A global coupled Eulerian-Lagrangian model and 1 × 1 km CO₂ surface flux dataset for high-resolution atmospheric CO₂ transport simulations


1 Central Aerological Observatory, Dolgoprudny, Russia
2 Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
3 Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, USA
4 Department of Earth Sciences, Royal Holloway, University of London, London, UK
5 Meteorological Research Institute, Tsukuba, Japan
6 School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK
7 Department of Physics, Imperial College London, London, UK
8 A.N. Severtsov Institute of Ecology and Evolution, Russia
9 Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, Russia
* now at: Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, USA/National Oceanic and Atmosphere Administration, Earth System Research Laboratory, Boulder, USA
** now at: LSCE, Gif Sur Yvette, France
*** now at: CAT/ESSC, Indian Institute of Tropical Meteorology, Pune, India

Correspondence to: A. Ganshin (alexander.ganshin@gmail.com)

Received: 26 July 2011 – Published in Geosci. Model Dev. Discuss.: 24 August 2011
Revised: 23 January 2012 – Accepted: 27 January 2012 – Published: 15 February 2012

Abstract. We designed a method to simulate atmospheric CO₂ concentrations at several continuous observation sites around the globe using surface fluxes at a very high spatial resolution. The simulations presented in this study were performed using the Global Eulerian-Lagrangian Coupled Atmospheric model (GELCA), comprising a Lagrangian particle dispersion model coupled to a global atmospheric tracer transport model with prescribed global surface CO₂ flux maps at a 1 × 1 km resolution. The surface fluxes used in the simulations were prepared by assembling the individual components of terrestrial, oceanic and fossil fuel CO₂ fluxes. This experimental setup (i.e. a transport model running at a medium resolution, coupled to a high-resolution Lagrangian particle dispersion model together with global surface fluxes at a very high resolution), which was designed to represent high-frequency variations in atmospheric CO₂ concentration, has not been reported at a global scale previously.

Two sensitivity experiments were performed: (a) using the global transport model without coupling to the Lagrangian dispersion model, and (b) using the coupled model with a reduced resolution of surface fluxes, in order to evaluate the performance of Eulerian-Lagrangian coupling and the role of high-resolution fluxes in simulating high-frequency variations in atmospheric CO₂ concentrations. A correlation analysis between observed and simulated atmospheric CO₂ concentrations at selected locations revealed that the inclusion of both Eulerian-Lagrangian coupling and high-resolution fluxes improves the high-frequency simulations of the model. The results highlight the potential of a coupled Eulerian-Lagrangian model in simulating high-frequency atmospheric CO₂ concentrations at many locations worldwide. The model performs well in representing observations of atmospheric CO₂ concentrations at high spatial and temporal resolutions, especially for coastal sites and sites located close to sources of large anthropogenic emissions. While this study focused on simulations of CO₂ concentrations, the model could be used for other atmospheric compounds with known estimated emissions.
1 Introduction

The anthropogenic emissions of greenhouse gases could potentially change the global average temperature, leading to global warming. The latest assessment report of the Intergovernmental Panel on Climate Change (IPCC-AR4) states that climate models are capable of reproducing the temperature trends observed in recent decades if they are forced with increasing concentrations of anthropogenic greenhouse gas (IPCC, 2007). A major contributor to anthropogenic greenhouse gases in the atmosphere is carbon dioxide (CO\textsubscript{2}), which plays a key role in the global climate. Consequently, estimations of CO\textsubscript{2} emissions are important in assessing its influence on ongoing climate change. To understand the nature of CO\textsubscript{2} cycling between the land, atmosphere and ocean, it is necessary to calculate precisely the natural and anthropogenic fluxes of CO\textsubscript{2} and the temporal and spatial variability of CO\textsubscript{2} concentrations in the atmosphere.

Variations in CO\textsubscript{2} concentrations in the atmosphere are generally assessed using transport models with prescribed surface fluxes. Such modelling efforts are essential in order to inversely estimate the CO\textsubscript{2} fluxes between the atmosphere and land–ocean surfaces using model-predicted quantities of atmospheric CO\textsubscript{2} and corresponding observations (Gurney et al., 2002). The essential elements required for modelling of atmospheric CO\textsubscript{2} concentrations are a transport model, meteorological drivers, and surface fluxes. Consequently, the accuracy of simulated concentrations is strongly dependent on the ability of the transport model to represent the observed variability of atmospheric CO\textsubscript{2} concentrations, thereby resulting in large differences in the outputs of various models (Patra et al., 2008). Therefore, the faithful representation of CO\textsubscript{2} concentrations in transport models is an area of active research (e.g. Patra et al., 2008). In this study, we present a novel strategy to improve the performance of a transport model by coupling two essential components of modelling, as described below.

The concentrations of atmospheric constituents and their transport are simulated using Lagrangian or Eulerian models. In Lagrangian models, the trajectories of individual particles (representing the chemical constituents of the atmosphere) are calculated by following a predetermined atmospheric velocity. One example of a Lagrangian model is a trajectory model in which the locations of air masses are traced using a notional particles. Lagrangian particle dispersion models (LPDM) are typical examples of trajectory models extended to simulations of plume diffusion, utilizing multiple particles to represent turbulence and convection processes, in addition to the movement of paritciles along mean flow trajectory paths (e.g. Thomson, 1987). LPDMs can be run forward or backward in time. Forward simulations are generally employed to calculate the transport and dispersion of tracers from point sources, whereas backward simulations are used to estimate the potential contributions of pollutants from many sites to a single location. In this case, it is only necessary to perform one backward simulation from the receptor point, tracing the given number of particles back to the possible source locations. Many previous studies have employed backward trajectory analyses of atmospheric tracer transport (Seibert et al., 1994; Stohl, 1996) and have utilized Lagrangian particle dispersion models (Lin et al., 2003; Seibert and Frank, 2004; Stohl et al., 2005; Folini et al., 2008; Lin et al., 2011). A notable advantage of backward transport models is their ability to solve adjoint equations of atmospheric transport and to explicitly estimate a source-receptor sensitivity matrix, which is useful for inverse modelling at high resolution (Stohl et al., 2009). The equivalence of backward transport models to the adjoint of forward transports has been discussed by Marchuk (1995) and Hourdin and Talagrand (2006).

In the case of Eulerian models, the evolution of the concentration field is solved numerically using finite difference approximations to the partial differential equation of tracer transport on a fixed grid rather than along the trajectory path (Richtmyer and Morton, 1967). Such models are applied to global-scale simulations of the concentrations of atmospheric constituents and to the inverse modelling of surface fluxes (Gurney et al., 2002, 2004).

Each modelling approach has its advantages and disadvantages. For example, Eulerian models reproduce the seasonal cycle of atmospheric CO\textsubscript{2} concentrations reasonably well, but they suffer strongly from numerical diffusion, meaning that they perform poorly in representing synoptic, super-synoptic and hourly variations. Lagrangian models, in contrast, do not suffer from numerical diffusion and they perform reasonably well in reproducing synoptic and hourly variations; however, it is necessary to employ a very long backward trajectory simulation (up to 4 months or more) to reproduce the seasonal cycle. The use of a longer trajectory results in the accumulation of errors and is computationally expensive (Stohl et al., 1998).

Given the above limitations, it is reasonable to consider a hybrid model in which a Eulerian model is run to generate the global background concentrations of atmospheric constituents, which are then used as initial conditions for a Lagrangian model (e.g. Koyama et al., 2011). In the present study, we extended the approach introduced by Koyama et al. (2011) to a high-spatial-resolution case for simulating the concentrations of atmospheric CO\textsubscript{2} in the same coupled Eulerian-Lagrangian modelling framework as that used in the earlier study.

Several studies have reported the advantages of using coupled Eulerian-Lagrangian models, and some have used the WRF-STILT modelling system (Nehrkorn et al., 2010) to estimate the CO\textsubscript{2} surface fluxes over North America (Gourdji et al., 2010). Rigby et al. (2011) and Koyama et al. (2009) outlined a method for combining information on the emissions-mole fraction sensitivity from Eulerian and Lagrangian chemical transport models, for use in estimating emissions at a global scale. Previous studies have
also suggested the use of high-resolution coupled Eulerian-Lagrangian models, especially for regional-scale studies. Vermeulen et al. (1999) designed the nesting of a regional Lagrangian model into a global model and implemented a regional inversion. Trusilova et al. (2010) presented the results of coupled TM3-STILT model simulations based on the nested atmospheric inversion scheme developed by Rödenbeck et al. (2009).

In the present study, we coupled a Eulerian model (National Institute for Environmental Studies-Transport Model, herein NIES-TM; Maksyutov et al., 2008; Belikov et al., 2011) and a Lagrangian particle dispersion model (FLEXPART; Stohl et al., 2005). We did not apply the spatial coupling at the domain boundary employed by the regional models described above; instead, we implemented a coupling at temporal boundaries in the global domain. The coupled model is described in detail in Sect. 2.

A goal of this study is to demonstrate the merit of using a coupled model together with a newly developed high-resolution (1 × 1 km) global CO$_2$ flux dataset in simulating high-frequency variations of observed atmospheric CO$_2$ concentrations. The use of 1-km surface fluxes combined with low-resolution meteorological wind fields for driving the model is justified because the wind fields are expected to be in a state of well-mixed daytime conditions, especially for observation sites located in relatively flat areas, where large-scale geostrophic motions are generally expected to be dominant. A spatial radius of correlation on the order of 100 km or more is a commonly observed feature of wind and temperature fields (Buell, 1960, 1972; Gandin, 1965).

The remainder of this paper is organised as follows. The model, data and methods are described in Sect. 2, and CO$_2$ emissions data and observations are presented in Sect. 3. The model results and a discussion are presented in Sect. 4, and the conclusions are provided in Sect. 5.

2 Model and methods

2.1 Lagrangian-Eulerian coupled model

Here, we describe the principles of the proposed Lagrangian-Eulerian coupled model. The concentration simulated by the Lagrangian model at the receptor (observation location) is usually calculated as the integral of the residence time of all particles at each grid cell multiplied by the flux corresponding to that grid (Lin et al., 2003; Seibert and Frank, 2004). The concentration of CO$_2$ in the Lagrangian model at any receptor point (corresponding to an observation site) can be written as follows (Holzer et al., 2000; Lin et al., 2003):

\[
C(x, t) = \int_0^t \int dV \int V I(x, t) S(x, t) dx dt + \int dV \int V I(x, t) C(x, 0)
\]

where: $C(x, t)$ – concentration at receptor point $x$ at time $t$; $C(x, 0)$ – initial concentration field at time $t_0$, which is obtained from the background fields simulated by the Eulerian model:

\[
I(x, t) = \text{influence function or Green’s function linking sources and sinks } S(x, t) \text{ to the concentrations}; \quad dV \quad \text{– volume element.}
\]

The first term of Eq. (1) denotes the concentration change at the receptor from sources/sinks in domain $V$ during the time interval between initialization and observation. The second term refers to the contribution to the concentration at the receptor point by the advection of CO$_2$ from the background tracer field $C(x, 0)$, which is provided by a Eulerian model. From a Lagrangian viewpoint, the influence function corresponds to the transition probability $p(x, t | x, 0)$ along the air mass trajectories $x(n, t)$, calculated by the Lagrangian model as follows:

\[
p(x, t | x, 0) = \frac{1}{N} \sum_{n=1}^{N} \delta(x_n(t) - x)
\]

where: $N$ – number of air parcels emitted in the backward direction from the receptor point, and $\delta$ – delta function representing the presence or absence of parcel $i$ at location $x$. Hence, in discrete form, Eq. (1) can be written as follows:

\[
C(x, t) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{I} \sum_{k=1}^{K} \sum_{l=0}^{L} \sum_{m=0}^{M} f_{ijk}^l C_{ijk}^m + \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{I} \sum_{k=1}^{K} \sum_{l=0}^{L} \sum_{m=0}^{M} f_{ijk}^l C_{ijk}^m (x, 0)
\]

where: $i, j, k$ – indices of a grid cell; $l$ – time index; $C_{ijk}^m$ – initial background concentration from the Eulerian model; $f_{ijk}^l$ – parcel outside the $i, j, k$ cell; $T$ – duration of trajectories; and $L$ – number of steps when sources are sampled by trajectories. Here, we emphasise that the conditioning event involves the particles passing through the receptor point $(x, t)$, and it is customary in probability theory to write receptor after the random variable (e.g. $p(x(t) | x, 0)$ with $t_0 > t$). However, to be consistent with previous authors, we employ the notation introduced by Holzer et al. (2000) and Lin et al. (2003), who placed the receptor on the left side, as established in Green’s function notation, which is the opposite to probability notation. Furthermore, Flesch et al. (1994) showed the equivalence of forward and backward conditional probabilities; i.e. $p^f(x, t | x, 0) = p^b(x, 0 | x, t)$.

In the case of sampling surface sources $F(x, y, t)$, we can consider their significance to a height $h$ (e.g. 500 m, which is the typical height of the planetary boundary layer):

\[
S(x, t) = \begin{cases} 
F(x, y, t)m_{air} & 0 \leq h \\
0 & z > h
\end{cases}
\]

where: $m_{air}$ and $m_{CO2}$ – molar masses of air and CO$_2$, respectively; and $\bar{\rho}$ – average air density below $h$. 

www.geosci-model-dev.net/5/231/2012/

Finally, we obtain
\[ C(x_t, t_t) = \frac{T m_{at}}{hN_t}\sum_{i_1} \sum_{j_1} \sum_{l_1} F_{ij} \sum_{n_1} f_{ij}^{n_1} + \frac{I \sum_{i_2} C_{ij}^{n_2} \sum_{n_2} f_{ij}^{n_2}}{N} \]  
where the left term is associated with concentrations obtained from the Lagrangian model and the right term is the background CO₂ concentration from the global Eulerian transport model.

In our coupled model, we use FLEXPART (run in backward mode) as the Lagrangian particle dispersion model and NIES-TM as the Eulerian global transport model (see the Introduction for details). The background CO₂ values on the model grid are obtained by NIES-TM. We use a 2-day length of backward transport in FLEXPART. Gloor et al. (2001) found that a period on the order of 1.5 days is the timescale over it is still possible to discern, in the mixing ratio variations, the imprint of surface fluxes on the air parcel before its arrival at an observation point. In the present case, the background CO₂ values on the model grid points were provided by NIES-TM sampled 2 days prior to the observations.

### 2.2 Meteorological drivers

We used FLEXPART version 8.0 adapted to using JCDAS data (Onogi et al., 2007), which are provided on hybrid sigma-pressure levels and a Gaussian grid (40 model levels, T106 grid). The original model was designed to use ECMWF data on a regular latitude-longitude horizontal grid and on hybrid sigma-pressure vertical levels. Therefore, to adapt JCDAS data for the FLEXPART model, the required parameter values were obtained via bilinear horizontal interpolation from a Gaussian grid to a regular 1.25° × 1.25° grid. The vertical structure of JCDAS data was used without any modifications; thus, the FLEXPART source code was adjusted for new parameters describing JCDAS vertical levels. The temporal resolution of input data is 6 h. The numerical constraints of the model (i.e., the maximum time-step required for a smooth tracking of contributions of grid cell fluxes to the model concentrations) demand that the model time step (\( \tau \)) is below \( h/U \), where \( h \) is the size of the flux grid cells and \( U \) is the wind speed. At a wind speed of \( U = 50 \text{ km h}^{-1} \), we obtain \( \tau = 2, 0.2, \text{ and } 0.02 \text{ h for } h = 100, 10, \text{ and } 1 \text{ km, respectively. Therefore, we used a period of } 1 \text{ min as both the particle-transport time step and the flux-integration time step (i.e. } T/L \text{ of Eq. 5) for the 1-km flux setup.} \)

The input meteorological data for the NIES-TM model were taken from NCEP reanalyses with a spatial resolution of 2.5° × 2.5° on pressure levels and a temporal resolution of 6 h. In this version of NIES-TM, the advection terms are solved by a second-order moment scheme (Prather, 1986). The implementation is documented by Belikov et al. (2011). The spatial resolution of the simulated NIES-TM background concentration, as supplied for the FLEXPART model, was 2.5° × 2.5° on 15 sigma-levels; the corresponding temporal resolution was 1 h (the same as temporal boundaries of coupling). Background values were used in the model without interpolation to the particle position within the Eulerian grid box during coupling.

### 3 CO₂ fluxes and observations of atmospheric CO₂ concentrations

Three types of flux scenarios are commonly used in atmospheric CO₂ simulations: fossil fuel emissions, and biospheric and oceanic sources and sinks. These three fluxes collectively represent the major types of CO₂ sources and sinks. Forest fires should also be included as a component of biospheric flux. In the present simulations, however, we excluded the contribution from forest fires assuming that it is reasonably small at the present observational sites, especially for the period of analysis used in this study. The above three fluxes are typically available at a spatial resolution of 1° × 1°; however, this may be insufficient to represent strong, local emissions of CO₂. In previous coupled modelling studies (e.g. Koyama et al., 2011), the surface CO₂ fluxes at 1° × 1° resolution were used for simulations of global atmospheric CO₂ concentrations. However, as mentioned above, a flux dataset prepared at a high spatiotemporal resolution is sufficient for regional high-resolution simulations (Trusilova et al., 2010). In this study, we used 1° × 1° fluxes for global-scale Eulerian calculations of atmospheric CO₂ concentrations, and used 1 × 1 km fluxes for the Lagrangian model. These contributions are combined in the coupled model, as described in the previous section. The 1° × 1° fluxes used for Eulerian simulations are the same as 1-km fluxes, with the only difference being resolution; the one exception is biospheric fluxes, for which we used those prepared by Nakatsuka and Maksyutov (2009) using an optimized Carnegie-Ames-Stanford Approach (CASA) terrestrial ecosystem model (van der Werf et al., 2003). Different biospheric emissions are used for the Eulerian and Lagrangian models because each model should handle the appropriate fluxes. The Eulerian model describes seasonal variations in CO₂ whereas the Lagrangian part reconstructs short-term variations. However, because VISIT fluxes did not provide optimized seasonality for the Eulerian model at a given time, we used optimized CASA emissions to describe seasonal variations in concentrations in the biosphere in the Eulerian part of the model. VISIT performs well in reconstructing synoptic variations, and it can be used for simulations at high spatial resolutions. VISIT has prognostic phenology, whereas CASA is driven by Normalized Difference Vegetation Index (NDVI) observations from space. The VISIT model provides biospheric fluxes with a temporal resolution of 1 day, and it is not neutral on the annual scale.

The temporal resolution of the employed flux datasets is low compared with the spatial resolution. An emissions
database prepared at hourly time scale, for example, would be suitable for high-resolution simulations like those presented here. However, datasets at an hourly time scale are only available for certain regions and cities (e.g. Vulcan; Gurney et al., 2009), and there is no available global CO₂ emissions dataset prepared at an hourly time scale. Therefore, we used the best available emissions databases that are currently available with monthly spatial resolution for fossil fuel and oceanic fluxes, and 1-day variations for the biosphere.

### 3.1 Technique for handling fluxes on a 1° × 1 km grid

Here, we describe technical aspects of surface flux simulations at a very high resolution. Because each of the global 1° × 1 km flux fields (e.g. biosphere, fossil and ocean fluxes) requires ~3.5 gigabytes (GB) of computer memory, it is inconvenient to operate with multiple layers of data at this resolution. To reduce the memory and disk storage requirements, we propose the novel approach outlined below.

a. The 1° × 1 km surface flux fields at a given point are calculated in the model using a combination of data fields at high and medium resolutions. In the case of fossil fuel emissions, we multiply 1-km-resolution annual mean fluxes by a medium-resolution (1° × 1°) spatially varying factor that represents the seasonal cycle at a monthly time scale. This factor is derived from seasonally varying fossil-fuel-emissions data (Andres et al., 2011) at a resolution of 1° × 1°, normalized to the annual mean. We use a land-cover mask at 1-km resolution to spatially redistribute the biospheric fluxes given at medium resolution and simulated separately for each of the 15 vegetation types. In this way, the memory usage and computational time are reduced considerably. A description of each dataset is given in the following subsections, and a summary of the combinations of fluxes used in the simulations is given in Table 1. The flux data at each model time step are obtained by linearly interpolating between the monthly fields, except for biospheric fluxes, which are provided a daily time step.

b. Even with the flux treatment outlined above, the memory requirements remained high. Consequently, we used sparse matrix storage (a method of representing matrices populated primarily with zeros; Tewarson, 1973) to reduce the memory demand for storing the anthropogenic emission field at 1-km resolution, because only ~1% of the elements in the matrix for anthropogenic emissions have non-zero values. To speed up the element search in the sparse matrix, we employed 1-dimensional lookup tables that contain indices for the first and last non-zero elements for each longitude.

It is possible to use the same approach for biospheric fluxes; however, this has little effect on storage requirements because about 30% of the elements have non-zero values in this case. The combined application of the above two techniques (i.e. those described in a. and b.) reduces the memory demand to below 1.5 GB.

### 3.2 Fossil fuel CO₂ emissions

We used the Open source Data Inventory of Anthropogenic CO₂ emission (ODIAC) inventory as fossil fuel CO₂ emissions fields, which is a global 1° × 1 km fossil fuel CO₂ emission inventory based on country-level fuel consumption, a global power plant database and satellite observations of night lights (Oda and Maksyutov, 2011). National annual total CO₂ emissions were estimated using BP’s fuel consumption statistics for coal, oil and natural gas (BP, 2008). The spatial distribution of point emissions was determined using power plant locations included in the CARbon Monitoring and Action (CARMA) power plant database (available at http://www.carma.org), and nightlight distributions were used for emissions from sources other than power plants. For further details of the ODIAC inventory, see Oda and Maksyutov (2011). We used the US Department of Energy Carbon Dioxide Information Analysis Center (CDIAC) monthly emission inventory (1° × 1° resolution) to create monthly fields (Andres et al., 2011). The CDIAC monthly emission inventory was created using the monthly fuel consumptions of the top 21 emitting countries. For other countries, we used a proxy based on the data for the country among the top 21 with the most similar climate and economics. Further details of this approach are given in Andres et al. (2011). In
the present study, the CDIAC monthly inventory was used to distribute ODIAC annual totals into the 12 months of the year. Monthly values on 1° × 1° grids were divided by annual totals at the grids, yielding 1° × 1° normalized coefficient fields that were applied to 1 × 1 km ODIAC annual fields to derive monthly varying fields. For the Eulerian model simulation, we created 1° × 1° emission fields from the 1 × 1 km ODIAC (for the spatial pattern) and 1° × 1° CDIAC (for monthly variability) inventories.

3.3 Terrestrial biosphere fluxes

To estimate the vegetation CO₂ fluxes at 1-km resolution, the fluxes were first simulated with the terrestrial biospheric model VISIT (Ito et al., 2007) at a resolution of 0.5° × 0.5° and a daily time step for each of 15 vegetation types in the IGBP classification (Friedl et al., 2002) for each 0.5° × 0.5° grid, closely following the procedure described by Saito et al. (2011). Saito et al. (2011) simulated the ecosystem processes only for four dominant vegetation types using medium-resolution fluxes. However, for high-resolution fluxes the number of vegetation types at each grid point (on a 0.5° × 0.5° grid) should be extended to include all possible types. The fluxes simulated on a 0.5° × 0.5° grid were interpolated spatially to each pixel of the vegetation map. Vegetation cover was derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Product (Friedl et al., 2002). The present study used a global dataset in the Plate Carrée projection (IGBP vegetation classification) at 30 arc-seconds (~1 km). This product was derived from MODIS data for the year 2001 and is based on the reprojected mosaic exports from the MODIS MOD12Q1 v.004 dataset (available at http://duckwater.bu.edu/lc/datasets.html). A MODIS land cover map was used only for the interpolation to obtain a 1 × 1 km flux map, but was not used in the VISIT model.

MODIS-IGBP classification divides terrestrial land covers into 17 biome types and water body, whereas VISIT treats 15 biomes; consequently, the land cover classification was reconstituted to match that of VISIT algorithms (Saito et al., 2011).

3.4 Oceanic CO₂ fluxes

Oceanic fluxes were obtained from a 4-D-var assimilation system based on Valsala and Maksyutov (2010). In this system, an offline biogeochemical model was driven by reanalysis ocean currents and was used to simulate the surface ocean partial pressure of CO₂ (pCO₂) and air-sea CO₂ fluxes. The surface ocean pCO₂ in the model was constrained by ship-based observations of corresponding pCO₂ values and by climatological mean maps of pCO₂. The model employs a variational assimilation method to constrain the simulated surface ocean pCO₂ with reference to observations. Climatological monthly mean air-sea ocean CO₂ flux data were constructed from this assimilation system based on the period between 1996 and 2004. The air-sea CO₂ flux data were remapped onto a regular 1° × 1° grid. This dataset was extended to coastal areas using land–ocean mask data (~1-km resolution) derived from the MODIS global vegetation map described above.

3.5 Atmospheric CO₂ data

The model results were evaluated by comparison with atmospheric CO₂ data obtained at the following continuous observation sites, representative of both polluted and background environments:

1. Fyodorovskoye tower, Russia;
2. Meteorological Research Institute (MRI) tower, Tsukuba, Japan;
3. Queen’s Tower, Imperial College London, London, UK;
4. Royall Holloway University, Egham, London, UK.

The Fyodorovskoye site is located in the central part of western Russia, near Tver (56°27′ N, 32°55′ E), within the Central Forest State Biosphere Natural Reserve (Milyukova et al., 2002; Kurbatova et al., 2008). The reserve is located far from industrial or residential areas and is therefore largely free of air pollution. The measurement tower is 29 m high and is located on a flat surface surrounded by homogenous vegetation (spruce forest). The ambient CO₂ concentrations at heights of 0.20, 1.0, 2.0, 5.0, 11.0, 15.6, 25.0, 27.6 and 29.0 m are measured by a system comprising LiCor non-dispersive infrared gas analysers (Li-Cor 6262-3 and Li-Cor 6251; LI-COR, Lincoln, NB, USA), a pump (KNF, Neurberger, Germany), a switching manifold, BEV-A-Line tubing, a data logger (CR23X, Campbell Scientific Inc., Logan, UT, USA) and a laptop. The instrument was calibrated weekly using air of known CO₂ concentration (pressure bottle) and compressed pure nitrogen. CO₂ data were analysed for the year 2008.

The MRI meteorological tower, which was dismantled in 2011, was 213 m high and located at MRI, Tsukuba, Japan (36°04′N, 140°07′E). Ambient air was introduced from inlets installed at heights of 1.5, 25, 100 and 200 m using diaphragm air pumps, and the CO₂ concentration was measured using a non-dispersive infrared analyser (NDIR) (Inoue and Matsuueda, 1996, 2001). Calibration was performed every 3 h using four standard gases (330, 360, 400 and 450 ppm) in the MRI87 scale, which is comparable to the World Meteorological Organization (WMO) mole-fraction scale (Ishii et al., 2004). In this study, we analysed hourly data on CO₂ concentrations obtained at 200 m height for the year 2009. Hourly CO₂ concentrations were calculated by averaging the observations taken at 16-min intervals.

The other two measurement sites, Queen’s Tower and Royall Holloway University of London, are located within
28 km of each other (Rigby et al., 2008). The Queen’s Tower (height 80 m) is located on the Imperial College campus in South Kensington (51°30’N, 0°11’W), several kilometres west of the geographical centre of London. Royal Holloway University (51°26’N, 0°34’W) is located just outside of the Greater London area, bordered by “green-belt” countryside to the west and south. Measurements were made through a rooftop inlet located approximately 15 m above the ground. The site is ideally located to provide a “background” CO₂ mixing ratio for London, because the prevailing wind direction is from the southwest, blowing toward London. In this study, the hourly CO₂ concentration was calculated by averaging observational data sampled at 30-min intervals. We analysed data for 2006 and 2007. Table 2 lists a summary of the observation stations used in this study.

4 Results and discussion

We performed several simulations using the new model and emissions scenarios: a “synthetic” test to compare the influence of low- and high-resolution fluxes on concentration simulations, sensitivity simulations with coupled and uncoupled versions of the Lagrangian and Eulerian models, and sensitivity simulations with low- and high-resolution fluxes in the coupled model. The results are presented at continuous monitoring stations with and without filtering of seasonal-scale variability. As measures of the correspondence between models and observations, we chose correlation coefficients and the centred root-mean-square difference (calculated over a 1-yr period). The Fisher r-to-z transformation was used to obtain two-tailed p-values and to assess the statistical significance of the differences between correlation coefficients.

4.1 “Synthetic” test

To demonstrate the differences between the usage of low- and high-resolution CO₂ fluxes and between Eulerian and coupled models, we performed a “synthetic” test that examined the transport of CO₂ around the city of Moscow, where several large power plants are located, emitting strong plumes of CO₂ that are transported to the east of the city by winds. We selected three prospective observation sites (separated from each other by ~50 km) located east of Moscow and performed the calculations. The model results for the three sites are shown in Fig. 1. It is difficult to distinguish concentrations simulated by the Eulerian model (Fig. 1a). For the coupled model and fluxes at a resolution of 1° × 1° (Fig. 1b), the results are similar for all three sites, but a sharper structure is resolved. For a resolution of 1 × 1 km (Fig. 1c), however, the results differ among the sites, clearly showing the impact of the plumes. We note that this case study serves to demonstrate the effect of flux resolution on the model results; observation data are not available from these sites for verification of the results.

4.2 Comparison of simulated concentrations at continuous monitoring stations

To demonstrate the feasibility and advantage of very-high-resolution tracer transport simulations using the coupled Eulerian-Lagrangian approach, we chose several continuous monitoring stations and performed simulations with (a) NIES-TM, (b) the coupled model with 1° × 1° fluxes and (c) the coupled model with 1 × 1 km fluxes.

First we present simulation results at the station Fyodorovskoye, for which we used 3-hourly samples of observations and model results for the year 2008. We performed a correlation analysis between the observed and simulated concentrations, yielding correlation coefficients of 0.40 for
Table 2. General information on the stations.

<table>
<thead>
<tr>
<th>Station</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Instrument height, m</th>
<th>Data period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fyodorovskoye</td>
<td>32.9220</td>
<td>56.4615</td>
<td>27</td>
<td>2008</td>
</tr>
<tr>
<td>MRI, Tsukuba</td>
<td>140.1237</td>
<td>36.0526</td>
<td>200</td>
<td>2009</td>
</tr>
<tr>
<td>Queen’s Tower, London</td>
<td>−0.1768</td>
<td>51.4983</td>
<td>80</td>
<td>Aug 2006–Jun 2007</td>
</tr>
<tr>
<td>Egham, London</td>
<td>−0.5616</td>
<td>51.4266</td>
<td>15</td>
<td>2006–2007</td>
</tr>
</tbody>
</table>

NIES-TM, 0.43 for the coupled model with $1^\circ \times 1^\circ$ fluxes, and 0.46 for the coupled model with $1 \times 1$ km fluxes. In this case, the coupled model outperforms the Eulerian model alone, although the overall correlations are relatively weak.

Similar results were obtained for the station at Egham, London, using hourly samples of observed and modelled concentrations for the year 2007. The correlation coefficients are 0.47 for NIES-TM, 0.54 for the coupled model with $1^\circ \times 1^\circ$ fluxes, and 0.55 for the coupled model with $1 \times 1$ km fluxes. Again, the correlations are relatively weak.

Overall, the correlations between observed and simulated concentrations are not as high as we expected, possibly because the model mixed layer depth was not accurately estimated. To investigate this possibility, we restricted our analysis to observations and model results sampled at hourly intervals in the daytime, because estimations of mixed layer depth are expected to be reasonably accurate during the well-mixed daytime conditions. The data selection could also consider the wind fields and boundary layer height, but we did not perform such extensive sensitivity tests based on model meteorology. In addition, the use of biospheric fluxes with daily variations helps to constrain the calculation of night-time values.

In the case of daytime sampling only, we see an obvious increase in the correlation coefficients between model and observations. For example, in the case of Fyodorovskoye, the correlation coefficients are 0.65 for NIES-TM, 0.68 for the coupled model with $1^\circ \times 1^\circ$ fluxes, and 0.735 for the coupled model with $1 \times 1$ km fluxes. For Egham, the corresponding correlations are 0.66, 0.66, and 0.69, respectively.

We also compared the simulations results obtained using different durations of trajectories in the Lagrangian model. Using 7-day instead of 2-day trajectories, we obtained a minor increase in the correlation coefficient from 0.735 to 0.740 for the $1 \times 1$ km fluxes at Fyodorovskoye. This improvement (by 0.005) is not statistically significant because the Fisher r-to-z transformation gives a two-tailed p-value of 0.7414, which is much greater than 0.05 (the result is generally considered to be statistically significant if the p-value is less than 0.05). Therefore, relatively little is gained despite the much longer computation time required for 7-day trajectories.

For the remaining stations, we considered only daytime sampling of observations. The results are presented in Table 3, and a Taylor diagram (Taylor, 2001) for all of the simulated stations is shown in Fig. 2a. The results show that...
the coupled model is superior to the Eulerian model alone in terms of reproducing the observations. The use of $1 \times 1$ km surface fluxes (instead of $1^\circ \times 1^\circ$ fluxes) results in a higher correlation coefficient between simulations and observations. For most of the stations considered here, the combined model and $1 \times 1$ km emissions inventory resulted in an improvement in the correlation and in the centred root-mean-square (RMS) difference.

4.3 Comparison of modelled and observed high-frequency variability

The high-resolution model performs better than the low-resolution model in representing the high-frequency variability of observed concentrations. The skill of simulations of CO$_2$ concentrations is affected by the quality of atmospheric transport, which dominates the variability at the synoptic scale, and the accuracy of surface fluxes, which affects the seasonal cycle (Patra et al., 2008). Therefore, to evaluate the quality of high-resolution simulations it is advisable to separate the high-frequency synoptic-scale variability from the low-frequency variability related to the seasonal cycle of surface fluxes.

The high-frequency variability was extracted from the observations and from the model results by employing the following equation:

$$C_d = C_i - \frac{1}{n+1} \sum_{j=-n/2}^{i+n/2} C_j$$

where: $C_d$ – de-seasonalized CO$_2$ concentration; $C_i$ – original CO$_2$ concentration; and $C_j$ – average CO$_2$ concentration over a 30-day period. Fifteen-day averaging was performed at the beginning and end of the analysis period.

Table 4 summarizes of correlation coefficients and RMS differences. The results for individual stations are shown in Figs. 3–6 for a period of 4 months. A Taylor diagram (Taylor, 2001) is presented in Fig. 2b. The coupled model outperforms the Eulerian model for all stations. The use of $1 \times 1$ km fluxes further improves the correlations and the RMS differences, although in some cases there is no clear advantage over the use of $1^\circ \times 1^\circ$ fluxes. For the site at Fyodorovskoye and the MRI tower, the correlations are relatively weak for all flux resolutions, possibly due to seasonal variations (which could potentially affect the total correlation).

The increasing misfit at higher resolutions may be explained by the rather coarse meteorology. Our version of the FLEPART model is driven by meteorological fields at a spatial resolution of $1.25^\circ \times 1.25^\circ$ (about 125 km), which is close to the observed horizontal scale of variability in atmospheric winds and temperatures, often expressed as the correlation radius. It is important to use analysed data and emissions at the highest possible resolution, and we are presently working on this problem. The present model is unable to resolve local phenomena such as sea breezes, which can be handled locally if wind fields generated with regional models (e.g. the Weather Research & Forecasting Model, WRF;
Table 3. Information on correlation coefficients (statistical significance between correlation coefficients (for 1°: comparison between NIES-TM and 1°; for 1 km: comparison between 1° and 1 km) and two-tailed p-value between current and previous (from the cell to the left of the current cell in table) correlation coefficients)/centred root-mean-square difference between simulations and observations.

<table>
<thead>
<tr>
<th>Station</th>
<th>NIES-TM</th>
<th>Coupled model (1° surface fluxes)</th>
<th>Coupled model (1 km surface fluxes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fyodorovskoye</td>
<td>0.65/8.6</td>
<td>0.68 (0.11)/8.3</td>
<td>0.73 (0.00)/7.7</td>
</tr>
<tr>
<td>MRI, Tsukuba</td>
<td>0.38/8.2</td>
<td>0.42 (0.14)/8.0</td>
<td>0.47 (0.06)/7.9</td>
</tr>
<tr>
<td>Queen’s Tower, London</td>
<td>0.66/8.9</td>
<td>0.79 (0.00)/7.3</td>
<td>0.81 (0.07)/6.8</td>
</tr>
<tr>
<td>Egham, London, 2006</td>
<td>0.68/10.3</td>
<td>0.78 (0.00)/9.0</td>
<td>0.78 (0.70)/8.8</td>
</tr>
<tr>
<td>Egham, London, 2007</td>
<td>0.66/15.6</td>
<td>0.66 (0.71)/15.4</td>
<td>0.69 (0.06)/14.4</td>
</tr>
</tbody>
</table>

Table 4. Information on correlation coefficients (statistical significance between correlation coefficients (for 1°: comparison between NIES-TM and 1°; for 1 km: comparison between 1° and 1 km) and two-tailed p-value between current and previous (from the cell to the left of the current cell in table) correlation coefficients)/centred root-mean-square difference between simulations and observations. The results have been deseasonalized.

<table>
<thead>
<tr>
<th>Station</th>
<th>NIES-TM</th>
<th>Coupled model (1° surface fluxes)</th>
<th>Coupled model (1 km surface fluxes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fyodorovskoye</td>
<td>0.17/6.1</td>
<td>0.34 (0.00)/5.7</td>
<td>0.34 (0.95)/5.7</td>
</tr>
<tr>
<td>MRI, Tsukuba</td>
<td>0.14/6.8</td>
<td>0.23 (0.01)/6.8</td>
<td>0.33 (0.00)/6.7</td>
</tr>
<tr>
<td>Queen’s Tower, London</td>
<td>0.55/5.5</td>
<td>0.66 (0.00)/4.9</td>
<td>0.69 (0.06)/4.7</td>
</tr>
<tr>
<td>Egham, London, 2006</td>
<td>0.52/7.2</td>
<td>0.73 (0.00)/5.8</td>
<td>0.71 (0.21)/6.0</td>
</tr>
<tr>
<td>Egham, London, 2007</td>
<td>0.52/9.0</td>
<td>0.55 (0.17)/8.7</td>
<td>0.63 (0.00)/8.2</td>
</tr>
</tbody>
</table>

Fig. 6. Deseasonalized results for Egham, London (representative 4-month time series).

available at http://www.wrf-model.org) are used. However, the model works well in the case of stronger winds or if the motion is close to geostrophic. Despite the problems related to coarse wind resolution, the model shows improvements that can be explained by the abovementioned properties of the large-scale circulation. Further progress requires models that provide high-resolution fields. On the other hand, we wish to stress that the main innovation in the present model is the focus on how to efficiently represent and handle kilometre-scale fluxes at the global scale.

5 Conclusions

We demonstrated the feasibility of a very-high-resolution tracer transport simulation with a coupled Eulerian-Lagrangian model extended to a global flux-field resolution of 1 × 1 km. We prepared and tested a high-resolution flux dataset and model framework, and performed simulations of atmospheric CO₂ at selected observational points using (a) a grid-based Eulerian transport model running at a medium resolution of 2.5°, (b) a Lagrangian plume dispersion model and (c) surface fluxes at a resolution of 1 × 1 km. A comparison of modelled and observed CO₂ simulations revealed that the coupled model outperforms the Eulerian model alone. The use of surface fluxes at a resolution of 1 × 1 km has a clear advantage over lower-resolution (1° × 1°) fluxes in reproducing the high-concentration spikes caused by anthropogenic emissions. In some cases, the use of high-resolution fluxes does not result in an improved simulation, yielding a weaker correlation with observed variability as compared with low-resolution fluxes. The simulated concentration variation is higher at high resolutions, resulting in larger misfit in cases where the timing and amplitude of simulated high-concentration events do not match observations. This misfit may explain the misrepresentation of the wind fields and fluxes at high resolutions.

We propose a technique to represent fluxes as a combination of fixed spatial patterns at a high resolution and temporal variability at medium resolution, thereby...
significantly reducing memory and computational demands. This model can be efficiently used to analyse continuous observation data for sites downwind of a large emitting source. The model explicitly treats areas of sharp discontinuities in CO₂ flux, such as coastlines, where many background monitoring sites are located. It is also possible to use the model to simulate and analyse observations from satellites and aircraft, and to perform inverse modelling and validation of high-resolution emission datasets.

Acknowledgements. We thank A. Stohl for providing the FLEXPART model and the JRA-25 long-term reanalysis cooperative research project carried out by the JMA and CRIEPI. V. Valsala acknowledges generous support by the GSAT project at NIES, Tsukuba, Japan. R. J. Andres was sponsored by the US Department of Energy, the Office of Science, and the Biological and Environmental Research (BER) program, and worked at Oak Ridge National Laboratory (ORNL) under US Department of Energy contract DE-AC05-00OR22725.

Edited by: P. Jöckel

References


Koyama, Y., Maksyutov, S., Mukai, H., Thoning, K., and Tans, P.:


Oda, T. and Maksyutov, S.: A very high-resolution (1 km × 1 km) global fossil fuel CO₂ emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543–556, doi:10.5194/acp-11-543-2011, 2011.


Valsala, K. V. and Maksyutov, S.: Simulation and assimilation of global ocean ρCO₂ and air-sea CO₂ fluxes using ship ob-