Atmos. Chem. Phys., 22, 15287–15312, 2022
https://doi.org/10.5194/acp-22-15287-2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Chemistry and Physics
Open Access
EGU

Research article

Towards monitoring the CO_2 source–sink distribution over India via inverse modelling: quantifying the fine-scale spatiotemporal variability in the atmospheric CO_2 mole fraction

Vishnu Thilakan[1,2], Dhanyalekshmi Pillai[1,2], Christoph Gerbig[3], Michal Galkowski[3,4], Aparnna Ravi[1,2], and Thara Anna Mathew[1]

[1]Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research Bhopal (IISERB), Bhopal, India
[2]Max Planck Partner Group (IISERB), Max Planck Society, Munich, Germany
[3]Department of Biogeochemical Signals, Max Planck Institute for Biogeochemistry, Jena, Germany
[4]Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Kraków, Poland

Correspondence: Dhanyalekshmi Pillai (dhanya@iiserb.ac.in, kdhanya@bgc-jena.mpg.de)

Received: 17 March 2022 – Discussion started: 24 June 2022
Revised: 15 November 2022 – Accepted: 15 November 2022 – Published: 1 December 2022

Abstract. Improving the estimates of CO_2 sources and sinks over India through inverse methods calls for a comprehensive atmospheric monitoring system involving atmospheric transport models that make a realistic accounting of atmospheric CO_2 variability along with a good coverage of ground-based monitoring stations. This study investigates the importance of representing fine-scale variability in atmospheric CO_2 in models for the optimal use of observations through inverse modelling. The unresolved variability in atmospheric CO_2 in coarse models is quantified by using WRF-Chem (Weather Research and Forecasting model coupled with Chemistry) simulations at a spatial resolution of 10 km × 10 km. We show that the representation errors due to unresolved variability in the coarse model with a horizontal resolution of 1° (∼ 100 km) are considerable (median values of 1.5 and 0.4 ppm, parts per million, for the surface and column CO_2, respectively) compared to the measurement errors. The monthly averaged surface representation error reaches up to ∼ 5 ppm, which is even comparable to half of the magnitude of the seasonal variability or concentration enhancement due to hotspot emissions. Representation error shows a strong dependence on multiple factors such as time of the day, season, terrain heterogeneity, and changes in meteorology and surface fluxes. By employing a first-order inverse modelling scheme using pseudo-observations from nine tall-tower sites over India, we show that the net ecosystem exchange (NEE) flux uncertainty solely due to unresolved variability is in the range of 3.1 % to 10.3 % of the total NEE of the region. By estimating the representation error and its impact on flux estimations during different seasons, we emphasize the need to take account of fine-scale CO_2 variability in models over the Indian subcontinent to better understand processes regulating CO_2 sources and sinks. The efficacy of a simple parameterization scheme is further demonstrated to capture these unresolved variations in coarse models.

Published by Copernicus Publications on behalf of the European Geosciences Union.

1 Introduction

Accurate assessment of the sources and sinks of CO_2 is essential in planning and implementing mitigation strategies for greenhouse gas emissions and associated climate change. However, estimations of CO_2 fluxes contain significant uncertainties, which increase even more with finer spatial scales such as those required for the climate change mitigation policies at regional and national levels (e.g. Ciais et al., 2014; Li et al., 2016; Cervarich et al., 2016). By using atmospheric CO_2 concentration measurements, the CO_2 fluxes can be estimated by a multi-constrained observation–modelling approach, often referred to as a top-down approach or inverse modelling (Enting, 2002). For about 2 decades, these top-down approaches have been widely used to understand the modifications in the carbon cycle through natural and anthropogenic-induced environmental changes (Bousquet et al., 2000; Schimel et al., 2001; Rödenbeck et al., 2003; Patra et al., 2005). In addition to the observations, the inverse modelling system makes use of an atmospheric transport model (forward model), which determines the distribution of CO_2 concentration. Thereby, the inverse optimization approach derives the surface fluxes that are consistent with the measured concentration. The United Nations Framework Convention on Climate Change (UNFCCC) has acknowledged the increasing capability of inverse modelling to systematically monitor greenhouse gas (GHG) concentrations (Bergamaschi et al., 2018).

Most of the inverse modelling systems rely on global atmospheric transport models with a coarse horizontal resolution (often greater than 1°) (Rödenbeck et al., 2003; Peters et al., 2007; Rödenbeck et al., 2018a, b; Inness et al., 2019). These global data assimilation systems play an important role in the study of continental or subcontinental fluxes at annual or sub-annual scales. However, the regional estimation of fluxes using global models is hindered by the inability of these transport models to represent the observed CO_2 variability. The observed variability, as seen from the spatial and temporal distribution of atmospheric CO_2, is highly correlated with the space and timescales of weather systems (Parazoo et al., 2011). This explains the presence of large model–data mismatches in regions where mesoscale circulation is predominant (Ahmadov et al., 2007). Wind speed, wind direction, and the height of the planetary boundary layer (PBL) are the critical variables that determine the atmospheric CO_2 variability. Strong wind normalizes other small-scale variations in observed concentration due to mixing. In such cases, the CO_2 variability is expected to be dominated by the variations in background concentration; hence, the predictability can be higher during these conditions (Sarrat et al., 2007). The height of the PBL is an essential variable since the atmospheric CO_2 is subjected to rapid mixing up to this altitude. Hence, for a given location with a negative gradient in CO_2 vertical distribution, an overestimation of the PBL height leads to an underestimation of CO_2 concentration, and vice versa (Gerbig et al., 2008).

Another important variable that impacts the CO_2 variability is the heterogeneous topography. Variations in topography influence the transport of the tracers. When the small-scale orographic details are not adequately represented in the models, they can lead to representation errors in CO_2 simulations as large as 3 ppm (parts per million) at scales of 100 km (Tolk et al., 2008; Pillai et al., 2010). Horizontal gradients in CO_2 concentrations can go up to values of 30 ppm within a spatial scale of 200 km, depending on the land surface heterogeneity (van der Molen and Dolman, 2007). Furthermore, variations in land use patterns between neighbouring regions can cause considerable variability in the CO_2 surface fluxes. Thus, a proper representation of land use patterns is also important in terms of simulating CO_2 variability. Previous studies based on airborne measurements reported that transport models need a spatial resolution smaller than 30 km to be able to represent CO_2 spatial variability in the continental boundary layer (Gerbig et al., 2003). Significant efforts have been invested in deriving fluxes by taking into account these fine-scale variations (e.g. Gerbig et al., 2003; Lauvaux et al., 2009a; Carouge et al., 2010; Pillai et al., 2011, 2012; Broquet et al., 2013) over North American and Eurasian domains in the past decade. However, there still exists lower confidence in estimates over other regions where there is a lack of both advanced modelling systems at relevant spatiotemporal resolutions and good coverage of ground-based monitoring stations.

In the context of the Indian subcontinent, the inverse-based estimation of fluxes at fine scales is essentially new; hence, many questions remain. A number of monitoring sites measuring atmospheric greenhouse gases have become available in India during the last decade (Tiwari et al., 2011; Lin et al., 2015; Nomura et al., 2021). Aside from the ongoing progress in augmenting observational data streams, it remains challenging to assimilate these data for deducing process-specific information effectively (e.g. McKain et al., 2012; Bréon et al., 2015; Pillai et al., 2016). The limitation of coarse global models in representing observations over the Indian subcontinent is reflected in the analysis made by Patra et al. (2011).

The seasonally reversing South Asian monsoon system is a prominent meteorological phenomenon affecting the Indian subcontinent, which is also expected to influence the terrestrial–atmosphere flux exchanges. Various studies have demonstrated the role of Indian monsoon circulations on regional atmospheric transport by strong southwesterly winds during the summer monsoon (June to September) and by northeasterly winds during the winter monsoon (October to November; e.g. Goswami and Xavier, 2005; Krishnamurthy and Shukla, 2007). Monsoon convection transports the boundary layer air into the upper troposphere. Subsequently, air parcels are slowly uplifted by diabatic heating to higher altitudes (e.g. Vogel et al., 2019). An accurate representation of convective vertical transport is very chal-

lenging and an important source of uncertainty in current transport models (Willetts et al., 2017). Note that the Asian summer monsoon anticyclone (ASMA) active during the Indian summer monsoon period plays a key role in uplifting trace gases to the upper troposphere and lower stratosphere (e.g. Park et al., 2007). Moreover, a significant component of flux variations can arise from biospheric fluxes (Schimel et al., 2014), which is influenced by variables such as rainfall, radiation, and temperature (Chen et al., 2019). Several studies showed that the monsoon system substantially impacts vegetation growth, generating distinct spatiotemporal patterns of the biogenic fluxes (e.g. Gadgil, 2003; Valsala and Maksyutov, 2013; Ravi Kumar et al., 2016; Kunchala et al., 2022). It is noteworthy that the cropping patterns over India have a strong dependence on seasons and are mainly determined by dry and wet seasons for nearly 65 % to 70 % of the country's area, except over the northeastern and southwestern (Western Ghats) regions of India. In India, wet-season crops (kharif crops cultivated from June to November), including rice, millets, and maize, mainly depend on monsoon rain. Dry-season crops (rabi crops, e.g. wheat, barley, and mustard, cultivated from November to April) are less water-dependent and primarily rely on irrigation (DAC/MA, 2015). Therefore, employing higher-resolution modelling over the Indian subcontinent is desirable to better account for fine-scale variations generated by both mesoscale transport processes and surface flux patterns.

This study focuses on accounting for unresolved sub-grid-scale variability when employing current generation global models. The assimilation of observations in an inverse framework requires the characterization of these error structures at relevant scales that can be utilized to retrieve the source–sink distribution over India. The main objectives of this paper are to describe and quantify the expected spatiotemporal variability in atmospheric CO_2 that is not resolved by the current generation global models, quantify to what extent these variations cause uncertainty in flux estimations, and assess how these uncertainties can be minimized by modelling the sub-grid variations in the global models. Specifically, we address the following questions: (1) how good is the level of agreement among global transport models that are used in current generation inversion systems for predicting atmospheric CO_2 concentrations over the Indian subcontinent? (2) How large are the variations in atmospheric CO_2 that are unresolved by global and regional models, which operate at different spatial scales from $4° \times 4°$ to $0.5° \times 0.5°$? (3) What is the role of seasonal changes in generating different patterns in these sub-grid variations in CO_2? (4) How much is the uncertainty in the inverse-based flux estimation caused by these unresolved variations in the coarse models when utilizing a given network of surface observations over the domain? (5) How effectively can we capture the key aspects of the variability and account for it in flux estimations? Information from observations can be better utilized if we improve the atmospheric transport models to resolve the observed variability as accurately as possible. As a result, the data assimilation system gains significantly (e.g. with increasing weights on observations and performing minimal data filtering) from this for improving the flux estimates.

In this article, we present results based on the analyses of high-resolution simulations at a spatial resolution of $10\,km \times 10\,km$ for the months of July and November 2017. The year 2017 was characterized by neutral Indian Ocean Dipole conditions over the Indian Ocean with the beginning of a mild La Niña over the Pacific by the end of the year (NOAA/ESRL, 2022a, b). The month of July represents a monsoon period when both biospheric and convective activity are strong. July is also characterized by strong low-pressure system activity over the Bay of Bengal, which results in large rainfall over central India (Krishnamurthy and Ajayamohan, 2010). In contrast, the month of November is more representative of post-monsoon wintertime over the Indian subcontinent. We quantify the expected sub-grid variability using our high-resolution simulations. We have also utilized optimized CO_2 products at global scales to provide a more comprehensive overview of the typical mismatch among the existing model simulations over the Indian subcontinent even at monthly and annual scales. These intermodel mismatches arise due to various reasons such as differences in input datasets (e.g. prior fluxes), transport, and model configuration. Part of these mismatches can also arise due to the inability of coarse-resolution global models to simulate the sub-grid-scale processes which can lead to representation errors; thus, there is uncertainty in inverse estimations. By designing a pseudo-surface observation network over the domain, we investigate the impact of these unresolved variations on the regional flux estimations and assess how a simple parameterization scheme can help to reduce these errors in the global model. To our knowledge, until now, there is no comprehensive published study of this kind over the Indian subcontinent assessing the magnitude and impact of temporal and spatial variability exhibited by atmospheric CO_2.

The outline of the paper (see Fig. S1 in the Supplement) is as follows: Sect. 2 describes our modelling system, data, and the methods used for reporting intermodel mismatches and quantifying the sub-grid-scale variability CO_2. In Sect. 3, we present our analysis, demonstrating the impact of unresolved sub-grid-scale variability on CO_2 flux estimation over India. Finally, we discuss the implications of our findings in Sect. 4, highlighting possible ways forward to yield an improved estimation of CO_2 budgets over India.

2 Data and methodology

In the following subsections, we describe our high-resolution modelling system (Sect. 2.1), the existing optimized global CO_2 simulations used in the study (Sect. 2.2), the quantification of the representation error (Sect. 2.3), and the observation system simulation experiment (OSSE) designed

to estimate the impact of the derived sub-grid-scale variations in flux estimations over India via inverse optimization (Sect. 2.4).

2.1 WRF-Chem-GHG modelling system

We use the modelling system WRF-Chem-GHG in which the Weather Research and Forecasting model (coupled with Chemistry) version 3.9.1.1 (Skamarock et al., 2008) is coupled with the greenhouse gas module (WRF-Chem-GHG; Beck et al., 2011) that is implemented as part of the WRF-Chem distribution (WRF-Chem; Grell et al., 2005). For simulating the atmospheric transport, the model uses fully compressible Eulerian non-hydrostatic equations on an Arakawa C staggered grid, conserving mass, momentum, and scalars (Skamarock et al., 2008). In the WRF-Chem-GHG (hereafter referred to as WRF-GHG), we use the passive tracer chemistry option to simulate changes in CO_2 mixing ratios associated with surface fluxes and atmospheric transport. We utilize a biospheric model and emission inventory data to simulate atmospheric CO_2 enhancements associated with biogenic and emission fluxes, as described in Sect. 2.1.1 and 2.1.2. Table 1 summarizes the model configuration, including physics parameterizations and input data used in this study.

The model domain covers a region spanning from 65 to 100° E and 5 to 40° N, configured in a Lambert conformal conic (LCC) projection with 307 × 407 grid points. The spatial resolution of the grid is 10 km × 10 km, and the model time step is 60 s. We have used a model output with a temporal resolution of 1 h for this study. The simulations are performed using 39 vertical levels, with the model top at 50 hPa and 10 levels within the lowest 2 km. WRF-GHG simulations are performed for the entire months of July and November 2017. Implementation of the WRF-GHG system over the Indian subcontinent enables us to customize it according to the domain features and build a state-of-the-art modelling system, which eventually estimates CO_2 fluxes through regional inverse systems. The potential of the WRF-GHG model in simulating the fine-scale spatial variability was also established in previous studies (Ahmadov et al., 2009; Pillai et al., 2011; Park et al., 2018).

2.1.1 Representation of biospheric fluxes

We use the Vegetation Photosynthesis and Respiration Model (VPRM) in the modelling system to calculate net ecosystem exchange (NEE) representing the biospheric fluxes (Mahadevan et al., 2008). VPRM is a diagnostic biosphere model, which utilizes the following remote sensing products: Enhanced Vegetation Index (EVI) and Land Surface Water Index (LSWI) derived from reflectance data of the Moderate Resolution Imaging Spectroradiometer (MODIS), in addition to meteorological data, namely solar radiation and air temperature. In this study, these hourly NEE calculations are performed within WRF-GHG and simultaneously with the meteorology simulations in which NEE is calculated as a sum of gross ecosystem exchange (GEE) and ecosystem respiration (R_{eco}). VPRM, in this case, uses the meteorological data provided by WRF-GHG. VPRM uses the SYNMAP vegetation classification (using the tile approach; Jung et al., 2006) together with EVI and LSWI from MODIS surface reflectance data at a resolution of 1 km and 8 d. We aggregate these indices specific to different vegetation types onto the LCC projection for the entire domain at the model's spatial resolution. A number of studies have used VPRM for other regions around the world in which derived NEE shows good prediction skills for hourly to monthly timescales (Ahmadov et al., 2009; Pillai et al., 2011; Liu et al., 2018; Park et al., 2018).

2.1.2 Representation of emission fluxes

Anthropogenic CO_2 emission fluxes are prescribed from the Emissions Database for Global Atmospheric Research (EDGAR) dataset, version 6.0, provided at a horizontal resolution of 0.1° × 0.1° (Crippa et al., 2021). We disaggregate the available annual emission data into hourly emissions using the temporal distribution CO_2 profiles (Steinbach et al., 2011; Kretschmer et al., 2014). To represent biomass burning emission, we have used data from the Global Fire Assimilation System (GFAS) with a spatial resolution of 0.1° × 0.1° and a temporal resolution of 1 d. GFAS is based on satellite data, which provides the fire emissions by assimilating fire radiative power (FRP) observations from MODIS instruments (Kaiser et al., 2012). All these flux data are gridded and projected to WRF-GHG's model domain.

2.1.3 Initial and boundary conditions

Meteorological and chemical initial and boundary conditions are required in WRF-GHG to account for the initial state and inflow or background flow. The initial and lateral boundary conditions for the meteorological variables include horizontal wind components, pressure, specific humidity, and sea surface temperature (SST). These fields and other necessary surface initialization fields are obtained from the ERA5 reanalysis dataset of the European Centre for Medium-Range Weather Forecasts (ECMWF) and extracted at a horizontal resolution of 25 km and a temporal resolution of 1 h (Hersbach et al., 2020). The initial and lateral boundary conditions of CO_2 tracers are obtained from the Copernicus Atmosphere Monitoring Service (CAMS) global greenhouse gas forecast products (currently in development; see Massart et al., 2016; Agustí-Panareda et al., 2019). Namely, we have used the dry-air mole fractions of CO_2 from the CAMS greenhouse gas experiment analysis (gqiq), with a temporal resolution of 6 h, horizontal resolution of 0.5° × 0.5° (original resolution 9 km × 9 km), and 137 vertical levels. Note that the CAMS product at 9 km × 9 km resolution is in the developmental

Table 1. WRF-GHG Model setup. NA stands for not available.

Domain						
Configuration	Single domain with a horizontal resolution of 10 km; 39 vertical levels; 307 × 407 grid points					
Vertical coordinates	Terrain-following hydrostatic pressure vertical coordinates					
Basic equations	Non-hydrostatic; compressible					
Grid type	Arakawa C grid					
Time integration	Third-order Runge–Kutta split explicit					
Spatial integration	Third- and fifth-order differencing for vertical and horizontal advection, respectively; both for momentum and scalars					
Time step	60 s					
Physics schemes						
Radiation	Rapid Radiative Transfer Model (RRTM) for longwave and Dudhia for shortwave					
Microphysics	WRF single-moment three-class (WSM3) classic simple ice scheme					
PBL	Yonsei University (YSU) scheme					
Surface layer	Monin–Obukhov					
Land surface	NOAH land surface model (LSM)					
Cumulus	Grell–Freitas ensemble scheme					
Emission fields						
Flux type	Product	Version	Spatial resolution	Temporal resolution	Source/website	Reference
Anthropogenic	EDGAR	v6.0	10 km	Annual	https://edgar.jrc.ec.europa.eu/ (last access: 20 May 2022)	Crippa et al. (2018)
Biomass burning	GFAS	v1.2	10 km	Daily	http://apps.ecmwf.int/datasets/data/cams-gfas/ (last access: 15 March 2020)	Kaiser et al. (2012)
Biospheric	VPRM		Adapted to model	Adapted to model		Mahadevan et al. (2008)
Initial and lateral boundary conditions						
Field	Product	Version	Spatial resolution	Temporal resolution	Source/website	Reference
Meteorology	ERA5	NA	25 km	1 h	https://cds.climate.copernicus.eu/cdsapp#!/home (last access: 18 March 2020)	Hersbach et al. (2020)
Tracer	ECMWF/CAMS	gqiq	50 km	6 h	http://atmosphere.copernicus.eu (last access: 11 March 2020)	Agustí-Panareda et al. (2019)

phase and not yet available to the general public (contact anna.agusti-panareda@ecmwf.int for more information).

We have utilized a simulation strategy to update the initial meteorological conditions to take advantage of assimilated meteorological fields from ECMWF. The model is reinitialized each day with ECMWF assimilated data at the model starting time of 12:00 UTC (day + 0) and runs for 30 h until 18:00 UTC of the next day (day + 1). The first 6 h are considered to be the meteorological spin-up, and the remaining 24 h (from day + 0, 18:00 UTC to day + 1, 18:00 UTC) are used for the analysis. The initialization of CO_2 is done at the beginning of the first hour of the model simulation, which is at 00:00 UTC (e.g. Ahmadov et al., 2012; Pillai et al., 2011).

2.2 Global model products

We have used optimized products at global scales to examine the differences in the representation of CO_2 variability over the Indian subcontinent at monthly and annual scales. These global model outputs are derived from inverse model simulations, which estimate the source–sink distributions of CO_2 and then generate three-dimensional CO_2 concentration fields that are consistent with the optimized posterior fluxes. Four global inverse modelling products – CarbonTracker, CarboScope, LSCE v18r3, and LSCE FT18r1 – available for the year 2017 are used for our analysis (see Table 2 for more details). The Laboratoire des Sciences du Climat et l'Environnement (LSCE) model version v18r3 (hereafter LSCE) utilizes surface observations, and the model version FT18r1 (hereafter LSCE FT) uses satellite retrievals from the Orbiting Carbon Observatory (OCO-2) for the optimization of CO_2 fluxes (Chevallier et al., 2005, 2010; Chevallier, 2013). All of these above models differ in terms of the model formulations and configuration (e.g. transport and the employed inversion methodology), observational datasets that were assimilated (e.g. data from surface monitoring stations, aircraft missions, ship cruises, AirCore balloon soundings, and the satellite's total column retrievals), prior datasets, and spatiotemporal resolutions. None of these products used ground-based observations from the Indian subcontinent for their optimization, which raises concerns about the reliability of the optimized flux estimations over the region. Hence, it can be assumed that part of the inter-model differences in predicting the variability is related to the paucity of the CO_2 observations over the region. To represent the daytime, we have used the concentration fields for the local time (LT) ranging from 11:30 to 16:30 LT from all these models for the analysis.

2.3 Quantification of spatial variability

For quantifying the spatial variability due to sub-grid-scale processes that cannot be resolved by the coarse-resolution models, we follow the approach as described in Pillai et al. (2010). The term "representation error" indicates the mismatch between the scales of model simulations and observations collected (Pillai et al., 2010; Janjić et al., 2017). In other words, the representation errors arise due to unresolved scales not captured by the model. Here we calculate the representation errors in the coarse-resolution models, which can be resolved by implementing a high-resolution model at 10 km resolution. It is assumed that the high-resolution simulation captures the majority of the sub-grid-scale variability, even though it cannot be expected to resolve all observed variability. Most of the current global model simulations are performed at coarse resolutions of several degrees. But, with the recent advancement in computational capacity and numerical techniques, a horizontal resolution of $1^\circ \times 1^\circ$ is quite likely achievable for the global data assimilation systems. For estimating the representation error in a coarse model with a typical spatial resolution of $1^\circ \times 1^\circ$, we have calculated the standard deviation of the CO_2 dry-air mole fraction simulated by the WRF-GHG model within the coarse grid boxes of $1^\circ \times 1^\circ$ as follows:

$$\sigma_{CO_{2(tot)}} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(m_j - \overline{m}\right)^2}, \quad (1)$$

where $\overline{m} = \frac{1}{n}\sum_{j=1}^{n} m_j$. n is the number of 10 km boxes inside the coarser grid cell of $1^\circ \times 1^\circ$, m is the CO_2 dry-air mole fraction corresponding to 10 km boxes, and $\overline{m}$ is the average within the coarser grid cell. So, the estimated values represent the sub-grid-scale variability within the coarse model grid cell with a horizontal resolution of $1^\circ \times 1^\circ$. The representation errors are calculated at corresponding vertical model levels to represent the impact of surface influence and mesoscale transport adequately, as predicted by the high-resolution model. As mentioned before, we assume that the high-resolution simulations represent the realistic distribution of CO_2. Furthermore, we assume that the coarse-resolution model also has a terrain-following vertical coordinate system and also has the same vertical grid spacing of high-resolution model. As spaceborne instruments measure total columns rather than near-surface concentrations, we extend the analysis to the column-averaged dry-air mole fraction (XCO_2) as measured by the satellite instrument, i.e. m represents either CO_2 at a given model level or XCO_2. In order to assess the dependence of the representation errors in the horizontal resolution of the employed model, we have computed representation errors for multiple resolutions ranging from $0.5^\circ \times 0.5^\circ$ to $4^\circ \times 4^\circ$, in addition to $1^\circ \times 1^\circ$, which would encompass the resolutions of both present and near-future global inverse modelling systems.

The surface representation errors are calculated using the model simulations from the second model level (mean height is ~ 200 m from sea level) to avoid the inconsistency that can be generated from inputting emission fluxes at the first model level. Representation errors are calculated separately for daytime (11:30 to 16:30 LT) and nighttime (23:30 to 04:30 LT)

Table 2. Specifications of different global model products used in this study.

Data availability								
Product	Version	Spatial resolution	Vertical levels	Temporal resolution	Source/website		Reference	
CarbonTracker	CT2019B	3 × 2	25	3 h	http://carbontracker.noaa.gov (last access: 10 December 2020)		Jacobson et al. (2020)	
CarboScope	s10oc_v2020	5 × 3.8	19	6 h	http://www.bgc-jena.mpg.de/CarboScope/ (last access: 20 July 2020)		Rödenbeck et al. (2003)	
LSCE	v18r3	3.7 × 1.8	39	3 h	http://atmosphere.copernicus.eu (last access: 22 July 2020)		Chevallier et al. (2019)	
LSCE	FT18r1	3.7 × 1.8	39	3 h	http://atmosphere.copernicus.eu		Chevallier et al. (2019)	
Data used in the inverse model simulations								
Product	Version	Forward model	Meteorology	Observation data	Anthropogenic emission fields	Biospheric emission	Fire emission	Oceanic emission
CarbonTracker	CT2019B	TM5	ECMWF	Ground based	Miller and ODIAC	CASA	GFED and GFED CMS	OIF and Takahashi et al. (2009)
CarboScope	s10oc_v2020	TM3	NCEP	Ground based	EDGAR	LPJ biosphere model	CDIAC	SOCAT
LSCE/PyVar	v18r3	LMDz6A	ECMWF	Ground based	EDGAR, CDIAC, and GCP	ORCHIDEE 4.6.9.5	GFED and GFAS	Denvil-Sommer et al. (2019) with updates described in Friedlingstein et al. (2019)
LSCE/PyVar	FT18r1	LMDz6A	ECMWF	Satellite (OCO-2 NASA)	EDGAR, CDIAC, and GCP	ORCHIDEE 1.9.5.2	GFED and GFAS	Landschützer et al. (2017)

to account for the difference in sub-grid-scale process during these times. The representation error presented in Eq. (1) varies from one model time step to the next. In order to obtain a typical (average) representation error, we compute the monthly average representation error (σ_{CO_2}) using Eq. (2).

$$\sigma_{CO_2} = \frac{1}{T}\sum_{t=1}^{T}\sigma_{CO_{2(tot)}}, \tag{2}$$

where T is the total number of simulations in a month during daytime or nighttime. Furthermore, we have calculated the representation error ($\sigma_{\overline{CO_2}_{(mon)}}$) using Eq. (3), which describes the systematic component of representation error and provides important constraints for inversions using both ground-based and satellite observations over India.

$$\sigma_{\overline{CO_2}_{(mon)}} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(M_j - \overline{M}\right)^2}, \tag{3}$$

where $\overline{M} = \frac{1}{n}\sum_{j=1}^{n} M_j$. n is the number of 10 km boxes inside the coarser grid cell of $1° \times 1°$, M_j is the monthly averaged CO_2 dry-air mole fraction at a 10 km spatial scale, and $\overline{M}$ is the corresponding average within the coarser grid cell of 1°. The difference between Eqs. (1) and (3) is that we use monthly averaged CO_2 concentration values in Eq. (3) instead of hourly values as in Eq. (1). Both July and November are used to understand the differences in the variability during summer and winter.

Due to the paucity of adequate ground-level observations over India, satellite observations play an essential role in the estimation of CO_2 fluxes via inverse modelling (e.g. Philip et al., 2022). Satellite observations can provide the column-averaged CO_2 (XCO_2) concentration with a precision of 1 to 1.5 ppm (O'Dell et al., 2012; Wunch et al., 2017; Liang et al., 2017). In order to utilize these satellite observations, the transport models being used in the inverse estimation must be highly accurate. Since satellite footprints are smaller ($\sim$2–20 km^2) than the current model grid size (> 100 km), using these measurements for optimization via inverse modelling introduces spatial representation errors and associated uncertainties in the inferred fluxes. Note that the spatial biases of a few 10ths of parts per million in column-averaged CO_2 can potentially alter even the annual subcontinental fluxes in the range of 10ths of a gigaton of carbon fluxes (Chevallier et al., 2007; Miller et al., 2007; Chevallier et al., 2010). To quantify these systematic transport errors when representing satellite measurements in inverse models, we calculate the spatial representation errors for XCO_2 that coarse inverse modelling would suffer from using highly precise and accurate satellite measurements.

We have selected monsoon (July) and post-monsoon (November) periods for our analysis to identify the seasonal differences in the sub-grid variability over India. In July, many low-pressure systems were active in the monsoon trough region (India Meteorological Department, IMD, weather reports; https://mausam.imd.gov.in, last access: 20 November 2020). In general, tropical cyclones in the Asian monsoon region can cause fast uplift of air masses into the upper troposphere and lower stratosphere (e.g. Li et al., 2021), which may increase the modelling error due to the misrepresentation of the associated convective activity that is only parameterized in global models. The presence of enhanced biospheric activity during July can reduce the CO_2 concentration in the lower troposphere (e.g. Patra et al., 2011). Also, the strong vertical and horizontal mixing due to the monsoon circulation dilutes the CO_2 concentration in the atmosphere during July compared to November. The convective activity associated with the Indian summer monsoon was absent during November; however, the convection caused by synoptic systems such as tropical cyclones was still present. Such a low-pressure system activity was found over the Bay of Bengal and over the Lakshadweep area ($\approx$8° N, 74° E) from 22 November onwards. One of these low-pressure systems in the Bay of Bengal further developed and intensified as a deep depression and moved to the southeastern Arabian Sea and evolved into a severe cyclonic storm (Ockhi) by 30 November.

2.4 Estimation of representation-error-induced flux uncertainty using pseudo-surface measurements

In order to quantify the impact of representation errors in flux estimations when utilizing surface measurements, we have devised the following strategy. We used nine CO_2 surface monitoring sites representing various geographical regions in India (Fig. 1). Not all of these observation stations are currently fully operational or have continuous measurements. We have performed an observation system simulation experiment (OSSE) using high-resolution CO_2 simulations generated by the WRF-GHG model for each of these stations. We focus on the biospheric flux component (NEE). The simulated values of coarse models to compare with the observations are obtained from the nine grid cells of the coarse model covering these sites. The pseudo-observations for these sites correspond to the values simulated by the WRF-GHG model at one of the fine grid cells contained in one cell of the coarse model. Since there are 100 fine grid cells per coarse grid cell, 100 different time series are generated, and 100 corresponding inversions are performed to obtain robust results. For deducing the contribution of the representation error to the biospheric flux uncertainty, we have made the following assumptions: (1) the hourly WRF-GHG simulations at 10 km ($\sim$0.1°) spatial scale represents actual variations in CO_2 mixing ratios of the measurement site, (2) there are no model or observation errors other than representation error, (3) the model captures the spatial and temporal patterns of fluxes correctly, and (4) the contribution from other surface fluxes and background mixing ratio (in ppm) is known. As a

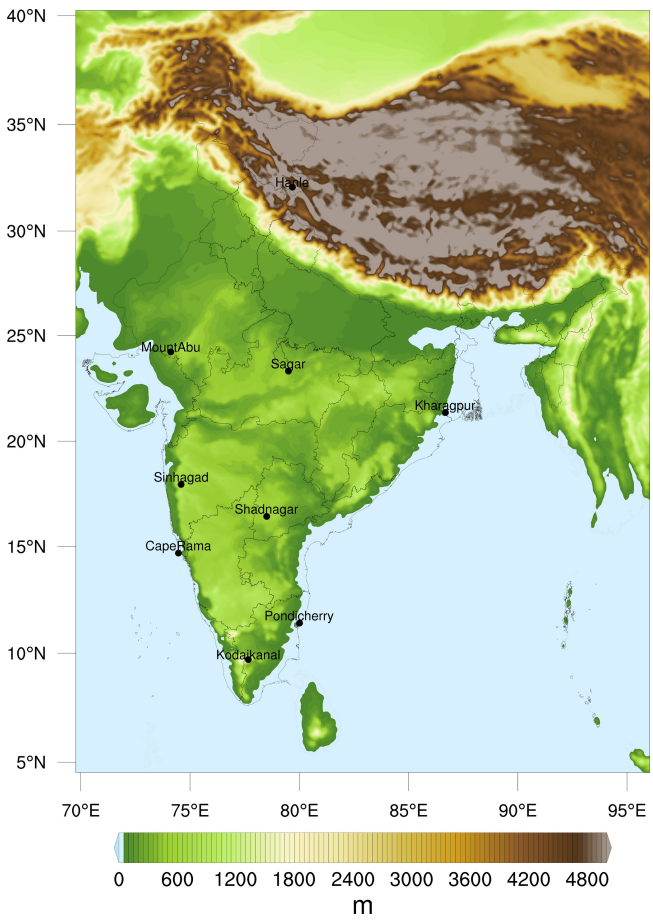

Figure 1. The WRF-GHG model domain used in this study, showing the topography. The CO_2 monitoring sites over India used for the OSSE experiments are marked. Not all these observation stations are currently fully operational. The colour scale is restricted to 5000 m for a better visualization of terrain details over the Indian subcontinent.

first-order simplification for the inversion, we assume that the footprints of each observation site span a radius of 200 km around the site based on our analysis using the Stochastic Time-Inverted Lagrangian Model (STILT; Lin et al., 2003). STILT footprints indicate that 50 % of the sensitivity of a site to fluxes over India is located in a region that has about the same area as a circle with a radius of 200 km. For nine stations, this footprint area covers around 35 % of the total area of India. The STILT is driven with ECMWF IFS (Integrated Forecasting System) meteorological fields and the trajectories are calculated based on 100 virtual particles that are released for each time interval and location. The residence time of the particles in the surface layer is weighted by the atmospheric density to derive the footprints of each location.

In our inversion set-up, we have used the hourly biospheric contribution of the atmospheric CO_2 mixing ratios simulated by WRF-GHG over the coarse grid cell of $1° \times 1°$ surrounding the location of each measurement site as OSSE observations ($m_{i,j}(t)$).

$$y_{i,j}(t) \equiv m_{i,j}(t) = \mathbf{H}_{i,j}(t) \cdot \mathbf{F}(\boldsymbol{\lambda}), \tag{4}$$

where $\mathbf{H}$ is the transport operator, and $\mathbf{F}(\boldsymbol{\lambda})$ is the flux model in which a subset of parameters $\boldsymbol{\lambda}$ out of total model parameters $\boldsymbol{p}$ will be optimized in the inversion. Here, i ($i = 1$ to 9) represents the nine observation sites and j ($j = 1$ to 100) is the number of WRF-GHG pixels inside the coarser grid cell of $1° \times 1°$.

The modelled biospheric CO_2 signal ($\overline{m}_i$) for the inversion is given by the following:

$$\overline{m}_i(t) = m_{i,j}(t) + \varepsilon_{i,j}(t). \tag{5}$$

The modelled values deviate from the observations by a representation error $\varepsilon_{i,j}(t)$. Since the modelled values ($\overline{m}_i$) correspond to the mean of the 100 fine grid cells, the simulated values at site i are given as follows:

$$\overline{m}_i(t) = \frac{1}{100} \sum_{1}^{100} m_{i,j}(t). \tag{6}$$

Here, $\mathbf{F}(\boldsymbol{\lambda})$ is taken as being linearly dependent on $\boldsymbol{\lambda}$; hence, it can be expressed as follows:

$$\mathbf{F}(\boldsymbol{\lambda}) = \boldsymbol{\Phi} \cdot \boldsymbol{\lambda}, \tag{7}$$

where $\boldsymbol{\Phi}$ is the biospheric flux (NEE) distribution over the region.

In the inversion, we retrieve the monthly NEE by utilizing hourly $m_{i,j}(t)$ and $\overline{m}_i(t)$ over a month. For OSSE and uncertainty flux estimation, we use the VPRM-derived NEE fluxes as the "true" fluxes (see Sect. 2.1.1). By using this inverse modelling design, we require the performance of 100 inversions per site, each of which uses a realization of the representation error to estimate the corresponding realization of the resulting uncertainty in the retrieved fluxes.

Both the observation and simulation vector have 6480 ($= 9 \times 30 \times 24$) elements for a month having 30 d, and the state vector has nine elements corresponding to scaling factors of fluxes for that month over regions around the nine sites (see Fig. 1). In other words, each site has been assigned with one scaling factor for NEE, and there is a total of nine scaling factors for a given month. We use a unit vector $\boldsymbol{\lambda}$ as the prior scaling factor. The prior uncertainty is neglected here, as the expected impact of the representation error in the retrieved fluxes is significantly smaller than the typical prior uncertainties assumed in Bayesian inversions (on the order of 50 %–100 % for biospheric fluxes). Hence, neglecting this prior uncertainty does not have a large impact on our results. The inversion retrieves optimized scaling factors $\boldsymbol{\lambda}_{\mathrm{retr}}$.

We have performed 100 inversions per site, and the scaling factors are retrieved by minimizing the cost function for each observation station as follows:

$$J\left(\boldsymbol{\lambda}_{i,j}\right) = \frac{1}{T} \sum_{t=1}^{T} \left(m_{i,j}(t) - \overline{m}_i(t)\boldsymbol{\lambda}_{i,j}\right)^2, \tag{8}$$

where T is the number of observations for a month. Minimizing these cost functions results in an optimized estimate of scaling factors $\boldsymbol{\lambda}_{\text{retr}}$, which is a vector of scaling factors with nine elements ($\boldsymbol{\lambda}_{\text{retr},i}$) for each of the 100 inversion cases.

By this inverse design, the deviation of posterior fluxes from the true fluxes over India is thus the uncertainty in retrieved fluxes, $\mathbf{S}_{\text{rep}}$, that arises solely due to the contribution from the representation error. The standard deviation of the scaling factors from these 100 inversions ($\boldsymbol{\sigma}_{\boldsymbol{\lambda}_{\text{retr}}}$) are used to retrieve flux uncertainty. $\mathbf{S}_{\text{rep}}$is obtained as follows:

$$\mathbf{S}_{\text{rep}} = \sqrt{\sum_{k=1}^{K} \left(\mathbf{S}_{\boldsymbol{\lambda}_{\text{retr},k}} \boldsymbol{\Phi}_{\text{true},k}\right)^2}, \tag{9}$$

where $\boldsymbol{\Phi}_{\text{true}}$ is the monthly VPRM biospheric flux (NEE) over the Indian region, and k is the number of pixels (33 141 pixels) over the Indian region. Here, $\mathbf{S}_{\boldsymbol{\lambda}_{\text{retr}}}$ has the dimension of Indian region at a 10 km spatial resolution and is defined in such a way that all the grid cells (at 10 km spatial resolution) other than the grid cells within the influence region (200 km radius around the station) of each station is given with zero values (21 335 pixels) and the grids in the influence region of each station (11 806 pixels) is given with the corresponding values of $\boldsymbol{\sigma}_{\boldsymbol{\lambda}_{\text{retr},i}}$. In this way, the approach does not depend on Eqs. (1) to (3) but shows the impact of difference between m_j and $\overline{m}$ on retrieved fluxes.

Any temporal correlations in the representation error are not considered for this experiment. We have performed the inversion separately for daytime and nighttime values to identify the impact of diurnal variations in representation errors in flux uncertainty. Note that, by following the above inversion design and assumptions, there is a high likelihood of underestimating the impact of the modelling error in flux estimations, since we have not considered other sources of uncertainties such as model transport uncertainty and inappropriate prior assumptions. Thus, the quantification of flux uncertainty using this approach may be considered to be the lower bound of the uncertainty (i.e. the minimum flux uncertainty one may expect while estimating fluxes using a model with a grid cell of $1^\circ \times 1^\circ$ and nine stations with the representativeness of 200 km).

3 Results and discussions

3.1 Agreement among global models

We first analyse the level of agreement among current-generation global transport models in simulating CO_2 concentration over the Indian subcontinent. Note that a mere agreement among the coarse models does not guarantee a good model performance over the region due to their plausibly large model errors in common and interdependency in terms of data sources. We restrict this analysis to daytime-only values, since different processes control the variability in CO_2 concentration at daytime and nighttime, and simulating nighttime variability is more complicated than the daytime (Lauvaux et al., 2009a). For a consistent comparison among global models, all the products are sampled at the same time for the region spanning from 67 to 98° E and 7 to 38° N. Figure 2a depicts the annual vertical profiles of CO_2 concentration, showing models' discrepancy in simulating the vertical gradients in concentration values including the boundary layer and the free troposphere. A notable difference is observed in the simulation of the gradient within the boundary layer. The magnitude and the height up to which this positive gradient is observed are different for these models. LSCE (both versions) has the largest positive gradient among these models (∼ 1 ppm). It shows the maximum concentration at around 700 m height and then a decrease in concentration. CarbonTracker also shows this positive gradient in the surface layers up to a height of 900 m. But the gradient is much smaller compared to the other two models. Among these four models, CarboScope does not exhibit this tendency in the lower atmosphere. Its concentration decreases linearly from the surface as the height increases.

The seasonal variability in CO_2 uptake through photosynthesis, release through ecosystem respiration, and vertical transport is seen when analysing the monthly averaged CO_2 concentration profiles over the Indian subcontinent (Figs. 2b and 3). Comparatively lower surface CO_2 concentrations are found during months with an active biosphere (June to October) than the rest of the period, owing to the higher ecosystem productivity over the Northern Hemisphere and particularly over the Indian subcontinent in response to the availability of monsoon rainfall. Also, the presence of strong southwesterly monsoon winds from June to September may result in bringing CO_2-depleted air from the Southern Hemisphere and thereby lowering the CO_2 concentration over the domain. While comparing the seasonal maximum (May) and minimum (September) of CO_2 concentrations measured at the Mauna Loa Observatory (MLO) located in Hawaii, Fig. 2b shows a temporal shift of around 1 month for exhibiting the seasonal maximum (April) and minimum (August) CO_2 concentrations. This temporal shift is attributed to the differential impacts of anthropogenic and terrestrial ecosystem activities on atmospheric concentration in addition to the long-distance transfer of atmospheric carbon dioxide to a remote location (Nomura et al., 2021). MLO observations are generally representative of global mean CO_2 due to the minimal influence of terrestrial ecosystems and anthropogenic activities at this remote location. The seasonal variation in monthly averaged CO_2 seen over the Indian subcontinent is mostly dominated by terrestrial carbon fluxes, i.e. net ecosystem exchange (NEE) as seen from the VPRM simulations (see Fig. S2) and as, e.g., in Tiwari et al. (2013).

Furthermore, we see a CO_2 vertical profile with a small vertical gradient (∼ 0.5 ppm within an altitude range of ∼ 500 to 4000 m) from June to October (Fig. 3). This is likely linked to the increased convective activities associated with

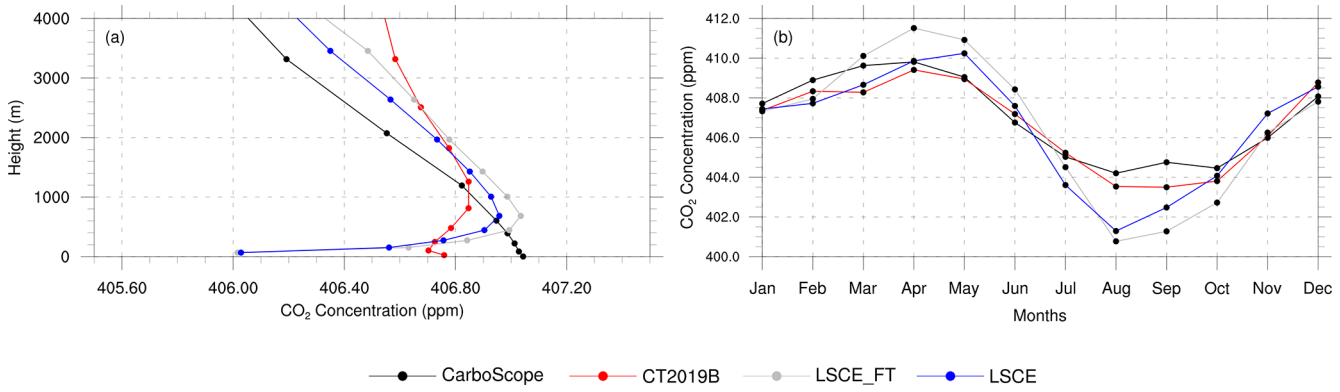

Figure 2. Comparison of global models over the model domain during daytime (11:30 to 16:30 LT) in 2017. **(a)** Annually averaged vertical profiles of CO_2 concentration in the lower troposphere. **(b)** Time series of monthly averaged CO_2 concentration at surface (∼ 100 m above surface).

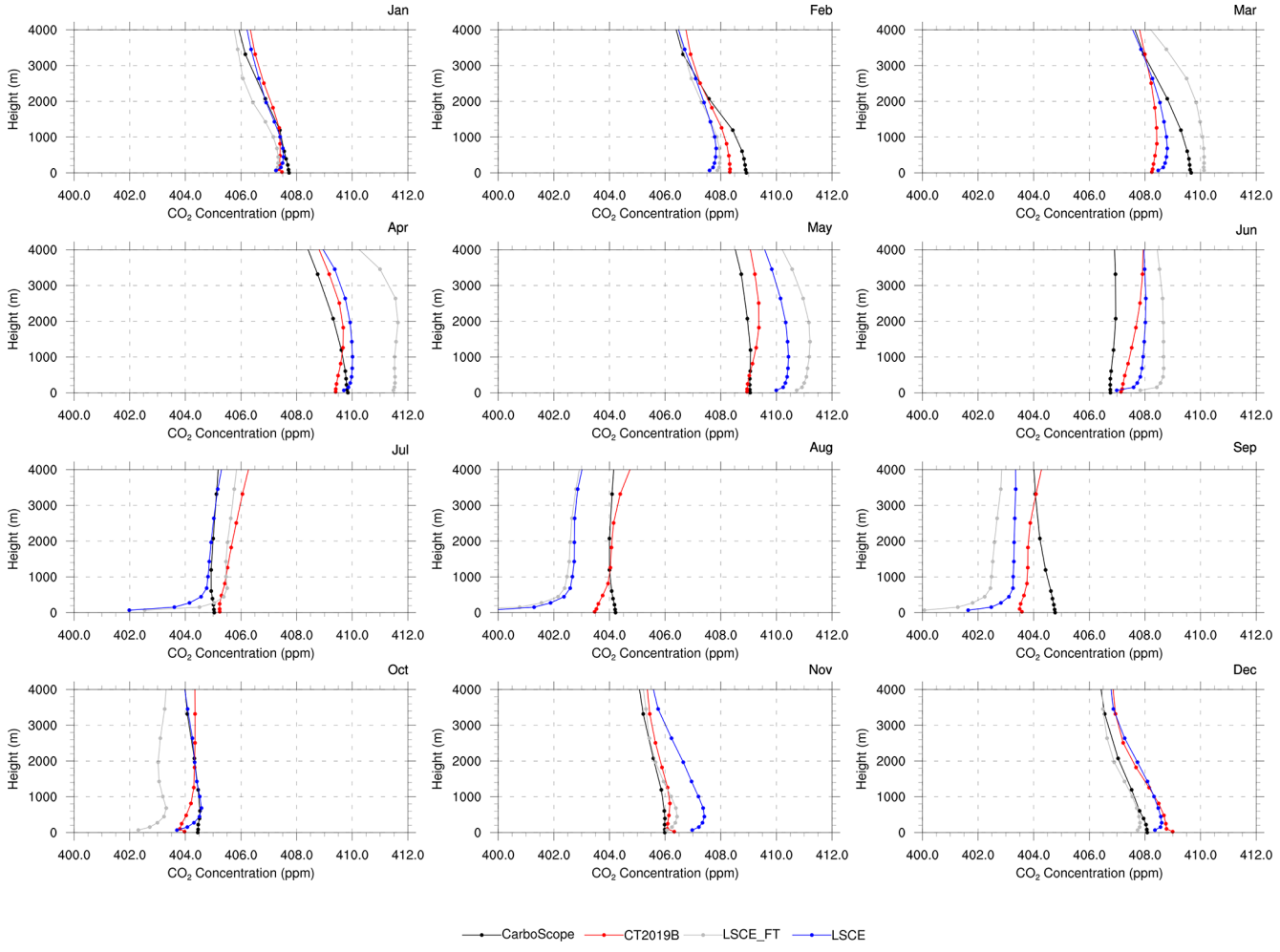

Figure 3. Comparison of average monthly vertical profiles of CO_2 concentration from global atmospheric transport models over the model domain during daytime (11:30 to 16:30 LT) in 2017. Panels show data for the respective months, as indicated at the top of each panel.

the monsoon. The strong vertical gradient in the surface levels as simulated by the LSCE model during the monsoon period is little plausible given the strong vertical mixing expected due to the presence of strong wind and convection. The considerable intermodel variation in monthly averaged CO_2 concentration profiles, as predicted by different global models, is problematic as it indicates significant uncertainties in flux estimations over India. Part of this discrepancy can come from the coarse-resolution global model's inability to represent transport processes like convection and vertical mixing, in addition to the strength and distribution of anthropogenic sources and ecosystem activities, that operate at fine scales. The extent of this unresolved variability in existing global models is further explored in Sect. 3.2. The spatial distribution of CO_2 concentration shows structural differences among these models (see Fig. S3), which indicates a substantial knowledge gap in models for representing atmospheric CO_2 variability over the Indian subcontinent. As a consequence, the country's carbon budget estimations inferred via inverse modelling can be unreliable.

3.2 Representation errors in global transport models

The spatiotemporal variability in representation error and the influence of various factors in creating this variability are examined here. The larger the variations that are caused due to sub-grid processes within the grid box of $1^\circ \times 1^\circ$, the larger the representation error. The derived seasonal differences in structural patterns of the sub-grid variability facilitate to (1) quantify what would be typical representation errors associated with incorporating observations from different seasons into atmospheric models, (2) determine what drives the seasonality in sub-grid variability and ultimately, and (3) design a possible parameterization of representation error with a seasonal component in the inverse modelling framework, as well as identify periods or seasons in which the use of this parameterization would be valid to improve our estimations of CO_2 fluxes. Furthermore, the seasonal spatial variability analysis of column averages can provide useful information to gap-fill the satellite-based products over India when large data gaps are present, which can be utilized for applications that do not demand high-precision observations (e.g. Hammerling et al., 2012).

3.2.1 Spatiotemporal patterns

Representation errors in the surface CO_2 concentrations of a global model at a spatial resolution of $1^\circ \times 1^\circ$ for July and November are shown in Fig. 4. The representation error at $1^\circ \times 1^\circ$ spatial scale reaches values ranging from 0.5 to 5 ppm, which are comparable to the magnitude of variability at hotspot emission regions or half of the seasonal variability in CO_2 over the region (see Fig. 2b). The median representation error is 1.2 ppm at the surface, which is considerably larger than the measurement errors. In the case of high-accuracy in situ measurements, the typical uncertainty for CO_2 measurements is of the order of 0.1 ppm (Andrews et al., 2014; Zellweger et al., 2016). A remarkable feature is the presence of very high representation error over the Northeastern and Western Ghats regions, where biosphere activity is very prominent. The heterogeneous distribution of biosphere fluxes generates significant sub-grid-scale variability that leads to high representation error. Also, we can find high representation error along the foothills of the Himalayas. In addition to the complex terrain, the region over the Ganges basin is characterized by increased anthropogenic activity, which contributes to a larger representation error surrounding in this region. High representational error is also found in the coastal regions, ranging from 2 to 5 ppm (median of 4 ppm), due to the temporal covariance between the coastal meteorology and exchange fluxes. The CO_2 fluxes from coastal regions can be transported over the ocean and accumulated in the shallow boundary layer over the ocean. The shallow boundary layer is a characteristic of the marine atmosphere due to less vertical mixing as compared to land regions. Horizontal CO_2 gradients can also be generated by the influence of highly varying biospheric fluxes under different advection patterns at the boundary between land and ocean. A similar mechanism is applicable to mountain regions, where the temporal covariance of mountain–valley circulation and respired CO_2 fluxes are regulated by atmospheric radiation. The terrain-following coordinates, as used in the model, may also result in spurious tracer concentration gradients over the steep mountain terrain (Beck et al., 2020; Skamarock et al., 2021; Park et al., 2019). Though the mesoscale models are expected to perform better in simulating CO_2 variations over the complex terrain than the coarse models (e.g. Engelen et al., 2002; Gerbig et al., 2003; Ahmadov et al., 2007; Corbin et al., 2008; Lauvaux et al., 2009b; Pillai et al., 2011; Uebel et al., 2017; Agustí-Panareda et al., 2019), they may also suffer from the inadequate representation of complex weather features and associated variability. We can also find individual cells with high representation errors associated with point emission sources such as cities, mining sites, and coal-fired power plants at different parts of the domain. The daily variations in surface representation errors are small (within a month), although there exists a clear distinction between daytime and nighttime values (figure not shown). The nighttime representation error is higher (e.g. a median value of 1.5 ppm for surface during November) compared to the daytime representation error (e.g. a median value of 1.1 ppm for surface during November) throughout the analysed domain. This is expected due to the coupling between nocturnal shallow transport and different flux processes accentuating local effects. During the nighttime, photosynthesis is absent, and respiration is the major biospheric activity, leading to an increase in CO_2 concentration in the atmosphere. The large heterogeneity in flux distribution that is mostly from respired CO_2 fluxes, the shallow boundary layer processes, and the weak nocturnal turbulence cause CO_2 to be accumulated lo-

cally near the surface with large variations. Compared to July, we find higher representation error in November owing to the wintertime transport with decreased vertical mixing and heterogeneous biospheric uptake (see Fig. 4).

In the case of XCO_2, the magnitude of sub-grid-scale variability is much smaller than that of surface CO_2 (Fig. 5), but it follows a similar spatial pattern. This confirms the dominance of surface-level processes in causing sub-grid variability in column averages. The sub-grid-scale variability in XCO_2 reaches up to 2 ppm in some parts of the region, especially where there are high variations in topographic features or point emission sources. The estimated column representation errors over these regions are thus capable of causing significant biases in the satellite inferred CO_2 fluxes as regional biases of a few 10ths of parts per million in column-averaged CO_2 can create a bias of a few 10ths of a gigaton of carbon fluxes (Chevallier et al., 2007). Also, the representation error for a large part of the domain is found to be above 0.5 ppm, which is around half of the typical precision of current satellite measurements. Note that the representation error reported here is different from satellite measurement errors (e.g. spectroscopic retrieval error or sampling biases) and tends to be systematic in nature.

Figure 6 shows the statistical distribution of the representation error (σ_{CO_2}) sampled over India, during July and November, separated by daytime and nighttime. July shows a median surface representation error of 0.9 and 1.1 ppm during daytime and nighttime, respectively, while November shows a median value of 1.1 and 1.4 ppm for daytime and nighttime, respectively. In July, 95 % of the representation error is less than 2.1 ppm for daytime (3.9 ppm for nighttime), while it is 3 ppm for daytime (4.2 ppm for nighttime) for November. For the column average, median values for representation error are 0.3 and 0.4 ppm for July daytime and November daytime, respectively.

To further reduce the effect of random error that might be introduced by short-term weather phenomena, the representation errors ($\sigma_{\overline{CO_2}_{(mon)}}$) are calculated from the monthly averaged CO_2 field and are denoted as a systematic error (Fig. 6). Uncorrelated errors are expected to decrease when averaging over a sufficiently long period. As expected, the median values of the systematic representation errors are smaller for all cases, showing the effect of random errors. Especially for November, when the cyclonic event was present, the values of the systematic errors (in the 95th percentile) for the surface CO_2 are considerably lower than total errors, reducing from 3 ppm (daytime) and 4.2 ppm (nighttime) to 2.2 ppm (daytime) and 3 ppm (nighttime). In the case of column CO_2, this reduction is from 1.1 ppm (daytime) and 0.9 ppm (nighttime) to 0.8 ppm (daytime) and 0.7 ppm (nighttime) in the 95th percentile. In contrast to surface representation error (Fig. 6a), median values of nighttime representation errors are found to be slightly lower than the daytime representation error for the column average (Fig. 6b). To assess the dependence of representation error in possible horizontal resolutions of the global models, we have further derived the representation errors for different spatial resolutions between 0.5° and 4°. As expected, we see reductions in representation errors for both surface and column-averaged CO_2 with the increasing horizontal resolution of the model (see Figs. 7 and S4). During July, the median surface representation error reduced from 1.6 ppm (2 ppm) to 0.6 ppm (0.7 ppm) during the daytime (nighttime), while increasing the horizontal resolution from 4° to 0.5°. This increment in the spatial resolution has also resulted in similar error reductions in November during which the median of surface representation error shows a reduction from 2.4 ppm (2.8 ppm) to 0.7 ppm (0.9 ppm) during daytime (nighttime). In the case of the column-averaged values, the median representation error decreased from 0.7 ppm (0.6 ppm) to 0.25 ppm (0.2 ppm) during the July daytime (nighttime) and from 0.95 ppm (0.9 ppm) to 0.25 ppm (0.2 ppm) during the November daytime (nighttime). The spatial distribution of the representation errors for a model with a horizontal grid resolution of $0.5° \times 0.5°$ (e.g. regional models) is provided in Figs. S5 and S6. On average, we find a ~ 33 % to 36 % decrease in daytime representation errors for both months when increasing the model grid resolution from 1 to 0.5°. There exists a similar spatial pattern of representation errors for both resolutions of 0.5 and 1°. Though our results indicate a reduction in the representation error for regional models with a typical resolution of 0.5°, compared to global models with 1° spatial resolution, the emission hotspots and point sources are still pronounced with high sub-grid-scale variability, especially during nighttime. The above analyses indicate that the sub-grid variability alone can produce significantly higher errors compared to the measurement errors (e.g. 0.1 ppm as per World Meteorological Organization standards for surface measurements), which necessitates a proper treatment of these errors in models for the optimal estimation of CO_2 fluxes.

3.2.2 Vertical distribution

Figure 8 shows the vertical profile of representation error distribution within different altitude bins. We find that the maximum representation error is in the surface layer, and most of the higher values are found to be within the lowest 4–6 km bins. Also, sub-grid-scale variability decreases sharply with increasing altitude. This dominance of the variability in the surface concentration can be explained by surface flux heterogeneity influencing mole fractions in lower atmospheric layers (PBL), as described in van der Molen and Dolman (2007) and Pillai et al. (2010). There is a slight increase in the representation error in the upper tropospheric levels near the 12 to 14 km altitude range. This may be associated with the presence of strong circulations in the upper troposphere and lower stratosphere, such as subtropical westerly jets or the Asian summer monsoon anticyclone (e.g. Chandra et al., 2017).

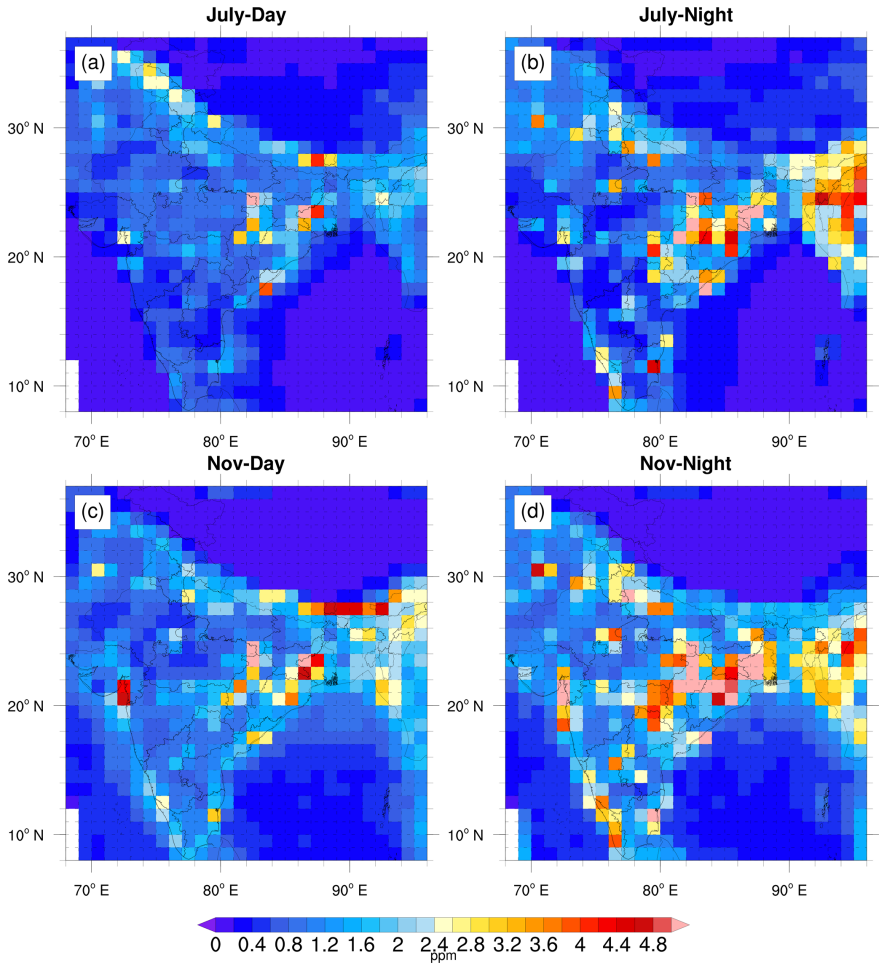

Figure 4. Monthly averaged values of representation error estimated for surface CO_2 concentration (second model level; mean height is ∼200 m from sea level) over the region 8 to 37° N and 68 to 96° E during 2017. **(a)** July daytime (11:30 to 16:30 LT). **(b)** July nighttime (23:30 to 04:30 LT). **(c)** November daytime. **(d)** November nighttime.

3.3 Influence of terrain heterogeneity and flux variability on representation errors

Here we explore the factors influencing the size and patterns of the representation error in coarse models. For this, statistical relationships between the representation error and possible explanatory variables are examined for both surface and column-averaged CO_2. Identifying these factors influencing representation errors and quantifying their local effects allow us to further investigate on how these biases in retrieved fluxes can be minimized in global models (see Sect. 3.5).

We find a significant influence of terrain heterogeneity on the representation error, which is evident from the spatial maps in Figs. 4 and 5, where the largest sub-grid-scale variations are found in the Himalayan regions. Spatial variations in topography produce mesoscale circulation patterns and corresponding variations in atmospheric CO_2 at fine scales. At the same time, there is a plausible additional error in global model simulations related to the insufficient resolution of vertical grids necessary to account for different surface influences (e.g. mountain vs. valley). This effect of the coarse vertical resolution is excluded in our representation error estimates by preserving the vertical grids used for the high-resolution simulations. To further explore the importance of using the high-resolution topography data to represent the CO_2 variability, we analyse the dependence of terrain variations (as derived from the standard deviation of the terrain height) on the distribution of the representation error. We have estimated the statistical dependence (R^2) of the representation error in topographic variability within the corresponding global climate models' grids to understand the relation between them. Topographic variability within a 1° × 1° spatial box is estimated as being the standard deviation of topography (m) for all 10 km × 10 km boxes within the larger grid and is denoted as σ_{topo}. Bins are created based on the values of this topographic variability, in which different points from different parts of the domain are binned together on the basis of their standard deviation of topography. Each bin is

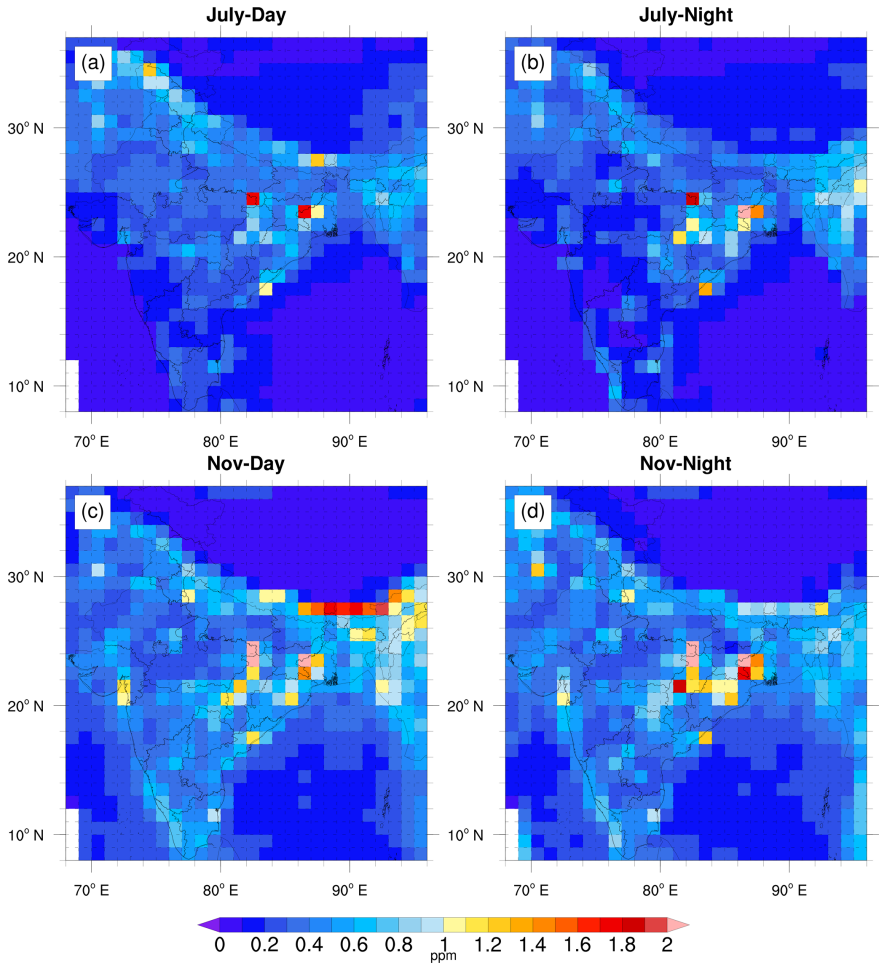

Figure 5. Monthly averaged values of representation error estimated for column-averaged CO_2 concentration over the region 8 to 37° N and 68 to 96° E during 2017. **(a)** July daytime (11:30 to 16:30 LT). **(b)** July nighttime (23:30 to 04:30 LT). **(c)** November daytime. **(d)** November nighttime.

created with a size of 50 m variation in terrain height. The linear fit is estimated between the average value of topographic variability within a bin and the average value of the representation error in the corresponding points in the bin. Our results show that the terrain heterogeneity alone can explain about 20 %–48 % ($R^2 = 0.20$ to 0.40) of the surface representation errors over the domain. In a similar way, we have estimated the influence of topographic variability on the representation error in the column-averaged model simulations. It is found that topography alone can explain 45 %–52 % ($R^2 = 0.45$ to 0.52) of the representation errors in the column-averaged simulations.

Furthermore, we estimate the statistical relationship (R^2) between the surface flux heterogeneity and representation error. The surface representation error is strongly linked to the biosphere flux variability, and the relationship between heterogeneity in biospheric surface flux (as derived from the standard deviation of VPRM-derived NEE fluxes, denoted as σ_{bio}) and representation errors depends on the time of the day and season. During daytime, when there is strong ecosystem activity, the dependence of representation error (σ_{CO_2}) on σ_{bio} of surface and column CO_2 is found to be ∼ 75 %–80 % and ∼ 66 %–74 %, respectively. σ_{bio} explains about 62 % of the surface CO_2 variability and 48 % of the column variability during the July nighttime. However, σ_{CO_2} and σ_{bio} are less correlated (23 % for surface and 19 % for column) during the November nighttime. The diurnal difference in the dependence of representation error in σ_{bio} can be explained by the increased magnitude and spatial variability in daytime biospheric fluxes in the growing season (primarily due to photosynthesis activities) compared to nighttime fluxes. Moreover, poor vertical mixing under the stable nocturnal atmospheric conditions with more advection and drainage flow reduces the influence of surface fluxes on spatial variability in mixing ratios. The dependence of the representation error in the anthropogenic flux heterogeneity (as derived from the standard deviation of EDGAR fluxes, denoted as σ_{ant}) is found to be negligible, except for the nighttime (13 %–30 %). We find

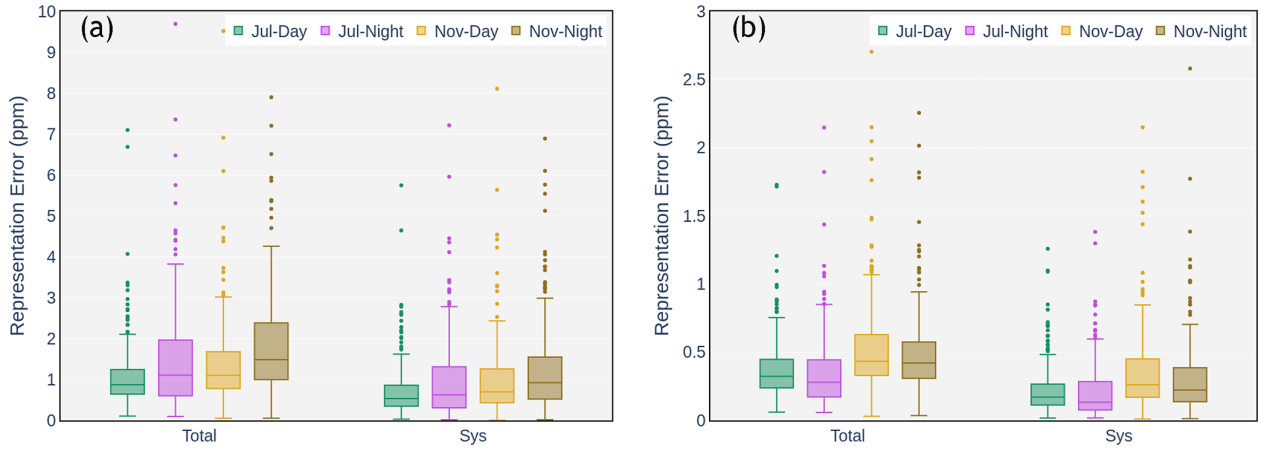

Figure 6. Variability in derived representation error over India in July and November 2017 (both during daytime and nighttime). Boxes indicate the central 50 %, the bar across the box is the median value, and the whiskers indicate the values between the 5th and 95th percentiles. Individual data points shown are the outliers. **(a)** Representation error estimated for the surface CO_2. **(b)** Representation error estimated for the column-averaged CO_2.

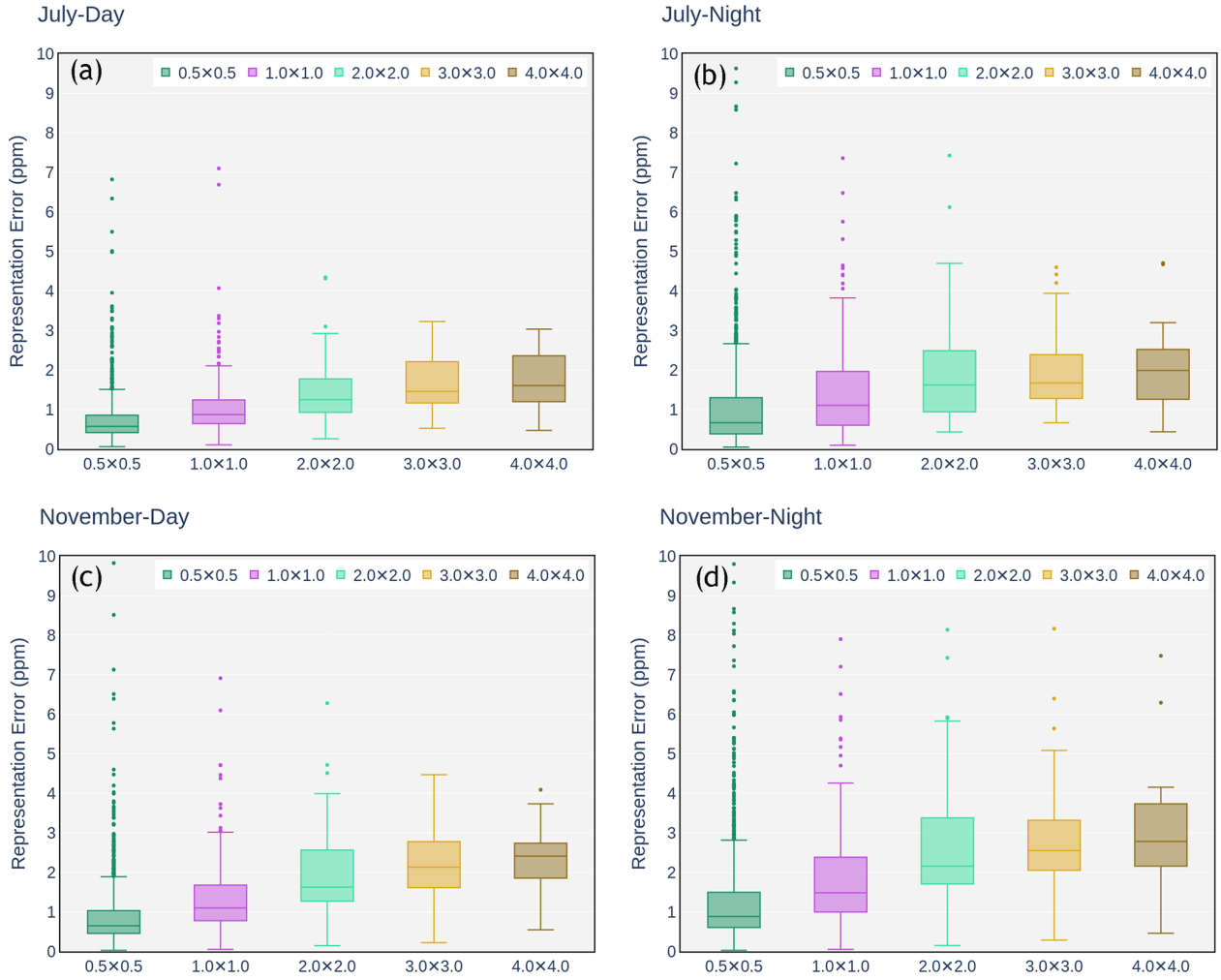

Figure 7. Variability in derived surface representation error over India for different horizontal resolutions. Boxes indicate the central 50 %, the bar across the box is median value, and the whiskers indicate the value between the 5th and 95th percentiles. Individual data points shown are the outliers. **(a)** Representation error estimated for July daytime. **(b)** July nighttime. **(c)** November daytime. **(d)** November nighttime.

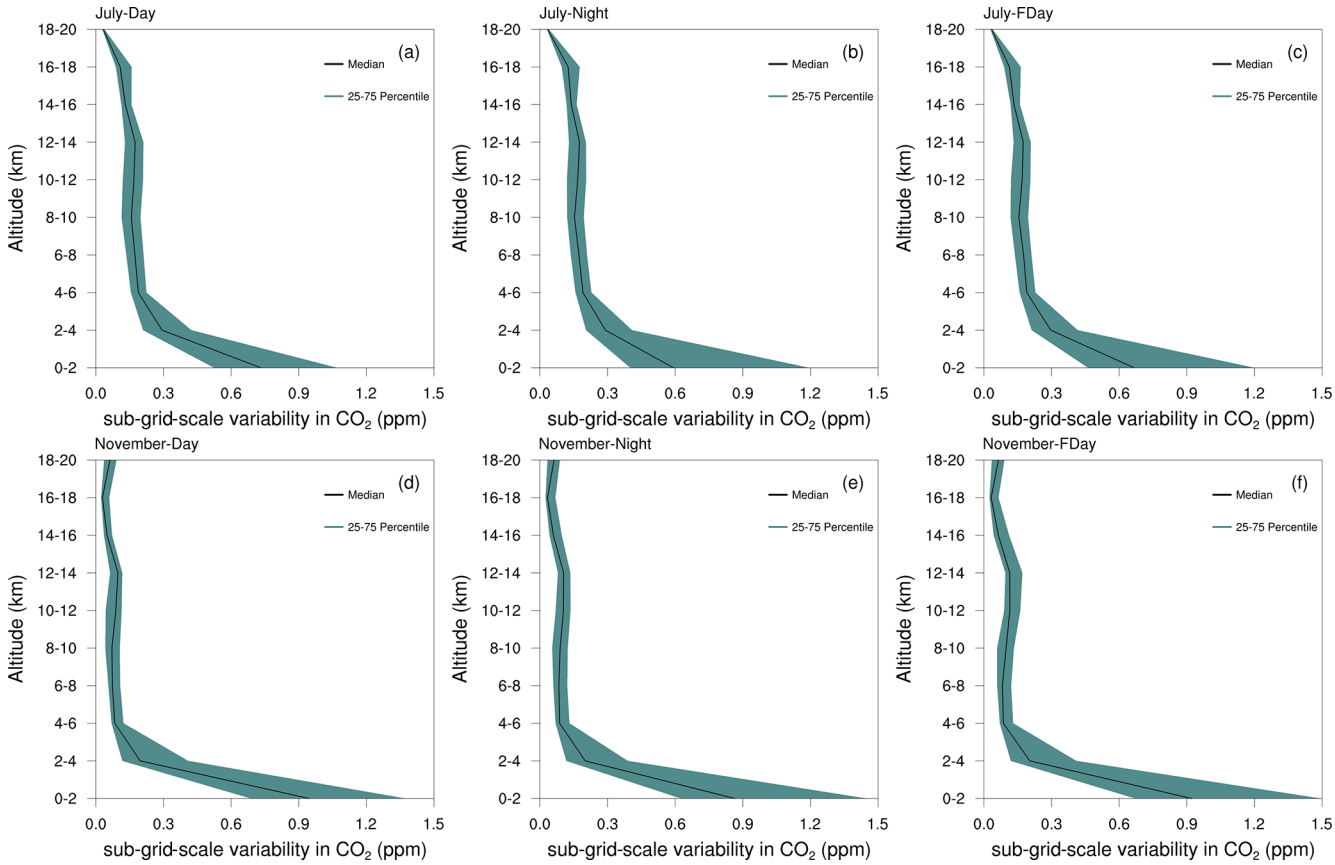

Figure 8. Variability in the representation error over India with altitude for July and November 2017. **(a)** July daytime, **(b)** July nighttime, **(c)** July full time, **(d)** November daytime, **(e)** November nighttime, and **(f)** November full time. Median values are plotted with black curves, and the shaded region indicate the 25th to 75th percentiles of data.

a smaller influence of seasonality on the relationship between anthropogenic surface flux heterogeneity and representation errors (see Table S1). Similar to the above analysis with σ_{bio}, the combined effect of atmospheric stability and flux heterogeneity can explain the diurnal differences in the relationship between σ_{ant} and σ_{CO_2}.

In the case of the variability in monthly averages, we see that $\sigma_{\overline{CO_2}_{(mon)}}$ is well explained by σ_{bio} during the daytime (see Table S2), as expected. A similar strong correlation can be seen between $\sigma_{\overline{CO_2}_{(mon)}}$ and σ_{bio} (23 %–69 %) during nighttime for surface variability in CO_2, while there exists only a lesser dependence of the nocturnal column CO_2 variability on local fluxes. This shows the decoupling of the mixing ratios in other parts of the column from the surface during the night, due to less vertical mixing, combined with more drainage flow in the nocturnal boundary layer, which reduces the effect of surface flux variability on the column CO_2 variability.

In general, the above analysis underlines the need for using digital elevation models (DEMs) at a high resolution to take into account the terrain-induced mesoscale atmospheric flows adequately in atmospheric transport models. Furthermore, the results indicate the importance of utilizing high-resolution surface fluxes in atmospheric CO_2 simulations.

3.4 Estimation of NEE flux uncertainty due to representation error

By following the assumptions and approach as given in Sect. 2.4, we have estimated the NEE flux uncertainty resulting from the representation errors. The results based on the OSSEs for nine observation sites are given in Table 3. The scaling factors, which are calculated separately for each site by adjusting the prior fluxes using pseudo-observations, are applied to the VPRM monthly fluxes. The total NEE flux for India estimated by VPRM for July and November are −373.3 and −417.1 Mt CO_2 per month, respectively. The flux uncertainties over India that arise solely due to the contribution from the representation error are estimated to be 38.59 (daytime observations) to 30.14 Mt CO_2 per month (nighttime observations; 10.33 % to 8.07 %) for July and 18.42 (daytime observations) to 13.34,Mt CO_2 per month (nighttime observations; 4.4 % to 3.1 %) for November, while utilizing data from nine observation stations. The

spatial representativeness of measurement stations used in this study is assumed to cover only 35 % of the country's total area (see Sect. 2.4). Consequently, the impact of representation error in flux uncertainty, as reported in this study, is an underestimation when measurements from more regionally representative sites or a dense observation network are utilized in inversions. The maximum flux uncertainty was found for July due to the enhanced biosphere activity and unresolved convection activities. The estimated uncertainties are considerable for the carbon budget assessment, especially given that these errors arise solely from the global models' representation error. Note that calculated representation error does not include other transport error sources such as advection, convection, or vertical mixing.

3.5 Possible treatment of representation error in the global model

The simplest possible way to minimize the uncertainty in the flux estimation using a coarse model is to construct a parameterization model that can account for the representation error using explanatory variables. For this, we create a multivariate model to capture spatial patterns in the representation error. Employing this parameterization in a global model would thus redefine the likelihood of better estimates (improving the state of knowledge) with a variance greater than that of the measurement error in the inverse framework by minimizing the modelling error. The multivariate linear model, with explanatory variables that includes sub-grid variations in terrain (σ_{topo}), biospheric (σ_{bio}) and anthropogenic (σ_{ant}) fluxes, remarkably captures the derived column representation error all over the Indian region during the July daytime, with a R^2 value of 0.96 (Fig. 9). The difference between the modelled and derived representation error is found to be well below 0.5 ppm in most parts of the domain. Similarly, we have modelled the surface representation error using the linear model with these three explanatory variables and found that the proposed model could capture the derived surface representation error well ($R^2 = 0.89$) with a deviation of less than 1 ppm in most of the regions (see Fig. S7 and Tables S1 and S2). More work is needed to demonstrate the extent of applicability of this method to minimize the flux uncertainties while utilizing actual observations. Nevertheless, the above finding provides a possibility for a parameterization that can be further developed in inverse models or data assimilation systems, which defines the degrees of freedom for describing the posterior states. Applying this parameterization scheme to the specific problem requires a high-resolution map of the terrain and prior information on anthropogenic and biogenic fluxes as the uncertainties in the topography and surface fluxes can significantly impact flux estimation. The caveat of this linear model is that the uncorrelated spatial variability in the prior and true states of the fluxes is ignored in the present form, which cannot be the case for the real inverse problems. This assumption obviously hampers the system in achieving the maximum reduction in uncertainty, and further study is needed to refine this model from a practical perspective. We emphasize, however, that the above parameterization does not require a high-resolution simulation of transport, which has high computational costs.

4 Conclusion

Given the upcoming availability of atmospheric observations over India, significant effort is required to critically enhance the modelling capabilities to derive carbon budgets over India within the definite confidence intervals and at scales relevant to the ecosystem and countrywide policy-making. The misrepresentation of mesoscale transport phenomena and unresolved flux variations in modelling systems operating on coarse grids hinders the optimal utilization of observations. In this context, the present study quantifies the spatial variability in atmospheric CO_2 mixing ratio over India that is not resolved by the coarse models and assesses their impact on flux estimations. We demonstrate the potential of a simple parameterization scheme to model these unresolved variations in the coarse models for minimizing the uncertainty in retrieved fluxes.

A large spread among existing global model simulations in representing the monthly averaged CO_2 concentration profiles indicates a considerable knowledge gap in the estimations of fluxes even at a monthly scale. It can be argued that a significant part of these differences arises due to the lack of observational constraints over India, which leads to a possible compensatory model artefact over this region in order to match the global mass constraint. At the same time, it is also expected that the spatial variability in the observed atmospheric CO_2 mole fractions can be large, so that these coarse models fail to represent them adequately. For instance, we find that the unresolved variations (representation error) in global models with a spatial resolution of $1^\circ \times 1^\circ$ can be ~ 1.5 ppm on average for the surface CO_2, which is even larger than the currently reported differences between global models (~ 1 ppm). Similarly, the average representation error estimated for the column-averaged CO_2 is ~ 1.1 ppm. These estimated values are larger than the corresponding measurement errors, which cause the inverse optimization to infer a state that is not close to the truth, as is required in the regional CO_2 budget for various applications.

Coastal areas and mountains have particularly high representation errors (≈ 4 ppm for surface CO_2). Emission hotspots can also lead to significant CO_2 variability near the surface as large as ≈ 8 ppm. Larger values are typically associated with the nocturnal shallow boundary layer dynamics and the stronger respiration signal with considerable flux variability. These findings are consistent with Pillai et al. (2010), which show that there exist spatial differences in the sub-grid variability for both surface and column CO_2. Although the magnitude of the sub-grid variability in the

Table 3. Flux uncertainty over India calculated from the OSSE experiments using a pseudo-observation network of surface observations. The time filter indicates the time of the data sampled for estimation of the scaling factors. A full day is 24 h in each day, daytime is from 11:30 to 16:30 LT, and nighttime is from 23:30 to 04:30 LT. Note: the fraction of uncertainty to the total NEE is given in parentheses.

Month	Time filter	True flux, aggregated over India $\sum_{k=1}^{K} \Phi_{true}$ (Mt CO_2 per month)	Flux uncertainty $\mathbf{S}_{rep}$ (Mt CO_2 per month) (%)
July	Daytime observations	−373.31	38.59 (10.33)
July	Nighttime observations		30.14 (8.07)
July	Full-day observations		23.20 (6.21)
November	Daytime observations	−417.12	18.42 (4.4)
November	Nighttime observations		13.34 (3.1)
November	Full-day observations		13.48 (3.2)

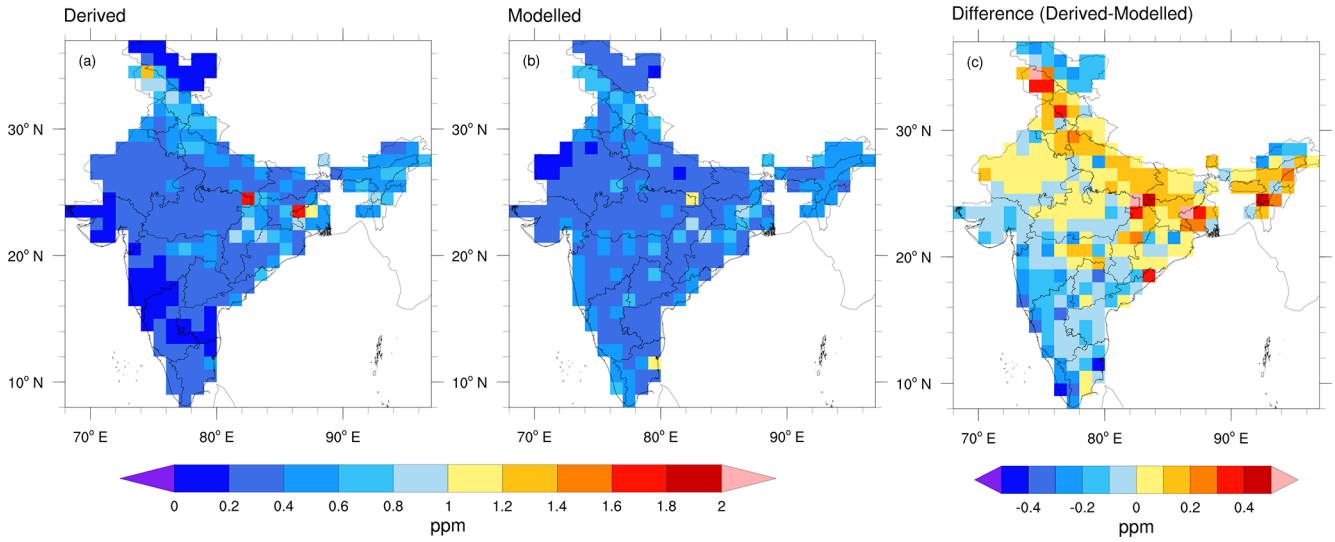

Figure 9. Monthly averaged values of representation error estimated for column-averaged CO_2 concentration during the July daytime (11:30 to 16:30 LT) in 2017. **(a)** Representation error derived from WRF-GHG simulations, as explained in Sect. 2.3. **(b)** Representation error calculated from the multivariate linear model, as described in Sect. 3.5. **(c)** Difference between panels **(a)** and **(b)**.

total column is significantly smaller than the variability at the surface, the spatial pattern remains similar for both, owing to the dominance of surface heterogeneity in topography and fluxes. With the underlying assumptions, the total uncertainty in optimized fluxes solely due to the unresolved sub-grid variations is estimated to be 3.1 % to 10.3 % of the total NEE, while utilizing pseudo-data from nine observation stations over India. Increasing the spatial and temporal resolutions of the transport models can generally capture the mesoscale features and associated CO_2 gradients, thereby reducing the representation error. Increasing the model resolution from 1 to 0.5° has shown an improvement in capturing variability with representation error reduction of 33 % and 36 % for summertime and wintertime, respectively. By showing the existence of the unresolved variability in 0.5° resolution with a similar spatial pattern of the error as of 1° spatial resolution, we demonstrate the need for a much finer resolution than 0.5° for representing the atmospheric CO_2 variability over India. However, merely increasing the resolution without having a realistic representation of terrain heterogeneity and flux (both natural and anthropogenic) variability would not be beneficial. The uncertainties in the high-resolution fluxes can worsen the model's skills, whose effect would not be more pronounced at coarser resolutions due to the diffusive nature of fluxes, as seen in Agustí-Panareda et al. (2019).

A parameterization scheme with explanatory variables of sub-grid variations in terrain, biospheric, and anthropogenic fluxes is shown to capture a considerable fraction of expected representation error in the global model. The proposed method is easy to implement in the coarse models, as it does not require computationally expensive transport simulations at high resolution. As we see a significant dependence of the distribution of sub-grid variability on terrain varia-

tions, our results reinforce the requirement for using high-resolution DEMs in the atmospheric transport models. The biosphere flux variability explains about 62 % to 80 % of the surface representation errors over the Indian region, indicating the need for using precise high-resolution surface fluxes.

Overall, we show that the mesoscale transport mechanisms and flux variability contribute to fine-scale CO_2 variations that the current-generation models cannot resolve. Our findings indicate that the models need to be critically improved to capture mesoscale variations associated with horizontal and vertical transport and fine-scale flux variability to maximize the potential of highly precise and accurate measurements. Our results provide a baseline for overcoming the shortcomings mentioned above and account for the realistic distribution of atmospheric CO_2 to improve the estimation of surface fluxes through inverse modelling.

Code and data availability. The WRF-Chem source code is publicly available at https://doi.org/10.5065/D6MK6B4K (NCAR, 2017). The CarbonTracker (CT-2019B) products are available online at https://doi.org/10.25925/20201008 (Jacobson et al., 2020). The data from the CarboScope inversion system are available online at http://www.bgc-jena.mpg.de/CarboScope/ (last access: 20 July 2020, MPI-BGC, 2020; Rödenbeck et al., 2003). The data from the LSCE modelling system used in this study are available at https://apps.ecmwf.int/datasets/data/cams-ghg-inversions/ (last access: 22 July 2020, CAMS/LSCE, 2019a, b; Chevallier et al., 2019). The sub-grid variability products based on the WRF-GHG model simulations can be accessed from https://doi.org/10.5281/zenodo.6616466 (Thilakan and Pillai, 2022). The WRF-GHG model CO_2 simulations used for this study are available upon request to the corresponding author, Dhanyalekshmi Pillai (dhanya@iiserb.ac.in; kdhanya@bgc-jena.mpg.de). The EDGAR data used in this study are publicly available at http://data.europa.eu/89h/97a67d67-c62e-4826-b873-9d972c4f670b (last access: 20 May 2022, Crippa et al., 2021, 2018). The GFAS data are publicly available at https://apps.ecmwf.int/datasets/data/cams-gfas/ (last access: 15 March 2020, CAMS, 2018; Kaiser et al., 2012). The ERA5 data are available at https://doi.org/10.24381/cds.adbb2d47 (Hersbach et al., 2018a, 2020) and https://doi.org/10.24381/cds.bd0915c6 (Hersbach et al., 2018b, 2020).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/acp-22-15287-2022-supplement.

Author contributions. DP designed the study and performed the model simulations. VT performed the analysis and wrote the paper. VT and DP interpreted the results. CG, MG, AR, and TAM provided significant input to the interpretation and the improvement of the paper. All authors discussed the results and commented on the paper.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Atmospheric Chemistry and Physics*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. This study has been supported by funding from the Max Planck Society allocated to the Max Planck Partner Group at IISERB and the Science and Engineering Research Board (SERB) through an Early Career Research Award (grant no. ECR/2018/001111) to Dhanyalekshmi Pillai. Thara Anna Mathew acknowledges the financial support provided by a SERB grant (Junior Research Fellowship). We acknowledge the support of IISERB's high-performance cluster system, for computations, data analysis, and visualization. The WRF-Chem simulations were done on the high-performance cluster Mistral of the Deutsches Klimarechenzentrum GmbH (DKRZ). We thank the editor and both referees, for their constructive comments.

Financial support. This research has been supported by the Max-Planck-Gesellschaft (Max Planck Partner Group, IISER Bhopal) and the Science and Engineering Research Board (SERB) through an Early Career Research Award (grant no. ECR/2018/001111).

The article processing charges for this open-access publication were covered by the Max Planck Society.

Review statement. This paper was edited by Rolf Müller and reviewed by Prabir K. Patra and one anonymous referee.

References

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., Law, R. M., Loh, Z., Morguí, J. A., Parrington, M., Peuch, V.-H., Ramonet, M., Roehl, C., Vermeulen, A. T., Warneke, T., and Wunch, D.: Modelling CO_2 weather – why horizontal resolution matters, Atmos. Chem. Phys., 19, 7347–7376, https://doi.org/10.5194/acp-19-7347-2019, 2019.

Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neininger, B., Dolman, A. J., and Sarrat, C.: Mesoscale covariance of transport and CO_2 fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model, J. Geophys. Res., 112, D22107, https://doi.org/10.1029/2007jd008552, 2007.

Ahmadov, R., Gerbig, C., Kretschmer, R., Körner, S., Rödenbeck, C., Bousquet, P., and Ramonet, M.: Comparing high resolution WRF-VPRM simulations and two global CO_2 transport models with coastal tower measurements of CO_2, Biogeosciences, 6, 807–817, https://doi.org/10.5194/bg-6-807-2009, 2009.

Ahmadov, R., McKeen, S. A., Robinson, A. L., Bahreini, R., Middlebrook, A. M., de Gouw, J. A., Meagher, J., E.-Y., Edgerton, E., Shaw, S., and Trainer, M.: A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006, J. Geophys. Res.-Atmos., 117, D06301, https://doi.org/10.1029/2011JD016831, 2012.

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647–687, https://doi.org/10.5194/amt-7-647-2014, 2014.

Beck, J., Brown, J., Dudhia, J., Gill, D., Hertneky, T., Klemp, J., Wang, W., Williams, C., Hu, M., James, E., Kenyon, J., Smirnova, T., and Kim, J.: An Evaluation of a Hybrid, Terrain-Following Vertical Coordinate in the WRF-Based RAP and HRRR Models, Weather Forecast., 35, 1081–1096, https://doi.org/10.1175/WAF-D-19-0146.1, 2020.

Beck, V., Koch, T., Kretschmer, R., Marshall, J., Ahmadov, R., Gerbig, C., Pillai, D., and Heimann, M.: The WRF Greenhouse Gas Model (WRF-GHG), Tech. Rep. 25, Max Planck Institute for Biogeochemistry, Jena, Germany, http://www.bgc-jena.mpg.de/bgc-systems/index.shtml (last access: 10 August 2019), 2011.

Bergamaschi, P., Danila, A., Weiss, R. F., Ciais, P., Thompson, R. L., Brunner, D., Levin, I., Meijer, Y., Chevallier, F., Bovensmann, H., Crisp, D., Basu, S., Dlugokencky, E., Engelen, R., Gerbig, C., Günther, D., Hammer, S., Henne, S., Houweling, S., Peylin, P., Pinty, B., Ramonet, M., Reimann, S., Röckmann, T., Schmidt, M., Strogies, M., Sussams, J., Tarasova, O., van Aardenne, J., Vermeulen, A. T., and Vogel, F.: Atmospheric monitoring and inverse modelling for verification of greenhouse gas inventories, Publications Office of the European Union, https://doi.org/10.2760/759928, 2018.

Bousquet, P., Peylin, P., Ciais, P., Le Quéré, C., Friedlingstein, P., and Tans, P. R. P.: Regional Changes in Carbon Dioxide Fluxes of Land and Oceans Since 1980, Science, 290, 1342–1346, https://doi.org/10.1126/science.290.5495.1342, 2000.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, https://doi.org/10.5194/acp-15-1707-2015, 2015.

Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039–9056, https://doi.org/10.5194/acp-13-9039-2013, 2013.

Carouge, C., Bousquet, P., Peylin, P., Rayner, P. J., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 1: Potential of the 2001 network, Atmos. Chem. Phys., 10, 3107–3117, https://doi.org/10.5194/acp-10-3107-2010, 2010.

Cervarich, M., Shu, S., Jain, A. K., Arneth, A., Canadell, J., Friedlingstein, P., Houghton, R. A., Kato, E., Koven, C., Patra, P., Poulter, B., Sitch, S., Stocker, B., Viovy, N., Wiltshire, A., and Zeng, N.: The terrestrial carbon budget of South and Southeast Asia, Environ. Res. Lett., 11, 105006, https://doi.org/10.1088/1748-9326/11/10/105006, 2016.

Chandra, N., Hayashida, S., Saeki, T., and Patra, P. K.: What controls the seasonal cycle of columnar methane observed by GOSAT over different regions in India?, Atmos. Chem. Phys., 17, 12633–12643, https://doi.org/10.5194/acp-17-12633-2017, 2017.

Chen, H. W., Zhang, L. N., Zhang, F., Davis, K. J., Lauvaux, T., Pal, S., Gaudet, B., and DiGangi, J. P.: Evaluation of Regional CO_2 Mole Fractions in the ECMWF CAMS Real-Time Atmospheric Analysis and NOAA CarbonTracker Near-Real-Time Reanalysis with Airborne Observations from ACT-America Field Campaigns, J. Geophys. Res.-Atmos., 124, 8119–8133, https://doi.org/10.1029/2018jd029992, 2019.

Chevallier, F.: On the parallelization of atmospheric inversions of CO_2 surface fluxes within a variational framework, Geosci. Model Dev., 6, 783–790, https://doi.org/10.5194/gmd-6-783-2013, 2013.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24309, https://doi.org/10.1029/2005jd006390, 2005.

Chevallier, F., Bréon, F. M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework, J. Geophys. Res., 112, D09307, https://doi.org/10.1029/2006JD007375, 2007.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, D21307, https://doi.org/10.1029/2010jd013887, 2010.

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions, Atmos. Chem. Phys., 19, 14233–14251, https://doi.org/10.5194/acp-19-14233-2019, 2019.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S.,

Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, https://doi.org/10.5194/bg-11-3547-2014, 2014.

Corbin, K. D., Denning, A. S., Lu, L., Wang, J.-W., and Baker, I. T.: Possible representation errors in inversions of satellite CO_2 retrievals, J. Geophys. Res., 113, D02301, https://doi.org/10.1029/2007JD008716, 2008.

Copernicus Atmosphere Monitoring Service (CAMS): Global Fire Assimilation System (GFAS) v1.2, European Centre for Medium-Range Weather Forecasts (ECMWF) [data set], https://apps.ecmwf.int/datasets/data/cams-gfas/ (last access: 15 March 2020), 2018.

Copernicus Atmosphere Monitoring Service (CAMS)/Laboratoire des Sciences du Climat et l'Environnement (LSCE): Greenhouse Gases Flux Inversions Version v18r3, European Centre for Medium-Range Weather Forecasts (ECMWF) [data set], https://apps.ecmwf.int/datasets/data/cams-ghg-inversions/ (last access: 22 July 2020), 2019a.

Copernicus Atmosphere Monitoring Service (CAMS)/Laboratoire des Sciences du Climat et l'Environnement (LSCE): Greenhouse Gases Flux Inversions Version FT18r1, European Centre for Medium-Range Weather Forecasts (ECMWF) [data set], https://apps.ecmwf.int/datasets/data/cams-ghg-inversions/ (last access: 22 July 2020), 2019b.

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

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J., and Vignati, E.: EDGAR v6.0 Greenhouse Gas Emissions, European Commission, Joint Research Centre (JRC) [data set], http://data.europa.eu/89h/97a67d67-c62e-4826-b873-9d972c4f670b (last access: 20 May 2022), 2021.

Denvil-Sommer, A., Gehlen, M., Vrac, M., and Mejia, C.: LSCE-FFNN-v1: a two-step neural network model for the reconstruction of surface ocean pCO_2 over the global ocean, Geosci. Model Dev., 12, 2091–2105, https://doi.org/10.5194/gmd-12-2091-2019, 2019.

Department of Agriculture & Cooperation, Ministry of Agriculture (DAC/MA): Agricultural Statistics at a Glance 2014, Directorate of Economics and Statistics, Department of Agriculture and Cooperation (DAC), Ministry of Agriculture (MA), Government of India, OUP, New Delhi, India, ISBN 0-19-945965-7, https://eands.dacnet.nic.in/PDF/Agricultural-Statistics-At-Glance2014.pdf (last access: 11 February 2021), 2015.

Engelen, R. J., Denning, A. S., Gurney, K. R., and TransCom3 modelers: On error estimation in atmospheric CO_2 inversions, J. Geophys. Res., 107, 4635, https://doi.org/10.1029/2002JD002195, 2002.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Atmospheric and Space Science Series, Cambridge University Press, https://doi.org/10.1017/CBO9780511535741, 2002.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

Gadgil, S.: The Indian Monsoon and its Variability, Annu. Rev. Earth Planet. Sc., 31, 429–467, https://doi.org/10.1146/annurev.earth.31.100901.141251, 2003.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, J. Geophys. Res.-Atmos., 108, D244756, https://doi.org/10.1029/2002jd003018, 2003.

Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591–602, https://doi.org/10.5194/acp-8-591-2008, 2008.

Goswami, B. N. and Xavier, P. K.: Dynamics of "internal" interannual variability of the Indian summer monsoon in a GCM, J. Geophys. Res., 110, D24104, https://doi.org/10.1029/2005jd006042, 2005.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled online chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Hammerling, D. M., Michalak, A. M., and Kawa, S. R.: Mapping of CO_2 at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, J. Geophys. Res., 117, D06306, https://doi.org/10.1029/2011JD017015, 2012.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.adbb2d47, 2018a.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on pressure levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.bd0915c6, 2018b.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Mu noz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G. O., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diaman takis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G. A., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Andrews, A., Mund, J., Ott, L., Collatz, G. J., Aalto, T., Afshar, S., Aikin, K., Aoki, S., Apadula, F., Baier, B., Bergamaschi, P., Beyersdorf, A., Biraud, S. C., Bollenbacher, A., Bowling, D., Brailsford, G., Abshire, J. B., Chen, G., Chen, H., Chmura, L., Sites Climadat, Colomb, A., Conil, S., Cox, A., Cristofanelli, P., Cuevas, E., Curcoll, R., Sloop, C. D., Davis, K., Wekker, S. D., Delmotte, M., DiGangi, J. P., Dlugokencky, E., Ehleringer, J., Elkins, J. W., Emmenegger, L., Fischer, M. L., Forster, G., Frumau, A., Galkowski, M., Gatti, L. V., Gloor, E., Griffis, T., Hammer, S., Haszpra, L., Hatakka, J., Heliasz, M., Hensen, A., Hermanssen, O., Hintsa, E., Holst, J., Jaffe, D., Karion, A., Kawa, S. R., Keeling, R., Keronen, P., Kolari, P., Kominkova, K., Kort, E., Krummel, P., Kubistin, D., Labuschagne, C., Langenfelds, R., Laurent, O., Laurila, T., Lauvaux, T., Law, B., Lee, J., Lehner, I., Leuenberger, M., Levin,I., Levula, J., Lin, J., Lindauer, M., Loh, Z., Lopez, M., Luijkx, I. T., Myhre, C. L., Machida, T., Mammarella, I., Manca, G., Manning, A., Marek, M. V., Marklund, P., Martin, M. Y., Matsueda, H., McKain, K., Meijer, H., Meinhardt, F., Miles, N., Miller, C. E., Mölder, M., Montzka, S., Moore, F., Morgui, J. A., Morimoto, S., Munger, B., Necki, J., Newman, S., Nichol, S., Niwa, Y., O'Doherty, S., Ottosson-Löfvenius, M., Paplawsky, B., Peischl, J., Peltola, O., Pichon J. M., Piper, S., Plass-Dölmer, C., Ramonet, M., Reyes-Sanchez, E., Richardson, S., Riris, H., Ryerson, T., Saito, K., Sargent, M., Sasakawa, M., Sawa, Y., Say, D., Scheeren, B., Schmidt, M., Schmidt, A., Schumacher, M., Shepson, P., Shook, M., Stanley, K., Steinbacher, M., Stephens, B., Sweeney, C., Thoning, K., Torn, M., Turnbull, J., Tørseth, K., Bulk, P. V. D., Dinther, D. V., Vermeulen, A., Viner, B., Vitkova, G., Walker, S., Weyrauch, D., Wofsy, S., Worthy, D., Young, D., and Zimnoch, M.: CarbonTracker CT2019B, Model published by NOAA Earth System Research Laboratory, Global Monitoring Division [data set] , https://doi.org/10.25925/20201008, 2020.

Janjić, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the representation error in data assimilation, Q. J. Roy. Meteor. Soc., 144, 1257–1278, https://doi.org/10.1002/qj.3130, 2017.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modelling, Remote Sens. Environ., 101, 534–553, https://doi.org/10.1016/j.rse.2006.01.020, 2006.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, https://doi.org/10.5194/bg-9-527-2012, 2012.

Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, A., Vogel, F., Hammer, S., and Totsche, K. U.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO2, Atmos. Chem. Phys., 14, 7149–7172, https://doi.org/10.5194/acp-14-7149-2014, 2014.

Krishnamurthy, V. and Ajayamohan, R. S.: Composite Structure of Monsoon Low Pressure Systems and Its Relation to Indian Rainfall, J. Climate, 23, 4285–4305, https://doi.org/10.1175/2010JCLI2953.1, 2010.

Krishnamurthy, V. and Shukla, J.: Seasonal persistence and propagation of intraseasonal patterns over the Indian monsoon region, Clim. Dynam., 30, 353–369, https://doi.org/10.1007/s00382-007-0300-7, 2007.

Kunchala, R. K., Patra, P. K., Kumar, K. N., Chandra, N., Attada, R., and Karumuri, R. K.: Spatio-temporal variability of XCO_2 over Indian region inferred from Orbiting Carbon Observatory (OCO-2) satellite and Chemistry Transport Model, Atmos. Res., 269, 106044, https://doi.org/10.1016/J.ATMOSRES.2022.106044, 2022.

Landschützer, P., Gruber, N., and Bakker, D. C. E.: An observation-based global monthly gridded sea surface pCO_2 product from 1982 onward and its monthly climatology (NCEI Accession 0160558), NOAA National Centers for Environmental Information [data set], https://doi.org/10.7289/v5z899n6, 2017.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO_2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, https://doi.org/10.5194/bg-6-1089-2009, 2009a.

Lauvaux, T., Gioli, B., Sarrat, C., Rayner, P. J., Ciais, P., Chevallier, F., Noilhan, J., Miglietta, F., Brunet, Y., Ceschia, E., Dolman, H., Elbers, J. A., Gerbig, C., Hutjes, R., Jarosz, N., Legain, D., and Uliasz, M.: Bridging the gap between atmospheric concentrations and local ecosystem measurements, Geophys. Res. Lett., 36, L19809, https://doi.org/10.1029/2009GL039574, 2009b.

Li, D., Vogel, B., Müller, R., Bian, J., Günther, G., and Riese, M.: Tropical cyclones reduce ozone in the tropopause region over the western Pacific: An analysis of 18 years ozonesonde profiles, Earths Future, 9, e2020EF001635, https://doi.org/10.1029/2020EF001635, 2021.

Li, W., Ciais, P., Wang, Y., Peng, S., Broquet, G., Ballantyne, A. P., Cana dell, J. G., Cooper, L., Friedlingstein, P., Le Quéré, C., Myneni, R. B., Peters, G. P., Piao, S., and Pongratz, J.: Reducing uncertainties in decadal variability of the global carbon budget with multiple datasets, P. Natl. Acad. Sci. USA, 113, 13104–13108, https://doi.org/10.1073/pnas.1603956113, 2016.

Liang, A., Gong, W., Han, G., and Xiang, C.: Comparison of Satellite-Observed XCO_2 from GOSAT, OCO-2, and Ground-Based TCCON, Remote Sens., 9, 1033, https://doi.org/10.3390/rs9101033, 2017.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res.-Atmos., 108, 4493, https://doi.org/10.1029/2002JD003161, 2003.

Lin, X., Indira, N. K., Ramonet, M., Delmotte, M., Ciais, P., Bhatt, B. C., Reddy, M. V., Angchuk, D., Balakrishnan, S., Jorphail, S., Dorjai, T., Mahey, T. T., Patnaik, S., Begum, M., Brenninkmeijer, C., Durairaj, S., Kirubagaran, R., Schmidt, M., Swathi, P. S., Vinithkumar, N. V., Yver Kwok, C., and Gaur, V. K.: Long-lived atmospheric trace gases measurements in flask samples from three stations in India, Atmos. Chem. Phys., 15, 9819–9849, https://doi.org/10.5194/acp-15-9819-2015, 2015.

Liu, Y., Yue, T., Zhang, L., Zhao, N., Zhao, M., and Liu, Y.: Simulation and analysis of XCO_2 in North China based on high accuracy surface modeling, Environ. Sci. Pollut. Res., 25, 27378–27392, https://doi.org/10.1007/s11356-018-2683-x, 2018.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochem. Cy., 22, GB2005, https://doi.org/10.1029/2006gb002735, 2008.

Massart, S., Agustí-Panareda, A., Heymann, J., Buchwitz, M., Chevallier, F., Reuter, M., Hilker, M., Burrows, J. P., Deutscher, N. M., Feist, D. G., Hase, F., Sussmann, R., Desmet, F., Dubey, M. K., Griffith, D. W. T., Kivi, R., Petri, C., Schneider, M., and Velazco, V. A.: Ability of the 4-D-Var analysis of the GOSAT BESD XCO_2 retrievals to characterize atmospheric CO_2 at large and synoptic scales, Atmos. Chem. Phys., 16, 1653–1671, https://doi.org/10.5194/acp-16-1653-2016, 2016.

Max Planck Institute for Biogeochemistry (MPI-BGC): Jena CarboScope Version s10oc_v2020, Max Planck Institute for Biogeochemistry [data set], http://www.bgc-jena.mpg.de/CarboScope/, last access: 20 July 2020.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, P. Natl. Acad. Sci. USA, 109, 8423–8428, https://doi.org/10.1073/pnas.1116645109, 2012.

Miller, C. E., Crisp, D., DeCola, P. L., Olsen, S. C., Randerson, J. T., Michalak, A. M., Alkhaled, A., Rayner, P., Jacob, D. J., Suntharalingam, P., Jones, D. B. A., Denning, A. S., Nicholls, M. E., Doney, S. C., Pawson, S., Boesch, H., Connor, B. J., Fung, I. Y., O'Brien, D., Salawitch, R. J., Sander, S. P., Sen, B., Tans, P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., and Law, R. M.: Precision requirements for space-based XCO_2 data, J. Geophys. Res., 112, D10314, https://doi.org/10.1029/2006JD007659, 2007.

National Center for Atmospheric Research (NCAR): Weather Research and Forecasting Model Version 3.9.1.1, NCAR [code], https://doi.org/10.5065/D6MK6B4K, 2017.

NOAA/ESRL: Dipole Mode Index, https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/DMI/, last access: 11 February 2022a.

NOAA/ESRL: Multivariate ENSO Index, https://www.esrl.noaa.gov/psd/enso/mei/, last access: 11 February 2022b.

Nomura, S., Naja, M., Ahmed, M. K., Mukai, H., Terao, Y., Machida, T., Sasakawa, M., and Patra, P. K.: Measurement report: Regional characteristics of seasonal and long-term variations in greenhouse gases at Nainital, India, and Comilla, Bangladesh, Atmos. Chem. Phys., 21, 16427–16452, https://doi.org/10.5194/acp-21-16427-2021, 2021.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, https://doi.org/10.5194/amt-5-99-2012, 2012.

Parazoo, N. C., Denning, A. S., Berry, J. A., Wolf, A., Randall, D. A., Kawa, S. R., Pauluis, O., and Doney, S. C.: Moist synoptic transport of CO_2 along the mid-latitude storm track, Geophys. Res. Lett., 38, L09804, https://doi.org/10.1029/2011gl047238, 2011.

Park, C., Gerbig, C., Newman, S., Ahmadov, R., Feng, S., Gurney, K. R., Carmichael, G. R., Park, S.-Y., Lee, H.-W., Goulden, M., Stutz, J., Peischl, J., and Ryerson, T.: CO_2 Transport Variability, and Budget over the Southern California Air Basin Using the High-Resolution WRF-VPRM Model during the CalNex 2010 Campaign, J. Appl. Meteorol. Clim., 57, 1337–1352, https://doi.org/10.1175/jamc-d-17-0358.1, 2018.

Park, M., Randel, W. J., Gettelman, A., Massie, S. T., and Jiang, J. H.: Transport above the Asian summer monsoon anticyclone inferred from Aura Microwave Limb Sounder tracers, J. Geophys. Res., 112, D16309, https://doi.org/10.1029/2006JD008294, 2007.

Park, S., Klemp, J. B., and Kim, J.: Hybrid Mass Coordinate in WRF-ARW and Its Impact on Upper-Level Turbulence Forecasting, Mon. Weather Rev., 147, 971–985, https://doi.org/10.1175/MWR-D-18-0334.1, 2019.

Patra, P. K., Ishizawa, M., Maksyutov, S., Nakazawa, T., and Inoue, G.: Role of biomass burning and climate anomalies for land-atmosphere carbon fluxes based on inverse modeling of atmospheric CO_2, Global Biogeochem. Cy., 19, GB3005, https://doi.org/10.1029/2004gb002258, 2005.

Patra, P. K., Niwa, Y., Schuck, T. J., Brenninkmeijer, C. A. M., Machida, T., Matsueda, H., and Sawa, Y.: Carbon balance of South Asia constrained by passenger aircraft CO_2 measurements, Atmos. Chem. Phys., 11, 4163–4175, https://doi.org/10.5194/acp-11-4163-2011, 2011.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, https://doi.org/10.1073/pnas.0708986104, 2007.

Philip, S., Johnson, M. S., Baker, D. F., Basu, S., Tiwari, Y. K., Indira, N. K., Ramonet, M., and Poulter, B.: OCO-2 satellite-imposed constraints on terrestrial biospheric CO_2 fluxes over South Asia, J. Geophys. Res.-Atmos., 127, e2021JD035035, https://doi.org/10.1029/2021JD035035, 2022.

Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., and Karstens, U.: High resolution modeling of CO_2 over Europe: implications for representation er-

rors of satellite retrievals, Atmos. Chem. Phys., 10, 83–94, https://doi.org/10.5194/acp-10-83-2010, 2010.

Pillai, D., Gerbig, C., Ahmadov, R., Rödenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B., and Lavrié, J. V.: High-resolution simulations of atmospheric CO_2 over complex terrain – representing the Ochsenkopf mountain tall tower, Atmos. Chem. Phys., 11, 7445–7464, https://doi.org/10.5194/acp-11-7445-2011, 2011.

Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B., and Heimann, M.: Comparing Lagrangian and Eulerian models for CO_2 transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, Atmos. Chem. Phys., 12, 8979–8991, https://doi.org/10.5194/acp-12-8979-2012, 2012.

Pillai, D., Buchwitz, M., Gerbig, C., Koch, T., Reuter, M., Bovensmann, H., Marshall, J., and Burrows, J. P.: Tracking city CO_2 emissions from space using a high-resolution inverse modelling approach: a case study for Berlin, Germany, Atmos. Chem. Phys., 16, 9591–9610, https://doi.org/10.5194/acp-16-9591-2016, 2016.

Ravi Kumar, K., Valsala, V., Tiwari, Y. K., Revadekar, J. V., Pillai, P., Chakraborty, S., and Murtugudde, R.: Intraseasonal variability of atmospheric CO_2 concentrations over India during summer monsoons, Atmos. Environ., 142, 229–237, https://doi.org/10.1016/j.atmosenv.2016.07.023, 2016.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, https://doi.org/10.5194/acp-3-1919-2003, 2003.

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: How does the terrestrial carbon exchange respond to interannual climatic variations? A quantification based on atmospheric CO_2 data, Biogeosciences, 15, 2481–2498, https://doi.org/10.5194/bg-15-2481-2018, 2018a.

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: History of El Ni no impacts on the global carbon cycle 1957–2017: a quantification from atmospheric CO_2 data, Philos. T. Roy. Soc. B, 373, 20170303, https://doi.org/10.1098/rstb.2017.0303, 2018b.

Sarrat, C., Noilhan, J., Dolman, A. J., Gerbig, C., Ahmadov, R., Tolk, L. F., Meesters, A. G. C. A., Hutjes, R. W. A., Ter Maat, H. W., Pérez-Landa, G., and Donier, S.: Atmospheric CO_2 modeling at the regional scale: an intercomparison of 5 meso-scale atmospheric models, Biogeosciences, 4, 1115–1126, https://doi.org/10.5194/bg-4-1115-2007, 2007.

Schimel, D., Stephens, B. B., and Fisher, J. B.: Effect of increasing CO_2 on the terrestrial carbon cycle, P. Natl. Acad. Sci. USA, 112, 436–441, https://doi.org/10.1073/pnas.1407302112, 2014.

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., Braswell, B. H., Apps, M. J., Baker, D., Bondeau, A., Canadell, J., Churkina, G., Cramer, W., Denning, A. S., Field, C. B., Friedlingstein, P., Goodale, C., Heimann, M., Houghton, R. A., Melillo, J. M., Moore, B., Murdiyarso, D., Noble, I., Pacala, S. W., Prentice, I. C., Raupach, M. R., Rayner, P. J., Scholes, R. J., Steffen, W. L., and Wirth, C.: Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems, Nature, 414, 169–172, https://doi.org/10.1038/35102500, 2001.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Duda, D. M. B. M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, NCAR Tech. Note, NCAR/TN-468+STR, Tech. rep., National Center for Atmosphere Research, Boulder, Colorado, USA, https://doi.org/10.5065/D68S4MVH, 2008.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker D. M., and Huang, X. Y.: A Description of the Advanced Research WRF Model Version 4.3 (No. NCAR/TN-556+STR), Technical report, NCAR Technical Notes, https://doi.org/10.5065/1dfh-6p97, 2021.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., and Mukai, H.: The CO_2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmos. Chem. Phys., 11, 6855–6870, https://doi.org/10.5194/acp-11-6855-2011, 2011.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R., and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans, Deep-Sea Res. Pt. II, 56, 554–577, https://doi.org/10.1016/j.dsr2.2008.12.009, 2009.

Thilakan, V. and Pillai, D.: Representation error in global model CO_2 simulations over India, Zenodo [data set], https://doi.org/10.5281/zenodo.6616466, 2022.

Tiwari, Y. K., Patra, P. K., Chevallier, F., Francey, R. J., Krummel, P. B., Allison, C. E., Revadekar, J. V., Chakraborty, S., Langenfelds, R. L., Bhattacharya, S. K., Borole, D. V., Kumar, K. R., and Steele, L. P.: CO_2 observations at Cape Rama, India for the period of 1993–2002: implications for constraining Indian emissions, Current Science, Indian Academy of Sciences, 101, 1562–1568, 2011.

Tiwari, Y. K., Revadekar, J. V., and Kumar, K. R.: Variations in atmospheric Carbon Dioxide and its association with rainfall and vegetation over India, Atmos. Environ., 68, 45–51, https://doi.org/10.1016/j.atmosenv.2012.11.040, 2013.

Tolk, L. F., Meesters, A. G. C. A., Dolman, A. J., and Peters, W.: Modelling representation errors of atmospheric CO_2 mixing ratios at a regional scale, Atmos. Chem. Phys., 8, 6587–6596, https://doi.org/10.5194/acp-8-6587-2008, 2008.

Uebel, M., Herbst, M., and Bott, A.: Mesoscale simulations of atmospheric CO_2 variations using a high-resolution model system with process-based CO_2 fluxes, Q. J. Roy. Meteor. Soc., 143, 1860–1876, https://doi.org/10.1002/qj.3047, 2017.

Valsala, V. and Maksyutov, S.: Interannual variability of the air–sea CO_2 flux in the north Indian Ocean, Ocean Dynam., 63, 165–178, https://doi.org/10.1007/s10236-012-0588-7, 2013.

van der Molen, M. K. and Dolman, A. J.: Regional carbon fluxes and the effect of topography on the variability of atmospheric CO_2, J. Geophys. Res., 112, D01104, https://doi.org/10.1029/2006jd007649, 2007.

Vogel, B., Müller, R., Günther, G., Spang, R., Hanumanthu, S., Li, D., Riese, M., and Stiller, G. P.: Lagrangian simulations of the transport of young air masses to the top of the Asian monsoon

anticyclone and into the tropical pipe, Atmos. Chem. Phys., 19, 6007–6034, https://doi.org/10.5194/acp-19-6007-2019, 2019.

Willetts, P. D., Marsham, J. H., Birch, C. E., Parker, D. J., Webster, S., and Petch, J.: Moist convection and its upscale effects in simulations of the Indian monsoon with explicit and parametrized convection, Q. J. Roy. Meteor. Soc., 143, 1073–1085, https://doi.org/10.1002/qj.2991, 2017.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO2 measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Zellweger, C., Emmenegger, L., Firdaus, M., Hatakka, J., Heimann, M., Kozlova, E., Spain, T. G., Steinbacher, M., van der Schoot, M. V., and Buchmann, B.: Assessment of recent advances in measurement techniques for atmospheric carbon dioxide and methane observations, Atmos. Meas. Tech., 9, 4737–4757, https://doi.org/10.5194/amt-9-4737-2016, 2016.