Variations in modelled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions

R. M. Law[1], P. J. Rayner[1], A. S. Denning[2], D. Erickson[3], M. Heimann[4], S. C. Piper[5], M. Ramonet[6], S. Taguchi[7], J. A. Taylor[8], C. M. Trudinger[9] and I. G. Watterson[9]

Received RECEIPT DATE; revised REVISION DATE; accepted ACCEPT DATE.

Short title: TRACER TRANSPORT MODEL COMPARISON

[1]Cooperative Research Centre for Southern Hemisphere Meteorology, Monash University, Clayton, Australia.

[2]Department of Atmospheric Science, Colorado State University, Fort Collins, USA.

[3]National Center for Atmospheric Research, Boulder, USA.

[4]Max Planck Institut für Meteorologie, Hamburg, Germany.

[5]Scripps Institution of Oceanography, La Jolla, USA.

[6]Centre des Faibles Radioactivites, Laboratoire Mixte CNRS/CEA, Gif-sur-Yvette Cedex, France.

[7]National Institute for Resources and Environment, Onogawa Tsukuba Ibaraki, Japan.

[8]Centre for Resource and Environmental Studies, Australian National University, Canberra, Australia.

[9]CSIRO Division of Atmospheric Research, Aspendale, Australia.

Abstract. Carbon dioxide concentrations due to fossil fuel burning and CO_2 exchange with the terrestrial biosphere have been modelled with twelve different three-dimensional atmospheric transport models. The models include both online and offline types and use a variety of advection algorithms and sub-grid scale parameterisations. A range of model resolutions is also represented.

The modelled distributions show a large range of responses. For the experiment using the fossil fuel source, the annual mean meridional gradient at the surface varies by a factor of two. This suggests a factor of two variation in the efficiency of interhemispheric exchange. In the upper troposphere, zonal mean gradients within the northern hemisphere vary in sign.

In the terrestrial biotic source experiment, the structure of the amplitude of the seasonal cycle of CO_2 concentration at the surface is largely conditioned by the position of the sources. The amplitudes, however, vary similarly to the fossil case. The annual mean response to the seasonal source also shows large differences in magnitude.

We discuss the implications of these results for carbon budget studies and suggest some methods for reducing the apparent large uncertainties in transport.

Introduction

This paper presents a summary of some of the major results from the CO_2 Transport Comparison Project, TransCom. The project was initiated at the 4th International CO_2 Conference in Carqueiranne in 1993. Its initial aim was to provide a qualitative understanding of the importance of differences in atmospheric tracer transport models used in CO_2 budget studies. The need for the project arose from the considerable variation in global budgets arising from such studies as *Keeling et al.* [1989], *Tans et al.* [1990] and *Enting et al.* [1995]. The uncertainties represented by the range of proposed budgets propagate into the future as uncertainties in the evolution of CO_2 concentration [*Enting et al.* 1994].

To understand how differences in transport characteristics of atmospheric tracer models appear as differences in CO_2 budgets, it is necessary to review how such models are used. The input data is a record of atmospheric trace gas concentrations (mainly CO_2) at a range of sites around the world. The record commences with the measurements of C.D. Keeling in the late 1950s at South Pole and Mauna Loa [*Keeling et al.*, 1995]. Spatial coverage increased gradually through the 1960s and 1970s, then more rapidly through the 1980s and early 1990s. The spatial distribution of concentration contains information about the spatial structure of sources and sinks of CO_2 which is accessed using a model of tracer transport. Atmospheric transport determines the relationship between the space-time structure of net sources and the space-time structure of concentration.

In the language of inverse theory, the current concentrations are related to past sources by a linear operator representing transport. 'Concentrations' may include multiple species as in the work of *Enting et al.* [1995]. The problem of estimating CO_2 sources, then, is one of inverting the linear operator representing transport. Such an estimate should yield not only an expectation for the source but also an error estimate e.g. *Enting et al.* [1993, 1995] and *Ciais et al.* [1995]. These authors considered the impact on uncertainties in source components due to uncertainties in the input data. They did not consider the impact of errors in the transport operator. As yet there is no computationally feasible way of characterizing errors in

three-dimensional atmospheric tracer transport models, nor data with which these errors can be established. In the absence of such data, model intercomparison provides one estimate of the uncertainty. In addition, the emergence of consensus on any aspect of transport should reinforce our confidence in results which depend on it.

While it is difficult to characterise model error, there are methods by which models can be validated. Various trace gases can be used to calibrate modelled transport, typical choices being krypton-85, CFCs and radon [e.g. *Prather et al.,* 1987, *Jacob et al.,* 1987, *Heimann and Keeling,* 1989, *Jacob and Prather,* 1990]. However, the limited data available and uncertainties in the sources means that no calibration is comprehensive. It is hoped that a combination of consensus and calibration could reduce the uncertainties in source estimation due to model error.

To provide a quantitative measure of differences in model transport, twelve atmospheric tracer transport models have been run with prescribed sources. The models either have been or are likely to be used in CO_2 source studies and the chosen sources are relevant to the seasonal and annual mean spatial distribution of CO_2. This is by no means a complete characterization of model differences, let alone model error, but it is a first step in this direction. It should also aid in interpreting any inversion studies by particular models.

Method

Sources

Two sources of CO_2 were chosen for this comparison. The first was the emissions of CO_2 due to fossil fuel burning and cement production. This is one of the best known components of the CO_2 budget and makes a good test of a model's interhemispheric transport since 95% of the fossil fuel emissions occur in the northern hemisphere. The source data used were provided by Fung and have been previously used by *Tans et al.* [1990]. They are based on country estimates derived by *Marland et al.* [1989] which have been distributed within

countries according to population density by Fung. They include no temporal variation. The data were provided on a 1° grid (Fig. 1) with modellers aggregating this to their own model resolution. The difference in resolution among the models introduces small-scale differences into the input sources, for example, the maximum source strength for a grid point in the high resolution GFDL model is 912 $gCm^{-2}yr^{-1}$ compared to only 515 $gCm^{-2}yr^{-1}$ in the lower resolution GISS model.

The second source used was the exchange of CO_2 with the biosphere. The data were compiled by *Fung et al.* [1987], combining satellite estimates of photosynthesis with local measurements of respiration and net primary productivity. The sources were validated by comparing modelled (using the GISS model) and observed seasonal cycles of CO_2 concentration. This source is the major contributor to the observed seasonal cycle of CO_2, at least in the northern extra-tropics. Thus some comparisons can be made between modelled and observed seasonal cycles. There is also considerable interest in the annual mean CO_2 field which results from the combination of seasonal sources and seasonal variation in transport. One recent study, *Denning et al.* [1995], has demonstrated strong correlations between a modelled planetary boundary layer and sources of CO_2 leading to substantial annual mean meridional gradients.

The time-latitude distribution of the vegetation source is shown in Fig. 2. The largest seasonality occurs in the northern mid-high latitudes. There are two periods when respiration dominates, indicated by large positive fluxes (May and October). This is balanced by a larger and shorter-lived negative spike during the peak growing season of summer. At other latitudes the fluxes are smaller and more sinusoidal in nature.

It was recommended that the experiments, referred to here as the fossil and biosphere experiments, be run for at least three years from an initial atmosphere with uniform CO_2. This provides sufficient time for the model atmosphere to establish 'equilibrium' (annually repeating) concentration distributions determined by the surface sources.

Models

The experiments were run with twelve different tracer transport models. An identifier is allocated to each model in Table 1a and the major characteristics of each model are summarised in Table 1b. The models included full General Circulation Models (GCMs) which transport trace gases along with other meteorological variables (so-called online models) and offline models, in which the tracer transport is effected by pre-existing wind fields. These wind fields can be either model generated or from observed analyses and are applied to the offline models at various frequencies. The models encompassed a range of horizontal and vertical resolutions and a range of advection schemes. Some models resolve the diurnal cycle. Each model also includes parameterisations of sub-grid scale tracer mixing. More detailed descriptions of the models used are given in *Rayner and Law* [1995] Appendix C.

Results

Contributing modellers supplied concentration fields for the surface layer, 500 and 200 hPa. In addition, zonal mean cross-sections were also analysed. Each dataset has been normalised such that the January global three-dimensional mean is zero. This accounts for differences in the number of years for which experiments were performed and to a large extent, any slight differences in net source. We present here only the key results arising out of the submitted data. *Rayner and Law* [1995] give a more extensive presentation of the results.

Fossil experiment

Differences in interhemispheric transport between models can be seen in the zonal annual mean surface concentrations which are shown in Fig. 3. While each model gives a broadly similar distribution, with maximum concentrations around 50°N and relatively small gradients through the southern hemisphere, there are large differences in the maximum and minimum concentrations. The range of concentrations found for the northern mid-latitudes

can be reduced by almost half if the CSIRO9 and GFDL results are excluded. It is likely that this smaller range is more realistic since there have been reported calibrations of meridional mixing using krypton-85 for the GISS and TM1 models (*Jacob et al.* [1987], *Heimann and Keeling* [1989]) which lie in this group.

The variation among models can be summarised by the interhemispheric concentration difference (northern minus southern hemisphere mean concentration). This is listed alongside each model identifier in the figure key. The interhemispheric differences vary by a factor of two. It is important to note that these differences are surface values and the variation between models reflects differences in both vertical and cross-equatorial transport. The CSIRO9 model produces the largest difference (4.7 ppmv) and the MU models the smallest (2.4 ppmv). In order to understand this difference better, MUTM (an offline model) was run with winds taken from the CSIRO9 model. This simulation produced an interhemispheric difference of 3.6 ppmv which indicates that, in this case, the large-scale winds account for about half the difference between the model results with the sub-grid scale parameterisations accounting for the other half. We discuss later the exchange times implied by these concentration differences and compare them to vertically-integrated exchange times and exchange times calculated from 'clean-air' rather than zonal mean data.

An example of the surface distribution of CO_2 is given for the TM2 model in Fig. 4. Every model produces high concentrations around Europe, North America and China associated with regions of high fossil emissions. The major differences among the models occur in the maximum values of these high concentration regions. For example, the GFDL model gives a European maximum of 25.2 ppmv compared to only 6.3 ppmv in the GISS model. Away from the source regions, the variation between models is smaller; concentrations at the locations where CO_2 is monitored generally range over about 1 ppmv (excluding CSIRO9 which lies about 1 ppmv higher than any other model in the northern mid and high latitudes). In general, the higher maximum values are produced by the higher resolution models, reflecting their ability to resolve better both the peak values in the source distributions as well as smaller

scale features in the modelled concentrations. Another factor is the strength of vertical mixing. CSIRO9 appears to have weak vertical mixing throughout the troposphere while GFDL has weak mixing out of its surface layer. In both models, this results in higher surface concentrations.

Vertical gradients in the lower troposphere can be characterised by the surface to 500 hPa concentration difference. The zonal mean of this difference is plotted in Fig. 5. There is good qualitative agreement between models but a large range in the magnitude of the vertical gradient. All models, except ANU, produce a region of increasing concentration with height in the southern low latitudes. The increase with height is largest around 10-20°S and is associated with high concentration air from the northern hemisphere which is transported southwards in the upper troposphere. The largest lower tropospheric gradients occur around 40-60°N with a model range of 0.7-3.0 ppmv. The larger values tend to occur for those models with higher surface concentrations. The zonal mean 500 hPa concentrations are relatively similar between models (not shown). The zonal mean concentration increases from south to north with a region of more rapid increase through the mid and low latitudes of both hemispheres. Longitudinal variations in concentration at 500 hPa are also small compared to the surface variations.

There is more variety in the model responses at 200 hPa (Fig. 6). Approximately half the models produce maximum concentrations around 0–30°N while the remainder have mid to high northern latitude maxima. *Nakazawa et al.* [1991] measured CO_2 concentration in the upper troposphere on flights between Tokyo and Sydney (36°N to 30°S) and found maximum annual mean concentrations around 0–10°N. This would be more consistent with those models that produce low latitude maxima at 200 hPa. However, it is important to note that the observed values are for CO_2 from all sources whereas the modelled results are for the fossil source only. Also, the model data at 200 hPa may include stratospheric air whereas this has been excluded from the observed data. The ANU model produces a more uniform distribution than the other models. This suggests that there is rapid horizontal mixing acting to reduce the meridional gradient. Weak vertical mixing could also contribute but this is less likely because

the ANU 200 hPa global mean concentration is similar to those from other models.

The global distribution of CO_2 at 200 hPa is illustrated for the MUGCM case (Fig. 7). This is one of the models that gave a low latitude maximum in the zonal mean. The figure shows maximum concentrations in the West Pacific-South East Asian region and around Central America. This is typical of all the models that produced low latitude maxima and may be associated with rapid mixing of lower tropospheric air (with high concentrations) to 200 hPa by convection. The models that produced maximum zonal mean concentrations at higher latitudes tended to produce a more zonally uniform global distribution.

Biosphere experiment

The seasonal nature of the biospheric source provides many options for characterising the models' responses. We choose here to focus on the amplitude of the seasonal cycle and the surface annual mean response.

The peak to peak (ptp) amplitudes are calculated as the difference between the maximum and minimum monthly mean concentration at each grid point. The zonal mean ptp amplitude at the surface (Fig. 8) increases from around 1–2 ppmv in the southern mid and high latitudes to between 22 and 52 ppmv (32 without CSIRO9 or GFDL) around 65° N. The larger values produced by CSIRO9 and GFDL are probably due to slow mixing out of the surface layer. This would be consistent with the larger meridional gradients produced by these models in the fossil experiment. As in the fossil experiment, the results from the CSIRO9 and MU models are very different. We have also performed the biosphere experiment forcing MUTM with the CSIRO9 winds. The ptp amplitudes that result are almost identical to those produced by MUTM (forced with MUGCM winds), particularly in the northern mid and high latitudes. This suggests that the sub-grid scale parameterisations control the amplitude of the seasonal cycles at these latitudes.

Most models produce surface amplitude distributions that are broadly similar. The distribution produced by the CSU model is shown in Fig. 9. Land regions have higher

amplitudes than ocean regions at the same latitude (except North Africa). The maximum amplitudes occur in regions of large seasonality in the sources (Northern Eurasia, Alaska, China, India, tropical Africa and south America). There is some disagreement between models on the relative magnitudes of these maxima but the Eurasian amplitudes are usually the largest. These Eurasian maxima range from 30 ppmv in the GISS model to 112 ppmv in the GFDL model. The larger amplitudes tend to occur in the higher resolution models but, in contrast to the fossil experiment, vertical resolution and associated vertical sub-grid scale processes appear to be more important than horizontal resolution. This is not surprising since the biosphere sources were given on a $4\times5°$grid which should be well resolved by all but the lowest resolution models. The importance of the vertical resolution and processes is supported by an experiment by Taguchi (pers.comm.) with the NIRE model. Vertical diffusion strength was increased by increasing the depth of the model's planetary boundary layer by 50 hPa. The resulting maximum ptp amplitude over China was 43 ppmv compared to 94 ppmv in the original simulation.

While there are difficulties associated with comparing modelled and observed amplitudes, such comparisons can assist in model evaluation, especially in the northern extra-tropics where the biosphere is the major contributor to seasonality. We focus on this region in a comparison of modelled and observed amplitudes at monitoring sites. As each model uses a different grid, the four nearest grid points to a monitoring site are used to interpolate to the actual location. This method does not take any account of the practice at continental coastal sites of measuring CO_2 from marine air, so comparison with the observations at these sites must be done with caution. We compare (Fig. 10) modelled ptp amplitude for 15 NOAA/CMDL monitoring sites, listed in Table 2, with observed amplitudes given by *Conway et al*. [1994a].

The model results range over 4 to 15 ppmv depending on the latitude but this is smaller than the range seen in the zonal means. With the exception of TM2 and TM2Z, all models overestimate the amplitude (taken here to mean they lie above the upper observation value) at one or more of the four northernmost sites. Some sites are close to land and this may have

an impact since we are comparing model data for all times with observed data selected for 'background' conditions. A data selection test using MUTM data indicated that the amplitude at Barrow could be reduced by 2 ppmv but that amplitudes at Mould Bay and Alert did not change significantly. This suggests that there may be some problems with the input source in this region as it seems unlikely that transport errors would consistently overestimate the amplitudes across so many models.

The models perform reasonably well for Cold Bay; with the exception of the CSIRO9 and TM2Z models, all the amplitudes are within about 1 ppmv of the observed range. At Shemya most models underestimate the amplitude while at Cape Meares the modelled amplitude is generally greater than observed. This site is one at which data selection could be expected to reduce the modelled amplitude. At the mid and low latitude sites the models normally span the observations. The NIRE model is typically near the top of the range of model results (possibly because it does not include tracer transport due to convection) while the ANU and CSU models are at the low end of the range.

The largest range of amplitudes in Fig. 10 occurred at Barrow. Fig. 11 shows the surface monthly mean concentration for each model at this site. Also shown is the observed seasonal cycle represented by the first two harmonics fitted to detrended data from *Conway et al.* [1994b]. In general there is a good agreement between models in the structure of the seasonal cycle although there are large discrepancies in the magnitudes of the maximum and minimum concentrations. The comparison with the observed seasonal cycle is not as good. All models produce a maximum value which is too large and occurs too late. The winter concentrations are too low but the August-November period is reasonably simulated. Since all the models are producing similar errors, this suggests an error with the input sources rather than with the model transport, particularly as similar discrepancies are seen at most of the high northern latitude sites. Fung (pers. comm.) has indicated that low light and long pathlengths result in errors in the NDVI (and hence CO_2 fluxes) at high latitudes in spring. There may also be errors in the respiration estimates through the use of air temperatures rather than soil

temperatures. This illustrates the potential extra information that can be gained by running a range of transport models: had only one result been available it would be more difficult to distinguish between source and transport errors.

While the annual mean biospheric source is zero everywhere, this is not true of the spatial distribution of annual mean concentration resulting from this source. The interaction of seasonal variations in transport with the seasonal source produces non-zero annual mean concentrations. The zonal annual mean (Fig. 12) shows that for some models (CSIRO9, CSU, GFDL, NCAR and NIRE) the north-south gradient is around half that produced in the fossil experiment. Most models produce small positive concentrations around the equator. These result from the seasonal shift of the ITCZ. The positive concentrations in the northern mid-latitudes appear to be more associated with seasonality in vertical transport. For example, experiments with MUTM have shown that seasonality of convection is important. Three of the models with large concentrations in this region have explicit planetary boundary layer formulations and it is likely that the seasonality of the PBL depth contributes to the positive annual means. The ability to resolve changes in vertical mixing due to the diurnal cycle may also be important. The ANU model annual mean is rather different to the other models. Taylor (pers. comm.) has indicated that the negative concentrations result from the use of 1980 winds at only 7 levels; small positive concentrations were obtained in subsequent experiments when winds from the 1990s at 14 or 15 levels were used.

Many of the differences between models seen in the surface layer are attenuated with height. At 500 hPa the models give similar results for the amplitude of the seasonal cycle. The zonal mean amplitude (not shown) is small and relatively uniform through the southern hemisphere and then increases to 12–14 ppmv at the north pole. The 500 hPa distributions show a number of common features. Higher amplitudes are found in the Himalayan region presumably reflecting the proximity of 500 hPa to the surface. A second common feature is higher amplitudes to the east of the southern African and South American continents indicating transport of continental air over the oceans.

At 200 hPa, differences between models re-emerge. The zonal mean ptp amplitudes (Fig. 13) are similar through the southern hemisphere but vary widely in the northern hemisphere. *Nakazawa et al*. [1991] observed upper tropospheric amplitudes of 2–3 ppmv between 30°S and the equator increasing to almost 8 ppmv at 36°N. These amplitudes are generally larger than those produced by the models which may indicate that the models are sampling significant amounts of stratospheric air at 200 hPa.

The 200 hPa global distributions (not shown) produce some differences and some common features. Large amplitudes are found in the India-China region in many models. All models give some regions, predominantly tropical, where the amplitudes are higher at 200 hPa than at 500 hPa. This may result from rapid lifting of lower tropospheric air by convection followed by horizontal advection.

Discussion

Large-Scale Transport and Source Estimates

A simple way to characterize the large-scale transport properties of the various models is by their exchange times. The exchange time is defined as the difference in inventory of two boxes divided by the flux between them. In these experiments, where sources and growth rates (and hence fluxes) are equal for all models, we can use the equivalent definition of the difference in concentration between two boxes divided by half the difference in source strength between them. There is a wide variety of partitions of the atmosphere available for a calculation of exchange times. In this case we choose hemispheric surface means and hemispheric three-dimensional means. The surface mean is calculated in two ways. The first (column 1) is to use the true mean from all supplied data. The second (column 2) is calculated by fitting a smoothing spline [*Enting* 1987] to the annual mean concentrations interpolated to the sites in Table 2. The hemispheric means are then the area-weighted means of these splined values. The fitting is performed in the sine of latitude and the 50% attenuation length-scale

is 1. The exchange time calculated by the second method is perhaps more relevant to carbon cycle inversions in which the station data is the input. It should avoid those model differences which arise from sampling the strong point sources evident in Figure 1. All times are derived from the fossil experiment and use annual mean concentration differences. The exchange times, for those models which provided sufficient data for their calculation, are shown in Table 3.

The first thing to note about Table 3 is the wide range of transport efficiencies demonstrated throughout. This reinforces the impression from Figure 3. Comparing columns 1 and 2 shows that, while using data at sampling locations reduces the large-scale gradient, it does not reduce the differences in these gradients among models. Column 2 does suggest a slight reordering of model exchange times with ANU replacing MUTM as the model with the fastest exchange. Also the GFDL model is now the slowest of a cluster of models rather than a clear outlier. Finally, CSIRO9, under this measure, has an exchange time almost 30% larger than the next model. Comparison of columns 1 and 3 indicates great variety in the ratio of surface and vertically integrated exchange. For example, the GFDL model, which shows the second slowest exchange between surface boxes, has the third quickest exchange between vertically integrated hemispheric boxes. Also, among the three models which supplied cross-section data and used analyzed winds from ECMWF, the surface exchange times are more closely grouped than the vertical mean exchanges. Both of these suggest considerable differences in rates of vertical transport, which is supported by the range of surface to 500 hPa differences (as shown in Fig. 5).

It is worth relating the differences in large-scale horizontal exchange rates evident in Table 3 to carbon budgets derived from such models. At the simplest and broadest level we can write a first order two box model such as

$$\frac{\partial Q_n}{\partial t} = S_n - \frac{Q_n - Q_s}{\tau} \quad (1)$$

$$\frac{\partial Q_s}{\partial t} = S_s - \frac{Q_s - Q_n}{\tau} \quad (2)$$

where Q refers to mixing ratios, S to sources, τ is the exchange time and the subscripts refer to northern and southern hemisphere.

Differencing these two equations and assuming steady state yields

$$S_{-} = \frac{2Q_{-}}{\tau} \tag{3}$$

where the subscript refers to interhemispheric differences.

Since the measured concentrations are at the surface, the mixing ratios and exchange times are those for the surface boxes. Calculating τ from the fossil experiment as enumerated in column 1 of Table 3 and then applying the range of values to a nominal observed interhemispheric difference of 2 ppmv gives a range in the distribution of the net source of 1.9–3.8 GtCyr^{-1}, i.e. roughly 3 $\pm$1 GtCyr^{-1}. The sum of the two hemispheric sources must agree with the change in global inventory, approximately 3 GtCyr^{-1} through the 1980s. These two constraints combine to give

$$S_n = 3 \pm 0.5 \text{ GtCyr}^{-1} \tag{4}$$

and

$$S_s = 0 \pm 0.5 \text{ GtCyr}^{-1} \tag{5}$$

The uncertainty in the estimates arises only from uncertainties in transport for a perfectly known gradient; uncertainties in the CO_2 distribution will compound it.

It is difficult to compare the uncertainties arising from transport with those from other sources. This is because the error is of a different form. Uncertainties in transport, like those in gradient, affect the ascription of net sources to particular regions. *Enting et al.* [1995], in their synthesis inversion, found that uncertainties in gradient did not greatly affect calculated errors in *global* fluxes. Put another way, knowledge of gradients was not a strong constraint on global fluxes. The above result further weakens this constraint. This analysis does not hold for regional budgets.

One other difficulty for comparison is that our transport error is quoted as a full range not a standard deviation. Numerically, our error in hemispheric source due to transport

uncertainty is comparable to that quoted by *Schimmel et al.* [1995] for the input due to fossil fuel combustion (0.5 $GtCyr^{-1}$). If the error in perturbation ocean uptake (0.8 $GtCyr^{-1}$) is distributed evenly between hemispheres, it also produces errors comparable to those arising from transport. Quoted errors in the terrestrial fluxes are considerably larger than all of these.

Fig. 12 highlights another uncertainty which will impact on source estimation. The zonal mean annual mean response to the purely seasonal terrestrial biotic source shows large differences from model to model. If a set of sources is required to match the seasonality of the observed CO_2 concentrations, it will contain seasonal components somewhat like the terrestrial biotic source used in this project. Fig. 12 suggests a large difference in the implied large-scale gradient from such a source.

While the budget analysis would seem to cast doubt on results using atmospheric transport models in CO_2 inversion studies, several ameliorating factors should strengthen our confidence. The most important is that the models used for the major studies in CO_2 inversions thus far, the TM1 model [*Keeling et al.* 1989] and the GISS model [*Tans et al.* 1990, *Enting et al.* 1993, 1995] have been calibrated against observed tracers. Also, as seen from both Figure 3 and Table 3, both models lie at the centre of a cluster of models. This cluster defines a rather smaller range, hence less uncertainty.

Reducing Uncertainties

Previous sections have highlighted the range of responses to equivalent sources among a range of transport models. We have also seen that this may have significant impact on our ability to distribute net sources spatially. This naturally raises the question of what to do to reduce the uncertainty. Three nonexclusive paths seem apparent. The first is to use a consensus of models as some guidance to the behaviour of the real atmosphere. It is noteworthy that about half the range seen in surface exchange times in column 1 of Table 3 or gradients in Fig. 3 is supplied by two models, GFDL and CSIRO9. Also, model consensus casts doubt on the reversed gradient seen in the annual mean surface response to the biotic source in the

ANU model. The same approach is implicit in our comments on the biotic source itself; the inability of the seasonal cycle of any model to match the observations suggests problems with the veracity of the source estimate.

Another approach is to use tracers with known sources and structure. Typically CFCs and krypton-85 have been used to assess interhemispheric transport while radon-222 has been used to assess vertical transport. Both are important as can be seen from comparison of columns 1 and 3 of Table 3. The two outlying models, CSIRO9 and GFDL, produce the largest ratios of surface to vertically integrated exchange times, suggesting their different behaviour arises from differences in vertical transport. Thus a good calibration of vertical transport would be a strong discriminator among models.

For any proposed calibration tracer, uncertainties in the sources and the relatively sparse observations, spatially or temporally, means that calibration is not an easy task. For example, *Prather et al*. [1995] report an intercomparison of 20 two and three-dimensional model simulations of radon and find a good agreement between 'established' 3d models but comment on the lack of observations available to test the model predictions. Despite these difficulties, calibration tracers should prove useful in validating model transport.

A third approach is to isolate and verify those aspects of model performance thought to be influential on CO_2 transport. Such calculations have been carried out by, for example, *Fung et al*. [1983], *Taguchi* [1994] and *Denning et al*. [1995]. It is intended to combine approaches two and three in a second phase of the TransCom project in which a calibration tracer, SF_6, will be used and detailed flux diagnostics will be retained for comparison. This should allow a mechanistic understanding of the differences in modelled transport.

Conclusions

We have compared the simulation of CO_2 concentration due to fossil fuel emissions and biospheric exchange by 12 atmospheric tracer transport models. While each model produces broadly similar concentration distributions there is a large range in the efficiency of transport

among models. For example, surface interhemispheric exchange times varied by a factor of two, although this range can be significantly reduced by removing a few outlier responses.

The implications of these results for CO_2 budget studies are substantial. In addition to the range in meridional gradient produced by the fossil experiment, there is no clear consensus among models on the annual mean response to the biosphere exchange. The uncertainties in transport produce uncertainties in regional carbon budgets comparable to those from other elements of the inversion. Uncertainties in transport can also be expected to impact modelling of other chemical species in the atmosphere, whenever the lifetime is long enough that transport affects the distribution of the species. More detailed observations of CO_2 and other species, particularly aloft and over the continents, can play a major role in constraining both transport and net sources.

Acknowledgments. This study was carried out with support from the Australian Government Cooperative Research Centres Programme. P. Rayner thanks the Program in Atmospheric and Oceanic Sciences at Princeton University for support in the early stages of this study. We have appreciated helpful comments on this manuscript from I. Enting and I. Smith.

References

Ciais, P., P.P. Tans, J.W.C. White, M. Trolier, R.J. Francey, J.A. Berry, D.R. Randall, P.J. Sellers, J.G. Collatz and D.S. Schimel. Partitioning of ocean and land uptake of CO_2 as inferred by $\delta^{13}C$ measurements from the NOAA Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network. *J. Geophys. Res.*, *100*, 5051-5070, 1995.

Conway, T.J., P.P. Tans and L.S. Waterman: Atmospheric CO_2 records from sites in the NOAA/CMDL air sampling network. *Trends '93: A Compendium of Data on Global Change*, T. Boden, D. Kaiser, R. Sepanski and F. Stoss Eds., Carbon Dioxide Information Analysis Center, Oak Ridge, U.S.A, 41-119, 1994a.

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, 1994b.

Denning, A.S., I.Y. Fung and D.A. Randall. Gradient of Atmospheric CO_2 due to Seasonal Exchange with Land Biota. *Nature*, *376*, 220-223, 1995.

Enting, I.G. On the use of smoothing splines to filter CO_2 data. *J. Geophys. Res.*, *92D*, 10977-10984, 1987.

Enting, I.G., C.M. Trudinger, R.J. Francey and H. Granek: Synthesis Inversion of Atmospheric CO_2 Using the GISS Tracer Transport Model. CSIRO Div. Atmos. Res. Tech. Paper No. 29, 44pp, 1993.

Enting, I.G., T.M.L. Wigley and M. Heimann: Future Emissions and Concentrations of Carbon Dioxide: Key Ocean/Atmosphere/Land Analyses. CSIRO Div. Atmos. Res. Tech. Paper No. 31, 118pp, 1994.

Enting, I.G., C.M. Trudinger and R.J. Francey. A synthesis inversion of the concentration and δ^{13}C of atmospheric CO_2. *Tellus*, *47B*, 35-52, 1995.

Fung, I., K. Prentice, E. Matthews, J. Lerner and G. Russell. Three-dimensional tracer model study of atmospheric CO_2 : response to seasonal exchanges with the terrestrial biosphere. *J. Geophys. Res.*, *88*, 1281-1294, 1983.

Fung, I.Y., C.J. Tucker and K.C. Prentice. Application of Advanced Very High Resolution Radiometer Vegetation Index to Study Atmosphere-Biosphere Exchange of CO_2. *J. Geophys. Res.*, *92*, 2999-3015, 1987.

Heimann, M., and C.D. Keeling: A Three-dimensional Model of Atmospheric CO_2 transport based on observed winds: 2. Model description and simulated tracer experiments. *Aspects of climate variability in the Pacific and the Western Americas, Geophysical Monograph 55*, D. Peterson Ed., AGU, Washington (USA), 237-275, 1989.

Heimann, M.: The global atmospheric tracer model TM2. Technical report no. 10, deutsches klimarechenzentrum, hamburg, germany (issn 0940-9327), 51pp, 1995.

Jacob, D.J., and M.J. Prather. Radon-222 as a test of convective transport in a general circulation model. *Tellus*, *42B*, 118-134, 1990.

Jacob, D.J., M.J. Prather, S.C. Wofsy and M.B. McElroy. Atmospheric Distribution of ^{85}Kr Simulated With a General Circulation Model. *J. Geophys. Res.*, *92*, 6614-6626, 1987.

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. *Aspects of climate variability in the Pacific and the Western Americas, Geophysical Monograph 55*, D. Peterson Ed., AGU, Washington (USA), 305-363, 1989.

Keeling, C.D., T.P. Whorf, M. Wahlen and J. van der Plicht. Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature*, *375*, 666-670, 1995.

Law, R., and I. Simmonds: Estimating the Southern Hemisphere sink of carbon dioxide with a three-dimensional atmospheric tracer transport model. *Proceedings of the Fourth International Conference on Southern Hemisphere Meteorology and Oceanography, Hobart, Australia, March 1993*. American Meteorological Society, 1993, 431-432.

Law, R., I. Simmonds and W.F. Budd. Application of an atmospheric tracer model to the high southern latitudes. *Tellus*, *44B*, 358-370, 1992.

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.

Marland, G., T.A. Boden, R.C. Griffin, S.F. Huang, P. Kanciruk and T.R. Nelson: Estimates of CO_2 emissions from fossil fuel burning and cement manufacturing, based on the U.S. Bureau of Mines cement manufacturing data. ORNL/CDIAC-25, NDP-030, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, 1989.

Nakazawa, T., K. Miyashita, S. Aoki and M. Tanaka. Temporal and spatial variations of upper tropospheric and lower stratospheric carbon dioxide. *Tellus*, *43B*, 106-117, 1991.

Prather, M., M. McElroy, S. Wofsy, G. Russell and D. Rind. Chemistry of the Global Troposphere: Fluorocarbons as Tracers of Air Motion. *J. Geophys. Res.*, *92*, 6579-6613, 1987.

Prather, M., R. Derwent, D. Ehhalt, P. Fraser, E. Sanhueza and X. Zhou: Other Tracer Gases and Atmospheric Chemistry. *Climate Change 1994: Radiative Forcing of Climate Change and An Evaluation of the IPCC IS92 Emission Scenarios*, J. Houghton, L. M. Filho, J. Bruce, H. Lee, B. Callander, E. Haites, N. Harris and K. Maskell Eds., Cambridge University Press, 73-126, 1995.

Ramonet, M., J.C.L. Roulley, P. Bousquet and P. Monfray. Radon-222 Measurements During the TROPOZ II Mission, and Comparison with a Global Atmospheric Transport Model. *J. Atmos. Chem*, 1996. In Press.

Rayner, P.J., and R.M. Law: A Comparison of Modelled Responses to Prescribed CO_2 Sources. CSIRO Div. of Atmos. Res. Tech. Paper No 36, 82pp, 1995.

Taguchi, S. Cross-Tropical Trajectories in the Troposphere. *J. Meteor. Soc. Japan*, *72*, 531-553, 1994.

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

Taylor, J.A. A stochastic Lagrangian atmospheric transport model to determine global CO_2 sources and sinks - a preliminary discussion. *Tellus*, *41B*, 272-285, 1989.

Watterson, I.G., M.R. Dix, H. Gordon and J.L. McGregor. The CSIRO 9-level atmospheric general circulation model and its equilibrium present and doubled CO_2 climates. *Aust. Meteorol. Mag.*, 1995.

Figure 1. Fossil emissions in $gCm^{-2}yr^{-1}$. Each $1^\circ \times 1^\circ$ square with non-zero source is shaded. The darker greys indicate higher emissions. The values used to determine the grey shade are 1, 50, 100, 200, 500, 1000 and 5000 $gCm^{-2}yr^{-1}$.

Figure 2. Zonal monthly mean biosphere source. The contour interval is 20 $gCm^{-2}yr^{-1}$ and the -10 and 10 contours are also shown.

Figure 3. Zonal annual surface mean concentration in ppmv due to fossil emissions. The key indicates the model to which the results apply and also gives the concentration difference between the northern and southern hemispheres for that model.

Figure 4. Surface annual mean concentration due to fossil emissions for the TM2 model. The contour interval is 1 ppmv.

Figure 5. Difference (in ppmv) between the zonal annual surface and 500 hPa concentrations for the fossil experiment. The key indicates the model which applies to each curve.

Figure 6. Zonal annual 200 hPa mean concentration (in ppmv) due to fossil emissions. The key indicates the model which applies to each line.

Figure 7. 200 hPa annual mean concentration for the fossil experiment for MUGCM. The contour interval is 0.4 ppmv.

Figure 8. Surface zonal mean peak to peak amplitude (in ppmv) for the biosphere experiment. The key indicates the model that applies to each curve.

Figure 9. Surface peak to peak amplitude for the biosphere experiment from the CSU model. The contours are 5, 10 , 20, 30 and 50 ppmv and values greater than 30 and 50 ppmv are shaded in light and dark grey respectively.

Figure 10. Peak to peak amplitude (in ppmv) at 15 northern hemisphere monitoring sites in the NOAA/CMDL network. The sites are identified at the top of the graph using the codes given in Table 2. The letter that applies to each model is indicated by the key. The dashed lines represent the $\pm$ one standard deviation spread of the observed values (taken from Fig. 10 of Conway et al. (1994a)). Lines have been used only for clarity not because it is realistic to connect the values from different sites. Also for clarity, the Ragged Point (RPB) data have been shifted slightly to the south so that they do not overlay the Guam (GMI) data.

Figure 11. Monthly surface mean concentration (in ppmv) at Barrow (71°N, 203°E) for the biosphere experiment. The curve that applies to each model is indicated by the key. The observed seasonal cycle is shown by the circles.

Figure 12. Zonal annual mean concentration (in ppmv) for the biosphere experiment. The curve that applies to each model is indicated by the key.

Figure 13. 200 hPa zonal mean peak to peak amplitude (in ppmv). The curve that applies to each model is indicated by the key.

Table 1a. Model identification

Model	Modeller	Reference
ANU	J. Taylor	*Taylor* [1989]
CSIRO9	I. Watterson	*Watterson et al*. [1995]
CSU	A.S. Denning	*Denning et al*. [1995]
GFDL	P. Rayner	*Mahlman & Moxim* [1978]
GISS	C. Trudinger	*Fung et al*. [1983]
MUGCM	R. Law	*Law* [1993]
MUTM	R. Law	*Law et al*. [1992]
NCAR	D. Erickson	Erickson et al. (sub)
NIRE	S. Taguchi	Taguchi (sub)
TM1	S. Piper	*Heimann & Keeling* [1989]
TM2	M. Heimann	*Heimann* [1995]
TM2Z	M. Ramonet	*Ramonet et al*. [1996]

Table 1b. Model details

Model	Type	HRes	VRes	Adv	Wind	Freq	Hdif	Vdif	Con	Tur	PBL
ANU	Off	2.5°	7	lagrang	ECMWF(80)	Stats	Y	Y	···	···	···
CSIRO9	On	3.3×5.6°	9	semi-lag	···	···	···	···	Y	Y	···
CSU	On	4×5°	17	2nd ord	···	···	···	···	Y	Y	Y
GFDL	Off	265 km	11	2,4 ord	GFDL GCM	6 hr	Y	Y	···	···	···
GISS	Off	8×10°	9	slopes	GISS GCM	4 hr	Y	···	Y	···	···
MUGCM	On	R21	9	spectral	···	···	Y	Y	Y	···	···
MUTM	Off	R21	9	spectral	MU GCM	24 hr	Y	Y	Y	···	···
NCAR	On	2.8°	18	semi-lag	···	···	···	Y	Y	···	Y
NIRE	Off	2.5°	15	semi-lag	ECMWF(92)	6 hr	···	···	···	···	Y
TM1	Off	8×10°	9	slopes	ECMWF(79)	12 hr	Y	···	Y	···	···
TM2	Off	4×5°	9	slopes	ECMWF(86)	12 hr	···	Y	Y	···	···
ALTTM2Z	Off	2.5°	9	slopes	ECMWF(90)	12 hr	···	Y	Y	···	···

The model type is either online (a full GCM) or offline (driven with externally derived winds) with horizontal resolution, HRes and vertical resolution, Vres. All models use sigma levels in the vertical with the exception of the ANU model which uses pressure levels. The advection schemes used (Adv) include lagrangian (lagrang), semi-lagrangian (semi-lag), spectral and finite difference (2nd and/or 4th order, slopes [Russell and Lerner, 1981]) schemes. For offline models the wind source and frequency is given. Where the ECMWF analyses have been used the year is also given. The ANU model uses a statistical representation of sub-bimonthly wind variability. The remaining columns indicate whether a model includes a parameterisation of the sub-grid scale processes of horizontal diffusion (Hdif), vertical diffusion (Vdif), convection (Con), turbulence (Tur) and the planetary boundary layer (PBL). The classification is based on the modellers' own choice of terminology and so there is some overlap between the processes listed.

Table 2. Site locations used for comparing modelled and observed seasonal cycles

Code	Station	Latitude	Longitude
	Alert	82.5°N	297.7°E
AVI	St Croix	17.8°N	295.3°E
AZR	Azores	38.8°N	332.9°E
BME	Bermuda East	32.4°N	295.4°E
BRW	Barrow	71.3°N	203.4°E
CBA	Cold Bay	55.2°N	197.3°E
CMO	Cape Meares	45.0°N	236.0°E
GMI	Guam	13.4°N	144.8°E
KEY	Key Biscayne	25.7°N	279.8°E
KUM	Kumakahi	19.5°N	205.2°E
MBC	Mould Bay	76.2°N	240.7°E
MID	Sand Is	28.2°N	182.6°E
RPB	Ragged Point	13.2°N	300.6°E
SHM	Shemya	52.8°N	174.1°E
STM	Station M	66.0°N	2.0°E

Table 3. Interhemispheric exchange times in years

Model	Surface	Station	Vertical
ANU	1.16	0.91	···
CSIRO9	2.14	1.92	1.16
CSU	1.54	1.29	0.99
GFDL	1.76	1.53	0.82
GISS	1.26	1.14	0.90
MUGCM	1.06	1.00	0.58
MUTM	1.06	0.99	0.66
NCAR	1.38	1.21	···
NIRE	1.50	1.38	0.92
TM1	1.58	1.47	1.20
TM2	1.52	1.38	···
TM2Z	1.46	1.36	1.06

The exchange time is calculated, from the fossil experiment, as the difference in mean concentration between two boxes divided by half the difference in source strength between them.