Inversion of short-lived pollutants in the global atmosphere using remote sensing data

Dissertation zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich für Physik und Elektrotechnik
der Universität Bremen

Vorgelegt von

Johann Rasmus Nüß

- 24.01.2024 -

Gutachter:

Prof. Dr. Mihalis Vrekoussis

Prof. Dr. Maarten C. Krol

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Claus Lämmerzahl

Prof. Dr. Maria Kanakidou

Dr. Nikos Daskalakis

Fabian Günther Piwowarczyk

Datum des Kolloquiums:

15.03.2024

Abstract

In the atmosphere, carbon monoxide is a trace gas with a relatively short lifetime in the order of a few months. On a global scale, it affects the climate, because most carbon monoxide is eventually oxidized to the greenhouse gas carbon dioxide. That reaction is also the largest sink of hydroxyl radicals and, therefore, prolongs greenhouse gas lifetimes. Close to the sources of carbon monoxide, concentrations can be high and adversely impact local air quality, because it is a precursor for tropospheric ozone.

For these reasons, knowledge about the global distribution of carbon monoxide and its sources is important. Inverse modeling is a powerful top-down technique to constrain trace gas emissions, or refine existing bottom-up source estimates, based on observational data. However, inverse modeling of atmospheric chemistry is not trivial and requires the use of sophisticated systems, which combine elaborate models with a plethora of input data, including remote sensing observations and boundary conditions for the model, e.g. meteorology or prior emission estimates. In this work, such an inverse modeling system will be improved by testing, updating, and revising most of its components. Most prominently, observations from a new satellite instrument, the TROPOspheric Monitoring Instrument (TROPOMI), are introduced into the system. In recent years, the data quality and resolution of satellite instruments have been steadily improving. These improvements inevitably also lead to an increase in the amount of data to be handled. For inverse modeling systems, large observational datasets can become problematic due to computational constraints.

In this work, methods for handling those datasets are developed. To investigate the capabilities and limitations of the new observational dataset, multiple inversion experiments are conducted. These experiments target carbon monoxide emissions from three categories, biomass burning, fossil fuel, and secondary production, for the second half of the year 2018 on a global scale. The results suggest that the emissions, especially in the southern hemisphere, are well constrained by the TROPOMI observations. However, the inversion experiments also reveal biases in the optimized emissions, especially in the northern tropics. These biases are linked to an imbalanced prior budget, i.e. to the boundary conditions of the model before the observations are considered. The budget and the biases are improved in multiple steps, most notably by revising the assumed hydroxyl radical distribution and the meteorology.

Contents

CONTENTS

CONTENTS

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

Chapter 1

Introduction

This chapter introduces the background, motivation, and basic concepts underlying this thesis. Section 1.1 gives a brief overview of the chemical composition of the atmosphere and explains why even gases that occur at low concentrations matter. In Section 1.2, the properties of one such gas, carbon monoxide, and its interactions in the atmosphere are presented. Section 1.3 then introduces the challenges, capabilities, and limitations of atmospheric chemistry and transport models, which can simulate those interactions. In Section 1.4, the concepts of bottom-up and top-down approaches are explained, which make use of models to estimate emissions. Finally, Section 1.5 provides the aim and structure of this thesis.

1.1 Atmospheric composition

The atmosphere consists primarily of nitrogen (N_2, 78.08 % dry air mole fraction), oxygen (O_2, 20.95 %), and argon (Ar, 0.93 %) [BORZENKOVA & TURCHINOVICH, 2009]. Another significant, but highly variable constituent is water vapor (H_2O), which makes up only about 0.25 % of the mass of the atmosphere; however, locally it can vary between almost 0 and 5 % by mole fraction [WALLACE & HOBBS, 2006]. Further, there is a wide variety of gases and aerosols that occur in smaller quantities, so called tracers. These tracers, despite existing in small quantities, can have a significant impact on the world surrounding us.

Some of them, for example H_2O and carbon dioxide (CO_2), are capable of trapping radiation within the atmosphere. This so called greenhouse effect brings the mean near-surface temperature of the atmosphere from well below the freezing point of water to temperatures that are capable of sustaining life as we know it [BERGER & TRICOT, 1992]. However, over the past century, anthropogenic emissions of greenhouse gases have amplified this effect, leading to a rise in global mean temperature and climate change [FORSTER et al., 2021]. Anthropogenic greenhouse gases are usually long-lived tracers and most prominently include CO_2 (417 ppm atmospheric mean mixing ratio

in 2022 [FRIEDLINGSTEIN et al., 2022]), methane (CH_4, 1857 ppb in 2018 [SAUNOIS et al., 2020]), and nitrous oxide (N_2O, 329 ppb in 2016 [PRINN et al., 2018]).

Moreover, trace gases can also impact our surroundings locally and on shorter time scales by affecting air quality. Here, the relevant tracers and aerosols are usually shorter lived and, therefore, have more variable concentrations. The largest group of these tracers are the (Non-Methane) Volatile Organic Compounds ((NM)VOCs), which are several thousand chemical species of varying complexity. The majority of NMVOCs are hydrocarbons (C_xH_y), of which isoprene (C_5H_8) makes up around 50 % [WALLACE & HOBBS, 2006]. Another important group of tracers, especially in urban environments [LEE et al., 1997], are nitrogen oxides ($NO_x = NO + NO_2$), which, depending on the atmospheric conditions, can lead to the formation of nitric acid (HNO_3) or tropospheric ozone (O_3) [BRASSEUR & JACOB, 2017]. The latter is detrimental for humans and plants alike, even at low mixing ratios (> 120 ppb for an exposure of 1 h; or less for a longer exposure [MCKEE, 1993]). These processes will be described in more detail in Section 1.2.

This work focuses on another trace gas with a relatively short lifetime, carbon monoxide (CO), which affects both climate and air quality through its interactions with other tracers.

1.2   Carbon monoxide and its chemistry

Roughly half of atmospheric CO comes from the oxidation of NMVOCs and CH_4, i.e. from secondary CO production. The rest is emitted directly and comes mostly from incomplete combustion of fossil fuels and biomass (e.g. wildfires or domestic wood burning), but also, in smaller quantities, from direct natural emissions from plants (biogenic CO) and the oceans [ZHENG et al., 2019]. Generally, CO is toxic [RYTER et al., 2018] at high mixing ratios (> 9 ppm for an exposure of 8 h; much shorter at higher mixing ratios, according to the World Health Organization [WHO, 1999]). However, in the atmosphere, such high mixing ratios are rarely observed and mostly only occur in direct proximity to combustion sources. Usually, the mean atmospheric CO background mixing ratios are in the order of 54–115 ppb [PRINN et al., 2018], which is low enough that the toxicity of CO and the resulting direct health effects are overshadowed by the indirect effects of CO on air quality. Most notably, CO is an O_3 precursor (see Reaction (1.5)) in the presence of NO_x and solar radiation [HOLLOWAY et al., 2000]. Figure 1.1 outlines the tropospheric CO chemistry described in more detail in the following. The reaction sequence that eventually leads to O_3 formation is initiated via the reaction of CO with hydroxyl radicals (OH) to create CO_2 [BRASSEUR & JACOB, 2017]:

$$CO + OH \xrightarrow{O_2} CO_2 + HO_2 \tag{1.1}$$
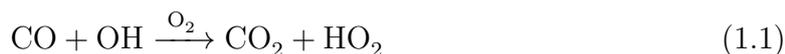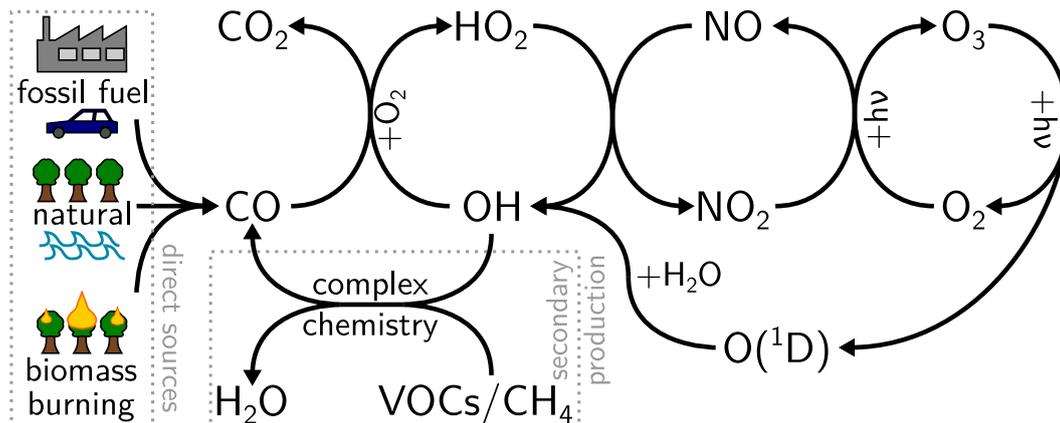
Figure 1.1: Carbon monoxide related chemistry in the troposphere. Note the formation of CO_2 and O_3 and the cycling of NO_x and OH.

which is the main sink [$> 90\,\%$ of the total CO loss; LOGAN et al., 1981; STEIN et al., 2014] for CO in the atmosphere. The O_2 above the reaction sign denotes that while the O_2 is part of the reaction, it does not limit its rate, since it is in abundance in the surrounding air. Through its reaction with OH, CO reduces the oxidative capacity of the atmosphere and both directly (by formation of CO_2) and indirectly (through the reduced OH abundance and thus longer CH_4 lifetime) increases greenhouse gas loads [RAUB & MCMULLEN, 1991; DANIEL & SOLOMON, 1998; HEILMAN et al., 2014]. The remaining CO loss is due to dry deposition to soil [STEIN et al., 2014]. Overall, the atmospheric lifetime of CO is in the order of 2 months [RAUB & MCMULLEN, 1991], with a strong dependence on photochemical conditions, with shorter lifetimes (< 1 month) in the tropics and much longer lifetimes (> 12 months) during polar winters [HOLLOWAY et al., 2000]. This large variation in CO lifetime is linked to the dependence of OH formation on photochemistry in combination with the very short lifetime of OH (in the order of seconds, LELIEVELD et al. [2016]), which prevents it from being transported over long distances. Primary formation of OH mainly happens during daytime when excited oxygen radicals ($O(^1D)$) created through photolysis of ozone

$$O_3 + h\nu(\lambda < 330\,\text{nm}) \longrightarrow O(^1D) + O_2 \qquad (1.2)$$
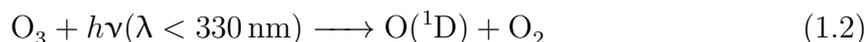
react with water vapor [BRASSEUR & JACOB, 2017]:

$$O(^1D) + H_2O \longrightarrow 2\,OH \qquad (1.3)$$

However, most of the OH in the troposphere comes from secondary sources through various OH recycling pathways, which depend on the local pollution levels (NO_x, (Biogenic) Volatile Organic Compounds ((B)VOCs), etc.), as detailed in LELIEVELD et al. [2016]. While these pathways still rely on photochemistry, some of the precursors are

longer lived than OH, which allows for transport to occur and, therefore, higher OH in regions with low primary production.

Aside from being the main CO sink, OH also plays a large role in the secondary CO production source by oxidizing VOCs and CH_4. In the atmosphere, most VOCs get oxidized, in some cases over many intermediate species, to eventually end up as CO and CO_2. The main oxidant in this process is OH, but O_3, nitrate (NO_3, during nighttime), halogen atoms, and, in some cases, photolysis also play a role. In the following, secondary CO production will refer to the sum over any CO that was chemically produced in the atmosphere, regardless of the initial reactant (VOCs or CH_4) or the source of that reactant (biogenic, anthropogenic, etc.).

Depending on the ambient atmospheric conditions, the oxidation of either CO or VOCs can lead to the production of tropospheric O_3, which further compromises air quality. Continuing with the HO_2 from Reaction (1.1), in the presence of NO_x, the following sequence starts:

$$HO_2 + NO \longrightarrow OH + NO_2 \tag{1.4}$$

$$NO_2 + h\nu \xrightarrow{O_2} NO + O_3 \tag{1.5}$$

which has O_3 as a product [Brasseur & Jacob, 2017]. Note that effectively no NO_x is used up by this process, because any NO used up by Reaction (1.4) is regenerated in Reaction (1.5) and vice versa for NO_2. Similarly, the sequence as a whole preserves OH, which is consumed by Reaction (1.1), but regenerated in Reaction (1.4). An equivalent sequence exists for VOCs that were (partly) oxidized by OH (RO_2). In that sequence, Reaction (1.4) is replaced by [Brasseur & Jacob, 2017]:

$$RO_2 + NO \longrightarrow RO + NO_2 \tag{1.6}$$
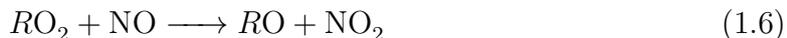
Either sequence depends on the available supply of NO_x on the one hand and VOCs or CO on the other hand. The vast majority of NO_x in the atmosphere is produced during high-temperature combustion processes (mostly fossil fuel, but also biomass burning) and from lightning and soil microbes [Lee et al., 1997]. As such, most of the troposphere, especially in remote regions, is usually NO_x-limited, which leaves sufficient OH radicals to react with CO or VOCs. In this case, the O_3 production scales linearly with NO_x, but is mostly independent of CO or VOC concentrations. However, when the NO_x concentration is high, the atmosphere becomes VOC-limited and OH not only reacts with CO and VOCs, but also with NO_2 to produce nitric acid. In that case, O_3 production scales linearly with CO and VOCs concentrations, but inversely with NO_x.

Overall, knowledge of sources, sinks, and distributions of CO and other tracers is crucial to assess their impact on our environment.

1.3   Atmospheric chemistry modeling

Atmospheric tracer distributions at various spatial scales can be estimated using chemical transport models (CTMs). In CTMs, the complex chemical and physical systems of the atmosphere and their interactions are simplified into a mathematical description. As the name suggests, CTMs model the chemical interactions of tracers and their transport through the atmosphere, which requires solving coupled systems of continuity equations. As the chemistry may include photochemical reactions (e.g. Reactions (1.2) and (1.5)), radiative processes also need to be represented in the model. Another example of interactions between chemistry and physics is the evolution of aerosols and their impact on cloud formation. Clouds, in turn, affect radiative processes and chemistry. On a more fundamental level, CTMs require the sources and sinks of their tracers as boundary conditions. Part of the sources and sinks may be chemical transformation, e.g. CO is formed via the oxidation of VOCs and CH_4 and lost via reaction with OH (Reaction (1.1)). This part may be handled explicitly by the chemistry scheme of the model. However, sources from direct emission require a representation of processes that are usually not explicitly treated in CTMs, for example, the combustion of vegetation in wildfires. Instead, CTMs handle these processes implicitly through emission inventories, which contain information on the emitted species directly. Additionally, non-chemical sink processes need to considered, i.e. wet deposition, when tracers are scavenged by clouds and precipitation, and dry deposition, when tracers are taken up by vegetation and other surfaces, especially the oceans. Any process that is not explicitly included in the CTM, but interacts with a process that is, needs to be parameterized in some form.

CTMs can be useful for a number of reasons. Since they represent (a subset of) our current understanding of atmospheric processes, they may be used to test and improve our knowledge by verifying the model results against observations. Once verified, the CTM may be employed to make predictions by creating a forecast or fill observational gaps by predicting tracer distributions in areas where no observations are available. Additionally, CTMs may help with the interpretation of observational data by extracting information from complex systems with many parameters and processes.

However, since they contain so many components, CTMs also contain many potential error sources. Firstly, computational limitations dictate that the spatial and temporal resolution of the CTM cannot be arbitrarily fine. By limiting the resolution, the representativeness of the CTM for the real world is also limited. The representativeness can affect many processes. For example, the limited temporal resolution may not be able to properly resolve very fast chemical processes, while the limited spatial resolution may cause the emissions from point sources to be dispersed too quickly. Further, there may be biases in the boundary conditions for some tracers, e.g. in the

assumed initial conditions or sources and sinks, which then may impact other tracers as well through their interactions. Additionally, CTMs need to limit the complexity of their chemical mechanism, again, due to computational limitations. This limitation usually happens either by considering hydrocarbons up to a certain complexity or by employing a chemical scheme that only represents a tracer or group of tracers of interest accurately. Any reaction that is not represented explicitly, could potentially introduce biases. Even for those tracers that are considered explicitly, the reaction rates may not be perfectly accurate. Another source of inaccuracies is the meteorology, which via temperature, wind fields, humidity, and various other variable affects most parts of the model in some form. More generally, any simplification applied in the mathematical description of the model processes can introduce errors. This especially concerns processes that are not explicitly included, but only parameterized. A prominent example of a simplification is the use of pre-calculated meteorological fields in so called offline models, as opposed to online models which simulate their own meteorology. In offline models, many interactions between, on the one hand, the chemistry and physics, e.g. aerosols, and, on the other hand, the meteorology will be neglected. Lastly, computational precision presents an upper limit for the achievable precision.

Overall, the model error must be kept sufficiently small for the model to be useful. As pointed out before, whether this is the case may be verified by observations. However, this process may also be turned around, by combing models and observations to constrain the boundary conditions of the model, e.g. to estimate the emissions of tracers.

1.4   Top-down and bottom-up emission estimates

As outlined before, sources of tracers in the atmosphere can broadly be divided into two types, direct emissions and secondary production, i.e. chemical formation within the atmosphere. Sources of each type may be further divided by source category, e.g. direct emissions from fossil fuel and direct emissions from biomass burning.

This process of dividing emission rates up by source category is not trivial, especially when estimating local emission rates at a global scale. Current remote sensing techniques can inform on global tracer distributions at relatively high spatial and moderate temporal resolutions. However, inferring the underlying emissions by source category requires additional information. This information can be incorporated by employing either bottom-up or top-down approaches, which estimated source rates indirectly.

Either approach requires a model, i.e. a mathematical description of the respective system, to extract the emission information from the observations. For bottom-up estimates, the model describes the process that caused the emissions, and observations of this process are used. For example, wildfire CO emissions may be extrapolated

based on observational knowledge about the burnt vegetation and the intensity of the fire [e.g. Wiedinmyer et al., 2023]. Conversely, for top-down estimates, the model describes the effect of the emissions. This effect is then measured and traced back to its source. Following again the example of wildfire CO emissions, the effect is an elevated CO mixing ratio in the atmosphere, which can be observed and then traced back and attributed to its source using atmospheric modeling [e.g. Krol et al., 2013]. A different and more direct kind of top-down approach is to prescribe atmospheric concentrations directly from observations, rather than extracting the emission rates. Such an approach can be meaningful for species where the global tracer distributions are much better constrained from observations than the emission rates, which is the case for some longer-lived species [Brasseur & Jacob, 2017].

Both bottom-up and top-down approaches are affected by various sources of errors. Bottom-up estimates usually require direct observations of the source event and some additional assumptions about the source itself, for example, fuel characterization (ecosystem type, fuel loadings, and fuel consumption rates) and emission factors in the case of biomass burning. Top-down estimates have more loose observational requirements but require a set of potentially more elaborate assumptions for the atmospheric modeling, for example, about chemistry and atmospheric transport. Overall, there is little overlap between the error sources and, therefore, one approach may be used to reduce the uncertainties of the other.

Inverse modeling, as described in detail in Section 2.1, is such a top-down approach which combines a CTM with observations of the atmosphere to retrieve tracer emissions.

1.5   Objective and outline

The aim of this work is to use the high-quality and relatively high resolution observations from the state-of-the-art TROPOMI satellite instrument to investigate the global atmospheric sources of carbon monoxide, which requires the creation of novel methods for handling the observation in an existing inverse modeling framework, TM5-4dvar. Further, various aspects of the model and its boundary conditions were refined and updated throughout this work to improve the source estimates. The thesis is structured as follows:

Chapter 2   consists of an in-depth introduction into inverse modeling, followed by the description of a concrete implementation of that technique, the TM5-4dvar model, and an introduction into the usage of satellite observations in inverse modeling.

Chapter 3   describes how the TM5-4dvar model is used to conduct a series of inverse modeling experiments, in which the global carbon monoxide emissions are optimized based on TROPOMI observations. This application requires the introduction of new methods for handling datasets with a large number of observations and a high spatial resolution relative to the model. This chapter showcases multiple shortcomings of the current setup, which are tackled in the following chapters.

Chapter 4   investigates inconsistencies between, on the one hand, the prior fields for secondary CO production from VOCs and CH_4 and, on the other hand, the prescribed OH climatology used to simulate chemical CO loss, by introducing alternatives for either into TM5-4dvar. Additionally, the biomass burning prior emissions are updated and explicit handling of direct natural CO emission is introduced into the model setup.

Chapter 5   elaborates the handling of the TROPOMI observations by suggesting a refined method for treating the error correlations within the high-resolution observations. Additionally, the satellite product is updated.

Chapter 6   discusses the benefits of the updated ERA5 meteorology and showcases its capacity in reducing biases in the results of inverse modeling experiments.

Chapter 7   combines the findings of the previous chapters and makes recommendations for future inversion experiments with TROPOMI observations in TM5-4dvar.

Chapter 2

Methods

This chapter covers the theoretical background and mathematical basis on which the rest of the thesis is built. In Section 2.1, inverse modeling is introduced and the 4DVAR approach is derived from first principles. Section 2.2 describes the TM5-4dvar inverse modeling framework. Section 2.3 first presents the general strengths and weaknesses of satellite observations in the context of inverse modeling and then introduces the basic interfaces between models and satellite observations.

2.1 Inverse modeling

Inverse modeling describes the concept of utilizing observations of a system in conjunction with a mathematical description, i.e. model, of the system to constrain the values of a subset of the variables that are driving the system, in a way that the difference between the observations and the model is minimal. This section introduces the concept in more detail and provides a derivation of the mathematical basis of one specific implementation of this concept, the 4DVAR approach, and is based on the more extensive description by BRASSEUR & JACOB [2017].

2.1.1 Formal statement of the inverse problem

Consider a model operator $\mathbf{F}$ that links the set of variables we are interested in, the state $\vec{x}$, to a set of observations $\vec{y}$. This relationship can be expressed as

$$\vec{y} = \mathbf{F}(\vec{x}, \vec{p}) + \vec{\varepsilon}_{\mathrm{O}} \tag{2.1}$$

where $\vec{p}$ is the set of additional parameters required by the model but not constrained by the observations and $\vec{\varepsilon}_{\mathrm{O}}$ is the observational error, which combines the uncertainties of the model, the observations, and the parameters.

For illustrative purposes, in the following, we take the state $\vec{x}$ to be emissions of

9

a chemical tracer, the model $\mathbf{F}$ to be a chemical transport model (CTM), and the observations $\vec{y}$ to be mixing ratio measurements from satellites and surface stations. Note that the model operator $\mathbf{F}$ describes our full understanding of the relationship between $\vec{x}$ and $\vec{y}$ and as such includes the mapping from the model to the observations. Hence, any temporal and spatial sampling of the CTM at the time and place of the observations will also be part of $\mathbf{F}$. However, depending on the problem at hand, $\vec{x}$, $\mathbf{F}$, and $\vec{y}$ can stand for various other things, even when only considering atmospheric systems. For example, a satellite retrieval can be formulated as an inverse problem with the sought after tracer column as the state, the radiance spectra from the satellite instrument as the observations, and a radiative transfer model as the model connecting the two.

Regardless of the specific quantities represented by $\vec{x}$, $\mathbf{F}$, and $\vec{y}$, the corresponding mathematical basis remains the same and so does the general problem: The model operator $\mathbf{F}$ allows estimation of the observed quantity $\vec{y}$, given the driving variables $\vec{x}$. In an inversion of $\mathbf{F}$ we are now interested in obtaining the optimal $\hat{\vec{x}}$ given the observation y, so that the difference between $\mathbf{F}$ and $\vec{y}$ is minimized. Because the observational error $\vec{\varepsilon}_O$ is generally non-zero, finding the optimal $\hat{\vec{x}}$ is not trivial. Additionally, proper handling of those errors is important. Assuming no errors, a state vector $\vec{x}$ with n variables would be fully constrained by an equal number $m = n$ of measurements in $\vec{y}$. However, we know that errors for atmospheric measurements are usually large and many observations ($m \gg n$) are required to properly constrain atmospheric states. Furthermore, to prevent overfitting, additional constraints on the optimal state may need to be considered. For example, biomass burning emissions from the oceans should be considered unrealistic and, therefore, should not be part of an optimal state. Mathematically, all these conditions may be tackled by formulating the inverse problem using Bayes' theorem.

2.1.2 Bayes' theorem

Let $P(\vec{x})$ be the probability density function (PDF) of the vector $\vec{x}$, so that integrating $P(\vec{x})$ over any range of possible values for $\vec{x}$ results in the probability of $\vec{x}$ being in that range. Similarly, let $P(\vec{y})$ denote the PDF of the vector $\vec{y}$. Additionally, we define $P(\vec{x}|\vec{y})$ and $P(\vec{y}|\vec{x})$ as the conditional PDFs describing the probability to find a certain realization of $\vec{x}$ given a fixed value for $\vec{y}$, and vice versa, the probability to find $\vec{y}$ given $\vec{x}$. Bayes' theorem connects these four PDFs:

$$P(\vec{x}|\vec{y}) = \frac{P(\vec{y}|\vec{x})P(\vec{x})}{P(\vec{y})} \tag{2.2}$$

Bayes' theorem can be used to formally state the inverse problem by identifying the PDFs with the quantities introduced in the previous section. First, we consider $\vec{x}$ and $\vec{y}$ again as the state and observation vectors, respectively. Then, $P(\vec{x}|\vec{y})$ provides the probability that a certain state $\vec{x}$ describes the system, given a fixed set of observations $\vec{y}$. $P(\vec{x}|\vec{y})$ is also called the posterior PDF, because the optimal $\vec{\hat{x}}$ that maximizes $P(\vec{x}|\vec{y})$ is the solution to the inverse problem, which can formally be expressed as:

$$\nabla_{\vec{x}} P(\vec{x}|\vec{y}) = 0 \tag{2.3}$$

with $\nabla_{\vec{x}}$ as the gradient operator with respect to $\vec{x}$. The counterpart of the posterior PDF $P(\vec{x}|\vec{y})$, is the prior PDF $P(\vec{x})$, which defines the initial distribution of $\vec{x}$ assumed before considering the measurements. In practice, $P(\vec{x})$ is obtained by an a priori guess $\vec{x}_A$ with an uncertainty $\vec{\varepsilon}_A$. $P(\vec{x}_A)$ constrains the solution to realistic values by making large deviation from the prior very unlikely. Next, $P(\vec{y}|\vec{x})$ provides the likelihood of a set of observations $\vec{y}$ given our current best guess for $\vec{x}$, distributed by the model and measurement errors $\vec{\varepsilon}_O$. Now, by combining Equations (2.2) and (2.3), we obtain

$$\nabla_{\vec{x}} \left[\frac{1}{P(\vec{y})} P(\vec{y}|\vec{x}) P(\vec{x}) \right] = 0 \tag{2.4}$$

where $P(\vec{y})$ defines the probability to make a certain observation $\vec{y}$. However, since $P(\vec{y})$ is independent of $\vec{x}$, it will remain constant while finding the optimal $\vec{\hat{x}}$ and can be ignored. This simplifies Equation (2.4) and leads to the Bayesian optimal estimate solution to the inverse problem:

$$\nabla_{\vec{x}} [P(\vec{y}|\vec{x}) P(\vec{x})] = 0 \tag{2.5}$$

2.1.3   Derivation of the cost function

In this section, the PDFs $P(\vec{y}|\vec{x})$ and $P(\vec{x})$ will be stated explicitly, in order to derive a mathematical expression, the cost function, that allows application of Equation (2.5) to solve concrete inverse problems. First, let $E[\,]$ be the expected value operator, which provides the mean values of the bracketed quantity after infinitely many realizations. By applying the expected value operator to the observational ($\vec{\varepsilon}_O$) and a priori ($\vec{\varepsilon}_A$) errors, the general error statistics of $\vec{y}$ and $\vec{x}_A$, respectively, may be represented in the form of their corresponding error covariance matrices (ECMs) $\mathbf{S}_O$ and $\mathbf{S}_A$:

$$\mathbf{S}_O = E\left[\vec{\varepsilon}_O \vec{\varepsilon}_O^T \right] \tag{2.6}$$

$$\mathbf{S}_A = E\left[\vec{\varepsilon}_A \vec{\varepsilon}_A^T \right] \tag{2.7}$$

2 METHODS

The ECMs have the variances of the respective vectors as diagonal elements and the covariances between different entries of the vectors as off-diagonal elements. In real world applications, neither $\mathbf{S}_A$ nor $\mathbf{S}_O$ are generally known with high accuracy; still, at the very least rough estimates are required. For the inversions presented in this thesis, the diagonal terms of $\mathbf{S}_A$ were assumed to be constant, while its off-diagonal terms, which represent spatial correlations in emissions, were set based on exponentially decaying correlation lengths, leading to a band matrix-like structure of $\mathbf{S}_A$. More details can be found in Section 3.2.3.1. For $\mathbf{S}_O$, the diagonal terms included contributions from both the instrument and model error. The off-diagonal terms, however, were assumed to be zero to simplify calculations. This assumption does not always hold, as discussed in more detail in Chapter 5.

Assuming normal distributed prior and observational errors, Gaussian PDFs for vectors may be used to describe $P(\vec{x})$ and $P(\vec{y}|\vec{x})$:

$$P(\vec{x}) = \frac{1}{c_A} \exp\left[-\frac{1}{2}(\vec{x} - \vec{x}_A)^{\mathrm{T}}\mathbf{S}_A^{-1}(\vec{x} - \vec{x}_A)\right] \tag{2.8}$$

$$P(\vec{y}|\vec{x}) = \frac{1}{c_O} \exp\left[-\frac{1}{2}(\vec{y} - \mathbf{F}(\vec{x}))^{\mathrm{T}}\mathbf{S}_O^{-1}(\vec{y} - \mathbf{F}(\vec{x}))\right] \tag{2.9}$$

where c_A and c_O are constants of the structure $(2\pi)^{\frac{n}{2}}|\mathbf{S}|^{\frac{1}{2}}$, with n as the dimension and $|\mathbf{S}|$ as the determinant of the matrices $\mathbf{S}_A$ and $\mathbf{S}_O$, respectively. However, as for the $\frac{1}{P(\vec{y})}$ in Equation (2.4), they do not depend on $\vec{x}$ and, therefore, do not matter for the solution. Considering this, we can use Bayes' theorem (Equation (2.2)) to obtain:

$$P(\vec{x}|\vec{y}) \propto \exp\left[-\frac{1}{2}(\vec{x} - \vec{x}_A)^{\mathrm{T}}\mathbf{S}_A^{-1}(\vec{x} - \vec{x}_A) - \frac{1}{2}(\vec{y} - \mathbf{F}(\vec{x}))^{\mathrm{T}}\mathbf{S}_O^{-1}(\vec{y} - \mathbf{F}(\vec{x}))\right] \tag{2.10}$$

As established in Section 2.1.2, the optimal state $\hat{\vec{x}}$ maximizes $P(\vec{x}|\vec{y})$. Maximizing the exponential function in Equation (2.10) is equivalent to maximizing its exponent. Additionally dividing by a constant factor of -2 further simplifies the expression and turns the maximization problem into a minimization problem. Overall, the inverse problem can now be expressed through the cost function

$$J(\vec{x}) = (\vec{x} - \vec{x}_A)^{\mathrm{T}}\mathbf{S}_A^{-1}(\vec{x} - \vec{x}_A) + (\vec{y} - \mathbf{F}(\vec{x}))^{\mathrm{T}}\mathbf{S}_O^{-1}(\vec{y} - \mathbf{F}(\vec{x})) \tag{2.11}$$

which is to be minimized to find the maximum of $P(\vec{x}|\vec{y})$ at the optimal estimate $\hat{\vec{x}}$.

As mentioned in Section 2.1.1, the observational error $\vec{\varepsilon}_O$, and thereby also its corresponding ECM $\mathbf{S}_O$, is made up from error contributions from both the observation vector $\vec{y}$ and the model operator $\mathbf{F}$. For the former, that contribution is the instrument error $\mathbf{S}_I$. For the latter, the contribution is split in the model error $\mathbf{S}_M$ (e.g. uncertainties introduced by the parameters $\vec{p}$, incomplete physics, or numerics) and the

representation error $\mathbf{S}_{\mathrm{R}}$ (uncertainties in the comparison to the actual measurements e.g. due to limited model resolution or a discrete grid in space and time). All three contributions can generally be assumed to be uncorrelated so that total observational error covariance matrix is

$$\mathbf{S}_{\mathrm{O}} = \mathbf{S}_{\mathrm{I}} + \mathbf{S}_{\mathrm{R}} + \mathbf{S}_{\mathrm{M}} \tag{2.12}$$

With this, the size of the second term in the cost function (Equation (2.11)) depends simultaneously on the errors from both the model and the measurements. Therefore, if either the model or the measurements are of poor quality, regardless of the quality of the other, then the optimal state will be poorly constrained.

Furthermore, the weighting between $\mathbf{S}_{\mathrm{A}}$ and $\mathbf{S}_{\mathrm{O}}$ may impact the result. Note that for this comparison not only the respective amplitudes of the ECMs matter, but also their dimensions. If the number of observations is very large compared to the size of the state vector $(m \gg n)$, the number of observational terms in cost functions might overwhelm the number of prior terms, making the inversion insensitive to the prior. Similar results are obtained when the measurement is very precise (small $\mathbf{S}_{\mathrm{O}}$), but the prior is not (large $\mathbf{S}_{\mathrm{A}}$). The other way around, when the measurements have very low confidence or are very low in number compared to the size of the state vector, simply the prior is returned. As pointed out before, both $\mathbf{S}_{\mathrm{A}}$ and $\mathbf{S}_{\mathrm{O}}$ are usually just rough estimates. One method to improve their relative weighting could be scaling the observational term of the cost function by a factor. By repeatedly solving the inverse problem while varying this factor, an improved solution may be found. From a physical standpoint, this factor is intended to capture autocorrelations within the observations that occur for dense datasets, but are often ignored in $\mathbf{S}_{\mathrm{O}}$. These autocorrelations would otherwise give too much weight to the observational part of the cost function. This approach is further elaborated for a concrete application in Section 3.3.2.2.

2.1.4 Solving the inverse problem

2.1.4.1 Analytical solution

The minimum of the cost function (Equation (2.11)) can be found by setting its gradient with respect to $\vec{x}$ equal to zero:

$$\nabla_{\vec{x}} J(\vec{x}) = 2\mathbf{S}_{\mathrm{A}}^{-1}(\vec{x} - \vec{x}_{\mathrm{A}}) + 2\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{O}}^{-1}(\vec{y} - \mathbf{F}(\vec{x})) = 0 \tag{2.13}$$

with

$$\mathbf{K} = \nabla_{\vec{x}}\mathbf{F} = \frac{\partial \vec{y}}{\partial \vec{x}} \tag{2.14}$$

as the derivative of the forward model $\mathbf{F}$, which is called its Jacobian matrix. The quantity $\mathbf{K}^{\mathrm{T}}(= \nabla_{\vec{x}}\mathbf{F}^{\mathrm{T}})$ is the adjoint of the forward model, which is equal to the transpose of its Jacobian. Assuming a linear forward model, or at least a model that can be linearized around $\vec{x}_{\mathrm{A}}$, we can replace $\mathbf{F}(\vec{x})$ by $\mathbf{K}\vec{x}$ and find the solution to Equation (2.13) as

$$\hat{\vec{x}} = \vec{x}_{\mathrm{A}} + \mathbf{G}(\vec{y} - \mathbf{K}\vec{x}_{\mathrm{A}}) \tag{2.15}$$

with

$$\mathbf{G} = \mathbf{S}_{\mathrm{A}}\mathbf{K}^{\mathrm{T}}(\mathbf{K}\mathbf{S}_{\mathrm{A}}\mathbf{K}^{\mathrm{T}} + \mathbf{S}_{\mathrm{O}})^{-1} \tag{2.16}$$

as the gain matrix, which gives the sensitivity of the optimal estimate $\hat{\vec{x}}$ to the observations $\vec{y}$. Finally, the error of the optimal estimate can now be fully characterized by the posterior error covariance matrix $\hat{\mathbf{S}}$, which can be calculated as

$$\hat{\mathbf{S}} = (\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{O}}^{-1}\mathbf{K} + \mathbf{S}_{\mathrm{A}}^{-1})^{-1} \tag{2.17}$$

The quality of the solution can be roughly evaluated by checking the following: Firstly, the optimal estimate should be an improvement over the prior estimate, i.e. $J(\hat{\vec{x}}) < J(\vec{x}_{\mathrm{A}})$. Secondly, the consistency of the assumed errors may be verified with $J(\hat{\vec{x}}) \approx m + n$, i.e., on average, each of the n elements of the prior and each the m elements of the observational vector should be reproduced within their respective variances. If, however, the cost of the optimal state $J(\hat{\vec{x}})$ is notably larger than the sum of those dimensions, at least some of the errors were likely assumed to be too small and vice versa for a too small cost and too large errors. Thirdly, the mismatch between the model and the observations should be improved, which can be verified by comparing $\mathbf{F}(\hat{\vec{x}}) - \vec{y}$ and $\mathbf{F}(\vec{x}_{\mathrm{A}}) - \vec{y}$. The former should overall be smaller and, ideally, uniform white noise around zero. Otherwise, the model might be biased, the characterization of errors might be faulty, or the choice of state might be improper for the system at hand.

Even if all these checks are successful, the analytical solution to the inverse problem presented above still has a number of limitations. Firstly, any a prior information that should be included must be expressible through the $P(\vec{x})$ part of Bayes' theorem. Secondly, the assumption of Gaussian distributed errors on the prior and the observations may not always hold and can potentially lead to negative entries in the optimized state vector, which may not be desirable if the state represents concentrations or emissions. A way to circumvent the issue of negative state elements will be discussed at the end of the next section. Thirdly, the analytical solution of the inverse problem requires explicit construction of the Jacobian matrix $\mathbf{K}$ for the forward model. This process

is computationally expensive for large state vectors and may be impractical for non-linear models, where it potentially needs to be repeated many times before a solution is found. For problems where the state vector is large, but the observational vector is not, it might be easier to construct the adjoint of the forward model $\mathbf{K}^{\mathrm{T}}$ instead. This construction still requires linearization of the model, which, again, is not always trivial. Generally, the larger the observational and state vectors are, the harder the Jacobian and its transpose are to estimate. If both vectors are very large and explicit construction of either matrix is no longer feasible, the adjoint model may be used instead, which implicitly models their impact given a state.

2.1.4.2   Adjoint-based inversion and 4DVAR

As pointed out in the previous section, analytical inversions require a linear(ized) forward model $\mathbf{F}$ and are limited by computational power. Sufficiently complex problems, therefore, require solving the inverse problem numerically instead, by employing variational methods, which imply a reduced computational cost when considering large state and observational vectors. Simply put, the gradient of the cost function $\nabla J(\vec{x})$ is repeatedly evaluated for different guesses for the state $\vec{x}$, while theses guesses are improved with each iteration by some steepest-decent algorithm. In the following, one specific implementation of this process, the adjoint-based inversion, will be introduced, which efficiently computes the gradient of the cost function using the adjoint of the forward model.

In an adjoint-based inversion, the transpose of the Jacobian $\mathbf{K}^{\mathrm{T}}$ is never actually constructed. Instead, as illustrated in Figure 2.1, in each iteration, the forward model $\mathbf{F}$ runs forwards in time, followed by the adjoint model, i.e. the adjoint of the tangent linear model, running backwards in time. In case of a linear problem, the tangent linear model is identical to the forward model. Each forward model run produces a set of adjoint forcings, i.e. the error weighted mismatch between the model and the observations $\mathbf{S}_{\mathrm{O}}^{-1}(\vec{y} - \mathbf{F}(\vec{x}))$. The adjoint forcings serve as input to the adjoint model, which then directly returns the gradient of the cost function. Based on that gradient, the steepest-descent algorithm then updates the state for the next iteration.

To express this process mathematically, consider Equation (2.14) with a CTM with discrete time steps t_i as the forward model $\mathbf{F}$. The sensitivity of the ensemble of observations $\vec{y}_{(p)}$ at time t_p to the state $\vec{x}_{(0)}$ at time t_0 can be expressed as $\mathbf{K} = \left(\frac{\partial \vec{y}_{(p)}}{\partial \vec{x}_{(0)}} \right)$. By applying the chain rule and transposing, $\mathbf{K}^{\mathrm{T}}$ can be expressed as a series of forcings for each time step between t_0 and t_p:

$$\mathbf{K}^{\mathrm{T}} = \left(\frac{\partial \vec{y}_{(0)}}{\partial \vec{x}_{(0)}} \right)^{\mathrm{T}} \prod_{i=1}^{p} \left(\frac{\partial \vec{y}_{(i)}}{\partial \vec{y}_{(i-1)}} \right)^{\mathrm{T}} \tag{2.18}$$

15
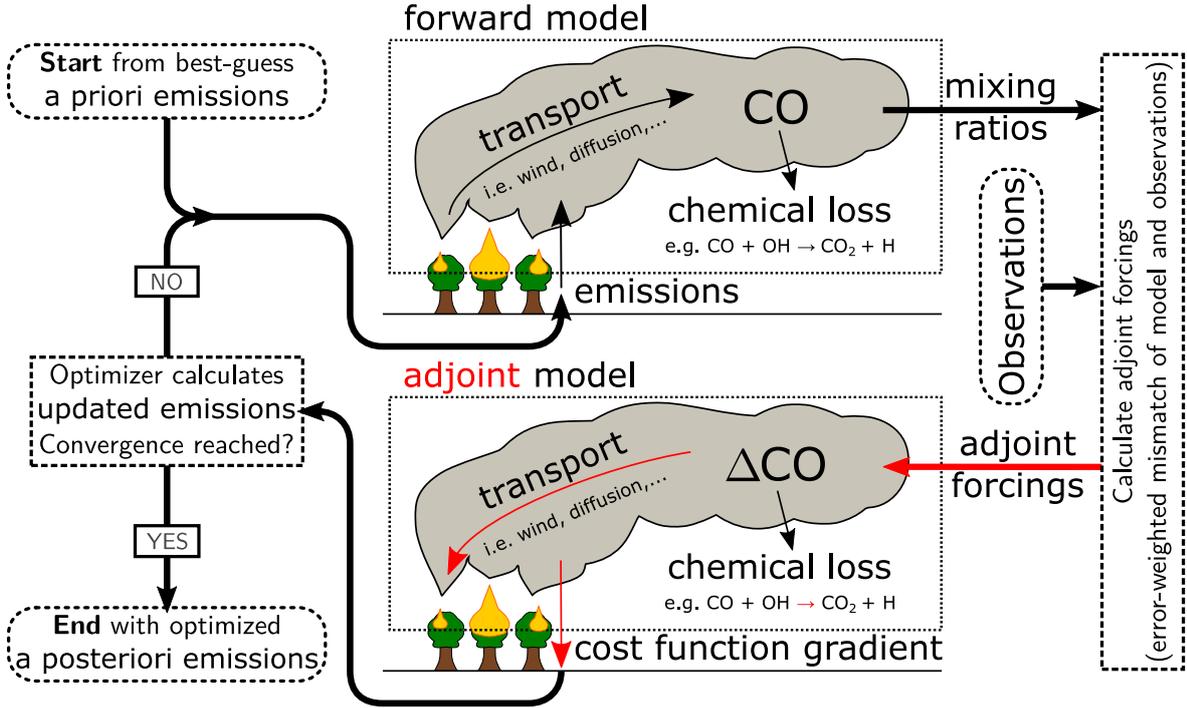
Figure 2.1: Schematic representation of an inverse modeling framework to find optimized emissions. Note that the chemistry operator is self-adjoint, i.e. chemical loss in the forward model does not turn in into production in the adjoint.

Each $\left(\frac{\partial \vec{y}_{(i)}}{\partial \vec{y}_{(i-1)}}\right)$ corresponds to a time step in the adjoint model. With this expression for $\mathbf{K}^{\mathrm{T}}$, the gradient of the cost function from Equation (2.13) may be evaluated. The initial iteration is solely driven by the prior $\vec{x}_{\mathrm{A}}$, which causes the first term of the gradient to be zero:

$$\nabla_{\vec{x}} J(\vec{x}_{\mathrm{A}}) = 2\mathbf{K}^{\mathrm{T}} \mathbf{S}_{\mathrm{O}}^{-1}(\vec{y} - \mathbf{F}(\vec{x}_{\mathrm{A}})) \tag{2.19}$$

After running the forward model (from time t_0 to t_p), the adjoint model runs backwards in time (from time t_p to t_0). Mathematically, the latter is expressed by, starting with $i = p$, applying the adjoint operator $\left(\frac{\partial \vec{y}_{(i)}}{\partial \vec{y}_{(i-1)}}\right)^{\mathrm{T}}$ to the adjoint forcing $\mathbf{S}_{\mathrm{O}}^{-1}(\vec{y}_{(i)} - \mathbf{F}(\vec{x}_{\mathrm{A}}))$, which results in the adjoint variables for the pth time step. This process is then repeated successively for the next and all following time steps ($i = p - 1$ to 1), however, each time before applying the adjoint operator, the adjoint variables for the previous step are added to the adjoint forcings of the current time step. The adjoint operator for the final time step ($i = 0$) is $\left(\frac{\partial \vec{y}_{(0)}}{\partial \vec{x}_{(0)}}\right)^{\mathrm{T}}$ to yield the full $\mathbf{K}^{\mathrm{T}} \mathbf{S}_{\mathrm{O}}^{-1}(\vec{y} - \mathbf{F}(\vec{x}_{\mathrm{A}})) \propto \nabla_{\vec{x}} J(\vec{x}_{\mathrm{A}})$, which leads a value for the derivative of the Jacobian (as stated above in Equation 2.19). This value is then used in the steepest-decent algorithm to find an updated guess $\vec{x}_1$ for the state.

In the next iteration, since $\vec{x}_1 \neq \vec{x}_{\mathrm{A}}$, the first term of gradient of the cost function

is also non-zero:

$$\nabla_{\vec{x}} J(\vec{x}_1) = 2\mathbf{S}_A^{-1}(\vec{x}_1 - \vec{x}_A) + 2\mathbf{K}^T \mathbf{S}_O^{-1}(\vec{y} - \mathbf{F}(\vec{x}_1)) \qquad (2.20)$$

However, the evaluation of the second term proceeds as in the initial iteration and a further improved state $\vec{x}_2$ can be obtained. This process is repeated until the optimal estimate $\hat{\vec{x}}$ is found. Practically, the gradient will virtually never reach exactly zero. Therefore, some other convergence criterion is usually defined at which the current best guess $\vec{x}_i$ is accepted as the optimal estimate. The convergence criterion could be, for example, a fixed number of iterations. Details on the concrete implementations used in this thesis can be found in Section 3.2.2. One shortcoming of the adjoint-based inversion is that it does not provide the posterior error covariance matrix $\hat{\mathbf{S}}$ for the optimal estimate.

Adjoint-based inverse modeling as outlined above may be implemented using the strategy of 4-D variational data assimilation (4DVAR). The 4 dimensions (4D) are the three spatial dimension of the forward model and time, emphasizing that, under this strategy, the observations that are assimilated are spread out over future time steps after the state they constrain, and that the model is sampled for each individual observation at its time of measurement. This strategy is in contrast to, for example, 3DVAR, where the state is optimized based on observations only at discrete time steps and only observations that occur at the same time as the state they update are considered.

The initial description of the 4DVAR approach can be found in TALAGRAND & COURTIER [1987], who applied it to meteorological assimilations. FISHER & LARY [1995] later extended the approach to assimilate atmospheric chemistry and ESKES et al. [1999] included satellite observations as powerful constraints. All these applications were strongly limited by computational power. However, with rising computational capabilities and more extensive datasets, the field flourished since.

Overall, the solution for 4DVAR is mathematically the same as for the general adjoint-based inversion shown above. In the following chapters, we will consider the specific case of emissions, including their temporal evolution, as the state vector. In this case, the state is not only the initial condition ($\vec{x}_{(0)}$), but also has entries spread over future time steps ($\vec{x}_{(j)}$) and each entry is updated by any observations for future time steps t_j to t_p, only limited by the lifetime of the tracer.

Since in this section the same cost function (Equation (2.13)) as for the analytical solution presented in the previous section has been used, the limitations imposed by assuming Gaussian distributed error statistics for the prior and observational vectors still hold. Most notably, Gaussian error statistics allow for negative entries in the optimal state, which can be problematic if the state is meant to represent emissions or

concentrations. To circumvent this issue, in the following, the approach of a 'semiexponential description of the PDF' as outlined in BERGAMASCHI et al. [2009] will be used. Instead of optimizing the emission vector $\vec{e}$ directly, a parameter vector $\vec{x}$ of the same length is optimized. The semiexponential description connecting $\vec{x}$ and $\vec{e}$ is

$$\vec{e} = \begin{cases} \vec{e}_{\mathrm{A}} \odot \exp(\vec{x}), & \text{for } x < 0 \\ \vec{e}_{\mathrm{A}} \odot (\vec{1} + \vec{x}), & \text{for } x \geq 0 \end{cases} \tag{2.21}$$

where $\vec{e}_{\mathrm{A}}$ are the a priori emissions, $\vec{1}$ is a vector where every entry is 1 with the same length as $\vec{x}$ and $\odot$ is the element-wise vector multiplication. The parameter $\vec{x}$ is then optimized as outlined above, with Gaussian error statistics and an a priori of zero. This approach ensures that $\vec{e}$ will never be smaller than zero, however, it also makes the forward model operator $\mathbf{F}$ non-linear. While the tangent linear model, and with it the adjoint model, can still be obtained, the implementation for the steepest-decent algorithm must be capable of handling non-linear problems.

2.2 TM5-4dvar model description

Throughout this thesis, the state-of-the-art inverse modeling framework TM5-4dvar is used as a concrete implementation of the 4DVAR inverse modeling approach outlined in the previous section. The original TM5-4dvar inverse modeling framework, which is based on TM4-4dvar [MEIRINK et al., 2006], was first applied in MEIRINK et al. [2008a] and is described in detail in MEIRINK et al. [2008b]. The first inversion studies based on the global atmospheric chemistry transport model TM5 [KROL et al., 2003] and the extended TM5-zoom [KROL et al., 2005] and their respective adjoint versions have been performed by GROS et al. [2003, 2004] for methyl chloroform and CO, and by BERGAMASCHI et al. [2005, 2007] for methane. The CO branch of the TM5-4dvar model used throughout this thesis was initially described in KROL et al. [2008] and has been applied in multiple studies since [HOOGHIEMSTRA et al., 2011, 2012a,b; KROL et al., 2013; NECHITA-BANDA et al., 2018].

Conceptually, in a model based inversion, one or more sets of observations of atmospheric mixing ratios are used to optimize a set of prior emissions (a priori) to find the set of posterior emissions that agrees best with these observations (a posteriori). The a priori emissions are usually taken from bottom-up inventories. Inverse modeling is then used to reduce the uncertainties of these inventories, by including information from additional observations, which can range from spatially and temporally sparse surface flask data [BERGAMASCHI et al., 2000; PÉTRON et al., 2002; BUTLER et al., 2005; PISON et al., 2009; HOOGHIEMSTRA et al., 2011], over local aircraft measurements [PALMER et al., 2003; HEALD et al., 2004], to global satellite observations [PÉTRON

et al., 2004; ARELLANO et al., 2004; FORTEMS-CHEINEY et al., 2009; HOOGHIEMSTRA et al., 2012a], or even combinations of multiple such datasets [HOOGHIEMSTRA et al., 2012b; KROL et al., 2013; JIANG et al., 2017; NECHITA-BANDA et al., 2018].

TM5-4dvar is an offline model which handles atmospheric transport using prepro-cessed meteorological fields. For advection, the slope scheme developed by RUSSELL & LERNER [1981] is employed. Simply put, in addition to the tracer mass, each box is assigned three slope values to capture the gradients in noroth-south, east-west, and up-down directions. The slopes level out over time and increase whenever tracer mass enters or leaves the cell.

The TM5-4dvar model allows for zooming [KROL et al., 2002] based on the technique described in BERKVENS et al. [1999], where only the region of interest is simulated at a high resolution (up to $1° \times 1°$; longitude $\times$ latitude), while the rest of the globe is simulated at a reduced resolution ($6° \times 4°$). The regions are two-way nested, so the coarser region does not only provide boundary conditions for the finer region in each time step but is, in return, itself also updated by the more precise results of the finer region.

This thesis, specifically, makes use of the simplified CO-only chemistry version of TM5-4dvar described in HOOGHIEMSTRA et al. [2011], where only the direct and indirect (via O_2) reactions of CO with OH (combined in Reaction (1.1)) are considered explicitly. As for the meteorology, the OH concentrations are prescribed either from climatological fields or via fields taken from a full chemistry model. As a second loss process, dry deposition is considered based on the parameterization from GANZEVELD et al. [1998], adapted for TM5.

2.3   Satellite observations

Observations from satellite instruments provide valuable input for global inverse model-ing. For one, satellites usually feature significantly higher spatial coverage (e.g. global) than in-situ measurements. Secondly, satellites provide long, consistent time series, covering years to decades, which is comparable to what surface stations provide and better than e.g. aircraft campaigns. However, those advantages come at a large com-putational cost because both imply large amounts of data, which needs to be stored, read, and processed. Notably, the processing includes the application of the averaging kernels of the satellite retrieval, see Section 2.3.1. There are ways to reduce the com-putational cost introduced by large numbers of observations. One such approach, the use of super-observations, is introduced in Section 2.3.2.

There are further short-comings of satellite observations in general; however, these are less significant for inverse modeling in particular. For example, global satellite in-struments tend to lack temporal resolution, with revisiting times in the order of days

or longer. In contrast, especially ship- or aircraft-borne in-situ campaigns often feature much higher temporal resolutions for the regions they investigate. However, one of the key features of (inverse) modeling is the ability to fill in gaps in the knowledge about the atmospheric state by observing the same air parcels multiple times over longer periods. Another short-coming of satellite observations are the relatively large uncertainties of individual measurements, especially with respect to the vertical tracer distribution. These uncertainties can be, at least partly, compensated by the large number of observations, and by using satellite data in conjunction with in-situ measurements, which add further constraints to the vertical distribution in case of co-located measurements.

2.3.1   Averaging kernels and comparisons to model data

Inverse modeling with the 4DVAR method (Section 2.1.4.2) requires sampling of the model at the time and place of an observation. This sampling process is relatively straight forward for most in-situ measurements, where simple spatial or temporal interpolations suffice. However, for satellite observations, the sampling is not as trivial, since how those observations are derived from a spectral measurement of the true atmosphere depends on the retrieval process. Part of the retrieval is obtaining an averaging kernel $\vec{A}$ for each measurement, which signifies the sensitivity of the satellite instrument to the true atmosphere over altitude and depends on various environmental parameters, e.g. the solar zenith angle. To derive and apply the averaging kernels, assumptions about the pressure profile $\vec{p}$ and a priori tracer profile $\vec{x}_{\mathrm{apr}}$ need to be made.

Therefore, to properly sample the model and obtain a quantity that is comparable to the satellite observation, we need to simulate what the satellite would have seen if it observed the modeled atmosphere rather than the true atmosphere. The first step is the same as for the in-situ measurements and consists of sampling the model profile at the time and place of each satellite observation and potentially applying spatial or temporal interpolations. The next step is rebinning the model profile from the M pressure levels of the model to the N pressure levels used during the retrieval, which may differ in number, thickness, distribution, and boundaries. The rebinned model profile $\vec{x}_{\mathrm{mod}}$ is combined with the information from the retrieval as the final step to yield the value of the model operator $\mathbf{F}$ (see Section 2.1.1) for the ith observation as

$$\mathbf{F}(\vec{x})_i = \sum_{l=1}^{N} \frac{p_{l-1,i} - p_{l,i}}{p_{0,i}} \left[A_{l,i} \left(x_{\mathrm{mod},l,i} - x_{\mathrm{apr},l,i} \right) + x_{\mathrm{apr},l,i} \right] \tag{2.22}$$

where p_0 is the surface pressure. Note that the pressure $\vec{p}$ is given at the boundaries of the retrieval layers and, therefore, has $N + 1$ entries.

2.3.2   Super-observations

Super-observations are obtained by combining multiple adjacent observations into a single observation. This process reduces the resolution of the observational dataset, which may or may not be detrimental in an inversion, depending on the relation between the model resolution and the resolution of the super-observations. A benefit of super-observations is that the smaller dataset leads to a reduced computational cost (reduced storage requirements, less I/O, less data to process). Further, each super-observation has a smaller observational error than the individual measurements it was created from. Therefore, super-observations can be useful to save time and storage, especially if the original (satellite) dataset had a much higher spatial resolution than the model. There are many methods to obtain super-observations and in the scope of this thesis a new method was developed, as presented in Section 3.3.2.1.

2   METHODS

Chapter 3

Efficacy of high-resolution satellite observations

This chapter presents a series of inversion experiments as a concrete implementation of the concepts presented in the previous chapters. Specifically, in the following, a new satellite dataset, observations from the TROPOMI instrument, is introduced into an existing inverse modeling setup, the impact of that dataset on the inversion results is investigated, and the novel methods developed to handle the dataset are described. Parts of this chapter have been submitted for publication.

Section 3.1 introduces the topic and goal of this chapter. In Section 3.2, short descriptions of the used model and the inverse modeling approach are provided, followed by an in-depth description of the model setup and the conducted experiments. Section 3.3 introduces the used observational datasets, including descriptions of the novel super-observation and error inflation techniques applied to the satellite observations. Section 3.4 consists of an in-depth analysis of the inversion results obtained with the new satellite dataset. Section 3.5 summarizes the results from this chapter.

3.1   Background and objective

Previous studies with the TM5-4dvar model (Section 2.2) employed satellite observations from the Measurements of Pollution in the Troposphere (MOPITT) instrument [Hooghiemstra et al., 2012a,b], the Infrared Atmospheric Sounding Interferometer (IASI) instrument [Krol et al., 2013] or both [Nechita-Banda et al., 2018]. Here, a new satellite dataset is introduced into the TM5-4dvar inverse model, by using combined data from (a) the high-resolution TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor (S5P) satellite and (b) the NOAA surface CO flasks from the ESRL Global Monitoring Laboratory and proposing an iterative process to more rigorously weight both datasets against each other in the in-

version. TROPOMI features several differences to, and advantages over MOPITT and IASI. Most notably, the TROPOMI CO retrievals are performed solely in the short-wavelength infrared [SWIR, around 2.3 µm; VEEFKIND et al., 2012] range, as opposed to IASI's mid-wavelength infrared [MWIR, around 4.76 µm; DE WACHTER et al., 2012] range. MOPITT uses mostly the thermal MWIR bands around 4.6 µm, assisted by the solar SWIR band around 2.3 µm [DRUMMOND et al., 2010]. By using shorter wavelengths, the TROPOMI retrievals exhibit less interference from Earth radiation and are, therefore, more sensitive to CO that resides close to the surface compared to MOPITT and IASI. Overall, TROPOMI has high sensitivity throughout the atmosphere, whereas MOPITT and IASI are most sensitive to the middle and upper troposphere. However, the combination with the SWIR band increases MOPITT's surface-level sensitivity under specific conditions [e.g. WORDEN et al., 2010]. Furthermore, TROPOMI procures CO observations at a high spatial resolution of up to $7 \times 7\,\mathrm{km}^2$ [VEEFKIND et al., 2012], which is roughly 10 times higher than the resolution of MOPITT [up to about $22 \times 22\,\mathrm{km}^2$; DRUMMOND et al., 2010] and the spatial sampling of IASI [up to about $25 \times 25\,\mathrm{km}^2$; CLERBAUX et al., 2009]. Additionally, TROPOMI takes one day to reach global coverage, which is comparable to IASI, whereas the MOPITT instrument takes about five days to achieve the same.

However, the TROPOMI observations correspond to a large data volume due to their high resolution and high coverage, which implies a large computational cost when using these data in the TM5-4dvar inversion suit. One established way to reduce the computational cost of global inversions is through zooming, where only a limited region is simulated at a fine resolution, while the rest of the globe is simulated at a coarser resolution. This way, the everlasting trade-off, where increasing the model resolution implies not only rising precision but also rising computational cost, can be partly overcome. This method has been proven to yield very similar results within the limited fine resolution region compared to simulations with fine resolution globally, while significantly reducing run times. Therefore, the coarser global simulation is still sufficient to provide meaningful boundary conditions to the finer region of interest. Intermediate regions may be used to provide more fluent transitions between the coarse and the fine region. Such nested grids can be found for example in TM5-4dvar [BERKVENS et al., 1999; KROL et al., 2005], and GEOS-Chem [WANG et al., 2004; CHEN et al., 2009].

Similarly, the resolution of satellite observations can be reduced by defining a grid and aggregating all observations within each cell of this grid into a single so-called super-observation [ESKES et al., 2003; MIYAZAKI et al., 2012; BOERSMA et al., 2016]. Here, a modified version of this super-observation approach is introduced, which is capable of reducing the number of observations in the dataset, which in turn reduces the computational cost they introduce in the inversion.

This chapter evaluates the added value of the new TROPOMI data in constrain-

ing global CO emissions. An investigation focusing on the emissions in the northern hemisphere in the second half of 2018 serves as a proof of concept. This investigation is split into a series of experiments, in which the same inversion is run multiple times, each time with slightly different settings. Firstly, CO emissions are optimized simultaneously towards TROPOMI satellite observation gridded to $0.5° \times 0.5°$ and NOAA surface flask measurements. This inversion will be used as a reference case, against which all other inversions are compared. For this reference inversion, an analysis of the increments to the a priori emissions at the global scale will be conducted, to identify short-comings in either the model or the bottom-up inventories that serve as a priori emissions. Secondly, by comparing the reference inversion to two inversions where the inventories used as biomass burning a priori emissions are varied, the influence of the a priori emissions is investigated. While biomass burning makes up less than a quarter of the total CO source in most years, those emissions come with the largest uncertainty (see Section 3.2.3.1 for more details), linked to their high spatial and temporal variability compared to the other sources. Thirdly, the inversion is repeated with the same a priori emissions as in the reference case two more times, once with only the TROPOMI satellite observations (and no flask data) and once with only the NOAA flasks (and no satellite observations). Comparing the results of those inversions with the reference inversion gives insight into the impact of the TROPOMI observation on the inversion results by highlighting areas where satellite observations and station measurements carry unique, redundant or even conflicting information. Finally, by also running the inversion with the full resolution satellite observations (up to $7 \times 7\,\mathrm{km}^2$) in combination with the NOAA surface flasks, the influence of gridded satellite observations on the model at its relatively coarse resolution of $3° \times 2°$ is analyzed.

3.2   Materials and methods

3.2.1   Model description

The Cycle 3 TM5-4dvar model (Section 2.2) as of revision 899b16 from the official code repository of the model[1] is used. In the scope of this chapter, the existing code is extended to handle the high-resolution TROPOMI observations. Additionally, support for anthropogenic emissions based on CMIP6 is implemented, the capabilities to use the output from the full-chemistry model TM5-MP as initial conditions and as a priori for the secondary sources of CO are extended, and some minor compatibility issues are resolved. The specific code version used here is available at Nüss et al. [2022a]. All other analysis and plotting scripts used throughout this chapter as well as any relevant model in- and outputs are collected and available at Nüss et al. [2023]

---

[1] `https://sourceforge.net/p/tm5/cy3_4dvar/`, last accessed 2024-01-09
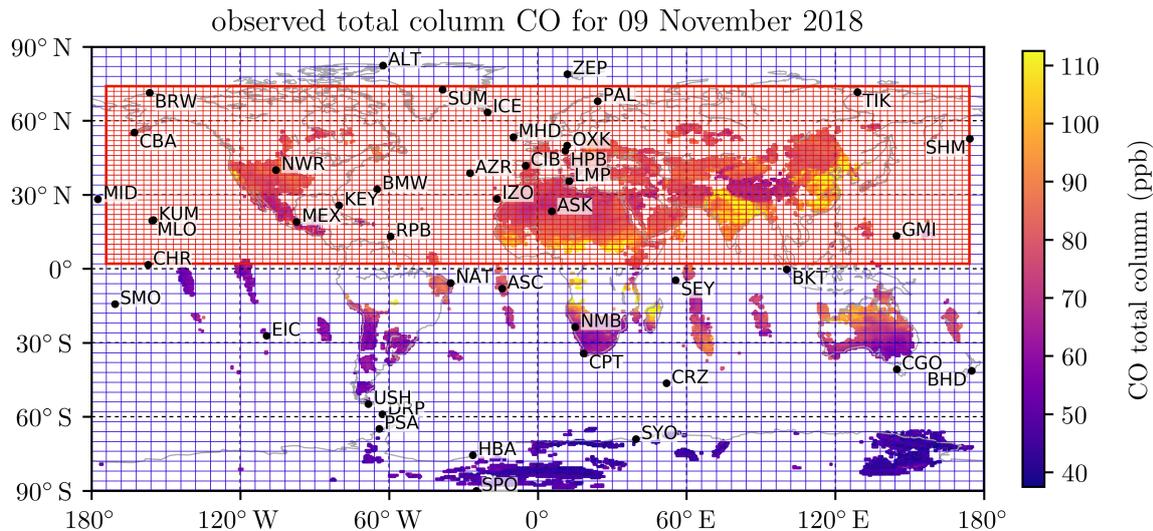
Figure 3.1: Used zooming setup, with $6° \times 4°$ grid globally (blue) and nested $3° \times 2°$ grid over the northern hemisphere (red). The locations of the background stations where the NOAA CO flask measurements are collected are shown as black dots and labeled with their respective station ID. The colormap shows the used global TROPOMI satellite observations for one day (9 November 2018) as an example of the daily coverage and resolution they provide. Due to strict quality filtering during the retrieval process [SCHNEISING et al., 2019], many places have no valid TROPOMI observations, despite every location on Earth being visible to the instrument at least once per day. A more comprehensive overview of the TROPOMI CO data coverage for all of 2018 can be found in Figures A.1 and A.2 in the appendix.

Throughout this thesis, the zooming capabilities of TM5-4dvar are used with a region of interest simulated at only a medium resolution of $3° \times 2°$, but covering a very large area. The region of interest is placed over the northern hemisphere, spanning $2°$ N–$74°$ N and $174°$ W–$174°$ E and captures all major land masses, as shown in Figure 3.1.

The meteorological fields used to drive the inversion experiments presented in this chapter are taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis project [ERA-Interim meteorology; DEE et al., 2011] coarsened to the lateral model resolution and 34 altitude levels (from surface pressure to the top of the atmosphere (0 hPa), with the highest resolution in the Upper Troposphere-Lower Stratosphere (UTLS)). Similarly, the OH concentrations used to simulate chemical loss of CO are prescribed by the widely used monthly climatological fields from the TransCom-CH$_4$ project described in PATRA et al. [2011], in which tropospheric OH is based on the OH fields from SPIVAKOVSKY et al. [2000], which are scaled by 0.92 as suggested in HUIJNEN et al. [2010].

3.2.2   4DVAR approach

An in-depth introduction into the mathematical basis of inverse modeling in general and the 4DVAR approach in particular can be found in Section 2.1.

Overall, in 4DVAR, the model is sampled temporally and spatially for each individual data point, and each point provides its own contribution to the cost function. As such, this approach is well capable of assimilating multiple datasets with different spatial and temporal resolutions at once and co-sampling of observations across datasets is neither necessary nor detrimental. One caveat is that the observations of different datasets need to be weighted properly against each other. On the one hand, this implies proper measurement error estimation. On the other hand, some form of error inflation (Section 3.3.2.2) might be required if datasets with vastly different numbers of observations are used, or if some datasets have a much higher resolution than the model.

Throughout this thesis, the inversions are carried out using the non-linear M1QN3 optimizer described in GILBERT & LEMARÉCHAL [1989]. This optimizer is capable of handling a semi-exponential description of the probability density function for the a priori emissions, which in turn avoids negative emissions [BERGAMASCHI et al., 2009], as detailed at the end of Section 2.1.4.2. As a convergence criterion, a reduction of the gradient norm of the cost function of 10^3 is chosen, i.e. the iterations are stopped once the cost function is one thousand times less steep. This criterion was suggested in MEIRINK et al. [2008b] to be sufficient to converge the emissions. With this criterion, it takes the model around 35 iterations to converge, whereas the budget terms are near constant for the last few iterations.

3.2.3   Model setup

The TM5-4dvar model, as described in Section 3.2.1, is used to perform multiple inversions of the CO emissions in the year 2018, with a specific focus on the northern hemisphere.

3.2.3.1   Inventories and emission categories

CO production from three distinct source categories – anthropogenic, biomass burning, and secondary CO production through chemistry – is considered. Since the contributions of oceanic and biogenic CO to the overall source are small compared to the aforementioned categories, they have been neglected in the experiments in this chapter, but will be considered explicitly in Chapter 4. Additionally, no daily cycles in emissions or chemistry were considered, mostly due to limitations of the OH climatology (see Section 3.2.1) and the secondary CO production a priori (introduced further down in this section).

3 EFFICACY OF HIGH-RESOLUTION SATELLITE OBSERVATIONS

The biomass burning a priori emissions are taken from FINN2.4+VIIRS [WIEDIN-MYER et al., 2023], i.e. the Fire INventory from NCAR version 2.4 including additional small fire detection via satellite observations from the Visible Infrared Imaging Radiometer Suite (VIIRS) and NMVOCs speciated to the Model for OZone And Related chemical Tracers (MOZART-T1) chemical mechanism [EMMONS et al., 2020]. FINN is based on three data products from the Moderate Resolution Imaging Spectroradiometer (MODIS) for active fires, land cover type, and vegetation continuous fields, which are used to infer burned area and fire emissions. Compared to the original FINNv1 [WIEDINMYER et al., 2011], the FINNv2 used in this thesis features an improved representation of large fires by merging overlapping fire pixel areas. Additionally, rather than using a single static vegetation map for all years, the respective MODIS land cover type and vegetation continuous field data from the previous year are used. Also, the fuel loadings and emission factors have been updated.

As a sensitivity study, additional inversion experiments are conducted where the FINN2.4+VIIRS inventory is replaced as the biomass burning a priori with (1) FINN2.4 (without VIIRS), and (2) emissions from the Global Fire Emissions Database version 4, including small fire boost [GFED4.1s; RANDERSON et al., 2017]. The inversion experiments are introduced in more detail in Section 3.2.3.4.

GFED4.1s is based on satellite observations of burned area from MODIS, and fire activity from both the Visible and Infrared Scanner (VIRS) and the Along Track Scanning Radiometer [ATSR; GIGLIO et al., 2013]. These observations are combined with datasets on vegetation characteristics and meteorology to infer burned area and fire emissions on monthly scales, along with scaling factors to receive higher (daily or 3-hourly) temporal resolutions [VAN DER WERF et al., 2017]. The small fire boost includes estimates for biomass burning emissions from fires that are below the detection limit of the burned area product (MODIS), but are still visible as thermal anomalies [RANDERSON et al., 2012]. While these estimates have fairly large errors on a local scale [ZHANG et al., 2018], including them leads to more realistic total biomass burning emissions on the regional to global scale of the model used in this thesis.

Both GFED and FINN are coarsened to the resolutions of the zooming regions and aggregated into daily bins to serve as global priors for the biomass burning emissions. After applying the emission factors, all fire types are lumped together into a single biomass burning fire type. Since both inventories only provide 2D surface level emissions, they are used in conjunction with injection heights from the IS4FIRES Integrated Monitoring and Modelling System for wildland fires developed at FMI [SOFIEV et al., 2012, 2013].

For calculating the contribution to the cost function, a grid-scale a priori error of 100 % is assumed globally for the biomass burning emissions. This error is constructed from the error of at least 50 % provided in VAN DER WERF et al. [2017] for the regional

carbon emissions in GFED4.1s, combined with the error of the emission factors that are used to convert the total (carbon) emissions of each fire type into distinct species (e.g. CO). These are fixed per fire type and are reported with an estimate of their natural variation in the order of one-third of the reported value [AKAGI et al., 2011]. Since GFED4.1s and FINN2.4(+VIIRS) are fairly similar in terms of spatial distribution and amplitude of wildfire emissions (see Figure A.3 in the appendix, note the logarithmic scale) and to keep the inversion results comparable, we assume an a priori error of 100 % for FINN2.4(+VIIRS) as well. Additionally, to prevent erroneous biomass burning emissions in the inversion result, the a prior error is set to zero over the oceans. While this implies fixed biomass burning emissions for some smaller islands, for example Hawaii, emissions from larger islands, for example Indonesia, are still optimized.

TM5-4dvar allows for spatial and temporal correlations for each emission category to be set. As introduced in Section 2.1.3, these reduce the effective number of degrees of freedom of the inversion, which can help to prevent overfitting of the observations and lead to more realistic results, while also reducing the number of iterations needed to reach convergence [MEIRINK et al., 2008b]. The numeric values for the spatial correlation lengths and temporal correlation times stated in the following are empirical and follow the values provided in KROL et al. [2013] and NECHITA-BANDA et al. [2018], who used a similar setup with the same model. Biomass burning events are usually fairly temporary, so a short exponentially decreasing correlation time of 0.1 months for emissions at different times in the same grid cell is used. The usually small spatial extent of biomass burning events (compared to the coarse model resolution) is accounted for by using an exponentially decreasing correlation length of only 200 km for emissions at the same time in neighboring grid cells. The biomass burning emissions in the state are optimized at a daily resolution (i.e. the optimizer can change the biomass burning emissions for each day separately) to best capture the high temporal frequency of the burning events and therefore maximize the distinction between the biomass burning emissions and the other categories. Previous studies [e.g. KROL et al., 2013; NECHITA-BANDA et al., 2018] used a 3-daily resolution and in KROL et al. [2013] a sensitivity study with daily resolution was conducted with mixed results.

Secondary CO production from the oxidation of CH_4 and other VOCs is based on 3D production fields from a simulation of the full chemistry model TM5-MP with the extended MOGUNTIA chemical scheme described in MYRIOKEFALITAKIS et al. [2020] for the year 2018. This source is optimized with a fairly conservative a priori error of only 20 %. Fairly gradual changes in time may be expected for this source. Therefore, an exponentially decreasing correlation time of 9.5 months is used for the secondary CO production from the same cell, but at different times. Note that this rather restrictive correlation time does not limit the model's ability to capture the seasonality of short lived VOCs like isoprene, since that seasonality is already included

in the prior production fields. Instead, it only limits how much the deviations from those prior fields may vary from month to month. Similarly, spatial emission changes are also expected to be gradual for secondary production, due to the well-mixed CH_4 background, leading to an exponentially decreasing correlation length of $1000\,\text{km}$. A monthly resolution in the state is chosen for the secondary CO production, i.e. the optimizer can change it only once per month and the production is constant over the course of that month. Choosing this much coarser state resolution compared to the daily resolution for biomass burning emissions, makes it cheaper, with respect to the cost function, for the optimizer to capture the usually short time scale biomass burning events with the intended emission category. All of this combined, the low a priori error, low state resolution, and large temporal and spatial correlation, reduces aliasing between the smooth fields of this category and the more patchy biomass burning emissions. Conversely, since NMVOC oxidation can be quite local occasionally, this approach bears the risk of capturing part of the secondary production in the biomass burning emissions, specifically when the NMVOCs are emitted by fire activity.

Anthropogenic CO emissions are taken from the Climate Model Intercomparison Project 6 (CMIP6) inventory [EYRING et al., 2016], specifically the SSP370 [FUJIMORI et al., 2017; RIAHI et al., 2017; GIDDEN et al., 2019] projection dataset [GIDDEN et al., 2018]. Due to the low interannual variation of anthropogenic emissions compared to secondary CO production or biomass burning emissions and the fairly up to date inventory (with historical data up to 2014 and projected data from 2015 onwards), a conservative a priori error of $10\,\%$ is assumed, with the same monthly state resolution as for the secondary production. Following the same argument as for secondary CO production, an exponentially decreasing correlation time of 9.5 months is used. Similarly, spatial changes in anthropogenic emission are expected to occur on the level of countries or economic zones, leading to an exponentially decreasing correlation length of $2000\,\text{km}$. As for the biomass burning emissions, changes to these anthropogenic emissions are restricted to land. Thus, shipping emissions are included in the inventory, but not optimized.

3.2.3.2   Simultaneous inversion of multiple emission categories

As mentioned in the previous section, anthropogenic emissions, biomass burning emissions, and the secondary CO production are optimized simultaneously, i.e. they are all part of the state vector $\vec{x}$ (Section 2.1) and the optimizer could adjust any of them to minimize the cost function. This approach will inadvertently lead to some aliasing between the categories, despite the rigid choices for the a priori error, correlation length and time, and state resolution for the secondary production category. However, optimizing the biomass burning emissions on their own is not an option either, since this will force the model to represent any mismatches by adjusting the biomass burn-

ing emissions, even if these mismatches actually stemmed from flaws in the chemical production or anthropogenic a priori. This extreme form of aliasing leads to very poor convergence at the background stations, even when extremely high a priori errors are assumed. Using not only sparse flask data, but also the high coverage, high resolution TROPOMI observations, might help the optimizer to better distinguish between the emission categories.

3.2.3.3   Initial conditions, spin-up and main inversions

The initial tracer distribution is an important part of the inversion. Close to the starting date of the inversion period, the initial tracer distribution must fit the total columns and horizontal distribution of the observational datasets reasonably well. If there are considerable over- or under-estimations, the emission increments will be dominated by the efforts of the model to fix the offset in the mixing ratios. These additional emission will mask the actual signal of the observations, i.e. by how much the a priori emissions differ from the true emissions. Additionally, the initial vertical CO distribution must be realistic, since the CO depletion and transport vary with altitude. Therefore, assuming too high initial mixing ratios in a layer with low transport and low loss will affect the model for a long time.

To accommodate this, the period of interest (the year 2018) is split into two separate inversions. The first period is a spin-up inversion to harmonize the global distribution of CO mixing ratios in the model with the observational datasets (see Section 3.3). Harmonizing the model and the observations, especially in remote regions where transport is slow, requires the model to be run over a prolonged period of time. Therefore, the spin-up inversion is run over multiple months, from 1 January 2018 to 1 July 2018. The second period is the main inversion, which uses the harmonized mixing ratios from the spin-up inversion as initial conditions. It is run from 1 June 2018 over seven months until 1 January 2019 and leads to the results of scientific interest presented in Section 3.4.

Note the one month of overlap in the inversion periods. This overlap is necessary, because emissions close to the end of the inversion period are verified by very few observations. Therefore, the final month of the spin-up inversion is considered as its spin-down period, during which confidence in the generated emissions and the resulting mixing ratios is diminished. Similarly, the final month of the main inversion, December 2018, should be considered as its spin-down period. The duration of this period was chosen based on the lifetime of CO of around two months [RAUB & MCMULLEN, 1991; HOLLOWAY et al., 2000]. Hence, a snapshot of the mixing ratios from the final iteration of the spin-up inversion from 1 June 2018 is used as initial conditions for the main inversion.

The spin-up inversion itself is started with tracer fields taken from the chemistry

Table 3.1: A priori emissions and observational setup for the experiments conducted in Chapter 3.

		Inversion	A priori emissions			Observations	
			biomass burning	anthrop.	secondary	satellite	flasks
		spin-up	FINN2.4+VIIRS			gridded	yes
Main inversions	Set 1	*reference*	FINN2.4+VIIRS			gridded	yes
		noVIIRS	FINN2.4			gridded	yes
		GFED	GFED4.1s	CMIP6	TM5-MP	gridded	yes
	Set 2	*satellite only*	FINN2.4+VIIRS			gridded	no
		stations only	FINN2.4+VIIRS			none	yes
		full satellite	FINN2.4+VIIRS			full	yes

transport model TM5-MP, which employed the MOGUNTIA chemistry scheme. In Myriokefalitakis et al. [2020] and sources therein, a detailed description of the model, setup, and chemistry scheme, alongside extensive validation versus observational data can be found. In addition to the simulation analyzed and described there, the TM5-MP model was run with the same settings for a longer time frame, including 2018. Here, instantaneous concentrations from that simulation are used as initial conditions for the spin-up, and monthly chemical budget terms for the secondary source of CO from VOC oxidation.

The validations in Myriokefalitakis et al. [2020] have shown that the TM5-MP model generally produces reasonably realistic tracer fields in terms of both vertical and horizontal distributions. However, some offsets to the observations still remain. For CO specifically, Myriokefalitakis et al. [2020] found too low mixing ratios in the northern hemisphere and too high mixing ratios in the southern hemisphere. The spin-up inversion is necessary to confidently remove those offsets.

Additionally, the spin-up inversion facilitates a smooth transition between the different emission datasets used by Myriokefalitakis et al. [2020] in TM5-MP and in the experiments presented here in TM5-4dvar. While the experiments presented here also use CMIP6 for anthropogenic CO and the same meteorology, Myriokefalitakis et al. [2020] use CMIP6 also for biomass burning, while here FINN2.4 or GFED4.1s is used. This discrepancy is accepted, because, for 2018, both these inventories provide historical data rather than projections and inversions strongly benefit from realistic lateral a priori distributions, which cannot be obtained from projection data as those in CMIP6. Another important difference is the handling of OH. While Myriokefal-itakis et al. [2020] calculate their OH online, here prescribed OH fields, as described in Section 3.2.1, are used. Both, the biomass burning prior and OH fields, are further investigated and exchanged in Chapter 4.

3.2.3.4   Inversion experiments

Table 3.1 gives an overview of the experimental setups for the inversions analyzed in this chapter. The main inversion period (1 June 2018 to 1 January 2019) is chosen based on the availability of the used input data and computational constraints. Regarding the input data, TROPOMI was in its commissioning until March 2018 and the ERA-Interim meteorology dataset ends in August 2019. The latter constraint will be lifted for future studies by switching to ERA5 meteorology [Hersbach et al., 2020], as described in Chapter 6. Still, the large zooming region over most of the northern hemisphere, which is chosen to gain deeper insight into the general anthropogenic emission patterns, combined with the long inversion period come at a high computational cost. Each inversion takes about five real-world days to run (even longer with the full resolution satellite observations). Therefore, the inversion period does not extend into 2019. In this chapter, emissions for this period are optimized a total of six times with different settings, split into two sets.

In the first set, the biomass burning a priori emissions are varied, while using the same observations (global gridded TROPOMI observations in conjunction with flask measurements from the NOAA background stations) to constrain the emissions. More details on the a priori emission inventories and the observations used, including the gridding process, can be found in Sections 3.2.3.1 and 3.3, respectively. These inversions are intended to investigate the sensitivity of the optimized emissions to the a priori, by introducing a new and updated version of FINN into the model and applying a significantly lower grid-scale biomass burning a priori error compared to previous studies. The first set includes (1) the *reference* inversion with FINN2.4+VIIRS, (2) the *noVIIRS* inversion with regular FINN2.4 and (3) the *GFED* inversion with GFED4.1s.

In the second set, the biomass burning emissions are kept fixed to the *reference* case (FINN2.4+VIIRS) and the observational datasets are varied. This way, it is possible to assess the information content in the different datasets and the loss of information through gridding. The second set includes (4) the *full satellite* inversion using the full resolution satellite data in conjunction with the NOAA surface flasks, (5) the *satellite only* inversion using only the gridded satellite observations but no surface flasks and (6) the *station only* inversion using no satellite observations at all, where the inversion is driven solely by the surface flasks.

For the *spin-up* inversion (1 January 2018 to 1 July 2018) the same setup as for the *reference* inversion is used, i.e. FINN2.4+VIIRS as biomass burning a priori and gridded satellite observations in conjunction with NOAA surface flasks. All of the main inversions are started from this one *spin-up*, to ensure comparability of the results.

3.3   Observations

3.3.1   In situ measurements

The in situ observations used here are the NOAA surface flask CO measurements from various stations assembled by the Carbon Cycle Greenhouse Gases (CCGG) group [Petron et al., 2020]. For filtering out non-background stations, the algorithm described in Hooghiemstra et al. [2012a] is applied to the 54 stations active between January and December 2018. Following this, only the 44 stations shown in Figure 3.1 are classified as background and subsequently used. This filtering is necessary to avoid the large representation error introduced by non-background stations. On the one hand, the model has a fairly low resolution and will not be able to capture local sources that might affect the stations. On the other hand, it also has a relatively short time-step compared to the weekly or even bi-weekly station measurements, which is why a daily cycle may be caught by the model but not by the stations. Therefore, any station where the model shows a large diurnal cycle is excluded. The criterion is a mean daily standard deviation of more than 3.5 ppb, following the example of Hooghiemstra et al. [2012a]. However, background stations and those affected by seasonal biomass burning signals are kept; in other words, large annual standard deviations are allowed. Using only background stations comes with the implied assumption that air masses reaching them are well-mixed and, therefore, even the coarse resolution of the model ($6° \times 4°$) is sufficient to capture the remaining spatial and temporal variation, allowing for a proper direct comparison of the model to the point observations. To account for any discrepancies from this assumption, the model estimates a representation error for each station based on the slopes (slope scheme introduced in Section 2.2) in the box that contains the station.

For the station data, in addition to the representation error of the model, a sampling error of 2 ppb is assumed. This error is composed of the instrument precision of 1.5 ppb given in Gerbig et al. [1999] for the fast-response vacuum-UV resonance fluorescence CO (VURF) instrument used at all stations in 2018 and the reproducibility of the measurements of 0.5 ppb provided in the readme of the dataset [Petron et al., 2020].

3.3.2   Satellite observations

The second assimilated dataset consists of the CO total columns from the TROPOspheric Monitoring Instrument (TROPOMI) on-board Sentinel-5 Precursor (S5P) satellite launched in October 2017 [Veefkind et al., 2012]. TROPOMI provides daily global coverage with a local overpass time at 13:30. The retrieved CO columns also feature a high spatial resolution of up to $7 \times 7 \, \text{km}^2$ at a swath width of 2600 km. Compared to that resolution, even the finest resolution of the model of $1° \times 1°$ might seem very

coarse. However, using high resolution observations not only implies a reduced aggregated observational error if multiple observations are available in a single model grid box, but it also gives a chance of at least some cloud-free pixels, i.e. some information, in cloudy model grid boxes.

This thesis employs the TROPOMI/WFMD version 1.2 product from the Carbon and Greenhouse Gas Group at the Institute of Environmental Physics (IUP) of the University of Bremen, retrieved with the Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm, which is described and validated in SCHNEISING et al. [2019]. In Chapter 5, the implications of using the latest product version, version 1.8, are investigated. The retrieval makes use of the TROPOMI observations in the shortwave infrared (SWIR) 2.3 µm spectral range to provide column-averaged dry-air mole fractions of methane and CO. The resulting total columns feature nearly constant sensitivity with respect to altitude. Notably, this includes the troposphere and boundary layer, which is especially useful when investigating biomass burning events and tropospheric air quality. In addition, observations in the SWIR spectral range, unlike those based on visible light, are capable of seeing through smoke plumes to some degree, making them critically valuable for investigating biomass burning events. The latter works for smoke but not clouds due to vastly different particle sizes, as demonstrated in SCHNEISING et al. [2020].

As detailed in SCHNEISING et al. [2019], the retrieval employs a fairly strict quality filter, especially with regard to cloudiness, surface brightness, and solar zenith angle ($< 75°$). This selection implies a clear sky bias in the observations, resulting in an overestimation of photochemical conditions, as well as very sparse data over the oceans due to their low albedo. The latter can be seen in Figure 3.1, where over the oceans observations are only possible due to sun glint, which occurs almost exclusively in the center of the orbits (i.e. in a nadir viewing geometry), while the sun is at the zenith. This implies that the sparse observations over the oceans are mostly clustered together.

3.3.2.1 Gridding

Above, inversions with gridded satellite observations were referenced. These so-called super-observations were created following the approach outlined in MIYAZAKI et al. [2012]. The implementation of the gridding approach used in this thesis is available at NÜSS et al. [2022b]. As shown in Figure 3.2, for each orbit, the intersection areas w_i of the footprint of each observation $\hat{y}_i^o$ with the cells of a regular $0.5° \times 0.5°$ grid is calculated. This grid resolution was chosen based on sensitivity studies (unpublished data) conducted in the LAMOS group at the IUP of the University of Bremen, which have shown that at the coarse model resolutions used throughout this thesis, inversions based on observations gridded to $0.5° \times 0.5°$ lead to almost the same optimized emissions as those based on the full satellite data, but with a significantly reduced
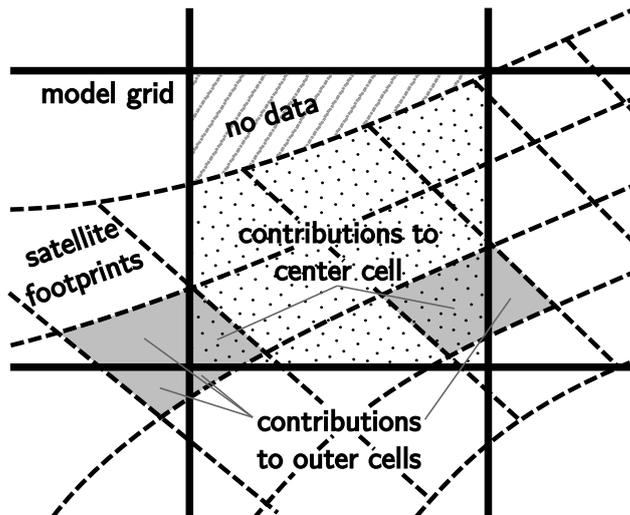
Figure 3.2: Schematic representation of several satellite footprints (outlined with dashed lines) intersecting with cells of a regular grid (thick, solid lines). The dotted areas show the portion w_i of each footprint that contributes to the center grid cell with area A_{cell}. For footprints that intersect with more than one grid cell (two examples highlighted in grey), their contributions are further deweighted based on the ratio between their respective intersecting area w_i (i.e. the part that is both dotted and grey) and their total area A_i (the entire grey area). For the striped area no observations are available, hence, the coverage α for the center cell is < 1.

computational cost (using full satellite data entails roughly $25\,\%$ longer computation times per iteration).

According to MIYAZAKI et al. [2012], a representative super-observation for each orbit and grid cell can be calculated as an area-weighted average:

$$\hat{y}_{\text{o}} = \frac{\sum_{i=1}^{m} w_i \hat{y}_i^{\text{o}}}{\sum_{i=1}^{m} w_i} \tag{3.1}$$

where m observations contribute to this super-observation. Notably, this average is not weighted by the retrieval error, which stems from the nature of the retrieval, where larger values have larger (absolute) errors, and, therefore, an error-weighted average would be skewed towards low values, as explained in BOERSMA et al. [2016]. The same process of calculating area-weighted averages is also applied to the measurement time, the a priori profile, the pressure levels of the retrieval, and the averaging kernel, level-wise for the latter three.

Unlike MIYAZAKI et al. [2012], before calculating the super-observation error as an area-weighted average, in the approach introduced here, the error corresponding to each individual intersection w_i is first inflated, so that its weight in the cost function (Eq. (2.11)) does not depend on the number of grid cells the corresponding footprint intersects with. This independence can be achieved with a factor $\sqrt{\frac{A_i}{w_i}}$, where A_i is the total area of the satellite pixel's footprint, which contains the i-th intersection. The

area A_i is equal to w_i if the footprint intersects exactly one grid box. Otherwise it will be larger, as exemplified in Figure 3.2, where the areas A_i, highlighted in grey, are larger than the areas w_i that are simultaneously grey and dotted for the two example footprints. The root stems from the least-squares nature of the cost function, while the rest is simply the inverse of the fraction of the footprint that intersects with the current grid cell. Taken together this yields an area-weighted error:

$$\sigma = \frac{\sum_{i=1}^{m} \sqrt{\frac{A_i}{w_i}} w_i \sigma_i^{\mathrm{o}}}{\sum_{i=1}^{m} w_i} = \frac{\sum_{i=1}^{m} \sqrt{A_i w_i} \sigma_i^{\mathrm{o}}}{\sum_{i=1}^{m} w_i} \tag{3.2}$$

Further following MIYAZAKI et al. [2012], this σ is then deflated by the number n of observations that contribute to the super-observation in that grid cell. However, this deflation is limited by the correlation c between errors of the individual observations (i.e. systematical errors from e.g. the albedo assumed in the retrieval are correlated in space and do not average out) as suggested in ESKES et al. [2003], and therefore, the super-observation error can be estimated as:

$$\sigma_{\mathrm{o}} = \sigma \sqrt{\frac{1-c}{n} + c} \tag{3.3}$$

Exact values for c are difficult to obtain, however, an upper bound may be found by considering the ratio of the systematic error of the TROPOMI observations versus its random error. From the validations against other observational datasets in SCHNEISING et al. [2019], this ratio can be estimated to be roughly 30 %. As not all systematic error sources from observations within each $0.5° \times 0.5°$ grid box are correlated, $c = 15\%$ is assumed here. It should be noted that the exact value of c has nearly no influence on the final inversion results, because a larger (smaller) c leads to overall larger (smaller) errors, which, for the most part, will be canceled out by a larger (smaller) error inflation (Section 3.3.2.2).

However, this σ_{o} does not yet include the representativeness error, which accounts for potential differences between the true average tracer concentration (which includes the parts of the cell that are not covered by observations) and the $\hat{y}_{\mathrm{o}}$ calculated above. For example, if the satellite observes a pristine background in one part of the grid cell, but there is also a plume with high tracer concentrations obscured by clouds in the remaining area, $\hat{y}_{\mathrm{o}}$ would be too low. The more of the grid cell area is covered, the smaller this representativeness error becomes.

MIYAZAKI et al. [2012] suggest a method to estimate this effect. First, the initial mean observation in a cell and the coverage $\alpha = \frac{\sum_{i=1}^{m} w_i}{A_{\mathrm{cell}}}, 0 \leq \alpha \leq 1$, where A_{cell} is the total area of the grid cell, are calculated. In Figure 3.2 the $\sum_{i=1}^{m} w_i$ is the total dotted area, whereas the A_{cell} is the total cell area enclosed by the thick, solid lines. Next, for
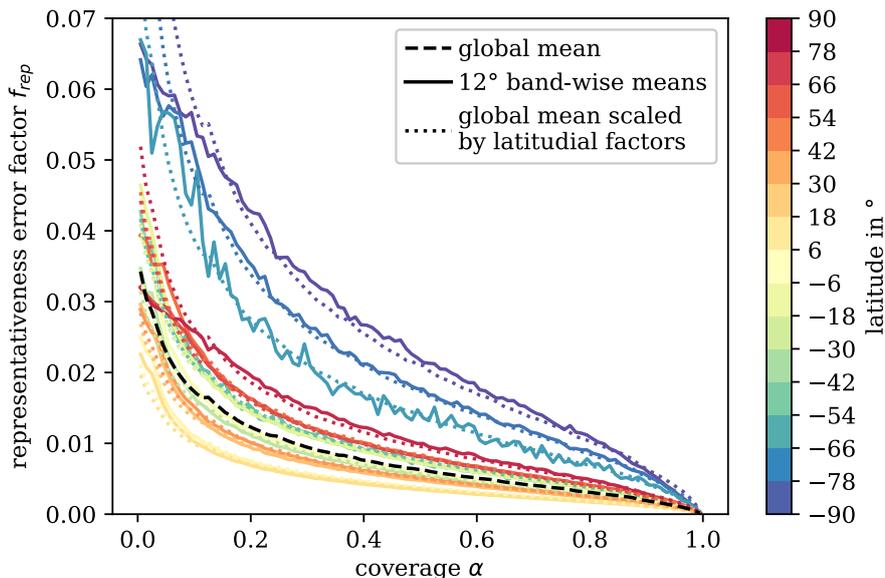
Figure 3.3: The dashed black line shows the global mean representativeness error factors over the satellite coverage in a given grid cell. This factor is zero for full coverage ($\alpha = 1$) and sharply increases at low coverage vales. The colored lines show the mean representativeness error factors over 12° bands. As these are quite noisy, they are instead used to obtain a single scaling factor for each band. These factors are then multiplied onto the global mean representativeness error factors, which leads to the much smoother colored dotted lines.

well covered grid cells ($\alpha > 90\,\%$ in MIYAZAKI et al. [2012]), the coverage α is artificially reduced by randomly removing observations. For each observation removed, the mean and coverage of the remaining observations are recalculated. The new mean is then compared to the original value to yield a relative deviation. By repeating this process for many grid cells, a mean relative deviation $f_{\mathrm{rep}}(\alpha)$ can be calculated. Multiplying this relative deviation with the super-observation value $\hat{y}_{\mathrm{o}}$ gives the representativeness error for that cell. In MIYAZAKI et al. [2012], the mean observations are calculated as a simple arithmetic mean, whereas here the area-weighted average, as introduced above, is used:

$$f_{\mathrm{rep}}(\alpha_k) = \left| \frac{\hat{y}_{\mathrm{o}} - \frac{\sum_{l=1}^{m-k} w_l \hat{y}_l^{\mathrm{o}}}{\sum_{l=1}^{m-k} w_l}}{\hat{y}_{\mathrm{o}}} \right|, 0 < k < m \qquad (3.4)$$

where k are the removed observations. For the sake of this analysis, the initial observations in each grid cell, i.e. before removing any of them, are treated as if they fully covered the cell. Therefore, $\alpha_k = \frac{\sum_{l=1}^{m-k} w_l}{\sum_{i=1}^{m} w_i}$ is the coverage compared to the initially covered area, rather than the full grid cell area.

Here, to estimate the representativeness error, 31 days of data, evenly spread over the available observations for 2018, are analyzed. Additionally, the coverage requirement is relaxed to $50\,\%$ to have a larger set of eligible observations, especially when considering coarser grids (not shown in this thesis). As α_k is a continuous variable, it
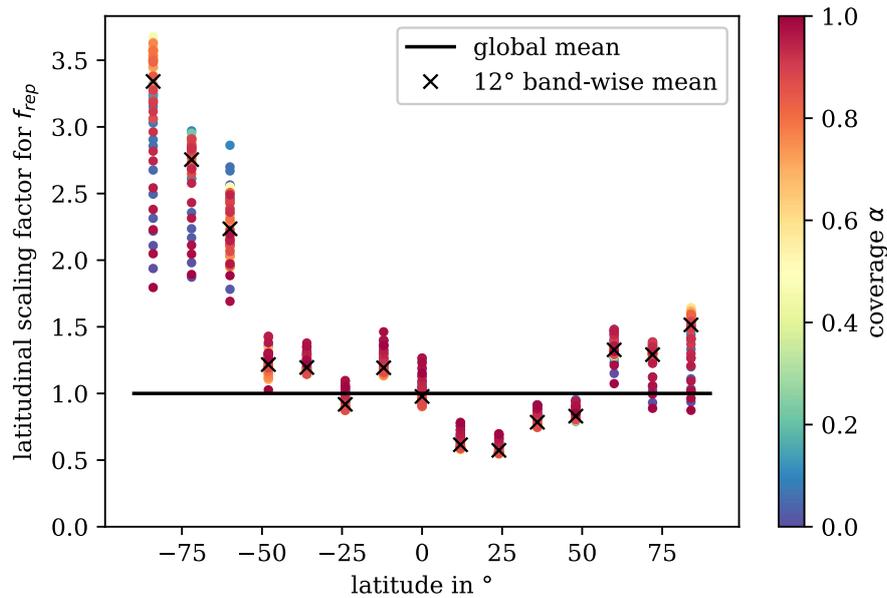
Figure 3.4: The black crosses are the 12° band-wise scaling factors for the global mean (black line) representativeness error factors, as shown in Figure 3.3. Clearly, representativeness errors rise towards the poles, especially in the southern hemisphere where there is less land-cover. Additionally, the band-wise scaling factors for each 1 % coverage bin, normalized over the respective global mean for that bin, are shown as colored dots.

had to be aggregated into 1 % bins for the sake of calculating the mean $f_{\mathrm{rep}}(\alpha)$ over the entire analyzed data. The resulting global mean representativeness error is shown as the black dashed line in Figure 3.3.

The TROPOMI satellite observations show a weak intra-annual variation in the representativeness error factor, with generally slightly larger error values in the northern hemispheric summer. However, its magnitude is smaller than the temporal variation on a daily basis. Therefore, the representativeness error is kept fixed in time.

In latitudinal direction, the analysis had to disregard the very few observations with a center point beyond 89.93° north/south, as these might touch and reach beyond the poles, which is problematic for area calculations in the used latitude-longitude projection. Additionally, as can be seen exemplified by the colored lines in Figure 3.3, there seems to be a strong latitudinal dependence of the representativeness error, with larger values towards the poles and in the southern hemisphere. This latitude dependence is likely caused by the poorer measurement quality over the oceans and in high latitudes, and smaller grid cell sizes towards the poles. Notably, while the magnitude of the representativeness error increases, the general dependence on the coverage α does not change. To capture this behavior, the representativeness error factor was additionally averaged over α for each latitudinal 12° band to obtain another scaling factor $\bar{f}_{\mathrm{rep}}(\phi)$, with ϕ as latitude. In Figure 3.4, these band-wise factors are plotted before (colored dots) and after (black crosses) averaging over α, all normalized over the global mean.

3 EFFICACY OF HIGH-RESOLUTION SATELLITE OBSERVATIONS

With this, the total representativeness error factor used in this thesis is:

$$f_{\mathrm{rep}}(\alpha, \phi) = \bar{f}_{\mathrm{rep}}(\phi) \cdot f_{\mathrm{rep}}(\alpha) \tag{3.5}$$

The resulting latitude-wise representativeness error factors are shown as colored dotted lines in Figure 3.3. The representativeness error can now be obtained for a given mean observation $\hat{y}_{\mathrm{o}}$, coverage α and latitude ϕ as

$$\sigma_{\mathrm{r}} = f_{\mathrm{rep}}(\alpha, \phi) \cdot \hat{y}_{\mathrm{o}} \tag{3.6}$$

This leads to the total error of the super-observations

$$\sigma_{\mathrm{s}} = \sqrt{\sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{r}}^2} \tag{3.7}$$

The super-observations are always assumed to be located at the center of their corresponding cells. This might lead to a spatial bias, because observation within an arbitrary grid cell cannot generally be assumed to be evenly distributed.

3.3.2.2 Error inflation

The uncertainties provided for the individual satellite observations (for the *full satellite* inversion) and the total error of each of the super-observations (for the inversions that use gridded satellite observations) are inflated with a global factor that depends on the specific inversion setup. For each inversion, this inflation factor is chosen so that the satellite and station observations each make up roughly half of the total observational cost, as suggested in HOOGHIEMSTRA et al. [2012a]. The intent of this inflation factor is to capture the spatial correlation between the individual satellite footprints and to prevent them from suppressing the signal of the surface stations by their sheer number.

In previous studies, this inflation factor has only been roughly estimated. For example, an empirically chosen variance inflation of 2 was used in CHEVALLIER [2007] for Orbiting Carbon Observatory (OCO) CO_2 observations gridded to $3.75° \times 2.5°$, an inflation of 50 was used in HOOGHIEMSTRA et al. [2012a] for MOPITT V4 level 3 CO observations gridded to $1° \times 1°$, and an inflation of again 50 was used in both KROL et al. [2013] and NECHITA-BANDA et al. [2018] for IASI CO observations at their native sampling resolution of up to about $25 \times 25\,\mathrm{km}^2$. This section suggests a more rigorous approach to finding the inflation that fulfills the condition of having each dataset make up an equal part of the observational cost.

Finding the inflation factor at which this condition is fulfilled is in itself an iterative process, where each iteration is a complete inversion. A close look at the cost function (Eq. (2.11)) reveals that for an attempted inflation I, the inflation I' for the next

iteration can be calculated as

$$I' = \sqrt{\frac{J_{\mathrm{obs,sat}}}{J_{\mathrm{obs}} - J_{\mathrm{obs,sat}}} \cdot I^2} \qquad (3.8)$$

where J_{obs} is the total observational cost of the attempt, $J_{\mathrm{obs,sat}}$ is the part of J_{obs} contributed by the satellite observations, and the inflation factors I, I' are a factor applied to the observational errors (standard deviations). It should be noted, however, that Eq. (3.8) will always underestimate the change in inflation needed. For example, if the initial inflation was too large, the formula will suggest an improved, but still slightly too large inflation for the next iteration. This happens, because reducing the inflation will increase the cost attributed to the satellite observations, which in turn causes the inversion to improve their fit. However, a closer fit to the satellite observations usually implies degradation of the fit to the flask observations, which will increase their contribution to the cost function. That way, the total cost increases and a slightly smaller inflation is needed so that the contribution of the satellite observations makes up half of that cost. In the opposite case, if the inflation was too small, the next guess will be better but still slightly too small.

It may seem that the inflation is solely a parameter of the observational datasets involved and, therefore, fixed for a given set of observations. However, experience has shown that the inflation also depends on the time of year, the error and temporal resolution of the a priori emissions, and the a priori datasets used. Both, a larger a priori error or a higher temporal resolution of the emissions, especially for the biomass burning emissions, enable the model to fit the satellite observations more easily (lower cost) without degrading the station fit, leading to lower required inflation factors to fulfill the criterion.

The setup outlined above results in significantly different inflation factors for the individual inversions. Inflation factors are generally larger for the main inversions compared to the *spin-up* inversion (45). They are also slightly larger for both of the FINN2.4 based inversions (72 for *reference* and 70 for *noVIIRS*) than for the *GFED* inversion (68), possibly due to better a priori fits at the stations. These (standard deviation) inflation values are larger than the aforementioned variance inflation factors used in HOOGHIEMSTRA et al. [2012a] for gridded MOPITT observation, and in KROL et al. [2013] and NECHITA-BANDA et al. [2018] for full resolution IASI observations. The larger values are expected, because of the higher grid resolution when compared to MOPITT, and the better coverage of TROPOMI when compared to IASI. Due to the much larger number of observations, the largest inflation is required for the *full satellite* inversion (182). This number is an indication of the higher spatial correlation within the individual observations compared to within the gridded observations, since the latter are, by definition, further apart.

The concentrations at the locations of the surface stations depend only relatively weakly on the exact value of the inflation factor, because the well-mixed background concentrations show much broader patterns, which are captured by either dataset to some extent. However, very small inflation factors will still cause the station fits to degrade heavily, because the satellite data will drown out the flasks. Conversely, for very large inflation factors the model approaches the *station only* inversion. This emphasizes the need for the inflation factor to properly weigh both datasets against one another.

However, some issues remain with the condition of having the observational cost equally distributed between the stations and the satellite observations. This condition implies that satellite observations with higher coverage or lower errors are assigned higher inflation values, i.e. higher quality data gets a lower weight in the cost function. Inadvertently, this will lead to overfitting of the surface flasks with increasing quality of the satellite instruments used. Additionally, while a somewhat larger inflation factor is to be expected at higher coverage due to increased correlation between the individual pixels, the current blanket approach of assigning a constant inflation factor to all footprints ignores the actual density and correlation of the observations. This implies that dense observations over the Sahara are inflated just as much as the sparse observations over the oceans. In Chapter 5, an alternative inflation strategy is introduced, that potentially resolves some of these issues.

3.4 Results

3.4.1 Mixing ratio mismatch at the surface stations

3.4.1.1 Set 1: Inversions using different biomass burning priors

In Figure 3.5, the modeled mixing ratios at 6 out of the 44 total ground-level stations are shown before and after the inversions from the first set of experiments (*reference*, *noVIIRS*, and *GFED*), where the biomass burning inventories were varied. Additionally, the corresponding flask measurement values as well as their assigned uncertainties are indicated. During the *spin-up* inversion (not pictured), many stations initially exhibit considerable under- or overestimations. The model corrects most of these within the first one or two months and the mixing ratios at the stations start to closely follow the observations. This way, during the main inversions (e.g. as shown in Figure 3.5), the modeled mixing ratios at all stations are initially close to the observations. At most stations, the mixing ratios simulated based on the optimized emissions remain close to the observations over the whole period of the main inversion. This can be seen for example at Mauna Loa (Figure 3.5d) and Rapa Nui (Figure 3.5f) in the northern and
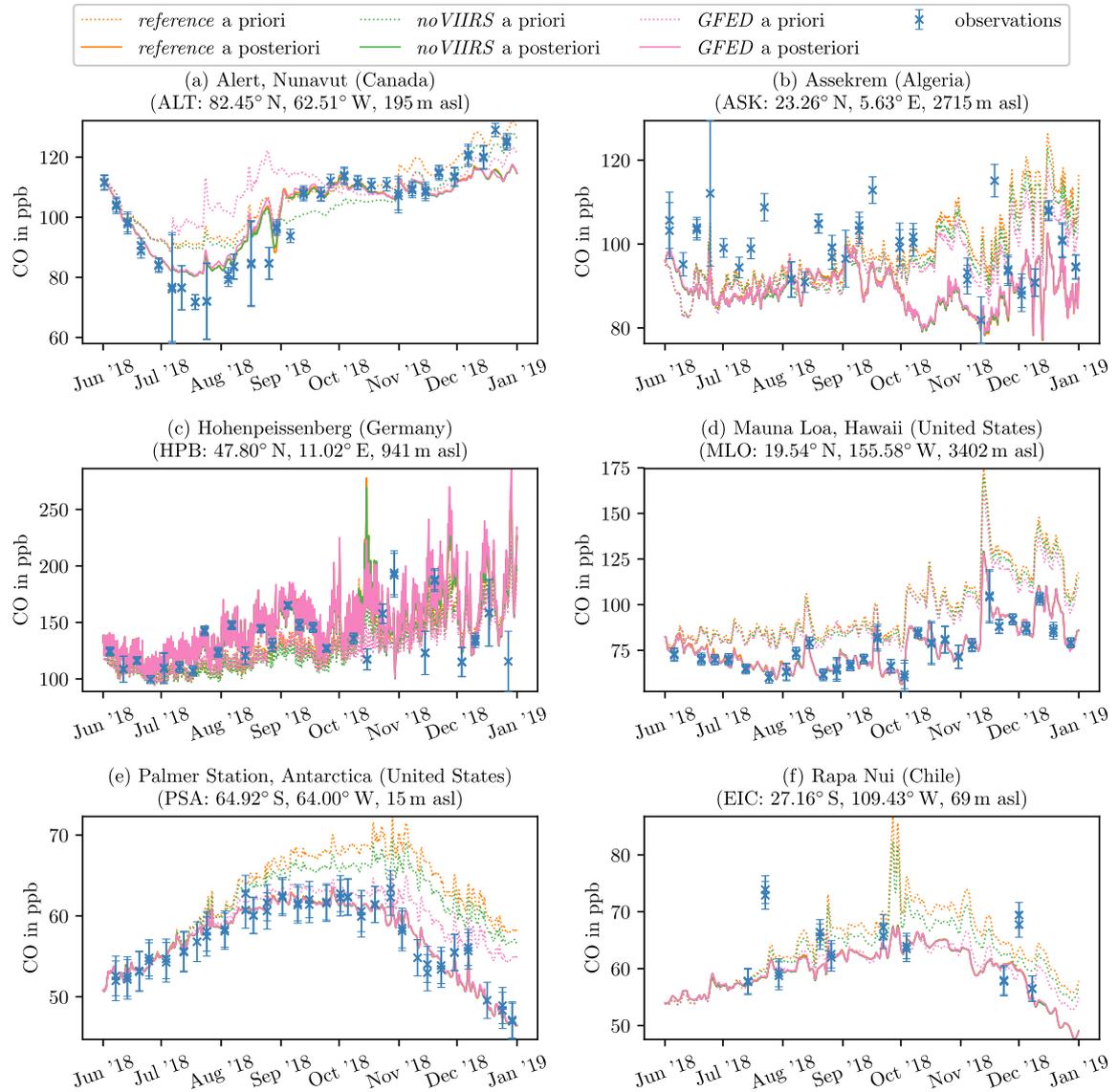
Figure 3.5: Modeled a priori (dotted lines) and a posteriori (solid lines) mixing ratios sampled at the locations of the stations as well as the flask observations (blue crosses) for six example stations and the three different biomass burning a priori inventories. For each observation, the corresponding measurement error is indicated as well. Lines are color-coded based on the a priori used: FINN2.4+VIIRS (*reference*) in orange, FINN2.4 (*noVIIRS*) in green and GFED4.1s (*GFED*) in pink. Unlike the first four, the bottom two stations ((e) PSA and (f) EIC) are in the southern hemisphere and, therefore, in the low resolution global region.

southern Pacific, respectively, but also at stations close to the South Pole, like Palmer Station in Figure 3.5e, despite their very remote nature.

However, at a few stations, the posterior mixing ratios diverge from the measurements to some degree. This effect is mostly limited to high ($> 55°$ N) northern latitudes. For example at Alert, as shown in Figure 3.5a, mixing ratios in July and August do not drop far enough, while towards the end of the year they do not rise high enough. Another problematic station is Assekrem, plotted in Figure 3.5b, where the flask observations are systematically underestimated by the model.

Generally, the a priori mixing ratios feature a global accumulation of ground-level CO over time not supported by the observations. This indicates an unbalanced budget, with either too large sources (overestimations in the a priori), or a too small sink (underestimations in the OH climatology). Given the setup of the inversions, the model resolves this by reducing the emissions in either case. However, there are stations where this does not hold and the a priori underestimates the observations. For example at Hohenpeissenberg in Figure 3.5c, the model finds a fairly strong diurnal cycle and generally too low a priori mixing ratios. The former is likely a result of the station being located at the top of a mountain, where upslope conditions cause surface CO to be transported up to the station during daytime and away during night. Even though not clearly visible in Figure 3.5c, where the full time series is shown, the model is only sampled at the time of the measurement, which would alleviate this issue to some degree. The too low a priori mixing ratios, however, could point to the relative proximity of the station to emission sources in Central Europe, and possibly indicate that the lateral model resolution is not fine enough to properly capture this station.

In the first eight rows of Table 3.2, the mean error-weighted mismatch $\bar{J}_{\text{flask}}$ between flasks and model for all main inversions is calculated as

$$\bar{J}_{\text{flask}}(\vec{x}) = \frac{\sum_{i=1}^{N_{\text{flask}}} \left[\frac{\left(y_{\text{flask},i} - \mathbf{F}(\vec{x})_i \right)^2}{\varepsilon_{\text{O},i}^2} \right]}{N_{\text{flask}}} \tag{3.9}$$

where N_{flask} is the total number of flask measurements y_{flask} with observational errors $\varepsilon_{\text{O},i}$, and $\mathbf{F}(\vec{x})_i$ is the model sampled at that measurement. The observational errors include the representation error of the model and the sampling error of the flasks. If the model is capable of capturing the variability of the observations, the unit-less quantity $\bar{J}_{\text{flask}}$ should be close to one. Larger values could point to an underestimated observational error, systematic errors in the model itself or a model with too few degrees of freedom to capture the variability in the observations, i.e. an underestimated model representation error. When comparing two inversions, lower values represent a better fit. As can be seen for all three experiments of the first set (*reference*, *noVIIRS*, and *GFED*), the fit after the inversion is vastly improved compared to the prior fit. Con-

Table 3.2: Error-weighted mismatches between observations and model for each main inversion. The first eight rows give the mean mismatches to different subsets of the flask measurements. There, even in the *satellite only* inversion, where the flasks did not constrain the emissions, the overall fit at the stations improves, although less compared to the other experiments. The mismatch for the *satellite only* inversion decreases significantly if only stations south of 55° N are considered (i.e. excluding ALT, BRW, CBA, ICE, PAL, SUM, TIK, and ZEP), while it stays roughly the same for all other experiments. A considerable portion of the remaining mismatch stems from the stations ASK, HPB, and OXK, where the model generally has problems capturing the modeled variation. The last two rows contain the total mismatch to the satellite observations, scaled down by 10^3 for readability. Similarly to the *satellite only* inversion above, even in the *station only* inversion, the overall fit to TROPOMI improves, despite those observations not constraining the inversion.

observations			reference	noVIIRS	GFED	satellite only	station only	full satellite
	all	prior	21.91	18.89	15.88	21.91	21.91	21.91
		posterior	3.61	3.83	3.99	9.12	3.26	3.50
stations	< 55° N	prior	24.33	20.98	16.52	24.33	24.33	24.33
		posterior	3.59	3.79	3.98	7.84	3.35	3.51
	excluding ASK, HPB, OXK	prior	22.21	18.72	15.63	22.21	22.21	22.21
		posterior	3.36	3.57	3.66	7.62	3.14	3.28
	< 55° N and excl. ASK, HPB, OXK	prior	24.92	20.94	16.26	24.92	24.92	24.92
		posterior	3.27	3.45	3.55	5.80	3.20	3.23
satellite		prior	88.08	75.00	62.22	88.08	88.08	72.05
		posterior	7.92	8.46	8.88	6.93	19.58	7.74

45

sidering how well the model captures the variability at most stations (e.g. Figure 3.5), the a posteriori $\bar{J}_{\text{flask}}$ values of 3 to 4 most likely indicate underestimated errors, rather than systematic model errors. Table A.1 in the appendix provides the individual mean error-weighted a priori and a posteriori mismatches for all 44 stations across all six main inversions. The same information is also plotted in Figure A.4, ordered by the latitude of the station.

For most stations, the choice of the biomass burning a priori has very little influence on the final fit, as evident from the orange, green, and pink lines in Figure 3.5 coinciding almost everywhere. Moreover, the a priori mixing ratios from the different inventories themselves are fairly similar. In general, a priori mixing ratios are lowest before the *GFED* inversion and highest before the *reference* inversion based on FINN2.4+VIIRS, though this does not allow for any conclusions regarding the quality of the inventories. With all three, the a priori mixing ratios are clearly overestimated. While GFED4.1s generates the lowest a priori mixing ratios which are, therefore, closest to the observations ($\bar{J}_{\text{flask}} = 15.88$ is the smallest prior mismatch out of all experiments), this could be coincidental.

3.4.1.2   Set 2: Inversions based on different observational datasets

For the same stations as in Figure 3.5, the modeled mixing ratios for the second set of experiments (*satellite only*, *station only*, and *full satellite*) based on different observational input datasets are shown in Figure 3.6. At the resolution of the model employed in this thesis, even within the zooming region (up to $3° \times 2°$), only minor differences in a posteriori mixing ratios are found between the *full satellite* inversion (green lines) versus the *reference* inversion (orange lines), i.e. for the sake of this analysis, those datasets are equivalent. This equivalence is also emphasized by very similar mismatch values in Table 3.2. In the *station only* inversion, where the satellite observations are excluded altogether (brown lines), the fit to the flask measurements gets slightly better (lowest $\bar{J}_{\text{flask}}$ in Table 3.2), though changes are mostly minimal. Larger changes are found when comparing the former three inversions to the *satellite only* inversion (pink lines), in which the model is not driven by the flasks at all. In Table 3.2, this leads to a significantly larger $\bar{J}_{\text{flask}}$, compared to all the other experiments, yet the mismatch is still lower than for the a priori. This shows that the error inflation factors introduced in Section 3.3.2.2 have been chosen to meaningful values, because the station fits do not significantly degrade due to the satellite observations in the combined inversions.

Stations at high ($> 55°$) northern latitudes, like Alert in Figure 3.6a, exhibit a poor fit quality for the *satellite only* inversion. During northern hemispheric summer, mixing ratios stay close to the a priori and much higher than the flasks, while in northern hemispheric winter they fall too low, diverging from the a priori and the flasks. This implies that these stations systematically have large mismatches. To illustrate that
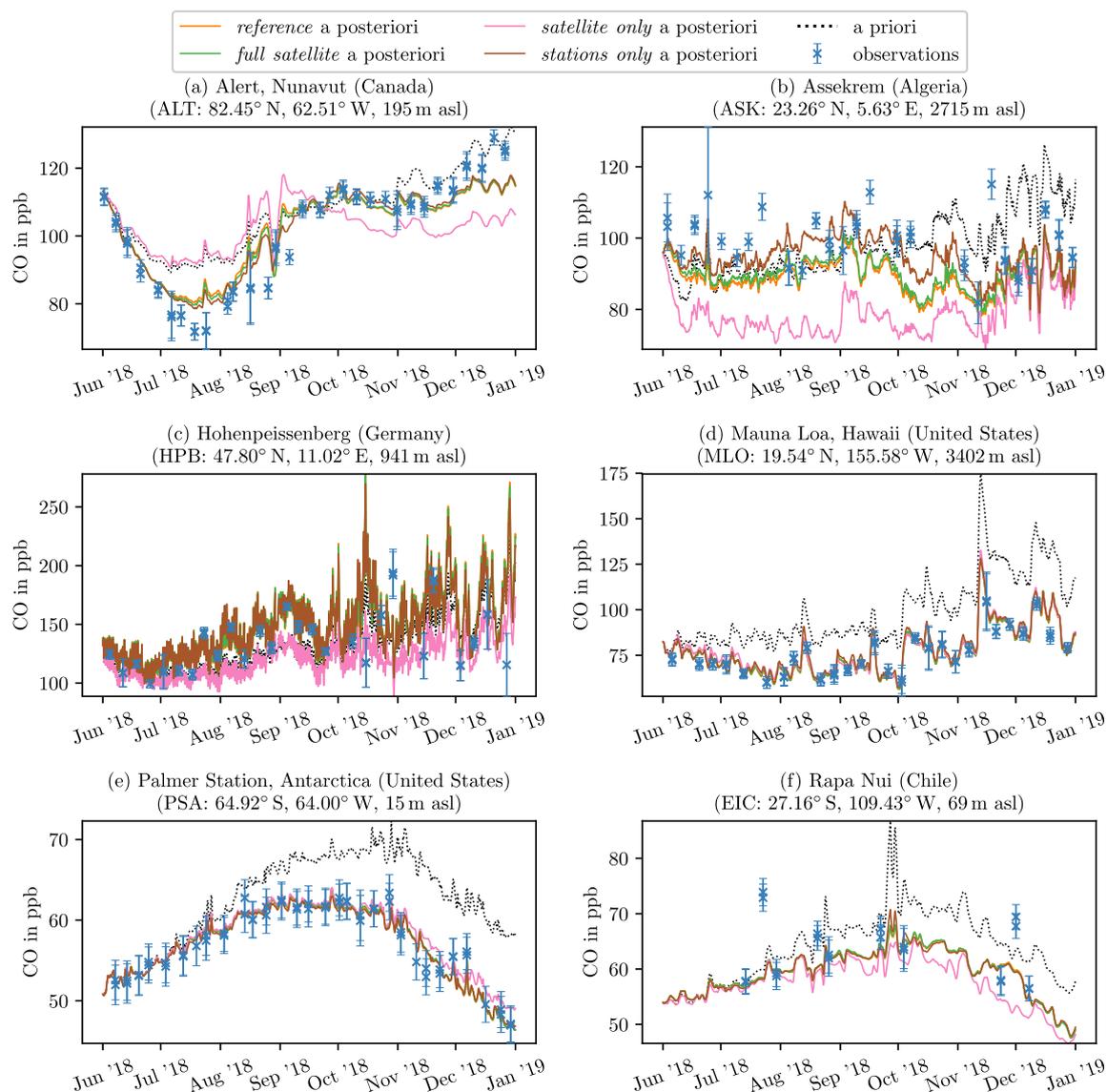
Figure 3.6: Modeled a priori (dotted line) and a posteriori (solid lines) mixing ratios sampled at the locations of the stations as well as the flask observations (blue crosses) for six example stations and four inversions with different observational datasets. For each observation, the corresponding measurement error is indicated as well. Lines are color-coded based on the observations used: The orange lines represent the *reference* inversion and are identical to the orange lines from Figure 3.5. In green the *full satellite* inversion is shown, which also uses a combination of satellite and flask observations. The pink and brown lines represent the *satellite only* and *station only* inversions, respectively. Note that because all inversions are based on the same a priori emissions, the single dotted black line holds for all four inversions.

the fit at other stations is better, $\bar{J}_{\text{flask}}$ is calculated only for stations south of $55°\,\text{N}$ in the third and fourth row of Table 3.2. While $\bar{J}_{\text{flask}}$ is significantly reduced for the *satellite only* inversion, it stays almost constant for all other experiments. This implies that the satellite observations specifically are insufficient to constrain these stations at high northern latitudes, while the model itself is well capable of capturing them. In the *satellite only* inversion, during northern hemispheric wintertime, there are very few observations in this region, due to little light and high cloud coverage. Therefore, the divergence from the a priori is likely driven by the unbalanced budget in the northern tropical and subtropical regions, where emissions all year round are heavily reduced as shown in Section 3.4.3 below. It is cheaper for the model, in terms of the cost function, to diffuse the decrements over a larger area and shift part of them to higher northern latitudes, than to have even deeper localized decrements in the tropics.

Aside from the northern stations, there are a few other stations that are problematic for the model to capture. The most extreme example of these issues is the station in the Assekrem (ASK) shown in Figure 3.6b, where the satellite drives the model to much lower mixing ratios than the flasks. This underestimation can be clearly seen by the very low a posteriori mixing ratios for the *satellite only* inversion (pink line), and by the *reference* inversion (orange line) ending up consistently lower than the *station only* inversion (brown line), which is seldom the case for other stations. For this specific station, this effect is likely amplified by its positioning within the Sahara desert, where satellite observations are plentiful due to high albedo and little cloud cover, but might also be adversely affected by dust. This oversampling causes the satellite observations to gain a relatively large weight in the cost function compared to the flasks at that location, causing the *reference* inversion to slightly diverge from the flask observations. Assekrem is also a high-altitude site, which could potentially be problematic with the limited representation of topography in the model. When considering the resulting emission increments (Section 3.4.3) it appears that the model is not capable of capturing this station properly. Another problematic station is Hohenpeissenberg (HPB), shown in Figure 3.6c, where the *satellite only* inversion, again, suggests much lower mixing ratios. Note the larger range on the vertical axis. Similar, albeit less pronounced results are found for Ochsenkopf station (OXK), which is relatively close to Hohenpeissenberg station geographically. Both are located on mountains at high altitudes. Therefore, as mentioned earlier, the coarse resolution of the model and its limited representation of topography might adversely affect the results there. This misrepresentation will also be further discussed in Section 3.4.3 below, where these specific stations are found to lead to unrealistically high emission increments, similar to Assekrem station. As for the stations at high northern latitudes, these three stations (ASK, HPB, and OXK) degrade the global mean error-weighted mismatch exceptionally strongly. To illustrate this, in the fifth and sixth row of Table 3.2 $\bar{J}_{\text{flask}}$ is calculated for all but these stations.

Again, $\bar{J}_{\text{flask}}$ for the *satellite only* is reduced strongly. However, there are also slight decreases for the other experiments, suggesting that the model overall has an issue with properly representing these stations.

Nonetheless, most other stations, regardless of geographical location, show good fits for all four investigated combinations of observational input. As examples for northern tropics, high southern latitudes, and southern tropics, Mauna Loa, Palmer Station, and Rapa Nui, respectively, are shown in Figures 3.6d–f. Most notably, the *satellite only* inversion manages to closely follow the flask measurements, despite them being not assimilated. This can be seen in the seventh and eighth row of Table 3.2, where both, the stations north of 55° N and the problematic stations (ASK, HPB, OXK) are excluded from the calculation and $\bar{J}_{\text{flask}}$ for the *satellite only* inversion gets much closer to the other experiments. These good fits suggest that inversions of current events driven solely by TROPOMI observations are feasible, as long as the region of interest is well south of around 55° N.

3.4.2   Mixing ratio mismatch to the satellite observations

In the final two rows of Table 3.2, the total error-weighted mismatch J_{sat} between satellite observations and model for all main inversions is calculated as

$$J_{\text{sat}}(\vec{x}) = \sum_i \left[\frac{\left(y_{\text{sat},i} - \mathbf{F}(\vec{x})_i\right)^2}{\varepsilon_{\text{O},i}^2} \right] \tag{3.10}$$

where $y_{\text{sat},i}$ are the satellite observations with observational errors $\varepsilon_{\text{O},i}$, and $\mathbf{F}(\vec{x})_i$ is the model sampled at that measurement, with the averaging kernel applied. Figure A.5 in the appendix shows the temporal (monthly) and spatial ($12° \times 12°$ grid) distribution of the total error-weighted mismatches for all main inversions. Unlike for the mean error-weighted mismatch $\bar{J}_{\text{flask}}$ between the flasks and the model introduced in the previous section, there is no division by the number of observations here, hence the total instead of the mean mismatch is calculated. Considering the total mismatch was necessary, because the number of observations in the *full satellite* inversion is much larger than in all other inversions that use the gridded super-observations. Therefore, the mean error-weighted mismatch for the non-gridded observations is much smaller, i.e. each single observation bears a smaller weight in the inversion. By design, the super-observations have smaller error than each single observation they are made up from (Section 3.3.2.1) and the error of satellite observations in the *full satellite* inversion is inflated the strongest (Section 3.3.2.2). Overall, the total mismatch leads to comparable numbers, in this case, while the mean mismatch would not. Again, as for the stations in the previous sections, more detailed data can be found in the appendix, where Figures A.6 and A.7 show the latitudinal distribution of the mean a priori and a

posteriori mismatch between the model and the satellite observation in 12° bands for all six main inversions.

Generally, the results are similar to the ones for the stations above. When considering the first set of inversions (*reference*, *noVIIRS*, *GFED*), the a priori mismatch is again smallest for *GFED* and largest for *reference*, and for the a posteriori mismatch this is inverted again. For the second set, the *satellite only* inversion results in the best fit to the satellite observations, while the *station only* inversion results in the worst. This is akin to the results from the previous section, where the *station only* inversion had the best fit to the station data and the *satellite only* inversion had the worst fit. As outlined above, the mismatch for the *full satellite* inversion is special, because it is calculated with respect to the non-gridded dataset. Regardless, the mismatch reduction is comparable to the *reference* inversion.

The mismatches mainly originate from regions known for biomass burning, such as central and southern Africa, northern South America, eastern North America, Indonesia, and Siberia. Even the $0.5° \times 0.5°$ grid of the super-observations is fine compared to the model resolution of $3° \times 2°$ or $6° \times 4°$. Therefore, any biomass burning event that leads to steep gradients in the observations cannot be resolved in the model and will lead to mismatches between the modeled and observed mixing ratios.

The global a posteriori mismatches also vary in time and are largest in August during the height of the burning season. More details on this can be found in the appendix in Figures A.8 and A.9, which show the global total prior and posterior mismatch between the satellite observations and the model for each month of each of the main inversions. This spike in August is especially pronounced in the *station only* inversion, where the mismatches already rise in July and slowly taper off over the following months. For this inversion, in addition to the coarse model resolution, the station measurements are too sparse in time and space to properly capture individual biomass burning events and only constrain the increases in the resulting well-mixed background mixing ratios. Similar as for the stations, the a priori mismatches are initially low in June and steeply rise over the following three months. The good initial fit shows that the *spin-up* inversion manages to properly harmonize the modeled mixing ratio with the observations, as intended. The following rise in mismatches also illustrates the suspected unbalanced budget that causes CO to accumulate in the model.

Figure A.10 in the appendix provides a closer look at the monthly lateral distribution of the total a posteriori mismatch between the satellite observations and the model for each inversion compared to the *reference* inversion, i.e. when and where each inversion preformed better or worse than the *reference* inversion. For the first set of inversions, it becomes apparent that, while the *GFED* inversion leads to worse mismatches overall, the mismatches in Indonesia are slightly smaller compared to the *reference* inversion. Additionally, *noVIIRS* and *GFED* perform slightly better than

reference in central Africa in the beginning of the burning season in August to October, but the *reference* inversion performs better there for the rest of the year.

Further analysis of the second set shows that for the *satellite only* inversion the lower mismatch originates mostly from the northern hemisphere. Curiously, the mismatch towards the satellite observations around Rapa Nui in the southern Pacific is significantly increased (roughly +50 %) in the *satellite only* inversion for the period October to December compared to the *reference* inversion, i.e. in that region, the additional use of flask measurements in the *reference* inversion leads to a better fit to the satellite observations than using the latter on their own. This apparent contradiction can be resolved by considering that the mixing ratios at such remote locations are, on the one hand, only weakly constrained by the sparse satellite observations over the oceans and, on the other hand, are strongly influenced by transport from distant, land-bound source regions [DASKALAKIS et al., 2022], which are much stronger constrained by the satellite observations. The addition of the high-confidence flask measurements from the Rapa Nui station causes the model to diverge from the a priori towards higher emissions around that station, which also better fit the (sparse) satellite observations in that region.

For the *station only* inversion, especially large mismatches are observed over northern Africa during the full inversion period. This is most likely related to the issues with the station in the Assekrem outlined in the previous section. During the burning season (July–September) the mismatches in the *station only* inversion are most pronounced over continental Asia, northern and central Africa, northern South America, eastern North America, and the oceans in between those regions. Towards the end of the year, large mismatches are also found around Indonesia. Notably, the *station only* inversion shows a degrading fit to the satellite observations in high northern latitudes ($> 55° \text{N}$), i.e. the a posteriori mismatch there is worse than the a priori mismatch (see also Figure A.6). This is the only place and time where a degrading fit occurs. As mentioned, all of this behavior is to be expected from the *station only* inversion, since the sparse station network cannot capture the full spatial and temporal variation of all biomass burning events globally.

While the mismatches for the *full satellite* inversion are problematic to compare directly to the other inversions due to the much larger number of observations and the error inflation, the mismatches appear to be smaller in remote regions and larger in active biomass burning regions, compared to the *reference* inversion. This mismatch distribution is expected, because the higher resolution of the full satellite observations implies finer and more pronounced structures from the individual biomass burning events, which the model can resolve even less.

Interestingly, the mismatches from all main inversions converge in the southern hemisphere, i.e. even the *station only* inversion fits the satellite observation just as

good as the *reference* or even the *satellite only* inversion. This shows that not only is each dataset on its own sufficient to constrain the (remote) southern hemisphere, but they also end up at roughly the same result there.

3.4.3   Optimized global emission fields

3.4.3.1   Secondary production
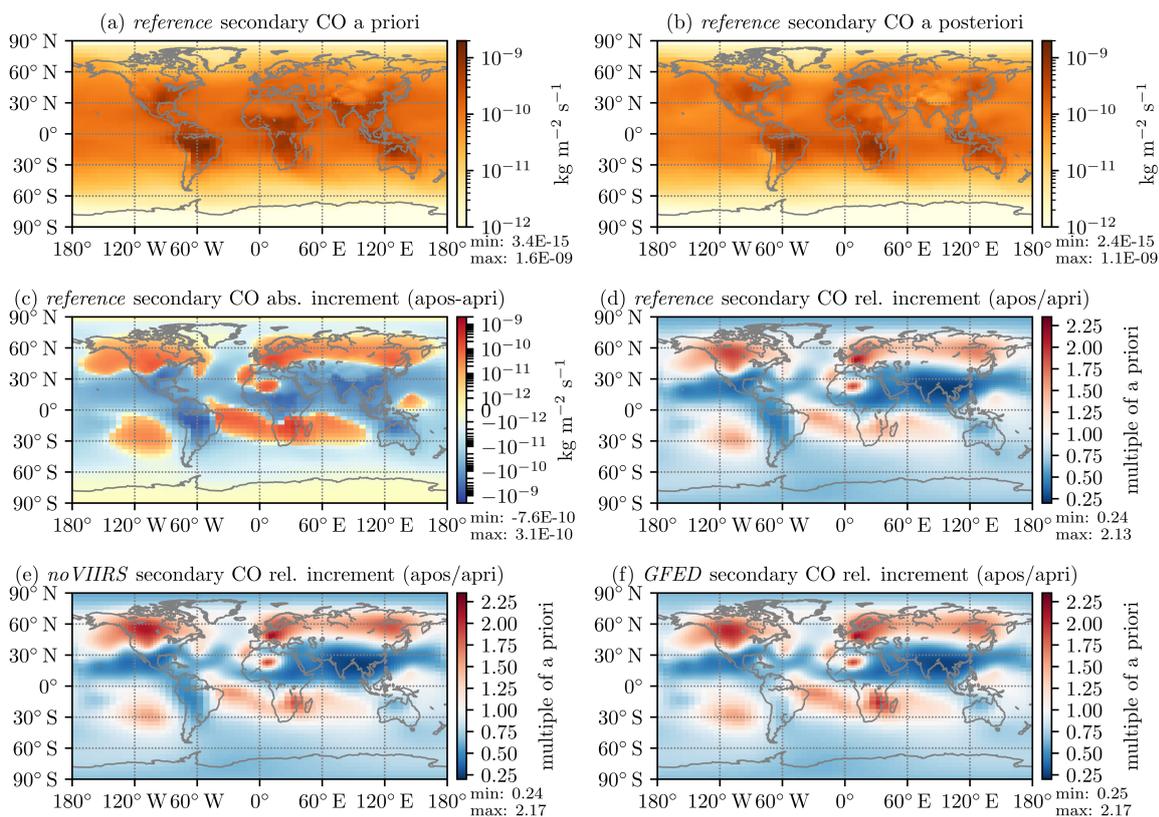
Figure 3.7: Global secondary CO production for September 2018 for the first set of experiments. The first four panels belong to the *reference* inversion (based on FINN2.4+VIIRS) and show (a) the a priori emissions, (b) the a posteriori emissions, (c) their absolute difference, and (d) the factor by which the emissions increased. Panels (e) and (f) show this factor for the *noVIIRS* and *GFED* inversions, respectively. Note the logarithmic color scales in the first three panels.

Figure 3.7 provides a global overview of the optimized secondary CO production from VOCs including CH_4 for September 2018 and a comparison to the a priori emissions for the *reference* inversion. In Panels (c) and (d), the absolute and relative differences between the a priori (Panel (a)) and a posteriori (Panel (b)) are shown. For comparison, the relative emission increments for the *noVIIRS* and *GFED* inversions can be found in Panels (e) and (f), respectively. September was arbitrarily chosen, because it is in the center of the inversion period and the results found for the other months are fairly similar. The differences that occur over time are small and limited to

variations in amplitude, but not in space. This is to be expected, considering the strict temporal correlation times and spatial correlation lengths introduced in Section 3.2.3.1. Figures A.11 to A.13 in the appendix provide a brief overview of the relative secondary CO increments resulting from the *reference* inversion for the remaining six months of the main inversion period and comparisons of those increments to the ones shown in Figure 3.7.

All main inversions result in large decrements in a band roughly between the Equator and 40° N. These decrements are especially deep over China and India, as can be seen in the relative increments in Figures 3.7d–f. In the later months of the inversion period, this region of large decrements shifts eastwards towards China for all experiments. This northern tropical decrement will be analyzed in more detail later on in Section 3.4.3.2, in the context of anthropogenic emission increments.

The band of decrements is accompanied by increased emissions north of 40° N, especially over central Europe, North America, and Siberia. Additional positive increments can be found between the Equator and 40° S, over the oceans, and in southern Africa. These appear to occur in biomass burning outflow regions, and could point to a systematic error in the lifetime of CO in the model. Due to the band-like structure of the positive and negative increments, this error is possibly caused by inaccurate OH values. Further evidence for such issues with OH values can be found in MYRIOKEFALITAKIS et al. [2020], where they compare their online calculated OH to the climatological fields from SPIVAKOVSKY et al. [2000] used here and find significant differences in those regions. Notably, in the full chemistry simulation, higher OH concentrations not only imply higher CO loss rates, but also higher secondary CO production. Here, those production rates are paired with loss rates base on the climatological OH, as pointed out in Section 3.2.3.3. Since in the inversion experiments throughout this thesis, the loss rates are fixed, the model can only compensate for this mismatch by, in some places considerably, changing the secondary CO source.

Overall, the a posteriori secondary CO source is lower than the a priori production flux in all experiments, as can be seen in the global budgets provided in Table 3.3, where the posterior masses at the end of the inversion period (final masses) are consistently lower than the prior final masses. All fluxes have been extrapolated to annual budget terms in Tg CO yr^{-1}, which might be misleading, because the inversion period of the main inversions includes the biomass burning season, but excludes the increased anthropogenic emissions due to heating during most of the northern hemispheric winter. Regardless, the extrapolated annual a posteriori budget terms found here are much closer to the ones found in literature [e.g. ZHENG et al., 2019] than the a priori terms, implying that the a posteriori terms are more realistic. However, as expected, the partitioning of there emission terms here is slightly different compared to ZHENG et al.

Table 3.3: Global prior and posterior budgets for each inversion, as a sum over the global and the zooming regions. The zooming column combines masses going into and coming from the communication cells between the zooming regions. For the main inversions, the 3° × 2° region perceives this as a net loss through advection into these cells, while the global region perceives it as a net gain through emissions within the cells. Only the net effect is shown here. Note that the annual rates (Tg CO yr⁻¹) are extrapolated from the emissions during the respective inversion periods, January to June (6 months) for the spin-up inversion and June to December (7 months) for the main inversions.

experiment		masses in Tg CO		losses in Tg CO yr⁻¹		zooming in Tg CO yr⁻¹	total	emitted in Tg CO yr⁻¹		
		initial	final	chemical	deposition			secondary	biomass	fossil fuel
reference	prior	555	747	-3035	-220	118	3464	2179	666	618
	posterior		585	-2487	-187	22	2702	1616	572	514
noVIIRS	prior	555	728	-2938	-210	107	3336	2179	618	618
	posterior		585	-2486	-186	22	2701	1681	500	520
GFED	prior	555	699	-2816	-199	88	3172	2179	618	618
	posterior		585	-2481	-184	21	2694	1770	368	556
satellite only	prior	555	747	-3035	-220	118	3464	2179	666	618
	posterior		581	-2481	-185	17	2692	1608	574	510
station only	prior	555	747	-3035	-220	118	3464	2179	666	618
	posterior		595	-2584	-194	26	2819	1674	636	508
full satellite	prior	555	747	-3035	-220	118	3464	2179	666	618
	posterior		587	-2502	-188	23	2722	1627	580	514
spin-up	prior	646	677	-2955	-215	30	3202	1991	575	637
	posterior		521	-2351	-182	-78	2359	1326	400	632

[2019], with lower anthropogenic/fossil fuel CO emissions, but higher secondary CO production.

As for the stations in Section 3.4.1, the differences in the emission increments between the inversions in the first set (different biomass burning a priori) are rather small. The most striking differences are the much larger increments (up to $+70\,\%$ higher final emissions) over southern Africa in the *GFED* inversion (Figure 3.7f). These are likely related to a known underestimation of African CO emissions in GFED4.1s as described in NGUYEN & WOOSTER [2020] and references therein. More subtle differences are found in South America, where the *GFED* inversion only leads to minor corrections (relative increments close to 1), while the *reference* and *noVIIRS* inversions show clear decrements (final emissions reduced by up to $-50\,\%$). These decrements could be coincidental, considering the importance of OH-chemistry and secondary CO production in that region. In the northern hemisphere, *noVIIRS* (Figure 3.7e) and *GFED* (Figure 3.7f) feature slightly higher increments over eastern Europe (*noVIIRS* $< +10\,\%$, *GFED* up to $+30\,\%$), North America (*noVIIRS* $< +10\,\%$, *GFED* $< +20\,\%$), and Siberia (*noVIIRS* $< +15\,\%$, *GFED* $< +5\,\%$) compared to the *reference* inversion. These differences could point to aliasing of the secondary production emission category to the biomass burning category. FINN2.4+VIIRS, which is used as biomass burning a priori in the *reference* inversion, has generally the highest emissions, mostly due to capturing small fires, which are common in these regions. For the other two, the model attempts to capture these missing sources, in part, through increasing the emissions in the other categories. Again, this misattribution can also be seen in the budgets in Table 3.3, where the posterior total emitted mass is very similar for all experiments of the first set, but the distribution over the three emission categories varies considerably.

In Figure 3.8, one month of the relative increments for the CO production from VOCs and CH_4 are shown for the second set of inversions. Figure 3.8a is from the *reference* inversion based on a combination of gridded satellite observations and surface flasks. As such, the content of Figure 3.7d above is repeated there for ease of comparison. Very similar results (Figure 3.8b) are obtained with the *full satellite* inversion, as already shown at the surface stations in Section 3.4.1. Minor differences are visible over North America and Siberia, likely due to less aliasing to the biomass burning category. When the higher resolution observations are used, the short term and local biomass burning events are more distinct, which makes it easier for the model to capture them in the appropriate category.

For the *satellite only* inversion (Figure 3.8c) many regional features are much less pronounced. However, the broader distribution of emission increments remains the same: There are still negative increments in a band between the Equator and $40°\,$N and over South America, and positive increments over southern Africa and the adjacent oceans. The positive increments over North America, Europe, and Siberia are weaker
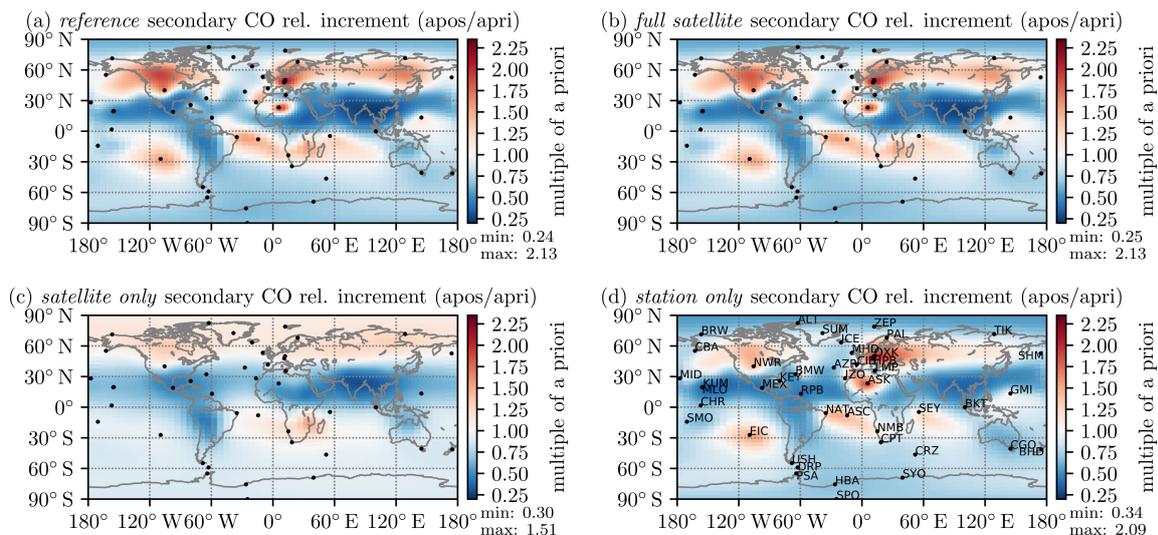
Figure 3.8: Global secondary CO relative emission increment for September 2018 for the second set of inversions, based on different observational datasets. The panels show the factor by which the emissions increased for (a) the *reference* inversion, (b) the *full satellite* inversion, (c) the *satellite only* inversion, and (d) the *station only* inversion. The locations of the surface stations are indicated with dots for easier orientation, in the last panel additionally with their station code. Note that Panel (a) of this figure is the same as Figure 3.7d.

and appear to be spread out over the whole northern hemisphere north of around 45° N, including over the oceans. These weaker features are likely linked to the different spatial distributions of observations in the two datasets; while there are many maritime stations and stations in the remote northern hemisphere, satellite observations there are more sparse and mostly found in continental regions. Additionally, towards the end of the year, i.e. the second half of the main inversion period, there are no more satellite observations at high northern latitudes, as exemplified in Figure 3.1 for one day in early November. All of this, in combination with the spatial correlations given to the optimizer, causes the model to prefer smooth, broad patterns to fill in any gaps.

These differences in information content between the two observational datasets stress the importance of the error inflation (Section 3.3.2.2). If the error on the satellite observations is not inflated, the optimized emissions end up very close to the ones from the *satellite only* inversion, because the signal from the sparse flask measurements is overshadowed. However, the current inflation may be too large, which causes the optimizer to "overfit" certain stations that are not well captured by the model. As can be seen in Figure 3.8d for the *station only* inversion, some stations clearly drive the model away from these broad patterns and towards strong positive regional increments. This overestimation is especially apparent for Assekrem (ASK) and Izana (IZO) stations, which lead to large increments over north-west Africa, and Hohenpeissenberg (HPB) and Ochsenkopf (OXK) stations, which drive emissions over central Europe up strongly. Neither of these increments are observed or supported by the satellite ob-

servations. Notably, all of these stations are at high altitudes, potentially pointing to short-comings in the representation of topography in the model. However, there are mountainous stations, like Mauna Loa (MLO), that are captured well by the model.

Less pronounced examples of overfitted stations are Rapa Nui (EIC) and Tutuila (SMO), which cause positive and negative increments over the southern Pacific, respectively. However, it should be noted that for the satellite the number of observations over oceans to constrain those emissions is very limited and, as shown for Rapa Nui in Figure 3.6f, the *satellite only* inversion still manages to fit these stations reasonably well.

Another factor that could play a role in the context of overfitted stations is the strength of the vertical transport in TM5, which KROL et al. [2018] find to be somewhat faster than in other models. This implies low vertical gradients in the troposphere and that modeled tracer mass might be transported upwards before the model can be sampled at the location of the station for comparison to the real observations. This is especially problematic for remote stations with limited surface sources in the vicinity, such as Rapa Nui (EIC) in the south-eastern Pacific. There, the model is forced to introduce unrealistic increments to the secondary CO source in the middle of the Pacific. Furthermore, due to the way those emissions are handled within the model, this will introduce additional CO over the whole column (and not only at the surface), which then hampers the comparison to the satellite observations. Similarly, for the station in the Assekrem, in the inversions that include station data, the low vertical gradients cause the optimizer to introduce unrealistically high secondary CO emissions over the Sahara. In contrast, those increments do not occur in the *satellite only* inversion, because the satellite observes the total column with a very limited vertical resolution and is, therefore, less affected by the vertical gradient in the model.

Finally, even in the *station only* inversion (Figure 3.8d), some station driven features appear weaker compared to the *reference* inversion (Figure 3.8a). For example, the positive increments over North America are much weaker and the spikes around the Assekrem and in central Europe are more spread out. These weaker features are again caused by a combination of the prescribed spatial correlations and the distribution of the available observations. While in the *station only* inversion the model prefers broader patterns to follow the prescribed spatial correlation of the emissions, in the *reference* inversion there are satellite observations all around the landlocked stations, which drive the model towards lower increments. Overall, the *station only* inversion is driven to the largest emitted mass of all experiments as shown in the budgets in Table 3.3. This is in line with the increased emissions around surface stations postulated in the context of the (too) fast vertical transport in TM5 above.

3.4.3.2 Anthropogenic emissions
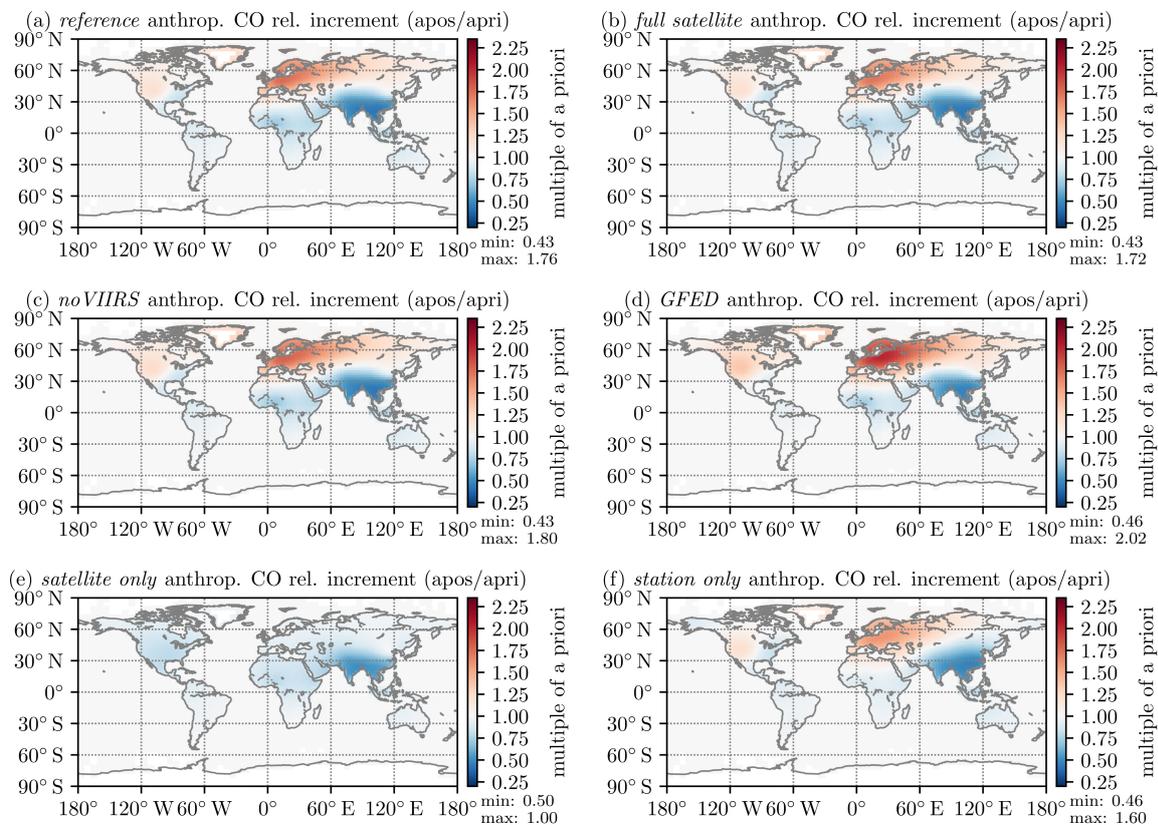
Figure 3.9: Relative global anthropogenic CO emission increments for September 2018 for all six inversion experiments. Panel (a) shows the *reference* inversion with FINN2.4+VIIRS as biomass burning a priori, and gridded satellite observations and surface flasks as observational input. The variations are (b) *full satellite* observations instead of gridded, (c) *noVIIRS* with FINN2.4 as biomass burning a priori, (d) *GFED* with GEFED4.1s as biomass burning a priori, and (e) *satellite only* and (f) *station only* to drive the inversion.

To better identify the aliasing between the emission categories, Figure 3.9 provides an overview of the relative increments in the optimized anthropogenic emissions for all six inversion experiments from this chapter. In Figure 3.9a the relative emission increments are shown for the *reference* inversion based on FINN2.4+VIIRS, and a combination of gridded satellite observations and surface flasks. The largest changes are positive increments over Europe, and negative increments over China and India. To investigate these increments further, one should consider that the anthropogenic a priori emissions taken from CMIP6 are projections for 2018, rather than historical data. For China, these projections predict relatively constant emissions. However, China managed to significantly reduce its CO emissions in recent years [KANAYA et al., 2020] in the scope of air quality policies, like the Coal to Gas policy only implemented in 2013 [LIU et al., 2020]. Additionally, the effect of most of these policies was somewhat offset by strong biomass burning years up until 2015 [ZHANG et al., 2020], making their effect harder to assess in advance. Regardless, reduced CO concentrations have been

now observed all over China, both at surface stations [LIU et al., 2019; ZHAI et al., 2019; LI et al., 2020] and from satellites [ZHANG et al., 2020]. This observed reduction has been linked to a decrease in emissions as calculated using inverse modeling [ZHENG et al., 2018]. The reduced emissions are most likely due to anthropogenic rather than natural factors [KANG et al., 2019]. By 2018, the year analyzed here, all of this adds up to at least part of the significant offset in CO emissions found above.

Unlike for China, there is no clear explanation for the negative increments over India. These might be an artifact due to spatial correlation, where India's proximity to China implies that it is cheaper in terms of the cost function to reduce emissions over a larger region, rather than strongly reducing only China's emissions. This could be compounded by low observational coverage, especially with regard to surface stations, and an OH climatology not appropriate for recent years.

When compared to the *full satellite* inversion shown in Figure 3.9b, again, the increments are almost the same, further justifying the usage of gridded satellite observations on a global scale to reduce the computational cost.

The *noVIIRS* (Figure 3.9c) and *GFED* (Figure 3.9d) inversions are slightly worse at capturing the small fires in Europe and North America compared to the *reference* inversion. The missing small fires lead to apparent anthropogenic increments, especially for *GFED*, over Europe and western Russia to close the CO budget. Further evidence for this aliasing is provided in Table 3.3, where the total a posteriori emissions for the inversions of the first set are almost identical, but the partitioning over the emission categories differs significantly. As such, *GFED* has over 35 % lower biomass burning emissions compared to *reference*, but almost 10 % and 8 % higher secondary production and anthropogenic emissions, respectively.

For the *satellite only* inversion, the relative anthropogenic emission increments are pictured in Figure 3.9e. They stay relatively close to, but below, 1 globally, i.e. the inversion mostly agrees with the a priori. Over India and China, again, a clear decrement is visible. Notably, there is no increment over Europe, in contrast to what is found when flask observations are included. In Section 3.4.1, this smaller increment caused the station at Hohenpeissenberg (Figure 3.6c) to be considerably underestimated in the *satellite only* inversion.

The *station only* inversion shown in Figure 3.9f leads to very similar results in terms of anthropogenic increments compared to the *reference* inversion. This shows how well the NOAA station network on its own is capable of constraining the global broad-scale background emission patterns. Differences include smaller increments over Europe and smaller decrements over Africa and an apparent shift of the decrement over India and China towards the East. The latter may be explained by a lack of background stations and, therefore, a lack of observations in that region, causing the decrement to be smoothed out due to spatial correlation.

Overall, the anthropogenic increments shown in Figure 3.9 compared to the ones for the secondary CO production in Figures 3.7 and 3.8 show similar general structures, with decrements in China and India and increments in Europe. However, there are noticeable differences both in finer scale spatial details, for example, the anthropogenic increments over Europe are more spread out towards Eastern Europe, and large scale patterns, with much smaller relative increments for North America. Generally, the ratios of a priori to a posteriori emissions, i.e., the relative emission increments, are not the same for all three categories. In other words, while there is some aliasing, the inversion setup is still capable of simultaneously optimizing multiple emission categories, which is ensured in the following ways:

Firstly, because of the different a priori errors, even in regions with similar spatial structures, the amplitudes of the relative emission increments differ significantly. Secondly, the different correlation lengths and times for each emission category, as introduced in Section 3.2.3.1, ensure that only the biomass burning category is capable of capturing short and local events. Conversely, long-lasting, large-scale mismatches could still lead to aliasing across all categories, as is the case, for example, over China. Thirdly, the a priori emissions of all three categories feature different spatial structures. These a priori structures, combined with enforcing spatial and temporal correlation, imply that it is cheapest for the model to change emissions following the 'spatial signature' of the correct source category, rather than evenly distributing the increments over all categories. An example for this can be found over North America, where the anthropogenic emissions are barely changed, while there are significant changes in the secondary CO production.

3.4.3.3   Biomass burning

An in-depth analysis of the optimized biomass burning emissions is not included in this thesis, because the low model resolution is not sufficient to capture individual burning events. This promotes aliasing between the emission categories, where the biomass burning emissions are in- or decreased in large regions co-located to the patterns observed in the secondary CO production. As an example of this, Figure A.3 shows the absolute biomass burning increments for 15 September 2018, the day in the center of the period analyzed above. Because the temporal variability in the secondary CO production is low, the biomass burning emissions also remain relatively constant in time.

3.5   Conclusions

TROPOMI satellite observations were introduced into the TM5-4dvar inverse modeling suit to optimize global CO emissions from three distinct emission categories (biomass burning, anthropogenic, and secondary production) in a set of six inversion experiments. The model ran at a relatively coarse resolution of up to $3° \times 2°$, which allowed for the use of satellite super-observations gridded to $0.5° \times 0.5°$ to reduce the computational cost. Compared to the inversion based on the full-resolution (up to $7 \times 7 \, \text{km}^2$) satellite observations, differences in the final mixing ratios and optimized emission fields were minimal. Yet, the computation time per iteration was around 25% longer for the full resolution inversion. However, at $3° \times 2°$ resolution, the model could not properly resolve the spatial scale of individual biomass burning events. This resulted in heavy aliasing of the biomass burning emissions to the other emission categories. In future studies, using additional observations to further constrain emissions from specific sources or by employing a finer zooming region could improve model performance. With the latter, such an inversion could make use of the full potential of the TROPOMI observations.

The comparison of model results and observations is vastly improved by the inversion and the a posteriori mixing ratios closely follow the observed values. Notably, this even holds true in regions like China and the North Pacific, where the a priori strongly overestimated the mixing ratio and very large emission decrements are required to reach a good a posteriori fit. The overestimated a priori mixing ratios in those regions reveal inconsistencies between the OH climatology used to simulate chemical loss, and the secondary CO production terms taken from the TM5-MP model. This will be further investigated in Chapter 4. For the inversion based only on satellite observations, sizable mismatches between model results and flask measurements remain for stations at high northern latitudes. These mismatches can be explained by considering that mixing ratios at high northern latitudes, on the one hand, are poorly constrained by the satellite observations, especially towards the end of the year, and, on the other hand, are governed by transport from the (well-constrained) mid latitudes, which leaves little leeway for the optimizer. Additionally, in the inversions based on flask measurements, there are very large increments around high-altitude stations. These increments are most likely linked to the coarse model orography that comes with the overall coarse model resolution and limitations in the representation of meteorology in the model. The latter will be further investigated in Chapter 6. Despite good coverage in those regions, the inversion based only on satellite observations neither confirms nor reproduces those strong increments. As such, for future inversions in this framework, an increased model representation error should be applied to those specific stations, to avoid biasing results by overfitting.

In the southern hemisphere, results are very similar across all inversions, regardless of the observational dataset(s) (satellite, stations, or both) used. This indicates that, in the southern hemisphere, either dataset is equally capable of and sufficient for constraining the background emissions and leads to the same mixing ratios. Potentially, these promising results could allow for inversions based solely on TROPOMI observations, so long as the region of interest is sufficiently far south of $55°$ N. There, as well as for validation, bias correction, and overall confidence in the optimized emissions, the surface flasks still play a crucial role in the inversion. By using the TROPOMI observations on their own, the long analysis cycle of the surface flasks could be circumvented and specific events could be investigated using this model in a more timely manner (within weeks rather than months), and only be verified against and adjusted by the flasks at a later stage.

Overall, the most reliable results are found from inversions using both datasets, because they complement each other in multiple ways. Firstly, their spatial coverage differs slightly – while the satellite observations are mostly valid over land but sparse over the oceans, most background stations are located on remote islands or in coastal settings. Secondly, both datasets on their own have very limited information on the vertical tracer distribution, where the flasks probe only the surface layer and the satellite observations provide only total column mixing ratios. Combining those datasets can yield better constraints on the vertical tracer distribution in places where in situ and satellite observations are co-located. Finally, in a joint inversion, the satellite observations are implicitly verified versus the flask measurements and it becomes possible to identify potential biases in the satellite observations. However, when using both datasets at once, the technical limitations of both apply, i.e. the high computational cost from using the satellite observations, and the long analysis cycle of the flask measurements.

Additionally, especially in a joint inversion, properly inflating the uncertainty of the satellite observations is crucial to weight them against the surface measurements. Where in previous studies such an inflation factor was empirically estimated, a more rigorous approach was introduced in this chapter. This approach still has a number of limitations, some of which will be tackled with the further developments presented in Chapter 5.

Chapter 4

Towards a consistent OH climatology

The inversion experiments conducted in the previous chapter resulted in significant biases in the optimized CO emission, especially in the northern tropics. This chapter presents how these biases are successively reduced by updating the boundary conditions for biomass burning, OH concentrations, direct natural CO emissions, and secondary CO production from VOCs and methane. Parts of this chapter are currently being prepared for publication.

In Section 4.1, an introduction to the topic and an overview of the aims of this chapter are given. Section 4.2 shortly introduces the used model setup and the conducted experiments and includes as in-depth comparison of the used OH fields. In Section 4.3, the results are presented, showcasing the achieved improvements. Section 4.4 then provides a summary over the findings of this chapter.

4.1   Background and objective

The principle function of an inversion is to find a state, e.g. a set of emissions, that reconciles a model with a set of observations. This becomes problematic if the deviations to the observations are caused by a process that is not part of the state. In such a case, the model might still be able to reach 'realistic' a posteriori mixing ratios that fit the observations reasonably well and recover trends in the emissions, however, the underlying absolute a posteriori emissions may not be meaningful [Yin et al., 2015]. As a more concrete example, in Chapter 3 an unbalanced prior budget in the TM5-4dvar model is found, which requires large emission corrections, especially in the northern tropics, and points to inherent flaws in the initial assumptions. These flaws could be benign if they are limited to the a priori emissions, since correcting those is the purpose of the inversion. However, if a significant portion of those flaws originates from

another part of the model, such as the chemistry parameterization, the dry deposition scheme, transport and meteorology, or, in this case, the prescribed hydroxyl radical (OH) fields, the inversion cannot properly represent that portion of the flaws and will instead capture it by introducing superfluous emission increments.

The distribution and abundance of OH has a significant impact on carbon monoxide (CO) inversions by modifying both the sources and the sinks of CO and can introduce notable biases in the optimized emissions [KOPACZ et al., 2010; MIYAZAKI et al., 2015; MÜLLER et al., 2018]. However, due to its short atmospheric lifetime ($\approx 1\,$s), in-situ measurements of OH are challenging [PATRA et al., 2014]. Instead, OH concentrations are estimated via proxy measurements, predominantly using the man-made solvent methylchloroform (CH_3CCl_3, MCF; e.g. NAUS et al. [2021]). However, NAUS et al. [2019] show OH to be currently under-constrained by the sparse available MCF measurements. Hence, finding MCF alternatives is subject to ongoing research (e.g. KROL et al. [2008]; LIANG et al. [2017]). Compared to these observation-based OH estimates, models generally find higher OH concentrations, especially in the northern hemisphere. However, the differences between model and observation based OH distribution estimates are overall within the uncertainty of the measurements [NAIK et al., 2013]. Explanations for the overestimated model OH include model biases in ozone (O_3) and water vapor (H_2O), both of which affect OH primary production [NAIK et al., 2013], and an underestimation of model CO in the northern hemisphere, which affects OH loss and is possibly caused by underestimated anthropogenic surface emission [SHINDELL et al., 2006; STEIN et al., 2014] or overestimated loss to dry deposition during northern hemispheric winter [STEIN et al., 2014]. Additionally, modeled OH estimates vary strongly between models. For example, MURRAY et al. [2021] find the tropospheric mean OH concentrations across multiple ACCMIP models, which use the same emissions, to vary by $\pm 30\,\%$ around the multi-model mean, and link this variation to differences in the lifetimes of nitrogen oxides (NO_x), which affect OH regeneration.

There are attempts to tackle the issue of the under-determined OH distribution through multi-sensor and multi-species inversions. For example, in their formaldehyde (HCHO) study, FORTEMS-CHEINEY et al. [2012] use satellite observations of CO and HCHO together with surface measurements of MCF and methane (CH_4) to constrain OH, among other species. YIN et al. [2015] use satellite observations of CO and station measurements of MCF and CH_4 with similar target species as FORTEMS-CHEINEY et al. [2012]. Both find CO to be well constrained, while the optimized OH fields depend heavily on the prior assumptions. MIYAZAKI et al. [2015] employ satellite observations of nitrogen dioxide, O_3, nitric acid (HNO_3), and CO and constrain OH indirectly by optimizing, among others, species that govern the chemical processes related to OH, namely O_3 for primary production, CO for removal, and NO_x for regeneration. While

this approach reduces the gap between model- and observation-based OH estimates, they conclude that possibly some errors remain in their modeled OH.

In this chapter, the aim is the investigation of how those model biases in CO and OH impact inversions in the TM5-4dvar inverse modeling framework. As introduced in Chapter 3, a substantial part of the prior emissions for the TM5-4dvar model, namely the secondary CO production terms from the oxidation of volatile organic compounds (VOCs) and CH_4, are taken from the chemical transport model TM5-MP [Myriokefalitakis et al., 2020], in which the aforementioned northern hemispheric low biases in CO and high biases in OH also exist. Currently, TM5-4dvar uses climatological OH fields based on Spivakovsky et al. [2000], which do not match the OH field underlying secondary CO production terms. Since OH and CO are estimated based on different methodologies, which are in turn based on different meteorologies, inconsistencies in (precursor) transport patterns may cause issues [compare e.g. Jiang et al., 2011]. Additionally, in regions where the OH in TM5-MP is high, the resulting high secondary CO production carries over to TM5-4dvar, but unlike in TM5MP, in TM5-4dvar that CO may be subject to smaller loss due to lower climatological OH. This discrepancy leads to the accumulation of CO in TM5-4dvar, which the inversion can only correct by lowering the emissions to unrealistic values. Such large increments can be problematic considering the linearized chemistry scheme, where changes in CO do not feedback onto the prescribed OH field, causing an inversion to generally overestimate changes [Stavrakou & Müller, 2006].

The following describes multiple attempts to find a more realistic prior by modifying the boundary conditions of the inverse model in multiple steps, so that the inversion can produce meaningful increments that reflect actual emission changes, rather than biases in OH. The impact of those changes is tested by performing a series of CO inversions constrained by satellite observations and surface station measurements and comparing their results. The biggest modification consists of replacing the climatological OH by monthly mean OH fields from a chemistry transport model. Similar approaches of comparing the impact of climatological and modeled OH can be found in Jiang et al. [2015] for optimized CO and Zhao et al. [2019] for modeled CH_4. Further, the biomass burning a priori is updated and climatological biogenic and oceanic CO emissions are explicitly include. To round out the comparison, an additional inversion is performed, where the prescribed model-based secondary CO production fields, see Section 3.2.3.1, are replaced by the corresponding fields that Nechita-Banda et al. [2018] employ in their inversion study.

4.2 Materials and methods

All inversions analyzed in this chapter were performed using the Cycle 3 TM5-4dvar model (Section 2.2) with the same basic setup as in Section 3.2.3, to ensure comparability between the results of both chapters. For this chapter, the base code was updated to revision c71f31 of the official code repository of the model. Additionally, support for biogenic and ocean CO emissions based on MEGAN and MACC, respectively, was added, and the capability to use the output from the full-chemistry model TM5-MP to obtain OH fields to simulate chemical loss was implemented.

Table 4.1: A priori emissions and OH used for the conducted inversion experiments. Spiv-OH and TM5-MP-OH refer to monthly climatological OH fields based on SPIVAKOVSKY et al. [2000] and monthly mean OH taken from the TM5-MP model, respectively. FINN2.4 and FINN2.5 refer to versions 2.4 and 2.5 of the FINN inventory including small fires from VIIRS. The first two experiments, *spin-up* and *FINN2.4*, are identical to the *spin-up* and *reference* inversions from Chapter 3, respectively.

Inversion	A priori emissions				OH
	biomass	anthro.	secondary	bio+ocean	
spin-up	FINN2.4			-	Spiv-OH
FINN2.4	FINN2.4		TM5-MP MOGUN-TIA	-	Spiv-OH
reference	FINN2.5	CMIP6		-	Spiv-OH
MOGOH	FINN2.5			-	TM5-MP-OH
MEGAN	FINN2.5			MEGAN +MACC	TM5-MP-OH
CB4prod	FINN2.5		TM5-MP CBM4	-	Spiv-OH

In this chapter, as outlined in Table 4.1, the setup of the *reference* inversion experiment from Chapter 3 was successively refined in three steps:

First, the biomass burning a priori was updated to version 2.5 of the Fire INventory from NCAR (FINNv2.5), which is available at WIEDINMYER & EMMONS [2022] and described in WIEDINMYER et al. [2023]. This defines the new *reference* inversion. Note that the *reference* inversion from the previous chapter is referred to as *FINN2.4* in this chapter and different to the *reference* inversion from this chapter.

Second, updated OH fields are introduced. Previously, the inversions used monthly climatological OH fields (Spiv-OH) from SPIVAKOVSKY et al. [2000], scaled by 0.92, as suggested in HUIJNEN et al. [2010]. These were replaced with monthly-mean OH fields (TM5-MP-OH) taken from the full chemistry model TM5-MP with the extended MOGUNTIA chemical scheme described in MYRIOKEFALITAKIS et al. [2020]. Specifically, the TM5-MP-OH came from the same simulation for the year 2018 as the 3D fields used to represent secondary CO production from the oxidation of CH_4 and other VOCs introduced in Chapter 3. An in-depth comparison of the two OH fields can be found in Section 4.2.1. The inversion experiments in this chapter still used the simplified CO-only chemistry version of TM5-4dvar described in HOOGHIEMSTRA et al. [2011].

In this scheme, both the OH fields and secondary CO production are prescribed and not calculated online.

Third, an emission category for direct natural CO emissions, i.e. CO from biogenic and ocean sources, was added. Due to their relatively small contribution to the total CO budget in comparison to the prior errors on the other categories, they were previously neglected. Specifically, the direct natural CO emissions were split into ocean CO from the Monitoring Atmospheric Composition and Climate (MACC) inventory and biogenic CO emissions from the Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGANv2.1) inventory from MEGAN-MACC distributed by the Emissions of atmospheric Compounds and Compilation of Ancillary Data [Guenther et al., 2012, ECCAD] repository. As in the TM5-MP simulations that generated the secondary CO production and OH fields described above, both natural emission sources were fixed to the fields for the year 2000. Because of weak observational constraints over the oceans, this new category was not optimized. Additionally, according to Zheng et al. [2019], little variation is to expected in the biogenic and ocean CO sources.

To further investigate the impact of the updated secondary CO production terms, an additional inversion was run based on the production fields previously used in Nechita-Banda et al. [2018] in conjunction Spiv-OH as described above. The production fields in Nechita-Banda et al. [2018] were taken from a 2010 simulation of the TM5-MP model with CBM4 chemistry, which is described in Huijnen et al. [2010]. To ensure a straightforward comparison, the settings were otherwise the same as for the *reference* experiment. For the same reason, all inversions started from tracer fields from the same *spin-up* inversion as in Chapter 3.

Overall, the optimizer settings for each category, i.e. a priori error, temporal correlation of emissions in the same cell, spatial correlation of emissions in neighboring cells, and temporal resolution of the state, remained unchanged to Chapter 3 and are summarized in the appendix in Table A.2.

4.2.1   OH field comparison

The two OH fields used in this chapter feature several notable differences. The following will provide an analysis of the tropospheric mean OH concentrations, the interhemispheric concentration ratios, and the spatial distributions of both fields and compare each to values reported in previous literature.

To calculate the volume- and airmass-weighted tropospheric annual mean OH concentrations, the volumes and the airmasses were based on the 3D pressure and temperature fields for 2018 from the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis project [ERA-Interim meteorology; Dee et al., 2011]. For both OH fields the same tropopause cutoff was applied at 150 ppb ozone (O_3) based

on 3D monthly mean fields from the same 2018 TM5-MP simulation that produced the TM5-MP-OH. Calculations for the Spiv-OH resulted in a volume-weighted mean of $9.9 \times 10^5 \frac{molec}{cm^3}$ and an airmass-weighted mean of $10.4 \times 10^5 \frac{molec}{cm^3}$. This tropospheric mean is on the lower end compared to the $(11.1\pm1.7)10^5 \frac{molec}{cm^3}$ [SHINDELL et al., 2006], $(11.1\pm1.6)10^5 \frac{molec}{cm^3}$ [NAIK et al., 2013], and $(11.7\pm1.0)10^5 \frac{molec}{cm^3}$ [VOULGARAKIS et al., 2013] reported as airmass-weighted tropospheric multi model mean concentrations for the year 2000 for different sub-sets of ACCMIP simulations. Furthermore, LIANG et al. [2017] present a trend-base two-box model approach, with which they estimate the mean global mass-weighted tropospheric OH concentration at $11.2 \times 10^5 \frac{molec}{cm^3}$ using methyl chloroform (MCF) observations. In comparison, the mean tropospheric concentrations for the TM5-MP-OH were significantly larger, at a volume-weighted mean of $13.6 \times 10^5 \frac{molec}{cm^3}$ and an airmass-weighted mean of $13.4 \times 10^5 \frac{molec}{cm^3}$. This overestimation of OH points to a known issue in chemical transport models. For example, in their multi-model comparison, LIANG et al. [2017] point out that the models that calculated OH online would have 20–30 % higher OH than those that used climatological OH fields based on SPIVAKOVSKY et al. [2000]. Similar ranges are found by ZHAO et al. [2019] who investigate the impact of using OH fields from different models in the LMDz model. They report volume-weighted mean tropospheric OH between $8.7 \times 10^5 \frac{molec}{cm^3}$ and $12.8 \times 10^5 \frac{molec}{cm^3}$, with a mean of $(10.5\pm1.1)10^5 \frac{molec}{cm^3}$ or an airmass-weighted mean of $11.3 \times 10^5 \frac{molec}{cm^3}$. As detailed in Section 4.1, these overestimations in modeled OH are likely linked to an underestimation of CO in the northern hemisphere.

To further describe the discrepancies in OH between models, another quantity that is often used is the interhemispheric ratio of the tropospheric mean concentrations [e.g. NAIK et al., 2013; PATRA et al., 2014; ZHAO et al., 2019, and sources therein]. While it would be sensible to split the hemispheres based on the location of the Inter-Tropical Convergence Zone (ITCZ), for practical reasons the Equator is usually used instead. Using the tropospheric mean concentration for each hemisphere calculated as above, resulted in an annual mean interhemispheric OH ratio of:

$$\frac{10.2 \times 10^5 \frac{molec}{cm^3}(NH)}{9.7 \times 10^5 \frac{molec}{cm^3}(SH)} = 0.96 \qquad (4.1)$$

for the Spiv-OH if using volume-weighted means and 0.99 based on airmass-weighted means. As for the global tropospheric mean concentration, the north-south ratio of the Spiv-OH field is in good agreement to observation based ratios, e.g. from PATRA et al. [2014], who use MCF observations from 2 station networks and an aircraft campaign in an atmospheric chemical transport model to infer OH for 2004 to 2011 and found a ratio of 0.97±0.12.

For the TM5-MP-OH field, the ratio was significantly lager, at:

$$\frac{15.1 \times 10^5 \, \frac{\text{molec}}{\text{cm}^3} (NH)}{12.1 \times 10^5 \, \frac{\text{molec}}{\text{cm}^3} (SH)} = 1.25 \tag{4.2}$$

for volume-weighted means and 1.28 for airmass-weighed means. This agrees very well with the ratios the inter-model comparisons by NAIK et al. [2013] and ZHAO et al. [2019] find, of 1.28±0.1 and 1.3, respectively, each based on airmass-weighted means. However, NAIK et al. [2013] consider it likely that this ratio is overestimated in the models due to biases in O_3 and CO. STRODE et al. [2015] follow a similar reasoning and investigate possible causes for the too low modeled CO that leads to too high modeled OH. They find the largest impact on interhemispheric OH ratio from anthropogenic nitrogen oxides, followed by CO emissions, and that, while biases in O_3 and H_2O might also play a role, the latter mostly affects global mean concentration, but not the interhemispheric ratio.

Finally, the more detailed spatial distributions of the two OH fields differ substantially in some regions. However, those differences are not represented by the broad quantities considered above. While both fields peak in the tropical mid-troposphere, this peak is at a greater altitude (around 600 hPa) in Spiv-OH than in TM5-MP-OH (around 700–800 hPa, depending on season). In combination with vertical transport, this difference affects chemistry, since with TM5-MP-OH air reaches higher OH concentrations sooner. Additionally, while the TM5-MP-OH is larger in most places, this is particularly true close to the surface in the northern tropics, where concentrations are larger by well over a factor of 2, which has significant implications for VOC chemistry. While TM5-MP-OH is likely too large overall, this feature is in agreement with NAUS et al. [2021], who conduct a 3D inversion of MCF to constrain OH, which effectively shifted OH towards the northern tropics compared to Spiv-OH. Spiv-OH is significantly larger than TM5-MP-OH at higher altitudes in the southern extratropics. However, since OH concentrations are overall low in that region, the impact on CO chemistry should be minimal. As a side note, SPIVAKOVSKY et al. [2000] additionally claim that their southern extratropical OH might be roughly 25 % too large, making up for some of the difference. Similar differences between modeled OH and climatological OH are shown in LIANG et al. [2017] and YIN et al. [2015].

A more detailed look at the two OH fields may be found in the appendix, where the zonal means of the vertical distribution, per season, are shown in Figures A.14 to A.17 and the lateral distribution for the surface layer, per season, can be found in Figures A.18 to A.21.

4.2.2 Observations

All inversion experiments analyzed in his chapter assimilated the same two datasets. Detailed descriptions of the datasets and any preprocessing steps, for example filtering or gridding, can be found in Section 3.3. This section provides a short recap.

The first dataset consists of NOAA surface flask CO measurements from various stations assembled by the Carbon Cycle Greenhouse Gases (CCGG) group [PETRON et al., 2022]. After filtering for background stations, these were assimilated with a sampling error of 2 ppb, in addition to the representation error calculated by the model.

The second dataset is the TROPOMI/WFMD version 1.2 product from the Carbon and Greenhouse Gas Group at the Institute of Environmental Physics (IUP) of the University of Bremen, retrieved with the Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm [SCHNEISING et al., 2019], gridded to $0.5° \times 0.5°$.

While this satellite product has been superseded by a more recent version, the experiments in this chapter still used the version 1.2 product to maintain direct comparability to the results of Chapter 3. Differences to the latest version (1.8) are presented in Chapter 5. The satellite error inflation values used for the inversion experiments presented here, were similar to the 72 used for the *FINN2.4* inversion. The *reference* and *MOGOH* inversions both also used 72, while *MEGAN* required a slightly larger inflation factor of 74 and *CB4prod* a slightly smaller at 70.

4.3 Results

4.3.1 Updated biomass burning a priori

Compared to FINNv2.4, used in the *FINN2.4* inversion, FINNv2.5, used in the *reference* inversion, has regionally up to 20 % lower emissions, which leads to globally around 8 % lower biomass burning a priori emissions in the *reference* inversion. This is equivalent to a reduction of < 2 % in the total a priori emissions. The reduced emissions caused minor reductions in the modeled mixing ratios, which in turn resulted in marginally smaller prior mismatches, as can be seen in Table 4.2. However, the posterior mixing ratios, and with them the posterior mismatches, remained almost the same.

Table 4.3 shows the budgets for the inversions analyzed in this chapter. Between the *FINN2.4* and the *reference* inversions, the total emissions were unchanged, but the partitioning of the emission categories had shifted from biomass burning (−4.4 %) to secondary CO (+1.5 %). However, as a caveat, it should be noted that the current setup constrained the partitioning only relatively weakly due to coarse model resolution. While the different temporal state resolutions allowed the model to distinguish

Table 4.2: Averaged error-weighted prior and posterior mismatches between model and observations for each inversion experiment. The first 4 rows give the mean mismatches to different subsets of the flask measurements. Using Spiv-OH in conjunction with secondary CO production fields from TM5-MP with MOGUNTIA chemistry leads to poor prior fits in the northern tropics (*FINN2.4* and *reference*). When TM5-MP-OH is used instead (*MOGOH* and *MEGAN*) the prior fit north of 23° N strongly degrades, but improves on the rest of the globe and the posterior fit improves. Using Spiv-OH with the old secondary CO production from TM5-MP with CB4 chemistry leads to the best prior fit overall, but also to the poorest posterior fit. Note that for satellite mismatches shown in the last row, the values were scaled up by 10^3 for readability. Because of the very large observation count, the mismatch, i.e. cost contribution, for each individual satellite observation is very small.

	observations		*FINN2.4*	*reference*	*MOGOH*	*MEGAN*	*CB4prod*
	all	prior	21.91	20.59	32.34	26.84	13.79
		posterior	3.61	3.63	3.43	3.31	4.03
stations	> 23° N	prior	20.95	19.93	60.80	49.62	21.56
		posterior	5.40	5.41	4.97	4.78	5.94
	northern tropics	prior	47.93	45.55	10.41	8.36	11.91
	(0 to 23° N)	posterior	2.90	2.90	3.02	2.97	3.23
	southern	prior	9.67	8.48	2.87	3.73	3.59
	hemisphere	posterior	1.42	1.46	1.43	1.38	1.69
	satellite	prior	11.05	10.15	3.39	3.26	3.65
		posterior	0.99	1.00	0.97	0.94	1.06

short pulses from biomass burning from the long-term changes in secondary CO production and anthropogenic emissions, it could not resolve local features on the scale of individual biomass burning events.

4.3.2 New OH field

4.3.2.1 Budget analysis

Changing the OH field from Spiv-OH to TM5-MP-OH had a major impact on the inversion results. A broad overview of this impact can be found in Table 4.3, which shows the global prior and posterior budget terms for all inversion experiments analyzed in this chapter. Due to the much higher tropospheric mean OH concentrations, as discussed in Section 4.2.1, the final masses after the initial 7 month forward run based solely on the prior emissions were significantly lower, i.e. the atmosphere was cleaner, in the *MOGOH* inversion compared to the *reference* inversion. Notably, this cleaner atmosphere was, in terms of total masses, much closer to the observation-based posterior atmosphere. As pointed out before, the prior budgets in the *FINN2.4* and *reference* were unbalanced, but appeared to be mostly closed in the *MOGOH* inversion.

Further evidence for an unbalanced prior budget was also visible in the more detailed budget terms. In the *reference* inversion, the emissions were reduced across the board, which also implied notable reductions in the loss terms. The secondary production in particular was lowered by 25 %. In contrast, in the *MOGOH* inversion, the total

Table 4.3: Global prior and posterior budgets for each inversion in Chapter 4, as a sum over the global and the zooming regions. The zooming column contains the net mass exchange between the zooming regions via the communication cells. The initial and final masses denote the total amount of CO in the modeled atmosphere at the beginning (1 June 2018, as taken from the *spin-up* inversion) and end (31 December 2018) of the inversion period, respectively. Note that the annual rates (Tg CO yr⁻¹) are extrapolated from the emissions during the 7-month inversion period from June to December 2018.

experiment		masses in Tg CO		losses in Tg CO yr⁻¹		zooming in Tg CO yr⁻¹	emitted in Tg CO yr⁻¹				
		initial	final	chemical	deposition		total	secondary	biomass	fossil fuel	natural
FINN2.4	prior	555	747	-3035	-220	118	3464	2179	666	618	-
	posterior	555	585	-2487	-187	22	2702	1616	572	514	-
reference	prior	555	739	-2995	-216	114	3411	2179	613	618	-
	posterior	555	585	-2487	-187	22	2703	1641	547	515	-
MOGOH	prior	555	558	-3200	-195	-12	3411	2179	613	618	-
	posterior	555	576	-3200	-198	15	3418	2264	532	622	-
MEGAN	prior	555	574	-3287	-204	-5	3527	2179	613	618	116
	posterior	555	576	-3199	-203	15	3422	2160	529	617	116
CB4prod	prior	555	654	-2723	-198	52	3037	1806	613	618	-
	posterior	555	584	-2504	-189	20	2722	1611	559	552	-

emissions and loss terms were virtually unchanged. However, the source attribution had shifted, with reductions in the biomass burning source, while the fossil fuel emissions and secondary production were slightly increased. As noted in Section 4.3.1, this shift in emissions away from biomass burning was likely linked to the coarse model resolution rather than an actual signal from the observations. The increase in secondary CO production of around $100\,\text{Tg CO yr}^{-1}$ could be related to aliasing to the direct natural CO emission sources, which were neglected in the prior of the *MOGOH* inversion. The impact of direct natural emissions will be further investigated in Section 4.3.3.

Instead of increasing the loss, another way to close the budget would be to reduce the emissions. In Section 4.3.4 below, this is investigated further by, instead of changing the OH, reverting the CO production from VOCs and CH_4 to an older version.

Overall, the budget terms from the *MOGOH* inversion were not in line with previous estimates found in the literature anymore, indicating that the addition of the TM5-MP-OH field might balance too high chemical production by introducing too high chemical loss. The posterior budget terms for the *reference* inversion, however, agreed fairly well with previous estimates. STAVRAKOU & MÜLLER [2006] find a total posterior CO source of around $2900\,\text{Tg CO yr}^{-1}$ in an inversion of MOPITT CO satellite observations for May 2000 to April 2001 using online OH. Their estimate is only around $7\,\%$ larger than the estimate from the Spiv-OH based *reference* inversion experiment presented here, but more than $15\,\%$ smaller than the TM5-MP-OH based *MOGOH* experiment. The inversion by KOPACZ et al. [2010] finds an only slightly lower total posterior source of $2630\,\text{Tg CO yr}^{-1}$ compared to the *reference* inversion, however, their source attribution is different with a $22\,\%$ lower secondary CO source ($1280\,\text{Tg CO yr}^{-1}$), balanced by $27\,\%$ more direct CO emissions ($1350\,\text{Tg CO yr}^{-1}$). Similarly, the CO inversion by FORTEMS-CHEINEY et al. [2011] finds even higher posterior total direct emissions of $1430\,\text{Tg CO yr}^{-1}$ for the years 2000 to 2010. In contrast, SHINDELL et al. [2006] report for the forward models analyzed in their multi-model comparison a larger mean secondary CO production of $1505\,\text{Tg CO yr}^{-1}$, which is close ($< 10\,\%$ difference) to the Spiv-OH based estimate from the *reference* inversion. Lower total emissions are found by JIANG et al. [2017], with a total CO source of only around $2400\,\text{Tg CO yr}^{-1}$, however, they also use a modeled OH field with a relatively low mass-weighted global annual tropospheric mean of $9.9 \times 10^5\,\frac{\text{molec}}{\text{cm}^3}$, which emphasizes the impact boundary condition choices can have.

4.3.2.2   Comparison to NOAA surface station measurements

The first four rows in Table 4.2 show the mean error-weighted prior and posterior mismatches to the flask measurements for each inversion, for all stations globally (first row) and for three subsets, split by zonal bands (second to fourth row). The individual mismatches for each station can be found in the appendix in Figure A.22 and Table A.3.

Considering all stations globally, the prior fit was worse for the *MOGOH* inversion compared to the *reference* inversion, however, the posterior fit was slightly improved. The worse prior fit originated from the underestimated CO and overestimated OH in the northern hemisphere, as discussed at length in Section 4.2.1. This bias in the northern hemisphere can be seen clearly when comparing the mean prior mismatches only for the stations in the northern extratropics shown in the second row of Table 4.2 to the other two regions. Both for the northern tropics and the southern hemisphere, prior fits were vastly improved with the TM5-MP-OH field. This is notable for the northern tropics, which were captured poorly with the Spiv-OH as detailed in Chapter 3. While the posterior fit was improved globally for the *MOGOH* inversion compared to the *reference* inversion, it was degraded slightly in the northern tropics. This could be traced mostly to the station on the Mariana Islands, Guam (GMI), which was overall captured poorly by the model. Notably, with TM5-MP-OH, even the prior fit in the southern hemisphere was better than the posterior fit in either of the northern regions.

To provide a closer look at the temporal evolution of the mismatches to the stations, six stations have been picked as examples. In Figure 4.1, for each of those stations, the measured CO mixing ratio and the modeled prior and posterior mixing ratios are shown for each of the inversion experiments.

Figure 4.1a represents Alert station in northern Canada as an example for high northern latitudes. The *reference* prior was too large during northern hemispheric summer compared to the measurements. In contrast, the *MOGOH* prior was too low over the whole 7 month period, but especially during northern hemispheric winter. As expected, the CO low bias from the forward model [MYRIOKEFALITAKIS et al., 2020] was reproduced in the prior. The posterior mixing ratios were similar and neither experiment manages to capture the low mixing ratios measured in July 2018 properly.

As an example for northern mid latitudes, Figure 4.1b shows Terceira Island in the North Atlantic. There, the unbalanced prior budget in the *reference* is clearly visible, with CO accumulating over the period. As for Alert station, the *MOGOH* prior underestimated the measurements throughout the inversion period, but by a significantly smaller margin.

Figures 4.1c and 4.1d represent Mauna Loa station on Hawaii and Bukit Kototabang in Indonesia, as examples for the northern tropics up to the Equator. For both, the *reference* prior strongly overestimated the measurements (note the larger mixing ratio range compare to the other panels in Figure 4.1), while the *MOGOH* prior closely followed the measurements and only minor corrections were needed to reach the best fit posterior.

The southern tropics and high southern latitudes are represented by Tutuila station in the Central Pacific in Figure 4.1e and Palmer station on Antarctica in Figure 4.1f, respectively. Similarly to the northern tropics, in the southern hemisphere the *reference*
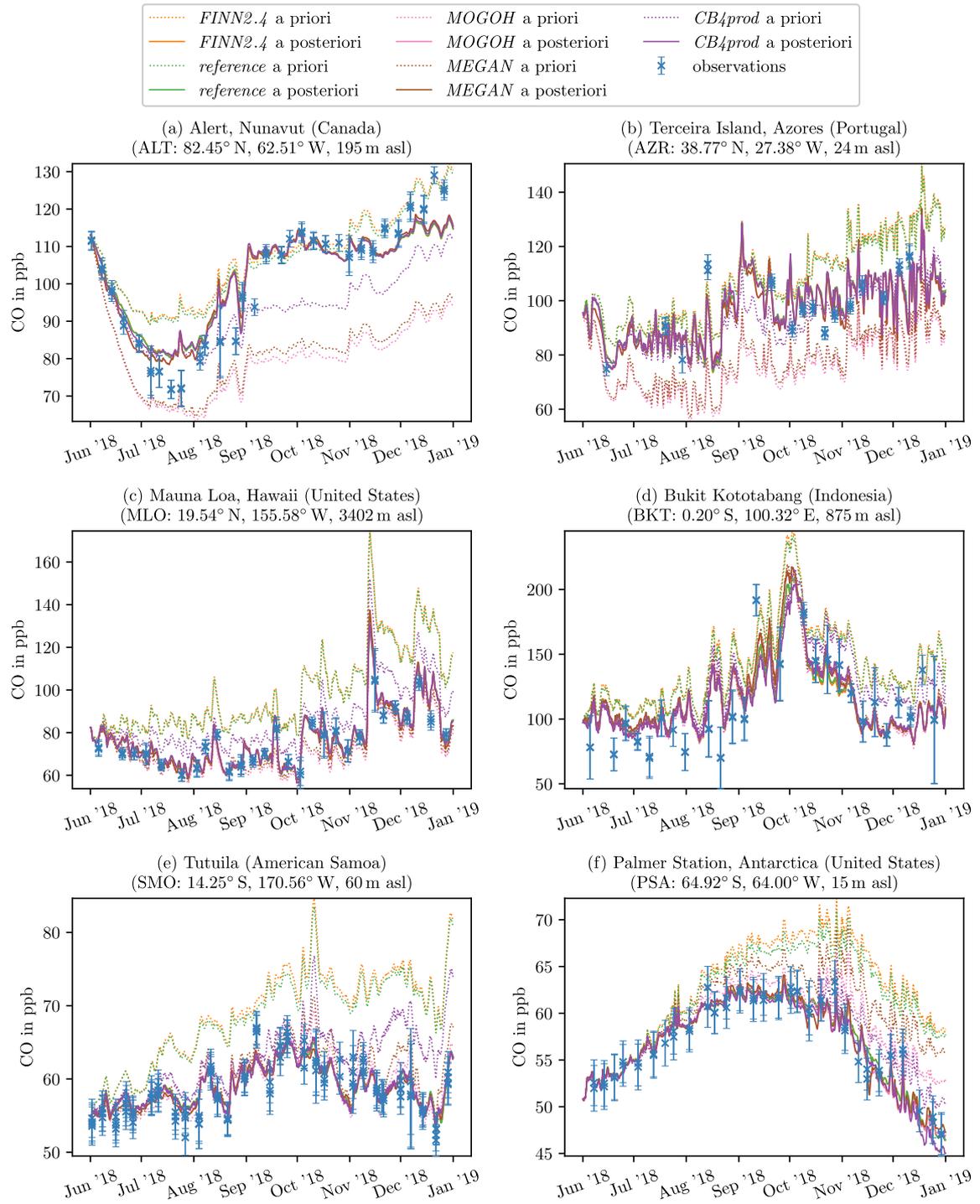
Figure 4.1: Modeled a priori (dotted lines) and a posteriori (solid lines) mixing ratios sampled at the locations of the stations as well the flask observations (blue crosses) for 6 example stations for each of the inversion experiments. For each observation, the corresponding measurement error is indicated as well.

prior continued to overestimate the measurements. While the *MOGOH* prior still improved over this, it also tended to slightly overestimate mixing ratios compared to the measurements, a trend that increased towards the South Pole.

Overall the *MOGOH* prior maintained the too low north-south gradient in CO of TM5-MP. Regardless, based on either prior the model reached an optimized emission state where the measurements were fit with similar quality. However, the emissions that were required to reach that state were vastly different, as detailed in Section 4.3.2.4 below.

4.3.2.3 Comparison to TROPOMI satellite observations

The final row in Table 4.2 shows the global prior and posterior mismatches to the TROPOMI satellite observations. While the prior mismatch was vastly improved for the *MOGOH* inversion compared to the *reference* inversion, the changes in the posterior fit were minor. The improvements in both the prior and posterior fit again originated mostly from the northern tropics to mid-latitudes. However, the prior fit was worse with TM5-MP-OH compared to Spiv-OH at high northern latitudes ($> 50\,^\circ$), similar to what was found for the surface measurements in the previous section. With respect to its temporal evolution, the prior mismatch no longer degraded strongly over time when using TM5-MP-OH, due to the budget being no longer unbalanced (Section 4.3.2.1). However, it still spiked in September, during the peak of the biomass burning season, as discussed in Chapter 3. In the appendix, the zonally (Figures A.22 and A.23) and monthly (Figures A.24 and A.25) averaged satellite-model mismatches can be found, to provide further details.

4.3.2.4 Emission increments

The principal results of the presented inversion experiments are emission increments. Analyzing those provides a deeper insight into how and where the changed OH field affected the inversion than the budgets shown in Section 4.3.2.1.

Figure 4.2a shows the relative increments in secondary CO production, i.e. the optimized posterior emissions divided by the a priori emissions, for the *reference* inversion for September 2018. In Section 4.3.1 above, strong similarities were found between the secondary CO production increments from the *reference* inversion and the ones from the *FINN2.4* inversion, which were analyzed in detail in Chapter 3. There, one very notable finding were the unrealistically large decrements in the northern tropics, in a band roughly between the Equator and 40° N, and especially for India and China. This finding is in contrast to what was found for the *MOGOH* inversion here. To illustrate the differences, Figure 4.2b is the same as Figure 4.2a but for the *MOGOH* inversion, i.e. using TM5-MP-OH instead of Spiv-OH. Further, Figure 4.2c shows the relative
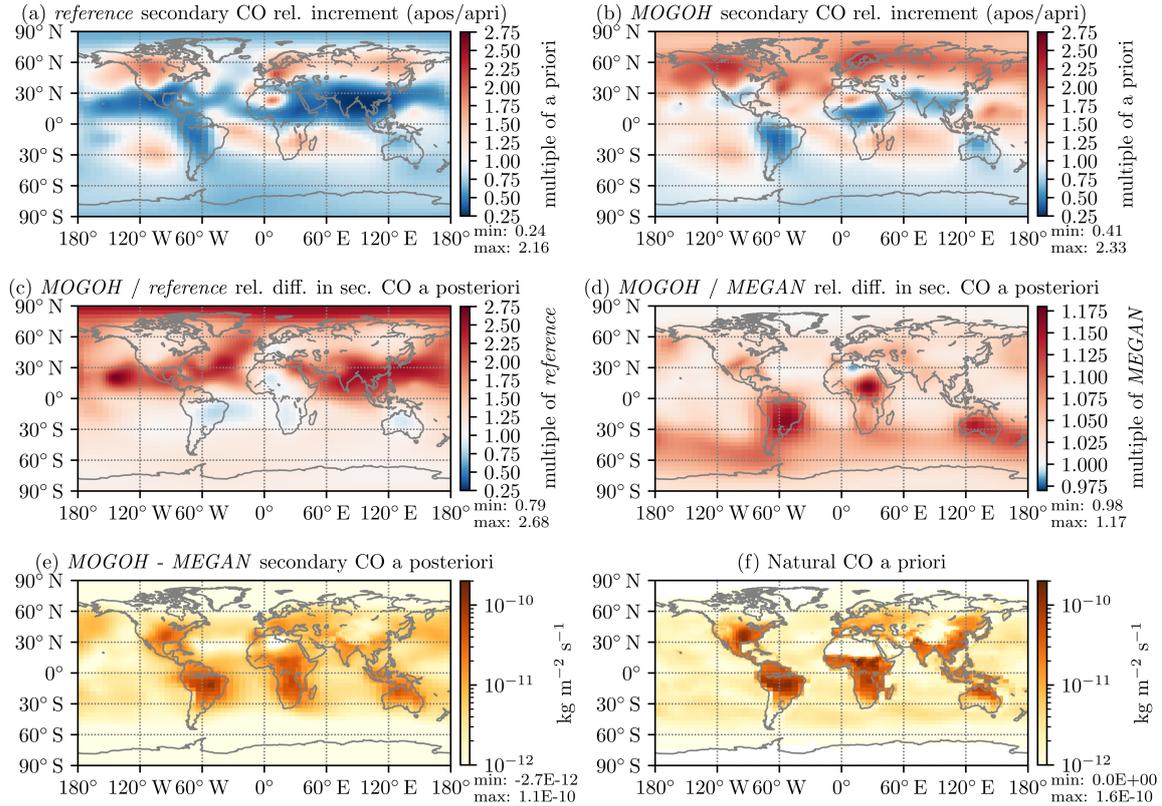
Figure 4.2: Global secondary CO production (a–e) and natural CO emissions (f) for September 2018. Panels (a) and (b) show the relative emission increments for the *reference* and *MOGOH* inversion, respectively. The differences originate from the different OH fields used. Panel (c) shows the relative differences between (a) and (b). Panel (d) shows the relative differences between the *MOGOH* and *MEGAN* inversion posteriors; note the significantly smaller range on the colorbar. Panel (e) is the absolute emission difference corresponding to Panel (d). For comparison, Panel (f) shows the natural CO emission a priori that was added in the *MEGAN* inversion compared to the *MOGOH* inversion.

difference between Figures 4.2a and 4.2b, which cancels out the influence of the prior. By comparison, the decrements in the northern tropics were much more moderate in the *MOGOH* inversion and could actually point to lowered anthropogenic emissions, as discussed in Chapter 3. Additionally, the increments compared to the a priori in the northern extratropics were larger in the *MOGOH* inversion, to make up for the CO low bias in TM5-MP in that region. Notably, the largest increments (by area and amplitude) were found over North America and northern Europe, potentially pointing to underestimated anthropogenic emissions in those regions [STEIN et al., 2014]. Secondary CO production in the southern hemisphere was mostly unchanged, with slightly less production over land and slightly more over the oceans.

4.3.3   Addition of natural CO emissions

Adding a dedicated emission category for direct CO emissions from natural sources, i.e. biogenic and ocean CO emissions, brought the TM5-4dvar model further in line with the emission setup of the full-chemistry forward model, TM5-MP. As can be seen in Table 4.3, this additional source left fossil fuel and biomass burning posterior emissions almost unchanged ($< 1\%$ change each) but slightly reduced the posterior secondary CO production (-4.6%), which roughly corresponded to the amount in Tg CO yr^{-1} added in natural emissions. However, the secondary CO production increment was now smaller, indicating a more balanced prior budget for the secondary CO. Overall, the total posterior emissions and loss were almost unchanged to *MOGOH* and only the attribution to the emission categories changed.

Notably, as shown in Table 4.2, the *MEGAN* inversion featured lower prior and posterior model-observation mismatches compared to the *MOGOH* inversion. In fact, the *MEGAN* inversion had the best posterior fit to the station measurements and satellite observations across all inversion experiments presented in this chapter. Compared to the *MOGOH* inversion, the a priori mismatches were improved in the northern hemisphere, but slightly worse in the southern hemisphere.

Figure 4.2d shows the relative difference between the optimized secondary CO production from the *MOGOH* and *MEGAN* inversions for September 2018. Note that the range of the colorbar is significantly narrower than in Figure 4.2c. The largest changes appeared to be in regions where natural CO emissions were expected. When comparing the absolute difference between the optimized secondary CO production from the *MOGOH* and *MEGAN* inversions in Figure 4.2e to the natural CO a priori in Figure 4.2f, striking similarities were found, even though the former appears more spread out.

4.3.4   Sensitivity to secondary CO production prior

Figures 4.3a and 4.3b show the relative secondary CO production increments for September 2018 for the *reference* and *CB4prod* inversions, respectively. Those two inversion experiments only differed by their secondary CO production prior, with one using 2018 rates from TM5-MP with MOGUNTIA chemistry and the other using 2006 rates from TM5-MP with CBM4 chemistry. Therefore, the relative increments were no longer directly comparable, since the respective posterior emissions were divided by different priors. To circumvent this, Figure 4.3d instead shows the relative difference between the posterior emissions of the two experiments. As can be seen, the *reference* and *CB4prod* inversions led to similar posterior emission fields, with relative differences close to 1 for most regions globally. The seemingly large differences in the remote southern extratropics occurred in regions with low absolute emissions. These
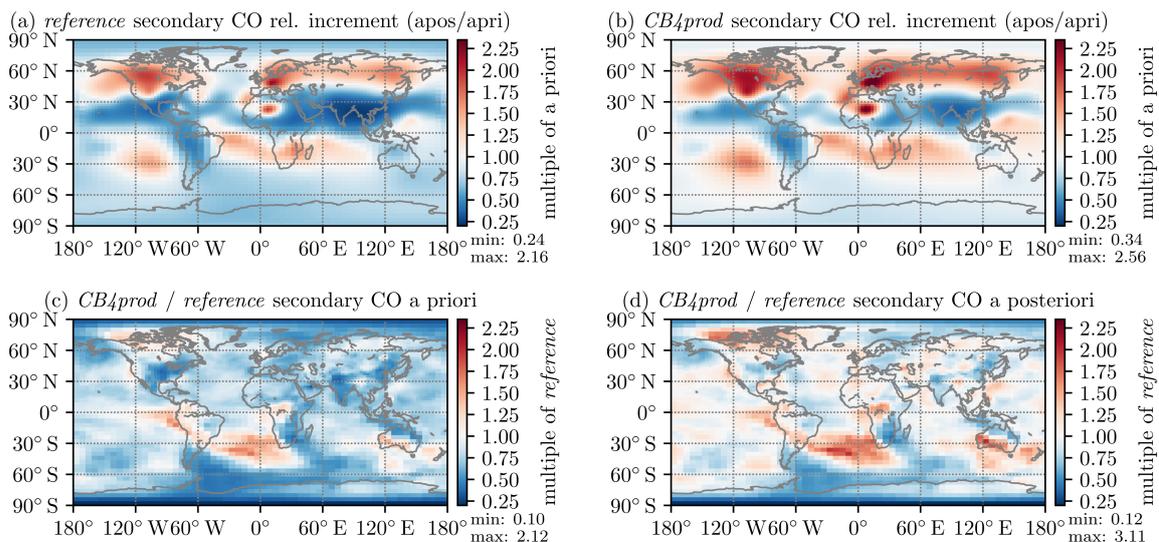
Figure 4.3: Relative increments of global secondary CO production for (a) the *reference* (same as Figure 4.2a, for convenience) and (b) the *CB4prod* inversion for September 2018. Note that these are relative to different priors and can, therefore, not be compared directly. Instead, Panels (c) and (d) to show the relative differences in the prior and posterior emissions, respectively, between the two inversions.

similarities in posterior emissions were as expected since the CO loss was driven by the same OH field, while the same atmospheric CO mixing ratios were well constrained by the observations, leaving little freedom for variations in the emissions. In contrast, Figure 4.2c shows a similar comparison between the *MOGOH* and *reference* inversions, which used different OH fields, and featured overall much larger differences.

Figure 4.3c shows the relative difference between the prior emissions of the *reference* and *CB4prod* inversions, again for September. Notably, the spatial distribution of the remaining differences in Figures 4.3d relatively closely follows the spatial distribution of the differences in those prior emissions. These similarities were expected, because the optimization operates by "scaling the priors", while the quadratic background term discourages large deviations. In other words, in terms of the cost function, it is very expensive for the optimizer to introduce large increments where prior emissions are low and it is cheaper to instead more moderately change emission over a larger area. Both inversion experiments lead to roughly the same total emissions needed to reach the observed atmospheric mixing ratios. For example, as shown in Figure 4.3c, TM5-MP with CBM4 chemistry predicts significantly lower VOC concentrations over India compared to the MOGUNTIA chemistry, which, as shown in Figure 4.3d, resulted in smaller posterior secondary CO production in *CB4prod* compared to *reference* for India, but slightly larger emissions in the surrounding regions. As for the OH fields presented in Section 4.2.1, a more detailed comparison of the secondary CO production fields can be found in the appendix, where Figures A.27 to A.30 show the zonal means of the

vertical distribution per season and Figures A.31 to A.33 show the lateral distribution for the surface layer per season.

The similarities in the posterior emissions are further emphasized in the budgets listed in Table 4.3. Compared to the *reference* inversion, the total emissions and losses ended up at similar values ($< 1\%$ difference) for the *CM4prod* inversion. Some minor differences exist in the attribution of emissions, with a shift away from secondary production and towards direct emissions, especially from fossil fuels. An overestimation of a posteriori secondary CO production in the *reference* experiment was expected, since the corresponding prior was high biased compared to the prescribed Spiv-OH field, as detailed in Section 4.3.2.1. While this bias was much smaller in the *CB4prod* experiment, its prior budget was still not closed and the model arrived at a too-high prior final mass. The *CB4prod* inversion then closed the budget by reducing emissions for all categories by around 10% on a global scale.

Regardless of the still slightly unbalanced prior budget, the setup for the *CB4prod* experiment led to the smallest global prior mismatch to the station measurements, as shown in the final column of Table 4.2. However, it also led to the overall worst posterior fit to the station measurements and satellite observations across all inversion experiments presented in this chapter. Regionally, there were large differences between the setups.

In the zone north of $23°$ N, which contains almost half of the stations, the *CB4prod* prior performed only slightly worse than the *reference* prior. Consequently, the *CB4prod* prior performed much better there than those based on TM5-MP-OH. At the level of individual stations, the *CB4prod* prior performed well in the northern mid latitudes (e.g. Terceira Island station in Figure 4.1b). However, the *CB4prod* prior underestimated the annual cycle at high northern latitudes (e.g. Alert station in Figure 4.1a), with too large prior mixing ratios during northern hemispheric summer and too low mixing ratios during northern hemispheric winter.

In the northern tropics and southern hemisphere, the prior mismatches from the *CB4prod* experiment were only slightly larger compared to the TM5-MP-OH based experiments. In turn, in both regions the prior performance of the *CB4prod* setup was much better than that of the *reference* setup. At the level of individual stations, this translated to generally mediocre prior performance for most of the rest of the globe (e.g. Figures 4.1c–e). By virtue of the lowest prior emissions, the *CB4prod* experiment had the closest prior fit to the stations close to the South Pole (e.g. Palmer station in Figure 4.1f), however, that could be coincidental.

The mismatches to the satellite observations, as shown in the appendix in Figures A.23 to A.26, featured very similar spatial prior and posterior distributions as the stations. Therefore, they are not analyzed in more detail here.

4.4   Conclusions

A series of inversion experiments driven by global satellite observations and surface measurements was conducted to investigate the impact of various boundary conditions on the optimized CO emissions for the second half of 2018. In the initial setup as used in Chapter 3, the prior budget was imbalanced, resulting in unreasonably large posterior emission decrements in the northern tropics. This imbalance was likely caused by inconsistencies between the OH climatology, which dominates the CO loss, and the secondary CO production terms taken from a CTM, which constitutes the largest CO source. Therefore, in this chapter, a special focus was set on the prescribed OH distributions and balancing the prior budget to obtain more reliable emission estimates.

By replacing the OH climatology with OH fields from the same CTM as the secondary CO production terms, a nearly closed prior budget was achieved. The a priori atmospheric total CO masses were close to those expected based on the observations. The results were further improved, and aliasing to the secondary CO production was reduced, by treating direct natural CO emissions from biogenic sources and the oceans explicitly. Notably, the updated OH fields lead to significantly more moderate posterior increments in the northern tropics compared to the inversions presented in Chapter 3. In the literature, the uncertainties on the global distribution of OH, and with that on the individual CO budget terms, are large. Still, the results presented in this chapter suggest that the CTM derived OH fields and secondary CO production rates tend to be too high. If those fields are used as boundary conditions in inverse modeling, the prior and posterior total CO sources and total CO sinks will possibly be too high. Additionally, compared to observations, there are spatial biases in the model derived prior CO and OH mixing ratios. Generally, the north-south gradient in modeled CO tends to be too low, with especially too low mixing ratios in the northern hemisphere. In turn, the low bias in CO leads to a high bias in modeled OH in the northern extratropics. Since the TM5-4dvar model lacks feedback between CO and OH, this high bias will result in a high bias in the optimized CO emissions.

Instead of replacing the climatological OH fields, another attempt to close the budget was using a different set of secondary CO production rates previously used in this inverse modeling framework by NECHITA-BANDA et al. [2018], which are based on a 2006 simulation of an older version of the same CTM mentioned above. While this approach did not manage to entirely close the prior budget, it led to the lowest globally aggregated prior mismatches to the observations out of all the inversion experiments considered in this chapter. However, except for the far southern hemisphere, this does not hold regionally. Additionally, the posterior mismatches to the observations are notably larger compared to the other experiments, both globally and regionally. Still, regardless of the choice of secondary CO production terms, the inversions that used the

same OH climatology resulted in very similar posterior emissions. Similarly, the exact choice of biomass burning prior emissions had little impact on the inversion results. As for the model-based OH field, the climatological OH is still subject to the uncertainties in OH as found in previous literature. The climatological OH required fine-tuning (scaling) to come into agreement with MCF measurements and there is at best limited evidence supporting the 3D distribution suggested by the now over two decades old climatology. On the contrary, there are hints that it is, for example, likely too low close to the surface in the tropics.

Overall, all inversions arrived at similar posterior mixing ratios, which fit both the satellite observations and the station measurements reasonably well. On the one hand, this could show how well the TROPOMI satellite observations in conjunction with the NOAA surface flask measurements constrain global CO. On the other hand, it could also indicate that the a priori error and correlation settings were too loose and the model overfit the observations. The latter is subject of further investigation. First steps are presented in Chapter 5. Notably, the inversion experiments arrive at these similar posterior mixing ratio through different posterior emissions. Differences are especially large if different OH fields are used, which emphasizes that the choice of OH is very important when trying to optimize absolute CO emissions. However, OH is currently too poorly constrained and there is a need to update the OH distributions used in literature. Regardless, the results of this chapter allow for a deeper understanding of potential biases, which can serve as a basis for the interpretation of future higher resolved inversion experiments e.g. of emissions from biomass burning events.

Chapter 5

Adaptive Error inflation for high-resolution observations

Chapter 3 introduced TROPOMI satellite observations into the TM5-4dvar inverse modeling framework and proposed a more rigorous approach for estimating the satellite error inflation factor, which was historically estimated empirically in studies using the TM5-4dvar model. However, limitations of this new approach were also found and discussed. This chapter introduces an updated satellite retrieval product version and investigates a novel method for handling the observational error of fine observations in a coarse-resolution inverse model.

Section 5.1 introduces the problem, prior work in literature, and the aims of this chapter. Section 5.2 present the new adaptive inflation approach. Section 5.3 gives a short description of the model setup and the conducted inversion experiments. Section 5.4 presents the results, split into showcasing the latest version of the satellite observations and the implications from using the adaptive error inflation approach. Section 5.5 provides a short summary of the findings of this chapter.

5.1   Background and objective

Equation (2.11) introduced the cost function, which may be split into two parts, the background cost

$$J_\mathrm{b}(\vec{x}) = (\vec{x} - \vec{x}_\mathrm{A})^\mathrm{T} \mathbf{S}_\mathrm{A}^{-1} (\vec{x} - \vec{x}_\mathrm{A}) \tag{5.1}$$

caused by deviations from the prior, and the observational cost

$$J_\mathrm{O}(\vec{x}) = (\vec{y} - \mathbf{F}(\vec{x}))^\mathrm{T} \mathbf{S}_\mathrm{O}^{-1} (\vec{y} - \mathbf{F}(\vec{x})) \tag{5.2}$$

caused by mismatches between the model and the observations. At the end of Section 2.1.3, it was already discussed that the relative weighting between those two parts matters for the quality of the inversion result. In principle, the weighting within each part, i.e. from one observation to the next or one element of the prior to another, is granted by the respective error covariance matrices (ECM), as outlined at the beginning of Section 2.1.3. The inverse of the ratio of any pair of variances found on the diagonal of the ECM gives the relative weight of the corresponding pair of elements of the observational or state vector, while the error covariances, i.e. the off-diagonal terms, inform on how much the weight of one element depends on the values of each of the other elements.

However, the relative weighting between the observational cost and the background cost as a whole, additionally depends on the size of the respective vectors (i.e. number of observations and dimension of the state), the precision of the observations compared to the precision of the prior, and whether these precision estimates are sufficiently accurate. For example, if there are many observations or if their error is underestimated, the regularization through the a priori may be lost, which can lead to slow convergence, potentially non-unique solutions, or a solution that is no longer anchored in reality and could be non-physical due to overfitting.

The problem regarding the relative weighting between components of the cost function is even more pronounced in inversion experiments that are based on multiple observational datasets, such as the ones presented in the previous chapters. For those experiments, the observational cost may be subdivided further into components per instrument, e.g. a satellite cost and a station cost, which only include contributions from those respective datasets. To obtain meaningful results from those experiments, a proper relative weighting between all three parts of the cost function (the background, satellite, and station cost) is required. So long as the problem is limited to two components and both are over- or underestimated simultaneously, the inversion still works. With three parts, if the contribution from one part is overestimated, it will overwhelm both of the other parts, even if those are properly defined. Even with proper estimates for the observational and prior errors, problems can arise if one instrument provides significantly more observations than the other(s). An example of such a problematic setup can be found in KOPACZ et al. [2010], where observations from three different satellite instruments (MOPITT, SCIAMACHY, and AIRS) are used and the joint inversion is dominated by the AIRS observations due to their much larger count.

Issues like that are bound to become more prevalent considering the steady rise in observation density with recent instruments. For example, the TROPOMI observations used throughout this thesis (described in Section 3.3.2) feature a significant increase in observation density in comparison to the observations from the Measurements of Pollution in the Troposphere (MOPITT) instrument or the Infrared Atmospheric Sound-

ing Interferometer (IASI) instrument used previously in TM5-4dvar [Hooghiemstra et al., 2012a,b; Krol et al., 2013; Nechita-Banda et al., 2018]. The origin of this increase is twofold. For one, TROPOMI CO observations feature a spatial resolution of up to $7 \times 7\,\mathrm{km}^2$ [Veefkind et al., 2012], which is about 10 times higher than the resolution of MOPITT of up to about $22 \times 22\,\mathrm{km}^2$ [Drummond et al., 2010] and the spatial sampling of IASI of up to about $25 \times 25\,\mathrm{km}^2$ [Clerbaux et al., 2009]. Additionally, TROPOMI reaches global coverage within a day, similar to IASI, while MOPITT requires about five days to reach global coverage. A more in-depth comparison of the mentioned instruments can be found in Section 3.1. In addition to using lower density observations, the previous TM5-4dvar studies also limited the use of satellite observations to their respective region of interest. This limitation circumvents some of the issues mentioned before. Overall, the contribution of the satellite is much smaller compared to using global data, which makes it less likely to overwhelm the other parts of the cost function. However, by limiting the satellite observations to the region of interest, potentially good and useful data are discarded.

For the experiments presented throughout this thesis, the much higher coverage of the TROPOMI satellite observations compared to the surface flask measurements would cause the satellite cost to overwhelm the station cost, if only the instrument errors were to be considered. In that case, the high confidence station measurements would no longer be able to constrain CO background mixing ratios. Hence, as described in Section 3.3.2.2, the standard deviations provided by the satellite retrieval are multiplied by inflation factors. These inflation factors are intended to bring the cost contribution of the high-resolution observations back in line with the rest of the cost function. Mathematically, the inflation is supposed to capture those parts of the observational error covariance matrix $\mathbf{S}_O$ (see Eq. (2.12)) that are not caused by instrument or retrieval errors (i.e. $\mathbf{S}_I$), but by the model representation errors (i.e. $\mathbf{S}_R$). For high density observations, the model representation errors are, in part, caused by correlations within the instrument and retrieval errors, which are often neglected in 4DVAR inverse modeling, i.e. the off-diagonal terms of $\mathbf{S}_O$ are assumed to be zero, so that each observation can be handled on its own. For example, Hilton et al. [2009] describe inverse modeling setups to optimize atmospheric tracer concentrations based on IASI radiance measurements, where the instrument error is inflated to account for error correlations, which, in their case, are mostly caused by uncertainties in atmospheric water vapor content. Additionally, the model representation errors also represent the limits to the information content the model can gain from the observations, e.g. the limited resolution of the model may not be able to resolve the finest features within the observations.

The method of using a single constant global inflation factor, as described in Section 3.3.2.2, has multiple shortcomings, some of which were already discussed in Chap-

ter 3. Compared to the TM5-4dvar studies based on MOPITT and IASI observations within their respective region of interest mentioned above, using global TROPOMI observations requires much larger error inflation factors to reach convergence. This issue persists even if super-observations at a coarser resolution (e.g. 0.5° × 0.5°, see Section 3.3.2.1) are used. These very large inflation factors imply that observations in regions with sparse coverage (e.g. over the oceans) have next to no weight in the cost function. The same applies to observations in regions with detailed structures and spikes from potentially interesting events (e.g. biomass burning), which are smoothed towards the background. Additionally, the very low weight of individual satellite data points could cause overfitting of the station measurements. Such issues of large inflation factors would be exacerbated with improving observational data, because, in the current setup, a lower observational error or improved coverage would result in an increase in the inflation factor, i.e. better data gets a lower weight. Note, however, that only improving the resolution of the observations would also imply an increase in autocorrelation within them, in which case their information content does not necessarily grow and an increase in inflation could be justified. Another issue of the approach outlined in Section 3.3.2.2 is that the inflation factor for a specific setup is not trivial to find, but requires a process that is in itself iterative and, therefore, computationally expensive. Additionally, each setup having its own inflation factor makes their results potentially harder to compare directly.

As pointed out in Liu & Rabier [2003], increasing the observation density beyond a certain threshold is not meaningful, especially if error correlations are ignored. However, there are approaches beyond variance inflation that may still make use of high density observations. One such approach are super-observations, see Section 2.3.2, where several observations that are in close proximity spatially, are aggregated into a single observation with a reduced aggregated error. An implementation of a super-observation approach based on the works of Eskes et al. [2003], Miyazaki et al. [2012] and Boersma et al. [2016] can be found in Section 3.3.2.1. Another widely used approach to treating high density observations is observation thinning. As described in Lahoz et al. [2007], sufficiently thinned data can be assumed to be uncorrelated, however, information on small scale structures within the data will also be lost. To overcome this shortcoming, Simonin et al. [2019] suggest an approach where the full error statistics are implemented to properly account for error correlations. Their approach allows for preservation of the full quantity of observations and, therefore, their full information content and impact on the forecast.

In this chapter, a similar approach is implemented. The approach had to be modified, because for 4DVAR inversions the full construction of the error covariance matrices is usually not feasible. Instead, individual error inflation factors for each satellite observation will be defined based on the local observation density. This adaptive error

inflation approach is possible as a pre-processing step and thereby computationally cheap compared to the rest of the inversion. As opposed to the super-observation approach described in Section 3.3.2.1, this new approach may retain any information contained in the full resolution TROPOMI observations. The approach also adds an opportunity to tune the sensitivity of the inversion towards potentially interesting patterns in the observations. In the wake of changing the handling of the observational error, the TROPOMI observations (see Section 3.3.2) are also updated to the latest TROPOMI/WFMD product, version 1.8 [SCHNEISING et al., 2023].

5.2   Adaptive error inflation

If no error correlations are assumed for the observations, the off-diagonal terms of the observational error covariance matrix $\mathbf{S}_\mathrm{O}$ are zero and only the diagonal terms remain, which are the variances σ_O^2 of the individual observations. Under this assumption, the observational part of cost function (Eq. (2.11)) may be simplified. For the inflation approach outlined in Section 3.3.2.2, the satellite cost is then

$$J_{\mathrm{O,sat}}(\vec{x}, M) = \sum_{i=1}^{M} \frac{(F_i(\vec{x}) - y_i)^2}{I^2 \sigma_\mathrm{O}^2} \tag{5.3}$$

with constant I as the inflation factor for the standard deviations, i as the index of each of the M observations y_i, and $F_i(\vec{x})$ as the model $\mathbf{F}(\vec{x})$ sampled for the ith observation, given a state $\vec{x}$. A constant inflation of $I = 1$ implies that the data are used as they are provided by the retrieval.

For the adaptive error inflation, the globally constant inflation I is now replaced by individual inflation factors I_i for each observation. Since observation thinning may be used to reduce error correlations, the observation density may be used as an indicator of those error correlations. Therefore, the adaptive error inflation factors I_i should depend on the number n_i of observations in 'close proximity' (to be defined later) to the ith observation. Furthermore, I_i should fulfill two basic conditions. First, increasing the number n_i of observations in 'close proximity' to the ith observation to n_i+1, should increase the inflation I_i so that

$$I_i(n_i) < I_i(n_i + 1) \tag{5.4}$$

i.e. each additional observation reduces the weight of each individual observation. Second, this reduction should happen in a way so that the value of the cost function still increases:

$$J_{\mathrm{O,sat}}(\vec{x}, n_i) < J_{\mathrm{O,sat}}(\vec{x}, n_i + 1) \tag{5.5}$$

i.e. each additional observation should still further constrain the inverse problem. For a constant inflation, as in Section 3.3.2.2, $J_{\mathrm{O,sat}}$ increases, on average, linearly with the number of observations, because each observation simply adds another independent term to the sum in Equation (5.3). As such, a constant inflation factor fulfills the second condition, but not the first, which causes the satellite cost to become very large for large numbers of observations and overwhelm the other parts of the cost function.

To achieve weighting by the observation density, the number n_i of observations in 'close proximity' to the observation y_i needs to be defined. The simplest approach would be to define a circle (or box) with a certain radius (or edge length) around the observation in question and count the number of observations enclosed within. While such an approach is easy to implement and computationally cheap, it has distinct disadvantages. The clear cut-off at a certain distance implies that any observations just beyond that distance get no weight at all. Additionally, since all observations before the cut-off get assigned the same weight (of one), their distribution is not taken into account, i.e. an observation with many observations close by does not suffer more from error correlation than one with the same number of observations just before the cut-off. A way to circumvent the issues introduced by the cut-off is to instead employ a weighing function w that depends on the actual distance between each of the observations and integrate that function over all observations. Such an approach has been used in SIMONIN et al. [2019] who use an exponentially decaying weighting function $w_{i,j} = \exp\left(-\frac{d_{i,j}}{L}\right)$ to estimate the correlations between observations, with $d_{i,j}$ as the distance between two observations y_i and y_j, and L as e-folding length. This e-folding or de-correlation length L is roughly equivalent to the radius of the circle in the simple approach outlined before. In fact, given a sufficiently finely resolved and evenly distributed field of observations, both methods would arrive at the same weighting. To avoid the spike-like structure of the exponential function for very short distances d and to promote a clearer cut-off at longer distances, here, a Gaussian function is used as the weighting function instead:

$$w_{i,j} = \exp\left(-\frac{d_{i,j}^2}{2L^2}\right) \tag{5.6}$$

At the resolutions of the experiments presented in this thesis, the de-correlation length L is dominated by the representativeness error of the model. Therefore, the de-correlation length will be estimated via the average spatial scale at which the model can resolve changes in the state. While the smallest scale at which the optimizer may introduce changes into the emission state is given by the 200 km spatial correlation length assigned to the biomass burning category (see Section 3.2.3.1), those emissions only make up a fraction of roughly 20 % of the total emissions. In contrast, the other two emission categories, fossil fuel combustion and secondary CO production from

VOCs and CH_4, were assigned much longer spatial correlation lengths of 2000 km and 1000 km, respectively. Since those two categories make up the remaining 80 % of the total emissions, a substantial amount of the changes the optimizer can potentially introduce will happen on those much longer spatial scales. The average scale L at which the state changes may be computed from the spatial correlation length L_{cat}, the a priori error σ_{cat}, and the fraction of the total emissions f_{cat} for each category as

$$L = \frac{\sum_{cat} L_{cat}\sigma_{cat}f_{cat}}{\sum_{cat} \sigma_{cat}f_{cat}} \approx 600 \, km \tag{5.7}$$

where both sums go over the three CO emission categories, biomass burning, fossil fuel, and secondary CO production. Based on the budgets presented in the previous chapters, e.g. in Table 3.3, the fractions of the total emissions were estimated as 20 %, 20 %, and 60 %, respectively. For the sake of simplicity, this source apportionment and the resulting de-correlation length was used globally. However, in a future development of this method, it might make sense to consider the local source apportionment at the location of the ith observation instead, because close to a biomass burning event the scale at which the optimizer can introduce changes into the state is shorter than over the oceans.

To minimize border effects when evaluating the weighing functions $w_{i,j}$ for a given observation y_i, all observations y_j within the same orbit and within the preceding and following orbits are considered. While close to the poles even observations from further orbits could be close spatially, they are at least 3 hours away in time, which is long enough for transport and chemistry to sufficiently reduce error correlations. The sum over all those weighting functions gives an approximation for the distance-weighted number of observations in 'close proximity' to the observation y_i:

$$n_i = \sum_{j=1}^{M'} w_{i,j} \tag{5.8}$$

where M' is the total number of observations in the preceding, current, and following orbit.

In the version of the adaptive error inflation applied below, for each observation, this distance-weighted observation count is used directly as the variance inflation factors $I_i^2 = n_i$. While this approach successfully reduces the satellite cost to no longer overwhelm the other parts of the cost function, it should be noted that it only weakly fulfills the second condition stated in Equation (5.5), i.e. especially at short distances, additional observations only marginally increase the average cost. If necessary in the future, fulfillment of the second condition could be more strongly ensured by applying some function that grows slower than a linear function, e.g. a logarithm or a root, to

the n_i, instead of using them as inflation factors directly. For example, $I_i^2 = \sqrt{n_i}$ would strongly fulfill both conditions stated above. However, if such a square root approach was used with the n_i calculated as described above, it would also lead to overall too low inflation factors and the satellite cost would overwhelm the station cost again.

5.3   Materials and methods

The model code used for the experiments presented below is mostly identical to the version described in Section 4.2, with only minor changes needed to handle the updated TROPOMI retrieval version and the observation-wise error inflation factors. The TM5-4dvar model itself and the basic inversion settings are as described in detail in Sections 2.2 and 3.2, respectively. The updated TROPOMI observations were also gridded to super-observations using the method described in Section 3.3.2.1. Aside from the updated observations and changed inflation factors, the only other change in the basic setup is that the biomass burning a priori emissions were fully changed to FINNv2.5 (from FINNv2.4, see Section 4.2 for details) for all inversion experiments. Unlike for the *FINN2.5* inversion experiment from the previous chapter, the spin-up inversion was also rerun with FINNv2.5. This change was made for the sake of consistency, despite the small impact on the inversion results caused by the differences between FINNv2.4 and FINNv2.5 as reported in Section 4.3.1.

Table 5.1: A priori emissions and observational setup for the experiments conducted in Chapter 5. All experiments used monthly climatological OH fields based on SPIVAKOVSKY et al. [2000].

Inversion	A priori emissions			Observations		Inflation
	biomass	anthrop.	secondary	satellite	flasks	factor
spin-up				v1.2, gridded	yes	45
sat. v1.2				v1.2, gridded	yes	72
full sat. v1.2	FINNv2.5 +VIIRS	CMIP6	TM5-MP MOGUNTIA	v1.2, full	yes	182
sat. only v1.2				v1.2, gridded	no	72
sat. v1.8				v1.8, gridded	yes	65
full sat. v1.8				v1.8, full	yes	165
sat. only v1.8				v1.8, gridded	no	65
adaptive				v1.8, full	yes	adaptive

Table 5.1 gives an overview of the experiments conducted for this chapter. Note the slightly smaller error inflation for the *sat. v1.8* compared to the *sat. v1.2* inversion. This discrepancy is caused by the slightly larger mean observational error $(+3\,\%)$ in the version 1.8 retrieval compared to version 1.2, in combination with a slightly reduced number of valid observations $(-2\,\%)$ during the period of the main inversions (June to December). Both the larger error and the lower observation count reduce the overall

satellite cost. Therefore, the inflation had to be adjusted down to fulfill the requirement introduced in Section 3.3.2.2 that the satellite and station costs should be equal. Effectively, the larger errors got compensated by a lower inflation, which illustrates one of the weaknesses of the previous inflation approach. Additionally, due to the different inflation factors, the results from those inversion are harder to compare.

5.4   Results

5.4.1   Retrieval version 1.2 versus 1.8

Optimizing emissions based on data from either retrieval version leads to very similar results. Therefore, it is not meaningful to shows the optimized emissions or the corresponding relative increments directly for each experiment, as done in the previous chapters. Relative secondary CO production increments are only included in Figure 5.1a for the *sat. v1.8* inversion for reference. For the comparison of the retrieval versions, instead, Figures 5.1b to 5.1d show the relative differences in posterior secondary CO production for September 2018 for the three pairs of inversion experiments, where the experimental setups in each pair only differ in using either data from product version 1.8 or 1.2, but are otherwise identical. This direct comparison is possible, because all inversion experiments presented in this chapter use the same prior setup. Note that all inversion experiments analyzed in the current section use the constant error inflation approach, as shown in Table 5.1.

Figure 5.1b compares the *sat. v1.2* and *sat. v1.8* inversions, i.e. the experiments that use gridded satellite observations in combination with surface flask measurements. Most differences globally are in the order of $\pm 1\%$. However, there are some more considerable differences in northern tropics, all of which are within the prior error assumed for the secondary CO production source category. The largest difference is an almost up to 20 % higher secondary production in West Africa obtained when using data from the updated satellite retrieval. This increment is accompanied by smaller positive increments stretching eastward from West Africa until north-west India. These increments in secondary CO production are compensated by lowered secondary production over the tropical Atlantic and the northern Middle East. While the increments may be attributed to changes in the retrieval process as outlined in Schneising et al. [2023], the decrements over the Atlantic, over which the satellite has poor coverage, are likely simply introduced by the optimization process to compensate. The budgets shown in Table A.4 in the appendix support this claim, as all terms for the *sat. v1.2* and *sat. v1.8* inversions agree within 1–2 ‰. The same strong agreement in budgets holds for the other two pairs of inversion experiments that only differ in the satellite product version and which will be analyzed in the following.
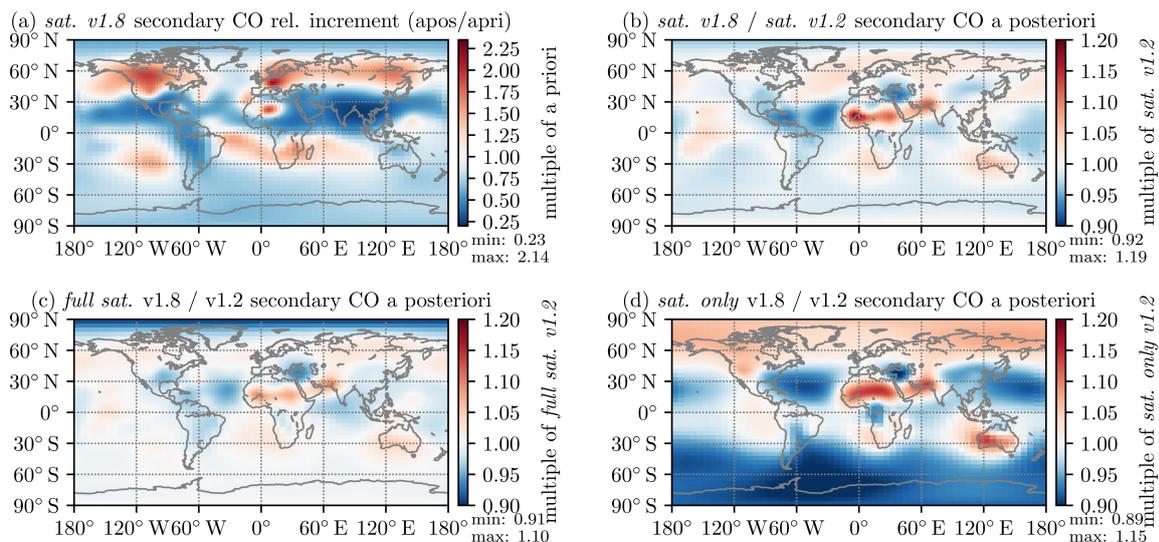
Figure 5.1: Comparison of the global posterior secondary CO production for September 2018 for inversion experiments based on data from different TROPOMI retrieval versions (1.2 and 1.8). Panel (a) shows the relative emission increments, i.e. the factor by which the emissions increased, for an inversion based on data from the latest WFMD satellite retrieval product (version 1.8). Panels (b) to (d) show the relative differences in posterior secondary CO production between using data from retrieval version 1.2 and version 1.8, when using (b) gridded satellite observations together with station measurements, (c) full resolution satellite observations and station measurements, and (d) only the gridded satellite observations and no station measurements. Note the different value ranges and asymmetry of the colorbars.

As Figure 5.1c shows, similar, albeit less pronounced patterns as for gridded satellite observations are found for the comparison of the *full sat. v1.2* and *full sat. v1.8* inversions, which use the full resolution satellite observations instead. As discussed in Section 3.3.2.2, using the full resolution satellite observations requires a very large inflation factor when applying the approach described there. This large inflation factor is required to get the satellite cost low enough to match the station cost, due to the very large number of observations. However, the large inflation factor also implies that each individual observation bears very little weight. Additionally, the model, due to its low spatial resolution, cannot capture the variability caused by the fine structures resolved by the observations. Consequently, in regions with high observational variability, the value of the cost function will be high with only a weak dependence on the exact model value, since there will always simultaneously be considerably over- and underestimated observations. In those cases, the best way for the optimizer to reduce the cost function is to reduce the station cost as much as possible, while mostly disregarding the satellite cost. Therefore, the minor differences in the different versions of the full resolution satellite products are translated more weakly into differences in the inversion results, compared to its gridded counterparts analyzed before.

Figure 5.1d reveals that the opposite holds for the comparison of the *sat. only v1.2* and *sat. only v1.8* inversions, which are driven only by the gridded satellite observa-

tions. Here, any difference in the retrieval versions is translated fully to the inversion result, because no other observations are considered that could direct the inversion to a different result. In the tropics, similar structures to Figure 5.1b can be seen, but more pronounced. Aside from the amplitude, these structures also differ slightly in lateral shape, which can be attributed to the absence of surface measurements to drive local emissions. Major differences between the *sat. only v1.2* and *sat. only v1.8* inversions occur at higher latitudes. In the northern hemisphere, north of $50°\,N$, widespread, but relatively low positive differences in the order of 5–$10\,\%$ can be found. The results presented here fall into northern hemispheric autumn. During that time, satellite coverage in that high-latitude region becomes increasingly sparse. Additionally, the updated satellite product features significantly fewer observations close to the North Pole due to improved quality filtering [SCHNEISING et al., 2023]. Therefore, these widespread differences are likely not constrained by any observations. Instead, they are introduced by the optimizer to resolve systematic mismatches between the model and the observations. Similarly, in the southern hemisphere, widespread decrements of a similar magnitude (note the asymmetric colorbar) are found. While sunlight is technically available for satellite soundings in this region, satellite coverage is still sparse due to most of the area being covered in dark oceans. Additionally, secondary CO production in this region is generally low. Therefore, these differences in the order of $-5\,\%$ to $-10\,\%$ have little impact on the budget and the comparison to the observations.

The first eight rows of Table A.5 list the mean error-weighted mismatch $\bar{J}_{\text{flask}}$ between flasks and model as introduced in Equation (3.9) for the seven inversion experiments presented in this chapter. For each pair of experiments that only differ by the version of the satellite product, the mismatches to the station measurements are slightly higher with the version 1.8 product. These larger mismatches are expected due to the lower satellite error inflation factors used. Furthermore, in line with the interpretations for the relative differences in posterior emissions analyzed above, the *full sat. v1.2* and *full sat. v1.8* inversions lead to the smallest mismatches and the *sat. only v1.2* and *sat. only v1.8* inversions lead to the largest.

The final two rows of Table A.5 provide the total error-weighted mismatch J_{sat} between satellite observations and model as introduced in Equation (3.10) for the seven inversion experiments presented in this chapter. As outlined in Section 3.4.2, it is necessary to consider the total mismatch rather than the mean mismatch, as done for the stations, because of the varying number of satellite observations between the experiments. The total mismatch for an observational dataset is equivalent to its contribution to the cost function. Overall, the total mismatch between the satellite observations and the model is larger for the prior when using version 1.8 of the satellite product compared to version 1.2. This larger total mismatch (i.e. cost) is, again, related to the lower inflation factors. However, the total posterior mismatch is lower

when using the updated satellite observations, which, notably, implies a larger prior over posterior cost ratio, i.e. the updated observations are more consistent with what the model can represent.

A more in-depth comparison of inversions based on different datasets (i.e. gridded satellite observations and stations versus gridded satellite observations alone versus full resolution satellite observations and stations) can be found in Section 3.4. The results for the TROPOMI product version 1.2 presented there still apply to the updated version 1.8 introduced into TM5-4dvar here.

5.4.2   Adaptive error inflation
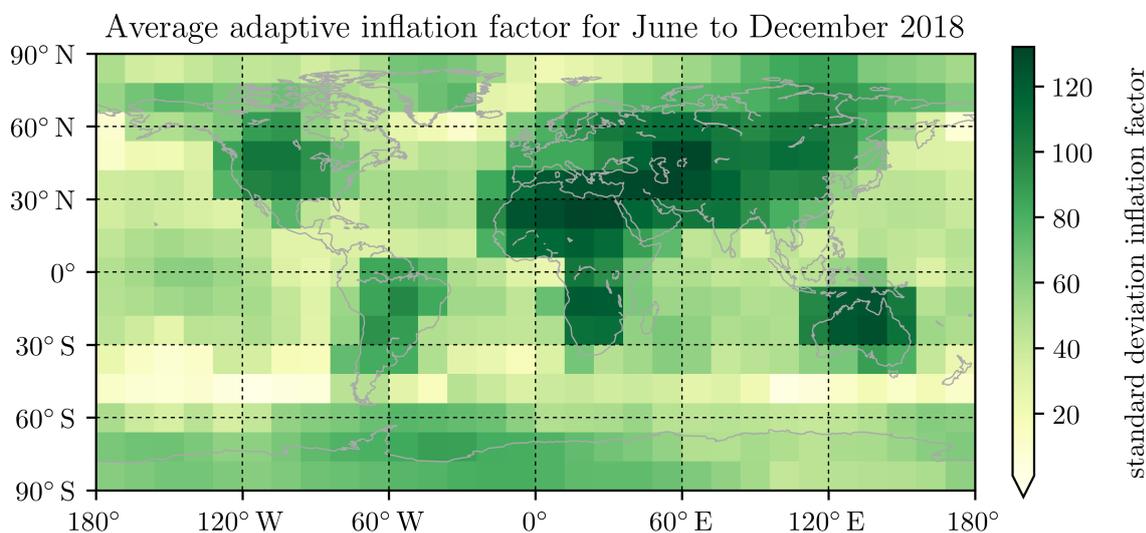
Figure 5.2: Adaptive error inflation factors aggregated in $12° \times 12°$ boxes and averaged over the inversion period (June to December 2018). For ease of comparison to the values reported throughout this work for the constant inflation approach, the inflation factors I_i for the standard deviation are shown, rather than the variance inflation factors I_i^2. Note the generally larger values over land and smaller values over the oceans.

As can be seen in Figure 5.2, the adaptive error inflation approach generally assigns a reduced weight (i.e. larger inflation factors) to the dense land-bound observations, but an increased weight to the sparse observations over the oceans, compared to the constant inflation approach as described in Section 3.3.2.2. In total, the posterior cost contribution from the satellite observations is more than 3.2 times larger than the station cost with the new adaptive approach. In contrast, the old constant inflation approach required that the satellite and station costs should be equal. Therefore, with the new approach, the satellite observations have a larger impact on the inversion results. Conversely, giving more weight to the satellite observations degrades how well the surface measurements are captured by the posterior state. This degradation can be seen in Table A.5, where the posterior error-weighted mean mismatches between

the station measurements and the model are 19 % larger for the *adaptive* inversion compared to the *sat. v1.8* inversion on a global scale. Notably, in the extratropical southern hemisphere (south of 23° S) the mismatch increases significantly less, by only 3 %. However, the impact of this smaller mismatch on the global mean is limited, because only about one quarter of the surface stations are located within that region. For the satellite observations, the total mismatches presented in Table A.5 are harder to compare directly, because the reduced overall inflation leads to a much larger satellite cost. Regardless, the relative improvement from the prior to the posterior mismatch is still around a factor of 10, similar to what the inversions based on the old inflation approach achieved.
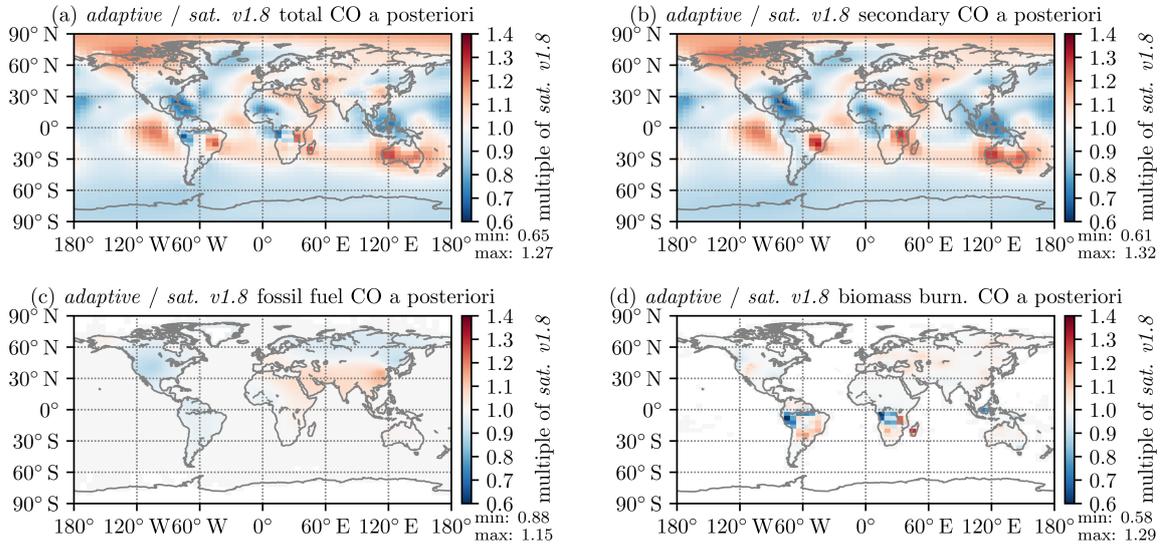
Figure 5.3: Relative differences in posterior CO source terms for September 2018 between two inversion experiments using different error inflation strategies. The compared source terms are (a) the total emissions, (b) secondary CO production from VOCs and CH_4, (c) direct CO emissions from burning fossil fuels, and (d) direct CO emissions from biomass burning. The *sat. v1.8* inversion uses the old globally constant inflation approach and the *adaptive* inversion uses the new adaptive inflation approach. The total and biomass burning emissions in Panels (a) and (d), respectively, are monthly mean emissions.

In Figure 5.3, the optimized emissions from the *adaptive* inversion are compared by source category to those from the *sat. v1.8* inversion, which uses the same setup except for the inflation approach. The changes in the lateral distribution of the total emissions are depicted in Figure 5.3a. While locally differences reach up to around $\pm 30\%$, they mostly stay well below 10 %. Additionally, as evident from the budget shown in Table A.4 in the appendix, at a global scale, the total emissions for the *adaptive* inversion are only less than 1 % lower compared to the *sat. v1.8* inversion. Still, some of the local differences are worth discussing in detail.

For example, the differences in the total emissions in Figure 5.3a contain dipole-like structures in South America and southern Africa. Considering the changes in

the secondary CO production and in the direct biomass burning CO emissions shown in Figures 5.3b and 5.3d, respectively, reveals a re-attribution from biomass burning emissions to secondary CO production in those regions. As can be seen in the budgets in Table A.4, on a global scale, this re-attribution leads to 3 % lower biomass burning CO emissions in the *adaptive* inversion compared to the *sat. v1.8* inversion. However, note that the *adaptive* inversion only reinforces strong local decrements in the biomass burning CO emissions in South America and southern Africa that were already found for the *sat. v1.8* inversion. The overall decrease in biomass burning emissions in either inversion experiment could be related to the model not being able to properly resolve biomass burning events due to its low resolution of only $6° \times 4°$ in those regions.

Another regional difference is found in Asia, where CO from the secondary production (Figure 5.3b) is shifted to direct anthropogenic emissions (Figure 5.3c). Combined, theses shifts result in a slightly larger CO source from East China and India (Figure 5.3a). This shift in source category could be a sign of aliasing. On a global scale, as the budgets in Table A.4 show, the anthropogenic CO emissions from fossil fuels are consequently increased by 2 % from the *sat. v1.8* to the *adaptive* inversion. Note that either inversion experiment still finds significant emission decrements in and around India and China, as reported in the previous chapters, and that these decrements are just marginally smaller for the *adaptive* inversion.

The comparison of the total emissions from the *sat. v1.8* and the *adaptive* inversions, as shown in Figure 5.3a, also features notable positive differences over Australia and the eastern tropical Pacific. These differences are attributed to changes in the secondary CO production by the optimizer, as shown in Figure 5.3b. However, in those regions, the *adaptive* inversion merely keeps the total emission increments closer to the prior, rather than it introducing or removing additional emissions. In total, on a global scale, there is almost no change (1 %) in the secondary CO production.

Combined, all these differences in CO emissions lead to only minor changes in the modeled mixing ratios at the locations of the surface stations. Overall, the mixing ratios from the *adaptive* inversion closely follow the *sat. v1.8* inversion. The largest deviations, in the order of a few ppb, occur for the stations that were found to be problematic in Chapter 3, i.e. those at high northern latitudes and mountainous stations. The deviations usually push the mixing ratios towards the *sat. only v1.8* inversions. This behavior is expected, because the satellite observations receive an overall higher weight in the cost function when the adaptive inflation approach is applied. However, the *adaptive* inversion still remains overall significantly closer to the *sat. v1.8* inversion than to the *sat. only v1.8* inversion, which signifies that the surface flask measurements are not overwhelmed by the satellite observations in the cost function.

5.5   Conclusions

In this chapter, the impact of using an updated satellite product version was analyzed and then, based on that new product, an extended satellite error inflation approach was introduced and tested. Changing the TROPOMI product version from 1.2 to 1.8 lead to very similar inversion results. Using version 1.8 observations produces an overall more consistent posterior fit between model, flask measurements, and satellite observations. On a global scale, the budget terms are almost identical. On a regional scale, moderate lateral shifts in emissions can be found in the northern tropics, which are attributable to changes in the retrieval process. Further differences occur at high latitudes due to the retrieval employing an updated cloud filtering algorithm. Furthermore, this analysis has shown that the translation of features in the observations into the inversion results can depend on which additional datasets are used.

The new adaptive satellite error inflation approach behaves as intended, by assigning more weight to stand-alone observations in remote regions, especially over the oceans, while de-weighting the highly correlated observations in densely covered regions. Notably, the new inflation approach is no longer an iterative process, but can be applied as a pre-processing step, which significantly reduces its impact on the computational time. Overall, the new approach increases the weight of the high-quality TROPOMI observations in the cost function without overwhelming the surface flask measurements. For the latter, their mismatches to the model are degraded only moderately, with almost no change in the southern hemisphere, where, according to Chapter 3, the TROPOMI satellite instrument and the surface stations are equally capable of constraining the emissions. Still, since increasing the weight of the satellite observations causes the mismatches of the surface measurements to the model to degrade, there appear to be biases between the satellite observations and the station measurements. In part, these biases originate from meteorology, as shown in Chapter 6. In the future, a bias correction scheme might be useful to obtain more consistent emission estimates.

For a more in-depth analysis of the merits and short-comings of the adaptive error inflation approach, a higher resolution inversion experiment, for example, one analyzing a specific biomass burning event, would be useful. However, as detailed in the previous chapters, several limitations of the inversion setup used throughout this thesis currently prevent such an application of TM5-4dvar. The most notable limitation is the discrepancy between the secondary CO production a priori fields and the climatological OH fields, as described in Chapter 4.

Furthermore, the adaptive error inflation approach, as described in the current chapter, only represents a basic concept, with several opportunities for further development. As mentioned at the end of Section 5.2, a more refined function to retrieve

the inflation factors from the number of observations in close proximity to each observation might improve results. Additionally, combining the adaptive inflation approach with super-observations could be useful to reduce the computational cost. Another future extension could be using the correlation in observation values in addition to their distance in the weighting function, which is used to count the number of close by observations. By giving higher weight when the correlation is low, potentially interesting patterns would be promoted. Additionally, including the observational error of the individual satellite soundings in the weighting function could maintain the impact of high-quality observations on the cost function, even if they happen to be surrounded by many low-quality observations.

Moreover, the error correlations of the satellite observations are expected to be partly flow-depended, since part of the retrieval relies on results from a chemical transport model. Therefore, the de-correlation length should also be flow-dependent. While not trivial to implement, such flow-depended weights would allow for a better separation of e.g. observations within a plume and the surrounding background.

Chapter 6

Benefits of using an updated meteorology

The inversion experiments in the previous chapters were driven by ERAI meteorology. In Chapter 3, biases in the optimized emissions around the locations of the surface stations, seemingly from overfitting of their measurements, were, in part, attributed to limitations in the representation of (vertical) transport in the TM5-4dvar model. This chapter investigates the efficacy of using the new ERA5 meteorology to drive the simulation of atmospheric transport, dry deposition, and chemistry rates for global inverse modeling experiments.

Section 6.1 introduces the topic and the meteorological fields in question. Section 6.2 shortly introduces the conducted inversion experiments. Section 6.3 showcases how the inversion results improve with the new meteorology, compensating for some of the biases found in the previous chapters. Section 6.4 contains the conclusions that can be drawn from this chapter for future inversion experiments.

6.1   Background and objective

As introduced in Section 3.2.1, TM5-4dvar is an offline model and as such does not generate its own meteorology. Instead, transport, temperature and pressure dependent chemistry, and dry deposition velocities are simulated based on preprocessed meteorological fields from a different model. Meteorological fields cover a wide range of variables describing the atmosphere and are not limited to precipitation, wind, temperature, and air pressure fields. They also include parameters describing humidity, clouds, atmospheric heat transport, and various auxiliary parameters pertaining to the surface, for example, for its orography, roughness, and vegetation.

Offline models can be used when the modeled processes do not impact the meteorology and have the advantage of a reduced computational cost compared to online

models, which need to calculate their own meteorology with each simulation. This advantage is especially relevant for 4DVAR inversions of atmospheric tracer emissions, where the same period is potentially simulated many times. However, to use an offline model, the complete set of preprocessed meteorological fields must be available beforehand, which requires large amounts of storage.

The specific model setup used throughout this thesis is capable of 'zooming' (see Section 3.2.1), a technique where different nested regions are simulated at different spatial and temporal resolutions. To minimize the required computation times and I/O, the meteorological fields are coarsened for each zooming region to its respective spatial resolution. This process needs to be done only once per region and time span, since the coarsened fields are then stored in intermediate files. These files are then reused in each iteration of the inversion, which further capitalizes on both the advantages (low run times) and disadvantages (high storage requirements) of offline models.

In the experiments presented in the previous chapters, the meteorological fields were taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA) project, specifically the ERA-Interim (ERAI) meteorology [Dee et al., 2011]. However, the ERAI meteorology has recently been superseded by the 5th generation ERA (ERA5) meteorology, which includes a wide range of improvements in meteorological modeling and data acquisition [Hersbach et al., 2020]. Compared to ERAI, ERA5 features increased spatial and temporal resolutions. With ERA5, the spatial resolution is improved to 31 km (from 79 km in ERAI), the vertical resolution is improved to 137 levels (from 60 in ERAI), and the temporal resolution is improved to up to hourly (from up to 3-hourly in ERAI). However, it should be noted that due to computational constraints the TM5-4dvar model used throughout this thesis is currently run at only up to $1° \times 1°$ (longitude $\times$ latitude) with 34 vertical layers and either meteorology data set is coarsened to that resolution. Consequently, the hourly temporal resolution ERA5 offers is also not required to accurately capture transport and coarsened to 3-hourly resolution as well. Regardless, unlike ERAI, ERA5 bears the possibility to improve the model resolution of TM5-4dvar beyond $1° \times 1°$ in the future. Another significant advantage of ERA5 over ERAI is that ERAI only has meteorology available up until August 2019 (starting 1979), whereas ERA5 receives continuous updates almost up until present day (starting 1950). ERA5 has a near real time (NRT) product that lags 5 days behind the present, while the final product has a lag of 2–3 months.

In the following, the impact on the inversion results of using the updated and upgraded ERA5 meteorology instead of the older ERAI meteorology will be investigated.

6.2   Materials and methods

For the evaluation of the efficacy of ERA5 meteorology in TM5-4dvar inversions, an additional inversion was performed. This additional inversion uses the same setup as the *sat. v1.2* inversion presented in the previous chapter (detailed in Section 5.3), but with the ERAI meteorology being replaced by ERA5. In short, the basic inversion settings are as described in detail in Section 3.2, with the code base updated to revision c71f31 of the official repository as described in Section 4.2, and using FINNv2.5 as the biomass burning a priori for both the spin-up and the main inversion as in the previous chapter. This new inversion experiment will henceforth be referred to as *ERA5*, while the *sat. v1.2* experiment from the previous chapter will be referred to as *ERAI* to make them easier to attribute to their respective meteorology.

Compared to the *ERAI* experiment, a noticeably larger satellite error inflation factor of 80 (up from 72) was required for the *ERA5* inversion to fulfill the inflation criterion detailed in Section 3.3.2.2. The change in inflation can be attributed to a larger relative improvement for the surface measurements than for the satellite observations in terms of the mismatches between the measurements and the concentrations that are modeled based on the posterior emissions. The origins of this discrepancy will be further explained in the next section.

6.3   Results

Figure 6.1 compares the posterior emission increments for the secondary CO production category for the *ERAI* and *ERA5* inversions. Overall, the increments are smoother for the *ERA5* inversion (Figure 6.1b) than for the *ERAI* inversion (Figure 6.1b), with a reduction in localized emission increments around the locations of certain surface stations. In the *ERAI* inversion, considerable increments in the secondary CO production can be found over the oceans upwind of these remote stations, e.g. Tudor Hill, Bermuda, (BMW) in the Atlantic, Mariana Islands, Guam, (GMI) in the Pacific, or Mahe Island, Seychelles, (SEY) in the Indian Ocean. There is no probable explanation why a sudden rise in either OH concentrations or CO precursors should occur in these remote regions with no sources nearby. Therefore, the *ERA5* inversion leads to more realistic results. As can be seen in Figures 6.1c and 6.1d, the *ERA5* inversion tends to shift the secondary CO production from remote regions towards land, where the CO precursor sources can be found. Another promising detail is the strongly reduced spike over Central Europe. This spike is driven mostly by Ochsenkopf (OXK) and Hohenpeissenberg (HPB) stations, both in Germany, and was previously (Chapter 3) linked to the limited representation of mountainous stations in the model. As shown
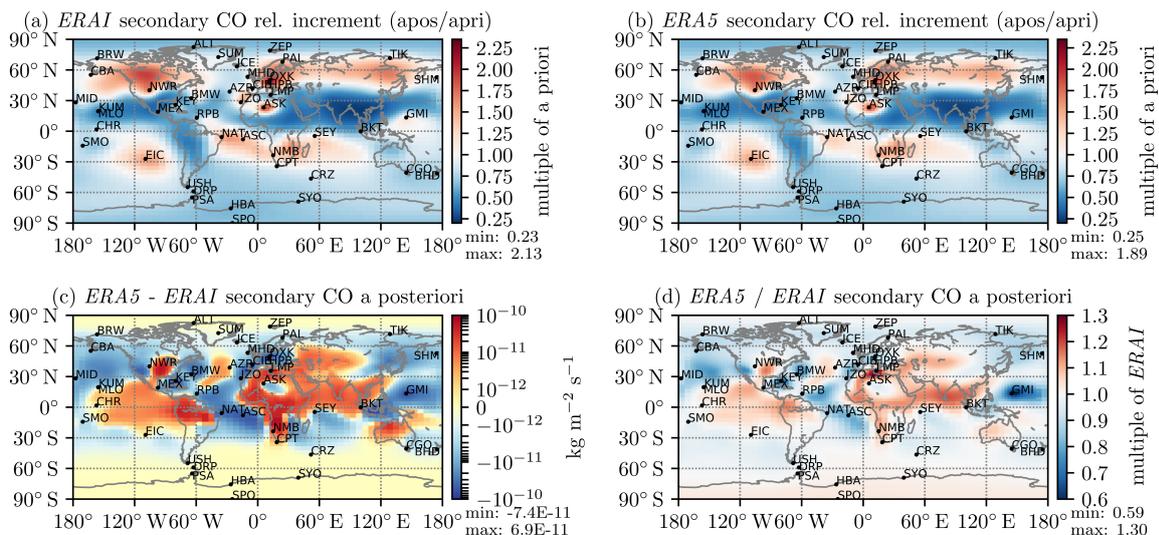
Figure 6.1: Comparison of the posterior global secondary CO production for September 2018 for two otherwise identical inversion experiments based on different meteorological datasets. Panels (a) and (b) show the relative emission increments, i.e. the factor by which the emissions increased, for the inversions based on ERAI and ERA5, respectively. Panels (c) and (d) show the absolute and relative differences between (a) and (b), respectively. The locations of the surface stations are indicated with dots and their station codes for easier orientation.

below, using ERA5 allows the TM5-4dvar model to capture those stations significantly better.

Figure 6.2 shows the measured and modeled (prior and posterior) mixing ratios at the locations of four example stations for the *ERAI* and *ERA5* inversions. For ease of comparison, the example stations shown here are a subset of the stations presented in Section 3.4.1. The number of stations was reduced, because the prior and posterior comparisons are very similar across the globe. The remaining stations are sufficient to represent the variability from the remote northern latitudes (Alert, Canada (ALT), Figure 6.2a), over the northern tropics (Mauna Loa, Hawaii (MLO), Figure 6.2b) and the Equator (Bukit Kototabang, Indonesia (BKT), Figure 6.2c), to the remote southern latitudes (Palmer Station, Antarctica (PSA), Figure 6.2d). The stations in Germany mentioned before (OXK, HPB) are not shown, because this representation is not meaningful for them due to their high temporal variability.

With ERA5 meteorology, the simulation based on the prior emissions leads to generally higher mixing ratios close to the surface, compared to the simulation based on ERAI meteorology. This effect is especially pronounced in the tropics (Figure 6.2c), where differences in the order of tens of ppb occur. Note that both experiments used an identical emission setup. Therefore, the prior differences must be caused either by differences in transport, in the reaction rate of CO with OH (via changes in temperature or pressure), or in dry deposition velocities (which depend on a slew of meteorological variables). Differences in the modeled mixing ratios based on the posterior emissions
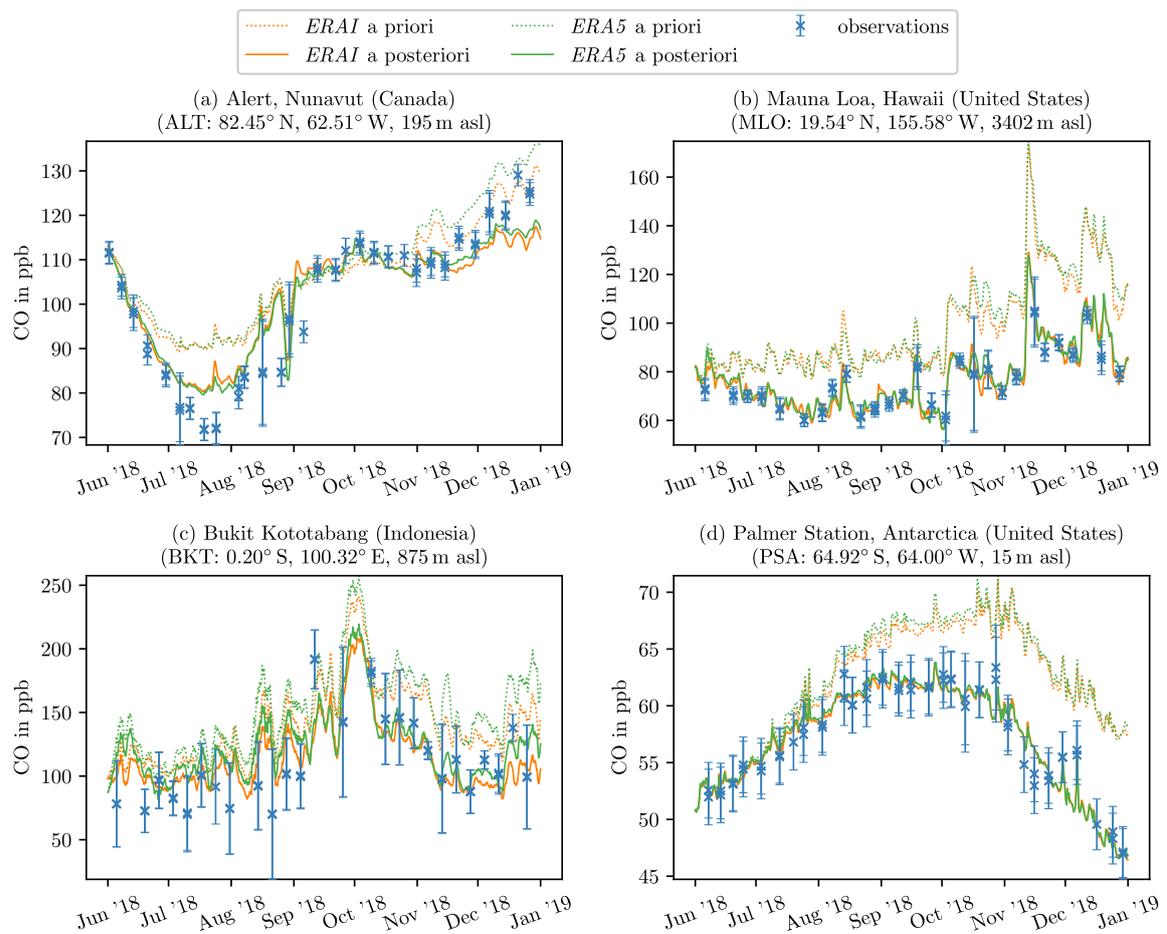
Figure 6.2: Modeled a priori (dotted lines) and a posteriori (solid lines) mixing ratios sampled at the locations of the stations as well the flask observations (blue crosses) for 4 example stations for two otherwise identical inversion experiments based on different meteorological datasets. The lines are color-coded with the ERAI-based inversion in orange and the ERA5-based inversion in green. For each observation, the corresponding measurement error is indicated as well. Note how for ERA5 the prior concentrations are usually slightly higher, while the posterior concentrations are similar, and how the largest deviations occur close to the Equator, as visible in Panel (c).

of the *ERAI* and *ERA5* inversions are noticeably more moderate and rarely exceed a few ppb. Yet, *ERA5* still leads to overall slightly higher mixing ratios close to the surface.

To further quantify the impact of these differences on the inversion results, the mean error-weighted mismatch $\bar{J}_{\text{flask}}$ between flasks and model as introduced in Equation (3.9) will be considered for the *ERAI* and *ERA5* inversions. As can be seen in Figure A.34 in the appendix, the prior mismatches between model and surface measurements is improved for some stations when using ERA5 instead of ERAI, while it is degraded for others. The numeric values displayed in Figure A.34 can be found in Table A.6 in the appendix. Notably, the prior fit is vastly improved for the previously very problematic OXK and HPB stations. Overall, the global mean error-weighted

prior mismatch between model and observations is slightly improved from 20.6 for
ERAI to 20.3 for *ERA5*. In contrast, the posterior mismatches are improved for all
stations expect TIK. On a global level, the mean posterior mismatch is noticeably
improved from 3.7 for *ERAI* to 2.7 for *ERA5*.
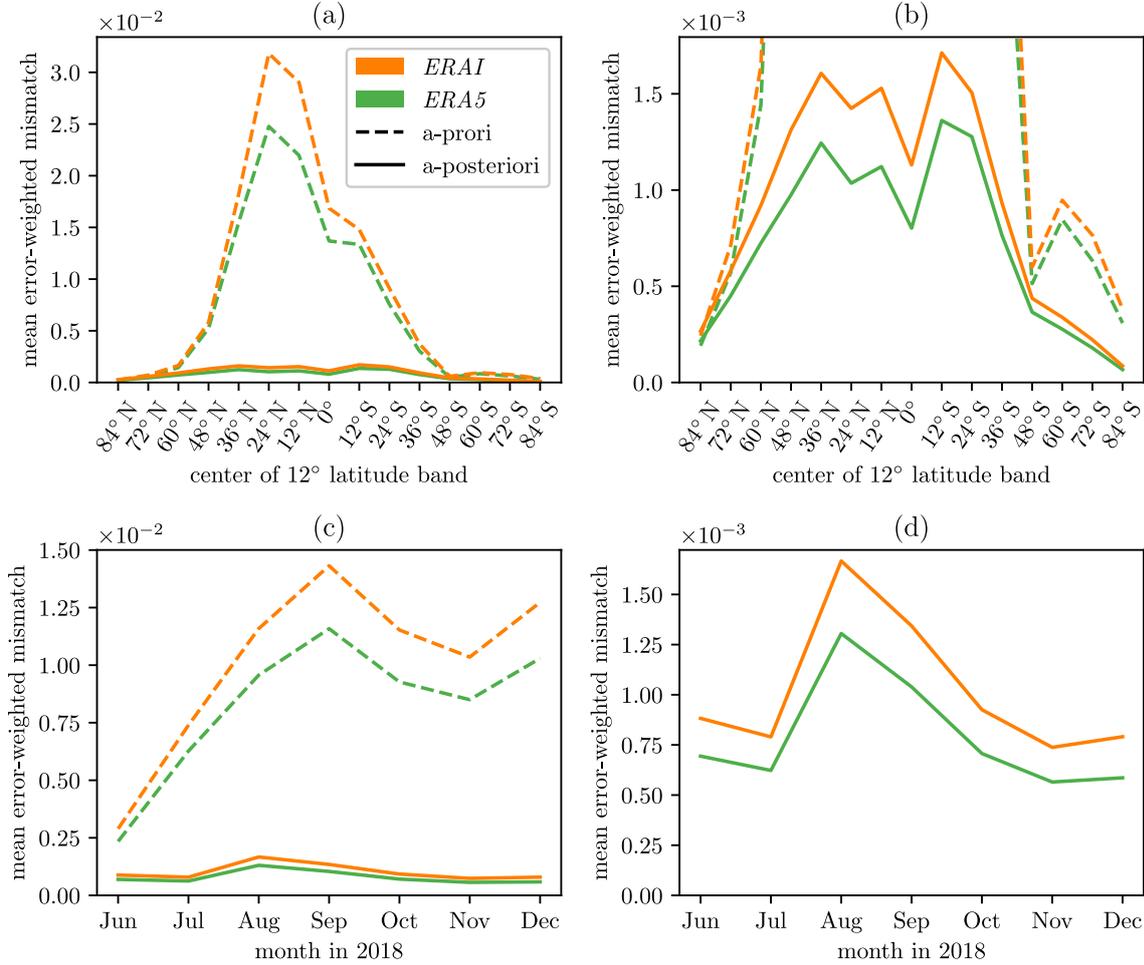
Figure 6.3: Mean error-weighted a priori (dashed lines) and a posteriori (solid lines) mis-
matches between the satellite and the model for two otherwise identical inversion experiments
based on different meteorological datasets. This metric is equivalent to the average contribu-
tion to the cost function of each satellite sounding. The values are very small, because the
large error inflation factors strongly reduce the weight of each individual observation. Panels
(a) and (b) show the mismatches over the whole inversion period aggregated in 12° zonal
bands. Note how the mismatches are largest in the tropics, but improved for ERA5. Panels
(c) and (d) show the global monthly mean mismatches. Note how the relative improvements
due to ERA5 are consistent over time. Panels (b) and (d) show the same data as Panels
(a) and (c), respectively, but with the vertical axis scaled only by the a posteriori graphs, to
make them more easily discernible.

In contrast to the stations, the mismatches to the satellite observations, calculated
using Equation 3.10, are already improved for the simulations based on the prior emis-
sions. This improvement is shown in Figures 6.3a and 6.3c, for the monthly global
mean mismatches and mean mismatches over the whole period divided by zonal bands,

respectively. As can be seen, the improvements from *ERAI* to *ERA5* are largest in the tropics, where accurate vertical transport is very important due to strong convection. Additionally, the improvement appears to be persistent over time. Overall, the global error-weighted mean prior mismatch is reduced by almost 20% from 10.2×10^{-3} for *ERAI* to 8.3×10^{-3} for *ERA5*.

As for the stations, the posterior mismatches to the satellite observations are significantly improved when using ERA5 instead of ERAI. Figures 6.3b and 6.3d are the same as Figures 6.3a and 6.3c, but zoomed in on the posterior mismatches. As for the prior mismatches, the largest improvements from *ERAI* to *ERA5* are found in the tropics, but there is little variability over time. Globally, the error-weighted mean posterior mismatch is reduced by more than 20% from 1.00×10^{-3} for *ERAI* to 0.77×10^{-3} for *ERA5*.

The posterior budgets of both experiments as shown in Table A.7 in the appendix reveal only minor changes in the chemical loss of CO to OH of less than $+1\%$. This signifies small differences in the temperature and pressures fields, implying that differences in the results of the two experiments likely originate from transport, boundary layer mixing, or dry deposition. The latter experienced a significant reduction, as shown in the budget, in the order of almost -20% compared to the inversion based on ERAI. Note that while dry deposition makes up a much smaller portion of the total loss, the difference in dry deposition is larger than in loss to OH, even when considering absolute values in Tg yr^{-1}. The overall lowered loss is counterbalanced by a minor reduction in overall emissions of roughly -0.5%, which is mostly attributed to a change in CO from fossil fuel combustion of around -2.1%. Since the modeled posterior mixing ratios at the locations of the surface stations remain mostly unchanged despite lowered upwind emissions when using ERA5 instead of ERAI, the reduction in total loss due to dry deposition is likely caused by differences in the meteorological surface fields, rather than by overall lowered surface mixing ratios.

The improvement in global mean mismatches from *ERAI* to *ERA5* is larger for the surface stations (around 27%) compared to the satellite observations (around 23%). This discrepancy occurs, in part, because the effect of the changes in dry deposition rates boundary layer mixing is stronger on the surface measurements. Since the satellite samples the total column, it is less affected by changes that predominantly impact the surface layer. Additionally, most valid satellite observations occur over land and are, therefore, close to the sources and less affected by transport and loss.

6.4   Conclusions

The meteorological fields driving the offline TM5-4dvar model were updated from ERAI to ERA5 and the impact of this change was investigated by comparing the results from

two otherwise identical inversion experiments. The ERA5 based inversion features overall improvements in the mismatches between the model and both the surface flask measurements and the satellite observations compared to the ERAI based inversion. For the surface stations, two issues already pointed out as shortcoming of the ERAI based setup in Chapter 3 were noticeably improved. One issue were improbable emission increments in remote regions around the locations of surface stations. These increments were significantly reduced, which is likely linked to a more realistic representation of dry deposition and boundary layer mixing when using ERA5 surface parameters. The other issue was a strong spike in emission increments in Central Europe, previously linked to shortcomings in the representation of mountainous stations in the model. This spike is considerably smaller when using ERA5 meteorology, with both prior and posterior mismatches improved by a factor of two each. Furthermore, both the prior and posterior mismatches of the model to the satellite observations were improved globally, but especially in the tropics, indicating a better representation of vertical transport in ERA5.

Chapter 7

Conclusions

In this thesis, an inverse modeling setup that makes use of the, compared to the model, high resolution satellite observations from the TROPOMI instrument, was created and refined to investigate global atmospheric carbon monoxide emissions.

A major technical contribution of this thesis was the development of methods for handling the TROPOMI observations. The issues that come along with their high observation density and, therefore, large observation count, were tackled by the introduction of novel methods for the creation of super-observations and for the estimation of satellite error inflation factors. None of the new methods are specific to TROPOMI and may be applied in other setups and with other instruments as well. Through the use of super-observations, the computation time could be reduced substantially (-20%), with minimal impact on the inversion results. A new method for estimating a globally constant satellite error inflation factor, which was based on the approach used in previous studies, was shown to be inadequate for the very large number of observations provided by the TROPOMI instrument. Consequently, another method was proposed, where each observation is assigned an inflation factor based on the local observation density. While preliminary results look promising, further research is required, in particular to investigate the efficacy of the new inflation method for capturing biomass burning events at higher model resolutions. Because of its relatively low lateral resolution, the current setup has only a limited capability for simultaneously optimizing emissions from multiple categories and suffers from aliasing, specifically from the biomass burning emissions to the other source categories. Therefore, a future study with a high resolution zooming region, to focus on specific biomass burning events, could be useful.

Based on the conducted inversion experiments, this work shows that the observations from the TROPOMI instrument on their own are sufficient to constrain the background CO emissions in the southern hemisphere without having to rely on the additional information provided by the surface flask measurements. This finding indicates that inversions based only on TROPOMI satellite observations are feasible, as long as

the region of interest is sufficiently far south of 55° N, e.g. in the southern hemisphere. Those inversion experiments could circumvent the long analysis cycle of the surface flask measurements and analyze specific events with a reduced time lag. In contrast, the representation of CO at high northern latitudes ($> 55°$ N) appears to be generally problematic in the model, leading to poor prior performance. There, the posterior is very poorly constrained by the TROPOMI satellite observations. Notably, this is the case even during northern hemispheric summer time, when satellite observations at high northern latitudes are plentiful.

These first experiments using the TM5-4dvar inverse modeling framework to optimize CO emissions based on TROPOMI satellite observations and surface flask measurements revealed that the initial setup had an imbalanced budget, which introduced biases in the optimized emissions. Multiple ways to reduce those biases were investigated:

To address the budget imbalance, first, the climatological OH field was replaced with one that is consistent with the secondary CO production rates that were taken from a full chemistry model. This approach lead to a nearly closed prior budget and strongly reduced the largest biases, which occurred in the northern tropics. While there is evidence for biases in the spatial distribution of the climatological OH fields, at the same time, the model based OH fields appear to be biased in amplitude and north-south gradient. Overall, OH remains problematic because, while it is under-constrained by current measurements, the OH distribution is crucial to obtain meaningful optimized CO emissions. As shown by this work, any bias in OH will directly translate to a bias in the emissions. In contrast, the exact choice of prior emissions has a much weaker impact on the inversion results, because emissions are optimized as part of the state. Therefore, further research beyond CO inversions is needed to find a proper OH distribution field for the global atmosphere.

Second, by updating to ERA5 meteorology, biases in the emissions around the surface stations and the overall mismatch to the satellite observations, especially in the tropics, could be reduced. Combined, an overall improved posterior fit to both observational datasets, the satellite observations and the surface measurements, was achieved. This improved posterior fit shows that even in the simplified model used here, ERA5 leads to a better representation of dry deposition and vertical transport.

Third, while only a minor factor, biases in the southern hemisphere were reduced (locally up to 17 % difference in optimized emissions) through explicit handling of direct natural CO emissions instead of considering them as part of the secondary CO source.

Appendix A

Additional materials

A.1 Additional figures and tables for Chapter 3
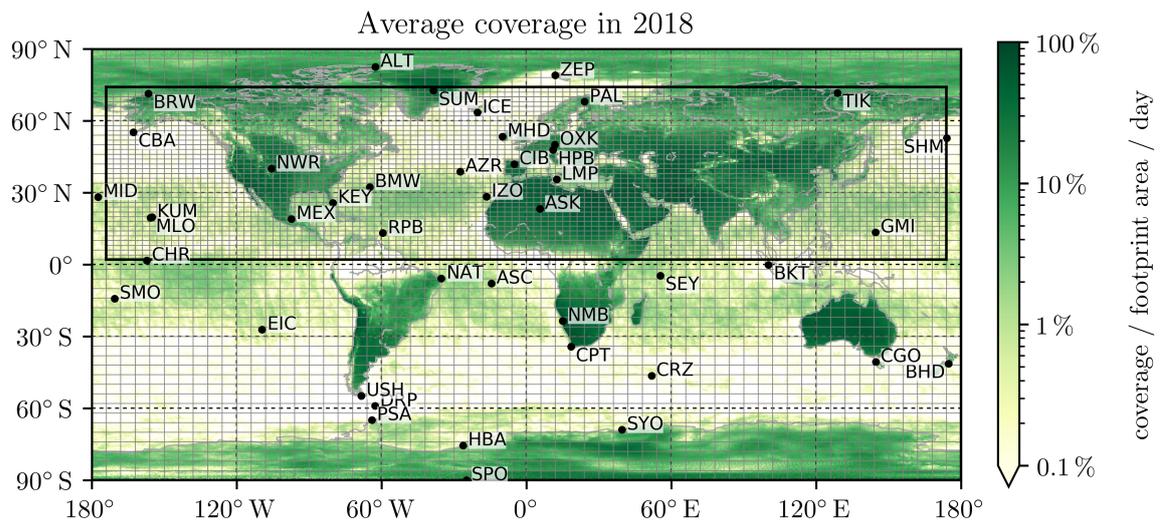
Figure A.1: Daily average TROPOMI CO data coverage for 2018 after quality filtering, normed on the mean footprint area per observation ($\approx 84\,\mathrm{km}^2$). As in Figure 3.1 from the main manuscript, the nested zoom regions are shown as grids and the locations of the surface stations as dots.
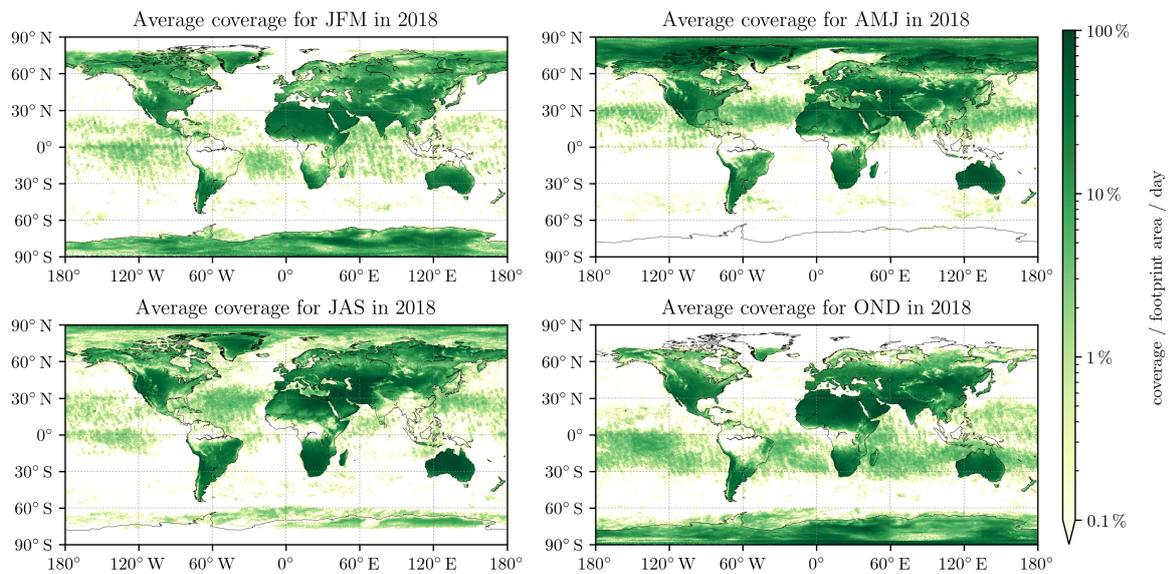
Figure A.2: Same as Fig. A.1, but split in 3-month periods to show seasonal variations.
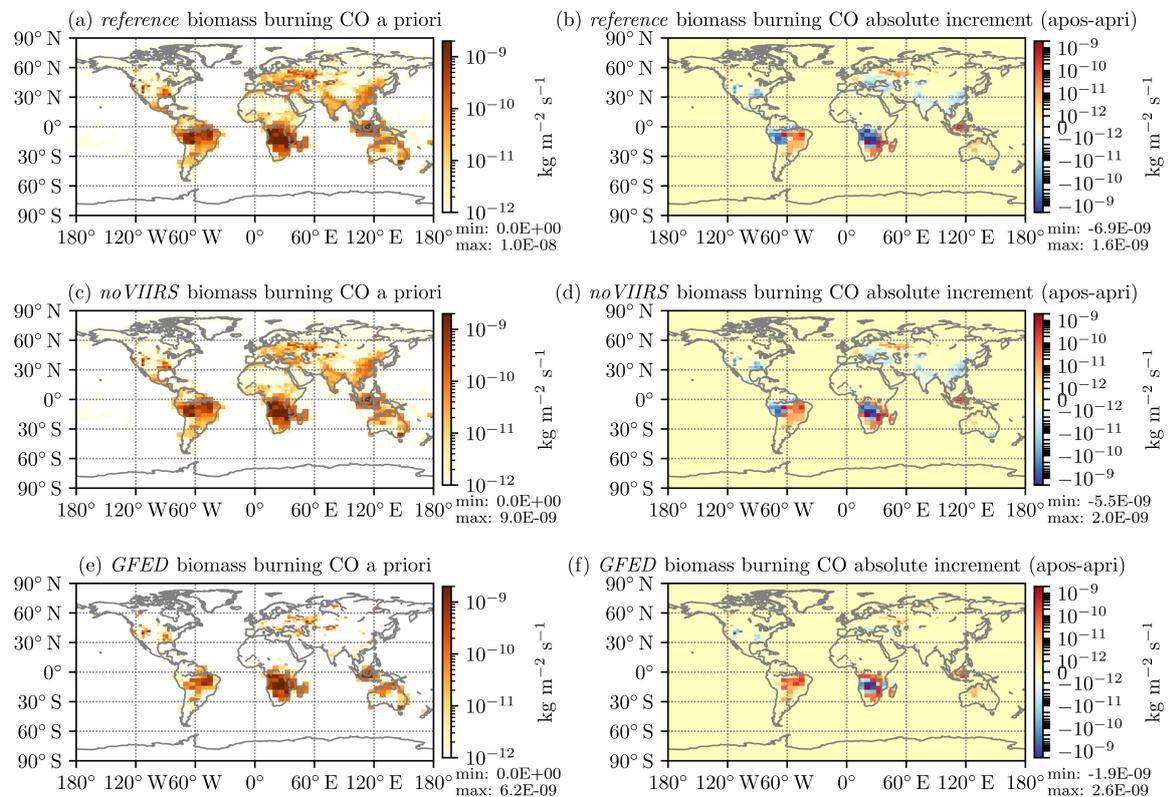
Figure A.3: Global biomass burning CO emission for a single day, 15 September 2018, for the first set of inversions. The panels show the a priori emissions (left) and absolute emission increment (right) for the FINN2.4+VIIRS ((a) and (b)), FINN2.4 ((c) and (d)), and GFED ((e) and (f)) inventories.

Figure A.4: Station-wise mean error-weighted mismatch between flasks and model for all main inversions. Stations are ordered by latitude. Starred (*) stations are the ones analysed in more detail in Chapter 3. The dashed line denotes the 55° N cutoff.
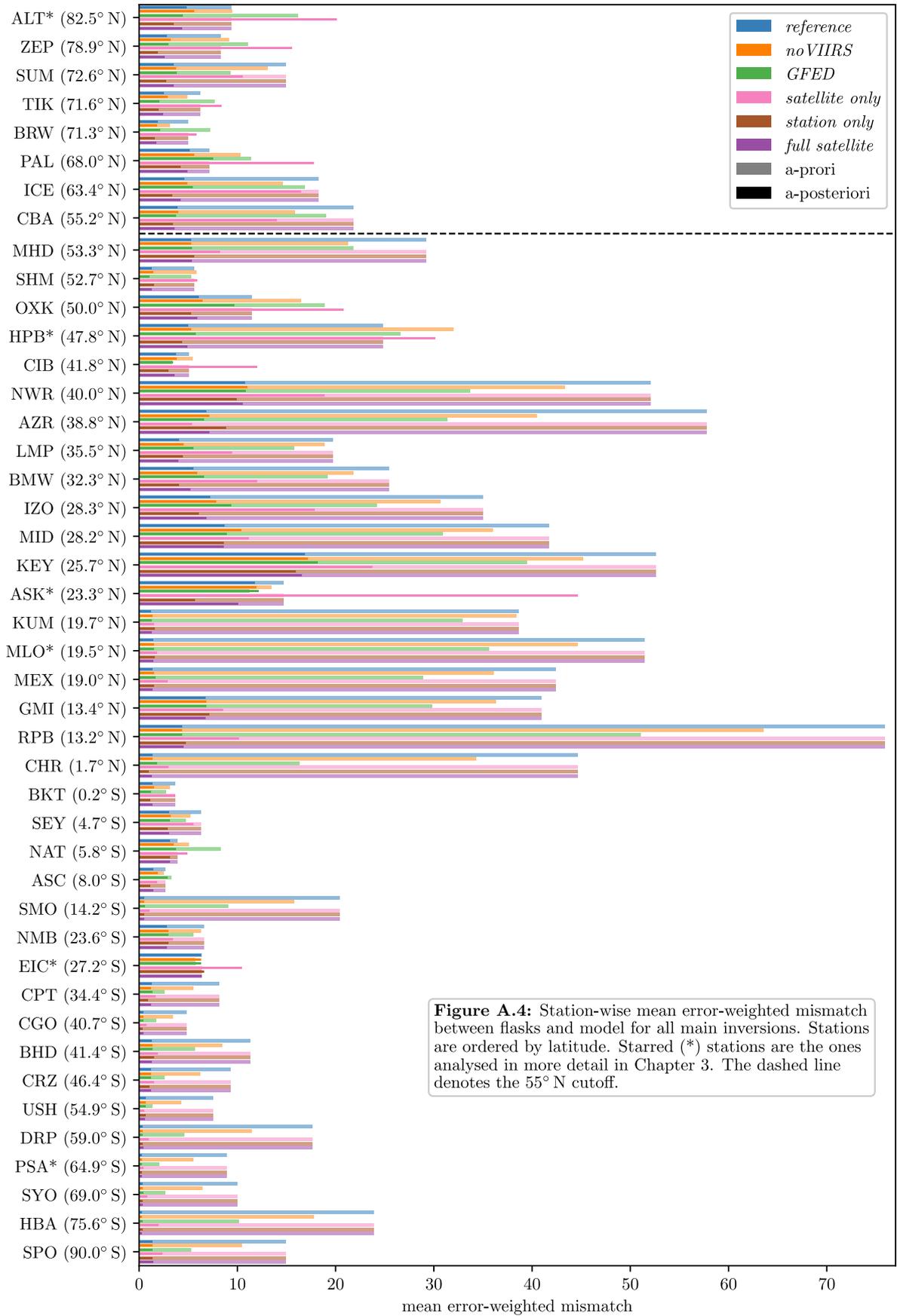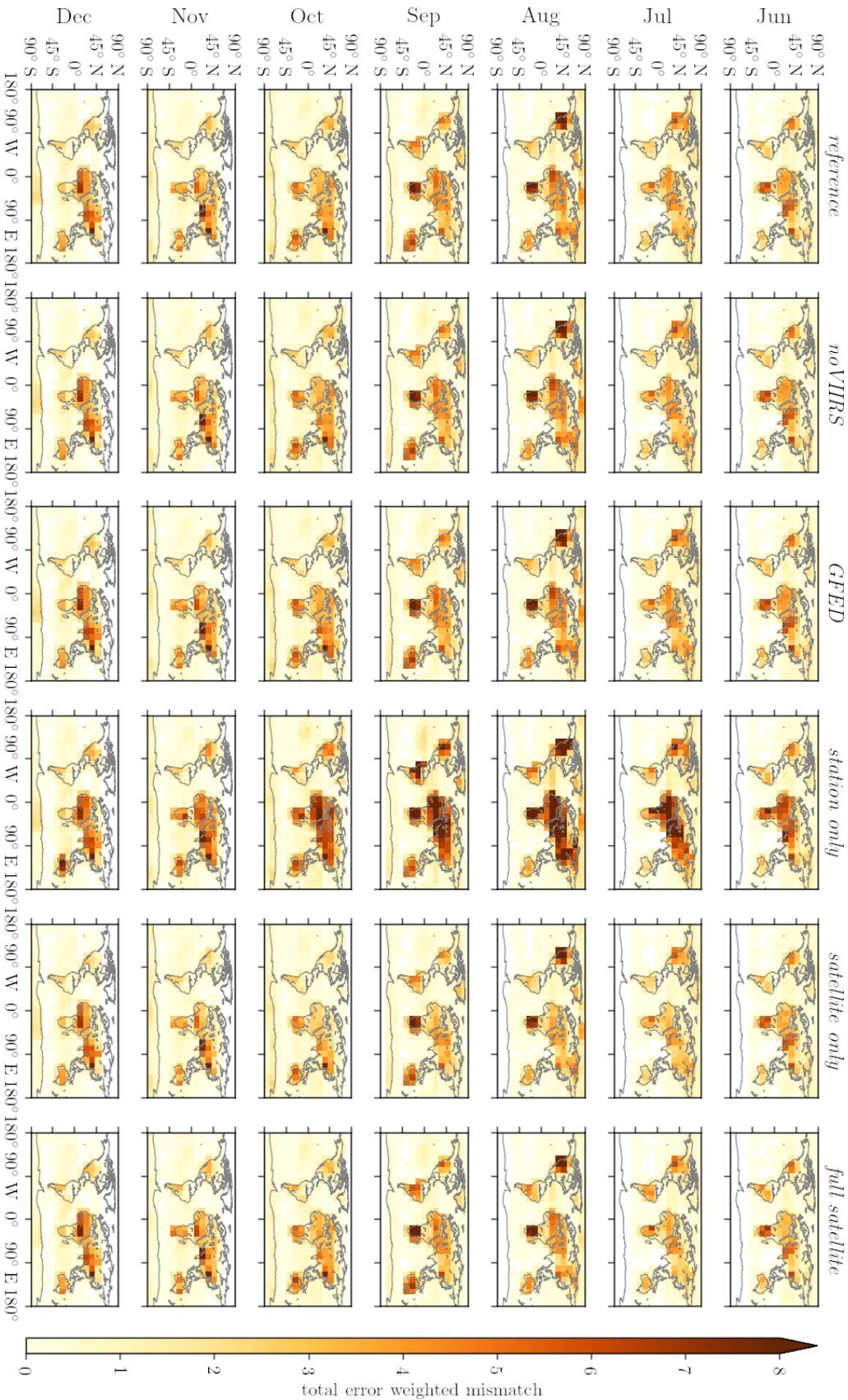
Figure A.5: Total error-weighted mismatch between model and satellite observations aggregated in 12° × 12° boxes and per month. This is equivalent to the contribution to the cost function from all observation in each box. The mismatches over the oceans are low because on the one hand there are fewer observations and on the other hand those observations represent the background mixing ratios, which are captured well by the model.
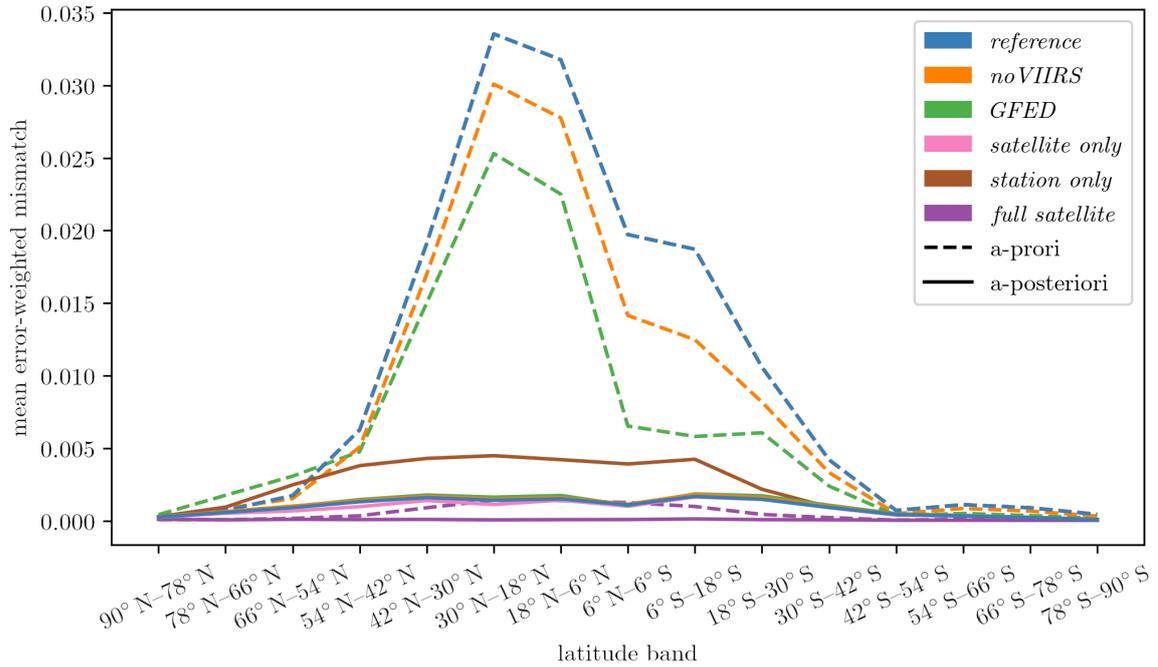
Figure A.6: Mean a priori and a posteriori mismatch between the satellite and the model for all main inversions in Chapter 3 over the whole period aggregated in 12° latitudinal bands. A priori is larger, expect *station only* in high northern latitudes.
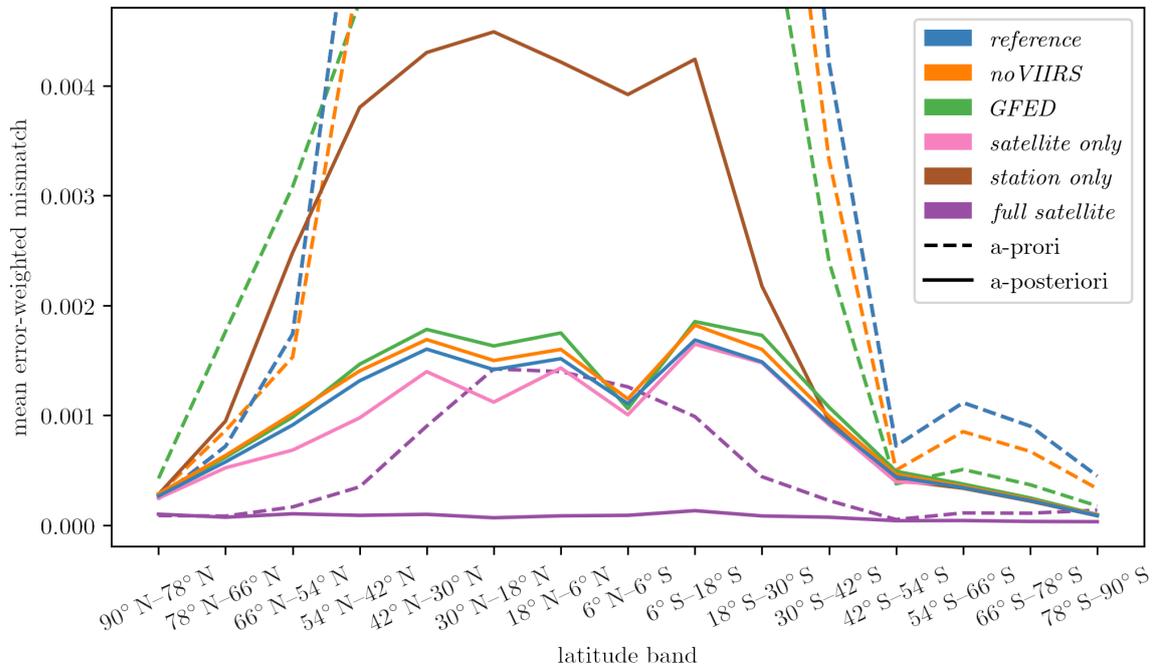
Figure A.7: Same as above, but with the vertical axis scaled only by the a posteriori graphs, to make them more easily discernible. Clearly, all curves converge in the southern hemisphere.

A ADDITIONAL MATERIALS
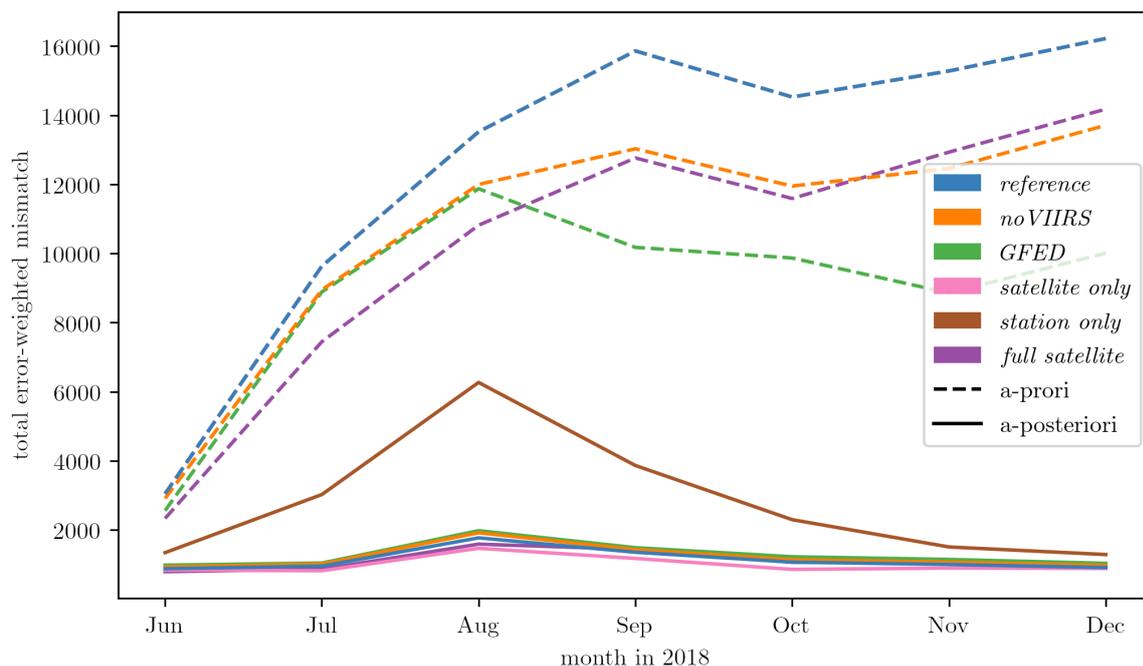
Figure A.8: Total a priori and a posteriori mismatch between the satellite and the model for all main inversions in Chapter 3. A priori mismatch only rises for the first two month of the inversion period and the reaches a plateau, pointing to the budget not being closed properly.
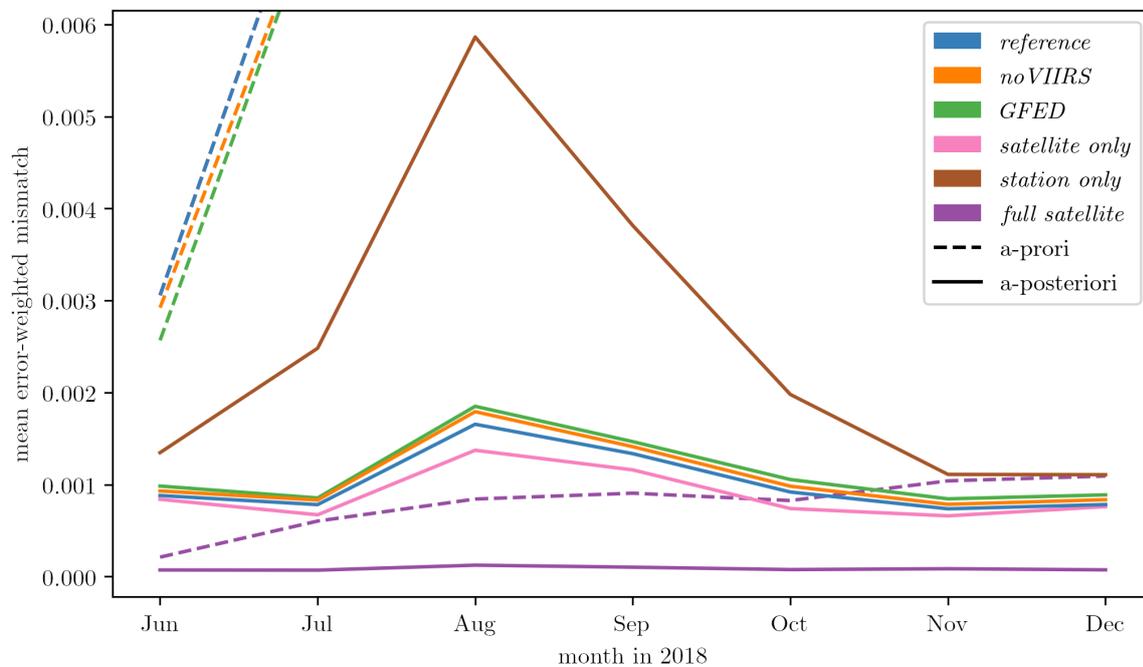
Figure A.9: Same as above, but with the vertical axis scaled only by the a posteriori graphs, to make them more easily discernible. The increased mismatch during the main burning season (Jul–Aug–Sep) due to the models inability to properly capture local biomass burning events is clearly visible, especially in the *station only* inversion.
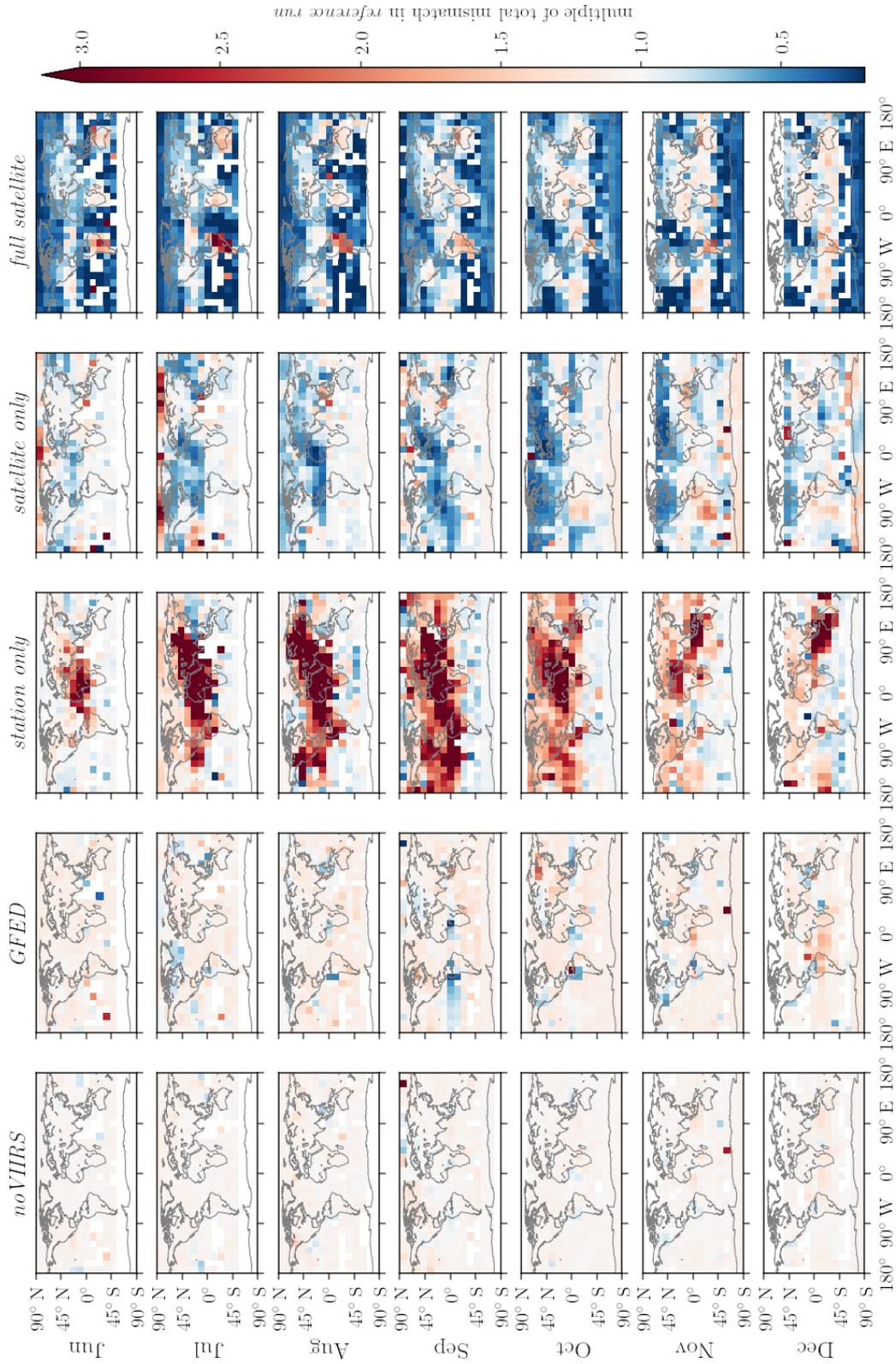
Figure A.10: Total error-weighted mismatch between model and satellite observations aggregated in 12° × 12° boxes and per month, divided by the corresponding mismatch of the *reference* inversion for that month. Red areas signify a worse fit than the *reference* inversion and vice versa for blue.
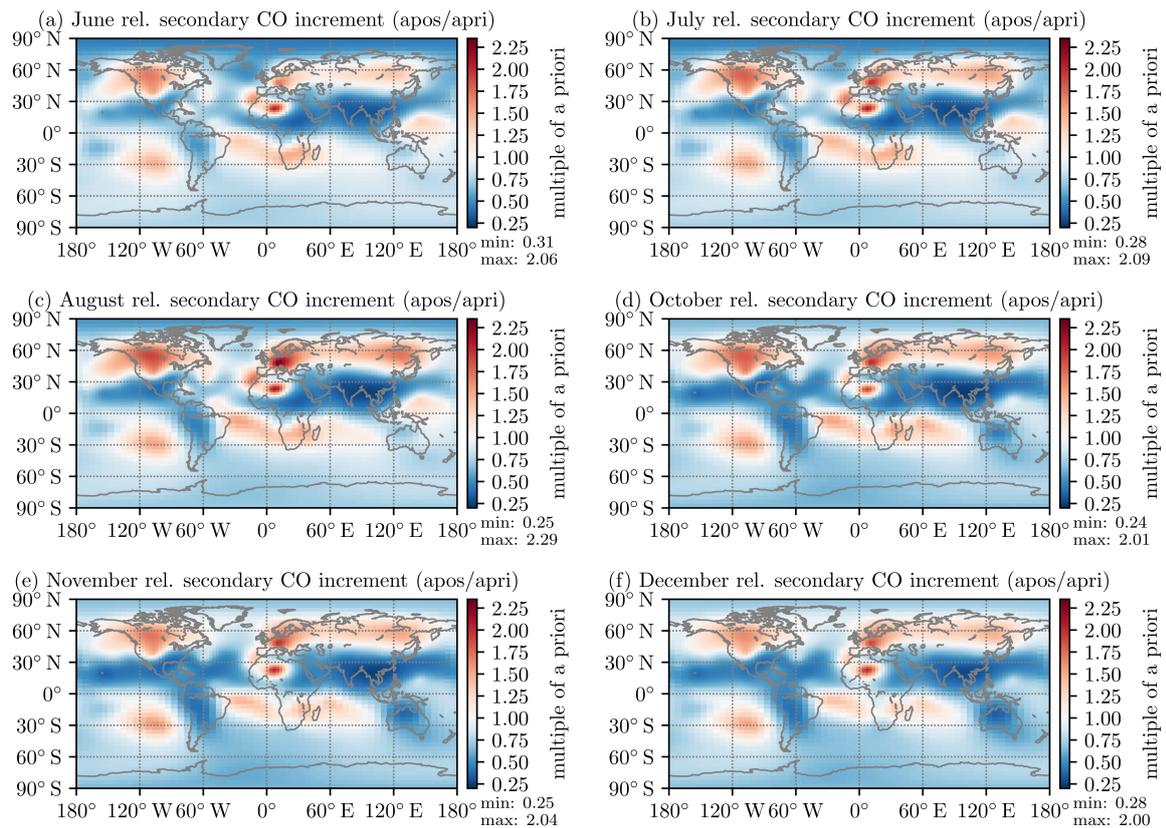
Figure A.11: Global relative secondary CO emission increments resulting from the *reference* inversion for the remaining six months not shown in the main text. Overall, the global pattern of increments and decrements are preserved over time, only the amplitudes differ slightly. Most notably, the decrements in the remote northern hemisphere are more pronounced in the beginning of the inversion period, while those in the southern hemisphere are more pronounced towards the end of the period. The differences are shown in more detail in Fig. A.12.
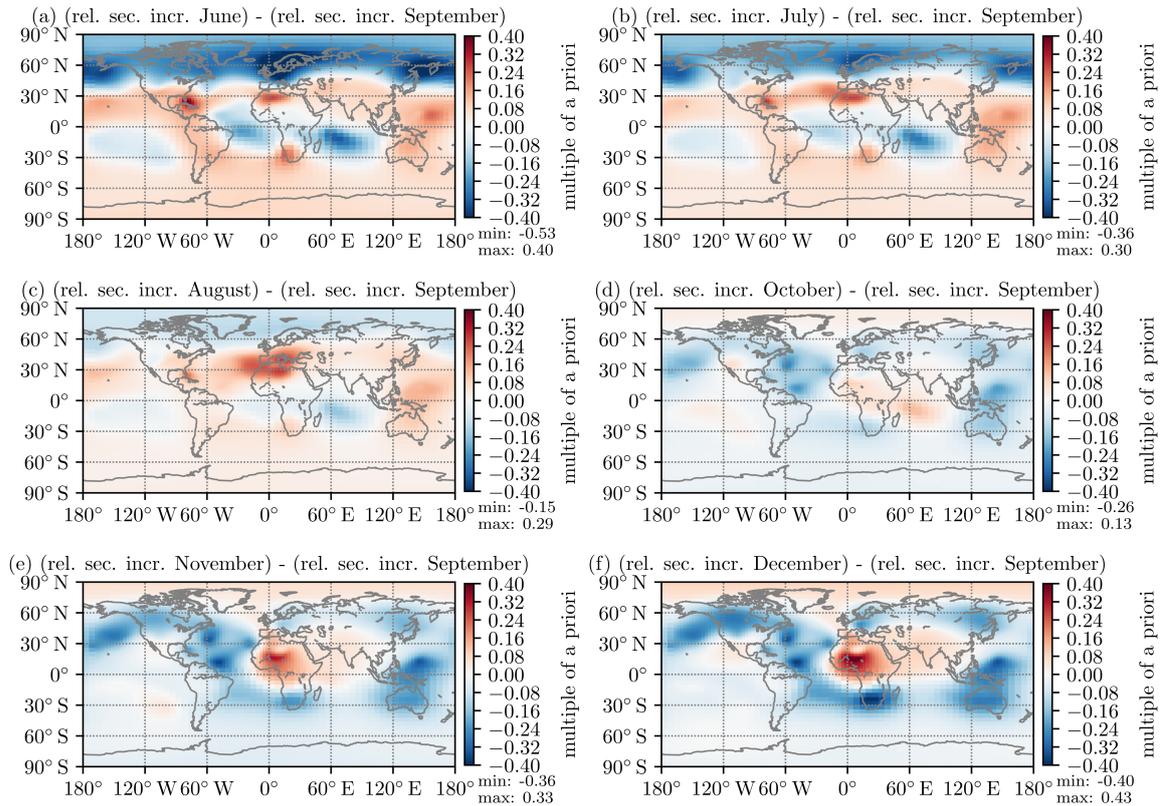
Figure A.12: Differences in relative secondary CO emission increments resulting from the *reference* inversion between September (the month shown in Fig. 3.7 of the main text) and the remaining six months not shown in the main text (see also Fig. A.11 above). Note the smaller range of the colorbar compared to Fig. A.11. Interpretation of these plots is not trivial, since positive differences potentially imply a smaller/larger deviation from the prior, if the relative increments are smaller/larger than 1, but so do negative differences, if the relative increments are larger/smaller than one. More details for the deviation from the prior are shown in Fig. A.13.
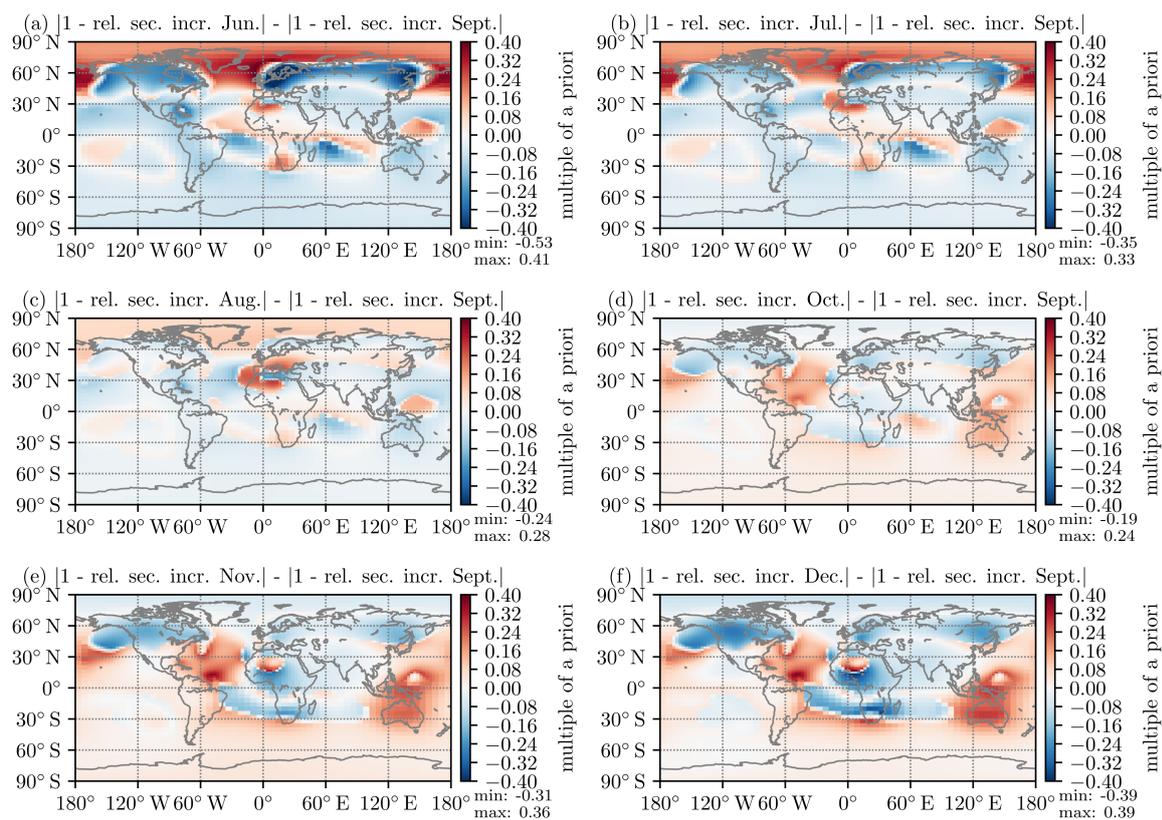
Figure A.13: Differences in relative deviation from the prior for the secondary CO production for the *reference* inversion for September (the month shown in Fig. 3.7 of the main text) and the remaining six months not shown in the main text (see also Fig. A.11 above). Red signifies areas were the relative increments deviate more from the prior than in September and vice versa for blue areas. Note that by taking the absolute value of 1 minus the relative increments, this does not contain information about whether the emission were increasing or decreasing anymore.

Table A.1: Station-wise mean error-weighted mismatch between flasks and model for the inversion experiments described in Chapter 3. The prior (apri) mismatches are omitted for the *station only*, *satellite only*, and *full satellite* inversion, because they are identical to the ones from the *reference inversion*.

station	reference		station only	satellite only	full satellite	noVIIRS		GFED	
	apri	apos	apos	apos	apos	apri	apos	apri	apos
ALT	9.35	4.82	3.49	20.10	4.35	9.46	5.61	16.16	4.42
ASC	2.65	1.43	1.10	1.82	1.41	2.50	1.86	3.26	2.89
ASK	14.70	11.76	5.69	44.62	10.08	13.43	11.94	11.20	12.12
AZR	57.80	6.82	8.84	5.36	7.13	40.49	7.13	31.39	6.57
BHD	11.31	1.27	1.49	1.91	1.24	8.48	1.35	5.65	1.34
BKT	3.69	1.39	1.13	3.65	1.35	3.15	1.47	2.72	1.20
BMW	25.42	5.55	4.05	12.00	5.22	21.78	5.89	19.18	6.58
BRW	4.95	1.86	1.59	5.84	1.74	3.12	1.84	7.21	2.15
CBA	21.77	3.87	3.46	14.00	3.61	15.90	3.96	19.05	3.77
CGO	4.83	0.39	0.35	0.72	0.42	3.47	0.39	1.74	0.42
CHR	44.66	1.37	0.94	2.99	1.29	34.30	1.37	16.29	1.78
CIB	5.09	3.78	2.95	12.01	3.61	5.45	3.83	3.34	3.42
CPT	8.17	1.24	0.90	1.66	1.22	5.55	1.23	2.61	1.37
CRZ	9.34	1.16	1.04	1.47	1.19	6.18	1.20	2.59	1.20
DRP	17.63	0.36	0.38	0.95	0.41	11.48	0.36	4.63	0.39
EIC	6.40	6.32	6.58	10.50	6.35	5.67	6.32	5.77	6.28
GMI	40.97	6.76	7.15	8.53	6.77	36.30	6.84	29.82	6.80
HBA	23.89	0.27	0.32	1.96	0.28	17.83	0.28	10.16	0.31
HPB	24.81	4.95	4.40	30.17	4.86	32.01	5.31	26.61	5.75
ICE	18.22	4.60	3.34	16.47	4.20	14.65	4.87	16.90	5.43
IZO	34.98	7.24	6.07	17.84	6.86	30.64	7.81	24.20	9.38
KEY	52.60	16.88	15.91	23.77	16.58	45.17	17.18	39.48	18.15
KUM	38.65	1.20	1.61	1.48	1.27	38.44	1.34	32.90	1.27
LMP	19.73	4.05	4.47	9.46	3.99	18.86	4.48	15.80	5.55
MEX	42.44	1.39	1.51	2.86	1.33	36.12	1.47	28.88	1.70
MHD	29.21	5.27	5.59	8.25	5.33	21.24	5.31	21.80	5.35
MID	41.72	8.72	8.59	11.16	8.64	36.04	10.40	30.94	8.90
MLO	51.42	1.40	1.60	1.82	1.43	44.68	1.52	35.65	1.49
NAT	3.88	3.15	3.11	4.89	3.11	5.06	3.52	8.32	3.72
NMB	6.61	2.80	2.94	3.45	2.82	6.28	2.97	5.50	2.98
NWR	52.03	10.78	9.92	18.83	10.53	43.35	10.99	33.67	10.86
OXK	11.46	6.04	5.26	20.78	5.94	16.48	6.44	18.90	9.69
PAL	7.11	5.11	4.19	17.75	4.93	10.33	5.59	11.37	7.54
PSA	8.90	0.27	0.26	0.45	0.28	5.48	0.27	2.04	0.27
RPB	75.91	4.36	4.77	10.18	4.48	63.58	4.33	51.05	4.35
SEY	6.32	3.09	2.88	5.53	3.06	5.20	3.18	4.71	3.10
SHM	5.58	1.29	1.52	5.90	1.28	5.85	1.42	5.31	1.07
SMO	20.44	0.49	0.50	1.03	0.48	15.75	0.52	9.07	0.57
SPO	14.92	1.37	1.36	2.35	1.44	10.45	1.37	5.28	1.38
SUM	14.90	3.48	2.77	10.55	3.47	13.12	3.71	9.29	3.84
SYO	9.99	0.37	0.35	0.84	0.38	6.47	0.38	2.68	0.39
TIK	6.23	2.53	1.93	8.38	2.41	4.89	2.88	7.70	2.02
USH	7.52	0.64	0.64	0.47	0.60	4.29	0.63	1.38	0.63
ZEP	8.30	2.84	1.93	15.52	2.58	9.18	3.17	11.10	2.99

A ADDITIONAL MATERIALS

A.2 Additional figures and tables for Chapter 4
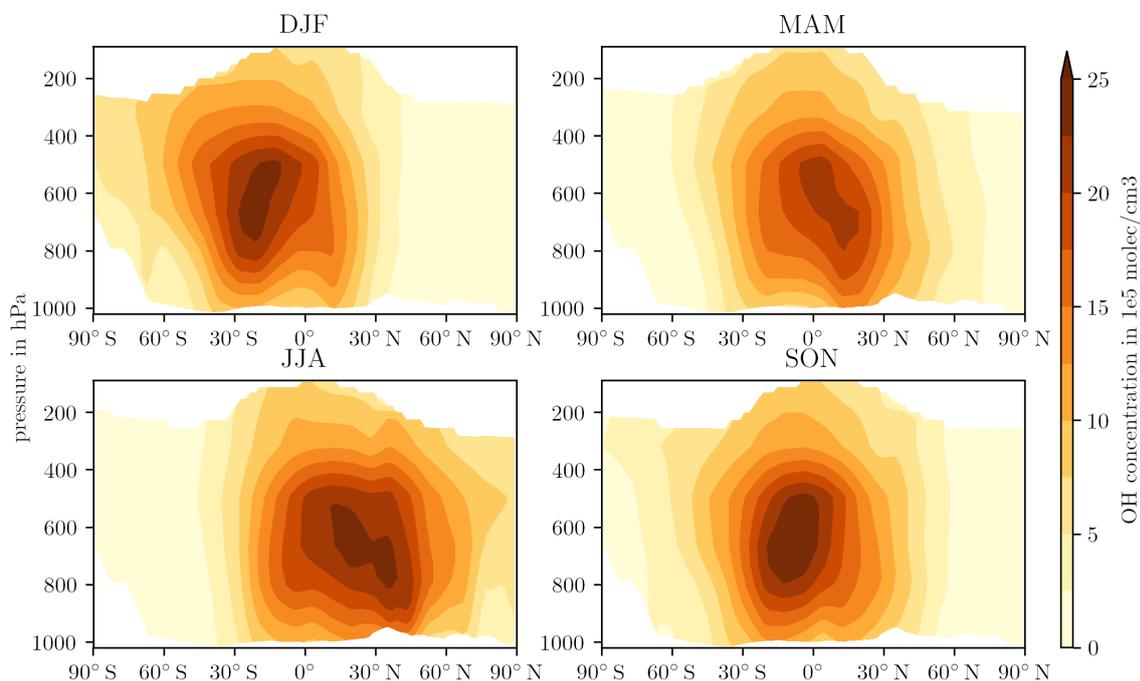
Figure A.14: Seasonal, volume-weighted tropospheric (cutoff at 150 ppb O$_3$) zonal mean OH concentrations for the climatological Spiv-OH field. From the maximum at $\approx$600 hPa concentrations strictly decrease towards the surface.
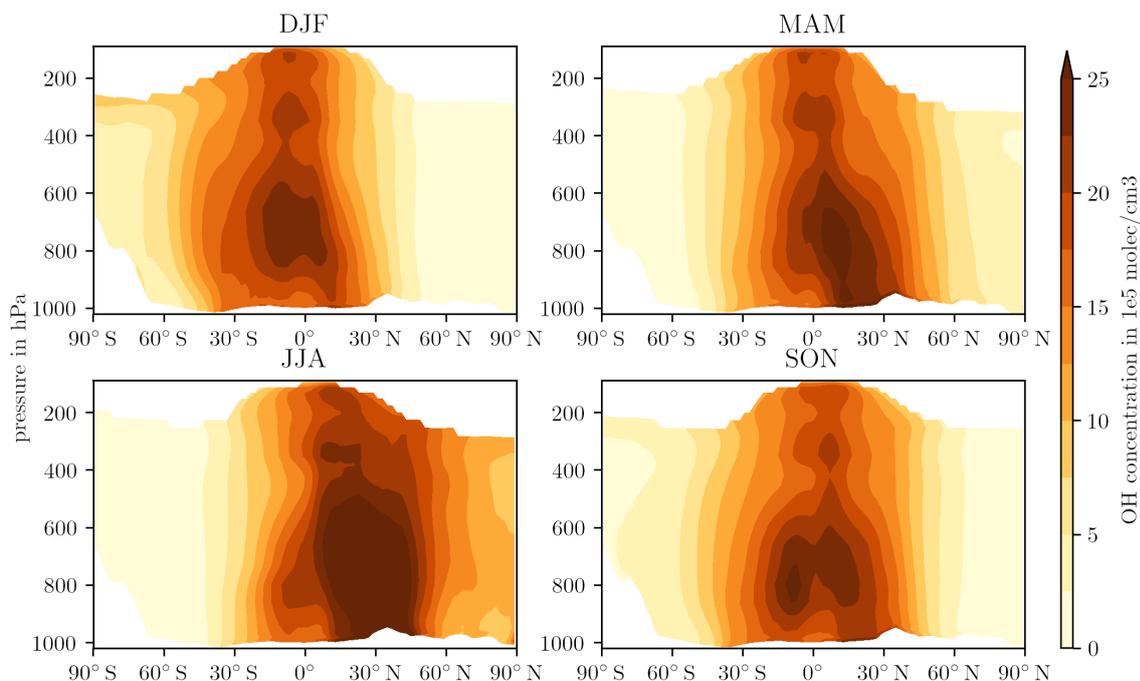
Figure A.15: Same as Fig. A.14 but for the modeled TM5-MP-OH field. Note the maximum closer to the surface (800 hPa) compared to Fig. A.14 and the secondary maximum at the surface in the northern (sub)tropics.
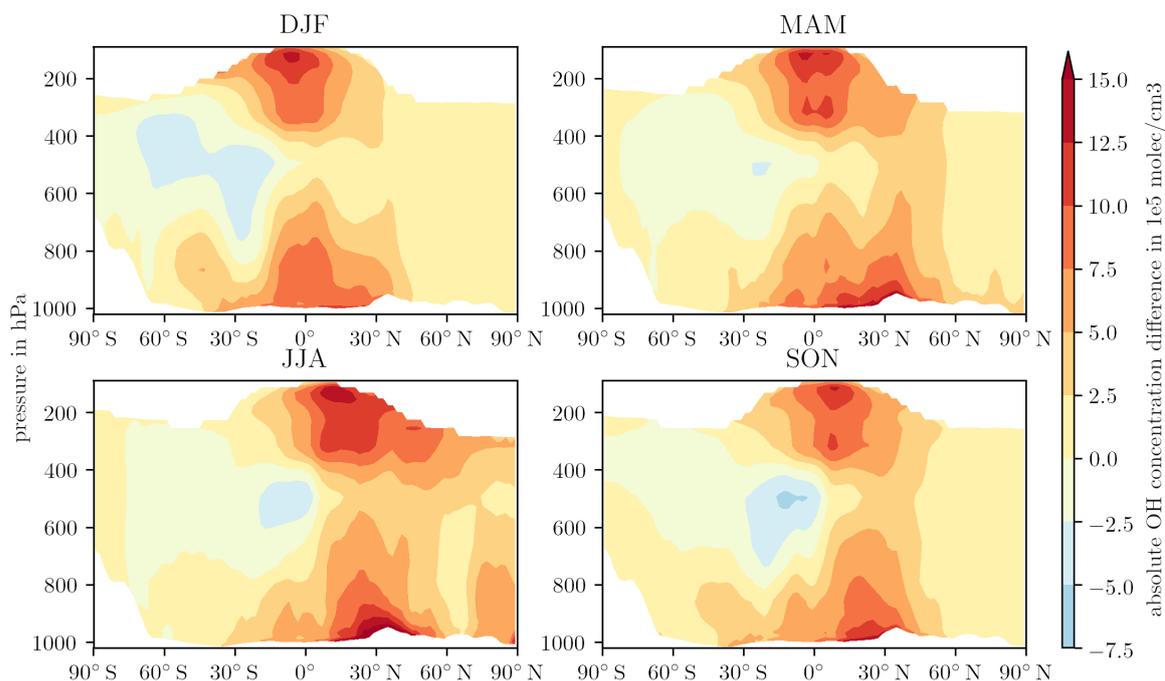
Figure A.16: Absolute difference between Figs. A.15 and A.14, highlighting the much larger OH concentrations close to the surface in TM5-MP-OH.
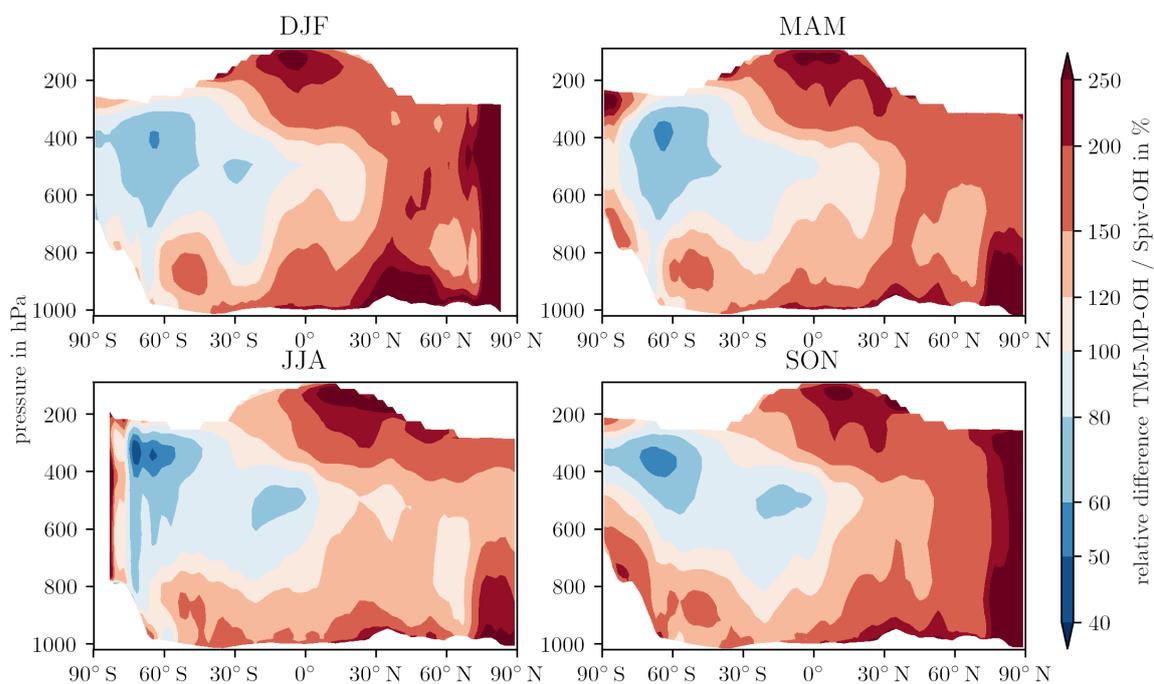
Figure A.17: Relative difference between Figs. A.15 and A.14. Surface OH is roughly twice as large in TM5-MP-OH in the northern hemisphere throughout the year. Note that only because OH concentrations are very low towards the Poles, there appear to be large differences.
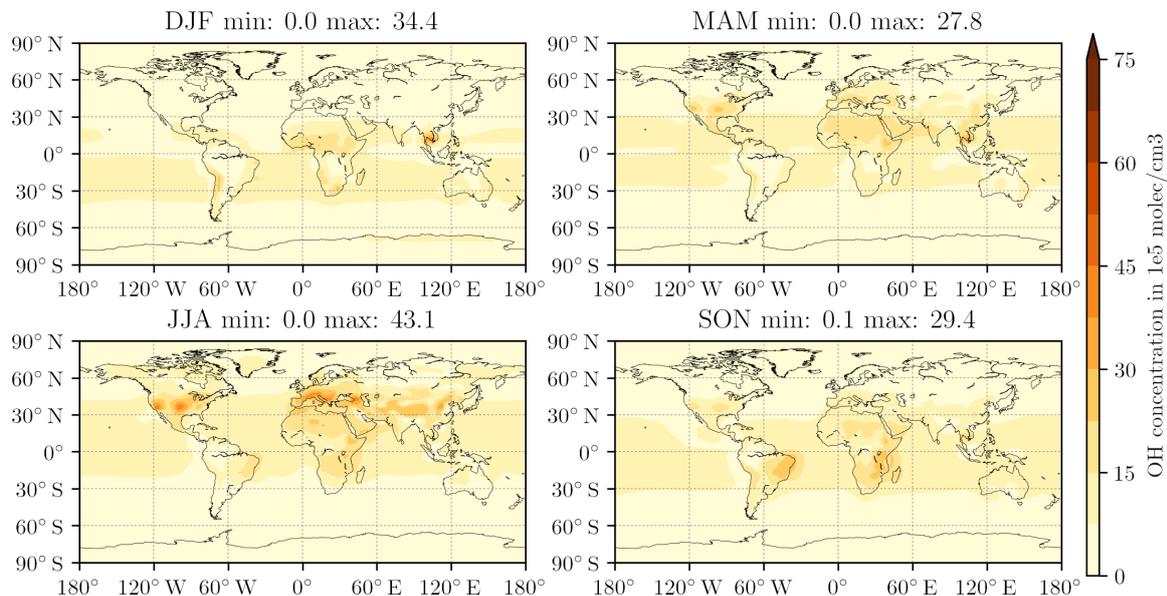
Figure A.18: Seasonal mean of surface layer OH concentration for the climatological Spiv-OH field. Patterns are broad and concentrations low overall.
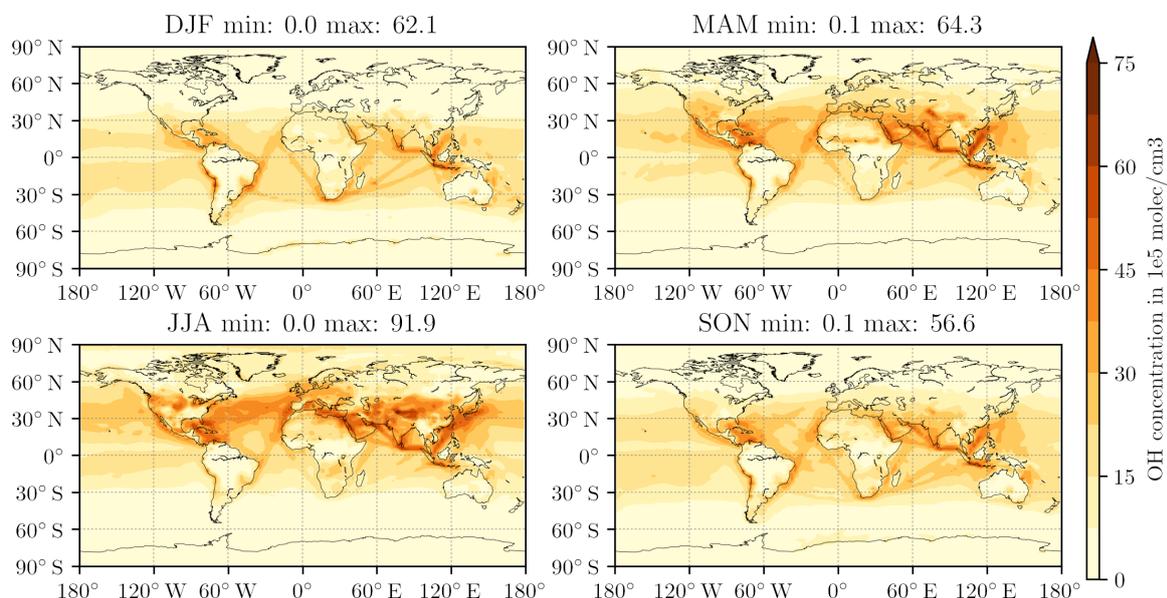
Figure A.19: Same as Fig. A.18, but for the modeled TM5-MP-OH field. Note the much higher concentrations compared to Fig. A.18, the much sharper patterns, e.g. from shipping lines, and the significantly larger concentrations in Asia during northern hemispheric spring and summer. The finer structures are visible because of the much higher model resolution of TM5MP compared to the Spivakovsky climatology.
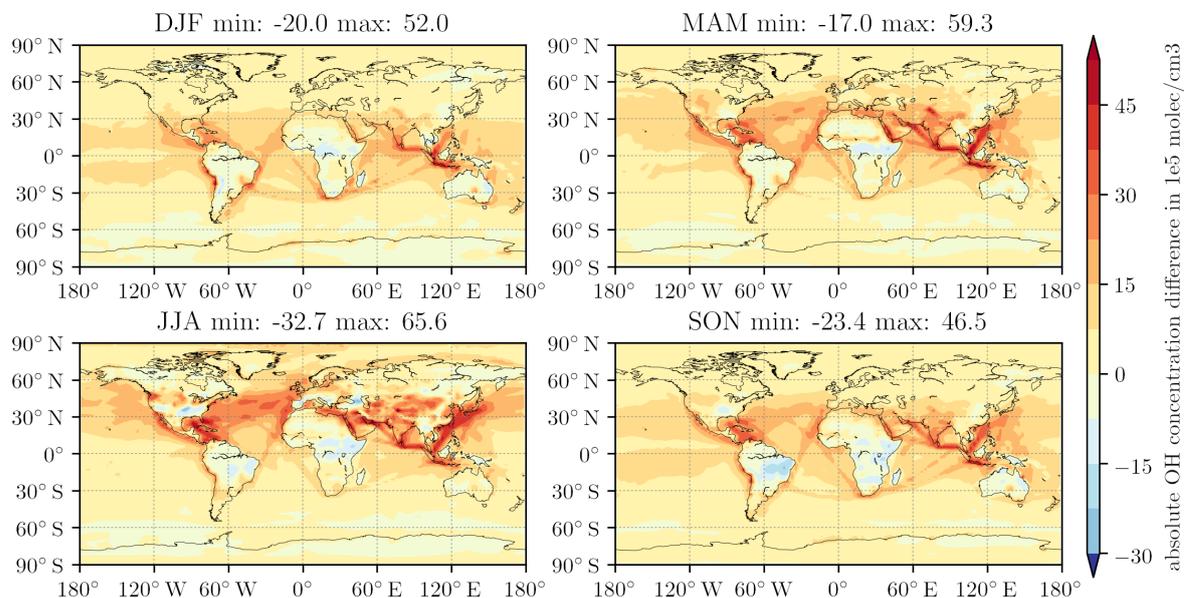
Figure A.20: Absolute difference between Figs. A.19 and A.18. While TM5-MP-OH is much larger over the oceans and Asia, Spiv-OH is slightly in many other regions over land.
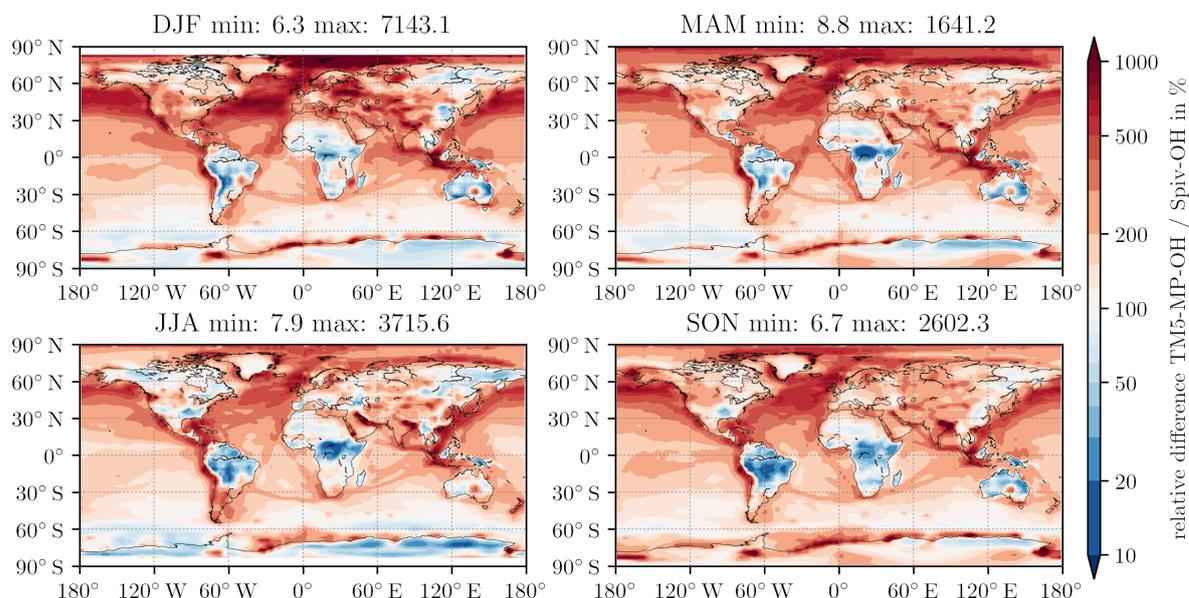
Figure A.21: Relative difference between Figs. A.19 and A.18. Note that the logarithmic scale implies that each factor of e.g. 2 in either direction results in the same color change and that the scale effectively goes from TM5-MP-OH is a tenth of Spiv-OH to it being ten times larger. This highlights how in many regions the two OH fields behave very differently.
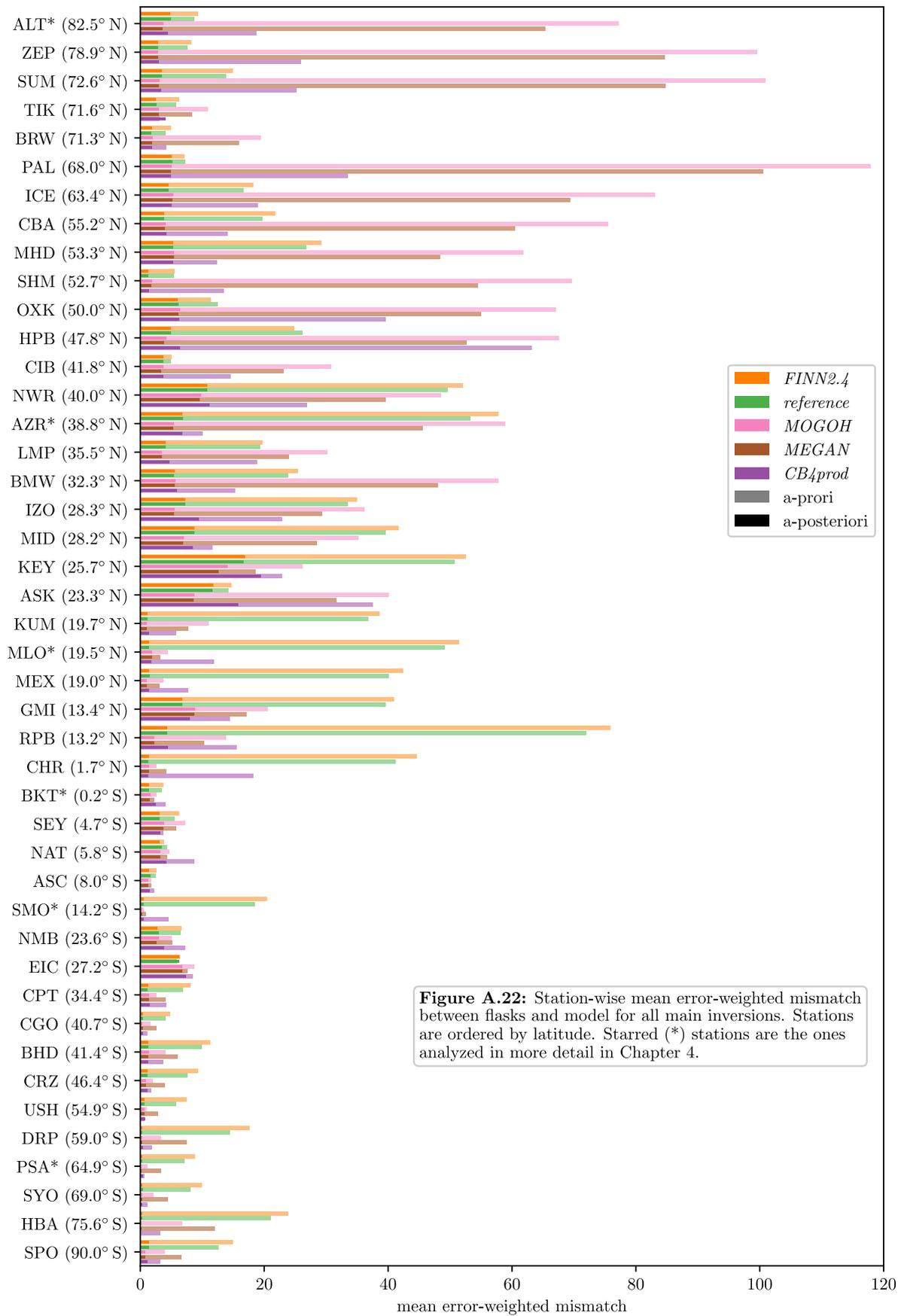
Figure A.22: Station-wise mean error-weighted mismatch between flasks and model for all main inversions. Stations are ordered by latitude. Starred (*) stations are the ones analyzed in more detail in Chapter 4.
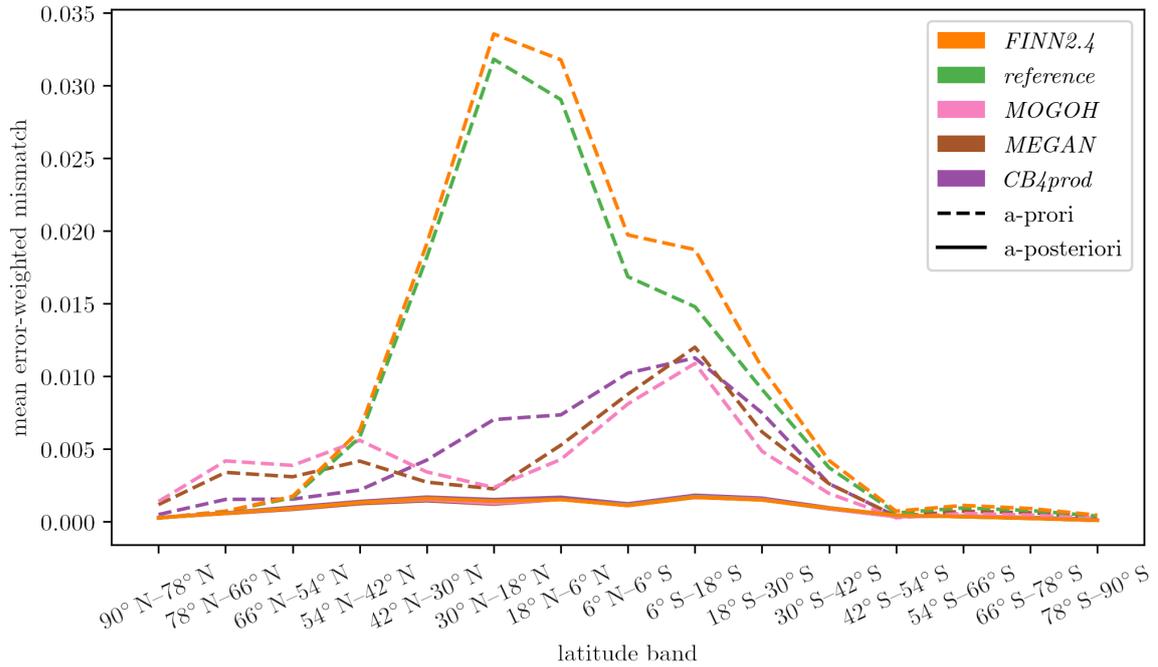
Figure A.23: Mean a priori and a posteriori mismatch between the satellite and the model for all main inversions in Chapter 4 over the whole period aggregated in 12° latitudinal bands. Priors of *MOGOH* and *MEGAN* perform exceptionally well in northern tropics, but worth than the Spivakovsky OH based inversions at higher northern latitudes.
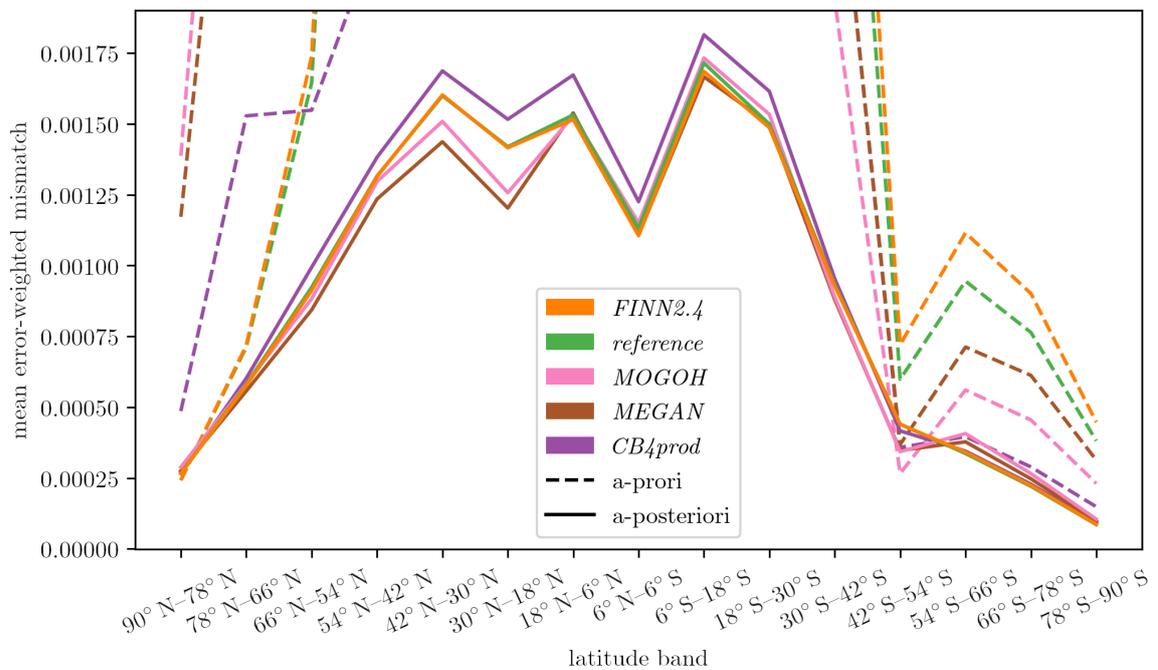
Figure A.24: Same as above, but with the vertical axis scaled only by the a posteriori graphs, to make them more easily discernible.

A   ADDITIONAL MATERIALS

Figure A.25: Mean a priori and a posteriori mismatch between the satellite and the model for all main inversions in Chapter 4. A priori mismatch is notably smaller in *MOGOH* and *MEGAN*. Prior mismatch for *oldprod* starts out small as well, but diverges for the final few months.

Figure A.26: Same as above, but with the vertical axis scaled only by the a posteriori graphs, to make them more easily discernible. The increased mismatch during the main burning season (Jul–Aug–Sep) due to the models inability to properly capture local biomass burning events is clearly visible. Notably, the overall mismatch is smallest for *MEGAN* and largest for *oldprod*, except in the first month.

Figure A.27: Seasonal, tropospheric (cutoff at 150 ppb O₃) zonal mean secondary CO production rates taken from TM5-MP with CBM4 chemistry for 2006.
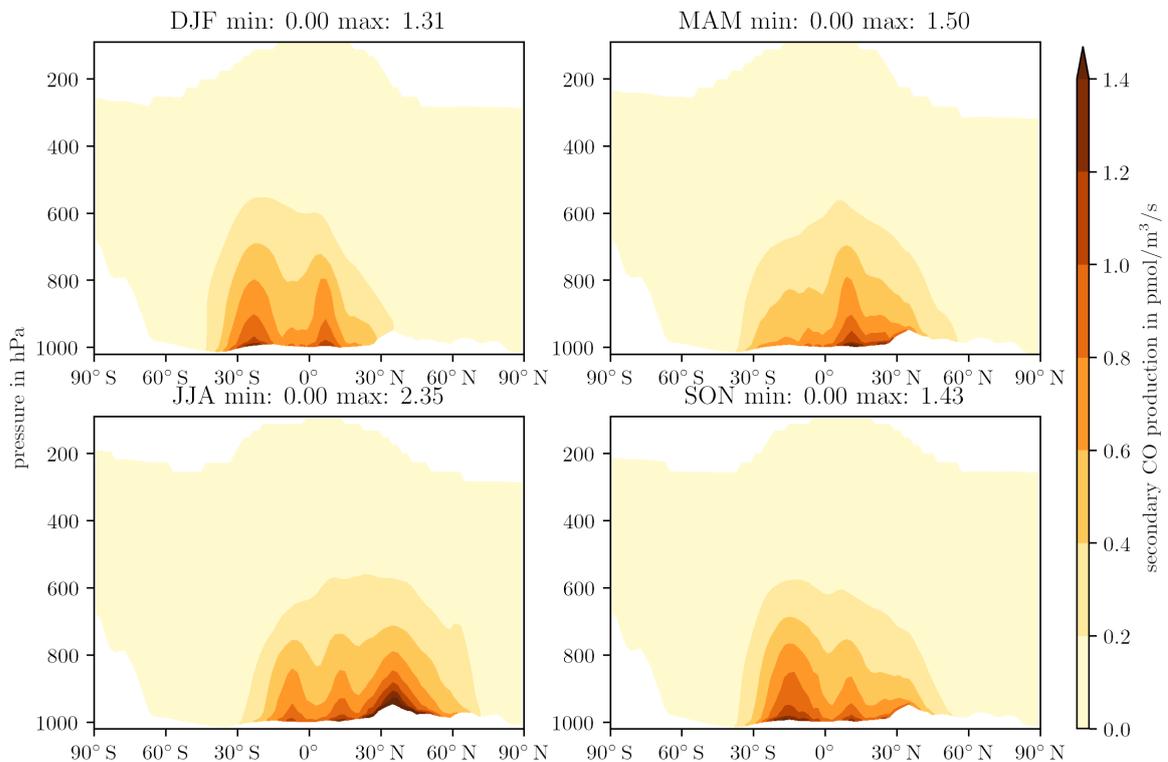
Figure A.28: Same as Fig. A.27, but taken from TM5-MP with MOGUNTIA chemistry for 2018.

A ADDITIONAL MATERIALS
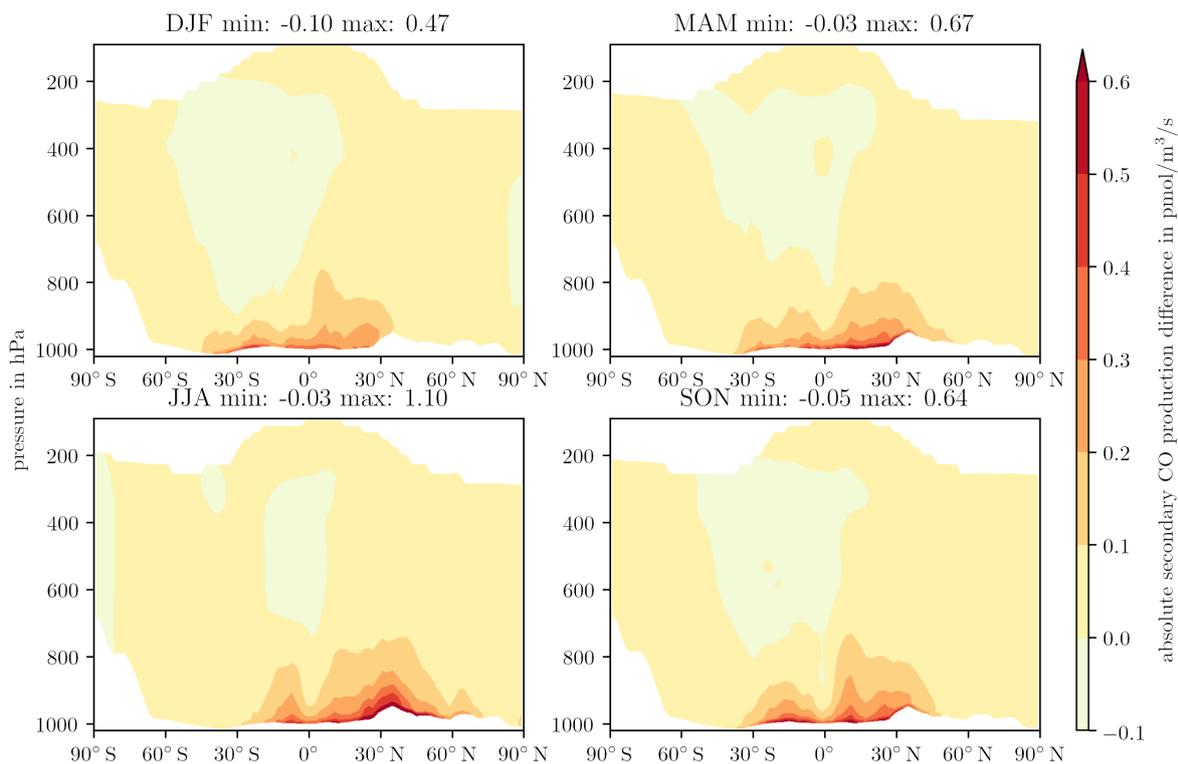
Figure A.29: Absolute difference between Figs. A.28 and A.27. Production rates are much higher with the MOGUNTIA chemistry close to the surface in the tropics and northern extratropics.
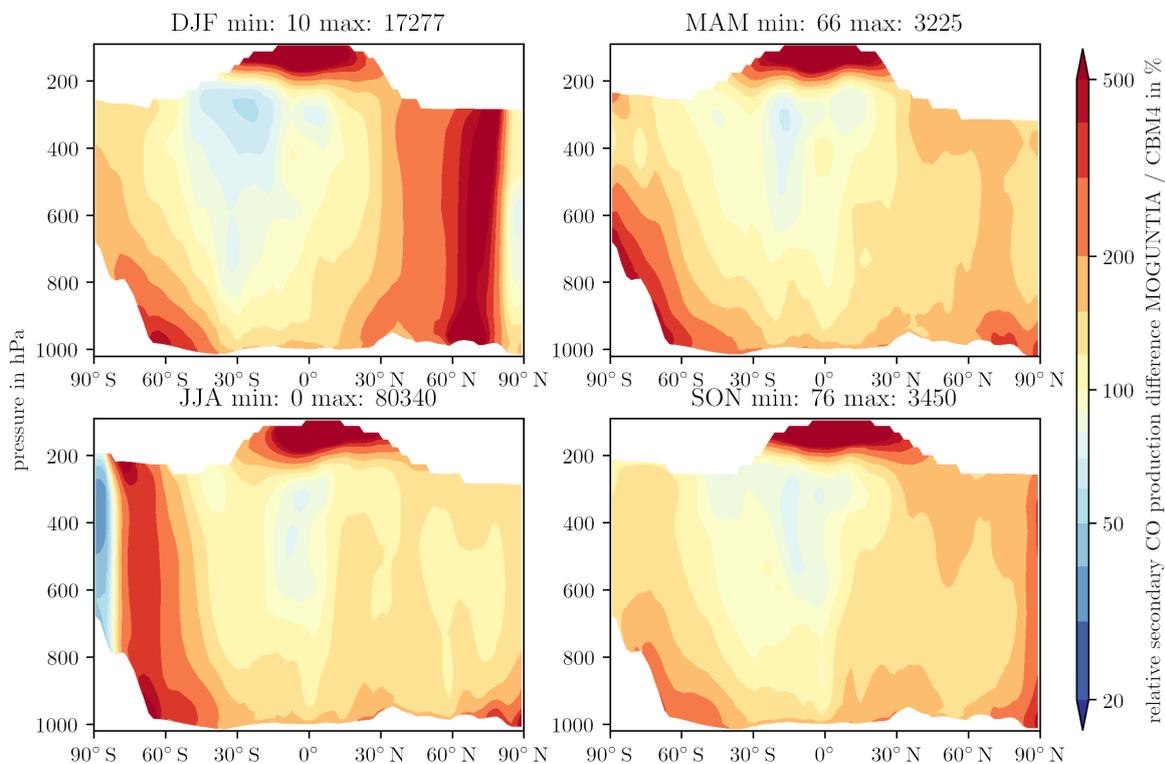
Figure A.30: Relative difference between Figs. A.28 and A.27. Aforementioned (Fig. A.29) differences equate to 50–100 %. Differences appear large close to the Poles only because of very low absolute values there.
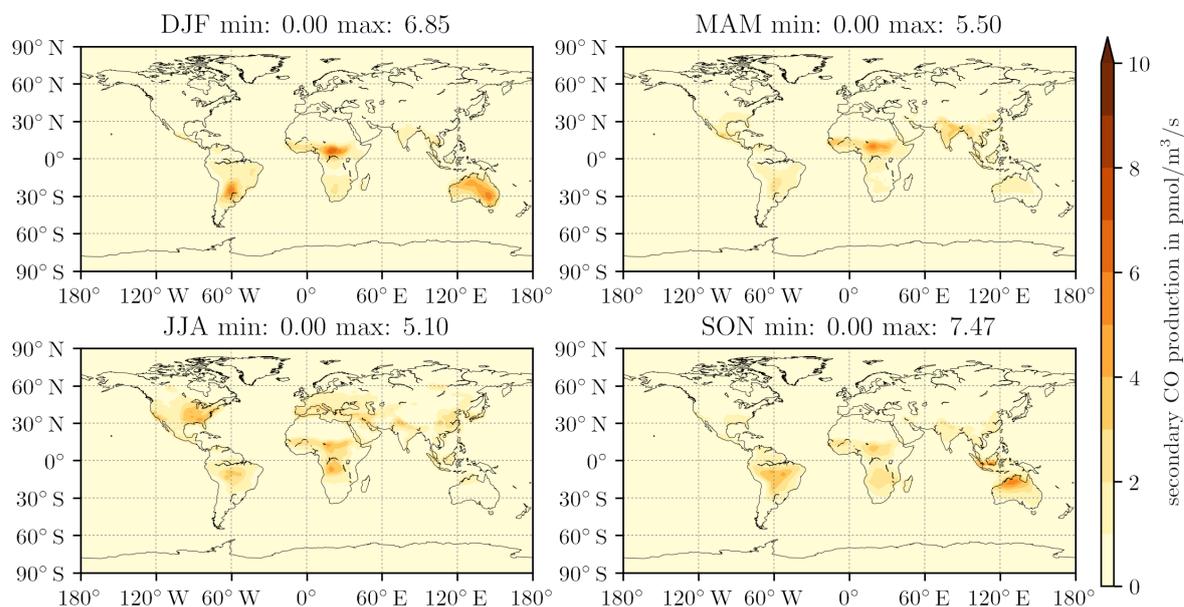
Figure A.31: Seasonal mean of surface layer secondary CO production rates taken from TM5-MP with CBM4 chemistry for 2006.
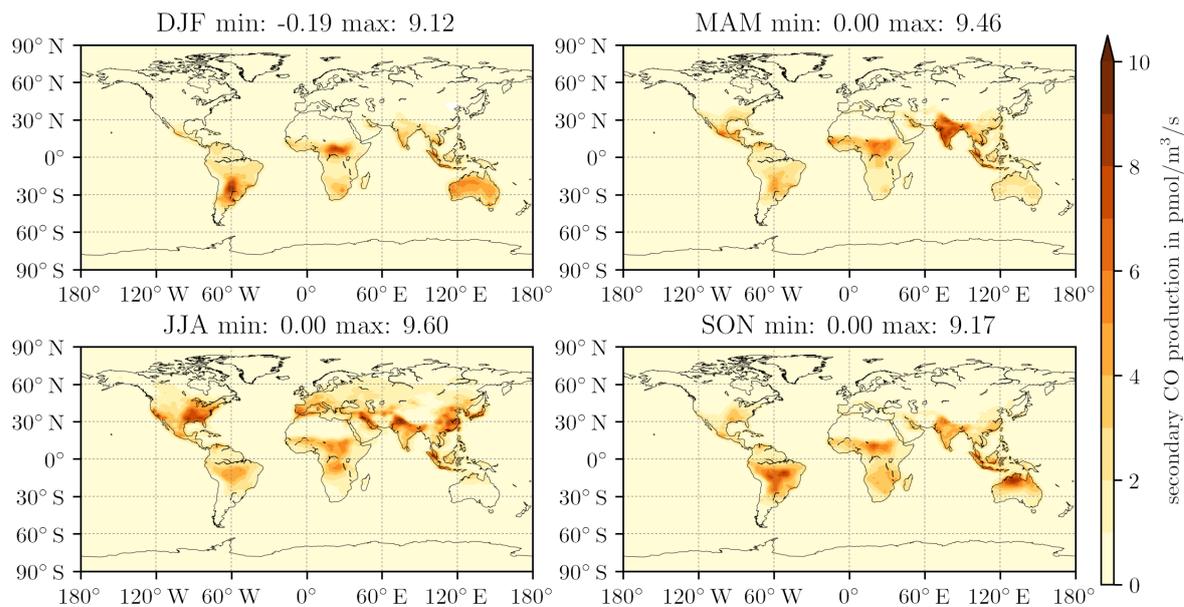
Figure A.32: Same as Fig A.31, but taken from TM5-MP with MOGUNTIA chemistry for 2018.

Figure A.33: Absolute differences between Figs. A.32 and A.31, highlighting the vastly higher production rates with the MOGUNTIA chemistry, especially in India during northern hemispheric spring and all over Asia and in southern North America during northern hemispheric summer.

Table A.2: Optimizer settings for the emission categories. These settings are used in all inversion experiments presented in this thesis.

category	a priori error	temporal correlation	spatial correlation	state resolution
secondary	20 %	9.5 months	1000 km	monthly
anthropogenic	10 %	9.5 months	2000 km	monthly
biomass burning	100 %	0.1 months	200 km	daily

Table A.3: Station-wise mean error-weighted mismatch between flasks and model for all Chapter 4 inversions.

station	FINN2.4		reference		MOGOH		MEGAN		CB4prod	
	apri	apos	apri	apos	apri	apos	apri	apos	apri	apos
ALT	9.35	4.82	8.78	4.91	77.20	3.68	65.38	3.58	18.78	4.51
ASC	2.65	1.43	2.50	1.60	1.83	1.31	1.72	1.29	2.26	1.58
ASK	14.70	11.76	14.21	11.72	40.15	8.79	31.64	8.59	37.50	15.82
AZR	57.80	6.82	53.36	6.88	58.90	5.49	45.55	5.33	10.13	6.75
BHD	11.31	1.27	9.90	1.31	4.05	1.43	6.03	1.28	3.69	1.23
BKT	3.69	1.39	3.44	1.41	2.59	1.60	2.22	1.55	4.11	2.44
BMW	25.42	5.55	23.87	5.45	57.78	5.68	48.02	5.55	15.32	5.90
BRW	4.95	1.86	4.06	1.79	19.47	2.03	15.88	1.95	4.26	1.94
CBA	21.77	3.87	19.68	3.87	75.54	4.14	60.54	3.94	14.09	4.18
CGO	4.83	0.39	4.08	0.39	1.61	0.36	2.66	0.37	1.22	0.39
CHR	44.66	1.37	41.27	1.33	2.64	1.39	4.15	1.37	18.30	1.29
CIB	5.09	3.78	4.97	3.70	30.86	3.73	23.10	3.34	14.64	3.72
CPT	8.17	1.24	6.84	1.21	2.67	1.37	4.12	1.35	4.21	1.54
CRZ	9.34	1.16	7.68	1.16	1.99	0.94	3.92	0.94	1.71	1.16
DRP	17.63	0.36	14.45	0.35	3.30	0.27	7.47	0.27	1.84	0.38
EIC	6.40	6.32	5.95	6.31	8.79	6.84	7.58	6.78	8.54	7.35
GMI	40.97	6.76	39.64	6.77	20.62	8.85	17.17	8.68	14.46	8.01
HBA	23.89	0.27	21.10	0.28	6.76	0.19	12.05	0.18	3.29	0.07
HPB	24.81	4.95	26.14	4.94	67.64	4.17	52.75	3.89	63.20	6.35
ICE	18.22	4.60	16.69	4.59	83.04	5.33	69.45	5.17	19.04	5.01
IZO	34.98	7.24	33.54	7.21	36.18	5.62	29.39	5.45	22.86	9.42
KEY	52.60	16.88	50.74	16.70	26.23	14.14	18.67	12.64	22.96	19.46
KUM	38.65	1.20	36.82	1.20	11.03	1.08	7.78	1.08	5.78	1.37
LMP	19.73	4.05	19.37	4.06	30.26	3.51	23.97	3.46	18.92	4.70
MEX	42.44	1.39	40.14	1.50	3.74	1.09	3.12	1.06	7.74	1.35
MHD	29.21	5.27	26.78	5.28	61.85	5.46	48.43	5.38	12.34	5.34
MID	41.72	8.72	39.62	8.79	35.17	7.05	28.58	6.88	11.62	8.45
MLO	51.42	1.40	49.18	1.42	4.51	1.88	3.24	1.84	11.89	1.72
NAT	3.88	3.15	4.31	3.44	4.69	3.28	4.32	3.18	8.68	4.24
NMB	6.61	2.80	6.58	2.94	5.05	2.94	5.15	2.66	7.30	3.88
NWR	52.03	10.78	49.70	10.80	48.48	9.86	39.62	9.63	26.88	11.15
OXK	11.46	6.04	12.46	6.17	67.08	6.46	55.06	6.19	39.58	6.30
PAL	7.11	5.11	7.29	5.16	117.91	5.07	100.58	4.98	33.48	4.91
PSA	8.90	0.27	7.14	0.27	1.12	0.21	3.34	0.18	0.66	0.40
RPB	75.91	4.36	71.96	4.28	13.85	2.30	10.38	2.29	15.54	4.42
SEY	6.32	3.09	5.56	3.14	7.33	3.81	5.79	3.73	3.73	3.18
SHM	5.58	1.29	5.44	1.33	69.72	1.86	54.58	1.74	13.47	1.35
SMO	20.44	0.49	18.48	0.50	0.59	0.30	0.96	0.31	4.62	0.50
SPO	14.92	1.37	12.63	1.35	3.92	0.79	6.70	0.79	3.20	1.15
SUM	14.90	3.48	13.85	3.48	100.93	3.08	84.80	3.04	25.18	3.37
SYO	9.99	0.37	8.15	0.37	2.18	0.35	4.44	0.32	1.19	0.33
TIK	6.23	2.53	5.80	2.66	10.90	3.00	8.36	2.96	3.13	4.15
USH	7.52	0.64	5.82	0.64	1.03	0.72	2.89	0.67	0.64	0.81
ZEP	8.30	2.84	7.65	2.86	99.61	2.92	84.71	2.86	25.91	2.98

A.3 Additional figures and tables for Chapter 5

Table A.4: Global prior and posterior budgets for the inversion experiments described in Chapter 5, as a sum for the global and the zooming regions. The zooming column combines masses going into and coming from the communication cells between the zooming regions. Note that the annual rates (Tg CO yr^{-1}) are extrapolated from the emissions during the inversion period, June to December (7 months). Since, except for the observations and inflation factors used, the basic setup is identical for all main inversion experiments presented here, they share the same prior budget.

		masses in Tg CO		losses in Tg CO yr^{-1}		zooming in Tg CO yr^{-1}	emitted in Tg CO yr^{-1}			
		initial	final	chemical	dry dep.		total	secondary	biomass	fossil fuel
spin-up	prior	646	670	-2925	-212	26	3159	1991	532	637
	posterior	646	521	-2351	-182	-78	2359	1340	383	636
main inversions	prior	556	739	-2995	-216	114	3411	2179	613	618
sat. v1.2	posterior	556	585	-2488	-187	22	2703	1639	546	518
full sat. v1.2	posterior	556	588	-2503	-188	23	2722	1652	555	516
sat. only v1.2	posterior	556	581	-2483	-184	17	2693	1635	548	510
sat. v1.8	posterior	556	584	-2486	-187	21	2700	1637	544	519
full sat. v1.8	posterior	556	586	-2501	-188	22	2718	1649	552	517
sat. only v1.8	posterior	556	578	-2477	-184	16	2683	1626	545	512
adaptive	posterior	556	578	-2471	-187	18	2677	1621	527	530

Table A.5: Error-weighted mismatches between observations and model for the inversion experiments before and after updating to the latest satellite retrieval version and for the experiment that uses the adaptive error inflation approach. The first eight rows give the mean mismatches to different zonal subsets of the flask measurements. The last two rows contain the total mismatch to the satellite observations, scaled down by 10^3 for readability.

stations		sat. v1.2	full sat. v1.2	sat. only v1.2	sat. v1.8	full sat. v1.8	sat. only v1.8	adaptive
all	prior	20.60	20.60	20.60	20.60	20.60	20.60	20.60
	poste	3.65	3.52	9.29	3.68	3.53	9.30	4.39
north of 23° N	prior	19.94	19.94	19.94	19.94	19.94	19.94	19.94
	poste	5.44	5.18	15.76	5.48	5.19	15.79	6.64
tropics (23° S − 23° N)	prior	29.13	29.13	29.13	29.13	29.13	29.13	29.13
	poste	2.33	2.33	3.92	2.38	2.34	4.00	2.81
south of 23° S	prior	7.92	7.92	7.92	7.92	7.92	7.92	7.92
	poste	1.33	1.34	1.95	1.33	1.33	1.78	1.37
satellite	prior	80.95	65.88	80.95	87.08	70.97	87.08	312.90
	poste	7.97	7.78	7.01	7.90	7.70	6.83	31.43

A.4   Additional figures and tables for Chapter 6
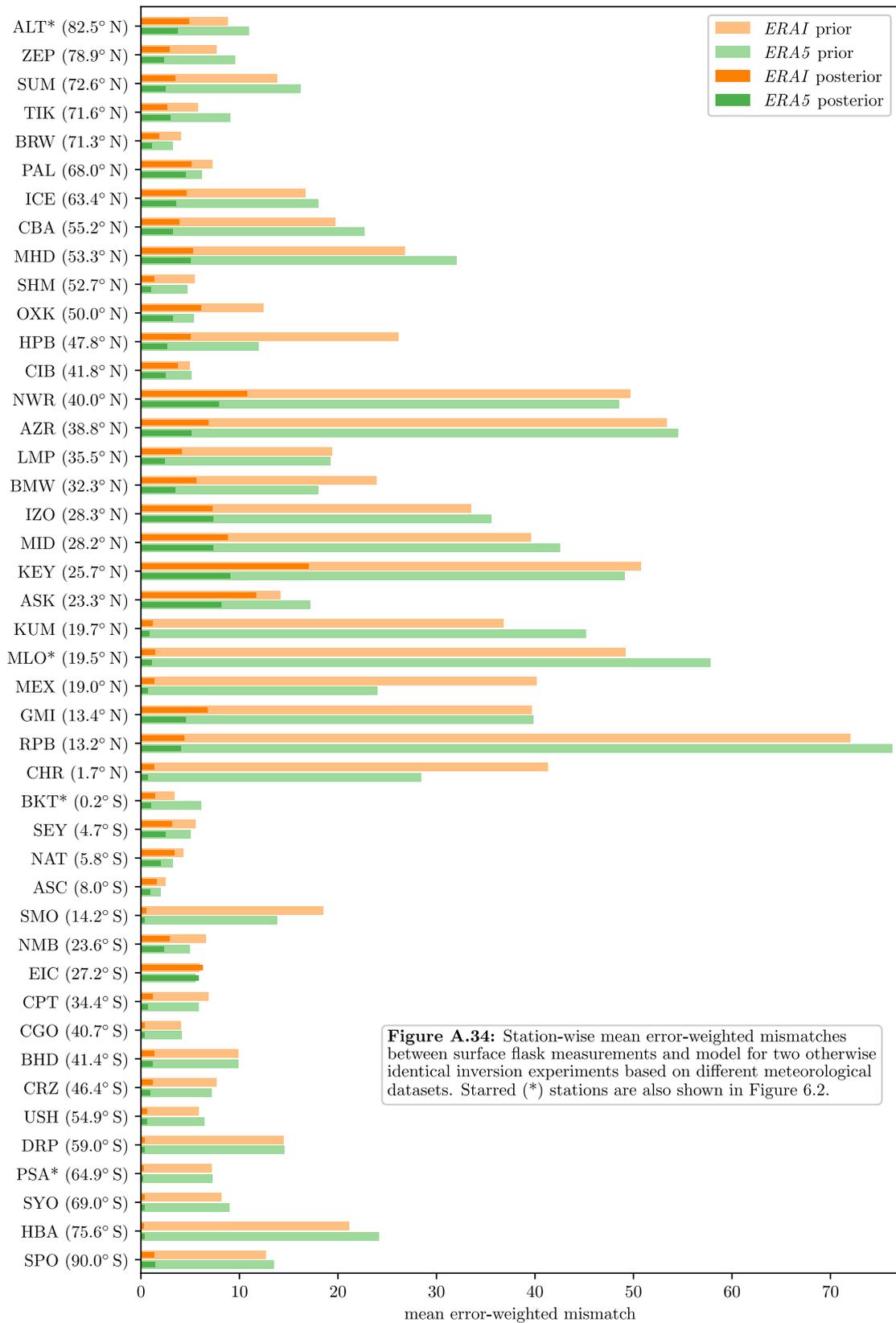
Figure A.34: Station-wise mean error-weighted mismatches between surface flask measurements and model for two otherwise identical inversion experiments based on different meteorological datasets. Starred (*) stations are also shown in Figure 6.2.

Table A.6: Station-wise mean error-weighted mismatch between flasks and model for two otherwise identical inversion experiments based on different meteorological datasets. A graphical representation of this data can be found in Figure A.34.

station	ERAI prior	ERAI posterior	ERA5 prior	ERA5 posterior
ALT	8.79	4.92	10.96	3.75
ASC	2.50	1.61	2.01	0.94
ASK	14.19	11.68	17.17	8.19
AZR	53.38	6.88	54.57	5.11
BHD	9.91	1.32	9.89	1.21
BKT	3.44	1.41	6.15	1.01
BMW	23.89	5.59	18.00	3.49
BRW	4.06	1.83	3.27	1.15
CBA	19.70	3.87	22.72	3.23
CGO	4.09	0.40	4.18	0.40
CHR	41.28	1.38	28.44	0.73
CIB	4.96	3.76	5.09	2.47
CPT	6.85	1.21	5.87	0.71
CRZ	7.69	1.18	7.15	0.95
DRP	14.46	0.36	14.57	0.34
EIC	5.95	6.28	5.57	5.83
GMI	39.64	6.76	39.83	4.56
HBA	21.11	0.29	24.19	0.33
HPB	26.14	5.01	11.91	2.65
ICE	16.70	4.61	17.99	3.55
IZO	33.54	7.25	35.61	7.31
KEY	50.75	17.00	49.12	9.07
KUM	36.83	1.20	45.22	0.85
LMP	19.38	4.18	19.25	2.44
MEX	40.15	1.38	23.98	0.68
MHD	26.79	5.31	32.03	5.07
MID	39.63	8.82	42.55	7.36
MLO	49.20	1.40	57.85	1.08
NAT	4.31	3.41	3.22	1.98
NMB	6.58	2.95	4.99	2.33
NWR	49.70	10.83	48.56	7.92
OXK	12.46	6.14	5.38	3.24
PAL	7.29	5.16	6.17	4.55
PSA	7.15	0.27	7.22	0.18
RPB	72.00	4.38	76.28	4.09
SEY	5.58	3.14	5.06	2.48
SHM	5.45	1.35	4.71	1.06
SMO	18.49	0.50	13.79	0.37
SPO	12.64	1.37	13.54	1.44
SUM	13.86	3.51	16.18	2.49
SYO	8.15	0.37	8.98	0.34
TIK	5.81	2.63	9.06	2.99
USH	5.83	0.64	6.44	0.58
ZEP	7.66	2.88	9.55	2.37

Table A.7: Global prior and posterior budgets for two otherwise identical inversion experiments based on different meteorological datasets, as a sum for the global and the zooming regions. The zooming column combines masses going into and coming from the communication cells between the zooming regions. Note that the annual rates (Tg CO yr^{-1}) are extrapolated from the emissions during the inversion period, June to December (7 months).

		masses in Tg CO		losses in Tg CO yr^{-1}		zooming in Tg CO yr^{-1}	emitted in Tg CO yr^{-1}			
		initial	final	chemical	deposition		total	secondary	biomass	fossil fuel
ERAI spin-up	prior	646	670	-2925	-212	26	3159	1991	532	637
	posterior	646	521	-2351	-182	-78	2359	1340	383	636
ERA5 spin-up	prior	646	674	-2958	-173	28	3159	1991	532	637
	posterior	646	520	-2373	-147	-78	2345	1345	391	608
ERAI main	prior	556	739	-2995	-216	114	3411	2179	613	618
	posterior	556	585	-2488	-187	22	2703	1639	546	518
ERA5 main	prior	555	744	-3027	-178	115	3411	2179	613	618
	posterior	555	583	-2508	-154	19	2690	1637	546	507

Bibliography

AKAGI, S. K., YOKELSON, R. J., WIEDINMYER, C., ALVARADO, M. J., REID, J. S., KARL, T., CROUNSE, J. D., & WENNBERG, P. O. (2011). Emission factors for open and domestic biomass burning for use in atmospheric models. *Atmospheric Chemistry and Physics*, 11(9), 4039–4072.

ARELLANO, A. F., KASIBHATLA, P. S., GIGLIO, L., VAN DER WERF, G. R., & RANDERSON, J. T. (2004). Top-down estimates of global CO sources using MOPITT measurements. *Geophysical Research Letters*, 31(1).

BERGAMASCHI, P., FRANKENBERG, C., MEIRINK, J. F., KROL, M., VILLANI, M. G., HOUWELING, S., DENTENER, F., DLUGOKENCKY, E. J., MILLER, J. B., GATTI, L. V., ENGEL, A., & LEVIN, I. (2009). Inverse modeling of global and regional CH_4 emissions using SCIAMACHY satellite retrievals. *Journal of Geophysical Research: Atmospheres*, 114(D22).

BERGAMASCHI, P., FRANKENBERG, C., MEIRINK, J. F., KROL, M. C., DENTENER, F., WAGNER, T., PLATT, U., KAPLAN, J. O., KÖRNER, S., HEIMANN, M., DLUGOKENCKY, E. J., & GOEDE, A. (2007). Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations. *Journal of Geophysical Research*, 112(D2).

BERGAMASCHI, P., HEIN, R., HEIMANN, M., & CRUTZEN, P. J. (2000). Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratios. *Journal of Geophysical Research: Atmospheres*, 105(D2), 1909–1927.

BERGAMASCHI, P., KROL, M. C., DENTENER, F., VERMEULEN, A., MEINHARDT, F., GRAUL, R., RAMONET, M., PETERS, W., & DLUGOKENCKY, E. J. (2005). Inverse modelling of national and European CH_4 emissions using the atmospheric zoom model TM5. *Atmospheric Chemistry and Physics*, 5(9), 2431–2460.

BERGER, A. & TRICOT, C. (1992). The greenhouse effect. *Surveys in Geophysics*, 13(6), 523–549.

BIBLIOGRAPHY

BERKVENS, P. J. F., BOTCHEV, M. A., LIOEN, W. M., & VERWER, J. G. (1999). A zooming technique for wind transport of air pollution. *Report MAS-R9921 (CWI, Amsterdam, 1999)*.

BOERSMA, K. F., VINKEN, G. C. M., & ESKES, H. J. (2016). Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV–vis tropospheric column retrievals. *Geoscientific Model Development*, 9(2), 875–898.

BORZENKOVA, I. I. & TURCHINOVICH, I. Y. (2009). History of atmospheric composition. In G. V. Gruza (Ed.), *Environmental structure and function: climate system*, volume 2 (pp. 184–205). Eolss Publishers Co. Ltd.

BRASSEUR, G. P. & JACOB, D. J. (2017). *Modeling of Atmospheric Chemistry*. Cambridge University Press.

BUTLER, T. M., RAYNER, P. J., SIMMONDS, I., & LAWRENCE, M. G. (2005). Simultaneous mass balance inverse modeling of methane and carbon monoxide. *Journal of Geophysical Research*, 110(D21).

CHEN, D., WANG, Y., MCELROY, M. B., HE, K., YANTOSCA, R. M., & LE SAGER, P. (2009). Regional CO pollution and export in China simulated by the high-resolution nested-grid GEOS-Chem model. *Atmospheric Chemistry and Physics*, 9(11), 3825–3839.

CHEVALLIER, F. (2007). Impact of correlated observation errors on inverted CO_2 surface fluxes from OCO measurements. *Geophysical Research Letters*, 34(24).

CLERBAUX, C., BOYNARD, A., CLARISSE, L., GEORGE, M., HADJI-LAZARO, J., HERBIN, H., HURTMANS, D., POMMIER, M., RAZAVI, A., TURQUETY, S., WESPES, C., & COHEUR, P.-F. (2009). Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmospheric Chemistry and Physics*, 9(16), 6041–6054.

DANIEL, J. S. & SOLOMON, S. (1998). On the climate forcing of carbon monoxide. *Journal of Geophysical Research: Atmospheres*, 103(D11), 13249–13260.

DASKALAKIS, N., GALLARDO, L., KANAKIDOU, M., NÜSS, J. R., MENARES, C., RONDANELLI, R., THOMPSON, A. M., & VREKOUSSIS, M. (2022). Impact of biomass burning and stratospheric intrusions in the remote South Pacific Ocean troposphere. *Atmospheric Chemistry and Physics*, 22(6), 4075–4099.

DE WACHTER, E., BARRET, B., LE FLOCHMOËN, E., PAVELIN, E., MATRICARDI, M., CLERBAUX, C., HADJI-LAZARO, J., GEORGE, M., HURTMANS, D., COHEUR, P.-F.,

NEDELEC, P., & CAMMAS, J. P. (2012). Retrieval of MetOp-A/IASI CO profiles and validation with MOZAIC data. *Atmospheric Measurement Techniques*, 5(11), 2843–2857.

DEE, D. P., UPPALA, S. M., SIMMONS, A. J., BERRISFORD, P., POLI, P., KOBAYASHI, S., ANDRAE, U., BALMASEDA, M. A., BALSAMO, G., BAUER, P., BECHTOLD, P., BELJAARS, A. C. M., VAN DE BERG, L., BIDLOT, J., BORMANN, N., DELSOL, C., DRAGANI, R., FUENTES, M., GEER, A. J., HAIMBERGER, L., HEALY, S. B., HERSBACH, H., HÓLM, E. V., ISAKSEN, L., KÅLLBERG, P., KÖHLER, M., MATRICARDI, M., MCNALLY, A. P., MONGE-SANZ, B. M., MORCRETTE, J.-J., PARK, B.-K., PEUBEY, C., DE ROSNAY, P., TAVOLATO, C., THÉPAUT, J.-N., & VITART, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597.

DRUMMOND, J. R., ZOU, J., NICHITIU, F., KAR, J., DESCHAMBAUT, R., & HACKETT, J. (2010). A review of 9-year performance and operation of the MOPITT instrument. *Advances in Space Research*, 45(6), 760–774.

EMMONS, L. K., SCHWANTES, R. H., ORLANDO, J. J., TYNDALL, G., KINNISON, D., LAMARQUE, J.-F., MARSH, D., MILLS, M. J., TILMES, S., BARDEEN, C., BUCHHOLZ, R. R., CONLEY, A., GETTELMAN, A., GARCIA, R., SIMPSON, I., BLAKE, D. R., MEINARDI, S., & PÉTRON, G. (2020). The chemistry mechanism in the community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(4).

ESKES, H. J., PITERS, A. J. M., LEVELT, P. F., ALLAART, M. A. F., & KELDER, H. M. (1999). Variational assimilation of GOME total-column ozone satellite data in a 2d latitude–longitude tracer-transport model. *Journal of the Atmospheric Sciences*, 56(20), 3560–3572.

ESKES, H. J., VAN VELTHOVEN, P. F. J., VALKS, P. J. M., & KELDER, H. M. (2003). Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model. *Quarterly Journal of the Royal Meteorological Society*, 129(590), 1663–1681.

EYRING, V., BONY, S., MEEHL, G. A., SENIOR, C. A., STEVENS, B., STOUFFER, R. J., & TAYLOR, K. E. (2016). Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958.

BIBLIOGRAPHY

Fisher, M. & Lary, D. J. (1995). Lagrangian four-dimensional variational data assimilation of chemical species. *Quarterly Journal of the Royal Meteorological Society*, 121(527), 1681–1704.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., & Zhang, H. (2021). The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 923–1054). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Carouge, C., Clerbaux, C., Coheur, P.-F., George, M., Hurtmans, D., & Szopa, S. (2009). On the capability of IASI measurements to inform about CO surface emissions. *Atmospheric Chemistry and Physics*, 9(22), 8735–8743.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., & Fried, A. (2012). The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system. *Atmospheric Chemistry and Physics*, 12(15), 6699–6721.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Szopa, S., Deeter, M. N., & Clerbaux, C. (2011). Ten years of CO emissions as seen from measurements of pollution in the troposphere (MOPITT). *Journal of Geophysical Research*, 116(D5).

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., Mc-Grath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I.,

Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., & Zheng, B. (2022). Global carbon budget 2022. *Earth System Science Data*, 14(11), 4811–4900.

Fujimori, S., Hasegawa, T., Masui, T., Takahashi, K., Herran, D. S., Dai, H., Hijioka, Y., & Kainuma, M. (2017). SSP3: AIM implementation of shared socioeconomic pathways. *Global Environmental Change*, 42, 268–283.

Ganzeveld, L., Lelieveld, J., & Roelofs, G.-J. (1998). A dry deposition parameterization for sulfur oxides in a chemistry and general circulation model. *Journal of Geophysical Research: Atmospheres*, 103(D5), 5679–5694.

Gerbig, C., Schmitgen, S., Kley, D., Volz-Thomas, A., Dewey, K., & Haaks, D. (1999). An improved fast-response vacuum-UV resonance fluorescence CO instrument. *Journal of Geophysical Research: Atmospheres*, 104(D1), 1699–1704.

Gidden, M., Riahi, K., Smith, S., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., & Takahashi, K. (2018). input4MIPs.CMIP6.ScenarioMIP.IAMC.IAMC-AIM-ssp370-1-1, version 20180620.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., & Takahashi, K. (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geoscientific Model Development*, 12(4), 1443–1475.

Giglio, L., Randerson, J. T., & van der Werf, G. R. (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences*, 118(1), 317–328.

Gilbert, J. C. & Lemaréchal, C. (1989). Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, 45(1-3), 407–435.

BIBLIOGRAPHY

GROS, V., WILLIAMS, J., LAWRENCE, M. G., VON KUHLMANN, R., VAN AARDENNE, J., ATLAS, E., CHUCK, A., EDWARDS, D. P., STROUD, V., & KROL, M. C. (2004). Tracing the origin and ages of interlaced atmospheric pollution events over the tropical Atlantic Ocean with in situ measurements, satellites, trajectories, emission inventories, and global models. *Journal of Geophysical Research: Atmospheres*, 109(D22).

GROS, V., WILLIAMS, J., VAN AARDENNE, J. A., SALISBURY, G., HOFMANN, R., LAWRENCE, M. G., VON KUHLMANN, R., LELIEVELD, J., KROL, M. C., BERRESHEIM, H., LOBERT, J. M., & ATLAS, E. (2003). Origin of anthropogenic hydrocarbons and halocarbons measured in the summertime European outflow (on Crete in 2001). *Atmospheric Chemistry and Physics Discussions*, 3(2), 1893–1923.

GUENTHER, A. B., JIANG, X., HEALD, C. L., SAKULYANONTVITTAYA, T., DUHL, T., EMMONS, L. K., & WANG, X. (2012). The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, 5(6), 1471–1492.

HEALD, C. L., JACOB, D. J., JONES, D. B. A., PALMER, P. I., LOGAN, J. A., STREETS, D. G., SACHSE, G. W., GILLE, J. C., HOFFMAN, R. N., & NEHRKORN, T. (2004). Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide. *Journal of Geophysical Research: Atmospheres*, 109(D23).

HEILMAN, W. E., LIU, Y., URBANSKI, S., KOVALEV, V., & MICKLER, R. (2014). Wildland fire emissions, carbon, and climate: Plume rise, atmospheric transport, and chemistry processes. *Forest Ecology and Management*, 317, 70–79.

HERSBACH, H., BELL, B., BERRISFORD, P., HIRAHARA, S., HORÁNYI, A., MUÑOZ-SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R., SCHEPERS, D., SIMMONS, A., SOCI, C., ABDALLA, S., ABELLAN, X., BALSAMO, G., BECHTOLD, P., BIAVATI, G., BIDLOT, J., BONAVITA, M., CHIARA, G., DAHLGREN, P., DEE, D., DIAMANTAKIS, M., DRAGANI, R., FLEMMING, J., FORBES, R., FUENTES, M., GEER, A., HAIMBERGER, L., HEALY, S., HOGAN, R. J., HÓLM, E., JANISKOVÁ, M., KEELEY, S., LALOYAUX, P., LOPEZ, P., LUPU, C., RADNOTI, G., ROSNAY, P., ROZUM, I., VAMBORG, F., VILLAUME, S., & THÉPAUT, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049.

HILTON, F., COLLARD, A., GUIDARD, V., RANDRIAMAMPIANINA, R., & SCHWAERZ, M. (2009). Assimilation of IASI radiances at European NWP centres. In *Workshop on the assimilation of IASI data in NWP* (pp. 39–48).: ECMWF.

HOLLOWAY, T., LEVY, H., & KASIBHATLA, P. (2000). Global distribution of carbon monoxide. *Journal of Geophysical Research: Atmospheres*, 105(D10), 12123–12147.

HOOGHIEMSTRA, P. B., KROL, M. C., BERGAMASCHI, P., DE LAAT, A. T. J., VAN DER WERF, G. R., NOVELLI, P. C., DEETER, M. N., ABEN, I., & RÖCKMANN, T. (2012a). Comparing optimized CO emission estimates using MOPITT or NOAA surface network observations. *Journal of Geophysical Research: Atmospheres*, 117(D6), 1–23.

HOOGHIEMSTRA, P. B., KROL, M. C., MEIRINK, J. F., BERGAMASCHI, P., VAN DER WERF, G. R., NOVELLI, P. C., ABEN, I., & RÖCKMANN, T. (2011). Optimizing global CO emission estimates using a four-dimensional variational data assimilation system and surface network observations. *Atmospheric Chemistry and Physics*, 11(10), 4705–4723.

HOOGHIEMSTRA, P. B., KROL, M. C., VAN LEEUWEN, T. T., VAN DER WERF, G. R., NOVELLI, P. C., DEETER, M. N., ABEN, I., & RÖCKMANN, T. (2012b). Interannual variability of carbon monoxide emission estimates over South America from 2006 to 2010. *Journal of Geophysical Research, D: Atmospheres*, 117(D15), 1–25.

HUIJNEN, V., WILLIAMS, J., VAN WEELE, M., VAN NOIJE, T., KROL, M. C., DENTENER, F., SEGERS, A., HOUWELING, S., PETERS, W., DE LAAT, J., BOERSMA, F., BERGAMASCHI, P., VAN VELTHOVEN, P., LE SAGER, P., ESKES, H., ALKEMADE, F., SCHEELE, R., NÉDÉLEC, P., & PÄTZ, H.-W. (2010). The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0. *Geoscientific Model Development*, 3(2), 445–473.

JIANG, Z., JONES, D. B. A., KOPACZ, M., LIU, J., HENZE, D. K., & HEALD, C. (2011). Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations. *Journal of Geophysical Research*, 116(D15).

JIANG, Z., JONES, D. B. A., WORDEN, H. M., & HENZE, D. K. (2015). Sensitivity of top-down CO source estimates to the modeled vertical structure in atmospheric CO. *Atmospheric Chemistry and Physics*, 15(3), 1521–1537.

JIANG, Z., WORDEN, J. R., WORDEN, H., DEETER, M., JONES, D. B. A., ARELLANO, A. F., & HENZE, D. K. (2017). A 15-year record of CO emissions constrained by MOPITT CO observations. *Atmospheric Chemistry and Physics*, 17(7), 4565–4583.

KANAYA, Y., YAMAJI, K., MIYAKAWA, T., TAKETANI, F., ZHU, C., CHOI, Y., KOMAZAKI, Y., IKEDA, K., KONDO, Y., & KLIMONT, Z. (2020). Rapid reduction in black carbon emissions from China: evidence from 2009–2019 observations on Fukue Island, Japan. *Atmospheric Chemistry and Physics*, 20(11), 6339–6356.

BIBLIOGRAPHY

Kang, H., Zhu, B., van der A, R. J., Zhu, C., de Leeuw, G., Hou, X., & Gao, J. (2019). Natural and anthropogenic contributions to long-term variations of SO_2, NO_2, CO, and AOD over East China. *Atmospheric Research*, 215, 284–293.

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., & Nedelec, P. (2010). Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES). *Atmospheric Chemistry and Physics*, 10(3), 855–876.

Krol, M. C., de Bruine, M., Killaars, L., Ouwersloot, H., Pozzer, A., Yin, Y., Chevallier, F., Bousquet, P., Patra, P., Belikov, D., Maksyutov, S., Dhomse, S., Feng, W., & Chipperfield, M. P. (2018). Age of air as a diagnostic for transport timescales in global models. *Geoscientific Model Development*, 11(8), 3109–3130.

Krol, M. C., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., & Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chemistry and Physics*, 5(2), 417–432.

Krol, M. C., Lelieveld, J., Oram, D. E., Sturrock, G. A., Penkett, S. A., Brenninkmeijer, C. A. M., Gros, V., Williams, J., & Scheeren, H. A. (2003). Continuing emissions of methyl chloroform from Europe. *Nature*, 421(6919), 131–135.

Krol, M. C., Meirink, J. F., Bergamaschi, P., Mak, J. E., Lowe, D., Jöckel, P., Houweling, S., & Röckmann, T. (2008). What can ^{14}CO measurements tell us about OH? *Atmospheric Chemistry and Physics*, 8(16), 5033–5044.

Krol, M. C., Peters, W., Berkvens, P., & Botchev, M. (2002). A new algorithm for two-way nesting in global models: Principles and applications. In *Air Pollution Modelling and Simulation* (pp. 225–234). Springer Berlin Heidelberg.

Krol, M. C., Peters, W., Hooghiemstra, P., George, M., Clerbaux, C., Hurtmans, D., McInerney, D., Sedano, F., Bergamaschi, P., Hajj, M. E., Kaiser, J. W., Fisher, D., Yershov, V., & Muller, J.-P. (2013). How much CO was emitted by the 2010 fires around Moscow? *Atmospheric Chemistry and Physics*, 13(9), 4737–4747.

Lahoz, W. A., Errera, Q., Swinbank, R., & Fonteyn, D. (2007). Data assimilation of stratospheric constituents: a review. *Atmospheric Chemistry and Physics*, 7(22), 5745–5773.

Lee, D. S., Köhler, I., Grobler, E., Rohrer, F., Sausen, R., Gallardo-Klenner, L., Olivier, J. G. J., Dentener, F. J., & Bouwman, A. F. (1997). Estimations of global NO_x, emissions and their uncertainties. *Atmospheric Environment*, 31(12), 1735–1749.

Lelieveld, J., Gromov, S., Pozzer, A., & Taraborrelli, D. (2016). Global tropospheric hydroxyl distribution, budget and reactivity. *Atmospheric Chemistry and Physics*, 16(19), 12477–12493.

Li, Y., Ma, Z., Han, T., Quan, W., Wang, J., Zhou, H., He, D., & Dong, F. (2020). Long-term declining in carbon monoxide (CO) at a rural site of Beijing during 2006–2018 implies the improved combustion efficiency and effective emission control. *Journal of Environmental Sciences*.

Liang, Q., Chipperfield, M. P., Fleming, E. L., Abraham, N. L., Braesicke, P., Burkholder, J. B., Daniel, J. S., Dhomse, S., Fraser, P. J., Hardiman, S. C., Jackman, C. H., Kinnison, D. E., Krummel, P. B., Montzka, S. A., Morgenstern, O., McCulloch, A., Mühle, J., Newman, P. A., Orkin, V. L., Pitari, G., Prinn, R. G., Rigby, M., Rozanov, E., Stenke, A., Tummon, F., Velders, G. J. M., Visioni, D., & Weiss, R. F. (2017). Deriving global OH abundance and atmospheric lifetimes for long-lived gases: A search for CH_3CCl_3 alternatives. *Journal of Geophysical Research: Atmospheres*, 122(21).

Liu, S., Fang, S., Liang, M., Sun, W., & Feng, Z. (2019). Temporal patterns and source regions of atmospheric carbon monoxide at two background stations in China. *Atmospheric Research*, 220, 169–180.

Liu, Z., Chen, X., Cai, J., Baležentis, T., & Li, Y. (2020). The impact of "Coal to Gas" policy on air quality: Evidence from Beijing, China. *Energies*, 13(15), 3876.

Liu, Z. & Rabier, F. (2003). The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quarterly Journal of the Royal Meteorological Society*, 129(594), 3013–3035.

Logan, J. A., Prather, M. J., Wofsy, S. C., & McElroy, M. B. (1981). Tropospheric chemistry: A global perspective. *Journal of Geophysical Research: Oceans*, 86(C8), 7210–7254.

Mckee, D. J. (1993). *Tropospheric Ozone: Human Health and Agricultural Impacts*. CRC Press.

BIBLIOGRAPHY

MEIRINK, J. F., BERGAMASCHI, P., FRANKENBERG, C., D'AMELIO, M. T. S., DLU-
GOKENCKY, E. J., GATTI, L. V., HOUWELING, S., MILLER, J. B., RÖCKMANN,
T., VILLANI, M. G., & KROL, M. C. (2008a). Four-dimensional variational data
assimilation for inverse modeling of atmospheric methane emissions: Analysis of
SCIAMACHY observations. *Journal of Geophysical Research*, 113(D17).

MEIRINK, J. F., BERGAMASCHI, P., & KROL, M. C. (2008b). Four-dimensional
variational data assimilation for inverse modelling of atmospheric methane emis-
sions: method and comparison with synthesis inversion. *Atmospheric Chemistry and
Physics*, 8(21), 6341–6353.

MEIRINK, J. F., ESKES, H. J., & GOEDE, A. P. H. (2006). Sensitivity analysis of
methane emissions derived from SCIAMACHY observations through inverse mod-
elling. *Atmospheric Chemistry and Physics*, 6(5), 1275–1292.

MIYAZAKI, K., ESKES, H. J., & SUDO, K. (2012). Global NO_x emission estimates
derived from an assimilation of OMI tropospheric NO_2 columns. *Atmospheric Chem-
istry and Physics*, 12(5), 2263–2288.

MIYAZAKI, K., ESKES, H. J., & SUDO, K. (2015). A tropospheric chemistry reanalysis
for the years 2005–2012 based on an assimilation of OMI, MLS, TES, and MOPITT
satellite data. *Atmospheric Chemistry and Physics*, 15(14), 8315–8348.

MURRAY, L. T., FIORE, A. M., SHINDELL, D. T., NAIK, V., & HOROWITZ, L. W.
(2021). Large uncertainties in global hydroxyl projections tied to fate of reactive
nitrogen and carbon. *Proceedings of the National Academy of Sciences*, 118(43).

MYRIOKEFALITAKIS, S., DASKALAKIS, N., GKOUVOUSIS, A., HILBOLL, A., VAN
NOIJE, T., WILLIAMS, J. E., LE SAGER, P., HUIJNEN, V., HOUWELING, S.,
BERGMAN, T., NÜSS, J. R., VREKOUSSIS, M., KANAKIDOU, M., & KROL, M. C.
(2020). Description and evaluation of a detailed gas-phase chemistry scheme in the
TM5-MP global chemistry transport model (r112). *Geoscientific Model Development*,
13(11), 5507–5548.

MÜLLER, J.-F., STAVRAKOU, T., BAUWENS, M., GEORGE, M., HURTMANS, D., CO-
HEUR, P.-F., CLERBAUX, C., & SWEENEY, C. (2018). Top-down CO emissions
based on IASI observations and hemispheric constraints on OH levels. *Geophysical
Research Letters*, 45(3), 1621–1629.

NAIK, V., VOULGARAKIS, A., FIORE, A. M., HOROWITZ, L. W., LAMARQUE, J.-
F., LIN, M., PRATHER, M. J., YOUNG, P. J., BERGMANN, D., CAMERON-SMITH,
P. J., CIONNI, I., COLLINS, W. J., DALSØREN, S. B., DOHERTY, R., EYRING, V.,
FALUVEGI, G., FOLBERTH, G. A., JOSSE, B., LEE, Y. H., MACKENZIE, I. A.,

NAGASHIMA, T., VAN NOIJE, T. P. C., PLUMMER, D. A., RIGHI, M., RUMBOLD, S. T., SKEIE, R., SHINDELL, D. T., STEVENSON, D. S., STRODE, S., SUDO, K., SZOPA, S., & ZENG, G. (2013). Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the atmospheric chemistry and climate model intercomparison project (ACCMIP). *Atmospheric Chemistry and Physics*, 13(10), 5277–5298.

NAUS, S., MONTZKA, S. A., PANDEY, S., BASU, S., DLUGOKENCKY, E. J., & KROL, M. C. (2019). Constraints and biases in a tropospheric two-box model of OH. *Atmospheric Chemistry and Physics*, 19(1), 407–424.

NAUS, S., MONTZKA, S. A., PATRA, P. K., & KROL, M. C. (2021). A three-dimensional-model inversion of methyl chloroform to constrain the atmospheric oxidative capacity. *Atmospheric Chemistry and Physics*, 21(6), 4809–4824.

NECHITA-BANDA, N., KROL, M. C., VAN DER WERF, G. R., KAISER, J. W., PANDEY, S., HUIJNEN, V., CLERBAUX, C., COHEUR, P., DEETER, M. N., & RÖCKMANN, T. (2018). Monitoring emissions from the 2015 Indonesian fires using CO satellite data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1760), 20170307.

NGUYEN, H. M. & WOOSTER, M. J. (2020). Advances in the estimation of high spatio-temporal resolution pan-African top-down biomass burning emissions made using geostationary fire radiative power (FRP) and MAIAC aerosol optical depth (AOD) data. *Remote Sensing of Environment*, 248, 111971.

NÜSS, J. R., DASKALAKIS, N., PIWOWARCZYK, F. G., GKOUVOUSIS, A., SCHNEISING, O., BUCHWITZ, M., KANAKIDOU, M., KROL, M. C., & VREKOUSSIS, M. (2022a). TM5-4DVAR inverse modeling suit with extensions for TROPOMI CO observations.

NÜSS, J. R., DASKALAKIS, N., PIWOWARCZYK, F. G., GKOUVOUSIS, A., SCHNEISING, O., BUCHWITZ, M., KANAKIDOU, M., KROL, M. C., & VREKOUSSIS, M. (2023). Data and scripts for manuscript "Efficacy of high-resolution satellite observations in inverse modeling of carbon monoxide emissions".

NÜSS, J. R., PIWOWARCZYK, F. G., & HILBOLL, A. (2022b). Jupyternotebooks for regridding of satellite observations into super-observations.

PALMER, P. I., JACOB, D. J., JONES, D. B. A., HEALD, C. L., YANTOSCA, R. M., LOGAN, J. A., SACHSE, G. W., & STREETS, D. G. (2003). Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific. *Journal of Geophysical Research: Atmospheres*, 108(D21).

BIBLIOGRAPHY

PATRA, P. K., HOUWELING, S., KROL, M., BOUSQUET, P., BELIKOV, D., BERGMANN, D., BIAN, H., CAMERON-SMITH, P., CHIPPERFIELD, M. P., CORBIN, K., FORTEMS-CHEINEY, A., FRASER, A., GLOOR, E., HESS, P., ITO, A., KAWA, S. R., LAW, R. M., LOH, Z., MAKSYUTOV, S., MENG, L., PALMER, P. I., PRINN, R. G., RIGBY, M., SAITO, R., & WILSON, C. (2011). TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere. *Atmospheric Chemistry and Physics*, 11(24), 12813–12837.

PATRA, P. K., KROL, M. C., MONTZKA, S. A., ARNOLD, T., ATLAS, E. L., LINTNER, B. R., STEPHENS, B. B., XIANG, B., ELKINS, J. W., FRASER, P. J., GHOSH, A., HINTSA, E. J., HURST, D. F., ISHIJIMA, K., KRUMMEL, P. B., MILLER, B. R., MIYAZAKI, K., MOORE, F. L., MÜHLE, J., O'DOHERTY, S., PRINN, R. G., STEELE, L. P., TAKIGAWA, M., WANG, H. J., WEISS, R. F., WOFSY, S. C., & YOUNG, D. (2014). Observational evidence for interhemispheric hydroxyl-radical parity. *Nature*, 513(7517), 219–223.

PETRON, G., CROTWELL, A. M., CROTWELL, M. J., DLUGOKENCKY, E., MADRONICH, M., MOGLIA, E., NEFF, D., THONING, K., WOLTER, S., & MUND, J. W. (2022). Atmospheric carbon monoxide dry air mole fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1988-2021. Version: 2022-07-28.

PETRON, G., CROTWELL, A. M., CROTWELL, M. J., DLUGOKENCKY, E., MADRONICH, M., MOGLIA, E., NEFF, D., WOLTER, S., & MUND, J. W. (2020). Atmospheric carbon monoxide dry air mole fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1988-2020, version: 2020-08.

PÉTRON, G., GRANIER, C., KHATTATOV, B., LAMARQUE, J.-F., YUDIN, V., MÜLLER, J.-F., & GILLE, J. (2002). Inverse modeling of carbon monoxide surface emissions using Climate Monitoring and Diagnostics Laboratory network observations. *Journal of Geophysical Research: Atmospheres*, 107(D24).

PÉTRON, G., GRANIER, C., KHATTATOV, B., YUDIN, V., LAMARQUE, J.-F., EMMONS, L., GILLE, J., & EDWARDS, D. P. (2004). Monthly CO surface sources inventory based on the 2000-2001 MOPITT satellite data. *Geophysical Research Letters*, 31(21).

PISON, I., BOUSQUET, P., CHEVALLIER, F., SZOPA, S., & HAUGLUSTAINE, D. (2009). Multi-species inversion of CH_4, CO and H_2 emissions from surface measurements. *Atmospheric Chemistry and Physics*, 9(14), 5281–5297.

PRINN, R. G., WEISS, R. F., ARDUINI, J., ARNOLD, T., DEWITT, H. L., FRASER, P. J., GANESAN, A. L., GASORE, J., HARTH, C. M., HERMANSEN, O., KIM, J., KRUMMEL, P. B., LI, S., LOH, Z. M., LUNDER, C. R., MAIONE, M., MANNING, A. J., MILLER, B. R., MITREVSKI, B., MÜHLE, J., O'DOHERTY, S., PARK, S., REIMANN, S., RIGBY, M., SAITO, T., SALAMEH, P. K., SCHMIDT, R., SIMMONDS, P. G., STEELE, L. P., VOLLMER, M. K., WANG, R. H., YAO, B., YOKOUCHI, Y., YOUNG, D., & ZHOU, L. (2018). History of chemically and radiatively important atmospheric gases from the Advanced Global Atmospheric Gases Experiment (AGAGE). *Earth System Science Data*, 10(2), 985–1018.

RANDERSON, J. T., CHEN, Y., VAN DER WERF, G. R., ROGERS, B. M., & MORTON, D. C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, 117(G4).

RANDERSON, J. T., VAN DER WERF, G. R., GIGLIO, L., COLLATZ, G. J., & KASIBHATLA, P. S. (2017). Global fire emissions database, version 4.1 (gfedv4).

RAUB, J. A. & MCMULLEN, T. B. (1991). *Air Quality Criteria for Carbon Monoxide*. United States Environmental Protection Agency, Washington, DC.

RIAHI, K., VAN VUUREN, D. P., KRIEGLER, E., EDMONDS, J., O'NEILL, B. C., FUJIMORI, S., BAUER, N., CALVIN, K., DELLINK, R., FRICKO, O., LUTZ, W., POPP, A., CUARESMA, J. C., KC, S., LEIMBACH, M., JIANG, L., KRAM, T., RAO, S., EMMERLING, J., EBI, K., HASEGAWA, T., HAVLIK, P., HUMPENÖDER, F., DA SILVA, L. A., SMITH, S., STEHFEST, E., BOSETTI, V., EOM, J., GERNAAT, D., MASUI, T., ROGELJ, J., STREFLER, J., DROUET, L., KREY, V., LUDERER, G., HARMSEN, M., TAKAHASHI, K., BAUMSTARK, L., DOELMAN, J. C., KAINUMA, M., KLIMONT, Z., MARANGONI, G., LOTZE-CAMPEN, H., OBERSTEINER, M., TABEAU, A., & TAVONI, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42, 153–168.

RUSSELL, G. L. & LERNER, J. A. (1981). A New Finite-Differencing Scheme for the Tracer Transport Equation.

RYTER, S. W., MA, K. C., & CHOI, A. M. K. (2018). Carbon monoxide in lung cell physiology and disease. *American Journal of Physiology-Cell Physiology*, 314(2), C211–C227.

SAUNOIS, M., STAVERT, A. R., POULTER, B., BOUSQUET, P., CANADELL, J. G., JACKSON, R. B., RAYMOND, P. A., DLUGOKENCKY, E. J., HOUWELING, S., PATRA, P. K., CIAIS, P., ARORA, V. K., BASTVIKEN, D., BERGAMASCHI, P., BLAKE, D. R.,

BIBLIOGRAPHY

Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., & Zhuang, Q. (2020). The global methane budget 2000–2017. *Earth System Science Data*, 12(3), 1561–1623.

Schneising, O., Buchwitz, M., Hachmeister, J., Vanselow, S., Reuter, M., Buschmann, M., Bovensmann, H., & Burrows, J. P. (2023). Advances in retrieving XCH$_4$ and XCO from Sentinel-5 Precursor: improvements in the scientific TROPOMI/WFMD algorithm. *Atmospheric Measurement Techniques*, 16(3), 669–694.

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., & Burrows, J. P. (2020). Severe Californian wildfires in November 2018 observed from space: the carbon monoxide perspective. *Atmospheric Chemistry and Physics*, 20(6), 3317–3332.

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V. A., Warneke, T., & Wunch, D. (2019). A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor. *Atmospheric Measurement Techniques*, 12(12), 6771–6802.

Shindell, D. T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J.-F., Pétron, G., Dentener, F. J., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz,

L. W., Isaksen, I. S. A., Lawrence, M. G., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., Unger, N., van Noije, T. P. C., & Zeng, G. (2006). Multimodel simulations of carbon monoxide: Comparison with observations and projected near-future changes. *Journal of Geophysical Research*, 111(D19).

Simonin, D., Waller, J. A., Ballard, S. P., Dance, S. L., & Nichols, N. K. (2019). A pragmatic strategy for implementing spatially correlated observation errors in an operational system: An application to Doppler radial winds. *Quarterly Journal of the Royal Meteorological Society*, 145(723), 2772–2790.

Sofiev, M., Ermakova, T., & Vankevich, R. (2012). Evaluation of the smoke-injection height from wild-land fires using remote-sensing data. *Atmospheric Chemistry and Physics*, 12(4), 1995–2006.

Sofiev, M., Vankevich, R., Ermakova, T., & Hakkarainen, J. (2013). Global mapping of maximum emission heights and resulting vertical profiles of wildfire emissions. *Atmospheric Chemistry and Physics*, 13(14), 7039–7052.

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., & McElroy, M. B. (2000). Three-dimensional climatological distribution of tropospheric OH: Update and evaluation. *Journal of Geophysical Research: Atmospheres*, 105(D7), 8931–8980.

Stavrakou, T. & Müller, J.-F. (2006). Grid-based versus big region approach for inverting CO emissions using Measurement of Pollution in the Troposphere (MO-PITT) data. *Journal of Geophysical Research*, 111(D15).

Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., & Clerbaux, C. (2014). On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations. *Atmospheric Chemistry and Physics*, 14(17), 9295–9316.

Strode, S. A., Duncan, B. N., Yegorova, E. A., Kouatchou, J., Ziemke, J. R., & Douglass, A. R. (2015). Implications of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate models. *Atmospheric Chemistry and Physics*, 15(20), 11789–11805.

Talagrand, O. & Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478), 1311–1328.

BIBLIOGRAPHY

VAN DER WERF, G. R., RANDERSON, J. T., GIGLIO, L., VAN LEEUWEN, T. T., CHEN, Y., ROGERS, B. M., MU, M., VAN MARLE, M. J. E., MORTON, D. C., COLLATZ, G. J., YOKELSON, R. J., & KASIBHATLA, P. S. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, 9(2), 697–720.

VEEFKIND, J. P., ABEN, I., MCMULLAN, K., FÖRSTER, H., DE VRIES, J., OTTER, G., CLAAS, J., ESKES, H. J., DE HAAN, J. F., KLEIPOOL, Q., VAN WEELE, M., HASEKAMP, O., HOOGEVEEN, R., LANDGRAF, J., SNEL, R., TOL, P., INGMANN, P., VOORS, R., KRUIZINGA, B., VINK, R., VISSER, H., & LEVELT, P. F. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120, 70–83.

VOULGARAKIS, A., NAIK, V., LAMARQUE, J.-F., SHINDELL, D. T., YOUNG, P. J., PRATHER, M. J., WILD, O., FIELD, R. D., BERGMANN, D., CAMERON-SMITH, P., CIONNI, I., COLLINS, W. J., DALSØREN, S. B., DOHERTY, R. M., EYRING, V., FALUVEGI, G., FOLBERTH, G. A., HOROWITZ, L. W., JOSSE, B., MACKENZIE, I. A., NAGASHIMA, T., PLUMMER, D. A., RIGHI, M., RUMBOLD, S. T., STEVENSON, D. S., STRODE, S. A., SUDO, K., SZOPA, S., & ZENG, G. (2013). Analysis of present day and future OH and methane lifetime in the ACCMIP simulations. *Atmospheric Chemistry and Physics*, 13(5), 2563–2587.

WALLACE, J. M. & HOBBS, P. V. (2006). *Atmospheric science - An introductory survey*, volume 92 of *International Geophysics Series*. Elsevier, second edition.

WANG, Y. X., MCELROY, M. B., JACOB, D. J., & YANTOSCA, R. M. (2004). A nested grid formulation for chemical transport over Asia: Applications to CO. *Journal of Geophysical Research: Atmospheres*, 109(D22).

WHO (1999). *Environmental health criteria 213: carbon monoxide*. Geneva: World Health Organization.

WIEDINMYER, C., AKAGI, S. K., YOKELSON, R. J., EMMONS, L. K., AL-SAADI, J. A., ORLANDO, J. J., & SOJA, A. J. (2011). The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4(3), 625–641.

WIEDINMYER, C. & EMMONS, L. (2022). Fire Inventory from NCAR version 2 fire emission. (Updated irregularly.) Accessed 22 May 2023.

WIEDINMYER, C., KIMURA, Y., MCDONALD-BULLER, E. C., EMMONS, L. K., BUCHHOLZ, R. R., TANG, W., SETO, K., JOSEPH, M. B., BARSANTI, K. C., CARLTON, A. G., & YOKELSON, R. (2023). The Fire Inventory from NCAR version 2.5: an

updated global fire emissions model for climate and chemistry applications. *Geoscientific Model Development*, 16(13), 3873–3891.

WORDEN, H. M., DEETER, M. N., EDWARDS, D. P., GILLE, J. C., DRUMMOND, J. R., & NÉDÉLEC, P. (2010). Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals. *Journal of Geophysical Research*, 115(D18).

YIN, Y., CHEVALLIER, F., CIAIS, P., BROQUET, G., FORTEMS-CHEINEY, A., PISON, I., & SAUNOIS, M. (2015). Decadal trends in global CO emissions as seen by MOPITT. *Atmospheric Chemistry and Physics*, 15(23), 13433–13451.

ZHAI, S., JACOB, D. J., WANG, X., SHEN, L., LI, K., ZHANG, Y., GUI, K., ZHAO, T., & LIAO, H. (2019). Fine particulate matter ($PM_{2.5}$) trends in China, 2013–2018: separating contributions from anthropogenic emissions and meteorology. *Atmospheric Chemistry and Physics*, 19(16), 11031–11041.

ZHANG, T., WOOSTER, M., DE JONG, M., & XU, W. (2018). How well does the 'small fire boost' methodology used within the GFED4.1s fire emissions database represent the timing, location and magnitude of agricultural burning? *Remote Sensing*, 10(6), 823.

ZHANG, X., LIU, J., HAN, H., ZHANG, Y., JIANG, Z., WANG, H., MENG, L., LI, Y. C., & LIU, Y. (2020). Satellite-observed variations and trends in carbon monoxide over Asia and their sensitivities to biomass burning. *Remote Sensing*, 12(5), 830.

ZHAO, Y., SAUNOIS, M., BOUSQUET, P., LIN, X., BERCHET, A., HEGGLIN, M. I., CANADELL, J. G., JACKSON, R. B., HAUGLUSTAINE, D. A., SZOPA, S., STAVERT, A. R., ABRAHAM, N. L., ARCHIBALD, A. T., BEKKI, S., DEUSHI, M., JÖCKEL, P., JOSSE, B., KINNISON, D., KIRNER, O., MARÉCAL, V., O'CONNOR, F. M., PLUMMER, D. A., REVELL, L. E., ROZANOV, E., STENKE, A., STRODE, S., TILMES, S., DLUGOKENCKY, E. J., & ZHENG, B. (2019). Inter-model comparison of global hydroxyl radical (OH) distributions and their impact on atmospheric methane over the 2000–2016 period. *Atmospheric Chemistry and Physics*, 19(21), 13701–13723.

ZHENG, B., CHEVALLIER, F., CIAIS, P., YIN, Y., DEETER, M. N., WORDEN, H. M., WANG, Y., ZHANG, Q., & HE, K. (2018). Rapid decline in carbon monoxide emissions and export from East Asia between years 2005 and 2016. *Environmental Research Letters*, 13(4), 044007.

ZHENG, B., CHEVALLIER, F., YIN, Y., CIAIS, P., FORTEMS-CHEINEY, A., DEETER, M. N., PARKER, R. J., WANG, Y., WORDEN, H. M., & ZHAO, Y. (2019). Global atmospheric carbon monoxide budget 2000–2017 inferred from multi-species atmospheric inversions. *Earth System Science Data*, 11(3), 1411–1436.

BIBLIOGRAPHY

Acknowledgments

First and foremost, I want to thank Nikos Daskalakis, who introduced me into the wondrous world of atmospheric modeling and taught me a lot about life, programming, and what it is like to be scientist. Further thanks goes to my supervisor, Professor Mihalis Vrekoussis, whose continued support through its highs and lows made this thesis possible, and to Professor Maarten Krol, who would never hesitate to share his extensive knowledge about atmospheric chemistry and inverse modeling in general and the TM5-4dvar model in particular.

Next, I want to thank Sarah-Lena Seemann and Santiago Parraguez Cerda for all the sometimes more and sometimes less work-related chats, which leave me hard-pressed to decide which were more important for the success of this thesis. I am also grateful for getting to work with Andreas Hilboll, who changed my view on the world, not only by teaching me the pythonic way of thinking. Of course, the rest of the extended LAMOS group and TM community should not go unmentioned, who are the most random, yet welcoming bunch of people I have met so far. Additionally, I want to express my gratitude for the fruitful cooperation and helpful discussions over the years with Maria Kanakidou, Oliver Schneising, Angelos Gkouvousis, Micheal Buchwitz, and Fabian Piwowarczyk.

A special thanks goes out to Martin Prinzler, Miriam Nüß, and Jan Berges, who proofread this document and gave valuable comments on how to improve its comprehensibleness, language, and layout. On that note, I also want to thank Petra Renken and Sandra Smit, who kept the administrative tasks manageable, and Marius Dan, who, beside keeping the cluster happy, was always there for any IT-related problems beyond that.

There are a few people beyond academia, who played important roles in keeping me going until the very end, in particular Kerstin Lückemeyer, Uwe Klein, my parents Ruth and Hendrik Nüß, and Dominik Koslowsky. Last but surely not least, I also want to thank my wonderful wife Janina Nüß, who stuck with me all this time and kept me pushing and my back free when I needed it.