JOURNAL OF GEOPHYSICAL RESEARCH, VOL. , NO. , PAGES 1–42,

Estimating the Distribution of Terrestrial CO_2 Sources and Sinks from Atmospheric Measurements: Sensitivity to Configuration of the Observation Network

P. Suntharalingam,[1] C. M. Spivakovsky,[1] J. A. Logan,[1] and M. B. McElroy[1]

[1]Department of Earth and Planetary

Sciences, Harvard University, Cambridge, MA,

USA.

Abstract. We explore the sensitivity of terrestrial CO2 flux estimates from a specific inversion methodology, based on the configuration of *Fan et al.* [1998], to different configurations of the global observation network (GLOBALVIEW-CO2). Using diagnostics derived from the inversion equations, we focus on quantifying the relative influence of individual stations on the flux estimates. We also examine the impact of different assumptions for the data uncertainty values by contrasting weighted and unweighted inversions and presenting related sensitivity analyses. For this particular methodology, unweighted estimates of continental scale fluxes prove very sensitive to network configuration. The inclusion or omission of a few important stations in and around the northern continents can result in shifts in continental-scale flux estimates of up to 1.5 Gt C/year. The weighted estimates are less sensitive to network configuration. Diagnostics of relative station influence indicate that this results from the reduced roles of previously influential continental sites; i.e., those stations characterized by high levels of data uncertainty. In the weighted approach, stations on continental peripheries associated with lower levels of data uncertainty are the most important in determining terrestrial fluxes. Finally, using the diagnostics of relative station influence, we discuss potential sampling strategies for the determination of regional fluxes from surface measurements.

1. Introduction

The magnitude and geographical distribution of uptake of anthropogenic CO_2 by the terrestrial biosphere and ocean remain key unresolved issues in current studies of the global carbon budget. Quantification of the budget for the '80s suggests that of the total anthropogenic emissions of 7.1 Gt C per year, 3.3 Gt C remained in the atmosphere, and 1.9 Gt C and 1.9 Gt C were taken up by the ocean and terrestrial biosphere respectively [*Intergovernmental Panel on Climate Change (IPCC)*, 2001]. The margins of uncertainty in these latter two estimates are relatively large; furthermore, the processes responsible for net terrestrial uptake are not as yet well defined. These processes are believed, however, to include forest regrowth, as well as enhanced plant growth due to CO_2 and nitrogen fertilization and climate variability [*Schimel*, 1995].

Over the past decade, identification of the large-scale distributions of net carbon uptake has been addressed by analysis of spatial gradients in atmospheric concentrations of CO_2 and other trace substances. *Tans et al.* [1990], one of the pioneering studies of such methods, employed measurements of the north-south gradient in atmospheric CO_2, in conjunction with 3-D tracer transport model simulations and constraints provided by surface oceanic CO_2 measurements, to infer the presence of a significant carbon sink in the northern hemispheric terrestrial biosphere. This conclusion was confirmed subsequently by analyses of gradients in O_2/N_2 [*Keeling et al.* 1996] and $^{13}CO_2$ [*Ciais et al.* 1995]. The O_2/N_2 data and the $^{13}CO_2$ measurements also provide important independent means to separate the relative contributions of oceanic and terrestrial uptake to the global budget of CO_2.

More recently, investigators have attempted to use analysis of atmospheric concentration gradients along with models of atmospheric transport, in a "top down" inverse modeling approach, to identify longitudinal variations in the northern hemispheric terrestrial sink. Estimates of long-term mean fluxes from these studies have not always been mutually consistent nor do they concur with alternative methods of estimating terrestrial carbon uptake [*Fan et al.* 1998, *Rayner et al.* 1999, *Bousquet et al.* 1999a]. The study of *Fan et al.* [1998], for example, which calculated the relative contributions of North America and Eurasia in the period 1988-92, employed analysis of annual mean concentrations and attributed the majority of net terrestrial uptake to the North American continent; i.e., approximately 1.7 Gt C per year out of a global sink of about 2.2 Gt C per year, an amount comparable to US fossil fuel emissions in the early '90s. This result is in sharp contrast to estimates of North American carbon uptake based on "bottom up" methods that rely on measurements from forest inventories [*Birdsey and Heath*, 1995] or estimates based on changes in land-management practices [*Houghton et al.* 1999]; these latter methods suggest that North American carbon uptake is less than 0.35 Gt C per year or only 10-30% of US fossil emissions for the period. A recent study by *Pacala et al.*, [2001], estimating US carbon uptake based on the latter approach and employing land carbon inventories, land use change estimates and ecosystem models, reported a somewhat larger range (0.37 to 0.71 Gt C per year net carbon flux); this increased estimate is ascribed to the inclusion of factors neglected in previous analyses, such as woody encroachment due to fire suppression.

Results from other recent "top down" inverse analyses of atmospheric concentration data have not been in agreement with the longitudinal partition of the northern mid-

latitude carbon sink implied by *Fan et al.* [1998], nor are they in strict accord with one another, highlighting the uncertainty attached to aspects of the analyses. The study of *Bousquet et al.* [1999a], which analyzed carbon uptake for the 1985-95 period, implies a significant sink in Eurasia, while the work of *Rayner et al.* [1999] suggests that the northern hemisphere terrestrial sink was approximately equally partitioned between North America and Eurasia during 1980-95. Postulated causes for the differences among these studies include variations in atmospheric transport models used, differences in time period examined, uncertainties in underlying assumptions about the carbon cycle, as well as differences in details of the inverse methodology employed.

A notable difference among the three inversion studies cited above is the use of different configurations of observation stations for atmospheric CO_2; i.e., 63 stations in the *Fan et al.* [1998] study, a maximum of 25 for *Rayner et al.* [1999] and 77 sites in *Bousquet et al.* [1999a]. Most CO_2 inversion analyses employ measurements from the flask network of the National Oceanic and Atmospheric Administration's Climate Monitoring and Diagnostics Laboratory (NOAA/CMDL) [e.g., *Conway et al.* 1994] or the extended GLOBALVIEW-CO2 network compiled by the same laboratory [e.g., *GLOBALVIEW-CO2* 1999, *Masarie and Tans*, 1995]. This latter database of over 80 stations represents a compilation of data from several international networks and provides the most widespread coverage for atmospheric CO_2 currently available. The majority of the stations (over 60 %) form part of the NOAA/CMDL global air sampling network mentioned above, which has been measuring CO_2 since the late '70s. A primary objective in the original design of the CMDL network was to measure "background" conditions characteristic of large spatial scales and long term trends; this resulted in the positioning of most of the sites in remote

oceanic locations, isolated consequently from local continental signals. Only over the past decade have stations been added to the network for the purpose of monitoring continental signals (e.g., stations on Shemya Island, Alaska (SHM), Tae-ahn Peninsula, Korea (TAP) and Qinghai Province, China (QPC), *Conway et al. [1994]*). Overall, however, the global CO_2 monitoring network includes relatively few stations near the continental signals of relevance for identifying the distribution of terrestrial carbon uptake.

Atmospheric CO_2 inversion studies of the past decade have employed a variety of different assumptions to characterize the data uncertainty associated with model-observation mismatch.These terms are complex to define, as they are required to represent not only instrument precision but also to capture the uncertainty in relating a model gridbox concentration to a station measurement. Recent analyses range from an "unweighted" situation (assuming a constant level of data uncertainty at all stations, e.g., *Enting et al.* [1995], *Fan et al.* [1998], *Rayner et al.* [1999]) to "weighted" scenarios that account for variable levels of measurement uncertainty among the station sites [e.g., *Bousquet et al.* 1999ab, *Peylin et al.* 1999, *Baker*, 2001, *Gurney et al.* 2002].

In this analysis we employ a version of the GISS II atmospheric tracer transport model in conjunction with a simple least squares minimization method, similar to the analysis of *Fan et al.* [1998], to investigate the sensitivity of North American and Eurasian terrestrial flux estimates to varying configurations of the observation network. We also focus on identifying the relative influences of individual stations on the inversion estimates and on analyzing results from different network configurations in terms of the contributions from the most influential stations. Our methodology is described in Section 2 and results are presented in Section 3. In view of the range of characterizations of the data uncertainty in

recent inversion studies, we also examine the impact of different assumptions about this quantity on the estimated fluxes.

In Section 4, using diagnostics developed to identify the location of the most influential stations in the inversion analysis, we discuss potentially suitable measurement locations for determining the longitudinal partition of net northern hemisphere terrestrial uptake. Section 5 discusses limitations of the methodology employed in this study, and Section 6 summarizes our conclusions.

2. Methodology

The analysis is based on methods of synthesis inversion, i.e., we seek the linear combination of sources and sinks such that the sum of their modeled concentrations best matches measured concentrations at the observation stations [e.g., *Enting et al.* 1995, *Tans et al.* 1990, *Fan et al.* 1998]. The magnitudes of net regional fluxes are estimated from a least squares minimization of the metric,

$$\chi^2 = \sum_{i=1}^{N} \frac{(yobs_i - ymod_i)^2}{\sigma_i^2} \tag{1}$$

i.e., the sum over N stations of the squared differences between modeled concentrations ($ymod_i$) and observations ($yobs_i$) weighted by the data uncertainty (σ_i). The modeled concentration at station i for a set of M sources is given by,

$$ymod_i = \sum_j b_j x_{ij} \tag{2}$$

x_{ij} represents the response of model concentrations at station i to a source j of unit magnitude, and b_j is the magnitude of this source to be determined. The set of responses to source j over the N stations characterizes the response function $X_j = [x_{1j} x_{2j} \cdots x_{Nj}]^T$.

Solution of the least squares minimization via the normal equations yields regional flux estimates $\mathbf{b}$ (in vector form), and the associated variance-covariance matrix $\mathbf{C_b}$ defining the error estimates [e.g., *Myers*, 1990].

$$\mathbf{b} = (\mathbf{X^T R_0^2 X})^{-1} \mathbf{X^T R_0^2 y_{obs}} \tag{3}$$

$$\mathbf{C_b} = (\mathbf{X^T R_0^2 X})^{-1} \tag{4}$$

Here, $\mathbf{X}$ (dimension N $\times$ M) is the composite matrix of response functions X_j, and $\mathbf{R_0}$ (dimension N $\times$ N), a diagonal matrix of elements $[1/\sigma_i]$, characterizing the influence of data uncertainty (i.e., we assume the measurement errors at the stations are uncorrelated).

Since this study focuses primarily on the sensitivity of flux estimates to network configuration, we have employed only a single inversion methodology similar to that used in *Fan et al.* [1998]; namely, an investigation of long-term mean fluxes (for the period 1988 - 1992). A comparison of inversion methods is beyond the scope of this study; some problems associated with this methodology will be discussed, however, in section 5.

2.1. Observational Data

The observational CO_2 data used in this study are taken from GLOBALVIEW-CO2 [1999]. An aim of the GLOBALVIEW-CO2 database is to provide temporally continuous records of measurements for use in analyses of sources and sinks of atmospheric CO_2. *Masarie and Tans*, [1995] provide a discussion of the integration of measurements and methodology of data extension techniques involved. We have constructed a range of networks using the measurement sites of *GLOBALVIEW-CO2* [1999], and the main ones discussed in this study are listed in Table 1. Two of the networks are based on those used in recent analyses; i.e., F1 [*Fan et al.* 1998] and R1 (the "extended" network of *Rayner*

et al., [1999]). We also examine results from two more general networks: CMDL-ANY, a network consisting of NOAA/CMDL flask sites that reported any actual measurements (as opposed to values obtained from extension methods) during the 1988-1992 period; and GV-ANY, a network made up of any sites from the *GLOBALVIEW-CO2* [1999] database that reported actual measurements in 1988-1992. This latter network, GV-ANY, is, therefore, the largest network of the study (75 stations). The remaining networks are based on subsets of this station set, with the exception of the F1 network which also includes Station P, a Canadian fixed shipboard station at which measurements ceased in 1981. Since no actual observations are reported from this site in the period 1988-1992, it is not included in the GV-ANY network.

The majority of the stations in the *GLOBALVIEW-CO2* [1999] database report flask measurements sampled at approximately weekly frequency. Some sites, however, also report much higher frequency measurements, for example, based on hourly in-situ sampling. The criterion used to select among multiple measurements at a site was that if a flask measurement was available, it was used in preference to other sampling methods to maintain consistency with the remainder of the network. This becomes an issue at the following sites: AMS, BRW, ITN, IZO, MHD, MLO, PRS, SMO and SPO. Both the GV-ANY and F1 networks include a few stations that report only high-frequency measurements and we examine implications of the inclusion of these data in the section on weighted estimates (3.2). Unless explicitly stated in the analysis below, the modeled concentrations used in this study do not attempt to reproduce sampling practices at station sites (e.g., sampling based on wind-sector). Sensitivity tests were conducted at certain coastal and continental sites by shifting the location of the model gridbox sampled to reduce the impact of local

emissions or better capture signals of marine air. While the majority of flux estimates were relatively unaffected by this procedure, cases for which it made a significant difference are noted in section 3.1.1. Finally, we note that as with the *Fan et al.* [1998] analysis, the modeled and observed CO_2 concentrations employed here are those in excess of the value at the South Pole station (SPO); a possible problem with this approach is discussed in section 5.

2.2. Transport Model, Emissions Fields and Regional Aggregation

The atmospheric tracer transport model used is the GISS II Chemical Transport Model (CTM) which is based on the meteorological fields of the general circulation model of *Hansen et al.* [1983]. Interhemispheric exchange was investigated using observations of CFCs and ^{85}Kr (*Prather et al.* [1987], *Jacob et al.* [1987]). As noted in these studies, a parameterization of sub-grid scale diffusion was introduced for consistency with observed interhemispheric gradients. Convective and synoptic transport was examined using ^{222}Rn [*Jacob et al.* 1990, *Balkanski et al.* 1992]. The GISS II CTM has been used extensively for studies of tropospheric chemistry [e.g., *Spivakovsky et al.* 1990,2000, *Jacob et al.* 1993, *Chin et al.* 1996, *Horowitz et al.* 1998, *Wang et al.* 1998]. The model employed in this analysis has a horizontal resolution of 4^o latitude $\times$ 5^o longitude, with 9 layers in the vertical. The lowest three layers, defining the planetary boundary layer, have tops at approximately 500 m, 1200 m and 2600 m respectively. The CTM is run on a 4 hour timestep using a 1 year archive of GISS GCM meteorological fields which include 4 hour averages of winds, mixed layer heights and column frequencies of wet and dry convection, and 5 day averages of temperature, humidity and convective mass fluxes. A detailed description of the computation of mass transport is presented in *Prather et al.* [1987].

Earlier versions of this model were used in studies of CO_2 by *Tans et al.* [1990], *Enting et al.* [1995] and *Rayner et al.* [1999], in which the model was run at a lower horizontal resolution ($8^o \times 10^o$) with different parameterizations adopted to describe convective mixing and advection.

The magnitude and distribution of emissions for fossil fuel, net oceanic uptake and seasonal biospheric exchange (assumed balanced in the annual mean) are prescribed according to the following sources (annual averages for the period 1988 - 1992) : (a) fossil fuel (*Andres et al.* [1995]), (b) net oceanic uptake (*Takahashi et al.* [1997] (In the absence of information of oceanic uptake over this five year period, we use the 1990 estimate of oceanic uptake from this source.), (c) annually balanced seasonal biospheric exchange fluxes from the CASA model [*Potter et al.* 1993]. We have not included the source of CO_2 from atmospheric CO oxidation [e.g., *Enting and Mansbridge*, 1991] in this analysis, and note that a recent detailed inversion analysis evaluating this source concluded that it showed a relatively small impact on estimates for surface CO_2 fluxes [*Baker*, 2001].

The global annually averaged magnitudes of the prescribed CO_2 fluxes used in the model are (a) 6.05 Gt C for fossil fuel, (b) 1.2 Gt C for oceanic uptake (We note that a larger oceanic uptake has been estimated in the more recent study of Takahashi et al. [1999]; this latter analysis includes new measurements in the Southern Ocean and certain variations in the gas-exchange parameterization), and (c) 58 Gt C for net primary production from the CASA model. In addition, the global budget is constrained using data on the atmospheric CO_2 concentration increase over the period 1988 - 1992; i.e., an equivalent annual atmospheric increase of 2.76 Gt C [*Conway et al.* 1994]. Figure 1 depicts the latitudinal distribution of modeled concentrations for the prescribed emissions (i.e.,

without the net terrestrial uptake component) at the NOAA-CMDL sites and compares this with the observations for the period 1988 - 1992 (this model simulation will be referred to as FBO in the text). As can be seen, modeled concentrations are markedly higher than observational data at northern mid-latitudes; this constitutes the primary feature of the results leading to the inference of a significant northern mid-latitude sink [e.g., *Tans et al.* 1990].

As noted by *Denning et al.* [1995], an issue of possible significance for annual mean inversions of the type discussed here is the ability of the model to represent successfully the covariance of transport (e.g. seasonally varying vertical mixing) and biospheric exchange (e.g., seasonal variation in magnitudes of photosynthesis and respiration). This covariance is often referred to as the "atmospheric rectifier" [e.g., *Heimann et al.* 1989, *Denning et al.* 1996]. An inter-model comparison of this effect (TRANSCOM I, *Law et al.* [1996]) noted that models fell into two groups (weak vs. strong rectifiers) primarily on the basis of their treatment of vertical mixing in the boundary layer. The version of the GISS CTM used in this study produces values of less than 1 ppm for the zonally averaged seasonal rectifier in the northern mid-latitudes; i.e., the annual mean zonally averaged surface concentration for seasonal biospheric emissions alone, using either the seasonal biospheric exchange fluxes of the CASA model [*Potter et al.* 1993] or those of [*Fung et al.* 1987] (c.f., Figure 12 of *Law et al.* [1996]). The model falls therefore into the category of "weak gradient" models discussed by *Law et al.* [1996]. Since estimates of the northern mid-latitude sink from annual-mean inversions are based on the mismatch between modeled and observed concentrations, the strength of the mid-latitude rectifier has a direct impact on the magnitude of the uptake inferred for the local terrestrial sink; models with weak

rectifiers yield smaller estimates for mid-latitude terrestrial uptake. Implications for our analysis are discussed in section 3.1.

Consistent with the study by *Fan et al.* [1998], we examine the relative distribution of net terrestrial biospheric uptake over three large regions, (1) North America, (2) Eurasia (both northward of 20^o N) and (3) the tropical and southern hemisphere continents. We note, however, that some studies advocate a higher regional resolution in the initial inversion process, followed by aggregation up to continental scales post-inversion, due to possible biases introduced in the low resolution solution [e.g., *Enting et al.* 1995, *Peylin et al.* 1999, *Kaminski et al.* 2001]. Since our analysis is focused on the sensitivity of flux estimates to different combinations of observation stations using a specific inversion methodology, we have not examined the impact of higher regional resolution. We do, however, discuss the implications of low regional resolution in our analysis in Section 3.1 and Section 5.

3. Results

3.1. Unweighted case : Homogeneous variance of measurement errors

We examine first the sensitivity of regional flux estimates to different configurations of the observation network, under the assumption that measurement errors are independent and normally distributed with constant variance, as has been done in some recent studies (i.e., $\sigma_i = \sigma_0$ for all stations) [e.g., *Fan et al.* 1998, *Rayner et al.* 1999]. Note that under this assumption equations 3 and 4 imply that the flux estimates b_j are independent of measurement error, while the flux error estimates given by the elements of $\mathbf{C_b}$ vary in direct proportion to σ_0.

It is difficult to quantify the data uncertainty, σ_i, as it must encompass not only measurement instrumentation error but also the uncertainty in relating a model gridbox spatial- and time-averaged concentration and its associated systematic errors to the local measurements. *Gloor et al.* [1999] note, for example, that model associated systematic errors can be categorized in terms of (a) possible flaws in representations of transport, and (b) errors in the emissions distributions used to simulate the modeled response functions. The majority of the observational sites used in this study employ weekly flask measurements and some follow specific sampling protocols (e.g., based on wind direction, as in the case of many coastal sites [*Conway et al.* 1994]). Such sampling may represent a level of synoptic or smaller scale variability not captured by a model grid box average; however, the low frequency of sampling may also result in an under-representation of synoptic scale variability.

The impacts of different formulations of measurement uncertainty on flux estimates are examined further in section 3.2, where we assume heterogeneous variance of the measurement errors and present sensitivity analyses on aspects of this issue. In this section we note that previous studies (using a constant σ_0) attempting to account for such considerations as model-related systematic errors have employed such values as of $\sigma_0 = 0.6$ ppm [*Fan et al.* 1998] and $\sigma_0 = 0.85$ ppm [*Rayner et al.* 1999]. In contrast, round-robin intercomparisons by the laboratories that contribute measurements to GLOBALVIEW-CO2 report agreement of measurements of standard gases to within 0.2 ppm [*Peterson et al.* 1999]. For the results discussed in this section, we use a value of $\sigma_0 = 0.7$ ppm for the main discussion but also present results for assumptions of σ_0 varying from 0.2 to 1 ppm.

In order to evaluate the large-scale resolution of terrestrial uptake permissible in this model, we compare, in Figure 2, the estimates of North American and Eurasian net flux and the total northern mid-latitude terrestrial uptake for the observation networks listed in Table 1. Figure 2 also shows the variation in estimated flux errors under different assumptions for σ_0; note that the estimates for $\mathbf{b}$ do not change for such variation (equation 3) and that the flux error scales in direct proportion to σ_0.

As seen, estimates of the total northern mid-latitude terrestrial sink are well constrained by the observations; derived flux values of 1.1 to 1.35 Gt C per year are relatively robust to observation network configuration, differing by less than 0.3 Gt C (a result which concurs with previous studies, e.g., *Tans et al.* [1990]). In contrast, the partitioning of the sink into longitudinally distinct zones is not so well constrained. Estimates of North American net fluxes, for example, vary considerably depending on the observation network employed; these fluxes range from uptakes of 1.5-1.6 Gt C per year (for network F1 and a variant of CMDL-ANY), to efflux of -0.3 Gt C per year (for network CMDL-ANY omitting the South China Sea (SCS) stations). This sensitivity can be ascribed to the following factors. Atmospheric mixing in the east-west direction is rapid (on the order of weeks, in comparison to about a year for interhemispheric mixing [*Warneck*, 1988]) and produces weak concentration signatures from relatively strong flux gradients. In the inversion procedure, the deduced flux gradient, is, therefore, very sensitive to the concentration gradient.

The difficulty in extracting information on the longitudinal variation of the sink is exacerbated by the sparsity of measurements in regions sensitive to the terrestrial biospheric signal. Figure 3 depicts the difference between two separate model simulations of atmo-

spheric CO_2: NA1, in which the entire northern mid-latitude terrestrial sink is situated in North America, and EA1, in which it is located in Eurasia. Large magnitudes of the difference (EA1 - NA1) indicate regions where fluxes from the separate continents can be well distinguished. Levels close to zero indicate locations where the signals from both North America and Eurasia are small, or where the signals from the two regions are indistinguishable. Figure 3 indicates that although the largest signals are found within and closest to the continents, the majority of the stations are sited in remote oceanic locations (due to the original NOAA/CMDL objective of measuring background conditions), and hence are ill-suited for detection of continental signals. Estimates of the longitudinal partition of terrestrial fluxes are determined primarily, therefore, by the small number of stations in and around the northern hemisphere continents. As discussed below, the modeled concentrations at some of these stations may be subject to such errors as problems in model transport; the disparity in estimates of North American uptake seen in Figure 2 arises, therefore, when inversion estimates are derived from different subsets of these critical stations.

We can obtain a more quantitative measure of the relative influence of the individual stations by examining the defining equation for the estimated fluxes, $\mathbf{b}$, in the form

$$\mathbf{b} = \mathbf{Q}\mathbf{y_{obs}} \tag{5}$$

where

$$\mathbf{Q} = (\mathbf{X}^T\mathbf{R_o}^2\mathbf{X})^{-1}\mathbf{X}^T\mathbf{R_o}^2 \tag{6}$$

For assumptions of constant σ_i, this reduces to $\mathbf{Q} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, where $\mathbf{Q}$ has dimensions of M $\times$ N.

Matrix **Q** can be considered an "influence" matrix, since its elements q_{ji} define the impact of the concentration measurement at station i (i.e., y_i), on flux estimate b_j. The units of its elements are, accordingly, (Gt C per year)/ppm. In Figure 4 we plot the magnitude of this diagnostic (i.e., q_{ji}) at the various stations of the four main networks (i.e., GV-ANY, CMDL-ANY, F1 and R1) to illustrate their relative influence on the estimate of net North American terrestrial flux (i.e., $b_{NorthAmerica}$).

Note that the most influential stations are those closest to the continental signals in North America and Eurasia and that these sites are not always common to the four networks. As Figures 4a and 4b indicate, the CMDL-ANY and GV-ANY networks include similar combinations of the most influential stations and yield similar unweighted flux estimates for North American uptake, as shown in Figure 2. The F1 network (Figure 4c), in comparison, does not include some of these important stations (e.g., ITN, UUM, and TAP), and North American uptake estimates from this network are significantly higher, as seen in Figure 2. We also depict, in Figure 2, the flux estimate from a CMDL-ANY network that excludes stations ITN, UUM and TAP; as can be seen this result differs by more than 1.2 Gt C from the original CMDL-ANY estimate.

The much larger apparent influence of the individual stations in the R1 network in comparison, for example, to the GV-ANY network is simply a function of the number of stations in the network. Since the R1 network has only 25 stations in comparison to the 75 of GV-ANY, its individual stations are required to exert more influence per ppm of CO_2 observed than the counterparts in GV-ANY.

Another method of characterizing the influence of individual stations on inversion estimates is obtained through a process of column deletion for matrix **Q**. Figure 5 summarizes

the variation in the estimate of the North American terrestrial uptake obtained using the CMDL-ANY network, as individual stations (or combination of stations for the case of the South China Sea data, SCS) are omitted from the analysis. The value associated with each station represents the estimate from the remainder of the network upon elimination of the station, and only the results for the 25 most influential stations are shown in the figure. This particular diagnostic of the influence of individual observations on the inversion estimates is formally defined as DFBETAS; it is related to the index of influence q_{ji} of equation 6 via the studentized residual of the omitted data point [e.g., *Belsley* 1980, *Myers* 1990]. The estimates are seen to be most sensitive to eliminations of stations SCS, ITN, TAP, UUM, and BAL; i.e., leading to conclusions similar to those obtained using the q_{ji} diagnostic discussed above.

3.1.1. Discussion of Specific Stations. We now examine the influence of specific stations in the context of characteristics of the atmospheric transport model and the observational dataset. We examine their impact on derived flux estimates and will rely primarily on the CMDL-ANY network as an illustrative example.

The analysis of Section 3.1 indicates that the estimate of North American uptake, derived using network CMDL-ANY, is extremely sensitive to inclusion or omission of the data from the South China Sea stations; i.e., shifting from a net sink of 0.45 Gt C year^{-1} to a net source of 0.33 Gt C year^{-1}. This sensitivity to the SCS stations is also apparent when using other networks; for example, Figure 2 depicts results for the F1 network excluding the SCS observations. In this case, North American uptake estimates drop from 1.5 Gt C year^{-1} to levels close to zero.

As will be discussed in detail below, part of the sensitivity to these stations may be attributed to (a) a poor simulation of transport off the continent in the South China Sea region by the GISS II CTM, (b) an overrepresentation of this region in the observational data using the unweighted estimation procedure, and (c) the correlation between North American and Eurasian flux estimates in the three terrestrial region partition used in the inversion analysis. As seen from Figure 1, which contrasts annual mean concentrations from the model (excluding net terrestrial uptake) with the observations, for the majority of Northern Hemisphere stations, model values are higher than the observations (the major cause being the lack of the net terrestrial sink component in the forward simulation shown). The South China Sea stations are, however, an exception to this pattern; modeled values range from 2.4 to 5.1 ppm, in comparison to observations of 3.0 to 5.8 ppm, suggesting either that the observations may be registering a local CO_2 signal not captured in the model emissions field, or that model transport may not adequately reproduce the offshore flow in this region. A comparison of the monthly averaged winds at the 1000mb level from the GISS II CTM with streamlines from the NCEP meteorology suggests that the latter may be true. In December, for example, (one of the months of highest flow off the continent), the SCS locations in the model experience a predominantly easterly flow carrying a low oceanic CO_2 signal in contrast to the higher concentrations transported by the observed north-easterly flow from the continent.

In solving for regional fluxes, the higher levels of observed relative to modeled concentrations at the SCS stations, coupled with the paucity of other observational sites around Eurasia, drives the solution towards a Eurasian source. The North American flux is affected in turn, as there is a strong constraint on the total mid-latitude uptake imposed by

the majority of northern hemisphere stations (e.g., see Figure 1). In the simple three terrestrial region partition used in this inversion method, this forces North American uptake to adjust in compensation. Furthermore, since there are 7 SCS stations exhibiting approximately the same concentration difference between observation and model, this constraint is overrepresented in the unweighted estimate; e.g., a variant of the CMDL-ANY network with only a single SCS station yields North American and Eurasian fluxes of -0.08 $\pm$ 0.6 Gt C per year (i.e., small source) and 1.4 $\pm$ 0.7 Gt C per year (a large sink) respectively.

The sensitivity of the continental scale estimates to the observations from the SCS stations may also be attributed to the formulation of the inversion problem with only three large terrestrial regions. Errors in modeled or observational concentrations at a few stations (as in the case of the SCS stations) can cause the flux estimate of an entire continental region to respond. *Bousquet et al.* [1999a], for example, whose inverse analysis used 6 separate regions in Eurasia, found a significant local sensitivity to the elimination of the SCS stations; i.e., their estimate of the tropical Asian source was reduced from 0.8 $\pm$ 0.4 Gt C yr^{-1} to 0.2 $\pm$ 0.4 Gt C yr^{-1}. However, in view of their higher regional resolution within Eurasia, flux estimates in Europe and Northern Asia were less affected.

As shown in Figure 2, a version of the CMDL-ANY network that excludes ITN, UUM, and BAL results in a shift in the North American sink from 0.45 Gt C per year to 1.6 Gt C per year, a level comparable to the result of the F1 network. Station ITN in North Carolina (flask sample) is one of the few sites in continental North America and constitutes the single most influential station in network CMDL-ANY; the elimination of observations from this station increases the estimate of North American uptake by more then 0.4 Gt C per year. For the majority of the northern mid-latitude sites, modeled

values exceed observations by about 2 ppm; in contrast, modeled annual mean CO_2 at ITN is greater than the observations by less than 1 ppm (5.4 ppm vs. 4.5 ppm). The inclusion of this site in the inversion, therefore, forces a requirement of a relatively low value for the North American sink. Eliminating station ITN from the analysis allows North American uptake to increase significantly. As noted in section 2.2, the GISS II CTM has a small modeled rectifier effect. If this effect is underestimated in the region of ITN, it could be a contributory factor in the low modeled value and the resulting low estimate of North American uptake.

Measurements at BAL (Baltic Sea site) are made aboard a ferry (midway between Poland and Sweden) in the southern Baltic Sea. The gridbox corresponding to this location in the transport model has a large land fraction and is subject to significant local fossil and biospheric emissions. The sampling location in the model does not, therefore, provide a reliable representation of observed sampling conditions; e.g., back trajectory analysis at BAL for the period 1994 - 1996 indicates that while less than 20% of the trajectories originate in the local regions just south of the site (and hence are influenced by western European emissions) the majority come in from the north and originate in the oceanic regions of the North Atlantic and Arctic (e.g., refer to http://www.cmdl.noaa.gov/ccgg/flask/bal.html). Modeled results for annual mean concentrations for the FBO simulation at BAL are almost 3 ppm higher than the observational values. In order to reduce the influence of local emissions on modeled concentrations at this site, we moved the sampling location one gridbox north-east into the Baltic Sea. This resulted in concentrations more than 1.5 ppm lower than results obtained at the previous

'on-shore' location; our estimate for the Eurasian sink dropped correspondingly by over 0.3 Gt C per year.

In summary, we note that under the assumption of constant levels of measurement uncertainty, this inversion method yields a longitudinal partition of terrestrial flux estimates that is very sensitive to the inclusion or omission of a handful of stations in and around the continental regions of the northern hemisphere. Since the locations of some of these stations may be subject to model-related errors, North American and Eurasian flux estimates based on varying combinations of the most influential stations show shifts of up to 1.5 Gt C per year.

3.2. Weighted case : Heterogeneous variance of measurement errors

In this section we explore the effect of incorporating varying levels of data uncertainty at the observation sites when solving for regional fluxes. To account for the heterogeneous variance of data uncertainty, we employ a weighted least squares estimator instead of the ordinary least-squares estimator of section 3.1. The data uncertainty matrix, $\mathbf{R_0}$, accounts now for the differing levels of uncertainty at the measurement sites [e.g., *Myers* 1990]; thus (assuming again that errors are uncorrelated),

$$\mathbf{R_0} = \mathrm{diag}[\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, ..., \frac{1}{\sigma_N}] \tag{7}$$

Estimates of regional fluxes, $\mathbf{b}$ and the variance-covariance matrix $\mathbf{C_b}$, are given, as before, by equations 3 and 4, incorporating the new definition of $\mathbf{R_0}$. It can be seen that in this case, when $\mathbf{R_0}$ is no longer equal to $(1/\sigma_0)\mathbf{I}$, the flux estimates, $\mathbf{b}$, are now a function of $\mathbf{R_0}$ and hence of the various σ_i, and that the error estimates for fluxes (derived from the diagonal elements of $\mathbf{C_b}$) no longer have a simple linear dependence on

σ_0, as they do in the unweighted scenario. In the section below, we discuss the use of the standard error of the mean measurement as a possible index to characterize the range of data uncertainty among observation sites.

3.2.1. Data uncertainty via standard errors of the measurements. The standard error of a set of data is an indication of the accuracy with which the sample mean estimates the population or true mean [e.g., *Press et al.* 1995]; its magnitude, therefore, provides a measure of the level of statistical uncertainty of the measurement. It is calculated as the residual standard deviation divided by the square root of the number of data points (i.e., $rsd/\sqrt{N}$). For the annual means of the NOAA/CMDL data, the residual standard deviation is derived from the deviation of the individual flask measurements from a fitted curve incorporating a time trend and seasonal components [*Conway et al.* 1994]. The networks in our study, however, include several stations that do not form part of the NOAA/CMDL network; we therefore constructed a proxy for the standard error based on information provided by the *GLOBALVIEW-CO2* [1999] database on residual standard deviation and annual numbers of observations at each site. A comparison of our estimated standard errors with those calculated for the NOAA/CMDL flask measurements for the period 1988-1992 (T. Conway, pers. comm) indicated that, in almost all cases, agreement was to better than 0.01 ppm.

As shown in Figure 6, for network CMDL-ANY, the magnitude of the estimated standard error varies by more than a factor of 15 among stations; highest values (greater than 0.5 ppm, e.g., BAL) are associated with northern hemisphere continental and coastal stations that are subject to a large degree of variability from local biospheric and fossil emissions or meteorology. Smallest values (less than 0.04 ppm, e.g., MLO, CGO) are

found at remote ocean sites and in the southern hemisphere, and can be ascribed to the lower levels of emissions and meteorological variability in these regions.

Using this distribution of the standard error to represent the data uncertainty at stations, we derive the weighted least squares estimates of the flux coefficients **b** for the various networks discussed in section 3.1. The results for the estimate of North American uptake are contrasted in Figure 7 with the unweighted estimates of section 3.1. An additional adjustment made to the standard error based data uncertainties employed in the weighted estimates of Figure 7 is the imposition of a minimum level of uncertainty of 0.1 ppm. This limit is set because it is unlikely that the data uncertainty, which is required to account for model-observation mismatch, can be characterized to such levels of accuracy as as 0.03 - 0.05 ppm, as is the case for such remote sites as MLO. Since this is an arbitrary threshold, the sensitivity of estimates to this and other characterizations of the measurement uncertainty is examined in section 3.2.3.

As seen in Figure 7, the weighted estimates are less sensitive to different configurations of the network; in particular, the high estimates for North American uptake are considerably reduced (e.g., to 0.35 ± 0.2 Gt C per year for the F1 network). Estimates obtained using different networks vary by about 0.8 Gt C per year, in contrast to over 1.5 Gt C per year for the unweighted scenarios. As discussed in Section 3.1, the large variation among unweighted estimates from different networks arose from the inclusion or omission of a few important stations in or around the northern hemisphere continents. In the case of the weighted estimates, the roles of many of these previously influential stations are reduced, since they are often characterized by higher levels of data uncertainty.

3.2.2. Influence of individual stations. In Figure 8 we illustrate the change in influence of individual stations on North American terrestrial flux estimates for unweighted vs. weighted estimates (network CMDL-ANY). Note that from equation 6, under the assumption of variations in σ_i among the stations, the influence matrix $\mathbf{Q}$ is no longer independent of σ_i; i.e., the impact of individual sites on the flux estimates $\mathbf{b}$ depend now on the local level of measurement uncertainty. In contrast, for the unweighted scenarios $\mathbf{Q}$ is independent of σ_i. The change in the influence of individual stations, illustrated in Figure 8, is defined as the the change in q_{ji} between the unweighted and weighted estimates; i.e., the elements Δq_{ji} that arise upon the incorporation of varying σ_i ($\Delta q_{ji} = \mathrm{q_{ji_{weighted}}} - \mathrm{q_{ji_{unweighted}}}$). We see that the sites that become less important in the weighted scenario are the northern hemisphere continental and offshore stations that also display higher levels of measurement uncertainty (e.g., TAP, BAL, and some of the SCS and POC cruise stations). In contrast, those that gain influence in the weighted estimate are sited in or close to the northern hemisphere continents and characterized by smaller measurement uncertainty (e.g., KEY, STM, SHM). In addition, certain remote ocean sites also display an increase in influence in the weighted estimate by virtue of their very low standard errors (e.g., MLO, CHR). The other networks, GV-ANY, F1 and R1 demonstrate similar results; for example, stations KEY and SHM figure among the most influential sites in the weighted estimates for all networks.

We also note that in the calculation of weighted estimates from network GV-ANY, the third most influential station becomes RYO; this is a station of the Japanese Meteorological Agency, which reports daily averages from continuous measurements, and which is not included in the other networks. This result stems from our definition of data uncer-

tainty as a function of the standard error of the station measurements; the uncertainty at a site thus depends not only on the measured variability (through the dependence on residual standard deviation) but also on the number of observations determining the mean measurement. Unlike networks CMDL-ANY and R1, which are composed only of flask samples (approximately weekly measurements), the GV-ANY and F1 networks also include a number of sites where measurements are made with much greater frequency (e.g., CMN, RYO, ZEP). Use of the standard error as an index of data uncertainty, therefore, results in the observations from these higher frequency sampling sites being identified as better determined than those from flask sampling sites with similar levels of variability. Station RYO, therefore, becomes an influential site in network GV-ANY due to its small standard error in combination with its location on continental Eurasia. It has a higher influence on flux estimates in comparison to a nearby station such as TAP (which has a similar residual standard deviation) as RYO is characterized by a larger number of measurements per year (more than six times as many as TAP). When using data from such high frequency observation sites it also becomes important to correctly characterize the frequency of sampling that will yield independent measurements. Although we have not accounted for this in our present analysis, autocorrelation procedures such as those discussed by *Prather* [1985] will provide an indication of the timescales over which measurements may be considered independent.

3.2.3. Sensitivity of weighted estimates to assumptions on measurement uncertainty. The use of the standard error as an index of uncertainty reflects strictly the error in measurements and cannot account for possible systematic errors arising from

modeled representation of the observations (e.g., through differences in spatial scale or errors in model transport or emissions).

However, we also note that for the flux estimates derived from the the weighted methodology of equations 3 and 4, it is not the magnitude of the individual uncertainty at the stations, σ_i, that is of significance but the relative distribution of σ_i among the various stations; i.e., if the uncertainty levels derived from the standard error index were scaled by a factor of α (under the assumption of non-correlated uncertainties as before), such that the new levels of measurement uncertainty were given by $\alpha\sigma_i$, the weighting matrix $\mathbf{R_0}$ could now be written as,

$$\mathbf{R_0} = \frac{1}{\alpha}\mathrm{diag}[\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, ..., \frac{1}{\sigma_N}] \tag{8}$$

Substituting this representation of the weighting matrix into equation 3 defining the flux estimates, we see that the fluxes $\mathbf{b}$ have no dependence on the factor α, but only on the relative weights of σ_i. The same is true for the influence matrix $\mathbf{Q}$. Errors in flux estimates (derived from the variance-covariance matrix $\mathbf{C_b}$ in equation 4) vary directly with α^2 (as was the case for the unweighted scenario). If accounting for model errors does little to alter the relative distribution of the standard error based σ_i (e.g., if it is equivalent to a scaling of the distribution), then flux estimates will not change significantly, though their associated errors will be affected. If, however, the added representation of model systematic error significantly alters the relative distribution of the weighting among stations, then equation 3 implies that flux estimates (and the influence matrix $\mathbf{Q}$) will also change. In the following section we examine the sensitivity of estimated fluxes and errors to some varying assumptions on the distribution of measurement uncertainty.

We vary levels of uncertainty in two ways: (a) we examine the impact of a minimum level of σ_i (a "threshold" value), so that remote ocean sites and southern hemisphere stations characterized by very low standard errors (e.g., less than 0.1 ppm) do not overly influence the inversion estimates. Minimum levels of σ_i are varied from 0.1 ppm to 1 ppm. (b) we contrast a measurement uncertainty distribution based on residual standard deviation with that based on the standard error.

Results of case (a) for the North American flux estimates are plotted in Figure 9 for the four main networks. The figure depicts the variation in estimates as the minimum σ_i value is varied (along the y-axes of all plots). Also included for comparison on the plots are the unweighted estimates from each network assuming a constant $\sigma_0 = 0.7$ ppm.

As would be expected, as the minimum threshold level of σ_i is increased, and the inter-station uncertainties are 'homogenized', the values of the weighted estimates approach that of the unweighted estimate. The fluxes from the F1 network (which has a large disparity between weighted and unweighted estimates) thus demonstrate the most sensitivity to the variations of the minimum threshold level. The greatest shift in estimates takes place as the threshold is varied from 0.1 to 0.3 ppm, since for most of the networks about two-thirds of the standard error values lie in this range. The main impact on estimates of the varying threshold for σ_i is on the increasing magnitude of the flux errors. Recall that these are derived from variance-covariance matrix $\mathbf{C_b}$, whose elements are driven by the magnitudes of the assumed σ_i (equation 4). The smaller values of the threshold (e.g., less than 0.2 ppm) are probably insufficient to account for possible model-derived systematic error, hence the low error values (0.2 to 0.3 Gt C per year) associated with the lower thresholds are likely unrealistic.

Some recent inversion analyses that assume heterogeneous variance of measurement uncertainty have relied on indices based on the residual standard deviations from the GLOBALVIEW-CO2 database [*Bousquet et al.* 1999ab, *Peylin et al.* 1999, *Baker* 2001]. Since the standard error is defined as the *residual standard deviation*/$\sqrt{N}$, and since the majority of stations have similar values of N (based on weekly flask sampling), a set of measurement uncertainties (σ_i) based on residual standard deviation has a similar relative distribution (though with different magnitudes) to one based on standard error. The main variations occur when N is significantly different, as in the case of sites reporting continuous measurements. In these latter cases, use of the residual standard deviation (instead of the standard error) to derive data uncertainty will not result in added influence for stations that measure more frequently. Results for a set of sensitivity analyses (similar to those of Figure 9) using a σ_i distribution based on residual standard deviation are shown in Figure 10. The main differences from the standard error based analysis of Figure 9 are (a) errors on flux estimates are much greater, since uncertainties derived from the residual standard deviation are a factor $\sqrt{N}$ larger than those from the standard error; (b) estimates demonstrate greatest sensitivity to the varying threshold through the range 0.2 to 1.2 ppm, since for most of the networks this is the range in which about two-thirds of the residual standard deviation values lie.

4. Location of potential measurement sites

We now explore the potential of the influence matrix, $\mathbf{Q}$, as a diagnostic to identify suitable measurement locations to help constrain the longitudinal partition of northern hemisphere net terrestrial fluxes. As seen in equation 6, the elements of $\mathbf{Q}$ do not depend on the observations $\mathbf{y_{obs}}$, but only on the measurement uncertainty encapsulated

in matrix $\mathbf{R_o}$. We can, therefore, use this property to investigate the value of a variety of measurement locations in hypothetical networks of stations, so long as we are able to define the measurement uncertainty at these sites.

We present results calculated for a hypothetical grid of surface stations spaced at intervals of 10 degrees longitude and 10 degrees latitude (although greater spacing may be required in some regions to ensure that measurements are spatially uncorrelated). The results of three assumptions on measurement uncertainty are examined. a) **Case CONST** : Measurement errors at all stations are normally distributed with constant variance ($\sigma_{allstations}$ = 0.3 ppm); b) **Case LAND/SEA** : Measurement errors at land-based stations have a higher variance than those at ocean stations ($\sigma_{landstations}$ = 0.4 ppm, $\sigma_{oceanstations}$ = 0.1 ppm); c) **Case CTMVAR** : Measurement uncertainty levels at stations are computed based on statistics of modeled concentration variability using the GISS II CTM for a scenario incorporating fossil fuel, terrestrial biosphere and ocean fluxes. The uncertainty then is calculated according to the following procedures: First, four-hourly concentration time-series are obtained at the specified gridpoints. In order to obtain independent measurements from the high frequency time-series, these data are binned into 5 day intervals; the determination of this timescale was based on autocorrelation analysis performed for pollution tracers in the GISS CTM [*Prather et al.*, 1987]. Using median values from the individual bins, the statistics for residual standard deviation and standard error of the modeled concentration variability are computed in a similar manner to the methods for NOAA/CMDL flask data; i.e., by analyzing the residual variability after removing a fitted linear time trend and seasonal components.

Using these three different scenarios of measurement uncertainty, we calculate matrix **Q** according to equation 6. The relative importance of the gridded measurement sites in constraining the North American terrestrial flux estimate, as quantified by the q_{ji} diagnostic, are presented in Figures 11 to 13.

Case **CONST**, which assumes a constant σ_i suggests that suitable observational sites are those closest to the signals; i.e. directly on the northern hemisphere continents (Figure 12). In contrast, case **LAND/SEA**, which assumes a higher level of measurement uncertainty associated with the land-based continental stations, indicates that determination of the longitudinal partition of the net terrestrial fluxes is best achieved by measurements around the continental margins, i.e., upstream and downstream of the continental signal (Figure 13). Finally, in Figure 14, we present results for case **Case CTMVAR**, in which uncertainty is determined using a measure of modeled concentration variability. As with the previous case, the most suitable measurement locations are positioned off-shore around the northern hemisphere continents. In contrast to case **LAND/SEA**, however, the distribution is predominantly focused around the southern peripheries of Eurasia and North America. The primary reason for the de-emphasis on measurement sites in the mid and high northern latitudes in this scenario, is the higher level of modeled concentration variability associated with these regions. The resulting high values of measurement uncertainty render these locations of less value as measurement sites.

We stress that the analysis presented in this section is intended more as an illustration of the potential of the influence matrix diagnostic (q_{ji}) than as a specific recommendation on measurement sites. Detailed analyses employing more realistic scenarios of measurement uncertainty must be carried out before such recommendations can be made. It is also

important to note that location of suitable stations identified via the influence matrix diagnostic depend very specifically on the question addressed, due to the dependence of matrix $\mathbf{Q}$ on the problem-specific response functions, X_j. A similar analysis of the land-sea partition of net CO_2 fluxes, for example, will yield a different set of suitable measurement locations.

5. Limitations of the inversion method

The inversion method of this study has certain limitations that influence derived flux estimates and may be partially responsible for the large sensitivity to network configuration demonstrated in the early sections. Some of these limitations and their impact on the results are discussed below.

(a) Solution for fluxes from continental scale regions: As was discussed in Section 3.1.1, part of the large sensitivity of terrestrial flux estimates to a few critical stations (most notably the SCS stations) resulted from the formulation of the inversion problem based on three large land regions as in the study of *Fan et al.* [1998]; i.e., estimation of continental scale fluxes was often determined by a few peripheral stations and affected by their associated systematic errors. *Fan et al.* [1998] note that this low level of resolution was employed because the sparse observation network did not permit accurate determination of fluxes from continental subdivisions and yielded impracticably large estimation errors. However, other recent inversion studies suggest that use of such large regions may result in biased estimates and advocate solving at a higher resolution followed by post-solution aggregation of fluxes. The rationale for this is that the resolution of the estimated fluxes should be determined by the data and not by an initial imposed distribution which may result in a restricted solution space [e.g., *Enting et al.* 1995, *Kaminski et al.* 2001].

(b) A predefined ocean emissions distribution : This inversion method employs fixed oceanic fluxes derived from *Takahashi et al.* [1997]; the derived flux estimates are, therefore, sensitive to the accuracy of this distribution. An underestimate of North Atlantic uptake, for example, is likely, in this three region inversion, to lead to a misallocation of an ocean flux component to a northern hemisphere continental region. We note that some recent inversion studies that solve for both oceanic and terrestrial fluxes (albeit with different inversion methods) yield results of larger North Atlantic uptake than the *Takahashi et al.* [1997] values; e.g., in their control case, *Bousquet et al.* [1999a] obtain a temperate North Atlantic (15^oN - 50^oN) uptake of over 0.5 Gt C per year in comparison to a *Takahashi et al.* [1997] a priori flux estimate of 0.3 Gt C per year. *Baker* [2001] (Table 3.6), who has conducted a series of inversions testing the sensitivity of flux estimates to a range of parameters, finds that total North Atlantic uptake increases significantly (to over 1 Gt C per year) and North American terrestrial uptake decreases (by 0.5 Gt C per year) when solving for oceanic fluxes in comparison to a scenario of fixed oceanic exchanges derived from *Takahashi et. al.* [1997].

(c) Independence of observations : Many recent inversion studies that employ observations from the GLOBALVIEW database or CMDL network assume that all measurement sites sample independent airmasses. Thus even sites in close proximity, for example, stations CMO and OPW, which are at approximately the same longitude and within three degrees latitude of each other, or the South China Sea ship track measurements, which are taken three degrees latitude apart, are treated independently in the inversion methodology. If, however, such stations are not sampling independent airmasses, this assumption will result in an overemphasis of the significance of such stations to the analysis. We also

note that the coarse grid resolution of the transport models (e.g., the horizontal resolution of the GISS II CTM is 4^o latitude $\times$ 5^o longitude) often results in modeled concentrations being taken from adjacent gridboxes or even the same gridbox when representing such sites. To gauge the magnitude of this assumption of independence of observation sites on the flux estimates, we conducted a very simple sensitivity study in which stations within three degrees of each other were represented by a single averaged value in the inversion analysis. The largest impact on flux estimates were for results from the CMDL-ANY and F1 networks; flux estimates for North American uptake dropped by 0.6 - 0.9 Gt C per year for the unweighted cases, and up to 0.3 Gt C per year for the weighted analyses. The main cause of this shift was the de-emphasis of the closely spaced SCS stations; as discussed in section 3.1.1, flux estimates, for the unweighted analysis in particular, show a large sensitivity to inclusion of these stations. We stress that this was a simple exercise in evaluating the assumption of independence of stations; a more sophisticated analysis would account for such factors as regional meteorology at the stations.

(d) Concentrations in excess of South Pole station value : As with the study of Fan et al. [1998], the concentrations employed in this analysis were those in excess of the value at the South Pole station (SPO). This site is often chosen for deriving such an offset as it is characterized by low variability in measurement and modeled concentrations. As noted by Baker [2001], however, this procedure may provide a source of error by propogating any measurement or model errors at the South Pole site to all other stations. A preferable approach would be to solve for a station independent offset, as outlined in Baker [2001].

6. Summary

We have examined the sensitivity of a specific inversion methodology, namely, a time-independent analysis similar to that employed by *Fan et al.* [1998], to varying configurations of the observational network. Using diagnostics derived from the inversion equations, we also identify the relative influence of individual stations on the continental scale fluxes. While estimates of the total northern mid-latitude terrestrial uptake are well constrained by the observations (1.1 to 1.35 Gt C per year), for the unweighted case, flux estimates for the separate northern hemisphere continents display a significant sensitivity to the makeup of the observation station network. Estimates of North American net fluxes, for example, range from uptake of 1.6 Gt C per year to efflux of -0.3 Gt C per year depending on the observation network employed. Use of the influence matrix diagnostic ($\mathbf{Q}$), which quantifies the influence of the individual stations on the flux estimates, indicates that the unweighted estimates are primarily determined by a few stations in and close to the northern hemisphere continents. Our analysis displays a particular sensitivity to the inclusion of the South China Sea (SCS) measurements.

We have also examined solutions to a weighted inversion method, where the measurement uncertainty associated with the stations is not assigned equally but instead varied to reflect confidence in measurement and model values at the sites. Estimates of the standard error of the measurements were used initially as a weighting approach. North American terrestrial uptake estimates derived from the weighted approach are lower and display less sensitivity to observation network configuration than for the unweighted inversion. In particular, the previously high uptake estimates are considerably reduced, as shown in Figure 8 (e.g., from 1.5 $\pm$ 0.7 Gt C per year to 0.35 $\pm$ 0.2 Gt C per year for

the *Fan et al.* [1998] network). This change is caused by the diminished role of several continental stations previously influential in the unweighted scenario; these are the sites that are also characterized by higher levels of data uncertainty.

As part of our analysis of weighted scenarios, we have also examined the influence of different characterizations of measurement uncertainty. It is important to note that for this inversion method, the derived flux estimates remain unaffected for variations in the magnitude of the measurement uncertainty terms, so long as the relative distribution of σ_i remains unchanged; i.e., the scaling of σ_i only affects the derived errors. Use of the standard deviation of the station measurements to characterize data uncertainty, as has been done in some in some recent studies, thus yields similar results for flux estimates as use of the standard error; this results from the fact that although the magnitude of the standard deviation is larger, its distribution across stations is similar to that of the standard error since sampling frequencies are often similar. Use of the standard deviation to characterize measurement uncertainty results, however, in larger values of the estimated errors, as these depend on the magnitude of σ_i.

Finally, we illustrate the potential of a diagnostic employed in this analysis, namely, the influence matrix $\mathbf{Q}$, to identify suitable measurement locations for achieving better constraints on the longitudinal partition of northern hemisphere terrestrial fluxes. Preliminary analyses, relying on distributions of measurement uncertainty derived from the CTM, suggest increased peripherial measurements around the northern hemisphere continents.

References

Andres et al. 1996 G. Marland, I. Fung, and E. Matthews, A 1x1 distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, *Global Biogeochemical Cycles, 10*, 419-429, 1996.

Baker, D., An inversion method for determining time-dependent surface CO2 fluxes, in Inverse Methods in Global Biogeochemical Cycles, P. Kasibhatla et al., eds., 279-293, American Geophysical Union, Washington, DC, 1999.

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

Balkanski, Y. J., D. J. Jacob, R. Arimoto, and M. A. Kritz, Long-range transport of radon-222 over the North Pacific Ocean: Implications for continental influence, *J. Atmos. Chem., 14*, 353-374, 1992.

Belsley, D. A., E. Kuh and R. E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, John Wiley and Sons, Inc., 1980.

Bousquet, P., Csias, P., 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., 104D*, 26,161-26,178, 1999a.

Bousquet, P., Csias, P., P. Peylin, M. Ramonet, and P. Monfray, Inverse modeling of annual atmospheric CO_2 sources and sinks : 2. Sensitivity study, *J. Geophys. Res., 104D*, 26,179-26,193, 1999b.

Bousquet, P., P. Peylin, P. Cias, C. Le Quere, P. Friedlingstein and P. Tans, Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science, 290*, 1342-

1346, 2000.

Birdsey, R.A., and L. S. Heath, in Productivity of America's Forest and Climatic Change, General Technical Report RM-GTR-271, 56-70, L.A. Joyce, Eds., U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 1995.

Ciais, P., P. Tans, M. Trolier, J. White and R.J. Francey, A large northern hemisphere terrestrial CO_2 sink indicated by $^{13}C/^{12}C$ of atmospheric CO_2, *Science, 269*, 1098-1102, 1995.

Chin, M., D. J. Jacob, G. M. Gardner, M. S. Forman-Fowler, P. A. Spiro and D. L. Savoie, A global three-dimensional model of tropospheric sulfate, *J. Geophys. Res., 101*, 18,667-18,690, 1996.

Conway, T.J., 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 National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res., 99D*, 22831-22855, 1994.

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

Denning, A.S., D.A. Randall, G.J. Collatz, and P.J. Sellers, Simulations of terrestrial carbon metabolism and atmospheric CO_2 in a general circulation model: Part 1, Surface carbon fluxes, *Tellus, 48B*, 521-542, 1996.

Enting, I. G. and J. V. Mansbridge, Latitudinal distribution of sources and sinks of CO_2; Results of an inversion study, *Tellus, 43B*, 156-170, 1991.

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

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 carbon dioxide data and models, *Science, 282*, 442-446, 1998.

Fung, I., C.J. Tucker and K. Prentice, Application of Advanced Very High Resolution Vegetation Index to study atmosphere-biosphere exchange of CO_2, *J. Geophys. Res., 92D*, 2999-3015, 1987.

GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA CMDL, Boulder, Colorado (Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path:ccg/co2/GLOBALVIEW), 1999.

Gloor, G., S-M. Fan, S. Pacala, J. Sarmiento and M. Ramonet, A model-based evaluation of inversions of atmospheric transport, using annual mean mixing ratios, as a tool to monitor fluxes of nonreactive trace substances like CO_2 on a continental scale, *J. Geophys. Res., 104D*, 14,245-14,260, 1999.

Gurney, K., Law. R., and TRANSCOM 3 authors, Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature, 415*, 626-630, 2002.

Hansen, J., G. Russell, D. Rind, P. Stone, A. Lacis, S. Lebedeff, R. Ruedy and L. Travis, Efficient three-dimensional global models for climate studies: Models I and II, *Mon. Weather Rev., 111*, 609-662, 1983.

Houghton, R.A., J.L. Hackler and K.T. Lawrence, *Science, 285*, 574-578, 1999.

Horowitz, L.W., J. Liang, G.M. Gardner, and D.J. Jacob, Export of reactive nitrogen from North America during summertime, *J. Geophys. Res., 103*, 13,451-13,476, 1998

Intergovenmental Panel on Climate Change (IPCC), Working Working Group I, Thrid Assessment Report, *Climate Change 2001: The Scientific Basis*, J. Houghton et al. eds, Cambridge University Press, 2001.

Jacob, D.J., J.A. Logan, R.M. Yevich, G.M. Gardner, C.M. Spivakovsky, S.C. Wofsy, J.W. Munger, S. Sillman, M.J. Prather, M.O. Rodgers, H. Westberg, and P. R. Zimmerman, Simulation of summertime ozone over North America, *J. Geophys. Res., 98*, 14797-14816, 1993.

Jacob, D.J., and M.J. Prather, Radon-222 as a test of convection in a general circulation model, *Tellus, 42*, 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.

Kaminski, T., P. Rayner, M. Heimann, and I. Enting, On Aggregation errors in atmospheric transport inversions, *J. Geophys. Res., 106D*, 2001.

Keeling, R.F., S.C. Piper and M. Heimann, Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration, *Nature, 381*, 218-221, 1996.

Law, R., et al., Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions, *Global Biogeochemical Cycles, 10*, 783-796, 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, D6*, 11,593-11,610, 1995.

Myers, 1989 Myers, R., Classical and Modern Regression with Applications, PWS-Publishing Company, 1990.

Pacala, S, G. C. Hurtt, D. Baker, P. Peylin, R. A. Houghton, R. A. Birdsey, L. Heath, E. T. Sundquist, R. F. Stallard, P. Ciais, P. Moorcroft, J. P. Caspersen, E. Shevliakova, B. Moore, G. Kohlmaier, E. Holland, M. Gloor, M. E. Harmon, S.-M. Fan, J. L. Sarmiento, C. L. Goodale, D. Schimel, and C. B. Field, Consistent Land- and Atmosphere-Based U.S. Carbon Sink Estimates, *Science, 292*, 2316-2320, 2001.

Peylin, P., P. Bousquet, P. Ciais, and P. Monfray, Differences of CO_2 flux estimates based on a time-independent versus a time-dependent inversion method, in *Inverse Methods in Global Biogeochemical Cycles*, P. Kasibhatla et al., eds., 295-309, American Geophysical Union, Washington, DC, 1999.

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

Prather, M. J., M. B. McElroy, S. C. Wofsy, G. Russell, and D. Rind, Chemistry of the global troposphere: Fluorocarbons as tracers of air motion, *J. Geophys. Res., 92*, 6579-6613, 1987.

Press, W. H., et al., *Numerical Recipes*, Cambridge University Press, New York, 1995.

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

Schimel, D.S., Terrestrial ecosystems and the carbon cycle, *Global Change Biology, 1*, 77-91, 1995.

Spivakovsky, C. M., J. A. Logan, S. A. Montzka, Y. J. Balkanski, M. Foreman-Fowler, D. B. A. Jones, L. W. Horowitz, A. C. Fusco, C. A. M. Brenninkmeijer, M. J. Prather, S. C. Wofsy, and M. B. McElroy, Three-dimensional climatological distribution of tropo-

spheric OH: Update and evaluation, *J. Geophys. Res., 105D*, 8,931-8,980, 2000.

Spivakovsky, C. M., R. Yevich, J. A. Logan, S. C. Wofsy, M. B. McElroy, and M. J. Prather, Tropospheric OH in a three-dimensional chemical tracer model: An assessment based on observations of CH^3CCl^3, *J. Geophys. Res., 95*, 18,441-18,471 1990.

Takahashi, T., R. Feely, R. Weiss, R. Wanninkhof, D. Chipman, S. Sutherland and T. Takahashi, Global air-sea flux of CO_2: An estimate based on measurements of sea-air pCO_2 difference, *Proceedings of the National Academy of Sciences, 94*, 8292-8299, 1997.

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

Thoning, K. W., P. P. Tans and W. D. Komhyr, Atmospheric carbon dioxide at Mauna Loa Observatory, 2, Analysis of the NOAA GMCC data 1974, 1985, *J. Geophys. Res., 94, D6*, 8549-8565, 1989.

Wang, Y., D.J. Jacob, and J.A. Logan, Global simulation of tropospheric O3-NOx-hydrocarbon chemistry, 1. Model formulation, *J. Geophys. Res., 103D*, 10,713-10,726, 1998.

Warneck, P., Chemistry of the Natural Atmosphere, Academic Press, 1988.

Network	**Description**	**Stations**
CMDL-ANY	Any NOAA CMDL stations reporting measurements in 1988-92	61
GV-ANY	Any GLOBALVIEW stations reporting measurements in 1988-92	75
F1	Station network of *Fan et al.* [1998]	63
R1	"Extended" network of *Rayner et al.* [1999]	25

Table 1. The main observation networks analyzed in this study.

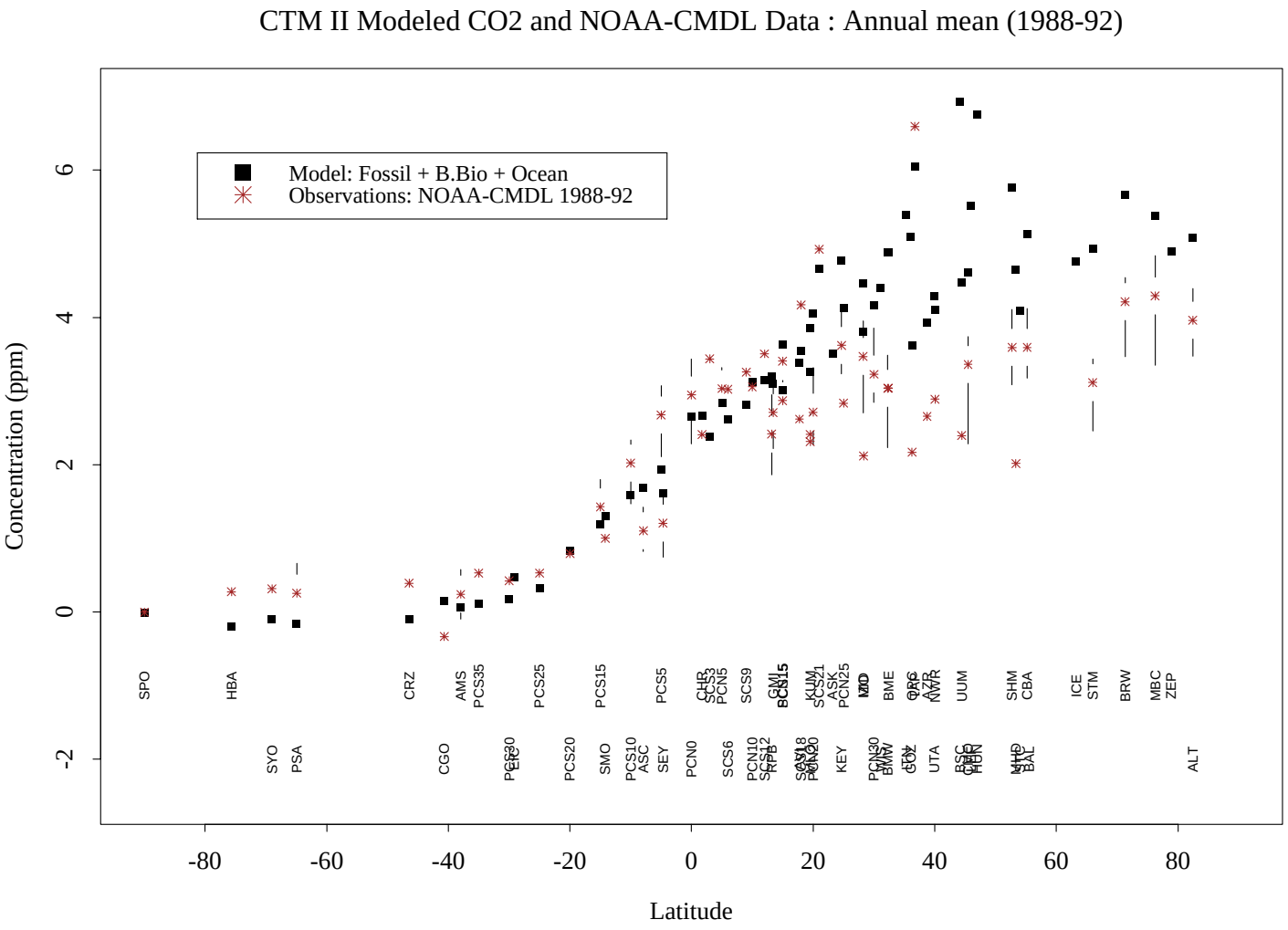

Figure 1 : Latitudinal distribution of modeled CO_2 and observations at the NOAA-CMDL sites (annual averages for the period 1988 - 1992, relative to the South Pole station (SPO)). Emissions are modeled for fossil fuel, annually balanced seasonal biospheric exchange and oceanic uptake (simulation FBO). Error bars represent maximum and minimum annual mean observations for the period.

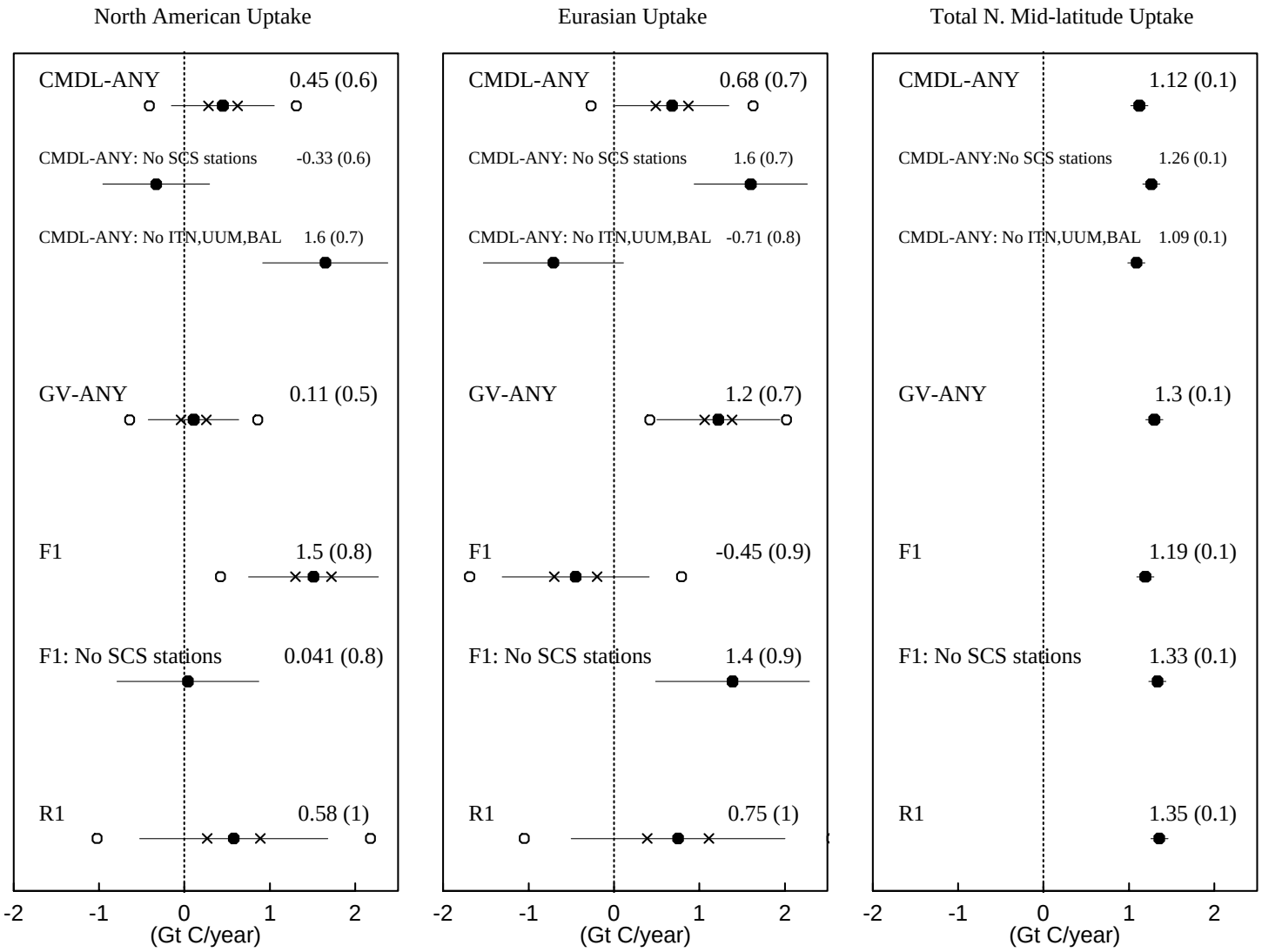

Figure 2 : Sensitivity of estimated net terrestrial fluxes to observation network. The left-hand and middle panels show results for North America and Eurasia (north of 20^o N), respectively. The right-hand panel shows results for the total northern mid-latitude terrestrial flux. Positive flux values indicate net uptake. Plotted values (filled circles) associated with each network denote the regional estimate and the lengths of the bars represent one standard error for a data uncertainty level of $\sigma_0 = 0.7$ ppm. Also shown for the main networks are the estimated errors for a data uncertainty level of $\sigma_0 = 1$ ppm (open circles), and a data uncertainty of $\sigma_0 = 0.2$ ppm (crosses).

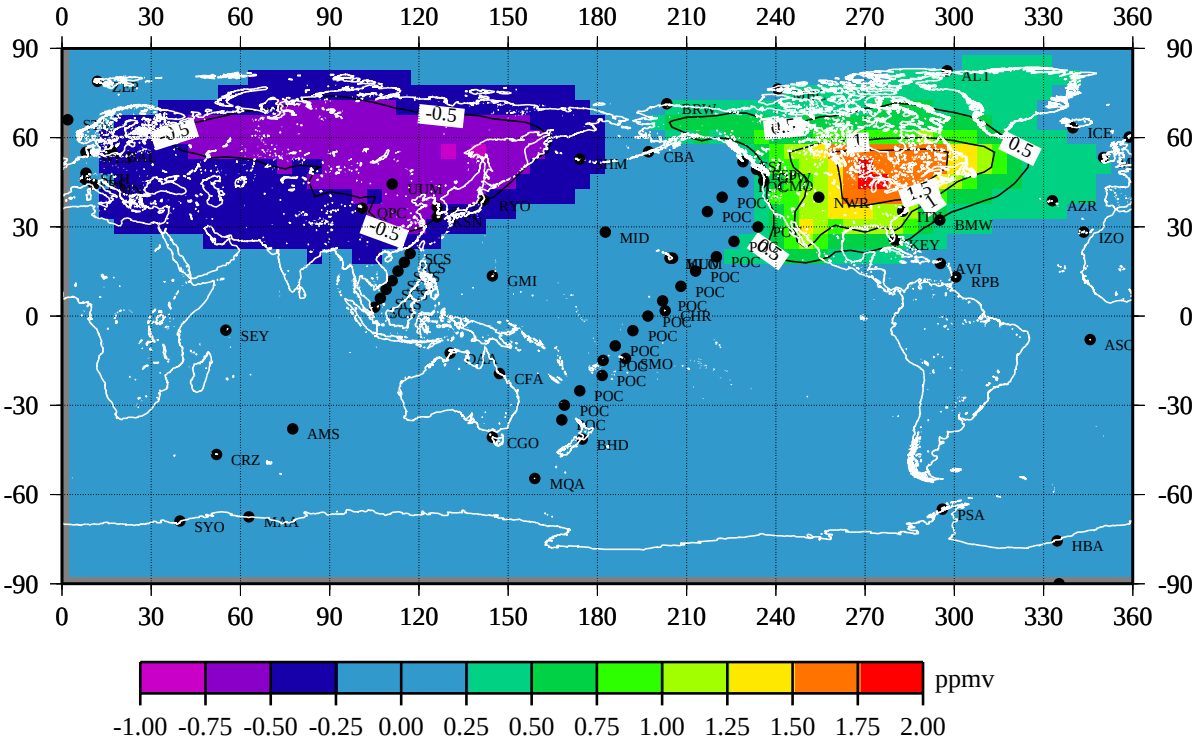

Figure 3: Regions most sensitive to biospheric uptake by North America and Eurasia, as illustrated by the difference between simulations in which the entire northern mid-latitude terrestrial sink is concentrated in North America (NA1) and Eurasia (EA1) respectively.

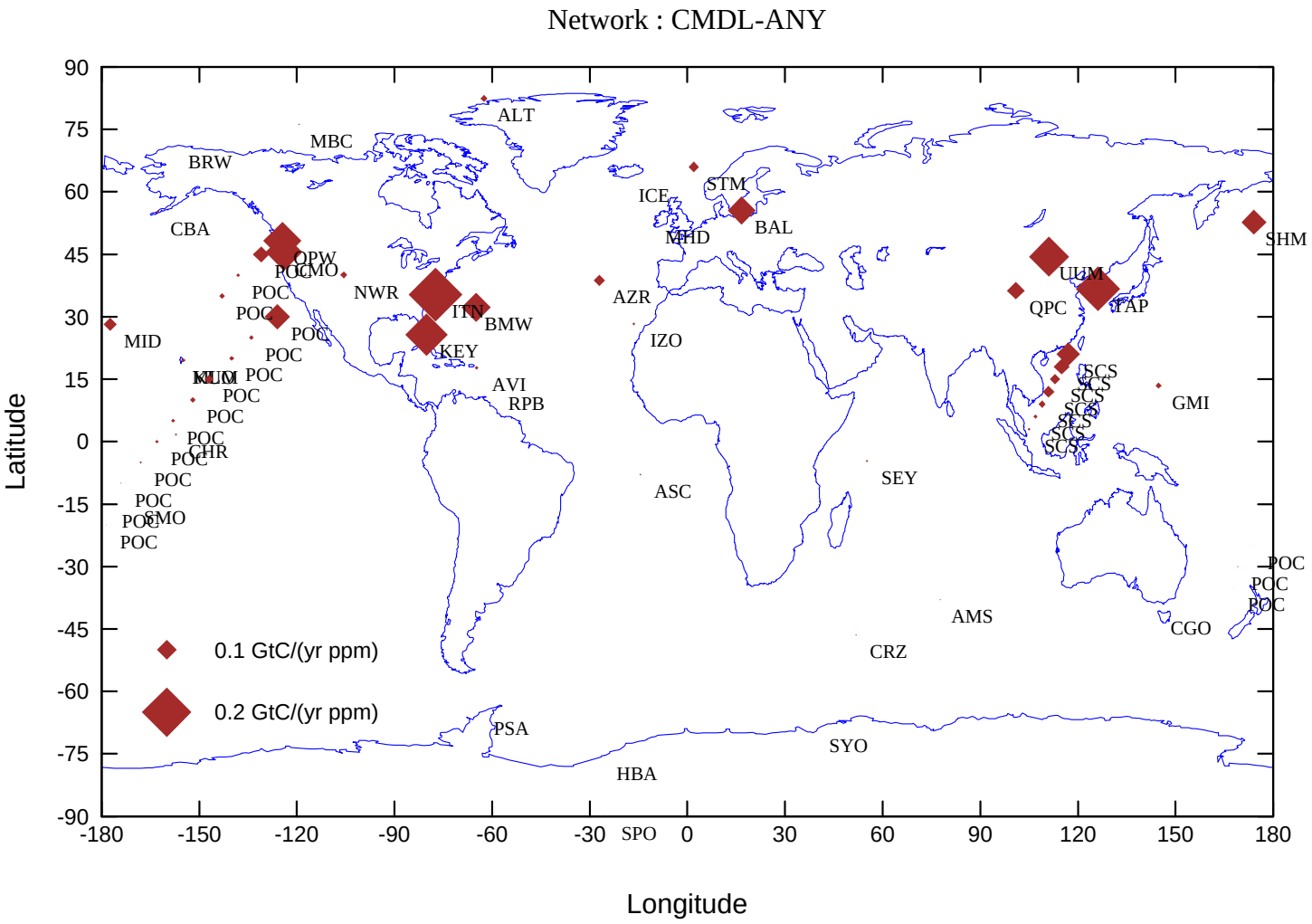

Figure 4a: Influence of individual stations on estimate of North American uptake via the diagnostic q_{jj}: Network CMDL-ANY. Unweighted case : homogeneous variance of the measurement errors.

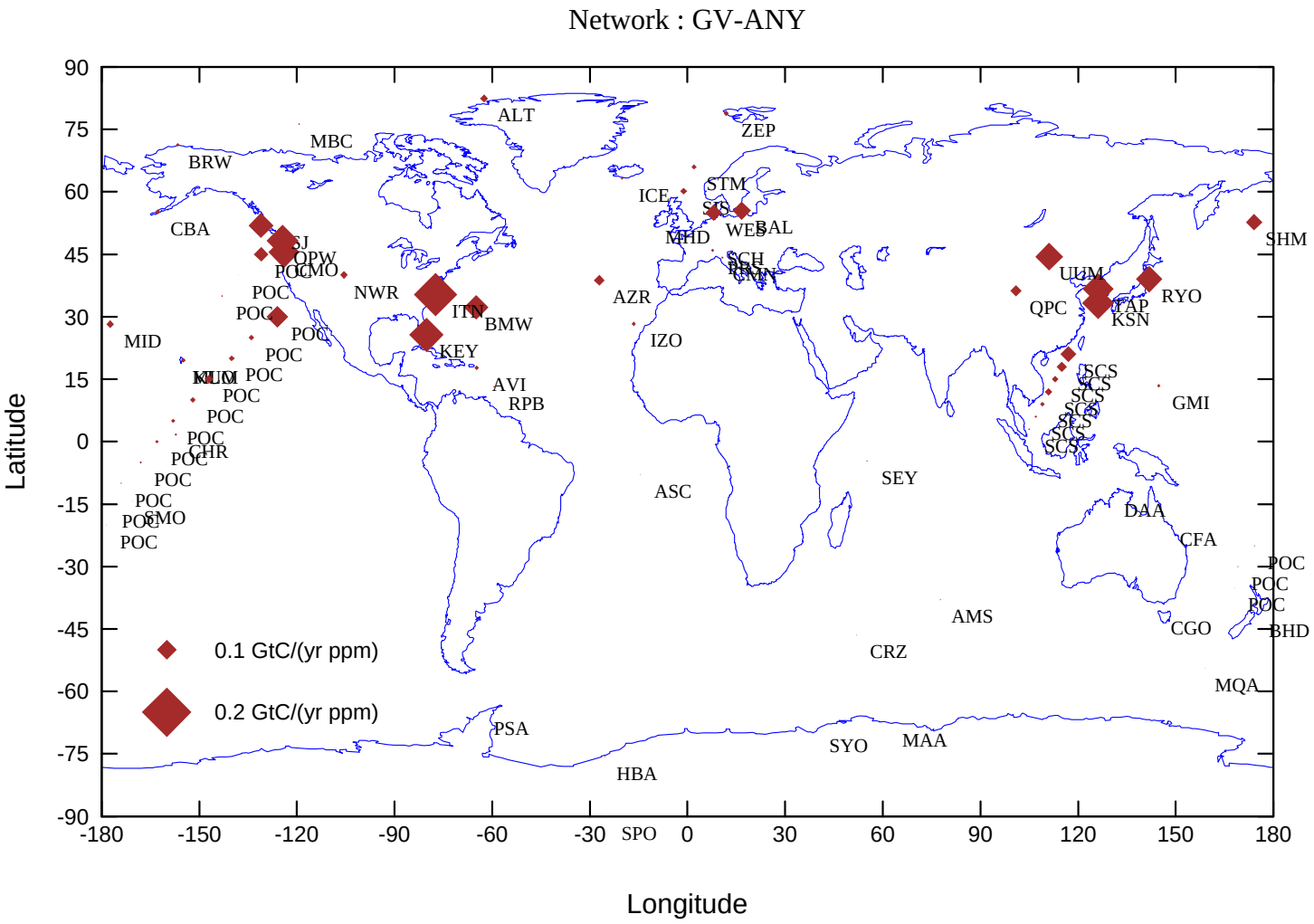

Figure 4b: Influence of individual stations on estimate of North American uptake via the diagnostic q_{ji}: Network GV-ANY. Unweighted case : homogeneous variance of the measurement errors.

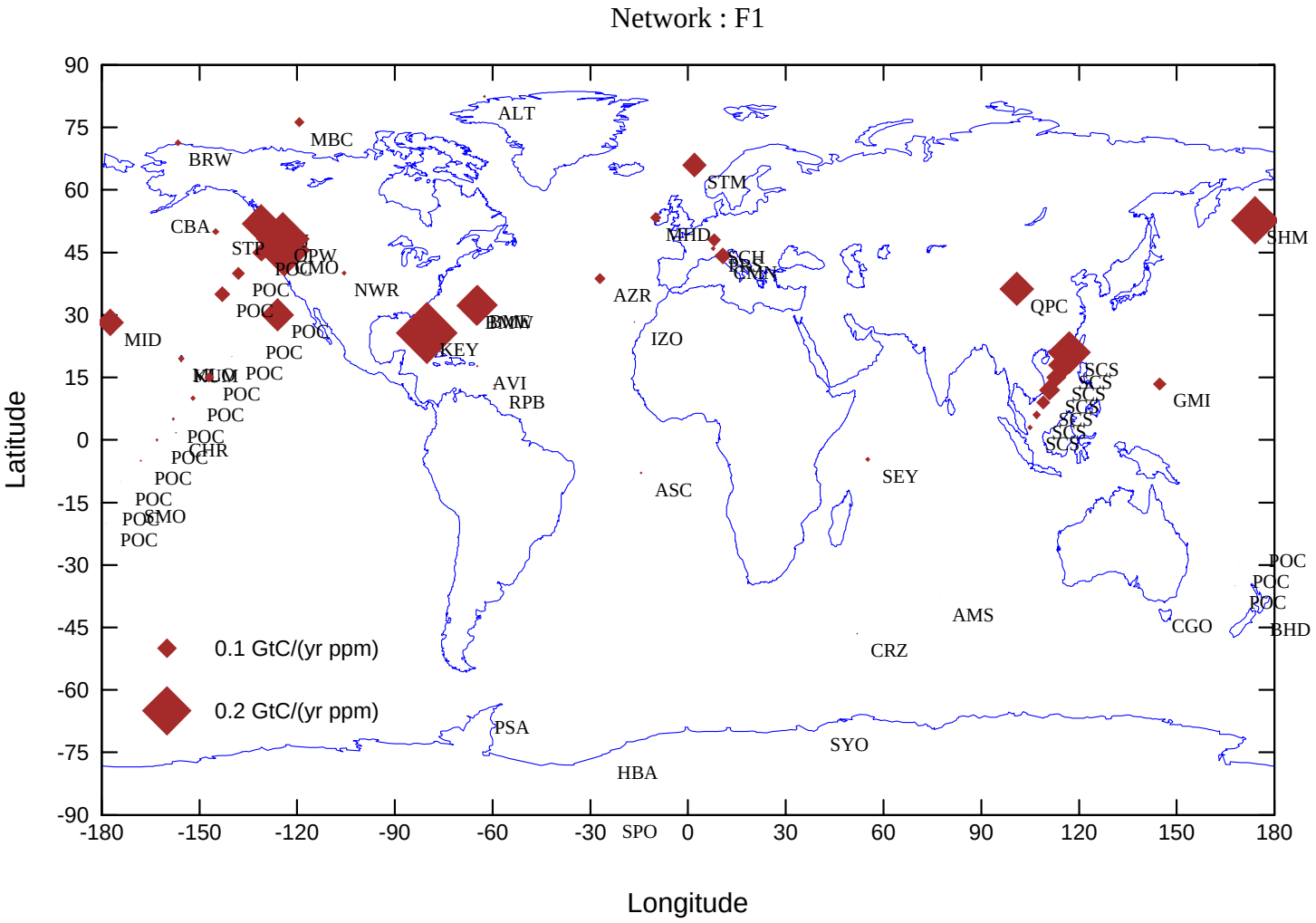

Figure 4c: Influence of individual stations on estimate of North American uptake via the diagnostic q_{ji}: Network F1. Unweighted case : homogeneous variance of the measurement errors.

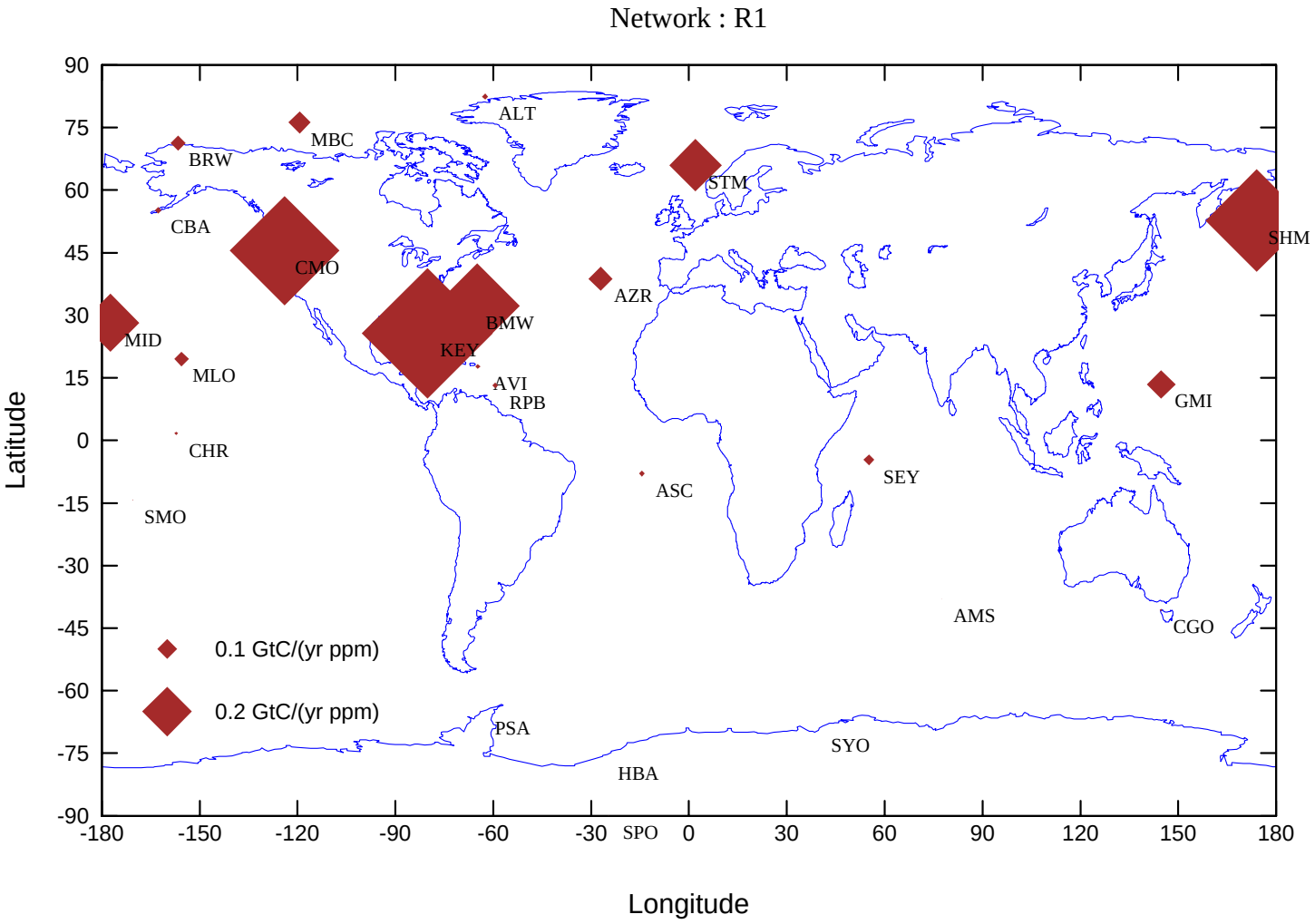

Figure 4d: Influence of individual stations on estimate of North American uptake via the diagnostic q_{ji}: Network R1. Unweighted case : homogeneous variance of the measurement errors.

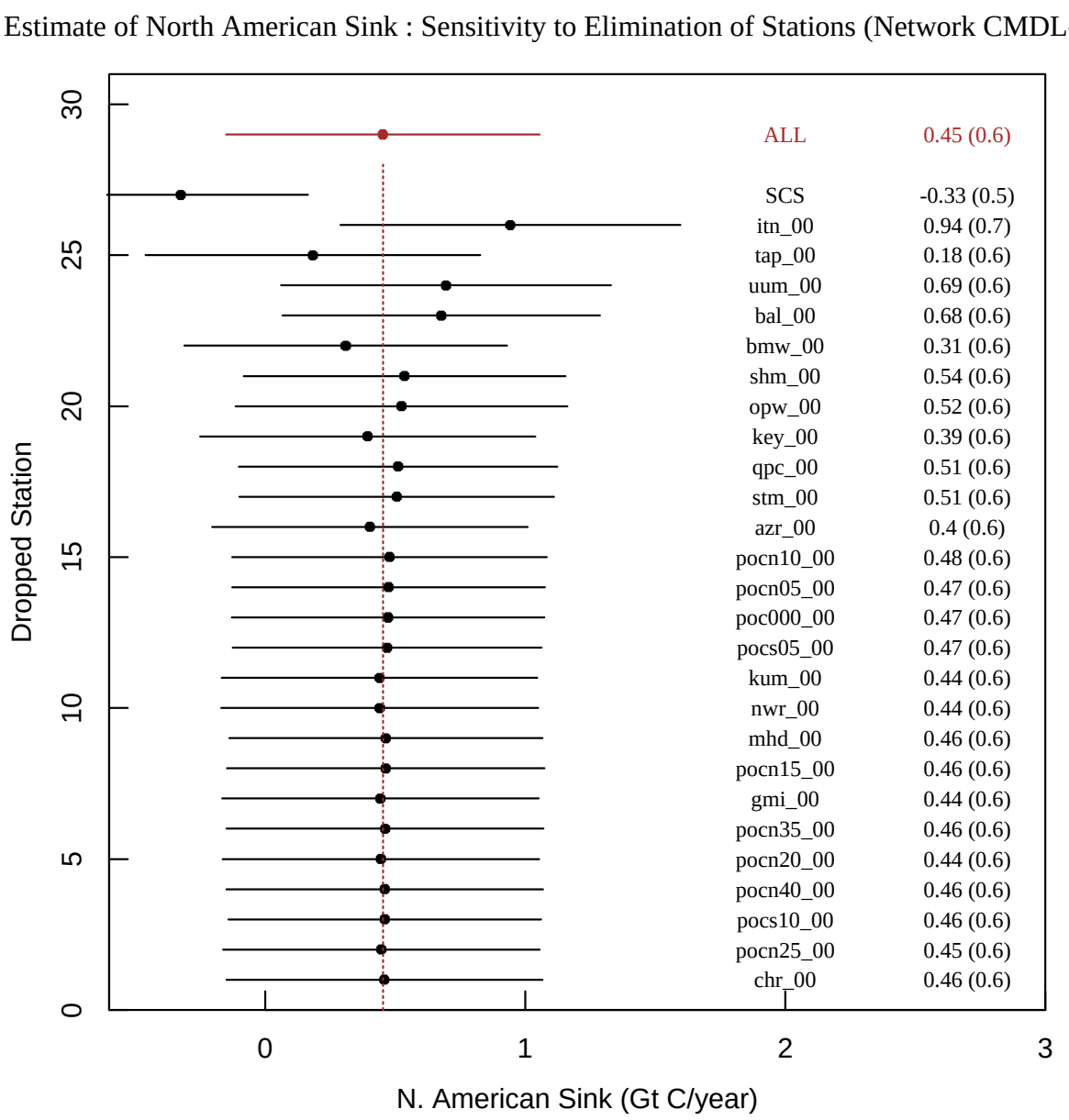

Figure 5: Sensitivity of estimate of North American terrestrial uptake to elimination of stations in the CMDL-ANY network. The plotted value associated with individual stations indicates the estimate from the remainder of the network when that station is dropped. Stations are ordered based on the impact of their elimination on the overall estimate ("ALL") and only the 25 most influential stations are shown. (Unweighted case : homogeneous variance of the measurement errors).

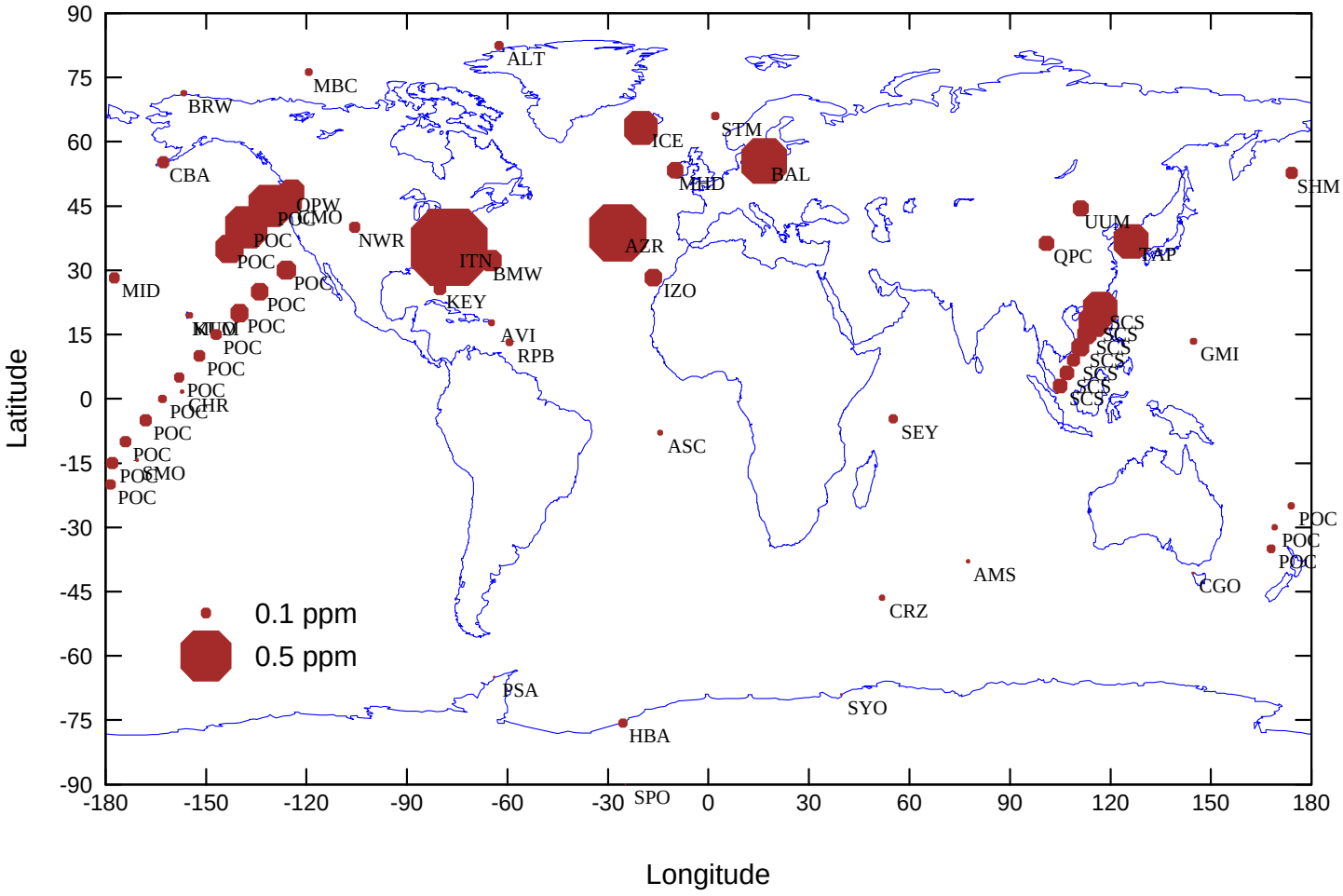

Figure 6: Magnitude of standard errors calculated for the CMDL-ANY network.

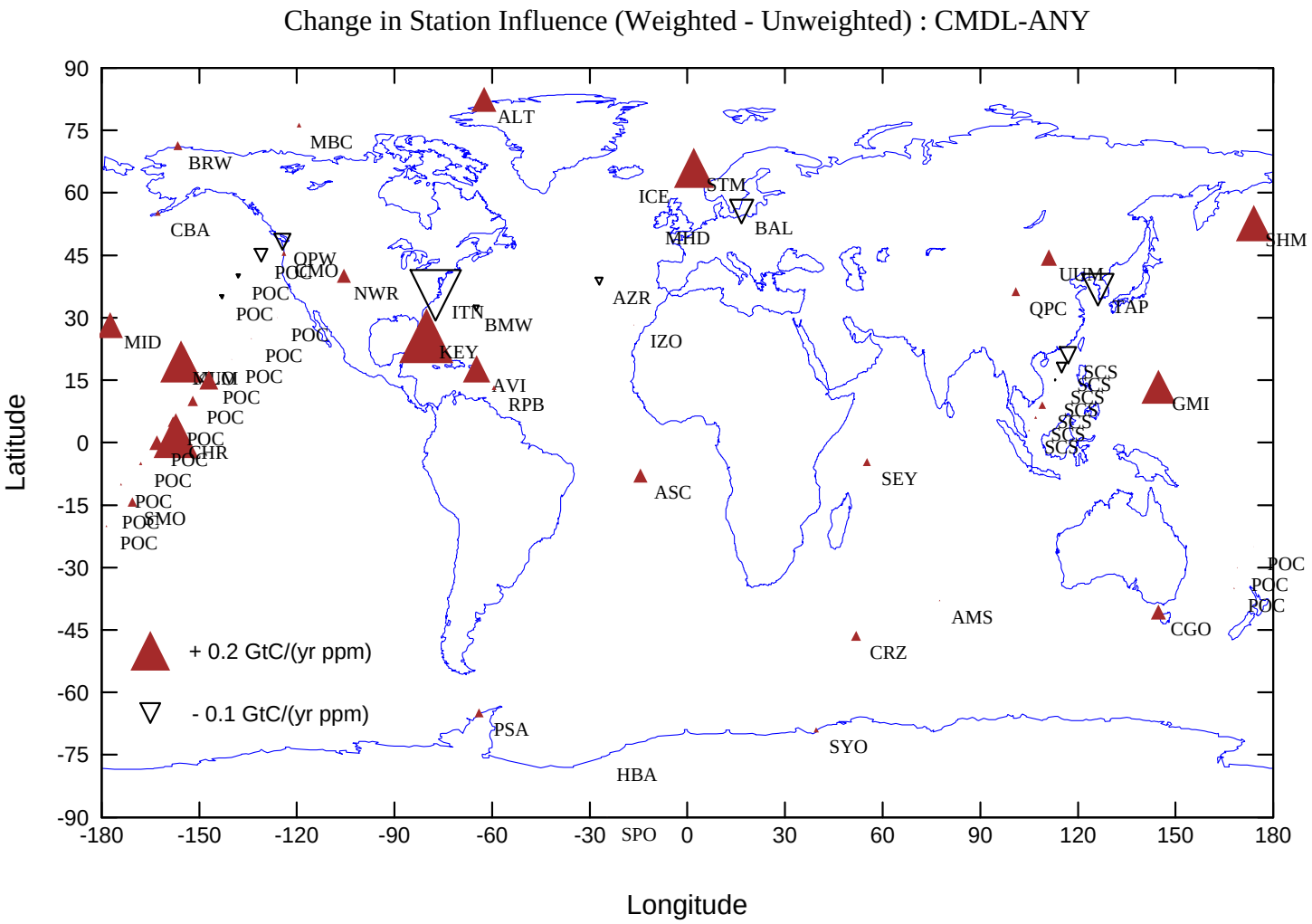

Figure 7: Change in the influence of individual stations for weighted vs. unweighted estimates for the CMDL-ANY network (via the diagnostic $\Delta q_{ji} = q_{ji_{weighted}} - q_{ji_{unweighted}}$). Upward pointing red triangles indicate an increase in the influence of the station and downward triangles indicate a decrease. The size of the symbol is proportional to the magnitude of the change.

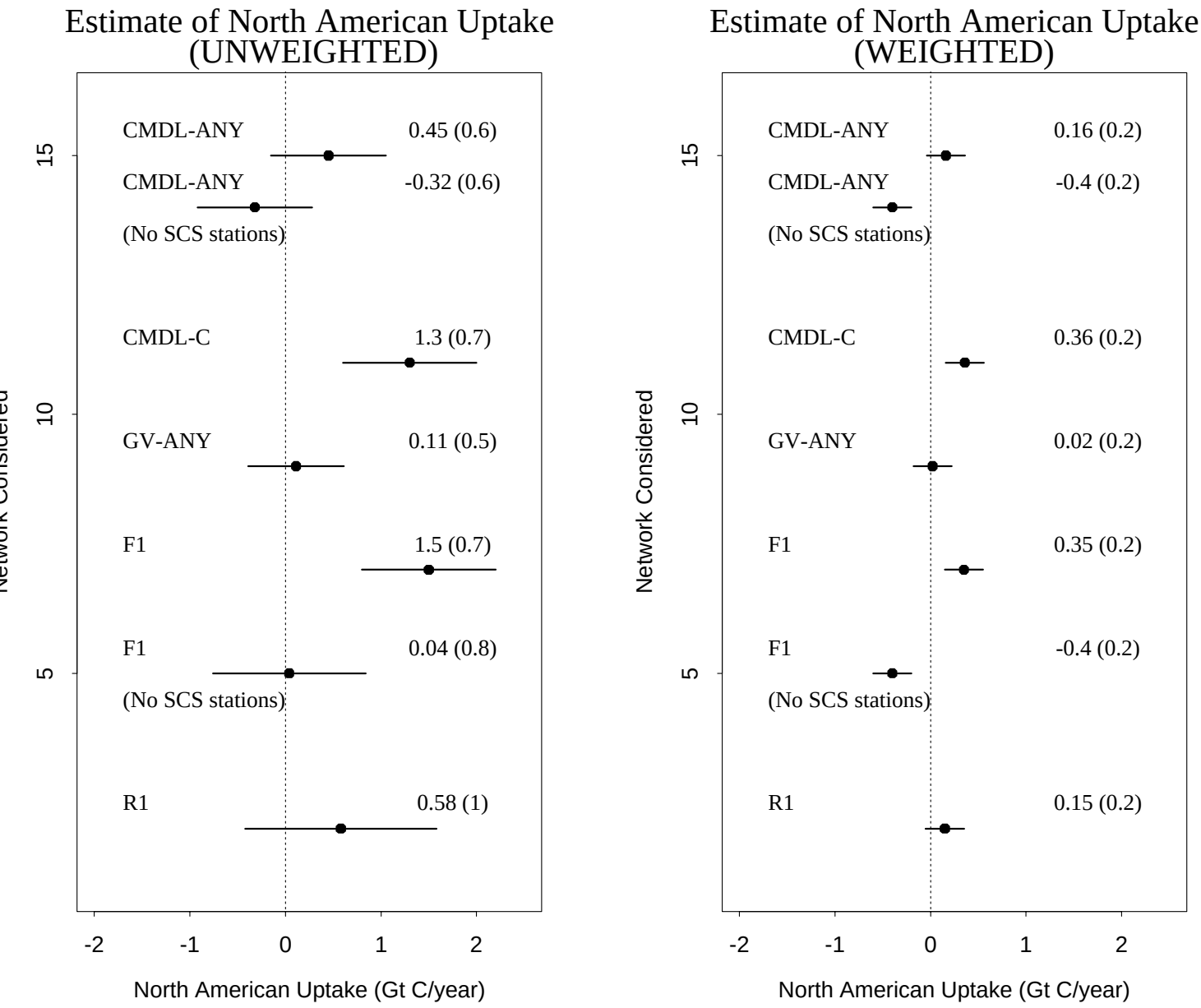

Figure 8: Estimates of North American Uptake : Weighted (right hand panel) vs. unweighted (left hand panel). Positive flux values indicate net uptake. Plotted values associated with each network denote the regional estimate and the lengths of the bars represent one standard error. Results for the unweighted estimates are shown for a measurement error level of $\sigma_0 = 0.7$ ppm.

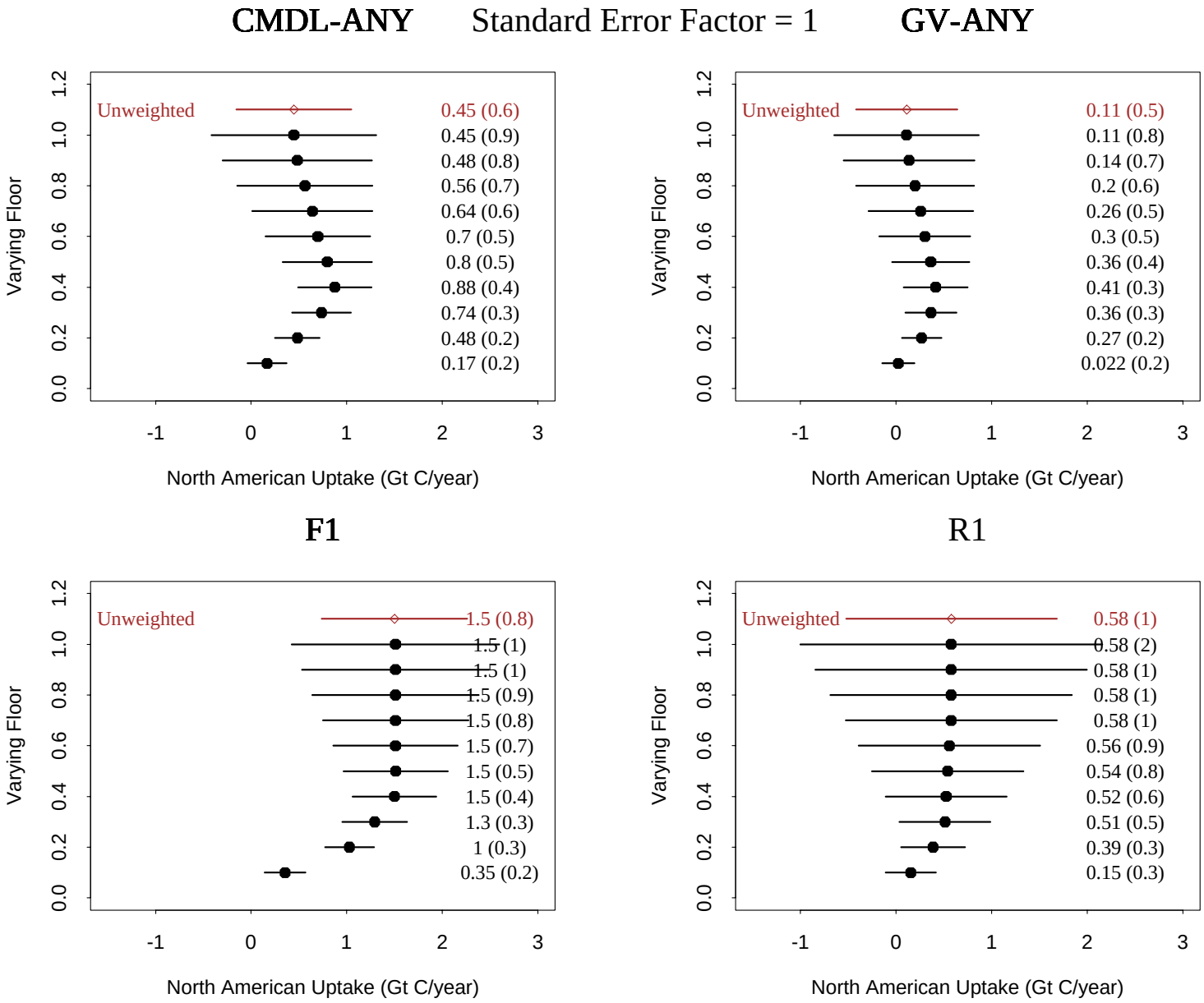

Figure 9: Sensitivity of weighted estimates to assumptions on measurement error. Results are shown for North American terrestrial flux for the four main networks. Measurement uncertainty levels depicted are based on the standard error with a varying minimum threshold level of 0.1 ppm to 1.0 ppm (on y-axis). Also shown are the estimates of the unweighted scenarios for each network for a measurement uncertainty level of 0.7 ppm.

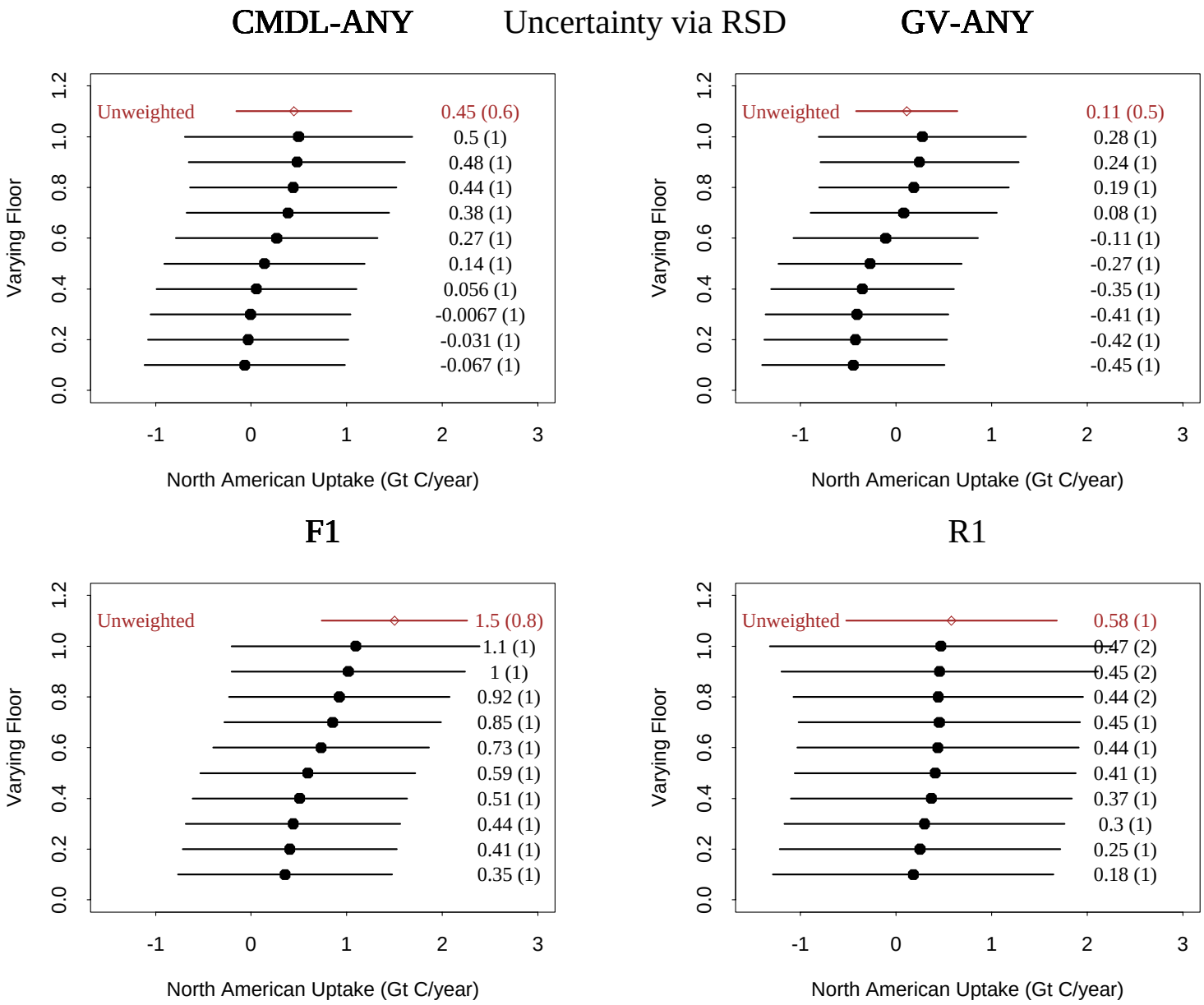

Figure 10: Sensitivity of weighted estimates to assumptions on measurement error. Results are shown for North American terrestrial flux for the four main networks. Measurement uncertainty levels depicted are based on the residual standard deviation with a varying minimum threshold level of 0.1 ppm to 1.0 ppm (on y-axis). Also shown are the estimates of the unweighted scenarios for each network for a measurement uncertainty level of 0.7 ppm.

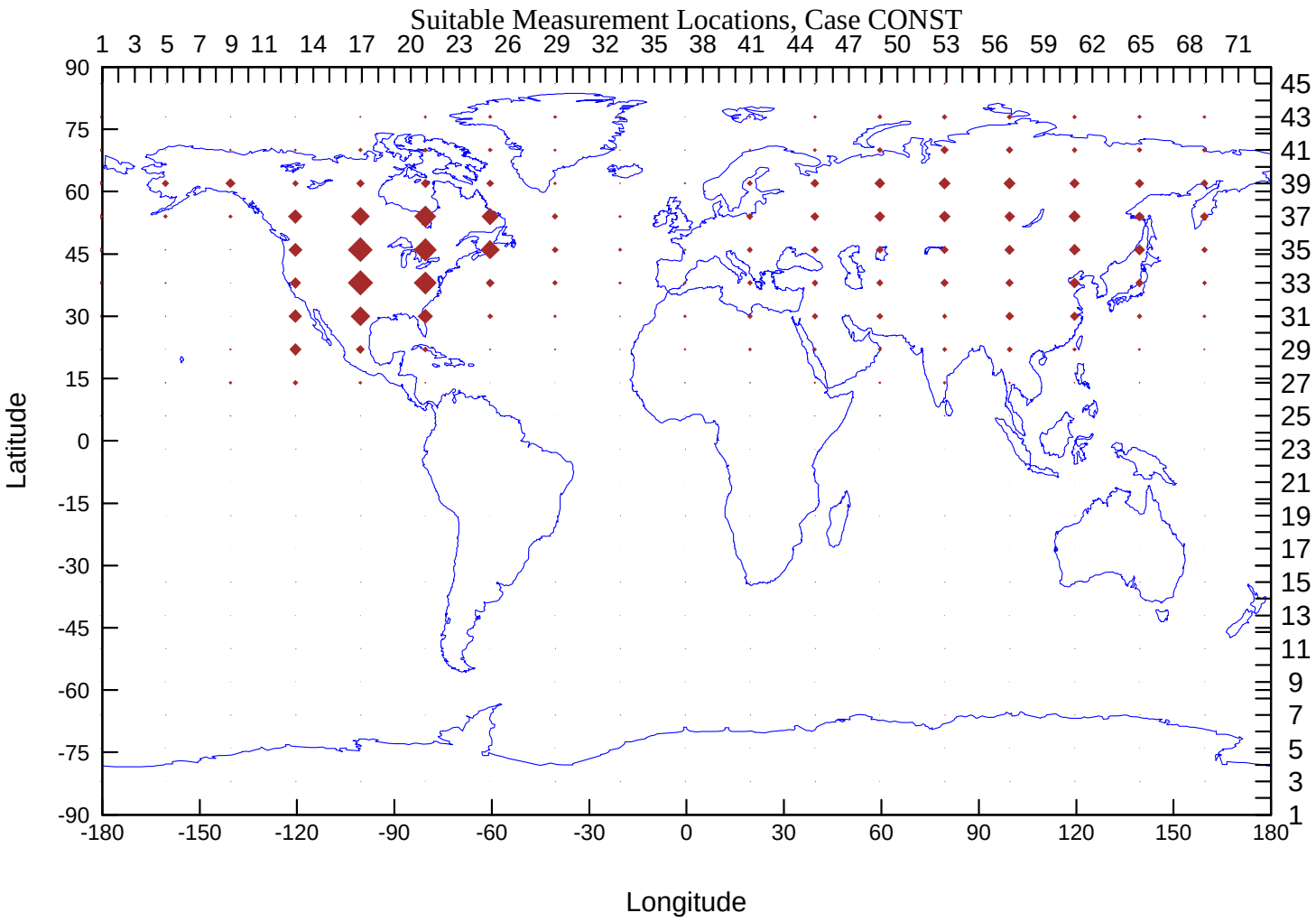

Figure 11: Location of suitable observational sites as predicted by the **Q** matrix diagnostic : Case CONST

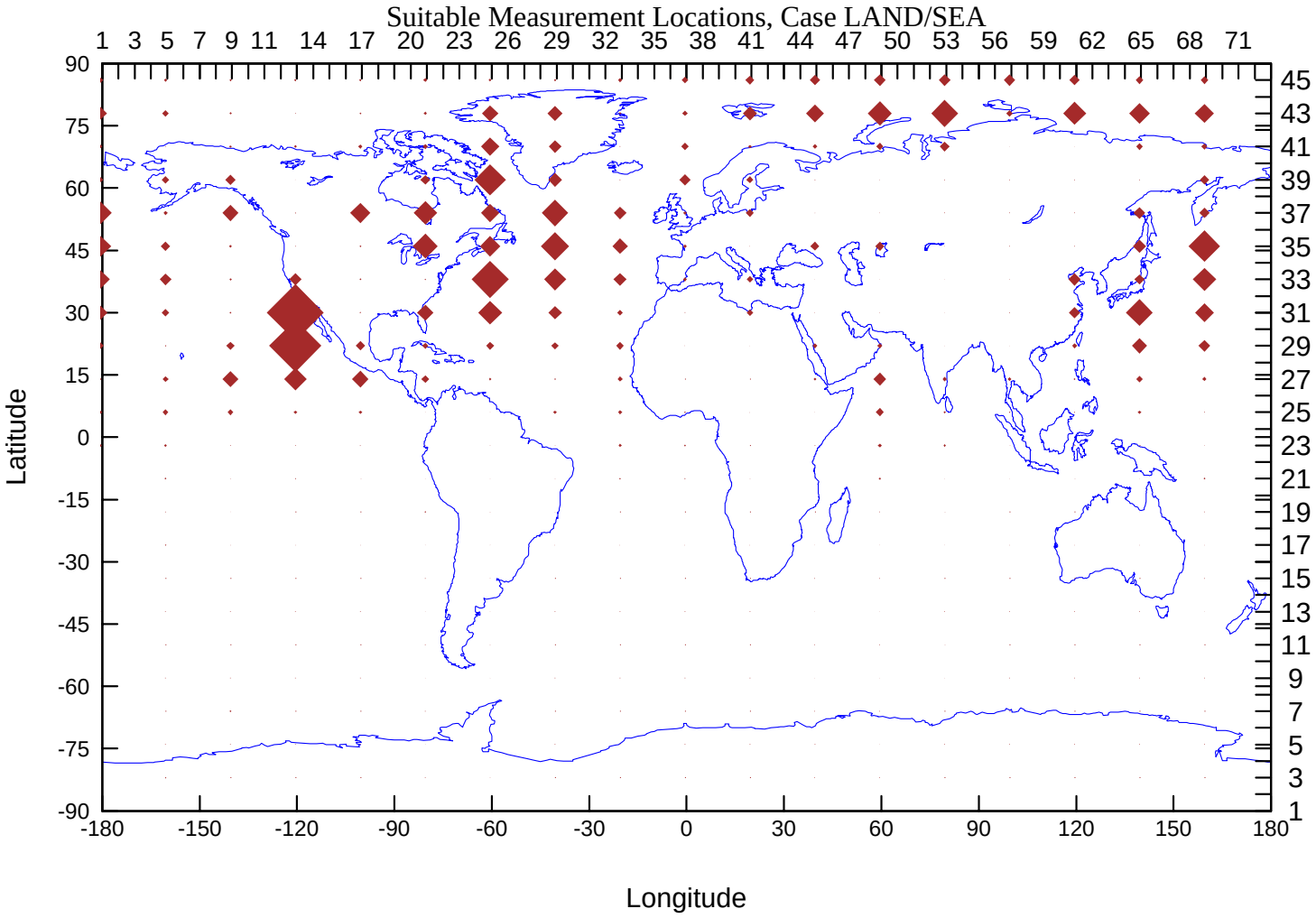

Figure 12: Location of suitable observational sites as predicted by the **Q** matrix diagnostic : Case LAND/SEA

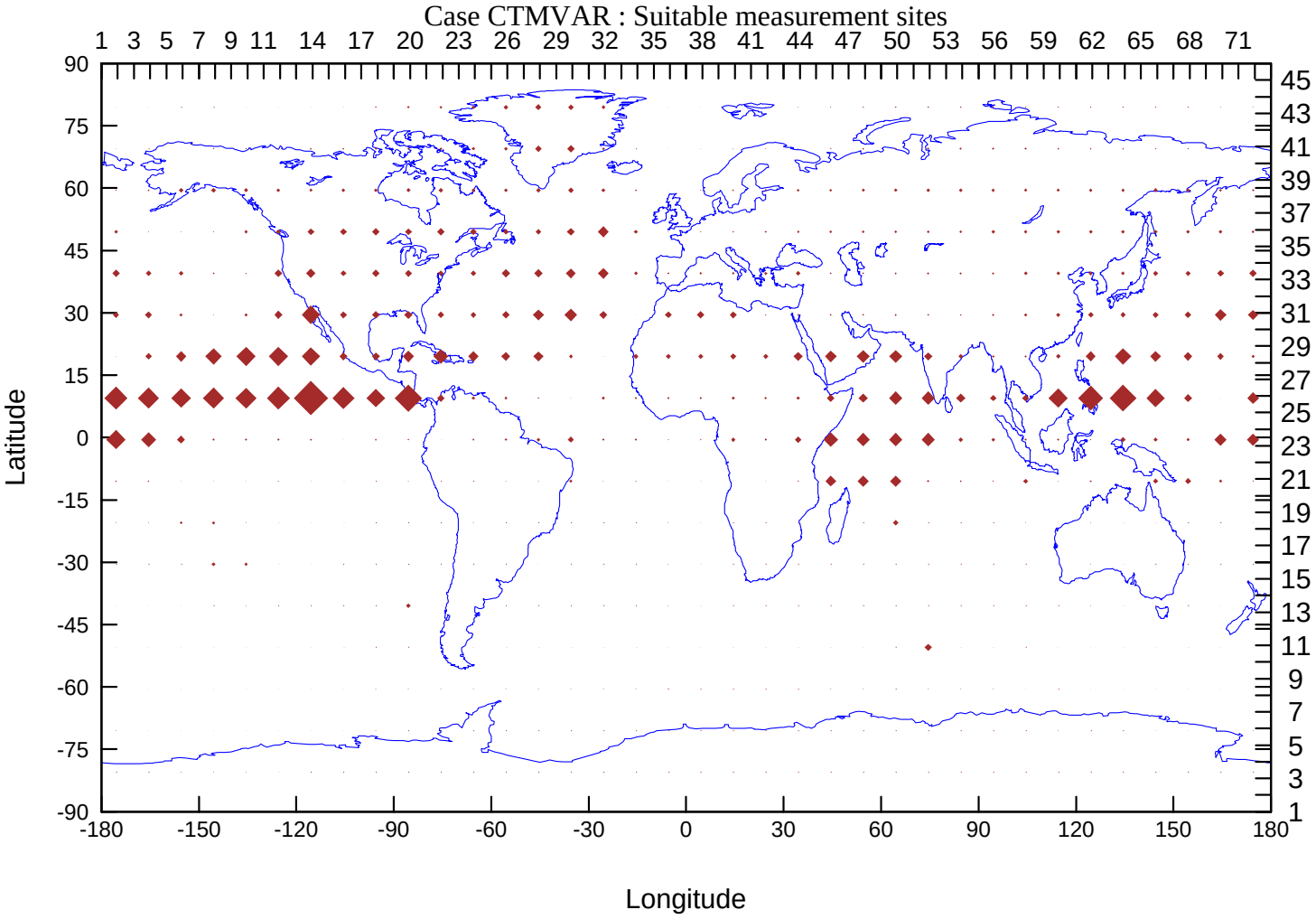

Figure 13: Location of suitable observational sites as predicted by the **Q** matrix diagnostic : Case CTMVAR