Adjoint of the Global Eulerian–Lagrangian Coupled Atmospheric transport model (A-GELCA v1.0): development and validation

D.A. Belikov[1,2,3], S. Maksyutov[1], A. Yaremchuk[4], A. Ganshin[3,5], T. Kaminski[6,*], S. Blessing[7], M. Sasakawa[1], Angel J. Gomez-Pelaez[8] and A. Starchenko[3]

[1]{National Institute for Environmental Studies, Tsukuba, Japan}

[2]{National Institute of Polar Research, Tokyo, Japan}

[3]{Tomsk State University, Tomsk, Russia}

[4]{N. Andreev Acoustic Institute, Moscow, Russia}

[5]{Central Aerological Observatory, Dolgoprudny, Russia}

[6]{The Inversion Lab, Hamburg, Germany}

[7]{FastOpt GmbH, Hamburg, Germany}

[8]{Izaña Atmospheric Research Center (IARC), Meteorological State Agency of Spain (AEMET), Izaña, 38311, Spain}

[*]{Previously at: FastOpt GmbH, Hamburg, Germany }

Correspondence to: D.A. Belikov (dmitry.belikov@nies.go.jp)

Abstract

We presented the development of the Adjoint of the Global Eulerian–Lagrangian Coupled Atmospheric (A-GELCA) model that consists of the National Institute for Environmental Studies (NIES) model as an Eulerian three-dimensional transport model (TM), and FLEXPART (FLEXible PARTicle dispersion model) as the Lagrangian Particle Dispersion Model (LPDM). The forward tangent linear and adjoint components of the Eulerian model were constructed directly from the original NIES TM code using an automatic differentiation tool known as TAF (Transformation of Algorithms in Fortran; http://www.FastOpt.com), with additional manual pre- and post-processing aimed at improving transparency and clarity of the code and optimizing the performance of the computing, including MPI (Message Passing Interface). The Lagrangian component did not require any code modification, as LPDMs are self-adjoint and track a significant number of particles backwad in time in order to calculate the sensitivity of the observations to the neighboring emission areas. The constructed Eulerian adjoint was coupled with the Lagrangian component at a time boundary in the global domain. The simulations presented in this work were performed using the A-GELCA model in forward and adjoint modes. The forward simulation shows that the coupled model improves reproducing of the seasonal cycle and short-term variability of CO_2. The adjoint of the Eulerian model was shown, through several numerical tests, to be very accurate compared to direct forward sensitivity calculations. The developed adjoint of the coupled model combines the flux conservation and stability of an Eulerian discrete adjoint formulation with the flexibility, accuracy, and high resolution of a Lagrangian backward trajectory formulation. A-GELCA will be incorporated into a variational inversion system designed to optimize surface fluxes of greenhouse gases.

Keywords: atmospheric transport and inverse modeling, adjoint model, carbon cycle

1. Introduction

Forecasts of CO_2 levels in the atmosphere and predictions of future climate depend on our scientific understanding of the natural carbon cycle (IPCC, 2007; Peters et al., 2007). To estimate the spatial and temporal distribution of carbon sources and sinks, inverse methods are used to infer carbon fluxes from geographically sparse observations of the atmospheric CO_2 mixing ratio (Tans et al., 1989). The first comprehensive efforts in atmospheric CO_2 inversions date back to the late 1980s and early 1990s (Enting and Mansbridge, 1989; Tans et al., 1989). With the increase in spatial coverage of CO_2 observations and the development of three-dimentional (3-D) tracer transport models, a variety of numerical experiments and projects have been performed by members of the so-called "TransCom" community of inverse modelers (e.g., Law et al., 1996, 2008; Denning et al., 1999; Gurney et al., 2002, 2004; Baker et al., 2006; Patra et al., 2011). A number of studies have proposed improvements to the inverse methods of atmospheric transport, i.e. the efficient computation of the transport matrix by the model adjoint proposed by Kaminski et al. (1999b), use of monthly mean GLOBALVIEW-CO_2 ground-based data (current version is for 2014) by Rödenbeck et al. (2003), development an ensemble data assimilation method by Peters et al. (2005), flux inversion at high temporal (daily) and spatial (model grid) resolution using for the first time of continuous CO_2 measurements over Europe by Peylin et al. (2005), use satellite data to constrain the inversion of CO_2 by Chevallier et al. (2005), develop of a new observational screening technique by Maki et al. (2010). Despite progress in atmospheric CO_2 inversions, a recent intercomparison (Peylin et al., 2013) demonstrated the need for further refinement.

In recent decades, the density of the observational network established to monitor greenhouse gases in the atmosphere has been increased, and more measurements taken onboard ships and aircraft are becoming available (Karion et al., 2013; Tohjima et al., 2015). However, on a global scale CO_2 observations do not exist for many remote regions not covered by networks. This lack of data is one of the main limitations of atmospheric inversions, which can be filled by monitoring from space (Rayner and O'Brien, 2001). The satellite observation data from current (GOSAT, Kuze et al., 2009; Yokota et al., 2009; OCO-2, Crisp et al., 2004) and future missions (CarbonSat/CarbonSat Constellation; Bovensmann et al., 2010; Buchwitz et al., 2013) offer enormous potential for CO_2 inverse modeling. Optimal application of large observed datasets requires expanding the inverse analysis of CO_2 to finer resolution, higher precision and faster performance.

To link surface fluxes of CO_2 to observed atmospheric concentrations, an accurate model of atmospheric transport and an inverse modeling technique are needed. Generally, the atmospheric constituents transport may be described in two different ways: the Lagrangian and the Eulerian approaches. The Eulerian method treats the atmospheric tracers as a continuum on a control volume basis, so it is more effective reproducing of long-term patterns, i.e. the seasonal cycle or the interhemispheric gradient. The Lagrangian Particle Dispersion Models (LPDMs) consider atmospheric tracers as a discrete phase and tracks each individual particle, therefore LPDMs are better for resolving synoptic and hourly variations.

To relate fluxes and concentrations of long-lived species like CO_2, a transport model must cover a long simulation period (e.g., Bruhwiler et al., 2005). Therefore, computing time is a critical issue and minimization of the computational cost is essential. For chemically inert tracers, the transport can be represented by a model's Jacobian matrix, because the simulated concentration at observational sites is a linear function of the flux sets. Theoretically, to compute such matrix the transport model is run multiple times with set of prescribed surface fluxes. However, this would require an extremely large number of forward model evaluations. The adjoint of the transport model is an efficient way to accelerate calculation of concentration gradient of the simulated tracer at observational locations (Kaminski et al., 1999). Marchuk (1974) first applied the adjoint approach in atmospheric science. After that, this method became widely used in meteorology. In the 1990s the use of this approach was expanded to the field of tracer transport modeling (Elbern et al., 1997; Kaminski et al., 1999).

Adjoint models have numerous applications, including the data assimilation of concentrations, inverse modeling of chemical source strengths, sensitivity analysis, and parameter sensitivity estimation (Enting, 2002; Haines et al., 2014). Recent studies have used this method to constrain estimates of the emissions of CO_2 using retrieved column integrals from the GOSAT satellite (Basu et al., 2013; Deng et al., 2014; Liu et al., 2015).

Using the adjoint model speeds up the process of inverse modeling. However, high CPU and memory demands prevent us from using Eulerian chemical transport models (CTMs) with high-resolution grids in inversions. It would be beneficial to increase the model resolution close to observation points, where the strong observation constraint can significantly improve the optimization of the resulting emission fluxes.

LPDM running in the backward mode can explicitly estimate a source–receptor sensitivity matrix by solving the adjoint equations of atmospheric transport (Stohl et al., 2009), which is mathematically presented by a Jacobian expressing the sensitivity of concentration at the observational locations. Marchuk (1995), and Hourdin and Talagrand

(2006) provided derivations proving equivalence of the adjoint of forward transport models to backward transport models.

In order to exploit the advantages of both methods, Lagrangian and Eulerian chemical transport models can be coupled to develop an adjoint, that is suitable for the simultaneous simulation of contributions from global and regional emissions. Coupling can be performed in several ways; e.g., a regional-scale LPDM can be coupled to a global Eulerian model at a regional domain boundary (Rödenbeck et al., 2009; Rigby et al., 2011), or a global-scale LPDM can be coupled to an Eulerian model at the time boundary (Koyama et al., 2011; Thompson and Stohl, 2014).

The goal of this study is to present the development and evaluation of an Adjoint of the Global Eulerian–Lagrangian Coupled Atmospheric model (A-GELCA), which consists of an Eulerian National Institute for Environmental Studies global Transport Model (NIES-TM; Maksyutov et al., 2008; Belikov et al., 2011, 2013a, 2013b) and a Lagrangian particle dispersion model (FLEXPART; Stohl et al., 2005). This approach utilizes the accurate transport of the LPDM to calculate the signal near to the receptors, and efficient calculation of background responses using the adjoint of the Eulerian global transport model. In contrast to previous works (Rödenbeck et al., 2009; Rigby et al., 2011; Thompson and Stohl, 2014), in which the regional models were coupled at the spatial boundary of the domain, we implemented a coupling at a time boundary in the global model domain (as described in Sect. 2.1). A-GELCA can be integrated into a variational inverse modeling system designed to optimize surface fluxes.

The remainder of this paper is organized as follows. An overview of the coupled model is provided in Sect. 2. In Sect. 3 we describe the variational inversion scheme. In Sect. 4 we address several problems regarding the coupled model that have not been covered previously (Ganshin et al., 2012). In Sect. 5 we describe the formulation and evaluation of the adjoint model. The computational efficiency of the adjoint model is analyzed in Sect. 6, and the conclusions are presented in Sect. 7.

2. Model and method

2.1. Global coupled Eulerian-Lagrangian model

In this paper we use a global Eulerian-Lagrangian coupled model, the principles of which are described by Ganshin et al. (2012). The coupled model consists of FLEXPART (version 8.0; run in backward mode) as the Lagrangian particle dispersion model, and NIES TM (version NIES-08.1i) as the Eulerian off-line global transport model. For concentration $C(x_r, t_r)$ (mole fraction) at receptor point x_r and time t_r we provide the equation in its discrete form, as implemented in the model for the case of surface fluxes:

$$C(x_r, t_r) = {}^{Tm_{air}}\!\Big/_{hNS\rho m_{CO_2}} \sum_{ij}^{IJ} \sum_{s=0}^{S} F_{ij}^{S} \sum_{n=1}^{N} f_{ij}^{sn} + {}^{1}\!/_{N} \sum_{ijk}^{IJK} C_{ijk}^{B} \sum_{n=1}^{N} f_{ijk}^{n}, \tag{1}$$

where *i, j,* and *k* are the indices that characterize the location of each grid cell; *s* is the time index; F_{ij}^{l} are the surface fluxes in $kg{\cdot}m^{-2}{\cdot}s^{-1}$; C_{ijk}^{B} are the background concentrations calculated by the Eulerian model at the coupling time; f_{ijk}^{n} equals unity if the particle is within cell *i, j, k*, otherwise it equals zero; *T* is the duration of the backward trajectory; *S* is the number of steps in time; *N* is the total number of particles; *h* is the height up to which the effect of the surface fluxes is considered significant; ρ is the average air density below height *h*; and m_{air} and m_{CO2} are the molar masses of air and carbon dioxide, respectively. The first term in this formula describes the contribution of the nearby sources of the considered component; these sources are located along the trajectories inside layer h (500 m). The value of the first term is proportional to the flux in each cell along the trajectory, and to the time during which the air particle is inside this cell (Ganshin et al., 2012). The background grid values of the concentrations (calculated by the Eulerian model), which are interpolated to the final points of the backward trajectories, are transferred to the observation point and are the second term in the right-hand side of Eq. (1). The FLEXPART model starts simulation at the observation point and calculates seven-day backward trajectories for 1000 air particles, which are dispersed under the influence of turbulent diffusion. The number of particles has been chosen to optimize the computational cost without compromising the quality of modeling by Ganshin et al., (2013). The scheme of concentration calculation for the given location includes coupling of two model approaches. NIES TM calculates global concentrations for the selected time period (usually 1 year to exclude spin-up effect), but stops 7 days before the time of the observations. To obtain the concentrations for the observation time we transport the background concentrations from NIES TM gridbox to the location of observation point along the trajectory ensemble calculated by FLEXPART model and add contribution from surface sources. Therefore we have implemented the coupling at a time boundary in the global domain of the NIES transport model, while nested regional modeling systems such as one by Rodenbeck et al (2009) have to couple at both region boundary and time boundary.

Since the first publication of the GELCA model in 2012, the NIES transport model has undergone significant updates. We provide a brief outline of the major features of the current model. NIES TM is a global three-dimensional CTM that simulates the global distribution of atmospheric tracers between the Earth's surface and a pressure level of 5 hPa. The model employs the standard horizontal latitude–longitude grid with reduced number of meshes towards the poles and a spatial resolution of 2.5° × 2.5° near the equator(Belikov et al., 2011). The vertical coordinate is a flexible hybrid sigma–isentropic (σ–θ) with 32 levels (Belikov et al., 2013b). To parameterize turbulent diffusivity we follow the method proposed by Hack et al. (1993), with a separate evaluation of transport processes in the free troposphere and the planetary boundary layer (PBL). The PBL heights are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis. The modified Kuo-type parameterization scheme is used for cumulus convection (Belikov et al., 2013a).

Inverse modeling assumes that the model reasonably well reproduces the relationship between atmospheric mixing ratio and surface fluxes, assuming that the biases between the simulated and observed concentrations are mostly due to the emission inventories errors. To ensure that this is the case, the NIES TM model has been evaluated extensively. Comparisons against SF_6 and CO_2 (Belikov et al., 2011, 2013b), CH_4 (Patra et al., 2011; Belikov et al., 2013b), and ^{222}Rn (Belikov et al., 2013a) measurements show the model ability to reproduce seasonal variations, interhemispheric gradient and vertical profiles of tracers.

2.2. FLEXPART

FLEXPART, like other LPDMs, considers atmospheric tracers as clouds of individual particles and tracks the pathway of each particle. The advantage of this approach is the direct estimation of the sensitivity of the measurements to the neighboring sinks and sources by tracking the particles backward in time. Usually it is sufficient to simulate for a limited number of days (2-10) to determine where particles intercept the surface layer before they spread vertically and horizontally.

2.3. Meteorological data

To run both models we use reanalysis dataset combining the Japanese 25-yr Reanalysis

(JRA-25) and the Japanese Meteorological Agency Climate Data Assimilation System (JCDAS) dataset (Onogi et al., 2007). The JRA-25/JCDAS dataset is distributed on a Gaussian T106 grid with horizontal resolution 1.25° × 1.25°, 40 sigma-pressure levels and in 6-hour time steps. The use of JRA-25/JCDAS data for Eulerian and Lagrangian models provides consistency in the calculated fields; however, some features of FLEXPART and NIES TM require different methods for processing the meteorological data.

2.3.1. Meteorological data processing for NIES TM

Isolation of the transport equations is an effective way to save a significant amount of CPU time during tracer transport simulation. At the preprocessing stage, the NIES TM core produced a static archive of advective, diffusive, and convective mass fluxes with time step similar to the one of the original JRA-25/JCDAS data (6 hour). After that the archive is used by an "offline" model specially designed only for passive transport of tracer. Intermediate fluxes are derived by interpolation.

Besides the mass fluxes, the static archives contain fields of temperature, pressure, humidity, vertical grid parameters (variation of the sigma-isentropic vertical coordinate over time), and others. The pre-calculated and stored data field can be used directly for any of the inert tracers. It is also possible to simulate chemically active tracers if the chemical reaction can be written in the linear decay form; e.g., for ^{222}Rn, CH_4. Approximately 20 3-D and 1-dimensional arrays are written to a hard disk for every record. This comprises around 10 GB of data per modelled month for the model's standard resolution of 2.5° × 2.5°.

2.3.2. Meteorological data processing for FLEXPART

Originally, FLEXPART was driven by ECMWF reanalysis dataset distributed on a grid with regular latitude–longitude horizontal structure and sigma–pressure vertical coordinate. The current version of the model was adapted to use JRA-25/JCDAS data, by horizontal bilinear interpolation of the required parameters from a Gaussian grid to a regular 1.25 × 1.25 grid. The vertical structure and temporal resolution of JRA-25/JCDAS data were used without modification.

Given the large differences in structure, resolution and parameter estimation methods used in different reanalysis dataset, the use of the same meteorology for both Eulerian and Lagrangian models provides significant benefit.

3. Inverse modeling for the flux optimization problem

Although the variational inversion method for minimizing the discrepancy between modeled and observed mixing ratios has been well described and published (i.e. Chevallier et al., 2005), we summarize it here.

The aim of the inversion problem is to find the value of a state vector **x** with *n* elements that minimizes the cost function *J(x)*:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\mathbf{Hx}-\mathbf{y})^T \mathbf{R}^{-1}(\mathbf{Hx}-\mathbf{y}), \quad (2)$$

where **y** is a vector of observations with *m* elements, and the matrix **H** represents the forward model simulation mapping the state vector **x** to the observation space. Here, **R** is the covariance matrix (size $m \times m$) for observational error, which includes instrument and representation errors. The matrix **R** also includes errors of the forward model **H**. **B** is the covariance matrix (size $n \times n$) of error for prior information of the state vector $\mathbf{x}_b$. The use of the cost function in the form of Eq. (2) assumes that all errors have Gaussian statistics and are unbiased (Rodgers, 2000).

The minimization of the cost function (Eq. 2) has an analytic solution that involves a matrix inversion. If the Jacobian **H** is available this analytic solution can implemented, unless the matrix sizes are too large for the available computing resources. Alternatively, Eq. 2 can be solved through an iterative minimization algorithm. In this case, the existence of the gradient of *J(**x**)* with respect to **x** allows using of powerful gradient algorithms for minimization. This gradient is efficiently provided by the adjoint (Giering and Kaminski, 1998; Kaminski et al., 1999; Chevallier et al., 2005).

4. Assessment of the coupled model

The effect of different horizontal resolutions on Eulerian models is discussed in detail by Patra et al. (2008). In general, higher resolution helps to resolve a more detailed distribution of the tracer. However, the use of a higher resolution grid leads to additional computational cost, which is not always justified by the resulting model output. Higher resolution does not produce better results largely due to the limited availability of high-resolution meteorology and tracer emission datasets.

The paper by Ganshin et al. (2012) describing the development of the GELCA model provides a model testing report. The advantage of GELCA in reproducing the high-concentration spikes and short-term variations caused mainly by anthropogenic emissions is more vivid when using high resolution (1 km × 1 km) surface fluxes compared to standard resolution (1° × 1°) fluxes. However those tests considered only short 4-month simulations for a limited number of locations.

We expanded the comparison undertaken by Ganshin et al. (2012) to a two-year period using an updated set of prescribed fluxes, which combines four components similar to the analysis performed by Takagi et al. (2011) and Maksyutov et al. (2013): (a) anthropogenic fluxes from the Open source Data Inventory of Anthropogenic CO_2 (ODIAC; Oda and Maksyutov, 2011) and the Carbon Dioxide Information Analysis Center's (CDIAC; Andres et al., 2009, 2011) datasets; (b) biosphere fluxes simulated by the Vegetation Integrative SImulator for Trace gases (VISIT) terrestrial biosphere model (Ito, 2010; Saito et al., 2011, 2013); (c) oceanic fluxes predicted by a data assimilation system based on the Offline ocean Tracer Transport Model (OTTM; Valsala and Maksyutov, 2010); and (d) biomass burning emissions from the Global Fire Emissions Database (GFED) version 3.1 (van der Werf et al., 2010). Biosphere fluxes have daily time step, while the others are monthly. The initial global CO2 distribution was obtained from GLOBALVIEW-CO2 (2014).

We considered several cases with different model resolutions. For NIES TM we tested grids at 10.0°, 2.5°, and 1.25° resolutions, with FLEXPART running at 1.0° (Table 1). The resolution of the input fluxes was matched to that of FLEXPART. Model results were compared with observations from the World Data Centre for Greenhouse Gases (WDCGG 2015) and the Siberian observations obtained by the Center for Global Environmental Research (CGER) of the National Institute for Environmental Studies (NIES) and the Russian Academy of Science (RAS), from six tower sites (JR-STATION) as described by Sasakawa et al. (2010). The selected site locations are shown in Fig. 1.

Although the total number of observational stations contributing to the WDCGG is about several hundreds, the set of sites conducting continuous (high temporal resolution is needed for the coupled model) observations is much smaller. We selected 19 sites (Table 2). Most of them are concentrated in the temperate latitudes of the northern hemisphere, where the variations in CO_2 concentration are most noticeable.

Siberia is assumed to be a substantial source and sink of CO_2, with high uncertainties in the fluxes describing them (McGuire et al., 2009; Hayes et al., 2011; Saeki et al., 2013). As a result, CTMs tend to reproduce the interannual variability of CO_2 quite poorly. We selected six tower JR-STATION sites to check the model performance in the Siberian region (Table 3).

The analyzed sites are divided into three groups. The first group includes remote and marine sites (ALT, AMS, BRW, CPT, IZO, JBN, MLO, MNM, ZEP) with very weak influence of local sources, so the seasonal variation of CO2 is controlled by global, large-scale variations. For these sites contribution by using the Lagrangian component is negligible (see Fig. 2-4 panel b to analyze the difference between the coupled and the Eulerian models).

The second group includes sites with domination of long term variability of CO_2 and relatively smooth and weak short term variations. Typically, these sites are located on the border of two regions with very different fluxes (AMY, CMN, MHD, PAL, PRS, YON).

The sites selected to the third group are strongly influenced by local emissions and global transport at the same time. Therefore the CO_2 concentration variation is controlled by the strength and direction of wind, the depth of the boundary layer and other factors. Such sites are mainly in the northern mid-latitudes (HUN, PUY, SSL, WSA) including all Siberian towers (DEM, IGR, KRS, NOY, VGN, YAK). For these locations contributions of the Eulerian and Lagrangian components are comparable. Therefore, the coupled model introduces the most significant improvement when simulating CO_2 for these sites.

Figures 5 compares the coupled and Eulerian model results with observations from the Igrim and Vaganovo towers. The recent modifications indicated in Sect. 2.2 have significantly improve the performance of NIES TM compared with the results reported by Ganshin et al. (2012). However, compared to the updated NIES TM the coupled model is better reproducing short term peaks of concentration. This explains the observed reduction of the mean bias and STD (up to 1.5 ppm), and the better simulation of the seasonal variation (in phase and amplitude). The improvements in the CO2 simulations due to the addition of the Lagrangian component to the Eulerian model are higher than those obtained by increasing the resolution of the Eulerian NIES transport model (Fig. 2-4). Given the huge difference in computation costs between NIES TM for low- and high-resolution grids (i.e. a difference by a factor of ~15 between grids with resolution 10.0° and 2.5°), the advantage of the GELCA model is clear. Performance is important, as the setup considered here is almost identical to that used in the inverse modeling of CO_2.

However, improvements in CO_2 simulation due to the implementation of the GELCA model were obtained not for all the considered sites. This shows that further modification of the setup (i.e. more detail meteorological data, switch to higher resolution) is necessary. Nevertheless, the coupled model is an effective way to improve simulation of CO_2 without increasing the resolution of the Eulerian model. We recognize that is quite problematic to use the highly uncertain surface fluxes to simulate the tracer concentrations and use these concentrations for estimating the quality of different model configurations. Nevertheless, we cannot improve our analysis, because we do not have concentration measurements for tracers whose surface fluxes are more accurately known, like SF6.

5. Construction and validation of the adjoint model

5.1. Construction

In this section, we present the development of the adjoint of the coupled model. The incorporation of the Lagrangian component does not require any modification to the code, as LPDMs are self-adjoint. The development of the adjoint of the Eulerian part is more complicated. We decided to develop a discrete adjoint of NIES TM in order to make it consistent with the forward model. An alternative approach is the construction of a continuous adjoint derived from the leading equations of the forward model (Giles and Pierce, 2000). The main advantage of the discrete adjoint model is that the resulting gradients of the numerical cost function are exact, even for nonlinear or iterative algorithms, and this makes easier to validate the adjoint model, which is an essential and complicated task.

The adjoint model for NIES TM was created manually to achieve maximum computational efficiency, while the adjoint of NIES TM to FLEXPART coupler was created using the Transformation of Algorithms in Fortran (TAF) software (http://www.FastOpt.com). However, the use of this tool required some manual treatment of the code. TAF successfully produces the tangent linear and adjoint code of individual procedures, but it gets confused when the model has complex structures (such as loops and conditional operators). Therefore we often manually redesigned and optimized the automatically generated adjoint code to optimize the efficiency, improve readability and clarity of the adjoint model and optimize the performance of computing using MPI, as the TAF code used here (version 1.5) do not fully support MPI routines.

The advantages of our coupled adjoint model are as follows.

1. Simple incorporation of the Lagrangian part, since no modification of the LPDM is required. Potentially, NIES TM can be coupled to any Lagrangian model.
2. Minimization of the simulation time can be obtained, as once calculated the output from the Lagrangian model is applicable for different long-lived tracers.
3. Reduction of aggregation errors can be achieved, as the sensitivity for small regions and even individual model cells near to observation sites is estimated using the LPDM part, while the sensitivity for large regions remote from the monitoring sites is derived using

the Eulerian part (Kaminski et al., 2001).

4. Minimization of the computational cost can be obtained, as high-resolution simulation are performed over a limited number of regions nearby to the observational sites using the LPDM part, while for the rest of the globe the coarse-resolution results are calculated by the Eulerian part.
5. High consistency of the tracer fields calculated by the Lagrangian and the Eulerian models due to the fact that both models use the same input meteorology.

The main drawback of the method is that the deriving of discrete adjoint of Eulerian model is a significant technical challenge.

5.2. Validation of the coupled adjoint

An essential stage of the adjoint model construction is its validation. A lack of accuracy in the adjoint model will likely degrade the performance of the cost function minimization (Eq. 2). Several different tests were carried out to evaluate the accuracy and precision of the constructed adjoint model. Considering the simple formulation of the Lagrangian part, we focused on testing the NIES TM adjoint.

5.2.1. Validation of the NIES TM adjoint

The discrete adjoint obtained through automatic differentiation can be easily validated by comparing the adjoint sensitivities with forward model gradients calculated using the finite difference approximation (Henze at al., 2007).

The forward model sensitivity, λ_F, is calculated using the one- or two-sided finite difference equation,

$$\lambda_F = \frac{M'(x+\varepsilon)-M'(x)}{\varepsilon} \quad (3)$$

$$\lambda_F = \frac{M'(x+\varepsilon)-M'(x-\varepsilon)}{2\varepsilon} \quad (4)$$

where *M*` denotes the tangent linear model. A range of ε = 0.1–0.01 was proved in most cases to give an optimal balance between truncation and roundoff error (Henze at al., 2007).

In the first test, adjoint simulations were carried out using an initial CO_2 distribution, zero surface flux for 2 days (1-2 January 2010) and a horizontal grid with resolution 2.5° × 2.5°. The adjoint gradient was then compared with that from the finite difference calculated using Eq. (3). This equation was selected in order to save CPU time by minimizing the number of forward model function calculations. For this test we used ε = 0.01.

To quantify the difference between the two calculations of the sensitivity λ, we define the local relative error

$$E(lon, lat) = \frac{|\lambda_A - \lambda_F|}{max\ \lambda_A}, \qquad (5)$$

where the subscripts A and F refer to adjoint and finite difference respectively, whereas *lon* and *lat* refer to longitude and latitude, respectively. Figure 6c shows *E(lon, lat)* with a logarithmic color scale. The sensitivities obtained for the adjoint have maximum relative error of order 10^{-16}, indicating that transport in the NIES TM adjoint is correct over short timescales. The overall comparisons did not seriously change if we select different grid cells or use other values of ε.

The definition of the adjoint of the tangent linear forward model M* requires that for an inner product $\langle\ ,\ \rangle$ and two random vectors **u** and **v**, the following expression should hold:

$$\forall \mathbf{u}, \forall \mathbf{v}\ \ \langle M'\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, M^{*}\mathbf{v}\rangle. \qquad (6)$$

For practical use the identity in Eq. (6) is rewritten as follows (Wilson et al., 2014):

$$\frac{\|M'(\mathbf{u})\|^2}{\left(\mathbf{u}, M^{*}\left(M'(\mathbf{u})\right)\right)} = 1. \qquad (7)$$

We use Eq. (7) to test the adjoint model initialized using several different random random vectors **u** and **v**. For all cases, Eq. (7) compares well within machine epsilon with mismatch between $-3e^{-14}$ to $6e^{-14}$.

5.2.2. Real case simulation

The next series of calculations was made for real measurements. We used data from the Siberian observation network (Table 3) for the period 1–4 January 2010. CO_2 initial conditions and fluxes were the same as those used for the CELGA forward simulations in Section 4. We run A-GELCA using grids of 10.0° and 2.5° for Eulerian part and of 1.0° for Lagrangian component (similar to Cs-1 and Cs-2 in Table 1) and considered several cases.

The sensitivities of CO_2 concentrations were calculated using the Eulerian component only in Figs. 7,8 a) (resolution of 2.5º), b) (resolution of 10.0º), using the Lagrangian component only in Figs. 7,8 c)(resolution of 1.0º), and d) (resolution of 1.0º, but aggregated on a grid with resolution of 2.5º), and using the coupled adjoint model in Fig. 7,8 e) (Eulerian component at a resolution of 2.5º and the Lagrangian component aggregated on the grid with a resolution of 2.5º), and f) (as for e) , but the resolution of the Eulerian adjoint model was 10.0º). Figure 7 corresponds to the 2-nd day of simulation, while Figure 8 is for 4-th day.

Above, we have already stated that the Eulerian part of the coupled model is more effective in reproducing of long-term patterns, while the Lagrangian part is better for resolving synoptic and hourly variations. This follows from the fact that the A-GELCA components have different footprints. The Eulerian adjoint has a wider footprint, with the greatest values in an area where the effect of all stations is summed. The Euler model monitors global and large-scale changes, although some stations can be outside this zone (i.e. YAK, at Fig. 7a,g or NOY, at Fig. 8a,b). These figures illustrate why the Eulerian model, even with a sufficiently detailed grid, fails to reproduce CO_2 variations (Sect. 4). The footprint width decreases when the NIES TM resolution is increased, but the value of the sensitivity increases.

The FLEXPART model sensitivity shows more irregular distributions, and higher values closer to the observational sites, thereby reflecting the model's ability to monitor small-scale changes (Fig. 7-8 panels c,d).

During coupling, the sensitivity is aligned due to the crosslinking of components (Fig. 7-8 panels e,f). Thus, the intensity has maximum near the stations and smoothly decreases when distance increases. The Eulerian and Lagrangian models employ different approaches and grid resolutions for the modeling of atmospheric tracers, and can thus resolve processes with different time and spatial scales, and underlying physics. By changing the Eulerian model resolution, it is possible to change size of the footprint. This system can utilize responses calculated at higher resolutions, such as 0.5° or 0.1°, but these setups require more accurate driving data and regular observations available for smaller time steps.

6. Computational efficiency

We tested several different methods to reduce the computational cost of the adjoint model. First, the Eulerian part of the adjoint model was driven by static archives of meteorological parameters, as described in Sect. 2.4.1. Second, the forward NIES model was altered so that at each model timestep it saved any variables that were also needed by the adjoint model. Therefore, these variables did not have to be recalculated for being used in the adjoint model. This was possible because we used a discrete version of the adjoint, which was fully compatible with the forward model. Third, the Lagrangian part of the adjoint model made use of pre-calculated response functions, as described in Sect. 2.4.2.

To run the adjoint model we used a Linux workstation with 8 Intel(R) Xeon(R) E5-4650 2.70 GHz processors and 64 GB of RAM. The CPU time of the adjoint model (backward only) was almost equal to CPU time required to run the forward model. It took about 1.3 min for a week-long iteration (forward and backward). The memory demand was about 1 GB. Henze et al. (2007) reports that the ratio between simulation time in backward and forward modes for adjoint models derived for other CTMs, as follows: GEOS-Chem: 1.5, STEM: 1.5, CHIMERE: 3–4, IM-AGES: 4, Polair: 4.5–7, and CIT: 11.75. Thus, the adjoint of the developed coupled model GELCA is quite efficient. To achieve this level of efficiency, a substantial amount of manual programming effort is required, despite the automatic code generated by TAF. The main disadvantage of TAF is that many redundant recomputations are often generated by the compiler. A crucial optimization of the adjoint code is required to eliminate these extra recomputations.

7. Summary

In this paper we have presented the construction and evaluation of an adjoint of the global Eulerian–Lagrangian coupled model GELCA that will be integrated into a variational inverse system designed to optimize surface fluxes. The coupled model combines the NIES three-dimensional transport model as its Eulerian part and the FLEXPART plume diffusion model as its Lagrangian component. The Eulerian and Lagrangian components are coupled at a time boundary in the global domain. The model was originally developed to study the carbon dioxide and methane atmospheric distributions.

The Lagrangian component did not require any code modification, as FLEXPART is a self-adjoint and tracks a significant number of particles backward in time in order to calculate the sensitivity of observations to the neighboring emissions areas.

For Eulerian part, the discrete adjoint was constructed directly from the original NIES TM code, instead of contrasting a continuous adjoint derived from the forward model basic equations. The tangent linear and adjoint models of the NIES TM to FLEXPART coupler were derived using the automatic differentiation software TAF (http://www.FastOpt.com), which significantly accelerated the development. However, considerable manual processing of forward and adjoint model codes was necessary to improve the transparency and clarity of the model and to optimize the computational performance of, including MPI, as the TAF code used here (version 1.5) does not fully support MPI routines.

The main benefit of the developed discrete adjoint is accurate calculation of the numerical cost function gradients, even if the algorithms are nonlinear. The overall advantages of the developed model also include relatively simple incorporation of the Lagrangian part and fast computation using the Lagrangian component, scalability of sensitivity calculation depending on distance to monitoring sites, thereby reducing aggregation errors, and computational efficiency even for high-resolution simulations.

The transport scheme accuracy of the forward coupled model was investigated using the distribution of the atmospheric CO_2. The GELCA components and the model itself had previously been validated using various tests and by comparison with measurements and with other transport models for CO_2 and other tracers. The analyses in the present paper have shown that CELGA is effective in capturing the seasonal variability of atmospheric tracer at observation sites. Decreasing of the Eulerian model resolution does not significantly distort the transport model performance; however, running the coupled model using NIES TM with low resolution grid can maximize simulation speed and use of data storage.

The Eulerian adjoint was validated using various tests in which the adjoint gradients were compared to gradients calculated with numerical finite difference. We evaluated each routine of the discrete adjoint of the Eulerian model and the adjoint gradients of the cost function. The precision obtained of the results of the considered numerical experiments demonstrates proper construction of the adjoint.

The CPU time needed by the adjoint model is comparable with those of other models, as we used several methods to reduce the computational cost. The forward NIES model was altered so that at each model time step it saved all variables that were also being needed by the adjoint model. These variables therefore did not have to be recalculated for use in the adjoint model. In addition, the adjoint simulation was isolated from the recalculation of NIES TM meteorological parameters and Lagrangian response functions. All supplementary parameters were pre-calculated before running the adjoint and were stored in static archives.

The developed A-GELCA model will be incorporated into a variational inversion system aiming studying greenhouse gases (mainly CH_4 and CO_2), by assimilating tracer measurements from *in situ*, aircraft and remote sensing observations. However, before performing real inverse modeling simulations it is necessary to select a proper minimization program and find the optimal values for the error covariance matrices **R** and **B**.

Code availability

All code in the current version of the NIES forward model is available on request. Any potential user interested in these modules should contact D. Belikov (dmitry.belikov@nies.go.jp) or S. Maksyutov (shamil@nies.go.jp), and any feedback on the modules is welcome. Note that that potential users may need help using the forward and adjoint model effectively, but open support for the model is not available due to lack of resources. The code of the adjoint part of the current NIES model is unavailable for distribution, as it was generated using the commercial tool TAF (http://www.FastOpt.com). However, we can provide the sources which were used as input for TAF.

The FLEXPART code was taken from the official web site http://flexpart.eu/. The procedures necessary to run FLEXPART with the JCDAS reanalysis are also available upon request.

Acknowledgments

The authors thank A. Stohl for providing the FLEXPART model. We also thank T. Machida for Siberian observation data (downloaded from http://db.cger.nies.go.jp/). The JRA-25/JCDAS meteorological datasets used in the simulations were provided by the Japan Meteorological Agency. The WDCGG observation data used in the comparisons were provided by The World Data Centre for Greenhouse Gases. The computational resources were provided by NIES. This study was supported by order of the Ministry for Education and Science of the Russian Federation No. 5.628.2014/K, by the Tomsk State University Academic D.I. Mendeleev Fund Program in 2014–2015 and by GRENE Arctic project.

References

Andres, R. J., Boden, T. A., and Marland, G.: Annual fossil-fuel CO2 emissions: Mass of emissions gridded by one degree latitude by one degree longitude. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A., doi 10.3334/CDIAC/ffe.ndp058.2009, 2009.

Andres, R. J., Gregg, J. S., Losey, L., Marland, G., Boden, T.: Monthly, global emissions of carbon dioxide from fossil fuel consumption, Tellus 63B, 309–327, 2011.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi:10.1029/2004GB002439, 2006.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO2 fluxes estimated from GOSAT retrievals of total column CO2, Atmos. Chem. Phys., 13, 8695-8717, doi:10.5194/acp-13-8695-2013, 2013.

Belikov, D., Maksyutov, S., Miyasaka, T., Saeki, T., Zhuravlev, R., and Kiryushov, B.: Mass-conserving tracer transport modelling on a reduced latitude-longitude grid with NIES-TM, Geosci. Model Dev., 4, 207–222, 2011.

Belikov, D. A., Maksyutov, S., Krol, M., Fraser, A., Rigby, M., Bian, H., Agusti-Panareda, A., Bergmann, D., Bousquet, P., Cameron-Smith, P., Chipperfield, M. P., Fortems-Cheiney, A., Gloor, E., Haynes, K., Hess, P., Houweling, S., Kawa, S. R., Law, R. M., Loh, Z., Meng, L., Palmer, P. I., Patra, P. K., Prinn, R. G., Saito, R., and Wilson, C.: Off-line algorithm for calculation of vertical tracer transport in the troposphere due to deep convection, Atmos. Chem. Phys., 13, 1093–1114, doi:10.5194/acp-13-1093-2013, 2013a.

Belikov, D., Maksyutov, S., Sherlock, V., Aoki, S., Deutscher, N. M., Dohe, S., Griffith, D., Kyro, E., Morino, I., Nakazawa, T., Notholt, J., Rettinger, M., Schneider, M., Sussmann, R., Toon, G. C., Wennberg, P. O., and Wunch, D.: Simulations of column-average CO_2 and CH_4 using the NIES TM with a hybrid sigma–isentropic (σ–θ) vertical coordinate, Atmos. Chem. Phys., 13, 1713–1732, doi:10.5194/acp-13-1713-2013, 2013b.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999.

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications, Atmos. Meas. Tech., 3, 781–811, doi:10.5194/amt-3-781-2010, 2010.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P. P.: An improved

Kalman Smoother for atmospheric inversions, Atmos. Chem. Phys., 5, 2691–2702, doi:10.5194/acp-5- 2691-2005, 2005.

Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V., Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO_2 and CH_4 retrieval errors by error parameterization, Atmos. Meas. Tech., 6, 3477–3500, doi:10.5194/amt-6-3477-2013, 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, doi:10.1029/2005JD006390, 2005.

Crisp, D., Atlas, R. M., Bréon, F.-M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. S., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, P., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, Adv. Space Res., 34, 700–709, 2004.

Deng, F., Jones, D. B. A., Henze, D. K., Bousserez, N., Bowman, K. W., Fisher, J. B., Nassar, R., O'Dell, C., Wunch, D., Wennberg, P. O., Kort, E. A., Wofsy, S. C., Blumenstock, T., Deutscher, N. M., Griffith, D. W. T., Hase, F., Heikkinen, P., Sherlock, V., Strong, K., Sussmann, R., and Warneke, T.: Inferring regional sources and sinks of atmospheric CO2 from GOSAT XCO2 data, Atmos. Chem. Phys., 14, 3703-3727, doi:10.5194/acp-14-3703-2014, 2014.

Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modeling, J. Geophys. Res., 102, 15,967–15,985, 1997.

Enting, I. G., and Mansbridge, J. V., Seasonal sources and sinks of atmospheric CO_2: Direct inversion of filtered data, Tellus B, 41B, 111–126, doi: 10.1111/j.1600-0889.1989.tb00129.x, 1989.

Enting, I. T.: Inverse problems in atmospheric constituent transport, Cambridge University Press, Cambridge, UK, 2002.

Ganshin, A., Oda, T., Saito, M., Maksyutov, S., Valsala, V., Andres, R. J., Fisher, R. E., Lowry, D., Lukyanov, A., Matsueda, H., Nisbet, E. G., Rigby, M., Sawa, Y., Toumi, R., Tsuboi, K., Varlagin, A., and Zhuravlev, R.: A global coupled Eulerian-Lagrangian model and 1 × 1 km CO_2 surface flux dataset for high-resolution atmospheric CO_2 transport simulations, Geosci. Model Dev., 5, 231–243, doi:10.5194/gmd-5-231-2012, 2012.

Giles, M. B., and Pierce, N. A.: An Introduction to the Adjoint Approach to Design, Flow Turbul. Combust., 65, 393–415, 2000.

Giering, R., and T. Kaminski, Recipes for adjoint code construction, Trans. Math. Software, 24(4), 437–474, doi:10.1145/293686.293695, 1998.

GLOBALVIEW-CO2 Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW], 2014.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwilerk, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J. R., Sarmiento, J., Taguchi, S., Takahashi, T. and Yuen, C.-W.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Peylin, P., Prather, M., and Taguchi, S.: Transcom 3 inversion intercomparison: model mean results for the estimation of seasonal carbon sources and sinks, Global Biogeochem. Cy., 18, GB1010, doi:10.1029/2003GB002111, 2004.

Hack, J. J., Boville, B. A., Briegleb, B. P., Kiehl, J. T., Rasch, P. J., and Williamson, D. L.: Description of the NCAR community climate model (CCM2), NCAR/TN-382, 108, 1993.

Haines, P. E., Esler, J. G., and Carver, G. D.: Technical note: Adjoint formulation of the TOMCAT atmospheric scheme in the Eulerian backtracking framework (RETRO-TOM), Atmos. Chem. Phys., 14, 5477–5493, 2014.

Hayes, D. J., McGuire, A. D., Kicklighter, D. W., Gurney, K. R., Burnside, T. J., and Melillo, J. M.: Is the northern high-latitude land-based CO_2 sink weakening?, Global Biogeochem. Cycles, 25, GB3018, doi:10.1029/2010GB003813, 2011.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413–2433, doi:10.5194/acp-7-2413-2007, 2007.

Hourdin, F., and Talagrand, O.: Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parametrization of subgid-scale transport, Q. J. Roy. Meteor. Soc., 132, 585–603, 2006.

IPCC 2007: Climate change 2007: the physical science basis, in: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., et al.), Cambridge University Press, Cambridge, pp. 135–145.

Ito, A.: Changing ecophysiological processes and carbon budget in East Asian ecosystems under near-future changes in climate: Implications for long-term monitoring from a process-based model, J. Plant Res., 123, 577–588, 2010.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport: 1. Adjoint model and Jacobian matrix, J. Geophys. Res., 104(D15), 18,535–18,553, doi:10.1029/1999JD900147, 1999a.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. Inversion of the transport of CO2 in the 1980s, J. Geophys. Res., 104(D15), 18,555-18,581, doi:10.1029/1999JD900146, 1999b.

Kaminski, T., Rayner, P., Heimann, M., and Enting, I.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106(D5):4703, 2001.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., Kofler, J., Neff, D., and Tans, P.: Long-term greenhouse gas measurements from aircraft, Atmos. Meas. Tech., 6, 511-526, doi:10.5194/amt-6-511-2013, 2013.

Koyama, Y., Maksyutov, S., Mukai, H., Thoning, K., and Tans, P.: Simulation of variability in atmospheric carbon dioxide using a global coupled Eulerian–Lagrangian transport model, Geosci. Model Dev., 4, 317–324, doi:10.5194/gmd-4-317-2011, 2011.

Kuze, A., Suto H., Nakajima M., and Hamazaki T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48, 6716–6733, doi:10.1364/AO.48.006716, 2009.

Law, R. M., Rayner, P. J., Denning, A. S., Erickson, D., Fung, I. Y., Heimann, M., Piper, S. C., Ramonet, M., Taguchi, S., Taylor, J. A., Trudinger, C. M., and Watterson, I. G.: Variations in modelled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions, Global Biogeochem. Cy., 10, 783–796, 1996.

Law, R. M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P. K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO_2: Experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cy., 22, GB3009, doi:10.1029/2007GB003050, 2008.

Liu, J., Bowman, K. W., and Henze D. K.: Source-receptor relationships of column-average CO2 and implications for the impact of observations on flux inversions. J. Geophys. Res. Atmos., 120, 5214–5236. doi: 10.1002/2014JD022914, 2015.

Maki, T., Ikegami, M., Fujita, T., Hirahara, T., Yamada, K., Mori, K., Takeuchi, A., Tsutsumi, Y., Suda, K., and Conway, T. J., New technique to analyse global distributions of CO_2 concentrations and fluxes from non-processed observational data, Tellus B, 62, 797–809, doi:10.1111/j.1600-0889.2010.00488.x, 2010.

Maksyutov, S., Patra, P. K., Onishi, R., Saeki, T., and Nakazawa, T.: NIES/FRCGC Global Atmospheric Tracer Transport Model: Description, validation, and surface sources and sinks inversion, J. Earth Simulator, 9, 3–18, 2008.

Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R., Ito, A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T.: Regional CO_2 flux estimates for 2009–2010 based on GOSAT and groundbased CO_2 observations, Atmos. Chem. Phys., 13, 9351–9373. http://dx.doi.org/10.5194/acp-13-9351-2013, 2013.

Marchuk, G.: Numerical solution of the problems of the dynamics of the atmosphere and the ocean (In Russian), Gidrometeoizdat, Leningrad, 303 pp., 1974.

Marchuk, G. I.: Adjoint equations and analysis of complex systems, Series: Mathematics and its applications, v. 295, Kluwer Academic Publishers, Dordrecht and Boston, 484 pp., 1995.

McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L. D., Hayes, D. J., Heimann, M., Lorenson, T. D., Macdonald, R. W., and Roulet, N.: Sensitivity of the carbon cycle in the Arctic to climate change, Ecol. Monogr., 79(4), 523–555, doi:10.1890/08-2025.1, 2009.

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

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 Reanalysis, J. Meteor. Soc. Japan, 85, 369–432, 2007.

Patra, P. K., Law, R. M., Peters, W., Rodenbeck, C., Takigawa, M., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Pieterse, G., River, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO_2: Analysis of synoptic-scale variations for the period 2002–2003, Global Biogeochem. Cy., 22, GB4013, doi:10.1029/2007GB003081, 2008.

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere, Atmos. Chem. Phys., 11, 12813–12837, doi:10.5194/acp-11-12813-2011, 2011.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304, doi:10.1029/2005JD006157, 2005.

Peters, W., Miller, J., Whitaker, J., Denning, S., Hirsch, A., Krol, M., Zupanski, D., Bruhwiler, L., and Tans, P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304,

doi:10.1029/2005JD006157., 2005. Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., and AEROCARB contributors: Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys., 5, 3173–3186, doi:10.5194/acp-5-3173-2005, 2005.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., and Zhang, X.: Global atmospheric carbon budget:

results from an ensemble of atmospheric CO_2 inversions, Biogeosciences Discuss., 10, 5301–5360, doi:10.5194/bgd-10-5301-2013, 2013.

Rayner P. J., and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Lett., 28, 175–178, 2001.

Rigby, M., Manning, A. J., and Prinn, R. G.: Inversion of long-lived trace gas emissions using combined Eulerian and Lagrangian chemical transport models, Atmos. Chem. Phys., 11, 9887–9898, doi:10.5194/acp-11-9887-2011, 2011.

Rodgers, C. D.: Inverse methods for atmospheric sounding, vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, World Scientific, Singapore, 2000.

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, doi:10.5194/acp-3-1919-2003, 2003.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys., 9, 5331–5342, doi:10.5194/acp-9-5331-2009, 2009.

Saito, M., Ito, A., and Maksyutov, S.: Evaluation of biases in JRA-25/JCDAS precipitation and their impact on the global terrestrial carbon balance, J. Climate, 24, 4109–4125, 2011.

Saito, M., Ito, A., and Maksyutov, S.: Optimization of a prognostic biosphere model in atmospheric CO_2 variability and terrestrial biomass, Geosci. Model Dev. Discuss., 6, 4243–4280, doi:10.5194/gmdd-6-4243-2013, 2013.

Saeki, T., Maksyutov, S., Sasakawa, M., Machida, T., Arshinov, M., Tans, P., Conway, T. J., Saito, M., Valsala, V., Oda, T., Andres, R. J., and Belikov, D.: Carbon flux estimation for Siberia by inverse modeling constrained by aircraft and tower CO_2 measurements, J. Geophys. Res. Atmos., 118, doi:10.1002/jgrd.50127, 2013.

Sasakawa, M., K. Shimoyama, T. Machida, N. Tsuda, H. Suto, M. Arshinov, D. Davydov, A. Fofonov, O. Krasnov, T. Saeki, Y. Koyama, and S. Maksyutov, Continuous measurements of methane from a tower network over Siberia, Tellus 62B, 403–416, 2010.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, doi:10.5194/acp-5-2461-2005, 2005.

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Greally, B. R., Lunder, C., Maione, M., Mhle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, Atmos. Chem. Phys., 9, 1597–1620, doi:10.5194/acp-9-1597-2009, 2009.

Takagi, H., Saeki, T., Oda, T., Saito, M., Valsala, V., Belikov, D., Saito, R., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., Yokota, T., and Maksyutov, S.: On the benefit of GOSAT observations to the estimation of regional CO_2 fluxes, SOLA, 7, 161–164, 2011.

Tans, P. P., Conway, T. J., and Nakazawa, T.: Latitudinal distribution of the sources and sinks of atmospheric carbon dioxide derived from surface observations and an atmospheric transport model, J. Geophys. Res., 94, 5151–5172, 1989.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, USA, 2005.

Thompson, R. L., and Stohl, A.: FLEXINVERT: an atmospheric Bayesian inversion framework for determining surface fluxes of trace species using an optimized grid, Geosci. Model Dev., 7, 2223–2242, doi:10.5194/gmd-7-2223-2014, 2014.

Tohjima, Y., Terao, Y., Mukai, H., Machida, T., Nojiri, Y., & Maksyutov, S.: ENSO-related variability in latitudinal distribution of annual mean atmospheric potential oxygen (APO) in the equatorial Western Pacific. Tellus B, 67. doi:http://dx.doi.org/10.3402/tellusb.v67.25869, 2015.

Valsala V. and Maksyutov S.: Simulation and assimilation of global ocean pCO_2 and air-sea CO_2 fluxes using ship observations of surface ocean pCO_2 in a simplified biogeochemical offline model, Tellus-B, 62B, 821–840, doi:10.1111/j.1600-0889.2010.00495.x, 2010.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.

Wilson, C., Chipperfield, M. P., Gloor, M., and Chevallier, F.: Development of a variational flux inversion system (INVICAT v1.0) using the TOMCAT chemical transport model, Geosci. Model Dev., 7, 2485–2500, doi:10.5194/gmd-7-2485-2014, 2014.

WDCGG: WMO World Data Centre for Greenhouse Gases, Japan Meteorological Agency, Tokyo, available at: http://ds.data.jma.go.jp/gmd/wdcgg/introduction.html, 2015.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO_2 and CH_4 retrieved from GOSAT: First preliminary results, SOLA, 5, 160–163, doi:10.2151/sola.2009-041, 2009.

Table 1. The coupled model setups analyzed in this study.

Case	Resolution, °		Flux combination
	NIES TM	FLEXPART	
Cs-1	10.0	1.0	VISIT + CDIAC + OTTM
Cs-2	2.50	1.0	VISIT + CDIAC + OTTM
Cs-3	1.25	1.0	VISIT + CDIAC + OTTM

Table 2. WDCGG continuous observation sites.

#	Identifying code	Location	Lat., °	Lon.,°	Height, m
1	ALT	Alert, Canada	82.45	-62.52	210
2	AMS	Amsterdam Island, France	-37.8	77.53	55
3	AMY	Anmyeon-do, Korea	36.53	126.32	47
4	BRW	Barrow, USA	71.32	-156.6	11
5	CMN	Monte Cimone, Italy	44.18	10.7	2165
6	CPT	Cape Point, South Africa	-34.35	18.48	230
7	HUN	Hegyhatsal, Hungary	46.95	16.65	248
8	IZO	Izana, Spain	28.3	-16.5	2367
9	JBN	Jubany, Argentina	-62.23	-58.67	15
10	MHD	Mace Head, Ireland	53.33	-9.9	8
11	MLO	Mauna Loa, USA	19.54	-155.58	3397
12	MNM	Minamitorishima, Japan	24.28	153.98	8
13	PAL	Pallas-Sammaltunturi, Finland	67.97	24.12	560
14	PRS	Plateau Rosa, Italy	45.93	7.7	3480
15	PUY	Puy de Dome, France	45.77	2.97	1465
16	SSL	Schauinsland, Germany	47.92	7.92	1205
17	WSA	Sable Island, Canada	43.93	-60.02	5
18	YON	Yonagunijima, Japan	24.47	123.02	30
19	ZEP	Zeppelinfjellet, Norway	78.9	11.88	475

Table 3. Tower network sites in Siberia (JR-STATION).

#	Identifying code	Location	Lat.,°	Lon.,°	Height, m
1	DEM	Demyanskoe	59.79	70.87	63
2	IGR	Igrim	63.19	64.42	47
3	KRS	Karasevoe	58.25	82.42	67
4	NOY	Noyabrsk	63.43	75.78	43
5	VGN	Vaganovo	54.50	62.32	85
6	YAK	Yakutsk	62.09	129.36	77

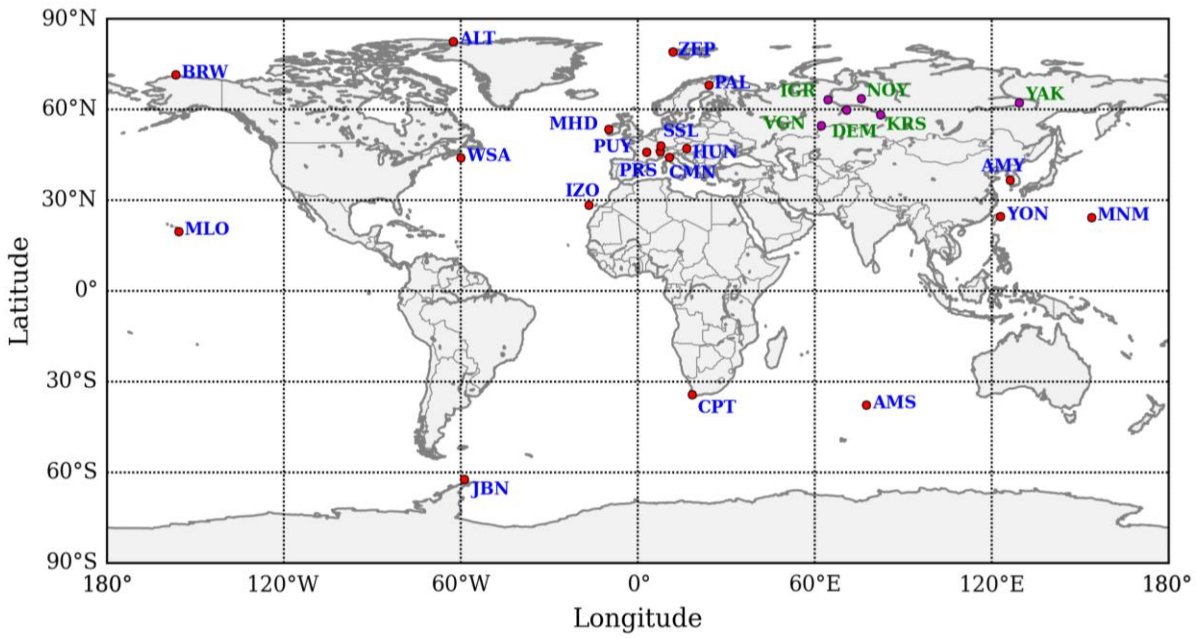


Fig. 1. Map showing the location of the 19 WDCGG sites (red dots, blue labels) and 6 tower network sites in Siberia (magenta dots, green labels) for which we have performed comparison using forward GELCA simulation.

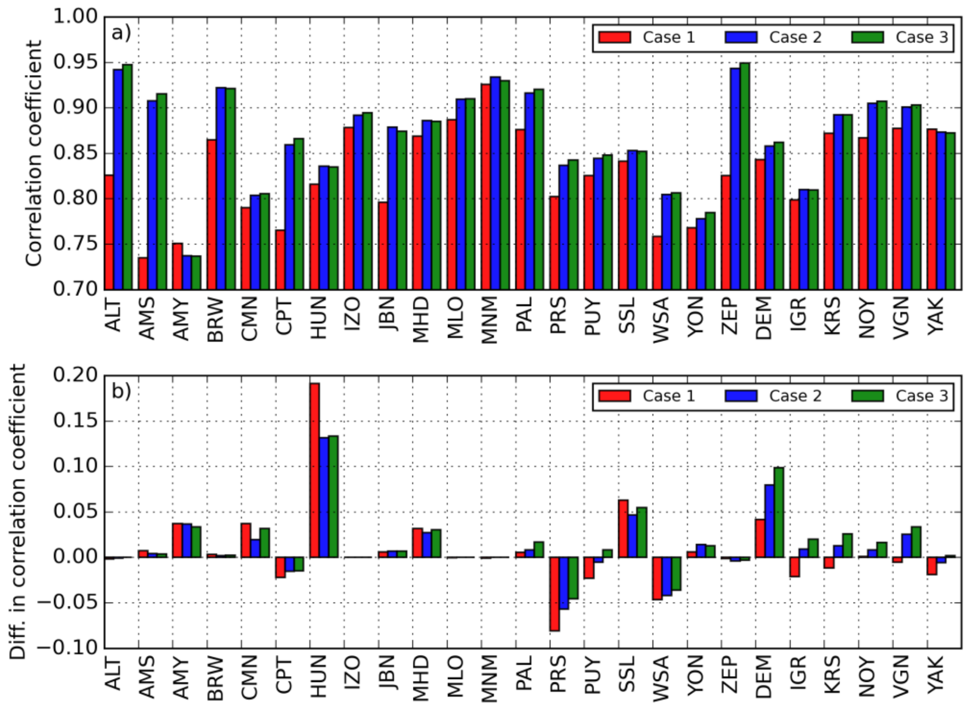


Fig. 2. a) Correlation coefficients between the CO_2 concentrations simulated with the coupled model and those observed, b) difference in correlation coefficients due to the application of the Lagrangian component (positive values mean the results of the coupled model are better than those of the Eulerian model alone) at the selected WDCGG and JR-STATION locations for 2009-2010.

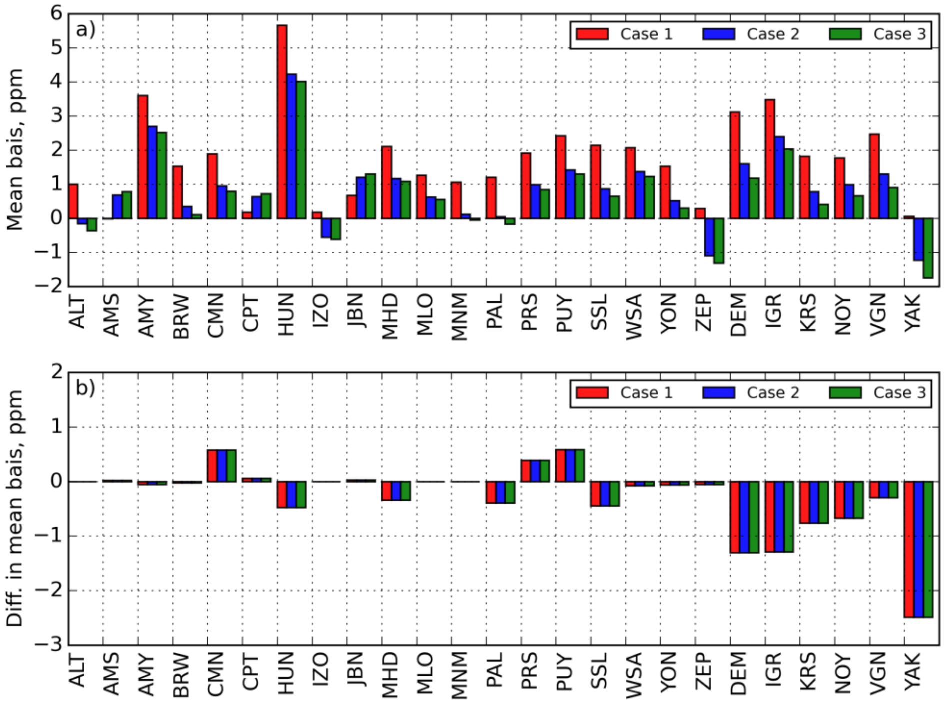


Fig. 3. a) Mean bias for the CO2 concentrations simulated with the coupled model, b) difference in mean bias due to the application of the Lagrangian component (for positive bias – the most usual case – negative values mean the results of the coupled model are better than those of the Eulerian model alone) at the selected WDCGG and JR-STATION locations for 2009-2010.

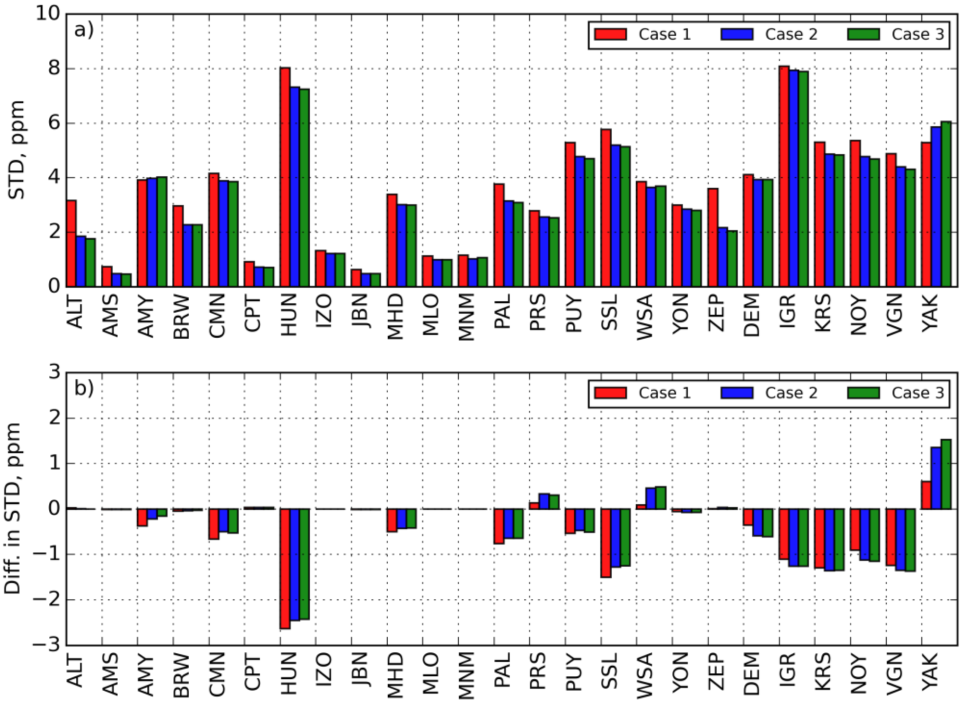
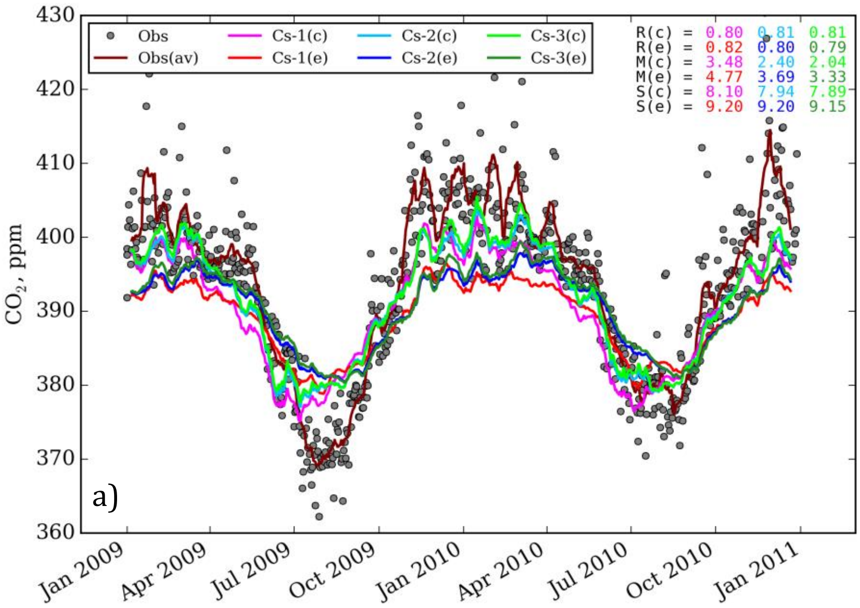


Fig. 4. a) Standard deviation (STD) for the CO_2 concentration model-observation mismatch when using the coupled model, b) difference in STD due to the application of Lagrangian component (negative values mean the results of the coupled model are better than of the Eulerian model alone) at the selected WDCGG and JR-STATION locations for 2009-2010.



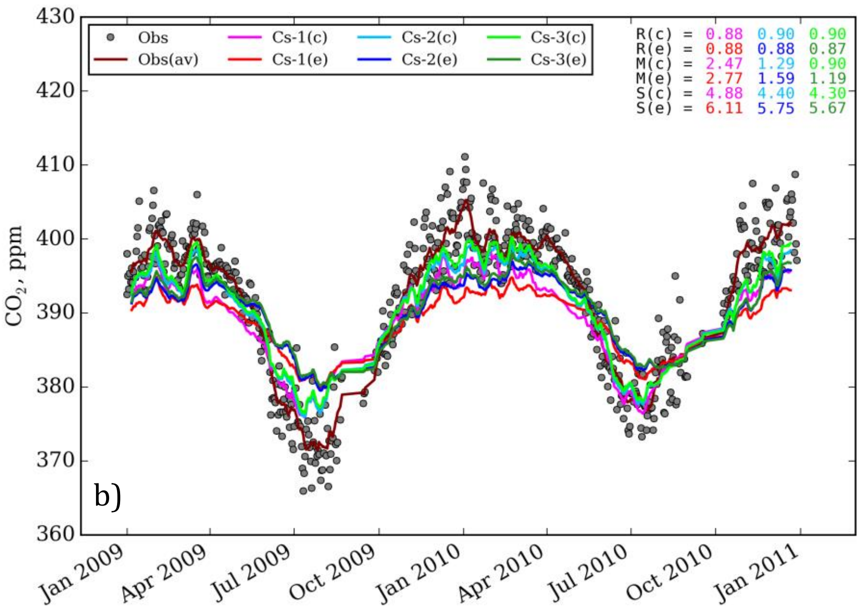


Fig. 5. CO_2 mixing ratios observed at a) the Igrim and b) Vaganovo towers, and simulated using the coupled (c) and Eulerian-only (e) models using the setups from Table 1 for 2009–2010. Symbols show individual observations; lines depict two-weeks running averages. Here, R, S and M mean the Pearson correlation, standard deviation and mean bias respectively.

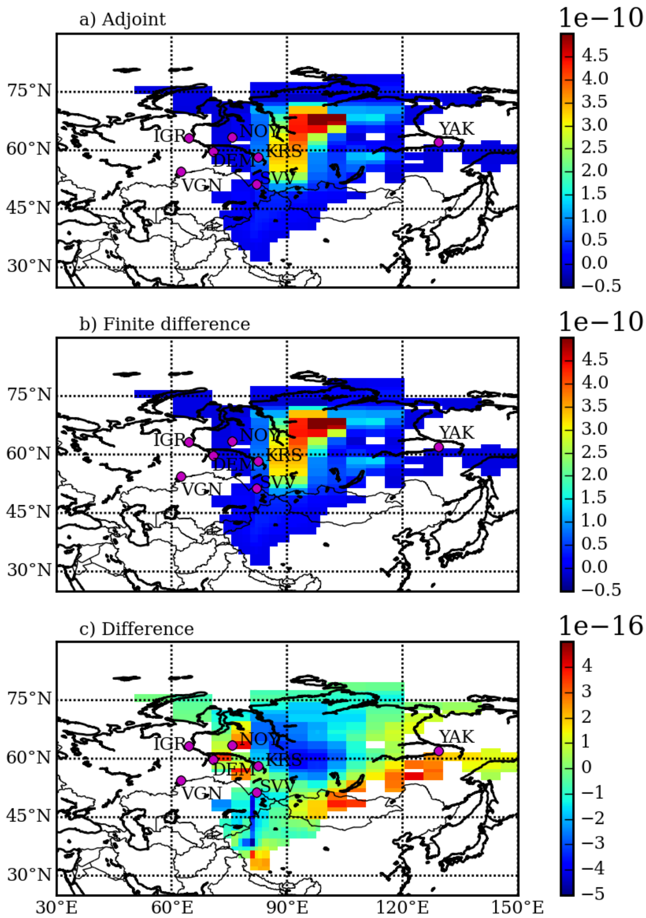


Fig. 6. Comparison of sensitivities of CO_2 concentrations (ppm/(μmol/m^{2}s)) for test 1: (a) sensitivity calculated considering only the Eulerian adjoint model at a resolution of 2.5°, (b) the same sensitivity calculated directly from NIES forward runs using the one-sided numerical finite difference method with perturbations of ε, and c) the relative difference between derived adjoint and the numerical finite difference gradients. Magenta dots with labels depicts the locations and names of the Siberian observation towers.

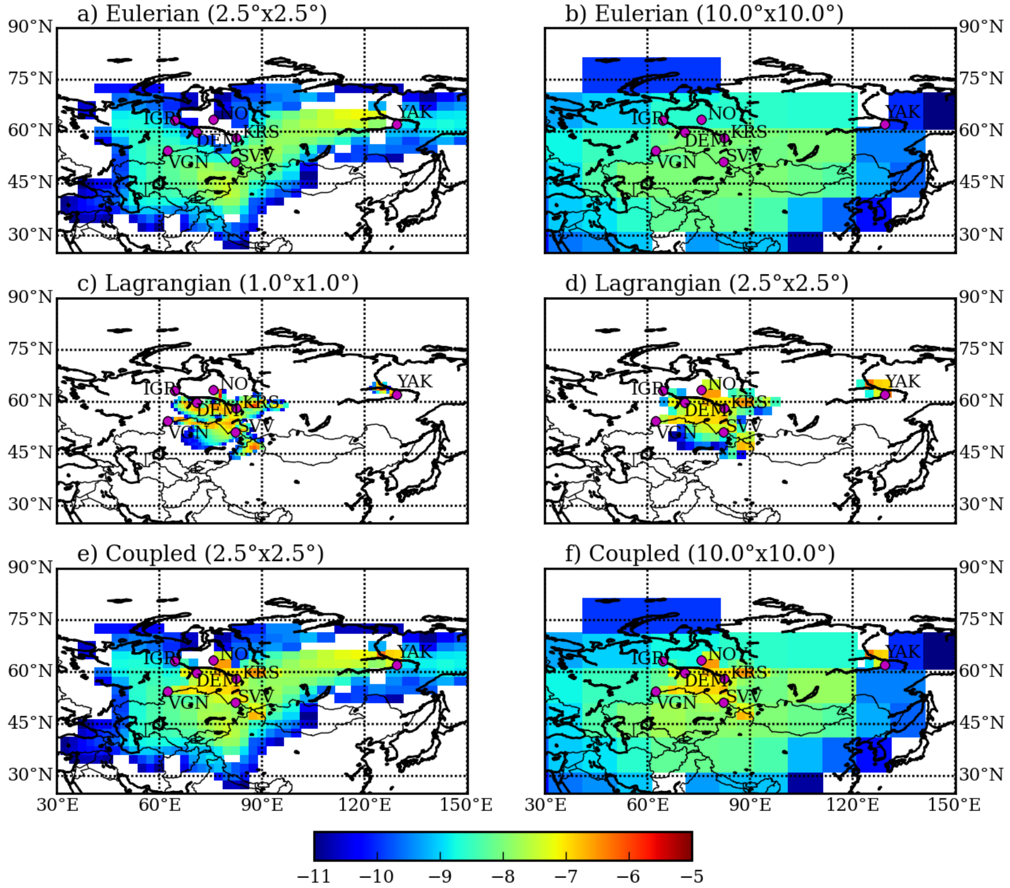


Fig. 7. Comparison of sensitivities of CO_2 concentrations [ppm/(µmol/m^{2}s)] at day 2 (see Sect. 5.2.2) calculated using: a) the Eulerian adjoint with a resolution of 2.5°, b) the Eulerian adjoint with a resolution of 10.0°, c) the Lagrangian model on the native model grid with a resolution of 1.0°, d) as for c), but aggregated on the grid with a resolution of 2.5°, e) the coupled adjoint model; results from the Lagrangian adjoint model were aggregated on the grid with a resolution of 2.5°, f) as for e), but the resolution of the Eulerian adjoint model was 10.0°. Note the logarithmic color scale for the plots.

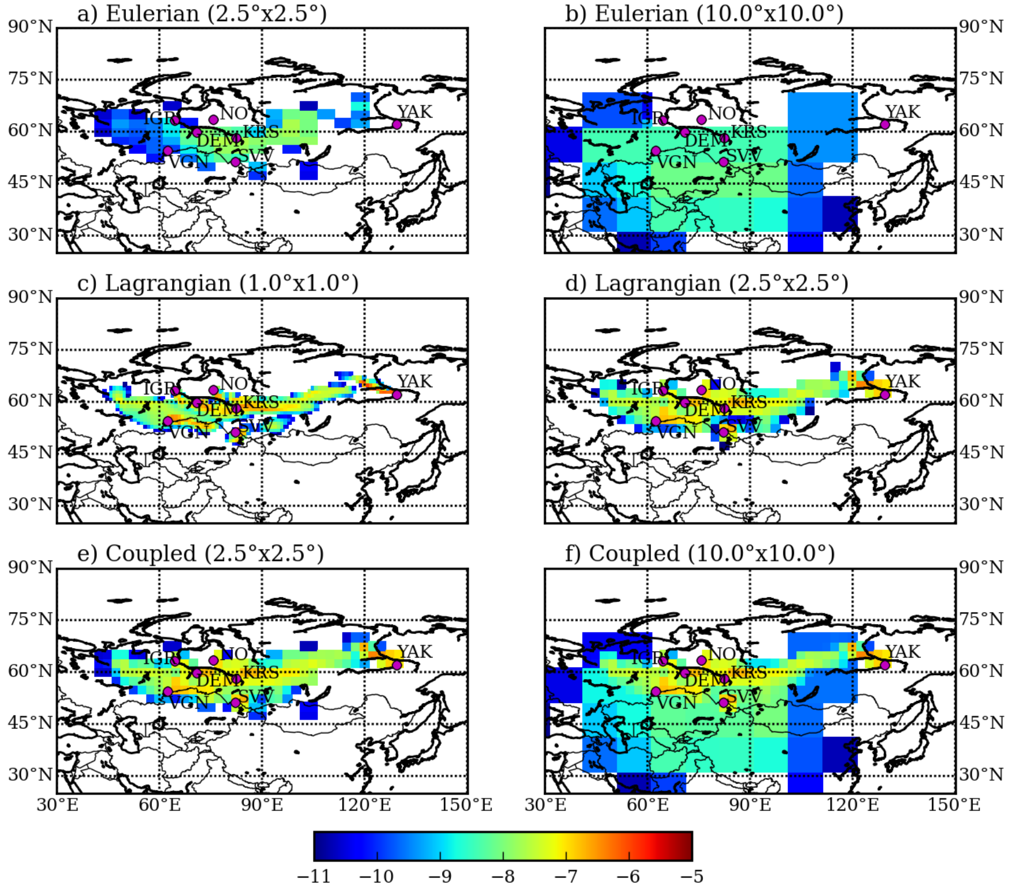


Fig. 8. As for Fig. 7, but for day 4.