CO_2 Human Emissions

Attribution Problem Configurations

Julia Marshall &
Tonatiuh Nuñez Ramirez

che-project.eu

D4.3 Attribution Problem Configurations

Dissemination Level:	Public
Author(s):	Julia Marshall and Tonatiuh Nuñez Ramirez (MPG)
Date:	17/10/2019
Version:	4.1
Contractual Delivery Date:	30/06/2019
Work Package/ Task:	WP4/ T4.2
Document Owner:	MPG
Contributors:	MPG, CEA, EMPA, NILU, TNO
Status:	Final

CHE: CO2 Human Emissions Project

Coordination and Support Action (CSA)
H2020-EO-3-2017 Preparation for a European capacity to monitor CO2 anthropogenic emissions

Project Coordinator: Dr Gianpaolo Balsamo (ECMWF)
Project Start Date: 01/10/2017
Project Duration: 39 months

Published by the CHE Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
gianpaolo.balsamo@ecmwf.int

The CHE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776186.

Table of Contents

Figures

Tables

1 Executive summary

A strategy for a European integrated observation system to monitor fossil carbon dioxide (CO_2), also referred to as a Fossil Fuel Data Assimilation System (FFDAS), has been proposed. This FFDAS includes space-borne high resolution and spatially explicit observations of the CO_2 total column abundance complemented by in situ air sampling networks, the development of state-of-the-art bottom-up emission inventories and the setup of several data assimilation systems. Yet the atmospheric signals of fossil fuel emissions can be often be dwarfed by natural signals. In consequence, it has been proposed to use additional trace species to separate the fossil CO_2 component from the natural CO_2 fluxes at regional scale. WP4 aims to describe an optimal configuration for a measurement network with respect to trace species and station location and density to better complement the proposed FFDAS. The tracers on which the work is focusing are radiocarbon CO_2 ($\delta^{14}CO_2$), carbon monoxide (CO), and atmospheric potential oxygen (APO). The target quantities will be the relative uncertainty reductions of fossil CO_2 emissions in a set of predetermined regions and time periods. Besides network design, WP4 already tests two additional objectives of the FFDAS: the use of state-of-the-art bottom-up emission inventories and the setup of several data assimilation systems.

The different modelling groups participating in WP4 will address its objectives with different focuses ranging spatially from urban areas, to regional, to country to continental scales, temporally from hourly, to weekly, to monthly, to yearly scales and thematically from differentiating only between fossil fuel and non-fossil fuel emissions to differentiating emissions processes explicitly. Consistent with the work in WP2 and the fluxes prepared therein, the year 2015 is chosen as the study period. Partners EMPA and MPG will perform European-wide modelling using COSMO and WRF-STILT respectively at ~5 km x 5 km spatial resolution and hourly time steps. Regional modelling will be carried out by NILU for the Oslo region at 2 km x 2 km and by CEA for northern France/Benelux/Western Germany (except Northern Netherlands) at 2 km x 2 km resolution.

This deliverable aims to ensure that all the modelling groups are working with a harmonized set of initial conditions and uncertainty assumptions in order to obtain comparable estimates of uncertainty reduction. This includes a common set of *a priori* flux estimates for each of the tracers involved, boundary conditions, comparable measurement and model error estimates, common *a priori* uncertainty estimates for prearranged regions, a common set of station arrangements and a similar simulation protocol.

2 Introduction

2.1 Rationale

The CO_2 report to the European Commission (Ciais et al., 2015) provided a vision and strategy for a European integrated observation system to monitor fossil carbon dioxide (CO_2) emissions within the Copernicus program. This observation system shall comprise three complementary components within a Fossil Fuel Data Assimilation System (FFDAS):

I. Column-integrated atmospheric CO_2 mole fraction measurements obtained from a dedicated space-borne sensor, with a high-resolution target (less than 3 km x 3 km) to distinguish individual emission regions and a wide swath (~500 km) for good spatial coverage (full coverage of Europe every three days). In situ air sampling networks shall complement the space-borne sensor.

II. The operational provision of bottom-up fossil CO_2 emission maps with high spatial (1 km x 1 km) and temporal resolution (hourly) and near-real-time production capability.

III. An operational data-assimilation system with a very good representation of regional atmospheric transport processes in atmospheric models to integrate atmospheric measurements with bottom-up information into consistent and accurate estimates of fossil CO_2 emissions.

Nevertheless, the atmospheric signals of fossil CO_2 emissions are diluted and often dwarfed by the large diurnal and annual variability in biogenic fluxes, related to photosynthesis and respiration. Furthermore, relatively significant influences of ocean processes and biomass burning emissions may be superimposed on the atmospheric CO_2 signals depending on the sampling location. Additionally, fossil CO_2 emissions are often point sources or highly concentrated over small (urban or industrial) areas. Therefore, Ciais et al., (2015) recommended the development of two complementary lines of approach to quantify fossil CO_2 emissions:

I. Focus on dense sampling of selected emission hotspots, e.g. megacities, major industrial areas, large power plants, e.g. urban networks.

II. Separate the fossil CO_2 component from the natural CO_2 fluxes at regional scale by measuring additional trace species such as, e.g. radiocarbon in CO_2 ($\delta^{14}CO_2$), carbon monoxide (CO) or atmospheric potential oxygen (APO).

These two approaches shall provide support to the spatially explicit space-borne measurements to improve the separation of fossil fuel and natural sources. Additionally on their own, these two approaches shall provide independent data to validate satellite-based emission estimates.

2.2 Objectives

WP4 aims to describe an optimal configuration for a measurement network with respect to trace species, station location and density as part of the proposed FFDAS to better complement spatially-explicit high resolution satellite measurements. The target quantities will be the relative uncertainty reductions of fossil CO_2 emissions in a set of predetermined regions and time periods. For this, detailed end-to-end simulations of the performances of an operational observation system of fossil CO_2 emissions for different air sampling network configurations will be provided. Moreover, making use of these additional trace species requires to define a clean background reference for each of them, and to quantify their non-fossil-fuel sources as well. The tracers on which the work is focusing are radiocarbon CO_2 ($\delta^{14}CO_2$), carbon monoxide (CO), and atmospheric potential oxygen (APO). Furthermore,

besides the *in situ* air sampling network configuration, WP4 addresses several of the design recommendations for the proposed FFDAS:

I. It will use a state-of-the-art bottom-up fossil CO_2 emission maps with high spatial (ranging between 1-7 km) and temporal resolution (hourly).

II. It will use several different high-resolution regional atmospheric transport models from different modelling groups at high spatial resolutions (1-50 km) to be able to represent small-to-mid scale transport patterns such as mountain-valley circulation or mountain wavers, the daily variability of the planetary boundary layer height, or sea breeze patterns or synoptic weather patterns.

III. The different modelling groups participating in WP4 will address its objectives with different focuses ranging spatially from urban areas, to regional, to country to continental scales, temporally from hourly, to weekly, to monthly, to yearly scales and thematically from differentiating only between fossil fuel and non-fossil fuel emissions to differentiating emissions processes explicitly.

2.3 Background

2.3.1 Radiocarbon CO_2 ($\delta^{14}CO_2$)

Radiocarbon (^{14}C), the radioactive isotope of carbon, is constantly being produced in the lower stratosphere and upper troposphere by the interaction of cosmic rays with the atmospheric molecules, which generate high-energy neutrons. These neutrons in turn collide mainly with nitrogen-14 ($^{14}_{7}N$) atoms and produce ^{14}C. ^{14}C quickly combines with atmospheric oxygen and produces ^{14}CO, which is ultimately oxidized to $^{14}CO_2$ by the reaction with the hydroxyl radical (OH). $^{14}CO_2$ diffuses through the atmosphere, dissolves into the ocean or it is taken up by plants via photosynthesis, fixed and distributed through the biosphere. In time, ^{14}C decays reverting back into $^{14}_{7}N$ with a half-life of ~5730 ± 40 years. Thus, fossil fuels, which have typically been buried for millions of years, are essentially void of ^{14}C. As a result, when CO_2 derived from fossil fuel combustion is emitted to the atmosphere, there is a measurable decrease in the relative abundance of ^{14}C with respect to ^{12}C. In consequence, the difficulty of separating fossil fuel emissions from biogenic fluxes in atmospheric inversions could be partially overcome by including measurements of $^{14}CO_2$. Nevertheless, the estimation of fossil fuel emissions is further complicated due to perturbations caused by emissions of ^{14}C from nuclear reactors and nuclear fuel reprocessing plants and ^{14}C releases due to heterotrophic respiration and ocean processes.

Atmospheric $^{14}CO_2$ measurements are usually reported in the $\Delta^{14}C$ notation, the deviation from the absolute radiocarbon reference standard normalized for isotope fractionation as given by Stuiver & Polach, (1977) in ‰ units:

$$\Delta^{14}C = \left(\frac{A_{SN}^{14}}{A_R^{14}} - 1\right) \times 1000‰, \qquad \text{Eq. (1)}$$

where A_R^{14} = 0.95 x 0.238 Bq g^{-1} is the standard ^{14}C activity based on 95% of the specific activity of the original NBS oxalic acid (OxA I) and A_{SN}^{14} is the measured fractionation-normalized ^{14}C activity of the sample. A_{SN}^{14} is calculated as:

$$A_{SN}^{14} = R^{14} \frac{\lambda N_A}{M_{^{14}C}} f_N^{13}, \qquad \text{Eq. (2)}$$

where R^{14} is the ^{14}C-to-C atom ratio of the respective sample, $M_{^{14}C}$ is the molar mass of carbon (14 g mol^{-1}), λ is the decay constant of ^{14}C (3.8332 x 10^{-12} s^{-1}), N_A is the Avogadro number (6.022 x 10^{26} atoms mol^{-1}), and f_N^{13} is the ^{13}C normalization based on the isotopic composition wood (set as 25‰), which itself is defined as:

$$f_N^{13} = 1 - 2\left(\frac{25 + \delta^{13}C}{1000}\right). \quad \text{Eq. (3)}$$

The purpose of this ^{13}C normalization is to correct for biochemical fractionation against the radiocarbon isotopologue $^{14}CO_2$ abundance during photosynthesis, which is assumed to be twice as strong as for $^{13}CO_2$ based on their respective departures in molecular mass from $^{12}CO_2$.

Thus, the ^{14}C-to-C atom ratio is:

$$R^{14} = \frac{M_{^{14}C} A_R^{14}}{\lambda N_A f_N^{13}} (\Delta^{14}C + 1). \quad \text{Eq. (4)}$$

The equation above can be simplified by defining the factor f:

$$f = \frac{M_{^{14}C} A_R^{14}}{\lambda N_A}. \quad \text{Eq. (5)}$$

For the OxA I standard, the factor f is 1.176×10^{-12} atoms^{-1}, which is then the ratio of the standard substance R_{STD}^{14}.

Thus, the ^{14}C-to-C atom ratio becomes:

$$R^{14} = \frac{R_{STD}^{14}}{f_N^{13}} (\Delta^{14}C + 1). \quad \text{Eq. (6)}$$

In this study, all stocks and fluxes of ^{14}C should be reported based on the ^{14}C-to-C atom ratio derived from Eq. (6).

However, when Δ^{14}C is used to calculate fossil fuel CO_2 content, the ^{13}C Suess Effect (Suess, 1955; Tans et al., 1979) is neglected introducing a significant biases for heavily polluted areas (Wang, 2016). Following Vogel et al. (2013, supplement), we use δ^{14}C notation to build our inversion system, which is defined as:

$$\delta^{14}C = \left(\frac{R^{14}}{R_{STD}^{14}} - 1\right) \times 1000‰. \quad \text{Eq. (7)}$$

However, in the following sections, we sometimes inevitably mention the Δ^{14}C notation and made some assumptions and approximations on the value of δ^{13}C to convert Δ^{14}C values into δ^{14}C values (Table 2-1), when the latter are not available. These assumptions and approximations will only slightly influence the estimate of $^{14}CO_2$ fluxes, as the variations in atmospheric δ^{13}C are relatively small (Vogel et al., 2013, supplement), but they will not impact our uncertainty reduction calculations.

Table 2-1 Recommended $\delta^{13}CO_2$ values for source types

Material	Typical δ^{13}C value [‰]
C3 plants	-27.5
C4 plants	-13.0

Atmospheric CO_2	-8.0
Ocean CO_2	-22.5

Moreover, despite its advantages, atmospheric $\delta^{14}CO_2$ measurements are so far expensive and difficult to make with sufficient precision because of its very low abundance (the accuracy at which the fossil fuel component of atmospheric CO_2 can be determined from ^{14}C measurements of CO_2 in air samples is about 1 ppm with Accelerator Mass Spectrometry). To be useful for inversions, the network, the number of ^{14}C observations need to be increased with a high data standard in terms of accuracy and availability (Levin & Karstens, 2007). The first simulations of a ^{14}C observation system based on a network of air sampling stations across Europe indicated that with weekly sampling, national fossil CO_2 emission budgets over a year could be independently verified with a typical accuracy of 10% for middle-sized countries (Wang, 2016). With daily sampling, a larger uncertainty reduction could be obtained. Ciais et al., (2015) recommended deploying $\delta^{14}CO_2$ measurements at approximately 50 atmospheric sampling sites across the European continent with higher density over regions with high emissions.

2.3.2 Carbon monoxide (CO)

CO is another potentially interesting tracer in the context of separating fossil fuel emissions, as any hydrocarbon oxidation process with CO_2 as an end product is to some extent associated with production of CO (Gamnitzer et al., 2006). Compared to $^{14}CO_2$ it is inexpensive and easy to measure with enough precision, and many more measurements are already available. Continuous (hourly frequency) CO_2 and CO measurements supplemented by high precision δ^{14}CO2 measurements (weekly to bi-weekly frequency) can be used to estimate regional hourly fossil fraction of the CO_2 mole fraction with uncertainties between 20% and 88% (Gamnitzer et al., 2006; Pickers, 2016). However, this method can only work with relatively correct information on CO/CO_2 ratio of the sources, which may vary by orders of magnitude even within the same source process (Olivier et al., 2005). Furthermore, this method also requires negligible influence of other sources. Importantly biofuel burning, which cannot be distinguished by the CO constraint, but which is part of anthropogenic activity and, therefore, it may be spatially and temporally correlated with fossil fuel combustion, adds significant uncertainty to the application of CO as a quantitative fossil CO_2 proxy (Gamnitzer et al., 2006). Other non-fossil-fuel CO emissions, including soil and ocean fluxes, the production of CO from the oxidation of methane and non-methane volatile organic compounds (NMVOC) and the varying lifetime of CO in the atmosphere, which ranges between a few weeks up to a year depending on season, further add uncertainty to the method.

2.3.3 Atmospheric potential oxygen (APO)

O_2 is not a trace gas, which means that its mole fraction is affected by small changes in other gases, such as CO_2 (Keeling et al., 1998a). Therefore, measurements of atmospheric O_2 are reported as changes in the O_2/N_2 molar ratio. Since these differences in the O_2/N_2 ratio are in the order of 10^{-6} to 10^{-4}, they are reported as $\delta(O_2/N_2)$ in permeg units (ratio expressed as a fraction of 10^6), where

$$\delta(O_2/N_2) = \frac{(O_2/N_2)_{sample} - (O_2/N_2)_{reference}}{(O_2/N_2)_{reference}} \cdot 10^6. \quad \text{Eq. (8)}$$

APO, introduced by Stephens et al., (1998), is a measure of the O_2 concentration that cancels out the effect of the land biosphere on the O_2/N_2 ratio. APO is calculated as

$$APO = \delta(O_2/N_2) + \frac{\alpha_L}{y_0^{O_2}}\left(y^{CO_2} - y_0^{CO_2}\right) \cdot 10^6, \qquad \text{Eq. (9)}$$

where y^{CO_2} is the CO_2 mole fraction in ppm units, α_L is the average O_2:CO_2 stoichiometric ratio for land photosynthesis or respiration (global average of -1.1), $y_0^{O_2}$ is a fixed reference O_2 mole fraction set to 209460 ppm, $y_0^{CO_2}$ is an arbitrary reference value of 350 ppm which has been introduced for convenience. Thus, high precision, continuous measurements of atmospheric O_2 and CO_2 are a valuable tool for gaining insight into carbon cycle processes, and for separating land biosphere and ocean fluxes.

Since CO_2 and O_2 fluxes from fossil fuel combustion are strongly anti- correlated, APO has been put forward as a potential tracer to detect fossil fuel signatures of CO_2, based on different oxidation ratios associated with different flux processes (Pickers et al., 2016). APO benefits from a smaller range of possible O_2:CO_2 emission molar ratios for fossil fuels (from -1.2 to -1.95, more typically in the range of -1.3 to -1.4) compared to CO:CO_2 (from <0.005 to >0.1, but typically in the range of 0.005 to 0.025), which results in a lower uncertainty estimation of the fossil CO_2 (Pickers, 2016). Hence, the APO method can likely be used independently of $\delta^{14}CO_2$ measurements (unlike the CO method), which are costly and are influenced by emissions from nuclear power plants, which are common in certain European regions. As in the case of CO, the estimation of fossil fuel CO_2 emissions with the APO constraint can be perturbed by biofuel combustion emissions. However, the influence is less than in the case of CO because the APO method is conservative with respect to solid biofuel and biomass burning ($\alpha \approx -1.1$). In addition, unlike CO, the APO constraint in not perturbed by biomass burning.

2.1. Scope of this deliverable

This deliverable provides the initial conditions and a common simulation framework for the different modelling groups in WP4 that will work on the attribution problem based on the use of *in situ* measurements.

2.2.1. Objectives of this deliverable

This deliverable aims to ensure that all the modelling groups are working with a harmonized set of initial assumptions. This includes assumptions related to fluxes (emissions of fossil and non-fossil sources for each tracer), concentrations (initial and boundary/background conditions), and uncertainties (prior flux uncertainties and correlations, measurement uncertainties, model-data mismatch, uncertainties in emission ratios, etc.). The assumed *a priori* uncertainties will not be identical, as different groups are solving for different quantities, but the statistical assumptions should be generally consistent.

2.2.2. Analysis of uncertainty

The comparison of the *a posteriori* covariance matrix with the *a priori* one corresponds to what is usually called the analysis of uncertainties: for each individual parameter one may

see how the initial uncertainty has been reduced (Tarantola, 2005). The *a posteriori* covariance matrix, which contains the uncertainties and correlations of the estimated parameter values, is determined by two factors (1) the *a priori* emission uncertainty relative to the sum in quadrature of the measurement and model uncertainty, and (2) the sensitivity of the simulated mole fraction at a station toward the fluxes (Houweling et al., 1999) given by the location and time of the measurements.

2.2.3. Work performed in this deliverable

This deliverable provides a documentation of the planned modelling approaches that will be taken by the different groups. This includes a description of the different spatial and temporal scales on which the fluxes will be optimized as well as a description of how the additional tracers will be implemented. This deliverable also compiles and documents state of the art emission estimates for not only the fossil fuel components relevant for each tracer but also their non-fossil-fuel sources, in order to simulate the background variability as realistically as possible.

2.2.4. Deviations and countermeasures

A delay has arisen in the generation of the APO fluxes. A member of staff of UEA, who focused on this tracer, but who has since left the CHE project, originally coordinated this. None of the other modeling groups have used this tracer previously in inverse modeling studies. The member of staff worked with TNO already last year to develop a product for CHE consistent with the CO_2 emissions being used. This involved supplying oxidative ratios specific to fuel type (not only process), which means it had to be done a level higher in the processing chain than the process-separated fluxes that are already supplied by TNO, and as such required more work.

Unfortunately, in the preparation of the proposal, APO fluxes were not included as a formal deliverable for TNO, and as such have not taken priority in the face of personnel limitations. APO fluxes were only available on August 6, 2019 and have since been over a number of iterations to adjust them better to the modeler's needs as well as correcting clear errors.

The implications of this delay are simply that the additional work in Task 4.2, namely the uncertainty reduction analysis of different sampling networks by the different modeling groups, will not begin in earnest until October. This is a delay of four months, which should still fit within the planned margins of the task. However, it reduces the safety buffer should additional complications arise. Nonetheless this still allows a full twelve months of work before the task is expected to be completed (month 36).

3 Inversion configurations

Consistent with the work in WP2 and the fluxes prepared therein, the year 2015 is chosen as the study period. Partners EMPA and MPG will perform European-wide modelling using COSMO and WRF-STILT respectively at 5 km x 5 km spatial resolution and hourly time steps. Regional modelling will be carried out by NILU for the Oslo region at 2 km x 2 km and by CEA for northern France/Benelux/Western Germany (except Northern Netherlands at 2 km x 2 km resolution. A common list of relevant tracers was agreed upon: CO_2, $^{14}CO_2$, CO

and APO. The modelling framework of each institution is described in turn in the following section.

3.1 MPG

We will provide detailed end-to-end analyses of the performances of hypothetical observation systems of fossil CO_2 emissions with the Jena Carboscope inversion system (Rödenbeck, 2005) [http://www.bgc-jena.mpg.de/CarboScope/].

3.1.1 Domain

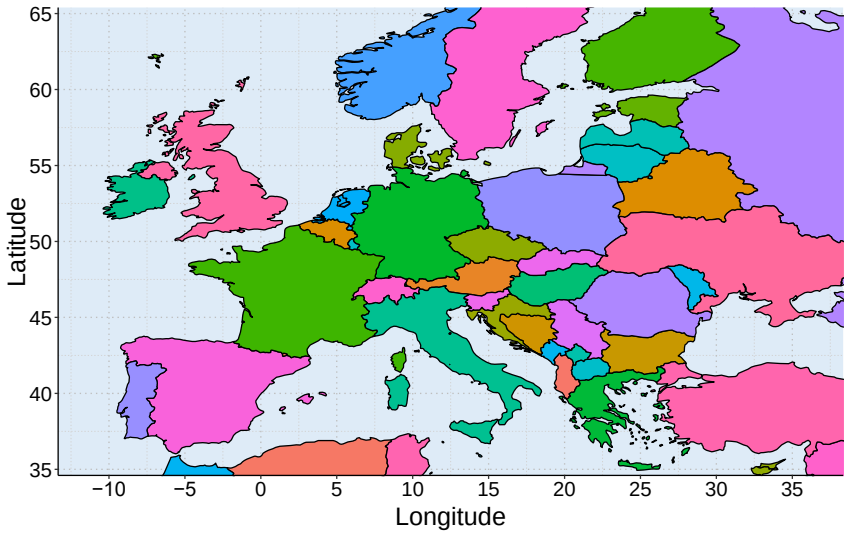

Figure 3.1 Simulation domain for the WRF/STILT simulation in WP4.

The regional STILT simulations will be conducted on a 1/24° × 1/16° (latitude, longitude) horizontal grid covering the minimal domain of the region of interest in the CHE project (11˚W – 36˚E, 36˚N – 64˚N). In WP2, we provided WRF-GHG simulations at 5-km resolution for the whole European domain. The wind fields generated by WRF-GHG will be reprojected and used in STILT to calculate the footprints (sensitivities to upstream surface-atmosphere fluxes) starting at a receptor location (Nehrkorn et al., 2010). The framework is thus analogous to a regional adjoint model in a Eulerian framework, providing an alternative to generating and implementing adjoint model code for a Eulerian transport model (Pillai et al., 2012).

3.1.2 Forward simulations

Forward simulations containing both CO_2, CO, APO and $^{14}CO_2$ tracers shall be carried out in order to estimate the strength of the signals at different locations using tagged tracers. The footprints calculated by STILT are then mapped to the high-resolution *a priori* fluxes and initial/lateral boundary conditions. This part of the framework – offline Lagrangian modelling – provides time series of CO_2, $^{14}CO_2$, O_2 and CO mole fractions mole fractions at the receptor location. $\delta^{14}CO_2$ and APO can be calculated in a post-processing step. In the resulting signals we will be able to distinguish the contribution from different processes to the

atmospheric signals sampled at different locations. An example of what can be produced is found at https://stilt.icos-cp.eu/viewer/.

3.1.3 Inversion setup

Typically, within the Jena Carboscope, regional-scale high-resolution inversions are performed with the coupled TM3-STILT system (Trusilova et al., 2010). TM3-STILT is based on a combination of the fine-scale regional Stochastic Time-Inverted Lagrangian Transport model STILT (Gerbig et al., 2003; Lin, 2003) and the coarse-grid global three-dimensional atmospheric offline transport model TM3 (Heimann & Körner, 2003), which are coupled according to the two-step nesting scheme of Rödenbeck et al. (2009). This nesting scheme allows the use of completely independent models for the representation of global and regional transport and hence facilitates an easy exchange of either component. The global inversion is used to calculate fluxes from the far field (outside of the regional domain of interest), and, subsequently, this information can be used to provide the boundary conditions for the regional inversion. Since we will not use real mole fraction data in this study, which means that the fluxes outside the domain of interest do not need to be compatible with the fluxes inside the domain, we will not use the initial and boundary conditions provided by the TM3 model but from other sources described later in this document.

The 4D-VAR variational inversion algorithm of the Jena Carboscope is described in detail in Rödenbeck (2005). Here we provide a brief summary. The primary input for flux estimations is a vector the observed tracer mole fractions $\boldsymbol{y_{obs}}$ that contains the total set of measurements at all times and locations. The modelled mole fractions $\boldsymbol{y_{mod}}$ results from the transport a discretized flux field **f**(x,y,t), which varies in time and space. This is formally expressed as:

$$\boldsymbol{y}_{mod} = \boldsymbol{A}\mathbf{f} + \boldsymbol{y}_{ini}, \qquad \text{Eq. (10)}$$

where $\boldsymbol{y}_{ini}$ is the initial conditions and $\boldsymbol{A}$ is the transport matrix. For regional inversions the transport matrix $\boldsymbol{A}$ has been pre-computed by the STILT model for all times and locations where a measurement exists. The inversion seeks to estimate the fluxes **f** that lead to minimum data-model mismatch $(\boldsymbol{y_{obs}} - \boldsymbol{y_{mod}})$ through the minimization of a cost function:

$$J = \frac{1}{2}(\boldsymbol{y_{obs}} - \boldsymbol{y_{mod}})^{\boldsymbol{T}}\boldsymbol{Q_m}(\boldsymbol{y_{obs}} - \boldsymbol{y_{mod}}), \qquad \text{Eq. (11)}$$

where the diagonal matrix $\boldsymbol{Q_m}$ weights the mole fraction values given their assumed measurement error, location-dependent modelling error and a data-density weighting (Rödenbeck, 2005). In regional scale inversions, the Jena Carboscope typically estimates three-hourly fluxes at gridcell scale resolution. Therefore, the problem is ill-posed, because the number of unknowns is larger than the number of measurements. The problem can be regularized by adding *a priori* information.

To simplify the structuring of the *a priori* information such that it reflects process understanding, all the a priori information for each element of the flux vector **f** is supplied in the form of a statistical linear flux model (Rödenbeck, 2005):

$$\mathbf{f} = \mathbf{f}_{fix} + \mathbf{Fp}. \qquad \text{Eq. (12)}$$

The statistical linear model represents the *a priori* probability distribution of flux vector **f**. The flux vector **f** represents net flux per gridcell per time step. It is composed of a fixed term $\mathbf{f}_{fix}$, which is the *a priori* expectation value $\langle \mathbf{f}_{pri} \rangle$, and an adjustable term, $\mathbf{Fp}$, which determines

the deviations around $\langle \mathbf{f}_{pri} \rangle$ (Gaussian distributed). The adjustable term is composed of matrix **F** and vector **p**. Vector **p** is a set of adjustable parameters assumed to be independent, with an a priori expectation value $\langle \mathbf{p}_{pri} \rangle = 0$ and a priori unit variance $\langle \mathbf{p}_{pri}\, \mathbf{p}_{pri}^{\mathrm{T}} \rangle = \frac{\mu}{2}\mathbf{I}$. Each of the elements of vector **p** acts as a multiplier to each one of the columns of the matrix **F**.

The matrix **F** comprises all the *a priori* information about flux uncertainties and correlations. Each column of matrix **F** represents an elementary spatiotemporal flux pattern or base function that is a building block of the total flux uncertainty (Rödenbeck et al., 2003). The extension in space and time of these elementary flux patterns determines the coherent behaviour or correlations of the flux elements. The matrix **F** is defined as

$$\mathbf{F} = \mathrm{f}_{sh}(x, y, t) \cdot \mathrm{g}_m^{time}(t) \cdot \mathrm{g}_m^{space}(x, y), \qquad \text{Eq. (13)}$$

where g_m^{time} and g_m^{space} are functions (range from zero to one) determining the temporal and spatial decomposition into statistically independent elements and $\mathrm{f}_{sh}(x, y, t)$ is a discrete spatiotemporal shape function, which determines the local/instantaneous *a priori* standard deviation of the flux **f** (Rödenbeck, 2005). Since the inversion algorithm will preferably project signals in the data into space and time locations with large *a priori* uncertainty, the shape function provides a spatiotemporal weighting of the flux adjustment and restricts the flux adjustments to prescribed source regions, e.g. land or ocean, or source periods, e.g. the growing season. The *a priori* covariance matrix is

$$\mathbf{Q}_F = \mathbf{F}\mathbf{F}^{\mathrm{T}}. \qquad \text{Eq. (14)}$$

Furthermore, the Jena Carboscope allows for the representation of the flux vector **f** as the sum of N_{comp} flux components, each of which is represented by its own independent statistical linear flux model:

$$\mathbf{f} = \sum_{i=1}^{N_{comp}} \mathbf{f}_{fix,i} + \mathbf{F}_i \mathbf{p}_i. \qquad \text{Eq. (15)}$$

Each flux component *i* may correspond to a physical source process, e.g. fossil fuels, gross primary productivity or ocean processes. Additionally, flux components may correspond to a particular process split into different temporal scales, e.g. mean seasonal cycle, interannual or short-term variability, or spatial scales. Since each component is independent, each component is assigned different *a priori* error covariance structure. Ideally, this approach would allow for the partitioning of the deviations from the *a priori* flux estimate for each flux component considered. This is the case for flux components that are geographically separate, e.g. land vs. ocean fluxes, or temporally separate, e.g. opposite seasonality of ecosystem respiration emissions and residential heating. However, due to the nature of CO_2 source and sink processes, there is considerable overlap among the processes contributing to the total flux in each gridcell even with good *a priori* knowledge of their spatiotemporal distribution. When flux components overlap within a data-model mismatch gradient, the deviations from the *a priori* flux estimate calculated by the inversion are attributed to the flux components proportionally to their relative contribution to the overall uncertainty $\mathbf{F}_i/\mathbf{F}$ and to the ratio between the *a priori* flux uncertainty and the data uncertainty (Rödenbeck, 2005).

Since the spatiotemporal distribution of the uncertainty of each component is not well known and the assumptions for the shapes $\mathbf{f}_{sh,i}$, although made based on process understanding, are rather arbitrary, the partitioning between different sources and sinks may not be realistic. In this study, we make use of the information provided by $\delta^{14}CO_2$, CO mole fraction and APO measurements to provide an additional constraint to the partitioning of the sources.

This means that the Jena Carboscope simultaneously optimizes CO_2, $^{14}CO_2$, CO and APO fluxes such that they are consistent with the observations of the four trace species. In order to represent $^{14}CO_2$, CO and APO fluxes, which are adjusted proportionally to the CO_2 fluxes, the Jena Carboscope allows for the definition of the so-called derived component. The derived component *j* requires a driving component *i*. In this study, the CO_2 fluxes are always the driving component and the $^{14}CO_2$, CO and APO fluxes are the derived components. The definition of the derived component varies from that of the general statistical flux model in such a way that, for driving component *i* and derived component *j*, the adjustable term of the derived component $\mathbf{F_j p_j}$ is replaced by the adjustable term of the driving component, $\mathbf{F_i p_i}$, scaled by a scalar proportionality factor $drvscale_{i,j}$ and by the shape function of the derived component $\mathbf{f}_{sh,j}$:

$$\mathbf{f}_j = \mathbf{f}_{fix,j} + drvscale_{i,j} \cdot \mathbf{f}_{sh,j} \cdot \mathbf{F_i p_i}. \qquad \text{Eq. (16)}$$

n this study, the factor $drvscale$ was set to one, and the shape function of the derived component $\mathbf{f}_{sh,j}$ represents the spatially and temporally explicit $^{14}CO_2/CO_2$, CO/CO_2 and O_2/CO_2 ratios (molar). This allows one to represent the spatial and temporal variability of the ratios. The relationship between the driving component and the derived component is fixed for every gridcell and time step. Deviations from these fixed ratios can only be represented by additional flux components of the type represented by Eq. (15), which are then optimized independently. In the case of error components $\mathbf{f}_{pri} = 0$ and **F** represents the instantaneous uncertainty of the $^{14}CO_2/CO_2$, CO/CO_2 and O_2/CO_2 ratios.

With this information, the inversion seeks to minimize the cost function that combines the observational and the *a priori* constraint:

$$J = \frac{1}{2}(\mathbf{y_{obs}} - \mathbf{y_{mod}})^{\mathbf{T}}\mathbf{Q_m}(\mathbf{y_{obs}} - \mathbf{y_{mod}}) + \frac{\mu}{2}\mathbf{p}^{\mathbf{T}}\mathbf{p} + C, \qquad \text{Eq. (17)}$$

where C is a constant that summarizes all parameter independent terms and μ is a tunable parameter that scales the impact of the *a priori* constraint on the Bayesian inversion with respect to atmospheric data constraint. The minimization of the cost function is done iteratively with respect to the parameters **p** using a Conjugate Gradient algorithm with re-orthogonalization (Press 2007).

3.1.4 Control vector

Eight emission processes will be considered (Table 3-1): fossil fuel combustion, biofuel combustion, non-combustion fossil emissions (e.g. cement production from reaction $\mathrm{CaCO_3} + heat \rightarrow \mathrm{CaO} + \mathrm{CO_2}$), gross primary productivity, ecosystem respiration, ocean processes, open biomass burning, nuclear power plants (only for $^{14}CO_2$). Anthropogenic fossil fuel and biofuel combustion emissions will be further divided into the components contained within the TNO-GHGco inventory. The three road transportation categories (F1: diesel, F2: gasoline and F3: LPG) will be aggregated into one. For biofuel combustion emissions, categories with no emissions are not considered. Additionally, road and off-road transport biofuel fuel emissions shall be considered as a derived component of the fossil fuel emission counterparts. Land biosphere and ocean fluxes will be separated into two temporal frequency components each with a component optimizing the yearly average and mean seasonal cycle, and a component optimizing short-term variability. In the case of CO, vegetation emissions and soil fluxes shall be coupled to the gross primary productivity and ecosystem respiration fluxes respectively, but with considerable degrees of freedom for their error components, which are independent. Unfortunately, the WRF-STILT inverse modelling framework is currently only able to represent surface fluxes. Therefore, cosmogenic and

photochemical sources will only be represented as included in the boundary conditions, and the *a priori* emissions will not be distributed vertically using the profiles provided by EMPA (Brunner et al., 2019). Additionally, the regional simulations require boundary conditions from global models. However, these boundary conditions cannot be optimized within the Jena Carboscope.

3.1.5 *A priori* error covariance matrix

Table 3-1 Processes and components considered in the control vector of the Jena Carboscope for WP4 as well as the source of their *a priori* estimate. Spatial and temporal correlation lengths represent exponential decay rate constants except for biogenic fluxes, where the spatial correlation is hyperbolic. LT stands for long-term average.

Process	Components	Spatial correlation length [km]	Temporal correlation length [d]
Fossil fuel combustion	Public power	92	35
	Industry	7	35
	Stationary combustion	7	35
	Fugitives	12	31
	Solvents	15	20
	Road transp.	18	26
	Shipping	34	12
	Aviation	47	13
	Off-road transp.	51	11
	Agriculture	83	26
	Waste	72	26
Biofuel combustion	Public power	92	35
	Industry	7	35
	Stationary combustion	7	35
	Road transp.	18	26
	Off-road transp.	51	11
	Agriculture	83	26
	Waste	72	26
Non-combustion fossil	Cement	7	35
	Other	7	35
Gross primary productivity	LT + Seasonal	35	21
	Short term	35	21
Respiration	LT + Seasonal	35	21
	Short term	35	21

Ocean	LT + Seasonal	Zonal: 200 Meridional: 150	32.2
	Short term	Zonal: 200 Meridional: 150	32.2
Biomass burning*		35	21
Nuclear power plants		92	35

*: spatial and temporal correlations will be convolve with specific event delimiter from the Global Fire Atlas (Andela et al., 2019)

The magnitude of the diagonal elements of the *a priori* error covariance matrix, which are equal to the shape function (f_{sh}) in Eq. (13), will be set to be relatively consistent with the assumptions that other groups are making (see section 5.5). Furthermore, spatial and temporal correlation functions were also defined in Eq. (13), which represent the off-diagonal terms of the *a priori* error covariance matrix. Defining temporal and spatial correlations between the flux elements reduces the number of degrees of freedom and the underdetermination of the problem (Rödenbeck et al., 2003). Additionally, because flux elements with long spatial or temporal correlations are constrained by a greater amount of data than the small scale, they filter out noise coming from measurement and model errors, and stabilize the calculation (Rödenbeck, 2005). Moreover, temporal and spatial correlations may be used to estimate flux processes that have a distinctly different time-varying behaviours or spatial distribution, e.g. seasonally varying in contrast to constant year-round emissions.

In the Jena Carboscope, temporal correlations are represented by the spectral weights of a low-pass filter (e.g. Figure 3.2) with component-specific cut-off frequencies (see Table 3-1). For anthropogenic emissions, the cut-off frequencies for each component were determined with the temporal correlation analyses in section A.4.3. For biogenic fluxes we assume a Gaussian temporal correlation structure with decay rate constants of 21.214 days (e-folding time of 30 days) based on Kountouris et al., (2015). For ocean fluxes we assume a Gaussian temporal correlation structure with a decay rate constants of 32.2 d based on Jones et al., 2012). For open biomass burning emissions, the temporal correlation will be the result of the convolution of the same Gaussian filter than for biogenic fluxes with a step function for each fire event contained in the Global Fire Atlas (Andela et al., 2019).

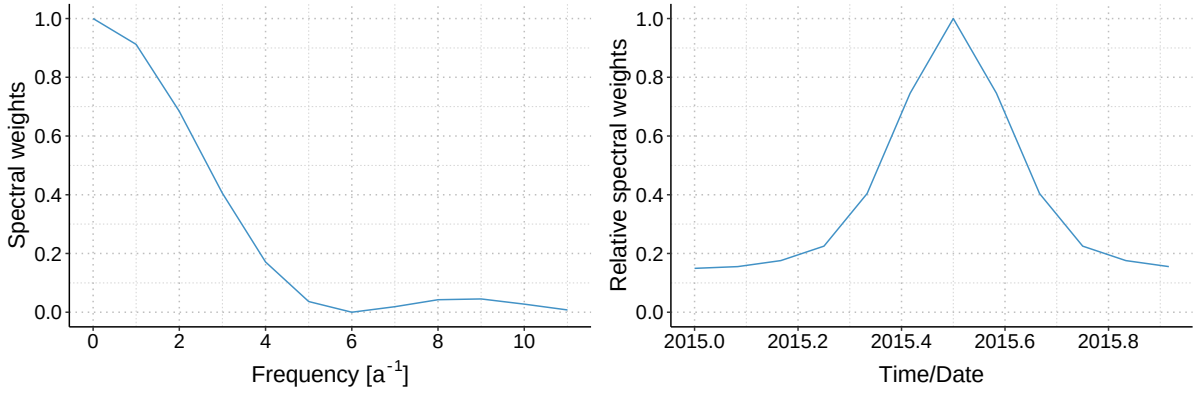

Figure 3.2 *Left*: spectral weights of a triangular low-pass filter with a cut-off frequency of 6 a^{-1} in the frequency domain. *Right*: relative spectral weights of the convolution of this filter with a unit pulse in the middle of the year.

Spatial correlations are the extent to which each of the elementary spatiotemporal flux patterns overlap each other in space and represent a spatial smoothing of previously uncorrelated a priori flux elements (Rödenbeck, 2005). For the anthropogenic emissions, the spatial correlation length was determined for each category based on the analyses in section A.4. In the case of biogenic fluxes, a hyperbolic spatial correlation r_{xy} similar to (Chevallier et al., 2012) was assumed:

$$r_{xy} = \frac{1}{1+{}^{d}/_{35\ \mathrm{km}}}, \qquad \text{Eq. (18)}$$

where *d* is the great circle distance between gridcell centres. The characteristic decay rate of 35 km was obtained from Kountouris et al. (2015). The hyperbolic spatial correlation from Chevallier et al. (2012) has a stronger impact from larger distances compared to the exponential shape leading to an aggregated uncertainty. For ocean fluxes, a Gaussian spatial correlation structure with a characteristic decay rate of 200 and 150 km in the zonal and meridional directions respectively based on Jones et al., (2012) was assumed. For open biomass burning emissions, the spatial correlation will be the result of the convolution of the same hyperpolic spatial correlation function with a two-dimensional step function for each fire event contained in the Global Fire Atlas (Andela et al., 2019).

3.1.6 Observation error covariance matrix

Going back to the cost function Eq. (17), the diagonal matrix $\mathbf{Q_m}$ weights the mole fraction values given their assumed measurement error, location-dependent modelling error and a data-density weighting (Rödenbeck, 2005). Thus, the diagonal elements of the matrix represent the sum in quadrature of measurement and modelling error for each observation:

$$\sigma_{tot} = \sqrt{\sigma_{obs}^2 + \sigma_{mod}^2}. \qquad \text{Eq. (19)}$$

The total error σ_{tot} acts as a weighting among the data values. Assumptions of the measurement and modelling uncertainties for each tracer are discussed in sections 5.1 and 5.2 respectively.

The off-diagonal elements of the error covariance matrix represent the correlations between the measurement errors (e.g. due to common influences of standard gases, and equipment biases) and between the modelling errors (e.g. due to conceptual deficiencies in the transport parameterizations, or errors in the driving meteorology that affect all measurements) (Rödenbeck, 2005). Nevertheless, due to the difficulties in characterizing error correlations, each observation is assumed independent so the off-diagonal elements are set to zero (Rödenbeck, 2005). This also means that observation and model error correlations between species are cannot be assigned explicitly. Nonetheless, since some measurements quantities are ratios, e.g. $\delta^{14}CO_2$ and APO, uncertainty propagation taking in consideration their correlations can be when transforming from $\delta^{14}CO_2$ and APO to $^{14}CO_2$ and O_2 mole fractions, for example:

$$\sigma_{^{14}C} = y_{^{14}C}\sqrt{\left(\frac{\sigma_{CO_2}}{y_{CO_2}}\right)^2 + \left(\frac{\sigma_{R_{14C/12C}}}{R_{14C/12C}}\right)^2 - \frac{\mathrm{COV}(y_{CO_2}, R_{14C/12C})}{y_{CO_2} \cdot R_{14C/12C}}}, \qquad \text{Eq. (20)}$$

where *y* and σ are the respective mole fraction and uncertainty values and $R_{14C/12C}$ is the ratio $^{14}CO_2/^{12}CO_2$ ratio.

Although the measurement error covariance is a diagonal matrix, we do consider for temporal correlations via a data density weighting. Looking back at the cost function, each element of the data-model mismatch vector is weighted by the inverse of its corresponding uncertainty such that the contribution of a particular measurement site to the cost function during a given period (here three hours) is

$$J^* = \frac{1}{2}\sum_{i=1}^{N^*} \frac{\left(y_{obs,i} - y_{mod,i}\right)^2}{\sigma_{tot,i}^2}, \qquad \text{Eq. (21)}$$

where N^* is the number of measurements within this period (Rödenbeck, 2005). This means the cost function and the cost function gradient increase proportionally to the number of measurements N^* within a particular time period and location (Rödenbeck, 2005). This would create unequal weighting between sites with continuous in-situ measurements (e.g. hourly) would have with respect to sites where air samples are taken every two weeks (CO_2 in comparison to $\delta^{14}CO_2$ observations). However, the Jena Inversion System allows the possibility to merge both discrete flask and continuous measurements by scaling the uncertainty of each value in the mole fraction time series:

$$\sigma_{tot,i} = \sigma_{tot,i} \cdot N^*. \qquad \text{Eq. (22)}$$

This way the implied data-model mismatch uncertainty averages over the given period and does not depend on the number of measurements any more, as long as there is at least one data value within the period (Rödenbeck, 2005). Although, through the data density weighting data streams with different data density have the same effect on the cost function, continuous data preserves its information on different areas of influence and correspondingly different signals among the modelled mole fractions that may help better determine the spatial structure of the fluxes (Rödenbeck, 2005). Furthermore, data-density weighting can mediate the impact of data gaps.

3.2 CEA/LSCE

The high dimensional inversion setup described in this section is built for the co-assimilation of CO_2, $^{14}CO_2$, and APO data. With this inversion setup, the potential of the combination of CO_2 space-borne imagers and ground-based networks to monitor emissions will be explored by assimilating both types of data. A proper CO and APO data assimilation requires controlling different emission sectors for each type of source (point, city and regional sources) separately in the inversion. However, we use such a high spatial control resolution for the co-assimilation of CO_2 and $^{14}CO_2$ that adding some sectorial resolution would make the control vector too large for the computations. Furthermore, the current expectations regarding the statistical constraint on the CO_2 emission estimates from the CO data assimilation are rather low. Therefore, CO is not co-assimilated with CO_2 and $^{14}CO_2$ in the high dimensional inversion set-up described in the following. Some tests will be carried out about the co-assimilation of CO and CO_2 data in a much lower dimension inversion system focused on Belgium (which is not described in the following). Although the expectations regarding the statistical constraint on the CO_2 emission estimates from the APO data assimilation are higher will still plan to use the same lower dimension inversion system to test it.

3.2.1 Domain

The domain of the inversion configuration covers the Western part of Europe (longitude: -6.82° to 19.18°; latitude: 42.0° to 56.39°) with a horizontal resolution that varies between 50

and 2 km. The 2 km × 2 km-resolution zoom covers Northern France, a large part of Benelux and Western Germany (longitude: -1.25° to 10.64°; latitude: 47.45° to 53.15°).

3.2.2 Inversion setup

The inversion system is based on an analytical inversion framework (Wu et al., 2016) with a dedicated configuration of the CHIMERE transport model (Menut et al., 2013). The analytical Bayesian inversion allows for the computation of the posterior uncertainty in the inverted flux budgets (its covariance matrix **A**) as a function of the observation operator **H** (connecting the flux budgets to the observation vector, and mainly built on the transport model), the covariance matrices of the prior uncertainties **B** and the model and observation errors **R** following (Tarantola, 2005): $\mathbf{A} = [\mathbf{B}^{-1}+\mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}]^{-1}$. In practice, to co-assimilate CO_2 and $^{14}CO_2$ data an iterative approach repeating analytical inversions will be used to deal with the non-linearity of the observation operator (Wang et al. 2016).

3.2.3 Control Vector

The inversion window will cover a few days, up to 14. Within this time window, the system will solve for (see also section 4.2, for the detail of these terms):

- Hourly budgets of the CO_2 fluxes associated with:
    - Fossil fuel combustion
    - Wood burning
    - Crop residue burning
    - Ecosystem net primary production (NPP)
    - Ecosystem heterotrophic respiration (HR)
- The daily $\delta^{14}C$ signature of the heterotrophic respiration of ecosystems (δ_{HR})
- The $\delta^{14}C$ signature of wood burning $\delta_{bf,wood}$ and of crop biofuels $\delta_{bf,crop}$ emissions (at a resolution that is not fixed yet)
- Three-hourly budgets of the nuclear $^{14}CO_2$ fluxes.

For anthropogenic (fossil fuels and biofuels) CO_2 fluxes (and potentially δ_{bf}), the control vector will distinguish cities (agglomerations) in Luxemburg, in all administrative regions of Belgium, in seven administrative regions of southern Netherlands, in three administrative regions in northern France and three administrative regions in western Germany (all comprised in the 2 km × 2 km-resolution zoom, see Figure 3.1Figure 3.3). In these administrative regions, the CO_2/$^{14}CO_2$ emissions from major point sources (e.g. power plants for CO_2 and nuclear power plants for $^{14}CO_2$) are also controlled separately. The control vector also controls the rest of the emissions separately for each of these administrative regions.

Outside this detailed area, and for France, Germany and Netherlands, the inversion controls the full budget of fossil fuel and biofuel CO_2 emissions and of nuclear $^{14}CO_2$ emissions over the administrative regions (Figure 3.4). Coarser areas of control for the anthropogenic emissions are used for the rest of the domain (Figure 3.5).

Biogenic fluxes (NPP, HR and δ_{HR}) are only controlled at the resolution of administrative regions and larger area, i.e., the spatial resolution of the control vector is the same as for

anthropogenic emissions except that it does not distinguish agglomerations and major point sources.

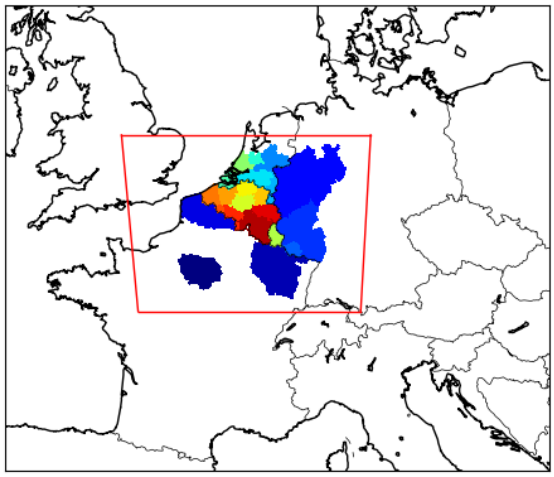

Figure 3.3 Administrative regions where agglomerations and point sources emissions are controlled separately in the inversion system at LSCE. The red line delimits the 2 km × 2 km-resolution zoom of the CHIMERE transport model.

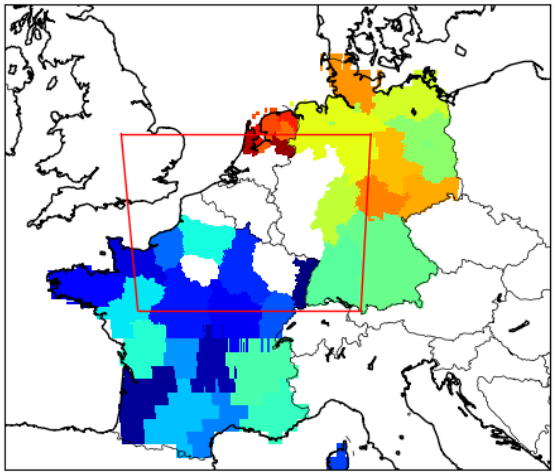

Figure 3.4 Administrative regions for which the full anthropogenic emission budgets are controlled in the inversion system at LSCE. The red line delimits the 2 km × 2 km-resolution zoom of the CHIMERE transport model.

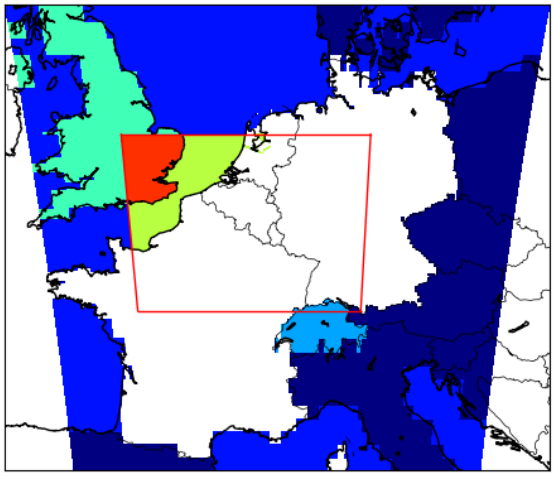

Figure 3.5 Coarser areas for which the emission budgets are controlled in the inversion system at LSCE. The red line delimits the 2 km × 2 km-resolution zoom of the CHIMERE transport model.

As mentioned in introduction of the presentation of this system, the proper co-assimilation of CO would likely require the separate control of anthropogenic emissions by main sectors of activities. However, we assume that it is much less critical for the co-assimilation of $^{14}CO_2$ and CO_2 while the dimension of the control vector defined above, when controlling the total fossil fuel or biofuel emissions for a given agglomeration, region, or area, is already very high and demanding in terms of computations. Therefore, we avoid the sectorial resolution in the control vector. However, of note is that we will conduct some tests of CO and CO_2 co-assimilation for a low dimensional inversion case focused on Belgium with only ten areas of control (nine in Belgium, and one for the rest of the domain), where six emission sectors are controlled separately.

3.2.4 Prior error covariance matrix

For the biogenic CO_2 fluxes (NPP, HR and δ_{HR}) and nuclear $^{14}CO_2$ fluxes, the setup of uncertainties in the prior estimate will rely on section 5 to be coherent with other partners. For the CO_2 fossil fuel and biofuel CO_2 anthropogenic fluxes, the setup of the statistics of uncertainty in the prior knowledge from inventories will rely on an analysis of the Monte Carlo ensemble of gridded inventories of emissions provided by TNO in WP4 as in Appendix A adapted to our control vector.

Meanwhile, we will still use a default setup with:

- A 1-σ prior uncertainty of 50% for the regional, city or point source hourly budgets of natural or anthropogenic fluxes from ecosystem models and inventories
- Temporal auto-correlation of this prior uncertainty with a 3-hour temporal scale
- No correlation of the prior uncertainties between different regions/cities/point sources, sectors and between natural and anthropogenic emissions.

For $\delta_{bf,wood}$ and $\delta_{bf,crop}$, the uncertainties will be derived from their spatial and temporal variabilities in Europe in 2015, using the ORCHIDEE-MICT emulator that provided the δ of the biomass (see section 4.2.2), these values may be updated for the sake of consistency with the data proposed in section 5.

3.2.5 Observation error covariance matrix

The errors associated with the CO_2 satellite observations will rely on satellite error simulations made by IUP Bremen within ESA projects (PMIF). The model error will be estimated from model comparisons in WP2 and analyses of the CHIMERE model forced by different meteorological forcings at LSCE. For CO_2 concentrations and $CO_2 \times \delta_{14C}$, the *in situ* observation-and-model-error will follow the guidelines of section 5 and will be complemented (if needed) by the diagnostics of Wang et al. (2017).

By default, we will first ignore temporal and spatial correlations in these errors. But we may finally include some of them to account for the systematic errors in the satellite data (simulated by IUP Bremen) in an appropriate way.

3.3 EMPA

3.3.1 Domain

EMPA will conduct forward simulations of tagged tracers of CO_2, CO, and $^{14}CO_2$ over Europe on a roughly 5 km x 5 km resolution domain at hourly time steps using the COSMO-GHG model.

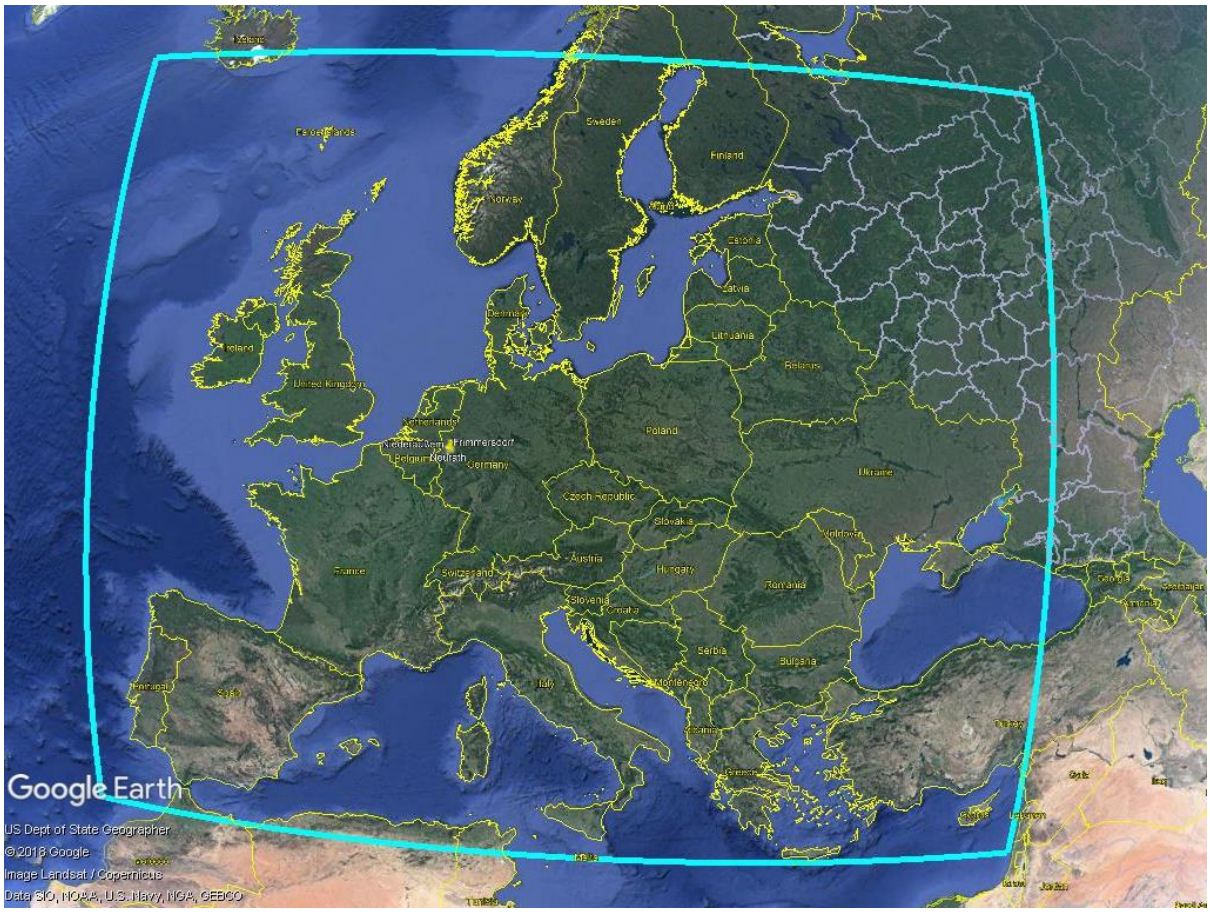

Figure 3.6 Domain of the COSMO-GHG simulations.

Table 3-2: Description of COSMO model grid.

Parameter	Zonal direction	Meridional direction
Rotated pole	-170.00º	43.00º
Start coordinates	-17.00º	-11.00º
Grid cell	0.05º	0.05º
Number of grid cells	760	610

3.3.2 Control vector

We aim to generate information that is complementary to that of MPG, which also simulates the entire European domain. We plan to include only few regions but with sector and fuel (fossil versus biofuel) split.

The European domain will be split in a few (approx. 5) coarse geographical regions. CO_2, CO and $^{14}CO_2$ tagged tracers will be simulated for different source types (traffic, power plants, others, for CO_2 and CO, GPP and R_{eco} for CO_2, anthropogenic, biogenic and nuclear for ^{14}C), also distinguishing between fossil and non-fossil fuels when needed.

Photochemical production of CO from NMVOCs will be considered through an additional three-dimensional source term as described in Section 4.4.5. Loss of CO by reaction with OH will be calculated using prescribed monthly mean OH fields averaged over an ensemble of chemistry-transport-models as described in Fiore et al., (2009) or the three-dimensional OH field of Spivakovski et al (2000).

Two one-month simulation periods in summer and winter 2015 will be selected. Monthly mean emissions for those species, categories and regions for the simulated time period will be estimated.

3.3.3 Prior error covariance matrix

No correlation between the different regions will be considered, unless the ensemble of emission scenarios from TNO indicates that such correlations are needed. The diagonal terms will be set consistently with the other modellers and the ensemble of emission scenarios from TNO (see section 5.5). For emissions from countries/regions that are not included in the TNO high-resolution European subdomain, the correlations will be estimated based on those available for the subdomain. For CO_2 and CO, the cross-species correlation terms will be derived from the TNO scenarios, though additionally considering the split between sectors. Similar covariance matrices will be produced for the CO_2+$^{14}CO_2$, CO_2+CO and CO_2+APO inversions, but we currently have too little information to provide any details at this point.

3.3.4 Observation error covariance matrix

A simple diagonal matrix will be considered in a first step. Since the resolution of the transport model is the same as that of WRF-GHG, these terms should be similar in both inversion setups. A more detailed matrix may be used in a further experiment.

3.4 NILU

3.4.1 Domain

The modelling domain consists of an inner domain (10.36°E - 10.94°E, 59.74°N - 60.02°N) around the Oslo centre including part of the Oslo fjord and an outer domain (8.87° E - 12.20° E, 58.95°N - 61.05° N) including surrounding urbanized areas is also included. Both the inner and outer domains have a resolution of ~1 km.

3.4.2 Control vector

Anthropogenic fluxes will be solved on a 3-hourly scale. Individual runs will be typically no longer than a week. Biogenic and oceanic fluxes will be considered for the background only and not resolved for. Synthetic runs for tracers representing CO, APO and ^{14}C will be also performed.

3.4.3 Prior error covariance matrix

For the diagonal terms in the prior error covariance matrix, we assume a relative uncertainty ranging between 50 – 100 %. The prior flux error is assumed to be 50 % for land grid cells and 100 % for water grid cells. Spatial correlation will be ignored for the off diagonal terms. The diagonal terms error covariance matrix for fluxes of CO, $^{14}CO_2$ and APO will be determined at a subsequent stage.

3.4.4 Observation error covariance matrix

The diagonal terms of the observation error covariance matrix for anthropogenic CO_2 range between 0.1 and 5 ppm. The diagonal terms of the observation error covariance matrix for CO, $^{14}CO_2$ and APO will be determined at a subsequent stage. Initially off-diagonal terms will be ignored. In supplementary observation system simulation experiments, temporal correlations of six hours and space correlation to represent transport model errors will be considered.

3.4.5 Forward simulations

FLEXPART driven by ECMWF meteorology at 0.1˚ x 0.1˚ resolution. Comparisons with output of FLEXPART-WRF planned.

4 Inversion inputs

4.1 CO2 simulations

4.1.1 Anthropogenic emissions

For WP2, TNO has produced the new TNO-GHGco inventory at ~5 km resolution over Europe for the year 2015, in which area and point sources are kept separate (TNO_GHGco_v1_1, longitude -30° to 60°, latitude 30° to 72°; 1/10° x 1/20° resolution). The inventories follow the G-NFR categories and are provided fuel type into fossil fuel and biofuel, and non-combustion fossil emissions. A high-resolution (~2 km) inventory over a smaller domain (TNO_GHGco_1x1km_v1_1, longitude -2° to 19°, latitude 47° to 56°; 1/60˚ x 1/120˚ resolution) is now available as well. This inventory includes ten different Monte Carlo realizations as an approximate estimate of uncertainty per point source or per gridbox. An analysis of the uncertainty is included in appendix A. These data can be downloaded from their ftp site, with access information available from Hugo Denier van der Gon (hugo.deniervandergon@tno.nl) upon request. Furthermore, EMPA has generated specific time functions for the different source sectors contained in the TNO-GHGco or EDGAR inventories in order to add temporal variability in emissions, e.g. residential heating increases with colder weather. Suggested vertical profiles depending on sector were also provided. The vertical profiles are only to be applied to point sources, while area emissions are released from the surface.

4.1.2 Biogenic fluxes

The biogenic fluxes for use in the both the WP2 forward simulations and the WP4 inversions are derived from the VPRM model (Vegetation Photosynthesis and Respiration Model). VPRM is a simple light-use-efficiency model driven by a combination of satellite reflectances from MODIS and meteorological input data. The MODIS reflectances are used to compute the indices EVI (enhanced vegetation index) and LSWI (land surface water index) at 8-day resolution. Due to noise in the signal and occasional missing pixels, these indices are loess filtered. The meteorological drivers are 2-m temperature and downward shortwave radiation at the surface. These were taken ECMWF's Tier 1 simulation produced within CHE at 0.1° spatial resolution and 3-hourly resolution. Because the VPRM fluxes are calculated at with an hourly time step while the meteorological input is only available in 3-hourly resolution, linear interpolation was employed. Fluxes are calculated for eight different land cover types (evergreen forest, deciduous forest, mixed forest, shrubland, trees and grasses, cropland, grassland, and other). Parameters for each land cover type determined through optimization with flux tower data from 47 sites for the year 2007, as described in Kountouris et al. (2018). The fractional land cover type per pixel is determined using the SYNMAP land cover product (Jung et al., 2006). VPRM CO_2 fluxes are split in two terms: Gross Primary Production ($F_{GPP,VPRM}$) and respiration ($F_{R,\ VPRM}$), which sums autotrophic and heterotrophic respiration. The model is more fully described in Mahadevan et al. (2008).

The VPRM simulations will be run at a 5 km x 5 km resolution for the whole European domain and at a 1 km x 1 km resolution for the regional domains (Oslo, Northern France and Benelux). In both cases, the simulations will be run at an hourly time step. Since the WRF-GHG simulations of MPG have the largest of domain, defined with a standard longitude of 12.5°E and true latitude of 51.604°N, with 962 longitudinal gridboxes and 776 latitudinal gridboxes at 5- km resolution. The 5-km data can be found in on the CHE ftp site under data-exchange/WP2/Biogenic_emissions/EU5km/. One level higher there is a netcdf file providing the latitudes and longitudes of the corners of each 5-km gridbox

(corners_EU5km.nc). There is one file per day, with 24 time steps per file. The fluxes corresponding to "0" are to be used for the hour between midnight and 1:00 UTC. For groups simulating higher resolution (~1 km) over a sub-domain, fluxes on the same grid as is used for the nested TNO domain have been produced as well.

4.1.3 Oceanic fluxes

Ocean fluxes CO_2 will be extracted from the NEMO-PlankTOM5 (Buitenhuis et al., 2010) which simulated gridded ocean fluxes for O_2, CO_2 and heat for 1959 to 2018. The model simulations start from observations in 1920s and are therefore not at equilibrium. Therefore the variability in these simulations is more robust than the mean fluxes. These fluxes will be adjusted using inverse modelling based fluxes (Rödenbeck et al., 2015) to avoid biases.

4.1.4 Biomass burning

Daily GFASv1.2 emissions at 0.1º x 0.1º spatial resolution, available from CAMS, should be used to represent realistic atmospheric variability. The GFASv1.2 emissions also include information on the plume injection height at daily time steps as provided by a plume rise model (mean altitude of maximum injection and altitude of plume top are provided).

4.1.5 Photochemical source

Photochemistry plays a non-negligible role in the CO budget. Since these photochemistry results in the production of CO_2, we shall include the three-dimensional CO_2 source derived from the reaction CO + OH → CO_2. For these we directly obtain the mass flux of CO_2 being generated this reaction from the CO inversion scenario 3 in (Zheng et al., 2019). This is a multispecies inversion assimilates measurements from MOPITT CO total column, surface methylchloroform (CH_3CCl_3) measurements, OMI formaldehyde total column and GOSAT CH_4 total column measurements. An estimate of the atmospheric signal generated by the photochemical production of CO_2 is shown in Appendix B.

4.1.6 Initial/Boundary conditions

The initial and boundary conditions for CO_2 should be taken from the Tier 1 simulation, described in D2.2. Ideally the initial conditions for CO_2 would include the CO_2 produced from CO oxidation. However, if this is not possible due to computational costs or model constraints, each group should decide to either to add the mole fraction fields in Appendix B to their initial conditions or to neglect this additional mole fraction.

4.2 $^{14}CO_2$ simulations

This section further on presents the required input data to simulate $^{14}CO_2$ transport. We follow the formulation of the $^{14}CO_2$ transport used in the inverse modelling study of Wang et al. (2016):

$$C_a\delta_a = H_{transp}\begin{bmatrix}\delta_{FF}F_{FF} + \delta_{bF}F_{bF} + \delta_{NPP}F_{NPP} + \delta_{HR}F_{HR} + \\ \delta_{oa}F_{oa} + \frac{1}{R_{std}}F^{14}_{Nucl} + \frac{1}{R_{std}}F^{14}_{cosm}\end{bmatrix} + H_{bc}[C_{bc}\delta_{bc}], \quad \text{Eq. (23)}$$

where:

- C_a is the CO_2 concentration and δ_a the $^{14}CO_2/^{12}CO_2$ ratio in the atmosphere, normalized by the $^{14}C/^{12}C$ ratio in the Modern Standard (R_{std}= 1.176 x 10^{-12}). Similarly, in the following, all δ are also normalized ratios.
- F_x terms correspond to different types x of CO_2 surface fluxes within the transport modelling domain:
    - Fossil fuel emissions (FF)
    - Bio-fuel emissions (bF)
    - Net primary production by vegetation (NPP)
    - Heterotrophic respiration from the soil (HR)
    - Emissions from the ocean to the atmosphere (oa)
- F^{14}_x terms correspond to $^{14}CO_2$ fluxes from nuclear power plants (Nucl) and cosmogenic production (cosm).
- δ_x coefficients correspond to the $^{14}CO_2$ abundance in the fluxes listed above.
- $C_{bc}\delta_{bc}$ are the boundary (top, lateral) and initial conditions of $^{14}CO_2$ concentrations.

In this section most of the δ in the fluxes and in the biomass ($\delta_{biomass}$) are calculated from products of an emulator of the ORCHIDEE-MICT model (Guimberteau et al., 2018). This emulator is forced by historical changes in global ^{14}C, rising CO_2, variable climate and land use change since 1850 (Wang et al., in prep). It provides daily data, per plant functional types (PFT), at 0.5° resolution, over Europe (longitude -15° to 35°; latitude 35° to 70°).

4.2.1 Anthropogenic emissions

F_{FF} and F_{bf}: fossil fuel and bio-fuel emissions are provided by the TNO inventories in 2015. The corresponding δ are applied to these fluxes:

- δ_{FF} = -1000 ‰, should be applied on the whole year and domain
- For biofuels (bF), we distinguish δ for wood burning fluxes ($\delta_{bf,wood}$) from the one for crop-fuel burning fluxes ($\delta_{bf,crop}$) because of their different growing periods and the resulting ^{14}C abundance difference. In a first approximation, we use, as δ_{bF}, the temporal (over the year 2015) and spatial mean (over the ORCHIDEE-MICT Europe domain) of the $\delta_{biomass}$ from the emulator of the ORCHIDEE-MICT model mentioned above. Such a computation of δ_{bF} relies on the strong hypothesis that the wood or cropfuel burnt in Europe comes from local and recently cut vegetation. Actually, imports, stocks and delays between cutting and burning wood or cropfuel are common.
    - $\delta_{bf,wood}$ = 84 ‰ for wood bio-fuel (GNRF categories A to C). We consider that bio-fuels burned in power plants, industry and residential sector, correspond to wood burning of European trees. Thus we use a spatial and temporal mean of $\delta_{biomass}$ for PFT corresponding to non-tropical trees, over Europe in 2015.
    - $\delta_{bf,crop}$ = 18 ‰ for crops bio-fuels (GNRF categories F and L). We make the hypothesis that these sectors (Road transport and Agriculture) burn local crops, harvest in 2015. Thus we use a spatial and temporal mean of $\delta_{biomass}$ for PFT corresponding to crops, over Europe in 2015.
    - Other sectors are considered to have negligible bio-fuel emissions.

4.2.2 Biogenic fluxes

F_{NPP} and F_{HR}: the VPRM simulations by MPI-Jena over Europe (latitude 31° to 68.7°; longitude -35.5° to 60.5°, 5 km x 5 km, 1 h, in 2015) provide estimates of gross primary production (GPP) and respiration (F_{GPP}^{VPRM} and F_{Resp}^{VPRM}). LSCE provides a daily partition coefficient (α_{HR}, calculated from the above mentioned ORCHIDEE-MICT simulations Guimberteau et al, 2018) to reallocate GPP and Respiration from VPRM into NPP and HR fluxes. This daily coefficient has the spatial resolution of VPRM simulation but on a restricted area (longitude -15° to 35°; latitude 35° to 68.7°) corresponding to the available ORCHIDEE-MICT simulation (longitude -15° to 35°; latitude 35° to 70°). It must be simply multiplied by F_{Resp}^{VPRM} to get F_{HR} and then we get F_{NPP} following these equations:

$$F_{HR} = \alpha_{HR} F_{Resp}^{VPRM}, \quad \text{Eq. (24)}$$

$$F_{NPP} = F_{GPP}^{VPRM} + (1 - \alpha_{HR}) F_{Resp}^{VPRM}. \quad \text{Eq. (25)}$$

δ_{NPP} monthly maps, for year 2015, are provided by LSCE. These maps have the spatial resolution of VPRM simulation but restricted to the area (longitude -15° to 35°; latitude 35° to 68.7°) depending on the domains of the 3 land cover maps used to process this dataset (VPRM and ORCHIDEE land cover maps and MIRCA2000 crop map (Portmann et al., 2010). To calculate δ_{NPP}, the following equation was used:

$$\delta_{NPP} = \delta_{a,surf} - \epsilon, \quad \text{Eq. (26)}$$

where

- $\delta_{a,surf}$ is the radiocarbon ratio in the surface atmospheric layer. $\Delta^{14}C$ monthly background measurements at Jungfraujoch and Schauinsland, in 2015 (Hammer et al., 2017) are used to characterize it. The mean of the 2 measurement sites is calculated and converted to δ following Stuiver et al. (1977). Here we neglect the impact of variations of this $\delta_{a,surf}$ at high resolution on the NPP fluxes themselves. Otherwise for a precise computation of the $\delta_{NPP,}$ and so $^{14}CO_2$ NPP fluxes, we should dynamically calculate it with $\delta_{a,surf}$ depending on $^{14}CO_2$ concentrations calculated by the transport model.
- ε is the sum of kinetic and enzymatic $^{14}CO_2$ fractionation with respect to $^{12}CO_2$ depending on the C3 or C4 photosynthesis pathway of the vegetation. ε corresponds to the double of the fractionation of $^{13}CO_2$ ($^{13}\varepsilon$) during photosynthesis: ε= 36 ‰ for C3 vegetation ($^{13}\varepsilon$ = 18 ‰, Degens, 1969) and 8 ‰ for C4 vegetation ($^{13}\varepsilon$ = 4 ‰, Farquhar et al, 1989). The C3/C4 monthly repartition on the VPRM grid relies on combination of three landcover maps: VPRM and ORCHIDEE land cover maps and monthly MIRCA2000 crop map.

δ_{HR} daily maps for year 2015 are provided by LSCE (e.g. Figure 4.1). These maps have the spatial resolution of VPRM simulation but on a restricted area (longitude -15° to 35°; latitude 35° to 68.7°). Indeed, their generation relies on simulations of the above-mentioned ORCHIDEE-MICT emulator. For each grid cell, the daily components of the CO_2 and $^{14}CO_2$ heterotrophic respiration were aggregated (litter respiration and 3 types of soil respiration) to calculate δ_{HR} and then project it on the VPRM grid.

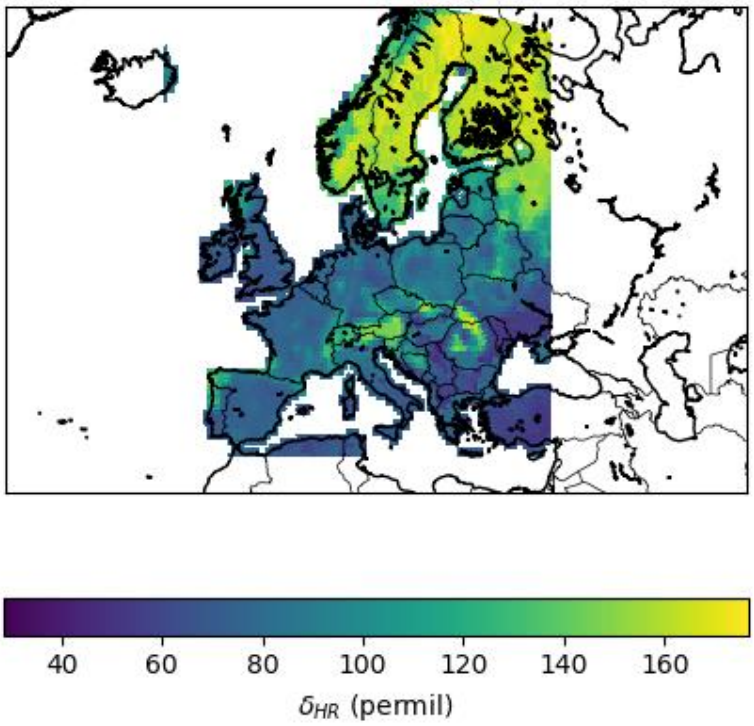

Figure 4.1 δ_{HR} (‰) on the VPRM grid on January 1, 2015. Values available only on a restricted area corresponding to the ORCHIDEE-MICT simulation domain.

4.2.3 Oceanic fluxes

Several of the existing stations are located near the shore and might be sensitive to oceanic $^{14}CO_2$ fluxes. For example, (Hsueh et al., 2007) found that while ocean exchange contributed only minimally to continental $^{14}CO_2$ patterns across North America, this flux induced variations in excess of 1‰ near coastal zones. The $\Delta^{14}CO_2$ source signature of ocean fluxes will be extracted from a NEMO-PlankTOM5 simulation (Khatiwala et al., 2018).

4.2.4 Biomass burning

Biomass burning emissions should be taken from GFASv1.2, as described in the section 4.1.4 for CO_2. LSCE does not provide biomass burning $^{14}CO_2$ emission but $\delta_{biomass}$ from ORCHIDEE-MICT will be used to derive the signal of derive these emissions.

4.2.5 Nuclear power plants

LSCE provides a $^{14}CO_2$ nuclear emission database as a listing of point source emissions. Nuclear $^{14}CO_2$ emissions are simply calculated following (Graven & Gruber, 2011) from the annual activity of each reactor, in 2015, reported in (Zazzeri et al., 2018). For each reactor, activity data A is converted into ^{14}C production in kg ^{14}C reactor^{-1} a^{-1} normalized by R_{std}:

$$^{14}C_{Nucl} = \mathrm{A} \cdot \alpha \cdot 10^{9}, \qquad \text{Eq. (27)}$$

with α = R_s/0.226, where R_s = 1.176 × 10^{-12} is the $^{14}C/C$ ratio in the Modern Standard and 0.226 Bq/gC is the conversion factor from activity to carbon production. Analysis of the data at LSCE and at EMPA could be used to define uncertainties in the temporal variability of these emissions.

4.2.6 Cosmogenic

We neglect cosmogenic production of ^{14}C in the upper atmosphere because we assume that it is a small source within the European domains considered in CHE. Cosmogenic fluxes would impact the atmosphere above typically 700 hPa, well above the planetary boundary layer (Turnbull et al., 2009), while we are interested in simulating $^{14}CO_2$ concentrations near the ground. Even though we use some high altitude stations, we can assume that most of the influence from cosmogenic production at these surface stations comes from the model lateral boundaries and that the cosmogenic production within the modelling domain can be neglected. Nevertheless, while we neglect cosmogenic production of ^{14}C in the upper atmosphere within our European domains, cosmogenic production is included in the simulation from which the initial and boundary conditions were derived.

4.2.7 Initial and boundary conditions

Three-dimensional monthly fields of δ_{14C} radiocarbon and CO_2 concentrations for 2015 are provided by LSCE at 2.5° x 3.75° (latitude x longitude) resolution horizontally and with 19 vertical layers (pressure levels, a and b pressure coefficients are also provided), over an area larger than the largest of the European modelling domains in WP4, so that each partner can extract and interpolate their own initial and boundary conditions. This product is derived from the global simulation (1998-2007) with LMDZ by Wang, (2016) with a trend added to approximate monthly fields of δ_{14C} and CO_2 for the year 2015.

4.3 APO simulations

The overall APO surface flux can be written as:

$$\boldsymbol{f}^{APO} = \boldsymbol{f}^{O_2} + \alpha \boldsymbol{f}^{CO_2} - \frac{y_0^{O_2}}{y_0^{N_2}} \boldsymbol{f}^{N_2}, \qquad \text{Eq. (28)}$$

where

$$\boldsymbol{f}^{O_2} = \boldsymbol{f}^{O_{2,FF}} + \boldsymbol{f}^{O_{2,BF}} + \boldsymbol{f}^{O_{2,BB}} + \boldsymbol{f}^{O_{2,OC}} \qquad \text{Eq. (29)}$$

$$\boldsymbol{f}^{CO_2} = \alpha_{FF}\boldsymbol{f}^{CO_{2,FF}} + \alpha_{BF}\boldsymbol{f}^{CO_{2,BF}} + \alpha_{BB}\boldsymbol{f}^{CO_{2,BB}} + \alpha_{OC}\boldsymbol{f}^{CO_{2,OC}} \qquad \text{Eq. (30)}$$

$$\boldsymbol{f}^{N_2} = \boldsymbol{f}^{N_{2,OC}} + \boldsymbol{f}^{N_{2,soil}} + \boldsymbol{f}^{N_{2,photochem}}. \qquad \text{Eq. (31)}$$

FF, BF, BB, and OC refer to fossil fuel, biofuel, biomass burning and ocean processes respectively. α_L is the stochiometric O_2/CO_2 ratio for each of the processes. $y_0^{N_2}$ and $y_0^{O_2}$ are the reference N_2 and O_2 mole fraction (780800 and 209460 ppm respectively). Notice that any contribution from the land biosphere is cancelling out to the extent that the stoichiometric ratio, α_L is realistic. Nitrogen fluxes from nitrification, denitrification and photochemistry are neglected ($f^{N_{2,soil}}, f^{N_{2,photochem}}$). To convert APO into an O_2 mole fraction for modelling:

$$y^{O_2} = y_0^{O_2}\left[\frac{ppm}{per\ meg}\right]APO. \quad \text{Eq. (32)}$$

4.3.1 Anthropogenic emissions

Initial work involved deriving the O_2 consumption fluxes directly from the TNO CO_2 emission inventory for complete consistency. The α_{FF} and α_{BF} will be assigned either using values from the literature (Steinbach et al., 2011) or from H:C ratios (based on the fact that for most fuels, H:C ratios and O_2/CO_2 ratios are strongly correlated). For the O_2/CO_2 ratios, TNO used the fuel type (solid, liquid, gaseous, biomass and other - mainly waste) as reported in the UNFCCC by each country. For fossil fuels, the O_2:CO_2 ratio (α_{FF}) ranges from ~-1.20 to ~-1.95. For biofuel combustion, we will use an oxidative ratio of 1.07 (Steinbach, 2010). Uncertainties in these ratios will also be provided where possible using the range of values in the literature, theoretical calculations, and existing observation data. For the calculation of the O_2:CO_2 ratios, the following assumptions are made:

a) Sulphur and nitrogen content has negligible impact for most fuels,

b) Fuels are burnt completely, and any CO produced is short-lived in the atmosphere and converted to CO_2 relatively quickly (Keeling et al. 1988),

c) Energy production is nearly proportional to O_2 consumption across a wide range of fuels, based on relatively similar net higher heating (NHHV) values (in kcal/mole) across different types of fuels (Keeling 1988).

Furthermore, APO offers the possibility to separate between fossil fuel combustion emissions and fossil non-combustion emission, which do not consume oxygen. Because the database underlying the TNO inventory contains industry sector emissions divided between combustion and non-combustion sources, we are able to represent this apart for emission from the cement production reaction, from other non-combustion processes including CO_2 emissions from blast furnaces in iron and steel, and from lime production.

4.3.2 Biogenic fluxes

Globally, the O_2:CO_2 ratio (α_L) of terrestrial biospheric exchange is -1.1 on a yearly scale. At the regional scale (i.e. Europe) on sub-annual timeframes, α_L may differ slightly from -1.1 (by up to ± 0.1). We will assume a O_2:CO_2 ratio for according to (Clay & Worrall, 2015; Worrall, Clay, & Macdonald, 2015) where an average α_L value is given per biome (for NPP) and soil type (for respiration). No diurnal, seasonal or latitudinal variability in α_L is assumed.

Table 4-1 Range and median values of the oxidative ratio (O_2/CO_2) of vegetation according to biome based on Clay & Worrall (2015)

Biome	O_2/CO_2 Median	O_2/CO_2 Range
Evergreen forest	1.07	1.03 – 1.08
Deciduous forest	1.08	1.03 – 1.13

Croplands	1.05	0.98 – 1.12
Woody savannah	1.06	1.03 – 1.11
Grasslands	1.03	0.79 – 1.30
Mixed forest	1.07	1.02 – 1.12
Cropland/natural mosaic	1.05	1.02 – 1.07
Closed shrublands	1.10	1.03 – 1.12
Open shrublands	1.02	
Savannah	1.09	1.05 – 1.12
Permanent wetlands	0.99	0.93 – 1.03
Urban	1.02	
Snow/ice	0.00	
Barren	0.00	
Unclassified	1.02	

Table 4-2 Range and median values of the oxidative ratio (O_2/CO_2) for each soil type of the USDA global soil based on Clay & Worrall (2015).

Biome	O_2/CO_2 Median	O_2/CO_2 Range
Alfisols	1.11	0.77 – 1.37
Andisols	1.03	
Aridisols	1.02	
Entisols	1.07	0.99 – 1.11
Histosols/Gelisols	1.08	0.84 – 1.24
Inceptisols	1.14	1.06 – 1.24
Mollisols	1.14	
Oxisols	1.00	0.80 – 1.06
Spodosols	1.08	
Ultisols	1.00	0.85 – 1.04
Vertisols	1.01	0.80 – 1.08

4.3.3 Oceanic fluxes

Gridded ocean CO_2, O_2 and heat fluxes will be extracted from the NEMO-PlankTOM5 model (Buitenhuis et al., 2010). The model simulations start from observations in 1920s and are therefore not at equilibrium and CO_2 fluxes have been manually tuned to match observations during the 1990. Therefore, there could be inconsistencies between the CO_2 and O_2 fluxes. As such, both APO fluxes shall be adjusted using inverse modelling based fluxes (Rödenbeck et al., 2015) to avoid biases. MPG has provided APO fluxes based on Jena Carboscope inversions using 9 background sites measuring O_2, based on the method described in Rödenbeck et al. (2008). These fluxes are from the apo99_v1.6 simulation documented (available under http://www.bgc-jena.mpg.de/CarboScope/). The fluxes have been uploaded to the CHE ftp site under data-exchange/WP4/APO/apo99_v1.6.nc.

Sea-air nitrogen fluxes arise mostly from temperature-induced solubility changes in the ocean. Therefore, ocean N_2 fluxes can be estimated from NEMO-PlankTOM5 heat fluxes and the Keeling et al. (1993) formula. Voluntarily some of the groups involved may want to include ocean N_2 fluxes in order to test their impact on model APO signals.

4.3.4 Biomass burning

Based on Steinbach (2010), a ratio of 1.11 corresponds to the typical biomass burning process where CO emissions account for ~10% of CO_2 emissions (Crutzen & Andreae, 1990). However, depending on the type of fire, the combustion process might be less complete and the CO/CO_2 emission ratio might go up to ~30%. For these cases, oxidative ratios up to 1.41 have been observed (Lueker et al., 2001). We will assume a linear relationship between the oxidative ratio and the CO/CO_2 ratio obtained from GFASv1.2:

$$\frac{O_2}{CO_2} = 1.55\ \frac{CO}{CO_2} + 0.945. \qquad \text{Eq. (33)}$$

4.3.5 Photochemistry

Stephens et al. (1998) introduced a correction term for the influence of photochemical CH_4 and CO oxidation because the reaction with OH represents a net loss of atmospheric O_2. This is represented as

$$\delta(O_2/N_2)_{corr} = \delta(O_2/N_2) - \frac{2}{y_0^{O_2}} y^{CH_4} - \frac{0.5}{y_0^{O_2}} y^{CO}, \qquad \text{Eq. (34)}$$

where y^{CH_4} and y^{CO} are the CH_4 and CO mole fractions in ppm. Since we are simulating CO in this study, we can account for this correction while neglecting the effect of CH_4.

4.3.6 Initial/Boundary conditions

For the concentration fields, initial and lateral boundary conditions can be taken from optimized concentration fields from the Jena Carboscope inversions using 9 background sites measuring O_2, based on the method described in Rödenbeck et al., (2008). These concentrations are based on a forward run of the optimized fluxes from the apo99_v1.6 simulation, as documented (available under http://www.bgc-jena.mpg.de/CarboScope/). The concentration fields have been uploaded to the CHE ftp site under data-exchange/WP4/APO/apo99_v1.6_mix_2015.nc.

4.4 CO simulations

4.4.1 Anthropogenic emissions

As for CO_2, CO emissions are available from the TNO-GHGco inventory at 5-km resolution over Europe and at 1-km resolution for a subset of the domain with an ensemble of 10 realizations. The inventory follows the G-NFR categories and the transport category is separated into fossil fuel and biofuel. A detailed analysis of the CO/CO_2 ratio is found in appendix A.

4.4.2 Vegetation

A small amount (up to 200 Tg CO a^{-1}) is directly generated from plant leaves. For biogenic CO emissions we will use emissions from the Model of Emissions from Gases and Aerosols from Nature (MEGANv2.1) (Guenther et al., 2012; 2006; Sindelarova et al., 2014) estimated within the Copernicus Atmosphere Monitoring Service (CAMS) (Granier et al., 2019). The parameterization for CO emissions was based on (Tarr et al., 1995). The MEGAN framework calculates ecosystem-specific emissions scaled to leaf area, light, and temperature. It is driven by ERA Interim meteorological reanalysis data (Dee et al., 2011). Emissions are provided at hourly time steps and 0.5º x 0.5º resolution.

4.4.3 Soils

Soils are generally considered as a sink of CO due to microbial oxidation processes. However, CO emissions have been reported from a wide range of soil–plant systems and it is not a result of metabolic processes (Conrad & Seiler, 1980). Soil production and consumption fluxes were obtained from (Liu et al., 2018).

4.4.4 Oceanic fluxes

Ocean processes are a minor source in the atmospheric CO budget (6.3 Tg CO a^{-1}, Conte et al., 2019). However, these emissions may play non-negligible role at coastal sampling sites. We account for ocean CO fluxes by including the estimates by the Nucleus for European Modelling of the Ocean - Pelagic Interaction Scheme for Carbon and Ecosystem Studies (NEMO-PISCES) in Conte et al., (2019). The NEMO-PISCES simulation accounts for ocean general circulation, biogeochemistry including the main CO sources and sinks (photoproduction and bacterial consumption), and sea-air exchange.

4.4.5 Biomass burning

The biomass burning fluxes for CO should be taken from GFASv1.2 as described for CO_2.

4.4.6 Photochemical

The oxidation of formaldehyde (HCHO), which itself results from the oxidation of CH_4 and NMVOCs and from biogenic emissions, is the single largest source term in the CO budget. Analogously, the reaction CO + OH → CO_2 is the largest sink term in the CO budget. Interested groups may include the three-dimensional CO source derived from the oxidation of HCHO. For this, the mass flux of CO from the CO inversion was obtained from the scenario 3 simulation in Zheng et al., (2019). This is a multispecies inversion, which assimilates measurements from MOPITT CO total column, surface methylchloroform (CH_3CCl_3) measurements, OMI formaldehyde total column and GOSAT CH_4 total column measurements. An estimate of the atmospheric signal generated by the photochemical production of CO is shown in Appendix B. A first order loss rate is applied to CO to account for its reaction with OH based on the rate constants determined by Liu & Sander (2015) and the OH climatological concentration fields from Spivakovsky et al., (2000) which were distributed within the TRANSCOM-CH4 experiment Patra et al. (2011). These reaction rate constants *k* [$cm^3\ molecule^{-1}\ s^{-1}$] are given by

$$k = k_{bi}(T) + \frac{k_0(T)[M]}{1+\frac{k_0(T)[M]}{k_\infty(T)}} \times 0.6^{\left\{1+\left[\log_{10}\left(k_0(T)[M]/k_\infty(T)\right)\right]^2\right\}^{-1}}, \quad \text{Eq. (35)}$$

where

$$k_{bi}(T) = Ae^{-\left(E_a/RT\right)}, \quad \text{Eq. (36)}$$

$$k_0(T) = k_0^{300}\left(\frac{T}{300}\right)^{-n}, \quad \text{Eq. (37)}$$

$$k_\infty(T) = k_\infty^{300}\left(\frac{T}{300}\right)^{-m}. \quad \text{Eq. (38)}$$

Here *T* is temperature [K], $[M]$ is the air number density [molecule cm^{-3}], k_{bi} [$cm^3\ molecule^{-1}\ s^{-1}$] is the pressure independent bimolecular rate constant, *A* [$cm^3\ molecule^{-1}\ s^{-1}$] is the Arrhenius constant (1.54×10^{-13}), E_a [kJ mol^{-1}] is the activation energy (13), *R* [kJ mol^{-1} K^{-1}] is the ideal gas constant, k_0 and k_∞ [$cm^6\ molecule^{-2}\ s^{-2}$] are the low- and high-pressure limiting rate constants derived from measurement values at 300 K (6×10^{-33} and 1.1×10^{-12} respectively). The parameters *m* and *n* are empirically determined (-1.3 and 1.9).

4.4.7 Initial/Boundary conditions

The initial and boundary conditions should be taken from the Tier 1 simulation, which is documented in D2.2. Ideally the initial and boundary conditions for CO_2 would include CO photochemistry (see appendix B). However, if this is not possible due to computational costs or model constraints, each group should decide to either to add the mole fraction fields in Appendix B to their initial conditions or to neglect this additional mole fraction.

4.5 Emission datasets summary

Table 4-3: Summary of source and sink processes considered, their relationship to the tracers simulated.

Process	CO_2	$\delta^{14}CO_2$	CO	APO

Fossil fuel combustion	TNO GHGco	-1000 ‰	TNO GHGco	TNO GHGco + COFFEE
Biofuel combustion	TNO GHGco	ORCHIDEE-MICT δ_{BF} Wood: 84 ‰ Crop-fuel: 18‰	TNO GHGco	TNO GHGco + COFFEE
Non-combustion fossil	TNO GHGco	-1000‰	0	0
Gross primary productivity	VPRM	ORCHIDEE-MICT δ_{NPP} + daily partitioning factor NPP/GPP	CAMS-BIO MEGANv2.10 Global biogenic emission	-1.07±0.04 (biome dependent)
Respiration	VPRM	ORCHIDEE-MICT δ_{HR}	Soil fluxes (Liu et al., 2018)	-1.07±0.04 (soil class dependent)
Ocean	NEMO-PlankTOM5 (scaled to Jena Carboscope global total)	NEMO-PlankTOM5 Khatiwala et al. (2018)	(Conte et al., 2019)	NEMO-PlankTOM5 (scaled to Jena Carboscope global total)
Biomass burning	GFASv1.2	ORCHIDEE-MICT δ_{BF} Wood: 84 ‰ Crop/grass: 18‰	GFASv1.2	-1.11 to -1.49
Nuclear power plants	0	LSCE database Graven et al. (2011), Zazzeri et al. (2018)	0	0
Photochemical	LMDZ 3D fields of CO destruction from inversion (Zheng et al., 2019)	LMDZ 3D fields of $\delta^{14}CO2$ and CO2 concentrations, trend corrected	LMDZ 3D fields of CO production from inversion (Zheng et al., 2019)	$\frac{0.5}{y_0^{O_2}} y^{CO}$
Cosmogenic	0	LMDZ 3D fields of $\delta^{14}CO2$ and CO2 concentrations, trend corrected	0	0
Initial/Boundary conditions	Tier 1 simulation	LMDZ 3D fields of $\delta^{14}CO2$ and CO2 concentrations, trend corrected	Tier 1 simulation	From APO inversion apo99_v1.6

5 Inversion comparability

In WP4, several groups will performed simulations of uncertainty reduction in fossil fuel CO_2 emissions with different models and targeting different regions, time resolutions and emission partitioning. However, it is necessary to have comparable quantities.

5.1 Measurement error

We set up comparable measurement errors. In Table 5-1, we have the typical and target measurement errors for the atmospheric species considered in this study. We recommend assigning these errors to all data and considering scenarios with the typical and with the target measurement uncertainties.

Table 5-1: Typical and target measurement error for atmospheric species considered in this study.

Species	Typical measurement error	Target measurement error	Unit
CO_2	0.10[a]	0.05[a]	ppm
$\delta^{14}CO_2$	2.00[b]	0.50[a]	‰
CO	5.00[a]	2.00[a]	ppb
$\delta(O_2/N_2)$	5.00	2.00	permeg
APO[c]	1.00	0.20	ppm[d]

[a]: Crotwell & Steinbacher (2018)
[b]: Gamnitzer et al. (2006)
[c]: Heimann, (2007)
[d]: To have a comparable mole fraction basis a conversion factor of 4.77 permeg ppm^{-1} (Keeling et al., 1998) was assumed.

If $\delta^{14}CO_2$ and APO measurements are modelled as independent CO_2, $^{14}CO_2$, O_2 and even N_2 mole fractions the uncertainty must be propagated. For $^{4}CO_2$, this is:

$$\sigma_{^{14}C} = y_{^{14}C}\sqrt{\left(\frac{\sigma_{CO_2}}{y_{CO_2}}\right)^2 + \left(\frac{\sigma_{R_{14C/12C}}}{R_{14C/12C}}\right)^2 - \frac{\mathrm{COV}(y_{CO_2}, R_{14C/12C})}{y_{CO_2} \cdot R_{14C/12C}}}, \qquad \text{Eq. (39)}$$

where y [ppm] is the mole fraction, σ [ppm] is the uncertainty and R is the $^{14}CO_2/CO_2$ ratio. In the case of APO, it must be taken into account the uncertainty of APO is the sum of the uncertainty of the $\delta(O_2/N_2)$ ratio and the CO_2 mole fraction,

$$\sigma_{APO} = \sqrt{\sigma^2_{\delta(O_2/N_2)} + \left(\left|\frac{\alpha_L}{y^{O_2}_0}\right| \sigma_{CO_2}\right)^2 + 2COV\left(\delta(O_2/N_2), \left|\frac{\alpha_L}{y^{O_2}_0}\right| y_{CO_2}\right)}. \qquad \text{Eq. (40)}$$

5.2 Model error

The model error reflects the ability of the transport model to realistically simulate the atmospheric mole fraction. It is expected that this ability markedly varies according to the complexity of the local circulation as well as that of the surface fluxes surrounding the individual sites. Initially, we propose that all sites should be classified according to their

location into the following classes: remote, shore/oceanic, mountain, tall tower, continental, and urban/strong sources. These site classes are then assigned different model uncertainties (without any time dependence) for the standard inversion set-up. The base value (Table 5-2) refers to remote station. Other station classes are assigned the multipliers in Table 5-3. Although the values are rather arbitrary (given the poor knowledge about the actual model errors), this scheme is simple and allows testing different values easily.

Table 5-2: Base model error assumed for the different species.

Species	Base model error	Unit	Base model error in alternative units[a]	Unit
CO_2	1.00	ppm		
$^{14}CO_2$	1.26×10^{-12}	ppm	3.00	‰
CO	7.00	ppb		
O_2	1.00	ppm	4.77	permeg
APO	1.00	ppm	4.77	permeg

[a]: Alternative units were calculated base on a $\delta^{14}CO_2$ signal of 17‰, a CO_2 mole fraction of 350 ppm and an oxygen mole fraction of 209500 ppm.

Table 5-3: Station class multipliers.

Station Class	Multiplier
Remote	1.0
Shore/Oceanic	1.5
Continental	3.0
Mountain	1.5
Tall tower	1.0
Urban/strong sources	5.0

The proposed values are not well known and rather arbitrary, especially for $\delta^{14}CO_2$, CO and APO. Furthermore, CO is not a well-mixed tracer. Thus, it may exhibit strong spatial gradients especially in the summer months. An alternative approach would be for one of the groups performing three-dimensional simulations of the entire European domain, or at least most of it, to produce a climatology (monthly or weekly resolution) of the standard deviation of the mole fraction per gridcell with the gridcells around it. This approach has been used in past publications (Rödenbeck et al., 2003)

5.3 Stations

The main objective of WP4 is to determine the optimal network to estimate fossil fuel combustion CO_2 emissions. Therefore, we shall perform uncertainty reduction calculations with different network configurations. We have created the two overall networks proposals: a more rural/background configuration (but which also includes some urban stations) and a configuration centered on large urban areas (Figure 5.1). These network configurations are based on the distribution of ICOS, NOAA and other GAW air sampling stations because we assume there is already infrastructure that would facilitate the establishment of equipment for continuous trace measurements and flask samples and that the stations were established on locations with good conditions to observe atmospheric properties. If no air sampling station was available other GAW stations with different types of measurements were taken, e.g. meteorology, UV radiation, deposition, etc. Meanwhile, urban stations were chosen to be around the largest urban areas in Europe. If no GAW station was available, local meteorological or air quality sampling stations were chosen, or, if none of these were available, the roofs of local science and engineering faculties were taken.

In the case of the rural/background network, several calculations shall be performed with varying network density shall be used for the calculations similar to what was done in (Wang et al., 2018). The design principal would be to include stations that collect signals from larger areas in the sparser network configurations, while more regionally or locally influenced stations are included for the denser networks. The main proposed network configurations are:

- The current ICOS network
- Network configuration based on a more evenly distribution in three densities
    - Sparse
    - Medium
    - Dense

Additionally, in the case of the simulations with $^{14}CO_2$, there shall be a configuration with dense measurements of the other tracers and a sparse network of $^{14}CO_2$ measurements due to the higher costs of this type of measurement. In the case of the urban configuration, besides what is shown in Figure 5.1, a dense urban network configuration around the key urban areas within regions of focus of groups involved in CHE shall be considered. These urban areas would be Paris, Berlin, Oslo, London, the Randstand metropolitan area, the Ruhr region, the Rhine-Main region (Frankfurt, Heidelberg, Mainz), the Lille European Metropolitan Area and Zurich. As for the rural network, in the case of the simulations with $^{14}CO_2$, there shall be a configuration with dense measurements of the other tracers and a sparse network of $^{14}CO_2$ measurements. We aim to have common observation vectors for all groups involved in WP4 only varying due to the different spatial and temporal domains considered.

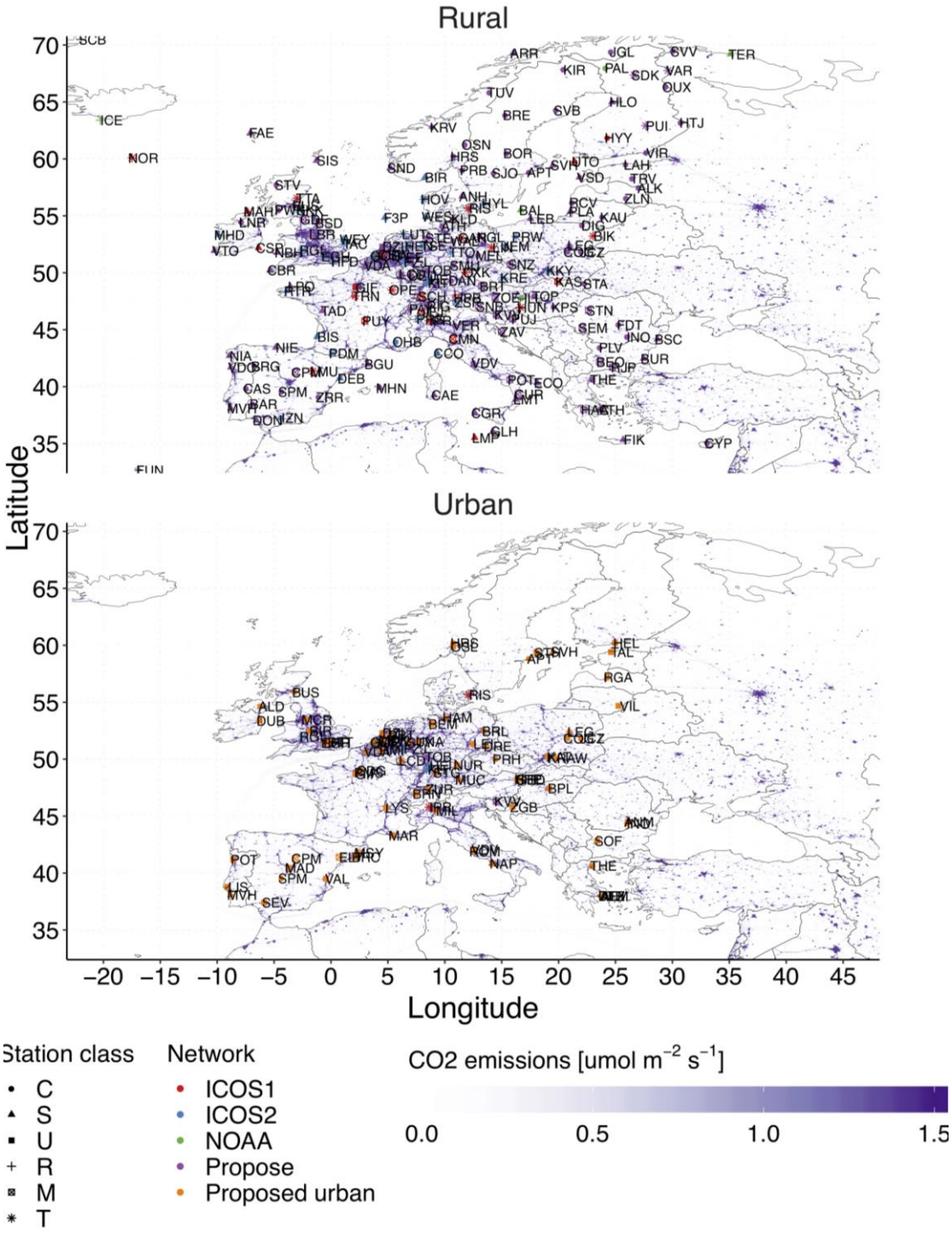

Figure 5.1 Proposed rural and urban network configurations. Station classes are C: Continental/Rural, S: Shore, R: Remote, M: Mountain, T: Tall tower, U: Urban.

5.4 Flux integration regions

While calculations will be performed by a variety of models over different domains, time spans and resolutions, we can agree that common integration regions are required in order to have comparable results. We have created a set of regions at sub-national level. These

regions mostly include defined states, provinces or statistical regions (Figure 5.2). However, in some cases the regions were selected because of transport, climate or land cover characteristics. The set of regions includes the largest urban areas in Europe and the maritime exclusive economic zones of each country (divided by basin, e.g. Mediterranean, Atlantic and English Channel for France). The files are available in an ESRI shapefile format or as netcdf binaries gridded with the resolution and domain extent from the TNO inventory (6 km and 1 km) as well as in the resolution we will use for the STILT runs of 5 and 10 km.

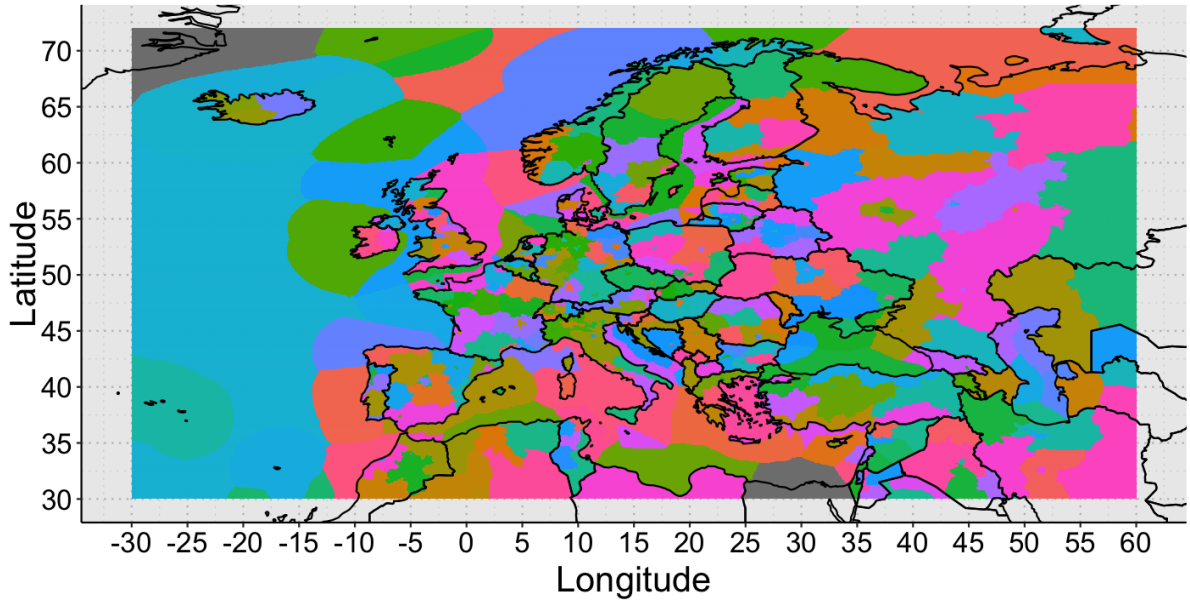

Figure 5.2 Integration regions created over the whole TNO GHGco emission inventory range at 6 km resolution.

These files can be retrieve provisionally from the FTP server ftp://ftp.bgc-jena.mpg.de (username "anonymous" and the password "none") under the path /pub/outgoing/tnunez/. The files have the names: CHE_WP4_integration_regions_20190814_*.nc

- In the files you will find the following variables:
    - country_name
    - a3_un_country_code: 3-letter code per country
    - region_name: name of each integration region
    - dxyp (gridcell area)
    - region (longitude x latitude x object number): region mask for individual regions
    - region_map (longitude x latitude): region with largest coverage per grid box

We selected a few regions, which should serve for comparison. The uncertainty per gridcell for the different emission categories shall be propagated to the scale of these regions and scaled by a global factor so that all groups use the same uncertainty assumptions. The regions of interest shall be chosen with the following criteria: 1) fully contained within the TNO-GHGco 1km x 1km ensemble inventory, 2) of interest for some specific groups, or 3) characteristic for some emission categories. The proposed categories are the following:

1. Paris urban area and/or Ile de France
2. London urban area and/or Greater London

3. Berlin urban area and/or Berlin state region
4. Oslo urban area and/or Oslo region
5. Zurich urban area and/or Zurich region
6. Randstad urban region
7. Lille European Metropolitan Area
8. Ruhr urban region
9. Frankfurt urban region
10. Munich urban region
11. Silesia, Poland
12. Switzerland
13. Netherlands
14. Belgium
15. Luxemburg
16. Germany
17. Czech Republic
18. Austria
19. North sea emissions from Netherlands, Belgium and Germany

5.5 Flux uncertainty assumptions

We set primarily the uncertainty per gridcell of CO_2 emissions based on the spread of different estimates, the mismatch of the priors against observations, expert knowledge, or an educated guess. The uncertainty of emissions of the other tracers is defined relative to the uncertainty of CO_2 emissions and the uncertainty of the respective $^{14}CO_2/CO_2$, O_2/CO_2 and CO/CO_2 ratios. Emission categories, which are not taken into account in the optimization, have no uncertainty, but some groups may choose to do sensitivity testing with different magnitudes of these emissions.

5.5.1 Fossil fuel combustion

5.5.1.1 CO_2

TNO has produced the new TNO-GHGco inventory at 1/60º x 1/120º resolution (approximately 1 km x 1 km) over Central Europe (ranging from 2ºW to 19ºE and from 47ºN to 57ºN) for year 2015. The inventory included ten different emission calculations per emitter for the categories with randomly varying parameters of the emission model for each species. The point emitters from the TNO-GHGco inventory were aggregated onto a 1/120º x 1/60º resolution grid for each process for each fuel type and for each of the ten scenarios (figure A.1). The main goal of providing ten different estimates or scenarios of the CO_2 and CO emissions was to provide an estimate of uncertainty.

We propose to use common domain-wide relative uncertainty (σ/μ) per gridcell to each emission category, which should be the median relative uncertainty (not counting gridcells with no emissions, see table 6). Additionally, each group may decide to do sensitivity tests with a lower and upper bound uncertainty estimate (25^{th} and 75^{th} percentile) according to the table in section 5.6. When aggregating emissions to a different grid, uncertainty propagation should be taken into account.

Table 5-4: Relative uncertainty [%] estimates for fossil fuel combustion CO_2 emissions per emission category.

Source	25th percentile	50th percentile	75th percentile
A: Public power	43.5	86.4	137.0
B: Industry	8.3	18.2	32.0
C: Other stationary combustion	12.2	15.4	19.8
D: Fugitive	25.5	34.1	41.1
E: Solvents	25.0	50.0	51.0
F1: Road transport (gasoline)	10.4	13.3	16.9
F2: Road transport (diesel)	8.1	9.9	12.0
F3: Road transport (LPG)	10.2	13.9	16.5
G: Shipping	7.9	7.9	7.9
H: Aviation	9.8	9.8	9.8
I: Off-road transport	15.2	22.4	34.4
J: Waste	92.0	156.4	321.0
L: Agriculture	14.1	65.3	285.4
Total	8.3	11.3	16.9

5.5.1.2 $\delta^{14}CO_2$

By definition of fossil CO_2 emissions, the uncertainty is zero.

5.5.1.3 CO

The ratio between emitted CO and CO_2 for fossil fuel combustion is related to the fuel type and instantaneous combustion efficiency and is quite uncertain. To represent this, the TNO-GHGco inventory also includes ten scenarios of CO emissions. After calculating the CO/CO_2 ratio per gridcell per emission category, we estimate as well a median, lower bound and upper bound (25th and 75th percentile) overall relative uncertainty (σ/μ) of the CO/CO_2 ratio of each emission category (table 7). The absolute uncertainty of the CO emissions would be the sum in quadrature of the relative uncertainties of the CO_2 emissions and the CO/CO_2 ratio multiplied times the magnitude of the CO emissions.

Table 5-5: Relative uncertainty [%] estimates for fossil fuel combustion CO/CO_2 ratio per emission category.

Source	25th percentile	50th percentile	75th percentile
A: Public power	46.7	58.4	94.8
B: Industry	23.8	38.9	74.5
C: Other stationary combustion	15.0	19.3	27.1

D: Fugitive	15.8	28.6	495.2
E: Solvents			
F1: Road transport (gasoline)	24.2	29.4	36.1
F2: Road transport (diesel)	23.3	30.9	34.1
F3: Road transport (LPG)	34.9	57.6	94.9
G: Shipping			
H: Aviation			
I: Off-road transport	29.7	41.3	78.1
J: Waste	306.2	368.3	788.9
L: Agriculture	67.5	32.2	42.8
Total	16.7	22.9	32.3

5.5.1.4 APO

The global average fossil fuel O_2:CO_2 combustion ratio, α_F, depends not only on the ratio of each fuel type, but also on the relative contribution of each fuel type to the global mix. The overall uncertainty of ±0.04 is dominated by uncertainty in the fuel relative contributions, as related to uncertainty in the production figures for each fuel type (Keeling, 1988; Manning & Keeling, 2006). Non-combustion sources do not consume oxygen and thus their APO uncertainty is zero.

5.5.2 Biofuel combustion

5.5.2.1 CO_2

The same strategy as for fossil fuel combustion is to be followed for biofuel combustion emissions. Table 7 shows domain-wide relative uncertainty estimates for biofuel combustion based of the spread in the ten scenarios of the TNO GHGco 1km x 1km inventory.

Table 5-6: Relative uncertainty [%] estimates for biofuel combustion CO_2 emissions per emission category.

Source	25th percentile	50th percentile	75th percentile
A: Public power	32.2	61.5	109.4
B: Industry	18.3	29.0	59.7
C: Other stationary combustion	17.1	22.1	30.1
D: Fugitive			
E: Solvents			
F1: Road transport (gasoline)	7.0	7.0	7.0
F2: Road transport (diesel)	7.0	7.0	7.0
F3: Road transport (LPG)	7.0	7.0	7.0

G: Shipping			
H: Aviation			
I: Off-road transport	15.2	22.4	34.4
J: Waste			
L: Agriculture	14.1	65.3	285.4

5.5.2.2 $\delta^{14}CO_2$

The model-data mismatch was of the same magnitude of the measurement precision of observed plant samples (3‰) (Bozhinova, 2015).

5.5.2.3 CO

As for fossil fuel inventory, the TNO-GHGco inventory also includes ten scenarios of CO emissions for biofuel emissions. Therefore, we also calculate the CO/CO_2 ratio per gridcell per emission category and obtain a median, lower bound and upper bound (25^{th} and 75^{th} percentile) overall relative uncertainty (table 9).

Table 5-7: Relative uncertainty [%] estimates for biofuel combustion CO/CO_2 ratio per emission category.

Source	25^{th} percentile	50^{th} percentile	75^{th} percentile
A: Public power	22.2	22.2	40.5
B: Industry	23.3	31.3	50.9
C: Other stationary combustion	22.7	37.9	44.8
D: Fugitive			
E: Solvents			
F1: Road transport (gasoline)	7.0	7.0	7.0
F2: Road transport (diesel)	7.0	7.0	7.0
F3: Road transport (LPG)	7.0	7.0	7.0
G: Shipping			
H: Aviation			
I: Off-road transport			
J: Waste			
L: Agriculture	14.1	65.3	285.4
Total			

5.5.2.4 APO

Biofuels have an oxidative ratio that corresponds to that terrestrial biospheric release (van der Laan-Luijkx et al., 2010). Therefore, we assume an oxidative ratio uncertainty of 0.05 (Manning & Keeling, 2006; Severinghaus, 1995).

5.5.3 Biogenic fluxes

5.5.3.1 CO_2

Kountouris et al. (2018) provided domain-wide yearly uncertainty estimates of net ecosystem exchange fluxes for different land cover classes (SYNMAP), which stem from the comparison of gridcell level fluxes with eddy covariance measurements. We propose to follow the same strategy but for partitioned gross primary productivity and ecosystem respiration fluxes separately. Additionally it would be of use to have an uncertainty estimate per month to account for seasonal misrepresentation of the fluxes.

5.5.3.2 $\delta^{14}CO_2$

Bozhinova et al. (2014) suggested that the resulting gradients in $^{14}CO_2$ in plant samples can vary ±3 - 7‰ from the atmospheric samples. We assume this uncertainty for biospheric fluxes.

5.5.3.3 CO

See section 5.5.6.

5.5.3.4 APO

We assume an oxidative ratio uncertainty of 0.05 (Manning & Keeling, 2006; Severinghaus, 1995).

5.5.4 Ocean fluxes

5.5.4.1 CO2

Similarly to the biogenic fluxes, we shall compare the prior fluxes with observations to derive an average relative uncertainty per gridcell for example from SOCAT. Wang (2016) assumed an uncertainty 0.3 g C m^{-2} d^{-1} for ocean fluxes everywhere.

5.5.4.2 $\delta^{14}CO_2$

We assume an uncertainty of $^{14}CO_2$ fluxes per gridcell of 15% based on Sweeney et al., (2007).

5.5.4.3 CO

See section 5.5.6.

5.5.4.4 APO

Similarly to CO_2 fluxes, the uncertainty of O_2 fluxes will be derived from measurements from Pickers (2016).

5.5.5 Biomass burning

We assume a base uncertainty of 26.6% for the fire radiative power estimate based on (Freeborn et al., 2014).

5.5.5.1 CO2

The uncertainty of for the CO_2 emission factors from Andreae & Merlet (2001) and it is added quadrature.

5.5.5.2 $\delta^{14}CO_2$

We assume the same uncertainty as for biospheric fluxes.

5.5.5.3 CO

The uncertainty of the CO/CO_2 ratio is obtained from emission factors from Andreae & Merlet (2001) and is added quadrature.

5.5.5.4 APO

We assume an oxidative ratio uncertainty of 0.05 (Manning & Keeling, 2006; Severinghaus, 1995).

5.5.6 Other processes

Other processes, e.g. for example nuclear power plants, cosmogenic or photochemical production (if represented) or soil processes for CO must be given an uncertainty in order to

have an impact on the uncertainty reduction calculation. Hence we shall assume 50% per gridbox.

5.5.7 Summary

To be comparable, a list of the yearly and monthly uncertainties per tracer for the regions listed in section 5.4 will be distributed so that each group can scale their uncertainty estimates to match the list.

Table 5-8: Summary of uncertainty assumptions. Values in % (also in referenced tables or sections) are with respect to the *a priori* emission magnitude per gridcell.

Process	CO_2	$\delta^{14}CO_2$ [‰]*	CO/CO_2	O_2/CO_2
Fossil fuel combustion	**Table 5-4**	0.00	**Table 5-5**	0.04
Biofuel combustion	**Table 5-6**	3.00	**Table 5-7**	0.05
Non-combustion fossil	Industry in **Table 5-4**	0.00	0.00	0.00
Gross primary productivity	See section 5.5.3.1	4.50	50 %	0.05
Respiration	See section 5.5.3.1	4.50	50 %	0.05
Ocean	0.3 g C m^{-2} d^{-1}	15 %	50 %	20 %
Biomass burning	26% + σ_{EF}**	4.50	σ_{EF}**	0.05
Nuclear power plants	0.00	50 %	0.00	0.00
Photochemical***	50 %	0.00	50 %	50 %
Cosmogenic***	0.00	50 %	0.00	0.00
Initial/Boundary conditions***	50 %	50 %	50 %	50 %
Nitrogen fluxes***	50 %	50 %	50 %	50 %

*: Values without unit are the absolute uncertainty of the $\delta^{14}CO_2$ signal of the process.

**: σ_{EF} is the uncertainty in the emission factors from Andreae & Merlet (2001)

***: optional component

5.6 Simulation protocol

To obtain a good characterization of the ability of different networks to estimate fossil fuel CO_2 emissions, we propose the following simulation protocol, independently of the spatial and temporal domain of focus.

Table 5-9 Proposed simulation protocol.

Tracers	Station network	Anthropogenic emission uncertainty	Biospheric flux uncertainty
CO_2	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$	Current ICOS/$^{14}CO_2$ sparse	Median	Standard
CO_2+CO	Current ICOS	Median	Standard
CO_2+APO	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$+CO	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$+CO	Current ICOS/$^{14}CO_2$ sparse	Median	Standard
CO_2+$^{14}CO_2$+APO	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$+APO	Current ICOS/$^{14}CO_2$ sparse	Median	Standard
CO_2+CO+APO	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Current ICOS	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Current ICOS/$^{14}CO_2$ sparse	Median	Standard
CO_2	Sparse	Median	Standard
CO_2+$^{14}CO_2$	Sparse	Median	Standard
CO_2+CO	Sparse	Median	Standard
CO_2+APO	Sparse	Median	Standard
CO_2+$^{14}CO_2$+CO	Sparse	Median	Standard
CO_2+$^{14}CO_2$+APO	Sparse	Median	Standard
CO_2+CO+APO	Sparse	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Sparse	Median	Standard
CO_2	Medium	Median	Standard
CO_2+$^{14}CO_2$	Medium	Median	Standard
CO_2+CO	Medium	Median	Standard
CO_2+APO	Medium	Median	Standard
CO_2+$^{14}CO_2$+CO	Medium	Median	Standard
CO_2+$^{14}CO_2$+APO	Medium	Median	Standard
CO_2+CO+APO	Medium	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Medium	Median	Standard
CO_2	Dense	Median	Standard
CO_2+$^{14}CO_2$	Dense	Median	Standard
CO_2+$^{14}CO_2$	Dense/$^{14}CO_2$ sparse	Median	Standard
CO_2+CO	Dense	Median	Standard

Tracers	Station network	Anthropogenic emission uncertainty	Biospheric flux uncertainty
CO_2+APO	Dense	Median	Standard
CO_2+$^{14}CO_2$+CO	Dense	Median	Standard
CO_2+$^{14}CO_2$+CO	Dense/$^{14}CO_2$ sparse	Median	Standard
CO_2+$^{14}CO_2$+APO	Dense	Median	Standard
CO_2+$^{14}CO_2$+APO	Dense/$^{14}CO_2$ sparse	Median	Standard
CO_2+CO+APO	Dense	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Dense	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Dense/$^{14}CO_2$ sparse	Median	Standard
CO_2	Dense	High	Standard
CO_2+$^{14}CO_2$	Dense/$^{14}CO_2$ sparse	High	Standard
CO_2+$^{14}CO_2$+CO	Dense/$^{14}CO_2$ sparse	High	Standard
CO_2+CO	Dense	High	Standard
CO_2+APO	Dense	High	Standard
CO_2+$^{14}CO_2$+CO+APO	Dense/$^{14}CO_2$ sparse	High	Standard
CO_2	Dense	Low	Standard
CO_2+$^{14}CO_2$	Dense/$^{14}CO_2$ sparse	Low	Standard
CO_2+$^{14}CO_2$+CO	Dense/$^{14}CO_2$ sparse	Low	Standard
CO_2+CO	Dense	Low	Standard
CO_2+APO	Dense	Low	Standard
CO_2+$^{14}CO_2$+CO+APO	Dense/$^{14}CO_2$ sparse	Low	Standard
CO_2	Dense	Median	Triple
CO_2+$^{14}CO_2$	Dense/$^{14}CO_2$ sparse	Median	Triple
CO_2+$^{14}CO_2$+CO	Dense/$^{14}CO_2$ sparse	Median	Triple
CO_2+CO	Dense	Median	Triple
CO_2+APO	Dense	Median	Triple
CO_2+$^{14}CO_2$+CO+APO	Dense/$^{14}CO_2$ sparse	Median	Triple
CO_2	Dense	High	Triple
CO_2+$^{14}CO_2$	Dense/$^{14}CO_2$ sparse	High	Triple
CO_2+$^{14}CO_2$+CO	Dense/$^{14}CO_2$ sparse	High	Triple
CO_2+CO	Dense	High	Triple
CO_2+APO	Dense	High	Triple
CO_2+$^{14}CO_2$+CO+APO	Dense/$^{14}CO_2$ sparse	High	Triple

Tracers	Station network	Anthropogenic emission uncertainty	Biospheric flux uncertainty
CO_2	Urban hotspots	Median	Standard
CO_2+$^{14}CO_2$	Urban hotspots	Median	Standard
CO_2+CO	Urban hotspots	Median	Standard
CO_2+APO	Urban hotspots	Median	Standard
CO_2+$^{14}CO_2$+CO	Urban hotspots	Median	Standard
CO_2+$^{14}CO_2$+APO	Urban hotspots	Median	Standard
CO_2+CO+APO	Urban hotspots	Median	Standard
CO_2+$^{14}CO_2$+CO+APO	Urban hotspots	Median	Standard

6 References

Andela, N., Morton, D. C., Giglio, L., Paugam, R., Chen, Y., Hantson, S., et al. (2019). The Global Fire Atlas of individual fire size, duration, speed and direction. *Earths System Science Data*, *11*, 529–552. http://doi.org/10.5194/essd-11-529-2019

Andreae, M. O., & Merlet, P. (2001). Emission of trace gases and aerosols from biomass burning. *Global Biogeochemical Cycles*, *15*(4), 955–966. http://doi.org/10.1029/2000GB001382

Bozhinova, D. N. (2015). *Interpreting plant-sampled $\Delta^{14}CO_2$ to study regional anthropogenic CO2 signal in Europe*. Wageningen University, Wageningen, NL.

Bozhinova, D., van der Molen, M. K., van der Velde, I. R., Krol, M. C., van der Laan, S., Meijer, H. A. J., & Peters, W. (2014). Simulating the integrated summertime Δ14CO2 signature from anthropogenic emissions over Western Europe. *Atmospheric Chemistry and Physics*, *14*(14), 7273–7290. http://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2

Buitenhuis, E. T., Rivkin, R. B., Sailley, S., & Le Quéré, C. (2010). Biogeochemical fluxes through microzooplankton. *Global Biogeochemical Cycles*, *24*(4), n/a–n/a. http://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2

Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., Altaf Arain, M., et al. (2012). What eddy-covariance measurements tell us about prior land flux errors in CO 2-flux inversion schemes. *Global Biogeochemical Cycles*, *26*(1), n/a–n/a. http://doi.org/10.1029/2006JD008371

Ciais, P., Crisp, D., Denier van der Gon, H., Engelen, R., Janssens-Maenhout, G., Heimann, M., et al. (2015). *Towards a European Operational Observing System to Monitor Fossil CO2 emissions* (pp. 1–68). Brussels, Belgium: European Commission.

Clay, G. D., & Worrall, F. (2015). Oxidative ratio (OR) of Southern African soils and vegetation: Updating the global OR estimate. *Catena*, *126*(C), 126–133. http://doi.org/10.1016/j.catena.2014.10.029

Conrad, R., & Seiler, W. (1980). Role of Microorganisms in the Consumption and Production of Atmospheric Carbon Monoxide by Soil. *Applied and Environmental Microbiology*, *40*(3), 437–445.

Conte, L., Szopa, S., Séférian, R., & Bopp, L. (2019). The oceanic cycle of carbon monoxide and its emissions to the atmosphere. *Biogeosciences*, *16*(4), 881–902. http://doi.org/10.5194/bg-16-881-2019-supplement

Crotwell, A. M., & Steinbacher, M. (2018). *19th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Measurement Techniques (GGMT-2017)* (No. GAW 242) (pp. 1–150). Dûbendorf, Switzerland: World Meteorological Organization Global Atmosphere Watch.

Crutzen, P. J., & Andreae, M. O. (1990). Biomass Burning in the Tropics - Impact on Atmospheric Chemistry and Biogeochemical Cycles. *Science*, *250*(4988), 1669–1678. http://doi.org/10.1126/science.250.4988.1669

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol SOC*, *137*(656), 553–597. http://doi.org/10.1002/qj.828

Degens, E. T. (1969). Biogeochemistryof Stable Carbon Isotopes. In G. Eglinton & M. T. J. Murphy (Eds.), *Organic Geochemistry* (pp. 1–26). Springer, Berlin, Heidelberg. http://doi.org/10.1007/978-3-642-87734-6_14

Farquhar, G. D., Ehleringer, J. R., & Hubick, K. T. (1989). Carbon isotope descrimination and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, *40*, 503–537.

Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., et al. (2009). Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. *Journal of Geophysical Research*, *114*(D4), 1716. http://doi.org/10.1029/2000JD900309

Freeborn, P. H., Wooster, M. J., Roy, D. P., & Chochrane, M. A. (2014). Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite-based active fire characterization and biomass burning estimation. *Geophysical Research Letters*, *41*, 1988–1994. http://doi.org/10.1002/(ISSN)1944-8007

Gamnitzer, U., Karstens, U., Kromer, B., Neubert, R. E. M., Meijer, H. A. J., Schroeder, H., & Levin, I. (2006). Carbon monoxide: A quantitative tracer for fossil fuel CO2? *Journal of Geophysical Research*, *111*(D22), 15915. http://doi.org/10.3402/tellusb.v48i4.15934

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., et al. (2003). Toward constraining regional-scale fluxes of CO 2with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research*, *108*(D24), n/a–n/a. http://doi.org/10.1034/j.1600-0889.2003.00030.x

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., et al. (2019). *The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version)* (pp. 1–54). Copernicus Atmosphere Monitoring Service.

Graven, H. D., & Gruber, N. (2011). Continental-scale enrichment of atmospheric $^{14}CO_2$ from the nuclear power industry: potential impact on the estimation of fossil fuel-derived CO_2. *Atmospheric Chemistry and Physics*, *11*(23), 12339–12349. http://doi.org/10.5194/acp-11-12339-2011-supplement

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. http://doi.org/10.5194/gmd-5-1471-2012

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics*, *6*(1), 3181–3210.

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., et al. (2018). ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation. *Geoscientific Model Development*, *11*(1), 121–163. http://doi.org/10.5194/gmd-11-121-2018-supplement

Heimann, M. (2007). Atmospheric Potential Oxygen (APO) - Distribution, Inversions and Data Needs (pp. 1–22).

Heimann, M., & Körner, S. (2003). *The Global Atmospheric Tracer Model TM3 Model Description and User's Manual Release 3.8a* (No. 5). *Max-Planck-Institut für Biogeochemie, Jena* (pp. 1–131). Max Planck Institute for Biogeochemistry.

Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., & Heimann, M. (1999). Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research*, *104*(D21), 26137–26160.

Hsueh, D. Y., Krakauer, N. Y., Randerson, J. T., Xu, X., Trumbore, S. E., & Southon, J. R. (2007). Regional patterns of radiocarbon and fossil fuel-derived CO_2 in surface air across North America. *Geophysical Research Letters*, *34*(2), 419. http://doi.org/10.1029/2005GL024213

Jones, S. D., Le Quéré, C., & Rödenbeck, C. (2012). Autocorrelation characteristics of surface ocean pCO2 and air-sea CO2 fluxes. *Global Biogeochemical Cycles*, *26*(2), n/a–n/a. http://doi.org/10.1016/j.csr.2009.11.009

Keeling, R. F. (1988, January). *Development of an Interferometric Oxygen Analyzer for Precise Measurement of the Atmospheric O_2 Mole Fraction*. Harvard University, Cambridge, Massachusetts.

Keeling, R. F. Manning, A. C, McEvoy, E. M., Shertz, S. R., (1998), Methods for measuring changes in atmospheric O_2 concentration and their application in southern hemisphere air. *Journal of Geophysical Research* 103(D3), 3381 - 3397

Khatiwala, S., Graven, H., Payne, S., & Heimbach, P. (2018). Changes to the Air-Sea Flux and Distribution of Radiocarbon in the Ocean Over the 21st Century. *Geophysical Research Letters*, *45*(11), 5617–5626. http://doi.org/10.1016/j.physd.2006.09.040

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., & Heimann, M. (2018). Atmospheric CO_2 inversions on the mesoscale using data-driven prior uncertainties: quantification of the European terrestrial CO₂ fluxes. *Atmospheric Chemistry and Physics*, *18*(4), 3047–3064. http://doi.org/10.5194/gmd-8-805-2015

Kountouris, P., Gerbig, C., Totsche, K. U., Dolman, A. J., Meesters, A. G. C. A., Broquet, G., et al. (2015). An objective prior error quantification for regional atmospheric inverse applications. *Biogeosciences*, *12*(24), 7403–7421. http://doi.org/10.1111/j.1469-8137.2010.03301.x

Levin, I., & Karstens, U. (2007). Inferring high-resolution fossil fuel CO2 records at continental sites from combined $^{14}CO_2$ and CO observations. *Tellus B*, *59*(2), 245–250. http://doi.org/10.1029/2005GL024213

Lin, J. C. (2003). A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research*, *108*(D16), ACH 2–1–ACH 2–17. http://doi.org/10.1029/2002JD003161

Liu, L., Zhuang, Q., Zhu, Q., Liu, S., van Asperen, H., & Pihlatie, M. (2018). Global soil consumption of atmospheric carbon monoxide: an analysis using a process-based biogeochemistry model. *Atmospheric Chemistry and Physics*, *18*(11), 7913–7931. http://doi.org/10.1002/gbc.20057

Liu, Y., & Sander, S. P. (2015). Rate Constant for the OH + CO Reaction at Low Temperatures. *The Journal of Physical Chemistry A*, *119*(39), 10060–10066. http://doi.org/10.1021/acs.jpca.5b07220

Lueker, T. J., Keeling, R. F., & Dubey, M. K. (2001). The Oxygen to Carbon Dioxide Ratios observed in Emissions from a Wildfire in Northern California. *Geophysical Research Letters*, *28*(12), 2413–2416. http://doi.org/10.1029/2000GL011860

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., et al. (2008). A satellite-based biosphere parameterization for net ecosystem CO 2exchange: Vegetation Photosynthesis and Respiration Model (VPRM). *Global Biogeochemical Cycles*, *22*(2). http://doi.org/10.1029/2006GB002735

Manning, A. C., & Keeling, R. F. (2006). Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network. *Tellus*, *58B*(2), 95–116.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., & Freitas, S. (2010). Coupled weather research and forecasting–stochastic time-inverted lagrangian

transport (WRF–STILT) model. *Meteorology and Atmospheric Physics*, *107*(1-2), 51–64. http://doi.org/10.1029/2008JD011671

Olivier, J. G. J., Van Aardenne, J. A., Dentener, F. J., Pagliari, V., Ganzeveld, L. N., & Peters, J. A. H. W. (2005). Recent trends in global greenhouse gas emissions: regional trends 1970-2000 and spatial distributionof key sources in 2000. *Journal of Environmental Sciences*, *2*(2-3), 81–99. http://doi.org/10.1080/15693430500400345

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., et al. (2011). TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere. *Atmospheric Chemistry and Physics*, *11*(24), 12813–12837. http://doi.org/10.5194/acp-11-12813-2011

Pickers, P. (2016, November 11). *New applications of continuous atmospheric O_2 measurements: meridional transects across the Atlantic Ocean, and improved quantification of fossil fuel-derived CO_2*. University of East Anglia.

Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B., & Heimann, M. (2012). Comparing Lagrangian and Eulerian models for CO₂ transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM. *Atmospheric Chemistry and Physics*, *12*(19), 8979–8991. http://doi.org/10.1029/2006JD007649

Portmann, F. T., Siebert, S., & Doell, P. (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*, –. http://doi.org/10.1029/2008GB003435

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). Numerical Recipes (Third Edition. New York). Cambridge University Press.

Rödenbeck, C. (2005). Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport. *Max-Planck-Institut Fur Biogeochemie*, *TECHNICAL REPORTS 6*, 1–61.

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., et al. (2015). Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM). *Biogeosciences*, *12*(23), 7251–7278. http://doi.org/10.5670/oceanog.2009.41

Rödenbeck, C., Gerbig, C., Trusilova, K., & Heimann, M. (2009). A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models. *Atmospheric Chemistry and Physics*, *9*, 5331–5342.

Rödenbeck, C., Houweling, S., Gloor, M., & Heimann, M. (2003). CO2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics*, *3*, 2003.

Rödenbeck, C., Le Quéré, C., Heimann, M., & Keeling, R. F. (2008). Interannual variability in oceanic biogeochemical processes inferred by inversion of atmospheric O_2/N_2 and CO_2 data. *Tellus B*, *60B*(5), 685–705. http://doi.org/10.2172/477735

Severinghaus, J. P. (1995). *Studies of the Terrestrial O2 and Carbon Cycles in Sand Dune Gases aind in Biosphere 2*. Columbia University.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., et al. (2014). Global dataset of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics*, *14*, 9317–9341. http://doi.org/10.5194/acp-14-9317-2014

Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., et al. (2014). On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations. *Atmospheric Chemistry and Physics*, *14*(17), 9295–9316. http://doi.org/10.5194/acp-14-9295-2014

Steinbach, J. (2010, September 7). *Enhancing the usability of atmospheric oxygen measurements through emission source characterization and airborne measurements*. Friedrich-Schiller-University Jena.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., & Mukai, H. (2011).

The CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios. *Atmospheric Chemistry and Physics*, *11*(14), 6855–6870. http://doi.org/10.5194/acp-11-6855-2011-supplement

Stephens, B. B., Keeling, R. F., Heimann, M., Six, K. D., Murnane, R., & Caldeira, K. (1998). Testing global ocean carbon cycle models using measurements of atmospheric O_2 and CO_2 concentration. *Global Biogeochemical Cycles*, *12*(2), 213–230.

Stuiver, M., & Polach, H. A. (1977). Discussion: Reporting of ^{14}C Data. *Radiocarbon*, *19*(3), 355–363.

Suess, H. E. (1955). Radiocarbon Concentration in Modern Wood. *Science*, *122*(3166), 415–417. http://doi.org/10.1126/science.122.3166.415-a

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G. A., Sarmiento, J. L., & Wanninkhof, R. (2007). Constraining global air-sea gas exchange for CO2 with recent bomb ^{14}C measurements. *Global Biogeochemical Cycles*, *21*(2), GB2015. http://doi.org/10.1029/2006GB002784

Tans, P. P., De Jong, A. F. M., & Dook, W. G. (1979). Natural atmospheric 14C variation and the Suess effect. *Nature*, *280*, 826–828. http://doi.org/10.1038/280826a0

Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation: SIAM. Society for Industrial and Applied Mathematics.

Tarr, M. A., Miller, W. L., & Zepp, R. G. (1995). Direct carbon monoxide photoproduction from plant matter. *Journal of Geophysical Research*, *100*(D6), 11403–11413.

Trusilova, K., Roedenbeck, C., Gerbig, C., & Heimann, M. (2010). Technical Note: A new coupled system for global-to-regional downscaling of CO2 concentration estimation. *Atmospheric Chemistry and Physics*, *10*(7), 3205–3213.

Turnbull, J., Rayner, P., Miller, J. B., Naegler, T., Ciais, P., & Cozic, A. (2009). On the use of 14CO 2as a tracer for fossil fuel CO 2: Quantifying uncertainties using an atmospheric transport model. *Journal of Geophysical Research*, *114*(D22), GB1002. http://doi.org/10.3402/tellusb.v48i4.15934

van der Laan-Luijkx, I. T., Karstens, U., Steinbach, J., Gerbig, C., Sirignano, C., Neubert, R. E. M., et al. (2010). CO_2, $\delta O_2/N_2$ and APO: observations from the Lutjewad, Mace Head and F3 platform flask sampling network. *Atmospheric Chemistry and Physics*, *10*(21), 10691–10704. http://doi.org/10.5194/acp-10-10691-2010

Wang, Y. (2016). *The potential of observations of radiocarbon in atmospheric CO2 for the atmospheric inversion of fossil fuel CO2 emission at regional scale*. Université Paris-Saclay, Paris.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., et al. (2018). Potential of European $^{14}CO_2$ observation network to estimate the fossil fuel CO_2 emissions via atmospheric inversions. *Atmospheric Chemistry and Physics*, *18*(6), 4229–4250. http://doi.org/10.5194/acp-18-4229-2018-supplement

Worrall, F., Clay, G. D., & Macdonald, A. (2015). The impact of fertilizer management on the oxidation status of terrestrial organic matter. *Soil Use and Management*, *32*(1), 45–52. http://doi.org/10.1007/s10533-013-9877-6

Zazzeri, G., Acuña Yeomans, E., & Graven, H. D. (2018). Global and regional emissions of radiocarbon from nuclear power plants from 1972 to 2016. *Radiocarbon*, *60*(4), 1067–1081. http://doi.org/10.1017/RDC.2018.42

Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., et al. (2019). Global atmospheric carbon monoxide budget 2000-2017 inferred from multi-species atmospheric inversions. *Earth System Science Data Discussions*, 1–42. http://doi.org/10.5194/essd-2019-61-supplement

A. Analysis of 1-km emission inventory ensemble for Central Europe

TNO has produced the new TNO-GHGco inventory at 1/60º x 1/120º resolution (approximately 1 km x 1 km) over Central Europe (ranging from 2ºW to 19ºE and from 47ºN to 57ºN) for year 2015. This inventory contains both a list of point emitters for CO_2 and CO as well as gridcell values for area sources, accompanied by their grid index, process (G-NFR categories), country, and by fuel type (fossil fuel or biofuel). The inventory included ten different emission calculations per point source or gridcell with randomly varying parameters of the emission model for each species (Monte Carlo approach). These data can be downloaded from the TNO ftp site, with access information available from Hugo Denier van der Gon (hugo.deniervandergon@tno.nl) upon request.

A.1. Emissions

For the year 2015, the TNO-GHG-co (1 km x 1km) inventory contains 1.92 ± 0.01 and 0.26 ± 0.01 Pg CO_2 a^{-1} emissions from fossil fuels and biofuels combustion respectively (where $\mu \pm \sigma$, are the mean and standard deviation for the emissions over the entire central European domain). Furthermore, the inventory contains 6.7 ± 0.31 and 2.49 ± 0.26 Tg CO a^{-1} emissions from fossil fuels and biofuels respectively. For CO_2, the largest fossil fuel emissions came from the power generation, industry, stationary combustion and road transport (diesel and gasoline) sectors with averages of 0.594 ± 0.008, 0.457 ± 0.002, 0.32 ± 0.003, 0.299 ± 0.005 and 0.126 ± 0.002 Pg CO_2 a^{-1} respectively (Figure A.1). Meanwhile, the lowest emissions were from waste, solvents, road transport (LPG) and agriculture, with averages of 0.002, 0.005, 0.005 and 0.01 Pg CO_2 a^{-1} respectively. For biofuel combustion, the largest emissions emission came from power generation, stationary combustion, industry, and road transport (diesel and gasoline) with averages of 0.103 ± 0.003, 0.088 ± 0.006, 0.046, 0.015 and 0.005 respectively. No biofuel emissions were provided for solvents, road transport (LPG), shipping, aviation and waste. In the case of CO, the largest emissions from fossil fuels were found for industry, road transportation (gasoline), off-road transportation, stationary combustion, road transportation (diesel) and power generation with averages of 2.591 ± 0.05, 1.738 ± 0.313, 0.792 ± 0.105, 0.788 ± 0.062, 0.323 ± 0.043 and 0.212 ± 0.003 Tg CO a^{-1} respectively. The agricultural sector was a small sink of CO (-0.001 ± 0.02 Tg CO a^{-1}). With respect to biofuels, the largest CO emissions were from stationary combustion, industry, agriculture, road transportation (gasoline), power generation and road transportation (diesel) with average emissions of 2.199 ± 0.267, 0.114 ± 0.002, 0.071, 0.061, 0.028 ± 0.002 and 0.16 Tg CO a^{-1} respectively.

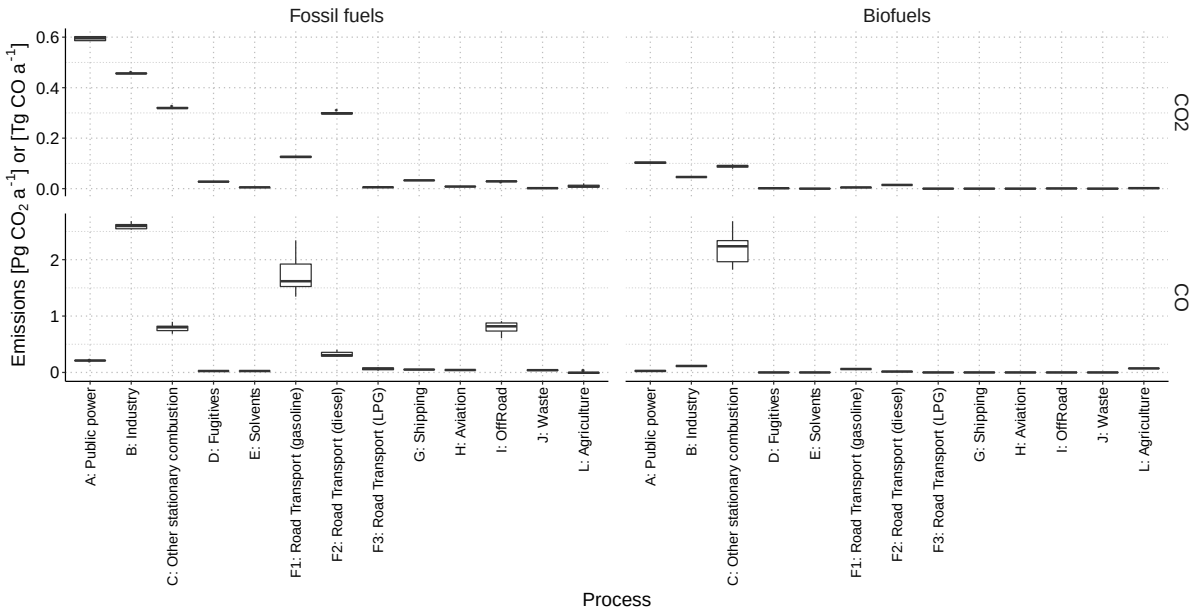

Figure A.1 Box plots showing the spread of the total emissions per for the ten emission scenarios.

The emissions from the TNO-GHGco inventory were aggregated onto a 1/120º x 1/60º resolution grid for each process for each fuel type and for each of the ten scenarios (Figure A.2). Clear spatial distribution differences emerge between CO_2 and CO with more CO_2 emissions at power stations, urban areas and roads and relative low emissions from biofuels. In contrast, CO emissions are relatively higher over roads, rural areas, and it is particularly strong for biofuel emissions.

A.2. Emission uncertainty

The main goal of providing ten different estimates or scenarios of the CO_2 and CO emissions was to provide an estimate of uncertainty. By gridding each scenario to the 1/60º x 1/120º resolution, it was possible to take the standard deviation per gridcell per process to obtain a measure of the uncertainty. While the spatial distribution of the absolute uncertainty reflected the spatial distribution of the emissions, the relative uncertainty (σ/μ) revealed that the parameters resulting in the variability between the scenarios are often country dependent (Figure A.3).

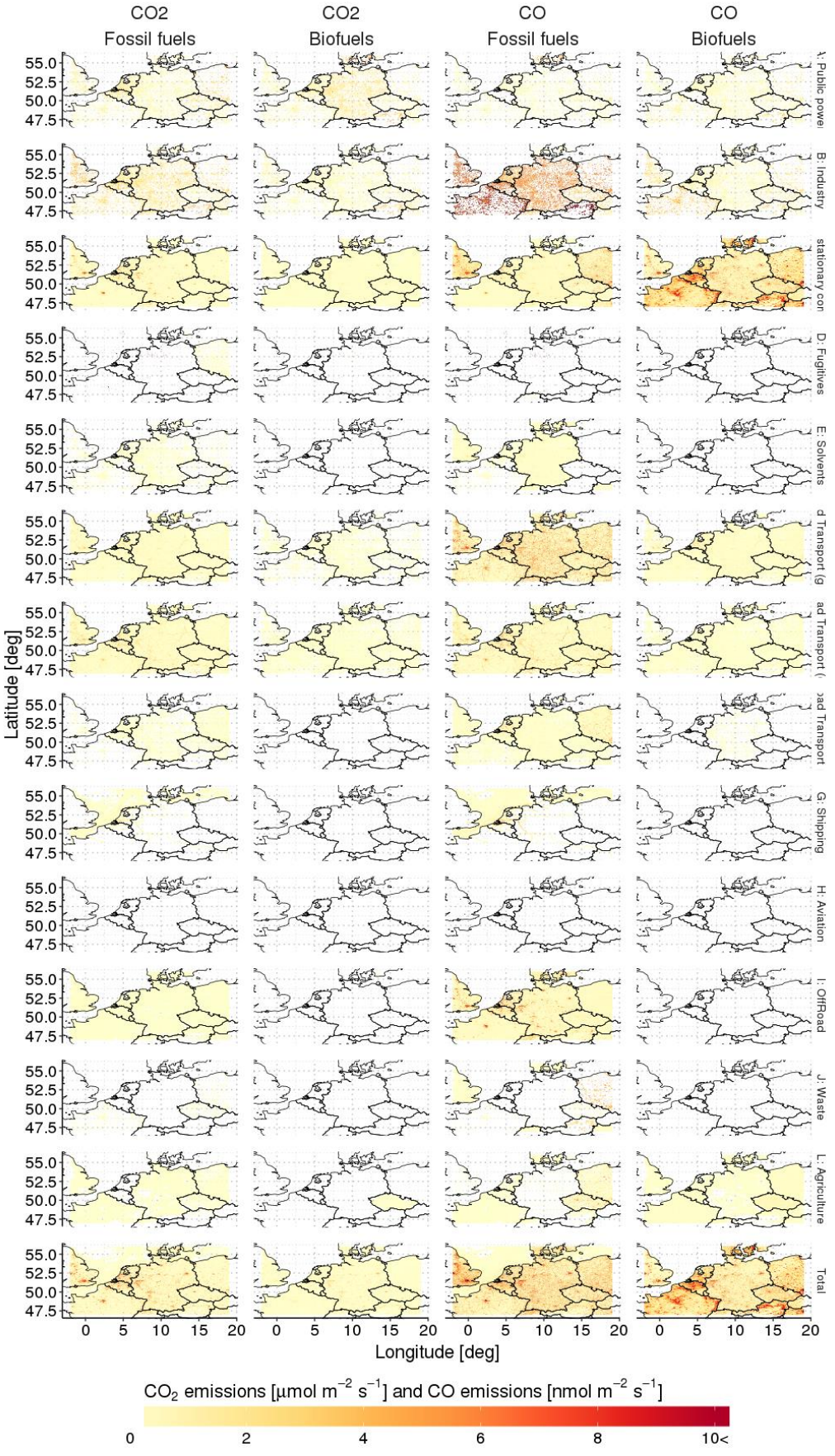

Figure A.2 Mean CO_2 and CO emissions per gridcell for each fossil fuel and biofuel combustion sector and total emissions per gridcell (bottom row).

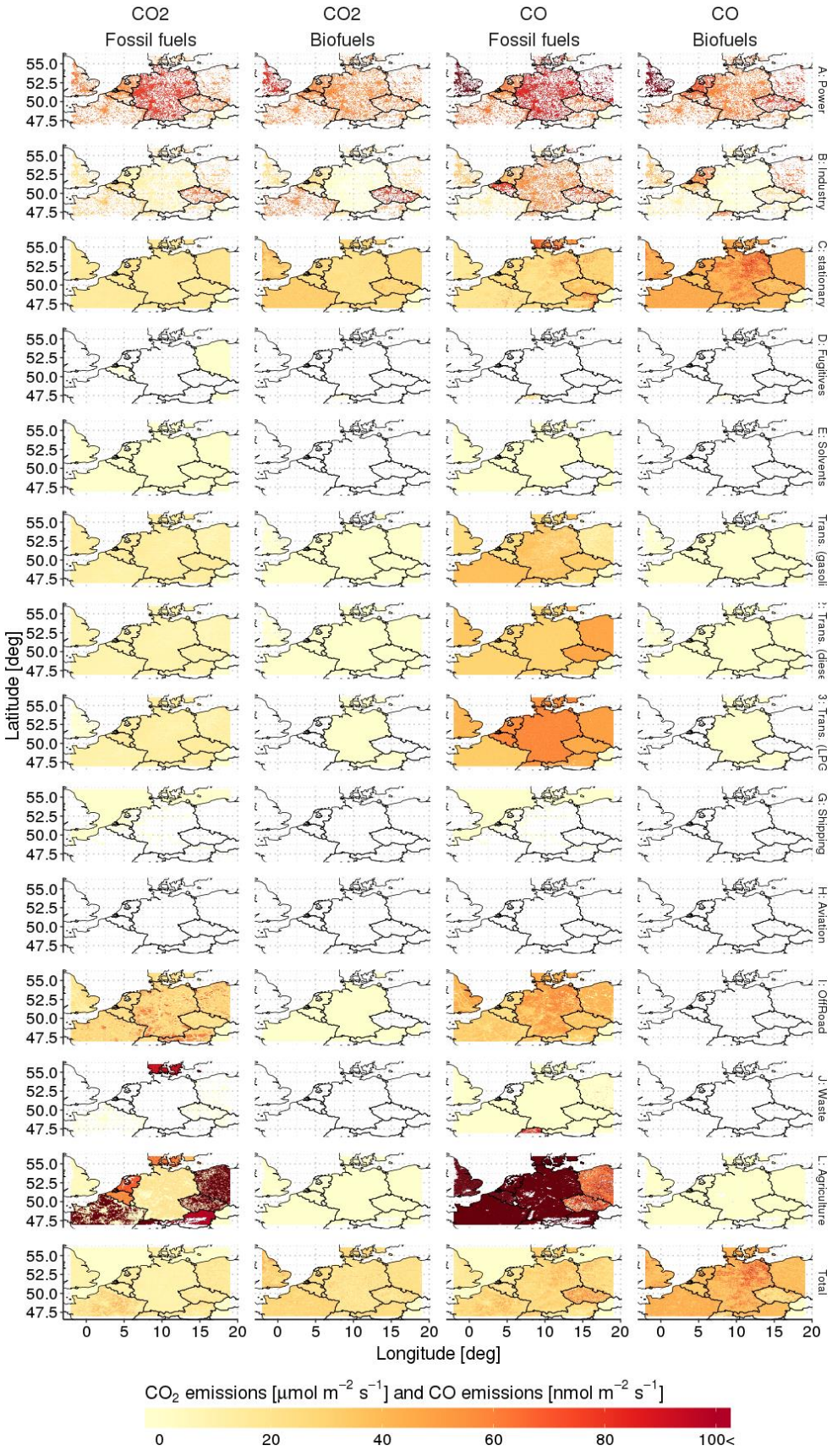

Figure A.3 Relative uncertainty (σ/µ, %) of CO_2 and CO emissions per gridcell for each fossil fuel and biofuel combustion sector and total relative uncertainty (bottom row).

A.3. CO/CO_2 ratios

For the setup of the a priori covariance matrix, it is of interest to know the uncertainty how the CO/CO_2 ratio varies between the different emission processes because this depends on the fuel type and the instantaneous combustion efficiency, and thus it is highly variable and uncertain. The overall average CO/CO_2 ratio (molar) of the total emissions was 0.0054 ± 0.0003 and 0.0148 ± 0.0014 for fossil fuels and biofuels respectively. However, we were able to distinguish typical CO/CO_2 ratios by process (Figure A.4 and Figure A.5). For fossil fuels, the largest CO/CO_2 ratios were found for off-road transportation, waste burning, road transportation (gasoline and LPG) and industry with averages of 0.0437 ± 0.0057, 0.029 ± 0.006, 0.0217 ± 0.004, 0.0187 ± 0.0059, and 0.0086 ± 0.0002 respectively. In contrast, emissions from power generation, fugitives and road transport (diesel) had particularly low CO/CO_2 ratios: 0.0006, 0.0015 and 0.0017 ± 0.0002. The agricultural sector had small negative CO/CO_2 ratios because of the overall negative CO emissions, and had a very large relative uncertainty of 1075%. For biofuels, the sectors with the largest CO/CO_2 ratios were agriculture, stationary combustion, road transport (gasoline and LPG), and industry with averages of 0.0608, 0.0391 ± 0.0036, 0.202, 0.0071 and 0.0035 ± 0.0001. Again, diesel motor vehicles had a much lower CO/CO_2 ratio than their gasoline counterparts, 0.0017. Power generation also had a very small CO/CO_2 ratio of 0.0004. No CO emissions from biofuels were given for fugitives and off-road transportation.

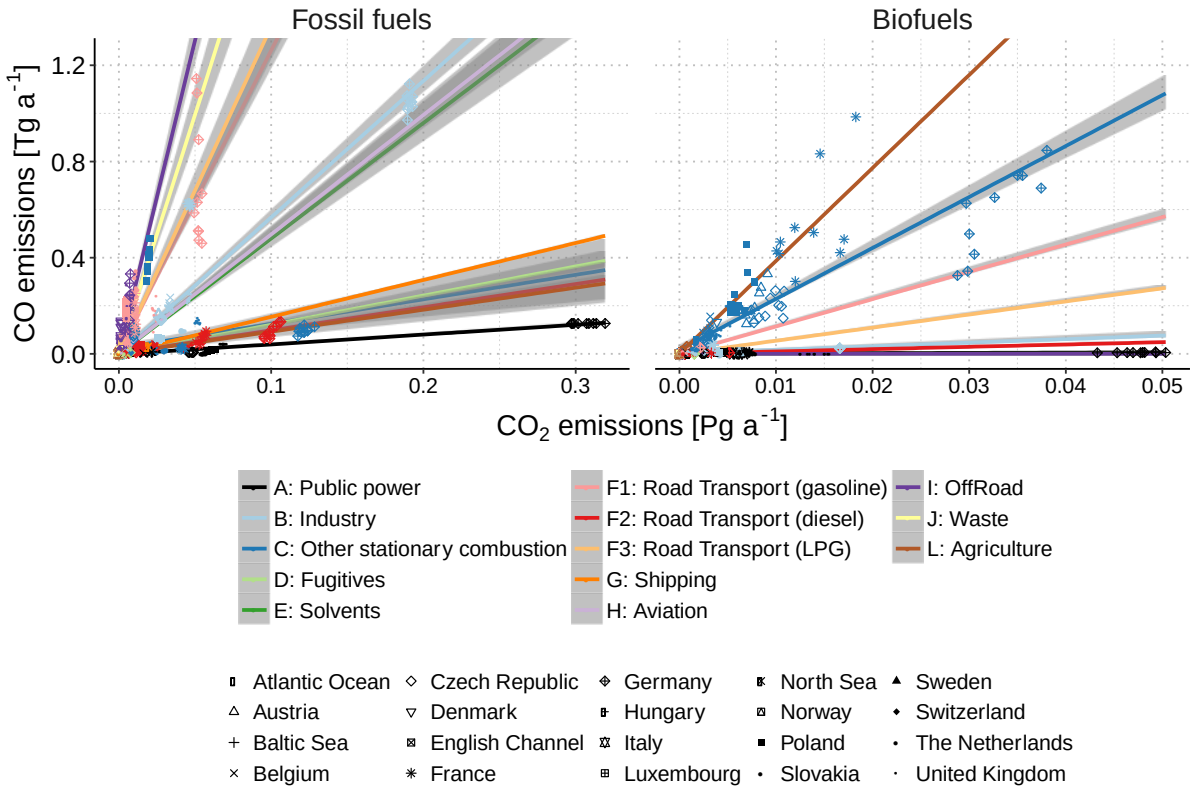

Figure A.4 CO_2 against CO emissions per process per country for fossil fuels (left) and biofuels (right).

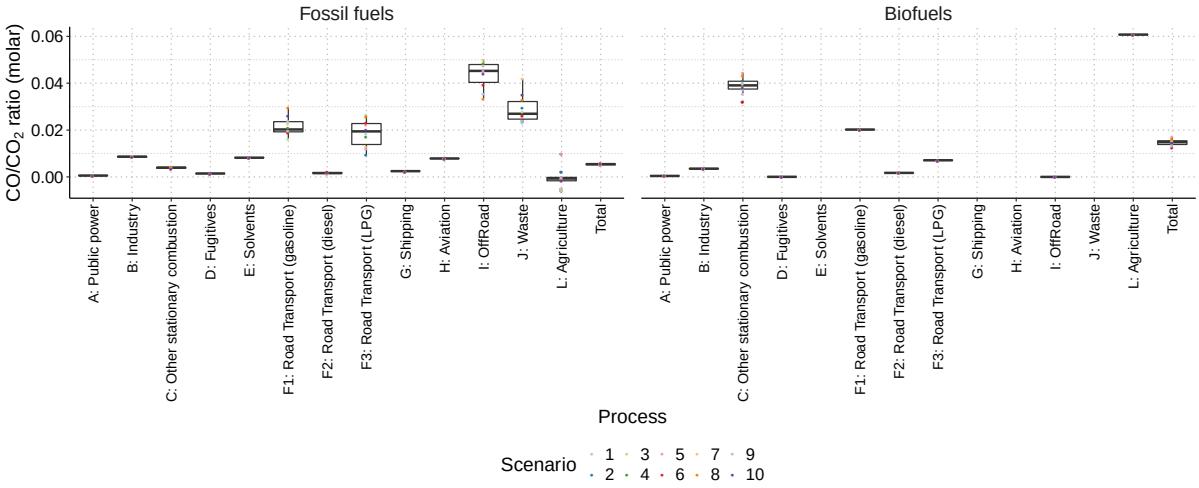

Figure A.5 Spread of average CO/CO_2 ratios (over entire central European domain) per process for the different emission scenarios.

The main driver of variations in the CO/CO_2 ratio is the combustion efficiency. In this case we defined the combustion efficiency, *CE*, as

$$CE = \frac{CO_2}{CO_2 + CO}$$

Under complete combustion (all fuel is converted to CO_2), the combustion efficiency is 100%. For fossil fuels, the most efficient processes (all above 99% efficiency) were power generation, fugitives, road transport (diesel), shipping, stationary combustion, industry and aviation (Figure A.6). In contrast, the most inefficient processes were waste combustion, off-road transport, and road transport (gasoline) with averages of 0.6524 ± 0.4119, 0.952 ± 0.0349, and 0.9777 ± 0.0127. The standard deviations between the estimates of combustion efficiency were generally inversely proportional to the combustion efficiency, which means the less efficient processes have a more uncertain CO/CO_2 ratio. For biofuels, the most efficient processes were off-road transport, power generation, road transport (diesel), industry and road transport (LPG), all above 99%. In contrast, the most inefficient combustion processes for biofuels were agriculture, stationary combustion and road transport (gasoline) with averages of 0.943, 0.959 ± 0.017 and 0.979 ± 0.011. Moreover, we found a strong dependency of the combustion efficiency on the country, but the proportion of the sources contained within it drives the overall combustion efficient of a country.

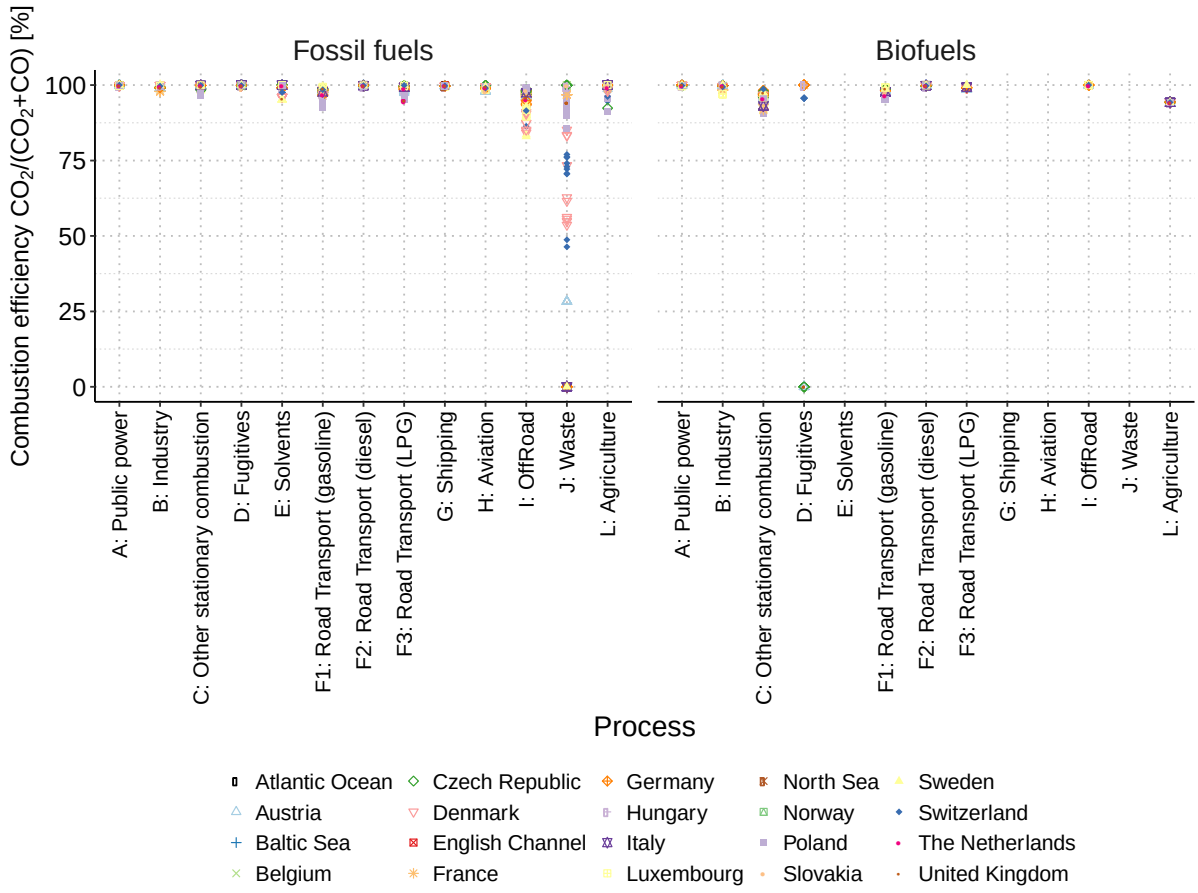

Figure A.6 Average combustion efficiency for fossil fuel (left) and biofuel (right) per process per country.

Surprisingly, we found that apart from process, the CO/CO_2 ratio in the TNO GHGco inventory is highly dependent on the country (figures A.4 and A.6). The relative uncertainty was also strongly determined by the country (figure A.5).

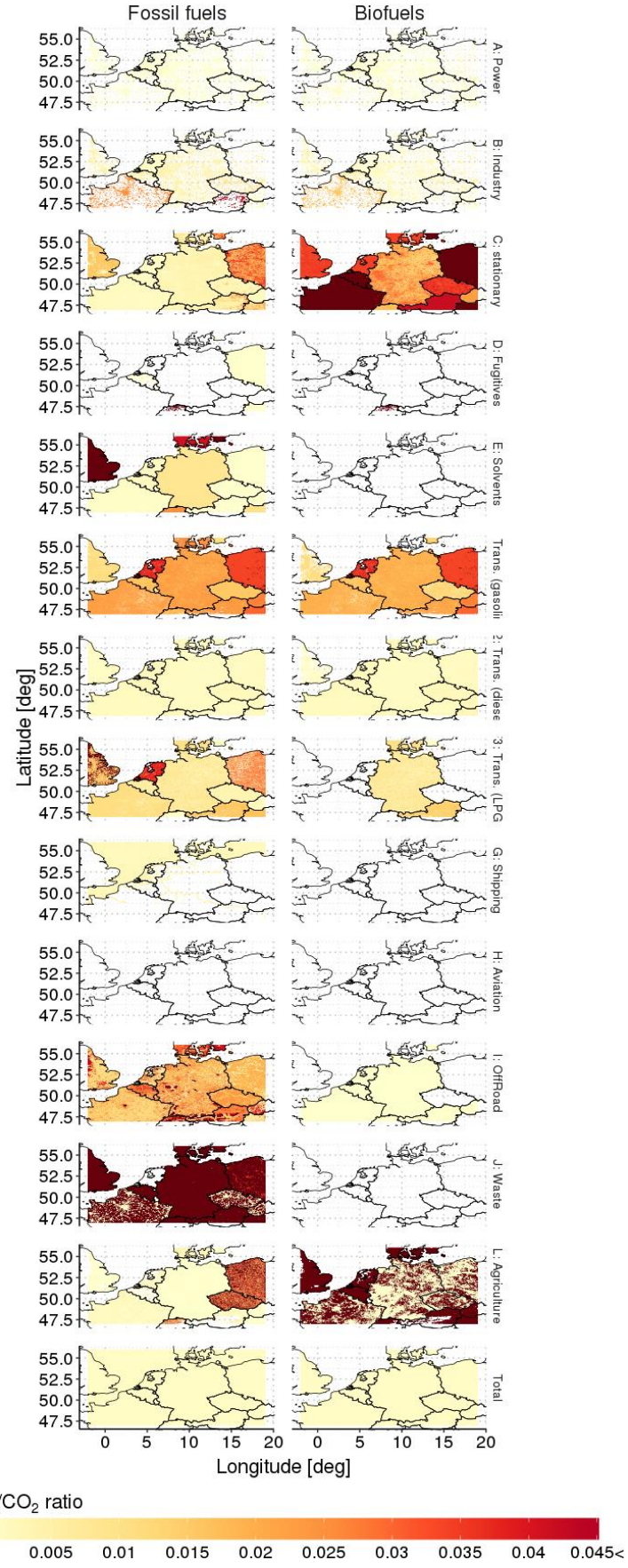

Figure A.7 Average CO/CO_2 ratio (molar) per gridcell for each emission process.

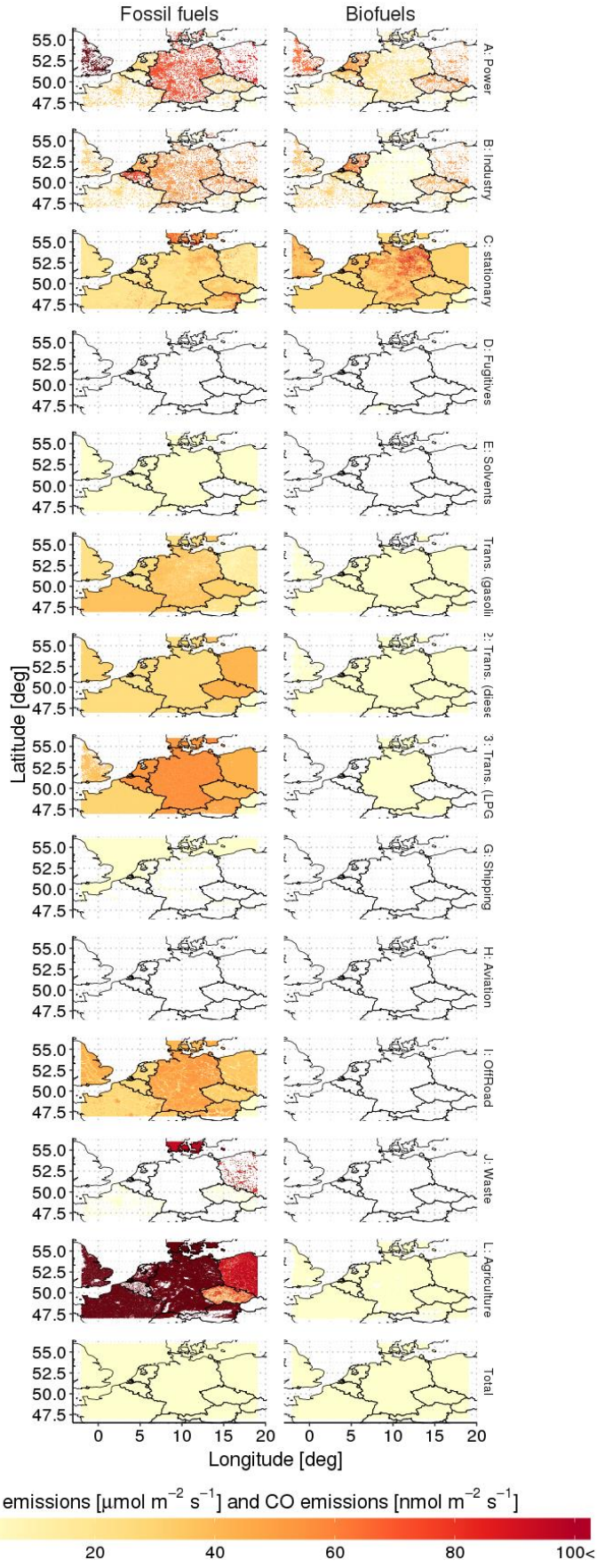

Figure A.8 Relative uncertainty (σ/μ, %) of the CO/CO_2 ratio (molar) per gridcell for each of the emission processes and the total emissions (bottom row).

A.4. Spatial correlations

The inversion uses can use the differences in the spatiotemporal distribution of emissions to distinguish the contribution of different processes. As such it is necessary to evaluate how the spatial distribution of the different process contained in the TNO GHGco inventory relates to each other.

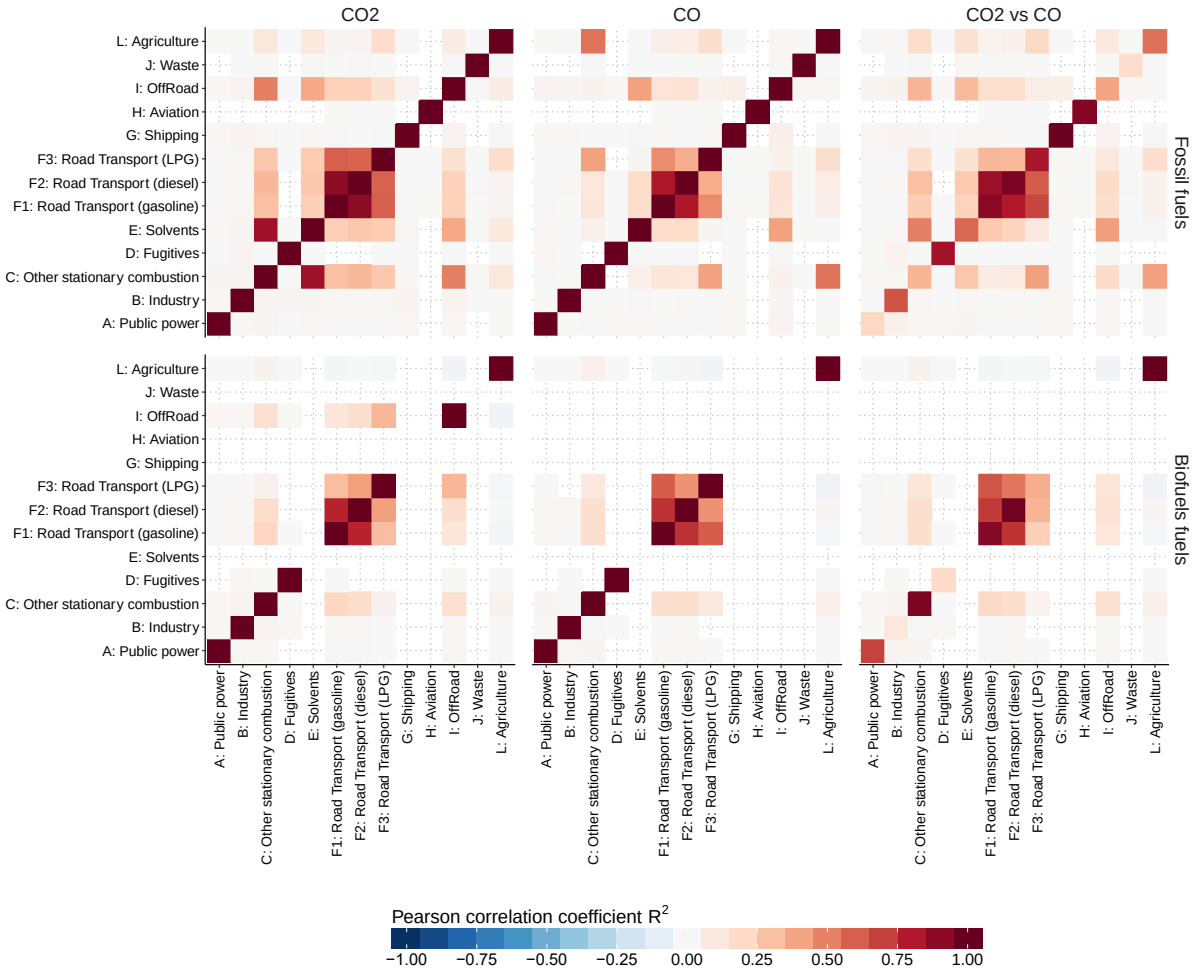

Figure A.9 Spatial correlation between emission process for CO_2 (left) and CO (middle) and between the emission processes for CO_2 and the emission process from CO (right).

A.4.1. Spatial correlations between CO_2 and CO emissions

Surprisingly, we found only moderate spatial correlations between the CO_2 and the CO emissions for many of the processes (Figure A.9, diagonal of matrices on the right column). For fossil fuels, the processes with the lowest correlations between the spatial distributions of their CO_2 and CO emissions were waste combustion, power generation, stationary combustion and off-road transportation with correlation coefficients, R^2, of 0.182, 0.209, 0.338 and 0.395 respectively. Remarkably, power generation and stationary combustion had

such a low correlation between their CO_2 and CO emissions due to their overall high combustion efficiency. In contrast, the process with the highest correlations between the spatial distribution their CO_2 and CO emissions were shipping, road transportation (diesel), aviation, and road transport (gasoline) with correlation coefficients, R^2, of 0.988, 0.939, 0.923 and 0.905 respectively. These components share two characteristics: they are in general confined to certain locations, e.g. oceans and water ways, roads and airports, and except for gasoline engines, they are characterized by high combustion efficiencies. For biofuel combustion, the processes with the lowest correlations between the spatial distributions of their CO_2 and CO emissions were industry, fugitives and road transport (LPG) (0.112, 0.191 and 0.366 respectively). In contrast, the processes with the highest correlation coefficients between the spatial distributions of their CO_2 and CO emissions were agriculture, road transport (diesel), stationary combustion and road transport (gasoline) with correlations coefficients, R^2, of 0.9999, 0.9692, 0.9476 and 0.9167 respectively.

A.4.2. Spatial correlations between emission processes

The spatial correlations between the emission processes are important to understand which processes can we aggregate to reduce the degrees of freedom of the optimization. Expectedly, for CO_2 fossil fuel emissions, the three road-transportation sectors are spatially highly correlated with each other (Figure A.9), particularly for diesel and gasoline vehicles (R^2 = 0.898), while the abundance of LPG powered vehicles may vary by region (R^2 of 0.602 and 0.591 with gasoline and diesel vehicles respectively). Furthermore, road transportation emissions are also correlated spatially with stationary combustion, solvent, and off-road transportation emissions (average R^2 of 0.299, 0.257 and 0.204 respectively). Stationary combustion has the strongest spatial overlap with other process, and it has a particularly high spatial correlation with solvents (R^2 of 0.851), a moderate spatial correlation with off-road transportation (R^2 of 0.511) and road transportation. In contrast, the aviation, waste, fugitives, power generation and industry emissions are spatially relatively independent to other processes. For biofuel CO_2 emissions, these correlation patterns are kept with the difference that there are no biofuel emissions for the solvents, shipping, aviation and waste sectors. With respect to CO emissions, the correlations between the processes are also similar to those found with the CO_2 emissions with the difference that there are no CO emissions from solvents (fossil fuels) and from off-road transportation (biofuels).

A.4.3. Spatial correlation lengths for the different emission processes

For the determination of the average spatial correlation length, it was necessary to subsample the emission data due to its high resolution. For this, the 13 gridcells with the largest emissions for each emission category were selected along with all gridcells within the same latitude and longitude band. With the selected gridcells, the spatial correlogram was estimated (*spline.correlog* function in R package *ncf*). A Gaussian curve was then fitted to the correlogram (*nls* function in R package *stats*). The decay constant of this Gaussian curve was assumed to represent the correlation length for each process in the entire domain (Figure A.10).

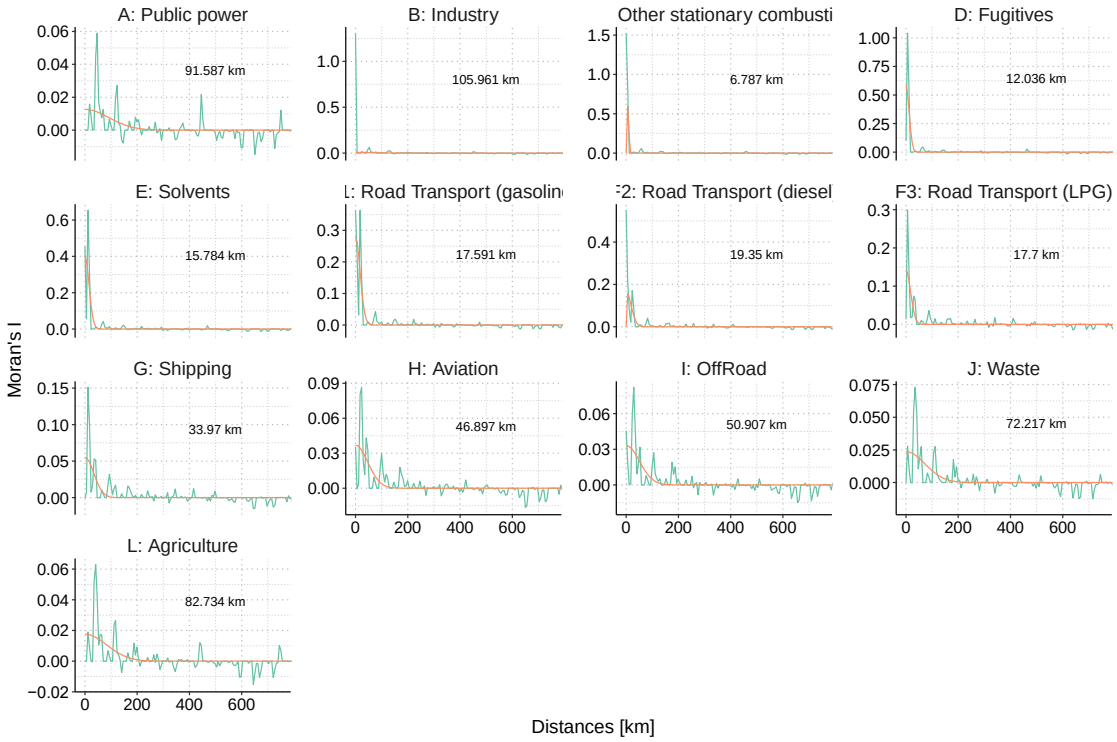

Figure A.10 Spatial correlation between emission process for CO_2 (left) and CO (middle) and between the emission processes for CO2 and the emission process from CO (right).

A.5. Temporal correlations

Emission processes may vary in time, e.g. residential heating increases with colder weather or road traffic emissions are higher within the weekdays. This variability may allow from atmospheric inverse modelling system to partition between emission processes involved. EMPA has generated specific daily, weekly and monthly scaling factors (Figure A.11, Figure A.12 and Figure A.13 respectively) for the different source processes contained in the TNO-GHGco inventory. To obtain hourly values, the annual TNO GHGco emissions were scaled with the specific hourly factors provided by EMPA (average temporal profiles of the ten scenarios). It is important to point out that we have scaled the temporal scaling functions such that they sum up to one for the whole year. Furthermore, EMPA also provided an ensemble of ten estimates of the scaling functions for each process to obtain a measure of uncertainty. The standard deviation of the temporal profiles from the ten scenarios can be added to the emission uncertainty.

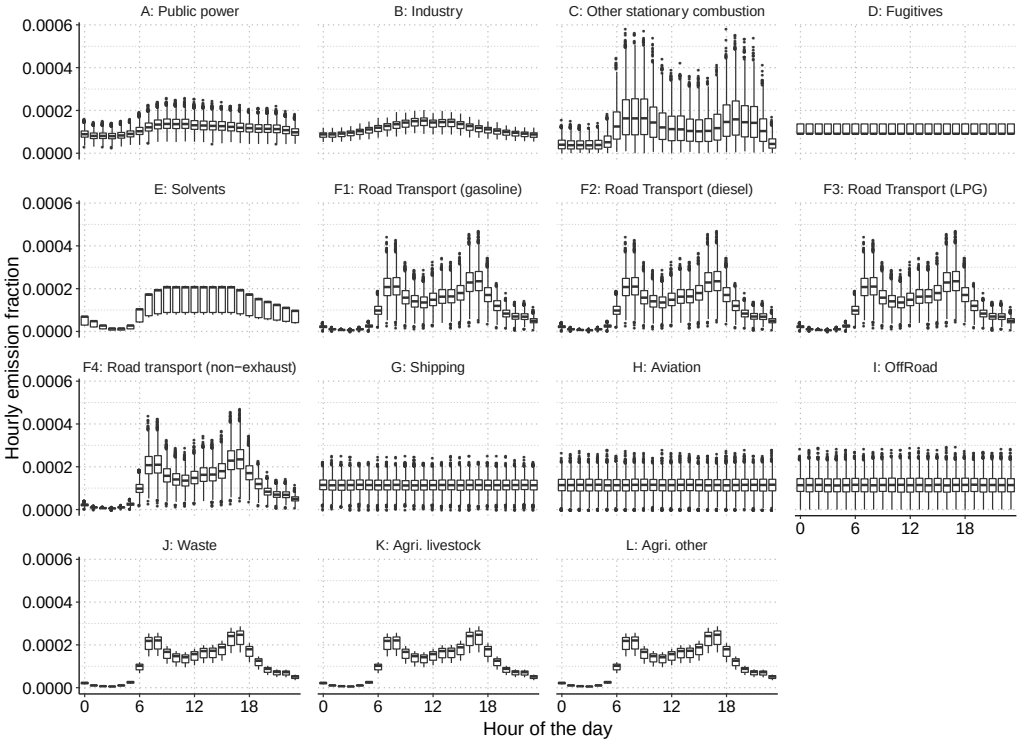

Figure A.11 Daily variability of the temporal scaling factors for each source process. The spread does not only represent the ensemble of ten estimates of the temporal variability provided by EMPA but also the overlapping weekly or monthly variability.

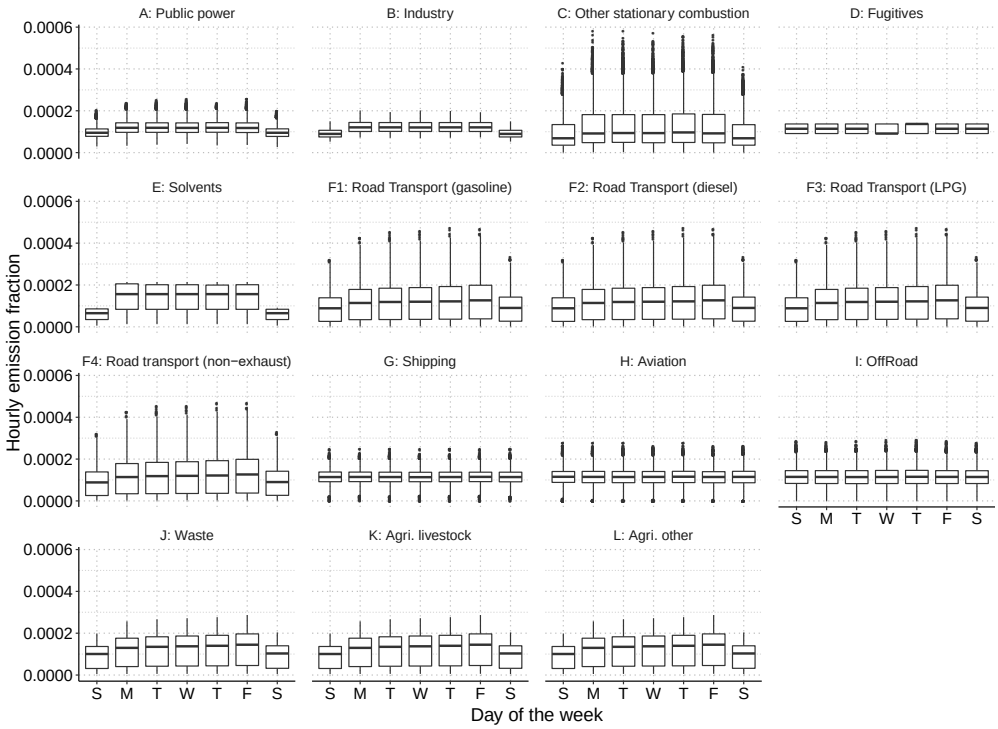

Figure A.12 Weekly variability of the temporal scaling factors for each source process.

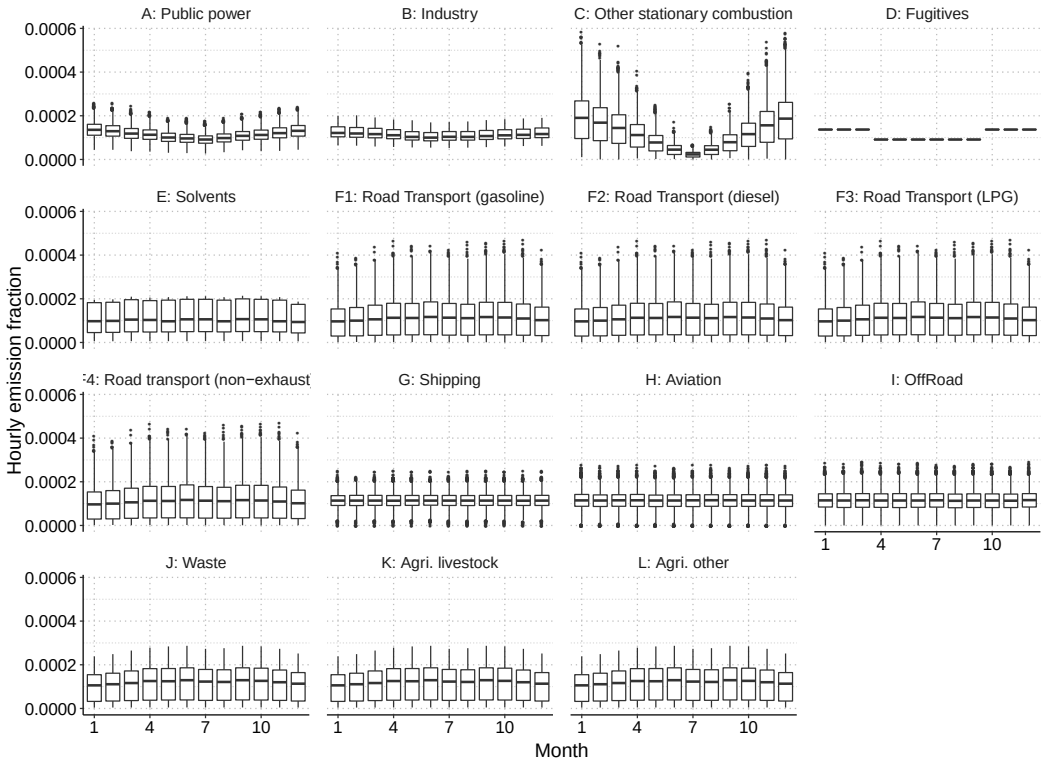

Figure A.13 Monthly variability of the temporal scaling factors for each of the source process.

A.5.1. Temporal correlations between emission processes

The temporal correlations between the emission processes are important to understand which processes we can aggregate to reduce the degrees of freedom of the optimization. Expectedly, the four road-transportation sectors are temporally fully correlated with each other (R^2 = 1, Figure A.14), but they were also highly correlated with waste burning and agriculture (R^2 = 0.923) and solvents (R^2 = 0.782). These processes were characterized with higher emissions during the summer months, during the weekdays and during the day, with peak emissions in the morning and afternoon. Moreover, (Stein et al., 2014) determined that increasing anthropogenic CO emissions during the winter (particularly from road transportation) improved the fit with observations in the Northern Hemisphere. Another cluster was conformed by power generation, industry and stationary combustion (R^2 values ranging from 0.476 to 0.676) with higher emissions during the winter, during weekdays, and during the day. The processes with no or little seasonal, weekly or diurnal variability, off-road transportation, shipping, aviation, and fugitive are not correlated with other processes.

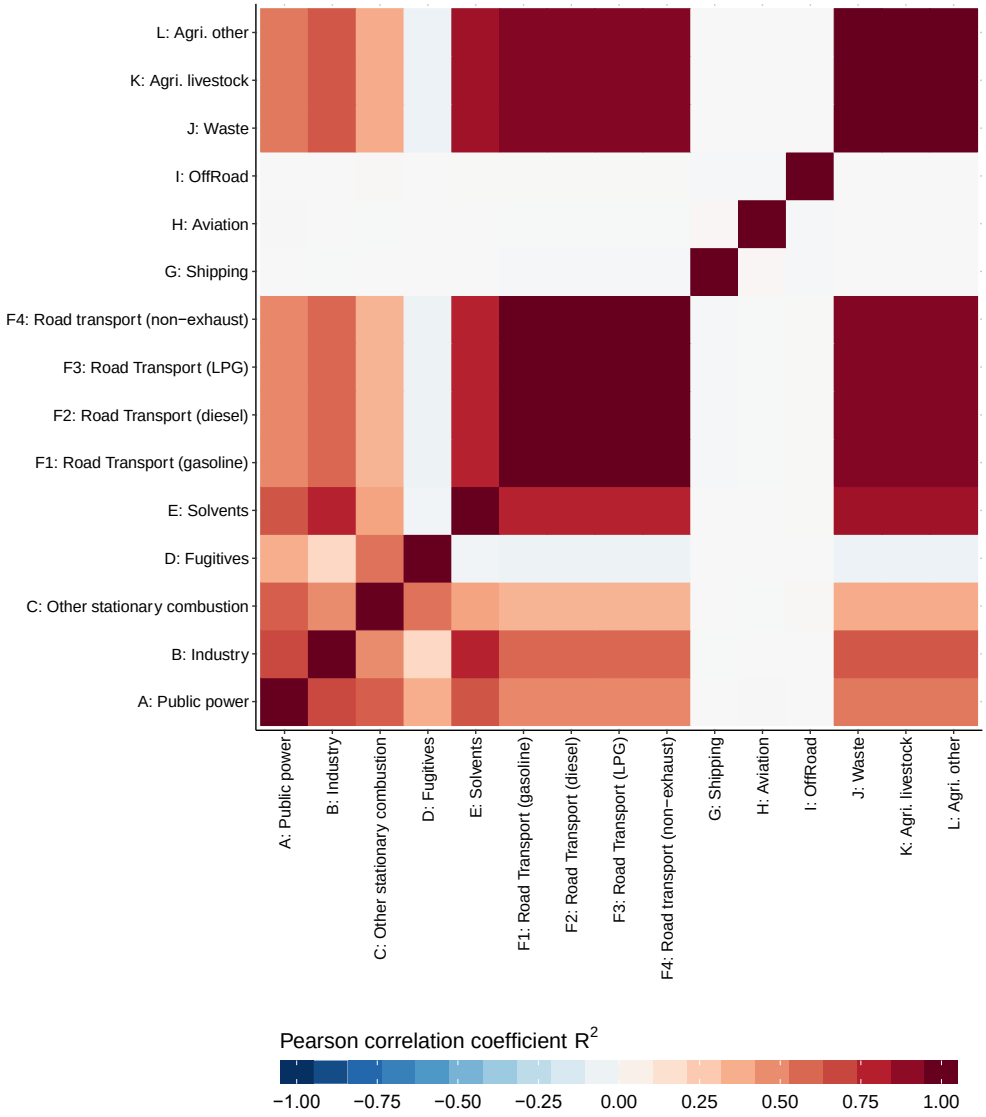

Figure A.14 Temporal correlation matrix between the processes included in the TNO GHGco inventory.

A.5.2. Temporal autocorrelations of the different emission processes

Because we are interested in optimizing yearly budgets, we filtered out daily and weekly variability from the time series using a simple box filter with a cut-off frequency of 28 days repeating the time series three times to avoid edge effects. We extracted the central time series and determined the autocorrelation function. We then fitted a Gaussian model to the autocorrelation function. The decay rate constants of the Gaussian model ranged from 11.86 to 34.98 days (Figure A.15). The longer decay rates belong to processes where the temporal variability is mainly determined by the seasonal cycle, while the shorter decay rates belong to process with little or no monthly variability but remnants of the weekly and daily variability in the filtered time series.

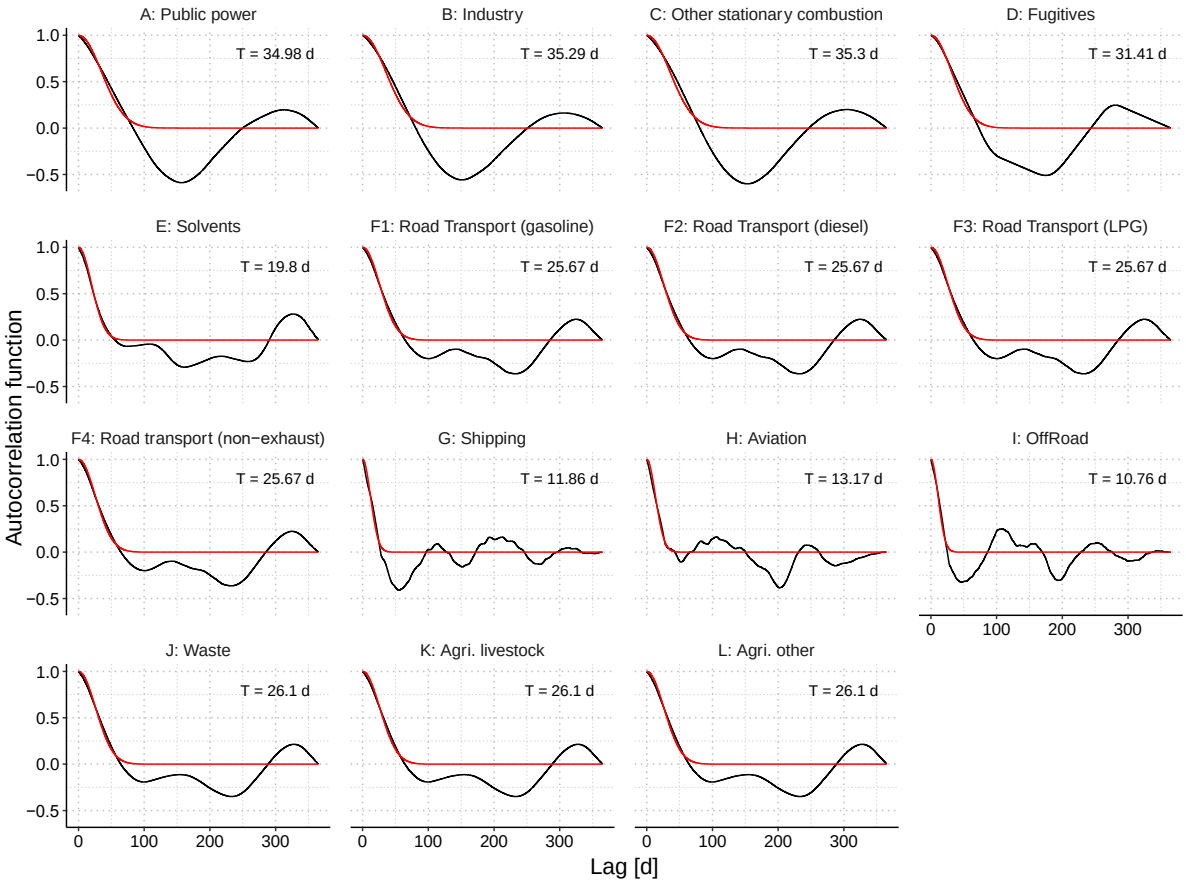

Figure A.15 Autocorrelation function per emission process (black line) and fitted Gaussian functions (red line). T represents decay rate constant of the Gaussian model.

B. Analysis of impact of CO photochemistry on regional CO and CO2 mole fractions

In this section, we analyse the necessity of including the photochemical generation of CO through the destruction of formaldehyde (HCHO) through the reaction HCHO+OH, and the destruction of CO through the reaction CO+OH, which results in the generation of CO_2. From a budget perspective, the reaction CO+OH generates 4.16 Pg CO_2 a^{-1} globally (Zheng et al., 2018). Within our full European and Central European domains, the reaction CO+OH generates 0.2 Pg CO_2 a^{-1} and 0.012 Pg CO_2 a^{-1} respectively. Putting these numbers into perspective (based on Le Quéré et al., 2018), the reaction CO+OH has a larger global contribution to atmospheric CO_2 than cement emissions (1.5 Pg CO_2 a^{-1}) and it is larger than the global CO_2 budget imbalance (1.9 Pg CO_2 a^{-1}). Moreover, within the European Union, cement production emissions accounted for 0.1 Pg CO_2 a^{-1} (Le Quéré et al., 2018), which is only half of the CO_2 produced from the CO+OH reaction within our European domain. Furthermore, the reaction HCHO+OH generated globally 1284.58 Tg CO a^{-1} (Zheng et al., 2018), and 54.6 and 3.40 Tg CO a^{-1} were generated within our European and Central European domains respectively. The chemical production of CO from the HCHO+OH reaction is the largest single source term in the global CO budget (total source of ~2600 Tg CO a^{-1}, Zheng et al., 2018).

To analyse the atmospheric signals generated by CO photochemistry, we obtained the

masses of CO and CO_2 generated from the reactions HCHO+OH and CO+OH respectively from the inverse modelling estimation in Zheng et al. (2018, inversion four). The mass fluxes were aggregated to TM3 fine grid resolution (Heimann and Körner, 2013) and transported forward in time accounting for the first order destruction of CO through the reaction with OH. For this, we used the OH climatological fields from Spivakovsky et al. (2000), which were distributed within the TRANSCOM-CH4 experiment (Patra et al., 2011), and the reaction from Liu and Sander (2015). We generated three scenarios, the first one included the fluxes from the reaction from the entire global, while the remaining only included the fluxes within our full European domain and our Central European domain.

The forward simulations revealed that there was a smooth latitudinal gradient in CO_2 and CO generation from photochemistry and these were more important towards the south of our European domain (Figure A.16 and Figure A.17). When looking at the global CO_2 and CO generation, there was a strong contribution of tracer coming from the south from the domain over southwestern Europe and northwestern Africa. Furthermore, CO_2 and CO generation over Europe peaks during the summer. When considering the full European domain we observed that the highest mole fractions were found over the Eastern Mediterranean Sea, Italy, the Balkans, Greece and Turkey. When considering the production within the Central European domain, the CO_2 and CO mole fractions are highest over the Alps and Southern Germany.

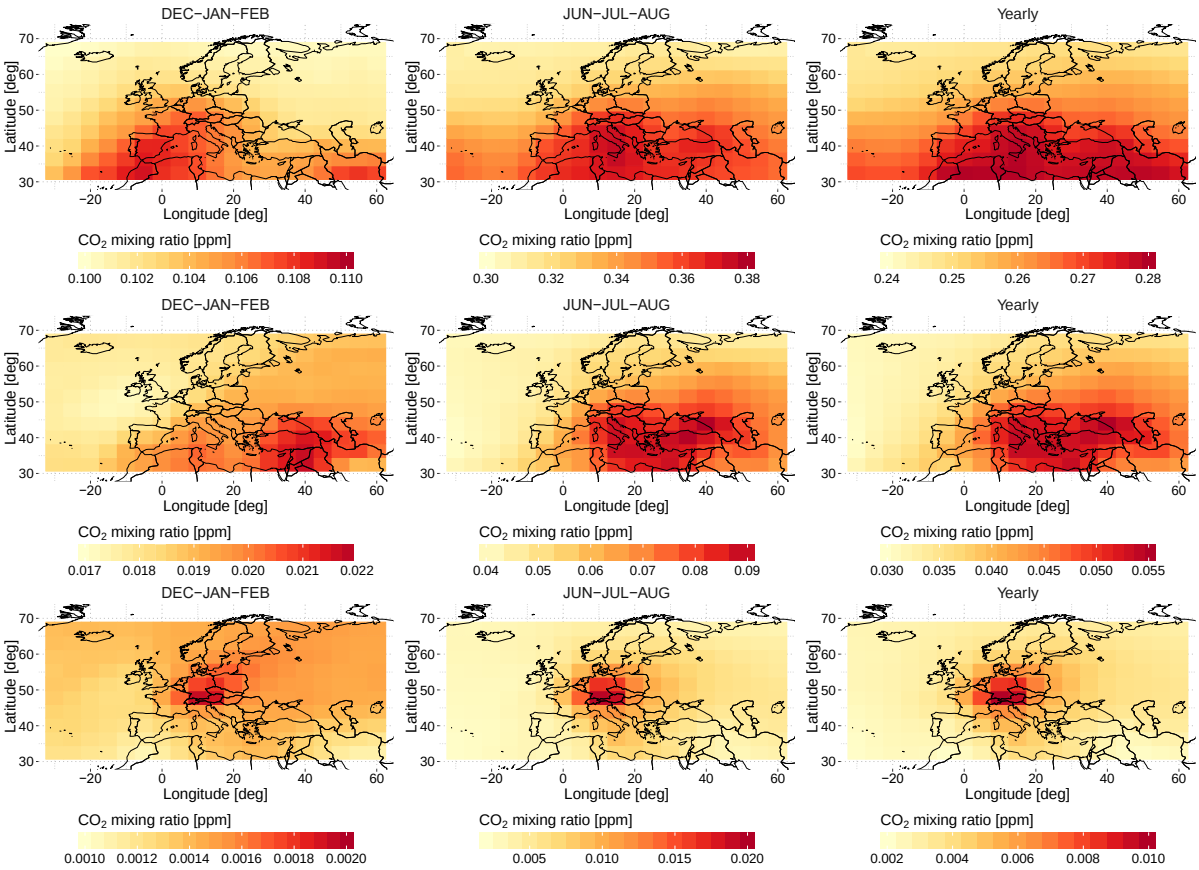

Figure A.16 Seasonal and yearly average CO_2 mole fraction in the surface-most TM3 model level which results from the reaction CO+OH globally (top row), within the full European domain (middle row) and within the Central European domain (bottom row).

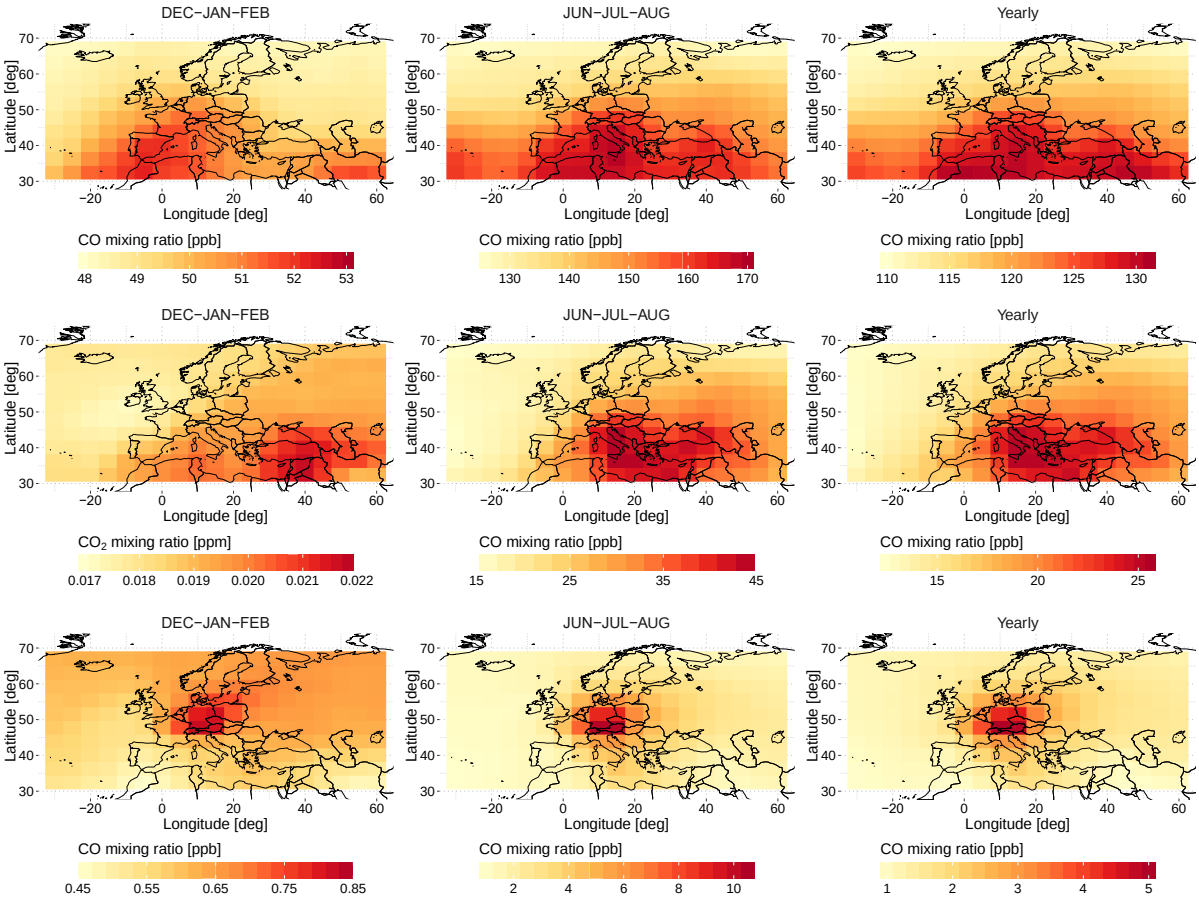

Figure A.17 Seasonal and yearly average CO mole fraction in the surface-most TM3 model level which results from the reaction HCHO+OH globally (top row), within the full European domain (middle row) and within the Central European domain (bottom row).

We sampled our four dimensional mole fraction fields at the locations of six sampling stations every three hours (Figure A.18 and Figure A.19) in order to look at evolution of the time series. When considering the global photochemical generation of CO_2 and CO, we observed a steep growth during the summer months, which continued at a lower rate into the winter. This happened because, despite the reduction of CO_2 and CO generation over Europe at the end of the summer, more of the tracer is being transported into the domain. At the end of the year mole fractions of ~0.5 ppm CO_2 and ~250 ppb CO were observed. In the case of CO_2 and CO generation within our full European and Central European domains, we observed that both tracers peak in early September and then decay smoothly toward the winter. For the full European domain, we observed peaked CO_2 and CO mole fractions of ~0.1 ppm and ~40 ppb respectively. For the Central European domain peak CO_2 and CO mole fractions reached ~0.03 ppm and ~10 ppb. The highest mole fractions peaks were observed at Bialystok, Poland (BIK, 53.23ºN, 23.02ºE) and Ochsenkopf, Germany (OXK, 50.03ºN, 11.81ºE). While typical the CO_2 and CO mole fraction measurement precisions lie within 0.1 ppm and 2 ppb respectively, we must remember we are considering only the differences with respect to the boundary conditions. Given the budget constraints, the emissions from anthropogenic sources likely produce similar magnitudes in mole fraction for these small domains.

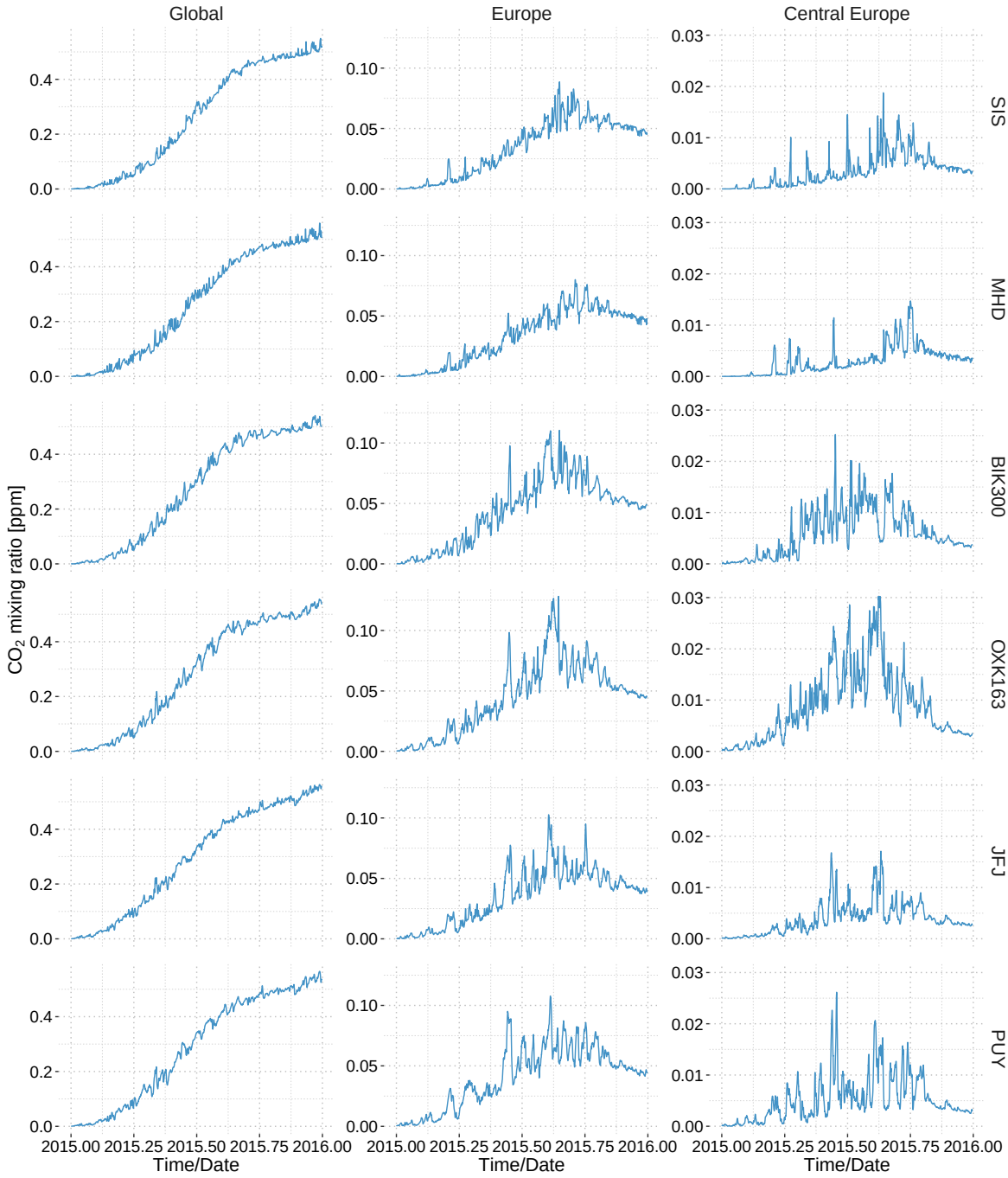

Figure A.18 CO_2 mole fraction signals resulting from the reaction CO+OH at six European stations (sorted by latitude). Signals are shown for the global CO2 generation as well as only within the European and Central European domains.

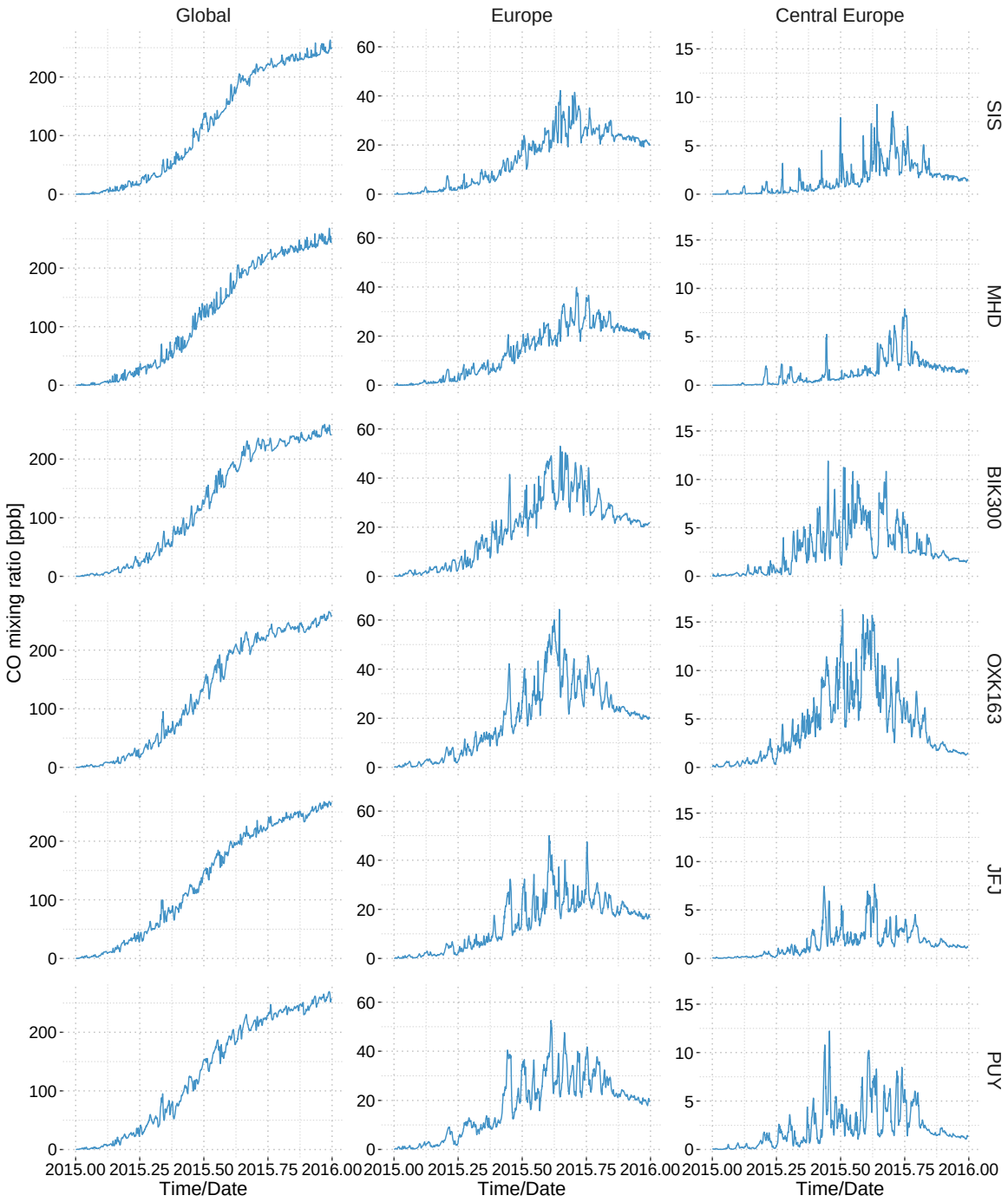

Figure A.19 CO mole fraction signals resulting from the reaction HCHO+OH at six European stations (sorted by latitude). Signals are shown for the global CO generation as well as only within the European and Central European domains.

In WRF initial and boundary conditions are determined from a global simulation. The photochemical source of CO_2 and CO should be added to the global simulation that produces initial and boundary conditions.

Document History

Version	Author(s)	Date	Changes
0-0	Tonatiuh Nuñez Ramirez (MPG) based on Confluence page input	2019-04-14	
1-0	Julia Marshall and Tonatiuh Nuñez Ramirez (MPG), updated based on internal iterations	2019-05-17	-Added APO products for ocean and boundary layer, plus anthropogenic emissions from COFFEE+EDGAR
1-1	Tonatiuh Nuñez Ramirez	2019-06-03	Appendix A first draft
1-2	Elise Potier, Gregoire Broquet, Frederic Chevalier	2019-07-10	Model setup for LSCE, $^{14}CO_2$ overall description, review of the document
1-3	Dominik Brunner	2019-07-10	Model setup for EMPA, review of the document
1-4	Ignacio Pisso	2019-08-22	Model setup for NILU
2-0	Tonatiuh Nuñez Ramirez	2019-09-10	Completion of all subsections and inclusion of subjects discussed in telecom of 2019-07-18.
2-1	Elise Potier, Gregoire Broquet, Frederic Chevalier	2019-09-30	Revision of consistency, review of the document
3-0	Tonatiuh Nuñez Ramirez	2019-10-09	Version for internal review
4-0	Tonatiuh Nuñez Ramirez	2019-10-10	Final
4-1	Tonatiuh Nuñez Ramirez	2019-10-17	Final + last minute reformulations and edits

Internal Review History

Internal Reviewers	Date	Comments
Michael Buchwitz (U. Bremen)	2019-09-24	
Andrew Manning (UEA)	2019-10-11	
Matthew Jones (UEA)	2019-10-10	

Estimated Effort Contribution per Partner

Partner	Effort
MPG	7
LSCE	1.5
EMPA	2.16
NILU	0.5
TNO	1.5
UEA	0.0041667
UB	0.12
Total	12.78417