TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices

Rachel M. Law[1], Yu-Han Chen[2], Kevin R. Gurney[3], and TransCom3 modelers[4]

[1]CSIRO Atmospheric Research, Aspendale, Victoria, Australia.

[2]Department of Earth, Atmospheric, and Planetary Science, Massachussetts Institute of Technology (MIT), Cambridge, Massachussetts, USA.

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

[4]D. Baker, P. Bousquet, L. Bruhwiler, P. Ciais, A. S. Denning, S. Fan, I. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, T. Maki, S. Maksyutov, K. Masarie, P. Peylin, M. Prather, B. Pak, J. Randerson, P. Rayner, J. Sarmiento, S. Taguchi, T. Takahashi, C-W. Yuen

Contact information:
Rachel M. Law
CSIRO Atmospheric Research,
PMB 1, Aspendale,
Victoria 3195,
Australia.
Tel: +61 3 9239 4427
Fax: +81 3 9239 4444
Email: rachel.law@csiro.au

Abstract. TransCom3 is an intercomparison project for CO_2 source inversions. Annual mean atmospheric CO_2 data are used to estimate CO_2 sources using 16 different atmospheric transport models. Here we test the sensitivity of the inversion to the atmospheric data including data network choice, time period of data, baseline data selection and the choice of data uncertainty used. We find that in most cases regional source estimates lie within the source uncertainty range given for the control inversion so that the inversion estimates are relatively insensitive to most data changes. A small number of regions are sensitive to changes in the data network used, with a small number of sites exerting large influences on the source estimates. These sites are often identified by an inability of the inversion to fit the data at these locations. The model mean inversion values are mostly insensitive to the time-period of data used, with the exception of a couple of regions. Data selection has a small impact on source estimates for the mean across models but individual model sensitivity can be large. The magnitude of data uncertainties controls the relative magnitude of the 'within-model' and 'between-model' source uncertainties. Smaller data uncertainties lead to larger between-model uncertainties.

1. Introduction

Gurney et al. (2002 and this issue) have introduced the TransCom3 inversion intercomparison project and have presented control annual mean inversion results. Their control case made a number of choices about the atmospheric CO_2 dataset that was inverted. They used 1992-1996 mean concentrations at 76 sites from the GLOBALVIEW-CO2 (2000) dataset. These sites were chosen because their records were 70% complete. Data uncertainties were scaled from the residual standard deviations given in GLOBALVIEW-CO2 (2000) and were modified to set a minimum uncertainty and to account for the uneven spatial distribution of sites. It is important to understand how sensitive the control inversion is to these choices.

There is also value in making a more systematic exploration of inversion sensitivity to various data issues since only relatively limited tests have been reported previously for the type of inversion performed here. Fan et al. (1998) and Bousquet et al. (1999b) both tested the impact of dropping single sites from the networks that they used and found mostly small impacts. Bousquet (1997) assessed the sensitivity of an inversion to data selection, again showing only a small impact. He also compared sources estimated for two two-year periods and compared this to his decadal results to explore interannual variability related to El Niño. Significant interannual variability was found. Pacala et al. (2001) also report source estimates for different time periods. Their sensitivity includes impacts due to changing data availability in time, since they de-weighted sites with little or no data during a given time period. Baker (2001) tested the sensitivity of his inversion results to whether constant or variable data uncertainties are used. Each of these tests were limited mostly to one or sometimes to two or three transport models. Here we are able to use all 16 transport models that submitted data to the TransCom3 project. By performing sensitivity tests across a range of transport models, a more reliable indication of sensitivity should be obtained than if one model alone was used.

A number of issues are relevant to testing the data sensitivity of the inversion. The first issue is the choice of which data are used. CO_2 is measured by a number of groups around the world, using either in-situ analysers or by filling flasks for later laboratory analysis. Most of the available CO_2 data are brought together in GLOBALVIEW-CO_2 (2000, 2001). Those performing inversions have typically made two different choices in regards to data usage. One choice has been to use almost all the available data in order to obtain better spatial coverage. The other choice has been to use only those sites with long (typically greater than 10 year) records and often from only one measuring program, particularly when performing time-dependent inversions. Finding a reasonable balance between maximising spatial coverage and minimizing missing data is a question that warrants more systematic treatment. Intercalibration questions also need to be addressed.

The second issue is associated with the time period over which data are used in the inversion. There are large interannual variations in CO_2 sources (e.g. Francey et al., 1995; Keeling et al., 1995). For those inversions that do not solve for the whole interannual time

history of sources, the atmospheric concentration data have typically been averaged over a number of years. Various time periods have been used. It would be helpful to know how much of the difference between inversions has been methodological and how much is contributed by the different time periods over which the data have been averaged in the inversion.

The third issue is that of data selection. CO_2 observations are usually analysed for their 'clean-air' or 'baseline' concentration. In the case of flask measurements, this is determined by filling flasks when the meteorological conditions, e.g. wind direction, imply that the air is representative of a large region. This usually means that the air sample is from a marine rather than a continental direction. In the in-situ case, data are available for all times but it is baseline air that is reported in GLOBALVIEW-CO_2 (2000). In either case, this means that in most inversions, model data that are available for all times, are being compared with selected observational data. Both Law (1996) and Ramonet et al. (1996) show that selecting model data in a manner similar to the selection of the observations, can have a significant impact at some sites. It is not known how these systematic differences between model and observational data impact on the sources estimated by an inversion.

Many inversion methods, such as the one that we will use here, require an uncertainty to be attached to each observational data record. The uncertainty does not just encompass measurement error, which is typically quite small (0.1-0.2 ppm). It also needs to provide some indication of a model's ability to represent a given site. For example, a model is likely to be more effective in representing a remote southern hemisphere site, than a northern hemisphere continental site. This is particularly true for sites that are surrounded by inhomogeneous and relatively unknown sources. Any misrepresentation of the assumed spatial structure of sources in a region will be seen in the inversion as a data error and hence needs to be incorporated into the data uncertainty (Kaminski et al., 2001). Past inversions have set these data uncertainties in different ways, typically using a fixed uncertainty at all sites (e.g. Rayner et al., 1999) or using uncertainties related to the measurement variability at a site (e.g. Bousquet et al., 1999a). Here we compare these cases and also consider how the inversion results are altered as the magnitudes of the data uncertainties are changed.

2. Method

There are three components involved in estimating CO_2 sources from atmospheric measurements; the inversion method, the transport model and the atmospheric data. Here we keep the inversion method fixed and perform a series of inversions across a range of transport models and data choices.

2.1. Inversion

We use a Bayesian synthesis inversion method. This type of inversion uses prior information about the sources and so the solution is a compromise between fitting the

observational data and not moving too far from the prior sources as determined by prior source uncertainties. The globe is divided into 11 land and 11 ocean regions for which sources are estimated. These regions are shown in Figure 1. The method, as applied in this case, is described in detail in Gurney et al. (this issue). We use the inversion presented in that paper as the control inversion here. The inversion returns source estimates for each region and their uncertainties. In most cases, our analysis focusses on the mean results across models. For each region, we calculate the mean and standard deviation of the sources from each model and the mean of the uncertainties from each model. Following Gurney et al. (2002), we refer to the mean uncertainty as the 'within-model' uncertainty and the source standard deviation (a measure of model spread) as the 'between-model' uncertainty.

2.2. Models

The inversions are performed with 16 transport models. These are described in the appendix of Gurney et al. (this issue). They cover a range of resolutions, advection schemes, sub-grid scale parameterizations and forcing winds. Three models are full general circulation models (GCMs) so that transport occurs 'online' while the others are all 'offline' calculations. Among the offline models, some use GCM winds and others use analysed winds.

2.3. Data

The atmospheric data used in these inversions are taken from GLOBALVIEW-CO_2 (2000) (or GLOBALVIEW-CO_2 (2001) for 1997 data). GLOBALVIEW contains data that have been processed using curve-fitting to provide continuous, synchronous records at all sites from 1979. Missing data are estimated using a data filling technique (Masarie and Tans, 1995). GLOBALVIEW also provides residual standard deviations, a product of the curve fitting and a measure of observational variability at a given site. No corrections have been made to the data to account for intercalibration differences between measurement programs or between in-situ and flask data.

2.3.1. Site choices. The control inversion in Gurney et al. (this issue) used 1992-1996 mean data from sites from all measuring programs that were at least 70% complete during this period. This was a total of 76 data records at 64 locations. (More than one measurement program has data at some sites). Here we compare this inversion with cases in which the criteria on the completeness of the record is varied from 30 to 100%. We also compare networks containing sites from all measuring programs (denoted 'ALL') with cases where the sites are restricted to National Oceanic and Atmospheric Administration, Climate Monitoring and Diagnostics Laboratory flasks only (denoted 'NOAA') or to NOAA/CMDL and Commonwealth Scientific and Industrial Research Organization flask programs only (denoted 'NOAA+CSIRO'). NOAA/CMDL runs the largest CO_2 monitoring program while CSIRO runs the second largest network and also has an extensive flask comparison project

with NOAA/CMDL (Masarie et al., 2001). The number of data records for each case, which ranges from 20 to 111 records, is given in Table 1, and site locations are shown in Figure 1. The table also indicates when each site first meets the missing data and measuring program criteria and is consequently introduced to the dataset.

2.3.2. Time period of data. We compare the control 5-year (1992-1996) inversion with three 3-year inversions using data from 1989-1991, 1992-1994, and 1995-1997. As in the control case, data from each of these time periods have been averaged into single values for the inversion. We use the same 76 sites with the same uncertainties as in the control. A consequence of using the same sites is that some sites may no longer satisfy the 70% data coverage that was required for the control inversion. For example, several of the 76 sites are inactive during 1989-1991. GLOBALVIEW extrapolated values were thus used more extensively for certain sites. However, using the same 76 control sites rather than following the 70% data criteria allows us to focus on inversion changes from the data averaging period and minimise impacts from changing the data network.

In addition to using different data in the inversion, we also change the global growth rate and the fossil-fuel emissions to values representative of the new time periods. At each site, the growth rate has been computed by fitting the CO_2 record, over the relevant time period, with a linear trend and two harmonics. We use the mean and standard deviation of the linear trends of the 76 sites to represent the global growth rate, using a conversion factor of 2.123 GtC ppm^{-1}. We keep the growth rate uncertainty unchanged from the control value of 0.074 GtC yr^{-1}. The global growth rates for the control, 1995-1997, 1992-1994, and 1989-1991 are 3.27, 3.15, 2.37, and 2.76 GtC yr^{-1} respectively. Each model generated two fossil-fuel responses, representing the years 1990 and 1995. For the control inversion, these responses were scaled in the proportions of 20/80% for 1990 and 1995, respectively. For the new inversions, we use proportions for 1995-1997, 1992-1994, and 1989-1991 of 0/100%, 40/60%, and 100/0%, respectively.

2.3.3. Data selection. Two tests of data selection are performed. The experimental protocol for TransCom3 required modellers to submit monthly mean response data at all gridpoints, and four-hourly responses at a number of specified monitoring locations. In most cases, modellers were expected to use model data from the nearest grid point of a requested land type (usually ocean) to the specified location. However for sites, such as those on coastlines where it is known that baseline data are only available from the marine sector, the sampling location in the model was moved further offshore. It is these offshore locations that have been used in the control inversion in Gurney et al. (this issue). 20 of the 76 sites used in that inversion were moved offshore. This means that some account of baseline selection has already been made in that inversion. The first test that we perform removes this site-placement accounting of selection by simply using data from the nearest model gridpoint to each site in the inversion. Coastal sites may be moved onto land depending on each model's grid. This case takes less account of data selection than the control case.

The second test is designed to mimic flask sampling, assuming that the flasks are taken to avoid air that has recently passed over land. We use four-hourly data from each model to select data using the following procedure at each site. Firstly we find which land region produces the largest annual mean response at that site. We call this the local land region. We assume that flask sampling would try to avoid air from this region and perform data selection as follows. The four-hourly timeseries for the local land region is detrended and divided into 73 five day segments. In each five day segment, we identify the third smallest concentration and the time at which it occurs. These times can be thought of as the flask sampling times. We select the third smallest rather than the minimum concentration because some advection schemes produce occasional unrealistic low concentrations that it is preferable to avoid. The inversion requires the annual mean response and this is calculated as the mean of the 73 selected concentrations for the local land region. For all the other responses (the remaining regions and the fossil, neutral biosphere and global ocean fields), the mean is calculated from the 73 sampling times identified for the local land region. Selection has been applied to all sites, even those that are distant from any land. This avoids applying an arbitrary criteria as to how distant a site should be before selection is no longer appropriate. Thus this case gives something approaching an upper-bound for the impact of selection. One issue that is not addressed in this selection is how to treat sites that are near land region boundaries. It is possible that while we have avoided air from the local land region, the selected times could be influenced by air from a nearby neighbouring land region.

2.3.4. Data uncertainty. The control inversion of Gurney et al. (this issue) used data uncertainties that were based on the variability of the observations at a given site. In generating the data uncertainties that were used, a scaling factor was applied to the GLOBALVIEW residual standard deviations, such that the inversion produced a χ^2 of approximately one. (See Gurney et al, this issue, for the definition of χ^2). Length of record, proximity to other sites and a minimum uncertainty were also taken into account. Here we compare the control case with inversions where the control data uncertainties are multiplied by $0.2, 0.5, 2.0$ and 5.0. We also perform inversions in which uncertainties smaller than 0.25 ppm are not increased to this value and in which uncertainties are not increased for sites that are located close to each other. Finally we perform inversions with very different spatial distributions of uncertainty; the first uses constant uncertainties of 0.3 ppm for all sites while the second 'reverses' the control spatial distribution by using uncertainties of $0.5/(control\ uncertainty)$. Thus, low variability sites that would have had an uncertainty of 0.25 ppm are assigned an uncertainty of 2.00 ppm while continental, noisy sites, may have uncertainties as low as 0.22 ppm. While this is not a realistic case, it is helpful for establishing patterns of response to uncertainty changes.

3. Results

3.1. Site choices

The results from running the inversion with different networks are summarised in Figure 2. Mean model sources are shown as differences from the control inversion (Gurney et al., this issue). The mean within-model uncertainty from the control inversion is shown by the box. This gives an indication on whether the spread in estimates from the different networks is large compared to the estimated uncertainty. The three rows of results in each box are the three sets of networks dependent on which laboratory's measurements are included. The '7' in the top row is the control case and hence always sits on the zero line. Overall most source estimates lie within the uncertainty range of the control case, which suggests that most regions are not too sensitive to the choice of network. Australasia is the most noticeable exception with estimates lying some way outside both ends of the uncertainty range. We will consider this region in more detail shortly.

Other regions also show some notable features. For Europe, when all sites are used the estimates are generally lower than when only NOAA/CSIRO sites are used. The ALL case includes sites in Italy and Germany, which would have large fossil signatures. Including them in the inversion requires a larger sink to compensate. Boreal Asia shows the opposite result with the ALL case giving larger estimates than the NOAA/CSIRO networks. It is likely that this region, which has few local sites, is responding to the changes in the European region.

Most networks give estimates that are higher than the temperate North American control result, indicating a tendency to smaller rather than larger sinks there. For tropical Asia the NOAA, NOAA+CSIRO and ALL cases give similar estimates for small networks (90 and 100 %) and for large networks (30-50%) that include the South China Sea data. For intermediate networks (60-80%) the estimates are more erratic as the cases are influenced by different more distant sites (e.g. Darwin for the NOAA+CSIRO case, Minamitorishima for the ALL case).

The ocean region with the largest spread of estimates is the South Pacific, mostly due to the gradual introduction of the Pacific ocean ship data. Easter Island (29°S, 109°W) comes into the network at 50% and the estimates drop significantly (by almost 1 GtC yr^{-1} for the NOAA only case). It is worth noting that the data mismatch at Easter Island is still around 1 ppm suggesting that an even larger sink would be required to fit this datum alone but a compromise with the higher concentration Pacific ship data is moderating the sink. One contributor to the data mismatch is that the inversion is estimating ocean sources that are differences from the Takahashi air-sea flux distribution (Takahashi et al., 1999). This set of fluxes gives higher concentrations at Easter Island than at the 25°S and 30°S ship locations, contrary to the observed concentration difference between these records. Hence, the inversion struggles to fit the data. If the (NOAA-only, 50%) inversion is run without including the Takahashi fluxes, there is little change (–0.08 GtC yr^{-1}) to the South Pacific source estimate but the data mismatches are smaller at all but one of the 10 sites between 10-60°S. This

suggests that the Takahashi air-sea flux distribution within the South Pacific region is not compatible with the atmospheric concentration data.

All but one estimate for the Southern Ocean region is larger than the control. This is noteworthy since the control was significantly larger than the prior estimate based on ocean data. This finding suggests that for most networks, the discrepancy with ocean measurements is larger rather than smaller than that indicated by the control. Gurney et al. (2002) discusses a possible reconciliation between atmospheric and oceanic derived flux estimates for this region.

3.1.1. Australasian case. The mean source estimate across models for each network for Australasia is shown in Figure 3. Within- and between-model uncertainty is also shown. The results for the smallest (100%) networks are all similar with relatively large within-model uncertainties and small between-model uncertainties. This uncertainty behaviour is expected when there is little data to constrain the source estimate. In the 90% cases, additional sites in the NOAA+CSIRO and the ALL cases shift the mean source, lower the within-model uncertainty and increase the between-model uncertainty. The data at Cape Ferguson (19°S, 147°E) is largely responsible for the change in estimate in the NOAA+CSIRO case while Amsterdam Island (38°S, 78°E) also influences the ALL case.

Since we know that we do not have perfect calibration between measurement laboratories, it is worth checking whether the change in source induced by the addition of Cape Ferguson data is affected by calibration issues. To test this we alter the Cape Ferguson observed annual mean by ± 0.3 ppmv. The change in Australasian source is ± 0.1 GtC yr^{-1}, about half the change obtained if Cape Ferguson is dropped from the inversion. This suggests that calibration differences alone are unlikely to account for the change in estimated source. The ability of the inversion to fit the input observations illustrates the compromise the inversion has to make. In the NOAA only case, the Cape Grim (41°S, 145°E) mean data mismatch was 0.13 ppm. In the NOAA+CSIRO case the mismatch is increased to 0.24 ppm (by the increased source), while the mismatch for Cape Ferguson is –0.19 ppm. The ability to fit both sites simultaneously is limited because the inversion solves for the whole of Australasia and the spatial distributions of sources within regions are predetermined. Here the data suggest that the sources in the northern part of Australia should be increased relative to those in the south.

The 80% cases include a number of new sites and the differences between the three estimates are larger than with the 90% networks. The increase in the estimate for the NOAA+CSIRO case is due to the inclusion of Darwin (12°S, 131°E), which has a concentration about 3 ppm higher than Cape Ferguson. Despite increasing the source, the inversion struggles to fit the data; the mean data mismatch is –2.48 ppm. The data at Darwin are often collected under light wind conditions and are consequently subject to local sources (R. Francey, pers. comm.). The inversion is unable to cope with this local signal and so it would be preferable to exclude this site from inversions using large regions. Another shift in source estimate occurs between the 80% NOAA+CSIRO and ALL cases. The Baring Head (41°S, 175°E) record makes the dominant contribution to the change in source. There is also

a 22% decrease in within-model uncertainty with the addition of Baring Head data. The reduction occurs because Baring Head receives a strong signal from New Zealand and, since New Zealand is treated as part of the Australasian region, the uncertainty on the whole region drops.

There is one more shift in the estimated sources that is of interest. The 50% NOAA case gives a source that is almost 0.6 GtC yr^{-1} larger than the 60% case and between-model uncertainty increases by about 250%. The Australasian results are influenced by the South Pacific region, which shows a decrease in source of about 1 GtC yr^{-1} between the 60 and 50% cases, i.e. Australasia appears to be partially compensating the South Pacific change. The South Pacific decrease is caused by the addition of Easter Island data but the Australasian increase is also influenced by the 35°S Pacific ship data, which is located between Australia and New Zealand. The increase in between-model uncertainty appears to be linked to the relative magnitude of a model's response at Cape Grim or at the 35°S ship location due to Australian sources. The seven models with a larger response at Cape Grim gave smaller increases in Australian source (0.42 GtC yr^{-1} on average) than the nine models with a larger response at the 35°S ship location (0.89 GtC yr^{-1} average source increase). This different sensitivity to the inclusion of the 35°S ship data results in the increased between-model uncertainty.

3.1.2. Individual model sensitivity. So far we have only looked at the network impact on the mean sources. We should also consider the impact on individual models. Of interest is whether all models give similar changes in sources with the introduction of sites or whether the changes are quite variable. Figure 4 shows, for each region, the standard deviation (across models) of the difference in sources between inversions as the network increases in size. Only the ALL cases have been used when calculating the source differences. A low standard deviation indicates that all models give similar changes in source as the network changes while high standard deviations indicate that the source changes are more variable. In general, the standard deviations are larger for land than ocean regions. There is also a general tendency to smaller standard deviations as the number of sites in the network increases. Large standard deviations tend to occur when there are larger mean source differences and vice versa. Thus the standard deviations are low for the 50-40 and 40-30 source differences because the sources themselves change little for these larger networks, to which relatively few new sites are being added.

The largest standard deviation is for the 90-80 case for temperate Asia. We examine this region in more detail shortly. Tropical Asia produces the second largest standard deviation, in the 60-50 case. This is associated with different model responses to the addition of South China Sea (SCS) ship data. While most models give large positive source changes (up to 1.2 GtC yr^{-1}) when the SCS data are added, three models give almost no change and one model gives a large negative change (–0.6 GtC yr^{-1}). The largest ocean standard deviation is for the tropical Indian region, 100-90 case and is associated with the addition of the Seychelles (5°S,

55°E) record. The South Pacific gives more consistently large standard deviations that, as discussed above, are associated with different model responses to the addition of Pacific ocean ship data and the Easter Island record.

Figure 5 shows the temperate Asian sources and within-model uncertainties for the eight ALL networks. There are two network transitions (90 to 80% and 60 to 50%) where the reductions in within-model uncertainty are larger and the source estimates change substantially and variably across models. These transitions correspond to the addition of Ulaan Uul in the 80% network and Mt. Waliguan in the 50% network. The within-model uncertainties are quite variable across models and are determined by the concentration responses at each site from the temperate Asian region sources. For example, the TM3 response at Ulaan Uul is 3.34 ppm while other models range from 2.23 to 2.91 ppm. Consequently the TM3 uncertainty is substantially lower (0.38 GtC yr^{-1}) than the other models (0.59-0.88 GtC yr^{-1}). This relationship is confirmed by the high correlation (0.96) between the Ulaan Uul minus Tae-Ahn (the nearest site in the 90% network) concentration response and the reduction in uncertainty between the 90 and 80% networks. Likewise there is a similarly high correlation (0.97) between the Mt. Waliguan minus Ulaan Uul response and the uncertainty reduction between 60 and 50% networks.

Most models give a decrease in source with the addition of the Ulaan Uul site (80% network), but the magnitude of the source change varies greatly from –2.3 to 0.0 GtC yr^{-1}. The change in source is greater than the uncertainty for at least a third of the models. While the range of source differences is large, the differences act to reduce the range of source estimates in the 80% network case (1.9 GtC yr^{-1}) compared to the 90% case (3.5 GtC yr^{-1}). By contrast, the addition of Mt. Waliguan as part of the 50% network, increases the source range across models to 2.7 GtC yr^{-1}. The change in source between the 60% and 50% networks is quite variable across models. This variable sensitivity to the addition of Mt. Waliguan data is in part due to the ‘rectifier’ contribution from the covariance of the seasonal cycles of biospheric fluxes and atmospheric transport. This is seen in Figure 6, which shows source estimates for inversions in which the neutral biosphere response is not included (see Gurney et al., this issue, section 6.1 for further discussion of the ‘no rectifier’ case). The range of source estimates is reduced, particularly in the 50% case, when an increased range is no longer produced by the Mt. Waliguan data. Clearly, the magnitude and spatial distribution of the rectifier is an important component of the model-model differences that we currently see in some regions. It will be important to check whether similar model differences are also seen in inversions from monthly mean data, when the rectifier does not have to be included as a separate response.

3.2. Time period of data

Figure 7 summarises the inversion sensitivity to different years of data by showing mean inversion results for all time periods and all 22 basis regions. The figure is similar in style to Figure 1 of Gurney et al. (2002), plotting mean source estimates and their within-model and between-model uncertainties. For each region, the leftmost value is the control run, followed by the 1995-1997, 1992-1994, and 1989-1991 data periods, respectively. For most regions, using different years for the inversion does not result in significant deviations from the control; the mean sources lie within the mean uncertainty of the control run. Only temperate North America for 1989-1991 and the tropical Indian ocean for 1995-1997 and 1989-1991 exceed a 0.5 GtC yr^{-1} difference from the control. We examine these two regions in more detail shortly. Since the control case encompasses all but 1997 in the 1992-1994 and 1995-1997 cases, the sources for these two time periods show deviations of opposite sign from the control case for all but two regions.

It is helpful to determine whether the differences in estimated sources are determined by the differences in CO_2 observations from the control time period or by the change in the global growth rate and the fossil fuel response to more closely match the actual years of data. To do this, a sensitivity check has been performed in which these two parameters are retained at their control values. Results indicate that source differences due to changes in the global growth rate constraint are generally spread over all regions, rather than influencing a particular region. Use of different fossil responses also does not have a large influence, as the distribution and magnitude of 1990 and 1995 fossil-fuel emissions are mostly similar. Only one of these sensitivity tests causes a regional change greater than 0.2 GtC yr^{-1} (1989-1991 Temperate Eurasia increases by 0.29 GtC yr^{-1} when the control fossil-fuel mixture is used). Thus the differences in the inversion results are mostly caused by interannual differences in the GLOBALVIEW-CO2 dataset.

As mentioned in the Methods section, the 1989-1991 time-period has the lowest amount of real data for the 76 inversion sites and therefore relies more heavily on the GLOBALVIEW interpolated and extrapolated data. The interpolation/extrapolation procedure fills data gaps by reference to a marine boundary layer latitudinal distribution created from existing observations. Thus extended data will contain any interannual variability in this reference latitudinal distribution but will not contain local interannual variability. Consequently, this procedure may underestimate the actual interannual variability that would otherwise have been observed. Thus, using the interpolated/extrapolated data likely decreases the differences in the inversion results between time-periods than if real observations had been fully available.

The results for temperate North America and the tropical Indian ocean are plotted as the solid and dashed lines, respectively, in Figure 8. All 16 model results have been plotted as source differences from the control case, as shown by departures from the zero line. The 1989-1991 period shows the greatest difference for both regions. The average of the

1992-1994 and 1995-1997 cases lie close to the control value, as indicated by the crossover point near the zero line. The 1992-1994 value is slightly closer to the control value because all three of its years are contained within the control time period, while only the first two years of the 1995-1997 period are contained by the control. The responses of the two regions also show different behaviours when compared to one another. The overall model spread is larger for temperate North America than for the tropical Indian ocean, particularly for the 1989-1991 time-period.

To determine which sites are responsible for the source differences from the control in Figure 8, we have performed sensitivity inversions by removing one or more sites. The control differences for the Tropical Indian Ocean are found to depend on the nearby Seychelles and Darwin sites. Removing these two sites collapses the 1995-1997, 1992-1994, and 1989-91 source differences from –0.56, 0.37, and 0.69 GtC yr^{-1} to 0.05, –0.05, and 0.07 GtC yr^{-1}, respectively. The 1989-1991 difference can also be eliminated by removal of just Seychelles from the inversion. No real observations were taken at Darwin during this period, so the inversion used extrapolated data that does not contain local interannual variability. The tropical Indian ocean is not sensitive to these particular extrapolated values. When Darwin becomes active for the later time periods, its anomolously positive values (Section 3.1.1) have a large impact on the tropical Indian ocean. For these periods, both Darwin and Seychelles must be removed for the control source difference to be removed. In contrast to the tropical Indian ocean, the source differences for temperate North America cannot be eliminated by removal of one or two nearby sites. Several sensitivity tests with multiple site removal fail to have a large impact on the control difference. This suggests that this region is sensitive to most of its nearby sites which happen to give consistent observational values. In this sense, the variations in source estimates for the different time periods for this region are more robust than for the tropical Indian ocean, which depends on only two stations. The dependence on multiple stations for temperate North America protects it from anomolous observations at a single site.

3.3. Data selection

Three sets of inversions have been run to compare cases with more (‘flask-like’) and less (‘nearest-grid’) data selection with a control case. The inversion configuration for the three cases is different, in two ways, to the standard case used above. Firstly, two models were excluded from these inversions. The CSU model was not used because the advection scheme used by this model is ??, which has no requirement to avoid ‘negative’ (i.e. smaller than initial condition) concentrations. Thus the scheme does not give a clean baseline signal and so the method used here for data selection cannot be successfully applied to this model. The SKYHI model was not used because high frequency time series were not submitted for this model so data selection could not be performed. Secondly, we have removed the Pacific ocean ship data

from these inversions as high frequency timeseries were not collected for these locations so baseline selection cannot be applied.

Figure 9 shows the mean results from the three inversions. Overall the results for the control with this inversion configuration are very similar to those from the previous control. The largest differences in source (up to 0.25 GtC yr^{-1}) are around the Pacific region and are associated with the removal of the Pacific ship data. Comparing the nearest-grid case with the control case, five regions show a change in mean source greater than 0.2 GtC yr^{-1}: temperate North America, temperate Asia, tropical America, tropical Asia and northern Africa. In these model mean results, none of the changes move the mean source outside of the range of sources given by the control case uncertainty. However these relatively small changes in the mean source hide much larger changes in some individual models, which are also reflected in increased between-model uncertainties for many regions. For example, for temperate North America, the nearest-grid case gives a larger sink than the control by over 1.0 GtC yr^{-1} for three models with the maximum difference being –2.8 GtC yr^{-1}. The Key Biscayne site is responsible for these large changes. If the inversion is run without this site then the mean model source difference is 0.08 GtC yr^{-1} instead of –0.43 GtC yr^{-1} and the range across individual models is –0.26 to 0.37 GtC yr^{-1}. The removal of Key Biscayne also reduces the mean source differences for tropical America and northern Africa, indicating that the changes in these regions were mostly a compensation for the change in temperate North America. It is worth commenting on the sign of the differences found; when Key Biscayne is moved onto or towards the land, the temperate North American sink tends to get larger. This is counter-intuitive. We might expect that moving Key Biscayne onto land would increase its concentration response and thus a smaller sink would be required to fit the observations. The opposite result occurs because the inversion also effectively uses fixed fossil, neutral biosphere and ocean responses to invert from. Shifting Key Biscayne towards land significantly increases the fossil contribution at this site, particularly in some models and so the inversion finds a larger sink to compensate for this extra fossil response. Similar reasoning can explain the change in the temperate Asian source but in this case the Tae-Ahn site is the dominant contributor to the change and tropical Asia and northern Africa are the regions that compensate changes in temperate Asia.

Comparing the flask-like selection with the control case, there are seven land regions with source differences greater than 0.2 GtC yr^{-1}: temperate North America, Europe, Australasia (positive differences) and boreal Asia, temperate Asia, northern Africa and South America (negative differences). It seems likely that the regions with the positive differences are directly influenced by the change in site responses due to selection while those regions with negative differences are probably compensating the positive changes. Given the ‘nearest-grid’ case, we might have expected temperate Asia to show a positive rather than negative difference. It appears that this does not occur because three of the sites in, or close to this region, Ulann Uul, Tae-Ahn and Ryori show larger responses (in almost all models) from boreal rather than

temperate Asia. Presumably this results from the combination of transport and the spatial distribution of source applied to each region so that a site sees more CO_2 from its neighbouring rather than local region. In the method chosen here, selection times are determined for the region with the largest concentration response and so any change in response due to selection is likely to be largest for this same region. This is found in most models with larger response changes in boreal Asia than temperate Asia.

The two regions with largest postive changes, temperate North America (0.49 GtC yr^{-1}) and Australasia (0.43 GtC yr^{-1}), are quite different in the consistency of the changes across models. Temperate North America shows much more variable changes, from –0.3 to 0.9 GtC yr^{-1} with one outlier at 2.4 GtC yr^{-1}, whereas the Australasian changes range from 0.2 to 0.6 GtC yr^{-1}. The smaller range appears to be mostly constrained by the relatively small (0.59 GtC yr^{-1}) prior uncertainty used for Australasia. If this uncertainty is increased to 1.5 GtC yr^{-1}, the magnitude used for temperate North America, then the mean change in source due to selection increases to 1.0 GtC yr^{-1}, with a range across models of 0.2-2.2 GtC yr^{-1}.

Overall the estimated source uncertainties change very little for the inversions with more or less selection. The change in within-model uncertainty for land regions is slightly larger and opposite in sign to that for ocean regions. The land change is negative in the less-selection case and positive in the more-selection case. This would be expected as selection is intended to remove the signal from land regions and thus uncertainty should increase as more selection is performed. The changes in between-model uncertainty are also larger for land than ocean but the change is positive, i.e. to larger between-model uncertaity, for both more and less-selection cases. This is not anticipated. It is possible that in the less-selection, nearest-grid case, the movement of sites onto land increases model differences caused by how the spatial distribution of sources is represented on the different grids of each model. In the more-selection case when the local land influence is minimised, this may make the inversion more dependent on transport differences between models as regions further from a site exert more influence.

3.4. Data uncertainty

Figure 10 shows the estimated sources and uncertainties for the inversions in which the data uncertainties were scaled up and down. The clearest feature is the change in relative magnitude of the within-model and between-model uncertainties. As data uncertainty is increased, the within-model uncertainty increases and between-model uncertainty decreases and vice versa. When data uncertainty is decreased the inversion is forced to fit the data more closely. This accentuates model transport differences and hence large between-model uncertainty results. The control case, which used data uncertainties scaled such that the inversion gave $\chi^2 \approx 1$, typically gives within-model uncertainties that are similar or slightly larger in magnitude that the between-model uncertainties. This would seem to be a desirable choice. It suggests that most of the time an individual model will give a source that is within

the range of sources determined by the mean source and uncertainty across models. The change in data uncertainty also produces a change in estimated source in many regions. However, in almost all cases, the mean source still lies within the uncertainty range of the control case. Exceptions are the 0.2 scaling case for the tropical and South Atlantic regions.

Figure 11 shows the estimated sources and uncertainties for inversions in which the data uncertainties were distributed differently. From left to right, the first case imposed no minimum on the data uncertainties so some sites could have data uncertainties less than the 0.25 ppm minimum used in the control. There is little change from the control inversion results (centre case in figure) except in the southern regions. This is anticipated since it is mostly sites in the southern mid-high latitudes that are affected by the minimum criteria. For these southern regions the within-model uncertainty tends to decrease and the between-model uncertainty tends to increase. This follows the general pattern seen in the previous figure when data uncertainty is reduced. There are changes in the mean source in all the southern regions except the South Atlantic and the Southern Ocean. Except for Australasia, the estimated source is within the control case uncertainty range. The changes in source are not clearly driven by just one or two sites. It is likely that the smaller data uncertainties used in this case mean that the inversion is trying to fit longitudinal concentration variations between sites within the Southern Ocean region and around the Antarctic coast. Since the Southern Ocean is treated as one large region, the improved data fit can only be accomodated by adjusting sources in the neighbouring regions to the north even though this is unlikely to be the real cause of the longitudinal variations in concentration. Rather, the longitudinal variations imply that the spatial distribution prescribed for the Southern Ocean region should be adjusted. This example illustrates the potential for biased source estimates due to spatial aggregation if the data uncertainties being used are too small.

The second leftmost case is for inversions in which the data uncertainty was not increased when sites were co-located (within 4° latitude and longitude and 1000m altitude). This colocation test impacts 36 of the 76 sites used in the inversion. Overall these changes in data uncertainties had little impact on the inversion results. The largest change in mean source is for Northern Africa (0.33 GtC yr^{-1}) and it seems most likely that this is compensating the small changes (around –0.1 GtC yr^{-1}) found for temperate North America, Europe and the South Indian ocean.

The fourth case from the left uses constant data uncertainties of 0.3 ppm at all sites. This is a smaller data uncertainty than the control for about two thirds of the sites. Consequently, in general, the results show reduced within-model source uncertainties relative to the control and increased between-model uncertainties. Europe is an extreme example. Europe is also one of the regions that shows a substantial change in mean estimated source. In the northern regions there is a shift in uptake from land to ocean. In the tropics, the Indian ocean changes from a small sink to a moderate source. This region appears to be quite sensitive to the data uncertainty applied to Cape Rama. In the southern regions, Australasia shows the largest

change in estimated source. This results from placing a low data uncertainty on the Darwin record, which we saw earlier is anomalously high. Thus, the Australasian source is increased to try and better fit the Darwin data. The increased sink in the South Pacific acts to compensate for the change in Australasia.

The rightmost case shows the results when the data uncertainties were small for land sites and large for ocean sites. For the southern regions (except Australasia) and half the tropical regions, this case gives estimates that are close to the prior sources, with little decrease in within-model uncertainties from their prior source uncertainties and small between-model uncertainties. The large data uncertainties placed on the southern and tropical ocean sites mean that the inversion is able to extract little information from the data. Australasia is the only southern region that does not follow this pattern. As in the previous case, its estimates are dominated by the Darwin data since here it has the lowest data uncertainty of all the sites in the vicinity of Australasia. Likewise the tropical Indian ocean result is dominated by the relatively low data uncertainty used for Cape Rama. For the northern regions, the impact of the small uncertainties used for continental sites is seen most clearly for temperate North America and Europe. In both cases, the within-model uncertainty is small, the between-model uncertainty is large and there are moderate changes in the estimated sources (to less uptake). The lost uptake is shifted to neighbouring land or ocean regions, which tend to also show increased within-model uncertainties.

4. Discussion

Much detailed information has been presented from the various sensitivity tests. Here we attempt to highlight the main findings particularly where these might provide guidance for how annual mean inversions are best set up. Overall, we have seen that for most of the sensitivity tests, the mean estimated sources lie within the range of sources determined by the within-model uncertainty from the control inversion. This would indicate that the mean sources and their uncertainties estimated from the control inversion are reasonably robust. It also suggests that the data-related choices made in the inversion set-up for the control case were good.

Data network choices remain one of the more difficult decisions when setting up an inversion. For annual mean inversions, such as used here, it is easy to perform many tests. The difficulty is in assessing whether the sensitivity obtained to any given site is representing new, valid information being added to the problem or not. A large sensitivity in a data sparse region may be quite appropriate where a similar sensitivity in a region where there are many sites suggests a problem with the new measurement or with the assumed spatial distribution of sources within regions. One indicator that a site should be examined more closely is the ability of the inversion to fit the data at that location. For example, if the data mismatch is more that twice the data uncertainty applied at that site, then a potential problem is implied.

Applying this criteria to the set of network inversions tested here, five sites have data mismatches that are too large in at least some of the cases: Black Sea, Cape Rama, Darwin, Easter Island and Guam. We have already highlighted the sensitivity of the inversion to Darwin and Easter Island. For Darwin, the major problem seemed to be with the representativeness of the measurements. Thus, we would suggest that Darwin data are not well suited to inversion work and should be flagged as such in the GLOBALVIEW dataset. For Easter Island, the problem appears to be more with the spatial distribution of fluxes within the South Pacific region. One solution is to split the South Pacific into more than one region. The other would be to increase the data uncertainties for all sites in the vicinity of the region. This aims to avoid biasing the source estimates due to the incorrect spatial distribution.

Of the other three sites, Black Sea has a short record (only making it into the 30% network) and hence is heavily dependent on the data extension method for determining the mean 1992-1996 concentration. In this continental location, it is probably unrealistic to be this reliant on data extension and the choice of a 70%-complete criteria for the control inversion seems to be reasonable. Both Cape Rama and Guam give reduced data mismatches when the inversion is run without the seasonal rectifier contribution. Cape Rama shows a low, positive correlation (0.22) and Guam, a higher correlation (0.75) between the data mismatch for individual models and the magnitude of the seasonal biospheric response for each model. Again this suggests a spatial distribution error, though it is not clear whether it is with the underlying biospheric sources or in the pattern of rectification produced by the transport models. Larger data uncertainties may be required to allow for this.

There has been little work until now on the choice of data uncertainties for annual mean inversions. The sensitivity that we have found suggests that data uncertainty choices are more important than had previously been assumed. Clearly variable uncertainties are preferable to constant ones to accomodate the different locations of data records and our ability to model those locations. Scaling the magnitudes of the data uncertainties to give $\chi^2 \approx 1$ seems to give a good balance between within-model and between-model uncertainty. Data uncertainties need to be large enough to absorb errors in source spatial distribution within regions. A large between-model uncertainty in a given region may act as an indicator that the inversion is 'over-fitting' the data in that region and that a larger data uncertainty is appropriate.

Perhaps surprisingly, data selection has a relatively small impact in these annual mean inversions, at least in the mean across models. The practice of moving sampling locations offshore for coastal sites provides a good, easy method for taking some account of selection issues. The small impact seen here may not extend to inversions that solve for the seasonal cycle. This will need to be considered in the future.

As for data selection, we have found that the mean inversion results are relatively insensitive to using different data time periods. While interannual variations in sources are suggested in some regions, the variations are almost always less than the estimated source uncertainties. Annual mean inversions are designed to estimate long-term mean fluxes rather

than interannual flux variability. As such, it is preferable to use as long a data period as possible given the constraints of data gaps and changing networks. One of the next stages of the TransCom3 project will be to perform inversions using monthly timeseries data to give interannual varying sources. It will be interesting to compare the sources obtained here for the different three year periods with those that will be obtained from the full interannual inversion.

5. Conclusions

We have found that there are clear advantages in being able to perform sensitivity tests over a range of transport models. Where the sensitivity is consistent across models, we can have more confidence in our conclusions. We are also able to ascertain when individual models give sensitivities that are large and unrepresentative and would have misled if we were reliant only on that model. For this reason, the TransCom3 dataset may have an ongoing role in providing a set of model responses that can be used to test inversion methodology choices. It is much easier to run multiple sensitivity tests with annual mean inversions than with interannual ones. Thus, the knowledege of sensitivities gained here will provide useful input to decisions about inversion set-up for the TransCom3 interannual cases.

Acknowledgments. Internal reviewer. TransCom3 was made possible through support from the National Science Foundation (OCE-9900310), the National Oceanic and Atmospheric Administration (NA67RJ0152, Amend 30) and the International Geosphere Biosphere Program/Global Analysis, Interpretation, and Modeling Project.

References

Baker, D. F., Source and sinks of atmospheric CO_2 estimated from batch least-squares inversions of CO_2 concentration measurements, Ph.D. thesis, Princeton University, 2001.

Bousquet, P., Optimisation des flux nets de CO_2: Assimilation des mesures atmosphériques en CO_2 et en $^{13}CO_2$ dans un modèle de transport tridimensionel, Ph.D. thesis, L'Univ. Paris VI, 1997.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray, Inverse modeling of annual atmospheric CO_2 sources and sinks 1. method and control inversion, *J. Geophys. Res.*, *104*, 26,161–26,178, 1999a.

Bousquet, P., P. Peylin, P. Ciais, M. Ramonet, and P. Monfray, Inverse modeling of annual atmospheric CO_2 sources and sinks 2. sensitivity study, *J. Geophys. Res.*, *104*, 26,179–26,193, 1999b.

Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans, A large terrestrial carbon sink in North America implied by atmospheric and oceanic CO_2 data and models, *Science*, *282*, 442–446, 1998.

Francey, R. J., P. P. Tans, C. E. Allison, I. G. Enting, J. W. C. White, and M. Trolier, Changes in oceanic and terrestrial carbon uptake since 1982, *Nature*, *373*, 326–330, 1995.

GLOBALVIEW-CO_2, Cooperative Atmospheric Data Integration Project - Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, 2000, [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

GLOBALVIEW-CO_2, Cooperative Atmospheric Data Integration Project - Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, 2001, [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

Gurney, K. R., et al., Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, 2002a, (in press).

Gurney, K. R., et al., TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to inversion set-up, *Tellus*, 2002b, (this issue).

Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting, On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, *106*, 4703–4715, 2001.

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., The selection of model-generated CO_2 data: a case study with seasonal biospheric sources, *Tellus*, *48B*, 474–486, 1996.

Masarie, K. A., and P. P. Tans, Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *J. Geophys. Res.*, *100*, 11,593–11,610, 1995.

Masarie, K. A., et al., NOAA/CSIRO flask air intercomparison experiment: A strategy for directly assessing consistency among atmospheric measurements made by independent laboratories, *J. Geophys. Res.*, *106*, 20,445–20,464, 2001.

Pacala, S. W., et al., Consistent land- and atmosphere-based U.S. carbon sink estimates, *Science*, *292*, 2316–2320, 2001.

Ramonet, M., and P. Monfray, CO_2 baseline concept in 3-d atmospheric transport models, *Tellus*, *48B*, 502–520, 1996.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. L. Langenfelds, Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, *Tellus*, *51B*, 213–232, 1999.

Takahashi, T., R. H. Wanninkhof, R. A. Feely, R. F. Weiss, D. W. Chipman, N. Bates, J. Olafsson, C. Sabine, and S. C. Sutherland, Net sea-air CO_2 flux over the global oceans: An improved estimate based on the sea-air pCO_2 difference, in *Extended abstracts of the 2nd International CO_2 in the Oceans Symposium, Tsukuba, Japan, January 18-22, 1999*, pp. 18–01, 1999.

Y.-H. Chen, Department of Earth, Atmospheric, and Planetary Science, Massachussetts Institute of Technology (MIT), Cambridge, MA 02141, USA

K. R. Gurney, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, USA.

R. M. Law CSIRO Atmospheric Research, PMB 1, Aspendale, Victoria 3195, Australia. (e-mail: rachel.law@csiro.au)

Figure 1. Regions (named in italics) for which sources are estimated by the inversion and sites used (numbers). The region boundaries and site locations are those from the MATCH-MACCM3 model. Other models will use slightly different regions and locations depending on their model grid. The numbers differentiate networks based on the completeness of the data record: 100% complete (0), greater than 90% (9), 80% (8), 70% (7), 60% (6), 50% (5), 40% (4), 30% (3). Where more than one data record exists at a single location, the number indicates the most complete record. Sites that are mentioned in the text are labelled (normal font).

Figure 2. Difference in mean estimated source between the control inversion and inversions with different networks as indicated by the numbers for (a) land and (b) ocean regions. The regions are arranged approximately in their north-south order. The bottom row in each box is NOAA networks, the middle row is NOAA+CSIRO networks and the top row is ALL networks. The number indicates the record-completeness criteria: 100% (0), 90% (9), 80% (8), 70% (7), 60% (6), 50% (5), 40% (4), 30% (3). The control inversion is the ALL case with 70% criteria so the '7' in the top row always lies on the zero difference line. The box shows the magnitude of the within-model source uncertainty estimated in the control case.

Figure 3. Australasian region source estimates for different networks, indicated by the record-completeness criteria (100 to 30%) (left to right) and the measuring programs included: NOAA only (N), NOAA+CSIRO (+C), ALL (A). The mean source across models is indicated by the x and the standard deviation across models (the between-model uncertainty) by the error bar. The circles indicate the mean within-model source uncertainty returned by the inversion.

Figure 4. Standard deviation, across models, of the difference in source between inversions with different networks. Only the ALL cases are shown and the differences are between a given network and the next largest in size, e.g. 100-90 indicates the difference between the 100% ALL network and the 90% ALL network. Four regions are shown by bold lines as given in the key and the remaining land regions are solid lines and the remaining ocean regions are dashed lines.

Figure 5. Source estimates (b) and uncertainties (a) for each model for temperate Asia for each of the ALL networks. The line identification is given in the key.

Figure 6. Source estimates for temperate Asia for each of the ALL networks for inversions calculated without the neutral biosphere annual mean response. The line identification for each model is given in the key.

Figure 7. Model mean inversion results for all 22 basis regions using different data from different time periods. The leftmost value in each box is the control case 1992-1996, followed by 1995-1997, 1992-1994, and 1989-1991 respectively. Mean source estimates (including the background ocean fluxes) are indicated by the x, mean within-model uncertainty by the o and between-model uncertainty (the standard deviation across model source estimates) by the bar. For each land (solid line) and ocean (dashed line) region, the boxes represent the prior source estimate (central horizontal line) and prior uncertainty. Regions are arranged to approximate their north-south and east-west relationship.

Figure 8. Inversion results for 16 models for the temperate North American (solid lines) and tropical Indian ocean (dashed lines) regions. The sources for the three time periods 1989-1991, 1992-1994 and 1995-1997 are plotted as deviations from the control source estimate for that model and region.

Figure 9. As Figure 7 but showing mean sources and uncertainties for three sets of inversions with different data selection strategies. Left is the 'nearest-grid' inversion, centre is the modified control inversion, right is the 'flask-like' inversion.

Figure 10. As Figure 7 but showing mean sources and uncertainties for five sets of inversions with scaled data uncertainties. The control data uncertainties were multiplied by 0.2 (far left), 0.5 (left), 2.0 (right), 5.0 (far right). The control inversion results are in the centre of each box.

Figure 11. As Figure 7 but showing mean sources and uncertainties for five sets of inversions that used different spatial patterns of data uncertainty: no minimum (far left), no account of co-location (left), control (centre), constant 0.3 uncertainty (right), reversed distribution (far right).

Table 1. Sites used for each record-completeness criteria and measuring-program case

Complete	Programs	Number	Sites
100%	NOAA	20	alt, asc, bme, brw, cba, cgo, crz, gmi, izo, kum, mbc, mhd mid, mlo, psa, smo, spo, stm, syo, tap
100%	+CSIRO	21	mqa_02D0
100%	all	29	alt_06C0, brw_00C0, cmn_17C0, izo_27C0, mlo_00C0 ryo_19C0, smo_00C0, spo_00C0
90%	NOAA	25	bmw, key, rpb, sey, shm
90%	+CSIRO	32	alt_02D0, cfa_02D0, cgo_02D0, maa_02D0, mlo_02D0, spo_02D0
90%	all	43	ams_11C0, itn496_00C3, sch_23C0
80%	NOAA	29	hba, itn, pocn15_00D1, uum
80%	+CSIRO	42	aia005_02D2, aia015_02D2, aia025_02D2, aia035_02D2 aia045_02D2, daa_02D0
80%	all	56	alt_06D0, bhd_15C0, prs_21D0
70%	NOAA	44	bal, car030_00D2, car040_00D2, car050_00D2, cmo, hun, ice poc000_00D1, pocn20_00D1, pocn25_00D1, pocn30_00D1 pocs05_00D1, pocs10_00D1, pocs15_00D1, uta
70%	+CSIRO	59	aia065_02D2, cri_02D0
70%	all	76	esp_06D0, mnm_19C0, prs_21C0
60%	NOAA	49	goz, pocn05_00D1, pocn10_00D1, pocs25_00D1, pocs30_00D1
60%	+CSIRO	65	aia055_02D2
60%	all	83	ksn_24D0
50%	NOAA	60	chr, eic, pocs20_00D1, pocs35_00D1, scsn03_00D1, scsn06_00D1 scsn09_00D1, scsn12_00D1, scsn15_00D1, wlg, zep
50%	+CSIRO	77	sis_02D0
50%	all	97	jbn_29C0, zep_31C0
40%	NOAA	65	lef, pocn35_00D1, pocn40_00D1, scsn18_00D1, scsn21_00D1
40%	+CSIRO	83	esp_02D0
40%	all	106	lef396_00C3, wes_23C0, wlg_33C0
30%	NOAA	68	azr, bsc, pocn45_00D1
30%	+CSIRO	86	
30%	all	109	

Table 1. (continued)

Complete	Programs	Number	Sites

The sites are listed when they first meet the record-completeness and measuring program criteria and are introduced to the dataset. Complete networks for all cases can therefore be constructed from the sites listed for the relevant smaller networks. The site codes are those used in GLOBALVIEW-CO_2 (2000) except that the three letter codes are appended _00D0. See http://wwwsrv.cmdl.noaa.gov/ccgg/globalview/co2/app_b.html for site information. The lowest levels of the itn and lef tower sites have not been used since these cannot be represented by the relatively low resolution global transport models used here.

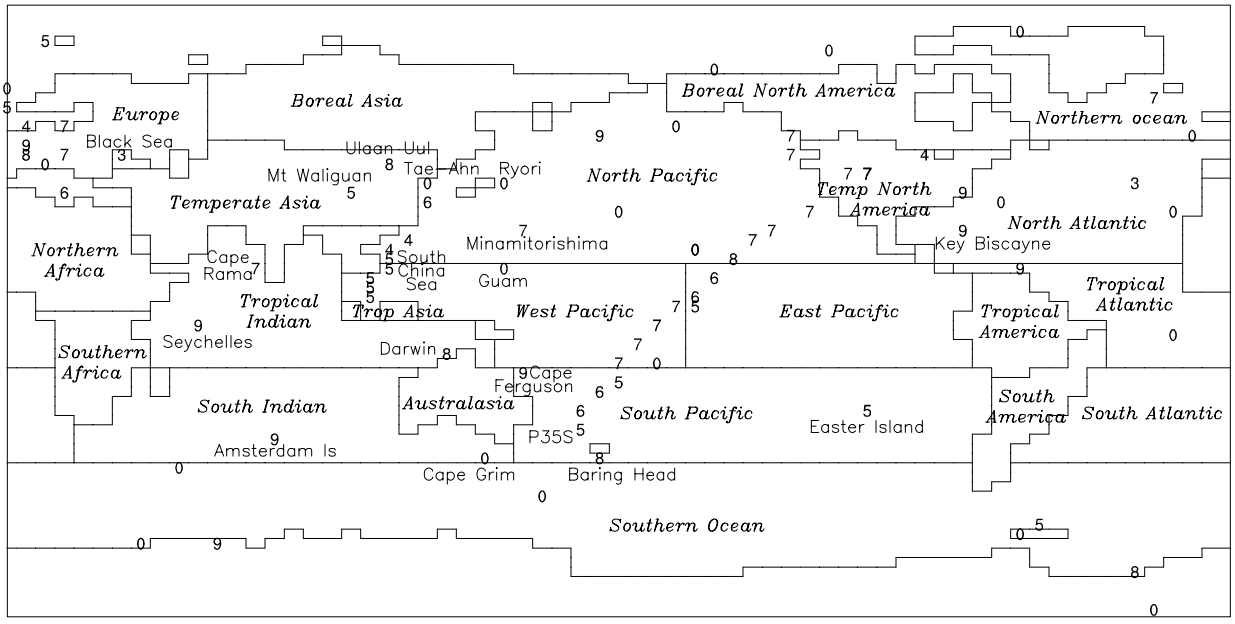

Figure 1. Regions (named in italics) for which sources are estimated by the inversion and sites used (numbers). The region boundaries and site locations are those from the MATCH-MACCM3 model. Other models will use slightly different regions and locations depending on their model grid. The numbers differentiate networks based on the completeness of the data record: 100% complete (0), greater than 90% (9), 80% (8), 70% (7), 60% (6), 50% (5), 40% (4), 30% (3). Where more than one data record exists at a single location, the number indicates the most complete record. Sites that are mentioned in the text are labelled (normal font).

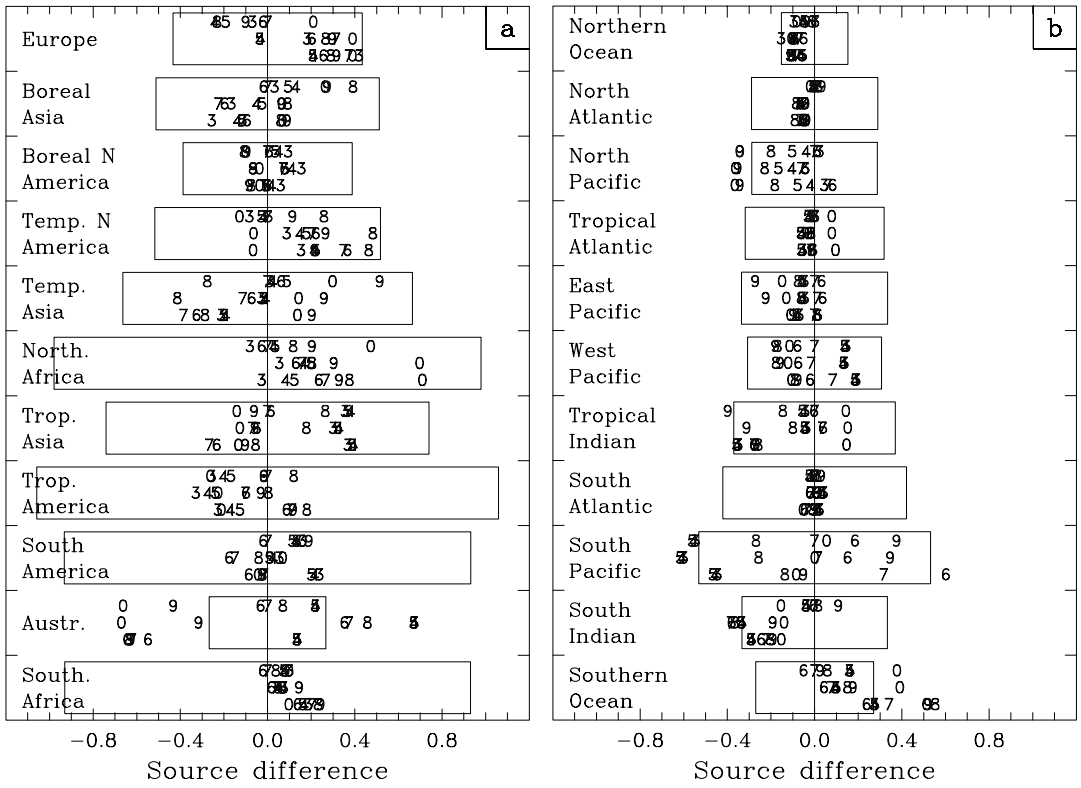

Figure 2. Difference in mean estimated source between the control inversion and inversions with different networks as indicated by the numbers for (a) land and (b) ocean regions. The regions are arranged approximately in their north-south order. The bottom row in each box is NOAA networks, the middle row is NOAA+CSIRO networks and the top row is ALL networks. The number indicates the record-completeness criteria: 100% (0), 90% (9), 80% (8), 70% (7), 60% (6), 50% (5), 40% (4), 30% (3). The control inversion is the ALL case with 70% criteria so the '7' in the top row always lies on the zero difference line. The box shows the magnitude of the within-model source uncertainty estimated in the control case.

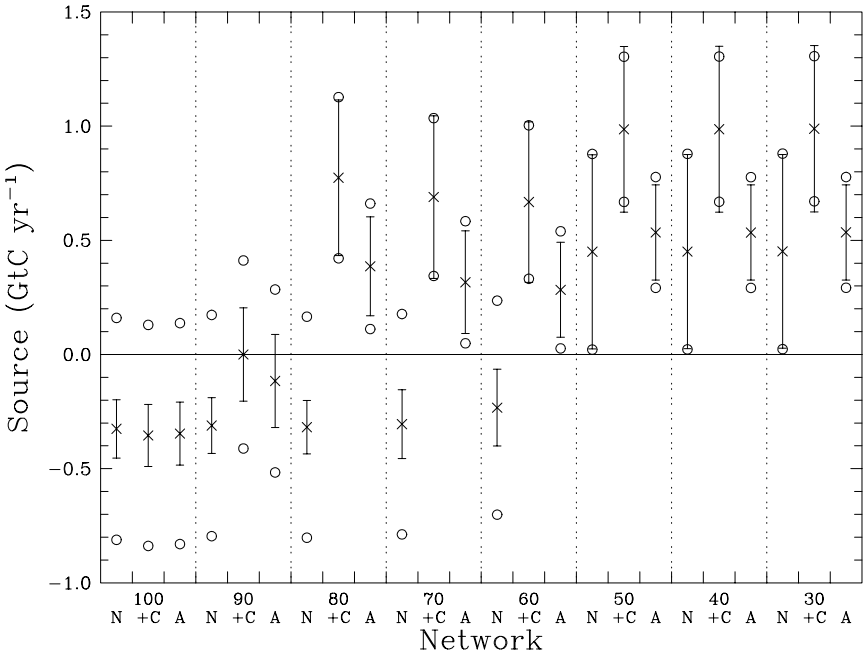

Figure 3. Australasian region source estimates for different networks, indicated by the record-completeness criteria (100 to 30%) (left to right) and the measuring programs included: NOAA only (N), NOAA+CSIRO (+C), ALL (A). The mean source across models is indicated by the x and the standard deviation across models (the between-model uncertainty) by the error bar. The circles indicate the mean within-model source uncertainty returned by the inversion.

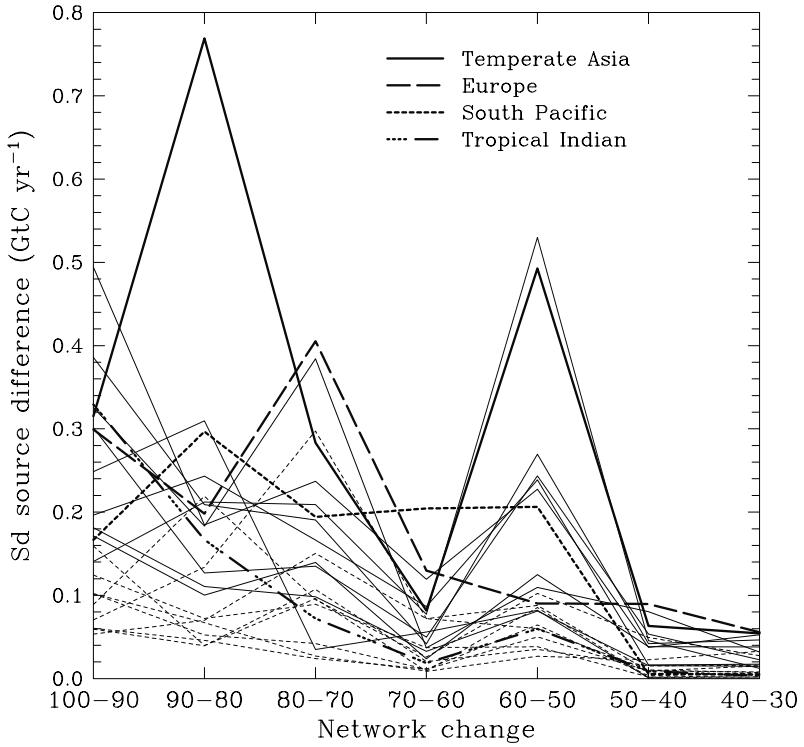

Figure 4. Standard deviation, across models, of the difference in source between inversions with different networks. Only the ALL cases are shown and the differences are between a given network and the next largest in size, e.g. 100-90 indicates the difference between the 100% ALL network and the 90% ALL network. Four regions are shown by bold lines as given in the key and the remaining land regions are solid lines and the remaining ocean regions are dashed lines.

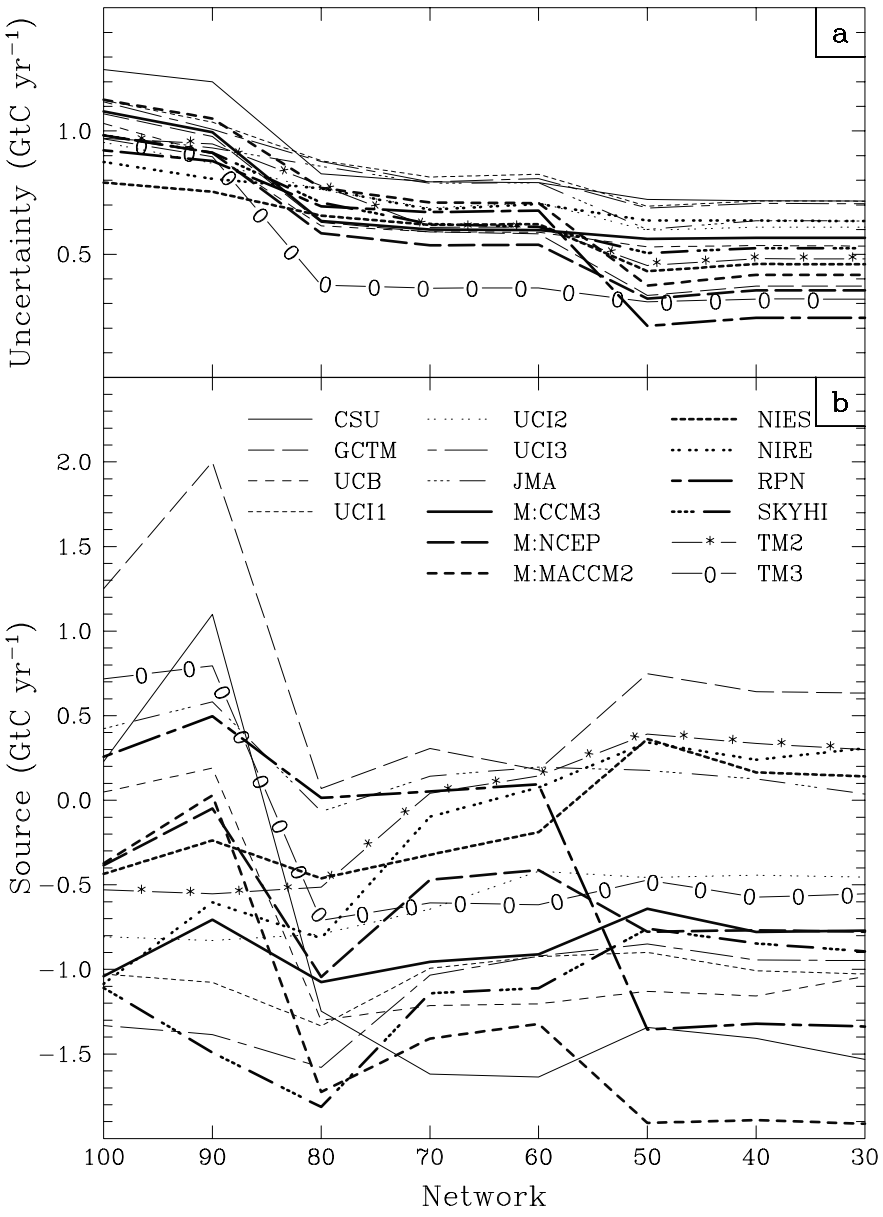

Figure 5. Source estimates (b) and uncertainties (a) for each model for temperate Asia for each of the ALL networks. The line identification is given in the key.

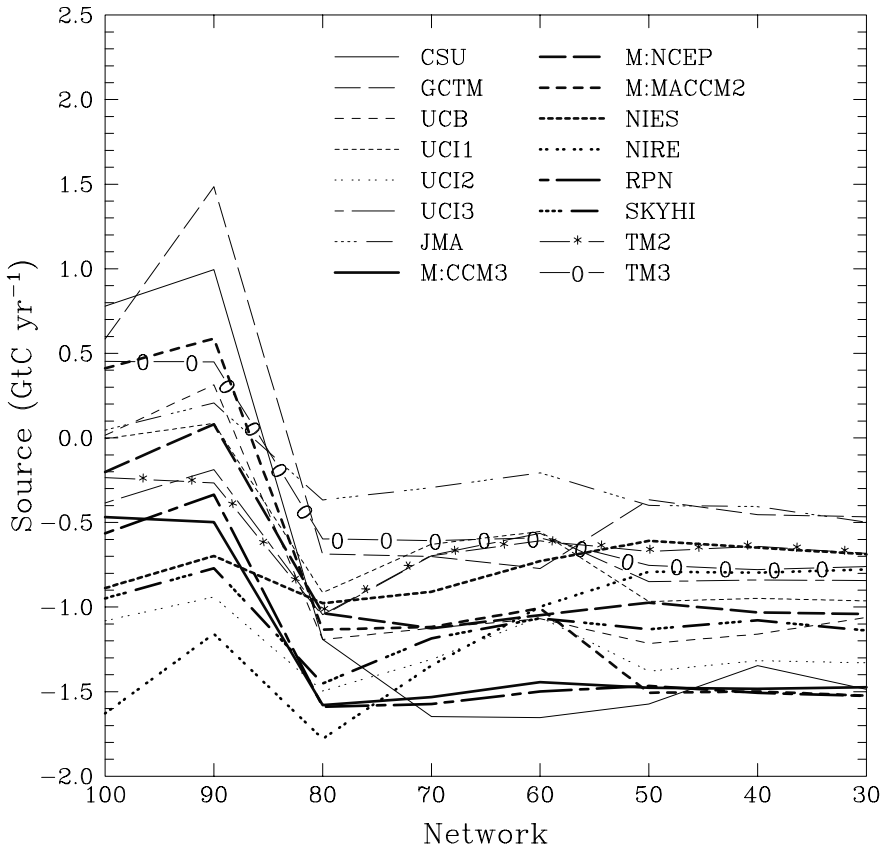

Figure 6. Source estimates for temperate Asia for each of the ALL networks for inversions calculated without the neutral biosphere annual mean response. The line identification for each model is given in the key.

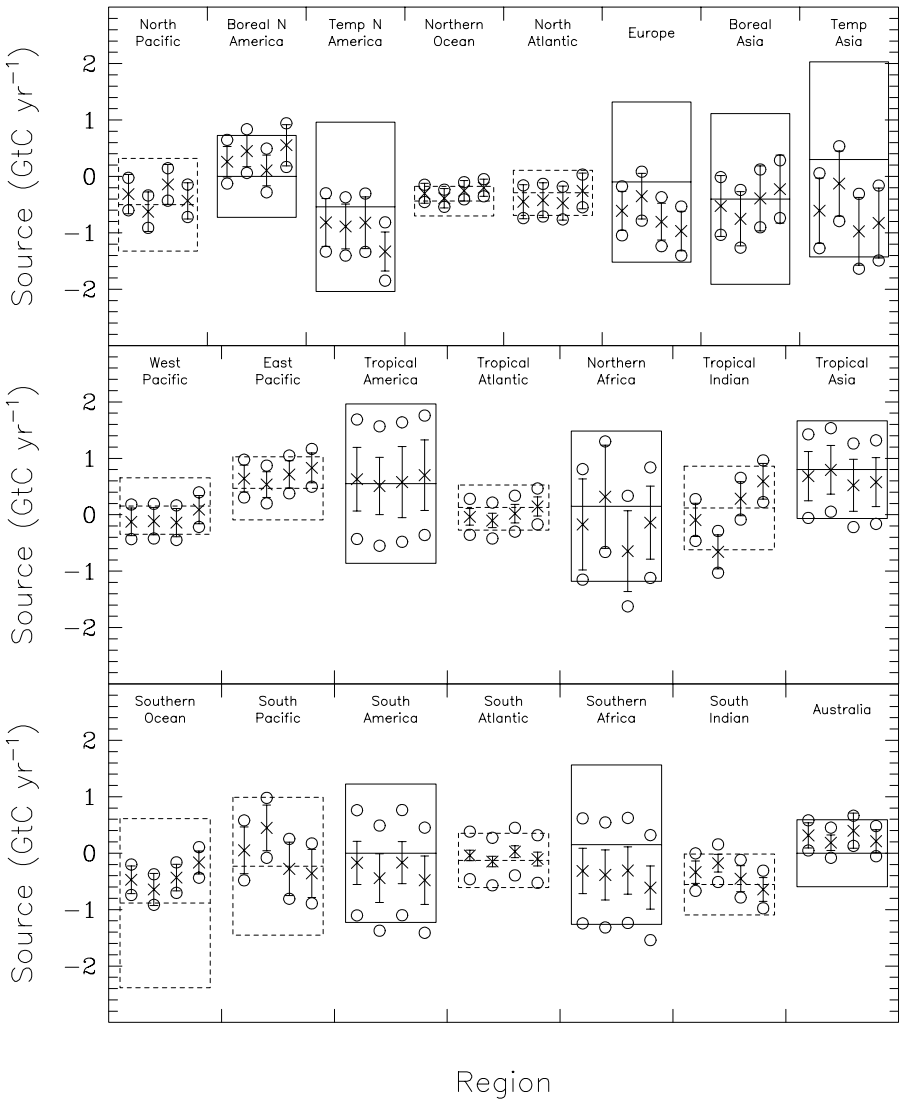

Figure 7. Model mean inversion results for all 22 basis regions using different data from different time periods. The leftmost value in each box is the control case 1992-1996, followed by 1995-1997, 1992-1994, and 1989-1991 respectively. Mean source estimates (including the background ocean fluxes) are indicated by the x, mean within-model uncertainty by the o and between-model uncertainty (the standard deviation across model source estimates) by the bar. For each land (solid line) and ocean (dashed line) region, the boxes represent the prior source estimate (central horizontal line) and prior uncertainty. Regions are arranged to approximate their north-south and east-west relationship.

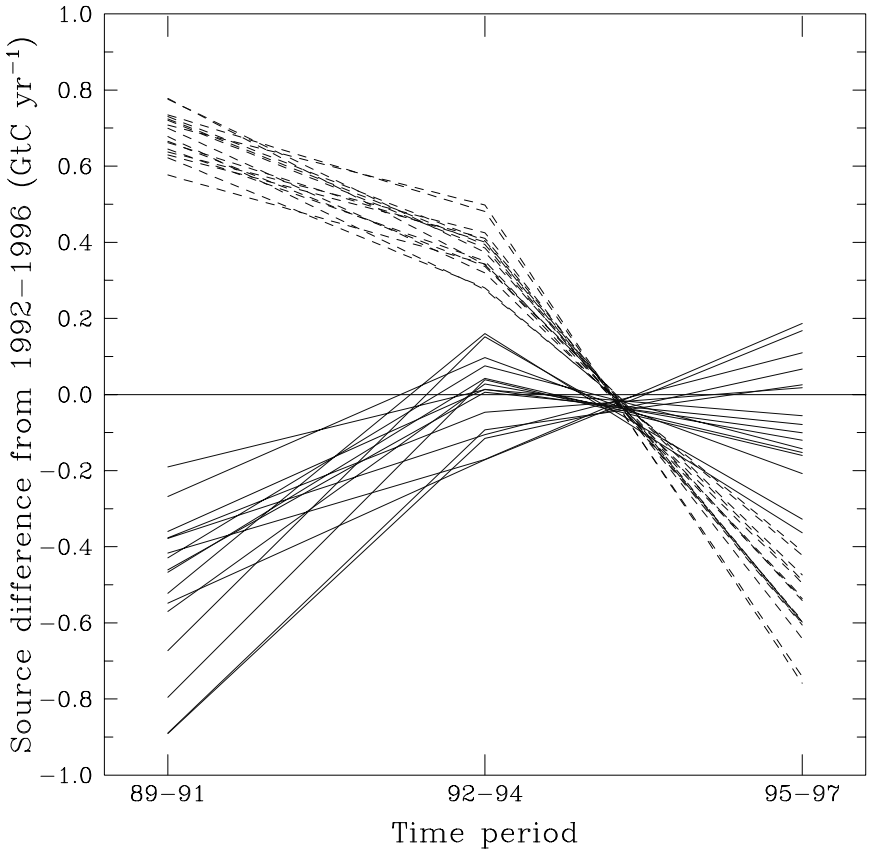

Figure 8. Inversion results for 16 models for the temperate North American (solid lines) and tropical Indian ocean (dashed lines) regions. The sources for the three time periods 1989-1991, 1992-1994 and 1995-1997 are plotted as deviations from the control source estimate for that model and region.

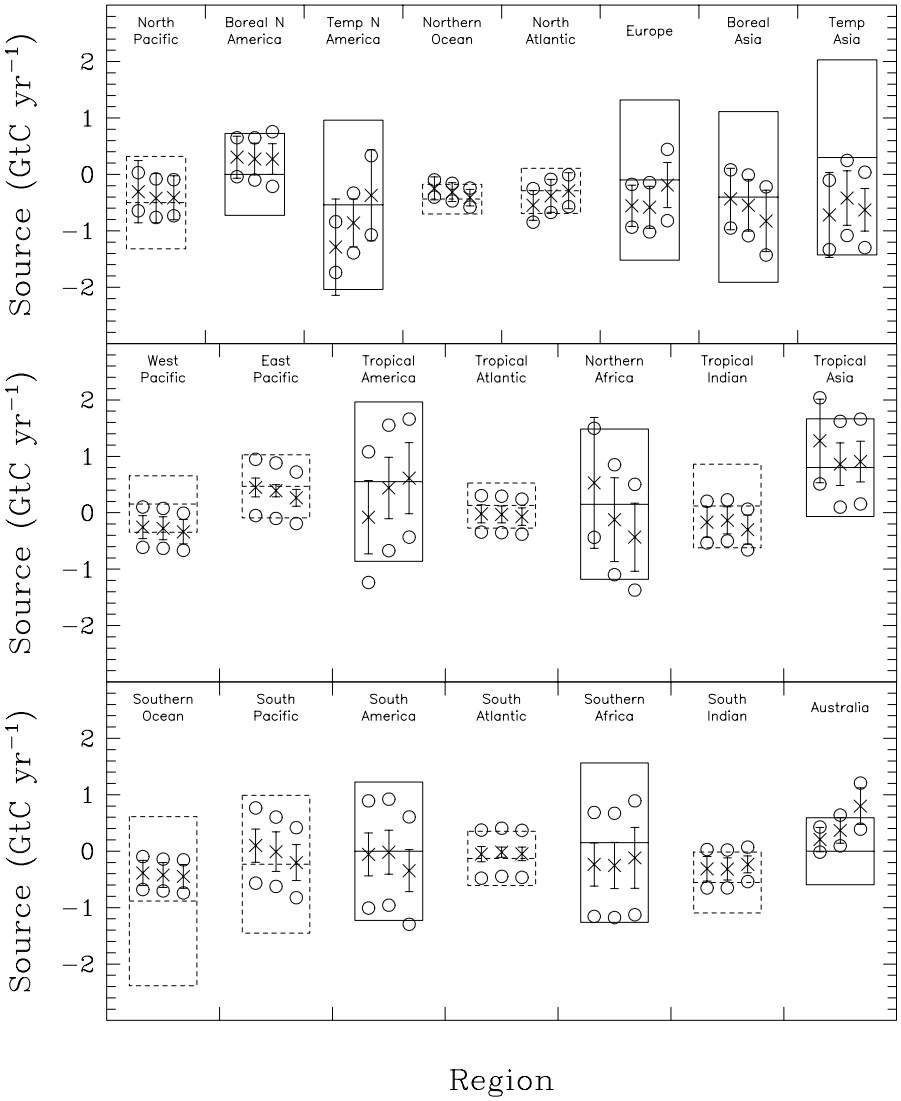

Figure 9. As Figure 7 but showing mean sources and uncertainties for three sets of inversions with different data selection strategies. Left is the 'nearest-grid' inversion, centre is the modified control inversion, right is the 'flask-like' inversion.

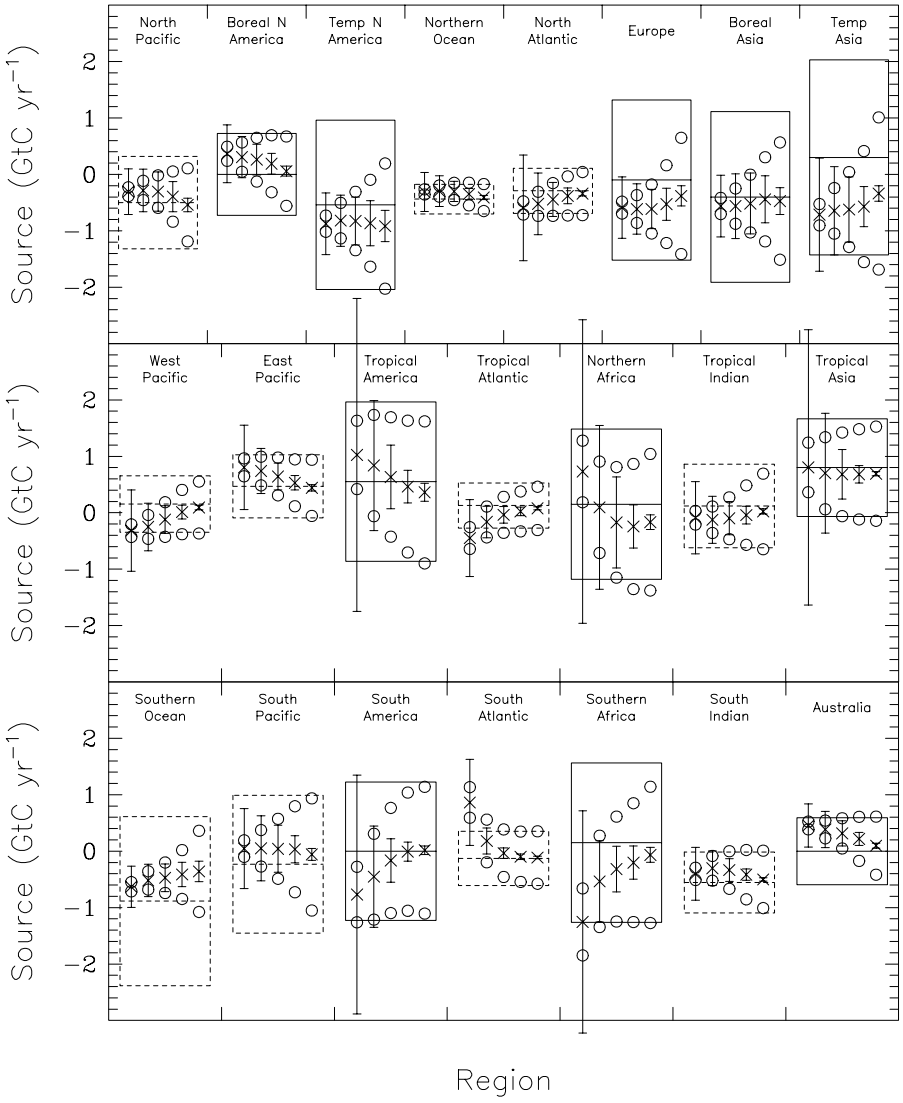

Figure 10. As Figure 7 but showing mean sources and uncertainties for five sets of inversions with scaled data uncertainties. The control data uncertainties were multiplied by 0.2 (far left), 0.5 (left), 2.0 (right), 5.0 (far right). The control inversion results are in the centre of each box.

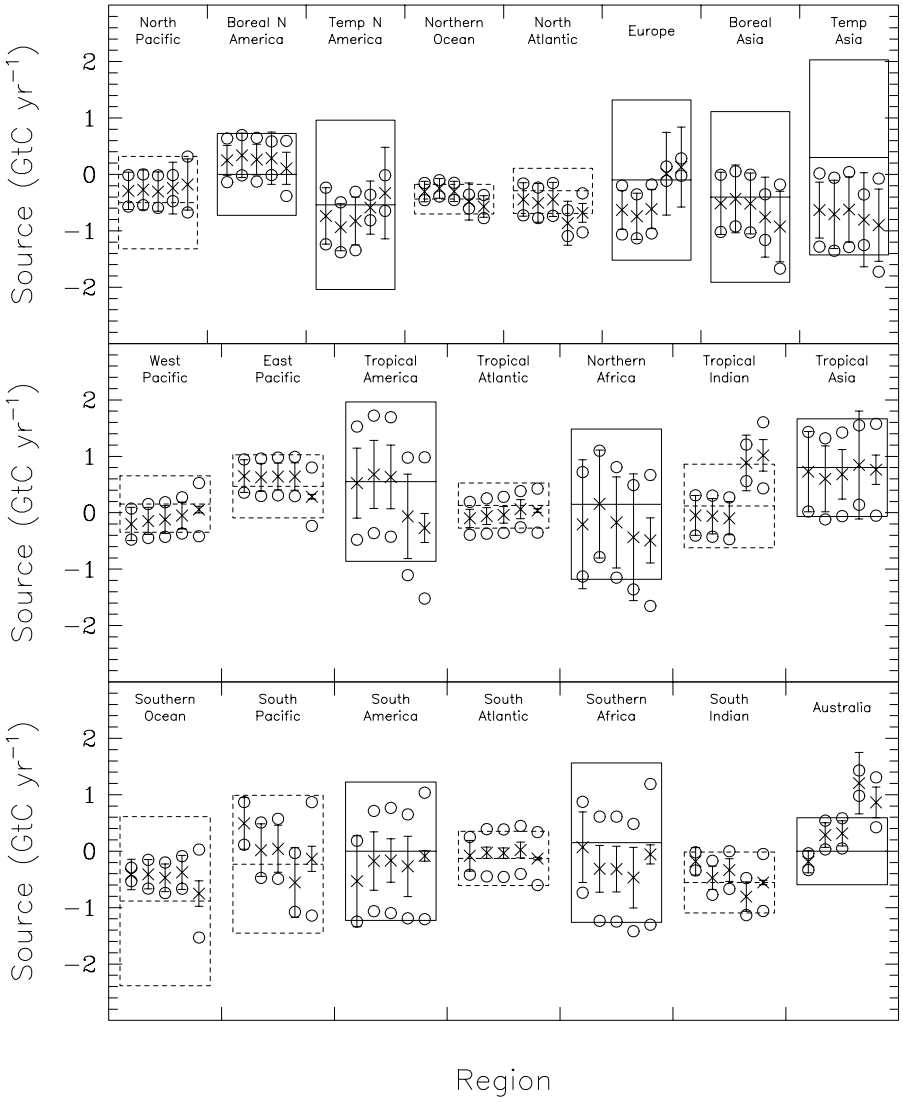

Figure 11. As Figure 7 but showing mean sources and uncertainties for five sets of inversions that used different spatial patterns of data uncertainty: no minimum (far left), no account of co-location (left), control (centre), constant 0.3 uncertainty (right), reversed distribution (far right).