

## Response to Reviewer

We thank the reviewer for the constructive and helpful suggestions. Below, we provide our responses to the reviewer's comments. We think our manuscript is much improved after addressing the reviewer's comments.

The main paper improvements are:

- The abstract has been modified;
- Proofreading and grammar check have been performed.

The reviewer's specific comments (shown in blue) are replied and addressed below.

## Response to Anonymous Referee #1

Submitted on 12 Jan 2016

The manuscript by Belikov et al. has been substantially revised. As requested in the initial review, the authors provided clarity in several areas, and expanded their evaluation of the forward model using additional observations. Overall, I find the revision to be a significant improvement. That being said, there are still a few rough spots that could use some clarification of the writing, where things are either not clear or simply vague. Also, the paper requires proof-reading and editing for grammar throughout. I provided a few comments for the abstract and introduction, but stopped after that point.

16: The authors did not comply with my previous request to be quantitative in their abstract. The state "However, we do not consider it is necessary to include any numbers in the introduction." This isn't the introduction, it is the abstract. As such, quantitative assessments of "improvements" are not only highly recommended but essentially a requirement. I strongly suggest they reconsider, if they want their work to be relevant and useful.

The abstract was revised to introduce quantitative assessments: "Mean bias and standard deviation for five of the six Siberian sites considered decrease roughly by 1 ppm when using the coupled model. The adjoint of the Eulerian model was shown, through several numerical tests, to be very accurate (within machine epsilon with mismatch around to  $\pm 6e^{-14}$ ) compared to direct forward sensitivity calculations."

4.25: It would be more correct to say high dimensional inverse modeling.

Revised.

Eq 1: This is difficult to read. Try putting the summation indices above and below the summation sign, rather than to the right. Use a horizontal bar for fractions rather than forward slash.

Revised

Section 2.1. Both reviewers requested clarification on the time-coupling described in this section. Despite the revisions, unfortunately it still isn't very clear. Could the authors consider some sort of schematic or equations to explain, given that the writing is a bit hard to follow? For example, statements such as "add contribution from surface sources" are vague.

The computational scheme of the coupled model was added. See Figure 1.

The sentence "... add contribution from surface sources." was revised as follows: "To obtain the concentrations for the observation time we transport the background concentrations from NIES TM gridbox and contribution from surface sources to the location of observation point along the trajectory ensemble calculated by FLEXPART model."

Figs 2-5: You might refer to Table 1 in the figure captions for definition of the cases.

We have added "The definitions of the cases 1-3 are in Table 1" in the captions of Figure 2-4. The caption of Figure 5 already had such kind of reference.

Fig 2-4 and associated text: The use of additional measurements to evaluate the forward model is appreciated. However, the discussion and analysis of these results is a bit lacking. Why are correlation degraded at some locations (e.g., PRS)? Why is the correlation better for some sites than others? Statements like "Most significant improvement" on 11.13 are quite vague. Please be more quantitative. Other statements aren't really correct, such as 11.21: The improvements....are higher than those obtained by increasing the resolution... It's easy to find cases that contradict this statement, for example there is a 50% reduction in STD between Case 1 and 2 for the ALT site (Fig 4a), but no apparent improvement between the Eulerian-only and coupled mode (Fig 4b). Recognizing, vaguely, that there are issues at some sites in the final paragraph of this section (11.28 - 12.3) doesn't really count; try merging this content into a single cohesive, precise, discussion. I'm not arguing that the coupled mode isn't better overall. Definitely it is! It's just that the authors need to be much more precise and careful with the writing and discussion of their results for a scientific paper.

Sentence at 11.13 revised as “Therefore, the simulation for CO<sub>2</sub> for these sites shows the full potential of the coupled model.”

Sentence at 11.15 – 12.3 revised as:

“Figure 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 improved 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). Generally, the improvements in the CO<sub>2</sub> 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, as seen for the third group of sites (Fig. 2-4).

However, improvements in CO<sub>2</sub> simulation due to the implementation of the GELCA model were obtained not for all the considered sites. There are several factors that limit the coupled model performance improvement. First, no significant improvement can be expected for the remote and marine sites since they are influenced by very distant emissions and/or nearby homogeneous emissions that are managed appropriately by the Eulerian model. The Lagrangian model introduces very significant improvements for sites influenced by relatively nearby inhomogeneous sources. Second, the use of the very rough Eulerian grid (10.0°) causes a wrong reproduction of the CO<sub>2</sub> seasonal cycle due to the large aggregation error, e.g., this happens for ALT and BRW. However, note that such low resolution is used in a rather synthetic case, which is unlikely to be used for actual simulations. Third, temporal irregularities in the observations and noise in the meteorological data bring erroneous signal to the Lagrangian model, causing spurious short term peaks of tracers, which cause degraded results at some locations (e.g., PRS, YAK). This shows that further modification of the setup (i.e. more detailed meteorological data, switch to higher resolution) is necessary. Fourth, the Lagrangian part is very sensitive to the local flux quality. Thus, it 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. However, we cannot improve our analysis, because we do not have concentration measurements for tracers whose surface fluxes are more accurately known, like SF<sub>6</sub>.

Given the large difference in computational costs running the NIES TM model when using the lower- and the higher-resolution grids (e.g., the computational cost increases by a factor of ~4 between Case 2 and 3), the coupled model is an effective way to improve the CO<sub>2</sub> simulation without changing the Eulerian model resolution.”

13.8: Another potential drawback is that discrete adjoints of nonlinear advection routines have been shown to have poorer performance for 4D-Var optimization than the continuous adjoints (Liu and Sandu, 2008, doi 10.1002/Fld.1547).

Added

15.29: In my understanding, checkpointing or recalculation of variables for an adjoint model are needed when nonlinearities arise. Yet your model is linear, no? So what is it that needs to be checkpointed?

Yes, the model is linear. The text has been revised and the following sentences have been removed: “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. “

Editorial:

2.16: reproducing —> reproduction

3.15: an —> of an

3.17: using for the first time of —> for the first time using

3.18: use —> using

3.19: develop —> , and development

4.4: effective —> effective in (or effective at)

4.4: remove “of”

4.5: The Lagrangian —> Lagrangian

4.6: tracks —> track

4.13: with set —> with a set

4.16: gradient —> gradients

4.18: 1990s,

4.20: data assimilation —> assimilation

5.4: adjoint, that —> adjoint that

All these editorial errors have been corrected

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

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

5 [1]{National Institute for Environmental Studies, Tsukuba, Japan}

6 [2]{National Institute of Polar Research, Tokyo, Japan}

7 [3]{Tomsk State University, Tomsk, Russia}

8 [4]{N. Andreev Acoustic Institute, Moscow, Russia}

9 [5]{Central Aerological Observatory, Dolgoprudny, Russia}

10 [6]{The Inversion Lab, Hamburg, Germany}

11 [7]{FastOpt GmbH, Hamburg, Germany}

12 [8]{Izaña Atmospheric Research Center (IARC), Meteorological State Agency of Spain  
13 (AEMET), Izaña, 38311, Spain}

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

15

16 Correspondence to: D.A. Belikov ([dmitry.belikov@nies.go.jp](mailto:dmitry.belikov@nies.go.jp))

17

18

## 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 backward 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 reproduction of the seasonal cycle and short-term variability of CO<sub>2</sub>. Mean bias and standard deviation for five of the six Siberian sites considered were decreased roughly by 1 ppm when using the coupled model. The adjoint of the Eulerian model was shown, through several numerical tests, to be very accurate (within machine epsilon with mismatch around to  $\pm 6e^{-14}$ ) 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 1. Introduction

2 Forecasts of CO<sub>2</sub> levels in the atmosphere and predictions of future climate depend on  
3 our scientific understanding of the natural carbon cycle (IPCC, 2007; Peters et al., 2007). To  
4 estimate the spatial and temporal distribution of carbon sources and sinks, inverse methods  
5 are used to infer carbon fluxes from geographically sparse observations of the atmospheric  
6 CO<sub>2</sub> mixing ratio (Tans et al., 1989). The first comprehensive efforts in atmospheric CO<sub>2</sub>  
7 inversions date back to the late 1980s and early 1990s (Enting and Mansbridge, 1989; Tans et  
8 al., 1989). With the increase in spatial coverage of CO<sub>2</sub> observations and the development of  
9 three-dimensional (3-D) tracer transport models, a variety of numerical experiments and  
10 projects have been performed by members of the so-called “TransCom” community of inverse  
11 modelers (e.g., Law et al., 1996, 2008; Denning et al., 1999; Gurney et al., 2002, 2004; Baker et  
12 al., 2006; Patra et al., 2011). A number of studies have proposed improvements to the inverse  
13 methods of atmospheric transport, i.e. the efficient computation of the transport matrix by the  
14 model adjoint proposed by Kaminski et al. (1999b), use of monthly mean GLOBALVIEW-CO<sub>2</sub>  
15 ground-based data (current version is for 2014) by Rödenbeck et al. (2003), development of  
16 an ensemble data assimilation method by Peters et al. (2005), flux inversion at high temporal  
17 (daily) and spatial (model grid) resolution ~~using~~ for the first time ~~using of~~ continuous CO<sub>2</sub>  
18 measurements over Europe by Peylin et al. (2005), ~~using~~ satellite data to constrain the  
19 inversion of CO<sub>2</sub> by Chevallier et al. (2005), ~~and development~~ of a new observational  
20 screening technique by Maki et al. (2010). Despite progress in atmospheric CO<sub>2</sub> inversions, a  
21 recent intercomparison (Peylin et al., 2013) demonstrated the need for further refinement.

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

33 To link surface fluxes of CO<sub>2</sub> to observed atmospheric concentrations, an accurate model

1 of atmospheric transport and an inverse modeling technique are needed. Generally, the  
2 atmospheric constituents transport may be described in two different ways: the Lagrangian  
3 and the Eulerian approaches. The Eulerian method treats the atmospheric tracers as a  
4 continuum on a control volume basis, so it is more effective at reproducing ~~of~~-long-term  
5 patterns, i.e. the seasonal cycle or the interhemispheric gradient. ~~The~~-Lagrangian Particle  
6 Dispersion Models (LPDMs) consider atmospheric tracers as a discrete phase and tracks each  
7 individual particle, therefore LPDMs are better for resolving synoptic and hourly variations.

8 To relate fluxes and concentrations of long-lived species like CO<sub>2</sub>, a transport model must  
9 cover a long simulation period (e.g., Bruhwiler et al., 2005). Therefore, computing time is a  
10 critical issue and minimization of the computational cost is essential. For chemically inert  
11 tracers, the transport can be represented by a model's Jacobian matrix, because the simulated  
12 concentration at observational sites is a linear function of the flux sets. Theoretically, to  
13 compute such matrix the transport model is run multiple times with a set of prescribed  
14 surface fluxes. However, this would require an extremely large number of forward model  
15 evaluations. The adjoint of the transport model is an efficient way to accelerate calculation of  
16 concentration gradients of the simulated tracer at observational locations (Kaminski et al.,  
17 1999). Marchuk (1974) first applied the adjoint approach in atmospheric science. After that,  
18 this method became widely used in meteorology. In the 1990s, the use of this approach was  
19 expanded to the field of tracer transport modeling (Elbern et al., 1997; Kaminski et al., 1999).

20 Adjoint models have numerous applications, including the ~~data~~-assimilation of  
21 concentrations, inverse modeling of chemical source strengths, sensitivity analysis, and  
22 parameter sensitivity estimation (Enting, 2002; Haines et al., 2014). Recent studies have used  
23 this method to constrain estimates of the emissions of CO<sub>2</sub> using retrieved column integrals  
24 from the GOSAT satellite (Basu et al., 2013; Deng et al., 2014; Liu et al., 2015).

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

30 LPDM running in the backward mode can explicitly estimate a source-receptor  
31 sensitivity matrix by solving the adjoint equations of atmospheric transport (Stohl et al.,  
32 2009), which is mathematically presented by a Jacobian expressing the sensitivity of  
33 concentration at the observational locations. Marchuk (1995), and Hourdin and Talagrand



1 (2006) provided derivations proving equivalence of the adjoint of forward transport models  
2 to backward transport models.

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

10 The goal of this study is to present the development and evaluation of an Adjoint of the  
11 Global Eulerian–Lagrangian Coupled Atmospheric model (A-GELCA), which consists of an  
12 Eulerian National Institute for Environmental Studies global Transport Model (NIES-TM;  
13 Maksyutov et al., 2008; Belikov et al., 2011, 2013a, 2013b) and a Lagrangian particle  
14 dispersion model (FLEXPART; Stohl et al., 2005). This approach utilizes the accurate transport  
15 of the LPDM to calculate the signal near to the receptors, and efficient calculation of  
16 background responses using the adjoint of the Eulerian global transport model. In contrast to  
17 previous works (Rödenbeck et al., 2009; Rigby et al., 2011; Thompson and Stohl, 2014), in  
18 which the regional models were coupled at the spatial boundary of the domain, we  
19 implemented a coupling at a time boundary in the global model domain (as described in Sect.  
20 2.1). A-GELCA can be integrated into a variational inverse modeling system designed to  
21 optimize surface fluxes.

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

28

## 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:

$$C(x_r, t_r) = \frac{T m_{air}}{h N S \rho m_{CO_2}} \sum_{ij}^I \sum_{s=0}^S F_{ij}^s \sum_{n=1}^N f_{ij}^{sn} + \frac{1}{N} \sum_{ijk}^{IJK} C_{ijk}^B \sum_{n=1}^N f_{ijk}^n, \quad (1)$$

where  $i, j$ , and  $k$  are the indices that characterize the location of each grid cell;  $s$  is the time index;  $F_{ij}^s$  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;  $\rho$  is the average air density below height  $h$ ; and  $m_{air}$  and  $m_{CO_2}$  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

1 the time of the observations. To obtain the concentrations for the observation time we  
2 transport the background concentrations from NIES TM gridbox and contribution from  
3 surface sources to the location of observation point along the trajectory ensemble calculated  
4 by FLEXPART model (Fig. 1)-and-add-contribution-from-surface-sources. Therefore we have  
5 implemented the coupling at a time boundary in the global domain of the NIES transport  
6 model, while nested regional modeling systems such as one by Rodenbeck et al (2009) have to  
7 couple at both region boundary and time boundary.

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

20 Inverse modeling assumes that the model reasonably well reproduces the relationship  
21 between atmospheric mixing ratio and surface fluxes, assuming that the biases between the  
22 simulated and observed concentrations are mostly due to the emission inventories errors. To  
23 ensure that this is the case, the NIES TM model has been evaluated extensively. Comparisons  
24 against  $\text{SF}_6$  and  $\text{CO}_2$  (Belikov et al., 2011, 2013b),  $\text{CH}_4$  (Patra et al., 2011; Belikov et al., 2013b),  
25 and  $^{222}\text{Rn}$  (Belikov et al., 2013a) measurements show the model ability to reproduce seasonal  
26 variations, interhemispheric gradient and vertical profiles of tracers.

## 27 **2.2. FLEXPART**

28 FLEXPART, like other LPDMs, considers atmospheric tracers as clouds of individual  
29 particles and tracks the pathway of each particle. The advantage of this approach is the direct  
30 estimation of the sensitivity of the measurements to the neighboring sinks and sources by  
31 tracking the particles backward in time. Usually it is sufficient to simulate for a limited  
32 number of days (2-10) to determine where particles intercept the surface layer before they  
33 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  $1.25^\circ \times 1.25^\circ$ , 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  $^{222}\text{Rn}$ ,  $\text{CH}_4$ . 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  $2.5^\circ \times 2.5^\circ$ .

### 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  $1.25 \times 1.25$  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

1 Lagrangian models provides significant benefit.

### 2 **3. Inverse modeling for the flux optimization problem**

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

6 The aim of the inversion problem is to find the value of a state vector  $\mathbf{x}$  with  $n$  elements  
7 that minimizes the cost function  $J(\mathbf{x})$ :

$$8 \quad J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}), \quad (2)$$

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

16 The minimization of the cost function (Eq. 2) has an analytic solution that involves a  
17 matrix inversion. If the Jacobian  $\mathbf{H}$  is available this analytic solution can implemented, unless  
18 the matrix sizes are too large for the available computing resources. Alternatively, Eq. 2 can be  
19 solved through an iterative minimization algorithm. In this case, the existence of the gradient  
20 of  $J(\mathbf{x})$  with respect to  $\mathbf{x}$  allows using of powerful gradient algorithms for minimization. This  
21 gradient is efficiently provided by the adjoint (Giering and Kaminski, 1998; Kaminski et al.,  
22 1999; Chevallier et al., 2005).

### 23 **4. Assessment of the coupled model**

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

30 The paper by Ganshin et al. (2012) describing the development of the GELCA model

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

6 We expanded the comparison undertaken by Ganshin et al. (2012) to a two-year period  
7 using an updated set of prescribed fluxes, which combines four components similar to the  
8 analysis performed by Takagi et al. (2011) and Maksyutov et al. (2013): (a) anthropogenic  
9 fluxes from the Open source Data Inventory of Anthropogenic CO<sub>2</sub> (ODIAC; Oda and  
10 Maksyutov, 2011) and the Carbon Dioxide Information Analysis Center's (CDIAC; Andres et  
11 al., 2009, 2011) datasets; (b) biosphere fluxes simulated by the Vegetation Integrative  
12 Simulator for Trace gases (VISIT) terrestrial biosphere model (Ito, 2010; Saito et al., 2011,  
13 2013); (c) oceanic fluxes predicted by a data assimilation system based on the Offline ocean  
14 Tracer Transport Model (OTTM; Valsala and Maksyutov, 2010); and (d) biomass burning  
15 emissions from the Global Fire Emissions Database (GFED) version 3.1 (van der Werf et al.,  
16 2010). Biosphere fluxes have daily time step, while the others are monthly. The initial global  
17 CO<sub>2</sub> distribution was obtained from GLOBALVIEW-CO<sub>2</sub> (2014).

18 We considered several cases with different model resolutions. For NIES TM we tested  
19 grids at 10.0°, 2.5°, and 1.25° resolutions, with FLEXPART running at 1.0° (Table 1). The  
20 resolution of the input fluxes was matched to that of FLEXPART. Model results were  
21 compared with observations from the World Data Centre for Greenhouse Gases (WDCGG  
22 2015) and the Siberian observations obtained by the Center for Global Environmental  
23 Research (CGER) of the National Institute for Environmental Studies (NIES) and the Russian  
24 Academy of Science (RAS), from six tower sites (JR-STATION) as described by Sasakawa et al.  
25 (2010). The selected site locations are shown in Fig. 12.

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

31 Siberia is assumed to be a substantial source and sink of CO<sub>2</sub>, with high uncertainties in  
32 the fluxes describing them (McGuire et al., 2009; Hayes et al., 2011; Saeki et al., 2013). As a  
33 result, CTMs tend to reproduce the interannual variability of CO<sub>2</sub> quite poorly. We selected six

1 tower JR-STATION sites to check the model performance in the Siberian region (Table 3).

2 The analyzed sites are divided into three groups. The first group includes remote and  
3 marine sites (ALT, AMS, BRW, CPT, IZO, JBN, MLO, MNM, ZEP) with very weak influence of  
4 local sources, so the seasonal variation of CO<sub>2</sub> is controlled by global, large-scale variations.  
5 For these sites contribution by using the Lagrangian component is negligible (see Fig. 23-45  
6 panel b to analyze the difference between the coupled and the Eulerian models).

7 The second group includes sites with domination of long term variability of CO<sub>2</sub> and  
8 relatively smooth and weak short term variations. Typically, these sites are located on the  
9 border of two regions with very different fluxes (AMY, CMN, MHD, PAL, PRS, YON).

10 The sites selected to the third group are strongly influenced by local emissions and  
11 global transport at the same time. Therefore the CO<sub>2</sub> concentration variation is controlled by  
12 the strength and direction of wind, the depth of the boundary layer and other factors. Such  
13 sites are mainly in the northern mid-latitudes (HUN, PUY, SSL, WSA) including all Siberian  
14 towers (DEM, IGR, KRS, NOY, VGN, YAK). For these locations contributions of the Eulerian and  
15 Lagrangian components are comparable. Therefore, the simulation of CO<sub>2</sub> for these sites  
16 showallows the fullfilling potential of the coupled model.

17 ~~introduces the most significant improvement when simulating CO<sub>2</sub> for these sites.~~

18 Figures 65 compares the coupled and Eulerian model results with observations from the  
19 Igrim and Vaganovo towers. The recent modifications indicated in Sect. 2.2 have significantly  
20 improved the performance of NIES TM compared with the results reported by Ganshin et al.  
21 (2012). However, compared to the updated NIES TM the coupled model is better reproducing  
22 short term peaks of concentration. This explains the observed reduction of the mean bias and  
23 STD (up to 1.5 ppm), and the better simulation of the seasonal variation (in phase and  
24 amplitude). Generally, the improvements in the CO<sub>2</sub> simulations due to the addition of the  
25 Lagrangian component to the Eulerian model are higher than those obtained by increasing the  
26 resolution of the Eulerian NIES transport model, as seen for (the third group of sites (Fig. 23-  
27 45). ~~Given the huge difference in computation costs between NIES TM for low- and high-~~  
28 ~~resolution grids (i.e. a difference by a factor of ~15 between grids with resolution 10.0° and~~  
29 ~~2.5°), the advantage of the GELCA model is clear. Performance is important, as the setup~~  
30 ~~considered here is almost identical to that used in the inverse modeling of CO<sub>2</sub>.~~

31 However, improvements in CO<sub>2</sub> simulation due to the implementation of the GELCA  
32 model were obtained not for all the considered sites. There are several factors that limit the  
33 coupled model performance improvement. First, no significant improvement can be expected

1 for the remote and marine sites since they are influenced by very distant emissions and/or  
2 nearby homogeneous emissions that are managed appropriately by the Eulerian model. The  
3 Lagrangian model introduces very significant improvements for sites influenced by relatively  
4 nearby inhomogeneous sources. Second, the use of the very rough Eulerian grid (10.0°) causes  
5 a wrong reproduction of the CO<sub>2</sub>-seasonal cycle due to the large aggregation error, e.g., i.e.  
6 this happens for ALT and BRW. However, Note that, such low resolution is used in a  
7 rather synthetic case, which is unlikely to be used for actual simulations. Third-Second, a  
8 temporal irregularities in the observations and a noise in the meteorological data bring  
9 erroneous signal to the Lagrangian model, causing spurious short term peaks of tracers,  
10 which causes degraded results degraded at some locations (e.g., PRS, YAK). This shows that  
11 further modification of the setup (i.e. more detailed meteorological data, switch to higher  
12 resolution) is necessary. FourthThird, the Lagrangian part is very sensitive to the local flux  
13 quality. Thus, it is quite problematic to use the highly uncertain surface fluxes to simulate the  
14 tracer concentrations and use these concentrations for estimating the quality of different  
15 model configurations. However, we cannot improve our analysis, because we do not have  
16 concentration measurements for tracers whose surface fluxes are more accurately known,  
17 like SF<sub>6</sub>.

18 Given the huge large difference in computational costs running the between NIES TM  
19 model when using the for lower- and the higher-resolution grids (e.g., the computational  
20 cost difference increases by a factor of ~15 between grids with resolution 10.0° and 2.5°~4  
21 between Case 2 and 3), the advantage of the GELCA model is clear. Performance is important,  
22 as the setup considered here is almost identical to that used in the inverse modeling of  
23 CO<sub>2</sub>. This shows that further modification of the setup (i.e. more detail meteorological data,  
24 switch to higher resolution) is necessary. Nevertheless, the coupled model is an effective way  
25 to improve the CO<sub>2</sub> simulation of CO<sub>2</sub> without changing increasing the resolution of the  
26 Eulerian model resolution. We recognize that is quite problematic to use the highly uncertain  
27 surface fluxes to simulate the tracer concentrations and use these concentrations for  
28 estimating the quality of different model configurations. Nevertheless, we cannot improve our  
29 analysis, because we do not have concentration measurements for tracers whose surface  
30 fluxes are more accurately known, like SF<sub>6</sub>.



## 1 5. Construction and validation of the adjoint model

### 2 5.1. Construction

3 In this section, we present the development of the adjoint of the coupled model. The  
4 incorporation of the Lagrangian component does not require any modification to the code, as  
5 LPDMs are self-adjoint. The development of the adjoint of the Eulerian part is more  
6 complicated. We decided to develop a discrete adjoint of NIES TM in order to make it  
7 consistent with the forward model. An alternative approach is the construction of a  
8 continuous adjoint derived from the leading equations of the forward model (Giles and Pierce,  
9 2000). The main advantage of the discrete adjoint model is that the resulting gradients of the  
10 numerical cost function are exact, even for nonlinear or iterative algorithms, and this makes  
11 easier to validate the adjoint model, which is an essential and complicated task.

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

22 The advantages of our coupled adjoint model are as follows.

- 23 1. Simple incorporation of the Lagrangian part, since no modification of the LPDM is  
24 required. Potentially, NIES TM can be coupled to any Lagrangian model.
- 25 2. Minimization of the simulation time can be obtained, as once calculated the output from  
26 the Lagrangian model is applicable for different long-lived tracers.
- 27 3. Reduction of aggregation errors can be achieved, as the sensitivity for small regions and  
28 even individual model cells near to observation sites is estimated using the LPDM part,  
29 while the sensitivity for large regions remote from the monitoring sites is derived using  
30 the Eulerian part (Kaminski et al., 2001).
- 31 4. Minimization of the computational cost can be obtained, as high-resolution simulation  
32 are performed over a limited number of regions nearby to the observational sites using

1 the LPDM part, while for the rest of the globe the coarse-resolution results are  
2 calculated by the Eulerian part.

3 5. High consistency of the tracer fields calculated by the Lagrangian and the Eulerian  
4 models due to the fact that both models use the same input meteorology.

5 The main drawback of the method is that the deriving of discrete adjoint of Eulerian  
6 model is a significant technical challenge. Another potential drawback is that discrete adjoints  
7 of nonlinear advection routines have been shown to have poorer performance for 4D-Var  
8 optimization than the continuous adjoints (Liu and Sandu, 2008).

## 9 **5.2. Validation of the coupled adjoint**

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

### 15 **5.2.1. Validation of the NIES TM adjoint**

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

19 The forward model sensitivity,  $\lambda_F$ , is calculated using the one- or two-sided finite  
20 difference equation,

$$21 \quad \lambda_F = \frac{M'(x+\varepsilon) - M'(x)}{\varepsilon} \quad (3)$$

$$22 \quad \lambda_F = \frac{M'(x+\varepsilon) - M'(x-\varepsilon)}{2\varepsilon} \quad (4)$$

23 where  $M'$  denotes the tangent linear model. A range of  $\varepsilon = 0.1-0.01$  was proved in most cases  
24 to give an optimal balance between truncation and roundoff error (Henze et al., 2007).

25 In the first test, adjoint simulations were carried out using an initial CO<sub>2</sub> distribution,  
26 zero surface flux for 2 days (1-2 January 2010) and a horizontal grid with resolution 2.5° ×  
27 2.5°. The adjoint gradient was then compared with that from the finite difference calculated  
28 using Eq. (3). This equation was selected in order to save CPU time by minimizing the number  
29 of forward model function calculations. For this test we used  $\varepsilon = 0.01$ .

30 To quantify the difference between the two calculations of the sensitivity  $\lambda$ , we define

1 the local relative error

$$2 \quad E(lon, lat) = \frac{|\lambda_A - \lambda_F|}{\max \lambda_A}, \quad (5)$$

3 where the subscripts A and F refer to adjoint and finite difference respectively, whereas *lon*  
4 and *lat* refer to longitude and latitude, respectively. Figure 67c shows  $E(lon, lat)$  with a  
5 logarithmic color scale. The sensitivities obtained for the adjoint have maximum relative error  
6 of order  $10^{-16}$ , indicating that transport in the NIES TM adjoint is correct over short  
7 timescales. The overall comparisons did not seriously change if we select different grid cells  
8 or use other values of  $\epsilon$ .

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

$$11 \quad \forall \mathbf{u}, \forall \mathbf{v} \langle M' \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, M^* \mathbf{v} \rangle. \quad (6)$$

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

$$13 \quad \frac{\|M'(\mathbf{u})\|^2}{(\mathbf{u}, M^*(M'(\mathbf{u})))} = 1. \quad (7)$$

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

### 17 **5.2.2. Real case simulation**

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

23 The sensitivities of CO<sub>2</sub> concentrations were calculated using the Eulerian component  
24 only in Figs. 7,8,9 a) (resolution of 2.5°), b) (resolution of 10.0°), using the Lagrangian  
25 component only in Figs. 8,9,7,8 c)(resolution of 1.0°), and d) (resolution of 1.0°, but  
26 aggregated on a grid with resolution of 2.5°), and using the coupled adjoint model in Fig.  
27 8,9,7,8 e) (Eulerian component at a resolution of 2.5° and the Lagrangian component  
28 aggregated on the grid with a resolution of 2.5°), and f) (as for e) , but the resolution of the  
29 Eulerian adjoint model was 10.0°). Figure 78 corresponds to the 2-nd day of simulation,  
30 while Figure 89 is for 4-th day.

1 Above, we have already stated that the Eulerian part of the coupled model is more  
2 effective in reproducing of long-term patterns, while the Lagrangian part is better for  
3 resolving synoptic and hourly variations. This follows from the fact that the A-GELCA  
4 components have different footprints. The Eulerian adjoint has a wider footprint, with the  
5 greatest values in an area where the effect of all stations is summed. The Euler model  
6 monitors global and large-scale changes, although some stations can be outside this zone (i.e.  
7 YAK, at Fig. 78a,g or NOY, at Fig. 89a,b). These figures illustrate why the Eulerian model, even  
8 with a sufficiently detailed grid, fails to reproduce CO<sub>2</sub> variations (Sect. 4). The footprint width  
9 decreases when the NIES TM resolution is increased, but the value of the sensitivity increases.

10 The FLEXPART model sensitivity shows more irregular distributions, and higher values  
11 closer to the observational sites, thereby reflecting the model's ability to monitor small-scale  
12 changes (Fig. 8-97-8 panels c,d).

13 During coupling, the sensitivity is aligned due to the crosslinking of components (Fig.  
14 78-89 panels e,f). Thus, the intensity has maximum near the stations and smoothly decreases  
15 when distance increases. The Eulerian and Lagrangian models employ different approaches  
16 and grid resolutions for the modeling of atmospheric tracers, and can thus resolve processes  
17 with different time and spatial scales, and underlying physics. By changing the Eulerian model  
18 resolution, it is possible to change size of the footprint. This system can utilize responses  
19 calculated at higher resolutions, such as 0.5° or 0.1°, but these setups require more accurate  
20 driving data and regular observations available for smaller time steps.

## 21 6. Computational efficiency

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

30 To run the adjoint model we used a Linux workstation with 8 Intel(R) Xeon(R) E5-4650  
31 2.70 GHz processors and 64 GB of RAM. The CPU time of the adjoint model (backward only)  
32 was almost equal to CPU time required to run the forward model. It took about 1.3 min for a

1 week-long iteration (forward and backward). The memory demand was about 1 GB. Henze et  
2 al. (2007) reports that the ratio between simulation time in backward and forward modes for  
3 adjoint models derived for other CTMs, as follows: GEOS-Chem: 1.5, STEM: 1.5, CHIMERE: 3–4,  
4 IM-AGES: 4, Polair: 4.5–7, and CIT: 11.75. Thus, the adjoint of the developed coupled model  
5 GELCA is quite efficient. To achieve this level of efficiency, a substantial amount of manual  
6 programming effort is required, despite the automatic code generated by TAF. The main  
7 disadvantage of TAF is that many redundant recomputations are often generated by the  
8 compiler. A crucial optimization of the adjoint code is required to eliminate these extra  
9 recomputations.

10

## 1 7. Summary

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

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

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

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

26 The transport scheme accuracy of the forward coupled model was investigated using the  
27 distribution of the atmospheric CO<sub>2</sub>. The GELCA components and the model itself had  
28 previously been validated using various tests and by comparison with measurements and  
29 with other transport models for CO<sub>2</sub> and other tracers. The analyses in the present paper have  
30 shown that GELCA is effective in capturing the seasonal variability of atmospheric tracer at  
31 observation sites. Decreasing of the Eulerian model resolution does not significantly distort  
32 the transport model performance; however, running the coupled model using NIES TM with  
33 low resolution grid can maximize simulation speed and use of data storage.

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

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

13           The developed A-GELCA model will be incorporated into a variational inversion system  
14 aiming studying greenhouse gases (mainly CH<sub>4</sub> and CO<sub>2</sub>), by assimilating tracer  
15 measurements from *in situ*, aircraft and remote sensing observations. However, before  
16 performing real inverse modeling simulations it is necessary to select a proper minimization  
17 program and find the optimal values for the error covariance matrices **R** and **B**.

18

## 1           **Code availability**

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

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

## 13           **Acknowledgments**

14           The authors thank A. Stohl for providing the FLEXPART model. We also thank T.  
15 Machida for Siberian observation data (downloaded from <http://db.cger.nies.go.jp/>). The JRA-  
16 25/JCDAS meteorological datasets used in the simulations were provided by the Japan  
17 Meteorological Agency. The WDCGG observation data used in the comparisons were provided  
18 by The World Data Centre for Greenhouse Gases. We appreciate cooperation of the WDCGG  
19 data providers listed in Table 2. The computational resources were provided by NIES. This  
20 study was supported by order of the Ministry for Education and Science of the Russian  
21 Federation No. 5.628.2014/K, by the Tomsk State University Academic D.I. Mendeleev Fund  
22 Program in 2014–2015 and by GRENE Arctic project.

23



## 1       **References**

- 2       Andres, R. J., Boden, T. A., and Marland, G.: Annual fossil-fuel CO<sub>2</sub> emissions: Mass of emissions  
3       gridded by one degree latitude by one degree longitude. Carbon Dioxide Information  
4       Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge,  
5       Tenn., U.S.A., doi 10.3334/CDIAC/ffe.ndp058.2009, 2009.
- 6       Andres, R. J., Gregg, J. S., Losey, L., Marland, G., Boden, T.: Monthly, global emissions of carbon  
7       dioxide from fossil fuel consumption, *Tellus* 63B, 309–327, 2011.
- 8       Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P.,  
9       Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov,  
10       S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion  
11       intercomparison: impact of transport model errors on the interannual variability of  
12       regional CO<sub>2</sub> fluxes, 1988–2003, *Global Biogeochem. Cy.*, 20, GB1002,  
13       doi:10.1029/2004GB002439, 2006.
- 14       Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P.,  
15       Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global  
16       CO<sub>2</sub> fluxes estimated from GOSAT retrievals of total column CO<sub>2</sub>, *Atmos. Chem. Phys.*,  
17       13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.
- 18       Belikov, D., Maksyutov, S., Miyasaka, T., Saeki, T., Zhuravlev, R., and Kiryushov, B.: Mass-  
19       conserving tracer transport modelling on a reduced latitude-longitude grid with NIES-  
20       TM, *Geosci. Model Dev.*, 4, 207–222, 2011.
- 21       Belikov, D. A., Maksyutov, S., Krol, M., Fraser, A., Rigby, M., Bian, H., Agusti-Panareda, A.,  
22       Bergmann, D., Bousquet, P., Cameron-Smith, P., Chipperfield, M. P., Fortems-Cheiney, A.,  
23       Gloor, E., Haynes, K., Hess, P., Houweling, S., Kawa, S. R., Law, R. M., Loh, Z., Meng, L.,  
24       Palmer, P. I., Patra, P. K., Prinn, R. G., Saito, R., and Wilson, C.: Off-line algorithm for  
25       calculation of vertical tracer transport in the troposphere due to deep convection,  
26       *Atmos. Chem. Phys.*, 13, 1093–1114, doi:10.5194/acp-13-1093-2013, 2013a.
- 27       Belikov, D., Maksyutov, S., Sherlock, V., Aoki, S., Deutscher, N. M., Dohe, S., Griffith, D., Kyro, E.,  
28       Morino, I., Nakazawa, T., Notholt, J., Rettinger, M., Schneider, M., Sussmann, R., Toon, G.  
29       C., Wennberg, P. O., and Wunch, D.: Simulations of column-average CO<sub>2</sub> and CH<sub>4</sub> using the  
30       NIES TM with a hybrid sigma–isentropic ( $\sigma$ – $\theta$ ) vertical coordinate, *Atmos. Chem. Phys.*,  
31       13, 1713–1732, doi:10.5194/acp-13-1713-2013, 2013b.
- 32       Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V.,  
33       and Goede, A. P. H.: SCIAMACHY: Mission objectives and measurement modes, *J. Atmos.*  
34       *Sci.*, 56, 127–150, 1999.
- 35       Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising,  
36       O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global  
37       monitoring of power plant CO<sub>2</sub> emissions from space and related applications, *Atmos.*  
38       *Meas. Tech.*, 3, 781–811, doi:10.5194/amt-3-781-2010, 2010.
- 39       Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P. P.: An improved

- 1 Kalman Smoother for atmospheric inversions, *Atmos. Chem. Phys.*, 5, 2691–2702,  
2 doi:10.5194/acp-5-2691-2005, 2005.
- 3 Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V.,  
4 Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon  
5 Monitoring Satellite (CarbonSat): assessment of atmospheric CO<sub>2</sub> and CH<sub>4</sub> retrieval  
6 errors by error parameterization, *Atmos. Meas. Tech.*, 6, 3477–3500, doi:10.5194/amt-  
7 6-3477-2013, 2013.
- 8 Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais,  
9 P.: Inferring CO<sub>2</sub> sources and sinks from satellite observations: method and application  
10 to TOVS data, *J. Geophys. Res.*, 110, D24309, doi:10.1029/2005JD006390, 2005.
- 11 Crisp, D., Atlas, R. M., Bréon, F.-M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S.  
12 C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P.,  
13 Salawitch, R. S., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P.  
14 O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, P., Pollock, R.,  
15 Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, *Adv. Space  
16 Res.*, 34, 700–709, 2004.
- 17 Deng, F., Jones, D. B. A., Henze, D. K., Bousserrez, N., Bowman, K. W., Fisher, J. B., Nassar, R.,  
18 O'Dell, C., Wunch, D., Wennberg, P. O., Kort, E. A., Wofsy, S. C., Blumenstock, T., Deutscher,  
19 N. M., Griffith, D. W. T., Hase, F., Heikkinen, P., Sherlock, V., Strong, K., Sussmann, R., and  
20 Warneke, T.: Inferring regional sources and sinks of atmospheric CO<sub>2</sub> from GOSAT XCO<sub>2</sub>  
21 data, *Atmos. Chem. Phys.*, 14, 3703–3727, doi:10.5194/acp-14-3703-2014, 2014.
- 22 Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry  
23 modeling, *J. Geophys. Res.*, 102, 15,967–15,985, 1997.
- 24 Enting, I. G., and Mansbridge, J. V., Seasonal sources and sinks of atmospheric CO<sub>2</sub>: Direct  
25 inversion of filtered data, *Tellus B*, 41B, 111–126, doi: 10.1111/j.1600-  
26 0889.1989.tb00129.x, 1989.
- 27 Enting, I. T.: *Inverse problems in atmospheric constituent transport*, Cambridge University  
28 Press, Cambridge, UK, 2002.
- 29 Ganshin, A., Oda, T., Saito, M., Maksyutov, S., Valsala, V., Andres, R. J., Fisher, R. E., Lowry, D.,  
30 Lukyanov, A., Matsueda, H., Nisbet, E. G., Rigby, M., Sawa, Y., Toumi, R., Tsuboi, K.,  
31 Varlagin, A., and Zhuravlev, R.: A global coupled Eulerian-Lagrangian model and 1 × 1 km  
32 CO<sub>2</sub> surface flux dataset for high-resolution atmospheric CO<sub>2</sub> transport simulations,  
33 *Geosci. Model Dev.*, 5, 231–243, doi:10.5194/gmd-5-231-2012, 2012.
- 34 Giles, M. B., and Pierce, N. A.: An Introduction to the Adjoint Approach to Design, *Flow Turbul.  
35 Combust.*, 65, 393–415, 2000.
- 36 Giering, R., and T. Kaminski, Recipes for adjoint code construction, *Trans. Math. Software*,  
37 24(4), 437–474, doi:10.1145/293686.293695, 1998.
- 38 GLOBALVIEW-CO<sub>2</sub> Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-  
39 ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to

- 1 ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW], 2014.
- 2 Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwilerk, L.,  
3 Chen, Y.-H., Ciais, P., Fan, S., Fung, I., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T.,  
4 Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J. R., Sarmiento, J.,  
5 Taguchi, S., Takahashi, T. and Yuen, C.-W.: Towards robust regional estimates of CO<sub>2</sub>  
6 sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, 2002.
- 7 Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., Bousquet, P.,  
8 Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov,  
9 S., Peylin, P., Prather, M., and Taguchi, S.: Transcom 3 inversion intercomparison: model  
10 mean results for the estimation of seasonal carbon sources and sinks, *Global*  
11 *Biogeochem. Cy.*, 18, GB1010, doi:10.1029/2003GB002111, 2004.
- 12 Hack, J. J., Boville, B. A., Briegleb, B. P., Kiehl, J. T., Rasch, P. J., and Williamson, D. L.: Description  
13 of the NCAR community climate model (CCM2), NCAR/TN-382, 108, 1993.
- 14 Haines, P. E., Esler, J. G., and Carver, G. D.: Technical note: Adjoint formulation of the TOMCAT  
15 atmospheric scheme in the Eulerian backtracking framework (RETRO-TOM), *Atmos.*  
16 *Chem. Phys.*, 14, 5477–5493, 2014.
- 17 Hayes, D. J., McGuire, A. D., Kicklighter, D. W., Gurney, K. R., Burnside, T. J., and Melillo, J. M.: Is  
18 the northern high-latitude land-based CO<sub>2</sub> sink weakening?, *Global Biogeochem. Cycles*,  
19 25, GB3018, doi:10.1029/2010GB003813, 2011.
- 20 Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, *Atmos.*  
21 *Chem. Phys.*, 7, 2413–2433, doi:10.5194/acp-7-2413-2007, 2007.
- 22 Hourdin, F., and Talagrand, O.: Eulerian backtracking of atmospheric tracers. I: Adjoint  
23 derivation and parametrization of subgrid-scale transport, *Q. J. Roy. Meteor. Soc.*, 132,  
24 585–603, 2006.
- 25 IPCC 2007: Climate change 2007: the physical science basis, in: Contribution of Working  
26 Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate  
27 Change (eds. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., et al.), Cambridge  
28 University Press, Cambridge, pp. 135–145.
- 29 Ito, A.: Changing ecophysiological processes and carbon budget in East Asian ecosystems  
30 under near-future changes in climate: Implications for long-term monitoring from a  
31 process-based model, *J. Plant Res.*, 123, 577–588, 2010.
- 32 Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse  
33 model of the atmospheric transport: 1. Adjoint model and Jacobian matrix, *J. Geophys.*  
34 *Res.*, 104(D15), 18,535–18,553, doi:10.1029/1999JD900147, 1999a.
- 35 Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse  
36 model of the atmospheric transport: 2. Inversion of the transport of CO<sub>2</sub> in the 1980s, *J.*  
37 *Geophys. Res.*, 104(D15), 18,555–18,581, doi:10.1029/1999JD900146, 1999b.
- 38 Kaminski, T., Rayner, P., Heimann, M., and Enting, I.: On aggregation errors in atmospheric  
39 transport inversions, *J. Geophys. Res.*, 106(D5):4703, 2001.

- 1 Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., Kofler, J., Neff, D., and  
2 Tans, P.: Long-term greenhouse gas measurements from aircraft, *Atmos. Meas. Tech.*, 6,  
3 511-526, doi:10.5194/amt-6-511-2013, 2013.
- 4 Koyama, Y., Maksyutov, S., Mukai, H., Thoning, K., and Tans, P.: Simulation of variability in  
5 atmospheric carbon dioxide using a global coupled Eulerian-Lagrangian transport  
6 model, *Geosci. Model Dev.*, 4, 317-324, doi:10.5194/gmd-4-317-2011, 2011.
- 7 Kuze, A., Suto H., Nakajima M., and Hamazaki T.: Thermal and near infrared sensor for carbon  
8 observation Fourier-transform spectrometer on the Greenhouse Gases Observing  
9 Satellite for greenhouse gases monitoring, *Appl. Opt.*, 48, 6716-6733,  
10 doi:10.1364/AO.48.006716, 2009.
- 11 Law, R. M., Rayner, P. J., Denning, A. S., Erickson, D., Fung, I. Y., Heimann, M., Piper, S. C.,  
12 Ramonet, M., Taguchi, S., Taylor, J. A., Trudinger, C. M., and Watterson, I. G.: Variations in  
13 modelled atmospheric transport of carbon dioxide and the consequences for CO<sub>2</sub>  
14 inversions, *Global Biogeochem. Cy.*, 10, 783-796, 1996.
- 15 Law, R. M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P.,  
16 Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S.,  
17 Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C.,  
18 Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P.  
19 K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R.,  
20 Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO<sub>2</sub>:  
21 Experimental overview and diurnal cycle results for 2002, *Global Biogeochem. Cy.*, 22,  
22 GB3009, doi:10.1029/2007GB003050, 2008.
- 23 [Liu, Z. and Sandu, A.: On the properties of discrete adjoints of numerical methods for the](#)  
24 [advection equation. \*Int. J. Numer. Meth. Fluids\*, 56: 769-803. doi:10.1002/fld.1547,](#)  
25 [2008.](#)
- 26 Liu, J., Bowman, K. W., and Henze D. K.: Source-receptor relationships of column-average CO<sub>2</sub>  
27 and implications for the impact of observations on flux inversions. *J. Geophys. Res.*  
28 *Atmos.*, 120, 5214-5236. doi: 10.1002/2014JD022914, 2015.
- 29 Maki, T., Ikegami, M., Fujita, T., Hirahara, T., Yamada, K., Mori, K., Takeuchi, A., Tsutsumi, Y.,  
30 Suda, K., and Conway, T. J., New technique to analyse global distributions of CO<sub>2</sub>  
31 concentrations and fluxes from non-processed observational data, *Tellus B*, 62, 797-  
32 809, doi:10.1111/j.1600-0889.2010.00488.x, 2010.
- 33 Maksyutov, S., Patra, P. K., Onishi, R., Saeki, T., and Nakazawa, T.: NIES/FRCGC Global  
34 Atmospheric Tracer Transport Model: Description, validation, and surface sources and  
35 sinks inversion, *J. Earth Simulator*, 9, 3-18, 2008.
- 36 Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R., Ito,  
37 A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T.: Regional CO<sub>2</sub> flux  
38 estimates for 2009-2010 based on GOSAT and groundbased CO<sub>2</sub> observations, *Atmos.*  
39 *Chem. Phys.*, 13, 9351-9373. <http://dx.doi.org/10.5194/acp-13-9351-2013>, 2013.

- 1 Marchuk, G.: Numerical solution of the problems of the dynamics of the atmosphere and the  
2 ocean (In Russian), Gidrometeoizdat, Leningrad, 303 pp., 1974.
- 3 Marchuk, G. I.: Adjoint equations and analysis of complex systems, Series: Mathematics and its  
4 applications, v. 295, Kluwer Academic Publishers, Dordrecht and Boston, 484 pp., 1995.
- 5 McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L. D., Hayes, D. J.,  
6 Heimann, M., Lorenson, T. D., Macdonald, R. W., and Roulet, N.: Sensitivity of the carbon  
7 cycle in the Arctic to climate change, *Ecol. Monogr.*, 79(4), 523–555, doi:10.1890/08-  
8 2025.1, 2009.
- 9 Oda, T., and Maksyutov, S.: A very high-resolution (1 km × 1 km) global fossil fuel CO<sub>2</sub>  
10 emission inventory derived using a point source database and satellite observations of  
11 nighttime lights, *Atmos. Chem. Phys.*, 11, 543–556, doi:10.5194/acp-11-543-2011, 2011.
- 12 Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T.,  
13 Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R.,  
14 Ose, T., Mannoji, N., and Taira, R.: The JRA-25 Reanalysis, *J. Meteor. Soc. Japan*, 85, 369–  
15 432, 2007.
- 16 Patra, P. K., Law, R. M., Peters, W., Rodenbeck, C., Takigawa, M., Aulagnier, C., Baker, I.,  
17 Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen,  
18 J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U.,  
19 Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y.,  
20 Onishi, R., Parazoo, N., Pieterse, G., River, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R.,  
21 Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO<sub>2</sub>:  
22 Analysis of synoptic-scale variations for the period 2002–2003, *Global Biogeochem. Cy.*,  
23 22, GB4013, doi:10.1029/2007GB003081, 2008.
- 24 Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-  
25 Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P.,  
26 Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G.,  
27 Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH<sub>4</sub> and related  
28 species: linking transport, surface flux and chemical loss with CH<sub>4</sub> variability in the  
29 troposphere and lower stratosphere, *Atmos. Chem. Phys.*, 11, 12813–12837,  
30 doi:10.5194/acp-11-12813-2011, 2011.
- 31 Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D.,  
32 Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO<sub>2</sub>  
33 surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.*, 110, D24304,  
34 doi:10.1029/2005JD006157, 2005.
- 35 Peters, W., Miller, J., Whitaker, J., Denning, S., Hirsch, A., Krol, M., Zupanski, D., Bruhwiler, L.,  
36 and Tans, P.: An ensemble data assimilation system to estimate CO<sub>2</sub> surface fluxes from  
37 atmospheric trace gas observations, *J. Geophys. Res.*, 110, D24304,  
38 doi:10.1029/2005JD006157., 2005. Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin,  
39 F., Heinrich, P., Ciais, P., and AEROCARB contributors: Daily CO<sub>2</sub> flux estimates over  
40 Europe from continuous atmospheric measurements: 1, inverse methodology, *Atmos.*

- 1 Chem. Phys., 5, 3173–3186, doi:10.5194/acp-5-3173-2005, 2005.
- 2 Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K.,  
3 Peters, W., Rayner, P. J., Rödenbeck, C., and Zhang, X.: Global atmospheric carbon budget:  
4 results from an ensemble of atmospheric CO<sub>2</sub> inversions, *Biogeosciences Discuss.*, 10,  
5 5301–5360, doi:10.5194/bgd-10-5301-2013, 2013.
- 6 Rayner P. J., and O'Brien, D. M.: The utility of remotely sensed CO<sub>2</sub> concentration data in  
7 surface source inversions, *Geophys. Res. Lett.*, 28, 175–178, 2001.
- 8 Rigby, M., Manning, A. J., and Prinn, R. G.: Inversion of long-lived trace gas emissions using  
9 combined Eulerian and Lagrangian chemical transport models, *Atmos. Chem. Phys.*, 11,  
10 9887–9898, doi:10.5194/acp-11-9887-2011, 2011.
- 11 Rodgers, C. D.: Inverse methods for atmospheric sounding, vol. 2 of Series on Atmospheric,  
12 Oceanic and Planetary Physics, World Scientific, Singapore, 2000.
- 13 Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO<sub>2</sub> flux history 1982–2001 inferred  
14 from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem.*  
15 *Phys.*, 3, 1919–1964, doi:10.5194/acp-3-1919-2003, 2003.
- 16 Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-  
17 resolution regional atmospheric trace gas inversions based on independent models,  
18 *Atmos. Chem. Phys.*, 9, 5331–5342, doi:10.5194/acp-9-5331-2009, 2009.
- 19 Saito, M., Ito, A., and Maksyutov, S.: Evaluation of biases in JRA-25/JCDAS precipitation and  
20 their impact on the global terrestrial carbon balance, *J. Climate*, 24, 4109–4125, 2011.
- 21 Saito, M., Ito, A., and Maksyutov, S.: Optimization of a prognostic biosphere model in  
22 atmospheric CO<sub>2</sub> variability and terrestrial biomass, *Geosci. Model Dev. Discuss.*, 6,  
23 4243–4280, doi:10.5194/gmdd-6-4243-2013, 2013.
- 24 Saeki, T., Maksyutov, S., Sasakawa, M., Machida, T., Arshinov, M., Tans, P., Conway, T. J., Saito,  
25 M., Valsala, V., Oda, T., Andres, R. J., and Belikov, D.: Carbon flux estimation for Siberia by  
26 inverse modeling constrained by aircraft and tower CO<sub>2</sub> measurements, *J. Geophys. Res.*  
27 *Atmos.*, 118, doi:10.1002/jgrd.50127, 2013.
- 28 Sasakawa, M., K. Shimoyama, T. Machida, N. Tsuda, H. Suto, M. Arshinov, D. Davydov, A.  
29 Fofonov, O. Krasnov, T. Saeki, Y. Koyama, and S. Maksyutov, Continuous measurements  
30 of methane from a tower network over Siberia, *Tellus 62B*, 403–416, 2010.
- 31 Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian  
32 particle dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5, 2461–2474,  
33 doi:10.5194/acp-5-2461-2005, 2005.
- 34 Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Grealley, B. R., Lunder, C., Maione, M.,  
35 Mhle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P.  
36 G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for  
37 determining regional and global emissions of greenhouse gases: Sensitivity studies and  
38 application to halocarbons, *Atmos. Chem. Phys.*, 9, 1597–1620, doi:10.5194/acp-9-1597-  
39 2009, 2009.

1 Takagi, H., Saeki, T., Oda, T., Saito, M., Valsala, V., Belikov, D., Saito, R., Yoshida, Y., Morino, I.,  
2 Uchino, O., Andres, R. J., Yokota, T., and Maksyutov, S.: On the benefit of GOSAT  
3 observations to the estimation of regional CO<sub>2</sub> fluxes, SOLA, 7, 161–164, 2011.

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

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

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

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

16 Valsala V. and Maksyutov S.: Simulation and assimilation of global ocean pCO<sub>2</sub> and air-sea CO<sub>2</sub>  
17 fluxes using ship observations of surface ocean pCO<sub>2</sub> in a simplified biogeochemical  
18 offline model, Tellus-B, 62B, 821–840, doi:10.1111/j.1600-0889.2010.00495.x, 2010.

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

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

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

28 Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global  
29 concentrations of CO<sub>2</sub> and CH<sub>4</sub> retrieved from GOSAT: First preliminary results, SOLA, 5,  
30 160–163, doi:10.2151/sola.2009-041, 2009.

31

1 **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

2

3 **Table 2.** WDCGG continuous observation sites.

#	Identifying code	Location	Lat., °	Lon., °	Height, m	<u>Contributor, contact person</u>
1	ALT	Alert, Canada	82.45	-62.52	210	<u>EC,</u> <u>Doug Worthy</u>
2	AMS	Amsterdam Island, France	-37.8	77.53	55	<u>LSCE,</u> <u>Michel Ramonet</u>
3	AMY	Anmyeon-do, Korea	36.53	126.32	47	<u>KMA,</u> <u>Haeyoung Lee</u>
4	BRW	Barrow, USA	71.32	-156.6	11	<u>NOAA/ESRL,</u> <u>Kirk W Thoning</u> <u>IAFMS,</u>
5	CMN	Monte Cimone, Italy	44.18	10.7	2165	<u>Centro Aeronautica</u> <u>Militare di</u> <u>Montagna</u>
6	CPT	Cape Point, South Africa	-34.35	18.48	230	<u>SAWS,</u> <u>Thumeka Mkololo</u>
7	HUN	Hegyhatsal, Hungary	46.95	16.65	248	<u>HMS,</u> <u>Laszlo Haszpra</u> <u>AEMET,</u>
8	IZO	Izana, Spain	28.3	-16.5	2367	<u>Angel J. Gomez-Pelaez</u>
9	JBN	Jubany, Argentina	-62.23	-58.67	15	<u>CNR-ICES, DNA-IAA,</u> <u>Claudio Rafanelli</u>
10	MHD	Mace Head, Ireland	53.33	-9.9	8	<u>LSCE,</u> <u>Michel Ramonet</u>
11	MLO	Mauna Loa, USA	19.54	-155.58	3397	<u>NOAA/ESRL,</u> <u>Kirk W Thoning</u> <u>JMA, Greenhouse</u>
12	MNM	Minamitorishima, Japan	24.28	153.98	8	<u>Gas observation</u> <u>section</u>
13	PAL	Pallas-Sammaltunturi, Finland	67.97	24.12	560	<u>FMI,</u> <u>Juha Hatakka</u>
14	PRS	Plateau Rosa, Italy	45.93	7.7	3480	<u>RSE,</u> <u>Francesco Apadula</u>
15	PUY	Puy de Dome, France	45.77	2.97	1465	<u>LSCE,</u> <u>Michel Ramonet</u>
16	SSL	Schauinsland, Germany	47.92	7.92	1205	<u>UBA,</u> <u>Karin Uhse</u>
17	WSA	Sable Island, Canada	43.93	-60.02	5	<u>EC,</u> <u>Doug Worthy</u> <u>JMA, Greenhouse</u>
18	YON	Yonagunijima, Japan	24.47	123.02	30	<u>Gas observation</u> <u>section</u>



19	ZEP	Zeppelifjellet, Norway	78.9	11.88	475	<u>ITM,</u> Birgitta Noone
----	-----	---------------------------	------	-------	-----	-------------------------------

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13

Here AEMET - Izana Atmospheric Research Center, Meteorological State Agency of Spain; CNR-ICES - International Center for Earth Sciences - CNR, Institute of Acoustics and Sensors; DNA-IAA - Direccion Nacional del Antartico- Istituto Antartico Argentino; EC - Environment Canada; HMS - Hungarian Meteorological Service; IAFMS - Italian Air Force Meteorological Service; ITM - Department of Applied Environmental Science, Stockholm University; JMA - Japan Meteorological Agency; KMA - Korea Meteorological Administration; LSCE - Laboratoire des Sciences du Climat et de l'Environnement; NOAA/ESRL - National Oceanic and Atmospheric Administration/Earth System Research Laboratory; RSE - Ricerca sul Sistema Energetico - RSE S.p.A.; FMI - Finnish Meteorological Institute; SAWS - South African Weather Service; UBA - Federal Environmental Agency Germany.

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

---

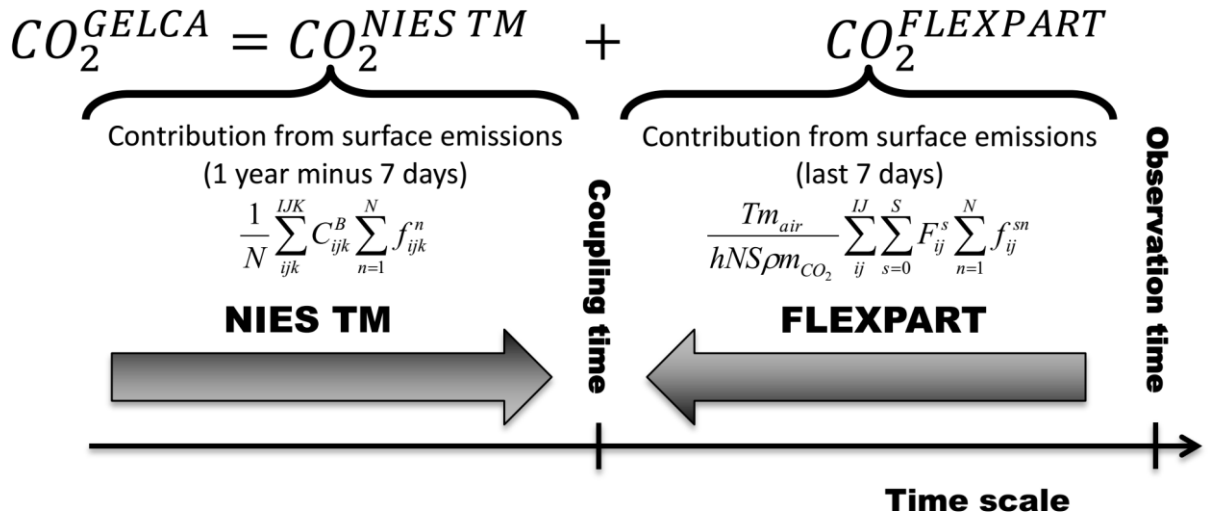
#	Identifying code	Location	Lat., °	Lon., °	Height, m
1	DEM	Demyanskoe	59.79	70.87	63
2	IGR	Igrim	63.19	64.42	47
3	KRS	Karasevoe	58.25	82.42	67
4	NOY	Noyabrsk	63.43	75.78	43
5	VGN	Vaganovo	54.50	62.32	85
6	YAK	Yakutsk	62.09	129.36	77

---

2

3

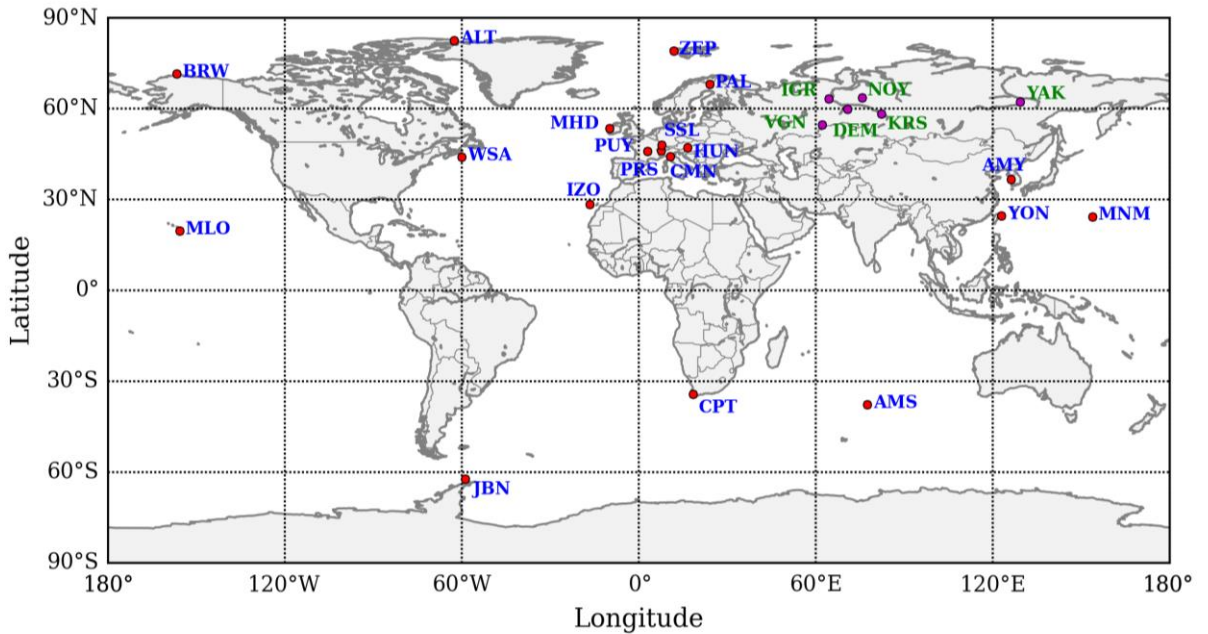
1



2

3

**Fig. 1.** The computational scheme of the coupled model.



4

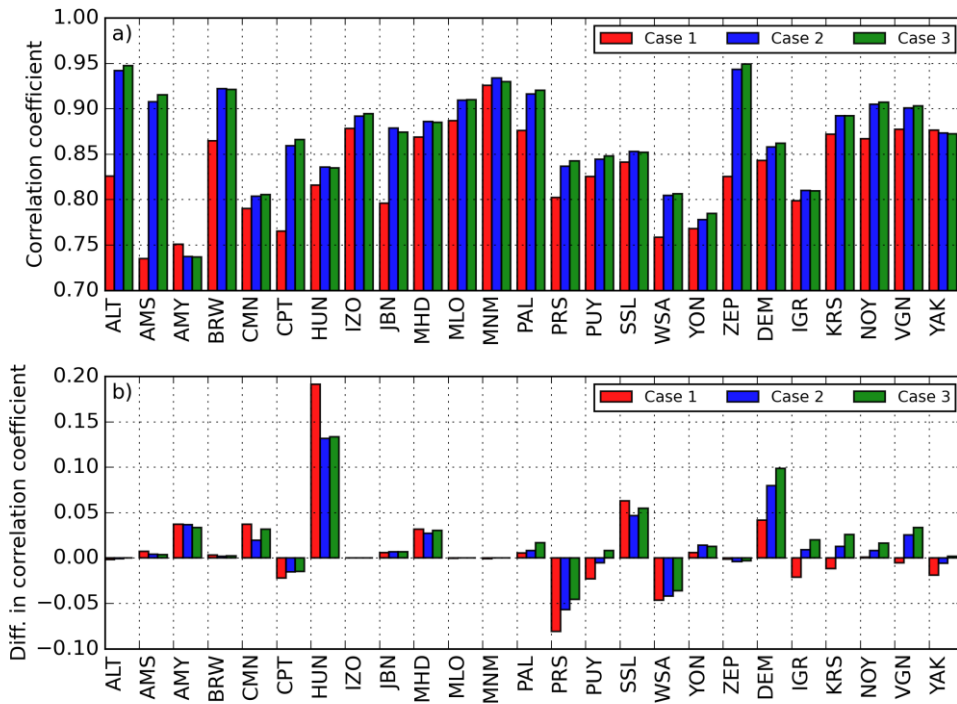
5

6

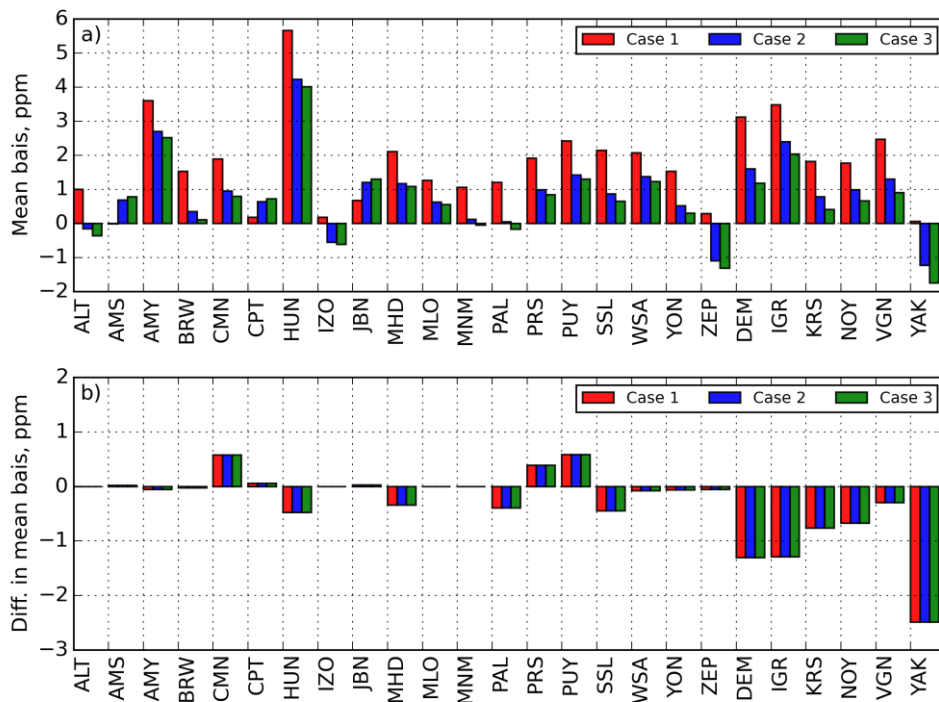
7

8

**Fig. 1.Fig. 2.** 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.

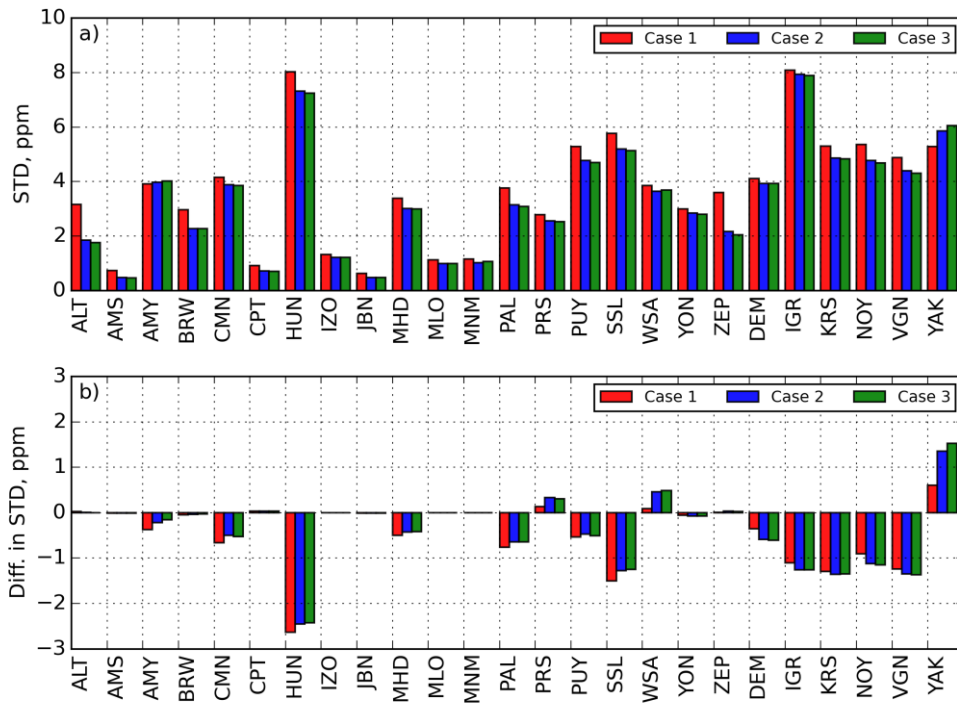


1  
 2 **Fig. 2.Fig. 3.** a) Correlation coefficients between the CO<sub>2</sub> concentrations simulated with the  
 3 coupled model and those observed, b) difference in correlation coefficients due to the  
 4 application of the Lagrangian component (positive values mean the results of the  
 5 coupled model are better than those of the Eulerian model alone) at the selected  
 6 WDCGG and JR-STATION locations for 2009-2010. The definitions of the cases 1-3  
 7 are in Table 1



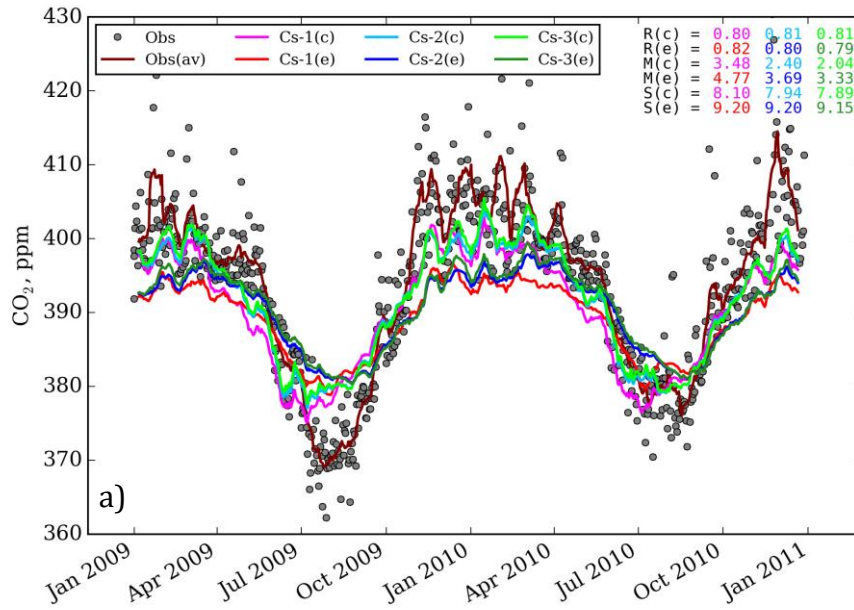
9  
 10 **Fig. 3.Fig. 4.** a) Mean bias for the CO<sub>2</sub> concentrations simulated with the coupled model, b)

1 difference in mean bias due to the application of the Lagrangian component (for  
2 positive bias – the most usual case – negative values mean the results of the coupled  
3 model are better than those of the Eulerian model alone) at the selected WDCGG and  
4 JR-STATION locations for 2009-2010. The definitions of the cases 1-3 are in Table 1  
5

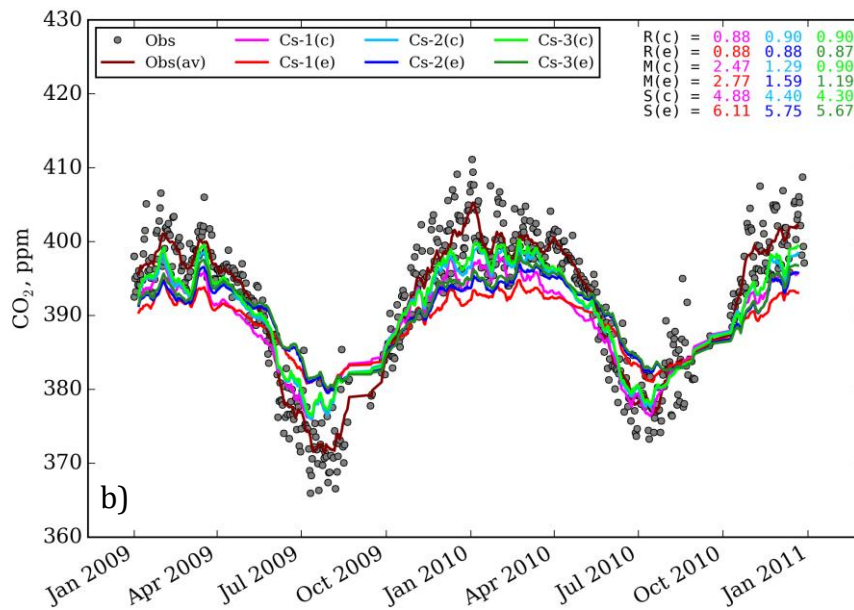


1  
2 **Fig. 4.** **Fig. 5.** a) Standard deviation (STD) for the CO<sub>2</sub> concentration model-observation mismatch  
3 when using the coupled model, b) difference in STD due to the application of  
4 Lagrangian component (negative values mean the results of the coupled model are  
5 better than of the Eulerian model alone) at the selected WDCGG and JR-STATION  
6 locations for 2009-2010. The definitions of the cases 1-3 are in Table 1.

7



1

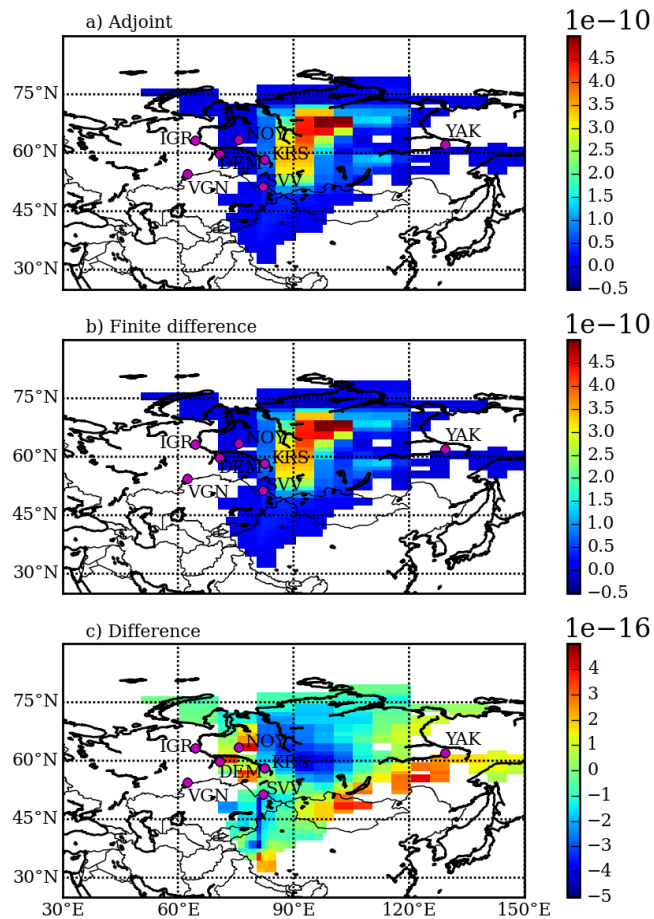


2

3

**Fig. 5-Fig. 6.** CO<sub>2</sub> 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.

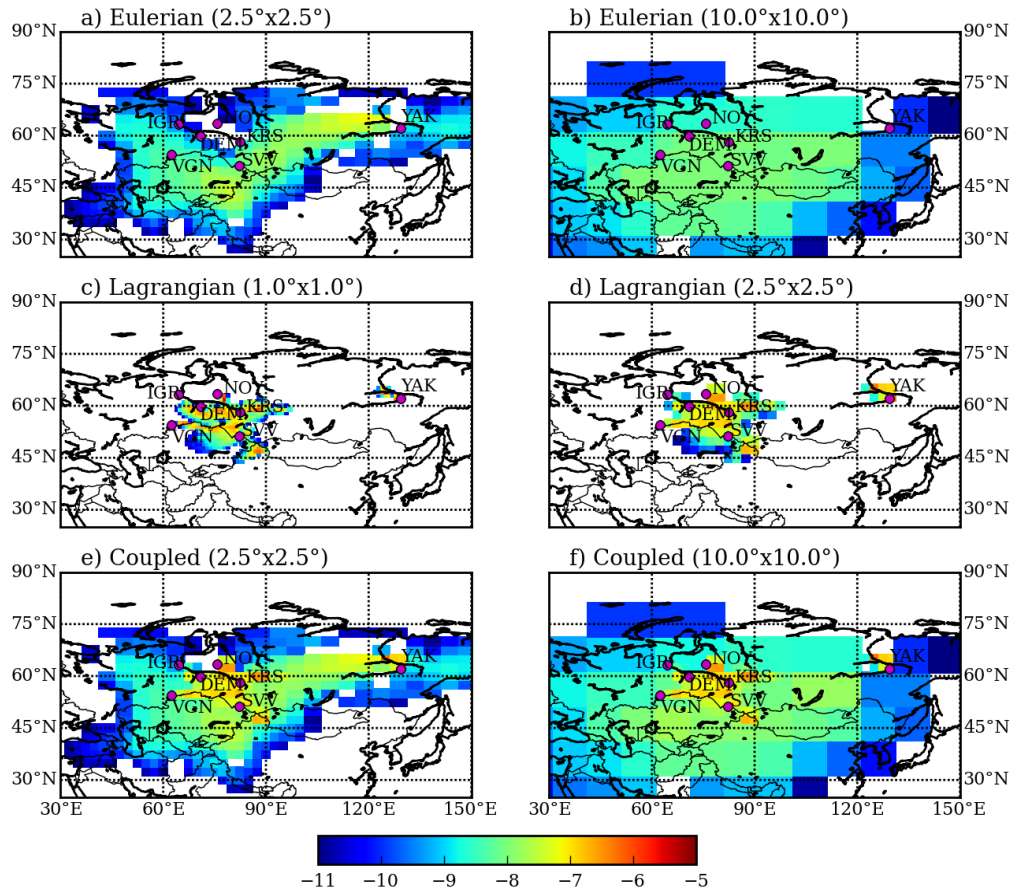
8



1  
 2 **Fig. 6. Fig. 7.** Comparison of sensitivities of CO<sub>2</sub> concentrations (ppm/(μmol/m<sup>2</sup>s)) for test 1: (a)  
 3 sensitivity calculated considering only the Eulerian adjoint model at a resolution of  
 4 2.5°, (b) the same sensitivity calculated directly from NIES forward runs using the one-  
 5 sided numerical finite difference method with perturbations of  $\epsilon$ , and c) the relative  
 6 difference between derived adjoint and the numerical finite difference gradients.  
 7 Magenta dots with labels depicts the locations and names of the Siberian observation  
 8 towers.

9

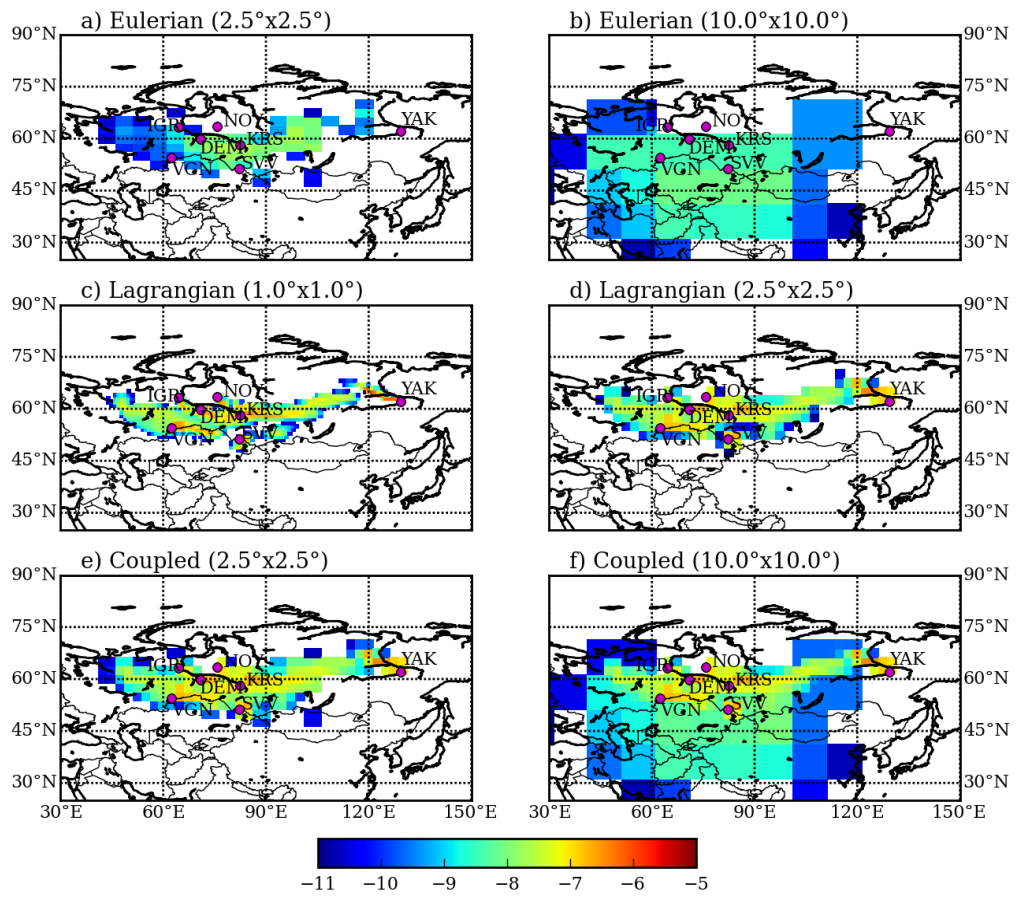




1  
2

3 **Fig. 7. Fig. 8.** Comparison of sensitivities of CO<sub>2</sub> concentrations [ppm/(μmol/m<sup>2</sup>s)] at day 2 (see  
 4 Sect. 5.2.2) calculated using: a) the Eulerian adjoint with a resolution of 2.5°, b) the  
 5 Eulerian adjoint with a resolution of 10.0°, c) the Lagrangian model on the native  
 6 model grid with a resolution of 1.0°, d) as for c), but aggregated on the grid with a  
 7 resolution of 2.5°, e) the coupled adjoint model; results from the Lagrangian adjoint  
 8 model were aggregated on the grid with a resolution of 2.5°, f) as for e), but the  
 9 resolution of the Eulerian adjoint model was 10.0°. Note the logarithmic color scale for  
 10 the plots.

11



1  
2  
3  
4

**Fig. 8.** **Fig. 9.** As for Fig. 78, but for day 4.