrounding area and that makes urban areas suitable for emission monitoring. Moreover, emission monitoring following exact country borders is difficult, while monitoring dominant emission hotspots can mostly be done without considering country borders and still constrains most of a countries emissions.

We choose to take the second strategy for two reasons. Firstly, because of the availability of unique in-situ measurements within a complex urban-industrial area. Secondly, we anticipated the possibility to concurrently monitor emissions of air pollutants to support source attribution. Air pollution is a major issue in urban areas and therefore the development of an urban monitoring system could potentially support the improvement of urban air quality as well. Previous scientific efforts to quantify emissions

and emission trends in urban areas have encountered several limitations, such as large errors in calculated atmospheric transport in the complex urban environment and difficulties isolating the fossil fuel signal from the observed concentrations (Boon et al., 2016; Bréon et al., 2015; Lac et al., 2013). Thus there is plenty of room for improvement and development of new methods, as described in this thesis.

The work discussed in this thesis aims to contribute to the current understanding of atmospheric monitoring systems at urban scales and enhance the amount of information that can be gained from these systems. This research can therefore be used to improve our knowledge of the urban carbon cycle, to set up effective monitoring frameworks, and to gain insight in the contribution of specific source sectors to support policy making. In a broader sense, better knowledge of the fossil fuel emissions and the trends therein can help to better understand the entire carbon cycle and how it is going to change in the future. For example, in inversions focusing on terrestrial fluxes the fossil fuel fluxes are often assumed to be relatively well-known and their uncertainty is included in the observation error (Broquet et al., 2013; Peters et al., 2010). This allows less of the residual to be attributed to terrestrial fluxes. Therefore, we argue that a better estimate of the fossil fuel CO<sub>2</sub> emissions also benefits the optimization of terrestrial fluxes.

In this chapter we discuss four themes (or sub-goals) that are important for atmospheric monitoring of urban fossil fuel CO<sub>2</sub> emissions and that cover all the work described in this thesis. The foundation of an atmospheric monitoring framework consists of an observation network and atmospheric transport models. The first theme therefore covers the value of different types of observation sites (Chapter 2 and 3) and the (dis)advantages of several observation network configurations (Chapter 5) in estimating urban emissions. The second theme revolves around the ability of different types of transport models to represent observed atmospheric concentrations (Chapter 3). The third theme is the quantification of urban fossil fuel CO<sub>2</sub> emissions. Within this theme several methods are discussed that can be used to quantify emissions, including a simple mass-balance approach (Chapter 2), the dynamic fossil fuel emission model (Chapter 4), and the advanced method of inverse modelling (Chapter 5). The final theme considers whether information can be gained by adding co-emitted species to the system and whether this information can help attribute CO<sub>2</sub> emissions to specific source sectors (Chapters 2-5). Together these themes help to understand the complexity of the urban emission landscape and what is needed to quantify it.

## **6.2 Monitoring fossil fuel emissions through atmospheric observations**

A main feature of an independent monitoring system is the presence of an observation network. In this thesis we have used several observation network set-ups that differ in the number of sites and their location to explore their use for monitoring urban fossil fuel emissions (RQ. 1). In Chapter 2 we have used an upwind-downwind gradient with only two

measurement sites close to the city border (semi-urban) and have shown that this is a promising set-up to estimate the overall fluxes of the intermediate area. In Chapter 3 we have compared sites at increasing distance (in-city to ~200 km) from the urban area. Here, we found that the further away from the city we measure, the smaller the observed signal. The smaller signal is more difficult to isolate from all other signals that are mixed in and the urban signal can easily be missed by a rural site depending on the wind direction. A remote site, however, does observe an integrated signal and is therefore useful for calculating the total flux from the city. In contrast, in-city sites are well-exposed to urban fossil fuel fluxes and can be used to separate between different source areas, such as the residential and industrial area in the case of Rotterdam. This is in line with previous findings (Boon et al., 2016; Bréon et al., 2015; Kort et al., 2013; Lac et al., 2013). Chapter 5 confirms these findings by giving a better constraint on source sector emissions with an urban network. The rural network shows to be relatively insensitive to urban emissions because the observed signals are comparable to the uncertainty caused by model errors. It should be noted that the sites in the rural network are further removed from the city than the downwind site used in Chapter 2.

An important notion from the above summary, which discretely runs like a thread through all the chapters, is the notion of scale. Processes affecting the urban carbon budget work at different spatial and temporal scales. The smaller the scale under consideration the more complex processes exist that play a role. For example, the complex atmospheric transport in a city affects the urban sites needed to constrain emissions from individual source sectors. However, in order to interpret the observed signals at remote sites (used to constrain the total urban emission) the effect of urban atmospheric transport can be neglected. Therefore, the purpose of the study determines the amount of detail that needs to be resolved.

Following this notion, we argue that the best monitoring strategy (type of instrument, location, period covered) is case dependent. In a data-sparse region where little is known about the emissions, a first general idea is needed of the overall CO<sub>2</sub> emissions. In this thesis continuous in-situ observations of CO<sub>2</sub> are used to estimate average fossil fuel fluxes (Chapter 2). We find that two monitoring sites for CO<sub>2</sub>, located up- and downwind of the city at about 15 km from the city border (semi-urban), are convenient to estimate the total emissions. With this method there is no need to account for stack emissions, which are already well-mixed by the time they reach the measurement sites (Chapters 2 and 3). In addition, a meteorological monitoring station is needed to provide information about at least the wind direction and wind speed, so that the most suitable data can be selected. Additional information, such as boundary layer height, are useful as well. Putting sites further away from the city (rural, more than ~50 km) is less helpful as the signals become small compared to the background noise. In this way we can improve our understanding of the carbon budget for the region with relatively few resources using the methodology described in Chapter 2. Note that in many cities an air quality monitoring

network is present that can also be used as basis to monitor CO<sub>2</sub> emissions. At existing sites the facilities are present to install an (additional) instrument and air quality measurements can be used in addition to CO<sub>2</sub> observations to gain more knowledge about dominant source sectors (see Sect. 6.5.5).

In a region where reliable emission data are available and the focus is on monitoring emission trends resulting from policy implementations, more detail is required to separate between source areas. Here, we recommend the use of an urban network monitoring CO<sub>2</sub> and co-emitted species (see Sect. 6.5.5). The hourly data used in this thesis provide a lot of detailed information on the short-term variability in emissions (Chapter 3). We have shown that a few weeks of data is sufficient to constrain the CO<sub>2</sub> emission landscape if a dynamic emission model is used as prior and a sufficiently large urban monitoring network is present. The optimized emission factors can be used to calculate the yearly total emissions from activity data. It should be noted, however, that emission ratios are variable and a few weeks of data is insufficient to correctly constrain the yearly emissions of co-emitted species. Moreover, a longer time series (at least several years) is needed to monitor changes in the emission landscape resulting from emission reduction policies. The size of the monitoring network depends on the size and complexity of the city and can be tested using an inversion experiment. Previous research has shown that a larger network of sensors can improve source attribution in the Paris metropolitan area (Wu et al., 2016), decreasing the uncertainty in optimized emissions by 18-33% for individual source sector when going from 10 to 70 measurement sites. In line with this we recommend extending the network to monitor the major stack emissions in detail since the optimization of power plant emissions remains a challenge. Relatively cheap instruments can be placed in the stacks. Although these instruments are less accurate, the additional information gained outweighs the reduced accuracy (Turner et al., 2016). Besides, concentrations in the exhaust are extremely large and a minor error in the observations becomes irrelevant.

Of course also practical limitations exist that determine the size and shape of the network. For example, one of our CO<sub>2</sub> measurement sites (Zweth) is not at an ideal location, situated in a small open area surrounded by high trees. The inlet height nearly reaches the height of the trees. Therefore, the trees affect the transport of pollutants by the wind, but they also cause a significant biogenic flux. This makes the interpretation of the observations more difficult. Although we have shown that this observation site is useful to constrain urban emissions due to its location relatively close to the city, a higher inlet or an open area could benefit our monitoring framework.

A major challenge in interpreting in-situ observations of atmospheric CO<sub>2</sub> concentrations is isolating the anthropogenic signal from the large advective flows of 'background' air. Throughout this thesis we have used different background definitions, which all have their pros and cons. In Chapter 2 we used an upwind-downwind gradient, where the upwind site is considered to give the background concentration. This method is only applicable for a specific wind direction. A small deviation from the exact perpendicular wind direction

already causes a large uncertainty in the background concentration. Besides, selecting for a specific wind direction greatly reduces the amount of observations that can be used. Similar limitations were found in previous studies (Bréon et al., 2015; Staufer et al., 2016). In Chapter 3 the background concentration is based on a smoothed function of the observed time series. Although this is a relatively easy method without need for extra observation sites it can have a large bias when above average clean or polluted air is advected because these relatively short-term features are not captured. The uncertainty in this method is also illustrated by the difference compared to other curve fitting methods shown by Pickers and Manning (2015). Finally, in Chapter 5 we use CO<sub>2</sub> concentrations from a large-scale model which gives us a grid box average value over a large area. Although the latter two methods give similar mean fossil fuel signals, the last method has more variability. Based on the difference between these methods we estimate the uncertainty in the background concentration (and thus in the remaining fossil fuel CO<sub>2</sub> signal) to be about 3 ppm. The background concentration is thus a significant source of uncertainty and each of the described methods has its limitations. Another method to identify the fossil fuel signals is by measuring radiocarbon (<sup>14</sup>C). This carbon isotope is lacking in fossil fuels and therefore fossil fuel burning results in a decrease in radiocarbon. A downside is that radiocarbon measurements are mostly very expensive, non-continuous and due to the need for laboratory analyses they are not available in real-time. Nevertheless, this method has been used in previous studies which have shown promising results (Basu et al., 2016; Bozhinova et al., 2014; Fischer et al., 2017; Levin and Karstens, 2007; Miller et al., 2012; Turnbull et al., 2006).

Although in this thesis we used in-situ observation of atmospheric concentrations there are other potentially interesting measurement techniques that could be of added value for urban monitoring studies. One example is aircraft flights, which have previously been used in combination with the mass-balance approach to estimate fluxes of methane and CO<sub>2</sub> from a specified area (Caulton et al., 2014; Karion et al., 2013; Mays et al., 2009; Peischl et al., 2015). With this method a vertical profile of concentrations can be made, which might be useful to detect plumes from elevated stacks. However, Mays et al. (2009) also report several challenges related to this method, such as the uncertainty in the background concentration and in the measured wind speed. A second alternative is remote sensing, which has the advantage of a large spatial coverage, also in data-sparse regions (Ciais et al., 2014). The applications are currently still limited for CO<sub>2</sub>, because a relatively high accuracy is required to distinguish the fossil fuel enhancement from the large background concentration. Nevertheless, some studies have examined the use of satellite data to constrain urban fluxes and have shown both opportunities and challenges (Silva et al., 2013; Worden et al., 2012). Finally, urban fluxes can be measured directly with eddy-covariance measurements (Kleingeld et al., 2017). The advantage of this method is that the measurements can be related directly to fluxes. However, the footprint of such



measurements is often limited and to constrain fluxes from a large urban area requires a lot of instruments.

### **6.3 Atmospheric transport modelling in an urban-industrial complex**

The interpretation of atmospheric observations is sometimes challenging due to the wide range of processes that affect the atmospheric concentrations. Atmospheric transport models can be used to explain and account for many of these processes, such as transport, mixing, and entrainment. Also biospheric fluxes can be included. The method of combining atmospheric observations and atmospheric transport models to constrain emissions is called inverse modelling and makes optimal use of both sources of information. The inversion method relies heavily on a correct representation of atmospheric transport such that differences between observed and modelled concentrations can be attributed to errors in the emissions. Therefore, we have examined the capability of several types of atmospheric transport models to represent the atmospheric transport and the importance of atmospheric transport for inverse modelling (RQ. 2 and 4). Also in this context the notion of scale is important.

Previous CO<sub>2</sub> inverse modelling studies were mainly focused on estimating large-scale (grid spacing of 50 to several hundred kilometres) weekly to monthly mean ecosystem fluxes (Basu et al., 2016; Broquet et al., 2013; He et al., 2017; Liu and Bowman, 2016; Meesters et al., 2012; Ray et al., 2014; Rödenbeck et al., 2009; Schuh et al., 2010; Tolk et al., 2011). The observations used in these studies capture biospheric signals from a large domain dominated by daily to weekly variations. Therefore, there was less need to resolve hourly and kilometre-scale transport and fluxes. Most of these studies used an Eulerian transport model. In Eulerian regional transport models each grid cell contains well-mixed air such that no spatial variability exists below the grid resolution. Relevant processes that occur at sub-grid scale therefore must be parameterized.

In contrast, urban fossil fuel fluxes vary at sub-hourly and street canyon scale (<100m), thus causing large variability in observed urban CO<sub>2</sub> concentrations (hourly averaged) that needs to be resolved. At these high resolutions the sub-grid processes in Eulerian models (such as turbulence) become important and cause significant errors in atmospheric transport (Lauvaux et al., 2016). Moreover, the complexity of build-up areas has an impact on the wind fields, turbulence, and radiative balance (Heusinkveld et al., 2014; Kent et al., 2018; Wicht et al., 2017), which in turn affect urban atmospheric transport and mixing. The model parameterizations of these processes in regional atmospheric transport models is not sufficient to correctly resolve small-scale transport. A special Urban Canopy Model (UCM) is available in WRF-Chem that describes the characteristics of an urban area, but we found no significant improvement using UCM. Chapter 3 shows that our Eulerian model WRF-Chem can reproduce the daytime average observed mixing ratios reasonably well. However, large errors occur in the simulated wind fields that cause incorrect

footprints. Consequently, hourly transport modelled with WRF-Chem ( $1 \times 1 \text{ km}^2$ ) is prone to large errors. We have shown in Chapter 5 that erroneous atmospheric transport limits the ability of the inversion system to correctly quantify the urban fossil fuel emissions. As such, the incorrect representation of atmospheric transport is an important limitation of our monitoring system, as was also suggested by previous studies (Boon et al., 2016; Deng et al., 2017; Vogel et al., 2013).

Additionally, due to the discrete nature of Eulerian models the presence of strong sources causes large errors in modelled concentrations at measurement sites close to these sources (Karamchandani et al., 2011; Tolk et al., 2009). This is especially true for point sources which have a small horizontal extent. We tried to overcome this problem by implementing a Lagrangian model, called the OPS model (Chapter 3). A Lagrangian model that follows particles through space and time is not limited by grid size, but uses sub-grid parameterizations similar to Eulerian models to represent turbulence at scales not resolved by the meteorological driver data. This indeed shows better results for the dispersion of point source emissions. An additional advantage is that the model can be driven by observed meteorological conditions. But also even more complex models are developed (such as LES) that can be used specifically in build-up areas and account for the impact of buildings and street orientation on the atmospheric transport (e.g. (Dezzutti et al., 2018; Tolias et al., 2018)). However, more detailed models require more input data which, if uncertain, add additional uncertainty to the results. For example, the Gaussian OPS model requires exact stack heights for point source emissions, which have a large impact on the modelled concentrations (30-50% change in mixing ratios when the WRF-Chem characteristics are used). If stack heights are unknown this would thus cause an error. Therefore, the added value of more complex models should always be weighed against the uncertainty caused by the additional input. In a data-rich area like Rotterdam we found most input data to be available and the increased complexity of the models not to be much of an issue.

Clearly, the choice of atmospheric transport model should be based on the purpose (and scale) of the study and be related to the observation network. To correctly represent urban observations a Lagrangian model that resolves small-scale features and dilution of point source emissions is needed. However, the difference between the Eulerian and Gaussian transport of point source emissions becomes negligible after about 15 km from the source area. As such, an estimate of the total urban emissions can be made with several semi-urban sites and an Eulerian model. The advantage of adding an atmospheric transport model is that the short-term variations in atmospheric processes can be accounted for compared to the case where only observations are used (such as in Chapter 2).

## 6.4 Quantification of urban fossil fuel CO<sub>2</sub> emissions

In the previous sections we have discussed the optimal monitoring strategies identified in this thesis, based on observations and models. In this section we discuss their ability to quantify urban fossil fuel CO<sub>2</sub> emissions (RQ. 3 and RQ. 4).

In Chapter 2 the urban fossil fuel CO<sub>2</sub> flux is estimated using only two observation sites close to the city border, providing a gradient in the CO<sub>2</sub> mixing ratio over the city. These gradients can be used to make a first estimate of the emissions in the city as long as the upwind site is representative for the background air. As we make long-term average flux estimates we only need to correctly represent the average transport processes. Although this mass-balance approach is very simplified, it shows reasonable flux estimates with errors of 15-23% compared to the emission inventory. This method even allows us to separate between three source areas while only accounting for the observed wind direction, except in the presence of elevated stack emissions. Nevertheless, the estimates can be influenced by biogenic fluxes and by atmospheric processes that are not taken into account (such as entrainment). Moreover, the flux estimates are made using a strict data selection based on well-mixed daytime conditions. Although such simplifications and selections are often done to minimize the model errors (Boon et al., 2016; Bréon et al., 2015; Lauvaux et al., 2013), the resulting emission estimates are not necessarily representative for hours outside the selection. Indeed, if we would include observations from all hours of the day in the flux estimate of Metropolitan Rotterdam (Table 2.2) the resulting hourly emission estimate would be about half as large (i.e. night time emissions are smaller). Besides this bias we find a large variability in emission estimates using individual concentration gradients, which can be 30% up to more than 100% of the average emission estimate for a specific source area. Similarly, Mays et al. (2009) found large short-term variability in emissions, which makes it difficult to extrapolate the results and calculate a yearly total emission based on a few concentration gradients.

In Chapter 4 we use a completely different method to estimate emissions, without using atmospheric observations. Instead, we explored the potential of several auxiliary data streams to predict high-resolution emissions based on general emission factors and activity data that can be used in the absence of local data. These data (consisting of emissions factors, emission ratios and activity data) are combined in a dynamic fossil fuel emission model that estimates emissions based on knowledge about the emission landscape. However, some data have a large uncertainty, especially emission factors of CO<sub>2</sub> and emission ratios of co-emitted species with CO<sub>2</sub>. This results in a significant uncertainty of about 15% in the overall emissions in the Netherlands, which is comparable to the uncertainty in the mass-balance approach. However, more specific knowledge about the region, such as which cleaning technologies are implemented in industrial stacks (e.g. desulphurization techniques), can reduce the uncertainties. Indeed, the uncertainty reported by the Dutch Pollutant Release and Transfer Register is only 1%

based on local emission factors and activity data. However, we argue that the reported uncertainty is likely underestimated when errors in the accounting methods are considered. Moreover, the emission model predicts emissions per source sector and therefore gives more detail than the mass-balance approach. Our analysis of uncertainties in the dynamic emission model also shows which parameters are most important and most uncertain. This knowledge can help policy makers to tackle those sectors and characteristics where most emission reduction can be gained. Moreover, the most uncertain parameters can be improved by inventory builders and inverse modellers to create a better emission estimate using new data streams and atmospheric constraints.

Finally, in Chapter 5 we use the dynamic fossil fuel emission model as prior for the inversion system. Different experiments are done which make use of different observation networks, time profiles or transport models. We optimized the emission factors of CO<sub>2</sub>, emission ratios and time profiles. We find that we can significantly reduce the uncertainty in the prior yearly emission estimate when using the urban network and only two weeks of CO<sub>2</sub> observations, leading to an error of only 5%. For comparison, Basu et al. (2016) were able to constrain the yearly US emission to within 1% using almost two years of CO<sub>2</sub> and <sup>14</sup>CO<sub>2</sub> pseudo-observations. Looking at the individual parameters of the emission model when including all tracers we find that most emission ratios and the emission factors of households, glasshouses and industry are indeed better constrained while the emission factors of power plants drift away from the truth.

In the inversions in Chapter 5 we increased the complexity compared to Chapter 2 by looking at shorter time scales, which requires resolving short-term transport processes as well. Due to the large transport errors found in Chapter 3 we decided to first use pseudo-observations instead of real observations in the inversions. By using the same atmospheric transport in the inversion that was used to create the pseudo-observations, the impact of erroneous transport is removed from the inversion. This allows us to examine the strengths and weaknesses of our system when not dominated by transport errors. Nevertheless, we do acknowledge the large impact of erroneous transport of inversions with real observations, as was shown before. In this thesis, however, we rather focus on the development of our inversion system and improving the prior definition of the emissions. This is, like reducing transport errors, a necessary step in the development of our capacities to constrain urban emissions.

Comparing these three methods (mass-balance, dynamic emission model, inverse modelling) we conclude that the optimization of a dynamic fossil fuel emission model (inverse modelling) is the preferred method to calculate fossil fuel emissions. The dynamic fossil fuel emission model gives details about the different source sectors, which is relevant in the context of emission reduction policies. However, in itself it contains large uncertainties when general emission factors and emission ratios are used, thus signifying the importance of optimizing the model parameters. Although the uncertainties in the different methods compared to the truth are difficult to determine, the dynamic emission

model has several advantages over the regular emission maps. Firstly, the dynamic emission model has the advantage that the optimized results have a physical meaning and can be compared to other cities, for example to verify whether the household emissions per degree day fall within the range of similarly developed cities in other countries. Secondly, the optimized emission ratios can also be used to improve the emission estimates of co-emitted tracers. Thirdly, the emission model can be built for any area using local information, whereas normal emission inventories cover large domains and are based on long-term average data and/or characteristics averaged over that domain. The potential of using local data is huge with emerging high-resolution data streams, such as real-time traffic and shipping monitoring using on-board computers and transponders or real-time energy consumption using smart meters. Unlocking the potential of such data could be a very nice step forward in dynamic emission modelling. Finally, the optimized parameters show little variability in the short-term. Therefore, the results can be more reliably extended to night time conditions, resulting in a smaller bias than the mass-balance approach given that the time profiles are relatively well-defined.

### **6.5 Source attribution using co-emitted species**

Previously, we mainly discussed the quantification of the total fossil fuel CO<sub>2</sub> emissions per area. However, the dynamic fossil fuel emission model estimates the emissions per source sector, which can be used to determine how much emission reduction can be gained from policies targeting a specific sector. Yet, all the emitted CO<sub>2</sub> looks exactly the same in our observations and it is impossible to separate CO<sub>2</sub> from road traffic and residential heating without further information. In this thesis we examined the use of co-emitted species to gain such information (RQ. 1).

Up to now most studies only used CO as an additional tracer (or in combination with <sup>14</sup>CO<sub>2</sub>) and the variability in the overall emission ratio has often shown to be a complicating factor in the use of CO for identifying fossil fuel CO<sub>2</sub> emissions (Levin and Karstens, 2007; Lopez et al., 2013; Turnbull et al., 2006; Turnbull et al., 2015; Van der Laan et al., 2010; Vogel et al., 2010). Moreover, this method requires subtracting a background concentration from the observed CO and CO<sub>2</sub> concentration as we are only interested in the ratio from the fossil signals (denoted  $\Delta\text{CO}:\Delta\text{CO}_2$ ). As discussed previously the definition of the background concentration introduces a large uncertainty which makes it difficult to get the exact value of the observed concentration ratios.

In Chapters 2 and 3 we also start with CO as co-emitted species that can potentially attribute a CO<sub>2</sub> signal to industrial or residential source areas. Large point sources, i.e. power plants and industry, emit relatively little CO compared to road traffic and residential heating. Therefore, a higher  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio indicates that the footprint of the measurement covers a residential and traffic-dominated area. We find that CO is indeed helpful to interpret the CO<sub>2</sub> signals when the two source areas are physically separated. If

this is not the case the  $\Delta\text{CO}:\Delta\text{CO}_2$  ratio would be more difficult to interpret. Moreover, the higher than expected  $\Delta\text{CO}:\Delta\text{CO}_2$  concentration ratios help us understand that the elevated stack emissions frequently pass by or over the measurement sites (Chapter 2). We also used this ratio to select cases where the industrial/power plant signal is dominant to compare the performance of the Eulerian and Gaussian models in representing the dispersion of stack emissions (Chapter 3).

In Chapters 4 and 5 we also included other co-emitted species ( $\text{NO}_x$  and  $\text{SO}_2$ ) and they prove useful to attribute  $\text{CO}_2$  emissions to different source sectors in the inverse modelling framework. Whereas the  $\text{CO}$  signal is often dominated by road traffic and can be used in our study to separate between industrial and residential sectors,  $\text{NO}_x$  and  $\text{SO}_2$  give constraints on the other source sectors as well. Therefore, to distinguish different source sectors there is less need for these sectors to be physically separated, but only if a source sector has a specific signature of emission ratios. For example, the industrial and coal-fired power plant emissions are often co-located and lack distinct signatures. This makes it difficult to separate them from atmospheric observations. Adding an additional tracer, such as  $\text{PM}_{2.5}$ , could solve this issue.

Without co-emitted species the inverse modelling framework is unable to attribute the model-data mismatch to the correct source sector (Chapter 5), although a good estimate of the overall  $\text{CO}_2$  emissions can also be made without the additional tracers. Moreover, our inversion framework can optimize the emission ratios per source sector, therewith also improving the emission estimates of co-emitted tracers. Therefore, we argue that co-emitted species are an important addition to an urban monitoring framework. Which co-emitted species are useful depends on the study area: the dominant source sectors, dominant fuel types, and applied filtering technologies. A first idea of relevant co-emitted species can be gained from an existing urban air quality monitoring network. With wind roses of the co-emitted species concentrations in Rotterdam we were able to identify which species peak when the wind blows from a specific source area.

## 6.6 Outlook and recommendations for future research

There are many ongoing projects that focus on quantification of urban greenhouse gas emissions and gaining more detail from co-emitted species or isotopes. Besides large research projects several (PhD) students work on such topics, amongst others at the Wageningen University and the University of Groningen. For example, research is done on the use of oxygen as a tracer for fossil fuel  $\text{CO}_2$  emissions. When fuels are burned  $\text{CO}_2$  is released and  $\text{O}_2$  is consumed according to an oxidative ratio that is specific for each fuel type, which makes  $\text{O}_2$  potentially interesting to track dominant fuel types. A first effort to include  $\text{O}_2$  in a regional atmospheric transport model has shown promising results, but further improvements are needed especially related to the selected oxidative ratios (Kuijpers, 2018).

An isotope that also merits attention is radiocarbon ( $^{14}\text{CO}_2$ ). For example, the RINGO (Readiness of ICOS for Necessities of Integrated Global Observations) project examines the ability of the Integrated Carbon Observation System (ICOS) measurement network to detect fossil fuel emissions by including  $^{14}\text{CO}_2$  measurements (ICOS, 2017). Other carbon isotopes ( $^{13}\text{CO}_2$ ) or oxygen isotopes ( $^{18}\text{O}$ ) might also provide valuable information about fuel or source types or about other sources and sinks of  $\text{CO}_2$  (Chen et al., 2017; Lopez et al., 2013; Popa et al., 2014). Yet, the use of isotopes requires further research to improve the definition of discrimination rates and interfering sources (such as nuclear facilities), which is needed to improve the model results. Additionally, the measurement uncertainty of some isotopes is relatively large, hampering the usefulness of isotope observations in inverse modelling studies.

Another project that follows up on our work is the H2020 project VERIFY that aims to provide independent verification of greenhouse emissions using observations in an inverse modelling framework (European Commission, 2018), starting from the dynamic emission model developed and described in this thesis. The inversions will also be used to examine the effectiveness of emission reduction policies and track whether emission reduction targets are met. This project aims to use real observations in an inverse framework with multiple tracers and isotopes, marking another step towards true observation-based monitoring of emission reductions. However, we argue that the representation of small-scale transport needs to be improved to be able to get reasonable emission estimates with inverse modelling when using real observations. Moreover, we think it is important to start monitoring urban emissions as soon as possible and not to wait until policies are implemented. This allows us to establish a baseline with which to compare the future emission landscape resulting from emission reduction policies.

What can be noticed from the listed initiatives is that they only consider greenhouse gas emissions. Co-emitted trace gases are merely used to improve the estimate of greenhouse gas emissions through attribution. Yet, the reason that trace gases can be used to monitor greenhouse gas emissions is that they come from the same sources. Therefore, emission reduction policies targeting greenhouse gases are likely to also affect the emissions of those co-emitted trace gases. Since some of these gases are important air pollutants (like  $\text{NO}_x$ ) it seems sensible to monitor changes in emissions of air pollutants simultaneously to avoid unintended adverse consequences. An additional advantage is that air pollution is more tangible and therefore higher on many political agendas. Moreover, some air pollutants also affect the climate or the uptake of  $\text{CO}_2$  by vegetation, such as ozone (Super et al., 2015). We therefore recommend to consider greenhouse gases and air pollutants from fossil fuel combustion in one system, not just for source attribution but also to monitor the implications of emission reduction policies on both air quality and climate.

Additionally, advances are made in earth observation to monitor a wide range of air pollutants and greenhouse gases at increasingly higher resolution, such as the recently launched TROPOMI instrument (Hu et al., 2018). The advantage of satellite observations is

that they cover a vast area, also where observations are currently scarce or even absent. Therefore, satellites are a useful tool for making emission estimates using a wide range of trace gases, such as CO (Dekker et al., 2017), or to provide a constraint on background concentrations. Also CO<sub>2</sub> can be measured from space. Yet, despite the large CO<sub>2</sub> concentration gradients at the surface due to the presence of fossil fuel emissions, the column-averaged mole fraction observed by a satellite instrument is relatively small. Therefore, only very large emission reductions at emission hotspots cause signals large enough to be detectable from space with the current limitations in accuracy and resolution (Liu et al., 2017). Nevertheless, satellites can be useful to constrain parts of the carbon cycle (biogenic, fossil fuel) for larger regions, as was previously shown (Fischer et al., 2017; Reuter et al., 2014; Silva et al., 2013).

For a PhD following up on my work a nice starting point would be improving the dynamic emission model. For example, an improved description of the industrial emissions is needed, including several sub-sectors with their own characteristic emission factors and ratios. Currently, the clustering of all industrial emissions is one of the major causes of uncertainty. Moreover, we currently use a simplified spatial distribution. But with emerging real-time data streams the description of the spatial and temporal distribution of emissions can be improved. Also the inversion system can be improved by making the system suitable for multiple species (possibly also isotopes), including the effects of chemistry. If we want to move towards an operational inversion system using real observations chemistry becomes an important factor of uncertainty. The inversion system could also be expanded with other types of observations, such as direct flux measurements or satellite data. These observations have their own advantages and together they are likely to better constrain the different aspects of the urban emission landscape. Finally, a major source of uncertainty is the atmospheric transport. Although this should not be the main focus for a PhD working on urban inversions, we believe that improving the description of the night time stable boundary layer height could benefit the inversions by allowing to use night time data with large concentration enhancements. This could be reached by using observed <sup>222</sup>Radon enhancements to scale the fossil fuel fluxes (Van der Laan et al., 2010; Van der Laan et al., 2014). With these adjustments some major steps are made towards an urban inversion using real measurements, which is the ultimate aim.





## Summary

### Background

Fossil fuel combustion causes an increase in atmospheric carbon dioxide (CO<sub>2</sub>) levels and is one of the major causes of climate change. Therefore, efforts are made to reduce CO<sub>2</sub> emissions from fossil fuel combustion through (inter)national agreements, with the most famous example being the Paris agreement. Each member state that ratified the agreement has to aim for pre-set emission reduction targets. In this collaborative effort it is important to keep track of the progress made towards these targets, but also to gain insight in which emission reduction policies are most effective to support future decision-making. Therefore, scholars have started developing atmospheric monitoring techniques, mainly focused on urban areas. Since about 70% of the anthropogenic CO<sub>2</sub> emissions takes place in urban areas, the largest emission reductions will take place here. This causes large atmospheric signals that are relatively easy to measure. However, scholars have faced some major challenges. For example, the transport within a built-up area is complex, making the interpretation of atmospheric observations difficult. Moreover, emission reduction policies often target specific source sectors (such as road traffic or industry). Hence, these sectors should be monitored separately to understand the effectiveness of individual measures. This source attribution is impossible with only CO<sub>2</sub> observations when source sectors are not spatially isolated.

### Aim

The overall aim of this thesis is to improve our understanding of the monitoring requirements to constrain urban fossil fuel CO<sub>2</sub> emissions per source sector. A key feature of a monitoring system is a network of observation sites. Therefore, the first research objective is to identify the most useful monitoring sites and network configurations. Besides CO<sub>2</sub> we also included measurements of trace gasses that are co-emitted with CO<sub>2</sub> during fossil fuel combustion. This happens in a ratio that is specific for a source sector and therefore these tracers have the potential to identify the source of a CO<sub>2</sub> signal. We examined this opportunity to use co-emitted species to attribute CO<sub>2</sub> signals to specific source sectors. Besides observations a good model representation of atmospheric transport is needed to interpret the observations. Therefore, the second research objective is to better understand the possibilities and limitations of atmospheric transport models in reproducing observed mixing ratios within/close to a city and find a useful modelling approach. The third objective is to predict high-resolution emissions in an urban area using proxy data and to gain insight in the uncertainties related to these emissions. Finally, we combine our insights related to measurements, models and emission modelling

into an inversion framework to estimate how well we can constrain urban CO<sub>2</sub> emissions per source sector (objective 4).

## Results and conclusions

In Chapter 2 we examined the effectiveness of two observation sites close to the city border of Rotterdam, providing a gradient in the CO<sub>2</sub> mixing ratio over the city from the upwind to the downwind site. The two sites provide one year of hourly mixing ratio gradients which are used to make a first estimate of the urban emissions. For this purpose we first examined whether the upwind site was representative for the composition of the background signal, which proved to be the case for specific wind directions. We found on average large enhancements at the downwind site compared to the upwind site for three major source areas: the city, the port and the glasshouse area. From the selected gradients we calculated emissions, accounting only for average biospheric fluxes, footprints, and boundary layer height. Although this approach is very simplified it shows reasonable flux estimates compared to the reported emissions. Nevertheless, we found that the estimates can be heavily influenced by local emissions and by transport processes that we could not take into account. For example, the presence of elevated stack emissions complicates the estimate of the emissions without detailed knowledge of the atmospheric transport. Finally, the results show that CO can potentially attribute a CO<sub>2</sub> signal to industrial or residential source areas. We conclude that observed mixing ratio gradients can be used to make a rough estimate of the urban emissions, in which CO is of added value to identify dominant source types.

In Chapter 3 we compared two atmospheric transport models: the Eulerian WRF-Chem model (1x1 km<sup>2</sup> resolution) and the Lagrangian OPS model. Atmospheric transport models are useful to account for the impact of transport, mixing, entrainment, and biospheric fluxes on the observed mixing ratios and can help interpret the observed signals. We examined the ability of these models to reproduce the observed mixing ratios at several measurement sites along a transect from an urban (Rotterdam) to rural location. On average, WRF-Chem gives good results, reproducing meso-scale features with the correct order of magnitude for the observed CO<sub>2</sub> mixing ratios. However, the timing of CO<sub>2</sub> mixing ratio enhancements is often incorrect, which is mainly the result of an incorrect representation of the wind direction causing the model to sample the wrong source area. Moreover, we found that the representation of point sources is problematic. In a Eulerian model emissions get instantly mixed throughout the grid box, which causes a large underestimation of local and downwind mixing ratios for sources with a small horizontal extent. Using the OPS model improves the representation of point sources, because it has no spatial discretization. The difference between OPS and WRF-Chem is only visible up to approximately 15 km from major stack emissions, such that point sources further away from observation sites can be represented by WRF-Chem as well. An additional advantage

of the OPS model is that it can be driven by locally observed meteorological data, such that it overcomes the wind direction issue from WRF-Chem. However, the OPS model is sub-optimal for area source emissions over a large domain and therefore we conclude that a combination of both models is the best option in Rotterdam. Finally, the results in Chapter 3 show that urban sites are well-exposed to urban fossil fuel fluxes and can be used to separate between different source areas (such as the residential and industrial area), especially if besides CO<sub>2</sub> also CO is included. Sites that are further removed from the city (semi-urban) provide a better constraint on the total flux.

Chapter 4 explored the potential of several data streams to predict high-resolution emissions. These data were combined in a dynamic fossil fuel emission model that estimates emissions based on additional knowledge about the emission landscape. First, we calculated the total yearly emissions for the Netherlands per source sector using activity data (such as Gross Domestic Product), emission factors (the amount of CO<sub>2</sub> emitted per amount of fuel consumed) and energy efficiency (amount of fuel consumed per amount of activity). Then the total yearly emissions were disaggregated to hourly and 1x1 km<sup>2</sup> scale using proxies and hourly activity data. In this way we created a dynamic emission map based on a wide range of parameters that are specified per source sector. One major advantage is that we can estimate the (unknown) uncertainty in the high-resolution emissions from the (better-known) uncertainty in the model parameters. We find that we can estimate the yearly emissions for the Netherlands with a 15% uncertainty when using generalized proxies (i.e. based on general, large-scale activity data and emission factors). Using more specific knowledge about the region (e.g. about technological advancement) and local activity data reduces this uncertainty. We can also use the emission model to calculate emissions of co-emitted species by multiplying the CO<sub>2</sub> emissions with the typical emission ratios for each source sector. These emission ratios are variable and uncertain and the emissions of co-emitted species have a larger uncertainty than the CO<sub>2</sub> emission. Finally, the model parameters have a physical meaning and can be linked to emission reduction policies, making it a useful tool for policy-makers.

With the dynamic emission model we identified the most important and uncertain parameters affecting the emissions (CO<sub>2</sub> emission factors, emission ratios and time profiles). In Chapter 5 we tried to optimize these parameters using a newly developed inverse modelling framework. The inversion system uses the multi-model framework described in Chapter 3 to translate the emissions calculated by the dynamic emission model into mixing ratios of CO<sub>2</sub>, CO, NO<sub>x</sub> (nitrogen oxides) and SO<sub>2</sub> (sulphur dioxide). We used the same modelling framework to create pseudo-observations, which are used to validate the model. The only difference is the values appointed to the parameters in the emission model (generalized data for the prior, local data for the pseudo-observations). We performed an experiment to explore the difference between an urban and a rural observation network, which shows that the CO<sub>2</sub> signals captured by the rural network are too small to contain relevant information. The urban network performs well and gives a

good estimate of the total yearly emissions for the Rotterdam area (5% error). When we included observations of the co-emitted tracers the emission estimate per source sector generally improves. Some sectors remain difficult to constrain, for example due to the lack of large enhancements or the lack of a clear emission ratio signature. The time profiles can also be constrained relatively well, at least the day-to-day variability. However, for households the error in the time profile gets aliased into the emission factor, causing the emission factor to be less well constrained. When we introduced erroneous atmospheric transport the results deteriorate drastically, especially for power plants and industry (i.e. point sources) which suffer most from the transport errors. We conclude that an inversion system with a dynamic emission model as a prior has great potential for monitoring urban emissions, but transport errors currently hamper its applicability to real observations.

This work contributed to a better understanding of the complexity of the urban fossil fuel emissions and what is needed to monitor this. Urban observations provide useful information and, depending on the size and shape of the monitoring network, can be used to constrain urban emissions in more or less detail. Observations of co-emitted species have the potential to attribute CO<sub>2</sub> emissions to specific source sectors and are an important addition to our inversion framework. The dynamic fossil fuel emission model has several major advantages over a regular emission map, being flexible and physically meaningful. Although several challenges remain, the work described in this thesis is an important step in the development of urban monitoring capacities.

## Samenvatting

### Achtergrond

De verbranding van fossiele brandstoffen zorgt voor een toename van koolstofdioxide ( $\text{CO}_2$ ) in de atmosfeer en is een belangrijke oorzaak van klimaatverandering. Daarom worden er allerlei (inter)nationale overeenkomsten gesloten om de  $\text{CO}_2$  uitstoot (emissies) terug te brengen, waarvan het Parijs-akkoord het bekendst is. Elk land dat deelneemt aan dit akkoord moet streven naar een vooraf vastgestelde reductie van de  $\text{CO}_2$  emissies. In zo'n samenwerkingsverband is het belangrijk om bij te houden hoeveel vooruitgang is geboekt in het behalen van de doelstellingen, maar ook om inzicht te verkrijgen in welke beleidsmaatregelen het meest effectief zijn in het terugdringen van de emissies om verdere besluitvorming te ondersteunen. Om dit te bewerkstelligen zijn wetenschappers begonnen met de ontwikkeling van methodes om de (verandering in de) emissies te monitoren. Deze methodes richten zich voornamelijk op steden. Bijna 70% van de door de mens veroorzaakte  $\text{CO}_2$  emissies vindt plaats in steden en hier zullen de grootste reducties in emissies plaatsvinden. Dit zorgt voor grote atmosferische signalen die relatief eenvoudig gemeten kunnen worden. Echter, monitoren in de stad brengt ook diverse uitdagingen met zich mee. Bijvoorbeeld het atmosferisch transport in een stad is complex vanwege de gebouwen, wat de interpretatie van gemeten concentraties lastig maakt. Ook richten beleidsmaatregelen zich vaak op specifieke economische sectoren (zoals verkeer of industrie). Het gevolg is dat de emissies van elke sector apart gemonitord moet worden om het effect van individuele maatregelen vast te stellen. Deze toekenning van  $\text{CO}_2$  signalen aan specifieke economische sectoren is onmogelijk met slechts  $\text{CO}_2$  metingen als de sectoren niet fysiek gescheiden zijn.

### Doelstelling

De voornaamste doelstelling van dit proefschrift is inzicht verkrijgen in wat er voor nodig is om fossiele  $\text{CO}_2$  emissies in steden te kwantificeren per economische sector. Een netwerk van metingen vormt hierbij een belangrijk uitgangspunt. We richten ons daarom eerst op de identificatie van de meest bruikbare types meetlocaties en netwerk configuraties. Naast  $\text{CO}_2$  maken we ook gebruik van andere gassen die tegelijkertijd met  $\text{CO}_2$  worden uitgestoten tijdens de verbranding van fossiele brandstoffen. Dit gebeurt in een bepaalde verhouding (ratio) die specifiek is voor een economische sector, waardoor deze gassen potentieel gebruikt kunnen worden om de bron van een  $\text{CO}_2$  signaal te achterhalen. We onderzoeken deze mogelijkheid om andere gassen te gebruiken voor de toekenning van  $\text{CO}_2$  signalen aan economische sectoren. Naast metingen is een goede modelweergave van atmosferisch transport nodig om de metingen te kunnen interpreteren. Dat leidt tot het tweede doel: het beter begrijpen van de mogelijkheden en

beperkingen van een atmosferisch transport model en de meest veelbelovende benadering identificeren. Het derde doel is het voorspellen van stedelijke CO<sub>2</sub> emissies op hoge resolutie op basis van diverse indicatoren en inzicht verkrijgen in de onzekerheden in de emissies. Als laatste combineren we de verkregen inzichten met betrekking tot de metingen en modellen in een inversiesysteem. Het doel hierbij is om te testen hoe goed de stedelijke CO<sub>2</sub> emissies per economische sector gekwantificeerd kunnen worden.

### Resultaten en conclusies

In hoofdstuk 2 hebben we de effectiviteit van twee meetlocaties vlakbij de stadrand van Rotterdam onderzocht. Die twee locaties geven een gradiënt in CO<sub>2</sub> concentraties over de stad van de bovenwindse naar de benedenwindse locatie, die we gebruiken om een eerste schatting te maken van de stedelijke emissies. Eerst hebben we gekeken of de bovenwindse locatie representatief was voor de samenstelling van het achtergrondsignaal, wat het geval bleek te zijn voor bepaalde windrichtingen. Op de benedenwindse locatie vonden we grote signalen voor drie brongebieden: de stad Rotterdam, de haven en het kassengebied. Met behulp van de geselecteerde concentratiegradiënten berekenden we de emissies van het brongebied, rekening houdend met gemiddelde CO<sub>2</sub> uitwisseling van de biosfeer, de voetafdruk van de metingen, en de grenslaaghoogte. Ondanks dat deze methode erg versimpeld is, geeft het een redelijke schatting van de emissies in vergelijking met gerapporteerde waarden. Toch zagen we dat de geschatte emissies sterk afhankelijk kunnen zijn van lokale bronnen en transportprocessen waar we geen rekening mee hebben kunnen houden. Bijvoorbeeld de aanwezigheid van schoorstenen bemoeilijkt de schatting zonder gedetailleerde kennis van het atmosferische transport. Tot slot laten onze resultaten zien dat koolstofmonoxide (CO) de potentie heeft om CO<sub>2</sub> signalen toe te kennen aan verschillende brongebieden die gekenmerkt worden door industrie of woningen. We concluderen dat gemeten CO<sub>2</sub> concentratiegradiënten gebruikt kunnen worden om een ruwe schatting te maken van stedelijke emissies, waarbij CO van toegevoegde waarde is voor het identificeren van dominante bronnen.

In hoofdstuk 3 vergeleken we twee atmosferische transport modellen: het Euleriaanse WRF-Chem model (1x1 km<sup>2</sup> resolutie) en het Lagrangiaanse OPS model. Atmosferische transport modellen berekenen onder andere het effect van transport en menging op de concentraties van gassen en helpen dus bij de interpretatie van metingen. We onderzochten hoe goed de modellen de gemeten concentraties kunnen reproduceren voor een aantal meetlocaties, van stedelijk gelegen (nabij Rotterdam) tot een landelijke locatie. Gemiddeld genomen geeft WRF-Chem goede resultaten. Het reproduceert mesoschaal variaties in de correcte orde van grootte voor de gemeten CO<sub>2</sub> concentraties. Echter, de tijdstippen van pieken in de CO<sub>2</sub> concentraties zijn vaak niet correct, wat vooral veroorzaakt wordt door fouten in de berekende windrichting waardoor het verkeerde

brongebied wordt opgepikt. Ook bleek de representatie van puntbronnen problematisch. In een Euleriaans model wordt de uitgestoten CO<sub>2</sub> onmiddellijk gemengd in een model “grid box”, wat zorgt voor een onderschatting van lokale en benedenwindse concentraties voor bronnen met een beperkte omvang. Het gebruik van het OPS model verbetert de representatie van puntbronnen doordat de ruimte continu is in plaats van discreet. Het verschil tussen OSP en WRF-Chem is alleen zichtbaar tot op ongeveer 15 kilometer van grote puntbronnen, waardoor puntbronnen die verder verwijderd zijn van de meetlocaties ook door WRF-Chem opgelost kunnen worden. Een bijkomend voordeel van het OPS model is dat het gedreven kan worden door lokaal gemeten atmosferische omstandigheden, zodat het probleem met de windrichting in WRF-Chem geen rol meer speelt. Aan de andere kant is het OPS model niet optimaal voor oppervlaktebronnen over een groot gebied en we concluderen dan ook dat een combinatie van beide modellen de beste optie is voor Rotterdam. Als laatste laten de resultaten in hoofdstuk 3 zien dat stedelijke meetlocaties blootgesteld zijn aan de stedelijke CO<sub>2</sub> emissies en geschikt zijn om verschillende brongebieden te onderscheiden, vooral als we naast CO<sub>2</sub> ook CO meenemen. Meetlocaties verder van de stad (landelijk gebied) zijn geschikter om de totale emissies van de stad te kwantificeren.

Hoofdstuk 4 beschrijft de potentie van allerlei datastromen om stedelijke emissies te bepalen op hoge resolutie. De data werden gecombineerd in een dynamisch emissiemodel dat de emissies schat op basis van extra kennis over het gebied. Eerst berekenden we de totale jaarlijkse emissies van Nederland per economische sector op basis van activiteit (zoals bruto nationaal product), emissiefactoren (de hoeveelheid CO<sub>2</sub> uitgestoten per hoeveelheid geconsumeerde brandstof) en energie-efficiëntie (de hoeveelheid geconsumeerde brandstof per hoeveelheid activiteit). Daarna werden de totale emissies opgesplitst naar emissies op 1x1 km<sup>2</sup> per uur door middel van diverse indicatoren en activiteit per uur. Op deze manier hebben we een dynamische emissiekaart geproduceerd op basis van een grote hoeveelheid parameters die specifiek zijn vastgesteld voor elke economische sector. Een belangrijk voordeel van het emissiemodel is dat we de (onbekende) onzekerheid in de hoge-resolutie emissies kunnen schatten uit de (beter bekende) onzekerheid in de gebruikte data. De onzekerheid in de geschatte jaarlijkse emissies voor Nederland is 15% wanneer we algemeen geldende indicatoren gebruiken. Als we specifieke kennis over de regio en lokale activiteit (zoals verkeerstellingen in de stad zelf) gebruiken wordt deze onzekerheid aanzienlijk kleiner. We kunnen het emissiemodel ook gebruiken om emissies van andere gassen te schatten door de CO<sub>2</sub> emissies te vermenigvuldigen met de typische emissieratio voor elke economische sector. Deze ratio's zijn variabel en onzeker, waardoor de emissies van andere gassen een grotere onzekerheid hebben dan de CO<sub>2</sub> emissies. Een laatste voordeel van het emissiemodel is dat de parameters een fysieke betekenis hebben en gerelateerd kunnen worden aan beleidsmaatregelen, waardoor het gebruikt kan worden door beleidsmakers.



Met behulp van het dynamische emissiemodel hebben we de belangrijkste en meest onzekere parameters geschat die van invloed zijn op de emissies ( $\text{CO}_2$  emissiefactoren, emissieratio's en tijdsprofielen). In hoofdstuk 5 hebben we geprobeerd de exacte waarden van deze parameters te schatten met een nieuw ontwikkeld inversiesysteem. Het inversiesysteem gebruikt de combinatie van WRF-Chem en OPS beschreven in hoofdstuk 3 om de emissies berekend met het dynamische emissiemodel te vertalen naar atmosferische concentraties van  $\text{CO}_2$ , CO,  $\text{NO}_x$  (stikstofoxiden) en  $\text{SO}_2$  (zwaveldioxide). We gebruikten dezelfde modellen voor het maken van pseudo-metingen, die gebruikt worden voor validatie van het model. Het enige verschil is de waarde die toegekend wordt aan elke parameter van het emissiemodel (algemene waarden als uitgangspunt voor de inversie, lokale data voor de pseudo-metingen). We hebben een experiment gedaan om het verschil tussen een stedelijk en ruraal netwerk van meetlocaties te onderzoeken. Daaruit blijkt dat de  $\text{CO}_2$  signalen gemeten op de rurale meetlocaties te klein zijn om daar informatie uit te kunnen halen. Het stedelijke netwerk geeft betere resultaten en geeft een goede schatting van de totale jaarlijkse emissies voor Rotterdam (met een fout van 5%). Als we metingen van de andere gassen toevoegen verbetert de emissieschatting per economische sector. De emissies van sommige sectoren blijft echter lastig te bepalen, bijvoorbeeld doordat er weinig signalen zijn die informatie bevatten over die sector of omdat de emissieratio's van die sector geen duidelijk signatuur hebben. De dag-tot-dag variaties in de emissies kunnen redelijk goed bepaald worden, hoewel voor sommige sectoren het verschil tussen gemeten en berekende concentraties onterecht toegekend wordt aan de emissiefactor in plaats van het tijdsprofiel. Als we fouten aanbrengen in het atmosferisch transport verslechteren de resultaten drastisch, vooral voor economische sectoren die vooral bestaan uit puntbronnen (zoals energiecentrales en industrie). We concluderen dat het inversiesysteem met als uitgangspunt het dynamisch emissiemodel potentie heeft voor het monitoren van stedelijke  $\text{CO}_2$  emissies, maar dat de huidige problemen met het berekenen van atmosferisch transport een belemmering zijn voor toepassingen met echte metingen.

Dit werk heeft bijgedragen aan een beter begrip van de complexiteit van stedelijke  $\text{CO}_2$  emissies door verbranding van fossiele brandstoffen en wat nodig is om dit te monitoren. Stedelijke meetlocaties bevatten veel informatie en kunnen, afhankelijk van de grootte en vorm van het netwerk, gebruikt worden om stedelijke emissies te schatten in meer of minder detail. Metingen van andere gassen hebben de potentie om  $\text{CO}_2$  signalen toe te kennen aan specifieke economische sectoren en zijn een belangrijke toevoeging in ons inversiesysteem. Het dynamische emissiemodel heeft een aantal belangrijke voordelen ten opzichte van standaard emissiekaarten, zoals de flexibiliteit en de relatie met beleidsmaatregelen. Ondanks dat er meerdere uitdagingen overblijven, is het werk beschreven in dit proefschrift een belangrijke stap in de ontwikkeling van methodes voor het monitoren van stedelijke emissies.

## Appendix A: Supplementary information to Chapter 2

### A.1 Instrumentation

We installed low-drift analysers based on cavity ring-down spectroscopy (CRDS) (Picarro Inc., CA, USA, type G2401) to measure atmospheric concentrations of CO<sub>2</sub>, CH<sub>4</sub> and CO. The CRDS technique uses optical absorption at wavelengths of 1603 nm, 1651 nm and ~1570 nm to simultaneously determine concentrations of CO<sub>2</sub>, CH<sub>4</sub> and CO, respectively. A laser beam at specific wavelength is directed into the cavity, which consists of three highly reflective mirrors, and travels through the cavity with an effective path length of 15-20 km. At wavelengths insensitive to absorption by gases, the decay time of light intensity is only determined by the small mirror transmittance (cavity-only ring down time). When changing to a wavelength sensitive to absorption by the gas sample an additional loss term is introduced, leading to a faster decay in light intensity (total ring down time). Comparing the total ring down time with the cavity-only ring down time provides information on the trace gas concentrations. A more detailed description of the instrument can be found in Crosson (2008).

Sampled air continuously flows through the cavity at regulated temperature and pressure, such that the measurements are insensitive to changes in ambient temperature and pressure (Crosson, 2008). Moreover, the CRDS analyser reports dry mole fractions based on the simultaneously measured water vapour concentrations and laboratory-derived water correction functions for CO<sub>2</sub> and CH<sub>4</sub> (Chen et al., 2010; Rella et al., 2013). Although a similar function has been established for CO, its generalizability for individual analysers is low (Chen et al., 2013). Therefore, the manufactory default water vapour correction is applied for CO.

### A.2 Calibration and data processing

The raw data for which the cavity pressure or temperature deviates more than 0.7 hPa (0.5 Torr) or 0.1 °C from the nominal settings of 186.2 hPa (140 Torr) or 45 °C, respectively, were flagged. We further corrected CO<sub>2</sub> and CH<sub>4</sub> mole fractions for the cavity pressure variations within 0.7 hPa from the nominal setting (Filges et al., 2015). The corrections are not necessary for normal monitoring conditions, but may be helpful when pressure variations occur during valve switching or short-term noise increase of the cavity pressure. Note that the magnitude of the corrections is up to 0.25 ppm for CO<sub>2</sub> and 4 ppb for CH<sub>4</sub>. Finally, the raw and temporally high-resolution (0.2 Hz) data were aggregated to minute averages.

The instruments are connected to 3 cylinders with different mole fractions, of which the exact mole fractions have been established in the Centre for Isotope Research in Groningen against the WMO X2007 (CO<sub>2</sub>) and X2004 (CO and CH<sub>4</sub>) scale (Zhao and Tans,

2006) (Table A1, values between brackets). Cylinder 2 was the reference used for calibration, whereas cylinders 1 and 3 were used as independent targets for quality control. Alternately, the contents of one of the cylinders is measured by the instrument for 20 minutes after 12 hours of measuring the sample (S). The measurement sequence is thus S (720 min.) – Targ1 (20 min.) - S (720 min.) – Ref2 (20 min.) - S (720 min.) – Targ3 (20 min.). This sequence is constantly repeated. Due to the instrument's linear response with concentration, using only two or three cylinders is sufficient (Crosson, 2008). The cylinders were used for final processing and calibration.

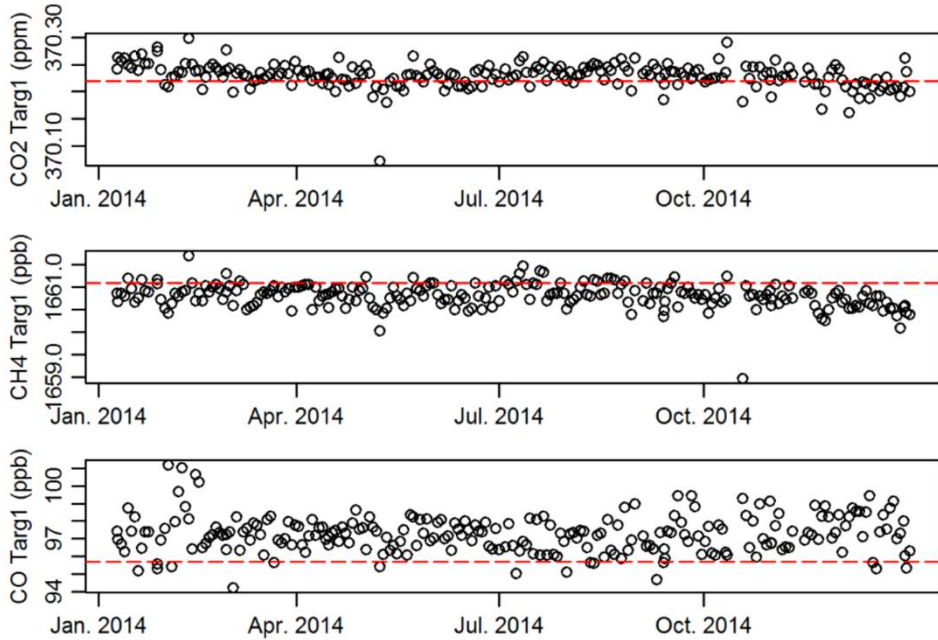
**Table A1: Average and 1 $\sigma$  of CO<sub>2</sub>, CH<sub>4</sub>, and CO mole fractions in the Targ1 and Targ3 cylinders as measured by the instruments at both measurement sites; between brackets the contents as established in the laboratory is given.**

	<b>Westmaas</b>			<b>Zweth</b>		
	CO <sub>2</sub> [ppm]	CH <sub>4</sub> [ppb]	CO [ppb]	CO <sub>2</sub> [ppm]	CH <sub>4</sub> [ppb]	CO [ppb]
<b>Targ1</b>	370.23±0.02 (370.22)	1660.8±0.27 (1661.1)	97.26±1.07 (95.7)	399.97±0.06 (399.98)	1952.7±0.34 (1952.5)	221.85±1.00 (223.2)
<b>Ref2</b>	(400.01)	(1957.4)	(229.2)	(410.55)	(1804.3)	(201.2)
<b>Targ3</b>	501.71±0.04 (501.70)	2099.6±0.27 (2099.9)	514.07±1.17 (512.8)	449.84±0.06 (449.85)	1956.7±0.30 (1956.9)	376.99±1.02 (377.7)

Based on the time series of the three cylinders we applied several corrections. First, the time series showed a systematic error compared to the concentrations listed in Table A1. So the first step was to correct for this bias. This offset-correction was also applied to the ambient air measurements. Secondly, there was a long-term analyser drift and more short-term tendencies. The data was corrected for these tendencies by applying a linear interpolation to the Ref2 calibration data, which means that every measured concentration from the Ref2 cylinder was put to exactly the value given in Table A1. The measurements of ambient air and the other target cylinder concentrations were then corrected using the interpolation. This single bias correction using only Ref2 allowed us to use the other cylinders for quality control. Finally, the minute-averages were aggregated to hourly averages. Similar methods are applied for CH<sub>4</sub> and CO.

### A.3 Data quality

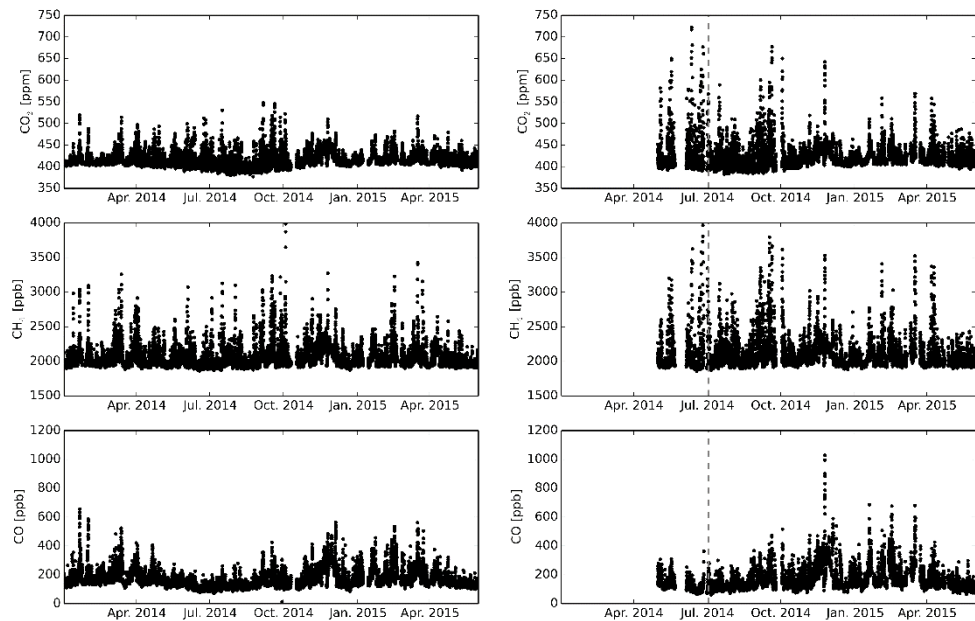
Figure A1 shows the CO<sub>2</sub>, CH<sub>4</sub> and CO mole fractions in target cylinder 1 measured at Westmaas after data processing. The mixing ratio of Targ1 as established in the laboratory is given by the red dashed line. All results for Targ1 and Targ3 at both locations are summarised in Table A1. The difference between the average Targ1 and Targ3 mole fractions measured by the instrument and the mole fractions determined in the laboratory gives us the bias in the observations. The CO<sub>2</sub> measurements are very accurate, with a bias of 0.01 ppm in all cases. The bias is largest for CO (up to 1.6 ppb). The precision of the measurements, based on the standard deviation, is satisfactory for all species. CO shows the largest relative variation.



**Figure A1:** Time series of CO<sub>2</sub>, CH<sub>4</sub>, and CO mixing ratio for target cylinder 1 after data processing at location Westmaas for the year 2014. The mole fraction of Targ1 as established in the laboratory is given by the dashed line.

The measurements at both locations are made continuously, although they are sometimes interrupted by technical malfunctions. For example, if the cavity temperature deviates too much from the regulated temperature, the measurements are shut down until the instrument has rebooted. Moreover, other activities near the measurement sites that could disturb the observations are written down in a log. Data affected by these malfunctions or activities are either corrected or flagged, such that the data quality is not affected. The resulting time series after all data processing are shown in Figure A2.

At the start of the measurements at Zweth a 10m mast was unavailable. Therefore, the sampling height was 3m until July 2014. After installing the 10m mast, we continued to take a few samples per day at 3m during one month (July). Previously observed vertical profiles show that the CO<sub>2</sub> mixing ratio can change with several ppm moving only a few tens of meters upwards (Vermeulen et al., 2011). This effect is especially clear during the night, but is dependent on the season due to the role of vegetation as CO<sub>2</sub> source and sink. Similarly, the CH<sub>4</sub> mixing ratio can vary with height due to the presence of local sources. Note that the inlet height is explicitly included in the data files provided for download. We examine the difference between the 3 and 10 m observations by binning observations by hour. This results in an average daily cycle from the irregular samples ( $N = 263$  for both heights). The daily cycles of the CO<sub>2</sub>, CH<sub>4</sub>, and CO mixing ratio at both heights and the differences between the two heights (given by  $\delta$ ) relative to the mixing ratio at 10m are shown in Figure A3.



**Figure A2:** Time series of ambient CO<sub>2</sub>, CH<sub>4</sub>, and CO mixing ratio at Westmaas (left) and Zweth (right). The vertical dashed lines indicate the transition from 3 to 10m sampling height.

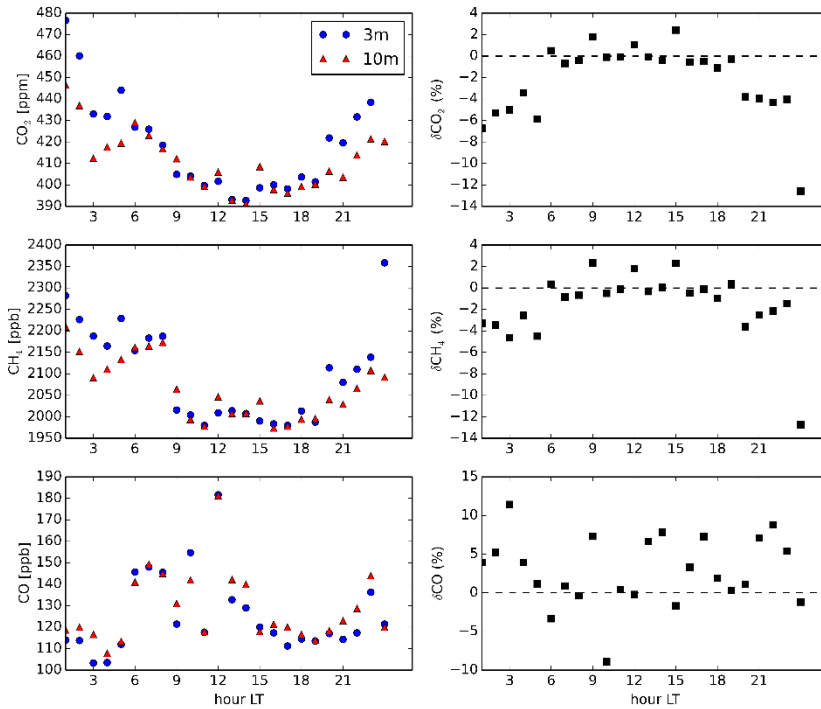
The diurnal cycle is mainly determined by atmospheric mixing and for CO<sub>2</sub> the diurnal cycle is stronger at 3m than at 10m. The mixing ratio is generally larger at 3m, except from about 8-15h UTC (10-17h LT) when  $\delta\text{CO}_2$  fluctuates around zero. For CO<sub>2</sub>, emissions from vegetation and the soil play an important role in this diurnal cycle. During the night respiration causes the CO<sub>2</sub> mixing ratio to increase. Since there is little mixing in a stable boundary-layer this increase is more pronounced closer to the surface. During daytime photosynthesis has the opposite effect, being an important sink of CO<sub>2</sub>. However, due to convective mixing the difference between the two heights is smaller than at night time.

CH<sub>4</sub> shows a similar difference between the two heights as CO<sub>2</sub>. However, vegetation does not affect CH<sub>4</sub>, so there are other local sources that cause this difference. One potentially important source is the peat soil, which is a CH<sub>4</sub> producer due to the high ground water level at this location (Le Mer and Roger, 2001; Schrier-Uijl et al., 2014; Smith et al., 2003). Estimates of CH<sub>4</sub> emissions are highly variable, as they depend on many environmental factors. Yet, emissions of  $10\text{-}30 \mu\text{mol m}^{-2} \text{hr}^{-1}$  have been estimated for Scottish wetlands with soil water depths of 0-10 cm (Smith et al., 2003), which could significantly enhance the CH<sub>4</sub> mixing ratio close to the surface. During daytime convection causes this CH<sub>4</sub> flux to be mixed quickly through the boundary-layer, causing  $\delta\text{CH}_4$  to be close to zero.

CO has no local sources and this mixing ratio is mainly determined by advection. This may explain the increase in CO in the morning, which has about the same timing as the

morning rush hour. Due to the absence of local influences we find no clear diurnal pattern in  $\delta\text{CO}$ . Yet the CO mixing ratio generally seems to be higher at 10m. This can be explained by the presence of large trees and bushes that surround the measurement area and potentially reduce the effect of advection at 3m. The 10m mast reaches to the tree tops and is therefore more exposed to advected air.

To conclude, users of this data set should be aware that data for May and June 2014 at location Zweth are more affected by local influences due to the lower sampling height. Considering this time series as homogeneous can therefore result in a bias. However, between 6 and 18h UTC the impact appears to be limited.



**Figure A3:** Left: Hourly averaged observed  $\text{CO}_2$ ,  $\text{CH}_4$ , and CO mixing ratio at 3 (blue circles) and 10m (red triangles) height at Zweth; Right: the difference between 10 and 3m  $\text{CO}_2$ ,  $\text{CH}_4$ , and CO mixing ratio relative to the mixing ratio at 10m. LT is the local wintertime (UTC+1).



## References

- Andres, R. J., Boden, T. A., Bréon, F. M., Ciais, P., Davis, S., Erickson, D., Gregg, J. S., Jacobson, A., Marland, G., Miller, J., Oda, T., Olivier, J. G. J., Raupach, M. R., Rayner, P., and Treanton, K.: A synthesis of carbon dioxide emissions from fossil-fuel combustion, *Biogeosciences*, 9, 1845-1871, 10.5194/bg-9-1845-2012, 2012.
- Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO<sub>2</sub>, CO, and CH<sub>4</sub> measurements from tall towers in the NOAA earth system research laboratory's global greenhouse gas reference network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, *Atmos. Meas. Tech.*, 7, 647-687, 10.5194/amt-7-647-2014, 2014.
- Andrews, G. E., Zhu, G., Li, H., Simpson, A., Wylie, J. A., Bell, M., and Tate, J.: The effect of ambient temperature on cold start urban traffic emissions for a real world SI car, in: SAE International, 2004.
- Angevine, W. M., Brioude, J., McKeen, S., and Holloway, J. S.: Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble, *Geosci. Model Dev.*, 7, 2817-2829, 10.5194/gmd-7-2817-2014, 2014.
- Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., and White, J. W. C.: Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years, *Nature*, 487, 70-72, 10.1038/nature11299, 2012.
- Basu, S., Miller, J. B., and Lehman, S.: Separation of biospheric and fossil fuel fluxes of CO<sub>2</sub> by atmospheric inversion of CO<sub>2</sub> and <sup>14</sup>CO<sub>2</sub> measurements: Observation System Simulations, *Atmos. Chem. Phys.*, 16, 5665-5683, 10.5194/acp-16-5665-2016, 2016.
- Bergeron, O., and Strachan, I. B.: CO<sub>2</sub> sources and sinks in urban and suburban areas of a northern mid-latitude city, *Atmos. Environ.*, 45, 1564-1573, 10.1016/j.atmosenv.2010.12.043, 2011.
- Berner, R. A.: The long-term carbon cycle, fossil fuels and atmospheric composition, *Nature*, 426, 323-326, 10.1038/nature02131, 2003.
- Bieser, J., Aulinger, A., Matthias, V., Quante, M., and Denier van der Gon, H. A. C.: Vertical emission profiles for Europe based on plume rise calculations, *Environ. Pollut.*, 159, 2935-2946, 10.1016/j.envpol.2011.04.030, 2011.
- Boersma, K. F., Geert, C. M. V., and Jean, T.: Ships going slow in reducing their NO<sub>x</sub> emissions: Changes in 2005–2012 ship exhaust inferred from satellite measurements over Europe, *Environ. Res. Lett.*, 10, 074007, 10.1088/1748-9326/10/7/074007, 2015.
- Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramonet, M., Paris, J. D., and Ciais, P.: Analysis of the potential of near-ground measurements of CO<sub>2</sub> and CH<sub>4</sub> in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model, *Atmos. Chem. Phys.*, 16, 6735-6756, 10.5194/acp-16-6735-2016, 2016.
- Bousquet, P., Gaudry, A., Ciais, P., Kazan, V., Monfray, P., Simmonds, P. G., Jennings, S. G., and O'Connor, T. C.: Atmospheric CO<sub>2</sub> concentration variations recorded at Mace Head, Ireland, from 1992 to 1994, *Phys. Chem. Earth*, 21, 477-481, 1996.



## REFERENCES

---

- Bozhinova, D., Van der Molen, M. K., Van der Velde, I. R., Krol, M. C., Van der Laan, S., Meijer, H. A. J., and Peters, W.: Simulating the integrated summertime  $\delta^{14}\text{CO}_2$  signature from anthropogenic emissions over Western Europe, *Atmos. Chem. Phys.*, **14**, 7273-7290, 10.5194/acp-14-7273-2014, 2014.
- Brand, C.: Beyond 'Dieselgate': Implications of unaccounted and future air pollutant emissions and energy use for cars in the United Kingdom, *Energy Policy*, **97**, 1-12, 10.1016/j.enpol.2016.06.036, 2016.
- Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area  $\text{CO}_2$  emissions from atmospheric concentration measurements, *Atmos. Chem. Phys.*, **15**, 1707-1724, 10.5194/acp-15-1707-2015, 2015.
- Briant, R., and Seigneur, C.: Multi-scale modeling of roadway air quality impacts: Development and evaluation of a Plume-in-Grid model, *Atmos. Environ.*, **68**, 162-173, 10.1016/j.atmosenv.2012.11.058, 2013.
- Brioude, J., Petron, G., Frost, G. J., Ahmadov, R., Angevine, W. M., Hsie, E. Y., Kim, S. W., Lee, S. H., McKeen, S. A., Trainer, M., Fehsenfeld, F. C., Holloway, J. S., Peischl, J., Ryerson, T. B., and Gurney, K. R.: A new inversion method to calculate emission inventories without a prior at mesoscale: Application to the anthropogenic  $\text{CO}_2$  emission from Houston, Texas, *J. Geophys. Res.-Atmos.*, **117**, 1-15, 10.1029/2011JD016918, 2012.
- Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S. W., Evan, S., McKeen, S. A., Hsie, E. Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: Assessing anthropogenic emissions of  $\text{CO}$ ,  $\text{NO}_x$  and  $\text{CO}_2$  and their impacts, *Atmos. Chem. Phys.*, **13**, 3661-3677, 10.5194/acp-13-3661-2013, 2013.
- Brophy, K., Graven, H., Manning, A. J., White, E., Arnold, T., Fischer, M. L., Jeong, S., Cui, X., and Rigby, M.: Characterizing uncertainties in atmospheric inversions of fossil fuel  $\text{CO}_2$  emissions in California, *Atmos. Chem. Phys. Discuss.*, 2018, 1-44, 10.5194/acp-2018-473, 2018.
- Broquet, G., Chevallier, F., Bréon, F. M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of  $\text{CO}_2$  ecosystem fluxes from atmospheric measurements: Reliability of the uncertainty estimates, *Atmos. Chem. Phys.*, **13**, 9039-9056, 10.5194/acp-13-9039-2013, 2013.
- Buckley, S. M., Mitchell, M. J., McHale, P. J., and Millard, G. D.: Variations in carbon dioxide fluxes within a city landscape: Identifying a vehicular influence, *Urban Ecosyst.*, **1**, 1-20, 10.1007/s11252-013-0341-0, 2014.
- Calabrese, E. J.: Ozone Risk Communication and Management, edited by: Gilbert, C. E., and Beck, B. D., CRC Press, London, 1990.
- Caulton, D. R., Shepson, P. B., Santoro, R. L., Sparks, J. P., Howarth, R. W., Ingraffea, A. R., Cambaliza, M. O. L., Sweeney, C., Karion, A., Davis, K. J., Stirr, B. H., Montzka, S. A., and Miller, B. R.: Toward a better understanding and quantification of methane emissions from shale gas development, in: *P. Natl. Acad. Sci. USA*, *P. Natl. Acad. Sci. USA*, 2014, 6237-6242, 2014.

- CBS: Energieverbruik huishoudens; energiedragers, 1975-2016, <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82375NED&D1=0&D2=6&D3=15,39&HDR=T,G1&STB=G2&VW=T>, access: 22-05-2018.
- Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: Development, evaluation, and applications to urban environmental problems, *Int. J. Climatol.*, **31**, 273-288, 10.1002/joc.2158, 2011.
- Chen, H., Winderlich, J., Gerbig, C., Hofer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO<sub>2</sub> and CH<sub>4</sub>) using the cavity ring-down spectroscopy (CRDS) technique, *Atmos. Meas. Tech.*, **3**, 375-386, 10.5194/amt-3-375-2010, 2010.
- Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and Tans, P. P.: Accurate measurements of carbon monoxide in humid air using the cavity ring-down spectroscopy (CRDS) technique, *Atmos. Meas. Tech.*, **6**, 1031-1040, 10.5194/amt-6-1031-2013, 2013.
- Chen, J. M., Mo, G., and Deng, F.: A joint global carbon inversion system using both CO<sub>2</sub> and <sup>13</sup>CO<sub>2</sub> atmospheric concentration data, *Geosci. Model Dev.*, **10**, 1131-1156, 10.5194/gmd-10-1131-2017, 2017.
- Ciais, P., Paris, J. D., Marland, G., Peylin, P., Piao, S. L., Levin, I., Pregarer, T., Scholz, Y., Friedrich, R., Rivier, L., Houwelling, S., and Schulze, E. D.: The European carbon balance. Part 1: Fossil fuel emissions, *Glob. Change Biol.*, **16**, 1395-1408, 10.1111/j.1365-2486.2009.02098.x, 2010.
- Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F. M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., Defries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J. D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., Van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, *Biogeosciences*, **11**, 3547-3602, 10.5194/bg-11-3547-2014, 2014.
- Coenen, P. W. H. G., Van der Maas, C. W. M., Zijlema, P. J., Baas, K., Van den Berghe, A. C. W. M., Te Biesebeek, J. D., Brandt, A. T., Geilenkirchen, G., Van der Hoek, K. W., Te Molder, R., Dröge, R., Montfoort, J. A., Peek, C. J., Vonk, J., and Van den Wyngaert, I.: Greenhouse Gas Emissions in the Netherlands 1990-2010. National Inventory Report 2012, Rijksinstituut voor volksgezondheid en milieu (RIVM), Bilthoven, 278, 2012.
- Cohn, S. E., Sivakumaran, N. S., and Todling, R.: A fixed-lag Kalman smoother for retrospective data assimilation, *Mon. Weather Rev.*, **122**, 2838-2867, 10.1175/1520-0493(1994)122<2838:afkfsf>2.0.co;2, 1994.
- Cooperative Global Atmospheric Data Integration Project: Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1968-2014;

## REFERENCES

---

- obspack\_co2\_1\_GLOBALVIEWplus\_v1.0\_2015-07-30, in, NOAA Earth System Research Laboratory, Global Monitoring Division, 2015.
- Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, *Appl. Phys. B*, 92, 403-408, 10.1007/s00340-008-3135-y, 2008.
- Daniels, E., Lenderink, G., Hutjes, R., and Holtslag, A.: Relative impacts of land use and climate change on summer precipitation in the Netherlands, *Hydrol. Earth Syst. Sc.*, 20, 4129-4142, 10.5194/hess-20-4129-2016, 2016.
- Dekker, I. N., Houweling, S., Aben, I., Röckmann, T., Krol, M., Martínez-Alonso, S., Deeter, M. N., and Worden, H. M.: Quantification of CO emissions from the city of madrid using MOPITT satellite retrievals and WRF simulations, *Atmos. Chem. Phys.*, 17, 14675-14694, 10.5194/acp-17-14675-2017, 2017.
- Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., Wu, K., Sarmiento, D. P., Hardesty, R. M., Bonin, T. A., Brewer, W. A., and Gurney, K. R.: Toward reduced transport errors in a high resolution urban CO<sub>2</sub> inversion system, *Elementa (Wash. D.C.)*, 5, 10.1525/elementa.133, 2017.
- Denier van der Gon, H., Beevers, S., D'Allura, A., Finardi, S., Honoré, C., Kuenen, J., Perrussel, O., Radice, P., Theloke, J., Uzbasic, M., and Visschedijk, A.: Discrepancies Between Top-down and Bottom-up Emission Inventories of Megacities: The Causes and Relevance for Modeling Concentrations and Exposure, in: *Air Pollution Modeling and its Application XXI. NATO Science for Peace and Security Series C: Environmental Security*, Springer Netherlands, Dordrecht, 199-204, 2012.
- Denier van der Gon, H. A. C., Hendriks, C., Kuenen, J., Segers, A., and Visschedijk, A.: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, TNO, Utrecht, 2011.
- Denier van der Gon, H. A. C., Kuenen, J. J. P., Janssens-Maenhout, G., Döring, U., Jonkers, S., and Visschedijk, A.: TNO\_CAMS high resolution European emission inventory 2000-2014 for anthropogenic CO<sub>2</sub> and future years following two different pathways, *Earth Syst. Sci. Data Discuss.*, 2017, 1-30, 10.5194/essd-2017-124, 2017.
- Derwent, R. G., Ryall, D. B., Manning, A. J., Simmonds, P. G., O'Doherty, S., Biraud, S., Ciais, P., Ramonet, M., and Jennings, S. G.: Continuous observations of carbon dioxide at Mace Head, Ireland from 1995 to 1999 and its net European ecosystem exchange, *Atmos. Environ.*, 36, 2799-2807, 10.1016/s1352-2310(02)00203-0, 2002.
- Dezzutti, M., Berri, G., and Venegas, L.: Intercomparison of atmospheric dispersion models applied to an urban street canyon of irregular geometry, *Aerosol Air Qual. Res.*, 18, 820-828, 10.4209/aaqr.2017.11.0489, 2018.
- Djuricin, S., Pataki, D. E., and Xu, X.: A comparison of tracer methods for quantifying CO<sub>2</sub> sources in an urban region, *J. Geophys. Res.-Atmos.*, 115, 1-13, 10.1029/2009JD012236, 2010.
- Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, *J. Atmos. Sci.*, 46, 3077-3107, 10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2, 1989.
- Duren, R. M., and Miller, C. E.: Measuring the carbon emissions of megacities, 2, 560-562, 10.1038/nclimate1629, 2012.
- Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental

- Prediction operational mesoscale Eta model, *J. Geophys. Res.-Atmos.*, 108, 10.1029/2002JD003296, 2003.
- European Commission: Observation-based system for monitoring and verification of greenhouse gases, [https://cordis.europa.eu/project/rcn/213924\\_en.html](https://cordis.europa.eu/project/rcn/213924_en.html), access: 14-5-2018.
- Filges, A., Gerbig, C., Chen, H., Franke, H., Klaus, C., and Jordan, A.: The IAGOS-core greenhouse gas package: A measurement system for continuous airborne observations of CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>O and CO, *Tellus B Chem. Phys. Meteorol.*, 67, 27989, 10.3402/tellusb.v67.27989, 2015.
- Finn, D. D., Clawson, K. L., Carter, R. G., Rich, J. D., Allwine, K. J., Flaherty, J., Biltoft, C., and Leach, M. J.: Atmospheric flow decoupling and its effects on urban plume dispersion, Seventh Symposium on the Urban Environment, San Diego, CA., 10-13 September 2007, 2007.
- Fischer, M. L., Parazoo, N., Brophy, K., Cui, X., Jeong, S., Liu, J., Keeling, R., Taylor, T. E., Gurney, K., Oda, T., and Graven, H.: Simulating estimation of California fossil fuel and biosphere carbon dioxide exchanges combining in situ tower and satellite column observations, *J. Geophys. Res.-Atmos.*, 122, 3653-3671, 10.1002/2016jd025617, 2017.
- Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, *Geosci. Model Dev.*, 2, 253-265, 10.5194/gmd-2-253-2009, 2009.
- Font, A., Grimmond, C. S. B., Kotthaus, S., Morgu , J. A., Stockdale, C., O'Connor, E., Priestman, M., and Barratt, B.: Daytime CO<sub>2</sub> urban surface fluxes from airborne measurements, eddy-covariance observations and emissions inventory in Greater London, *Environ. Pollut.*, 196, 98-106, 10.1016/j.envpol.2014.10.001, 2014.
- Ganshin, A., Oda, T., Saito, M., Maksyutov, S., Valsala, V., Andres, R. J., Fisher, R. E., Lowry, D., Lukyanov, A., Matsueda, H., Nisbet, E. G., Rigby, M., Sawa, Y., Toumi, R., Tsuboi, K., Varlagin, A., and Zhuravlev, R.: A global coupled Eulerian-Lagrangian model and 1 × 1 km CO<sub>2</sub> surface flux dataset for high-resolution atmospheric CO<sub>2</sub> transport simulations, *Geosci. Model Dev.*, 5, 231-243, 10.5194/gmd-5-231-2012, 2012.
- Gemeente Arnhem: Reversed Waste Collection: CO<sub>2</sub> emission reduction, [https://www arnhem.nl/Inwoners/wonen\\_en\\_milieu/afval/reversed\\_waste\\_collection/the\\_results/co2\\_emission\\_reduction](https://www arnhem.nl/Inwoners/wonen_en_milieu/afval/reversed_waste_collection/the_results/co2_emission_reduction), access: 2-5-2018.
- Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO<sub>2</sub> with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *J. Geophys. Res.-Atmos.*, 108, 10.1029/2003JD003770, 2003.
- G ckede, M., Turner, D. P., Michalak, A. M., Vickers, D., and Law, B. E.: Sensitivity of a subregional scale atmospheric inverse CO<sub>2</sub> modeling framework to boundary conditions, *J. Geophys. Res.-Atmos.*, 115, 10.1029/2010JD014443, 2010.
- Gratani, L., and Varone, L.: Daily and seasonal variation of CO<sub>2</sub> in the city of Rome in relationship with the traffic volume, *Atmos. Environ.*, 39, 2619-2624, 10.1016/j.atmosenv.2005.01.013, 2005.
- Graven, H., Fischer, M. L., Lueker, T., Jeong, S., Guilderson, T. P., Keeling, R. F., Bambha, R., Brophy, K., Callahan, W., Cui, X., Frankenberg, C., Gurney, K. R., LaFranchi, B. W., Lehman, S. J., Michelsen, H., Miller, J. B., Newman, S., Paplawsky, W., Parazoo, N. C., Sloop, C., and Walker, S. J.: Assessing fossil fuel CO<sub>2</sub> emissions in California using atmospheric observations and models, *Environ. Res. Lett.*, 13, 065007, 10.1088/1748-9326/aabd43, 2018.

## REFERENCES

---

- Griffin, R. J., Chen, J., Carmody, K., Vutukuru, S., and Dabdub, D.: Contribution of gas phase oxidation of volatile organic compounds to atmospheric carbon monoxide levels in two areas of the United States, *J. Geophys. Res.-Atmos.*, 112, 10.1029/2006jd007602, 2007.
- Grimmond, C. S. B., King, T. S., Cropley, F. D., Nowak, D. J., and Souch, C.: Local-scale fluxes of carbon dioxide in urban environments: Methodological challenges and results from Chicago, *Environ. Pollut.*, 116, S243-S254, 10.1016/s0269-7491(01)00256-1, 2002.
- Hausfather, Z.: Common Climate Misconceptions: Atmospheric Carbon Dioxide, <https://www.yaleclimateconnections.org/2010/12/common-climate-misconceptions-atmospheric-carbon-dioxide/>, access: August 3, 2017.
- He, W., Van der Velde, I. R., Andrews, A. E., Sweeney, C., Miller, J., Tans, P., Van der Laan-Luijkx, I. T., Nehrkorn, T., Mountain, M., Ju, W., Peters, W., and Chen, H.: CTDAS-Lagrange v1.0: A high-resolution data assimilation system for regional carbon dioxide observations, *Geosci. Model Dev. Discuss.*, 2017, 1-41, 10.5194/gmd-2017-222, 2017.
- Hendriks, D. M. D., Van Huissteden, J., Dolman, A. J., and Van der Molen, M. K.: The full greenhouse gas balance of an abandoned peat meadow, *Biogeosciences*, 4, 411-424, 10.5194/bg-4-411-2007, 2007.
- Heusinkveld, B. G., Steeneveld, G. J., Van Hove, L. W. A., Jacobs, C. M. J., and Holtslag, A. A. M.: Spatial variability of the rotterdam urban heat island as influenced by urban land use, *J. Geophys. Res.-Atmos.*, 119, 677-692, 10.1002/2012jd019399, 2014.
- Hogue, S., Marland, E., Andres, R. J., Marland, G., and Woodard, D.: Uncertainty in gridded CO<sub>2</sub> emissions estimates, *Earths Future*, 4, 225-239, 10.1002/2015ef000343, 2016.
- Hong, S. Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, *Mon. Weather Rev.*, 134, 2318-2341, 10.1175/mwr3199.1, 2006.
- Houweling, S., Dentener, F., Lelieveld, J., Walter, B., and Dlugokencky, E.: The modeling of tropospheric methane: How well can point measurements be reproduced by a global model?, *J. Geophys. Res.-Atmos.*, 105, 8981-9002, 10.1029/1999JD901149, 2000.
- Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., Butz, A., and Hasekamp, O.: Toward global mapping of methane with TROPOMI: First results and intersatellite comparison to GOSAT, *Geophys. Res. Lett.*, 45, 3682-3689, 10.1002/2018GL077259, 2018.
- Hudman, R. C., Murray, L. T., Jacob, D. J., Millet, D. B., Turquety, S., Wu, S., Blake, D. R., Goldstein, A. H., Holloway, J. S., and Sachse, G. W.: Biogenic versus anthropogenic sources of CO in the United States, *Geophys. Res. Lett.*, 35, 10.1029/2007gl032393, 2008.
- Huszar, P., Belda, M., and Halenka, T.: On the long-term impact of emissions from central European cities on regional air quality, *Atmos. Chem. Phys.*, 16, 1331-1352, 10.5194/acp-16-1331-2016, 2016.
- Hutchins, M. G., Colby, J. D., Marland, G., and Marland, E.: A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States, *Mitig. Adapt. Strat. Gl.*, 22, 947-972, 10.1007/s11027-016-9709-9, 2017.
- ICOS: RINGO. Results and impact, <https://www.icos-ri.eu/ringo/results-and-impact>, access: 14-5-2018.
- IEA: World Energy Outlook 2008, International Energy Agency, Paris, 2008.
- IPCC: Summary for policymakers, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J.,

- Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge, United Kingdom and New York, NY, USA, 2013.
- Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press, Princeton, NJ, 1999.
- Jacobs, A. F. G., Heusinkveld, B. G., and Holtslag, A. A. M.: Carbon dioxide and water vapour flux densities over a grassland area in the Netherlands, *Int. J. Climatol.*, 23, 1663-1675, doi:10.1002/joc.959, 2003.
- Järvi, L., Nordbo, A., Junninen, H., Riikonen, A., Moilanen, J., Nikinmaa, E., and Vesala, T.: Seasonal and annual variation of carbon dioxide surface fluxes in Helsinki, Finland, in 2006-2010, *Atmos. Chem. Phys.*, 12, 8475-8489, 10.5194/acp-12-8475-2012, 2012.
- Jiménez, P. A., Vilà-Guerau de Arellano, J., Dudhia, J., and Bosveld, F. C.: Role of synoptic- and meso-scales on the evolution of the boundary-layer wind profile over a coastal region: The near-coast diurnal acceleration, *Meteorol. Atmos. Phys.*, 128, 39-56, 10.1007/s00703-015-0400-6, 2016.
- Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.-Atmos.*, 106, 4703-4715, 10.1029/2000JD900581, 2001.
- Karamchandani, P., Vijayaraghavan, K., Chen, S. Y., Seigneur, C., and Edgerton, E. S.: Plume-in-grid modeling for particulate matter, *Atmos. Environ.*, 40, 7280-7297, 10.1016/j.atmosenv.2006.06.033, 2006.
- Karamchandani, P., Vijayaraghavan, K., and Yarwood, G.: Sub-grid scale plume modeling, *Atmosphere (Basel)*, 2, 389-406, 10.3390/atmos2030389, 2011.
- Karamchandani, P., Zhang, Y., and Chen, S. Y.: Development and initial application of a sub-grid scale plume treatment in a state-of-the-art online Multi-scale Air Quality and Weather Prediction Model, *Atmos. Environ.*, 63, 125-134, 10.1016/j.atmosenv.2012.09.014, 2012.
- Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R., and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, *Geophys. Res. Lett.*, 40, 4393-4397, 10.1002/grl.50811, 2013.
- Kent, C. W., Lee, K., Ward, H. C., Hong, J. W., Hong, J., Gatey, D., and Grimmond, S.: Aerodynamic roughness variation with vegetation: analysis in a suburban neighbourhood and a city park, *Urban Ecosyst.*, 21, 227-243, 10.1007/s11252-017-0710-1, 2018.
- Kim, Y., Seigneur, C., and Duclaux, O.: Development of a plume-in-grid model for industrial point and volume sources: Application to power plant and refinery sources in the Paris region, *Geosci. Model Dev.*, 7, 569-585, 10.5194/gmd-7-569-2014, 2014.
- Kleingeld, E., Van Hove, B., Elbers, J., and Jacobs, C.: Carbon dioxide fluxes in the city centre of Arnhem, a middle-sized Dutch city, *Urban Clim.*, 10.1016/j.uclim.2017.12.003, 2017.
- Koffi, E. N., Rayner, P. J., Scholze, M., Chevallier, F., and Kaminski, T.: Quantifying the constraint of biospheric process parameters by CO<sub>2</sub> concentration and flux measurement networks through a carbon cycle data assimilation system, *Atmos. Chem. Phys.*, 13, 10555-10572, 10.5194/acp-13-10555-2013, 2013.
- Korsakissok, I., and Mallet, V.: Subgrid-scale treatment for major point sources in an Eulerian model: A sensitivity study on the European Tracer Experiment (ETEX) and Chernobyl cases, *J. Geophys. Res.-Atmos.*, 115, 10.1029/2009JD012734, 2010a.

- Korsakissok, I., and Mallet, V.: Development and application of a reactive plume-in-grid model: Evaluation over Greater Paris, *Atmos. Chem. Phys.*, 10, 8917-8931, 10.5194/acp-10-8917-2010, 2010b.
- Kort, E. A., Frankenberg, C., Miller, C. E., and Oda, T.: Space-based observations of megacity carbon dioxide, *Geophys. Res. Lett.*, 39, 1-5, 10.1029/2012GL052738, 2012.
- Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO<sub>2</sub> emissions: Minimum site location requirements for the Los Angeles megacity, *J. Geophys. Res.-Atmos.*, 118, 1-8, 10.1002/jgrd.50135, 2013.
- Krol, M., Houweling, S., Bregman, B., Van den Broek, M., Segers, A., Van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications, *Atmos. Chem. Phys.*, 5, 417-432, 10.5194/acp-5-417-2005, 2005.
- Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC-II emission inventory; A multi-year (2003-2009) consistent high-resolution European emission inventory for air quality modelling, *Atmos. Chem. Phys.*, 14, 10963-10976, 10.5194/acp-14-10963-2014, 2014.
- Kuijpers, B.: Oxygen as a tracer for fossil fuel CO<sub>2</sub> emission sources, MSc thesis, Meteorology and Air Quality, Wageningen University, Wageningen, 2018.
- Lac, C., Donnelly, R. P., Masson, V., Pal, S., Riette, S., Donier, S., Queguiner, S., Tanguy, G., Ammoura, L., and Xueref-Remy, I.: CO<sub>2</sub> dispersion modelling over Paris region within the CO<sub>2</sub>-MEGAPARIS project, *Atmos. Chem. Phys.*, 13, 4941-4961, 10.5194/acp-13-4941-2013, 2013.
- Lacis, A. A., Schmidt, G. A., Rind, D., and Ruedy, R. A.: Atmospheric CO<sub>2</sub>: Principal control knob governing Earth's temperature, *Science*, 330, 356-359, 10.1126/science.1190653, 2010.
- LaFranchi, B. W., Pétron, G., Miller, J. B., Lehman, S. J., Andrews, A. E., Dlugokencky, E. J., Hall, B., Miller, B. R., Montzka, S. A., Neff, W., Novelli, P. C., Sweeney, C., Turnbull, J. C., Wolfe, D. E., Tans, P. P., Gurney, K. R., and Guilderson, T. P.: Constraints on emissions of carbon monoxide, methane, and a suite of hydrocarbons in the Colorado front range using observations of <sup>14</sup>CO<sub>2</sub>, *Atmos. Chem. Phys.*, 13, 11101-11120, 10.5194/acp-13-11101-2013, 2013.
- Lauvaux, T., Miles, N. L., Richardson, S. J., Deng, A., Stauffer, D. R., Davis, K. J., Jacobson, G., Rella, C., Calonder, G. P., and Decola, P. L.: Urban emissions of CO<sub>2</sub> from Davos, Switzerland: The first real-time monitoring system using an atmospheric inversion technique, *J. Appl. Meteorol. Clim.*, 52, 2654-2668, 10.1175/JAMC-D-13-038.1, 2013.
- Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO<sub>2</sub> emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), *J. Geophys. Res.-Atmos.*, 121, 5213-5236, 10.1002/2015jd024473, 2016.
- Le Mer, J., and Roger, P.: Production, oxidation, emission and consumption of methane by soils: a review, *Eur. J. Soil Biol.*, 37, 25-50, 10.1016/s1164-5563(01)01067-6, 2001.
- Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K.,

- Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S. I., Nojiri, Y., Padin, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., Van der Laan-Luijkx, I. T., Van der Werf, G. R., Van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global carbon budget 2017, *Earth Syst. Sci. Data*, 10, 405-448, 10.5194/essd-10-405-2018, 2018.
- Lee, T. R., and De Wekker, S. F. J.: Estimating daytime planetary boundary layer heights over a valley from rawinsonde observations at a nearby airport: An application to the page valley in Virginia, United States, *J. Appl. Meteorol. Climatol.*, 55, 791-809, 10.1175/jamc-d-15-0300.1, 2016.
- Levin, I., and Karstens, U.: Inferring high-resolution fossil fuel CO<sub>2</sub> records at continental sites from combined <sup>14</sup>CO<sub>2</sub> and CO observations, *Tellus B Chem. Phys. Meteorol.*, 59, 245-250, 10.1111/j.1600-0889.2006.00244.x, 2007.
- Levin, I., Hammer, S., Eichelmann, E., and Vogel, F. R.: Verification of greenhouse gas emission reductions: The prospect of atmospheric monitoring in polluted areas, *Philos. Trans. R. Soc. Lond. A*, 369, 1906-1924, 10.1098/rsta.2010.0249, 2011.
- Liander: Innovatie & Diensten: Open data, <https://www.liander.nl/over-liander/innovatie/open-data/data>, access: 22-02-2018.
- Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A. C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.-Atmos.*, 108, 10.1029/2002jd003161, 2003.
- Lindenmaier, R., Dubey, M. K., Henderson, B. G., Butterfield, Z. T., Herman, J. R., Rahn, T., and Lee, S. H.: Multiscale observations of CO<sub>2</sub>, <sup>13</sup>CO<sub>2</sub>, and pollutants at Four Corners for emission verification and attribution, in: *P. Natl. Acad. Sci. USA*, *P. Natl. Acad. Sci. USA*, 2014, 8386-8391, 2014.
- Liu, J., and Bowman, K.: A method for independent validation of surface fluxes from atmospheric inversion: Application to CO<sub>2</sub>, *Geophys. Res. Lett.*, 43, 3502-3508, 10.1002/2016gl067828, 2016.
- Liu, Y., Gruber, N., and Brunner, D.: Spatiotemporal patterns of the fossil-fuel CO<sub>2</sub> signal in central Europe: Results from a high-resolution atmospheric transport model, *Atmos. Chem. Phys.*, 17, 14145-14169, 10.5194/acp-17-14145-2017, 2017.
- Lopez, M., Schmidt, M., Delmotte, M., Colomb, A., Gros, V., Janssen, C., Lehman, S. J., Mondelain, D., Perrussel, O., Ramonet, M., Xueref-Remy, I., and Bousquet, P.: CO, NO<sub>x</sub> and <sup>13</sup>CO<sub>2</sub> as tracers for fossil fuel CO<sub>2</sub>: Results from a pilot study in Paris during winter 2010, *Atmos. Chem. Phys.*, 13, 7343-7358, 10.5194/acp-13-7343-2013, 2013.
- Mays, K. L., Shepson, P. B., Stirr, B. H., Karion, A., Sweeney, C., and Gurney, K. R.: Aircraft-based measurements of the carbon footprint of Indianapolis, *Environ. Sci. Technol.*, 43, 7816-7823, 10.1021/es901326b, 2009.
- McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, in: *P. Natl. Acad. Sci. USA*, *P. Natl. Acad. Sci. USA*, 2012, 8423-8428, 2012.
- Meesters, A. G. C. A., Tolk, L. F., Peters, W., Hutjes, R. W. A., Vellinga, O. S., Elbers, J. A., Vermeulen, A. T., Van der Laan, S., Neubert, R. E. M., Meijer, H. A. J., and Dolman, A. J.: Inverse carbon



## REFERENCES

---

- dioxide flux estimates for the Netherlands, *J. Geophys. Res.-Atmos.*, 117, 10.1029/2012JD017797, 2012.
- Meirink, J. F., Bergamaschi, P., Frankenberg, C., d'Amelio, M. T. S., Dlugokencky, E. J., Gatti, L. V., Houweling, S., Miller, J. B., Röckmann, T., Villani, M. G., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations, *J. Geophys. Res.-Atmos.*, 113, 10.1029/2007jd009740, 2008.
- Miller, J. B., Lehman, S. J., Montzka, S. A., Sweeney, C., Miller, B. R., Karion, A., Wolak, C., Dlugokencky, E. J., Southon, J., Turnbull, J. C., and Tans, P. P.: Linking emissions of fossil fuel CO<sub>2</sub> and other anthropogenic trace gases using atmospheric <sup>14</sup>CO<sub>2</sub>, *J. Geophys. Res.-Atmos.*, 117, 1-23, 10.1029/2011jd017048, 2012.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.-Atmos.*, 102, 16663-16682, 10.1029/97JD00237, 1997.
- Mues, A., Kuenen, J., Hendriks, C., Manders, A., Segers, A., Scholz, Y., Hueglin, C., Builtjes, P., and Schaap, M.: Sensitivity of air pollution simulations with LOTOS-EUROS to the temporal distribution of anthropogenic emissions, *Atmos. Chem. Phys.*, 14, 939-955, 10.5194/acp-14-939-2014, 2014.
- Nakicenovic, N., Alcamo, J., Davis, G., De Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Jung, T. Y., Kram, T., La Rovere, E. L., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., Van Rooijen, S., Victor, N., and Dadi, Z.: IPCC Special Report on Emissions Scenarios, Cambridge University Press, Cambridge, UK, 2000.
- Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., and Deng, F.: Improving the temporal and spatial distribution of CO<sub>2</sub> emissions from global fossil fuel emission data sets, *J. Geophys. Res.-Atmos.*, 118, 917-933, 10.1029/2012jd018196, 2013.
- National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999, updated daily, in, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, 2000.
- Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled weather research and forecasting-stochastic time-inverted lagrangian transport (WRF-STILT) model, *Meteorol. Atmos. Phys.*, 107, 51-64, 10.1007/s00703-010-0068-x, 2010.
- Netherlands PRTR: Netherlands Pollutant Release & Transfer Register, <http://www.emissieregistratie.nl/>, access: Mar 2014.
- Olivier, J. G. J., Brandes, L. J., and Te Molder, R. A. B.: Uncertainty in the Netherlands' greenhouse gas emissions inventory, Netherlands Environmental Assessment Agency (PBL), 68, 2009.
- Olsen, S. C., and Randerson, J. T.: Differences between surface and column atmospheric CO<sub>2</sub> and implications for carbon cycle research, *J. Geophys. Res.-Atmos.*, 109, 10.1029/2003JD003968, 2004.
- Oney, B., Henne, S., Gruber, N., Leuenberger, M., Bamberger, I., Eugster, W., and Brunner, D.: The CarboCount CH sites: Characterization of a dense greenhouse gas observation network, *Atmos. Chem. Phys.*, 15, 11147-11164, 10.5194/acp-15-11147-2015, 2015.

- Pan, X. L., Kanaya, Y., Wang, Z. F., Tang, X., Takigawa, M., Pakpong, P., Taketani, F., and Akimoto, H.: Using Bayesian optimization method and FLEXPART tracer model to evaluate CO emission in East China in springtime, *Environ. Sci. Pollut. R.*, 21, 3873-3879, 10.1007/s11356-013-2317-2, 2014.
- Peischl, J., Ryerson, T. B., Aikin, K. C., De Gouw, J. A., Gilman, J. B., Holloway, J. S., Lerner, B. M., Nadkarni, R., Neuman, J. A., Nowak, J. B., Trainer, M., Warneke, C., and Parrish, D. D.: Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions, *J. Geophys. Res.-Atmos.*, 120, 2119-2139, 10.1002/2014jd022697, 2015.
- Peters, W., Krol, M. C., Dlugokencky, E. J., Dentener, F. J., Bergamaschi, P., Dutton, G., Velthoven, P. v., Miller, J. B., Bruhwiler, L., and Tans, P. P.: Toward regional-scale modeling using the two-way nested global model TM5: Characterization of transport using SF<sub>6</sub>, *J. Geophys. Res.-Atmos.*, 109, D19314 19311-19317, 10.1029/2004jd005020, 2004.
- Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO<sub>2</sub> surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.-Atmos.*, 110, 1-18, 10.1029/2005JD006157, 2005.
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., Van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *P. Natl. Acad. Sci. USA*, 104, 18925-18930, 10.1073/pnas.0708986104, 2007.
- Peters, W., Krol, M. C., Van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., Di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., Van der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Glob. Change Biol.*, 16, 1317-1337, 10.1111/j.1365-2486.2009.02078.x, 2010.
- Picarro: Datasheet G2401 CO<sub>2</sub> + CO + CH<sub>4</sub> + H<sub>2</sub>O CRDS analyzer, [http://www.picarro.com/products\\_solutions/trace\\_gas\\_analyzers/co\\_co2\\_ch4\\_h2o](http://www.picarro.com/products_solutions/trace_gas_analyzers/co_co2_ch4_h2o), access: 23-05-2016.
- Pickers, P. A., and Manning, A. C.: Investigating bias in the application of curve fitting programs to atmospheric time series, *Atmos. Meas. Tech.*, 8, 1469-1489, 10.5194/amt-8-1469-2015, 2015.
- Popa, M. E., Vollmer, M. K., Jordan, A., Brand, W. A., Pathirana, S. L., Rothe, M., and Röckmann, T.: Vehicle emissions of greenhouse gases and related tracers from a tunnel study: CO:CO<sub>2</sub>, N<sub>2</sub>O:CO<sub>2</sub>, CH<sub>4</sub>:CO<sub>2</sub>, O<sub>2</sub>:CO<sub>2</sub> ratios, and the stable isotopes <sup>13</sup>C and <sup>18</sup>O in CO<sub>2</sub> and CO, *Atmos. Chem. Phys.*, 14, 2105-2123, 10.5194/acp-14-2105-2014, 2014.
- Pouliot, G., Pierce, T., Denier van der Gon, H. A. C., Schaap, M., Moran, M., and Nopmongkol, U.: Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, *Atmos. Environ.*, 53, 4-14, 10.1016/j.atmosenv.2011.12.041, 2012.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes in C (2nd ed.): The Art of Scientific Computing, Cambridge University Press, New York, NY, 1992.

## REFERENCES

---

- Pucher, G.: Deriving traffic-related CO<sub>2</sub> emission factors with high spatiotemporal resolution from extended floating car data, in: *Lecture Notes in Geoinformation and Cartography*, 55-68, 2017.
- Ramírez, A. R., De Keizer, C., and Van der Sluijs, J. P.: Monte Carlo analysis of uncertainties in the Netherlands Greenhouse Gas Emission Inventory for 1990 – 2004, Copernicus Institute for Sustainable Development and Innovation, Utrecht, 2006.
- Raupach, M. R., Marland, G., Ciais, P., Le Quééré, C., Canadell, J. G., Klepper, G., and Field, C. B.: Global and regional drivers of accelerating CO<sub>2</sub> emissions, *P. Natl. Acad. Sci. USA*, 104, 10288-10293, 10.1073/pnas.0700609104, 2007.
- Raupach, M. R., Rayner, P. J., and Paget, M.: Regional variations in spatial structure of nightlights, population density and fossil-fuel CO<sub>2</sub> emissions, *Energy Policy*, 38, 4756-4764, 10.1016/j.enpol.2009.08.021, 2010.
- Ray, J., Yadav, V., Michalak, A. M., Van Bloemen Waanders, B., and McKenna, S. A.: A multiresolution spatial parameterization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions, *Geosci. Model Dev.*, 7, 1901-1918, 10.5194/gmd-7-1901-2014, 2014.
- Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., and Koffi, E.: A new global gridded data set of CO<sub>2</sub> emissions from fossil fuel combustion: Methodology and evaluation, *J. Geophys. Res.-Atmos.*, 115, 10.1029/2009jd013439, 2010.
- Rayner, P. J., Utembe, S. R., and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: A theoretical study, *Atmos. Meas. Tech.*, 7, 3285-3293, 10.5194/amt-7-3285-2014, 2014.
- Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, *Atmos. Meas. Tech.*, 6, 837-860, 10.5194/amt-6-837-2013, 2013.
- Reuter, M., Buchwitz, M., Hilboll, A., Richter, A., Schneising, O., Hilker, M., Heymann, J., Bovensmann, H., and Burrows, J. P.: Decreasing emissions of NO<sub>x</sub> relative to CO<sub>2</sub> in East Asia inferred from satellite observations, *Nat. Geosci.*, 7, 792-795, 10.1038/ngeo2257, 2014.
- Ribeiro, I., Monteiro, A., and Lopes, M.: Potential effects of using biodiesel in road-traffic on air quality over the Porto urban area, Portugal, *Atmos. Environ.*, 125, 78-91, 10.1016/j.atmosenv.2015.11.006, 2016.
- Richardson, S. J., Miles, N. L., Davis, K. J., Crosson, E. R., Rella, C. W., and Andrews, A. E.: Field testing of cavity ring-down spectroscopy analyzers measuring carbon dioxide and water vapor, *J. Atmos. Ocean. Technol.*, 29, 397-406, 10.1175/jtech-d-11-00063.1, 2012.
- Riebeek, H.: The Carbon Cycle, <https://earthobservatory.nasa.gov/Features/CarbonCycle/>, access: August 3, 2017.
- Rissman, J., Arunachalam, S., Woody, M., West, J. J., Bendor, T., and Binkowski, F. S.: A plume-in-grid approach to characterize air quality impacts of aircraft emissions at the Hartsfield-Jackson Atlanta International Airport, *Atmos. Chem. Phys.*, 13, 9285-9302, 10.5194/acp-13-9285-2013, 2013.
- Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, *Atmos. Chem. Phys.*, 9, 5331-5342, 10.5194/acp-9-5331-2009, 2009.
- Royal Netherlands Meteorological Institute (KNMI): Determination of the mixing layer height from ceilometer backscatter profiles, <https://www.knmi.nl/kennis-en->

- datacentrum/achtergrond/determination-of-the-mixing-layer-height-from-ceilometer-backscatter-profiles, access: 29-04-2016.
- Sauter, F., Van Zanten, M., Van der Swaluw, E., Aben, J., De Leeuw, F., and Van Jaarsveld, H.: The OPS-model. Description of OPS 4.5.0, National Institute for Public Health and the Environment (RIVM) Bilthoven, 2016.
- Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., Prihodko, L., Suits, N., and Philpott, A.: Combined simple biosphere/Carnegie-Ames-Stanford approach terrestrial carbon cycle model, *J. Geophys. Res.-Biogeo.*, 113, 10.1029/2007jg000603, 2008.
- Schmidt, M., Lopez, M., Yver Kwok, C., Messenger, C., Ramonet, M., Wastine, B., Vuillemin, C., Truong, F., Gal, B., Parmentier, E., Cloué, O., and Ciais, P.: High-precision quasi-continuous atmospheric greenhouse gas measurements at Trainou tower (Orléans forest, France), *Atmos. Meas. Tech.*, 7, 2283-2296, 10.5194/amt-7-2283-2014, 2014.
- Schrier-Uijl, A. P., Kroon, P. S., Hendriks, D. M. D., Hensen, A., Van Huissteden, J., Berendse, F., and Veenendaal, E. M.: Agricultural peatlands: Towards a greenhouse gas sink - A synthesis of a Dutch landscape study, *Biogeosciences*, 11, 4559-4576, 10.5194/bg-11-4559-2014, 2014.
- Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Ullasz, M., Parazoo, N., Andrews, A. E., and Worthy, D. E. J.: A regional high-resolution carbon flux inversion of North America for 2004, *Biogeosciences*, 7, 1625-1644, 10.5194/bg-7-1625-2010, 2010.
- Silva, S. J., Arellano, A. F., and Worden, H. M.: Toward anthropogenic combustion emission constraints from space-based analysis of urban CO<sub>2</sub>/CO sensitivity, *Geophys. Res. Lett.*, 40, 4971-4976, 10.1002/grl.50954, 2013.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF version 3, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, CO, 2008.
- Smith, K. A., Ball, T., Conen, F., Dobbie, K. E., Massheder, J., and Rey, A.: Exchange of greenhouse gases between soil and atmosphere: Interactions of soil physical factors and biological processes, *Eur. J. Soil Sci.*, 54, 779-791, 10.1046/j.1351-0754.2003.0567.x, 2003.
- Srinivas, C. V., Hari Prasad, K. B. R. R., Naidu, C. V., Baskaran, R., and Venkatraman, B.: Sensitivity analysis of atmospheric dispersion simulations by FLEXPART to the WRF-simulated meteorological predictions in a coastal environment, *Pure Appl. Geophys.*, 173, 675-700, 10.1007/s00024-015-1104-z, 2016.
- Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO<sub>2</sub> emissions based on atmospheric inversion, *Atmos. Chem. Phys.*, 16, 14703-14726, 10.5194/acp-16-14703-2016, 2016.
- Steenefeld, G. J., Ronda, R. J., and Holtslag, A. A. M.: The challenge of forecasting the onset and development of radiation fog using mesoscale atmospheric models, *Bound.-Lay. Meteorol.*, 154, 265-289, 10.1007/s10546-014-9973-8, 2014.
- Su, H., Yang, Z. L., Niu, G. Y., and Wilson, C. R.: Parameter estimation in ensemble based snow data assimilation: A synthetic study, *Adv. Water Resour.*, 34, 407-416, 10.1016/j.advwatres.2010.12.002, 2011.
- Super, I., Vilà-Guerau de Arellano, J., and Krol, M. C.: Cumulative ozone effect on canopy stomatal resistance and the impact on boundary layer dynamics and CO<sub>2</sub> assimilation at the diurnal scale:

## REFERENCES

---

- A case study for grassland in the Netherlands, *J. Geophys. Res.-Biogeo.*, 120, 1348-1365, 10.1002/2015jg002996, 2015.
- Super, I., Denier van der Gon, H. A. C., Van der Molen, M. K., Sterk, H. A. M., Hensen, A., and Peters, W.: A multi-model approach to monitor emissions of CO<sub>2</sub> and CO from an urban-industrial complex, *Atmos. Chem. Phys.*, 17, 13297-13316, 10.5194/acp-17-13297-2017, 2017a.
- Super, I., Denier van der Gon, H. A. C., Visschedijk, A. J. H., Moerman, M. M., Chen, H., Van der Molen, M. K., and Peters, W.: Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam, *Atmos. Pollut. Res.*, 8, 174-187, 10.1016/j.apr.2016.08.008, 2017b.
- Terrenoire, E., Bessagnet, B., Rouil, L., Tognet, F., Pirovano, G., L  tinois, L., Beauchamp, M., Colette, A., Thunis, P., Amann, M., and Menut, L.: High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE, *Geosci. Model Dev.*, 8, 21-42, 10.5194/gmd-8-21-2015, 2015.
- Thoning, K. W., Tans, P. P., and Komhyr, W. D.: Atmospheric carbon dioxide at Mauna Loa Observatory. 2. Analysis of the NOAA GMCC data, 1974-1985, *J. Geophys. Res.-Atmos.*, 94, 8549-8565, 10.1029/JD094iD06p08549, 1989.
- Tohjima, Y., Kubo, M., Minejima, C., Mukai, H., Tanimoto, H., Ganshin, A., Maksyutov, S., Katsumata, K., Machida, T., and Kita, K.: Temporal changes in the emissions of CH<sub>4</sub> and CO from China estimated from CH<sub>4</sub>/CO<sub>2</sub> and CO/CO<sub>2</sub> correlations observed at Hateruma Island, *Atmos. Chem. Phys.*, 14, 1663-1677, 10.5194/acp-14-1663-2014, 2014.
- Tolias, I. C., Koutsourakis, N., Hertwig, D., Efthimiou, G. C., Venetsanos, A. G., and Bartzis, J. G.: Large Eddy Simulation study on the structure of turbulent flow in a complex city, *J. Wind Eng. Ind. Aerod.*, 177, 101-116, 10.1016/j.jweia.2018.03.017, 2018.
- Tolk, L. F., Meesters, A. G. C. A., Dolman, A. J., and Peters, W.: Modelling representation errors of atmospheric CO<sub>2</sub> mixing ratios at a regional scale, *Atmos. Chem. Phys.*, 8, 6587-6596, 10.5194/acp-8-6587-2008, 2008.
- Tolk, L. F., Peters, W., Meesters, A. G. C. A., Groenendijk, M., Vermeulen, A. T., Steeneveld, G. J., and Dolman, A. J.: Modelling regional scale surface fluxes, meteorology and CO<sub>2</sub> mixing ratios for the Cabauw tower in the Netherlands, *Biogeosciences*, 6, 2265-2280, 10.5194/bg-6-2265-2009, 2009.
- Tolk, L. F., Dolman, A. J., Meesters, A. G. C. A., and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, *Atmos. Chem. Phys.*, 11, 10349-10365, 10.5194/acp-11-10349-2011, 2011.
- Touma, J. S., Isakov, V., Ching, J., and Seigneur, C.: Air quality modeling of hazardous pollutants: Current status and future directions, *J. Air Waste Manag. Assoc.*, 56, 547-558, 10.1080/10473289.2006.10464480, 2006.
- Turnbull, J. C., Miller, J. B., Lehman, S. J., Tans, P. P., Sparks, R. J., and Southon, J.: Comparison of <sup>14</sup>CO<sub>2</sub>, CO, and SF<sub>6</sub> as tracers for recently added fossil fuel CO<sub>2</sub> in the atmosphere and implications for biological CO<sub>2</sub> exchange, *Geophys. Res. Lett.*, 33, 1-5, 10.1029/2005GL024213, 2006.
- Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., Patarasuk, R., and Razlivanov, I.: Toward quantification and source sector identification of fossil fuel CO<sub>2</sub> emissions

- from an urban area: Results from the INFLUX experiment, *J. Geophys. Res.-Atmos.*, 120, 292-312, 10.1002/2014jd022555, 2015.
- Turner, A. J., Shusterman, A. A., McDonald, B. C., Teige, V., Harley, R. A., and Cohen, R. C.: Network design for quantifying urban CO<sub>2</sub> emissions: assessing trade-offs between precision and network density, *Atmos. Chem. Phys.*, 16, 13465-13475, 10.5194/acp-16-13465-2016, 2016.
- UNFCCC: Paris Agreement, UNFCCC, 2015.
- US EPA: Air quality criteria for carbon monoxide, US Environmental Protection Agency, Office of Research and Development, Washington, DCEPA-600/B-90/045F, 1991.
- Van der Laan-Luijkx, I. T., Van der Velde, I. R., Krol, M. C., Gatti, L. V., Domingues, L. G., Correia, C. S. C., Miller, J. B., Gloor, M., Van Leeuwen, T. T., Kaiser, J. W., Wiedinmyer, C., Basu, S., Clerbaux, C., and Peters, W.: Response of the Amazon carbon balance to the 2010 drought derived with CarbonTracker South America, *Glob. Biogeochem. Cycles*, 29, 1092-1108, 10.1002/2014gb005082, 2015.
- Van der Laan-Luijkx, I. T., Van der Velde, I. R., Van der Veen, E., Tsuruta, A., Stanislawski, K., Babenhauerheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: Implementation and global carbon balance 2001–2015, *Geosci. Model Dev.*, 10, 2785-2800, 10.5194/gmd-10-2785-2017, 2017.
- Van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: Methane and nitrous oxide emissions in the Netherlands: Ambient measurements support the national inventories, *Atmos. Chem. Phys.*, 9, 9369-9379, 10.5194/acp-9-9369-2009, 2009a.
- Van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: A single gas chromatograph for accurate atmospheric mixing ratio measurements of CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, SF<sub>6</sub> and CO, *Atmos. Meas. Tech.*, 2, 549-559, 10.5194/amt-2-549-2009, 2009b.
- Van der Laan, S., Karstens, U., Neubert, R. E. M., Van der Laan-Luijkx, I. T., and Meijer, H. A. J.: Observation-based estimates of fossil fuel-derived CO<sub>2</sub> emissions in the Netherlands using  $\Delta^{14}\text{C}$ , CO and <sup>222</sup>Rn, *Tellus B Chem. Phys. Meteorol.*, 62, 389-402, 10.1111/j.1600-0889.2010.00493.x, 2010.
- Van der Laan, S., Van der Laan-Luijkx, I. T., Zimmermann, L., Conen, F., and Leuenberger, M.: Net CO<sub>2</sub> surface emissions at Bern, Switzerland inferred from ambient observations of CO<sub>2</sub>,  $\delta(\text{O}_2/\text{N}_2)$ , and <sup>222</sup>Rn using a customized radon tracer inversion, *J. Geophys. Res.-Atmos.*, 119, 1580-1591, 10.1002/2013JD020307, 2014.
- Van der Laan, S., Manohar, S., Vermeulen, A., Bosveld, F., Meijer, H., Manning, A., Van der Molen, M., and Van der Laan-Luijkx, I.: Inferring <sup>222</sup>Rn soil fluxes from ambient <sup>222</sup>Rn activity and eddy covariance measurements of CO<sub>2</sub>, *Atmos. Meas. Tech.*, 9, 5523-5533, 10.5194/amt-9-5523-2016, 2016.
- Van der Velden, N., and Smit, P.: Energiemonitor van de Nederlandse glastuinbouw 2014, LEI Wageningen UR, Wageningen, 2015.
- Van Jaarsveld, J. A.: The Operational Priority Substances model. Description and validation of OPS-Pro 4.1, National Institute for Public Health and the Environment, Bilthoven, 2004.
- Van Leuken, J. P. G., Van de Kastelee, J., Sauter, F. J., Van der Hoek, W., Heederik, D., Havelaar, A. H., and Swart, A. N.: Improved correlation of human Q fever incidence to modelled C. burnetii concentrations by means of an atmospheric dispersion model, *Int. J. Health Geogr.*, 10.1186/s12942-015-0003-y, 2015.

## REFERENCES

---

- Vardag, S. N., Gerbig, C., Janssens-Maenhout, G., and Levin, I.: Estimation of continuous anthropogenic CO<sub>2</sub>: model-based evaluation of CO<sub>2</sub>, CO,  $\delta^{13}\text{C}(\text{CO}_2)$  and  $\Delta^{14}\text{C}(\text{CO}_2)$  tracer methods *Atmos. Chem. Phys.*, 15, 12705–12729, 10.5194/acp-15-12705-2015, 2015.
- Velasco, E., Perrusquia, R., Jiménez, E., Hernández, F., Camacho, P., Rodríguez, S., Retama, A., and Molina, L. T.: Sources and sinks of carbon dioxide in a neighborhood of Mexico City, *Atmos. Environ.*, 97, 226–238, 10.1016/j.atmosenv.2014.08.018, 2014.
- VEMW: Industrie presenteert uniek plan voor 95 procent CO<sub>2</sub>-reductie door industriële vernieuwing, <https://www.vemw.nl/Nieuwsoverzicht/2017-04-10-Klimaatbeleid-vestigingsklimaat-energietransitie.aspx>, access: 2-5-2018.
- Vermeulen, A. T., Hensen, A., Popa, M. E., Van den Bulk, W. C. M., and Jongejan, P. A. C.: Greenhouse gas observations from Cabauw Tall Tower (1992–2010), *Atmos. Meas. Tech.*, 4, 617–644, 10.5194/amt-4-617-2011, 2011.
- Vilà-Guerau de Arellano, J., Gioli, B., Miglietta, F., Jonker, H. J. J., Baltink, H. K., Hutjes, R. W. A., and Holtslag, A. A. M.: Entrainment process of carbon dioxide in the atmospheric boundary layer, *J. Geophys. Res.-Atmos.*, 109, 1–15, 10.1029/2004jd004725, 2004.
- Vinken, G. C. M., Boersma, K. F., Jacob, D. J., and Meijer, E. W.: Accounting for non-linear chemistry of ship plumes in the GEOS-Chem global chemistry transport model, *Atmos. Chem. Phys.*, 11, 11707–11722, 10.5194/acp-11-11707-2011, 2011.
- Vogel, F. R., Hammer, S., Steinhof, A., Kromer, B., and Levin, I.: Implication of weekly and diurnal <sup>14</sup>C calibration on hourly estimates of CO-based fossil fuel CO<sub>2</sub> at a moderately polluted site in southwestern Germany, *Tellus B Chem. Phys. Meteorol.*, 62, 512–520, 10.1111/j.1600-0889.2010.00477.x, 2010.
- Vogel, F. R., Thiruchittampalam, B., Theloke, J., Kretschmer, R., Gerbig, C., Hammer, S., and Levin, I.: Can we evaluate a fine-grained emission model using high-resolution atmospheric transport modelling and regional fossil fuel CO<sub>2</sub> observations?, *Tellus B Chem. Phys. Meteorol.*, 65, 10.3402/tellusb.v65i0.18681, 2013.
- Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R., and Tao, S.: Potential of European <sup>14</sup>CO<sub>2</sub> observation network to estimate the fossil fuel CO<sub>2</sub> emissions via atmospheric inversions, *Atmos. Chem. Phys.*, 18, 4229–4250, 10.5194/acp-18-4229-2018, 2018.
- Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO<sub>2</sub> and CH<sub>4</sub> air monitoring sites, *Atmos. Meas. Tech.*, 6, 1217–1226, 10.5194/amt-6-1217-2013, 2013.
- Whitaker, J. S., and Hamill, T. M.: Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.*, 130, 1913–1924, 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2, 2002.
- Wicht, M., Wicht, A., and Osiska-Skotak, K.: Mapping urban porosity and roughness characteristics as a mean of defining urban ventilation corridors, *Proc. SPIE Int. Soc. Opt. Eng.*, 2017.
- Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance CO<sub>2</sub>/CH<sub>4</sub>/H<sub>2</sub>O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, *Atmos. Meas. Tech.*, 3, 1113–1128, 10.5194/amt-3-1113-2010, 2010.
- Worden, H. M., Cheng, Y., Pfister, G., Carmichael, G. R., Zhang, Q., Streets, D. G., Deeter, M., Edwards, D. P., Gille, J. C., and Worden, J. R.: Satellite-based estimates of reduced CO and CO<sub>2</sub>

- emissions due to traffic restrictions during the 2008 Beijing Olympics, *Geophys. Res. Lett.*, 39, 10.1029/2012GL052395, 2012.
- Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO<sub>2</sub> emissions?, *Atmos. Chem. Phys.*, 16, 7743-7771, 10.5194/acp-16-7743-2016, 2016.
- Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, *Geophys. Res. Lett.*, 36, 10.1029/2009gl039825, 2009.
- Zhang, Q. J., Beekmann, M., Freney, E., Sellegri, K., Pichon, J. M., Schwarzenboeck, A., Colomb, A., Bourrienne, T., Michoud, V., and Borbon, A.: Formation of secondary organic aerosol in the Paris pollution plume and its impact on surrounding regions, *Atmos. Chem. Phys.*, 15, 13973-13992, 10.5194/acp-15-13973-2015, 2015.
- Zhao, C. L., and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, *J. Geophys. Res.-Atmos.*, 111, 10.1029/2005jd006003, 2006.
- Zimnoch, M., Godłowska, J., Necki, J. M., and Rozanski, K.: Assessing surface fluxes of CO<sub>2</sub> and CH<sub>4</sub> in urban environment: A reconnaissance study in Krakow, Southern Poland, *Tellus B Chem. Phys. Meteorol.*, 62, 573-580, 10.1111/j.1600-0889.2010.00489.x, 2010.





## Dankwoord

Daar ligt ie dan, mijn proefschrift. Het resultaat van ruim vier jaar aan uitdagingen en acceptatie, doelloos ronddobberen en de weg weer vinden, huilen en lachen. Alle sterke en zwakke punten worden uitvergroot en in de schijnwerpers gezet. Niemand heeft ooit gezegd dat een PhD traject een 'walk in the park' zou zijn en dat is het ook zeker niet. Maar ik ben de uitdaging aangegaan en wat ik heb gekregen is een proces waarbij ik mijzelf op professioneel en persoonlijk vlak heb kunnen ontwikkelen. Ik kan met recht zeggen dat ik vandaag de dag een ander persoon ben dan toen ik aan dit avontuur begon. De verlegen, onzekere student is veranderd in een zelfbewuste onderzoeker. En daar ben ik ontzettend trots op! Gelukkig heb ik niet alles in mijn eentje moeten doorstaan. Tijdens de hele periode stonden er allerlei mensen voor mij klaar om samen mijn successen te vieren en mijn tegenslagen te helpen verdragen. En ik wil iedereen dan ook heel erg bedanken die onderdeel is geweest van dit spannende avontuur.

Als eerste wil ik mijn begeleiders bedanken zonder wie dit boekje niet tot stand had kunnen komen: Wouter, Michiel en Hugo. Dankzij jullie uiteenlopende persoonlijkheden en achtergronden is het project een succes geworden. Tijdens de uitvoerige discussies over welke richting het onderzoek op moest gaan kwamen altijd tal van interessante opties op tafel te liggen, van wetenschappelijk georiënteerd tot praktijkgericht. Vaak had ik na zo'n meeting een aantal dagen nodig om structuur aan te brengen en mijn gedachten te ordenen. Maar daar staat tegenover dat ik, nadat de orde was teruggekeerd, altijd barste van de energie en ideeën. Ook zat jullie feedback op mijn stukken lang niet altijd op één lijn, wat me soms wat stress opleverde. Maar dankzij die uitvoerige feedback zijn de hoofdstukken mooi en duidelijk leesbaar geworden. Bedankt daarvoor!

Michiel, voornamelijk tijdens de eerste twee jaar stond je altijd voor me klaar om te helpen als ik weer eens problemen had met WRF. Je hebt me echt op weg geholpen en richting gegeven aan de eerste periode. Het was fijn om te weten dat als ik ergens mee vast zat, je direct alles liet vallen om mijn model of script te debuggen. Vooral je enorme geduld en positieve instelling worden erg gewaardeerd. Ook al is de begeleiding daarna minder intensief geworden, jouw mental support is altijd belangrijk geweest. Jij hebt me vaak gezegd dat een goede balans tussen werk en privé ervoor zorgt dat je op je werk ook optimaal kunt presteren. Dus elke keer als ik het even niet meer zag zitten zei jij dat ik de middag vrij moest nemen om te gaan wandelen en het de volgende dag opnieuw moest proberen. En dat hielp! Tijdens het wandelen kreeg ik meestal weer inspiratie en motivatie. Bedankt!

Wouter, jij hebt juist in de tweede helft van het traject de begeleiding op je genomen. In het begin liep het af en toe wat stroef en begrepen we elkaar niet altijd. Maar naarmate de tijd vorderde leerden we elkaar beter kennen en verliep de samenwerking steeds beter. Het is erg fijn om een begeleider te hebben die je niet alleen verteld wat je niet

goed doet, maar die ook vaak benadrukt dat je juist wel goed bezig bent. Je eerlijkheid heeft mij daardoor veel inzicht gegeven in mijn sterke en zwakke punten. Je geduld bij het uitleggen (en nogmaals uitleggen, en nogmaals uitleggen...) van de principes achter data assimilatie wordt heel erg gewaardeerd, omdat ik nu eindelijk het gevoel heb dat ik weet waar ik mee bezig ben... geloof ik. Als laatste wil ik je bedanken dat je mij toegang hebt gegeven tot je enorme netwerk, waardoor ik veel contacten heb opgedaan waar ik de rest van mijn carrière voordeel van zal hebben.

Hugo, ik waardeer met name jouw enthousiasme over alles wat ik voorstelde en liet zien. Of het nou een doorbraak was met het emissiemodel of een aantal slecht leesbare figuren met de eerste resultaten. Ook je praktische instelling is voor mij van onschatbare waarde geweest. Je weet altijd feilloos de link te leggen tussen onze wetenschappelijk gefundeerde onderzoeksvragen en wat er speelt in de praktijk. Ik heb altijd een brede interesse gehad en ben graag bezig met dingen die ik beschouw als relevant voor de maatschappij. Daarom is de wetenschap dat mijn onderzoek van belang is voor de praktijk voor mij erg belangrijk. Daarnaast wil ik je bedanken dat je mij toegang hebt gegeven tot de middelen beschikbaar bij TNO: metingen, kennis, emissie data, en ga zo maar door. Deze zijn erg belangrijk geweest voor het slagen van het project.

I would also like to thank all of my colleagues that made my life at the department a real joy! Especially Arjan, for always joining me for coffee and lunch strolls around campus and for being a great friend and paranymph. Aris and Auke, thanks for being my supportive roommates for the last year! Peter, thanks for sharing your views on all topics imaginable during the coffee break. But also thanks to all other PhD's, current and past, for your joyful presence. I would also like to thank Ingrid for your support throughout my PhD. Whether I needed advice on inverse modelling, someone to proofread my paper or someone to be my paranymph, you were always willing to help. Every time I came by to ask a short question we ended up talking and laughing for a long time. Thanks a lot! And sorry to keep you from work...

Besides my supervisors and colleagues I received a lot of support from people at, amongst others, TNO (Arjo, Antoon, Jeroen, Marcel, Stijn, Richard, Sander), RIVM (Marina, Ferd), ECN (Daniëlle, Arjan), and Rijksuniversiteit Groningen (Huili and PhD's). I am very grateful for your willingness to provide me with necessary data, methods and ideas and to help me with data analysis. I would also like to thank my colleagues at LSCE. I really enjoyed my stay with you and learned a lot. It was very nice to have in-depth discussions about research methods and to end the day with a beer. The insights I gained during our discussions have guided my research. Thank you for letting me be part of your group for a few months. And finally, I would like to express my gratitude to my opponents: Albrecht Weerts, Harro Meijer, Gregoire Broquet and Julia Marshall. Thank you for your time and Gregoire and Julia for your willingness to travel to the Netherlands for the defence.

Then I would like to thank my friends. During my PhD I enjoyed playing ultimate at WAF and I played nice games with a lot of nice people. The trainings and games gave me a lot of

positive energy. So thanks to all (former) WAffies for sharing all these victories and losses! A special thanks to Jasper. Although we only manage to meet once or twice a year I really enjoy our lunches and talks. I would also like to thank Kai, Roel and Ilse. I am very grateful that we became friends and that you are always there for us. Sébastien and Fangqian, thank you for your friendship and support during the summer school. You really kept me going with your positive spirit and jokes, even when I had a difficult time. I have great memories of us lying under a truck in the streets of Paris to measure its exhaust and wandering through Valencia with Fangqian unintentionally ending up with a group of Dutch guys. I will never forget those moments!

Als laatste natuurlijk dank aan mijn familie. Papa en mama, jullie steun door dik en dun is heel erg belangrijk geweest. De wetenschap dat jullie er altijd voor mij zijn is onbetaalbaar! Het viel niet altijd mee om te praten over het project en vaak leverde mijn poging iets te vertellen alleen maar vragende gezichten op. Maar dat doet geen afbreuk aan jullie ondersteuning en het enthousiasme over die mooie figuren in mijn artikelen. Ook hebben jullie mij altijd gestimuleerd in mijn enthousiasme, ook al moesten jullie daarvoor met zware rugzakken vol stenen door de bergen sjouwen. De weekendjes weg met het gezin zijn altijd een fijne onderbreking van de dagelijkse sleur en ik heb er erg van genoten. Ik hoop dat we dat nog lang blijven volhouden! En dan als allerlaatste, maar zeker niet de onbelangrijkste persoon in mijn leven, Sabina. Als mijn lieve, trotse zus en beste vriendin ben je onmisbaar. Ook al kan ik soms lastig en koppig zijn, jij staat altijd voor mij klaar met goede adviezen, een ondersteunende knuffel of een noodzakelijke schop onder mijn kont. Qua inhoudelijke discussies over ons werk hebben we misschien niet veel aan elkaar gehad, maar qua mentale ondersteuning des te meer. Als twee PhD-zusjes weten we precies wat de ander doormaakt en voelen we elkaar goed aan. Dat heeft me enorm gesteund in de moeilijke periodes. Natuurlijk moet ik ook de goede periodes niet vergeten, waarin we allerlei leuke activiteiten hebben ondernomen die ervoor hebben gezorgd dat ik vol energie weer aan het werk kon. Papa, mama en Sabina, dank voor al jullie steun en positivisme!!!



## List of publications

**Super, I.**, Vilà-Guerau de Arellano, J., and Krol, M. C.: Cumulative ozone effect on canopy stomatal resistance and the impact on boundary layer dynamics and CO<sub>2</sub> assimilation at the diurnal scale: A case study for grassland in the Netherlands, *J. Geophys. Res.-Biogeo.*, 120, 1348-1365, 10.1002/2015jg002996, 2015.

**Super, I.**, Denier van der Gon, H. A. C., Visschedijk, A. J. H., Moerman, M. M., Chen, H., Van der Molen, M. K., and Peters, W.: Interpreting continuous in-situ observations of carbon dioxide and carbon monoxide in the urban port area of Rotterdam, *Atmos. Pollut. Res.*, 8, 174–187, 10.1016/j.apr.2016.08.008, 2017.

**Super, I.**, Denier van der Gon, H. A. C., Van der Molen, M. K., Sterk, H. A. M., Hensen, A., and Peters, W.: A multi-model approach to monitor emissions of CO<sub>2</sub> and CO from an urban–industrial complex, *Atmos. Chem. Phys.*, 17, 13297-13316, 10.5194/acp-17-13297-2017, 2017.



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

# D I P L O M A

*For specialised PhD training*

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

***Ingrid Super***

born on 13 January 1990 in Emmen, The Netherlands

has successfully fulfilled all requirements of the  
Educational Programme of SENSE.

Wageningen, 11 October 2018

On behalf of the SENSE board

Prof. dr. Huub Rijnaarts

the SENSE Director of Education

Dr. Ad van Dommelen

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



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



The SENSE Research School declares that **Ingrid Super** has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 43.5 EC, including the following activities:

#### SENSE PhD Courses

- o Uncertainty propagation in spatial and environmental modelling (2014)
- o Environmental research in context (2014)
- o Research in context activity: 'Co-organizing and coordinating follow-up for workshop on: Regional air quality modelling with WRF-Chem: successes, opportunities and challenges, 15-16 September 2016, Wageningen'

#### Other PhD and Advanced MSc Courses

- o Climate-KIC summer school, The Journey (2014)
- o Scientific publishing, Wageningen University (2014)
- o Essentials of scientific writing and presenting, Wageningen University (2015)
- o Career orientation, Wageningen University (2016)
- o Business Development Course: Create a business from your research in water technology, Climate-KIC, Wetsus & Water Campus (2017)

#### External training at a foreign research institute

- o Data assimilation, Centre Européen de recherche et de formation avancée en calcul scientifique (CERFACS), France (2017)

#### Management and Didactic Skills Training

- o Teaching of practicals in the BSc course 'Integration course soil, water, atmosphere' (2014-2015)
- o Teaching of practicals in the BSc course 'Chemical processes in soil, water and atmosphere' (2014-2016)

#### Oral Presentations

- o *A multi-model approach to constrain emissions from an urban-industrial complex.* EGU General Assembly, 18-20 April 2016, Vienna, Austria
- o *Monitoring fossil fuel CO<sub>2</sub> emissions from an urban-industrial complex.* Buys Ballot Research School (BBOS) Autumn symposium, 26-28 October 2016, Oranjestad, The Netherlands

SENSE Coordinator PhD Education

Dr. Peter Vermeulen

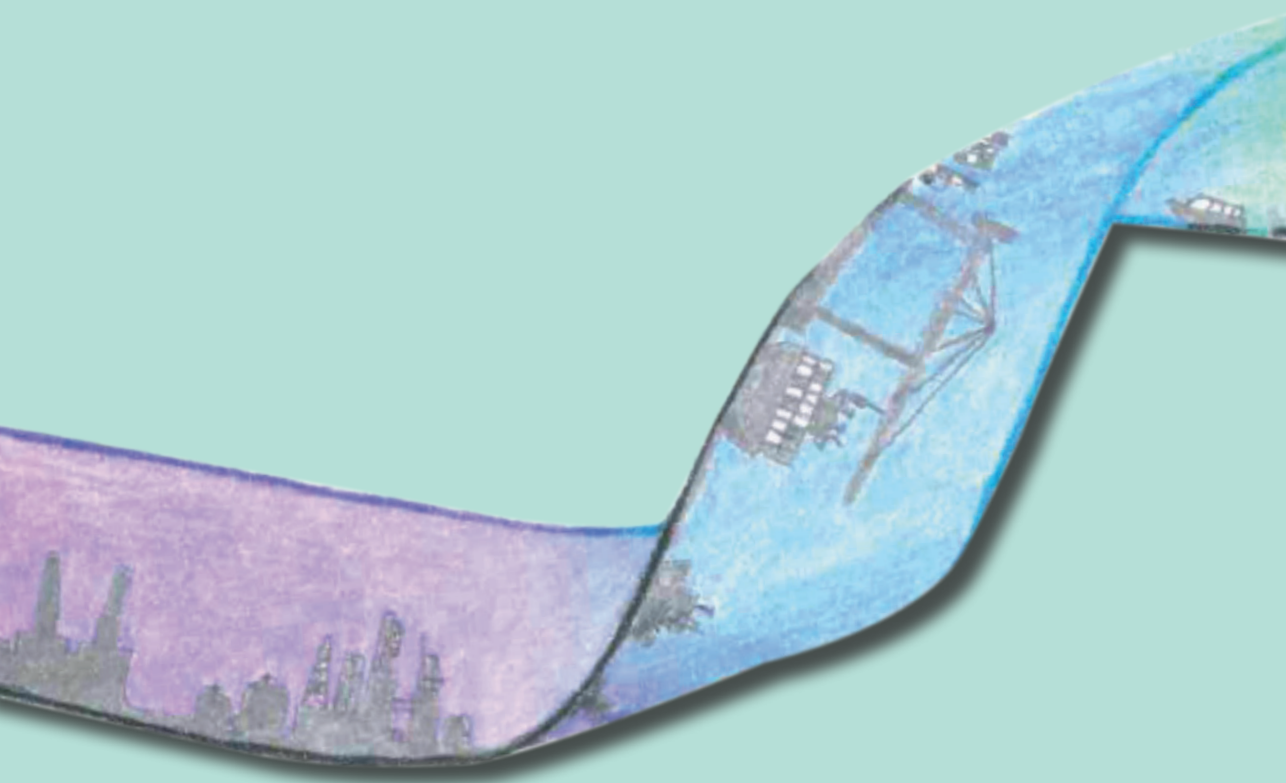


The research described in this thesis was financially supported by Climate-KIC. Climate-KIC is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Ingrid Super and ProefschriftMaken

Printed by DigiForce | | ProefschriftMaken, Vianen



## **Propositions**

1. Regional atmospheric transport models currently suffer from too large errors to be applicable for urban scale monitoring.  
(*this thesis*)
2. Atmospheric observations are needed to verify the properties of an urban emission landscape.  
(*this thesis*)
3. Modelling and coding requires an amount of patience and accuracy more abundant among monks than among scientists.
4. Considering greenhouse gases and air pollutants concurrently is necessary to push climate action forward.
5. The effort needed to build a strong foundation for model-based research is undervalued by the current system of publications and citations.
6. The moral obligation of the Netherlands to be frontrunner in the field of climate change mitigation is in conflict with the country's economic interest.
7. A challenging project not only supports a PhD in her scientific growth, but also in her personal growth.

Propositions belonging to the thesis, entitled:

Quantification and attribution of urban fossil fuel emissions through atmospheric observations

Ingrid Super

Wageningen, 11 October 2018