Karlsruher Institut für Technologie

# Estimating the Carbon Dioxide Emissions of Thessaloniki by Combining Observations and Simulations Based on an Inventory Rescaling Approach

Zur Erlangung des akademischen Grades einer

DOKTORIN DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Lena Feld, M. Sc.
aus Heidelberg

Tag der mündlichen Prüfung: 15. November 2024
Referent: Prof. Dr. Peter Braesicke
Korreferent: Priv.-Doz. Dr. Frank Hase
Betreuer: Dr. Roland Ruhnke

# Abstract

Our planet is warming at a rapid pace. The increase in global surface temperature is driven by the growing concentration of greenhouse gases in our atmosphere, primarily due to the extensive use of fossil fuels. Currently, the most important greenhouse gas that is emitted from human activities is carbon dioxide ($CO_2$). Climate change is already having a dramatic impact on our environment, and the effects are projected to intensify as emissions continue. To mitigate the worst effects of climate change, targeted action is needed to achieve a rapid reduction in the emissions of all greenhouse gases, not only $CO_2$. Precise monitoring methods are needed to improve current emission inventories for informed decision-making.

However, the currently available emission inventories contain considerable discrepancies. Especially at smaller scales, such as urban areas, the uncertainties are high. Concurrently, urban areas are major sources of anthropogenic greenhouse gases. Existing inventories are constructed using a bottom-up approach based on reported emission activities. Top-down estimation of emissions is an alternative approach, where emissions are calculated from atmospheric observations. The method can be used to validate and improve bottom-up estimates. It has been applied in several cases, mostly covering global to regional scales. Top-down estimation of urban emissions is challenging because of the dense and heterogeneous emission structure, and the difficulty to perform accurate simulations of atmospheric transport at small scales.

In this work, I perform a measurement-based estimation of the $CO_2$ emissions of the urban area of Thessaloniki, Greece. For Thessaloniki, the discrepancies between the bottom-up inventories are substantial: For 2019, the EDGAR inventory reports urban emissions of 3.1 Mt, which is significantly higher than the estimates by the CAMS (1.7 Mt) and ODIAC (1.8 Mt) inventories.

For the emission estimation, a measurement campaign was conducted in collaboration with partners from the Aristotle University of Thessaloniki. Two solar Fourier Transform Infrared spectrometers of the type EM27/SUN were operated over a period of three months, covering three weeks in October 2021 and ten weeks between May and July 2022. The EM27/SUN can record time series of the column-averaged dry-air molar fraction of $CO_2$ ($XCO_2$) with high precision and temporal resolution. One spectrometer was positioned centrally within the city. The second spectrometer was equipped with solar power supply for portability, enabling positioning at various locations. The collected dataset contains 179 hours of observation, excluding the observations for calibration.

To interpret the recorded dataset, a corresponding simulation of the emission sources and transport of $CO_2$ is required. The ICON-ART model – an operational model at the German Weather Service (DWD) – was used for this purpose. The ODIAC inventory was used as a starting point, and within the city it was separated into different sub-areas to allow scaling of the emissions during post-processing. An estimate of the net ecosystem exchange was constructed from different available datasets. The simulations show a good agreement when compared to wind and pressure observations in the city and water vapor columns co-observed by the EM27/SUN. However, the agreement is poor for $XCO_2$. The Pearson correlation coefficient between the simulated and observed time series of $XCO_2$ was only 0.1.

The agreement is enhanced significantly by rescaling the anthropogenic emissions inside the city area. To find an optimal scaling, a least-square approach is applied. Two different methods are compared: Firstly, the whole city is uniformly scaled, leaving just one degree of freedom. Secondly, the previously separated source regions of the city center are scaled individually, giving 30 degrees of freedom. Furthermore, two subsamples with good prior agreement are selected to evaluate the robustness of the results. The optimization leads to a significantly improved agreement between simulated and observed time series for both subsamples. The largest improvement is found in the smaller sample, where 5 days were selected from the full time series. Here, the correlation coefficient improves from 0.34 to 0.77. Still, discrepancies remain in the time series for all configurations. Possible reasons for this include the inaccurate representation of the biogenic sinks and sources, imperfect simulation of tracer transport and limitations of the model setup such as short simulation time and limited spatial resolution.

Despite the remaining discrepancies, the different scaling configurations show robust results: When looking at the configurations where all source regions are scaled individually, expected emission hot spots consistently receive higher weights, supporting the correct attribution of emissions by the optimization approach. For every configuration, the rescaling results in a distinct increase of Thessaloniki's total emissions. The estimates range from 2.9 to 4.4 Mt/yr. This indicates an underestimation of the actual emission in the ODIAC and CAMS inventories and supports the higher estimate from the EDGAR inventory. The results demonstrate the potential of measurement-based methods to enhance our knowledge about urban scale emissions.

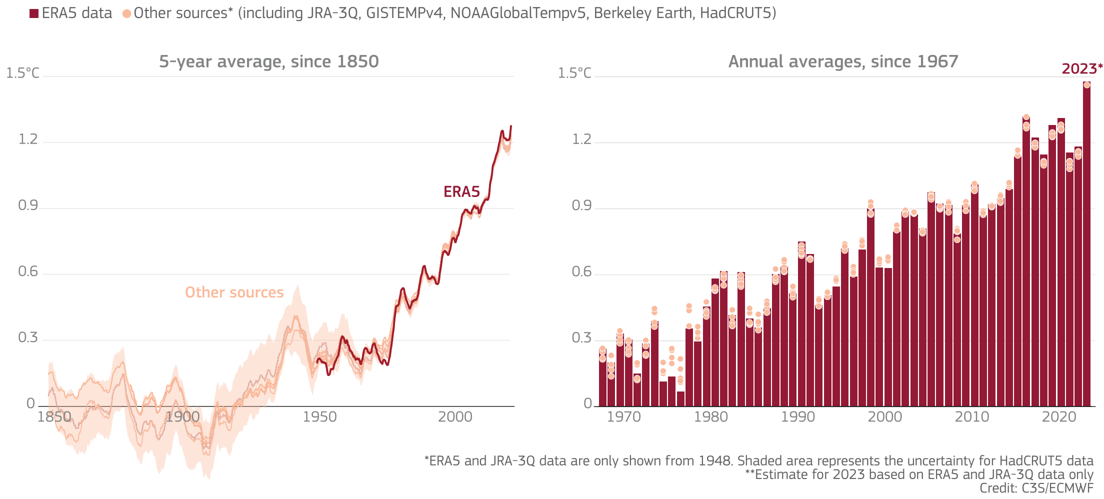
# Contents

# 1. Introduction

## 1.1. Anthropogenic Climate Change

We are facing a rapidly warming planet. The global surface temperature has risen significantly compared to the pre-industrial level (compare Figure 1.1). Last year – 2023 – was the warmest calendar year ever recorded in the global temperature records going back to 1850 (Copernicus Climate Change Service, 2024). The global mean surface temperature was 14.98°C, which is 1.48 °C warmer compared to the average of the pre-industrial reference period from 1850 to 1900.

The global warming is driven by the increase of greenhouse gases (GHGs) in the atmosphere. Carbon dioxide ($CO_2$) is the most important anthropogenic GHG. The increase of the atmospheric abundance of $CO_2$ is continuously recorded at the Mauna Loa observatory since 1958. During this period the mixing ratio increased from below 320 ppmv at



*ERA5 and JRA-3Q data are only shown from 1948. Shaded area represents the uncertainty for HadCRUT5 data
**Estimate for 2023 based on ERA5 and JRA-3Q data only
Credit: C3S/ECMWF

**Figure 1.1.: Global surface temperature: Increase above pre-industrial level (1850-1900).** Global surface air temperature (°C) increase above the average for 1850-1900, the designated pre-industrial reference period, based on several global temperature datasets shown as 5-year averages since 1850 (left) and as annual averages since 1967 (right). The figure was adapted from Copernicus Climate Change Service (2024).

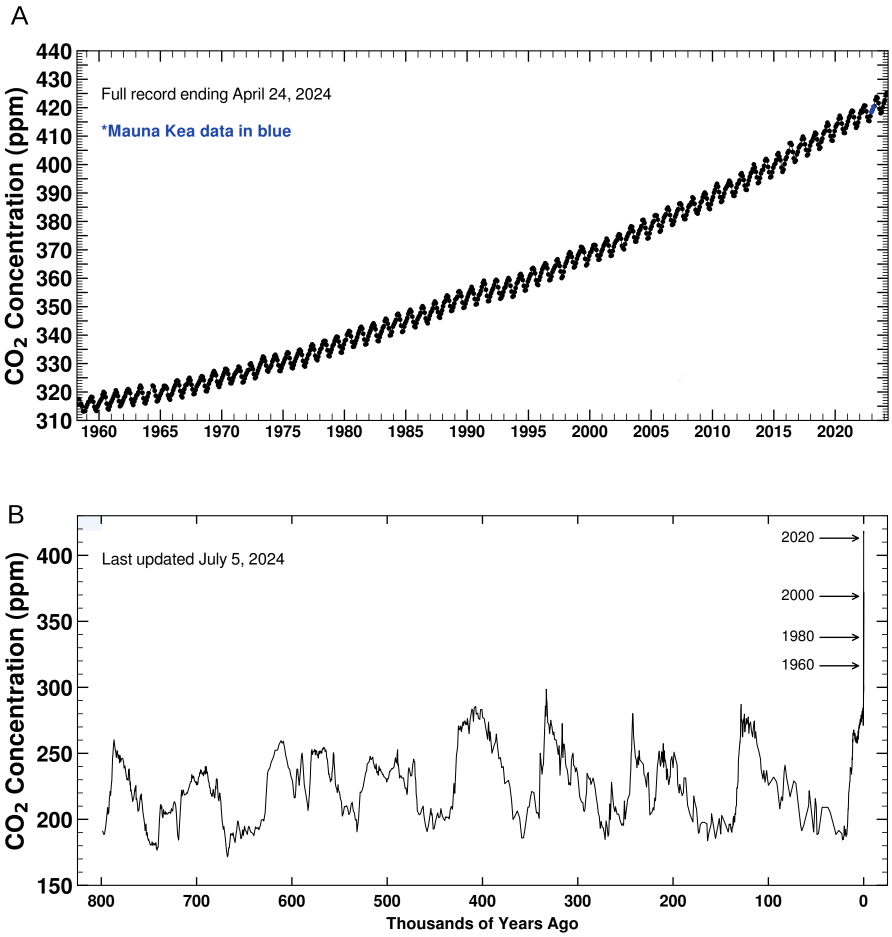


**Figure 1.2.: Rapid increase of carbon dioxide in our atmosphere.** Figure A shows the recorded data collected at the Mauna Loa observatory since 1958. Figure B additionally shows an estimate of the concentration for the last 800 thousand years, that was collected by investigation of ice cores. The figure was adapted from Monroe (2024), with data sources from Keeling and Keeling (2017), Rubino et al. (2019) and Lüthi et al. (2008).

the beginning of the recording to over 420 ppmv in 2024 (Keeling and Keeling, 2017), as displayed in Figure 1.2. This increase exceeds the average mixing ratio over the previous 800,000 years, as determined from measurements made from ice-core measurements (Rubino et al., 2019; Lüthi et al., 2008).

The mechanism that induces the warming is called greenhouse effect. The basic principle is that GHGs absorb light in the infrared range while being transparent in the visible range. This shifts the Earth’s equilibrium of incoming and outgoing radiation, as the outgoing radiation has a larger proportion of infrared light. Therefore an increase in GHGs leads to a net warming.

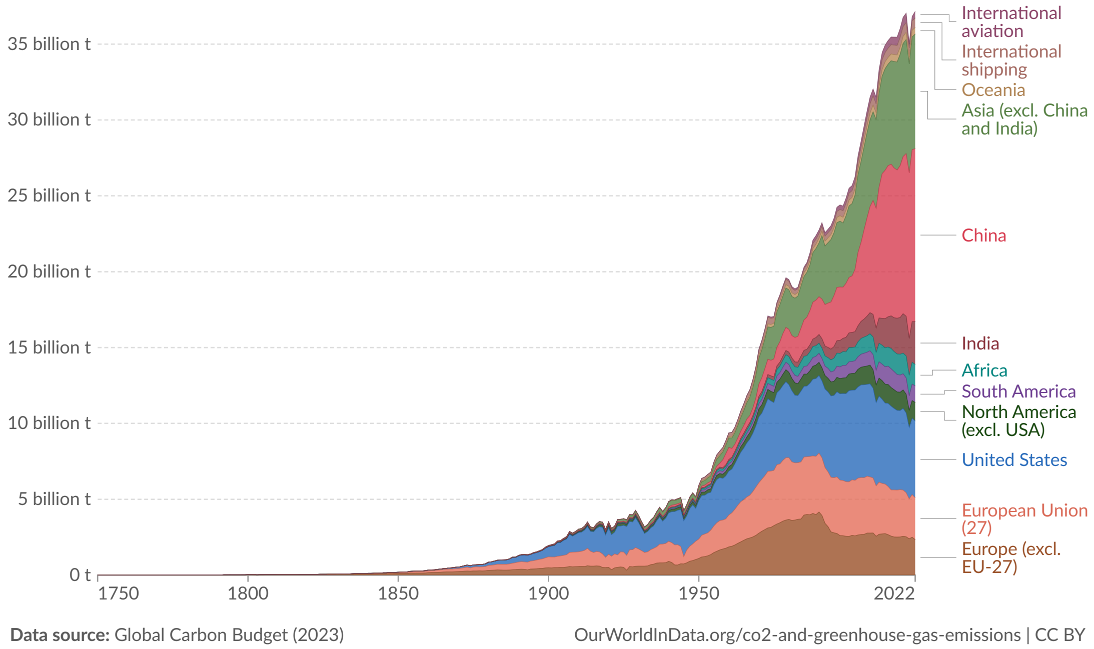


**Figure 1.3.: Annual $CO_2$ emissions by world region.** Emissions from fossil fuels and industry are included, but not land-use change emissions. International aviation and shipping are included as separate entities, as they are not included in any country's emissions. The figure is created by ourworldindata.org (Ritchie and Roser, 2024), with Data from the Global Carbon Budget (Friedlingstein et al., 2023).

The increase of GHGs in the atmosphere is caused by the extensive exploit of fossil carbon deposits since the industrialization. Figure 1.3 shows the global annual $CO_2$ emissions from fossil fuel burning since 1750, with the rate of increase being approximately exponential. The continuing emissions of GHGs will lead to a further warming. This has substantial effects on our environment. Glaciers and polar ocean ice sheet extents are shrinking significantly. For the period between 1979 and 2019, the Intergovernmental Panel on Climate Change (IPCC) reported a reduction of the arctic sea ice extent from 6.23 to 3.76 million $km^2$ for the month of September (IPCC, 2023b, p. 343). The melting of glaciers affects the human societies by altering water supplies and inducing sea level rise.

Climate change is leading to an increase in both the intensity and the frequency of extreme weather events (IPCC, 2023c). The effects are expected to intensify in the future, but can already be seen today. In recent years, amplified wildfire seasons, prolonged droughts and flooding events caused damage and death in unprecedented extents. Recent examples are the record-breaking 2023 wildfire season in Canada (Sottile et al., 2023) and the Mediterranean cyclone Daniel (Fleming, 2023) that caused thousands of deaths. Zachariah et al. (2023) found that an event like cyclone Daniel became 50 times more likely to occur and about 50% more intense over Libya in a 1.2 °C hotter climate.

Future impacts are projected to become more severe compared to present time. The extent of the increase depends on the amount of GHG emission reduction we achieve. In the report of Working Group 2 about the impacts, adaptation and vulnerability, the IPCC (2023a) stated that projected impacts of a continued warming induce an increased impact on ecosystems and biodiversity, food and water supply, and coastal communities through sea level rise. In a high-emission scenario, they project over nine million climate-related deaths per year by the end of the century. The projected fatalities are caused by an increase in heat-related mortality and undernutrition, a spread of transmission regions for vector-borne diseases as Malaria or Dengue fever, and an increase in diseases as diarrhea and cholera. "Under all global warming levels, some regions that are presently densely populated will become unsafe or uninhabitable" (IPCC, 2023a, p.64).

To mitigate the hazards caused by climate change, the global community agreed to limit the global warming to 1.5°C in the Paris Agreement in 2015 (UNFCCC, 2016). Despite 195 parties signed the treaty (United Nations Treaty Collection, 2024), the global emissions are still increasing. In 2023 a record of 37.55 billion tons of fossil fuels have been emitted (Tiseo, 2024). A turning point in this increase is urgently needed, as rapid decrease is mandatory to mitigate climate change.

## 1.2. Stocktaking of Greenhouse Gas Emissions

Limiting global warming requires a rapid reduction of GHG emissions. A reliable stocktaking of current GHG emissions and their evolution over time is needed for informed decision-making and to enforce compliance. Here, the activities to monitor GHG emissions are introduced.

### Reported Emissions and Inventories

The parties of the Paris Agreement agreed to establish a global stock take (GST) every five years, starting in 2023, to monitor the reduction of the emissions. The implications of the first GST were negotiated in December 2023 at the 28th Conference of the Parties (COP28), and the parties can submit their updated nationally determined contributions (NDCs) until December 2025. In the GST, the parties submit their emissions, mitigation and adaptation efforts to the United Nations Framework Convention on Climate Change (UNFCCC) (see UNFCCC (2016)). In the outcome of the first GST, it was recognized that the potential warming could be decreased to a range between 2.1 to 2.8 °C if all current climate action plans are implemented. But there is still a mitigation gap, as the world's emissions today are not in line with the target to limit the warming to 1.5 °C (UNFCCC, 2024). Industrialized countries committed to reporting their annual GHG emissions to the UNFCCC on a more frequent basis, covering the period from the base year 1990 to the present. Other countries can submit their reports less often depending on the funds they receive for their reporting (Granier et al., 2023) and their records do not go back to 1990. The reports to UNFCCC are prepared in written format.

Janssens-Maenhout et al. (2019) state that the data collected by UNFCCC does not provide a complete, consistent and comparable global dataset. For this reason, other emission inventories have been created. Many of these inventories use a gridded format to ensure the emission information more accessible to modeling communities. The following introduction of these inventories is based on Granier et al. (2023).

The data sources and methodologies of the inventories vary in detail, but all are based on a bottom-up (BU) concept. It is based on the following two components:

1. Activity data, quantifying the human activity (e.g. the amount of coal burned),
2. Emission factors – a linear factor describing the amount of a pollutant emitted by the respective activity (e.g. the amount of $CO_2$ emitted per unit mass of burned coal).

The activity data is combined with the emission factors to yield emission estimates. In the following, the inventories that are relevant for this work are introduced briefly.

- The **Emissions Database for Global Atmospheric Research (EDGAR)** (Janssens-Maenhout et al., 2019; Crippa et al., 2018) provides global gridded emissions with a spatial resolution of 0.1° x 0.1° and a monthly temporal resolution. The newest version, EDGARv8.0 (Crippa et al., 2022), encompasses the period from 1970 until 2022. The emissions are split into different sectors. They combine the data reported to UNFCCC with various other sources. For example the combustion-related activity factors are derived from the World Energy Balances of the International Energy Agency (IEA, 2024).
- The **Community Emissions Data System (CEDS)** (Hoesly et al., 2018) is an open-source framework with the goal to provide a sectoral gridded inventory in support of the Coupled Model Intercomparison Project (CMIP, 2023). For enabling pre-industrial runs, CEDS provides data since 1750. For this approach, activity data and emission factors are first collected and calibrated afterward to match with existing country-level inventories. Among others, CEDS also relies on the estimates of EDGAR for the default emission estimates. The newest version that is

provided in a gridded format is v_2021_04_21. It has a resolution of 0.1 ° x 0.1 ° and includes data up to 2019 (Ahsan et al., 2022).

- The gridded emission inventory by the **Copernicus Atmospheric Monitoring Service (CAMS)** (Granier et al., 2019) is used for their model predictions and reanalysis products. The CAMS inventory is based on emissions from EDGAR and CEDS. The anthropogenic emission inventory has a resolution of 0.1 ° x 0.1 ° and ranges from 2000 - 2020 with a monthly temporal resolution. It provides many different species of pollutants and GHGs.
- The **Carbon Dioxide Information and Analysis Center (CDIAC)** provides an independent long-time series of $CO_2$ emissions going back to 1751 (Gilfillan and Marland, 2021). Other than the CAMS and CDES inventories, CDIAC is a primary inventory that only relies on emission data and does not depend on data from other inventories. CDIAC is not provided in a gridded format.
- The **Open-Source Data Inventory for Anthropogenic $CO_2$ (ODIAC)** (Oda and Maksyutov, 2011; Oda et al., 2018) only provides $CO_2$ emissions, but it has an exceptionally high spatial resolution of 1 km. ODIAC relies on the emissions of CDIAC such that emissions from CDIAC are combined with satellite images of anthropogenic light emission observed at night as a proxy for spatial disaggregation. The newest version – ODIAC2022 – encompasses the time period from 2000 to 2021.

The global annual emission estimates have little discrepancies between the different inventories. Janssens-Maenhout et al. (2019) compares the total $CO_2$ emission from some of the above-mentioned inventories for the year 2010. They report total emissions of 30.5 ± 5.3 Pg for the EDGAR inventory with a confidence level of 95 % for fossil fuel combustion. The inventories ODIAC (33.4 Pg) and CDIAC (32.7 Pg) are within the uncertainty limit. However, there are large regional differences between all the above inventories, even though they are partly dependent on each other. While the carbon dioxide emissions are much better known than many other GHGs and pollutants (Maksyutov et al., 2022), the uncertainties are still significant, especially when looking at more localized sources.

### Emission Estimates from Atmospheric Observations

A way to improve the knowledge about emissions are top-down (TD) estimates, complementing the discussed BU inventories. In TD estimation, the emissions are calculated from atmospheric measurements, with the advantage that the observations are completely independent of the reporting. No emission source can be double-counted or missed.

In recent years the interest in TD approaches has increased. Manning et al. (2011) performed an inventory-independent TD estimation of the methane ($CH_4$) and nitrous oxide ($N_2O$) emission inventory of the United Kingdom, finding a lower reduction in $CH_4$ emissions than stated in the BU estimation due to the poor agreement in the 1990s. Switzerland added TD estimates of $N_2O$ and $CH_4$ in the appendix to the submission for UNFCCC (FOEN, 2024). The inversion approach that was applied for the TD estimation is described by Henne et al. (2016). The government of the United States of America formulated the objective to improve the atmosphere-based TD estimates in a strategy paper, naming concrete fields for improvements (The White House, 2023). The German government initialized the Integrated Greenhouse Gas Monitoring System (ITMS, for Integriertes Treibhausgas Monitoring System) to build up an operational system for TD emission monitoring in Germany (ITMS Germany, 2024). After the Global Methane Pledge at the COP26, The UN Environment Program (UNEP) implemented the International Methane Observatory (IMEO), an initiative with the goal to provide measurement-based emission data to catalyze reduction measures (see United Nations Environment Programme and International Methane Emissions Observatory (2023) for their 2023 report).

Many of the above-mentioned TD emission estimates or initiatives focus on non-$CO_2$ GHGs. This has several reasons. The reported emissions often have significantly larger uncertainties compared to the emissions due to fossil fuel burning. For example, $N_2O$ and $CH_4$ are released from agricultural soils, landfills or wastewater treatment plants, through microbiological processes. These processes are dependent on environmental conditions and, therefore, difficult to be quantified in an inventory (Maksyutov et al., 2022). For this reason, some TD analyses of $CO_2$ targeted the interactions with the biosphere and the ocean, assuming the fossil fuel emissions to be correct (see Peters et al. (2007); Friedlingstein et al. (2023)).

### Carbon Dioxide Emissions from Urban Areas

While many of the aforementioned activities focus on $N_2O$ or $CH_4$ and address global to regional scales, $CO_2$ emissions in cities are of special interest, as a large fraction of the anthropogenic $CO_2$ emissions are produced in urban areas. Crippa et al. (2021) estimated that a third of the global emissions were produced in urban areas in 2015, a fraction that has increased over the past decades. As mentioned in the previous chapter, the uncertainties of the global emissions are not particularly high. Andres et al. (2014) estimate an uncertainty of 8.4 % for the global $CO_2$ emissions from fossil fuel based on the CDIAC inventory. Solazzo et al. (2021) concluded a similar uncertainty of 7 % for the global estimate based on EDGAR, but when looking at individual pixels of the emission map, they found uncertainties up to 40 % for a central European domain. Urban areas are often represented by only one or a few pixels at the common resolution of the inventories, so these high local uncertainties are expected to apply to them as well.

Coincidently, cities are challenging to monitor because of their heterogeneous structure and emissions from a variety of sources. One of the major challenges of the TD emission estimation is that the observations of atmospheric concentrations are connected to the emissions in a non-trivial dependency. To obtain emission estimates, normally a complex meteorological model is used to calculate the transport of the emissions to the point of observation. Often the methods include a BU inventory as a prior, altering it with help of the observations. This puts challenges for the application to urban-scale studies, as many of the BU inventories do not have a sufficient resolution and the accurate simulation of atmospheric transport is also challenging at this scale.

In summary, while the urban emissions are particularly important, their contributions remain quite uncertain in the BU inventories. Simultaneously, the urban scale is posing challenges to TD emission estimation due to the heterogenic sources and small-scale transport. Through the HEPTA Project, a collaborative initiative between the Aristotle University of Thessaloniki (AUTh) and the Karlsruhe Institute of Technology (KIT), I had the opportunity to investigate Thessaloniki's emissions. In this work, I perform a TD estimate of the emissions of the urban area of Thessaloniki, Greece. The objectives of this work and a general overview of the text at hand are given in the following section.

## 1.3. Objective and Research Questions

The emissions of Thessaloniki are an example of the uncertainty of current emission inventories at the urban scale. For this purpose four of the previously introduced emission inventories, the CAMS (v4.2), CEDS (v-2021-04-21), ODIAC (ODIAC2022) and EDGAR (v8.0) inventories, are compared (see Figure 1.4). A reference area around Thessaloniki (40.5 to 40.8° N and 22.7 to 23.1° E) is defined for comparison. The total emission estimates of the different inventories in this area deviate significantly from each other. While the CAMS and ODIAC inventories report lower total emission estimates of 1.7 Mt/yr and 1.8 Mt/yr, the EDGAR and CEDS inventories report 3.1 Mt/yr and 3.2 Mt/yr. From these estimates, the EDGAR inventory reports 82 % more than stated by CAMS. This exceeds

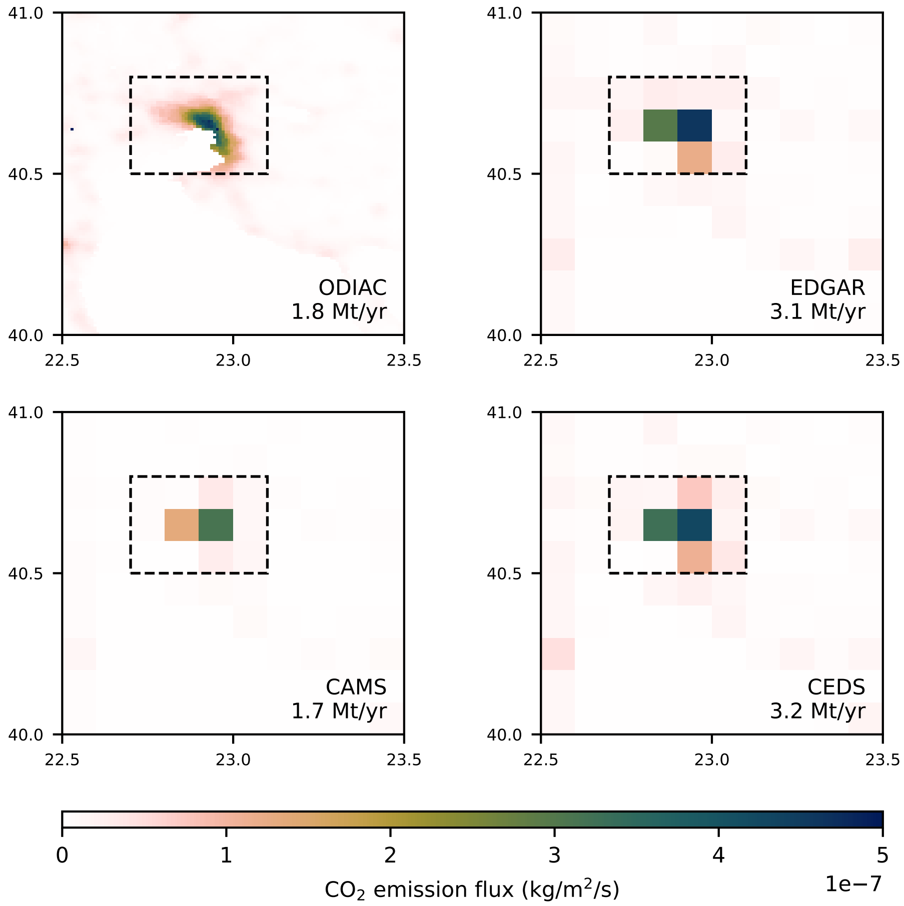


**Figure 1.4.: Comparing different emission inventories for Thessaloniki.** The average annual emissions in 2019 for the urban area of Thessaloniki are shown for the ODIAC, EDGAR, CAMS and CEDS inventories. The EDGAR, CAMS and CEDS inventories are spatially more aggregated than the ODIAC inventory. The total emissions of the urban area inside the dashed rectangle are displayed in the lower right corner of each frame.

the uncertainties of the EDGAR inventory for singular pixels of 40 % stated by Solazzo et al. (2021). No uncertainties are provided for the other inventories.

These reported discrepancies motivate the scientific objective to obtain a measurement-based emission estimate for the city of Thessaloniki. Furthermore, the deduction of the location of emission hot-spots from the observational data is tested. With this, the potential of atmospheric measurements combined with complex atmospheric tracer simulations for emission estimation at the urban scale is demonstrated.

To achieve this, a measurement campaign in Thessaloniki, Greece, has been performed using solar Fourier Transform Infrared (FTIR) spectrometers of the type EM27/SUN. Several studies have used solar FTIR observations for this purpose. In the context of the Collaborative Carbon Column Observing Network (COCCON), a framework for

EM27/SUN operation, several urban measurement campaigns have been conducted in the past (see Section 2.2.4). COCCON is a measurement framework that defines instrumental and data analysis standards (refer to Frey et al. (2019); Alberti et al. (2022) for further details). To estimate emissions from the observation, the recorded data is combined with corresponding simulations from the numerical weather prediction model ICON-ART.

An introduction to the principles of solar FTIR observations, the concept of emission estimation based on atmospheric observations, and an introduction to the structure of the ICON-ART model are given in Chapter 2. Chapter 3 introduces the implementation of the measurement campaign and gives an overview over the recorded dataset. The corresponding simulations are presented in Chapter 4, where the simulation setup is described and the quality of the simulated wind and water vapor are assessed. An emission estimate is derived from the joint interpretation of the observational dataset and the corresponding simulations in Chapter 5. The methods and results are summarized in Chapter 6.

# 2. Theoretical and Technical Background

A general understanding of the atmosphere is the basis to follow the methods used in this work. It is also necessary for explaining the underlying motivation: The problem of a changing climate due to greenhouse gas (GHG) emissions. A basic description of the atmosphere's properties is given in Section 2.1

The concepts of Fourier Transform Infrared (FTIR) spectroscopy will be explained in Section 2.2, covering the fundamental principles and the specific implementation in the Collaborative Carbon Column Observing Network (COCCON).

The atmospheric $CO_2$ column abundances that are measured by this method are affected by emissions (the subject of this work), but also by atmospheric transport. The problem of inverse emission estimation, the processes of atmospheric transport and how they can be captured by models is introduced in Section 2.3.

## 2.1. Properties of the Atmosphere

First, the general structure of the atmosphere will be explained (Section 2.1.1). This is followed by a description of the radiative properties of the atmosphere in Section 2.1.2, which are fundamental to both the greenhouse effect and the measurement principle.

### 2.1.1. Structure and Composition of the Atmosphere

The atmosphere is a thin gaseous layer surrounding the condensed part of the earth. The total mass of the atmosphere is $5.2{\cdot}10^{18}$ kg (Möller, 2019, p. 23), which is less than 0.0001% of the solid and liquid parts of earth with a total of $6.0{\cdot}10^{24}$kg (Williams, 2024). Along the vertical, the atmosphere consists of different layers, characterized by their behavior in transport, constituents, temperature and pressure. A schematic view of the different layers of the atmosphere can be seen in Figure 2.1.

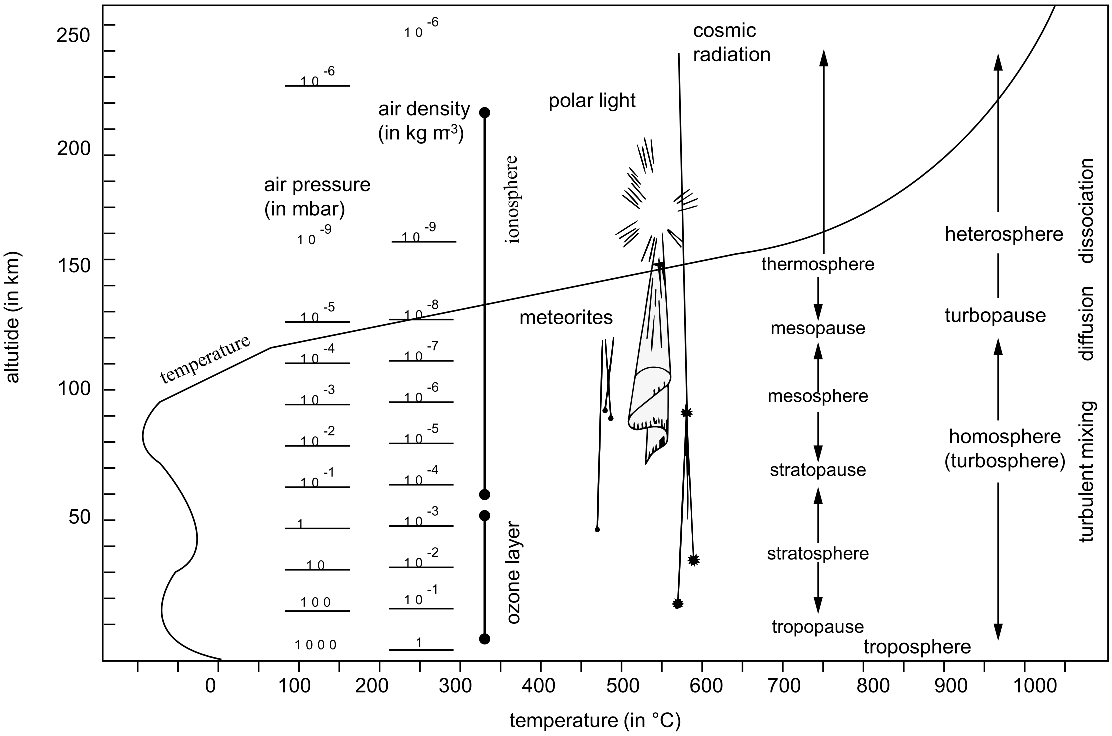


**Figure 2.1.: Layers of the atmosphere.** The temperature profile is displayed in dependency of the height. The corresponding pressure and air density is overlayed. The names of the different layers, the prevalent mode of transport and the hight of visible atmospheric phenomena are shown as well. The figure is taken from (Möller, 2019, p. 24).

Starting from the surface, the lowermost layer is called troposphere. It is characterized by turbulent convection and most cloud formation and precipitation processes happen in this layer. In the troposphere, the temperature decreases with increasing height. The top edge of the troposphere is called the tropopause. In the following layer, the stratosphere, the temperature increases with increasing height. The energy intake is related to the formation of the ozone layer by absorption of radiation in the ultraviolet range. The increase in temperature prevents air masses from vertical mixing, making the mass exchange between troposphere and stratosphere significantly lower than the mixing inside the troposphere. The upper boundary of the stratosphere, the stratopause, is characterized by a local maximum in temperature of about 0 °C. The layer above is called the mesosphere. There the temperature starts to decrease again and starts to rise again in the thermosphere and exosphere. With the interest of measuring near-surface emissions of $CO_2$ in mind, the relevant transport and variation of concentration happen in the troposphere. More distant layers play a minor role. The mechanisms of transport in the troposphere will be covered in Section 2.3.

**Table 2.1.: Composition of the atmosphere.** The mixing ratios of the most common constituents of the dry remote atmosphere (global mean concentrations 2015-2016) are shown. The table is adapted from Möller (2019).

| Substance | Formula | Mixing ratio (ppm) | |
|---|---|---:|---|
| Nitrogen | $N_2$ | 780,825 [a] | Constant |
| Oxygen | $O_2$ | 209,432 [a] | Constant |
| Argon | Ar | 9,339 [a] | Constant |
| Carbon dioxide | $CO_2$ | 404 | Increasing |
| Neon | Ne | 18.18 | Constant |
| Helium | He | 5.24 | Constant |
| Methane | $CH_4$ | 1.845 | Increasing |
| Krypton | Kr | 1.14 | Constant |
| Hydrogen | $H_2$ | 0.5 | Constant |
| Dinitrogen monoxide | $N_2O$ | 0.328 | Increasing |
| Carbon monoxide | CO | 0.12 | Increasing |
| Xenon | Xe | 0.087 | Constant |
| Ozone | $O_3$ | 0.03 | Variable |

[a] Related to “clean” atmosphere $O_2 + N_2 + Ar + CO_2$ (= 100 %)

The atmosphere consists of two main components, nitrogen ($N_2$) with 78.1% and oxygen ($O_2$) with 20.9%. The mixing ratios are given relative to dry air, as the content of water significantly varies. The concentration of GHGs is significantly lower. The mixing ratios for the most common atmospheric constituents are given in Table 2.1.

$CO_2$, as was stated in Chapter 1, is steadily increasing. The annual global average rose to a record of 419.3 ppmv in 2023 (NOAA Climate, 2024). Also, other important GHGs like methane ($CH_4$) and dinitrogen monoxide ($N_2O$) are increasing significantly. More detail on how rising GHGs are linked to rising temperatures is provided in Section 2.1.3. Before, the radiative properties of the atmosphere, which are the foundation of the greenhouse effect, are introduced.

### 2.1.2. Radiative Properties of the Atmosphere

The interaction of the atmosphere with electromagnetic radiation is an important prerequisite for understanding the atmosphere’s energy balance and thus the change of our climate. More importantly for this work, it is also the basis for the measurement principle used here – solar FTIR spectroscopy. For atmospheric remote sensing of $CO_2$ the infrared

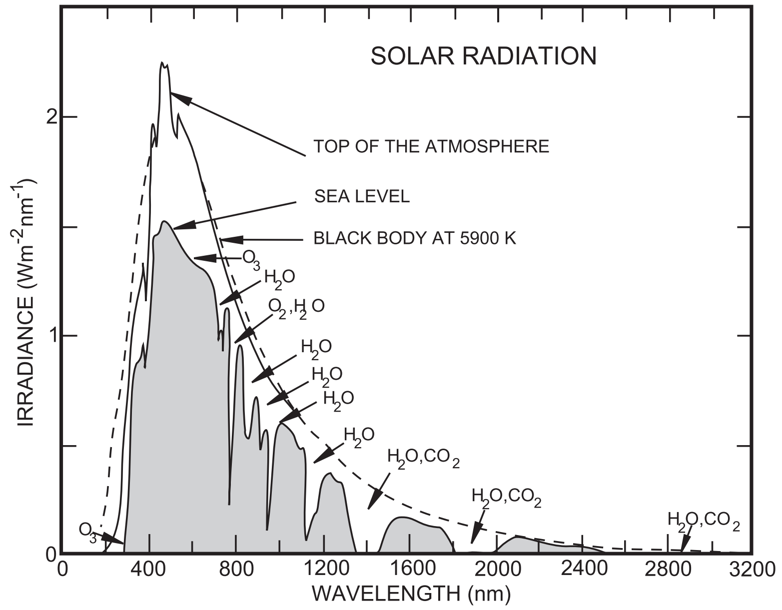


**Figure 2.2.: Solar spectrum.** The solar radiation reaching the top of the atmosphere is depicted with a black solid line. For comparison, the black body radiation spectrum at a temperature of 5900 K is shown as a dashed line. The gray area shows the solar radiation at the Earth's surface; the important areas of absorption from different atmospheric substances are annotated.

part of the electromagnetic spectrum is used, we will therefore focus on this part of the spectrum in the following. The interactions between the atmosphere (and matter in general) can be divided into three categories: emission, absorption and scattering. Scattering is not relevant for infrared light in the atmosphere and will therefore not be further considered. This section is based on Salby (2012), Haken and Wolf (2006) and Mayer-Kuckuk (1997).

### Black Body Radiation

Matter emits electromagnetic radiation, depending on its temperature. The intensity $F_\lambda(T)\,\mathrm{d}\lambda$ in an interval of wavelength $[\lambda, \lambda + \mathrm{d}\lambda]$ can be described by Planck's law (Roedel and Wagner, 2017)

$$F_\lambda(T)\,\mathrm{d}\lambda = \frac{2\,h\,c^2}{\lambda^5} \cdot \frac{\mathrm{d}\lambda}{\exp\left(\frac{h\,c}{\lambda\,k\,T}\right) - 1} \quad . \tag{2.1}$$

Here, $T$ is the body's temperature in Kelvin. The Planck constant $h$, the Boltzmann constant $k$ and the speed of light $c$ are physical constants. This emission behavior is called black body radiation.

Most of the incoming radiation at the top of the atmosphere originates from the Sun. Figure 2.2 shows that the incoming radiation at the top of the atmosphere is described well by a black body radiation spectrum with 5900 K, which is centered in the visible range of the electromagnetic spectrum. The different shape of the spectrum at the Earth's surface is caused by the absorption of radiation in the atmosphere and will be discussed further below. The second source of radiation in the atmosphere is the Earth itself. The emission spectra of the Earth's surface and atmosphere are in the infrared range. The black body spectra for the surface temperatures of the Earth and the Sun are shown in the top panel of Figure 2.3.

### Absorption and Line Shapes

One way that radiation interacts with the atmosphere is by absorption. The decrease of the intensity $I$ along a pencil beam in direction $s$ can be described by the differential equation

$$\frac{\mathrm{d}I_\lambda}{\mathrm{d}s} = -\beta_{a\lambda} I_\lambda \quad , \tag{2.2}$$

where $\beta_{a\lambda}$ is called the absorption coefficient. Equation (2.2) implies an exponential decrease of the intensity while passing through the pencil beam.

$$I_\lambda(s) = I_\lambda(0) e^{-\int_0^s \beta_{a\lambda} ds'} \quad , \tag{2.3}$$

where the integral in the exponent is also called the optical depth.

The degree of absorption significantly depends on the wavelength of the incident radiation. The absorption coefficient is high at specific wavelengths that correspond to transitions between different excitation states of the atmosphere's molecules. The excited states are quantized in the infrared, leading to singular absorption lines instead of a continuous spectrum.

In the infrared region, the relevant excitation states of molecules correspond to rotation and vibration modes of the molecules. In the atmosphere a variety of different molecules absorb parts of the electromagnetic radiation. The different species vary in concentration in the atmosphere. This results in a complex absorption or respectively transmission pattern. The percentage of absorption at the height of 11 km and at the ground level are shown in Figure 2.3. This wavelength-dependent absorption leads to the spectral shape of the solar radiation at the Earth's surface shown in Figure 2.2.

The excitation energy of a molecule consists of three components

$$E = E_{el} + E_{vib} + E_{rot} \tag{2.4}$$

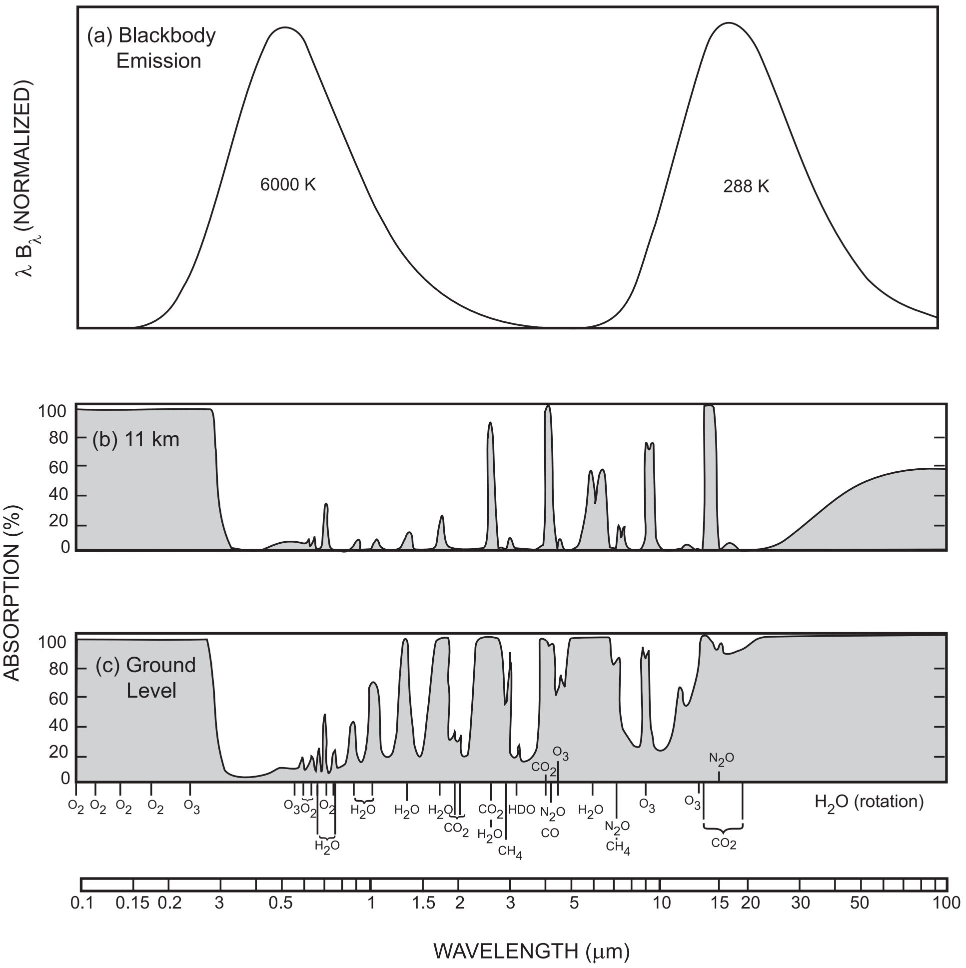


**Figure 2.3.: Absorption of radiation in the atmosphere.** (a) Blackbody emission spectra for the Sun's and Earth's surface temperatures. The percentage of absorption in dependency of the wavelength at the height of 11 km (b) and at the surface (c) (adapted by Salby (2012) from Goody and Yung (1995)).

Actually, the energy levels are defined by their quantum mechanical potentials in Schrödinger's equation, where also a coupling between the movement of the electrons and nuclei exists. In the Born-Oppenheimer approximation, the coupling is neglected. The motion of the electrons and nuclei have different time scales. This motivates the assumption to treat them independently. For atmospheric conditions, this approximation is in good agreement with the observed energy levels. The spectrum consists of the possible transitions $\Delta E$ between this excitation states. Changes in the states of the electrons correspond to wavelengths in the visible range and are not considered further here.

For the simplified assumption of a rigid rotor with a molecule of two atoms, the molecule has a parabolic potential. From this the quantized energy of the excitation states can be found to be

$$E_{rot} = \frac{\hbar^2}{2\Theta} J(J+1) \quad , \tag{2.5}$$

where $\Theta$ is the rotational inertia. The quantum number $J$ takes on natural numbers and corresponds to the state of excitation.

For the vibrational states, we can also make the simplified assumption of a parabolic potential. Analogue to the rotation this leads to the quantized energy levels

$$E_{vib} = \hbar\omega \left( v + \frac{1}{2} \right) \quad , \tag{2.6}$$

with the reduced Planck constant $\hbar$ and the oscillation frequency $\omega$. The quantum number $v$ takes on natural numbers. The derivation of these simplified energy levels can be found in Haken and Wolf (2006, Ch.9,10).

Actually, the spectrum is more complex due to a number of different reasons. In reality the vibrational potential is not parabolic, instead the finite dissociation energy is better approximated by a Morse potential. The rotational potential also differs from the assumption of a rigid rotor as the distance of the nuclei and thus the rotational inertia changes with the electrical excitation. The vibrational and rotational quantum numbers can change simultaneously and coupled ro-vibrational states exist. The quantum mechanical calculation of the transition probabilities gives specific transition rules, not all transitions occur. The line intensities depend on the probability of occupancy of the different, states which is temperature dependent. Finally, adding more atoms to the molecule also changes the spectrum. These effects result in a complex line structure.

The HITRAN database provides the spectral lines for different molecules including also the above-mentioned effects, to support spectroscopy applications in gaseous media. Ab initio calculations of the energy levels are combined with various experimental data. It is widely used, for example by satellite missions and for exoplanetary investigations. Two segments of the spectrum of $CO_2$ provided by HITRAN can be seen in Figure 2.4.

The absorption lines correspond to state transitions with a singular energy. However, the lines actually have a finite width due to various effects that broaden the line shape. The broadening depends on the composition of the atmosphere, as well as pressure and temperature. The shape of the individual lines is the key to retrieving GHG concentrations from measured spectra.

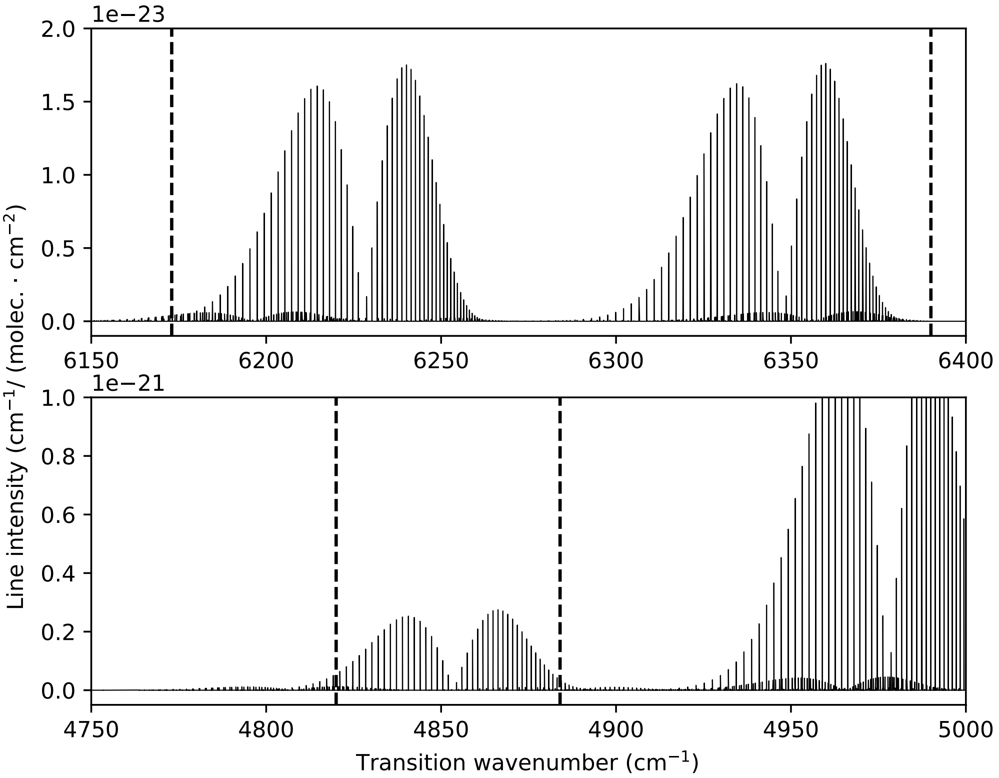


**Figure 2.4.: The line intensities for $CO_2$ as provided by HITRAN.** The intensity for the three most important isotopologues is multiplied by the isotopologue abundance and it is evaluated at T = 296 K (Gordon et al., 2022). The vertical dashed lines indicate the segments of the spectrum that are used for retrieving the vertical columns by COCCON. The upper panel shows the segment commonly used by COCCON, the lower panel shows an additional alternative segment that was recently implemented into the COCCON retrieval software.

The line shape cannot be narrower than the natural line width. It is limited by the Heisenberg uncertainty principle

$$\Delta E \Delta t > \frac{\hbar}{2} \quad , \tag{2.7}$$

linking a width in energy (and thus in wave number) to the decay time $\Delta t$ of an excited state. The resulting line width of the natural line shape is very small compared to other broadening effects (about $10^{-8}$ cm$^{-1}$) and is therefore negligible in the context of atmospheric spectroscopy (Efremenko and Kokhanovsky, 2021, p. 125f).

For high pressures, the decay time can be significantly larger than the interval between collisions of molecules. The collisions can trigger state transitions and therefore reduce the lifetime, which results in an increased line width, following Equation (2.7). Pressure broadening is described by a Lorentzian profile and dominates in low altitudes.

Doppler broadening is induced by the velocity distribution of the molecules, resulting in a Doppler shift of the emitted photon (Demtröder, 2018). If an atom has a velocity component $v_\mathrm{x}$ in direction of the emitted photon, the energy is shifted by

$$f = f_0 \left(1 - \frac{v_\mathrm{x}}{c}\right)^{-1} \quad . \tag{2.8}$$

The Maxwell distribution describes the number density $n(v_\mathrm{x})$ of molecules with a velocity component of $v_\mathrm{x}$ within the interval $[v_\mathrm{x},\, v_\mathrm{x} + \mathrm{d}v_\mathrm{x}]$ in a group of molecules with the temperature $T$

$$n(v_\mathrm{x})\mathrm{d}v_\mathrm{x} \propto \exp\left(-\frac{M v_\mathrm{x}^2}{2RT}\right)\mathrm{d}v_\mathrm{x} \quad , \tag{2.9}$$

where $M$ is the molecular weight and $R$ is the ideal gas constant. This leads to a Gaussian line shape

$$I(f) = I_0 \exp\left(-\frac{Mc^2}{2RT}\frac{(f - f_0)^2}{f^2}\right) \quad . \tag{2.10}$$

Under the assumption that the Doppler broadening is independent from the pressure broadening, the resulting combined line shape is a convolution of a Gaussian profile from the Doppler broadening and Lorentzian profile from the pressure broadening. This convolution is called Voigt profile. For comparison a Gaussian, a Lorentzian and a Voigt profile are shown in Figure 2.5.

In conclusion, the intensity of the absorption and the precise line shapes are dependent on the pressure, temperature and the distribution of GHGs in the atmosphere. How this absorption behavior is used for measuring the GHG abundances in the atmosphere will be covered in Section 2.2.

### 2.1.3. Earth's Energy Imbalance

The Earth's energy budget describes the sum of the energy arriving and leaving Earth. Ideally all components add up to zero – an equilibrium.

The solar radiation is the most relevant source of energy arriving at the Earth. Other sources as the heat transfer from the core of the Earth or nuclear processes are irrelevant in comparison. The relatively constant flux of incoming radiation is also referred to as *solar constant* with 1360 $\mathrm{W/m^2}$ (Roedel and Wagner, 2017). The geometrical consideration

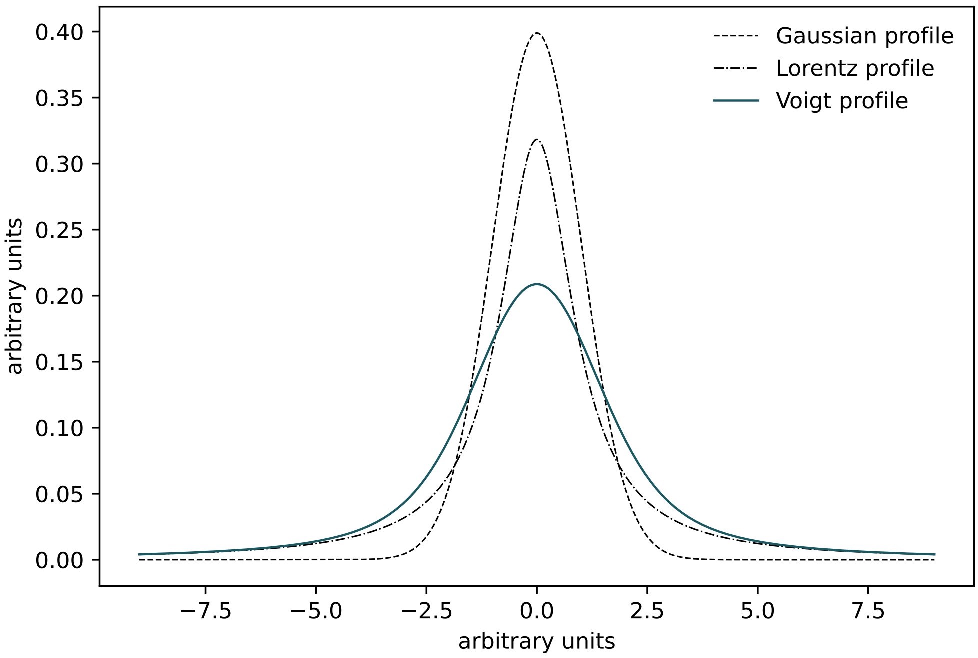


**Figure 2.5.: Shape of a Voigt profile.** The shapes of a Gaussian profile (Doppler broadening) and a Lorentzian profile (pressure broadening) are shown. The Voigt profile is a convolution of these profiles. All curves' areas are normalized to one.

that the projection of the Earth is illuminated, while the average heating of the whole surface is investigated, leads to the average incoming energy of

$$1360\,\frac{\mathrm{W}}{\mathrm{m}^2}\cdot\frac{\pi R_\mathrm{E}^2}{4\pi R_\mathrm{E}^2}=340\frac{\mathrm{W}}{\mathrm{m}^2} \tag{2.11}$$

The different pathways and transitions of the incoming energy are depicted in Figure 2.6 and will be explained in the following.

The solar radiation peaks at short wavelengths (compare Figure 2.3). Parts of the incoming radiation are directly reflected from the top of the atmosphere and the Earth's surface. The total effective planetary albedo (i.e. the fraction of reflection) is 29.4 % (Roedel and Wagner, 2017). Most of the remaining radiation is absorbed at the Earth's surface, while the atmosphere absorbs only a small fraction of the short wave radiation.

Simultaneously the Earth's surface and atmosphere emit radiation. The radiation emitted by the Earth's surface and the atmosphere is within the thermal infrared regime with wavelengths between 10 to 100 micrometers. While the atmosphere is relatively transparent for the Sun's incoming short wave radiation, its absorption is significantly

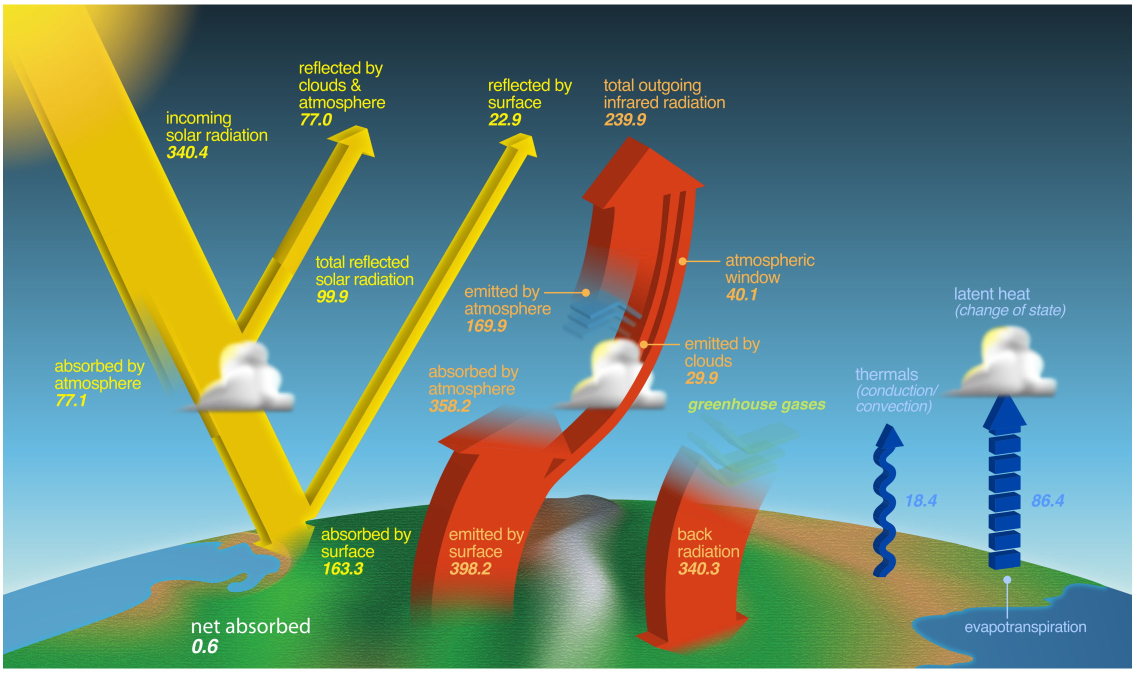


**Figure 2.6.: Earth's energy budget.** The incoming and outgoing radiation and the processes that are relevant for the global energy budget are depicted schematically. Short wave radiation originating from the sun is depicted with yellow arrows, infrared radiation emitted by the Earth's surface and atmosphere is marked by the orange arrows. Other processes that transport heat from the Earth's surface to the atmosphere are included with blue arrows. The energy flux that is transported by the different processes is annotated to the arrows in $\mathrm{W/m^2}$. The Figure is adapted from NASA (2016).

larger for the long wave radiation (compare Figure 2.3). The optical depth for the emitted radiation is in the order of 100 m. As a result, most of the emitted infrared radiation from the Earth's surface is absorbed by the atmosphere. The atmosphere itself emits radiation into all directions, therefore parts of it are re-emitted back towards the Earth's surface. This back radiation heats the Earth's surface with $340.4\,\mathrm{W/m^2}$, which is about the same amount as the energy income from solar radiation at the top of the atmosphere. As a consequence, the Earth is significantly warmer than without, or with an infrared-transparent atmosphere. This is effect is called *greenhouse effect*.

The absorption behavior of the atmosphere is altered by the increase of GHGs (as e.g. $CO_2$ and $CH_4$) in the atmosphere. As a result the amount of infrared radiation that is absorbed and re-emitted increases, leading to an effective heating. This effect is partly compensated by other emitted substance, that absorb radiation in the visible range, e.g. sulfur dioxide ($SO_2$), which leads to a net cooling effect. The change in the energy budget due to external drivers is called *radiative forcing* (unit $\mathrm{W/m^2}$). The radiative forcing of different species emitted by anthropogenic activities is depicted in Figure 2.7. It is linked to the estimated heating or cooling effect in the right panel of the plot.

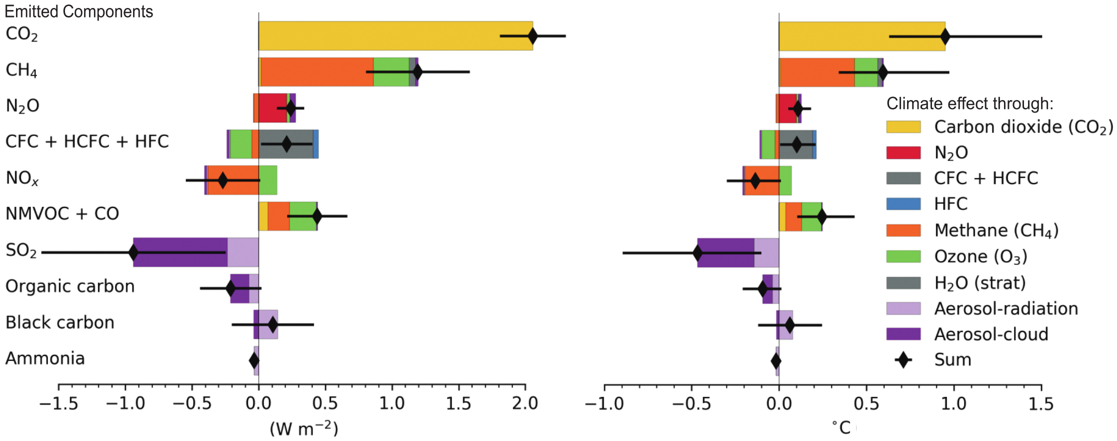


**Figure 2.7.: Radiative forcing and effect on surface temperature for emitted components.** The effective radiative forcing of different species emitted is shown in the left panel. The corresponding change in global surface temperature in the period between 1750 and 2019 is shown in the right panel. The Figure is adapted from IPCC (2023c).

The current energy budget is no longer in an equilibrium but in a state of energy imbalance. Loeb et al. (2021) determine the net energy uptake from satellite observations and in situ data between mid-2015 to mid-2019. The average energy uptake was $0.77 \pm 0.06\ \frac{\mathrm{W}}{\mathrm{m}^2}$ in this period, with an increasing trend that is consistent in both independent datasets.

The dramatic effects of the global warming were sketched out already in the introduction, underlining the need for a quick reduction of GHG emissions. We return now to the task of measurement-based emission estimation in the following section.

## 2.2. The Measurement Framework for Greenhouse Gas Emission Estimation

This section introduces the foundations of the atmospheric measurements performed in the present work. The basic principles of FTIR spectroscopy are explained in Section 2.2.1. After that, Section 2.2.2 shows the specific application of this measurement principle in the COCCON using EM27/SUN spectrometers. The processing of the raw measurement data collected by the EM27/SUN using the PROFFAST algorithms is explained in Section 2.2.3.

Former EM27/SUN measurement campaigns to estimate local emissions are summarized in Section 2.2.4. Finally, Section 2.2.5 presents a previous attempt to determine Thes-

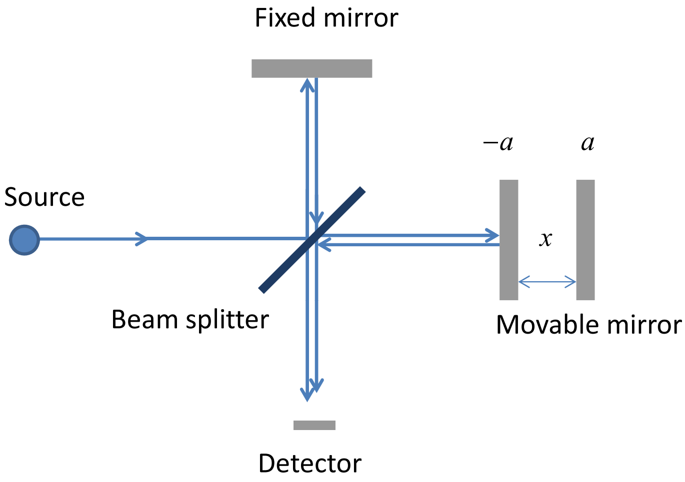


**Figure 2.8.: Schematic overview of a Michelson interferometer.** The incoming light is parted by a beam splitter and recombined at a detector. To one of the beams a variable path difference is added by a movable mirror. The figure is taken from Efremenko and Kokhanovsky (2021).

saloniki's emissions from the measurement campaign that is investigated also in this work.

### 2.2.1. Fourier Transform Infrared Spectroscopy

The main principle of FTIR spectroscopy is to collect a spectrum of infrared light using an interferometer. The interferometer splits the light to be examined into two beams. A variable path difference is added to one of the beams. After recombining the two beams, the so-called interferogram – the measured intensity as a function of the path difference – is recorded. A simple and illustrative example of an interferometer is the Michelson interferometer, which is depicted in Figure 2.8.

The interferogram $I(x)$, which is a function of the path difference $x$, can be converted into the spectrum $S(f)$, where $f$ is the frequency (Griffiths and De Haseth, 1986).

$$S(f) = \int_{-\infty}^{+\infty} I(x) e^{-i2\pi f} \mathrm{d}x \quad . \tag{2.12}$$

For the case of the actual measurement with a finite number of observation points, the Fourier transformation is discretized.

This technique finds application in many contexts. For remote sensing of the atmosphere, it is used in different configurations. In the present work the sun is used as external light source, and the absorption of the atmosphere is the basis of determining greenhouse gas abundances from the spectra. How the greenhouse gas abundances are technically obtained from the recorded interferograms is explained in more detail in Section 2.2.3.

**Figure 2.9.: Two EM27/SUN spectrometers operated side-by-side for calibration.** The measurements were taken in Thessaloniki in 2022 for calibration. The spectrometers can be carried by two persons. Photographed by Lena Feld.

### 2.2.2. The EM27/SUN Spectrometer and the COCCON Measurement Network

The EM27/SUN is a portable interferometer to record the atmosphere's absorption of sunlight for measurements of GHG abundances (Gisi et al., 2011; Hase et al., 2016). The EM27/SUN was developed in a cooperation between Bruker and the KIT. Apart from the interferometer, it is equipped with a solar tracker, that reflects the sunlight into the spectrometer. The CamTracker software continuously re-orients the mirrors using a camera imaging the field stop of the spectrometer. This allows automatic tracking of the Sun's position (see Gisi et al. (2011)). The interferometer is shielded by a robust box and the spectrometer has proven to be resilient against transportation. Herkommer et al. (2024) observed shocks of 16 g during transportation, letting the instrument characteristic unchanged in the limits of measurement precision. Two EM27/SUN spectrometers in operation are depicted in Figure 2.9.

An advantage of the EM27/SUN is its portability. It can be carried by two persons and is easy to transport with a small car. The portability makes inter-calibration between the different instruments practicable. The COCCON is a framework for the application of EM27/SUN (Frey et al., 2019). All spectrometers are calibrated against the reference

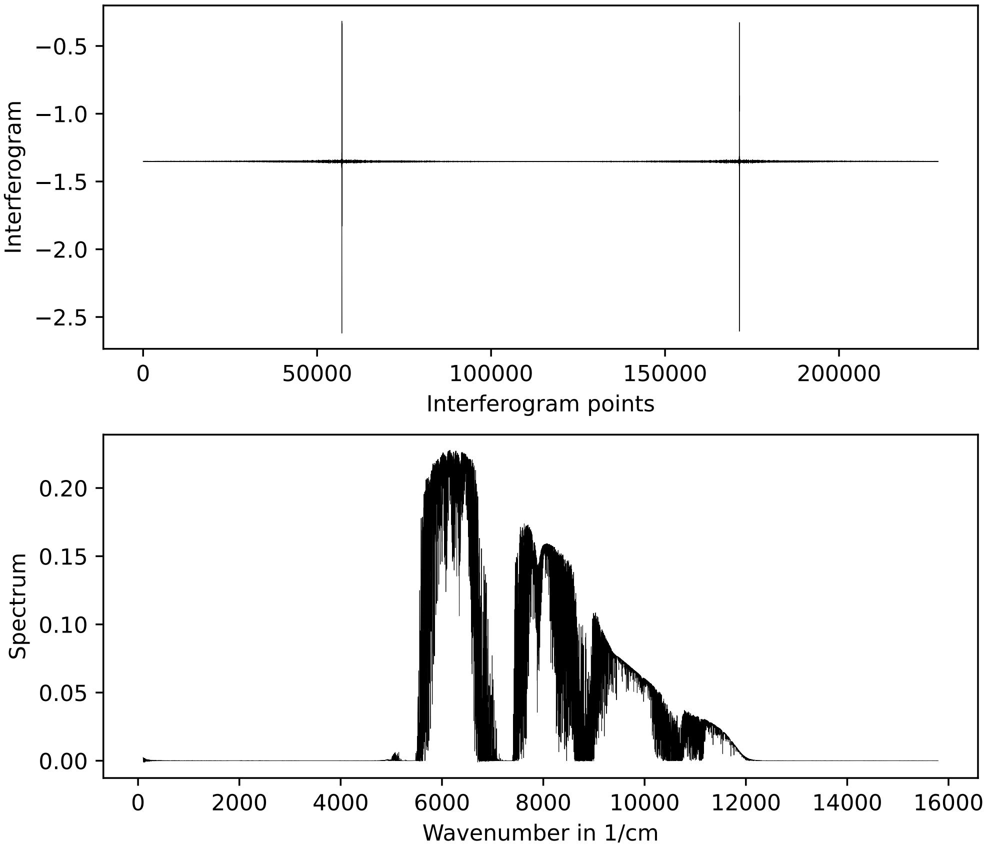


**Figure 2.10.: Recorded interferogram and derived spectrum.** The upper panel shows one interferogram recorded by the EM27/SUN with the serial number SN96 at Karlsruhe for calibration purposes. The corresponding spectrum derived from the interferogram is shown in the lower panel. The shown interferogram and spectrum are recorded with the first channel of the EM27/SUN. In addition, a second channel is available covering smaller wave numbers down to about 4000 $\mathrm{cm}^{-1}$.

instrument at KIT. In addition, important instrumental parameters are recorded and provided (see Alberti et al. (2022)). The KIT provides a set of processing standards, including algorithms for the retrieval, which will be summarized in more detail in Section 2.2.3. These aspects lead to a high consistency of data collected within the network, that is challenging to achieve with less portable instruments.

In the framework of the COCCON, continuous measurements are operationally conducted at 27 sites (e.g. in Karlsruhe, Sodankylä and Thessaloniki) (Dubravica, 2024). Additionally, many campaigns have been performed for emission estimation, which will be explained in Section 2.2.4.

Data with a spectral resolution of $0.5\,\mathrm{cm}^{-1}$ and a temporal resolution of 1 min is commonly recorded with the EM27/SUNs (Frey et al., 2019). As an example, one recorded interferogram and the corresponding spectrum are depicted in Figure 2.10. How the columnar abundances are derived from the spectra will be explained in the following.

### 2.2.3. Retrieval of Column-averaged Abundances

In order to obtain columnar abundances of GHGs from the recorded interferograms, a number of steps must be taken. From the interferogram the spectrum is obtained by Fourier transformation (see Section 2.2.1). From the absorption lines of $CO_2$, $CH_4$, CO, $H_2O$ and $O_2$ the vertical columns of the different species can be calculated. This requires a prior distribution of the gas in the atmosphere; the vertical column is derived by calculating a synthetic spectrum from a given prior, and iteratively scaling of the prior in order to minimize the difference between the synthetic and the observed spectra. The targeted column-averaged dry-air volume mixing ratios $X\mathrm{CO}_2$ are derived from the absolute vertical columns ($VC$).

$$X_{\mathrm{gas}}^{\mathrm{col.\ av.}} = \frac{VC_{\mathrm{gas}}}{VC_{\mathrm{dry\ air}}} \quad . \tag{2.13}$$

The vertical column of dry air is calculated from the vertical column of co-observed oxygen, for which the ratio is well known and constant throughout the atmosphere

$$X_{\mathrm{O}_2}^{\mathrm{col.\ av.}} = \frac{VC_{\mathrm{O}_2}}{VC_{\mathrm{dry\ air}}} \approx 0.2095 \quad . \tag{2.14}$$

For the trace gas this leads to

$$X_{\mathrm{gas}}^{\mathrm{col.\ av.}} = \frac{VC_{\mathrm{gas}}}{VC_{\mathrm{O}_2}} \cdot 0.2095 \quad . \tag{2.15}$$

$X_{\mathrm{CO}_2}^{\mathrm{col.\ av.}}$ is commonly referred to as $X\mathrm{CO}_2$. This nomenclature will be used in the following, if there is no danger of confusing it with the local volume mixing ratio of $CO_2$, which is called $X_{\mathrm{CO}_2}$.

#### Processing Steps of PROFFAST

In general, the inverse calculation of the spectra is based on a forward model. The target quantity (the spectrum) is simulated by transmission calculation from a prior assumption for the atmosphere's composition and pressure. Then the prior is altered to minimize the residuals between the simulated and observed spectra.

For the COCCON, the processing steps described above are conducted by a software package called PROFFAST which is implemented in Fortran.

The following program parts are performed sequentially

- PROFFASTpreprocess
- PROFFASTpcxs
- PROFFFASTinvers

PROFFASTpreprocess executes the Fourier transformation, after applying some corrections to the interferograms and performing quality checks to reject low-quality observations.

In PROFFAST, the quantitative trace gas analysis is split into two parts. First, PROFFASTpcxs (for pre-calculation of cross-sections) is performing a transmission calculation for selected solar zenith angles (SZAs) for the given prior distribution of gases in the atmosphere. This prior information originates from an atmospheric model and for each day, a prior representing the noon in local time is chosen. From the explicitly calculated cross sections, a dependency of the solar zenith angle is derived as a polynomial function, to derive the cross sections at any desired SZA. Additionally, the response to a change in temperature and pressure is evaluated. This way, the forward calculation of the transmission does not have to be repeated for each SZA which increases the speed of the calculation significantly. PROFFASTpcxs also calculates a sensitivity of the column to concentration changes in different heights. The column sensitivities are given for each model level of the internal grid-spacing used in the forward calculation of the transmission. With this, the response of the output to a change in concentration can be constructed. The column sensitivities can be used in combination with the prior to compare output from another model to the measurements.

In PROFFASTinvers the whole prior (in all heights) is scaled with a single factor. The results of PROFFASTpcxs can be used for efficiently finding an optimal scaling factor, that minimizes the residuals between the forward calculated spectrum and the observed spectrum. The result is an optimally-scaled prior distribution, from which the vertical column can be obtained. The calculation of the transmission critically depends on a correct surface pressure. Therefore, a precise pressure observation at the location of the observation is required as additional input. Tu (2019) found that an increase of 1 hPa resulted in an increase of 0.039 % in $X\text{CO}_2$, which corresponds to 0.16 ppmv at a concentration of 420 ppmv. This shows the need for including an accurate pressure observation.

To be consistent with results from the long-established sister network Total Carbon Column Observing Network (TCCON) (Wunch et al., 2011) empirical correction factors are applied to the vertical columns. The COCCON is tied to the TCCON network, that uses the IFS 125HR of Bruker. Therefore, a continuous side-by-side measurement between the TCCON station at KIT and a reference EM27/SUN (SN37) is recorded. The above-mentioned empirical correction factors are determined from this simultaneous time series. Additionally, side-by-side measurements with two EM27/SUN at the TCCON station in Sodankylä, Finland, are taken into account. Currently, two airmass-dependent correction factors and one airmass independent correction factor are used to correct the vertical columns. After correcting the columns, the column-averaged volume mixing ratios are obtained as described in Equation (2.15).

The column-averaged mole fractions retrieved from the COCCON reference spectrometer SN37 are assumed to be fully calibrated and in agreement to the Karlsruhe TCCON site. To correct for the bias between different EM27/SUNs, side-by-side measurements are performed relative to SN37. The bias originates from instrument specific imperfections as small differences e.g. in the optical alignment and detectors. The relative correction factors are applied to $X_{\text{gas}}^{\text{col. av.}}$ directly. The specific implementation of this relative calibration for the Thessaloniki campaign is explained in Section 3.1.1.

### 2.2.4. Urban Emission Estimation from COCCON Campaigns

In the framework of COCCON several measurement campaigns have been performed to estimate local emissions. Most were targeted on cities, for example, Berlin (Hase et al., 2016), Paris (Vogel et al., 2019) or Madrid (Tu et al., 2022).

To estimate emissions from campaign observations, a model of the transport of the tracers is essential to be able to link the observations to the observed time series of column-averaged molar fractions. Previous campaigns used models of different complexity for this purpose.

Simplified approaches as a box model (Makarova et al., 2021) or a Gaussian plume model (Tu et al., 2022) were used in city campaigns in St. Petersburg and Madrid. For the Berlin campaign, Hase et al. (2015) implemented a Lagrangian dispersion model, with a simplified wind field for the city. For the emission determination, the city was divided into five rectangular emission regions.

More complex models were used for the campaigns in Tokyo and Paris. In these campaigns, the observations were predicted by a numerical model, using a prior emission map and simulating the transport towards the observation sites. For Paris, the model

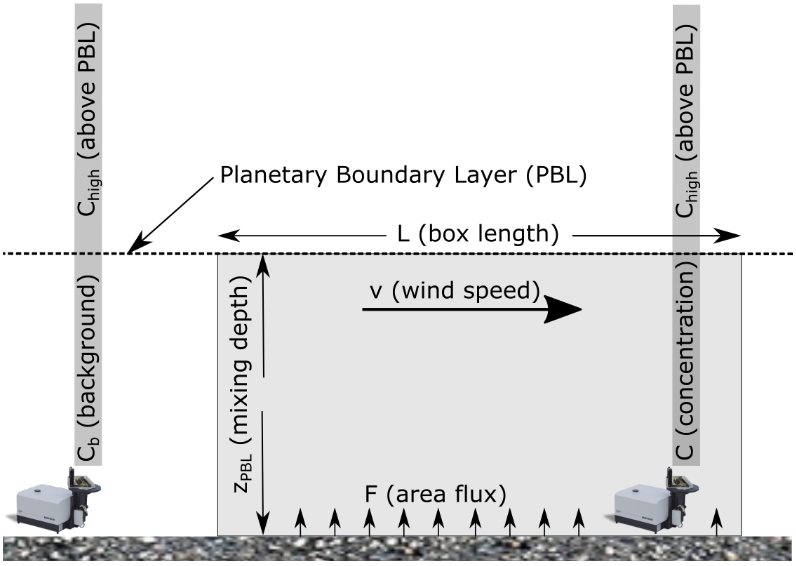


**Figure 2.11.: Box model with two EM27/SUN.** The principles of the box model are depicted schematically. The columnar observations with two EM27/SUN spectrometers are included in the image. Figure by Schmid (2023).

CHIMERE was used and the comparison between the observed and simulated time series of $XCO_2$ was discussed quantitatively (Vogel et al., 2019). For the Tokyo campaign, Ohyama et al. (2023) simulated trajectories using HYSPLIT. They used a Bayesian inversion approach to derive emission estimates from the observational data. A complementary approach was implemented by Viatte et al. (2017), for a campaign targeted at diaries in Los Angeles. Here, the numerical weather prediction (NWP) model WRF was used to simulate the emissions and transport of methane. Estimates based on a Bayesian inversion were derived in this work.

### 2.2.5. Estimation of Thessaloniki's Emissions Using a Box Model

A measurement campaign in Thessaloniki, Greece is the topic of this work. In the campaign a pair of EM27/SUN spectrometers was deployed. I operated one of the spectrometers with the help of Pablo Schmid, a master's student I supervised. The details of the campaign implementation will be the topic of Chapter 3.

This work focuses on the combined interpretation of the campaign results together with simulations with the NWP model ICON-ART to determine the emissions of the city. Schmid (2023) estimated the city emissions using the simpler box model approach. The results of his work will be explained in the following.

In the box model, the emission flux ($F$) is constant over a specific area. Within a box extending to the planetary boundary layer and horizontally limited by the emission area, the tracer is assumed to be well mixed. In addition the wind $v$ is assumed to be distributed uniformly over the box. The box model is sketched out schematically in Figure 2.11.

With the above mentioned assumptions the concentration of the tracer inside the box $C$ can be described by the following formula (Hanna et al., 1982)

$$L\, z_{\mathrm{PBL}} \frac{\partial C}{\partial t} = L\, F + v\, z_{\mathrm{PBL}} (C_{\mathrm{b}} - C) + L \frac{\partial z_{\mathrm{PBL}}}{\partial t} (C_{\mathrm{high}} - C) \quad . \tag{2.16}$$

Here, the concentration above (next to) the box is $C_{\mathrm{high}}$ ($C_{\mathrm{b}}$). The length of the box is $L$, the height of the planetary boundary layer $z_{\mathrm{PBL}}$ and the quantity of interest, the area flux of emissions is $F$.

With assumptions of a constant hight of the PBL, and a steady-state scenario ($\frac{\partial C}{\partial t} = 0$) the equation can be simplified to

$$F = \frac{v\, z_{\mathrm{PBL}}\, (C - C_{\mathrm{b}})}{L} \quad . \tag{2.17}$$

The difference in concentrations $\Delta C = C - C_{\mathrm{b}}$ can alternatively be expressed through a difference in column-averaged volume mixing ratios $X_{\mathrm{gas}}^{\mathrm{col.\ av.}}$ with

$$\Delta X_{gas}^{\mathrm{col.\ av.}} = \frac{N_A\, z_{\mathrm{PBL}}}{VC_{\mathrm{dry\ air}} \cdot M_{\mathrm{gas}}} \Delta C \quad . \tag{2.18}$$

Here, $M_{\mathrm{gas}}$ is the molar mass of the respective gas. This implies that the vertical column of dry air $VC_{\mathrm{dry\ air}}$ can be assumed to be equal for both observation sites. When inserting $\Delta C$ into Equation (2.17), $z_{\mathrm{PBL}}$ cancels out:

$$F = \frac{v\, M_{\mathrm{gas}}\, VC_{\mathrm{dry\ air}}}{L\, N_{\mathrm{A}}} \Delta X_{\mathrm{gas}}^{\mathrm{col.\ av.}} \quad . \tag{2.19}$$

The independence of $z_{\mathrm{PBL}}$ shows that observations of $X_{\mathrm{gas}}^{\mathrm{col.\ av.}}$ are less sensitive to changes in $z_{\mathrm{PBL}}$. This is an advantage of using column-averaged observations, as a correct prediction or measurement of the height of the PBL is challenging.

With this approach Schmid (2023) estimated the city emissions of $CO_2$ to be 59.9 $\frac{kt}{km^2\, yr}$, with an city area of 111 $\mathrm{km}^2$. This corresponds to total city emissions of 6.6 Mt/yr. However, he argues that the assumptions of the box model are not fulfilled for the observed dataset. In particular, the complex wind field is in contrast to the assumption of an evenly distributed wind over the box area. He estimates the uncertainty of the emission to be in the order of 100 %. Instead of the box model, a model approach that accounts for the complex wind field in Thessaloniki is required.

## 2.3. Modeling Atmospheric Transport of Trace Gases

This work aims to determine emissions based on atmospheric concentrations. This section explores the general relationship between atmospheric concentrations and emission sources and sinks.

In Section 2.3.1 the inverse problem of emission estimation is introduced along with two fundamental concepts for emission determination. Section 2.3.2 gives an overview about the dynamical processes in the atmosphere that drive the transport of tracers in the atmosphere. The Numerical Weather Prediction (NWP) model ICON-ART is introduced in Section 2.3.3. The Sections 2.3.1 and 2.3.2 are based on Enting (2002) and Roedel and Wagner (2017).

### 2.3.1. Inverse Emission Estimation

The atmospheric concentrations are not only determined by the emissions but also by the constituent transport. In general, the modeled concentration $m(\mathbf{r}, t)$can be expressed as

$$\frac{\partial}{\partial t} m(\mathbf{r}, t) = s(\mathbf{r}, t) + \mathcal{T}[m(\mathbf{r}, t), t] \quad , \tag{2.20}$$

with a source $s$ and the transport operator $\mathcal{T}[.,.]$.

This can be considered in a differential way

$$\hat{s}(\mathbf{r}, t) = \frac{\partial}{\partial t} \hat{m}(\mathbf{r}, t) - \mathcal{T}[\hat{m}(\mathbf{r}, t), t] \quad , \tag{2.21}$$

with the statistical estimates $\hat{s}$ and $\hat{m}$. This approach is often called ’mass-balance’ technique, since the equation (as Equation (2.20)) is based on the assumption of mass conservation.

Alternatively, Equation (2.20) can be rewritten in an integral form

$$m(\mathbf{r}, t) = m_0(\mathbf{r}, t) + \int_{t_0}^{t} G(\mathbf{r}, t, \mathbf{r}', t')\, s(\mathbf{r}', t')\, \mathrm{d}^3 r' \mathrm{d}t' \tag{2.22}$$

where $m_0$ describes the evolution of the system in the absence of sources and sinks. Formally $G$ is the Green’s function and the problem is also referred to as synthesis approach. In practice a discretized form of this formula is used

$$m(x_\mathrm{j}) = m_0(x_\mathrm{j}) + \sum_{\mu} G_\mu(x_\mathrm{j})\, s_\mu \quad , \tag{2.23}$$

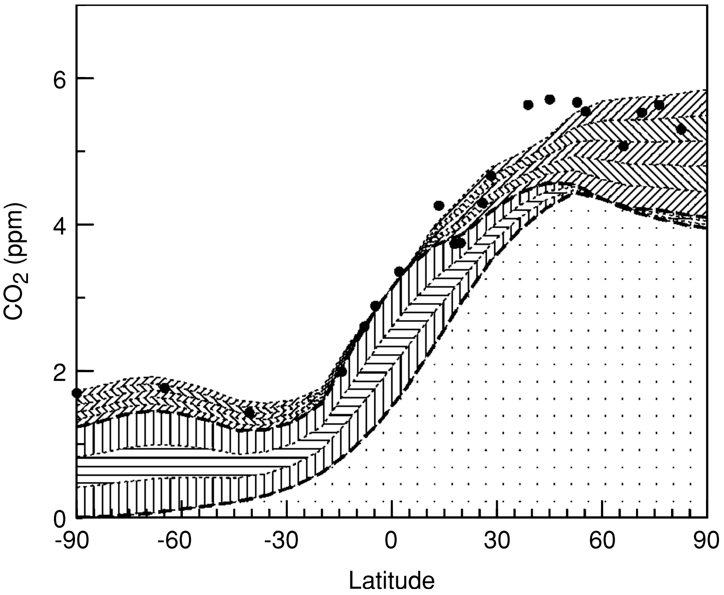


**Figure 2.12.: Synthesis approach example.** The measured longitudinal CO$_2$ variation in dependency of the latitude is indicated with black dots. These measured molar fractions can be approximated by modeled results. The model is a linear combination from the different components. The contribution from five different terrestrial areas are indicated above the upper dashed line with diagonal hatching. The ocean's influence is indicated by three components between the two dashed lines, filled with vertical and horizontal hatches. For the fossil fuel component (dotted area) only one component is included due to it's lower uncertainty. The image is taken from Enting (2002).

where discretized model prediction $m$ at the location $x_\mathrm{j}$ is a sum of the different components $G_\mu(x_\mathrm{j})$ that connect a set of sources $s_\mu$ to atmospheric concentrations.

For making this synthesis approach more tangible, Enting (2002) considers the global variation CO$_2$ in dependency of the latitude as an example (see Figure 2.12). The latitude-dependent variation of the CO$_2$ molar fraction can be simulated from a set of different sources and sinks, as the ocean and terrestrial contribution and fossil fuel combustion. The different components can be added, yielding an estimate for the total CO$_2$ variation. A regression is performed by fitting a set of observations $c$ at the locations $x_j$ to the function

$$m(x_\mathrm{j}) = \sum_\mu s_\mu G_\mu(x_\mathrm{j}) \overset{!}{\approx} c(x_\mathrm{j}) \quad . \tag{2.24}$$

The optimal coefficients $s_\mu$ are determined through the fitting procedure, representing the adapted source strength. The different components $G_\mu$, whose linear combination represents the total CO$_2$ concentration, are shown schematically in Figure 2.12. In this work a synthesis approach similar to this example is used and we will come back to it in Chapter 4.

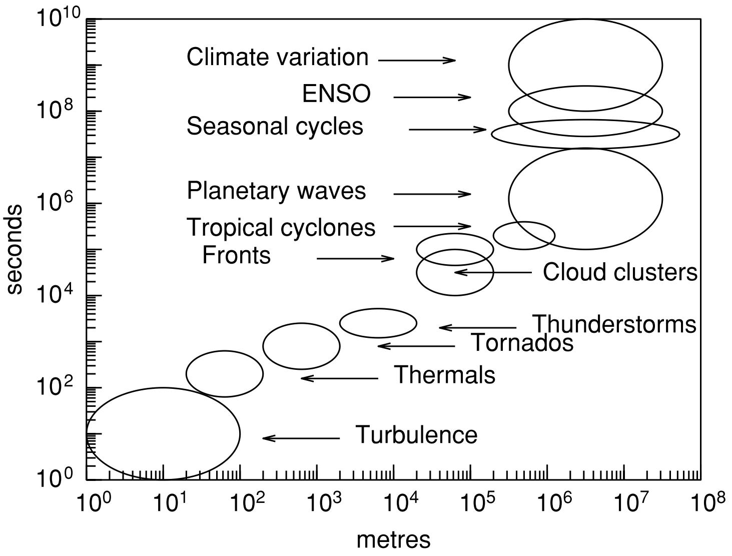


**Figure 2.13.: Dynamical processes of the Atmosphere.** The processes are ordered by their scale in space and time. The Figure is taken from Enting (2002) and based on Orlanski (1975)

### 2.3.2. Dynamical Processes of the Atmosphere

An accurate description of the transport, that links the sources to the atmospheric concentrations is mandatory for both inversion approaches from Equation (2.21) and Equation (2.22). Generally, particles can be transported by advection and molecular diffusion, where advection is a transport proportional to the velocity of the fluid parcel and molecular diffusion is a mixing resulting from the movement of the individual particles. For the task of estimating urban emissions, changes in concentrations are investigated in the troposphere. Here, diffusion can be neglected because the mean free path is small and advective mixing and transport are the dominant processes.

An accurate calculation of the wind field is important to correctly estimate the transport. The Eulerian equation of motion describes the movement of an air parcel

$$\rho \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\mathrm{grad}\, p - \rho \cdot \mathrm{grad}\, \Phi + \mathbf{F}_\mathrm{C} + \mathbf{F}_\mathrm{F} \quad . \tag{2.25}$$

Here, $\mathbf{v}$ is the velocity of the air parcel and $\rho$ is the density of air. The acceleration of the air parcel is determined by the gradient of pressure $p$ and the gradient of the geopotential $\Phi$. In addition the air parcel is accelerated by the Coriolis force $\mathbf{F}_\mathrm{C}$, a force of inertia rectangular to the wind direction originating from Earth's rotation and the non-inertial frame of reference of the terrestrial observer. Finally, the frictional forces $\mathbf{F}_\mathrm{F}$ affect the acceleration of the particle.

In addition, the continuity equation expressing the conservation of mass must be satisfied

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho\,\mathbf{v}) = 0 \quad . \tag{2.26}$$

From the conservation of energy the equation

$$\frac{\mathrm{d}T}{\mathrm{d}t} = Q_\mathrm{W} + Q_\mathrm{I} \quad , \tag{2.27}$$

where $Q_W$ is the temperature change due to aggregation state transitions (most importantly of water) and $Q_\mathrm{I}$ is the temperature change due to radiation. This equation resembles the first law of thermodynamics.

Finally, an ideal gas is assumed for the atmosphere, with the ideal gas equation

$$p\,V = n\,R\,T\,. \tag{2.28}$$

Here, $R$ is the ideal gas constant. The number of moles of air $n$ can be expressed as $n = N/N_\mathrm{A}$ with the total number of molecules $N$ and the Avogadro constant $N_\mathrm{A}$.

In principle, the atmospheric processes can be derived from these equations. In practice, however, knowledge of the initial conditions is always incomplete. An overview of the different dynamical processes in the atmosphere on different space and time scales is shown in Figure 2.13, ranging from large-scale phenomena that are present when looking at the global scale as seasonal cycles or the El Niño/Southern Oscillation (ENSO), up to turbulence – a seemingly chaotic movement of air on scales of millimeters to hunderets of meters.

The mass concentration $\rho_\mathrm{gas}$ of a transported gas and the velocity $\mathbf{v}$ of the air parcel can be separated into the average and fluctuations from that average.

$$\rho_\mathrm{gas} = \bar{\rho}_\mathrm{gas} + \rho'_\mathrm{gas} \tag{2.29}$$

and

$$\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}' \quad . \tag{2.30}$$

When looking at the average flux density $\mathbf{j} = \rho_\mathrm{gas}\mathbf{v}$ we find

$$\bar{\mathbf{j}} = \bar{\rho}_\mathrm{gas} \cdot \bar{\mathbf{v}} \;+\; \bar{\rho}_\mathrm{gas} \cdot \bar{\mathbf{v}}' \;+\; \bar{\rho}'_\mathrm{gas} \cdot \bar{\mathbf{v}} \;+\; \overline{\rho'_\mathrm{gas}\,\mathbf{v}'} \quad . \tag{2.31}$$

When integrating over a time span $T$ to find the averages, $\bar{\rho}_\mathrm{gas}$ and $\bar{\mathbf{v}}$ are constants, therefore the second and third term vanish, as $\bar{\rho}'_\mathrm{gas} = \bar{\mathbf{v}}' = 0$ by definition.

The continuity equation (2.26) must also hold for the air component $\rho_{\mathrm{gas}}$ if there are no chemical reactions destroying or producing the respecive substance. With the decomposition from above, the continuity equation describing the average transport can be re-written as

$$\frac{\partial \bar{\rho}_{\mathrm{gas}}}{\partial t} = -\mathrm{div}(\bar{\rho}_{\mathrm{gas}} \bar{\mathbf{v}}) - \mathrm{div}(\overline{\rho'_{\mathrm{gas}}\, \mathbf{v}'}) \quad . \tag{2.32}$$

The first term is the advective transport by the average wind motion. The second term is identified as the concentration change due to turbulent mixing (also referred to as turbulent diffusion).

For the task of determining urban emissions, transportation times from minutes to days and distances in the km-scale are relevant. Smaller variations are only relevant for their effect on the transport, but not the explicit variations at that scale. Numerical modeling of the wind fields and transport is only feasible up to a finite resolution. Transport due to processes smaller than the model resolution is typically parametrized by other model variables (e.g., wind speed) that represent the average at a given grid cell.

### 2.3.3. The Numerical Weather Prediction Model ICON-ART

The aim of Numerical Weather Prediction (NWP) is to forecast the state of the atmosphere from a given initial state. Therefore, the fluid dynamic equations are discretized, with parameterizations for small scale processes such as micro cloud physics or the turbulence mentioned above. In most cases real measurement data is used to construct an initial state of the atmosphere that best represents the real state of the atmosphere by data assimilation.

The Icosahedral Nonhydrostatic (ICON) model is the numerical weather prediction model that is operationally used by the German Weather Service (DWD). It is developed at DWD and the Max Planck Institute for Meteorology (MPI-M). Special in comparison to other models is the implementation of non-hydrostatic equations and the icosahedral grid structure. This results in a grid without singularities and similar grid sizes over all latitudes. The grid structure will be explained in more detail below. It is a global model, that can be operated on various resolutions. The model can be operated globally or regionally, and with various resolutions

The module for Aerosols and Trace gases (ART) (Schröter et al., 2018) includes the functionality to transport tracers, simulate their emissions and chemical interactions. ART is developed at the KIT.

As a whole, ICON-ART is a highly complex model with a variety of different setup possibilities. In the following, a few aspects of ICON that are relevant for the current work will be explained in more detail.

### Grid Structure

As mentioned above, ICON-ART uses an icosahedral grid structure, allowing a variety of different resolutions. The ICON grids are constructed by successive refinement of a spherical icosahedron, consisting of 20 triangles of equal size. To construct a refined grid, the circle arcs connecting the edges of a triangle are divided into n parts, that are additional edge points of the refined grid. All edges of the refined grid, can be found by connecting the new edges by arcs. This first step is called root division step. In a second step the triangles are recursively subdivided into 4 smaller triangles $k$ times. These subdivisions are called bisection steps. The so found grid version is named $RnBk$. The ICON grid structure is described by Zängl et al. (2015).

### Online and Offline Nesting

ICON-ART is capable of both, global and regional simulations. For regional simulations, often a high resolution is desired, but simulations at a high resolution are computationally costly for a global domain. Therefore, two methods are available for regional refinement of the grid. The first solution is online nesting. Therefore, in a defined area, grid cells of the parent grid are separated into four smaller cells with an effective resolution of $\frac{1}{2}$ compared to the parent grid. Different configurations are possible. For a configuration with so-called two-way-nesting, the meteorological variables from the coarser grid are fed into the simulations of the finer grid and vice-versa. A drawback is, that the resolution can be only halved in the nested domain, and the simulation is fixed to the same base grid.

A second possibility is the offline-nesting. This mode is also called limited area mode (LAM). Here, a coarser simulation is run independently from the fine resolution in advance. The output of this parent simulation is remapped to a different resolution at regular time steps. This is determining the borders of the the LAM model. This way the simulation is linked less closely to the parent simulation, since the borders of the fine grid are updated less frequently and also no feedback from the calculations in the finer domain can be given to the coarser simulation. The advantages are, that a enhancement of spatial resolution is possible, from the parent resolution to the nested resolution, and also a change in the base-grid is possible.

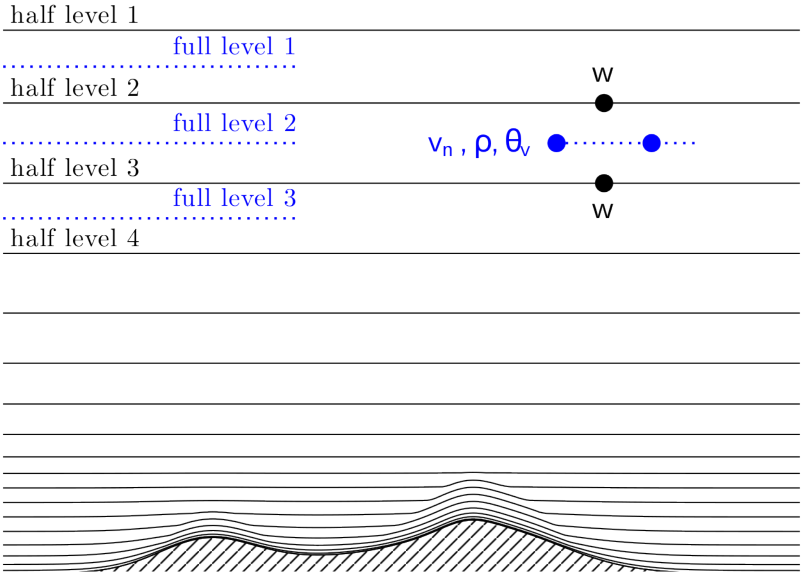


**Figure 2.14.: Model levels in ICON.** The model levels are depicted schematically. The numbering of the layers is annotated at the left hand side and the locations of the model variables are indicated in the center. The black lines show the limits of the model layers. The lowermost levels follow the terrain (indicated as hashed area). The image is taken from Reinert et al. (2024).

### Model Levels

ICON-ART is a non-hydrostatic model. Since the pressure is not fixed in the model, geometrical heights are more practical than a grid on pressure levels. The geometrically defined model levels are counted starting from the top level. The height levels, are not equidistant in height, but are densest near the surface.

The lowermost levels follow the topography. As a results, the height levels differ between the grid cells. The tracing of the topography fades out toward the top, so that the upper levels have the same horizontal heights. The horizontal structure is illustrated in Figure 2.14. The vertical grid can be chosen individually. The number of levels, top height, minimum layer thickness and how fast the topography fades out towards the top are adjustable. The operational global Setup of ICON has a model top height of 75 km (Reinert et al., 2024). For a nested domain a reduced number of model levels is possible compared to the parent grid.

### Transport

The transport module of ICON solves the simplified continuity equation (Reinert et al. (2024))

$$\frac{\partial}{\partial t}\bar{\rho}_{\mathrm{gas}} + \nabla(\bar{\rho}_{\mathrm{gas}}\mathbf{v}) = 0 \quad , \tag{2.33}$$

with the average mass density of the trace gas $\bar{\rho}_{\mathrm{gas}}$. The equation is simplified as the velocity $\bar{\mathbf{v}}$ represents the average at a given grid cell. The actual velocities vary over much smaller scales leading to turbulent mixing (compare Equation (2.32)). The turbulent component of the transport is not modeled explicitly but parametrized from the average velocity instead. A second simplification is the absence of a source term. The mass conservation is not necessarily fulfilled for an individual trace gas. Particles of the species can be created or destroyed through chemical or biological processes. In addition they can be released from or absorbed by the ocean or land surface. The emission and chemistry modules alter the state of the respective grid cells in between the iterative processing steps.

For the specific implementation in ICON-ART the vertical transport is separated from the horizontal transport. Different discretization approaches can be chosen in the model setup. For simplification, ART defines template sets of predefined parameters for transport. The transport template used in this work is the `hadv52aero` template, where for the horizontal transport a hybrid scheme based on Miura (2007) and Harris et al. (2011) is used. The vertical transport is following the PPM (piecewise parabolic method) scheme developed by Colella and Woodward (1984).

### Sinks and Sources

ICON-ART provides different modules to include atmospheric chemistry. This work is focused on $CO_2$ and considers mainly short transportation times. Therefore, $CO_2$ is simulated as a passive tracer, so there is no sink in the atmosphere in this model configuration. Three emission methods are available, the Online Emission Module (OEM) developed at EMPA, an area-based emission module by Weimer et al. (2017) and a method to define point sources.

The OEM is calculating the emissions at the processing step from a static emission grid, and an overlaid time dependency. For the area-based emission module, gridded emission maps are read in and interpolated for all time steps at the beginning of the simulation. The OEM has the option to specify a vertical profile for different emission sectors; the area emission module only handles surface emissions. For the point source emissions, the emission source can be selected variably in the volume of the atmosphere.

For all emissions modules the molar fraction in the respective grid cell $i$ is altered by

$$\Delta X_{\mathrm{i}} = \frac{\Delta n_{\mathrm{i}}}{n_{\text{wet air}}} \quad , \tag{2.34}$$

where $\Delta n_\mathrm{i}$ is the number moles of the emitted particles and $n_\mathrm{wet\ air}$ is the total number of moles in the grid cell.

For the OEM and the area emission module, $\Delta n_i$ is calculated from the emission flux of the grid cell $E_\mathrm{i}$,

$$\Delta n_i = \frac{E_\mathrm{i} A_\mathrm{i} \Delta t}{M_\mathrm{i}} \quad , \tag{2.35}$$

with the molar mass of the emitted substance $M_\mathrm{i}$, the surface area $A_\mathrm{i}$, and the model time step $\Delta t$.

### ICON D2 Operational Setup

Because being a prototype for the simulations performed in this work, the operational weather forecast setup of the German weather service will be explained briefly in the following. The operational setup has 2 steps, at first a global simulation with a resolution of 13 km on a R3B7 grid is performed with an online-nested domain over Europe (R3B8) with a resolution of 6.5 km. The global grid has 120 height levels and extends to a height of 75 km, the EU nest runs on 74 of these layers, extending to a height of 23 km. The simulated results for Europe are used to feed a LAM run with an offline-nested domain over Germany, with a higher resolution of 2.1 km and 65 height layers extending to a height of 22 km. Apart from a deterministic run, several ensemble members with perturbations in the initial conditions are performed to estimate the uncertainty of the prediction. For the global forecast 40 members and for the ICON-D2 nest 20 ensemble members are run, both at a lower resolution than the deterministic run. For the forecast, meteorological observations are assimilated, pulling the simulation towards the actual state of the atmosphere. The operational setup is described by Reinert et al. (2021).

The specific model setup that was chosen to simulate the emissions and observed trace gas concentrations corresponding to the Thessaloniki measurement campaign is explained in Chapter 4.

# 3. Implementation of the Thessaloniki City Campaign

This chapter gives an overview how the COCCON measurement campaign in Thessaloniki was implemented. The general strategy and positioning of the instruments is explained in Section 3.1. In this section, also an overview over the collected dataset will be given. Significant improvements were introduced to the workflow of data processing during the campaign (see Section 3.2).

## 3.1. Strategy of the Campaign and Dataset Collection

The Thessaloniki campaign took place in October 2021 and from May to July 2022. During this period two EM27/SUN spectrometers were available, one of which (*SN52*) has been in operation in Thessaloniki since 2019 as part of COCCON. Marios Mermigkas continued to collect a long-term time series during the period of the campaign (Mermigkas et al., 2021). A second spectrometer (*SN96*) was brought to Thessaloniki for the campaign. It was equipped for mobile deployment and the master student Pablo Schmid and I conducted the measurements.

The inter-calibration of the two instruments is presented in Section 3.1.1. Next, Section 3.1.2 explains the mobile operation of one of the spectrometers for the campaign. This allowed for probing different observation sites during the campaign. The positioning of the spectrometers is discussed in Section 3.1.3. An important prerequisite for the data processing is the pressure at the position of the instruments. Section 3.1.4 describes the methodology used to estimate the surface pressure at the observer's locations. This is followed by the results of a sensitivity study to test the local variability over the city on smaller scales (see Section 3.1.5). In the final part of this section (Section 3.1.6) an overview over the dataset that was compiled during the campaign period is given.

**Table 3.1.: Calibration factors.** The calibration factors derived from the side-by-side observations are listed for the spectrometers SN52 and SN96 for the respective species.

| | $k_{52}$ | $k_{96}$ |
|---|---|---|
| tracer | | |
| XCO2 | 1.001125 | 1.000135 |
| XCO | 1.009516 | 1.011334 |
| XCH4 | 1.000772 | 1.000329 |
| XH2O | 1.000640 | 1.000224 |

### 3.1.1. Calibration

For the calibration of the two EM27/SUN spectrometers, side-by-side measurements were conducted (compare Alberti et al. (2022)). In a first step, one of the spectrometers (SN96) was calibrated relative to the COCCON reference spectrometer (see Section 2.2.2). Afterwards, the two spectrometers that were deployed in the campaign are calibrated relative to each other.

With instrument SN96 one day of side-by-side measurements relative to the COCCON reference Instrument SN37 were performed at Karlsruhe before sending it to Thessaloniki. During the campaign, a total of 7 calibration days were performed, distributed over the whole campaign period with SN96 and SN52.

For the calibration between two spectrometers A and B, a calibration factor $k$ is determined so that.

$$1 = \text{mean}\left(\frac{XCO2_A^{raw} \cdot k_A}{XCO2_B^{raw} \cdot k_B}\right) \tag{3.1}$$

The calibration factor of the reference spectrometer SN37 is 1 by definition.

Since the recorded interferograms are not temporally synchronized, the time series of the two instruments need to be harmonized. For this purpose the data is averaged in bins of 5 minutes.

Coincident bins divided by each other are shown in Figure 3.1. The calibration factors are derived from the quotient of both measurements $q_{\mathrm{A;B}} = X\mathrm{CO_2}^A / X\mathrm{CO_2}^B$. With Equation (3.1) we get

$$q_{37;96} = k_{96}, \quad q_{96;52} = \frac{k_{52}}{k_{96}} \tag{3.2}$$

The calibration factors for $X\mathrm{CO_2}$, $\mathrm{CH_4}$, CO and $\mathrm{H_2O}$ were determined. All are listed in Table 3.1. Furthermore, the calibration factors were estimated individually for each day of the side-by-side observations to estimate the systematic uncertainty due to calibration

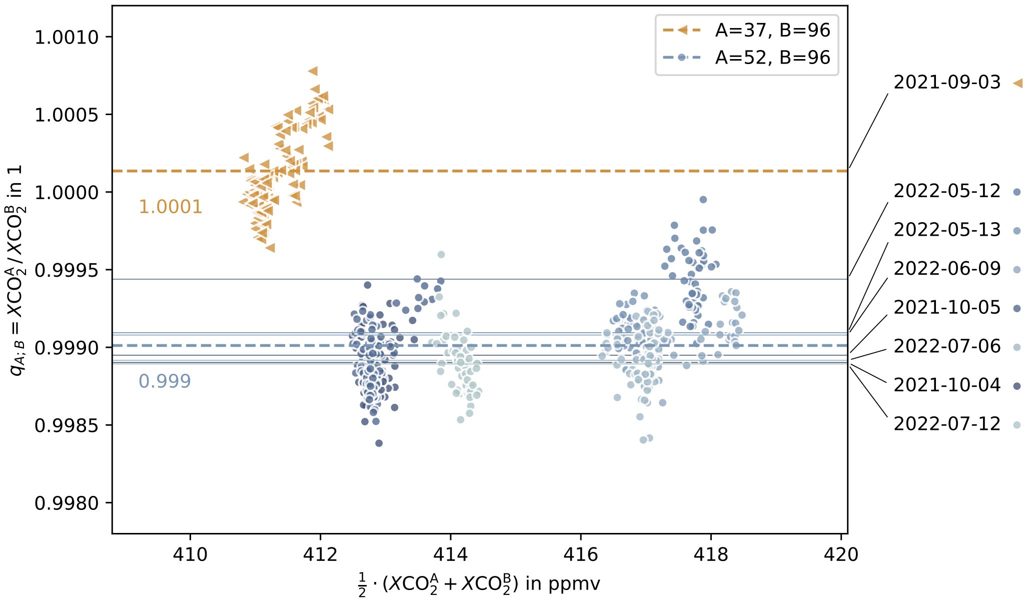


**Figure 3.1.: Side-by-Side observations for calibration.** The quotient of two coincident measurements is plotted in dependency of the pair of observations. The mean over all quotients is indicated as line. The orange triangles represent the measurements from one day of observations in Karlsruhe, where side-by-side measurements between SN37 (the COCCON reference spectrometer) and SN96 were performed. The observations from 7 days of side-by-side measurements between SN96 and SN52 are shown as blue dots, with different hues of blue for the different days of observation. For each individual day the mean of all quotients is marked a thin solid line, the mean over the whole sample of observations is indicated by a thicker dashed line.

for the side-by-side observations. These single-day calibration factors are shown together with the estimate using all observations in Figure 3.1. Here, only the calibration of $XCO_2$ is shown, as this work is focused on this greenhouse gas. The deviation of the singe-day calibration factors can be taken as estimate for the measurement uncertainty due to the calibration. There is a maximum difference between relative calibration factors of 0.05 % which corresponds to an uncertainty of 0.22 ppmv for the calibration period.

### 3.1.2. Portable Setup

Jochen Gross designed and constructed a portable setup to allow an autonomous operation of one of the spectrometers. This setup is shown in Figure 3.2 and the technical details will be explained in the following.

**Figure 3.2.: Portable setup of SN96.** The pictures show the setup that allowed mobile deployment of SN96 on a daily basis. The first image shows the whole setup, with the solar panels in the front, and the spectrometer in operation in the back. The spectrometer is shielded from sunlight with foam. The small box next to the spectrometer is shown in more detail in the second image. It contains the electronic control of the power supply and the buffer battery. Photography by Pablo Schmid and Lena Feld.

As the EM27/SUN is operated during sunny conditions, a solar power supply solution is therefore convenient. To power the spectrometer and the laptop for controlling the measurements an output of 100 W is sufficient. Two foldable solar panels with a peak output of 100 W respectively were selected for this purpose.

For being able to operate the spectrometer also in the morning and evening hours or during short intervals of shadow, a battery was included in the setup as buffer. A reasonably priced solution was a lead-gel battery with a voltage of 12 V and a capacity of 26 Ah. Lead-gel batteries are lighter than standard car batteries, making them easier to transport.

The centerpiece of the setup is the solar charge controller. A commercial solution of the type *MPPT plus 20A* with a sufficient input and output voltage, is used for this purpose. In addition, it enhances the efficiency by identifying the ideal operation of the solar panels and protects the battery from exhaustive discharge or overcharge.

### 3.1.3. Measurement Sites

As explained above, two spectrometers were available for the campaign in Thessaloniki: SN52 (in Thessaloniki since 2019) and SN96 (brought to Thessaloniki for the campaign).

For the campaign, the operation of SN52 was continued on the campus, which is in a central position of the city (see Figure 3.3). The portable setup for SN96 allowed a mobile positioning of the spectrometer during the campaign. Different sites were selected before the campaign; the following criteria were required for the sites:

- Accessible by car
- Open Area without surrounding trees or buildings to allow for all-day operation
- No disturbance of other people (as e.g. on a busy parking lot)

To be sensitive to a large area of the city, a set of observation sites around the border of the city was selected, covering several directions. This complements the central observation site. The selected observation sites are shown in Figure 3.3. SN96 was transported to the positions at the city boundary, while SN52 stayed on the campus for most measurement days. On some of the days SN52 was brought to Thermi due to safety issues.

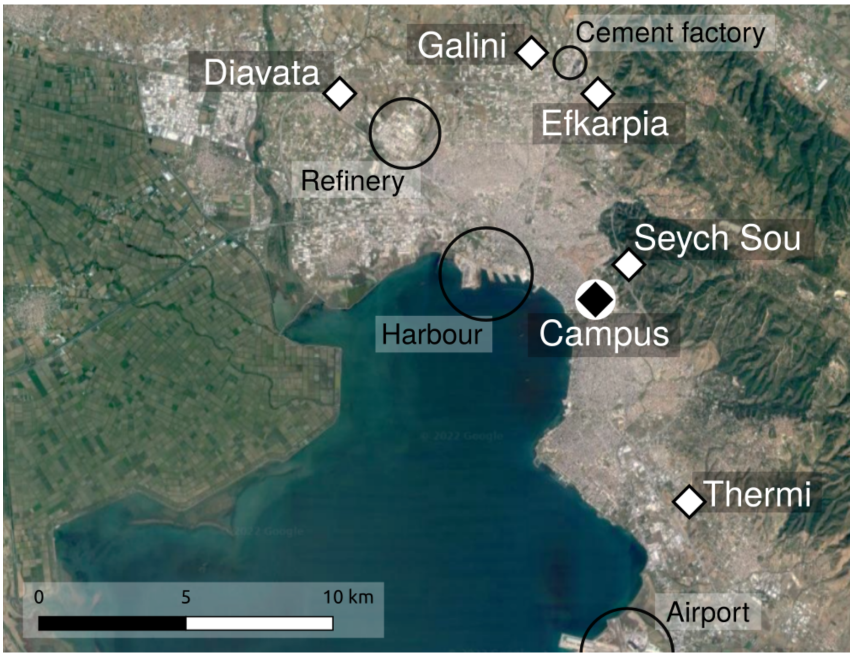


**Figure 3.3.: Observation sites for the Thessaloniki campaign.** The central observation site on the campus is shown as a black diamond, and the sites to which the portable spectrometer was transported are shown as white diamonds. Anticipated emission hot spots are depicted as black circles. The background satellite image is taken from Google Maps (2022).

### 3.1.4. Surface Pressure Estimation

As explained in Section 2.2.2, the pressure is an important prerequisite to obtain column abundances from the recorded interferograms. COCCON's sister network the TCCON requires an accuracy of 0.3 hPa (TCCON Wiki, 2021) and COCCON recommends the same quality. To achieve the recommended pressure data quality, a calibrated Vaisala PTB330 sensor with an accuracy of 0.15 hPa was positioned on the campus of the Aristotle University of Thessaloniki (AUTh), at the same height as the observations. Unfortunately there were gaps in the pressure record during the campaign period, therefore the pressure record from the AUTh meteorological station was used instead, and calibrated to the height of the Vaisala pressure sensor using overlapping data. This combined pressure will be called *reference pressure* in the following.

To also have a precise pressure at the various deployment sites of SN96, a height-correction is needed. To obtain these height corrections for the different observation sites, a portable pressure sensor was transported from the site of the reference pressure to the various observation sites. This way, a direct measurement of the required height correction could be performed.

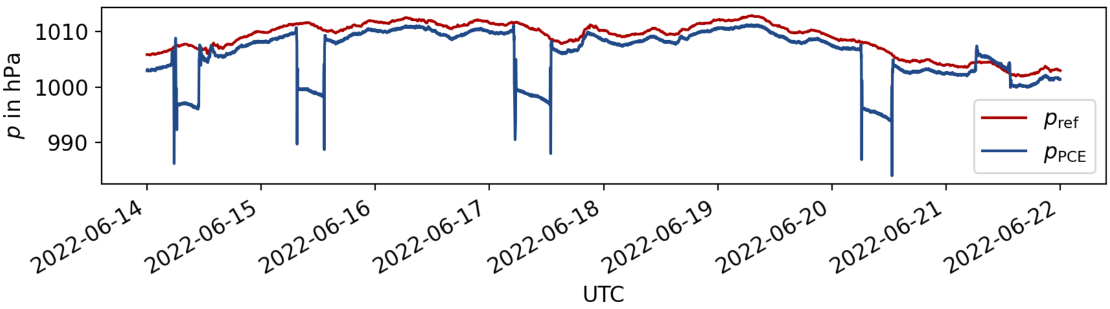


**Figure 3.4.: Pressure sensor transported to different sites in comparison to the reference pressure.** The red line shows the pressure record of the mobile pressure sensor that was transported to the different observation sites. In comparison the reference pressure is shown as blue line. The transportation events are clearly visible by sudden changes in the pressure. Note, that the offset between the red and blue curve changes at the transportation events. The figure is taken from Schmid (2023).

**Table 3.2.: Pressure factors determined using a mobile pressure sensor.** The pressure values determined by transporting a pressure sensor to the different sites are listed. Also the GPS heights are given. The values were calculated by Schmid (2023).

| Site | Pressure factor | GPS height |
|---|---:|---:|
| Galini | 0.99272 | 132.5 |
| Efkarpía | 0.99112 | 138.8 |
| Seych Sou | 0.98930 | 157.1 |
| Diavatá | 1.00461 | 17.2 |
| Thermi | 0.99777 | 84.2 |
| Campus (Meteorology) | 1.00123 | 53 |
| Campus (Physics) | 1.00000 | 62.9 |

A time series of the portable pressure sensor containing several transportation events is shown in comparison to the reference pressure in Figure 3.4. For several days a transportation to a different site (i.e. with a different offset) and a return of the device to the previous location can be recognized. Short periods with large variations are visible at the time of transportation. From these time series, scaling factors for the height correction of the mobile locations were determined. All pressure factors for the campaign were derived by Schmid (2023) and are listed in Table 3.2.

### 3.1.5. Small-distance Observations to Investigate Local Variability

Two of the measurement days were dedicated to small-distance observations, to test if the observed variability originates from sources in the proximity. The two spectrometers

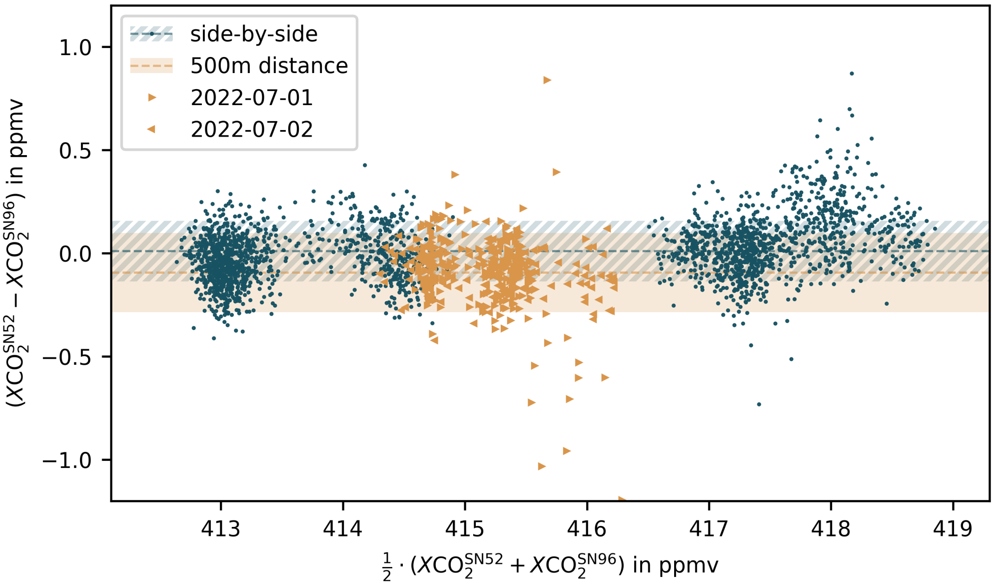


**Figure 3.5.: Comparison of calibration measurements with small-distance observations.** Side-by-side measurements (blue dots) are compared to measurements with a distance of 500 m (orange triangles). The difference between coincident measurements is shown against the mean. The mean of all values is shown as a line and the standard deviation as a transparent band. The seven days dedicated for side-by-side measurements are also displayed in Figure 3.1. Two measurement days (2022-07-01 and 2022-07-02) were dedicated for measurements with 500 m distance.

were deployed with a distance of 500 m. Both locations were on the campus in the central part of the city (see Figure 3.3). These measurements were compared to side-by-side observations. The comparison is depicted in Figure 3.5. The small-distance measurements have a slightly higher standard deviation (0.19 ppmv) compared to the side-by-side data (0.15 ppmv). In addition, a small bias of 0.1 ppmv is evident in this observation. This bias is within the estimated calibration uncertainty. In summary, the comparison indicates that the 500 m observations align with the side-by-side measurements, supporting the assumption that observed variations are dominated by larger-scale sources rather than small sources in the vicinity.

### 3.1.6. Overview over the Dataset

The dataset was published at ESA's EVDC Data Repository (Feld et al., 2024a). In the following, a short overview over the recorded dataset is presented.

During the campaign period from October 4 to 22 2021 and from May 12 to July 12, a total of 30 measurement days were collected with both instruments in parallel. An overview of the data is visualized in Figure 3.6. In Appendix A all measurement days are

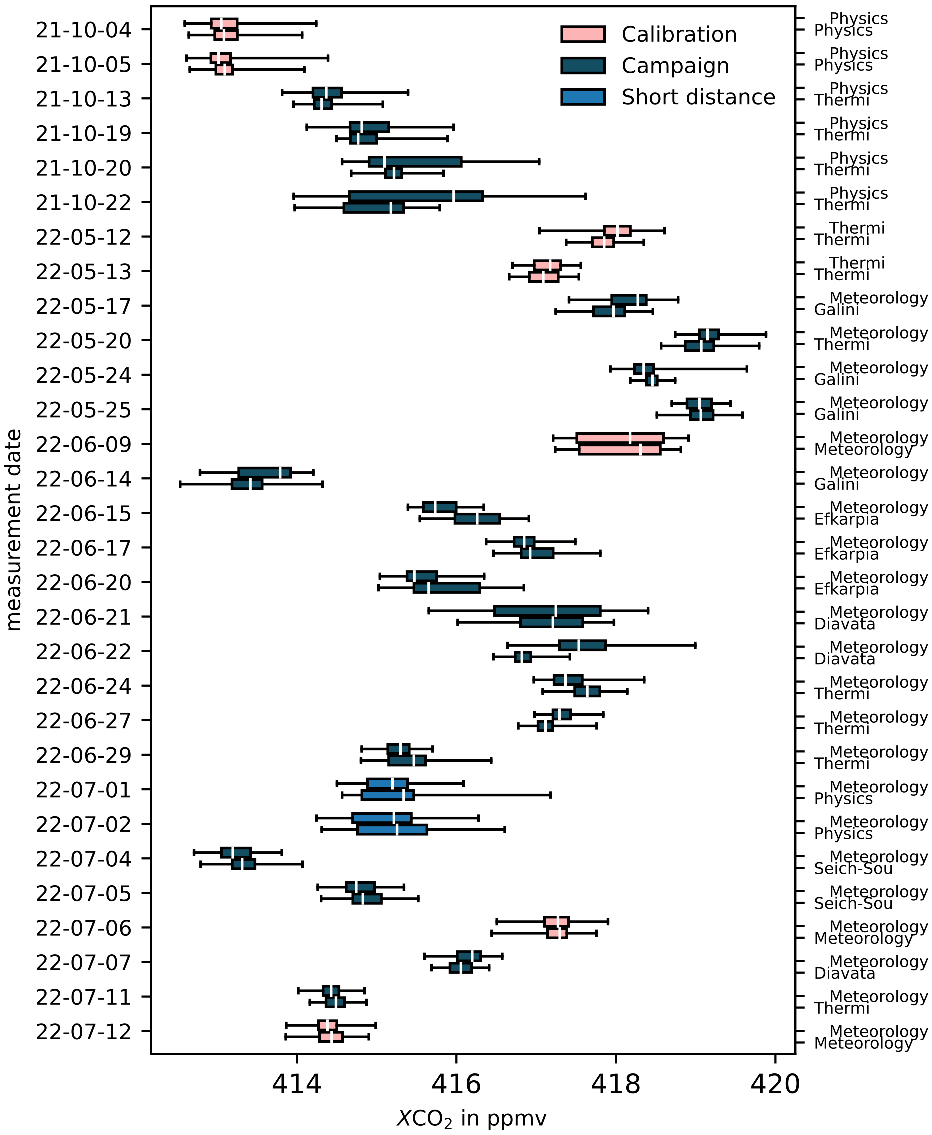


**Figure 3.6.: Overview over the collected dataset.** For each day and pair of box and whisker plots is shown, representing the distributions of $XCO_2$ observed at the given day. The left of two box plots always represents SN52 (central spectrometer), the right SN96 (portable spectrometer). Each box ranges from the 25 to the 75 percentile, the whiskers mark the maximum and minimum observed value. The positions of the instruments during the campaign are color coded, with the legend on the top left. In Addition, the abbreviations of the observation sites are also attached to the top axis for clarity.

listed. From the 30 days recorded, 7 were used for calibration by performing side-by-side measurements (see Section 3.1.1).

One of the remaining 23 days was excluded because no initial data was available for the corresponding simulations, which is explained in Chapter 4. The selected 22 days contain a total of 179 hours of observation. As described in Section 3.1.5, 2 of these days (2022-07-01 and 2022-07-02) were dedicated to test the small scale variability.

While the background level between different days distinctly varies, the diurnal variations are smaller. The difference between the maximum and minimum observed $XCO_2$ at a given date, ranged from 0.56 ppmv to 3.66 ppmv, with a median of 1.3 ppmv. The maximum observed difference between both spectrometers $\Delta XCO_2$ is 2.03 ppmv. The median of $|\Delta XCO_2|$ is only 0.17 ppmv. Compared to previous campaigns targeting larger cities, only small-scale emission signatures are present in the dataset (e.g. Ohyama et al. (2023) observed gradients up to 9.5 ppmv during the Tokyo campaign).

## 3.2. Atomization of the Retrieval

Processing the spectra with PROFFAST required considerable manual effort. The PROFFAST retrieval workflow is described in Section 2.2.3. For each day of recorded interferograms, several steps were needed.

Together with Benedikt Herkommer, I developed an atomization framework to minimize the workload of data processing. The Python interface PROFFASTpylot can handle a variety of different processing scenarios. It is a reimplementation and a comprehensive expansion of a previous set of Python scripts created by Qiansi Tu and Darko Dubravica that partly automatized the retrieval workflow.

We published an article about the interface in the *Journal of Open Sources Software (JOSS)*, see Feld et al. (2024b). The following is based on this article and the indented blocks are directly quoted from this work.

> The manual operation of PROFFAST has the following workflow: For each of the above described steps, the user has to create input files with the relevant parameters. A list of interferograms to be processed and specific input parameters are required for *preprocess*. Secondly, *pcxs* requires the specification of the atmospheric conditions. Finally, for *invers* the output generated by the previous steps has to be listed. PROFFAST creates several output files; only a single day can be processed at a time. The task repetition and file organization makes the processing of longer measurement series work intensive and prone to application errors. The following requirements are addressed by PROFFASTpylot:
>
> - Significant improvement of usability by enabling a single set of input parameters and the simultaneous processing of many measurement days.

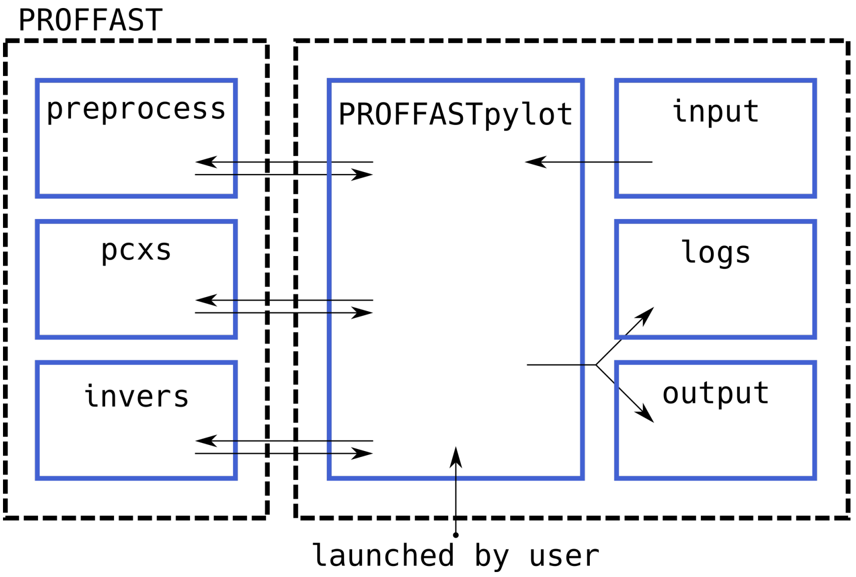


**Figure 3.7.: Schematic overview of the interactions of the PROFFASTpylot interface with PROFFAST.** The three parts of the PROFFAST algorithm are shown on the left hand side. The PROFFASTpylot takes over all direct interactions with the PROFFAST Fortran routines and organizes the produced logging and output files. The figure is taken from Feld et al. (2024b).

> - Untangling of raw data, processing files and output.
> - Reduction of application errors by introducing various cross-checks and user warnings.
> - Flexibility to allow experimental use cases besides the COCCON standard.
>
> Already during development we received many comments and questions from the global COCCON user community indicating the great interest in this tool. PROFFASTpylot has already been used by Schmid (2023) and Herkommer (2024).

The interface is widely used in the COCCON community and was also included in a larger automation framework (Makowski, 2024).

PROFFASTpylot replicates all prior user interactions with PROFFAST. The primary functionalities encompass extracting the parameters from a singular input file, interpolating supplementary data such as ground pressure records and generating input files for distinct processing stages of PROFFAST. PROFFASTpylot compiles a comprehensive log for each execution, gathers the outputs from each segment of PROFFAST and merges the final data from all processed days into a unified output file. The interactions are schematically shown in Figure 3.7.

To ensure a simple user experience and fast error detection several measures have been taken:

- The empirical instrumental parameters (ILS parameters) (Alberti et al., 2022) are taken automatically from an internal list.
- For auxiliary data, crosschecks are implemented that generate a warning or a controlled program stop (e.g. checking the correct location of atmospheric a-priori files).
- Automatic handling of different time zones in interferograms and auxiliary data.
- Correct handling of various pressure records (different sampling intervals or data formats).
- Different levels of logging ("warning", "info", and "debug") help readability and troubleshooting.

The program is technically organized as follows: The PROFFASTpylot consists of three layers that are interconnected, along with an independent fourth class. The initial layer, *prepare* is responsible for collecting of all the days to be processed and producing the necessary input files for PROFFAST. Within this layer, there is a function call to the independent *pressure* class, which facilitates the interpolation of the provided pressure data. The role of the *filemover* class is to ensure the provision of essential input data to each segment and to transfer intermediary files to the subsequent phase. The *pylot* class is designed to engage with the user, offering methods to initiate individual sections of PROFFAST or execute them sequentially upon a single command.

We added a comprehensible in-code documentation and also in depth user instructions (KIT IMK-ASF, 2024) to enable others to use and expand our work for their own projects. PROFFASTpylot is an open source project, licensed under the GNU General Public License version 3.

# 4. Corresponding Transport Simulations

To interpret the recorded data from the campaign, a corresponding ICON-ART simulation is performed, predicting high-resolution wind fields and calculating the transported trace gases starting from an emission inventory.

The simulation setup is explained in Section 4.1. In Section 4.2 observed wind and pressure data is compared to the simulation. Furthermore, the variability in water vapor is compared to the water columns co-observed by the EM27/SUn. In Section 4.3 the simulated increase in $X\mathrm{CO}_2$ is investigated.

## 4.1. Simulation Setup

In the following, the simulation setup will be presented. It is oriented at the operational forecast system for Germany that was explained earlier. The initialization of the domain is first explained (see Section 4.1.1), followed by an explanation of the inclusion of sources and sinks in the model (Section 4.1.2). Section 4.1.3 provides an overview of the derivation of observed quantities from available model variables.

### 4.1.1. Domain and Initialization

As in the operational D2 setup described in Section 2.3, the model setup uses a limited area mode (LAM) domain that is driven by initial and boundary conditions from the operational ICON-EU forecast. The ICON-EU domain has a resolution of 6.5 km.

In contrast to the ICON-D2 domain, the grid I generated for Thessaloniki is a R19B7 grid and has a higher resolution of 1 km. It is centered around Thessaloniki (40.6 °N, 23.0 °E) with a half width of 1 degree in longitudinal and 0.8 degree in latitudinal direction. The number of cells of the two-dimensional grid is 7250. The generated domain is visualized in Figure 4.1. The domain is simulated with 65 vertical layers covering the altitude range up to 22 km, analogous to the ICON-D2 setup. The required internal model time step is

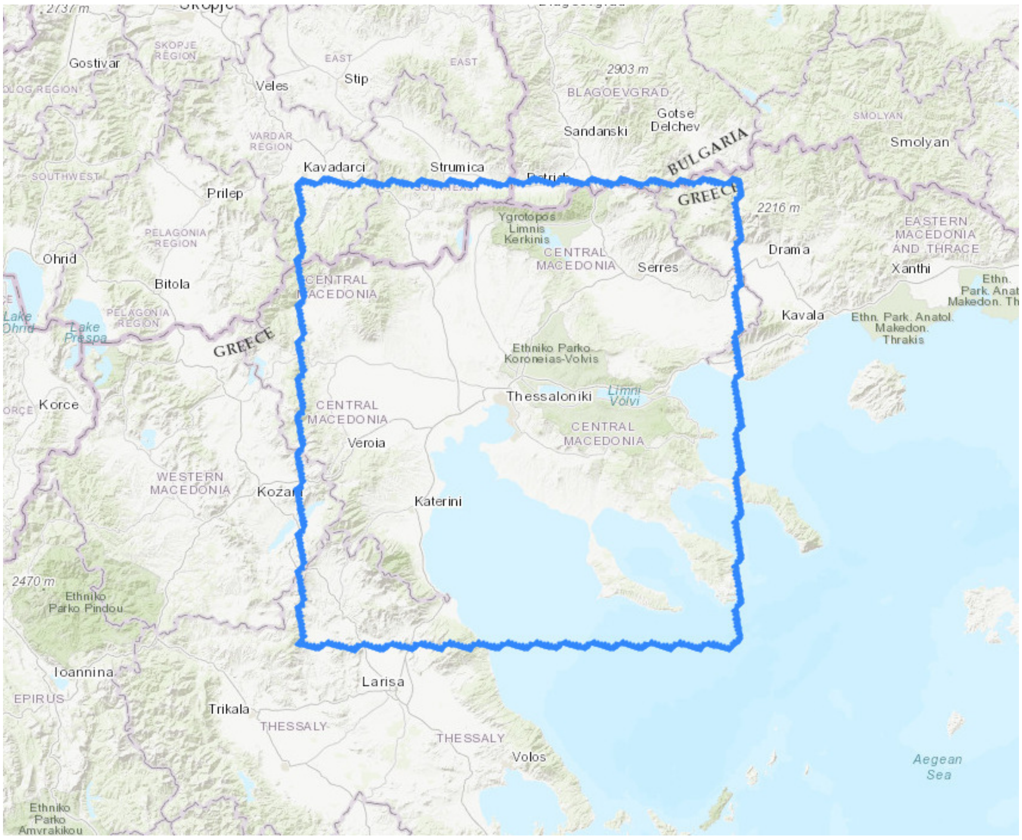


**Figure 4.1.: Simulation domain around Thessaloniki.** The borders of the 1 km resolution domain centered around Thessaloniki are shown in blue over a map of northern Greece. The image is generated by the DWD ICON tools (Prill, 2020).

determined by the resolution. The 1 km resolution used in the present setup requires a time step of 10 seconds.

For each day of measurements, a separate simulation is generated, starting at 3 UTC. The initial meteorological data for the domain is generated from the ICON-EU forecast, which was also used for the 3-hourly boundary condition files that are interpolated to a boundary region of the LAM grid. The domain and the initial and boundary conditions are generated with the DWD ICON tools (Prill, 2020).

$CO_2$ is initialized with 0 ppmv. Furthermore, no $CO_2$ is transported into the domain from the borders. Thereby, only the emissions from sources within the domain are simulated. The strategy for $CO_2$ background estimation will be explained in Section 5.1.

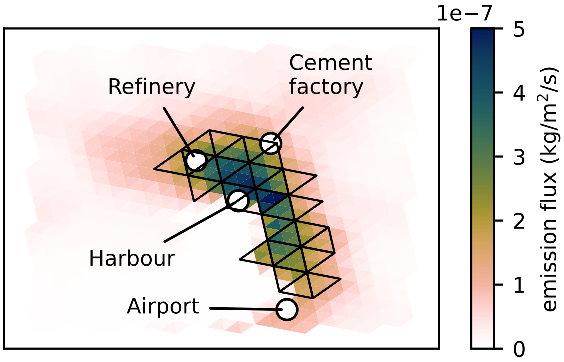


**Figure 4.2.: The ODIAC inventory remapped to the ICON grid.** The remapped inventory is shown as a map plot. The $CO_2$ emission strength is indicated by color. The black lines show the separation of the grid into different source areas to allow a scaling of the city emissions during the post-processing.

### 4.1.2. Sources and Sinks

The $CO_2$ sources and sinks were obtained from different external datasets.

#### Anthropogenic Source

For the anthropogenic source, different datasets are available. I chose the *ODIAC* inventory (Oda and Maksyutov, 2011) because of its high resolution. ODIAC is a global dataset reporting monthly averaged anthropogenic $CO_2$ emissions with a resolution of approximately 1 km. The inventory was remapped to the ICON grid using emiproc (C2SM-RCM, 2023). For the simulation the version ODIAC2020b for May 2019 was used.

To enable a post-scaling of the anthropogenic emissions in the city area, the ODIAC grid was separated over the city area into different source regions – called emission pixels in the following. For each of the emission pixels a separate tracer was simulated. Thereby, a combined total $CO_2$ could be derived from a linear combination of all tracers. This separation is the basis for the emission estimation using a synthesis approach (see Section 2.3.1). The scaling approach will be described in Section 5.2.

I chose 30 pixels to avoid making the simulations too computationally costly and to limit the degrees of freedom for a late optimization (Section 5.2). To cover a large area of the city, four grid cells were combined to one emission pixel. The separation is shown in Figure 4.2.

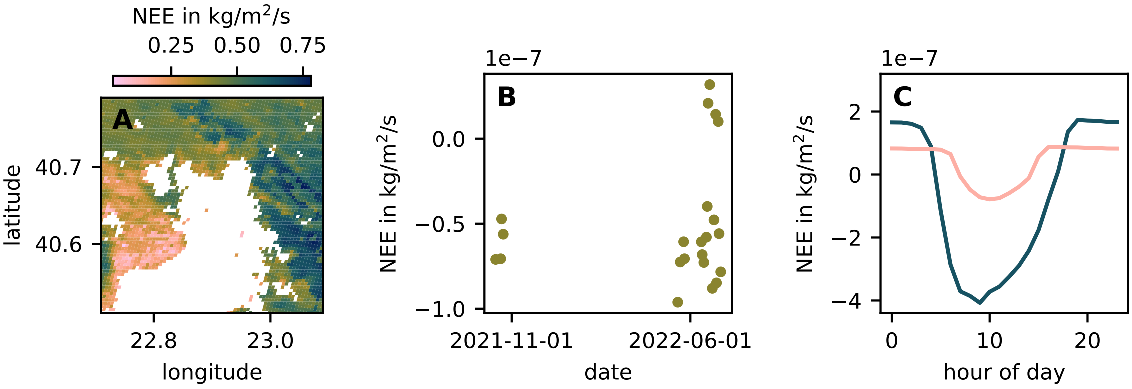


**Figure 4.3.: The NEE from three biogenic datasets.** The three datasets from which the NEE in the model is derived are shown in the Figure. The MODIS inventory (a) was used for its lateral resolution. The timely variation are deduced from two datasets. The daily variations in the SMAP dataset are shown for each day of observation in panel (b). The monthly averaged diurnal cycles from the FLUXCOM X-BASE dataset are shown for October (orange) and May (blue) in panel (c).

#### Biogenic Sources and Sinks

Coupling a biogenic model to the simulation is beyond the scope of this work. As an alternative approach, different available datasets were combined to construct a product with a high lateral resolution and a diurnal cycle of the net ecosystem exchange (NEE). A satellite product derived from the Moderate Resolution Imaging Spectroradiometer (*MODIS*) provides the yearly average of NEE with a resolution of 500 m. The product MOD17A3HGF Version 6.1 (Running and Zhao, 2021) is derived from space-borne observations of the Leaf Area Index and the fraction of photosynthetically active radiation. Scaling factors to add a timely variation to this dataset were derived from *SMAP* by Kimball et al. (2022), which provides daily averages on a 0.1 x 0.1 degree scale. The diurnal variation was derived from *FLUXCOM X-BASE* that provides a monthly averaged diurnal cycle (Weber, 2023). For SMAP and FLUXCOM X-BASE the daily variation was averaged over the whole domain. The three datasets can be seen in Figure 4.3.

### 4.1.3. Derived Observation Quantities

The target quantity that is derived from the EM27/SUN observations is the column-averaged dry-air molar fraction, denoted $X_{\text{gas}}^{\text{col. av.}}$ (see Equation (2.13)). This quantity will be expressed though model quantities in the following. The required variables that are available in ICON are listed in Table 4.1.

**Table 4.1.: Quantities to determine $X_{\mathrm{gas}}^{\mathrm{col.\ av.}}$ from the ICON-ART output.** The table lists the available model quantities available in ICON-ART that are used to determine $X\mathrm{CO}_2$ from the model output.

| Variable | Description |
|---|---|
| $T$ | Temperature in grid cell |
| $p$ | Pressure in grid cell |
| $X_{gas}$ | Volume mixing ratio with respect to wet air |
| $q_v$ | Specific Humidity |

The volume mixing ratio is defined as

$$X_{\mathrm{gas}} = \frac{V_{\mathrm{gas}}}{V_{\mathrm{wet\ air}}} \ . \tag{4.1}$$

Here, the partial volume of the trace gas is $V_{\mathrm{gas}}$. The total volume of air (including the water vapor content) is referenced as $V_{\mathrm{wet\ air}}$. The nomenclature is chosen for clarification, as the EM27/SUN records column averaged volume mixing ratios with respect to dry air while ICON-ART tracers are defined relative to wet air as in Equation (4.1).

The specific humidity is the mass maxing ratio of water vapor with respect to wet air.

$$q_v = \frac{m_{\mathrm{H_2O}}}{m_{\mathrm{wet\ air}}} \ , \tag{4.2}$$

with the total mass of the respective substance $m_{\mathrm{X}}$. Since all measurements and simulations were done in clear sky conditions, the other aggregation states of water are not considered.

Additionally, the ideal gas constant $R$, the Avogadro constant $N_A$ and the molar masses of dry air, water and the trace gases $M_x$ are known.

### Vertical Column of Trace Gases

The vertical column is defined as

$$VC_{\mathrm{gas}} = \int_0^{\infty} \rho_{\mathrm{gas}}^N \, \mathrm{d}z \ , \tag{4.3}$$

with the number density $\rho^N = \frac{N}{V_{\mathrm{wet\ air}}}$. From the ideal gas equation with partial volumes

$$p\, V_{\mathrm{gas}} = \frac{N_{\mathrm{gas}}}{N_{\mathrm{A}}} R\, T \tag{4.4}$$

the number density can be derived

$$\rho^N_{\text{gas}} = \frac{N_\text{A}}{R}\frac{p}{T}X_{\text{gas}} \; . \tag{4.5}$$

With this the vertical column of the gas can be expressed as

$$VC_{\text{gas}} = \frac{N_\text{A}}{R}\int_0^\infty \frac{p}{T}X_{\text{gas}}\,\mathrm{d}z \; . \tag{4.6}$$

#### Vertical Column of Water Vapor

To calculate the vertical column of water vapor we need to express $\rho^N_{\text{H}_2\text{O}}$ from model quantities. Therefore, the specific humidity is rewritten as

$$\begin{aligned} q_\text{v} =& \frac{m_{\text{H}_2\text{O}}}{m_{\text{wet air}}} \qquad (4.7)\\ =& \frac{M_{\text{H}_2\text{O}}\, n_{\text{H}_2\text{O}}}{\rho_{\text{wet air}}\, V_{\text{wet air}}} \qquad (4.8)\\ =& \frac{M_{\text{H}_2\text{O}}\, N_{\text{H}_2\text{O}}}{\rho V_{\text{wet air}}\, N_\text{A}} \qquad (4.9) \end{aligned}$$

which leads to

$$\rho^N_{\text{H}_2\text{O}} = \frac{N_{\text{H}_2\text{O}}}{V_{\text{wet air}}} = \frac{q_\text{v}\, N_\text{A}\, \rho}{M_{\text{H}_2\text{O}}} \; . \tag{4.10}$$

With this, the vertical column water vapor can be expressed as

$$\begin{aligned} VC_{\text{H}_2\text{O}} =& \int_0^\infty \rho^N_{\text{H}_2\text{O}}\mathrm{d}z \qquad (4.11)\\ =& N_A \int_0^\infty \frac{q_\text{v}\rho}{M_{\text{H}_2\text{O}}}\mathrm{d}z \; . \qquad (4.12) \end{aligned}$$

#### Vertical Column of Dry Air

With the vertical column of water, the vertical column of dry air can be calculated from the ground pressure.

$$\begin{aligned} p_{\text{sfc}} &= \frac{g}{A}m_{\text{wet air}} \qquad (4.13)\\ &= \frac{g}{A}(m_{\text{dry air}} + m_{\text{H2O}}) \qquad (4.14)\\ &= \frac{g}{A}(\frac{N_{\text{dry air}}}{N_\text{A}}M_{\text{dry air}} + \frac{N_{\text{H}_2\text{O}}}{N_\text{A}}M_{\text{H}_2\text{O}}) \qquad (4.15)\\ &= \frac{g}{N_\text{A}}(VC_{\text{dry air}}\, M_{\text{dry air}} + VC_{\text{H}_2\text{O}}\, M_{\text{H}_2\text{O}}) \qquad (4.16) \end{aligned}$$

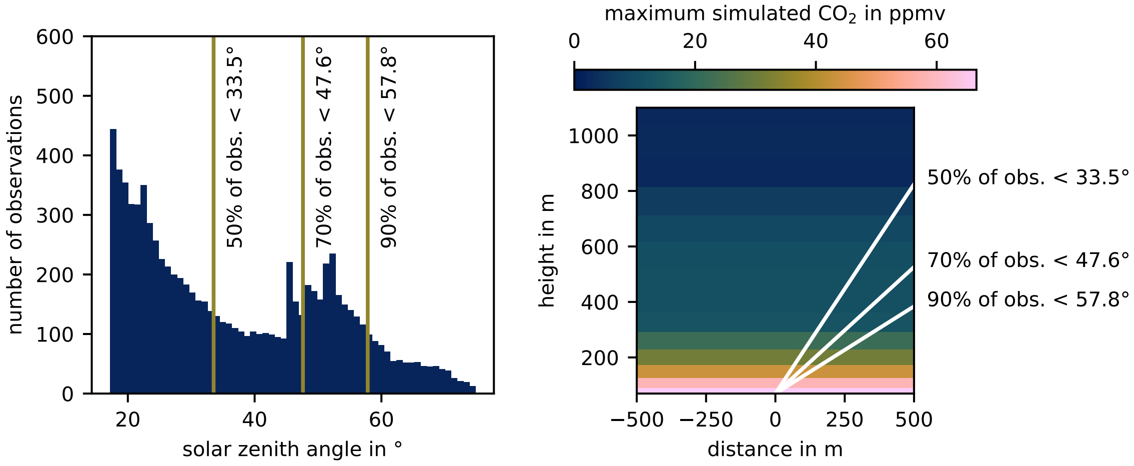


**Figure 4.4.: Distribution of the solar zenith angle.** The left panel shows the distribution of the solar zenith angle of the observations as histogram. The median, 70th and 90th percentile are indicated as vertical lines. The respective angles are shown geometrically in the right panel. The extent of the x-axis represents approximately the average resolution of the simulation. The maximum simulated value of $CO_2$ in the respective height is indicated by color.

using $VC = \frac{N}{A}$ and $M = N/N_A \cdot m$. Solving for $VC_{\text{dry air}}$ gives

$$VC_{\text{dry air}} = \frac{p_{\text{sfc}} \cdot N_A}{g \cdot M_{\text{dry air}}} - VC_{H_2O}\frac{M_{\text{H2O}}}{M_{\text{dry air}}} \,. \tag{4.17}$$

### Discretization of the Integrals

To calculate the integral as a sum over model levels, the following discretization is used

$$\int f(z)\,\mathrm{d}z \approx \sum_{\text{level=i}} (z_{\text{i+1}} - z_{\text{i}}) \cdot \frac{f_{\text{i}} + f_{\text{i+1}}}{2} \,. \tag{4.18}$$

### Neglecting the Inclination and Height-dependent Column Sensitivity

In the above derivation of the vertical column, the slope of the column is not taken into account. The simulation only has significant contributions near the surface, and most of the observation time is recorded at low solar zenith angles (SZA), which limits the effect of the slant observation angle. Figure 4.4 shows the distribution of the SZAs as a histogram. 90 % of the observation data is recorded and SZAs smaller than 57.8 ° and the median SZA of the observation is 33.5 °. The right panel of Figure 4.4 shown the maximum simulated value of $CO_2$ on the campus, which is indicated by color for the different heights. The extent of the x-axis represents the average resolution of one

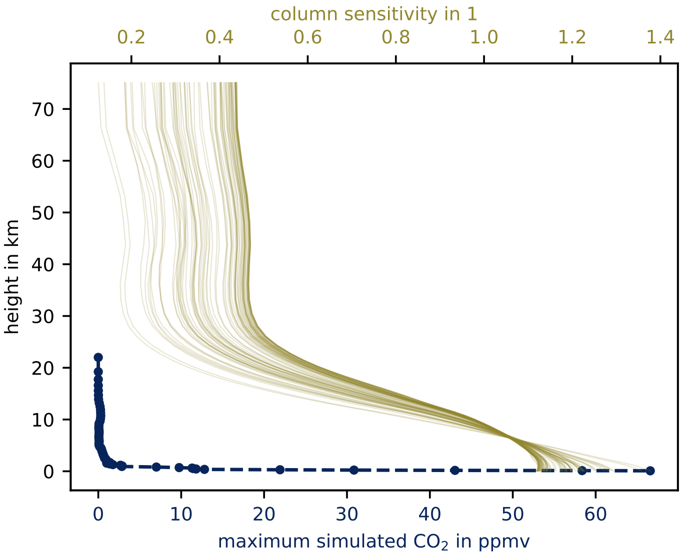


**Figure 4.5.: Column sensitivities of the campaign observation.** For each observation, the height dependent sensitivity of the column is shown as a yellow line with the respective axis shown at the top of the plot. For each height, the maximum simulated value for $CO_2$ is shown in blue, with the values at the bottom axis.

grid-cell. Even for the relatively shallow observation angle of 57.8 ° (the 90th percentile) the slant column passes the most relevant layers inside the same grid cell. Therefore, the observation angle is not considered in the calculation of $XCO_2$.

A second contribution that is not considered is the effect of the column sensitivities of the observation. Figure 4.5 shows the height-dependent sensitivity for all data points of the measurement together with the maximum simulated concentration in the model in the respective height. It is visible that only the lowermost layers of the simulated column contribute. Near the surface the column sensitivity is close to 1 for most of the observations; the median is 1.1 for the lowermost layer. Therefore, the column sensitivity is not taken into account when evaluating the simulated columns.

## 4.2. Simulated Meteorological Variables

In this section, the simulated results for wind (Section 4.2.1), pressure (4.2.2) and water vapor (4.2.3) are compared to observations to assess the model predictions.

**Table 4.2.: Available wind observations.** The table lists the three available stations providing wind observations. If known the instrument is listed. The link to the data is given if publicly available. In addition, the frequency of the data recording is listed.

| Site | Instrument | Data frequency |
|---|---|---|
| Thermi | Davis (Vantage Pro2) | 10 min |
| Campus [1] | | 10 min |
| Airport [2] | WMO (Station 16622) | 1 h |

[1] https://meteo.geo.auth.gr/en/data-availablity/

[2] https://meteostat.net/de/station/16622

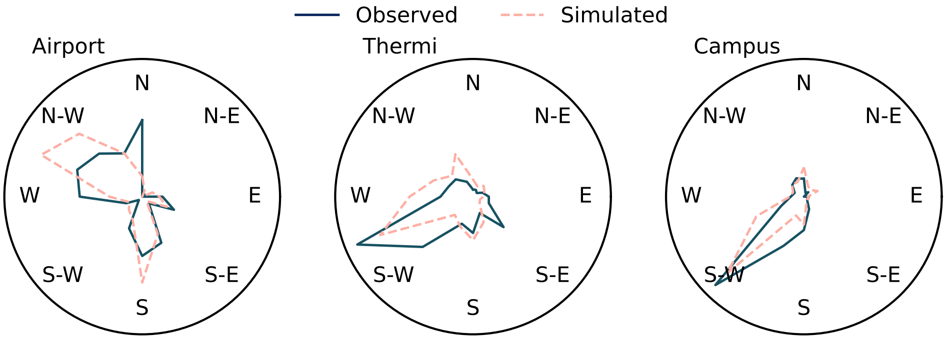


**Figure 4.6.: Wind distribution for the three observation sites.** The simulated an observed wind distributions are compared in a windrose plot, including all wind speeds into a single category. The distribution differs significantly between the sites, but in all cases the simulation (dashed orange line) matches the observation (solid blue line).

### 4.2.1. Assessing the Simulated Wind

The simulated wind field was interpolated to three locations where wind observations were available. The external meteorological measurements that were available are listed in Table 4.2. Figure 3.3 provides a reference for the site locations.

We will first look at the overall distribution of the wind over the simulated time period. Windrose plots for all three sites are shown in Figure 4.6. It is clearly visible that the wind rose at the airport differs significantly from the other two stations. At the airport, there are two prevailing wind directions (south and west-northwest), while at the Thermi and campus stations, there is only one prevailing wind direction from west to southwest. The model predictions reproduce these different distributions.

Next, we will look at the correlations between the measured values and the corresponding simulated values. The simulated and observed eastward wind $u$ and northward wind $v$ are depicted in Figure 4.7. The correlation is apparent, although there is a significant

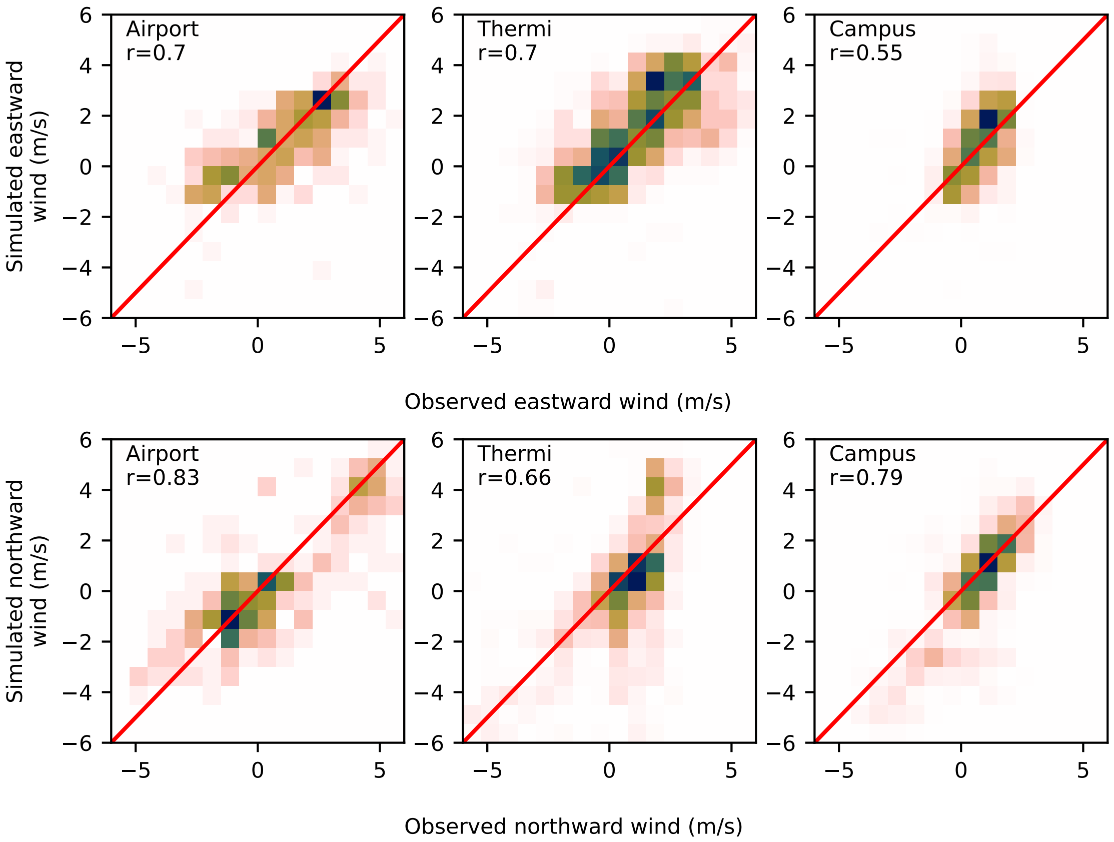


**Figure 4.7.: Comparing simulated and observed wind.** For the three available observation datasets, the comparison between simulated and observed wind is visualized as normalized two-dimensional histogram for each observation site. The eastward wind is shown in the upper panel and the westward wind is shown in the lower panel. The 1:1 line is shown in red and the Pearson correlation coefficient is annotated.

scatter of the individual points. Table 4.3 lists all Pearson correlation coefficients. With correlation coefficients between 0.55 and 0.83 the simulated wind field is clearly correlated to the observation at all sites. The large scatter becomes visible from the large standard deviation. A bias of 0.3 m/s exists for the campus and Thermi stations. Both observation sites are located in a built-up area. This possibly explains the bias in the wind observations.

### 4.2.2. Surface Pressure Observation on Campus

Analogous to the comparison between the simulated and observed wind speed, the reference pressure that was introduced in Section 3.1.4 is compared to the simulation. The high correlation coefficient of 0.99 is also visible in the two-dimensional histogram in Figure 4.8. The standard deviation of the difference between the simulated and observed datasets is also small; with 0.4 hPa it is less than 0.04 % of the average observed pressure

**Table 4.3.: Agreement between the simulated and observed wind.** The observed mean value $\bar{x}_{\mathrm{OBS}}$, the Pearson correlation coefficient $r$, the standard deviation of the difference between the observed and simulated wind $\sigma(x_{\mathrm{OBS}} - x_{\mathrm{SIM}})$, and the average bias between the simulated and observed datasets $\overline{x_{\mathrm{OBS}} - x_{\mathrm{SIM}}}$ are shown for the available observation sites and wind components.

| Site | Variable | $\bar{x}_{\mathrm{OBS}}$ in m/s | $r$ | $\sigma(x_{\mathrm{OBS}} - x_{\mathrm{SIM}})$ in m/s | $\overline{x_{\mathrm{OBS}} - x_{\mathrm{SIM}}}$ in m/s |
|---|---|---|---|---|---|
| Thermi | u | 1.23 | 0.70 | 1.46 | -0.22 |
| Thermi | v | 0.59 | 0.66 | 1.68 | 0.31 |
| Campus | u | 0.75 | 0.55 | 1.07 | -0.16 |
| Campus | v | 0.59 | 0.79 | 1.23 | 0.34 |
| Airport | u | 0.95 | 0.70 | 1.40 | -0.01 |
| Airport | v | 0.32 | 0.83 | 1.68 | 0.18 |

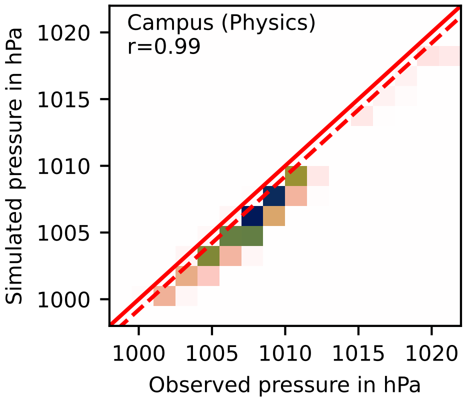


**Figure 4.8.: Comparison between simulated and observed pressure.** The reference pressure that is described in Section 3.1.4 is compared to the simulated surface pressure at the site of the observation. The correlation is visualized as normalized two-dimensional histogram. The 1:1 line is shown as a red solid line and the Pearson correlation coefficient is annotated. The red dashed line is scaled by a factor derived from the barometric formula using the height difference between the simulation and the observation as determined by GPS.

of 1007.1 hPa. However, a bias of 1.9 hPa is clearly visible. There might be an offset in height between the observation site and the height of the simulation. The surface at the interpolated location of the observation site is in a height of 69.8 m. The height of the observation determined by GPS (see Section 3.1.4) is 62.9 m. This height difference would imply a bias of 0.8 hPa. The larger actual bias is might be caused by an inaccuracy in the determined height difference by GPS, but also a bias in the simulation cannot be ruled out. Still, the pressure is very well-matched by the simulation.

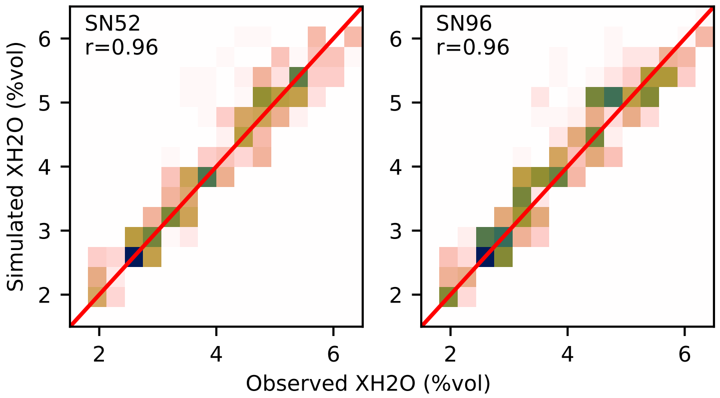


**Figure 4.9.: Comparison between simulated and observed water vapor.** The correlation between the simulated and observed column of water vapor is visualized as normalized two-dimensional histogram. The 1:1 line is shown in red and the Pearson correlation coefficient is annotated.

**Table 4.4.: Agreement between the simulated and observed $XH_2O$.** The observed mean value $\bar{x}_{\mathrm{OBS}}$, the Pearson correlation coefficient $r$, the standard deviation of the difference between the observed and simulated wind $\sigma(x_{\mathrm{OBS}} - x_{\mathrm{SIM}})$, and the average bias between the simulated and observed datasets $\overline{x_{\mathrm{OBS}} - x_{\mathrm{SIM}}}$ are shown for both spectrometers and the difference between the two.

| Site | Variable | $\bar{x}_{\mathrm{OBS}}$ in %vol | $r$ | $\sigma(x_{\mathrm{OBS}} - x_{\mathrm{SIM}})$ in %vol | $\overline{x_{\mathrm{OBS}} - x_{\mathrm{SIM}}}$ in %vol |
|---|---|---|---|---|---|
| SN52 | XH2O | 3.977 | 0.96 | 0.324 | -0.045 |
| SN96 | XH2O | 3.888 | 0.96 | 0.339 | -0.087 |
| SN52-SN96 | XH2O | 0.052 | 0.33 | 0.167 | 0.045 |

### 4.2.3. Comparison to Co-observed Water Vapor Columns

In addition to the observation of $X_{\mathrm{CO}_2}^{\text{col. av.}}$, also the column-averaged molar fraction of water vapor was observed with the EM27/SUN during the campaign. This quantity can be derived from the model output, as described in Section 4.1.3.

The comparison between simulated and observed column-averaged water vapor shows a very clear correlation with correlation coefficient of 0.96 for both, the central and the movable spectrometer (see Table 4.4). This is also clearly visible in Figure 4.7. But when looking at the difference between the two observing stations, the correlation is only 0.33, showing that the representation of small scale variations is not as good as the overall agreement.

In conclusion, the wind, pressure and water vapor observations are well-matched by the simulation, but when looking at small scale differences the correlation decreases significantly. This implies that a one-to-one agreement is not to be expected, when looking at the $XCO_2$ differences between the two stations.

## 4.3. Simulated Enhancements of $\mathrm{XCO_2}$

The simulated increases in $XCO_2$ at different locations are investigated in this section. First, the lateral structure is examined (Section 4.3.1). The relative contribution of different sectors and the influence of the distance to the emission source are discussed subsequently.

### 4.3.1. Complex Plume Structure

As we saw in the previous section, the wind has a high spatial variability. The wind field is influenced by the topography and coastline which also causes a high temporal variability. This impacts the plume over the city because the emissions are distributed in a complex structure. The simulation results for July 2, 2022 illustrate this as an example. The simulated increase in $XCO_2$ due to the emissions from the ODIAC inventory is shown as map plots in Figure 4.10 for hourly time steps. In addition, the wind field is displayed for each time step. When looking at the wind field, the high spatial variability is obvious, as there is no time step in which the wind is uniformly distributed in the displayed area. The temporal variability is also visible, as for example the wind directly over the sea in the southwest corner changes from northwest to southwest to east during the day. As a result the plume has a highly inhomogeneous structure. At the beginning of the day the wind speed is lower and the $XCO_2$ accumulates over the city in the center of the displayed area. From 7 UTC, the accumulation starts to be transported southwest and deformed. After 11 UTC it is not visible in the depicted domain anymore and the continuing emissions result in a inhomogeneous structure. This highlights the challenge of deriving emissions from an observed time series at a limited number of sites in complex terrain.

For one of the time steps, that are displayed in Figure 4.10 the vertical structure of the plume is visualized in Figure 4.11. Here, parts of the plume have been transported upwards, and towards south-western direction, although the surface wind field would have led to a transportation towards south-eastern direction. This illustrates the complexity of the wind field also in the vertical dimension. The plot also illustrates the advantage of observing the column averaged molar fraction in comparison to surface observations, because also the upper parts of the plume are visible in the observations.

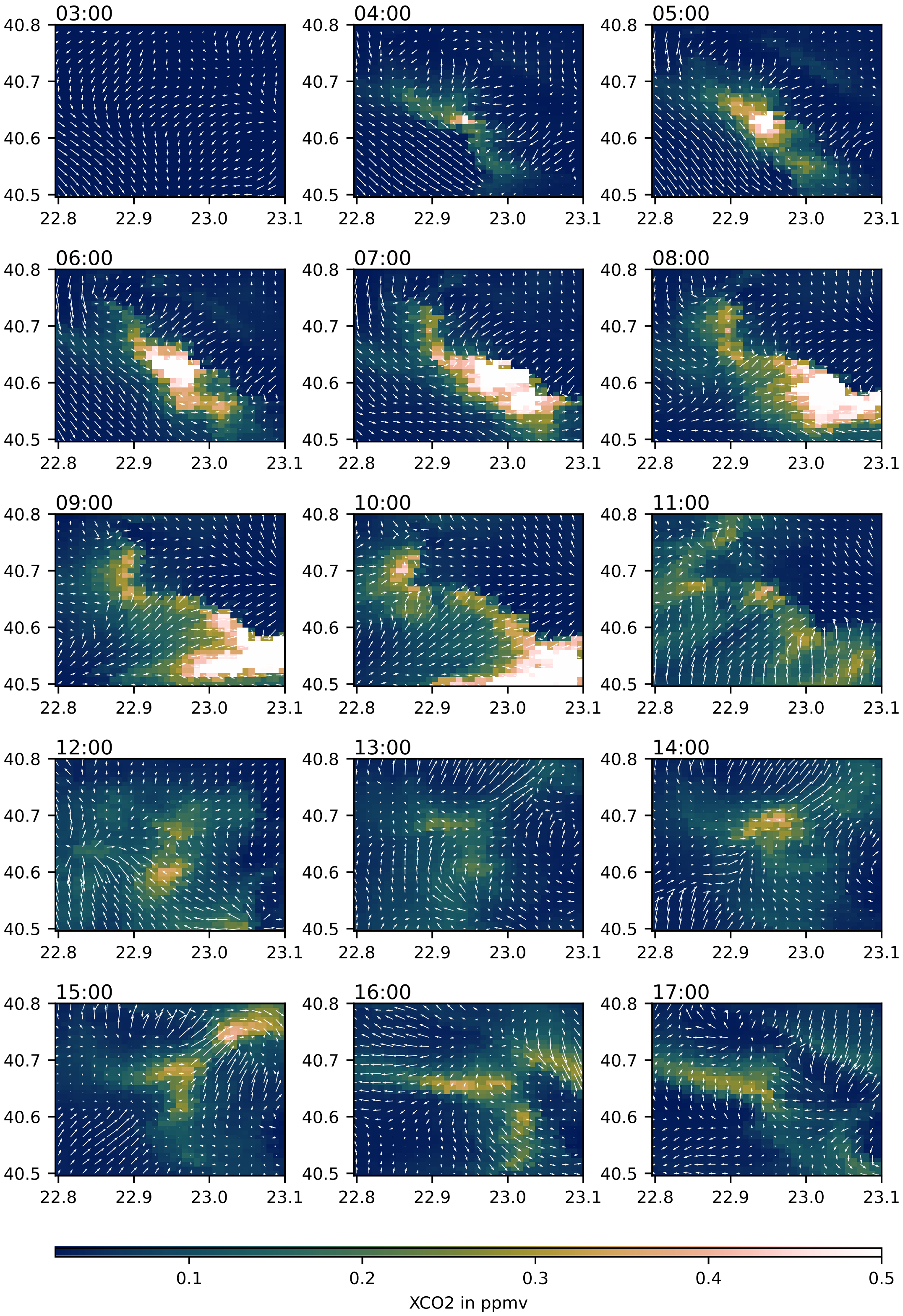


**Figure 4.10.: Anthropogenic $XCO_2$ plume for 2nd July 2022.** The lateral distribution of $XCO_2$ is shown as a map plot for different time steps. As an example, the data for July 2, 2022 is shown. The wind field is indicated by white arrows.

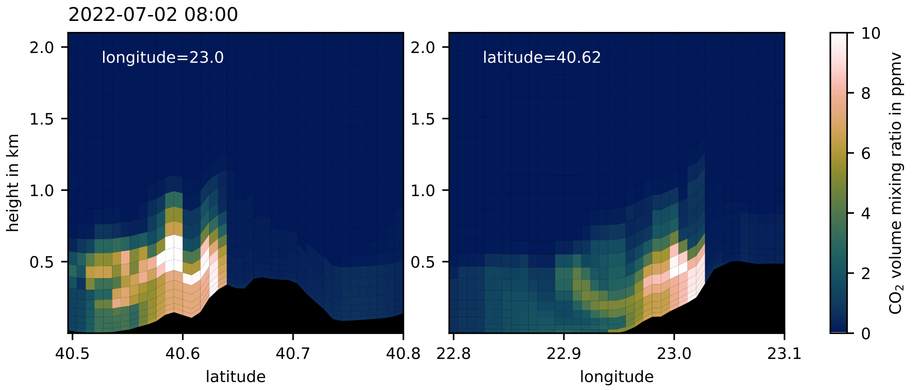


**Figure 4.11.: Vertical structure of the plume.** For 2022-07-02 8:00 UTC, the vertical structure of the plume is visualized by a latitudinal (left) and a longitudinal (right) cross section. The structure is shown up to a height of 2 km. The fixed longitude and latitude for the cross sections are given in the upper left corner of each plot. The black area represents the topography of the surface.

### 4.3.2. Influence of the Biogenic Sinks and Sources

To asses the relative influence of the biogenic sinks and sources, the simulated distributions of $XCO_2$ are investigated. The sectoral distributions of the combined biogenic component and the the anthropogenic emissions from the ODIAC inventory are shown in Figure 4.12 together with the sum of the two sectors. The three distributions are shown for the campus and for Galini – one of the sites at the city boundary. From the comparison of the three distributions can be seen that both, the biogenic and anthropogenic contribution are relevant contributors to the total enhancements as the sum of the sectors differs from the individual distributions. For the observation site in Galini, the influence of the biogenic sector seems to have a larger influence compared to the anthropogenic sector. In conclusion, the increase in $XCO_2$ is significantly influenced by the biogenic sinks and sources in the simulation.

### 4.3.3. Impact of Emissions Far from the Observer

To investigate the impact of emissions far from the city, the impact of the city emissions far from the city were reversely investigated. I defined four testing points, the edges of a 0.5 ° square centered around Thessaloniki. The four points have a distance of approximately 70 km from Thessaloniki. The testing points are depicted in Figure 4.13A.

For these points, $XCO_2$ increases due to emissions from the city are compared to the emissions from all sources in the ODIAC inventory. This is possible due to the separation of the emission inventory described in Section 4.1. Although Thessaloniki is the largest

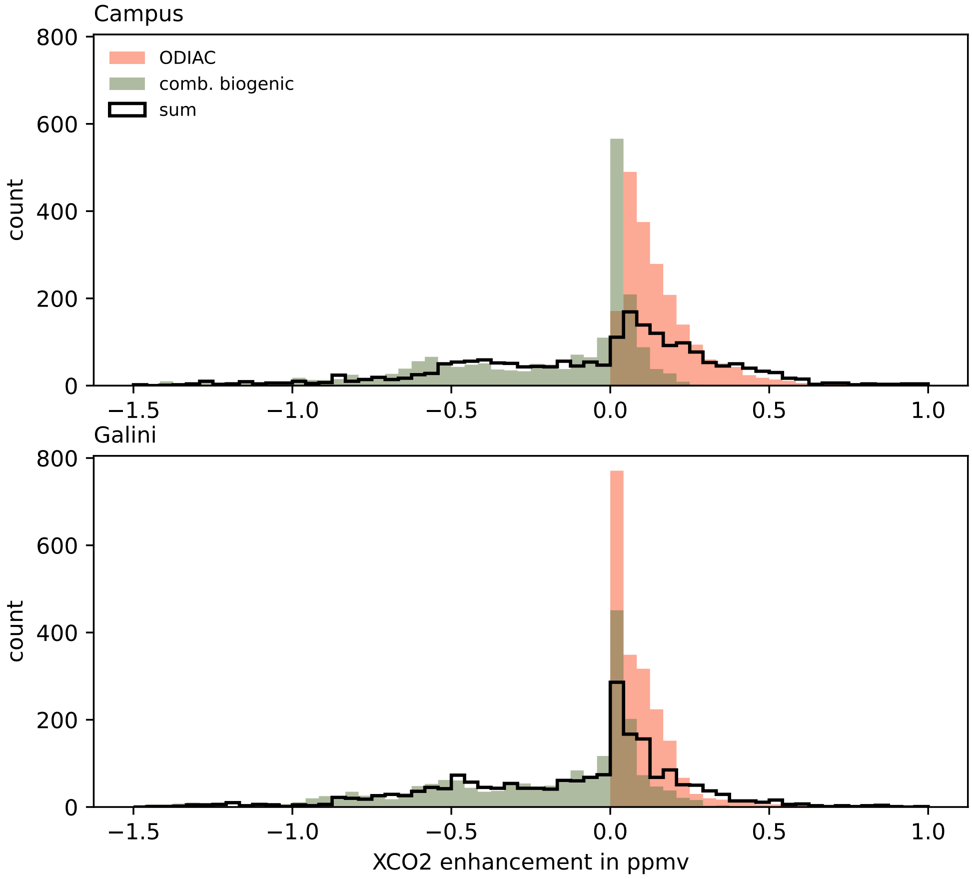


**Figure 4.12.: Increase in $XCO_2$ from different sectors.** The distribution of simulated increase in $XCO_2$ is shown as a histogram. The increase from the biogenic contribution is shown in green. The orange histogram shows the increase due to the anthropogenic emissions. The sum of both distributions is indicated by the solid black line.

emission source in the area, the enhancements from local sources are much larger, as can be seen in Figure 4.13B. In comparison, the emissions from Thessaloniki are negligible for the majority of the simulation time. This means that the observed temporal variations are mainly influenced by local effects in the simulation.

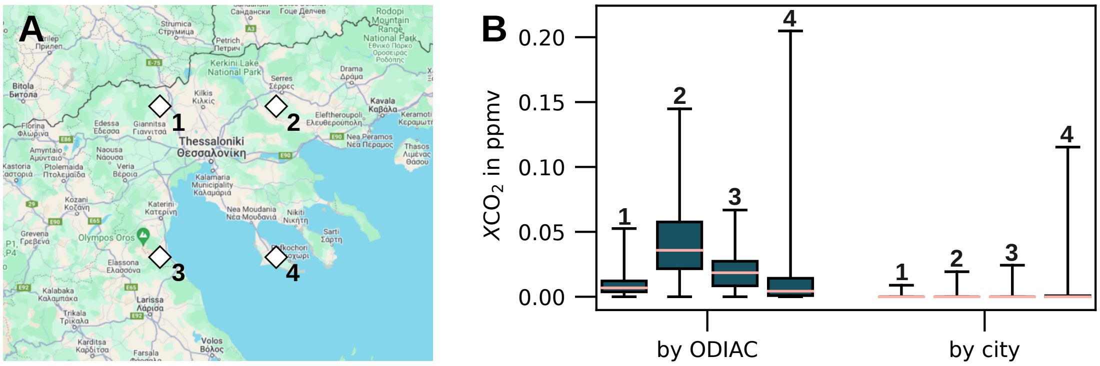


**Figure 4.13.: Testing the effect of far-distance emission sources.** The simulated increase in $XCO_2$ is examined at four points far from Thessaloniki. The points are shown in panel A. In panel B the increase is shown as box and whisker plots. The boxes extend from the 25th to the 75th percentile and the whiskers extend from the minimum to the maximum simulated enhancement. The enhancement from all sources in the ODIAC inventory is shown by the four left boxes. The four right boxes show only the emissions from the Thessaloniki urban area.

# 5. Joint Interpretation of the Observed and Simulated Data

In this chapter the simulated and the observed datasets are combined and conclusions for the anthropogenic emissions are derived from the joint interpretation of both datasets. First, the simulated and observed datasets need to be harmonized (Section 5.1). In Section 5.2 they are compared and the prior emissions are scaled to improve the agreement between both datasets. Finally, in Section 5.4, the limitation of the agreement and possible improvements are discussed. The resulting emission estimates are presented in Section 5.3. This chapter is based on Feld et al. (2024c).

## 5.1. Constructing a Harmonized Dataset

This section explains how the simulated and observed datasets were combined for further investigation. The interpolation step is explained in Section 5.1.1. As the simulation was initialized with a mixing ratio of 0 ppmv for $CO_2$, a background offset between the observed and simulated data is present. The procedure for removing this background from the observations is explained in Section 5.1.2.

### 5.1.1. Spatial Interpolation and Temporal Resampling

A spatial and temporal mismatch exists between the observed and simulated datasets. The model output represents the average over a square kilometer, while the observation is averaged over a cone of 0.3° and can be considered point-like in the horizontal dimension. The mismatch in the temporal dimension is less sensible. The model output represents a specific point in time and is not averaged over the output interval of 10 minutes. The observations are averaged over one minute, which are both relatively short intervals.

To overcome the mismatch in the spatial dimension, I applied temporal averaging to the observational data, as the lateral wind transport physically averages the concentrations in the wind direction by mixing and advecting small fluctuations. In a post-processing step, the minutely observation data were averaged in 10 min bins. For a median wind speed of 1.54 m/s observed at the campus, the wind traveled 924 meters in a period of 10 minutes, approximately matching the resolution of the simulation grid.

To correctly interpret the gradients in the observational dataset, a high precision is required in the model. Sometimes only a few grid cells separate the two observation sites, therefore a correct interpolation is needed. After the vertical column is derived from the datasets as described in Section 4.1.3, the two-dimensional field of vertical columns is interpolated to the observation sites using the `griddata` function from the NumPy python package (Harris et al., 2020).

For better data accessibility the columns interpolated to the measurement sites are stored in a separate netCDF file accessible at https://doi.org/10.5281/zenodo.12666197.

### 5.1.2. Background Removal

Previous measurement campaigns used background stations upwind of the city for removing a background offset between the observed and simulated datasets. No dedicated background observation is available for the Thessaloniki campaign.

As an alternative approach, the background is defined in the following way: For each measurement day, the 5th percentile from the joint dataset of both instruments is determined as background and subtracted, shifting the observation, so that enhancements over the defined background remain. The same procedure is applied to the simulated dataset, here the 5th percentile also takes on negative values for days with a strong influence of the biogenic sink, because the simulation was initialized with 0 ppmv. This way the distributions are shifted to the same level. The resulting quantity will be referenced as $X\mathrm{CO}_2'$ in the following. This strategy implies two assumptions

1. The background concentration is constant over the measurement period of one day.
2. The background is equal for both stations.

The observation times ranged from 4 to 7 hours a day. In this relatively short time, the assumption of a constant background is plausible. Regarding the second assumption, most of the background coming from closer sources is still included in the domain and therefore contained in the simulation dataset. There might be an effect of longer transportation times, but for these, the atmosphere is expected to be better mixed. The

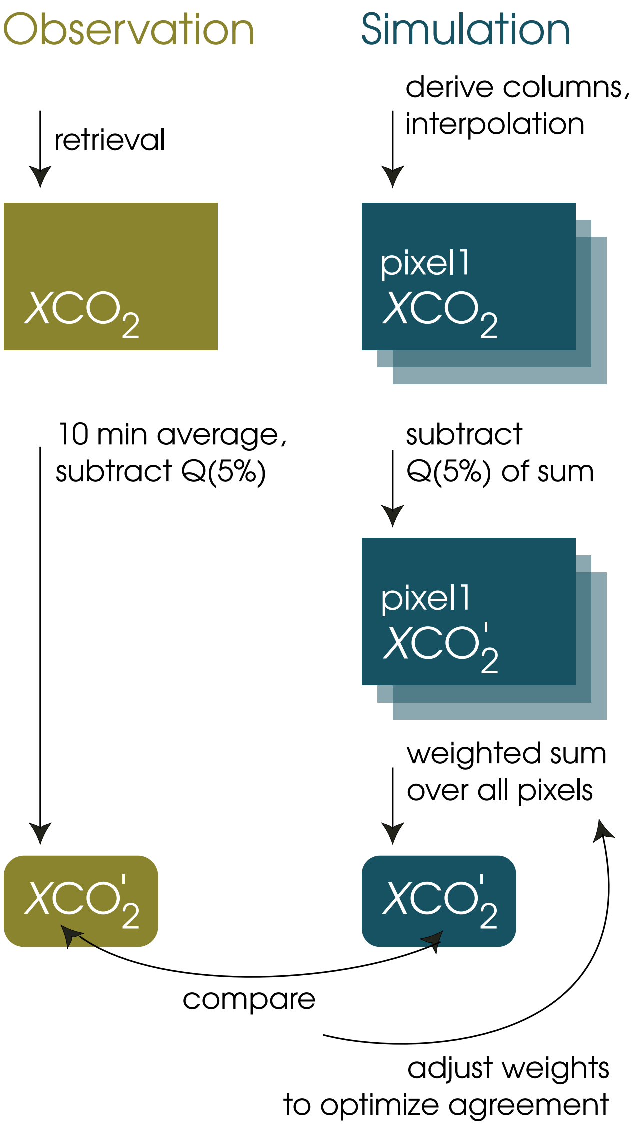


**Figure 5.1.: Creation of a harmonized dataset.** The processing steps to make the observed and simulated data comparable is illustrated. The left part of chart depicts the steps applied to the observational data, indicated in green. The right part shows the steps for the simulated data in blue. The loop in the lowermost part represents the fitting procedure.

approach is similar to the background determination by Ohyama et al. (2023). They used the 5th percentile of a dedicated background measurement as baseline for a given day.

All processing steps for both, the simulation and the observations are illustrated in Figure 5.1. With this, two comparable datasets are constructed, that will be investigated in more detail in the following section.

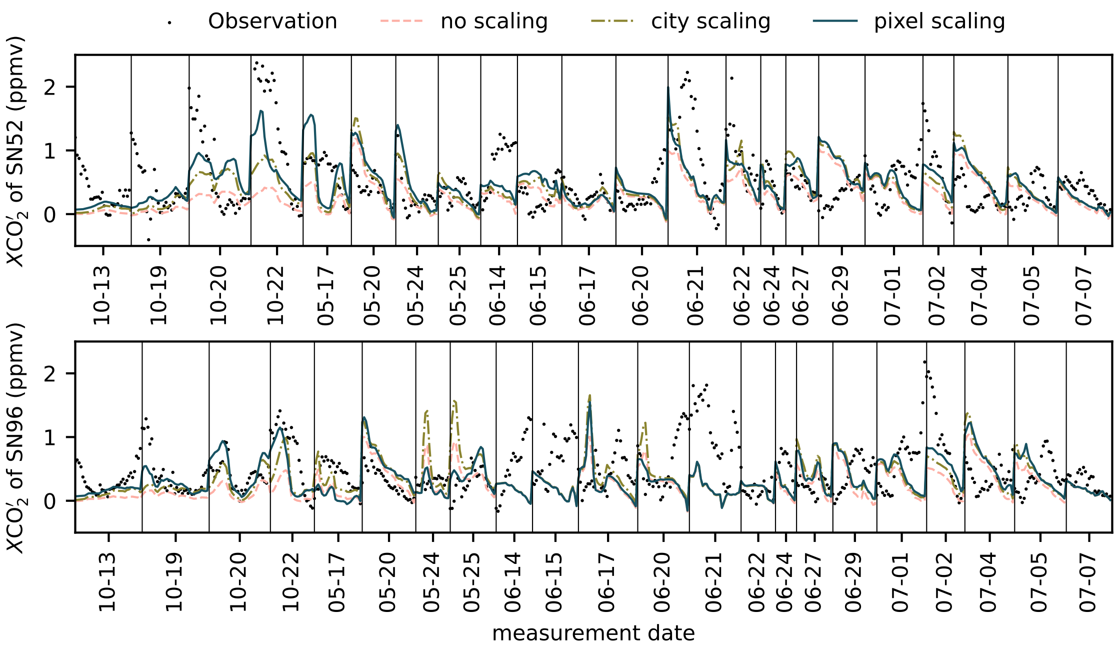


**Figure 5.2.: The agreement between the simulated and observed time series for the whole sample.** The observed time series is shown together with the simulated time-series. The rose dashed line represents the model result without further scaling of the emissions. The green dashdottet and the blue solid line represent the results of the fit using the pixel scaling approach and the city scaling approach. The results for SN52 are shown in the upper row, for SN96 in the lower row.

## 5.2. Improvements by Inventory Scaling

In this section, the improvement of the agreement between the simulation and observation by scaling the anthropogenic inventory is investigated. The previously limited agreement (Section 5.2.1) can be significantly improved by the scaling, as is shown in Section 5.2.2. The consistency using different subsets of the whole sample is tested in Section 5.2.3.

### 5.2.1. Pre-scaling Agreement of the Time Series

When comparing the observations to the corresponding simulation, without applying any scaling to the emission, the agreement is very limited. When looking at both sites simultaneously, the Pearson correlation coefficient is only 0.1. The disagreement is also clearly visible in Figure 5.2 where the time series are compared for the two observation sites. A number of reasons might cause this disagreement. One category of reasons is the representation of transport in the model. As discussed in Section 4.2 the simulated and observed wind is clearly correlated. Still the simulated values scatter around the observations. These mismatches might be a relevant factor, affecting the transport. Also, the parametrization of the turbulent transport itself might be a source of inaccuracy. The second possible origin are the sources and sinks. In addition to the uncertainty in the anthropogenic source, the background contributions that are not considered in the model and an incorrect dataset for the NEE might be significant contributors. In the following, the improvement of the agreement by scaling the anthropogenic emissions will be investigated. The remaining limitations of the model will be discussed in more depth in Section 5.4.

### 5.2.2. Improved Time Series by Optimizing the City Emissions

The possibility of post-scaling of individual emissions pixels in the anthropogenic inventory is explained in Section 4.1.2. To find the optimal scaling configuration, a least-squares approach is used. The two time series corresponding to the two observation sites are optimized simultaneously.

To find a re-weighting factor $x_i$ for each pixel, the following cost function is minimized (implemented in the `scipy` library in the function `optimize.lsq_linear` (Virtanen et al., 2020))

$$c = \frac{1}{2} \cdot || \sum_i x_i \mathbf{a_i} - \mathbf{b} ||^2 \quad , \tag{5.1}$$

where $\mathbf{b}$ is the observation vector and $\mathbf{a_i}$ is the corresponding vector originating from pixel $i$,

Two different scaling approaches were used. In one, only a single weight was determined, scaling all 30 pixels of the city. Here, the fit included one degree of freedom, this configuration will be referenced as *city scaling* in the following. In the other optimization approach, called *pixel scaling*, all pixels were allowed to be scaled with individual weights, leaving 30 degrees of freedom.

The improvement from the city scaling and pixel scaling are shown in Figure 5.2. Both show a significant improvement. The correlation factor increases from 0.1 to 0.29 (0.17) for the pixel (city) scaling.

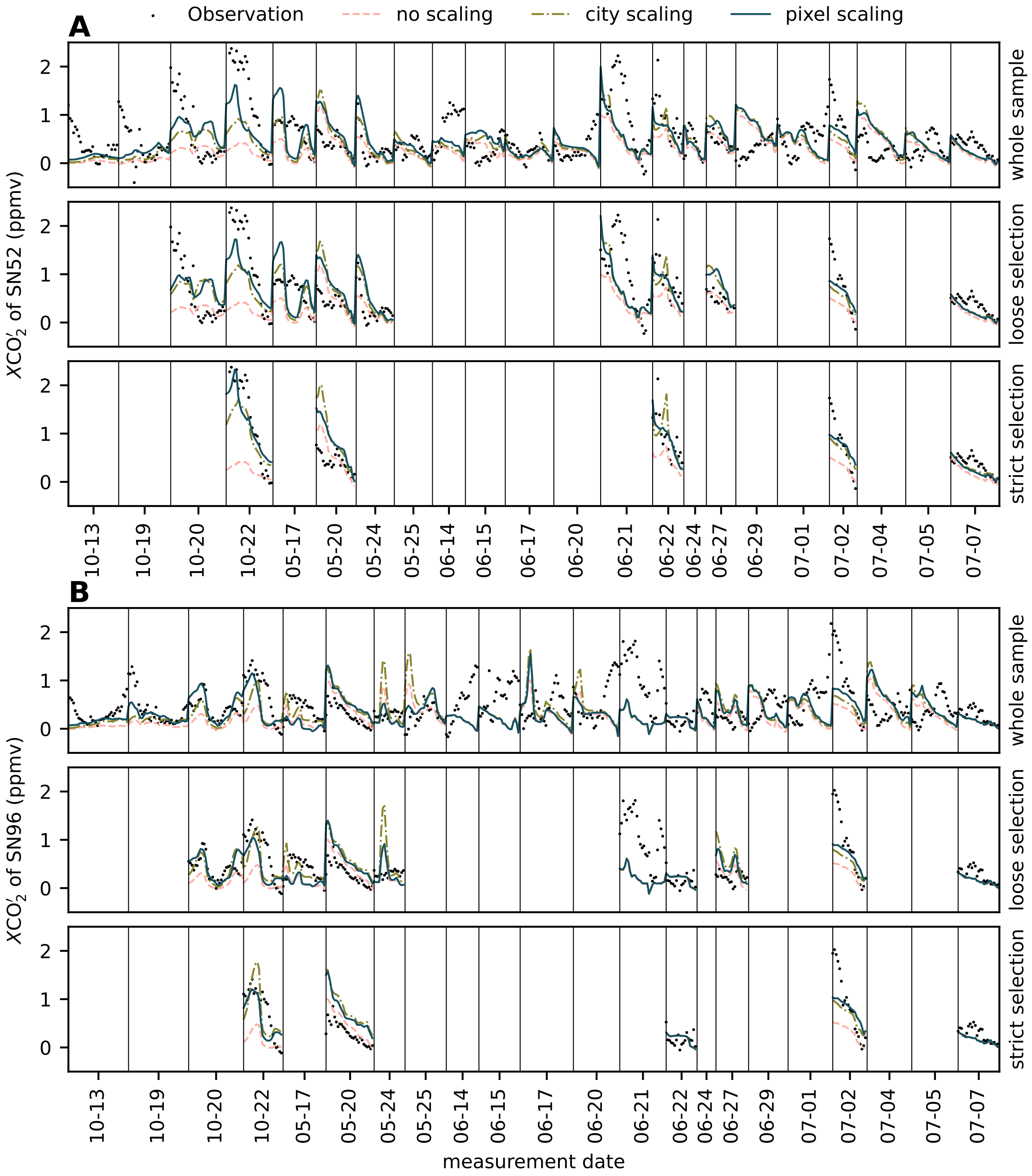


**Figure 5.3.: The agreement between the simulated and observed time series for the subsamples.** The observed time series is shown next to the simulated time-series. The rose dashed line represents the model result without further scaling of the emissions. The green dashdottet and the blue solid line represent the results of the fit using the pixel scaling approach and the city scaling approach. Panel A shows the results for SN52, Panel B for SN96. In the second and third row of each panel, the results for the two selected subsamples are shown. In the first row of each panel, the results for the whole sample from Figure 5.2 are shown for comparison.

**Table 5.1.: The fit results for the three tested configurations.** The total emissions of the city area, the correlation $r$ and the cost function $c$ (see Equation 5.1) are shown for the original ODIAC emission inventory, and two rescaled emission inventories. The results for the whole sample are given in the first part of the Table. The lower part gives the analogue results for two subsamples. Note that the emissions from version ODIAC2020b for May 2019 were used in the model. There, the urban emissions (1.9 Mt/yr) differ slightly from the annual average for 2019 in the version ODIAC2022 (1.8 Mt/yr), that is shown in Figure 1.4.

| | | emissions (Mt/yr) | r | c |
|---|---|---|---|---|
| whole sample | no scaling | 1.9 | 0.10 | 158.9 |
| | city scaling | 2.9 | 0.17 | 144.8 |
| | pixel scaling | 2.9 | 0.29 | 125.4 |
| loose selection | no scaling | 1.9 | 0.36 | 89.9 |
| | city scaling | 3.4 | 0.42 | 68.1 |
| | pixel scaling | 3.0 | 0.53 | 57.3 |
| strict selection | no scaling | 1.9 | 0.34 | 49.1 |
| | city scaling | 4.4 | 0.66 | 24.5 |
| | pixel scaling | 3.7 | 0.77 | 17.1 |

### 5.2.3. Investigation of Subsamples

When looking at individual days, the correlation varies significantly. To test the robustness of the results, I selected two subsamples where the unscaled simulation matched the observations more successfully. The following two selections with a threshold for the correlation of a given day, are discussed in the following.

1. *loose selection* ($r > 0.2$)
2. *strict selection* ($r > 0.6$)

For both subsamples the city and the pixel scaling were applied. As to be expected, the correlation of the unscaled simulated improved in comparison to the whole sample. However, for both subsamples the city and the pixel scaling were able to further improve the agreement. This can be seen in Table 5.1, where a comparison for all configurations is given. It is also visible in the time series. For comparison also the results from Section 5.2 for the whole sample are collected in Figure 5.3.

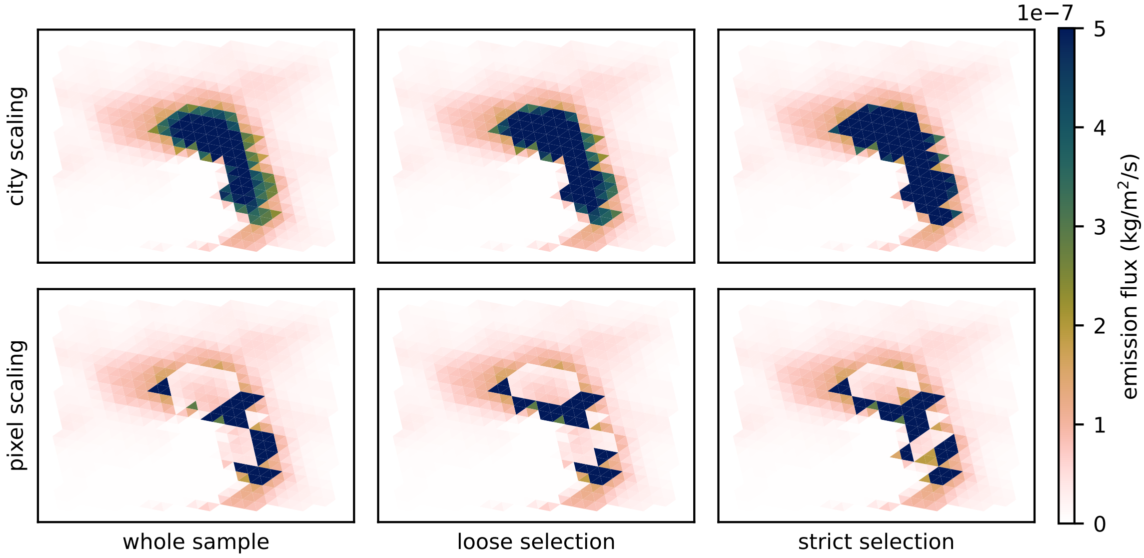


**Figure 5.4.: Spatial distribution of emissions resulting from the fit.** For each sample and each scaling aproach the resulting scaling factors imply an altered emissionmap, which is depicted in this plot. The two rows represent the two scaling approaches, the different samples are shown in the different columns.

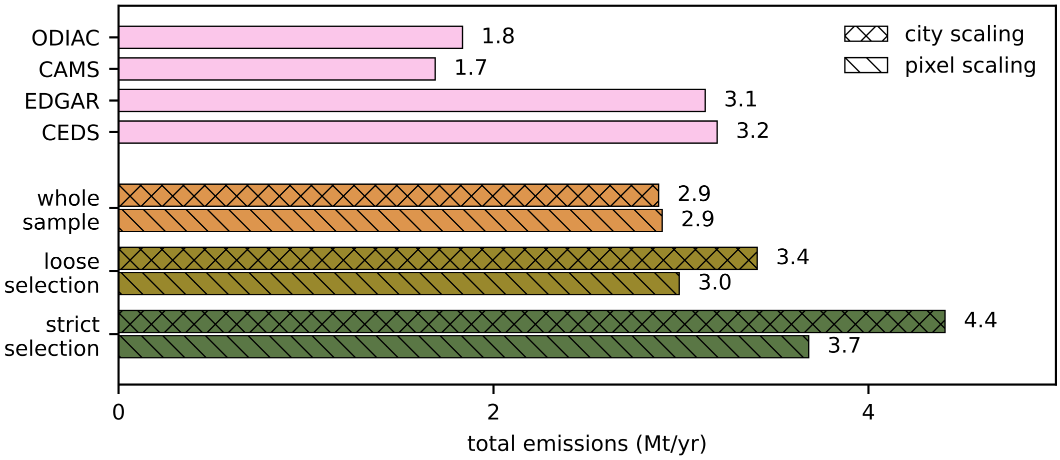


**Figure 5.5.: Yearly emissions for the city area resulting from the fit.** The sum of the city emissions is shown as bar chart for the different samples and scaling approaches. The estimates from the different inventories shown in Figure 1.4 are shown in the top for comparison. The three pairs of bars below represent the different samples. The scaling approaches are indicated by hashes.

## 5.3. Resulting Emission Inventories

The optimizations from the previous section result in a set of scaling factors for the city emissions. These scaling factors significantly improved the simulated agreement to the atmospheric observations, as discussed above. In this section the re-scaled emissions maps resulting from the fit, are discussed.

The emission maps from the different configurations are shown in Figure 5.4. The three upper plots show the city scaling with only one degree of freedom. The city weights are higher than one for all three samples, which can be seen from comparison with Figure 4.2. In the lower three plots, the pixel scaling is used. It is visible that some areas were up-weighted by the optimization while others received lower weights. Interesting is that, when comparing the three plots, some of the pixels receive high weights in all configurations, while others vary. When looking at the hot-spots indicated in Figure 4.2, the consistently high pixels in the center of the city are near to the harbor. The city is also most densely populated in this area. Also, the westernmost pixel consistently receives high weights. It is located close to the refinery – one of the expected emission hot-spots. The small spatial mismatch might be due to a wrong height of the emissions, as all emissions from the ODIAC inventory are emitted into the surface layer. In conclusion, a hot-spot detection by the re-weighting seems plausible.

When looking at the total emissions instead of the emission distribution, the fit results are consistent for all configurations. The total emissions after reweighting are in all cases significantly higher than those reported in the original ODIAC inventory. The reweighted emissions for the city area range from 2.9 Mt/yr for the whole sample to 4.3 Mt/yr for the strict selection with city scaling. The estimates for the total annual emissions of Thessaloniki are illustrated in Figure 5.5. The derived top-down estimates are considerably higher compared to the emissions in ODIAC of 1.8 Mt/yr. This indicates an underestimation of the actual emissions of the city of Thessaloniki in the ODIAC inventory. The fit results are in better agreement to the EDGAR and CEDS inventories, stating an emission estimate of 3.1 Mt/yr and 3.2 Mt/yr for Thessaloniki.

## 5.4. Limitations of the Agreement

As discussed above, without any scaling of the anthropogenic emissions, the agreement between the simulated and observed enhancements is very limited. Although the scaling of the anthropogenic emissions could distinctly improve it, significant discrepancies remain. Possible reasons for the remaining mismatch are discussed in the following.

### 5.4.1. Initialization, Background and Resolution

The systematic limitations of the model as its initialization and its resolution might be an important contributor the discrepancies. Starting with the resolution, the real-world

**Table 5.2.: Correlation in $XH_2O$ for the subsamples.** For each sample, the Pearson correlation factor between observation and simulation is investigated for both instruments individually and for the difference between the two observation sites. This table and the caption are taken from Feld et al. (2024c).

| **Instrument**<br>**Sample** | SN52 | SN96 | SN52-SN96 |
|---|---|---|---|
| whole sample | 0.962 | 0.958 | 0.330 |
| loose selection | 0.986 | 0.987 | 0.209 |
| strict selection | 0.995 | 0.994 | 0.664 |

sub-grid variability can be assumed to not play a major role. This was tested with small-distance observations (see Section 3.1.5). One factor contributing to uncertainty arises from the limited simulation time. Longer transportation times, such as the return of the Thessaloniki plume from a previous day, cannot be simulated in the model setup that was chosen.

### 5.4.2. Meteorological Conditions and Transport

In general, the wind and water vapor were well-matched by the simulations, as was discussed in Section 4.2. The observed absolute $XH_2O$ is very well represented by the simulation. However, the difference between the observation sites shows a considerably lower correlation. This suggests that the transport (indirectly from wrong wind and directly from turbulence parametrization) is a relevant source of uncertainty in the model.

The correlation of $XH_2O$ improves in the subsample with the strict selection to 0.64 (see Table 4.4). It is plausible that at days with a better agreement between observed and simulated $XCO_2$, the transport was better matched by the simulation, which results in a better agreement in $XH_2O$ as well.

A more accurate representation of the transport could result in a better simulation of $XCO_2$ in the future. To assess the accuracy of the transport for future simulations, measuring the wind field more accurately might be beneficial. For example a wind Lidar system could be used to discover inaccuracies in the simulated wind field that result in wrong transport. An improvement in the parametrization of the turbulent transport could enhance the accuracy of the simulations. Better input data, as for example the three-dimensional structure of the city are probably also relevant for the distribution.

### 5.4.3. Sources and Sinks

An assumption of the inventory scaling methodology was, that a major source of uncertainties comes from the anthropogenic emission inventory. The improved agreement after applying an optimized scaling makes this assumption plausible.

In the scaling only the lateral distribution and total strength of emissions was adjusted. Remaining degrees of freedom encompass the temporal change of the emission inventory and the biogenic source and sink.

The annual temporal variability of the emissions might have an influence as the measurement days cover different seasons of the year. The ODIAC inventory provides monthly averages, but the emissions fluxes range from 1.7 Mt/yr in summer to 1.9 Mt/yr in winter in version ODIAC2022. This variation is small compared to the discrepancies between the different inventories. The CEDS inventory is also provided as monthly averages, but the emission flux does not vary at all from month to month. I therefore assume that the largest uncertainty does not originate from the temporal variations. The impact of the diurnal variability can not be assessed here, as none of the used datasets contained a diurnal variability of the emissions.

The representation of the NEE is a large source of inaccuracy. This can be seen from the fact that many days in June and July show large discrepancies and at the same time there are only small differences between the time series without scaling and the scaled version (see Figure 5.3). From this can be seen that the main contribution to the daily variation of $XCO_2'$ comes from the biogenic contribution, which appears to be not well-matched for these days. Plants in the Mediterranean are able to close their stomata in hot and dry conditions to reduce water loss. This leads to a decrease in gas exchange rates and a decrease in the NEE (Lange et al., 1985). This effect is not present in the averaged diurnal cycles from the FLUXCOM X-BASE dataset (compare Figure 4.3C). The generally large influence of the biogenic sinks and sources is also discussed in Section 4.3.2.

In future research this can be improved by coupling a biogenic model to the simulations. Either an observation based solution like the Vegetation Photosynthesis and Respiration Model (VPRM) (Mahadevan et al., 2008) or a land process model as the Jena Scheme for Biosphere-Atmosphere Coupling in Hamburg (JSBACH) (Reick et al., 2021) could be used for this purpose. The coupling of a biosphere model to the simulation was outside the scope of this work.

# 6. Summary

In this work, I studied the $CO_2$ emissions in the urban area of Thessaloniki. The work was motivated by the significant discrepancies between different bottom-up inventories for the urban area of Thessaloniki. Thessaloniki's emissions from the EDGAR (3.1 Mt/yr) and CEDS (3.2 Mt/yr) inventories are significantly higher than the emissions stated by the CAMS (1.7 Mt/yr) and ODIAC (1.8 Mt/yr) inventories.

In the light of these discrepancies, the main objective was the measurement-based estimation of the total anthropogenic $CO_2$ emissions of Thessaloniki. In addition, the spatial distribution of the anthropogenic emissions inside the city area was investigated.

A measurement campaign with two solar Fourier Transform Infrared (FTIR) spectrometers of the type EM27/SUN was implemented for this purpose. Time series of column-averaged dry-air molar fraction of $CO_2$ ($X\mathrm{CO}_2$) were observed with a pair of EM27/SUN spectrometers. The campaign was performed in two parts: A first shorter visit for preparatory purposes took place in October 2021. During this period 6 measurement days were recorded. The second part of the campaign was performed between May 12 and July 12, 2024, with 24 days of measurements. During the second period, there was the possibility to transport one of the instruments to various locations while the other stayed at a fixed position in the central part of the city. A total of 30 days of measurements were recorded. 7 of these were side-by-side measurements for calibration. 22 days were selected for emission estimation, containing a $X\mathrm{CO}_2$ time series with a total of approximately 179 measurement hours. The median difference between the two instruments was only 0.17 ppmv, which is below the estimated systematic calibration uncertainty of 0.22 ppmv. The maximum observed difference was 2.03 ppmv. These differences are small compared to previous studies. This is in line with the expectations, since Thessaloniki is smaller than many cities previously investigated. However, the weaker signal makes it more challenging to estimate the anthropogenic emissions.

Corresponding simulations were conducted with the numerical weather prediction model ICON-ART in a limited area setup with a resolution of approximately 1 km x 1 km. For each day a separate simulation was initialized, using meteorological data from the German Weather Service (DWD). The anthropogenic emissions were taken from the ODIAC inventory, due to its high resolution. The inventory was separated into 30 different source regions over the urban area of Thessaloniki. Each source region was simulated as a separate tracer to enable scaling of the emissions during post-processing. Estimates of the biogenic sources and sinks were constructed from three different datasets to achieve a sufficient spatial and temporal resolution. I initialized $CO_2$ without background concentration, simulating only the enhancements from the emission sources within the domain. To assess the validity of the simulation, I compared the wind data of the simulation to wind observations from three meteorological stations. The simulation was clearly correlated to the observations, even though the general wind situation in the city was quite complex. The Pearson correlation coefficients for longitudinal and latitudinal wind components range from 0.55 to 0.83. Also the predicted water vapor columns, that were co-observed during the campaign agree very well, with correlation coefficients of 0.96 for both spectrometers. The highest correlation was found comparing the observed and simulated pressure, with a correlation coefficient of 0.99.

For comparing the observed $X\mathrm{CO}_2$ time series to the model results, the datasets were first harmonized. This included interpolation of the simulated molar fractions of $CO_2$ to the observation sites, calculation of column averaged molar fractions from the simulation, and temporal averaging of the observed data. To remove the background offset between the modeled and simulated data, the 5th percentile was subtracted from both datasets for each day. Without scaling of the emissions, the agreement between the simulated and observed data is poor: The time series have a correlation coefficient of only 0.1.

The agreement could be enhanced significantly by rescaling the anthropogenic emissions. An optimal scaling was determined through a least-squares approach. Two different scaling approaches were chosen: In the first approach, all pixels inside the city received the same weights, leaving just one degree of freedom. This approach is referenced as city scaling. Alternatively, all source regions (pixels) were allowed to be scaled differently. This pixel scaling approach had 30 degrees of freedom. The optimization resulted in an altered time series. The agreement between the rescaled time series and the observation data improved, which can be seen from the increase of the correlation coefficient to 0.17 for the city scaling and 0.29 for the pixel scaling.

To test the robustness of the fit two subsets were created, selecting days with better prior agreement. I required the correlation coefficient for a specific day to be greater than 0.2 for a loose selection, and greater than 0.6 for a strict selection. As expected, the correlation of the unscaled time series improved. However, in both cases the optimal scaling also improved the correlation visibly over the unscaled version. The correlation factor increased up to 0.77 for the pixel scaling when using the strict selection criterion.

Apparent discrepancies between the simulated and observed time series remains for all scaling configurations. I identify the following factors as relevant for the limited agreement: The short simulation time, limitations in the simulation of transport, missing time dependence of anthropogenic emissions, and large uncertainties in the estimation of the net ecosystem exchange. In particular, the inaccurate representation of biogenic interactions seems to significantly affect the agreement during certain periods of the observation.

The optimal rescaling increased the emissions of the city consistently compared to the original inventory for every configuration. The estimates range from 2.9 to 4.4 Mt/yr, which are significantly higher than the emissions reported by the ODIAC and CAMS inventories, but are better in line with the estimates from EDGAR and CEDS. When comparing the distribution of the emissions after applying the pixel scaling approach, it is apparent that some pixels consistently receive high weights. Some of these up-weighted pixels could be linked to expected emission hot spots, making a hot spot detection by the fit plausible.

In conclusion, this work shows that a dataset with limited signal strength still has the potential for emission estimation when combined with accurate transport simulations. It demonstrates that a simpler approach, such as scaling the inventory with a least-squares method, is also sufficient to extract insights about the emissions from the dataset.

To further advance top-down emission estimation at the urban scale in the future, the following improvements would be worth considering: First, and most importantly, a better representation of the biogenic sources and sinks could significantly improve the agreement. It might even be beneficial to target campaigns on urban atmosphere-biosphere interactions in a place with presumably low uncertainties in the anthropogenic contribution. When selecting targets for validating anthropogenic emission inventories, it might be useful to determine spots with high uncertainties. Comparing different inventories can provide insights for this. A provided inventory uncertainty determined from the underlying emission factors and activity data would be a very valuable foundation for identifying potential targets for future campaign observations. The development of measurement-based emission estimation methods, that are completely independent of bottom-up emission inventories could further enhance the importance of measurement-based inventory validation methods. Finally, there is great potential in automating measurements and improving data accessibility. This could enable scientists to better incorporate different data sources into the investigations. In particular the combination of different complementing datasets such as in-situ and remote-sensing measurements, could provide an essential gain in information, improving ongoing and future top-down emission estimation.

# Bibliography

Ahsan, H., O'Rourke, P., Smith, S., Mott, A., Mcduffie, E., Crippa, M., Klimont, Z., Mcdonald, B., Wang, S., Nicholson, M., Hoesly, R., and Feng, L.: CEDS v_2021_04_21 Gridded emissions data - 0.1° [Dataset], Pacific Northwest National Laboratory 2. doi:10.25584/PNNLDH/1854347, 2022.

Alberti, C., Hase, F., Frey, M., Dubravica, D., Blumenstock, T., Dehn, A., Castracane, P., Surawicz, G., Harig, R., Baier, B. C., Bès, C., Bi, J., Boesch, H., Butz, A., Cai, Z., Chen, J., Crowell, S. M., Deutscher, N. M., Ene, D., Franklin, J. E., García, O., Griffith, D., Grouiez, B., Grutter, M., Hamdouni, A., Houweling, S., Humpage, N., Jacobs, N., Jeong, S., Joly, L., Jones, N. B., Jouglet, D., Kivi, R., Kleinschek, R., Lopez, M., Medeiros, D. J., Morino, I., Mostafavipak, N., Müller, A., Ohyama, H., Palmer, P. I., Pathakoti, M., Pollard, D. F., Raffalski, U., Ramonet, M., Ramsay, R., Sha, M. K., Shiomi, K., Simpson, W., Stremme, W., Sun, Y., Tanimoto, H., Té, Y., Tsidu, G. M., Velazco, V. A., Vogel, F., Watanabe, M., Wei, C., Wunch, D., Yamasoe, M., Zhang, L., and Orphal, J.: Improved calibration procedures for the EM27/SUN spectrometers of the COllaborative Carbon Column Observing Network (COCCON), Atmospheric Measurement Techniques, 15, 2433–2463. doi:10.5194/amt-15-2433-2022, 2022.

Andres, R. J., Boden, T. A., and Higdon, D.: A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission, Tellus B: Chemical and Physical Meteorology, 66, 23 616. doi:10.3402/tellusb.v66.23616, 2014.

C2SM-RCM: emiproc, `https://github.com/C2SM-RCM/emiproc` (Last Access: 2024-02-01), 2023.

CMIP: CMIP Phase 7 (CMIP7), `https://wcrp-cmip.org/cmip7/` (Last Access: 2024-08-01), 2023.

Colella, P. and Woodward, P. R.: The Piecewise Parabolic Method (PPM) for gas-dynamical simulations, Journal of Computational Physics, 54, 174–201. doi:10.1016/0021-9991(84)90143-8, 1984.

Copernicus Climate Change Service: Global Climate Highlights 2023 - The 2023 Annual Climate Summary, `https://climate.copernicus.eu/global-climate-highlights-2023` (Last Access: 2024-04-29), 2024.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., and Janssens-Maenhout, G.: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2, Earth System Science Data, 10, 1987–2013. doi:10.5194/essd-10-1987-2018, 2018.

Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hutfilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, Environmental Research Letters, 16, 074 033. doi:10.1088/1748-9326/ac00e2, 2021.

Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Becker, W., Ferrario, F. M., Grassi, G., Rossi, S., Melo, J. B. D., Oom, D. J. F., Branco, A., San-Miguel, J., and Vignati, E.: Emissions Database for Global Atmospheric Research, version v8.0_FT_2022 [Dataset], European Commission, Joint Research Centre (JRC), `http://data.europa.eu/89h/809d7b72-55ef-4e52-8bd4-7d33f2f9916b` (Last Access: 2024-08-15), 2022.

Demtröder, W.: Atoms, molecules and photons: an introduction to atomic-, molecular- and quantum physics, Graduate texts in physics, Springer, Berlin; Heidelberg, 3 edn. ISBN 978-3-662-55523-1, 2018.

Dubravica, D.: KIT-IMK-ASF - About IMK-ASF - Research Groups - Ground-Based Remote Sensing - COCCON - Data Repository, `https://www.imk-asf.kit.edu/english/3884.php` (Last Access: 2024-06-10), 2024.

Efremenko, D. and Kokhanovsky, A. A.: Foundations of atmospheric remote sensing, Springer, Cham. ISBN 978-3-030-66745-0, 2021.

Enting, I. G.: Inverse problems in atmospheric constituent transport, Cambridge University Press, Cambridge; New York. ISBN 978-0-521-81210-8, 2002.

Feld, L., Dubravica, D., Balis, D., Hase, F., Mermigkas, M., Panou, T., and Schmid, P.: COCCON Version 1 dataset from atmospheric observatory of Thessaloniki available at the EVDC Data Handling Facilities covering start date Oct 4th 2021 to end date Jul 12th 2022 [Dataset], EVDC. doi:10.48477/COCCON.PF23.THESSALONIKI-CAMPAIGN.R01, 2024a.

Feld, L., Herkommer, B., Vestner, J., Dubravica, D., Alberti, C., and Hase, F.: PROFFASTpylot: Running PROFFAST with Python, Journal of Open Source Software, 9, 6481. doi:10.21105/joss.06481, 2024b.

Feld, L., Schmid, P., Mermigkas, M., Balis, D., Gross, J., Dubravica, D., Alberti, C., Herkommer, B., Versick, S., Ruhnke, R., Hase, F., and Braesicke, P.: XCO2 observations compared to km-scale ICON-ART simulations indicate an underestimation of Thessaloniki's emissions in the ODIAC inventory, In Preparation, 2024c.

Fleming, L.: Libya floods: Entire neighbourhoods dragged into the sea, `https://www.bbc.com/news/world-africa-66785466` (Last Access: 2024-07-09), 2023.

FOEN: Latest greenhouse gas inventory of Switzerland - Inventory data and documentation submitted to the UNFCCC in 2024 (covering the years 1990–2022), including related review reports., `https://www.bafu.admin.ch/bafu/en/home/themen/thema-klima/klima--daten--indikatoren-und-karten/daten--treibhausgasemissionen-der-schweiz/`

`swiss-climate-reporting-under-the-unfccc/swiss-greenhouse-gas-inventories/swiss-greenhouse-gas-inventory-1990-2013.html` (Last Access: 2024-07-12), 2024.

Frey, M., Sha, M. K., Hase, F., Kiel, M., Blumenstock, T., Harig, R., Surawicz, G., Deutscher, N. M., Shiomi, K., Franklin, J. E., Bösch, H., Chen, J., Grutter, M., Ohyama, H., Sun, Y., Butz, A., Mengistu Tsidu, G., Ene, D., Wunch, D., Cao, Z., Garcia, O., Ramonet, M., Vogel, F., and Orphal, J.: Building the COllaborative Carbon Column Observing Network (COCCON): long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer, Atmospheric Measurement Techniques, 12, 1513–1530. doi:10.5194/amt-12-1513-2019, 2019.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tans, P. P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2023, Earth System Science Data, 15, 5301–5369. doi:10.5194/essd-15-5301-2023, 2023.

Gilfillan, D. and Marland, G.: CDIAC-FF: global and national CO2 emissions from fossil fuel combustion and cement manufacture: 1751–2017, Earth System Science Data, 13, 1667–1680. doi:10.5194/essd-13-1667-2021, 2021.

Gisi, M., Hase, F., Dohe, S., and Blumenstock, T.: Camtracker: a new camera controlled high precision solar tracker system for FTIR-spectrometers, Atmospheric Measurement Techniques, 4, 47–54. doi:10.5194/amt-4-47-2011, 2011.

Goody, R. M. and Yung, Y. L.: Atmospheric radiation: theoretical basis, Oxford Univ. Press, New York, 2. edn. ISBN 978-0-19-510291-8, 1995.

Google Maps: Satellite Image of Thessaloniki, `https://www.google.com/maps/` (Last Access: 2022-06-16), 2022.

Gordon, I. E., Rothman, L. S., Hargreaves, R. J., Hashemi, R., Karlovets, E. V., Skinner, F. M., Conway, E. K., Hill, C., Kochanov, R. V., Tan, Y., Wcisło, P., Finenko, A. A., Nelson, K., Bernath,

P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Coustenis, A., Drouin, B. J., Flaud, J. M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Mlawer, E. J., Nikitin, A. V., Perevalov, V. I., Rotger, M., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G., Adkins, E. M., Baker, A., Barbe, A., Canè, E., Császár, A. G., Dudaryonok, A., Egorov, O., Fleisher, A. J., Fleurbaey, H., Foltynowicz, A., Furtenbacher, T., Harrison, J. J., Hartmann, J. M., Horneman, V. M., Huang, X., Karman, T., Karns, J., Kassi, S., Kleiner, I., Kofman, V., Kwabia–Tchana, F., Lavrentieva, N. N., Lee, T. J., Long, D. A., Lukashevskaya, A. A., Lyulin, O. M., Makhnev, V. Y., Matt, W., Massie, S. T., Melosso, M., Mikhailenko, S. N., Mondelain, D., Müller, H. S. P., Naumenko, O. V., Perrin, A., Polyansky, O. L., Raddaoui, E., Raston, P. L., Reed, Z. D., Rey, M., Richard, C., Tóbiás, R., Sadiek, I., Schwenke, D. W., Starikova, E., Sung, K., Tamassia, F., Tashkun, S. A., Vander Auwera, J., Vasilenko, I. A., Vigasin, A. A., Villanueva, G. L., Vispoel, B., Wagner, G., Yachmenev, A., and Yurchenko, S. N.: The HITRAN2020 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 277, 107 949. doi:10.1016/j.jqsrt.2021.107949, 2022.

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version) [Dataset], Copernicus Atmosphere Monitoring Service. doi:10.24380/D0BN-KX16, 2019.

Granier, C., Liousse, C., McDonald, B., and Middleton, P.: Anthropogenic Emissions Inventories of Air Pollutants, in: Handbook of Air Quality and Climate Change, edited by Akimoto, H. and Tanimoto, H., pp. 3–52, Springer Nature, Singapore. doi:10.1007/978-981-15-2760-9_5, 2023.

Griffiths, P. R. and De Haseth, J. A.: Fourier transform infrared spectrometry, no. 83 in Chemical analysis, Wiley, New York. ISBN 978-0-471-09902-4, 1986.

Haken, H. and Wolf, H. C.: Molekülphysik und Quantenchemie: Einführung in die experimentellen und theoretischen Grundlagen, Springer, Berlin, 5. edn. ISBN 978-3-540-30315-2, 2006.

Hanna, S. R., Briggs, G. A., and Hosker, R. P.: Handbook on atmospheric diffusion, Technical Information Center, U.S. Dept. of Energy, Oak Ridge. ISBN 978-0-87079-127-7, 1982.

Harris, C. R., Millman, K. J., Walt, S. J. v. d., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. v., Brett, M., Haldane, A., Río, J. F. d., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362. doi:10.1038/s41586-020-2649-2, 2020.

Harris, L. M., Lauritzen, P. H., and Mittal, R.: A flux-form version of the conservative semi-Lagrangian multi-tracer transport scheme (CSLAM) on the cubed sphere grid, Journal of Computational Physics, 230, 1215–1237. doi:10.1016/j.jcp.2010.11.001, 2011.

Hase, F., Frey, M., Blumenstock, T., Groß, J., Kiel, M., Kohlhepp, R., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Application of portable FTIR spectrometers for detecting greenhouse gas emissions of the major city Berlin, Atmospheric Measurement Techniques, 8, 3059–3068. doi:10.5194/amt-8-3059-2015, 2015.

Hase, F., Frey, M., Kiel, M., Blumenstock, T., Harig, R., Keens, A., and Orphal, J.: Addition of a channel for XCO observations to a portable FTIR spectrometer for greenhouse gas measurements, Atmospheric Measurement Techniques, 9, 2303–2313. doi:10.5194/amt-9-2303-2016, 2016.

Henne, S., Brunner, D., Oney, B., Leuenberger, M., Eugster, W., Bamberger, I., Meinhardt, F., Steinbacher, M., and Emmenegger, L.: Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling, Atmospheric Chemistry and Physics, 16, 3683–3710. doi:10.5194/acp-16-3683-2016, 2016.

Herkommer, B.: Improving the consistency of greenhouse gas measurements from ground-based remote sensing instruments using a portable FTIR spectrometer, Ph.D. thesis. doi:10.5445/IR/1000168723, 2024.

Herkommer, B., Alberti, C., Castracane, P., Chen, J., Dehn, A., Dietrich, F., Deutscher, N. M., Frey, M. M., Groß, J., Gillespie, L., Hase, F., Morino, I., Pak, N. M., Walker, B., and Wunch, D.: Using a portable FTIR spectrometer to evaluate the consistency of Total Carbon Column Observing Network (TCCON) measurements on a global scale: the Collaborative Carbon Column Observing Network (COCCON) travel standard, Atmospheric Measurement Techniques, 17, 3467–3494. doi:10.5194/amt-17-3467-2024, 2024.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-i., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geoscientific Model Development, 11, 369–408. doi:10.5194/gmd-11-369-2018, 2018.

IEA: Data and statistics, `https://www.iea.org/data-and-statistics` (Last Access: 2024-08-01), 2024.

IPCC: Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 1 edn. doi:10.1017/9781009325844. ISBN 978-1-00-932584-4, 2023a.

IPCC: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 1 edn. doi:10.1017/9781009157896. ISBN 978-1-00-915789-6, 2023b.

IPCC: Technical Summary, in: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 35–144, Cambridge University Press, Cambridge. doi:10.1017/9781009157896.002, 2023c.

ITMS Germany: Integriertes Treibhausgas-Monitoringsystem für Deutschland (ITMS), `https://www.itms-germany.de/en` (Last Access: 2024-07-15), 2024.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, Earth System Science Data, 11, 959–1002. doi:10.5194/essd-11-959-2019, 2019.

Keeling, R. F. and Keeling, C. D.: Atmospheric Monthly In Situ CO2 Data - Mauna Loa Observatory, Hawaii. In Scripps CO2 Program Data [Dataset], UC San Diego Library Digital Collections. doi:10.6075/J08W3BHW, 2017.

Kimball, J. S., Jones, L. A., Endsley, A., Kundig, T., and Reichle, R.: SMAP L4 Global Daily 9 km EASE-Grid, Carbon Net Ecosystem Exchange, Version 7. doi:10.5067/3K9F0S1Q5J2U, 2022.

KIT IMK-ASF: PROFFASTpylot v1.3 documentation, `https://www.imk-asf.kit.edu/english/4261.php` (Last Access: 2024-06-21), 2024.

Lange, O. L., Tenhunen, J. D., and Beyschlag, W.: Effects of humidity during diurnal courses on the CO2- and light-saturated rate of net CO2 uptake in the sclerophyllous leaves of Arbutus unedo, Oecologia, 67, 301–304. doi:10.1007/BF00384305, 1985.

Loeb, N. G., Johnson, G. C., Thorsen, T. J., Lyman, J. M., Rose, F. G., and Kato, S.: Satellite and Ocean Data Reveal Marked Increase in Earth's Heating Rate, Geophysical Research Letters, 48, e2021GL093 047. doi:10.1029/2021GL093047, 2021.

Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., and Stocker, T. F.: High-resolution carbon dioxide concentration record 650,000–800,000 years before present, Nature, 453, 379–382. doi:10.1038/nature06949, 2008.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochemical Cycles, 22. doi:10.1029/2006GB002735, 2008.

Makarova, M. V., Alberti, C., Ionov, D. V., Hase, F., Foka, S. C., Blumenstock, T., Warneke, T., Virolainen, Y. A., Kostsov, V. S., Frey, M., Poberovskii, A. V., Timofeyev, Y. M., Paramonova, N. N., Volkova, K. A., Zaitsev, N. A., Biryukov, E. Y., Osipov, S. I., Makarov, B. K., Polyakov, A. V., Ivakhov, V. M., Imhasin, H. K., and Mikhailov, E. F.: Emission Monitoring Mobile

Experiment (EMME): an overview and first results of the St. Petersburg megacity campaign 2019, Atmospheric Measurement Techniques, 14, 1047–1073. doi:10.5194/amt-14-1047-2021, 2021.

Makowski, M.: EM27 Retrieval Pipeline, `https://em27-retrieval-pipeline.netlify.app/` (Last Access: 2024-06-21), 2024.

Maksyutov, S., Brunner, D., Turner, A. J., Zavala-Araiza, D., Janardanan, R., Bun, R., Oda, T., and Patra, P. K.: Chapter 14 - Applications of top-down methods to anthropogenic GHG emission estimation, in: Balancing Greenhouse Gas Budgets, edited by Poulter, B., Canadell, J. G., Hayes, D. J., and Thompson, R. L., pp. 455–481, Elsevier. doi:https://doi.org/10.1016/B978-0-12-814952-2.00006-X, 2022.

Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, Journal of Geophysical Research: Atmospheres, 116. doi:10.1029/2010JD014763, 2011.

Mayer-Kuckuk, T.: Atomphysik: Eine Einführung, Vieweg+Teubner Verlag, Wiesbaden, 5. edn. ISBN 978-3-663-01606-9, 1997.

Mermigkas, M., Topaloglou, C., Balis, D., Koukouli, M. E., Hase, F., Dubravica, D., Borsdorff, T., and Lorente, A.: FTIR Measurements of Greenhouse Gases over Thessaloniki, Greece in the Framework of COCCON and Comparison with S5P/TROPOMI Observations, Remote Sensing, 13, 3395. doi:10.3390/rs13173395, 2021.

Miura, H.: An Upwind-Biased Conservative Advection Scheme for Spherical Hexagonal–Pentagonal Grids, Monthly Weather Review, 135, 4038–4044. doi:10.1175/2007MWR2101.1, 2007.

Monroe, R.: The Keeling Curve, `https://keelingcurve.ucsd.edu` (Last Access: 2024-04-29), 2024.

Möller, D.: Chemistry of the climate system – Fundamentals and processes, De Gruyter, Berlin; Boston, 3 edn. ISBN 978-3-11-056126-5, 2019.

NASA: Earth's Energy Budget Poster, `https://science.nasa.gov/wavelength-acf/earths-energy-budget-poster/` (Last Access: 2024-09-20), 2016.

NOAA Climate: Climate Change: Atmospheric Carbon Dioxide, `http://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide` (Last Access: 2024-06-12), 2024.

Oda, T. and Maksyutov, S.: A very high-resolution (1 km×1 km) global fossil fuel $CO_2$ emission inventory derived using a point source database and satellite observations of nighttime lights, Atmospheric Chemistry and Physics, 11, 543–556. doi:10.5194/acp-11-543-2011, 2011.

Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic $CO_2$, version 2016 (ODIAC2016): a global monthly fossil fuel $CO_2$ gridded emissions data product for tracer transport simulations and surface flux inversions, Earth System Science Data, 10, 87–107. doi:10.5194/essd-10-87-2018, 2018.

Ohyama, H., Frey, M. M., Morino, I., Shiomi, K., Nishihashi, M., Miyauchi, T., Yamada, H., Saito, M., Wakasa, M., Blumenstock, T., and Hase, F.: Anthropogenic $CO_2$ emission estimates in the Tokyo metropolitan area from ground-based $CO_2$ column observations, Atmospheric Chemistry and Physics, 23, 15 097–15 119. doi:10.5194/acp-23-15097-2023, 2023.

Orlanski, I.: A Rational Subdivision of Scales for Atmospheric Processes, Bulletin of the American Meteorological Society, 56, 527–530, `http://www.jstor.org/stable/26216020` (Last Access: 2024-06-13), 1975.

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, Proceedings of the National Academy of Sciences, 104, 18 925–18 930. doi:10.1073/pnas.0708986104, 2007.

Prill, F.: DWD ICON Tools Documentation, `dwd_icon_tools/doc/icontools_doc.pdf` (Last Access: 2023-06-23), 2020.

Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjeld, S.: JSBACH 3 - The land component of the MPI Earth System Model: documentation of version 3.2, p. 4990986. doi:10.17617/2.3279802, 2021.

Reinert, D., Prill, F., Frank, H., Denhard, M., Baldauf, M., Schraff, C., Gebhardt, C., Marsigli, C., Förstner, J., Zängl, G., and Schlemmer, L.: DWD Database Reference for the Global and Regional ICON and ICON-EPS Forecasting System, `https://www.dwd.de/SharedDocs/downloads/DE/modelldokumentationen/nwv/icon_d2/icon_d2_dbbeschr_aktuell.html` (Last Access: 2024-06-13), 2021.

Reinert, D., Rieger, D., and Prill, F.: ICON Tutorial 2024: Working with the ICON Model. doi:10.5676/DWD_PUB/NWV/ICON_TUTORIAL2024, 2024.

Ritchie, H. and Roser, M.: CO2 emissions, Our World in Data, `https://ourworldindata.org/co2-emissions` (Last Access: 2024-07-09), 2024.

Roedel, W. and Wagner, T.: Physik unserer Umwelt: die Atmosphäre, Lehrbuch, Springer Spektrum, Berlin; Heidelberg, 5 edn. ISBN 978-3-662-54258-3, 2017.

Rubino, M., Etheridge, D., Thornton, D., Allison, C., Francey, R., Langenfelds, R., Steele, P., Trudinger, C., Spencer, D., Curran, M., Van Ommen, T., and Smith, A.: Law Dome Ice Core 2000-Year CO2, CH4, N2O and d13C-CO2 [Dataset], CSIRO. doi:10.25919/5BFE29FF807FB, 2019.

Running, S. and Zhao, M.: MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500 m SIN Grid V061 [Dataset], NASA EOSDIS Land Processes Distributed Active Archive Center. doi:10.5067/MODIS/MOD17A3HGF.061, 2021.

Salby, M. L.: Physics of the Atmosphere and Climate, Cambridge University Press, New York, `http://ebookcentral.proquest.com/lib/karlsruhetech/detail.action?docID=807160` (Last Access: 2022-12-07). ISBN 978-1-139-15740-7, 2012.

Schmid, P. L.: Quantification of Greenhouse Gas Emissions in Thessaloniki, Greece, Master's thesis, Karlsruher Institut für Technologie (KIT). doi:10.5445/IR/1000159168, 2023.

Schröter, J., Rieger, D., Stassen, C., Vogel, H., Weimer, M., Werchner, S., Förstner, J., Prill, F., Reinert, D., Zängl, G., Giorgetta, M., Ruhnke, R., Vogel, B., and Braesicke, P.: ICON-ART 2.1: a flexible tracer framework and its application for composition studies in numerical weather forecasting and climate simulations, Geoscientific Model Development, 11, 4043–4068. doi:10.5194/gmd-11-4043-2018, 2018.

Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., and Janssens-Maenhout, G.: Uncertainties in the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases, Atmospheric Chemistry and Physics, 21, 5655–5683. doi:10.5194/acp-21-5655-2021, 2021.

Sottile, J., Ward, Z., and Sutton, T.: Canadian wildfire smoke reaches Europe as Canada reports its worst fire season, `https://www.cnn.com/2023/06/26/americas/canada-wildfire-season-worst-2023/index.html` (Last Access: 2024-07-09), 2023.

TCCON Wiki: TCCON Requirements, `https://tccon-wiki.caltech.edu/Main/TCCONRequirements` (Last Access: 2024-06-25), 2021.

The White House: National Strategy to Advance an Integrated U.S. Greenhouse Gas Measurement, Monitoring, and Information System, `https://www.whitehouse.gov/ostp/news-updates/2023/11/29/national-strategy-to-advance-an-integrated-u-s-greenhouse-gas-measurement-monitoring-and-infor` (Last Access: 2024-07-15), 2023.

Tiseo, I.: Global CO2 emissions by year 1940-2023, `https://www.statista.com/statistics/276629/global-co2-emissions/` (Last Access: 2024-07-09), 2024.

Tu, Q.: Observation of atmospheric greenhouse gas abundances on regional scales in boreal areas using portable FTIR Spectrometers, Ph.D. thesis. doi:10.5445/IR/1000095901, 2019.

Tu, Q., Hase, F., Schneider, M., García, O., Blumenstock, T., Borsdorff, T., Frey, M., Khosrawi, F., Lorente, A., Alberti, C., Bustos, J. J., Butz, A., Carreño, V., Cuevas, E., Curcoll, R., Diekmann, C. J., Dubravica, D., Ertl, B., Estruch, C., León-Luis, S. F., Marrero, C., Morgui, J.-A., Ramos, R., Scharun, C., Schneider, C., Sepúlveda, E., Toledano, C., and Torres, C.: Quantification of

$CH_4$ emissions from waste disposal sites near the city of Madrid using ground- and space-based observations of COCCON, TROPOMI and IASI, Atmospheric Chemistry and Physics, 22, 295–317. doi:10.5194/acp-22-295-2022, 2022.

UNFCCC: The Paris Agreement - Publication, `https://unfccc.int/documents/184656` (Last Access: 2024-07-09), 2016.

UNFCCC: Why the Global Stocktake is Important for Climate Action this Decade, `https://unfccc.int/topics/global-stocktake/about-the-global-stocktake/why-the-global-stocktake-is-important-for-climate-action-this-decade#What-does-the-global-stocktake-tell-us` (Last Access: 2024-07-10), 2024.

United Nations Environment Programme and International Methane Emissions Observatory: An Eye on Methane: The road to radical transparency - International Methane Emissions Observatory 2023 Report, United Nations Environment Programme, `https://wedocs.unep.org/xmlui/handle/20.500.11822/44129` (Last Access: 2024-07-15). ISBN 978-92-807-4102-5, 2023.

United Nations Treaty Collection: Status of Treaties - 7. d Paris Agreement, `https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=_en` (Last Access: 2024-07-15), 2024.

Viatte, C., Lauvaux, T., Hedelius, J. K., Parker, H., Chen, J., Jones, T., Franklin, J. E., Deng, A. J., Gaudet, B., Verhulst, K., Duren, R., Wunch, D., Roehl, C., Dubey, M. K., Wofsy, S., and Wennberg, P. O.: Methane emissions from dairies in the Los Angeles Basin, Atmospheric Chemistry and Physics, 17, 7509–7528. doi:10.5194/acp-17-7509-2017, 2017.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272. doi:10.1038/s41592-019-0686-2, 2020.

Vogel, F. R., Frey, M., Staufer, J., Hase, F., Broquet, G., Xueref-Remy, I., Chevallier, F., Ciais, P., Sha, M. K., Chelin, P., Jeseck, P., Janssen, C., Té, Y., Groß, J., Blumenstock, T., Tu, Q., and Orphal, J.: $XCO_2$ in an emission hot-spot region: the COCCON Paris campaign 2015, Atmospheric Chemistry and Physics, 19, 3271–3285. doi:10.5194/acp-19-3271-2019, 2019.

Weber, U.: FLUXCOM-X-BASE [Dataset], ICOS ERIC – Carbon Portal. doi:10.18160/5NZG-JMJE, 2023.

Weimer, M., Schröter, J., Eckstein, J., Deetz, K., Neumaier, M., Fischbeck, G., Hu, L., Millet, D. B., Rieger, D., Vogel, H., Vogel, B., Reddmann, T., Kirner, O., Ruhnke, R., and Braesicke, P.: An emission module for ICON-ART 2.0: implementation and simulations of acetone, Geoscientific Model Development, 10, 2471–2494. doi:10.5194/gmd-10-2471-2017, 2017.

Williams, D. R.: NASA Planetary Fact Sheet, `https://nssdc.gsfc.nasa.gov/planetary/factsheet/` (Last Access: 2024-05-03), 2024.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. doi:10.1098/rsta.2010.0240, 2011.

Zachariah, M., Kotroni, V., Kostas, L., Barnes, C., Kimutai, J., Kew, S., Pinto, I., Yang, W., Vahlberg, M., Singh, R., Thalheimer, D., Marghidan Pereira, C., Otto, F., Philip, S., El Hajj, R., El Khoury, C., Walsh, S., Spyratou, D., Tezapsidou, E., and Salmela-Eckstein, S.: Interplay of climate change-exacerbated rainfall, exposure and vulnerability led to widespread impacts in the Mediterranean region, Tech. rep., Imperial College London. doi:10.25561/106501, 2023.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, Quarterly Journal of the Royal Meteorological Society, 141, 563–579. doi:10.1002/qj.2378, 2015.

# Appendix

# A. Measurement Notes

During the campaign, the meausurement days, locations, start and stop times and comments were noted. These notes are listed in the following.

| Date | SN52 | SN96 | Start | Stop | Comments |
|---|---|---|---|---|---|
| 2021-10-04 | Physics | Physics | 08:35:00 | 18:07:00 | |
| 2021-10-05 | Physics | Physics | 09:07:00 | 17:30:00 | |
| 2021-10-06 | Physics | Physics | 09:40:00 | 10:30:00 | Cloudy, nearly no data |
| 2021-10-13 | Physics | Thermi | 09:00:00 | 18:00:00 | |
| 2021-10-19 | Physics | Thermi | 09:05:00 | 18:00:00 | |
| 2021-10-20 | Physics | Thermi | 09:50:00 | 16:40:00 | Background not from Sea |
| 2021-10-22 | Physics | Thermi | 09:30:00 | 15:30:00 | |
| 2022-05-12 | Thermi | Thermi | 09:15:00 | 14:50:00 | SN52: At 10:45 fixed radius problem |
| 2022-05-13 | Thermi | Thermi | 09:45:00 | 15:30:00 | |
| 2022-05-17 | Meteorology | Galini | 10:00:00 | 15:20:00 | |
| 2022-05-20 | Meteorology | Thermi | 09:45:00 | 15:30:00 | SN52: Until 10:40 Camtracker problems |
| 2022-05-24 | Meteorology | Galini | 09:45:00 | 13:55:00 | SN96: Until 11:05 loose cable, 1/2h data loss |
| 2022-05-25 | Meteorology | Galini | 09:40:00 | 14:30:00 | |
| 2022-06-09 | Meteorology | Meteorology | 08:30:00 | 12:10:00 | |
| 2022-06-14 | Meteorology | Galini | 09:30:00 | 13:30:00 | SN96: 10:05 sun non centered, corrected morph |
| 2022-06-15 | Meteorology | Efkarpia | 10:50:00 | 15:45:00 | |
| 2022-06-17 | Meteorology | Efkarpia | 08:45:00 | 15:20:00 | SN96: Until 11:20 ca. 1/2h sun not centered |
| 2022-06-20 | Meteorology | Efkarpia | 09:40:00 | 15:15:00 | |
| 2022-06-21 | Meteorology | Diavata | 09:55:00 | 15:30:00 | |
| 2022-06-22 | Meteorology | Diavata | 09:00:00 | 12:50:00 | |
| 2022-06-24 | Meteorology | Thermi | 13:00:00 | 15:45:00 | |

| | | | | | |
|---|---|---|---|---|---|
| 2022-06-27 | Meteorology | Thermi | 09:40:00 | 13:25:00 | SN52: CH2 A x8 Signal and x4 Background possibly unreliable Data from 9.6 for Channel 2! |
| 2022-06-29 | Meteorology | Thermi | 11:45:00 | 16:25:00 | SN96: 16:00 sun not centered (not enough Power) |
| 2022-07-01 | Meteorology | Physics | 11:00:00 | 16:30:00 | |
| 2022-07-02 | Meteorology | Physics | 11:00:00 | 14:30:00 | SN52: clicking noise increases, end of spectra with artefacts |
| 2022-07-04 | Meteorology | Seich-Sou | 10:45:00 | 15:00:00 | |
| 2022-07-05 | Meteorology | Seich-Sou | 09:30:00 | 15:00:00 | SN52: Blackout at morning |
| 2022-07-06 | Meteorology | Meteorology | 09:45:00 | 16:15:00 | |
| 2022-07-07 | Meteorology | Diavata | 09:50:00 | 15:00:00 | SN96: 12:00 short break due to a tractor to trim the gras |
| 2022-07-11 | Meteorology | Thermi | 09:50:00 | 14:30:00 | |
| 2022-07-12 | Meteorology | Meteorology | 09:45:00 | | SN96: 12:00 Sun not Centered |

# B. List of Figures

# C. List of Tables

# Acknowledgments

Ich bin sehr dankbar, dass ich die Möglichkeit hatte, mit diesem spannenden Projekt zu promovieren. Das wurde von der Helmholtz European Partnership for Technological Advancement (HEPTA) ermöglicht. Mein Aufenthalt in Thessaloniki wurde außerdem von der Graduate School for Climate and Environment (GRACE) unterstützt. Die Simulationen für diese Arbeit wurden auf dem Supercomputer HoreKa durchgeführt, der vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg und vom Bundesministerium für Bildung und Forschung gefördert wird.

Mein besonderer Dank gilt Prof. Peter Braesicke, Priv.-Doz. Frank Hase und Dr. Roland Ruhnke für die wunderbare Betreuung. Danke für die unzählbaren, hilfreichen Ideen, Vorschläge, Fragen und Diskussionen, die meine Arbeit jedes Mal weitergebracht haben. Und genauso sehr für den Mut und das Selbstvertrauen, das ihr mir vermittelt habt. Aber vor allem bedanke ich mich dafür, dass ihr mir die Freiheit gegeben habt, meine eigenen Ideen auszuprobieren.

Danke Prof. Ulrich Husemann, dass Sie die Aufgabe des Mentors übernommen haben. Es ist gut im Notfall zu wissen, an wen man sich wenden könnte, auch wenn es in meinem Fall keinen Notfall gab.

I especially appreciate the opportunity for the research stay in Thessaloniki for the field campaign, which has been a wonderful experience. In particular the outstanding hospitality I experienced there. I thank Prof. Dimitrios Balis, Dr. Marios Mermigkas, Thomas Panou and Pablo Schmid for the good work together. I also thank the Greek HEPTAs for the warm welcome. Ευχαριστώ!

Pablo, vielen Dank für deine Motivation und dein Engagement während deiner Masterarbeit. Es hat mir viel Spaß gemacht und auch für mich war die Zeit in Thessaloniki „eine Gute“.

Ich möchte noch allen Mitgliedern der BOD und MOD Gruppen sowie allen anderen Mitarbeiter:innen des IMK-ASF für die vielfältige Unterstützung danken. Ihr habt mir immer das Gefühl gegeben willkommen zu sein. Danke für das gemeinsame Klettern, Musizieren, Konzerte Besuchen, Feierabendbier, Whisky Tasting, Quizzen und die vielen

gute Gespräche über Wissenschaft und nicht-wissenschaftliche Themen. Außerdem danke ich Philipp Dietz, Dr. Stefanie Falk, Valentin Hanft, Dr. Benedikt Herkommer, Nga Ying Lo, Yiling Ma, und Dr. Stefan Versick für die hilfreichen Anmerkungen und Verbesserungsvorschläge zur Thesis.

Ganz besonders möchte ich noch dir, Benedikt, danken. Der PROFFASTpylot war ein super Projekt; auch mir hat es viel Spaß gemacht. Das gemeinschaftlich produktive Arbeiten mit dir hat meinen Arbeitsalltag sehr bereichert. Ich hoffe deine zukünftigen Kollegen wissen ihr Glück zu schätzen!

Ich danke meiner Familie, besonders meinen Eltern Usch und Frank, meinen anderen Eltern Jan und Ela und meiner Schwester Zoey. Danke, dass ihr mich immer unterstützt habt! Ich finde schön, dass wir eine so gute Gemeinschaft sind. Ein großes Dankeschön geht auch an den Physikerchor, der immer eine zuverlässige Quelle von Aufheiterung war. Ganz besonders danke ich auch dir, Anika, für dein offenes Ohr für einfach alles.

Jan, ich danke Dir für alle Zeit die wir bisher teilen durften. Und für genauso viel Geduld und Verständnis.