Global-mean marine $\delta^{13}C$ and its uncertainty in a glacial state estimate

Geoffrey Gebbie[a], Carlye D. Peterson[b], Lorraine E. Lisiecki[b], Howard J. Spero[c]

[a]*Department of Physical Oceanography, Woods Hole Oceanographic Institution, 360 Woods Hole Rd., MS # 29, Woods Hole, MA, 02543 USA, ggebbie@whoi.edu*
[b]*Department of Earth Science, University of California, Santa Barbara, California, USA*
[c]*Department of Earth and Planetary Sciences, University of California, Davis, California, USA*

Abstract

A paleo-data compilation with 492 $\delta^{13}C$ and $\delta^{18}O$ observations provides the opportunity to better sample the Last Glacial Maximum (LGM) and infer its global properties, such as the mean $\delta^{13}C$ of dissolved inorganic carbon. Here, the paleo-compilation is used to reconstruct a steady-state water-mass distribution for the LGM, that in turn is used to map the data onto a 3D global grid. A global-mean marine $\delta^{13}C$ value and a self-consistent uncertainty estimate are derived using the framework of state estimation (i.e., combining a numerical model and observations). The LGM global-mean $\delta^{13}C$ is estimated to be $0.14‰ \pm 0.20‰$ at the two standard error level, giving a glacial-to-modern change of $0.32‰ \pm 0.20‰$. The magnitude of the error bar is attributed to the uncertain glacial ocean circulation and the lack of observational constraints in the Pacific, Indian, and Southern Oceans. Observations in the Indian and Pacific Oceans generally have 10 times the weight of an Atlantic point in the computation of the global mean. To halve the error bar, roughly four times more observations are needed, although strategic sampling may reduce this number. If dynamical constraints can be used to better characterize the LGM circulation, the error bar can also be reduced to 0.05 to

0.1‰, emphasizing that knowledge of the circulation is vital to accurately map $\delta^{13}C_{DIC}$ in three dimensions.

Keywords: Paleoceanography, Physical Oceanography, Carbon reservoirs, Last Glacial Maximum, Inverse methods

1. Introduction

Carbon-13 to carbon-12 ratios (i.e., $\delta^{13}C$) can chemically fingerprint different carbon reservoirs, and thus glacial-interglacial changes in $\delta^{13}C$ of oceanic dissolved inorganic carbon (i.e., $\delta^{13}C_{DIC}$) reflect the carbon partitioning between terrestrial, atmospheric, and marine reservoirs. Dramatic environmental changes during the Last Glacial Maximum (LGM, 23,000 to 19,000 years before present) altered the terrestrial biosphere, and some of the low isotopic signature of terrestrial carbon ($\delta^{13}C \approx -25‰$) was transferred to the glacial ocean, consistent with observations of benthic foraminiferal $\delta^{13}C$ lower than the modern-day (e.g., Shackleton, 1977; Curry et al., 1988; Duplessy et al., 1988). The glacial atmosphere held approximately 170 gigatons (Gt) less carbon (e.g., Monnin et al., 2001), leaving the ocean as the most readily available source of compensation for the other two reservoirs. Pollen records and vegetation models that more directly reflect terrestrial carbon change yield higher estimates of glacial-to-modern carbon transfer (e.g., 750 to 1900 Gt C, Crowley, 1995; Adams and Faure, 1998; Kaplan et al., 2002) than the marine-based estimates (e.g., 330 to 650 Gt C, Shackleton, 1977; Curry et al., 1988; Duplessy et al., 1988; Köhler et al., 2010), although an inert terrestrial carbon pool may reconcile the difference (Ciais et al., 2012). A recent compilation of benthic *Cibicidoides* spp. $\delta^{13}C$ has nearly twice the data points of previous compilations and coverage of the Atlantic, Pacific, and

Indian Oceans (Peterson et al., 2014), and thus motivates the re-investigation of the marine-based whole-ocean $\delta^{13}C$ estimates.

Determining the mean value of a spatially-distributed tracer field reduces to a linear operation in most cases (i.e., an inner vector product):

$$\overline{c} = \mathbf{w}^T \mathbf{y} + \overline{c}_0 = \sum_{i=1}^{N} w_i y_i + \overline{c}_0 \tag{1}$$

where $\overline{c}$ is the global mean value of a tracer c, $\mathbf{w}$ is a vector of weights with w_i for the ith element, ^{T} is the vector transpose, $\mathbf{y}$ is a vector containing N observations of y_i, and $\overline{c}_0$ is a constant included for full generality. If all observations are assumed to contain equal information about the global mean and no other information is available (i.e., $\overline{c}_0 = 0$), the optimal weights would all be $1/N$, and equation (1) reverts to the basic sample mean. The sparse, irregularly-spaced nature of glacial observations invalidates this assumption, of course. Originally, paleoceanographers best dealt with this issue by choosing cores from what was thought to be the most representative oceanic regions (e.g., Shackleton, 1977). As more data became available, basin-wide or regional means were computed as a preliminary step before global averaging (e.g. Curry et al., 1988; Boyle, 1992; Matsumoto and Lynch-Stieglitz, 1999; Peterson et al., 2014). This multi-step process naturally leads to non-uniform weights on the observations in equation (1).

When the global-mean oceanic $\delta^{13}C_{DIC}$ is computed as a succession of sub-averages, the result may be sensitive to the size and location of the chosen sub-domains, and only by producing $\delta^{13}C$ maps at higher spatial resolution will this sensitivity be reduced. The distance between LGM observations, however, is often greater than the decorrelation lengthscale of oceanic property fields, and thus the typical method of "objectively" mapping the observations onto a regular grid (e.g., optimal interpolation or objective mapping, Bretherton et al., 1976) reverts to a

first-guess estimate in many locations. In other words, large regions of the LGM ocean would be unconstrained by the data, especially at intermediate depths where little core coverage is available. Furthermore, the objectively-mapped estimate will leave local extrema in the estimated tracer field around the data points. Such features are undesirable because they are not physically sustainable in equilibrium when diffusion has sufficient time to act (e.g., for atmospheric momentum, Hide, 1969). It is not clear, however, how equilibrated the glacial ocean was and whether eddy processes can be accurately modeled as a diffusive process. Computation of an accurate global mean is challenging even for modern-day cases, such as sealevel rise (e.g., Wunsch et al., 2007). A new method is needed to create a map with sparse LGM observations that addresses these complications.

Here we suggest that a method originally developed for estimating the oceanic water-mass distribution from sparse observations (Gebbie, 2014) is also well-suited to make three-dimensional global maps. Specifically, we combine a tracer transport model (Section 2.1) with observations (Section 2.2) to produce an LGM state estimate. Rather than using the assumed statistics of circulation length-scales, like optimal interpolation, we illustrate that the circulation itself can be used to make a gridded field (Section 2.3). The numerical model serves a dual purpose: 1) a means to readily interpret the sources, sinks, and pathways of tracer, and 2) a kinematic interpolator and extrapolator that allows large-scale information to be extracted from the observations. Here we extend the state estimation framework by deriving a self-consistent formula for the global-mean uncertainty (Section 2.4).

This work has two major results: 1) an estimate of the LGM global-mean $\delta^{13}C_{DIC}$, and 2) its uncertainty within a explicit set of assumptions. To connect

these results to deglacial climate dynamics and the carbon cycle, we reconstruct a global map of LGM $\delta^{13}C_{DIC}$ and detect a largescale, coherent pattern of LGM-to-modern changes (Section 3). The glacial-mean $\delta^{13}C_{DIC}$ uncertainty is partially attributed to the sparsity and measurement error in the observations, but also due to the difficulty in accurately modeling the LGM circulation (Section 4). Our results are discussed in the context of previous observational techniques (Section 5), especially how the observational weights in the averaging equation (1) are modified by the assumed circulation regime. We conclude by emphasizing the importance of circulation knowledge in the goal of further reducing the global-mean $\delta^{13}C_{DIC}$ uncertainty (Section 6).

2. Global LGM state estimate

The global LGM state estimate is produced by combining a kinematic tracer transport model with a global array of benthic foraminiferal observations of $\delta^{13}C$ and $\delta^{18}O$. Global, three-dimensional gridded distributions are produced for multiple tracers: $\delta^{13}C_{DIC}$, seawater $\delta^{18}O$ (i.e., $\delta^{18}O_w$), potential temperature, practical salinity, and phosphate. The model, observations, and state estimation method are detailed next.

2.1. Model

The model is a statistically steady-state conservation equation that is assumed to hold for, C, a general tracer: $\nabla \cdot (\vec{F}C) = Q$, where $\vec{F}$ is the mass flux and Q is a local source or sink. In the statistical steady state, any temporal variability that has a net diffusive or advective effect is represented by the model used here.

In practice, the model equations are discretized on a global, three-dimensional grid. Here the grid is defined with $4^\circ \times 4^\circ$ horizontal resolution and 33 vertical

levels with enhanced resolution near the surface. Glacial ocean computations are undertaken on the same grid as a modern-day reference case, but gridcells shallower than 120 meters modern-day water depth are discarded due to the sealevel drop. After discretization, the equations are normalized by the sum of all mass fluxes into the gridcell, $\phi_i = \sum_{j=1}^{N} f_{ij}$, where f_{ij} is the flux from gridcell j to i, and there are N neighboring gridcells. Then the tracer transport equation at gridcell i becomes more similar to a water-mass mixing model (following Gebbie and Huybers, 2012):

$$\sum_{j=1}^{N} m_{ij} c_j - c_i = q_i \tag{2}$$

where m_{ij} is the ratio of the inward flux from j to the total flux ($m_{ij} = f_{ij}/\phi_i$), c_i is the tracer concentration in cell i, and q_i is equilibrium tracer source with units of the tracer concentration itself ($q_i = Q/\phi_i$). For conservative tracers, the source and sink vanishes ($Q = 0$). These algebraic manipulations lead to a well-conditioned set of equations that can be solved quickly, but with the tradeoff that information is lost regarding the absolute rate of circulation.

The isotope variables, $\delta^{13}C_{DIC}$ and $\delta^{18}O_w$, require some further consideration. In particular, the sink of $\delta^{13}C_{DIC}$ due to remineralization is assumed to be equal to $-0.95‰/(\mu\text{mol/kg})$ times the source of remineralized phosphate, which is adjusted relative to the modern ratio of $-1.1‰/(\mu\text{mol/kg})$ due to changes in whole-ocean $\delta^{13}C_{DIC}$ and upper-ocean biological fractionation (e.g., following Broecker and Maier-Reimer, 1992). Here we model the ratio (delta value) rather than the individual isotopes which incurs an error (e.g., Walker, 1991), but it is small because the $^{18}O/^{16}O$ ratio in Vienna Standard Mean Ocean Water (VSMOW) standard is about $1/500$, and the $^{13}C/^{12}C$ ratio in the Vienna Pee Dee Belemnite (VPDB) standard is about $1/90$. Furthermore, this error is damped in the vicinity of obser-

vations by the formal data constraints.

For reasons that should become clear below, the state vector, $\mathbf{x}$, is defined to contain both tracer and circulation information, i.e., $\mathbf{x}^T = [\mathbf{c};\mathbf{m}]^T$, where $\mathbf{c}$ is a vector that represents all of the global three-dimensional tracer distributions and $\mathbf{m}$ describes the circulation by concatenating all of the mass-flux ratios, m_{ij} (e.g., Gebbie and Huybers, 2010). This state vector definition is not unique, but it provides sufficient information to permit a steady-state tracer distribution to be computed, and thus is an acceptable definition of the state. All of the tracer transport equations are combined and symbolically represented as: $\mathcal{L}[\mathbf{x}] = \mathbf{q} + \mathbf{v}$, where $\mathcal{L}$ is a nonlinear operator due to the multiplication of the tracer concentration and flow field that encapsulates advective and diffusive processes, and $\mathbf{v}$ is the source deviation from the modern-day first-guess field, $\mathbf{q}$. The model equation includes surface concentration (i.e., Dirichlet) boundary conditions for completeness.

2.2. Observations

A major extension to the work of Gebbie (2014) is the use of the paleo-data compilation of Peterson et al. (2014) that includes observations in the Pacific and Indian sectors, rather than the Atlantic-only data used previously. This compilation contains 376 δ^{13}C and 369 δ^{18}O measurements from benthic foraminifera dated to the Last Glacial Maximum from 23,000 to 19,000 years before present (23-19 kyr BP) following a re-derived age model for many cores (Stern and Lisiecki, 2014). The need for inter-species and interlaboratory offsets is reduced by compiling only *Cibicidoides* spp. δ^{13}C data, although there are few data of this type in intermediate waters. Other data are added (e.g., personal communication, D.W. Oppo and W. Curry, Marchal and Curry, 2008; Makou et al., 2010; Hesse et al., 2011), including porewater salinity and δ^{18}O data points (Adkins et al., 2002).

The deglacial records indicate a group of outliers where the LGM-to-Holocene $\delta^{18}O$ change is less than 0.6‰ and can be traced to low temporal resolution in the cores. Therefore we have removed these values from the compilation. Observations were also culled when the phytodetritus effect was implicated by the original authors (e.g., Mackensen et al., 2000) and at locations that fall more than 200 km outside of the model grid. When combining the Peterson et al. (2014) compilation with the additional data, we have 492 LGM $\delta^{13}C$ and 492 $\delta^{18}O$ data points that constrain the model simulation (locations are later shown in Figure 8).

The collection of observational equations is expressed as one matrix equation: $\mathbf{y} = \mathbf{E}\mathbf{x} + \mathbf{n}$, where $\mathbf{y}$ is a vector of observations, $\mathbf{E}$ predicts the observations from the state, $\mathbf{x}$, and $\mathbf{n}$ is the observational error. We assume that benthic foraminiferal $\delta^{13}C$ refects $\delta^{13}C_{DIC}$ (e.g., Duplessy et al., 1984) and that the error in this assumption is normally distributed with a standard deviation of 0.2‰ (Marchal and Curry, 2008). We check this assumption by calculating the misfit between the Late Holocene core data and a projection of the modern-day $\delta^{13}C_{DIC}$ distribution of Gebbie (2014) onto the core sites. The standard deviation of the misfit is 0.27‰, suggesting that 0.2‰ is a reasonable guess of this unknown quantity. Systematic errors in this relationship will be addressed in Section 4.6.

2.3. State estimation method

The solution method is started with prior knowledge of the state as encapsulated in a first guess, $\mathbf{x}_0$. Here we use modern-day property distributions, but where $\delta^{18}O_w$ and practical salinity are adjusted higher by 1.1‰ and 1.1 on the practical salinity scale, respectively, to account for glacial sealevel drop. The first-guess constraint is equivalently written as an equation: $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$, where $\mathbf{u}$ is a deviation from the first guess that is permitted to be large (see the weight matrices

below). Then, the solution state is determined by minimizing a cost function that combines the constraints from this and the previous two sections (2.1 and 2.2)

$$J = \mathbf{u}^T \mathbf{S}^{-1} \mathbf{u} + \mathbf{n}^T \mathbf{W}^{-1} \mathbf{n} + \mathbf{v}^T \mathbf{Q}^{-1} \mathbf{v}, \tag{3}$$

where $\mathbf{S}, \mathbf{W}$, and $\mathbf{Q}$ are matrices that provide the relative weightings (see Appendix A for the chosen values relevant for equation 3). The three terms on the right hand side represent the three major constraints: 1) prior information about the tracer distribution and circulation, 2) proxy observations, and 3) the tracer transport model.

The complete solution method for the global $\delta^{13}C_{DIC}$ distribution was reported by Gebbie (2014) and includes a number of other constraints (and, hence, terms in the cost function), such as gravitational stability and the non-negativity of tracer concentrations. To recap, this weighted, tapered least-squares problem is solved by transforming equation (3) into a constrained cost function by appending Lagrange multiplier terms, and then hand-coding the adjoint equations that give sensitivity information. It is solved iteratively using a limited-memory quasi-Newton gradient descent routine (Nocedal, 1980). After solving for the $\delta^{13}C_{DIC}$ state, we seek a formula for the global-mean uncertainty that is valid in the neighborhood of the known solution. The final solution is gravitationally stable and has non-negative tracer distributions, indicating that the additional cost function terms are not actively constraining the solution at the final iteration. Thus we simplify the uncertainty problem by assuming that equation (3) represents all active constraints, and derive an uncertainty formula under such an assumption next.

2.4. Global-mean uncertainty of the state estimate

Many methods for computing error bars are available, and we explicitly define what is meant by uncertainty in this work by following the development of Wunsch (1996). The uncertainty, $P_{\overline{c}}$, is defined as the expected squared difference between the estimated global mean, $\tilde{\overline{c}}$, and the true value, $\overline{c}$ (i.e., $P_{\overline{c}} \equiv < (\tilde{\overline{c}} - \overline{c})^2 >$), where the brackets indicate the expected value operator. For an unbiased estimator, the true solution, $\overline{c}$, may be set equal to the expected value that would emerge from our estimation method over many different realizations, $< \tilde{\overline{c}} >$, but here we restrain from such an interpretation given the highly nonlinear nature of the problem and the simplified nature of the model. Thus we are restricted to solving for the expected range of solutions, here defined as the dispersion of $\tilde{\overline{c}} - < \tilde{\overline{c}} >$ or equivalently the covariance of the global mean, $C_{\overline{c}} \equiv < (\tilde{\overline{c}} - < \tilde{\overline{c}} >)^2 >$. The standard error, $\sigma_{\overline{c}}$, is then defined as the square root of the solution covariance: $\sigma_{\overline{c}} \equiv \sqrt{C_{\overline{c}}}$. We follow the convention of Peterson et al. (2014) by quoting twice the standard error as our estimate of the glacial-mean $\delta^{13}C_{DIC}$ uncertainty, which can naturally be interpreted as the 95% confidence interval of a normally-distributed process.

Before calculating the uncertainty of the global mean, we first define the global mean explicitly as

$$\overline{c} = \mathbf{r}^T \mathbf{x}, \tag{4}$$

where $\mathbf{r}$ is the appropriately-defined mass-weighting vector (i.e., $r_i = M_i / \sum_{j=1}^{J} M_j$ for all $i \leq J$ where M_i is the mass of gridcell i and the global ocean has J gridcells, and $r_i = 0$ for all $i > J$). By construction, the sum of elements of $\mathbf{r}$ is equal to one ($||\mathbf{r}||_1 \equiv 1$) as is usual in averaging equations. As three-dimensional fields are modeled for temperature and salinity, the mass of each gridcell is determined

by the product of the cell volume and the seawater density as calculated by the international thermodynamic equation of state (IOC, 2010). Although the LGM density field is not well constrained, the spatial range of density has variations no larger than 5%, and thus volume (set by the size of the gridcells) dominates the calculation.

Substituting the global-mean equation (4) into the definition of the global-mean covariance, the uncertainty of the global mean is clearly dependent upon the uncertainty over the entire globe:

$$\mathrm{C}_{\overline{c}} = \mathbf{r}^T \mathbf{C}_{\tilde{\mathbf{x}}} \mathbf{r}, \tag{5}$$

where $\mathbf{C}_{\tilde{\mathbf{x}}}$ is the full solution covariance. Thus it is necessary to determine the uncertainty of the three-dimensional tracer distribution to calculate the uncertainty of the global mean. Following Appendix B, the solution covariance is

$$\mathbf{C}_{\tilde{\mathbf{x}}} = (\mathbf{S}^{-1} + \mathbf{E}^T \mathbf{W}^{-1} \mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T \mathbf{Q}^{-1} \mathbf{L}_{\tilde{\mathbf{x}}})^{-1}, \tag{6}$$

where $\mathbf{L}_{\tilde{\mathbf{x}}}$ is the model linearized about the state, $\tilde{\mathbf{x}}$ (i.e., $\mathbf{L}_{\tilde{\mathbf{x}}} \equiv \partial \mathcal{L} / \partial \mathbf{x}]_{\tilde{\mathbf{x}}}$). Equation (6) assumes that the weight matrices are chosen to be the expected second-moment matrices of the residuals (i.e., $\mathbf{R}_{nn} = \mathbf{W}$, $\mathbf{R}_{qq} = \mathbf{Q}$, and $\mathbf{R}_{xx} = \mathbf{S}$, following Wunsch (1996)), as well as the assumption that our simplified cost function is valid.

Substitution of equation (6) into (5) permits the global-mean uncertainty to be written explicitly in terms of the known input variables in the problem:

$$\mathrm{C}_{\overline{c}} = \mathbf{r}^T (\mathbf{S}^{-1} + \mathbf{E}^T \mathbf{W}^{-1} \mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T \mathbf{Q}^{-1} \mathbf{L}_{\tilde{\mathbf{x}}})^{-1} \mathbf{r}. \tag{7}$$

Equation (7) illustrates that the three contributions to the uncertainty come from the three constraints: the first-guess uncertainty in the state ($\mathbf{S}^{-1}$), the uncertainty

related to the observations ($\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E}$), and the uncertainty related to the imperfect model ($\mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}}$). While these three terms have clear contributions to the uncertainty, they are under the inverse in the equation and thus the total uncertainty is not simply a linear combination of the three parts.

3. LGM δ^{13}C and deglacial change

3.1. Global-mean $\delta^{13}C$

The glacial-mean δ^{13}C and its uncertainty are calculated by adapting equations (4) and (7) for $\delta^{13}C_{DIC}$. The full solution uncertainty (equation 6) is never needed explicitly, thus avoiding the storage of a 1.8 terabyte matrix (the state has 475773 elements and $\mathbf{C}_{\tilde{\mathbf{x}}}$ has this dimension squared). As $\mathbf{r}$ is a column vector, we can break equation (7) into two parts, one a matrix-vector product and one a vector inner product, such that memory usage is minimized. We find that the LGM mean $\delta^{13}C_{DIC}$ is $\overline{\delta^{13}C_{DIC}}^{G} = 0.14‰ \pm 0.20‰$ (at the 2σ uncertainty level). This uncertainty estimate accounts for the observational sparsity and measurement error, but does not include errors incurred by the assumed δ^{13}C proxy equation (to be addressed later in Section 4.6).

We are motivated by the atmospheric CO_2 change of 80-100ppm from the LGM to the pre-industrial era, and thus we seek to estimate the deglacial marine δ^{13}C change, here denoted $\Delta_{MG}[\delta^{13}C_{DIC}]$ (i.e., modern, "M," minus glacial, "G"). A modern-day reference circulation that attempts to reconstruct the pre-industrial ocean was produced using the same method (Gebbie, 2014) and will be used for comparison here. The GLODAP and CARINA seawater $\delta^{13}C_{DIC}$ compilation (Schmittner et al., 2013) constrained the modern-day distribution, and observations in the upper 1 km of the water column were downweighted to ac-

count for the Suess effect (e.g., Olsen et al., 2006) and to produce a pre-industrial gridded $\delta^{13}C_{DIC}$ distribution. The result is a reference modern-day estimate of $\overline{\delta^{13}C_{DIC}}^{M} = 0.47‰ \pm 0.03‰$. Uncertainty due to the Suess effect is not included in this error bar, and is revisited in more detail in Section 4.3 where we find that it may increase the uncertainty by approximately $0.05‰$.

Under the assumption that the LGM and modern estimates are independent, the LGM-to-modern difference is $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.32‰ \pm 0.20‰$ (at the 2σ level after rounding). Our estimate provides corroborating evidence for both 1) recent data-based estimates of $0.34‰ \pm 0.26‰$ (Ciais et al., 2012) and $0.34‰ \pm 0.19‰$ (Peterson et al., 2014), and 2) a forward circulation model that reproduced about 100 observations and found a change of $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.31‰ \pm 0.20‰$ (Tagliabue et al., 2009). Note that the data-based estimates were actually glacial-to-Late Holocene differences (here distinguished by the notation, $\Delta_{HG}[\delta^{13}C_{DIC}]$) that may reflect a different quantity due to the core information ending over the last few thousand years of the Holocene. The level of agreement is surprising, especially considering that our estimate results from a spatially structured pattern of change, shown next.

3.2. Spatial distribution of LGM $\delta^{13}C$

The glacial state estimate faithfully reproduces the LGM $\delta^{13}C$ observations in all ocean basins. When comparing basin-wide zonally averaged $\delta^{13}C_{DIC}$ to the foraminiferal observations (squares versus the background colors in Figure 1), the spatial pattern is in broad agreement. Any visual differences are attributable either to an actual misfit of the data or an artifact of zonal variability (i.e., the collapsed third dimension). To eliminate the effect of zonal variability, the state estimate is projected onto the observational locations using a linear interpolation

or extrapolation from the 8 nearest gridpoints. The resulting $\delta^{13}C_{DIC}$ distribution fits the observations with a standard deviation of the misfit of 0.19‰ that is an acceptable fit within the expected uncertainty of 0.2‰ in the data. 50% of the points are fit within 0.02‰, although outliers larger than ±0.6‰ exist. While not being the primary variable of interest here, $\delta^{18}O$ is also fit well, with a standard deviation of 0.22‰ that is approximately equal to its expected value of 0.2‰.

The main characteristic of the LGM Atlantic $\delta^{13}C$ distribution is a depletion of $\delta^{13}C_{DIC}$ below 2500 meters depth (upper panel, Figure 1) that conforms with expectations (e.g., Duplessy et al., 1988; Curry and Oppo, 2005). The North Atlantic is characterized by a mixing zone from 2 to 4 km depth with a nearly constant vertical gradient from 0.8‰ to 0‰. The zonally-averaged South Atlantic, however, does not have a gradient as sharp as seen in the Brazil Margin data or the map of Curry and Oppo (2005), but can be explained by observations of lower $\delta^{13}C$ values in the eastern South Atlantic between 2 and 3 km depth. Of more concern are the undersampled regions, where this inversion sometimes disagrees with the Gebbie (2014) inversion. For example, the updated state estimate has a much higher $\delta^{13}C_{DIC}$ in the Arctic (1.2‰ compared to 0.6‰). The most depleted values in the Southern Ocean are about -0.8‰ in both inversions, but should not necessarily be interpreted as a robust result. The uncertainty of the estimate in these regions without data will be explicitly addressed in the next section.

The state estimate permits the inference of more $\delta^{13}C_{DIC}$ structure in the Pacific and Indian Oceans (middle and lower panels, Figure 1) than was previously mapped, due in large part to the larger data compilation and the smoothing capabilities of the model used here. The zonally-averaged Pacific picture masks significant zonal variability, including a deepening of the most depleted western

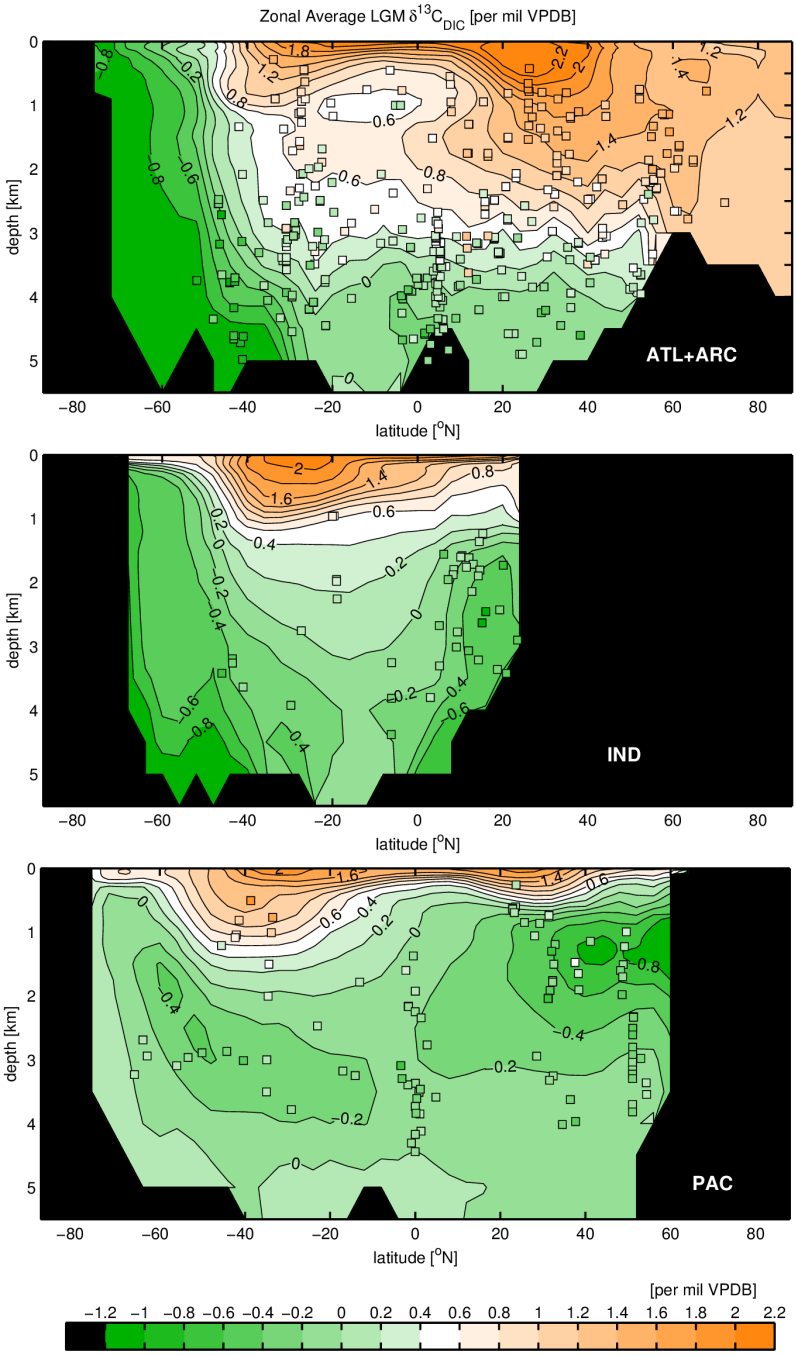

Figure 1: Zonally-averaged, LGM $\delta^{13}C_{DIC}$ divided into three sections: the Atlantic and Arctic (*upper*), Indian (*middle*), and Pacific (*lower*), with all observations from the particular basin (*colored squares*) and 3D gridded field (*background colors and contours*). The oceans are divided according to Figure 7 of Gebbie and Huybers (2010). The colored symbols are on the same color scale as the background field.

Pacific $\delta^{13}C_{DIC}$ from 1.5 km in the modern-day to almost 3 km during the LGM (e.g., Matsumoto et al., 2002; Herguera et al., 2010). There are significant differences with previous maps, however, including a closed region of depleted $\delta^{13}C_{DIC}$ due to remineralization north of 40°N in the North Pacific not seen by Herguera et al. (2010). The Indian Ocean is reconstructed as a hybrid of the other oceans, with low-$\delta^{13}C_{DIC}$ Southern Ocean bottom waters like the Atlantic, but a lack of high-$\delta^{13}C_{DIC}$ waters at mid-depths like the Pacific.

The bottom waters of the Southern Ocean have $\delta^{13}C_{DIC}$ differences approaching 1‰ between the Ross (0‰) and Weddell (-0.8‰) Seas that are masked by the zonal averages. These horizontal gradients are partially explained as the projection of the large vertical $\delta^{13}C_{DIC}$ gradients into the horizontal by the patchiness of northern source water incursion into the Southern Ocean. The inhomogeneities do not appear to reflect any major shift in the transport of the Antarctic Circumpolar Current, as northern-source waters continue to spread eastward into the Indian Ocean over a similar range as the modern day. While major circulation changes are not implicated, the model reconstructs a greater filling of the abyss by the Ross Sea rather than the Weddell Sea. Increased glacial abyssal mixing and a vigorous deep circulation could explain this feature (Wunsch, 2003; Arbic et al., 2004; Schmittner et al., 2015), but our current model lacks absolute rate information and cannot directly test this hypothesis.

Besides the $\delta^{13}C$ distribution, the state estimate permits the diagnosis of water-mass distributions consistent with the seawater properties. The inferred LGM Atlantic water-mass geometry is similar to the previous inversion of Gebbie (2014), where southern source waters dominate the ocean composition only below 4 km depth in the North Atlantic. In other words, the southern source waters contribute

50% or more of the water by mass in that limited region. Significant deep remineralization (with the addition of ^{12}C effectively causing a sink of $\delta^{13}C$) was originally reported in the deep glacial Atlantic (Gebbie, 2014), but is less pronounced in the updated state estimate due to the use of a more spatially-coherent $\delta^{18}O$ dataset (derived from LGM-to-Late Holocene differences). The estimated LGM Pacific has deeper North Pacific Water (down to 2 km instead of 800 m today, as defined by the 50% concentration line). Otherwise southern water masses fill the same part of the Pacific sector.

3.3. Spatial pattern of $\delta^{13}C$ change

The difference of the modern-day and LGM $\delta^{13}C_{DIC}$ permits the mapping of a global, three dimensional field of $\Delta_{MG}[\delta^{13}C_{DIC}]$, where large-scale coherent patterns emerge (Figure 2). In much of the world ocean above 1 km depth, $\delta^{13}C_{DIC}$ decreased over the deglaciation despite the increase in the global mean. Even the North Pacific, which was previously found to have little $\delta^{13}C_{DIC}$ change above 2 km (Matsumoto et al., 2002), is estimated to contain upper ocean regions with deglacial $\delta^{13}C$ decreases. In general, $\Delta_{MG}[\delta^{13}C_{DIC}]$ increases with depth and toward the south, with the biggest changes reserved for the Atlantic sector of the Southern Ocean. From these maps, it is clear that the competing influence of many different regions must be included in order to accurately assess the the global mean.

4. Contributions to uncertainty

4.1. Observational sensitivity

Obvious candidates to contribute to the uncertainty include the sparsity, measurement error, and representativeness of the observations. Here we test how the

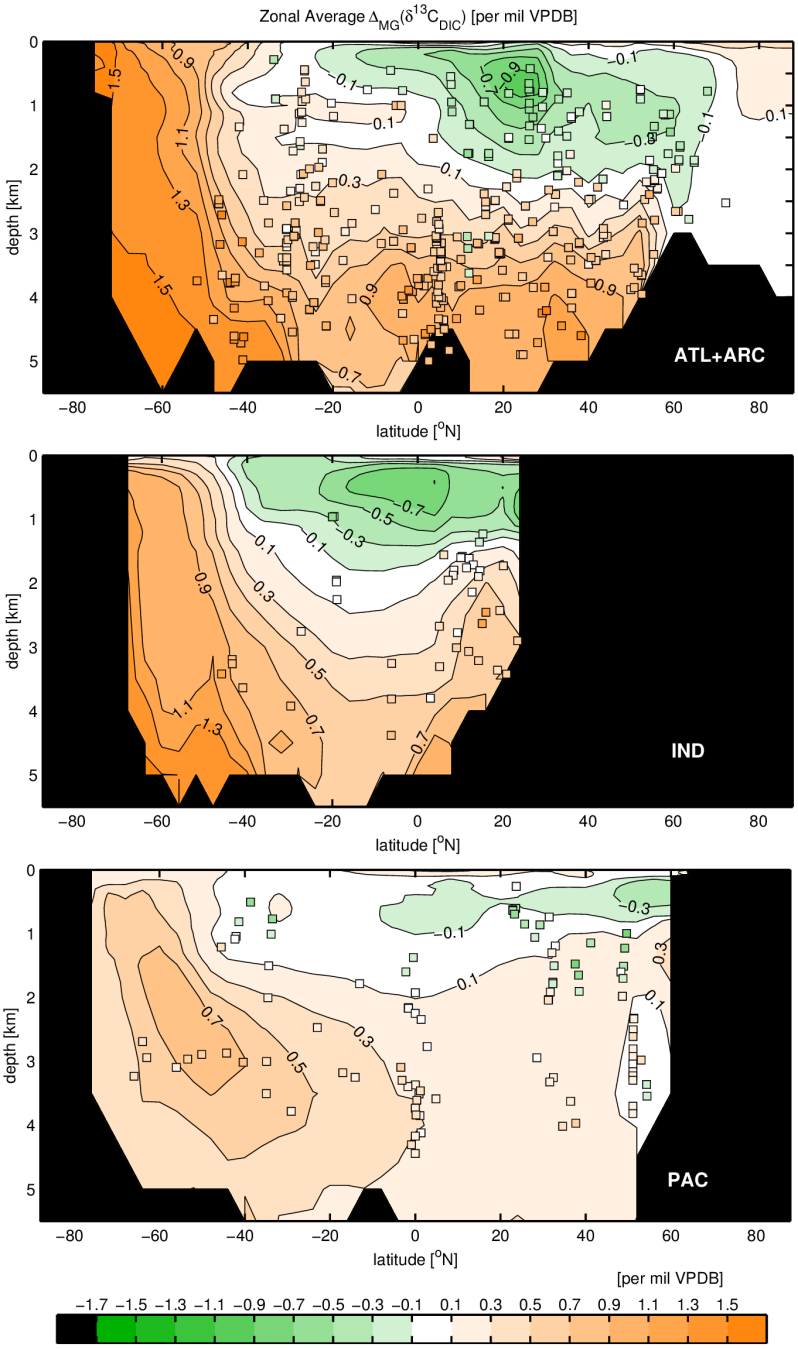

Figure 2: Similar to Figure 1, but for the Atlantic, Indian, and Pacific zonal-average difference in $\delta^{13}C_{DIC}$ between the modern-day and LGM (i.e., $\Delta_{MG}[\delta^{13}C_{DIC}] = \delta^{13}C^{M}_{DIC}$ - $\delta^{13}C^{G}_{DIC}$).

number of sediment core observations affects the resulting uncertainty by creating hypothetical cases with a range of observations from 10 to 15000. For the cases where the number of observations is less than observed in reality, we randomly select a subset of the true observations. This constraint is symbolically written: $\breve{\mathbf{y}} = \breve{\mathbf{E}}\mathbf{x} + \breve{\mathbf{n}}$, where the variables with the breve mark have rows deleted from the original $\mathbf{E}$ definition (second term on the right hand side of equation 7). The actual observational values do not matter in this calculation, but only that the observational term is modified. For hypothetical cases with more observations than reality, we augment the actual observations with additional observations taken randomly from the seafloor, although this is not entirely realistic due to issues with carbonate preservation, species habitats, and sediment availability. Five trials are performed with the modified $\breve{\mathbf{E}}$ matrix and the mean of the trials is reported to help make the results more robust.

The uncertainty decreases with an increasing number of observations according to an apparent power law ($\sigma_{\bar{c}} \approx 2‰/\sqrt{N}$), where N is the number of observations (Figure 3). This function is consistent with the uncertainty estimate for the actual number of observations ($2\sigma_{\bar{c}} = 0.20‰$ for $N = 492$). Between $N = 500$ and $N = 1000$, the slope of the power law increases, suggesting that the additional randomly-distributed data points potentially sample the ocean more efficiently than the irregularly-clustered 492 data points that are actually available, and indicating that a strategic sampling plan (e.g., see Section 4.5) could produce an even greater reduction in uncertainty. Limitations on the presence of *Cibicidoides* calcite on the seafloor would also provide a constraint for a sampling strategy. In tests where the number of observations is increased past 15000, the standard error does not decrease below 0.06‰, suggesting that seafloor observa-

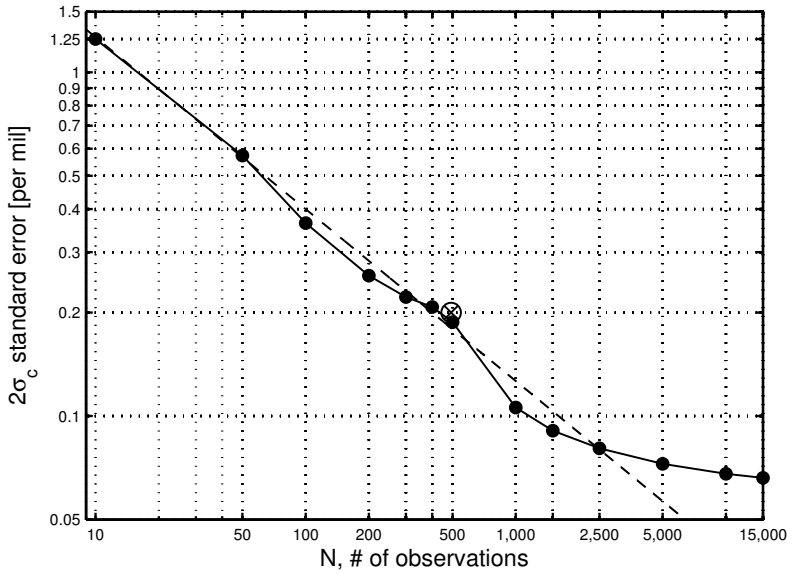

Figure 3: The $2\sigma_{\bar{c}}$ standard error of the glacial-mean $\delta^{13}C_{DIC}$ as a function of number of observations, on a log-log set of axes. Each dot represents the average of five trials. For $N = 10$ to $N = 400$, a subset of actual observational locations is used. For $N = 500$ to 15000, the 492 actual observations are augmented with observations randomly distributed along the lateral and bottom oceanic boundaries. A power law (*dashed line*) approximately fits the function: $\sigma_{\bar{c}} \approx 2‰/\sqrt{N}$ where N is the number of observations. The LGM state estimate is also plotted (*circle with an X*).

tions are eventually limited in their ability to record interior ocean signals due to their location and the remaining measurement error. In summary, we expect that the uncertainty on global-mean $\delta^{13}C_{DIC}$ would be halved with a fourfold increase in the number of observations.

4.2. Circulation dependence

A second major contributor to uncertainty is the circulation field because knowledge of the flow paths permits extrapolation over the large distances between paleo-data. In our steady-state scenario, the mass fluxes are expressed as

dimensionless mass flux ratios, **m**, that are bounded by 0 and 1 (recall the discussion of the model in Section 2.1). Here we calculate how the global-mean standard error depends upon the assumed uncertainty in the circulation. All input variables are kept constant in equation (7) except the part of the $\mathbf{S}^{-1}$ matrix corresponding to the circulation is adjusted. For hypothetical cases where the circulation is known as well or better than the modern-day ($\sigma_m < 0.05$), the resulting global mean uncertainty is small (0.03‰), but it cannot be further reduced due to the observational characteristics (Figure 4). This limit is lower than previously estimated (e.g., 0.26‰, Ciais et al., 2012), and might result from the improved reconstruction ability of the state estimate. For circulation uncertainties greater than 0.05, the standard error increases from 0.05‰ to 0.6‰ with increasing circulation uncertainty, and follows an approximate exponential relation (i.e., $\sigma_{\bar{c}} \approx 0.6‰\ \sigma_m^{0.9}$). The prior estimate of circulation uncertainty, σ_m, is a difficult quantity to estimate, and here we have suggested that a reasonable value is 0.3. The reported error of global-mean $\delta^{13}C_{DIC}$ is sensitive to this choice, however, as is shown in Figure 4. The sensitive dependence originates from the use of the circulation to interpolate and extrapolate over the data-void regions of the globe.

4.3. Regional analysis

A more traditional means of analysis is to break the ocean into subdomains to quantify the geographic contributions to global-mean uncertainty. Here, regional means are calculated by taking subdomains of the global domain and recalculating the mass-weighting vector, **r**, in equation (7). Specifically, the elements of **r** that correspond to locations outside the region of interest are set to zero, and the vector is renormalized such that the elements sum to one. Here we select 13 regions of interest in order to compare to the recent work of Peterson et al. (2014).

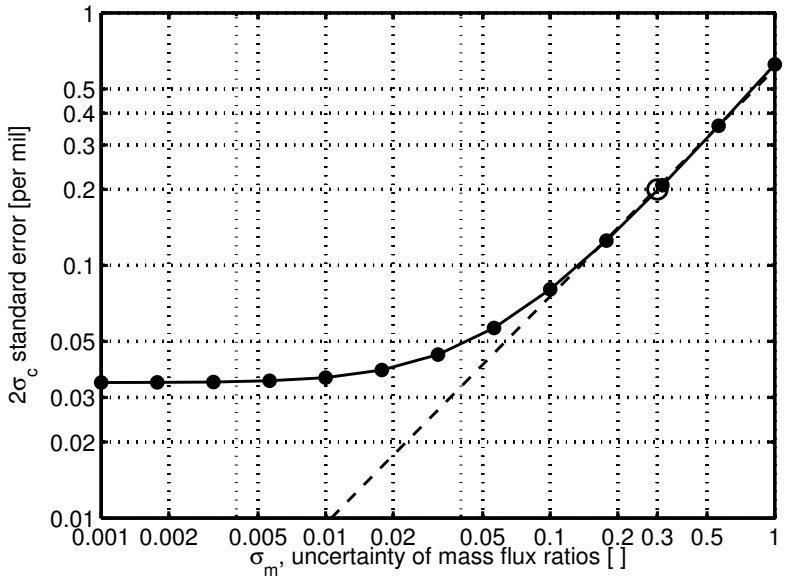

Figure 4: Similar to Figure 3, but the standard error ($2\sigma_{\bar{c}}$) of the glacial-mean $\delta^{13}C_{DIC}$ as a function of how well the circulation is known (in terms of the uncertainty of nondimensional mass-flux ratios, σ_m). An uncertainty of $\sigma_m = 1$ in the mass-flux ratios represents zero circulation knowledge. The glacial circulation uncertainty is chosen to be 0.3 (*circle*). The function roughly asymptotes to the function, $\sigma_{\bar{c}} \approx 0.3‰\ \sigma_m^{0.9}$ (*dashed line*).

Region	$\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$	$\overline{\delta^{13}C_{DIC}}^{M}$	N_M	μ_M	σ_M	$\overline{\delta^{13}C_{DIC}}^{G}$	N_G	μ_G	σ_G
Global	0.32 ± 0.20	0.47 ± 0.03	19922	0.09	0.27	0.14 ± 0.20	492	0.01	0.19
Deep	0.49 ± 0.23	0.37 ± 0.02	229	0.05	0.21	-0.12 ± 0.22	0	-	-
Surface	0.07 ± 0.71	1.06 ± 0.07	8453	0.21	0.36	1.01 ± 0.70	6	-0.01	0.05
NW Atlantic	0.16 ± 0.67	0.96 ± 0.06	321	0.11	0.20	0.80 ± 0.67	72	-0.02	0.18
NE Atlantic	0.25 ± 0.85	0.99 ± 0.25	540	0.10	0.15	0.74 ± 0.81	155	0.01	0.22
SW Atlantic	0.41 ± 0.31	0.77 ± 0.02	836	-0.03	0.14	0.36 ± 0.31	45	0.06	0.23
SE Atlantic	0.60 ± 0.21	0.72 ± 0.05	274	0.00	0.13	0.12 ± 0.20	79	0.04	0.23
South Atlantic	0.84 ± 0.40	0.62 ± 0.06	23	0.21	0.20	-0.22 ± 0.40	23	0.02	0.27
North Pacific	0.13 ± 0.27	-0.08 ± 0.02	1254	0.03	0.22	-0.21 ± 0.27	65	-0.01	0.11
South Pacific	0.33 ± 0.33	0.35 ± 0.01	4157	0.00	0.15	0.02 ± 0.33	36	0.00	0.08
Indian	0.25 ± 0.32	0.34 ± 0.01	2735	-0.04	0.12	0.09 ± 0.32	42	0.02	0.12
Southern (AI)	1.22 ± 0.85	0.54 ± 0.03	648	0.06	0.10	-0.68 ± 0.85	0	-	-
Southern (P)	0.44 ± 0.30	0.45 ± 0.02	401	0.01	0.10	0.01 ± 0.30	1	0.07	-

Table 1: Mean $\delta^{13}C_{DIC}$ with error estimates and statistics for 13 oceanic regions defined by Peterson et al. (2014). The quantities include (*from left column to right*): global-mean LGM-to-modern $\delta^{13}C$ change and 2σ uncertainty ($\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$), modern-day mean $\delta^{13}C_{DIC}$ and 2σ uncertainty ($\overline{\delta^{13}C_{DIC}}^{M}$), number of modern-day observations (N_M), mean modern-day model-data misfit (μ_M), standard deviation of modern-day model-data misfit (σ_M), LGM mean $\delta^{13}C_{DIC}$ and 2σ uncertainty ($\overline{\delta^{13}C_{DIC}}^{G}$), number of LGM observations (N_G), mean LGM model-data misfit (μ_G), and standard deviation of LGM model-data misfit (σ_G). Three large-scale regions are included: Global, Deep (everywhere below 5 km depth), and Surface (everywhere shallower than 500 m). The Atlantic is split into five regions: NW (west of 33°W, north of 0°), NE (east of 33°W, north of 0°), SW (west of 15°W, 0° to 55°S), SE (east of 15°W, 0° to 55°S), and South Atlantic (east of 22°W, 40°S to 55°S). The Pacific is split into two regions: North Pacific (0° to 60°N) and South Pacific (0° to 66°S). The Indian Ocean is defined as one region (north of 55°S, 30°E to 125°E). The Southern Ocean is split into two parts: Atlantic-Indian (AI) sector (south of 55°S) and the Pacific (P) sector (south of 66°S). See Figure 1 of Peterson et al. (2014) for complete boundaries.

The primary geographic contributors to global-mean uncertainty are the Surface and North Atlantic regions ($2\sigma_{\overline{c}} > 0.6‰$ in Table 1). This result does not straightforwardly proceed from a consideration of the number of observations in any given region. For example, both the Surface and Deep regions have less than 10 observations, yet the Deep region has much more moderate uncertainty (0.23‰ versus 0.70‰ in the Surface). Waters that enter the Deep region must pass through observations at shallower depths, and are therefore somewhat constrained by the tracer transport model and observations. The uncertainty in the North Atlantic occurs despite over 200 LGM observations in those regions, and results from the nearly-unobserved $\delta^{13}C_{DIC}$ values in the Nordic, Mediterranean, and Caribbean Seas. The center of the North Atlantic, on the other hand, has some of the lowest estimated errors, but this is masked by our choice of regional boundaries.

Up to this point, we have emphasized the agreement of our global-mean estimate with previous studies, and generally speaking, our regional results are also consistent with previous regional estimates (e.g., Oliver et al., 2010; Peterson et al., 2014). In a detailed investigation of the Deep ($>$ 5km) region, however, our model reconstructs a 0.49‰ change, suggesting that the 0.74‰ estimate of Peterson et al. (2014) is an overestimate owing to the extrapolation by their assumed linearly-varying vertical structure. In the Surface region, there are similarities between the estimates (here: $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.07‰ \pm 0.71‰$, Peterson et al. (2014): $\Delta_{HG}[\overline{\delta^{13}C_{DIC}}] = 0.02‰ \pm 0.40‰$), but our more sophisticated reconstruction technique yields larger uncertainty, suggesting that their ad-hoc error bar is an underestimate.

Our large Surface region error bars indicate that the best place to isolate a reservoir from benthic foraminiferal detection is not the bottom ocean, but instead

the upper thermocline waters that primarily recirculate in the subtropics and tropics. This result points to the need to compile planktonic $\delta^{13}C_{DIC}$ records (e.g., Broecker and McGee, 2013) so that they can be used in concert with benthic records. In this regard we note that there is convincing evidence indicating the $\delta^{13}C$ of planktonic foraminifera shells vary with seawater carbonate ion concentration (Spero et al., 1997; Russell and Spero, 2000; Peeters et al., 2002). Although this effect has not been identified in benthic foraminifera, its pervasive presence among many species of planktonic foraminifera suggests that deep water carbonate ion variations between the modern and glacial could shift benthic foraminifera shell $\delta^{13}C$ away from a 1:1 relationship with $\delta^{13}C_{DIC}$ and contribute to uncertainties in these modeled reconstructions (see Section 4.6).

4.4. $\delta^{13}C$ Suess effect

The model fits the observations to an acceptable level in almost all regions, as evidenced by the standard deviation of the modern and glacial model-data misfits, σ_M and σ_G, being less than or equal to 0.2‰. One exception is the modern-day surface ($\sigma_M = 0.36$‰) which may be symptomatic of seasonal variations not captured by the steady-state model. The model-data misfit also has a strong mean offset in the modern-day surface ocean ($\mu_M = 0.21$‰), where the state estimate has more positive $\delta^{13}C_{DIC}$ values than the GLODAP/CARINA observations. This upper ocean effect has consequences for the global statistics, as the modern-day state estimate is on average 0.09‰ more enriched in $\delta^{13}C_{DIC}$ than the observations over the entire world ocean. The sign of the misfit is consistent with the observations being contaminated by anthropogenically-derived $\delta^{13}C_{DIC}$.

Although the state estimation methodology appears able to filter the Suess effect by downweighting $\delta^{13}C_{DIC}$ datapoints in the upper ocean (i.e., expecting

larger errors for these points, recall Section 3.1), we still need to consider whether the implied magnitude of the effect is plausible. To quantify the effect of this downweighting, we additionally invert a case that represents the 1990s where all modern-day $\delta^{13}C_{DIC}$ observations are weighted equally. The data is still well-fit for the 1990s case ($\sigma_M = 0.11‰$), but the mean $\delta^{13}C_{DIC}$ is shifted to $\overline{\delta^{13}C_{DIC}}^M = 0.41‰$: 0.06‰ less than our original modern-day (pre-industrial) estimate. The spatial pattern of the difference (Figure 5) shows similarities to the expected Suess effect, such as enhanced uptake and negative values in the North Atlantic deep water formation sites and the subtropical gyres. Features in the deep Pacific are unlikely to be the Suess effect, but more likely reflect changes in how the data is extrapolated into data-sparse regions. The estimated LGM-to-modern $\delta^{13}C_{DIC}$ change directly depends upon the handling of the Suess effect, and here the accounting for the contaminated $\delta^{13}C_{DIC}$ values (by downweighting) leads to an $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ value that is 0.06‰ higher than would otherwise have been estimated.

To check whether the magnitude of our Suess effect correction is reasonable, another inversion is produced where the anthropogenic $\delta^{13}C_{DIC}$ signal is removed from seawater measurements using the "FeL" simulation of Schmittner et al. (2013). To do so, we first project the modeled estimate of 1990s anthropogenic $\delta^{13}C_{DIC}$ onto the data locations and subtract it from the observations to make a corrected dataset. Then the state estimate is re-derived by weighting all data equally. By this method, the mean pre-industrial $\delta^{13}C_{DIC}$ is $\overline{\delta^{13}C_{DIC}}^M = 0.49‰$, not significantly different than our original estimate of $\overline{\delta^{13}C_{DIC}}^M = 0.47‰ \pm 0.03‰$. Thus, state estimation methods suggest that the mean Suess effect is 0.06‰ (from the previous paragraph) or 0.08‰ (this paragraph), both somewhat

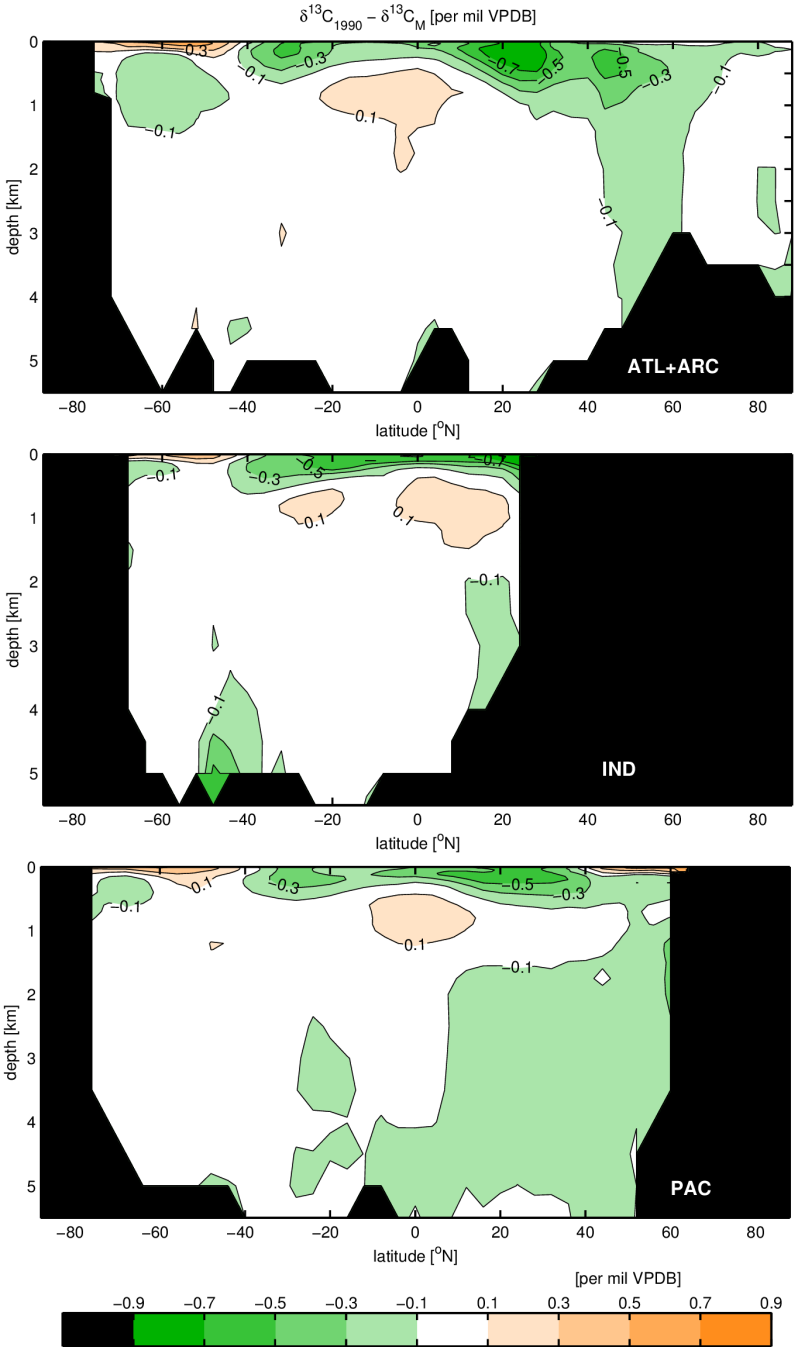

Figure 5: Similar to Figure 2, but for the Atlantic, Indian, and Pacific zonal-average difference between $\delta^{13}C_{DIC}$ in the 1990s and the modern (i.e., pre-industrial) case.

smaller than a forward model (Tagliabue and Bopp, 2008) that found 0.12‰ when constrained by the observations of Gruber et al. (1999). While we believe that our state estimation methodology has done a reasonable job in assessing the anthropogenic signal, an uncertainty of perhaps ±0.05‰ (loosely based on the spread of the three estimates) still remains in the magnitude of the $\delta^{13}C$ Suess effect and therefore should be added to the uncertainty of the LGM-to-modern $\delta^{13}C_{DIC}$ change.

4.5. Large $\delta^{13}C$ change scenario

In this section, we seek to confirm that the estimated error bars are plausible, and to visualize where the estimate glacial $\delta^{13}C_{DIC}$ field is least constrained. These issues can be addressed in greater detail by performing a test to determine whether a large $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ is consistent with the observations and a steady-state circulation. To perform this test, we add an additional "observation" that the LGM-to-modern $\delta^{13}C_{DIC}$ change is just above the upper limit as given by the 2σ error bar: 0.6‰. As in Section 4.1, this modification is handled by introducing a modified observational matrix, vector, and weighting: $\check{\mathbf{E}}$, $\check{\mathbf{y}}$, and $\check{\mathbf{W}}$, that enforces the additional constraint with a small error: $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.6‰ \pm 0.01‰$. The nonlinear solution method of Section 2.3 is then run with these additions and no other changes.

A second LGM state estimate (hereafter, LGM State Estimate 2) is indeed capable of fitting the data while producing a whole-ocean change of $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]=$ 0.59‰ (Figure 6). The spatial pattern of remaining model-data misfits do not suggest that the phytodetritus or carbonate ion effect are at play. The implied circulation leads to a deep (greater than 3 km) northern-southern water-mass interface in the Atlantic Ocean, similar to that in LGM estimate 1. The standard

deviation of the model-data misfit is actually smaller than that for LGM estimate 1 (LGM 1: $\sigma_G = 0.19‰$, LGM 2: $\sigma_{G2} = 0.17‰$). The model-data misfit statistics of LGM state estimate 2 are consistent with what is expected from a 2σ outlier of LGM state estimate 1; the estimated mean model-data misfit of $\mu_{G2} = -0.03‰$ is insignificant at the 5% level, but just so ($p = 0.06$ for $N = 492$). The larger whole-ocean change is due to increased changes in specific regions, such as the Atlantic sector of the Southern Ocean where $\Delta_{\mathrm{MG}}[\delta^{13}\mathrm{C}_{\mathrm{DIC}}] > 1.4‰$. In addition, the unconstrained Arctic Ocean and Mediterranean Sea have much increased $\Delta_{\mathrm{MG}}[\delta^{13}\mathrm{C}_{\mathrm{DIC}}]$ at intermediate levels.

The difference in $\delta^{13}\mathrm{C}_{\mathrm{DIC}}$ between the two LGM state estimates, $\Delta_{\mathrm{GG}}[\delta^{13}\mathrm{C}_{\mathrm{DIC}}]$ (the difference of two glacial "G" estimates), can be thought of as the observational null space and illustrates the ocean regions that are both unconstrained and important for setting the global mean. The biggest differences occur at the Pacific surface and the South Indian Ocean (Figure 7). Surface differences, especially in the Pacific Ocean, are as large as $1‰$, and are consistent with the large error bars previously detailed in Table 1. Should the planktonic $\delta^{13}\mathrm{C}$ records that indicate little change between the LGM and modern-day (e.g., Broecker and McGee, 2013) be representative of the entire tropics, LGM state estimate 1 (from Sections 3-5) would be considered more reasonable. Our map of the difference between the two LGM state estimates emphasizes the regions in which additional observations would be most useful.

LGM State Estimate 2 may be relevant to more than just checking the machinery for producing error bars, as there are a number of reasons to suspect that the LGM-to-modern $\delta^{13}\mathrm{C}_{\mathrm{DIC}}$ change might be larger than the recent marine-based consensus. For example, the state estimate reverts to modern-day conditions in

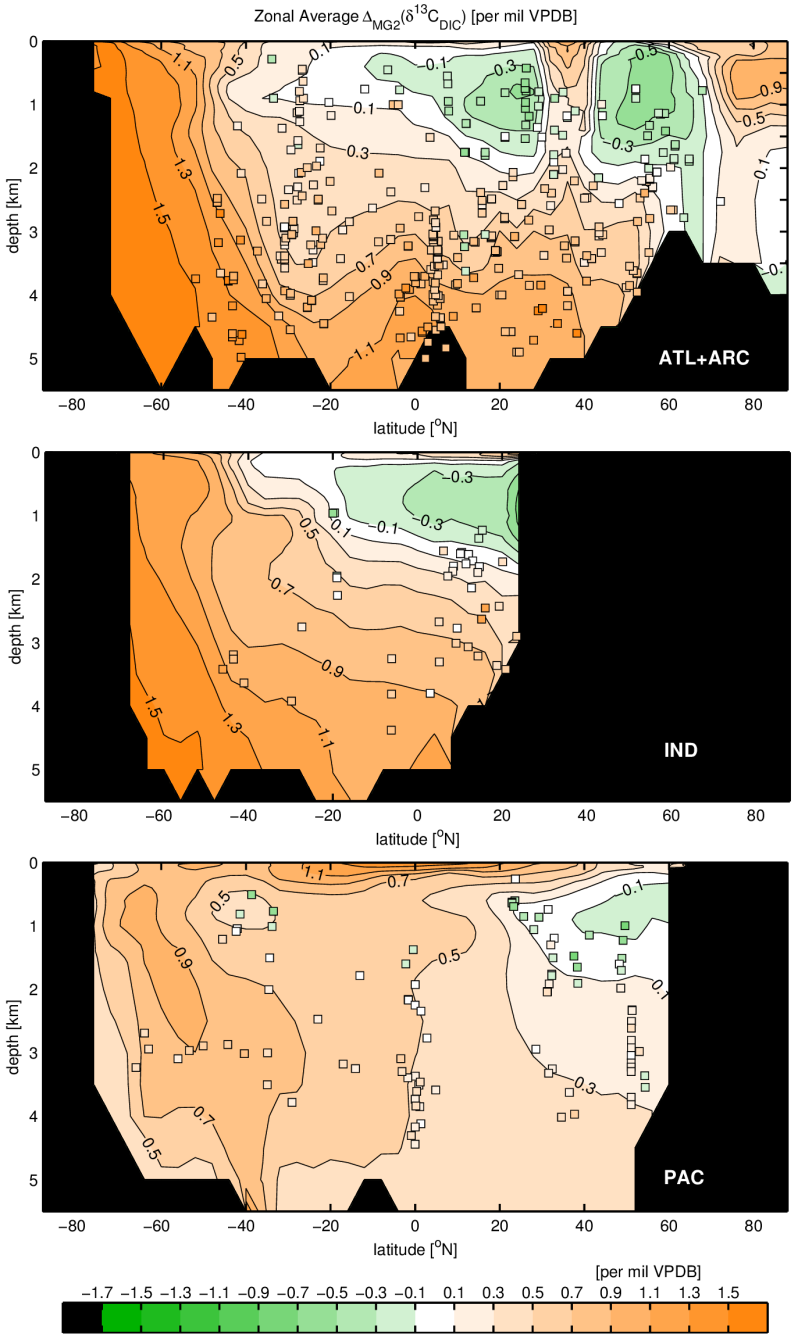

Figure 6: Atlantic (and Arctic), Indian, and Pacific zonal-average difference of $\delta^{13}C$ between LGM State Estimate 2 and modern-day (i.e., $\Delta_{MG}[\delta^{13}C_{DIC}] = \delta^{13}C^{M}_{DIC}$ - $\delta^{13}C^{G2}_{DIC}$).

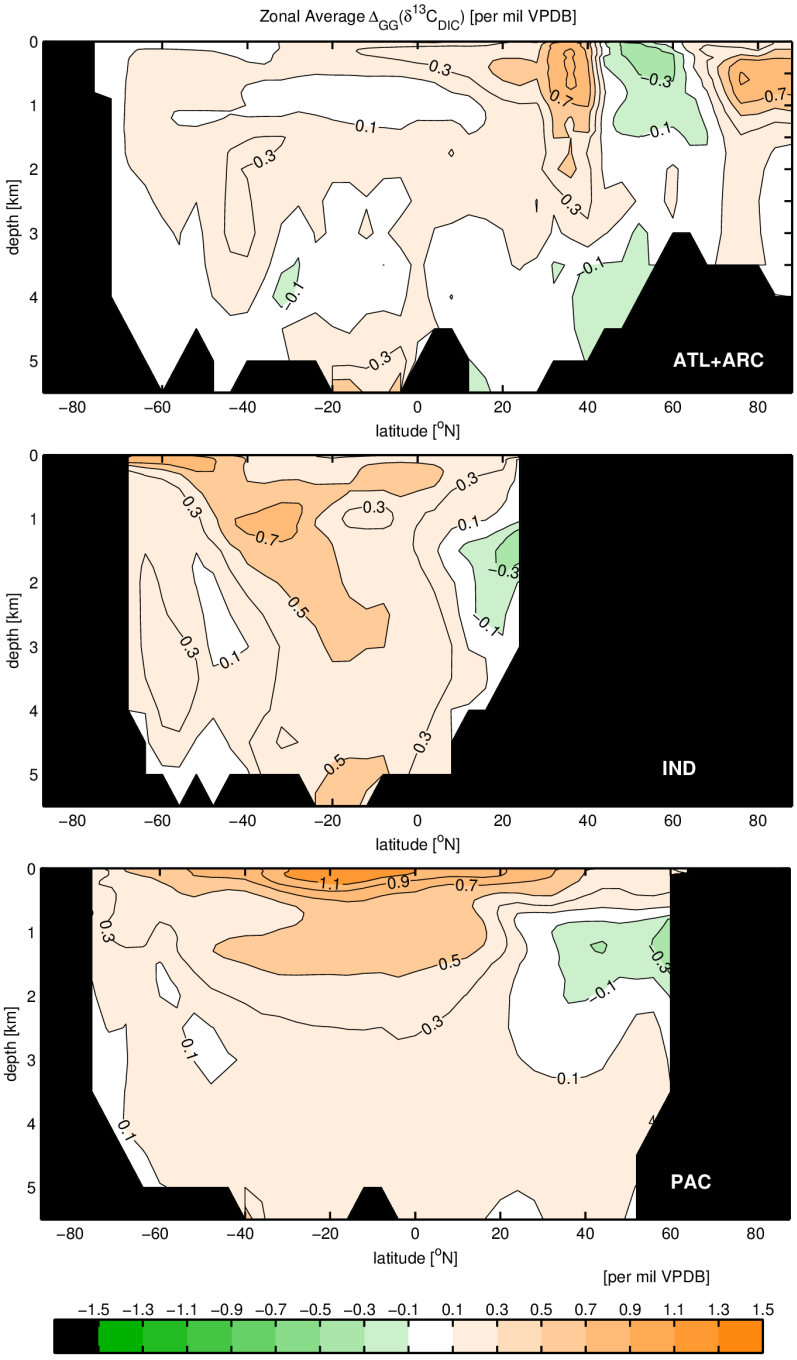

Figure 7: Atlantic and Pacific zonal-average difference of $\delta^{13}C_{DIC}$ between 2 LGM Estimates ($\Delta_{GG}[\delta^{13}C_{DIC}] = \delta^{13}C^{G}_{DIC}$ - $\delta^{13}C^{G2}_{DIC}$).

the absence of any paleo-data and thus, the $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ value from LGM State Estimate 1 could be biased low due to remnant modern-day constraints pulling the estimate toward no change. A low estimate may also be due to low temporal resolution in the cores, as high-resolution cores can have higher $\Delta_{MG}[\delta^{13}C_{DIC}]$ values (A. Mix, personal communication, 2014). Furthermore, if a $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ value of 0.6‰ is possible, then terrestrial-based (e.g., Crowley, 1995) and marine-based (e.g., Tagliabue et al., 2009) carbon partitioning estimates could be brought into consistency without the need to invoke an inert terrestrial carbon pool (Ciais et al., 2012).

4.6. Seawater-to-calcite relationship

Here we revisit our definition of $\Delta_{MG}[\delta^{13}C_{DIC}]$ and whether it significantly differs from $\Delta_{HG}[\delta^{13}C_{DIC}]$, defined as the LGM-to-Late Holocene difference. Information regarding $\Delta_{HG}[\delta^{13}C_{DIC}]$ is more directly available through the difference of LGM and Late Holocene core data, as interlaboratory offsets are canceled by the differencing. One tradeoff is that only 365 measurements of a Late Holocene and LGM value from the same core are available even when the Late Holocene time interval is extended from 0-6 kyr BP. Using the modern-day circulation to recreate a global $\delta^{13}C_{DIC}$ field using the Late Holocene coretop values, we estimate a Late-Holocene mean value of $\overline{\delta^{13}C_{DIC}}^{H} = 0.55‰ \pm 0.20‰$ that is 0.08‰ higher than the modern-day estimate. The difference directly affects the inferred LGM-to-modern change and is attributable to Holocene temporal variability (e.g., Oppo et al., 2003) and error in the seawater-to-calcite calibration, but it is difficult to separate the two. In particular, a systematic offset in the $\delta^{13}C$ calcite-to-seawater proxy relationship could occur due to a dependence of calcite $\delta^{13}C$ on carbonate ion or temperature (e.g., Spero et al., 1997; Hesse et al., 2014). To determine the

size of such an effect, an improved calibration between the Late Holocene $\delta^{13}C$ values and the modern-day seawater characteristics including $\delta^{13}C_{DIC}$, temperature, salinity, CO_3^{2-}, DIC, and pH should be pursued.

5. Discussion

The discussion aims to put the results of this work into the context of previous observational methods to estimate LGM-to-modern $\delta^{13}C_{DIC}$ change.

5.1. Optimal data weights

Given a circulation field, the state estimate formulation permits the coefficients and constants of equation (1) to be explicitly calculated. Following Appendix C, the optimal set of data weights is

$$\mathbf{w} = \mathbf{W}^{-1}\hat{\mathbf{E}}(\hat{\mathbf{S}}^{-1} + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\hat{\mathbf{E}} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})^{-1}\hat{\mathbf{r}}, \tag{8}$$

where the hat represents truncating the variables to the quantities related to the tracer field. Interestingly, the elements of $\mathbf{w}$ need not sum to one, nor are they necessarily non-negative, as occurs in linear extrapolation problems. The additional, usually-neglected constant in equation (1),

$$\overline{c}_0 = (\mathbf{c}_0^T\hat{\mathbf{S}}^{-1} + \mathbf{q}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})(\hat{\mathbf{S}}^{-1} + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\hat{\mathbf{E}} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})^{-1}\hat{\mathbf{r}}, \tag{9}$$

is related to the sampling bias. If all the observations have a value of zero, the best estimate of the global mean has a value of $\overline{c}_0$, indicating that prior information is being used to calculate the global mean. Thus, the degree to which the observations sample the global ocean in a biased way is quantified. In the case that no observations are available, $\overline{c}_0$ reverts to the first-guess global mean. The weights

and the constant are optimal in the sense that they solve a tapered, weighted least squares problem.

The map of observational weights is spatially heterogeneous with eastern Atlantic points generally having the smallest weights (Figure 8). A point in the Indian Ocean is upweighted the most, with $w \approx 10/N$ (or 10 times the weight that it would be given in the basic arithmetic mean). This map generally corresponds to the upweighting of Pacific data points (due to their relative sparsity) and downweighting of Atlantic points. Finer detail is also present, however, such as the upweighting of the few points in the Nordic Seas, but a downweighting of nearby points that are just south of the Greenland-Iceland-Scotland ridge. Other details depend on the flow patterns in the glacial circulation, as observations have influence both up- and downstream.

5.2. Interpreting data histograms

Important information about $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ is available by aggregating pointwise estimates of $\Delta_{MG}[\delta^{13}C_{DIC}]$ in a histogram. Pointwise values of $\Delta_{MG}[\delta^{13}C_{DIC}]$ are here inferred by projecting the modern-day state estimate $\delta^{13}C_{DIC}$ field onto the core sites by a linear interpolation of the 8 nearest gridpoints, then comparing with LGM sediment core values. In our 492 points of $\Delta_{MG}[\delta^{13}C_{DIC}]$, the median is 0.45‰ and the mean is 0.39‰ (upper left panel Figure 9). The mode of the distribution is 0.6‰, which suggests that Shackleton (1977) was more likely to estimate a number this high with a small number of cores. It is not clear from this analysis, however, whether the mean of the histogram is a good estimate for the global-mean $\delta^{13}C_{DIC}$.

To better interpret such a histogram, consider formulating a modern-day and LGM equation of the type of equation (1) and taking their difference. For a general

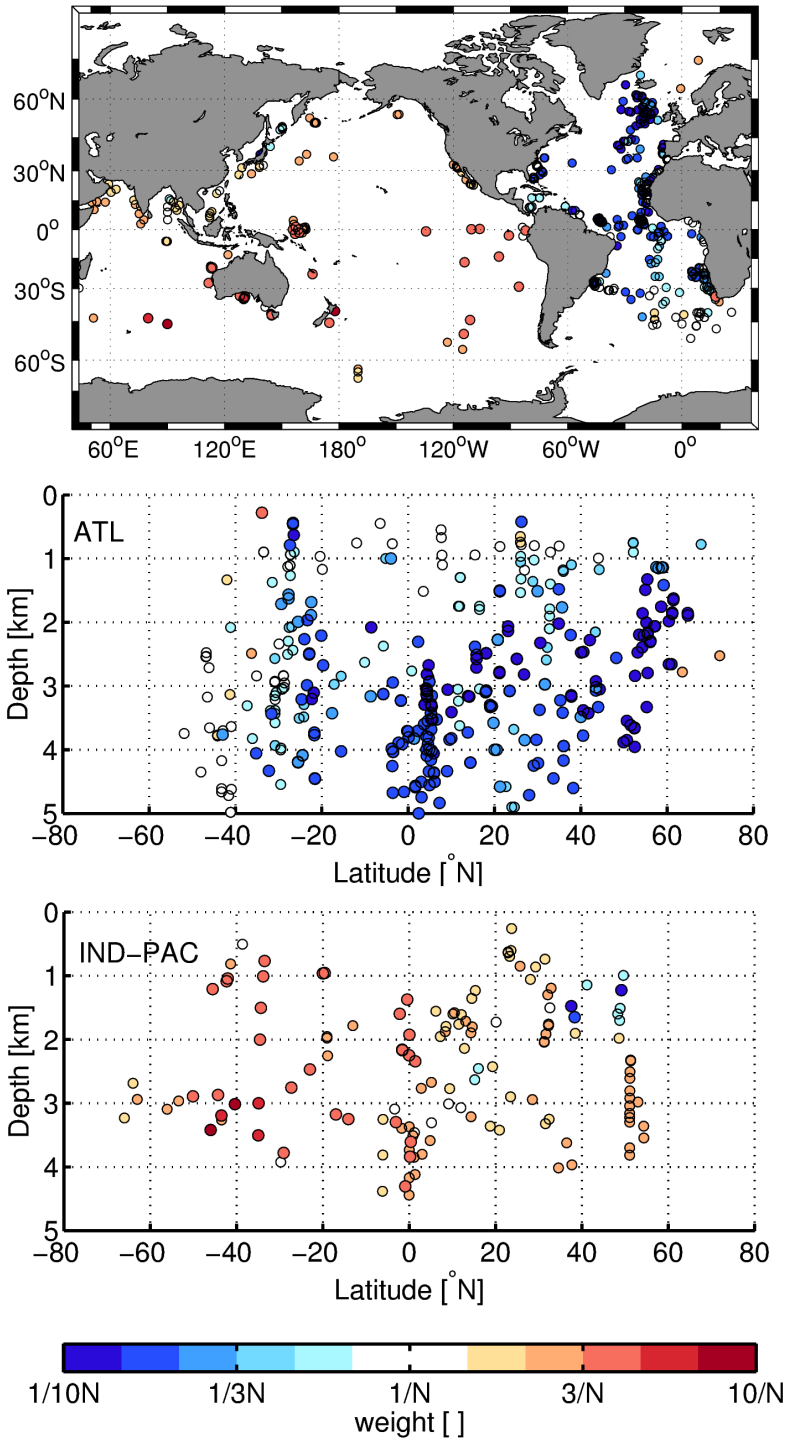

Figure 8: Optimal observational weights to construct a global mean given the glacial circulation (LGM State Estimate 1): plan view (*top*), Atlantic data (*middle*) and Indo-Pacific data (*bottom*). The colorscale is logarithmic. The weights vary from 10 times greater to 100 times lesser than the arithmetic-mean weight and thus some values are offscale low (saturated at a value of $1/10N$).

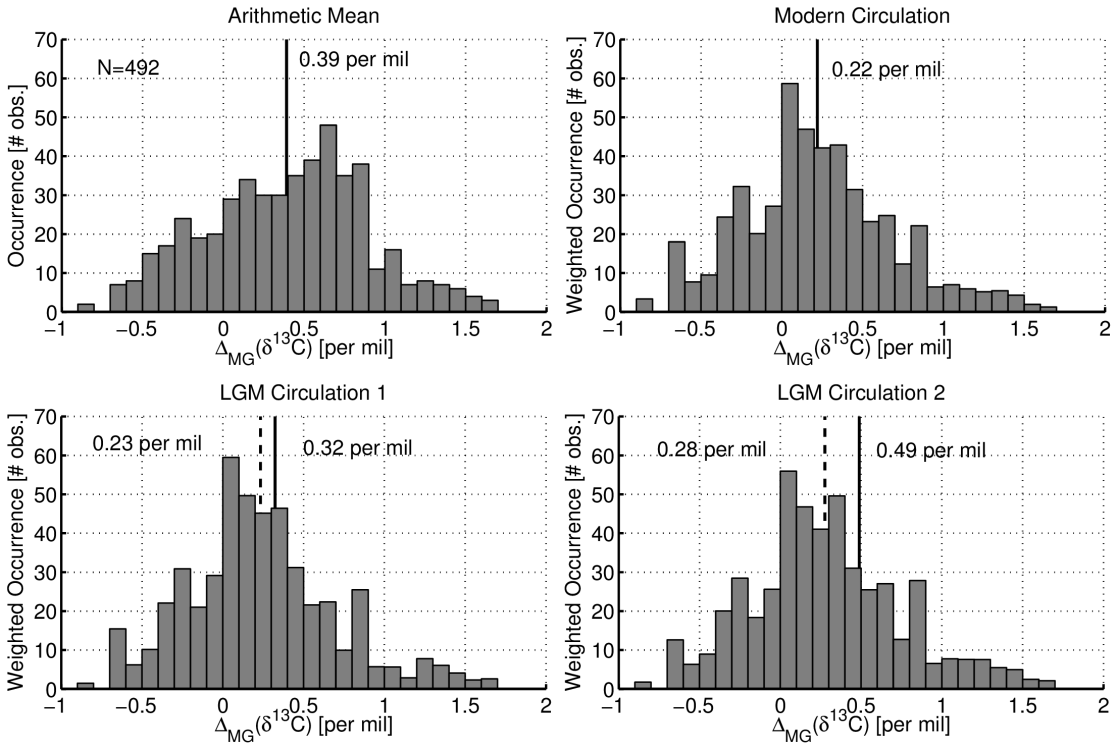

Figure 9: Inference of LGM-to-modern $\delta^{13}C_{DIC}$ change from pointwise measurements. *Upper left*: Histogram of the 492 observations of $\Delta_{MG}[\delta^{13}C_{DIC}]$. *Upper right*: Histogram modified by the optimal weights computed for the modern-day circulation. *Bottom row*: Histogram modified by the optimal weights computed for the glacial circulation (*lower left*: LGM State Estimate 1, *lower right*: LGM State Estimate 2). The inferred $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ (*solid lines*) results from the mean of the histogram in the top row. In the bottom row, the mean of the histogram (*dashed lines*) is corrected by the last two terms in equation (10) to produce the $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}]$ estimate (*solid lines*).

tracer, we obtain

$$\Delta_{\mathrm{MG}}(\overline{c}) = \overline{\overline{\mathbf{w}}}^{T}\Delta_{\mathrm{MG}}(\mathbf{y}) + \Delta_{\mathrm{MG}}(\mathbf{w})^{T}\overline{\overline{\mathbf{y}}} + \Delta_{\mathrm{MG}}(\overline{c}_0), \tag{10}$$

where the Δ_{MG} operator acts elementwise on each vector and the double overbar represents the temporal mean of modern-day and LGM conditions (to distinguish from the single overbar that is a global spatial average). The mean of the aforementioned histogram reflects the true global mean in the case that the data weights are based on the arithmetic mean ($\overline{\overline{w}}_i = 1/N$ for all i) and that the second and third terms of equation (10) vanish. As discussed in the Introduction, the spatially irregular distribution of observations makes these conditions improbable, and thus the mean of the histogram is not usually a good estimate of global-mean change.

When a circulation is available to compute the optimal weights of equation (8) and the constant of equation (9), the mean LGM-to-modern change can be better estimated. For illustration, consider a case where it is assumed that the modern-day circulation is representative of the LGM. Then we can use the calculated modern-day weights ($\mathbf{w}_M$) to better approximate the first term of equation (10). These weights shift the mode of the distribution to $\Delta_{\mathrm{MG}}[\delta^{13}\mathrm{C}_{\mathrm{DIC}}] = 0.05‰$ and the inferred global-mean decreases from $0.39‰$ to $0.22‰$ (visualized as a weighted histogram, upper right panel, Figure 9). In effect, the observations of smaller change are upweighted because they are located along modern-day circulation pathways that influence more of the ocean.

A full interpretation of the pointwise data should also account for LGM-to-modern circulation change, of course. In particular, we use both the circulations from LGM State Estimate 1 and 2 (lower row, Figure 9). The mean of the histogram is again modified, this time back toward larger values ($0.23‰$ and $0.28‰$, respectively). The two correction terms due to the changing ocean circulation

must also be considered (terms 2 and 3 of the right hand side of equation 10). For LGM State Estimate 1, the correction is 0.09‰ and the final estimate of $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.32‰$ is consistent with the results showcased in Sections 3 and 4. For LGM State Estimate 2, the correction is even larger: 0.21‰ (dominated by term 3: 0.15‰). A global-mean change of $\Delta_{MG}[\overline{\delta^{13}C_{DIC}}] = 0.49‰$ is then diagnosed from equation (10), smaller than the actual 0.59‰, which is symptomatic of the breakdown of the linear assumption (i.e., equation B.4). Unfortunately, the correction terms are poorly known due to the uncertainty in the LGM circulation. We emphasize that the correction cannot be determined from the data histogram alone.

6. Conclusion

The LGM-to-modern $\delta^{13}C_{DIC}$ change is explicitly estimated by first mapping benthic foraminiferal observations onto a global grid, and then taking a mass-weighted average of the gridded values. The mapping process, however, requires a method more sophisticated than typical interpolation because of the sparsity of the dataset and large spatial gaps. Here we demonstrate that an LGM state estimate derived from a recent compilation of 492 $\delta^{13}C$ data points combined with a tracer transport model can provide a reasonable globally-gridded field, as well as self-consistent uncertainty estimates. Our updated best estimate of LGM-to-modern global $\delta^{13}C_{DIC}$ change is $0.32 \pm 0.20‰$ at the 2σ uncertainty level. A coherent picture of the LGM $\delta^{13}C$ distribution emerges that is consistent with previous Atlantic estimates and fills in the missing details of the Pacific distribution. Maps of the LGM-to-modern difference in $\delta^{13}C$ also display coherent spatial patterns, with largest changes in the Atlantic sector of the Southern Ocean.

While previous investigators have used various sub-domains and multiple-step averaging techniques to determine how to best weight pointwise observations to obtain the global mean, here we show that determination of the optimal data weights requires knowledge of the ocean circulation. The diagnosed optimal weights conform to expectation in many ways, such as upweighting data in regions with sparse coverage. It is very difficult, however, to determine the global mean change from pointwise histograms of the local $\delta^{13}C_{DIC}$ change unless the concommitant circulation change is also known. Thus, much of the existing uncertainty in LGM-to-modern $\delta^{13}C$ change is due to the difficulty in determining the glacial ocean circulation.

Our glacial state estimate points toward future directions to reduce the considerable remaining global-mean $\delta^{13}C_{DIC}$ uncertainty. For example, the addition of randomly-distributed hypothetical data would reduce the uncertainty of the global mean as $N^{-1/2}$, where N is the number of observations. Furthermore, the greatest differences between two LGM estimates that both fit the data occur in the upper ocean, the Southern Ocean, and various marginal seas, pointing to regions where additional information would be most useful. In particular, a compilation of planktonic records would help reduce the uncertainty in surface regions. Provided that challenges with interpretation regarding DIC change and species offsets can be overcome (e.g., Spero et al., 1997), we expect that strategic sampling can reduce the global-mean uncertainty at a faster rate than the hypothetical randomly-sampled case. A two-pronged approach appears best suited to reduce the global-mean uncertainty: compilation of information from strategically-placed locations, including planktonic records, and the implementation of a more sophisticated dynamical model that can better constrain the circulation for making global maps.

Appendix A. Input variables

The input variables include the weighting matrices for the different constraints of equation (7). The $\mathbf{S}^{-1}$ matrix is chosen to reflect prior knowledge of the $\delta^{13}C_{DIC}$ field and mass flux ratios. For example, the surface $\delta^{13}C_{DIC}$ is assumed to vary by no more than 2.4‰ with lengthscales no smaller than 10° of latitude or longitude (following the nondiagonal weighting matrix method of Gebbie et al. (2006)). The mass flux ratios, $\mathbf{m}$, are given a prior uncertainty of 0.3 (relative to their nondimensional range of 0 to 1), chosen by the standard deviation of the $\mathbf{m}$ entries for the modern-day circulation. This choice reflects our desire to constrain the glacial circulation with little subjective prior information, but to require that the statistics of glacial transport should not fall outside the range of the modern-day statistics. For the modern-day circulation, on the other hand, we estimate that the uncertainty of the $\mathbf{m}$ values is about 0.05, based on the differences between various modern solutions. The $\mathbf{W}^{-1}$ matrix assumes that the observational uncertainty is 0.2‰, as discussed in the body of the text. The choice of the $\mathbf{Q}^{-1}$ matrix agnostically assumes that the glacial source of remineralized material is the same magnitude as the modern day source or sink.

Appendix B. Uncertainty derivation

Appendix B.1. Least-squares solution

Section 2.4 defines a cost function, $J = \mathbf{u}^T\mathbf{S}^{-1}\mathbf{u} + \mathbf{n}^T\mathbf{W}^{-1}\mathbf{n} + \mathbf{v}^T\mathbf{Q}^{-1}\mathbf{v}$, that is here written in a more complete form by substituting the equations for the first-guess adjustment, the observational constraint, and the model:

$$J = (\mathbf{x}-\mathbf{x}_0)^T\mathbf{S}^{-1}(\mathbf{x}-\mathbf{x}_0) + (\mathbf{E}\mathbf{x}-\mathbf{y})^T\mathbf{W}^{-1}(\mathbf{E}\mathbf{x}-\mathbf{y}) + (\mathcal{L}[\mathbf{x}]-\mathbf{q})^T\mathbf{Q}^{-1}(\mathcal{L}[\mathbf{x}]-\mathbf{q}). \tag{B.1}$$

The solution, $\tilde{\mathbf{x}}$, is at the minimum of J and thus satisfies the stationary assumption:

$$J_{\tilde{\mathbf{x}}} = 2\{\mathbf{S}^{-1}(\mathbf{x}-\mathbf{x}_0)+\mathbf{E}^T\mathbf{W}^{-1}(\mathbf{E}\mathbf{x}-\mathbf{y})+\mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}(\mathcal{L}[\mathbf{x}]-\mathbf{q})\} = 0, \tag{B.2}$$

where $J_{\tilde{\mathbf{x}}} \equiv \partial J/\partial \mathbf{x}]_{\tilde{\mathbf{x}}}$. We define a linearization of model in the neighborhood of the solution:

$$\mathcal{L}[\mathbf{x}] = \mathcal{L}[\mathbf{x}'] + \mathbf{L}_{\mathbf{x}'}(\mathbf{x}-\mathbf{x}') + \varepsilon, \tag{B.3}$$

where ε represents the higher-order terms in the expansion. The solution estimate, $\tilde{\mathbf{x}}$, satisfies

$$\tilde{\mathbf{x}} = (\mathbf{S}^{-1}+\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E}+\mathbf{L}_{\mathbf{x}'}^T\mathbf{Q}^{-1}\mathbf{L}_{\mathbf{x}'})^{-1}\{\mathbf{S}^{-1}\mathbf{x}_0+\mathbf{E}^T\mathbf{W}^{-1}\mathbf{y}+\mathbf{L}_{\mathbf{x}'}^T\mathbf{Q}^{-1}(\mathbf{q}+\mathbf{L}_{\mathbf{x}'}\mathbf{x}'-\mathcal{L}[\mathbf{x}'])\}, \tag{B.4}$$

where the higher-order terms become negligible as one approaches the solution and are dropped from this last equation.

Appendix B.2. Uncertainty of tracer distribution and circulation

Here we seek the expected solution in the hypothetical case that a perfect first-guess, observations, and tracer source are available ($\mathbf{x}_0^*$, $\mathbf{y}^*$, and $\mathbf{q}^*$, respectively). Defining

$$\mathbf{d}_0 = \mathbf{S}^{-1}\mathbf{x}_0^*+\mathbf{E}^T\mathbf{W}^{-1}\mathbf{y}^*+\mathbf{L}_{\mathbf{x}'}^T\mathbf{Q}^{-1}(\mathbf{q}^*+\mathbf{L}_{\mathbf{x}'}\mathbf{x}'-\mathcal{L}[\mathbf{x}']), \tag{B.5}$$

the expected solution is

$$<\tilde{\mathbf{x}}> = (\mathbf{S}^{-1}+\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E}+\mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}})^{-1}\mathbf{d}_0. \tag{B.6}$$

The dispersion of $\tilde{\mathbf{x}} - <\tilde{\mathbf{x}}>$ is equal to the solution covariance of $\tilde{\mathbf{x}}$:

$$\mathbf{C}_{\tilde{\mathbf{x}}} \equiv <(\tilde{\mathbf{x}}-<\tilde{\mathbf{x}}>)(\tilde{\mathbf{x}}-<\tilde{\mathbf{x}}>)^T>, \tag{B.7}$$

and substitution of equation (B.6) into (B.7) gives

$$\mathbf{C}_{\tilde{\mathbf{x}}} = (\mathbf{S}^{-1} + \mathbf{E}^T\mathbf{W}^{-1}\mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}})^{-1} < (\mathbf{d}-\mathbf{d}_0)(\mathbf{d}-\mathbf{d}_0)^T > (\mathbf{S}^{-1} + \mathbf{E}^T\mathbf{W}^{-1}\mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}})^{-1}. \tag{B.8}$$

Assuming that errors in the first guess, observations, and model are uncorrelated (i.e., $\mathbf{R}_{xn} = 0$, $\mathbf{R}_{xq} = 0$, etc.), the expected value in the right hand side of (B.8) is

$$< (\mathbf{d}-\mathbf{d}_0)(\mathbf{d}-\mathbf{d}_0)^T >= \mathbf{S}^{-1}\mathbf{R}_{xx}\mathbf{S}^{-1} + \mathbf{E}^T\mathbf{W}^{-1}\mathbf{R}_{nn}\mathbf{W}^{-1}\mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{R}_{qq}\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}}. \tag{B.9}$$

Substituting equation (B.9) into (B.8) and assuming that the weight matrices are equal to the expected second-moment matrices of the residuals (i.e., $\mathbf{R}_{nn} = \mathbf{W}$, $\mathbf{R}_{qq} = \mathbf{Q}$, and $\mathbf{R}_{xx} = \mathbf{S}$), we obtain the solution covariance in terms of the known input variables:

$$\mathbf{C}_{\tilde{\mathbf{x}}} = (\mathbf{S}^{-1} + \mathbf{E}^T\mathbf{W}^{-1}\mathbf{E} + \mathbf{L}_{\tilde{\mathbf{x}}}^T\mathbf{Q}^{-1}\mathbf{L}_{\tilde{\mathbf{x}}})^{-1}, \tag{B.10}$$

that is used in the main text as equation (6). The standard error is here defined as $\pm\sqrt{\mathbf{C}_{\tilde{\mathbf{x}}}}$ of the diagonal elements.

Appendix C. Calculation of the optimal observational weights

In the case that the circulation is known exactly, the cost function equation (B.1) can be simplified

$$J = (\mathbf{c}-\mathbf{c}_0)^T\hat{\mathbf{S}}^{-1}(\mathbf{c}-\mathbf{c}_0) + (\hat{\mathbf{E}}\mathbf{c}-\mathbf{y})^T\mathbf{W}^{-1}(\hat{\mathbf{E}}\mathbf{c}-\mathbf{y}) + (\mathbf{L}_\mathbf{c}\mathbf{c}-\mathbf{q})^T\mathbf{Q}^{-1}(\mathbf{L}_\mathbf{c}\mathbf{c}-\mathbf{q}), \tag{C.1}$$

where the hat indicates truncation of the $\mathbf{E}$ and $\mathbf{S}$ matrices to the parts related to the tracer field. Using the least-squares estimate of the tracer solution and the following definition of the global mean, $\overline{c} = \hat{\mathbf{r}}^T\mathbf{c}$, where $\hat{\mathbf{r}}$ is also a truncated vector,

we obtain a simplified equation

$$\tilde{\bar{c}} = \hat{\mathbf{r}}^T(\hat{\mathbf{S}}^{-1} + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\hat{\mathbf{E}} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})^{-1}(\hat{\mathbf{S}}^{-1}\mathbf{c}_0 + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\mathbf{y} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{q}). \quad \text{(C.2)}$$

Comparison of the equation (C.2) to equation (1) permits the identification of the optimal data weights:

$$\mathbf{w} = \mathbf{W}^{-1}\hat{\mathbf{E}}(\hat{\mathbf{S}}^{-1} + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\hat{\mathbf{E}} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})^{-1}\hat{\mathbf{r}}, \quad \text{(C.3)}$$

and the additional constant

$$\bar{c}_0 = (\mathbf{c}_0^T\hat{\mathbf{S}}^{-1} + \mathbf{q}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})(\hat{\mathbf{S}}^{-1} + \hat{\mathbf{E}}^T\mathbf{W}^{-1}\hat{\mathbf{E}} + \mathbf{L}_\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{L}_\mathbf{c})^{-1}\hat{\mathbf{r}}. \quad \text{(C.4)}$$

Acknowledgments

The authors are indebted to Olivier Marchal for thoroughly commenting on a version of the manuscript and the mathematics. The authors thank Ed Boyle, Thomas Chalk, David McGee, and Alan Mix for helpful discussions. GG is supported by NSF grants OIA-1124880 and OCE-1357121, the WHOI Ocean and Climate Change Institute, and The Joint Initiative Awards Fund from the Andrew W. Mellon Foundation.

References

Adams, J. M., Faure, H., 1998. A new estimate of changing carbon storage on land since the Last Glacial Maximum, based on global land ecosystem reconstruction. Global and Planetary Change 16, 3–24.

Adkins, J., McIntyre, K., Schrag, D., 2002. The salinity, temperature, and $\delta^{18}O$ of the glacial deep ocean. Science 298, 1724–1725.

Arbic, B. K., MacAyeal, D. R., Mitrovica, J. X., Milne, G. A., 2004. Palaeoclimate: Ocean tides and Heinrich events. Nature 432 (7016), 460–460.

Boyle, E. A., 1992. Cadmium and $\delta^{13}C$ paleochemical ocean distributions during the Stage 2 Glacial Maximum. Annual Review of Earth and Planetary Sciences 20 (1), 245–287.

Bretherton, F., Davis, R., Fandry, C., 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep-Sea Res 23, 559–582.

Broecker, W., Maier-Reimer, E., 1992. The influence of air and sea exchange on the carbon isotope distribution in the sea. Global Biogeochemical Cycles 6 (3), 315–320.

Broecker, W., McGee, D., 2013. The ^{13}C record for atmospheric CO_2: What is it trying to tell us? Earth and Planetary Science Letters 368, 175–182.

Ciais, P., Tagliabue, A., Cuntz, M., Bopp, L., Scholze, M., Hoffmann, G., Lourantou, A., Harrison, S., Prentice, I., Kelley, D., et al., 2012. Large inert carbon pool in the terrestrial biosphere during the Last Glacial Maximum. Nature Geoscience 5 (1), 74–79.

Crowley, T. J., 1995. Ice age terrestrial carbon changes revisited. Global Biogeochemical Cycles 9 (3), 377–389.

Curry, W., Duplessy, J., Labeyrie, L., Shackleton, N., 1988. Changes in the distribution of $\delta^{13}C$ of deep water $\sum CO_2$ between the last glaciation and the Holocene. Paleoceanography 3 (3).

Curry, W., Oppo, D., 2005. Glacial water mass geometry and the distribution of $\delta^{13}C$ of $\sum CO_2$ in the Western Atlantic Ocean. Paleoceanography 20, 10.1029/2004PA001021.

Duplessy, J., Shackleton, N., Fairbanks, R., Labeyrie, L., Oppo, D., Kallel, N., 1988. Deepwater source variations during the last climatic cycle and their impact on the global deepwater circulation. Paleoceanography 3 (3).

Duplessy, J., Shackleton, N., Matthews, R., Prell, W., Ruddiman, W., Caralp, M., Hendy, C., 1984. ^{13}C record of benthic foraminifera in the last interglacial ocean: Implications for the carbon cycle and the global deep water circulation. Quaternary Research 21 (2), 225–243.

Gebbie, G., 2014. How much did Glacial North Atlantic Water shoal? Paleoceanography 29 (3), 190–209.

Gebbie, G., Heimbach, P., Wunsch, C., 2006. Strategies for Nested and Eddy-Permitting State Estimation. J. Geophys. Res. 111, C10073, 10.1029/2005JC003094.

Gebbie, G., Huybers, P., 2010. Total matrix intercomparison: a method for resolving the geometry of water-mass pathways. J. Phys. Oceanogr. 40 (8), 1710–1728, DOI: 10.1175/2010JPO4272.1.

Gebbie, G., Huybers, P., 2012. The mean age of ocean waters inferred from radiocarbon observations: sensitivity to surface sources and accounting for mixing histories. J. Phys. Oceanogr. 42 (2), DOI: 10.1175/JPO-D-11-043.1, 291–305.

Gruber, N., Keeling, C. D., Bacastow, R. B., Guenther, P. R., Lueker, T. J., Wahlen, M., Meijer, H. A., Mook, W. G., Stocker, T. F., 1999. Spatiotemporal patterns

725 of carbon-13 in the global surface oceans and the oceanic Suess effect. Global
726 Biogeochemical Cycles 13 (2), 307–335.

727 Herguera, J., Herbert, T., Kashgarian, M., Charles, C., 2010. Intermediate and
728 deep water mass distribution in the Pacific during the Last Glacial Maximum
729 inferred from oxygen and carbon stable isotopes. Quaternary Science Reviews
730 29 (9), 1228–1245.

731 Hesse, T., Butzin, M., Bickert, T., Lohmann, G., 2011. A model-data compari-
732 son of δ^{13}C in the glacial Atlantic Ocean. Paleoceanography, 26, PA3220. 26,
733 PA3220.

734 Hesse, T., Wolf-Gladrow, D., Lohmann, G., Bijma, J., Mackensen, A., Zeebe,
735 R. E., 2014. Modelling δ^{13}C in benthic foraminifera: Insights from model sen-
736 sitivity experiments. Marine Micropaleontology 112, 50–61.

737 Hide, R., 1969. Dynamics of the atmospheres of the major planets with an ap-
738 pendix on the viscous boundary layer at the rigid bounding surface of an
739 electrically-conducting rotating fluid in the presence of a magnetic field. J. At-
740 mos. Sci. 26, 841–853.

741 IOC, S., 2010. IAPSO: The international thermodynamic equation of seawater–
742 2010: Calculation and use of thermodynamic properties, intergovernmental
743 oceanographic commission, manuals and guides no. 56.

744 Kaplan, J. O., Prentice, I. C., Knorr, W., Valdes, P. J., 2002. Modeling the dynam-
745 ics of terrestrial carbon storage since the Last Glacial Maximum. Geophysical
746 Research Letters 29 (22), 31–1.

747 Köhler, P., Fischer, H., Schmitt, J., 2010. Atmospheric $\delta^{13}CO_2$ and its relation
748 to pCO_2 and deep ocean $\delta^{13}C$ during the late Pleistocene. Paleoceanography
749 25 (1).

750 Mackensen, A., Schumacher, S., Radke, J., Schmidt, D., 2000. Microhabitat pref-
751 erences and stable carbon isotopes of endobenthic foraminifera: clue to quan-
752 titative reconstruction of oceanic new production? Marine Micropaleontology
753 40 (3), 233–258.

754 Makou, M. C., Oppo, D. W., Curry, W. B., 2010. South Atlantic intermediate
755 water mass geometry for the Last Glacial Maximum from foraminiferal Cd/Ca.
756 Paleoceanography 25 (4).

757 Marchal, O., Curry, W., 2008. On the abyssal circulation in the glacial Atlantic.
758 Journal of Physical Oceanography 38 (9), 2014–2037.

759 Matsumoto, K., Lynch-Stieglitz, J., 1999. Similar glacial and Holocene deep wa-
760 ter circulation inferred from Southeast Pacific benthic foraminiferal carbon iso-
761 tope composition. Paleoceanography 14 (2), 149–163.

762 Matsumoto, K., Oba, T., Lynch-Stieglitz, J., Yamamoto, H., 2002. Interior hydrog-
763 raphy and circulation of the glacial Pacific Ocean. Quaternary science reviews
764 21 (14-15), 1693–1704.

765 Monnin, E., Indermuhle, A., Dallenbach, A., Fluckiger, J., Stauffer, B., Stocker,
766 T. F., Raynaud, D., Barnola, J., 2001. Atmospheric CO_2 concentrations over the
767 last glacial termination. Science 291, 112–114.

768 Nocedal, J., 1980. Updating quasi-Newton matrices with limited storage. Mathe-
769 matics of computation 35 (151), 773–782.

770 Oliver, K., Hoogakker, B., Crowhurst, S., Henderson, G., Rickaby, R., Edwards,
771 N., Elderfield, H., 2010. A synthesis of marine sediment core δ^{13}c data over the
772 last 150 000 years. Climate of the Past 6, 645–673.

773 Olsen, A., Abdirahman, O., Bellerby, R., Johannessen, T., Ninnemann, U., Brown,
774 K., Olsson, K., Olafsson, J., Nondal, G., Kivimäe, C., et al., 2006. Magnitude
775 and origin of the anthropogenic CO_2 increase and the ^{13}C Suess Effect in the
776 Nordic Seas since 1981.

777 Oppo, D. W., McManus, J. F., Cullen, J. L., 2003. Palaeo-oceanography: Deep-
778 water variability in the Holocene epoch. Nature 422 (6929), 277–277.

779 Peeters, F. J., Brummer, G.-J. A., Ganssen, G., 2002. The effect of upwelling on
780 the distribution and stable isotope composition of *Globigerina bulloides* and
781 *Globigerinoides ruber* (planktic foraminifera) in modern surface waters of the
782 NW Arabian Sea. Global and Planetary Change 34 (3), 269–291.

783 Peterson, C. D., Lisiecki, L. E., Stern, J. V., 2014. Deglacial whole-ocean δ^{13}C
784 change estimated from 480 benthic foraminiferal records. Paleoceanography.

785 Russell, A. D., Spero, H. J., 2000. Field examination of the oceanic carbonate ion
786 effect on stable isotopes in planktonic foraminifera. Paleoceanography 15 (1),
787 43–52.

788 Schmittner, A., Green, J., Wilmes, S., 2015. Glacial ocean over-turning intensified
789 by tidal mixing in a global circulation model. Geophys. Res. Lett 42, 2.

790 Schmittner, A., Gruber, N., Mix, A. C., Key, R. M., Tagliabue, A., Westberry,
791 T. K., 2013. Biology and air-sea gas exchange controls on the distribution of
792 carbon isotope ratios ($\delta^{13}C$) in the ocean. Biogeosciences Discussions 10 (5),
793 8415–8466.

794 Shackleton, N., 1977. Carbon 13 in Uvigerina: Tropical rainforest history and the
795 equatorial Pacific carbonate dissolution cycles. Marine science.

796 Spero, H. J., Bijma, J., Lea, D. W., Bemis, B. E., 1997. Effect of seawater
797 carbonate concentration on foraminiferal carbon and oxygen isotopes. Nature
798 390 (6659), 497–500.

799 Stern, J. V., Lisiecki, L. E., 2014. Termination 1 timing in radiocarbon-dated re-
800 gional benthic $\delta^{18}O$ stacks. Paleoceanography.

801 Tagliabue, A., Bopp, L., 2008. Towards understanding global variability in ocean
802 carbon-13. Global Biogeochemical Cycles 22 (1).

803 Tagliabue, A., Bopp, L., Roche, D., Bouttes, N., Dutay, J., Alkama, R., Kageyama,
804 M., Michel, E., Paillard, D., 2009. Quantifying the roles of ocean circulation
805 and biogeochemistry in governing ocean carbon-13 and atmospheric carbon
806 dioxide at the Last Glacial Maximum. Clim. Past 5, 695–706.

807 Walker, J. C. G., 1991. Numerical adventures with geochemical cycles. Oxford
808 University Press New York.

809 Wunsch, C., 1996. The Ocean Circulation Inverse Problem. Cambridge University
810 Press.

811 Wunsch, C., 2003. Determining paleoceanographic circulations, with emphasis
812 on the Last Glacial Maximum. Quaternary Science Reviews 22 (2-4), 371–385.

813 Wunsch, C., Ponte, R., Heimbach, P., 2007. Decadal trends in sea level patterns:
814 1993-2004. Journal of Climate 20 (24), 5889–5911.