A model-based evaluation of inversions of atmospheric transport, using annual mean mixing ratios, as a tool to monitor fluxes of nonreactive trace substances like CO_2 on a continental scale

Manuel Gloor, Song-Miao Fan, Stephen Pacala, Jorge Sarmiento, and Michel Ramonet

Carbon Modeling Consortium, Atmospheric and Oceanic Sciences Program and Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey

Abstract. The inversion of atmospheric transport of CO_2 may potentially be a means for monitoring compliance with emission treaties in the future. There are two types of errors, though, which may cause errors in inversions: (1) amplification of high-frequency data variability given the information loss in the atmosphere by mixing and (2) systematic errors in the CO_2 flux estimates caused by various approximations used to formulate the inversions. In this study we use simulations with atmospheric transport models and a time independent inverse scheme to estimate these errors as a function of network size and the number of flux regions solved for. Our main results are as follows. (1) When solving for 10–20 source regions, the average uncertainty of flux estimates caused by amplification of high-frequency data variability alone decreases strongly with increasing number of stations for up to ~150 randomly positioned stations and then levels off (for 150 stations of the order of ± 0.2 Pg C yr^{-1}). As a rule of thumb, about 10 observing stations are needed per region to be estimated. (2) Of all the sources of systematic errors, modeling error is the largest. Our estimates of SF_6 emissions from five continental regions simulated with 12 different AGCMs differ by up to a factor of 2. The number of observations needed to overcome the information loss due to atmospheric mixing is hence small enough to permit monitoring of fluxes with inversions on a continental scale in principle. Nevertheless errors in transport modeling are still too large for inversions to be a quantitatively reliable option for flux monitoring.

1. Introduction

As a result of fossil fuel burning and land use changes, the CO_2 concentration in the atmosphere has been rising steadily since the mid-18th century (*Neftel et al.* [1985]; currently at the rate of ~1.5 ppm yr^{-1} or ~0.25% yr^{-1}) and is now at the highest level ever since modern humans appeared on Earth [*Barnola et al.*, 1987]. CO_2 is the largest contributor to greenhouse warming of all the greenhouse gases of anthropogenic origin (~50% of the effect according to *Solomon and Srinivasan*, [1995]). Its lifetime in the atmosphere is of the order of hundreds of years because it is unreactive and because the timescale for full equilibration with the ocean is determined by the ocean overturning timescale of the order of 1000 years. This raises serious concerns as to the possible consequences for Earth's climate and argues for the development of monitoring techniques, both to improve the understanding of the carbon cycle and for enforcing future emission controls.

A conceptually elegant monitoring method is the inversion of atmospheric transport, using measurements of atmospheric CO_2 mixing ratios [*Tans et al.*, 1989, 1990; *Enting and Mansbridge*, 1989; *Enting et al.*, 1993]. An inversion is adequate in a situation where one is interested in a cause and where (1) the effect of the cause is more readily accessible to observation

than the cause itself and where (2) one possesses a good conceptualization of the relation between cause and effect. In this case, one may simply apply the inverse of this relation to the effect to characterize the cause.

In the case of anthropogenic trace substances like CO_2 the cause is the fluxes (typically localized near Earth's surface), and the effect is the resulting spatiotemporal mixing ratio distribution in the atmosphere, the flux's "footprint." The long-lived anthropogenic trace substances like CO_2 or SF_6, for example, are currently emitted predominantly from the main industrial centers in the Northern Hemisphere (the North American East Coast, Western Europe, and Southern China and Japan). These emissions cause the well-known latitudinal tracer distribution in the atmospheric surface layer with high values in the Northern Hemisphere midlatitudes, a strong decrease toward the South Pole and a much lower decrease toward the North Pole.

To illustrate the principles of an inversion, let us consider a simple example. Let us conceptualize the atmosphere as a well-mixed Northern and Southern Hemisphere box with interhemispheric exchange parameterized via relaxation of the mixing ratio difference toward zero with an interhemispheric exchange time of the order of 1 year [e.g., *Tans*, 1997]. Let us further assume that a Northern Hemisphere and a Southern Hemisphere flux, ϕ_N and ϕ_S (Pg C yr^{-1}), is emitted at constant rate into each of the two boxes. After a time period of the order of 3–5 times the relaxation timescale, the difference of

the tracer mixing ratios between the two boxes will attain a stationary state. At a stationary state the surface fluxes and the difference between the Northern and Southern Hemisphere mixing ratio, χ_N and χ_S, are related according to

$$\phi_N - \phi_S = \frac{\mu_{tracer}}{\mu_{air}} \frac{1}{2} M_{atm}(\chi_N - \chi_S),$$

and the sum of the surface fluxes to the growth rate of the tracer inventory according to

$$\phi_N + \phi_S = \frac{\mu_{tracer}}{\mu_{air}} \frac{1}{2t} M_{atm}\{[\chi_N(t) + \chi_S(t)]$$
$$- [\chi_N(0) + \chi_S(0)]\},$$

where μ_{tracer} and μ_{air} are the molar masses of the tracer and air, respectively, and M_{atm} is the mass of the atmosphere. The mixing ratio difference between the two hemispheres and the growth rate of the atmospheric tracer inventory hence permits us, in principle, to infer the magnitude of the surface fluxes without the need to measure them directly.

The inversion method that we consider in this paper is a slight generalization of this example. The main differences are the consideration of more than two flux regions and the use of an atmospheric tracer transport model instead of a two-box model. We still assume that the spatial mixing ratio pattern with respect to a reference station is stationary.

To set up the inversion method, we partition Earth's surface into R regions. We then emit an annually repeating tracer flux $\varphi_r = \varphi_r(\vartheta, \varphi, t)$, $r = 1, \ldots, R$ (g m^{-2} s^{-1}) from each region, where ϑ is latitude, φ longitude, and t time. The spatial pattern of the fluxes would, in practice, be chosen as near as possible to the real (unknown) spatial pattern. To a good approximation, the annual mean of the mixing ratio difference with respect to an arbitrary reference station

$$\Delta\chi_r(\mathbf{x}, t) \equiv \chi_r(\mathbf{x}, t) - \chi_r(\mathbf{x}_{ref}, t)$$

reaches a stationary state $\Delta\chi_r(\mathbf{x})$ within a few years, when integrated forward in time from some initial distribution. To obtain flux estimates, we arrange the observations and the footprints sampled at the observation stations as vectors, $\Delta\boldsymbol{\chi}_{obs}$ and $\Delta\boldsymbol{\chi}_r$, respectively. We then determine the linear combination of the sampled footprints $\Delta\boldsymbol{\chi}_r$ that reproduces the observations most faithfully, by minimizing

$$\left| \Delta\boldsymbol{\chi}_{obs} - \sum_{r=1}^{R} \lambda_r \Delta\boldsymbol{\chi}_r \right|$$

with respect to the multipliers λ_r, where λ_r is the estimate of the flux from region r in units of the annual flux $\phi_r = \int_{1\,yr} \int_{region\,r} \varphi_r(\vartheta, \varphi, t)\, d\sigma\, dt$ (Pg C yr^{-1}) ($d\sigma$ is a surface element). If λ_r is positive, then the region r is a source; if λ_r is negative, it is a sink.

This approach allows strictly only the estimation of fluxes that are constant in time. To extend the method to fluxes with an annual cycle superposed on a linear trend, one has to account for the contribution to the annually averaged mixing ratio distribution that results from the covariation of the seasonally varying part of the flux with atmospheric transport. In carbon cycle research this contribution has gained much attention because of its large magnitude caused by the seasonality of biospheric exchange fluxes with the atmosphere. It is known as "rectification" [*Tans et al.*, 1990; *Denning et al.*, 1995]. As an

example for rectification, consider an annually balanced biosphere. If the mean transport during the drawdown season is in the opposite direction to that during respiration, the annual mean signal will not be zero [cf. *Taguchi*, 1996, Figure 1]. A possibility to correct for rectification is to simulate the effect with an atmospheric tracer transport model and a biospheric model, and to presubtract the effect from the observations before the inversion.

The inversion approach that we just described assumes that fluxes do not change from year to year. In the case of the carbon cycle, this is not a very realistic assumption, and the method may sensibly be applied to periods of several years only. In the case of CO_2, it is furthermore advantageous to subtract the simulated fossil fuel footprint from the observations before the inversion because fossil fuel emissions are the best known component of the carbon cycle.

Several CO_2 inversion studies have found that the currently available data constrain sources and sinks poorly [*Keeling et al.*, 1989; *Tans et al.*, 1990; *Fan et al.*, 1998]. For example, *Fan et al.* [1998] made use of weekly mean data from 63 ground-based observation stations, covering the period from 1988 to 1992. They also used estimates of oceanic exchange fluxes by *Takahashi et al.* [1997] and simulations with the biogeochemical ocean model of *Sarmiento et al.* [1995]. Based on these data, they were able to obtain robust estimates of exchange fluxes only for North America and Eurasia. Estimating fluxes from a larger number of regions would have increased uncertainties in the flux estimates to the same or even larger magnitude as the estimates themselves.

There are two principal reasons why inversions may fail. The first is of a purely dynamical nature: the dilution of surface CO_2 flux signals by atmospheric mixing might simply be too strong. Not only does the uncertainty of individual estimates of fluxes increase with an increasing level of mixing (amplification of high-frequency data variability), there is also an increasing chance for the appearance of pairs of strong, spurious, counteracting flux estimates, in particular, for regions located at the same latitude. As "high-frequency variability," we consider the fluctuations in observations with periods of the order of days or less that are caused, for example, by weather systems. As an instrumental definition, we adopt that followed by *Conway et al.* [1994], which is described in detail by *Masarie and Tans* [1995] (it is, in essence, the standard deviation divided by the square root of the number of observations of the residual distribution of the difference between the data and a model fit to the data.) The reason for the occurrence of counteracting flux estimates is that zonal mixing in the troposphere is much quicker than latitudinal mixing. Hence the footprints for fluxes from different regions within one zonal band may be almost indistinguishable. Consider, for example, the difference of the annual mean spatial mixing ratio distribution with reference to the South Pole for fluxes from the North America boreal and Eurasia boreal regions, respectively, in Figure 1. These two signals differ significantly from each other only within and next to the flux regions. Because of this, if CO_2 concentrations were observed along the dateline only, the signals from fluxes from these two regions would be almost indistinguishable. Under these circumstances, the inversion procedure would become unstable, and estimates would be of unrealistically large magnitude [*Tikhonov and Arsenin*, 1977; *Golub and Van Loan*, 1989].

The second possible reason for failure of inversions is the presence of systematic errors, which might be induced by the
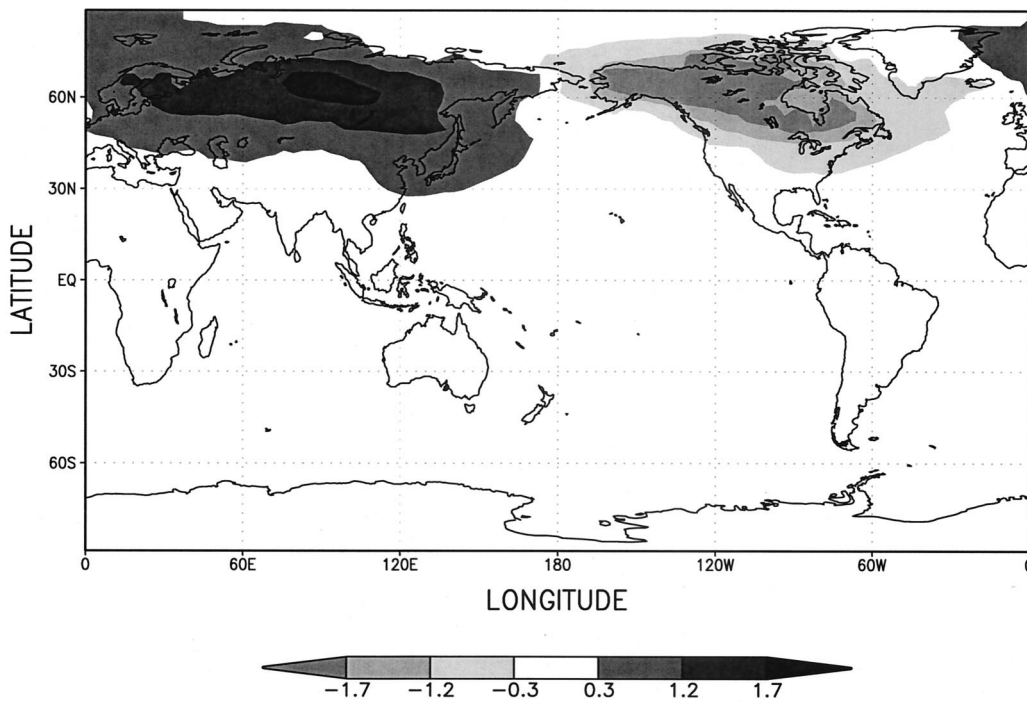
Figure 1. Difference of the mixing ratio distribution ("footprint") (ppm) for fluxes from Eurasia boreal and North America boreal.

assumptions and approximations needed to set up the problem. For an inversion method based on annual mean mixing ratios, these approximations include (1) the simulation of atmospheric transport by a model, which assumes that (a) transport processes like ventilation of the planetary boundary layer (PBL) and interhemispheric transport are correctly represented and (b) synoptic and interannual variability of transport affect estimates only marginally, and (2) the prescription of specific spatiotemporal flux patterns $\varphi_r = \varphi_r(\vartheta, \varphi, t)$ to simulate the footprints $\Delta\chi_r(\mathbf{x})$ for the inversions. For example, the spatial pattern of biospheric exchange fluxes may be modeled proportionally to satellite measurements (normalized difference vegetation index (NDVI)), which may be close to the truth but may also be in error.

These approximations may have two consequences. First, estimates may be biased. Unrealistically low ventilation of the PBL, for example, will cause a systematic underestimation of fluxes if use is made of surface stations only. Second, there will be additional uncertainty in the estimates over that caused by mixing in the atmosphere alone. If, for example, there is a mismatch between real spatiotemporal flux patterns and those used to simulate the footprints, the differences will be misinterpreted by the inversion. These differences will effectively add to and increase the estimated natural high-frequency variability of the data.

To assess the value of inversions for flux monitoring purposes, the magnitude of all these possible errors will be estimated here. One major emphasis of this investigation is to use a method which avoids the convolution of the effect of various error sources and which is independent from a specific observation network. To achieve this goal, we adopt a Monte Carlo approach that generates ensembles of observation networks with randomly positioned observation stations to obtain ensembles of estimates. From these ensembles, we calculate

mean estimates and standard deviations. We base our analysis on the inversion method that we have already discussed. To simulate the footprints $\Delta\chi_r(\mathbf{x})$, we use two atmospheric transport models (SKYHI and global chemical transport model (GCTM), both developed at Geophysical Fluid Dynamics Laboratory/National Oceanic and Atmospheric Administration (GFDL/NOAA)), and use either spatially uniformly distributed flux patterns or fossil fuel flux patterns. We then use these footprints to invert "pseudo-observations" obtained from simulations with various models of fossil fuel burning, oceanic fluxes, land biosphere net primary productivity and respiration, and SF_6 emissions.

In the first part of the paper, we concentrate on the relation between the number of measurement stations and the number of source flux regions which they permit to estimate to a given accuracy, if there were no systematic errors. To achieve this goal, we use the expression for the error propagation for a linear inversion problem and the measured high-frequency data variability of the measurements of NOAA/Climate Monitoring and Diagnostics Laboratory (CMDL) as reported by *Conway et al.* [1994].

In the second part, we quantify biases caused by all the approximations listed above that are needed for inversions based on annual mean observations. We select suitable simulations for the inversions and the pseudo-observations such that estimates are only affected by one source of error at a time. Table 1 summarizes our choice of combinations of inversion schemes and pseudo-observations. We estimate the magnitude of the error in modeled transport in two ways. First, the emissions of SF_6 are estimated from the simulations of all 12 models that participated in the model intercomparison study TransCom2 [*Denning et al.*, 1999]. We base the inversions on the GLOBALVIEW-CO2 network [*National Oceanic and Atmospheric Administration*, 1997] (Also available on Internet via

Table 1. Choice of Combinations of Footprints and Pseudodata to Evaluate the Magnitude of Errors Caused by the Error Sources in an Inversion Based on Annual Mean Observations

	Footprints		Pseudodata	
Error Source	Model	Flux Pattern	Model	Flux Pattern
Atmospheric model				
Transport	GCTM (five regions)	fossil fuel burning	participants of TransCom2: CCC, CSU-GCTM, GFDL-GCTM, GFDL-SKYHI, GISS, GISS-UVIC MUTM, NIRE, TM2, TM3	SF_6 emissions
	GCTM (17 regions)	spatially uniform	SKYHI (17 regions)	spatially uniform
Interannual variability	GCTM and SKYHI (17 regions)	spatially uniform	TM2, 1992 and 1993	biosphere (CASA) fossil fuel emissions
Flux pattern used in the inversion				
Spatial	GCTM (17 regions including North America)	spatially uniform	GCTM (a small subregion of North America)	spatially uniform
Temporal	SKYHI (17 regions)	spatially uniform	SKYHI	biosphere (CASA) (seasonal rectification)

CCC, Canadian Climate Centre general circulation model; CSU-GCTM, Colorado State University general circulation model; GFDL-GCTM, Geophysical Fluid Dynamics Laboratory global chemical tracer transport model (Princeton, New Jersey); GFDL-SKYHI, Geophysical Fluid Dynamics Laboratory general circulation model; GISS, NASA-GISS tracer transport model; GISS-UVIC, GISS-UVIC tracer transport model; MUTM, Melbourne University tracer model; NIRE, NIRE tracer transport model; TM2, tracer model version 2; TM3, tracer model version 3, CASA, Carnegie Ames Stanford Approach; and SF_6, sulfur hexafluoride.

anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/ GLOBALVIEW) and fossil fuel flux patterns. These are similar to the flux patterns of SF_6 because both are strongly tied to energy consumption [*Denning et al.*, 1999]. Second, we determine transport errors with the two conceptually identical inversion schemes of GCTM and SKYHI and relate them to differences in model transport. We address the errors caused by neglecting the interannual variability of atmospheric transport by inverting the simulations of the annually repeating Carnegie Ames Stanford Approach (CASA) biosphere and fossil fuel emissions from different model years and comparing the difference of the estimates (the emissions themselves do not vary from year to year.) We determine the errors caused by differences between the flux patterns used to simulate the footprints for the inversion and the flux pattern that caused the actual mixing ratio distribution by inverting the mixing ratio distribution that results from a highly localized source within North America. North America itself is one of the flux regions used to simulate the footprints for the inversion. Finally, to estimate the biases induced by the neglect of the covariation of the seasonal cycle of fluxes with transport in an inversion based on annual means, we invert the annual mean mixing ratio distribution caused by a balanced biosphere (i.e., a biosphere with zero annual flux) predicted by CASA. A balanced biosphere is suited for this purpose because of its strong seasonal cycle.

2. Methods for Error Estimation

2.1. Amplification of High-Frequency Data Variability by the Inversion Process

Dilution of flux signals by mixing in the atmosphere leads to uncertainty of flux estimates because of amplification of the natural high-frequency variability in the data. To quantify the uncertainty of flux estimates due to this error source, it is helpful to arrange the footprints of regional fluxes, sampled at a specific network, in a matrix $A = \{\Delta\chi_1, \ldots, \Delta\chi_R\}$. The ith component of the vector $\Delta\chi_r$ is the annually averaged mixing ratio (with respect to a reference station) observed at station i

for a tracer emitted from region r. The matrix A maps a specific combination of regional fluxes to the observed spatio-temporal pattern: $\Delta\chi = A\lambda$. Correspondingly, the flux contributions to an observed signal in units of ϕ_r are obtained from its pseudo-inverse by $\lambda = A^{-1}\Delta\chi$ [*Golub and Van Loan*, 1989]. The variance-covariance matrix of the flux estimates, C_φ, may now be expressed in terms of the pseudo-inverse of A and the covariance matrix of the data $C_{\Delta\chi}$ as $C_\varphi = A^{-1}C_{\Delta\chi}(A^{-1})^T$. The diagonal elements of C_φ and $C_{\Delta\chi}$ are the variances of the flux estimates and the data, respectively. The off-diagonal elements are the correlation of estimates from different regions and data from different observation stations to each other, respectively. There are hence two parts which contribute to the uncertainty of estimates: the data covariance $C_{\Delta\chi}$ and its amplification, which is given by the matrices A^{-1} and $(A^{-1})^T$ which bracket $C_{\Delta\chi}$. The matrix $\mathbf{ErrAmp} \equiv A^{-1}(A^{-1})^T$, with units $(Pg\ C\ yr^{-1}\ ppm^{-1})^2$, is called the error amplification matrix [e.g., *Menke*, 1989]. This matrix is independent of any assumptions on data uncertainties and hence reflects exclusively atmospheric transport properties and the choice of sites and number and location of source regions. It is one decisive piece of information for the determination of the detection limit of fluxes with an inversion method. Writing out its diagonal elements, $\mathbf{ErrAmp}_{rr} = \Sigma_{k=1}^{S} (\delta\phi_r/\delta\Delta\chi_k)(\delta\phi_r/\delta\Delta\chi_k)$, helps to clarify its meaning: they are the sum over all observation stations of the sensitivities of the flux estimate of region r to changes in data. Finally, an overall measure for the amplification of high-frequency data variability is the average over all regions: $(1/R)\ \Sigma_{r=1}^{R}\ \mathbf{ErrAmp}_{rr}$, which we will call mean error amplification.

2.2. Method to Estimate Systematic Errors and Amplification of High-Frequency Data Variability Independently From a Specific Network

Let us assume for the moment that the mixing ratio distribution observed in the atmosphere is the one from fossil fuel emissions, $\Delta\chi_{FF}(\mathbf{x})$, and that we would like to estimate the regional fluxes that caused $\Delta\chi_{FF}(\mathbf{x})$ with an inversion. Let us further assume that we have no preconception of the spatial
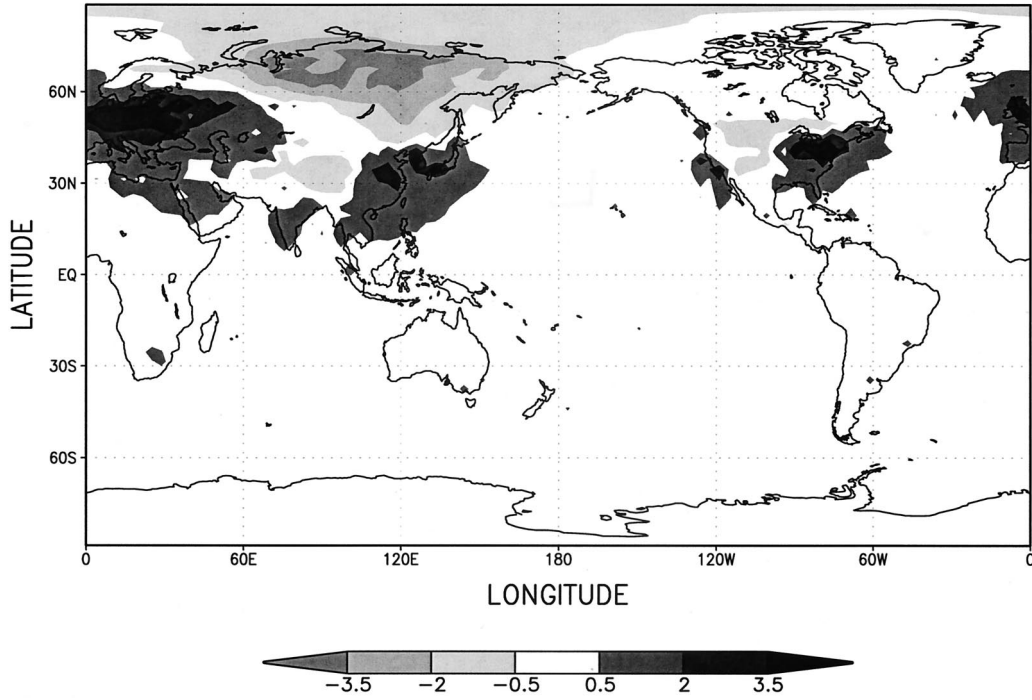
Figure 2. Difference between the mixing ratio distribution resulting from the linear superposition of the regional footprints (simulated using spatially uniform flux patterns) multiplied by the annual regional fossil fuel emissions ($\phi_{FF,r} = \int_{1 \text{ yr,region} r} \varphi_{FF} \, d\sigma \, dt$) and the mixing ratio distribution resulting from fossil fuel emissions ($\Delta\chi_{FF}(\mathbf{x})$): $\Sigma_{r=1}^{R} \, \phi_{FF,r} \Delta\chi_r(\mathbf{x}) \, - \, \Delta\chi_{FF}(\mathbf{x})$.

pattern of these fluxes so we choose spatially uniform fluxes to simulate the footprints for the inversion. The resulting smooth mixing ratio distributions contrast drastically with the fossil fuel mixing ratio distribution, which exhibits very large local maxima at the main industrial regions, since the emissions are strongly concentrated there. Accordingly, it is not possible to represent exactly the fossil fuel mixing ratio distribution in the atmosphere as a linear combination of the footprints, and the estimate of fossil fuel emissions will be biased. To illustrate this point, consider in Figure 2 the difference $\Sigma_{r=1}^{R} \, \phi_{FF,r} \Delta\chi_r(\mathbf{x}) - \Delta\chi_{FF}(\mathbf{x})$ between the mixing ratio distribution, which results from the combination $\Sigma_{r=1}^{R} \, \phi_{FF,r} \Delta\chi_r(\mathbf{x})$ of regional footprints generated by homogeneously distributed fluxes multiplied with the regional fossil fuel emissions ($\phi_{FF,r} \equiv \int_{\text{Region } r} \varphi_{FF}(\vartheta, \varphi) \, d\sigma \, dt$), and the actual fossil fuel footprint $\Delta\chi_{FF}(\mathbf{x})$. Near the industrial centers the approximation of the fossil fuel mixing ratio distribution with those based on homogeneously distributed sources is poor. As a consequence, stations positioned there will drastically misinterpret the local signals, and the estimates will be unprecise. Typically, estimates based on pseudodata sampled at a concrete observation network like GLOBALVIEW-CO2 indeed deviate strongly from true fluxes (here fossil fuel). Nevertheless these estimates convolute different causes of errors with each other. For example, it is unclear how much of the disagreement is due to the limited, possibly insufficient size of the network, how much due to the specific positions of the observation stations, and how much is really due to spatial mismatch. One needs a methodology that permits one to distinguish the contributions of these various error causes to a flux estimate, and furthermore, one would like to determine biases, like the one caused by spatial mismatch, independently from a specific observation network.

For this reason, we use here a Monte Carlo approach which positions observation stations randomly to generate ensembles of estimates. To determine biases caused by systematic errors, we specifically proceed as follows. First, we determine a mean flux estimate from a particular footprint $\Delta\chi(\mathbf{x})$ used as pseudo-observation (e.g., the one resulting from fossil fuel emissions) by taking the mean over an ensemble of N estimates based on N random networks: $\boldsymbol{\phi}_{\text{mean}} \equiv (1/N) \, \Sigma_{i=1}^{N} \, A_i^{-1} \Delta\chi_i$. Here A_i is the map between regional fluxes and observations for the ith random network and $\Delta\chi_i$ are the values of the pseudo-observations. The mean estimate $\boldsymbol{\phi}_{\text{mean}}$ will generally differ from the true fluxes because of the differences of the spatial structure of footprints and pseudodata. The difference between the mean estimate and the true fluxes then is the bias: $\boldsymbol{\phi}_{\text{bias}} \equiv \boldsymbol{\phi}_{\text{mean}} - \boldsymbol{\phi}_{\text{true}}$, where $\boldsymbol{\phi}_{\text{true}}$ are the "true" fluxes that were used to generate the pseudo-observations. Mean estimates, determined as described, are found to be essentially independent from the fixed number of stations used, as long as they encompass no less than 50 stations.

Similarly, we may measure the enhanced scatter of estimates caused by differences of the spatial structure by the standard deviation of the estimates of the ensemble of random networks:

$$STD_r = \left(\frac{1}{(N-1)} \sum_{i=1}^{N} (\phi_{i,r} - \phi_{\text{mean},r})^2 \right)^{1/2},$$

where $\phi_{i,r}$ is the estimate of the flux from region r based on the randomly chosen network from the ith draw. For the example above, for instance, one expects large scatter of the estimates of fluxes from the main industrial centers if based on networks with randomly positioned stations, because of the large dis-

agreement between the footprints and the fossil fuel mixing ratio there, and much less scatter of the estimates for the remaining regions.

We use a similar approach to determine how much estimates are affected by the neglect of interannual variability of transport in an inversion. We determine mean estimates (ϕ_{mean}) for two different years with different atmospheric transport but identical biospheric and fossil fuel flux patterns simulated by TM2 (which is based on analyzed winds). To determine the maps A_i, we use the footprints generated with source strengths that are spatially-uniformly-distributed. The difference between the mean estimates (ϕ_{mean}) from different years gives us an estimate of the magnitude of the differences of estimates caused by the interannual variability of model transport. Although this use of two specific years may be criticized for its lack of generality, we obtained similar conclusions using two different years, and our main point is to estimate the order of magnitude of the biases and which regions are most affected.

For the determination of systematic errors, we generally use randomly generated networks with 150 observation stations because (1) ~150 stations are necessary to estimate fluxes on a continental scale (for 10–20 regions), if positioned randomly, to suppress the errors caused by the amplification of high-frequency data variability; (2) because systematic biases manifest themselves strongly if inversions are based on small networks with 40–80 stations but there is a "saturation" level from which point on biases do not decrease much further with additional stations; for the estimation of 10–20 regions this saturation level is reached at ~150 stations, and (3) because for the large number of 150 stations, networks with randomly positioned stations are almost as efficient as optimized networks (cf. section 4.1).

Finally, to determine the relation between estimate uncertainty and number of observation stations, we average the mean amplification of high-frequency data variability over random draws of observation networks with a fixed number of stations (section 2.1):

$$\frac{1}{N} \sum_{i=1}^{N} \left(\sum_{r=1}^{R} \frac{1}{R} \, \mathrm{ErrAmp}_r^{(i)} \right)^{1/2} .$$

3. Description of Models, Simulations, and Footprints $\Delta\chi_r(\mathbf{x})$

3.1. Model Characteristics

The off-line atmospheric transport model GCTM [*Mahlman and Moxim*, 1978] developed at GFDL/NOAA is driven at a time step of 26 min by linearly extrapolated 6-hour time-averaged, annually repeating climatological winds. These were originally simulated by a modified, seasonally varying version of ZODIAK [*Holloway and Manabe*, 1971]. GCTM has no diurnal physics. It solves the tracer transport equation on 11 sigma levels, extending from Earth's surface to about 30 km height. The centers of the lowermost layers lie at 0.08, 0.5, 1.5, and 3.1 km height. Horizontally, an equal-area grid with an ~265 km × 265 km box size is used. Vertical subgrid-scale transport is parameterized by an eddy diffusion coefficient, which takes mixing due to velocity shear, as well as convection in case of unstable density profiles, into account. In addition, within the PBL, a mixing length scheme with decreasing mixing length from the lowest to the third model layer is added [*Levy et al.*, 1982, 1989]. The mixing lengths for these three levels

have been adjusted so as to match ^{222}Rn profiles compiled by *Liu et al.* [1984].

The atmospheric general calculation model (AGCM) SKYHI [*Fels et al.*, 1980], developed at GFDL/NOAA, calculates tracer transport on-line, has 40 vertical levels extending from the surface up to ~80 km, uses a hybrid pressure-sigma coordinate in the vertical, and has a diurnal cycle of solar radiation. The horizontal grid is regular, and the resolution used for this study is 3° longitude by 3.6° latitude. There are eight layers between the surface and 5.2 km height, the lowermost ones centered 0.08, 0.27, 0.74, and 1.38 km above Earth's surface. The parameterization of vertical mixing in the PBL in SKYHI differs from that of GCTM: in case of potential temperature inversions, the vertical diffusion coefficient is set to the maximal value that does not cause numerical instability. A detailed description of SKYHI and its climatology is given by *Hamilton et al.* [1995].

The transport core of the TM2 model derives from the Goddard Institute for Space Studies (GISS) tracer transport model version of *Russell and Lerner* [1981], and European Centre for Medium-Range Weather Forecasting (ECMWF) analyzed winds are used to drive tracer transport off-line [*Heimann*, 1996]. The model grid encompasses nine layers in the vertical, extending from the ground up to 1 mbar (lowermost layers 0.22, 0.8, 1.85, 3.56, and 5.86 km), and the horizontal resolution is 8° latitude by 10° longitude, which is much coarser than that of both GCTM and SKYHI. Subgrid-scale transport is simulated by a cumulus cloud convection scheme, and vertical diffusion in the PBL is parameterized in a similar way as in GCTM.

References for the models participating in TRANSCOM2 are given by *Denning et al.* [1999].

3.2. Footprint Simulations

We simulated two different types of footprints, $\Delta\chi_r(\mathbf{x})$, $r = 1, \ldots, R$, for the inversion schemes: one type with spatially uniform flux patterns (within a region) and another type with the flux pattern of fossil fuel burning. For the first type, we build the following sets by combining or partitioning the footprints of 17 basic regions (Figure 3): (1) 11 source regions: Eurasian boreal and Eurasian temperate combined, North American boreal and temperate combined, Indian Ocean tropical and temperate combined, Atlantic tropics and South Atlantic temperate combined, Australasia, Pacific tropics, and South Pacific temperate combined, and all remaining 17 regions by themselves; (2) 22 source regions, including five additional regions, which result from splitting South America, Africa, South Pacific temperate, Australasia, and North Pacific temperate in two parts (Figure 3); and (3) 25 regions (focus on North America): N. American boreal and temperate split up in 10 subregions. For the second type, we simulated footprints with the pattern of fossil fuel emissions from five regions: North America boreal and temperate, Eurasia boreal and temperate, and the combination of South America, Africa, and Australasia.

For use as pseudo-observations, we simulated mixing ratio distributions for the following flux patterns: fossil fuel emissions for the year 1990 [*Andres et al.*, 1996], estimated oceanic fluxes [*Takahashi et al.*, 1997], and estimates of land biosphere net primary productivity and respiration derived from satellite measurements (NDVI [*Potter et al.*, 1993]). Each of these pseudo-observations have been simulated with GCTM, SKYHI, and TM2. We also use the simulations of the distribution of
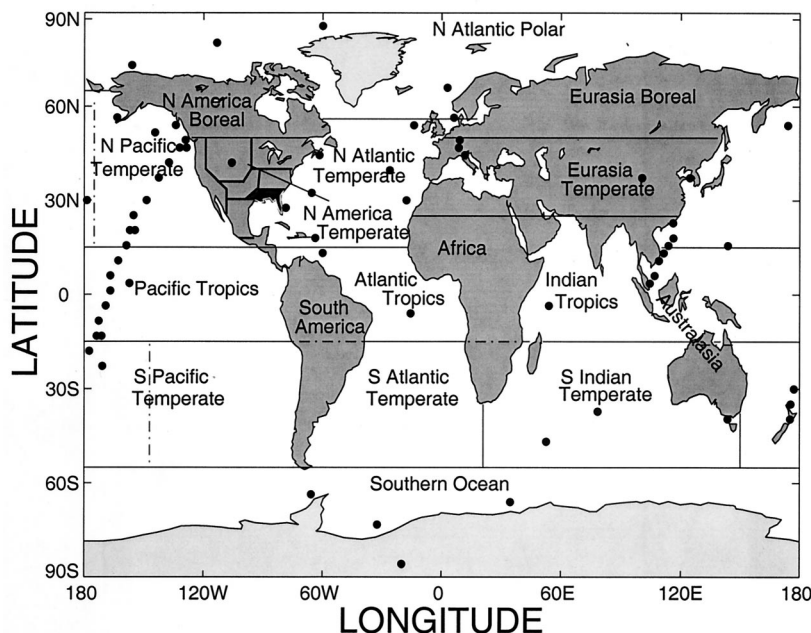
Figure 3. Partitioning of Earth's surface in regions used for the inversion schemes.

SF_6 in the atmosphere from the participants of TransCom2 [*Denning et al.*, 1999].

3.3. Relation Between Footprints $\Delta\chi_r(x)$ and Model Transport Properties

Annually averaged surface mixing ratio distributions due to regional fluxes (Figure 4) may be conceptualized as a superposition of an approximately zonally symmetric "background" field, reflecting the predominance of zonal winds and related rapid zonal mixing in the troposphere, and a strong, "regional" signal within the flux region itself, determined by the ventilation rate of the PBL. Background fields exhibit interhemispheric slopes in the surface layer between ± 2 ppm Pg C^{-1} yr^{-1}, and regional signals vary from 2 to 4 ppm Pg C^{-1} yr^{-1} for the northernmost continents to $1-1.2$ ppm Pg C^{-1} yr^{-1} in the midlatitude regions, and to 0.6 ppm Pg C^{-1} yr^{-1} in tropical, continental regions where convective ventilation of the PBL is largest.

The magnitude of the regional signal differs considerably between SKYHI and GCTM, but the spatial structure (i.e., form of isolines) is almost identical; the background fields are also similar. Some signals resulting from fluxes from a bounded region at Earth's surface as simulated by GCTM and SKYHI are shown in Figure 4. For most regions, regional annual-mean surface mixing ratios simulated by SKYHI are lower than those simulated by GCTM, with the exception of the three northernmost regions: North Atlantic polar, Eurasian boreal, and North American boreal. There, surface mixing ratios simulated by GCTM are smaller by about 50%. The reason for these discrepancies is that the PBL ventilation simulated by SKYHI is stronger than that by GCTM for most regions (for the reason explained in section 3.1). The exceptional regions are the northernmost ones, for which SKYHI predicts near-ground inversions during winter and hence surface fluxes are trapped, resulting in high surface concentrations. The resolution of the PBL of ZODIAK, from which the off-line winds for GCTM derive, is too coarse to resolve such wintertime inversions, and the PBL ventilation rate is therefore much larger there. Rawinsonde data from these regions support the simulations of SKYHI.

Results of inversions based on zonally averaged models rely strongly on the interhemispheric gradient resulting from asymmetric fluxes in both hemispheres. The magnitude of this interhemispheric gradient for fixed source strength depends both on the efficiency of PBL ventilation and interhemispheric mass exchange. We use a simulation of the CO_2 distribution resulting from fossil fuel emissions to compare the models. The zonally averaged interhemispheric gradient resulting from fossil fuel burning for 1990 [*Andres et al.*, 1996] as simulated by GCTM is larger than that by SKYHI by $\sim 10-15\%$, whereas the latitudinal structure is similar (the interhemispheric exchange times are 0.8 and 0.9 years, respectively). The zonally averaged interhemispheric gradient from pole to pole as simulated by TM2 is similar to that by GCTM, whereas in the northern midlatitudes the mixing ratios are smaller than in GCTM by $\sim 15-20\%$. An intercomparison study by *Law et al.* [1996], which encompassed 11 other models as well, showed that the interhemispheric gradient simulated by GCTM caused by fossil fuel emissions is larger in the midlatitudes of the Northern Hemisphere compared to most other models. A large part of the discrepancy is attributable to the higher spatial resolution of GCTM, which permits much larger spatially localized signals compared to the remaining models (compare with *Denning et al.* [1999, Table 1 and Figure 4]).

4. Evaluation of Error Sources

4.1. Relation Between Estimate Uncertainty and Number of Observation Stations

An essential concern in applying inversions to the monitoring of surface fluxes is a determination of the minimum number of stations needed to estimate fluxes to a reasonable accuracy for a given number of regions. Also of interest is the maximum number of regions for which surface fluxes may be
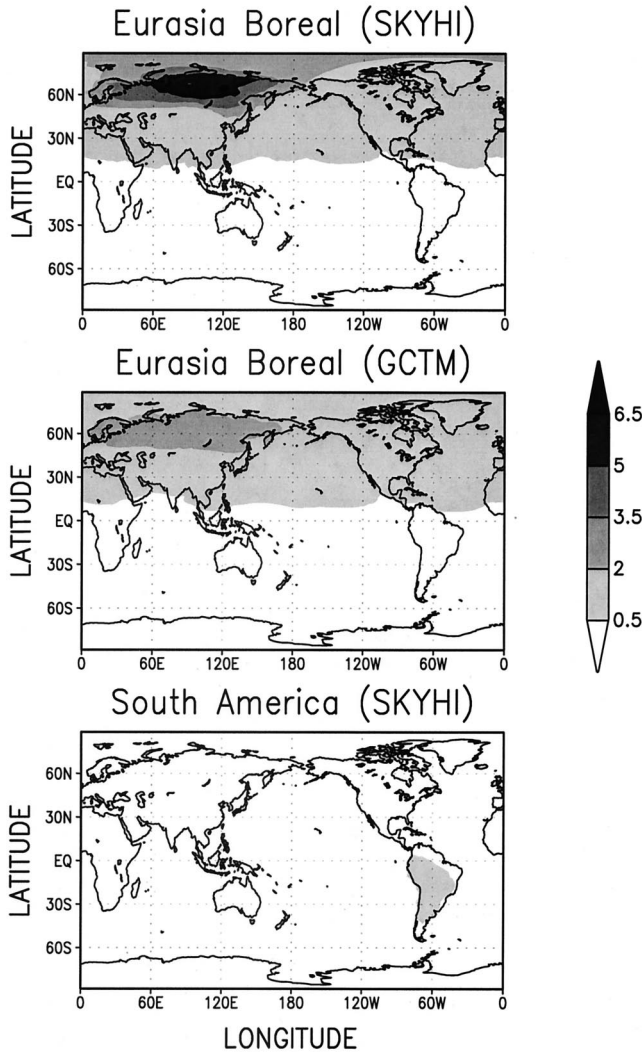
Figure 4. Carbon dioxide mixing ratio distribution (ppm) with reference to the South Pole for a flux from Eurasia boreal simulated by (top) SKYHI and (middle) GCTM, and (bottom) a flux from South America simulated by SKYHI.

estimated to satisfying accuracy, given a fixed number of observation stations (the "spatial resolvability" of the method).

To determine average flux estimate errors from amplification of high-frequency data variability (with the formula $A^{-1}C_{\Delta\chi}(A^{-1})^T$) (section 2.1)), one needs an estimate of high-frequency data variability (the diagonal elements of $C_{\Delta\chi}$) as a function of latitude and longitude. We cannot rely on our model simulations for these estimates because our biospheric model has no daily cycle of biospheric fluxes. Here we will instead use the variability in the available observations. Guided by the magnitude of observed variability, two types of observation stations, "continental" and "remote," positioned on islands in the oceans, are usually distinguished. High-frequency data variability for remote stations were reported by *Conway et al.* [1994]. Standard deviations of annual mean data from the year 1992 are mostly of the order of 0.1–0.2 ppm, with the exception of three stations (maximal standard error for Cape Meares, $\sigma \sim 0.3$ ppm, where winds loaded with pollutants occasionally blow from the land). Data variability at continental stations in temperate and boreal ecozones is larger. For

example, *Bakwin et al.* [1995] report mixing ratios from measurements on a very tall tower in North Carolina. Monthly standard deviations of daily mean CO_2 mixing ratios at 496 m height are of the order of 4 ppm. The standard error of the annual mean mixing ratio accordingly is ~1.2 ppm, the value that we adopt in the following for continental stations.

Consider in Plate 1 (left panel) the dependence of the amplification of high-frequency data variability on the number of observation stations and flux regions. For all surface partitionings the decrease of the mean error amplification with increasing amount of stations is dramatic for small networks of up to ~150 stations and modest for larger ones. Installation of more than 150 monitoring stations on a continental scale (i.e., solving for 10–20 flux regions) is hence not very helpful because uncertainties decrease only very marginally. The decrease follows ~$1/\sqrt{N}$ (i.e., like statistical counting error), where N is the number of observation stations. The level of amplification of high-frequency data variability for large networks is not sensitive to the number of flux regions used for the inversion.

Conversely, these curves show that trying to estimate fluxes from more regions than approximately a tenth of the number of observation stations is not possible (at least if stations are positioned randomly) because the flux errors caused by the amplification of high-frequency data variability alone (not to mention systematic errors) in that case reach uncertainties of 1 Pg C yr^{-1} region^{-1} (the product of the value of error amplification with the value of high-frequency data variability), which makes the solutions useless for most purposes.

Finally and most important, for an intermediate value of high-frequency data variability of 0.5 ppm for each of 150 stations used to estimate fluxes from 10–20 regions, the amplification of high-frequency data variability for the estimation of fluxes results in errors of the order of 0.2–0.3 Pg C yr^{-1} region^{-1} (the product of the mean error amplification ~0.5 Pg C yr^{-1} ppm^{-1} region^{-1} with the level of high-frequency data variability ~0.5 ppm). Note that a larger value of ~0.4 Pg C yr^{-1} region^{-1} is obtained if only remote stations are used (for which the high-frequency data variability is only 0.2 ppm, but the mean error amplification is 1.9 Pg C yr^{-1} ppm^{-1} region^{-1}). The ratio of the number of observation stations needed to the number of regions to be estimated, as far as amplification of high-frequency data variability is concerned, is therefore ~10; monitoring of CO_2 exchange fluxes would thus be feasible with a fairly reasonable measuring effort.

The amplification of high-frequency data variability increases strongly with height because information on fluxes fades away quickly with increasing distance to the ground. We find approximately a fivefold increase between 0 and 5 km height, and a 20-fold increase between 0 and 10 km height. The profile of error amplification with height over Earth's surface was determined for layers centered at 0.08, 0.5, 1.5, 3.1, 5.5, 8.7, and 12 km height and using the same method as described in section 2.2. The values are global averages. This result means that airplane transects at high altitudes (≥5 km) are not efficient in constraining inversions to estimate surface fluxes. It is important to notice what this result does not imply. First, high-altitude transects may well be of considerable value in constraining the modeling of atmospheric transport, as suggested for example, by the results of the model intercomparison study TransCom2 [*Denning et al.*, 1999]. Second, the result is restricted purely to high-altitude transects and is not valid for vertical profiles extending to the earth surface.

Note also in Plate 1 the convergence of optimal networks
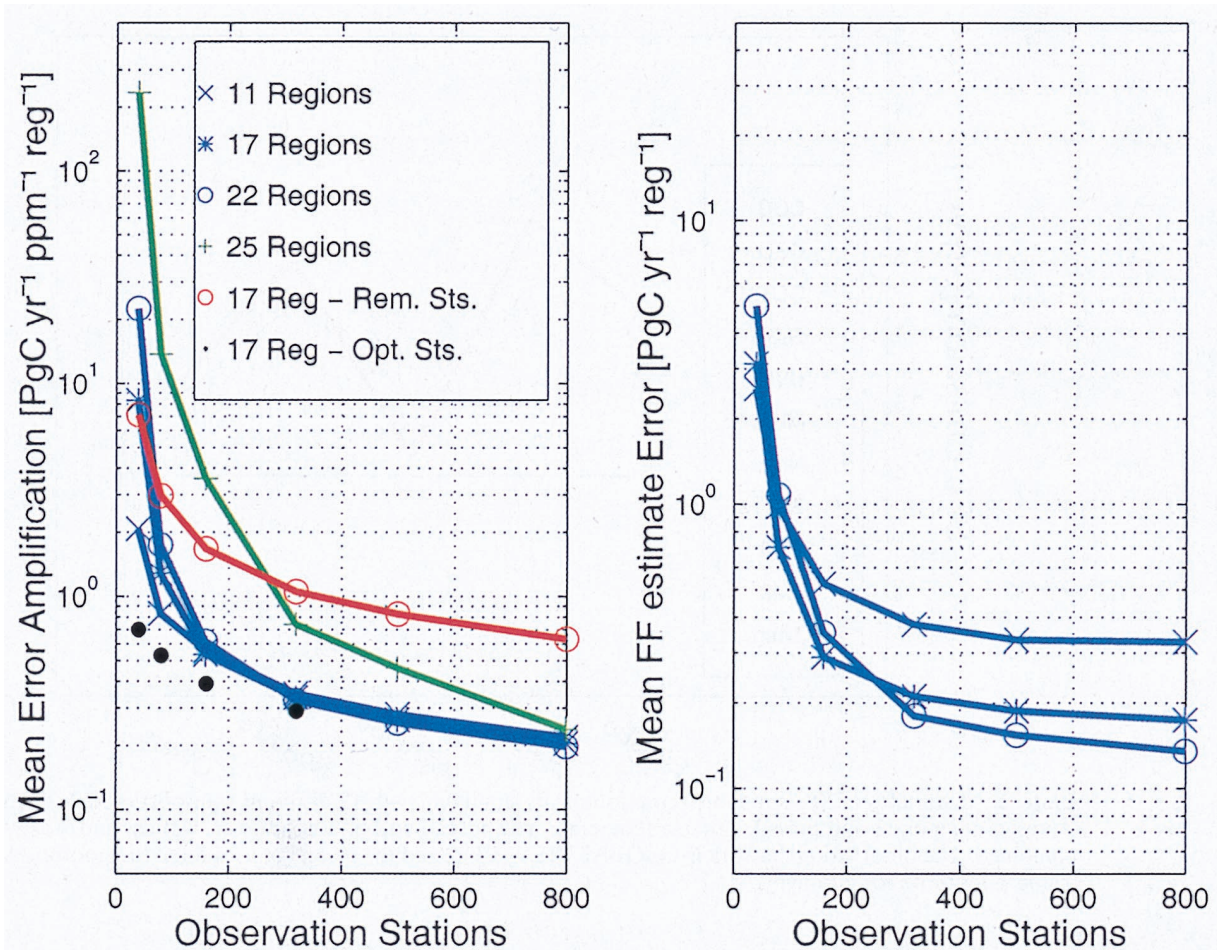
Plate 1. (left) Average error amplification ($(1/R)\ \Sigma_{r=1}^{R}\ \mathrm{ErrAmp}_{rr})^{1/2}$) and (right) average bias for the estimation of fossil fuel emissions for an inversion scheme based on a uniformly distributed flux from Earth's surface as a function of the number of observation stations for various partitionings of Earth's surface. The determination of the averages over random networks is described in section 2.2. The red curve was obtained permitting only stations in the oceans, and the black dots were determined with an algorithm that places stations optimally for the estimation of surface fluxes with minimal uncertainty.

toward random networks with increasing number of observation stations. An optimized network with 40 stations is ~10 times more efficient in constraining sources and sinks compared to a randomly positioned network for the estimation of 17 regions, whereas an optimal network with 160 stations (for which the mean error amplification is minimal) is only marginally superior to a random network (by ~25%). This is one of the justifications for the use of 150 stations for random networks as a basis to determine the magnitude of systematic biases. Optimal networks were determined by minimizing the uncertainty of estimates, given a best estimate of high-frequency variability. The optimizations were performed with the simulated annealing algorithm [*Kirkpatrick et al.*, 1983]. For a more detailed discussion, we refer to M. Gloor et al. (Optimal network design for the purpose of inverse modeling: A model study, submitted to *Global Biogeochemical Cycles*, 1999).

The decline of biases with the number of stations is illustrated in Plate 1 (right panel) for the same example as in section 2.2: the recovery of fossil fuel emissions with an inversion based on spatially uniform fluxes. The main feature here is that there is a saturation level for the number of stations, above which biases do not decrease any further with additional

stations. For the estimation of 10–20 regions, this level is approximately reached with 150 stations, which is the second of the justifications mentioned in the Introduction to base our analysis on 150 observation stations.

4.2. Relation Between Atmospheric Transport and Regional Flux Estimate Uncertainty

The Monte Carlo approach applied to regional error amplification permits one to identify the regions with largest information loss (Figure 5). For the 17-source region partitioning, these are South America, Africa, Australasia, and to a lesser extent, Eurasia temperate. Further subpartitioning of the first three regions into tropical and subtropical zones reveals that the weakest constrained regions are equatorial South America, equatorial Australasia, and equatorial Africa (the average error amplifications for Africa, South America, and Australasia are 0.31, 0.54, and 0.32 Pg C yr^{-1} ppm^{-1} region^{-1}, respectively). As mentioned above, fluxes from continental, equatorial regions, result in comparably small signals within the source region itself. This is due to strong convective ventilation of the PBL. The regional error amplification is hence inversely related to the regional signals. This simple result illustrates that what really determines the amplification of high-frequency
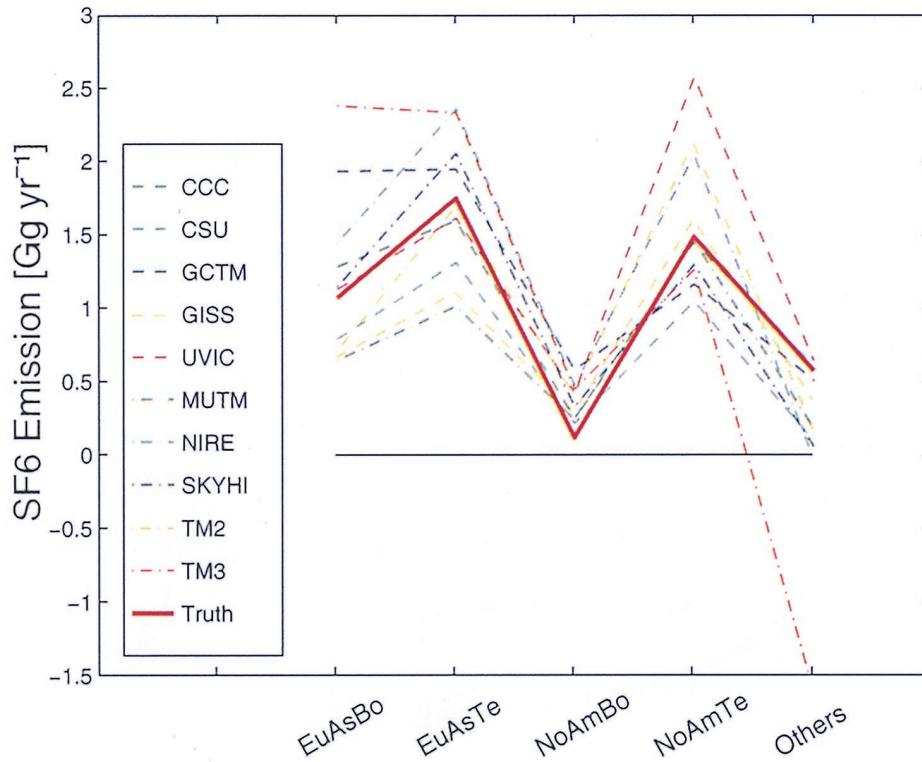
Plate 2. Estimates of SF$_6$ emissions from footprints simulated with 12 different transport models for North America temperate and boreal, Eurasia temperate and boreal, and South America, Africa, and Australasia combined. The observation network is GLOBALVIEW-CO2, and the inversion uses fossil fuel footprints. See Table 1 footnote for acronyms.
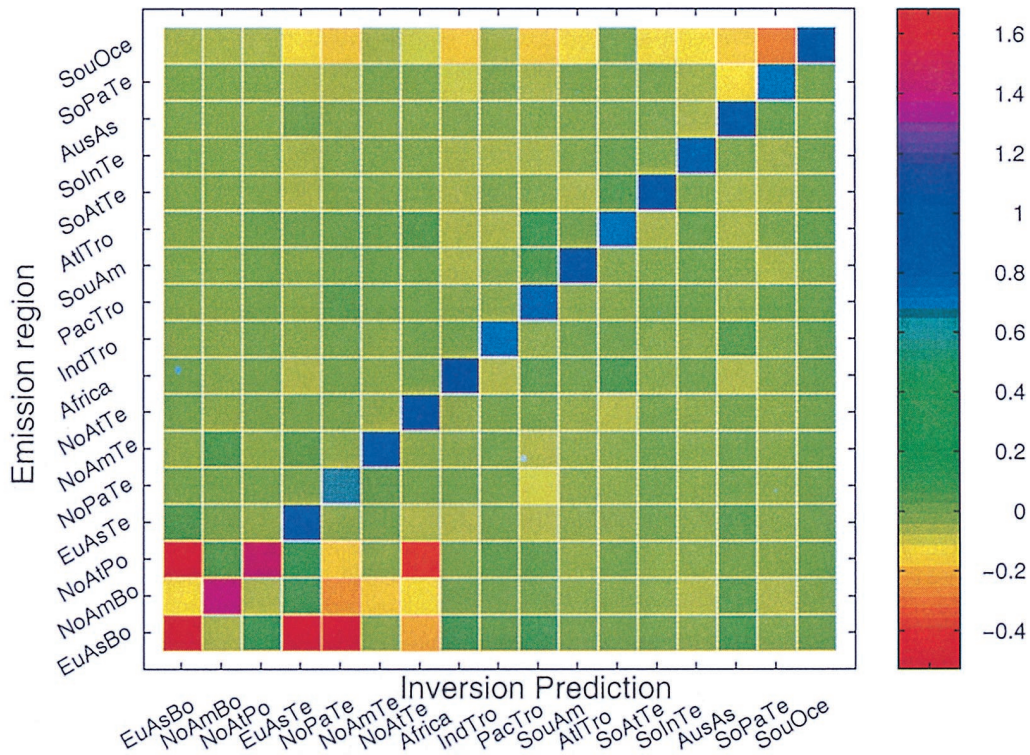
Plate 3. Average estimates of fluxes of 1 Pg C yr^{-1} region^{-1} simulated in SKYHI and estimated with the inversion scheme of GCTM and 150 observation stations. The flux patterns used for the inversion with GCTM are identical with those in SKYHI used to generate the data mixing ratio fields. Average estimates are aligned horizontally for a flux of 1 Pg C yr^{-1} for each of 17 regions listed on the y axis at a time and a flux of 0 Pg C yr^{-1} from the remaining regions.
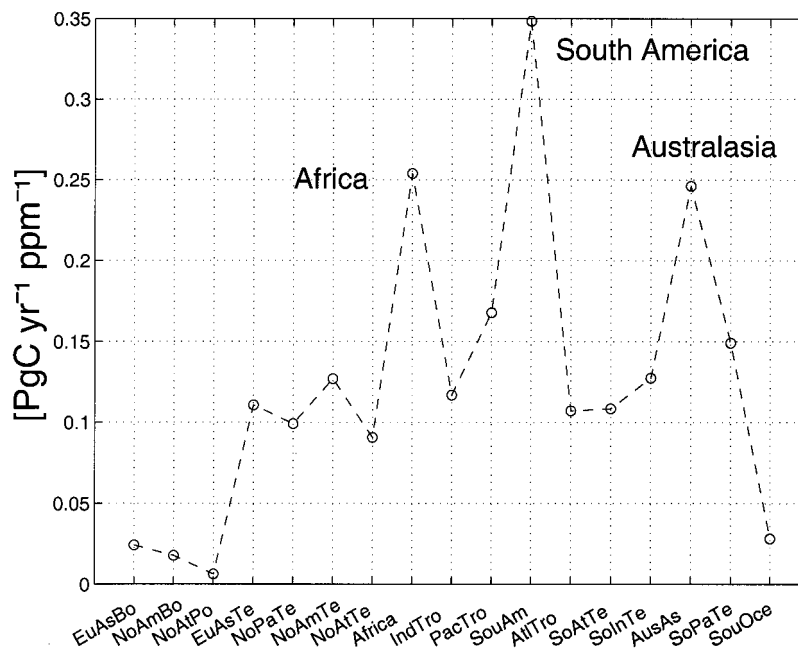
Figure 5. Average error amplification for each of the 17 source regions of Figure 1 and an observation network encompassing 500 observation stations. The determination of the average over random networks is explained in section 2.2.

data variability and hence a lower bound for the flux detection limit of the inversion is the intensity of tropospheric mixing. It also implies that the continental, equatorial regions need higher data coverage than others for equally trustworthy flux estimates.

4.3. Systematic Errors

**4.3.1. Errors caused by differences of modeled transport.** Estimates of fluxes from inversions are biased, in part, because modeled transport differs from the true transport. To get a handle on the magnitude of these errors, given the state of art of atmospheric transport modeling, we estimate SF_6 emissions from mixing ratio distributions simulated by 12 models, based on an identical flux pattern. We then analyze in greater detail the relation between errors and transport differences for the two models GCTM and SKYHI.

The simulations of SF_6 mixing ratio distributions were performed during TransCom2 [*Denning et al.*, 1999], a model transport intercomparison study. We estimated the SF_6 emissions using fossil fuel footprints calculated with GCTM for the following five regions: Eurasia boreal, Eurasia temperate, North America boreal, North America temperate, and all remaining continental regions combined. We use the simulations of emissions from fossil fuel burning with GCTM. As the observation network, we use GLOBALVIEW-CO2 [*NOAA*, 1997] (also available on Internet via anonymous FTP to ftp. cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW). the largest self-consistent data set of CO_2 observations currently available. The estimates of SF_6 emissions derived from each of the 12 TransCom2 model SF_6 results are shown in Plate 2. Note that differences between flux estimates and the true emissions are of the same order of magnitude as the emissions themselves. The reason for this is mainly attributable to different transport properties of the models because 66 stations (the number of observation stations of GLOBALVIEW-CO2) are enough to reduce the error from amplification of high-frequency data variability to an insignificant level (cf. section 4.1).

In the case of the GCTM SF_6 overestimate of the flux from Eurasia boreal, the reason cannot be transport because the SF_6 mixing ratio distribution and the fossil fuel footprints are both simulated with GCTM. In this case the error is caused by the difference of the SF_6 emission pattern, based on electrical energy consumption and population density [cf. *Denning et al.*, 1999] and the CO_2 emission pattern, based on fossil fuel energy consumption, cement manufacture, and population density [*Andres et al.*, 1996]. Differences are particularly large at the Westerland monitoring station (8°E, 55°N) next to the North Sea, where the fossil fuel emission pattern predicts comparably smaller values than the SF_6 emission pattern, which in turn, results in an overestimation of the Eurasian boreal flux.

We next use SKYHI to generate pseudo-observations for constant fluxes from each of 17 regions, and the footprints simulated with GCTM, based on exactly the same spatially uniform flux patterns, for the inversion to estimate the emissions in SKYHI. This procedure isolates the effect of model transport differences completely from all other sources of error. In addition, to get error estimates independently from a specific observation network, we follow again the Monte Carlo approach described in section 2.2. We obtain annual-average estimates for the flux from each of the 17 regions from the inversion of each of the 17 mixing ratio distributions simulated by SKYHI for a spatially-uniformly-distributed flux of 1 Pg C yr^{-1}. For a given emission region these average estimates are aligned horizontally in Plate 3. For example, the estimate of emissions in SKYHI from the Atlantic tropical region (line six) attributes a flux of ~0.7 Pg C yr^{-1} to Atlantic tropics, ~0.15 Pg C yr^{-1} to Pacific tropics, and approximately ±0.05 Pg C yr^{-1} to all the remaining regions. If there would be no transport error, fluxes from the Atlantic tropics would be estimated to be 1 Pg C yr^{-1}, and 0 Pg C yr^{-1} would be estimated for all other regions. More generally, if there would be no difference in the transport properties between the two models, all diagonal el-
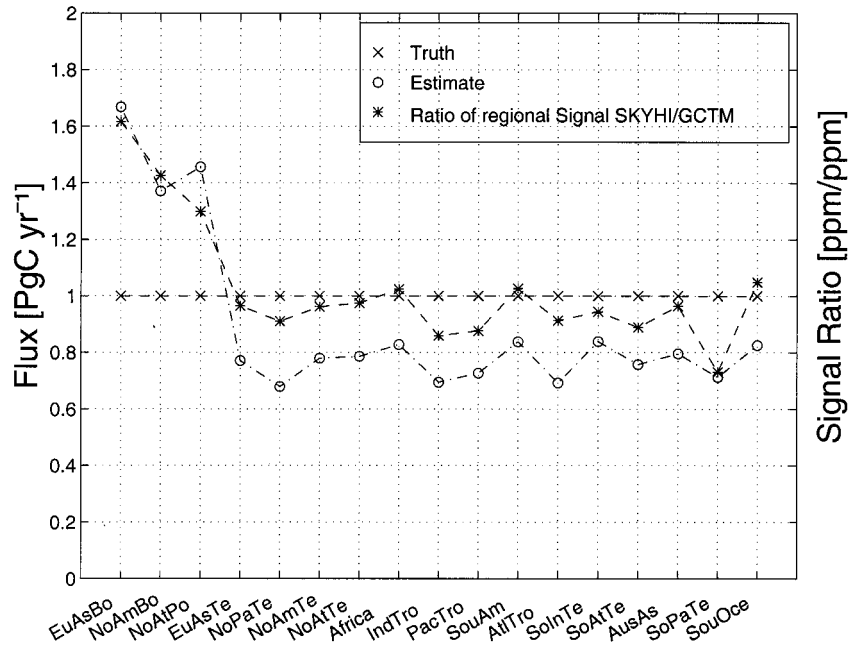
Figure 6. Mean estimates of SKYHI fluxes from 17 regions with the GCTM inversion scheme (diagonal elements of the matrix in Plate 2) and ratio of local signals of SKYHI and GCTM (cf. section 4.3.1).

ements in Plate 3 would be 1, whereas all the off-diagonal elements would be zero. Most flux regions hence are correctly localized by the inversion scheme, with the exception of the three northernmost and the southernmost ones.

A comparison of the ratio of the integrated "regional signals" (the signals within the flux regions) as simulated by SKYHI and GCTM ($\int_{\text{region } r} \Delta\chi_{\text{SKYHI}}(\mathbf{x}) \, d\sigma / \int_{\text{region } r} \Delta\chi_{\text{GCTM}}(\mathbf{x}) \, d\sigma$) for the case above (Figure 6) reveals the following.

1. As long as the ratio of regional signals due to a regional flux is "small" (of the order of 10–20%), transport errors affect only the estimate of the flux from the flux region itself.

2. If differences of regional signals are of the order of 50%, as for the three northernmost regions, fluxes are misattributed to several regions. It is noteworthy, however, that the misattribution is only to other regions within the same zonal band as the flux region itself.

3. Differences of estimates for the remaining regions may partially be explained by the ratio of the "regional signals" between GCTM and SKYHI (Figure 6) (~40% of the discrepancy) which are a result of stronger PBL ventilation in SKYHI and correspondingly smaller regional signals. The remaining discrepancy is attributable to different interhemispheric exchange rates: SKYHI's exchange is more rapid by ~10%. Fluxes from the regions besides the northernmost ones are underestimated by approximately the same percentage. This result suggests that too rapid interhemispheric exchange will cause underestimates of the magnitude of fluxes for all regions and that the estimates scale inversely with interhemispheric exchange rates.

4.3.2. Errors caused by the neglect of interannual variability of atmospheric transport. We estimate the magnitude of biases and standard deviations of estimates caused by interannual variability by using footprints from one atmospheric transport model (SKYHI) for the estimation of fluxes from the atmospheric patterns simulated by two different model years of

another model (TM2). As footprints for the inversions, we use spatially uniformly distributed fluxes from 17 regions simulated by SKYHI; as pseudo-observations, we use the atmospheric patterns from fossil fuel emissions and biospheric exchange fluxes simulated by TM2. Note that TM2 is driven by analyzed winds, which differ from year to year while the fluxes are kept fixed from year to year. Differences of estimates from different years are not sensitive to the spatial patterns used for the inversions (here spatially uniform) because the estimates from the 2 years are affected in the same way by the specific choice of the spatial flux pattern, and their effect hence cancels out. As above, we base the inversions on randomly generated observation networks with 150 observation stations. The biases in this case are the difference between the mean estimates of the 2 years (as explained in section 2.2) and are tabulated in Table 2, together with the standard deviations of the estimates.

For biospheric fluxes, biases caused by interannual variability are of considerable magnitude (~0.2–0.3 Pg C yr^{-1} region^{-1}) for the regions where biospheric exchange fluxes are large: Eurasia temperate, North America temperate, and Africa (Table 2). Otherwise, biases are small ($\leq$0.05 Pg C yr^{-1} region^{-1}). The standard deviations of the ensemble of estimates, for a network with 150 stations, though, may be as large as 0.5 Pg C yr^{-1} region^{-1}. For fossil fuel emissions, biases are smaller than for the land biosphere; maximally ~0.2 Pg C yr^{-1} region^{-1} (Table 2). Again, the standard deviations of the estimates are somewhat greater, maximally 0.25 Pg C yr^{-1} region^{-1}, and their mean is 0.1 Pg C yr^{-1} region^{-1}. A possible way to reduce these biases and standard deviations might be the use of a model with assimilated winds combined with a time-dependent inversion scheme.

4.3.3. Errors caused by incorrect modeling of the spatial flux pattern. To quantify the magnitude of systematic errors introduced by spatial differences between the flux patterns $\varphi_r(\vartheta, \varphi, t)$ used for the inversion and the real ones, we use the inversion scheme based on fluxes from 17 regions and, as

Table 2. Systematic Errors (Biases and Standard Deviations (s.d.)) Caused by Interannual Variability of Model Transport and by Mismatch of Flux Patterns in Space and Time (Determined With a Monte Carlo Approach (Section 2.2)) Between Footprints and Pseudodata

	Pseudodata							
	Biosphere (TM2)		Fossil Fuel (TM2)		Focus USA (GCTM)		Rectification (SKYHI)	
Region	Bias	s.d.	Bias	s.d.	Bias	s.d.	Bias	s.d.
Eurasia boreal	0.04	0.06	0.04	0.0	0.00	0.06	0.45*	0.46*
N. America boreal	0.02	0.04	0.07	0.03	0.01	0.17*	0.05	0.08
N. Atlantic polar	−0.06	0.06	0.03	0.03	0.00	0.04	−0.01	0.04
Eurasia temperate	0.09	0.28*	0.09	0.07	0.00	0.18*	−0.12*	0.28*
N. Pacific temperate	−0.02	0.15*	0.08	0.06	0.00	0.24*	−0.18*	0.25*
N. America temperate	0.22*	0.27*	0.11*	0.12*	0.27*	1.17*	0.41*	0.46*
N. Atlantic temperate	−0.03	0.12*	0.07	0.16*	−0.01	0.49*	−0.06	0.13
Africa	−0.38*	0.51*	−0.01	0.15*	0.00	0.29	−0.37*	0.48*
Indian Ocean tropical	−0.07	0.51*	−0.03	0.25*	−0.01	0.10*	−0.15*	0.19*
Pacific tropical	0.01	0.14*	−0.10*	0.16*	0.00	0.14*	0.21*	0.25*
S. America	−0.05	0.31*	−0.23*	0.14*	0.01	0.11*	0.08	0.28*
Atlantic tropical	0.03	0.34*	0.02	0.14*	0.00	0.19*	0.26*	0.36*
S. Atlantic temperate	−0.01	0.07	−0.01	0.06	0.00	0.03	−0.07	0.1*
S. Indian Ocean temperate	0.03	0.06	−0.02	0.03	0.00	0.02	−0.04	0.07
Australasia	−0.03	0.21*	−0.14*	0.13*	0.00	0.02	−0.09	0.2*
S. Pacific temperate	0.08	0.09	0.06	0.07	0.00	0.05	−0.09	0.12*
Southern Ocean	0.10*	0.01	0.00	0.01	0.00	0.01	0.02	0.04
Mean	0.1	0.2	0.09	0.1	0.02	0.20	0.21	0.27

Values are in units of Pg C yr^{-1}. Regional fluxes used to generate the footprints for the inversion scheme are based on spatially-uniformly-distributed source strength and constant flux in time. Observation networks consist of 150 observation stations. As pseudo-observations to test for the effect of spatial mismatch, mixing ratio distributions resulting from fluxes within the United States are used. Similarly, to test for the influence of the temporal structure, the CASA biosphere is used.

*Values exceeding 0.1 Pg C yr^{-1}.

pseudo-observation, the mixing ratio distributions resulting from a flux from a subregion within North America temperate (Figure 3). The flux pattern underlying this pseudo-observation deviates strongly from the spatially uniform patterns used to generate the footprints for the inversion scheme. The difference between the mixing ratio distributions resulting from a flux from the entire North America temperate and a flux of the same magnitude emitted from the subregion in southeast North America displayed in black in Figure 3 is shown in Figure 7. The mixing ratio distributions differ strongly from each other within the source region with differences of the order of 5 ppm Pg C^{-1} yr, while the differences in the "far-field" are very small ($\leq$0.05 ppm).

The average flux estimate based on an ensemble of randomly generated networks is summarized in Table 2 (columns headed Focus USA): the estimate for the flux region is almost unbiased, but its standard deviation is very large (of the same order as the flux itself). The bias and standard deviation for the estimates for the other regions on the other hand are small ($\leq$0.2 Pg C yr^{-1} region^{-1}). This is clearly a reflection of the spatial distribution of the difference between the signals as described above (large differences within the flux region and small differences elsewhere). Randomly positioned stations within the flux region are likely to observe a signal that will be strongly misinterpreted by the inversion; since the far field is unaffected by the specific spatial structure of the fluxes, though, fluxes from other regions are correctly estimated (to be almost zero). The region for which the flux is the most affected, besides North America temperate, is North Atlantic temperate, its next region downstream.

On average, for networks encompassing 500 and 2000 stations the standard deviation of the estimate for the flux region

is still 0.64 and 0.4 Pg C yr^{-1} region^{-1}, respectively. Enlargement of the network does not help. One possible way to reduce this source of error would be to increase the number of regions, which however, is cost intensive with respect to the number of necessary observation stations ($\sim$10 stations region^{-1}), as already discussed.

The expected bias introduced in the estimate by spatial mismatch could be estimated roughly directly from the local differences of the mixing ratio signal (Figure 7) and the mean error amplification as a function of the network size (Plate 1): differences in local signals of the order of 1–2 ppm propagate to an error in estimates of $\sim$0.5–1 Pg C yr^{-1} region^{-1} for a network consisting of 150 stations.

4.3.4. Errors caused by improper modeling of the temporal flux pattern. The zonal average of the rectification of CO_2 biospheric exchange fluxes is of the order of 2 ppm according to the model intercomparison study reported by *Law et al.* [1996]. Since the zonally averaged signal from fossil fuel burning of $\sim$5 Pg C yr^{-1} is of the order of 4–5 ppm [*Law et al.*, 1996], the rectification signal suggests a spurious carbon source in the Northern Hemisphere of the order of 2–3 Pg C yr^{-1}. How large is this effect if we abandon the zonally averaged picture in favor of an approach that resolves sources and sinks by latitude and longitude?

To answer this question, we determine a mean estimate of the seasonal rectification of the CASA biosphere [*Potter et al.*, 1993] with atmospheric transport simulated with SKYHI. We use spatially uniformly distributed flux patterns and, again, SKYHI to simulate the footprints for the inversions. The mean estimates are displayed in Table 2. Their magnitude does not exceed 0.45 Pg C yr^{-1}, and their mean is only 0.2 Pg C yr^{-1}. It is of interest to compare the CASA/SKYHI rectification in
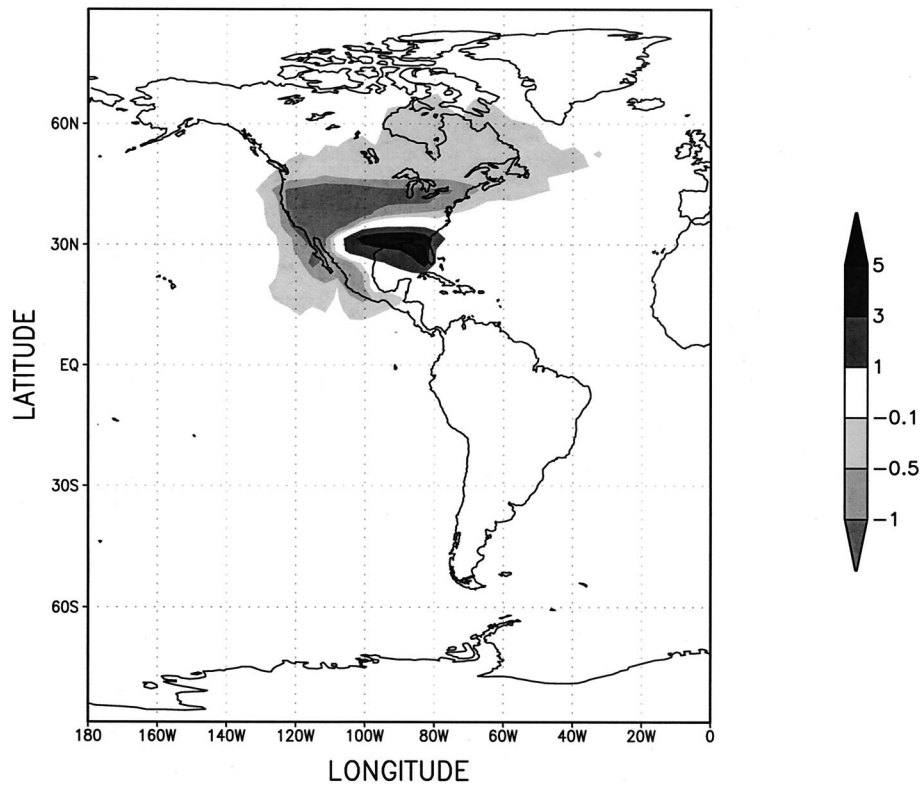
Figure 7. Mixing ratio difference (ppm) for a flux of 1 Pg C yr^{-1} from North America temperate and a southwestern subregion (dark subregion of the United States in Figure 1).

Figure 8 with the seasonal rectification of the CASA biosphere as simulated with the Colorado State University Model (CSU model) as published by *Denning et al.* [1995, Figure 1]. The effects are fairly similar both in structure and magnitude even though the magnitude simulated by the CSU model is tendentially somewhat larger. We conclude that if sources and sinks are resolved by longitude and latitude, then the seasonal rectification is significantly less dramatic than if sources and sinks
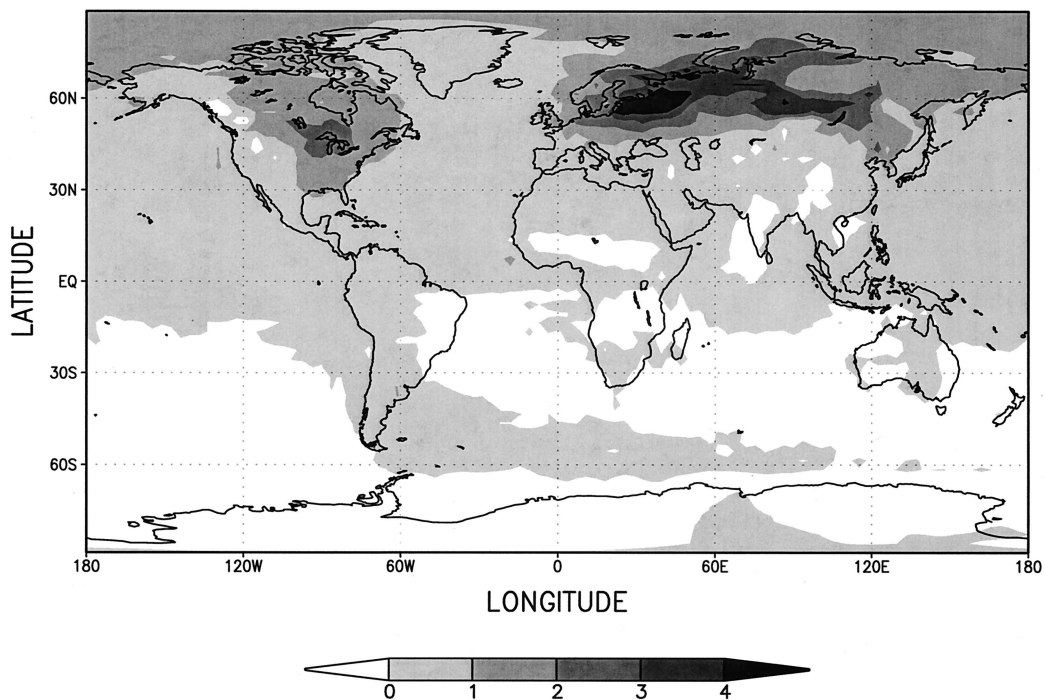
Figure 8. Seasonal rectification effect (ppm) as predicted by the CASA biosphere model and the atmospheric tracer transport model SKYHI.

are resolved by latitude only. The main reason is that the signal of rectification of the land biosphere is tied to the continents, whereas the mixing ratio distribution caused, e.g., by fossil fuel emissions is additionally large over the oceans.

If fluxes of tracers other than CO_2 are known to exhibit minor seasonality only, then rectification is of no concern for the inversion methodology based on annually averaged mixing ratios.

5. Summary and Conclusions

We have analyzed the possibilities and limitations of the inversion of atmospheric transport, using annual mean mixing ratios, as a tool to estimate CO_2-surface fluxes. The analysis is based on simulations with different transport models. Our approach avoids the choice of a particular observation network and the convolution of different error sources at a time. We distinguish two main categories of errors: (1) amplification of high-frequency data variability in an inversion and (2) systematic errors caused by inaccuracy of the inversion model.

The magnitude of the first error source is determined by the level of mixing in the atmosphere and varies with the ratio of observation stations to regions to be solved for. We find that ∼10 observation stations per region are needed in order to estimate fluxes within an uncertainty of ∼0.2 Pg C yr^{-1} (for the current magnitude of fossil fuel emissions and biospheric exchange fluxes, and for randomly positioned stations); equatorial, continental regions need comparably more data coverage, high-latitude regions comparably less. Even though we based our analysis on a particular inversion method (which uses annual means), we believe that this result applies similarly to other, e.g., recursive, inversion methods. The problem of information loss by mixing in the atmosphere is hence solvable with a feasible measurement effort, at least to estimate fluxes on a continental scale.

For the analysis of systematic errors, we distinguished between errors caused by (1) the inaccurate modeling of atmospheric transport processes, (2) the neglect of interannual variability of atmospheric transport in our annual-mean approach, and (3) the mismatch of the spatiotemporal flux pattern used for the inversion with the "true" flux pattern.

Of these sources of errors, the first is the largest. This claim is based on the inversion with GCTM of SF_6 distributions simulated by 12 different models, which all used the same flux pattern (TransCom2 [*Denning et al.*, 1999]). We estimated fluxes from five regions and based the inversion on the GLOBALVIEW-CO2 network (66 stations). Because 66 stations are enough to eliminate the errors caused by the amplification of high-frequency data variability, the difference of estimates is caused primarily by different transport properties of the models. We find that these differences are of the same magnitude as the fluxes themselves. We analyzed the relation between model transport properties and biases for the models SKYHI and GCTM in more detail and found that a major source for biases is the differences in PBL ventilation rates. Biases are up to 50% of the estimates and scale in this case inversely proportional to the ventilation rates. Another major well-known source of error is the difference among models in interhemispheric exchange rates; as expected, the magnitude of flux estimates scale inversely proportional to interhemispheric exchange rates. Collectively, these results show that misrepresentations of transport in models are currently a very serious concern for inversions. A most crucial transport pro-

cess, the ventilation of the PBL in models, needs further validation for the use of inversions of atmospheric transport as a monitoring tool.

The second largest error source for inversions based on annual means of CO_2 mixing ratios is the neglect of interannual variability of atmospheric transport. This affects primarily the estimates of regions with strong biospheric and fossil fuel fluxes and results in biases of up to 0.4 Pg C yr^{-1} region^{-1} (again for estimated fossil fuel emissions of 6 Pg C yr^{-1} for the year 1990). Also, estimate uncertainties increase considerably for these and neighboring regions (up to 0.5 and 0.2 Pg C yr^{-1} region^{-1}, respectively). This source of error is nonnegligible if inversions are to be used for emissions monitoring. The development of recursive, time-dependent inversion methods possibly based on analyzed winds might help to reduce this source of error.

We also investigated the errors caused by neglecting the covariation of temporal variability of sources with atmospheric transport by inverting the CASA biosphere. We find a smaller effect than expected from discussions in the literature [*Denning et al.*, 1995]. Regions with large biospheric exchange fluxes are most biased, with errors up to 0.45 Pg C yr^{-1} region^{-1}. This implies (1) that as long as fluxes are changing moderately in time only, then the method is robust and (2) that the seasonal rectification seems to be a less severe source of error than anticipated. Note that this result is based on a very similar rectification mixing ratio distribution as the one published by *Denning et al.* [1995].

To examine the effect of spatial heterogeneity of fluxes, we estimated spatially-highly-localized sources. These sources were correctly localized by the inversion, and almost no fluxes were wrongly attributed to other regions, but the magnitude of the estimates was highly uncertain. Highly localized sinks and sources are hence not properly estimable with this method. The only strategy that would ameliorate this problem is a subdivision of flux regions, which in turn, requires a higher coverage with observation stations.

We conclude that monitoring of surface fluxes with inversions of atmospheric transport on a continental scale is possible in principle because the damping of flux signals by mixing in the atmosphere may be largely overcome by positioning roughly 10 observation stations per region solved-for. The limiting factor then becomes primarily the use of transport models whose different transport properties lead to a scatter of estimates of the order of 100% of the fluxes and, secondarily, the neglect of interannual variability. Improvement of the calibration of transport models and possibly the development of time-dependent inversion schemes are needed for inversions to become a serious option for future monitoring.

Finally, it is of interest to put the recent inversion results by *Fan et al.* [1998] in perspective of the findings presented in this paper. The paper by *Fan et al.* [1998] restricts itself to the estimation of fluxes from Eurasia, North America, and the remaining continental regions. The findings in this paper on the latitudinal distribution of error amplification and on the number of observation stations needed to estimate fluxes, in combination with the factual bias of existing observation stations to the Northern Hemisphere, support the choice of *Fan et al.* [1998]. Similarly supported are their exclusion of the Westerland and Tae-Ahn Peninsula station from the inversion analysis. The strong dependence of estimates of SF_6 emissions on model transport, on the other hand, reaffirms the call in

Fan et al. [1998] for further corroboration with alternative transport models.

References

Andres, R. J., G. Marland, I. Fung, and E. Matthews, A 1° × 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990, *Global Biogeochem. Cycles*, *10*, 419–429, 1996.

Bakwin, P. S., P. P. Tans, C. Zhao, W. Ussler III, and E. Quesnell, Measurements of carbon dioxide on a very tall tower, *Tellus, Ser. B*, *47*, 535–549, 1995.

Barnola, J. M., D. Raynaud, Y. S. Korotkevich, and C. Lorius, Vostok ice core provides 160,000 year record of atmospheric CO_2, *Nature*, *329*, 408–414, 1987.

Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang, Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, *99*, 22831–22855, 1994.

Denning, A. S., I. Y. Fung, and D. Randall, Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota, *Nature*, *376*, 240–243, 1995.

Denning, A. S., et al., Three-dimensional transport and concentration of SF_6: A model intercomparison study (TransCom2), *Tellus*, in press, 1999.

Enting, I. G., and J. V. Mansbridge, Seasonal sources and sinks of atmospheric CO_2: Direct inversion of filtered data, *Tellus, Ser. B*, *41*, 111–126, 1989.

Enting, I. G., C. M. Trudinger, R. J. Francey, and H. Granek, Synthesis inversion of atmospheric CO_2 using the GISS tracer transport model, *Div. Atmos. Res. Technol. Pap. 29*, Commonwealth Sci. and Ind. Res. Organ., Melbourne, Australia, 1993.

Fan, S., M. Gloor, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans, A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models, *Science*, *282*, 442–446, 1998.

Fels, S. B., J. D. Mahlman, M. D. Schwarzkopf, and R. W. Sinclair, Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamical response, *J. Atmos. Sci.*, *37*, 2265–2297, 1980.

Golub, G. H., and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins Univ. Press, Baltimore, Md., 1989.

Hamilton, K. P., J. Wilson, J. D. Mahlman, and L. M. Umscheid, Climatology of the SKYHI troposphere-stratosphere-mesosphere model, *J. Atmos. Sci.*, *48*, 651–678, 1995.

Heimann, M., The Global Atmospheric Tracer Model TM2, *Tech. Rep. 100*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1996.

Holloway, J. L., and S. Manabe, Simulation of climate by a global general circulation model, I, Hydrological cycle and heat balance, *Mon. Weather Rev.*, *99*, 335–370, 1971.

Keeling, C. D., S. C. Piper, and M. Heimann, A three-dimensional model of atmospheric CO_2 transport based on observed winds, 4, Mean annual gradients and interannual variations, in *Aspects of Climate Variability in the Pacific and the Western Americas*, *Geophys. Monogr. Ser.*, vol. 55, edited by D. H. Peterson, pp. 305–363, AGU, Washington, D. C., 1989.

Kirkpatrick, S., C. D. Gelatt Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science*, *220*, 671–680, 1983.

Law, R. M., et al., Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions, *Global Biogeochem. Cycles*, *10*, 783–796, 1996.

Levy, H., II, and W. J. Moxim, Simulated global distribution and deposition of reactive nitrogen emitted by fossil fuel combustion, *Tellus, Ser. B*, *41*, 256–271, 1989.

Levy, H., II, J. D. Mahlman, and W. J. Moxim, Tropospheric N_2O variability, *J. Geophys. Res.*, *87*, 3061–3080, 1982.

Liu, S. C., J. R. McAfee, and R. J. Cicerone, Radon 222 and tropospheric vertical transport, *J. Geophys. Res.*, *89*, 7291–7297, 1984.

Mahlman, J. D., and W. J. Moxim, Tracer simulation using a Global General Circulation Model: Results from a Midlatitude Instantaneous Source Experiment, *J. Atmos. Sci.*, *35*, 1340–1374, 1978.

Masarie, K. A., and P. Tans, Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *J. Geophys. Res.*, *100*, 11593–11610, 1995.

Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, San Diego, Calif., 1989.

National Oceanic and Atmospheric Administration (NOAA), *GLOBALVIEW-CO2 Cooperative Atmospheric Data Integration Project—Carbon Dioxide* [CD-ROM], NOAA/Clim. Monit. and Diagnostics Lab., Boulder, Colo., 1997.

Neftel, A., E. Moor, H. Oeschger, and B. Stauffer, Evidence from polar ice cores for the increase in atmospheric CO_2 in the past two centuries, *Nature*, *315*, 45–47, 1985.

Potter, C. S., J. T. Randerson, C. B. Field, P. A. Mason, P. M. Vitousek, H. A. Mooney, and S. A. Klooster, Terrestrial ecosystem production: A process model based on global satellite and surface data, *Global Biogeochem. Cycles*, *7*, 811–841, 1993.

Russell, G., and J. Lerner, A new finite-differencing scheme for the tracer transport equation, *J. Appl. Meteorol.*, *20*, 1483–1498, 1981.

Sarmiento, J. L., R. Murnane, and C. LeQuere, Air-sea CO_2 transfer and the carbon budget of the North Atlantic, *Philos. Trans. R. Soc. London, Ser. B*, *348*, 211–219, 1995.

Solomon, S., and J. Srinivasan, Radiative forcing of climate change, *IPCC Report 1995*, pp. 109–118, Cambridge Univ. Press, New York, 1995.

Taguchi, S., A three-dimensional model of atmospheric CO_2 transport based on analyzed winds: Model description and simulation results for TransCom, *J. Geophys. Res.*, *102*, 15099–15109, 1996.

Takahashi, T., R. A. Feely, R. Weiss, R. H. Wanninkhof, D. W. Chipman, S. C. Sutherland, and T. Takahashi, Global air-sea flux of CO_2: An estimate based on measurements of sea-air pCO_2 difference, *Proc. Natl. Acad. Sci. USA*, *94*, 8292–8299, 1997.

Tans, P. P., A note on isotopic ratios and the global atmospheric methane budget, *Global Biogeochem. Cycles*, *11*, 77–83, 1997.

Tans, P. P., T. J. Conway, and T. Nakazawa, 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.

Tans, P. P., I. Y. Fung, and T. Takahashi, Observational constraints on the global atmospheric CO_2 budget, *Science*, *247*, 1431–1438, 1990.

Tikhonov, A. N., and V. Y. Arsenin, *Solution of Ill-Posed Problems*, John Wiley, New York, 1977.

S.-M. Fan, M. Gloor, S. Pacala, M. Ramonet, and J. Sarmiento, Department of Ecology and Evolutionary Biology, Sayre Hall, Forrestal Campus, Princeton, NJ 08544.