Inferring Likelihoods and Climate System Characteristics from Climate Models and Multiple Tracers

K. Sham Bhat, Murali Haran, Roman Tonkonojenkov, and Klaus Keller *

Abstract

Characterizing the risks of anthropogenic climate change poses considerable statistical challenges. A key problem is how to combine the information contained in large-scale observational data sets with simulations of Earth system models in a statistically sound and computationally tractable manner. Here we describe a statistical approach for improving projections of the North Atlantic Meridional Overturning Circulation (AMOC). The AMOC is part of the global ocean conveyor belt circulation and transfers heat between low and high latitudes in the Atlantic basin. The AMOC might collapse in a "tipping point" response to anthropogenic climate forcings. Assessing the risk of an AMOC collapse is of considerable interest since it may result in major impacts on natural and human systems. AMOC projections rely on simulations from complex climate models. One key source of uncertainty in AMOC projections is uncertainty about background ocean vertical diffusivity (K_v), an important model parameter. K_v cannot be directly observed but can be inferred by combining climate model output with observations on the oceans (so called "tracers"). Here, we combine information from multiple tracers, each observed on a spatial grid. Our two-stage approach emulates the computationally expensive climate model using a flexible hierarchical model to connect the tracers. We then infer K_v using our emulator and the observations via a Bayesian approach, accounting for observation error and model discrepancy. We utilize kernel mixing and matrix identities in our Gaussian process model to reduce considerably the computational burdens imposed by the large data sets. We find that our approach is flexible, reduces identifiability issues, and enables inference about K_v based on large data sets. We use the resulting inference about K_v to improve probabilistic projections of the AMOC.

Keywords: computer model calibration, Bayesian hierarchical modeling, Gaussian process, computer experiments, multivariate spatial data, climate change.
Short Title: Inferring Climate System Characteristics.

*K. Sham Bhat is a graduate student, Department of Statistics, Pennsylvania State University, University Park, PA 16802 (E-mail: *kgb130@psu.edu*). Murali Haran is Associate Professor, Department of Statistics, Pennsylvania State University, University Park, PA 16802 (E-mail: *mharan@psu.edu*). Roman Tonkonojenkov is a graduate student, Department of Geosciences, Pennsylvania State University, University Park, PA 16802 (E-mail: *rzt2-wrk@psu.edu*). Klaus Keller is Associate Professor, Department of Geosciences, Pennsylvania State University, University Park, PA 16802 (E-mail: *klaus@psu.edu*). This work was partially supported by the National Science Foundation and from the US Geological Survey. Any opinions, findings, and conclusions expressed in this work are those of the authors alone, and do not necessarily reflect the views of the NSF and USGS. The authors thank Andreas Schmittner for providing the output of the published runs. The authors also thank Nathan Urban, Rui Paulo, and Andreas Schmittner for helpful discussions.

1 Introduction

Anthropogenic greenhouse gas emissions are changing the climate (Alley et al., 2007). Assessing the risks of future climate change requires estimates of the probability of specific outcomes (Schneider, 2001). Consider, for example, the potential collapse of the North Atlantic meridional overturning circulation (AMOC), which is part of the global 'ocean conveyor belt' circulation. This circulation system transfers heat from low to high latitudes in the Atlantic basin as cold, salty, and dense water sinks leaving warm water at the ocean surface. Both observations (Bryden et al., 2005) as well as climate model simulations (Cubasch et al., 2001; Meehl et al., 2007) indicate that the AMOC may weaken in response to anthropogenic forcings. The collapse of the AMOC could, for example, result in abrupt climate change, in particular major temperature and precipitation changes, and a shift in terrestrial ecosystems (Schneider et al., 2007; Vellinga and Wood, 2008). Delivering an early and accurate prediction of an approaching AMOC collapse may improve the design of risk management strategies and can have considerable expected economic value of information (Keller et al., 2004, 2007).

One key source of uncertainty in AMOC projections is the uncertainty about the parameterization of the processes driving vertical transport in the models (e.g. Kuhlbrodt et al., 2007; Schmittner and Weaver, 2001). A central problem is that many exchange processes in the real oceans occur at spatial scales that are not resolved by the coarse grids of Earth system models (Wunsch and Ferrari, 2004). The model parameterizations hence depend on parameter values that cannot be measured directly (Toole et al., 1994). Rather, these parameters are typically inferred from the distribution of ocean tracers (e.g. Goes et al., 2010; Matear and Wong, 1997; Schmittner et al., 2009). The UVic model runs analyzed in this study (and described in more detail in Schmittner et al. (2009)) uses state-of-the-art representations of transport processes (e.g., a spatially variable tidally induced mixing over tough topography, (cf. Simmons et al., 2004)). Yet, this model implementation still requires to

account for other, thus far neglected mixing processes to result in realistic tracer distribtions. The effect of these unresolved processes on vertical mixing is parameterized in Schmittner et al. (2009) as a globally constant background vertical diffusivity (K_v). Adopting a globally constant pattern for K_v is clearly a strong approximation, but arguably a reasonable first step that allows us to focus on the statistical issues discussed below. More recent work (e.g. Sriver et al., 2010). shows that representing additional mixing processes using a spatially resolved pattern has the potential to improve model hindcasts.

Reducing the uncertainty of K_v may also result in better predictions of AMOC strength (Dijkstra, 2008) as well as other climate predictions (Schmittner et al., 2009). Our main objective is to characterize and (if possible reduce) uncertainty about K_v. While K_v cannot be measured directly, there is a large amount of information about K_v in the form of oceanic tracers. Oceanic tracers are observations that provide information about ocean transport processes. In climate models, these tracers are strongly affected by the value of K_v (Goes et al., 2010; Schmittner et al., 2009). We can therefore use observations of these tracers to infer K_v. For example, larger observed values of Δ^{14}C in the deep oceans suggest a higher intensity of vertical mixing of the ocean. In this work, we use the tracers trichlorofluoromethane (CFC11) and Δ^{14}C, for which observations are available at many longitudes, latitudes, and ocean depths (Key et al., 2004). A detailed description of the data is provided in Section 4. These observations are subject to measurement error and may be sparse and irregularly observed over space. Often the total number of observations is on the order of thousands to millions (Key et al., 2004; Levitus, 1998).

To perform statistical inference on climate system parameters, we need to establish a relationship between the observations and the climate parameters. We accomplish this by using an Earth System Model of intermediate complexity (EMIC). EMICs simulate the complex phenomenon of the atmosphere and the oceans to derive hindcasts and projections of quantities such as temperature, precipitation, or concentrations of carbon dioxide, carbon isotopes, or chlorofluorocarbons under specific forcings and parameter values. The climate models are

complex computer codes representing the solution to a large set of differential equations that approximate physical, chemical, and biological processes (Weaver et al., 2001). EMICs often take weeks to months to execute for any given climate parameter setting. Hence, obtaining the output at a large number of parameter setting is computationally costly. Computer model emulation is a powerful approach pioneered by Sacks et al. (1989) to approximate the computer model by a stochastic process. Emulation provides the advantage of obtaining an approximate output at any parameter setting at a small fraction of the computational burden compared to the full model. Depending on the resolution of the locations, the output may be on the order of thousands to tens of thousands of data points on a spatial field at each climate parameter setting. In this paper, we analyze a previously published ensemble of runs from the University of Victoria (UVic) Earth System Climate Model as described in Schmittner et al. (2009).

We are interested in determining the climate parameters for which the model is most likely to reproduce the observations (Δ^{14}C, CFC11 data). Previous efforts to approach this problem, including Sansò et al. (2008), Drignei et al. (2008), Han et al. (2009), Forest et al. (2008), and Goes et al. (2010) require either heavy spatial aggregation or a restrictive covariance structure to allow analysis to be computationally feasible. Building on the framework in Kennedy and O'Hagan (2001), Sansò et al. (2008) describes a fully Bayesian approach for the calibration of computationally intensive climate models. Their approach is elegant and provides a joint model for both climate model output and observations. However, extending their approach to allow for a flexible model for multivariate tracers and large data sets may be challenging. Higdon et al. (2008) uses a principal components approach, while Bayarri et al. (2007a) uses wavelets to obtain computationally tractable approach. However, it is difficult to utilize this framework when jointly modeling multiple tracers in a flexible, direct, and easily interpretable manner.

Here we develop a two-stage approach to the calibration problem that provides a potentially useful avenue to mitigate these problems. Specifically we first emulate the climate

model output and then perform inference for the climate parameter K_v using the emulator and the observations while accounting for model discrepancy and observation error. We use kernel mixing (Higdon, 1998) and matrix identities (cf. Cressie and Johannesson, 2008; Stein, 2008) to enable tractable computation for Gaussian process models of large spatial data sets compared to previous approaches. We connect the two tracers using a flexible piecewise linear relationship in a hierarchical modeling approach (Royle and Berliner, 1999). We reduce the number of parameters estimated simultaneously using the two-stage model by estimating some of the parameters in the first stage, and the remaining parameters in the second stage, thereby reducing parameter identifiability issues and enabling us to obtain useful inference for K_v. Our approach improves on previous work by (i) enabling computer model calibration with larger spatial data sets and (ii) allowing for a flexible model to combine information from multiple spatial fields for computer model calibration.

The study in Schmittner et al. (2009) also makes inferences about the parameter K_v using both one-dimensional and three-dimensional observations. We note briefly some significant advantages of our approach over their methods. Both approaches in Schmittner et al. (2009) (for the 3D and 1D cases) are based on several simplifying assumptions such as that model-data misfit terms are uncorrelated in space and ignores the cross-correlation between tracers. Schmittner et al. (2009) acknowledge these limitation and as a remedy use a method which relaxes some of these assumptions. Unfortunately, computational difficulties preclude them from using this new method on a 3D grid, and thus they apply it to 1D globally averaged profiles. Goes et al. (2010) provides a computationally inexpensive approach that avoids computer model emulation approach, while modeling the cross-correlation between tracers in a statistically sound framework. However, their approach utilizes several simplifying assumptions and is applied to 1D globally averaged profiles. Note that applying either method to even a 2D field does introduce nontrivial computational requirements. We introduce a method that uses an emulator and accounts for spatial dependence and model discrepancy, while still considerably reducing computational expenses using a 2D field, thereby making it

much more widely applicable than the approaches above.

The remainder of the paper is organized as follows. In Section 2 we discuss the features of our two-stage approach for combining model output and observations. In Section 3, we describe our approach for multiple tracers and for handling the computational challenges posed by the size of data. In Section 4, we describe implementation details and in Section 5 we summarize our results. Finally, we conclude with a discussion, caveats, and avenues for future research in Section 6.

2 Model Description

In this section, we describe our model for inferring climate characteristics from the observations and model output of Δ^{14}C and CFC11. We use the following two-stage approach to analyze the data. In the first stage we emulate the climate model by fitting a Gaussian process to the computer model output. We combine the information from the model output for two tracers using a hierarchical modeling approach. In the second stage, we use the emulator to connect the climate parameter to the observations, while allowing for additional sources of uncertainty, such as model discrepancy and observation error. This approach has much in common with Bayarri et al. (2007b) in that we split the inference into two stages. Our computer model emulation step is an approach to inferring a probability model connecting the parameters to the observations. Our approach of splitting inference into two-stages can be seen as a way of 'cutting feedback' (suggested by Nicky Best, see Rougier, 2008a) or modularization (see Liu et al., 2009). In particular, we model the emulator based only on the model output, and not the observations. Modularization and cutting feedback has the advantage of disassociating the portion of the statistical model that are known to be accurate from the parts of the model that are less accurate (Rougier, 2008a). By modeling the emulator separately, we also easily obtain diagnostics regarding the accuracy of the emulator. We note that, in a sense, we are inferring a likelihood for the climate parameter by using the

output from the computer climate model at different values of the climate parameter (see also Rappold et al., 2007, for another scenario where the likelihood is derived). For ease of exposition, we first describe our model for a single tracer below.

We begin with some notation. Let $Z(\mathbf{s})$ be the observation of a single tracer at location $\mathbf{s}$, where $\mathbf{s}$=(latitude, depth). Let θ be a climate parameter. In this paper, we consider only one climate parameter, but this framework can be easily expanded for multiple climate parameters, where θ may be a vector. $Y(\mathbf{s}, \theta)$ denotes the climate model output at the location $\mathbf{s}$, and at the climate parameter setting θ. The spatial data from the climate model grid may or may not coincide with the locations of the observations. The objective here is to infer a posterior distribution of θ given the observed data and climate model output.

Let $\mathbf{Y} = (Y_{11}, \cdots, Y_{n1}, Y_{12}, \cdots, Y_{n2}, \cdots, Y_{1p}, \cdots, Y_{np})^T$, obtained by stacking computer model output at all climate parameter settings, denote the climate model output for a single tracer. Y_{ik} corresponds to the model output for location $\mathbf{s}_i$ and climate parameter setting θ_k, and n is the number of model output locations and p is the number of climate parameter settings. Since a location consists of two dimensions, let s_{i1} and s_{i2} be the latitude and depth respectively for location $\mathbf{s}_i$. Similarly, $\mathbf{Z} = (Z_1, \cdots, Z_N)^T$ are the observations for the tracer, where N is the total number of observations.

2.1 Climate Model Emulation

We model the climate model output $\mathbf{Y}$ using a Gaussian process:

$$\mathbf{Y} \mid \boldsymbol{\beta}, \theta, \boldsymbol{\xi} \sim N(\mu_{\boldsymbol{\beta}}, \Sigma(\boldsymbol{\xi}_y)),$$

where we assume a linear mean function, $\mu_{\boldsymbol{\beta}} = X\boldsymbol{\beta}$, with X is a covariate matrix of dimension $np \times b$, where there are $(b-1)$ covariates. The covariates we use are latitude, depth, and the climate parameter. Note that the covariates consist of all the coordinates used in constructing the covariance of the Gaussian process, including settings of the climate parameter θ. $\boldsymbol{\xi}_y$ is a

vector of covariance parameters that specify the covariance matrix $\Sigma(\boldsymbol{\xi}_y)$ (a specific example is described later in Section 3.3.1), and $\boldsymbol{\beta}$ is a vector of regression coefficients. Let the maximum likelihood estimate of $(\boldsymbol{\xi}_y, \boldsymbol{\beta})$ be $(\hat{\boldsymbol{\xi}}_y, \hat{\boldsymbol{\beta}})$. Let $\mathbf{S}$ be the set of locations where the observations were collected. Following the standard kriging framework (Cressie, 1993; Stein, 1999), the multinormal predictive distribution for the computer model output at a new θ at $\mathbf{S}$ is obtained by substituting $(\hat{\boldsymbol{\xi}}_y, \hat{\boldsymbol{\beta}})$ in place of $(\boldsymbol{\xi}_y, \boldsymbol{\beta})$ and conditioning on $\mathbf{Y}$. We denote the random variable with this predictive distribution by $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ in the second stage.

2.2 Inference for Climate Parameters

To infer θ based on the observations $\mathbf{Z}$, we need a probability model connecting θ and $\mathbf{Z}$. The predictive distribution from Section 2.1 provides a model for climate model output at any θ and any set of new locations. We now model the observations $\mathbf{Z}$ as realizations from a stochastic process obtained by adding additional error to the climate model emulator from Section 2.1. Our model for the observations $\mathbf{Z}$ is therefore

$$\mathbf{Z} = \boldsymbol{\eta}(\mathbf{Y}, \theta) + \boldsymbol{\delta}(\mathbf{S}) + \boldsymbol{\epsilon}$$

, where $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ is as described in Section 2.1, $\boldsymbol{\epsilon} \sim N(0, \psi I_N)$, and where $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_N)^T$ is the observation error with $\psi > 0$ as the observation error variance. $\boldsymbol{\delta}(\mathbf{S})$, the model discrepancy (or model error), is modeled as a zero-mean Gaussian process. While the zero-mean assumption appears to be a strong assumption, our experience when including a non-zero mean term resulted in identifiability issues. Further, the zero-mean Gaussian process appears to be flexible enough to account for the model discrepancy. Hence, $\boldsymbol{\delta}(\mathbf{S}) \sim N(\mathbf{0}, \Sigma_d(\boldsymbol{\xi}_d))$, where $\boldsymbol{\xi}_d$ is a vector of covariance parameters that specify the covariance matrix $\Sigma_d(\boldsymbol{\xi}_d)$. We have in essence 'inferred a likelihood' for use in our Bayesian framework, since for any fixed $\mathbf{Z}$, we can obtain a value of the likelihood for any θ.

Following Bayarri et al. (2007b), we allow for κ_y, the magnitude of the emulator spatial

variance, to be re-estimated here. We can now perform inference on θ, ψ, κ_y, and $\boldsymbol{\xi}_d$ by specifying a prior for these parameters. Using Markov Chain Monte Carlo (MCMC), we estimate a posterior distribution for θ, 'integrating out' the other parameters ψ, κ_y, and $\boldsymbol{\xi}_d$. We discuss prior selection for θ, ψ, κ_y, and $\boldsymbol{\xi}_d$ in Section 4. Note that the computational complexity of the second stage of our approach is solely dependent on N, the size of $\mathbf{Z}$, and not $M = np$, where M is the size of the ensemble of model output $\mathbf{Y}$.

3 Multivariate Tracer Model

In this section, we discuss how our approach can be used to combine information from multiple tracers. We use a hierarchical approach following Royle and Berliner (1999) to model the relationship of the climate model output from the two tracers. We extend our notation to allow for two tracers. Let $\mathbf{Y}_1 = (Y_{11}, \cdots, Y_{1np})^T$, and $\mathbf{Y}_2 = (Y_{21}, \cdots, Y_{2np})^T$ denote the climate model output for the two tracers Δ^{14}C and CFC11 respectively. Similarly, $\mathbf{Z}_1 = (Z_{11}, \cdots, Z_{1N})^T$ and $\mathbf{Z}_2 = (Z_{21}, \cdots, Y_{2N})^T$ are the observations for Δ^{14}C and CFC11. We will often refer to $\mathbf{Y} = (\mathbf{Y}_1\ \mathbf{Y}_2)$ and $\mathbf{Z} = (\mathbf{Z}_1\ \mathbf{Z}_2)$ when both tracers are used.

We build a joint model for $\mathbf{Y}_1$ and $\mathbf{Y}_2$ hierarchically, first modeling $\mathbf{Y}_2$ as a Gaussian process, then modeling $\mathbf{Y}_1 \mid \mathbf{Y}_2$ also as a Gaussian process. We include a mean function that provides additional flexibility when modeling the relationship between $\mathbf{Y}_1$ and $\mathbf{Y}_2$. Modeling the relationship between the tracers in the mean term allows us to include a richer relationship between the tracers, avoid estimating complex covariance relationships and reducing the complexity and nonstationarity of the covariance (Royle and Berliner, 1999).

3.1 Stage 1: Emulation for Multivariate Spatial Fields

The climate model output for the tracers are modeled using a hierarchical framework according to equation (1).

$$\begin{aligned} \mathbf{Y}_1 \mid \mathbf{Y}_2, \boldsymbol{\beta}_1, \boldsymbol{\xi}_{y1}, \gamma &\sim N(\mu_{\boldsymbol{\beta}_1} + \mathbf{B}(\gamma)\mathbf{Y}_2, \Sigma_{y_{1.2}}(\boldsymbol{\xi}_{y1})), \text{ and} \\ \mathbf{Y}_2 \mid \boldsymbol{\beta}_2, \boldsymbol{\xi}_{y2} &\sim N(\mu_{\boldsymbol{\beta}_2}, \Sigma_{y_2}(\boldsymbol{\xi}_{y2})), \end{aligned} \tag{1}$$

where $\mu_{\boldsymbol{\beta}_i}$ is a function of the climate parameters, and $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ are the coefficient vectors for $\mathbf{Y}_1, \mathbf{Y}_2$ respectively. We assume $\mu_{\boldsymbol{\beta}_i} = X\boldsymbol{\beta}_i$ (for i=1,2) where X is the covariate matrix of dimension $M \times b$, with covariates (latitude, depth, and climate parameters) as specified in Section 2.1. $B(\boldsymbol{\gamma})$ is an $M \times M$ matrix that specifies the relationship between $\mathbf{Y}_1$ and $\mathbf{Y}_2$, and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \cdots, \gamma_5)$. Since a simple linear relationship between the two tracers is not sufficient, our hierarchical approach using the B matrix below, allows for a more flexible model in describing the relationship between the two tracers, while resulting in a model that is still computationally tractable. Our model, which is informed by a combination of a scientific understanding of the relationship between the two tracers and exploratory data analysis (see Figure 1) of the model output, is described below:

$$\begin{aligned} &B_{ii} = \gamma_1 \quad \text{if} \quad \mathbf{s}_{i2} < 200, \mathbf{Y}_{i2} < 2.1 && B_{ii} = \gamma_2 \quad \text{if} \quad \mathbf{s}_{i2} < 200, \mathbf{Y}_{i2} > 2.1 \\ &B_{ii} = \gamma_3 \quad \text{if} \quad 200 < \mathbf{s}_{i2} < 800, \mathbf{Y}_{i2} < 1.8, && B_{ii} = \gamma_4 \quad \text{if} \quad 200 < \mathbf{s}_{i2} < 800, \mathbf{Y}_{i2} > 1.8, \\ &B_{ii} = \gamma_5 \quad \text{if} \quad 800 < \mathbf{s}_{i2} < 2500 && B_{ii} = 0 \quad \text{if} \quad 2500 < \mathbf{s}_{i2}, \text{ and} \\ &B_{ij} = 0 \quad \text{if} \quad i \neq j. \end{aligned}$$

Note that the mean for the ith value of Δ^{14}C (Y_{i1}) (and B_{ii}) depends on the depth at which it is obtained (s_{i2}) and the CFC11 value at the same location (Y_{i2}). The depth and CFC11 breakpoints were obtained by exploratory data analysis including the consideration of several alternative breakpoints.

$\Sigma_{y_{1.2}}(\boldsymbol{\xi}_{y1})$ is the conditional covariance matrix of $\mathbf{Y}_1|\mathbf{Y}_2$, and $\Sigma_{y_2}(\boldsymbol{\xi}_{y2})$ is the marginal covariance matrix of $\mathbf{Y}_2$, both have dimensions $M \times M$. $\boldsymbol{\xi}_{y1}, \boldsymbol{\xi}_{y2}$ are vectors of covariance parameters for $\mathbf{Y}_1 \mid \mathbf{Y}_2$ and $\mathbf{Y}_2$ respectively. As in Section 2.1, we find MLEs for $\boldsymbol{\xi}_{y1}, \boldsymbol{\beta}_1, \boldsymbol{\gamma}, \boldsymbol{\xi}_{y2}, \boldsymbol{\beta}_2$ and obtain a predictive distribution at $\mathbf{S}$ by plugging in the MLEs and conditioning on $\mathbf{Y}_1$ and $\mathbf{Y}_2$. We note that the log-likelihood for $\ell(\boldsymbol{\beta}_1, \boldsymbol{\xi}_{y1}, \boldsymbol{\beta}_2, \boldsymbol{\xi}_{y2}, \boldsymbol{\gamma} \mid \mathbf{Y}_1, \mathbf{Y}_2)$ can be broken up into the sum of two functions, $\ell_1(\boldsymbol{\beta}_1, \boldsymbol{\xi}_{y1}, \boldsymbol{\gamma} \mid \mathbf{Y}_1, \mathbf{Y}_2)$ and $\ell_2(\boldsymbol{\beta}_2, \boldsymbol{\xi}_{y2} \mid \mathbf{Y}_2)$, which reduces computations (see Appendix B).

We follow the same approach as in Section 2.1 to obtain $\boldsymbol{\eta}(\mathbf{Y}, \theta)$, which has a multinormal predictive distribution for each θ at $\mathbf{S}$. Recall that $\mathbf{Y} = (\mathbf{Y}_1 \ \mathbf{Y}_2)$. We first obtain the multinormal prediction for the second tracer (CFC11), $\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta)$, by following the standard kriging framework as in Section 2.1. We then obtain the prediction for tracer 1 given tracer 2, $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$, using the kriging framework and the model for $\mathbf{Y}_1|\mathbf{Y}_2$ from equation (1). $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$ is also a multivariate normal distribution. We can easily compute the distribution of $\mathbf{Y}$ in equation (2) below by combining the conditional distribution of $\mathbf{Y}_1 \mid \mathbf{Y}_2$ and the marginal distribution of $\mathbf{Y}_2$ and applying the laws of iterated expectations and variances. $\Sigma_{y_{1.2}}$ and Σ_{y_2} are the covariance matrices for $\mathbf{Y}_1|\mathbf{Y}_2$ and $\mathbf{Y}_2$ using the MLE for $\boldsymbol{\xi}_{y1}$ and $\boldsymbol{\xi}_{y2}$ respectively. We can in a similar manner derive the predictive distribution $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ from the conditional and marginal distributions described above.

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} X_y\hat{\boldsymbol{\beta}}_1 + \mathbf{B}(\hat{\boldsymbol{\gamma}})X_y\hat{\boldsymbol{\beta}}_2 \\ X_y\hat{\boldsymbol{\beta}}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{y_{1.2}} + B(\hat{\boldsymbol{\gamma}})\Sigma_{y_2}B^T(\hat{\boldsymbol{\gamma}}) & B(\hat{\boldsymbol{\gamma}})\Sigma_{y_2} \\ \Sigma_{y_2}B^T(\hat{\boldsymbol{\gamma}}) & \Sigma_{y_2} \end{bmatrix} \right) \quad (2)$$

3.2 Stage 2: Inference for Multiple Tracers

We model the observed data $\mathbf{Z} = (\mathbf{Z}_1 \ \mathbf{Z}_2)$ as follows:

$$\mathbf{Z} = \boldsymbol{\eta}(\mathbf{Y}, \theta) + \boldsymbol{\delta}(\mathbf{S}) + \boldsymbol{\epsilon},$$

where $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ is as described earlier in Section 3.1, and $\boldsymbol{\epsilon} = (\epsilon_{11}, \cdots, \epsilon_{N1}, \epsilon_{12}, \cdots, \epsilon_{N2})^T$ is the observation error. $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \Sigma_{\boldsymbol{\epsilon}})$, with

$$\Sigma_{\boldsymbol{\epsilon}} = \begin{bmatrix} \psi_1 I_N & \mathbf{0} \\ \mathbf{0} & \psi_2 I_N \end{bmatrix},$$

where $\psi_1, \psi_2 > 0$ are the observation error variances for the two tracers. The model discrepancy $\boldsymbol{\delta}(\mathbf{S})$ is modeled as a vector of two independent zero mean Gaussian processes below:

$$\boldsymbol{\delta}(\mathbf{S}) \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_{d1}(\boldsymbol{\xi}_{d1}) & \mathbf{0} \\ \mathbf{0} & \Sigma_{d2}(\boldsymbol{\xi}_{d2}) \end{bmatrix} \right).$$

The above independence assumption reduces identifiability issues and allows for greater computational efficiency. We allow the emulator spatial variance parameters from the first stage, κ_{y1} and κ_{y2}, to be re-estimated instead of using MLEs. We perform inference on θ, ψ_1, ψ_2, κ_{y1}, κ_{y2}, $\boldsymbol{\xi}_{d1}$, and $\boldsymbol{\xi}_{d2}$ as discussed in Section 2.2 by writing the log-likelihood $\ell(\theta, \psi_1, \psi_2, \boldsymbol{\xi}_{d1}, \boldsymbol{\xi}_{d2} \mid \mathbf{Z})$ as the sum of two multivariate normal functions by using a standard result for partitioned matrices (see Appendix B for details, Anderson, 2003). This reduces the matrix computations from dimension $2N \times 2N$ to $N \times N$. Using the observations $\mathbf{Z}$ and a prior specification on the parameters (details are discussed in Section 4), we estimate the posterior distributions of θ, ψ_1, ψ_2, κ_{y1}, κ_{y2}, $\boldsymbol{\xi}_{d1}$, and $\boldsymbol{\xi}_{d2}$ given $\mathbf{Z}$ using MCMC.

3.3 Computationally Tractable Models for Large Data Sets

When the number of observations is small, we can apply the methods in Section 2 and 3.1-3.2 with relative computational ease. However, for larger data sets such as the ocean tracer data set considered here, computing becomes practically infeasible even for a high performance computational cluster. In this section, we describe an approach that provides significant computational gains. Our methods take advantage of matrix rank reduction methods and

identities for matrices with special structure, such as the Sherman-Morrison-Woodbury Theorem (cf. Golub and Van Loan, 1996), used in spatial modeling contexts by Cressie and Johannesson (2008), Stein (2008), and Rougier (2008b).

3.3.1 Kernel Mixing

We briefly describe the kernel mixing approach, following Higdon (1998). This approach is useful both for computer model emulation as well as for modeling spatial dependence. Kernel mixing uses the fact that a continuous process can be created by convolving a continuous white noise process w with a convolution kernel k, thereby creating a continuous spatial process over the region D defined at any location $\mathbf{s}$ by

$$z(\mathbf{s}) = \int_D k(\mathbf{s} - \mathbf{u}) w(\mathbf{u}) d\mathbf{u}.$$

We replace the continuous white noise process w by a finite sum approximation $\mathbf{w}$ defined on a lattice $\mathbf{u}_1, \ldots, \mathbf{u}_J$ that covers the relevant region, and refer to each $(\mathbf{u}_j)$ as a 'knot location'. In addition to the finite sum approximation, we can also include a non-zero mean $\mu(\mathbf{s})$:

$$z(\mathbf{s}) = \sum_{j=1}^{J} k(\mathbf{s} - \mathbf{u}_j) w(\mathbf{u}_j) + \mu(\mathbf{s}),$$

where $w(\mathbf{u}_j)$ is the value of the white noise process at location $\mathbf{u}_j$. We assume the following kernel:

$$k(\mathbf{u}) = \kappa \exp\left\{ -\frac{\|\mathbf{u}\|^2}{\phi} \right\}.$$

This kernel corresponds to a Gaussian covariance function. While Stein (1999) cautions against the use of this kernel due to its smoothness, this seems to be a tenable assumption for modeling output from the climate model.

3.3.2 Implementation of Kernel Mixing

We apply kernel mixing to a large data set with N locations in order to obtain large computational gains by reducing the dimension of the matrices to be inverted from $N \times N$ to $J \times J$. Our set of knots are $((u_1, v_1, l_1), \ldots, (u_J, v_J, l_J))^T$, where (u_j, v_j, l_j) is the latitude, depth, and climate parameter value of the jth knot location. These knots define a lattice over the entire region of interest. Let $w(u_j, v_j, l_j)$ be the process at the jth knot, and is normally distributed with zero mean and variance 1. We express the random process for the model output $Y(\mathbf{s}_i, \theta_i)$ at location $\mathbf{s}_i$=(s_{i1}, s_{i2}) and climate parameter θ_i as follows,

$$Y(\mathbf{s}_i, \theta_i) \mid \mathbf{w}, \kappa_m, \boldsymbol{\beta}, \phi_s, \phi_o, \phi_c, \zeta \sim N\left(X(\theta_i)\boldsymbol{\beta} + \sum_{j=1}^{J} K_{ij}(\phi_s, \phi_o, \phi_c, \kappa_m) w(u_j, v_j, l_j), \zeta\right).$$

In the above equation, $\mathbf{w} = (w(u_1, v_1, l_1), \ldots, w(u_J, v_J, l_J))$, $\kappa_m > 0$ is a precision parameter, $\zeta > 0$ is the nugget, $\boldsymbol{\beta}$ are regression coefficients and ϕ_s, ϕ_o, $\phi_c > 0$ are parameters that describe dependence across latitude, depth, and climate parameter space respectively. $X_i(\theta_i) = (1, s_{1i}, s_{2i}, \theta_i)$ is the ith row of the covariate matrix as we assume a first order mean trend in the locations, depth, and the climate parameters. The kernel function is:

$$K_{ij}(\boldsymbol{\phi}_m, \kappa_m) = \sqrt{\kappa_m} \exp\left(-\frac{|s_{i1} - u_j|^2}{\phi_s^2} - \frac{|s_{i2} - v_j|^2}{\phi_o^2} - \frac{|\theta_i - l_j|^2}{\phi_c^2}\right). \tag{3}$$

where $\boldsymbol{\phi}_m = (\phi_s, \phi_o, \phi_c)$, $\kappa_m, \phi_s, \phi_o, \phi_c > 0$. The kernel is separable over location, depth, and climate parameters, although a nonseparable kernel could be chosen if appropriate.

We approximate the model discrepancy using the following kernel:

$$(K_d)_{ij}(\boldsymbol{\phi}_d, \kappa_d) = \sqrt{\kappa_d} \exp\left(-\frac{|s_{i1} - u_j|^2}{\phi_{ds}^2} - \frac{|s_{i2} - v_j|^2}{\phi_{do}^2}\right), \tag{4}$$

where $\boldsymbol{\phi}_d = (\phi_{ds}, \phi_{do})$, and $\kappa_d, \phi_{ds}, \phi_{do} > 0$.

Stage 1: Computer Model Emulation. For the model output of a single tracer, $\mathbf{Y}_i$, we

model the knot process $\mathbf{w}_i \sim N(0, I)$, and the kernel function $K_{yi}(\boldsymbol{\phi}_{yi}, \kappa_{yi})$ is defined in equation (3). Hence, the model is as described below, where X_Y is a covariate matrix of dimension $M \times b$ for the model output:

$$\mathbf{Y}_i \mid \boldsymbol{\beta}_i, \zeta_i, \boldsymbol{\phi}_{yi}, \kappa_{yi} \sim N(X_Y \boldsymbol{\beta}_i, \zeta_i I_N + K_{yi} K_{yi}^T),$$

where K_{yi} is a $M \times J$ dimensional kernel matrix. We estimate MLEs of the regression and covariance parameters $\boldsymbol{\beta}_i, \boldsymbol{\xi}_{yi} = (\zeta_i, \boldsymbol{\phi}_{yi}, \kappa_{yi})$, respectively. To do so, we must invert an $M \times M$ matrix for each evaluation of the likelihood, which is computationally expensive since $M = 5926$ in our dataset. However, by using kernel mixing to obtain a covariance matrix with a specific structure, $\Sigma_{y_i} = \zeta_i I + K_{yi} K_{yi}^T$, where $\zeta_i > 0$, and rewriting the inverse of this matrix by using the Sherman-Woodbury-Morrison identity (see Appendix A), the matrix inversions reduces to a $J \times J$ matrix. We include the nugget term ζ_i since it allows us to write the matrix in the desired form and also account for microscale variation (cf. Cressie, 1993, p. 59). When we have multiple tracers, we use two knot processes, $\mathbf{w}_1$ and $\mathbf{w}_2$, and follow the approach in Section 3.1 by splitting the log-likelihood into two functions and maximizing each separately. Note that unlike Higdon (1998), we do not work with the latent knot processes $(\mathbf{w}_1, \mathbf{w}_2)$. Instead, we use the kernel mixing approach to obtain a covariance matrix with a specific structure. We estimate MLEs for the following parameters using the computer model output: $\mathbf{Y}_1, \mathbf{Y}_2$: $\zeta_1, \kappa_{y1}, \boldsymbol{\phi}_{y1}, \boldsymbol{\gamma}, \boldsymbol{\beta}_1, \zeta_2, \kappa_{y2}, \boldsymbol{\phi}_{y2}, \boldsymbol{\beta}_2$.

When the number of observations, N, is small enough to invert an $N \times N$ covariance matrix directly, we follow the approach described in Sections 3.1 and 3.2 to obtain the predictive distribution $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ and infer θ based on the observations, $\mathbf{Z}$. We use the equations in (5) for the covariance matrices in equation (2). We rewrite the inverses of these matrices in a manner that requires inversions only of $J \times J$ matrices (see Appendix A).

$$\Sigma_{y_{1.2}} = \zeta_1 I + K_{y1} K_{y1}^T, \quad \Sigma_{y_2} = \zeta_2 I + K_{y2} K_{y2}^T. \tag{5}$$

When N is large, computation in the second stage becomes intractable. We develop a computationally tractable approach to computing the predictive distribution and modeling $\mathbf{Z}$. As in Section 3.2, we compute in closed form the distribution of the emulator predictions, $\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta)$ and $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$. We express the distribution of $\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta)$ below:

$$\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta) \sim N(\boldsymbol{\mu}^*_{z_2}, \Sigma^*_{z_2}),$$

where we express the mean vector $\boldsymbol{\mu}^*_{z_2}$ and covariance matrix $\Sigma^*_{z_2}$ in the form below in Equations (6) and (7) using kriging equations and matrix simplification:

$$\boldsymbol{\mu}^*_{z_2} = X_z(\theta)\boldsymbol{\beta}_2 + K_{z2}K_{y2}^T\Sigma_{y_2}^{-1}(\mathbf{Y}_2 - X_y\boldsymbol{\beta}_1), \text{ and} \tag{6}$$

$$\Sigma^*_{z_2} = \zeta_2 I + K_{z2}(I_J - K_{y2}^T\Sigma_{y_2}^{-1}K_{y2})K_{z2}^T. \tag{7}$$

We express the distribution of $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$, in a similar manner in equation (8) with mean vector and covariance matrix in equations (9) and (10):

$$\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta) \sim N(\boldsymbol{\mu}^*_{z_{1.2}}, \Sigma^*_{z_{1.2}}), \tag{8}$$

$$\boldsymbol{\mu}^*_{z_{1.2}} = X_z(\theta)\boldsymbol{\beta}_1 + K_{z1}K_{y1}^T\Sigma_{y_1}^{-1}(\mathbf{Y}_1 - X_y\boldsymbol{\beta}_1), \text{ and} \tag{9}$$

$$\Sigma^*_{z_{1.2}} = \zeta_1 I + K_{z1}(I_J - K_{y1}^T\Sigma_{y_1}^{-1}K_{y1})K_{z1}^T, \tag{10}$$

where K_{z1} and K_{z2} are the kernel matrices for the locations of the observations $\mathbf{S}$ and parameters $\boldsymbol{\beta}_1, \boldsymbol{\xi}_{y1}$ and $\boldsymbol{\beta}_2, \boldsymbol{\xi}_{y2}$ substituted respectively. $X_z(\theta)$ is a covariate matrix of dimension $N \times b$ with $\mathbf{S}$ and climate parameter θ as covariates. In addition, the covariance matrices $\Sigma^*_{z_{1.2}}$ and $\Sigma^*_{z_2}$ can be rewritten in a manner that requires matrix inversions only of $J \times J$ matrices (see Appendix A). We derive the distribution of $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ from the conditional and marginal distributions above.

Stage 2: Inference for Climate Parameters We model the observed data $\mathbf{Z} = (\mathbf{Z}_1\ \mathbf{Z}_2)$ as follows:

$$\mathbf{Z} = \boldsymbol{\eta}(\mathbf{Y}, \theta) + \boldsymbol{\delta}(\mathbf{S}) + \boldsymbol{\epsilon},$$

where $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ is as described earlier in this section, and $\boldsymbol{\epsilon} = (\epsilon_{11}, \cdots, \epsilon_{N1}, \epsilon_{12}, \cdots, \epsilon_{N2})^T$ is the observation error. $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \Sigma_{\boldsymbol{\epsilon}})$, with

$$\Sigma_{\boldsymbol{\epsilon}} = \begin{bmatrix} \psi_1 I_N & \mathbf{0} \\ \mathbf{0} & \psi_2 I_N \end{bmatrix},$$

where $\psi_1, \psi_2 > 0$ are the observation error variances for the two tracers.

The model discrepancy (or model error) $\boldsymbol{\delta}(\mathbf{S})$ is modeled as a vector of two independent zero mean Gaussian processes below:

$$\boldsymbol{\delta}(\mathbf{S}) \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_{d1}(\boldsymbol{\xi}_{d1}) & \mathbf{0} \\ \mathbf{0} & \Sigma_{d2}(\boldsymbol{\xi}_{d2}) \end{bmatrix} \right).$$

Using the kernel mixing approach in this section, we can write the covariance matrices as

$$\Sigma_{di} = K_{di} K_{di}^T,$$

and the kernel function K_{di}, a $N \times J$ dimensional kernel matrix, is defined in equation (4). We allow the emulator spatial variance parameters from the first stage, κ_{y1} and κ_{y2}, to be re-estimated, rather than using MLEs. We infer θ, ψ_1, ψ_2, κ_{y1}, κ_{y2}, $\boldsymbol{\xi}_{d1}$, and $\boldsymbol{\xi}_{d2}$ as discussed in Section 3.2 by writing the log-likelihood as the sum of two functions that are of the form of multivariate normal distributions with means and covariance matrices below (see Appendix B):

$$\mu_{z_{1.2}} = \mu^*_{z_{1.2}},\ \mu_{z_2} = \mu^*_{z_2},\ \Sigma_{z_{1.2}} = \psi_1 I_N + K_{d1} K_{d1}^T + \Sigma^*_{z_{1.2}},\ \Sigma_{z_2} = \psi_2 I_N + K_{d2} K_{d2}^T + \Sigma^*_{z_2}.$$

Matrix inversions can be reduced to dimension $J \times J$ (see Appendix A). The full details of the computation of matrix inverses and determinant are provided in Appendix C. Using the observations $\mathbf{Z}$, we estimate the posterior distributions of θ, ψ_1, ψ_2, κ_{y1}, κ_{y2}, $\boldsymbol{\xi}_{d1}$, and $\boldsymbol{\xi}_{d2}$ given $\mathbf{Z}$ using MCMC.

4 Application to Ocean Tracer Data

We apply our method to a data set of two ocean tracers, CFC11 (Chlorofluorocarbon-11) and Δ^{14}C ('Delta C-14' Carbon isotope). ^{14}C (radiocarbon) is a radioactive isotope of carbon, which is both produced naturally and by detonation of thermonuclear devices. ^{14}C enters the oceans from the atmosphere by air-sea gas exchange and is transported from the ocean by advection, diffusion, and to a lesser degree by biological processes (Key et al., 2004). Change in oceanic radiocarbon is reported as Δ^{14}C [per mil], which is the activity ratio relative to a set standard with a correction applied for fractionation. CFCs are chemical compounds produced starting in the 1930s. Similar to Δ^{14}C, CFC11 enters the oceans by air-sea gas exchange and is transported within the ocean by advection and diffusion (McCarthy et al., 1977). Oceanic CFC11 distributions are largely a function of oceanic currents and mixing as well as historical atmospheric concentration patterns. Oceanic CFC11 observations have been used previously to diagnose the skill of ocean models (England, 1995), and to estimate K_v (Schmittner et al., 2009).

4.1 Ocean tracer data description

Δ^{14}C and CFC11 measurements were collected by a large number of oceanographers (cf. Key et al., 2004) for all oceanic basins in the 1990s, with locations denoted by a latitude, longitude, and depth. The data have been quality controlled and gridded by Key et al. (2004). We use the observations from the data synthesis project by Key et al. (2004). We average the observations zonally (i.e. aggregated over longitudes), resulting in a data set

with $N = 3,706$ locations of latitude and depth. Zonal averaging is a standard procedure in atmospheric and oceanic sciences. It allows an easier visualization of climate features and also reduces the data dimension (cf. Doney et al., 2004).

We analyze model output at $p = 6$ different values of K_v, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5 cm^2/s, on a grid of locations of latitude, longitude, and an ocean depth from the University of Victoria (UVic) Earth System Climate Model as described in Schmittner et al. (2009). The model output were zonally aggregated, providing a 'blurred' snapshot representing an average between 1990-2000. The total number of locations in the model output is $n = 988$, resulting in $M{=}np{=}5928$ model output values. We exclude observations from latitudes above 60° N and depths below 3000 m to minimize problems due to sparse sampling (Key et al., 2004) and model artifacts (Schmittner et al., 2009). Observations and model output are on different spatial resolutions. Often, this is corrected by regridding the data and model output to the same spatial resolution (e.g. Schmittner et al., 2009). However, our approach does not require this regridding step because we use an emulator to predict at the observation locations. One key difference is that our approach does include the effects of interpolation uncertainty. Approaches which model the difference between the observations and model output without using an emulator typically ignore the uncertainties due to regridding the model output to coincide with data. Hence, we believe using computer model emulation using Gaussian processes, as we have here, provides a very significant advantage over other approaches.

The model runs are set up such that ^{14}C model output can be compared to the observations of $\Delta^{14}C$ (Meissner, 2007; Schmittner et al., 2009). To convert model output of CFC11 [mol/m^3] into units of the observations of CFC11 [$pmol/kg^1$], in-situ temperature was calculated following Bryden (1973) and Fofonoff (1977), and using the 1980 UNESCO International Equation of State (IES80) (UNESCO, 1981). We calculate model densities and temperatures from the model ocean pressure field; depth and latitude for the seawater density are from simplified equations following Lovett (1978).

The difference between the observations and the model output at each value of K_v is shown in Figures 6.2 and 3 for CFC11 and Δ^{14}C respectively. The spatial structure of the residuals also indicates structural model errors that are dependent on the analyzed tracer. These spatial patterns and the differences across tracers can point to promising avenues to reduce structural errors in the Earth system model of intermediate complexity (cf. Schmittner et al., 2009).

4.2 Implementation details

Here we discuss some of the details of the application of our approach to the ocean tracer data. Due to the data and model output being large spatial fields, we use the computationally tractable approach described in Section 3.3. We account for the curvature of the earth by using a geodesic distance formula to determine the distance between locations (see Banerjee, 2005). There is a non-directional anisotropy between latitude and depth, for which Roemmich (1983) suggests a correction ratio of 100 km distance to 1000 m depth. Our approach accounts for the anisotropy is to estimate different range parameters for the two dimensions rather than combining them into a single distance metric. We implement our approach for the tracers as two separate univariate data sets, and as bivariate data by combining the information of the two tracers. To select the knots, we chose 7 equally spaced latitude locations, 7 depths, and 4 different climate parameter values. We selected each combination of these values as a knot, for a total of J=196 knots. Using these knots appear to produce a reasonable model based on cross-validation analysis. Also in analyses using more knots (say 360), we found virtually the same inferential results. We verify the emulator using a cross-validation approach, where we held out the model output for one climate parameter (Rougier, 2008b) and predicted at all locations for that climate parameter setting. We also use a spatial cross-validation approach, where we withheld approximately one-ninth of the locations for all climate parameter settings, and used the emulator to predict at those locations. In both cross-validation studies, we find that predictions at the withheld climate parameter values and hold out locations are visually

similar to the actual model output (Figure 4). The results of cross-validation show that we have enough information to learn about the computer model based on the relatively small number of runs we have.

In addition, the emulator we use in the first stage provides a better fit to the model output than several less flexible alternative approaches. In fact, our ability to easily study this highlights one of the advantages of using a two-stage approach to computer model calibration. Our results are not surprising given the fact that our model (for e.g. deriving an appropriate mean function relating the two tracers) is influenced by our extensive exploratory data analysis. For example, we found that our flexible conditional hierarchical emulator results in a higher maximized log likelihood than does the model where the two tracers are assumed to be independent. In addition, our model results in a higher maximized likelihood than an alternative model using a separable cross-covariance for the tracers. Similarly, our model results in a substantially higher maximized likelihood than the alternative model which assumes a simple linear function as the relationship between the two tracers in the mean function.

In the second stage, we use Markov Chain Monte Carlo (MCMC) methods to obtain the posterior distributions of θ. We use two different priors on θ, a Lognormal (-1.55, 0.59) cm^2/s and a Uniform (0, 0.60) cm^2/s. This lognormal prior puts most of its mass on smaller values of K_v (see Figure 5). We use a wide inverse gamma prior for the observation error and model discrepancy variances, specifically $\psi_1 \sim IG(2, 100)$ and $\kappa_{d1} \sim IG(2, 100)$ for $\Delta^{14}C$ and $\psi_2 \sim IG(2, 5)$ and $\kappa_{d2} \sim IG(2, 5)$ for CFC11. We use wide uniform priors for the range parameters for model discrepancy. We use narrow priors for the emulator spatial variance parameters from the first stage that are centered around their MLEs.

These priors were obtained after an exploratory analysis of the data suggested the approximate scale of these parameters. Our priors are fairly wide with infinite variance (except for the emulator spatial variances, κ_{y1} and κ_{y2}), and are weakly informative given the available data, as also verified by prior sensitivity analysis. The only sensitivity was for the priors

selected for κ_{y1} and κ_{y2}; we found that we needed tight priors on these parameters to obtain reasonable inferential results.

5 Results

5.1 Ocean tracer data results

In this section we present the results from our analyses using the tracers Δ^{14}C and CFC11. While there is substantial overlap among the posterior distributions of K_v obtained by using Δ^{14}C and CFC11, separately and then jointly, there are also clear differences (Figure 5). We calculated credible regions using the Highest Posterior Density (HPD) method (Chen et al., 2000). The 90% credible region for K_v using the single tracer CFC11 is between 0.17 and 0.40 cm^2/s, the 90% credible region for K_v using the single tracer Δ^{14}C is between 0.16 and 0.35 cm^2/s, and the 90% credible region for K_v using the tracers jointly is between 0.13 and 0.30 cm^2/s. We note that the posterior distribution of K_v using the tracers jointly is to the left of the posterior distribution using the single tracers. We also use the estimates of the parameters from the sample-based inference in the second stage to obtain posterior predictions of observations of Δ^{14}C and CFC11. The posterior predictions for the Δ^{14}C and CFC11 observations using the both the bivariate approach and the univariate approach appear to be similar to the observations (Figure 6). Posterior predictions of Δ^{14}C appear to differ slightly from the observations in shallow waters in the Southern Ocean, and differ both in moderate and deep tropical waters for CFC11.

A reduction in the uncertainty in the value of K_v reduces uncertainty for other climate characteristics (cf. Forest et al., 2002), including AMOC strength projections (cf. Dijkstra, 2008). The results shown in Figure 7 demonstrate that the model hindcasts and projections of AMOC strength are strongly affected by the K_v parameter; the AMOC strength projections are obtained from the same climate model for the six K_v values between 0.05 and 0.50 cm^2/s.

This finding is consistent with previous studies (e.g. Dalan et al., 2005a,b; Schmittner and Weaver, 2001). Our results, as well as the previous studies discussed above, suggest a smooth relationship between the K_v value and the AMOC intensity (Figure 8(a)). We hence map every value of K_v from our sample based posterior distribution to a value of AMOC strength in 2100 using a loess fit (Figure 8(a)), and thus obtain a distribution of AMOC strength in 2100. We note that this approach potentially ignores the uncertainty in the relationship between K_v and the AMOC intensity; however, assuming a smooth relationship between them appears to be reasonable and it may be difficult to capture additional uncertainties given the information at hand.

As is apparent from our results, there is a substantial reduction in the predicted uncertainty of AMOC strength in the model when the posterior distribution of K_v from tracers Δ^{14}C and CFC11 combined is used for prediction compared to using the Lognormal prior—the range of the 90 % credible interval is reduced by a factor of 2 (Figure 8(b)). A possible concern is the potential effect of the specific ordering chosen for the hierarchical model. To investigate this issue, we re-ran our entire analysis where Y_1 and Y_2 are CFC11 and $\delta^{14}C$ respectively (instead of the other way around.) We find that the resulting inference is virtually identical to the results we summarize here.

5.2 Computation

To ensure convergence of our MCMC based estimates in the second stage of our calibration approach, we obtained Monte Carlo standard errors for the posterior mean estimates of θ and other parameters computed by consistent batch means (Flegal et al., 2008; Jones et al., 2006). The posterior mean estimates of θ had MCMC standard errors below 10^{-4} for both the univariate and bivariate approaches. The MCMC standard errors for the other parameters were less than 10^{-3} for both the univariate and bivariate approaches. Computing the MLEs of covariance and regression parameters in the first stage required approximately 4.5 hours.

The computer programs were implemented in R (Ihaka and Gentleman, 1996) using a 3.0 GHz Intel Xeon on a Dell PowerEdge server with 32GB of RAM. We computed MLEs in the first stage using a differential evolution algorithm (Ardia, 2007; Storn and Price, 1997). Obtaining the 200,000 samples using MCMC in the second stage required approximately 15 hours and about 80 hours of computer time for the univariate and the bivariate approaches, respectively.

To obtain the posterior density of K_v using the Uniform prior, we used importance resampling (cf. Gelman et al., 2004, p. 450) from 1000 thinned, approximately independent, samples already obtained for the posterior distribution using the Log Normal prior. The distribution for K_v using the Uniform prior and the Log Normal prior were very similar for both the univariate and bivariate approaches.

One possible concern is our bivariate spatial model may be affected by the specific ordering chosen for the hierarchical specification, (Y_1 and Y_2 are the tracers $\Delta^{14}C$ and CFC11). In theory, the posterior distribution of θ are different for both orderings for the hierarchical specification. However, we applied our approach to the other hierarchical ordering (Y_1 and Y_2 are CFC11 and $\Delta^{14}C$), found that the posterior distribution of K_v was very similar regardless of the ordering for this data set. This however, does not imply that such a result would hold for all data sets.

6 Discussion

6.1 Summary

We develop and apply a novel approach for inferring climate parameters by combining information from observations and climate model output, while accounting for observation error and model discrepancy. We use our approach to learn about the climate model parameter K_v and are able to use the posterior distribution of K_v to make predictions about the strength of

the AMOC. Our methodology allows for a flexible model for relating multiple ocean tracers and accounts for spatial dependence among the observations. For a computationally tractable approach for large data sets, we model dependence using special covariance structures and take advantage of matrix identities that allow for fast computations. Our approach can, in principle, be extended to larger spatial datasets and to spatiotemporal data. Our hierarchical approach allows us to model the relationship between tracers utilizing domain expertise and exploratory data analysis, while still remaining computationally efficient. While we can apply our approach to modeling hierarchies of three or more tracers, this may prove challenging in practice as the number of tracers gets large, especially if there is relatively little expertise to guide the development of the hierarchy. In fact, this is a criticism of the hierarchical approach to modeling of multivariate spatial fields (following Royle and Berliner, 1999) in general. A potentially important advantage of our approach is that it allows us to model relationships that are not easily captured otherwise. In particular, the method enables the use of exploratory data analysis and knowledge of changing relationships between the spatial fields (our 'tracers') at different depths to determine a mean function that directly captures this non-linear relationship between them.

Note that the kernel mixing approach representing covariances allows for additional flexibility, for example the ability to model non-stationarity (Higdon et al., 1999; Paciorek and Schervish, 2006) or to allow the covariance to be nonseparable (Calder et al., 2002). This additional flexibility may be useful in developing more realistic models for a spatial process when different regions in the ocean have very different climate properties or when the interactions between the (multiple) climate parameters and spatial dimensions need to be included in the model.

6.2 Caveats

While including a model discrepancy term generally results in more accurate inference about the calibration parameter, caution is required in interpreting the inferential results of both the climate parameter as well as any other parameters (cf. Bayarri et al., 2007b; Goes et al., 2010; Liu et al., 2009). We now offer several caveats concerning our scientific conclusions. First, our analysis neglects the model uncertainty when we use the simple emulator connecting AMOC strength to K_v. Second, the sparse sampling of K_v may introduce considerable biases in the emulator. Having more samples of K_v or obtaining these samples using design of experiments techniques may reduce these errors. However, our cross-validation results suggests that this is not a major issue in our study. Third, we only use a single observational synthesis data set rather than comparing several data sets. In other words, we neglect the potential effects of structural uncertainty in deriving the synthesized data set. While we believe the data we use (Key et al., 2004) is the best available, there are likely considerable biases and uncertainties from any single data set based on relatively sparse observations. Fourth, the UVic climate model has a simplified representation of internal climate variability. However the current state-of-the-art climate models all have considerable shortcomings in representing important modes of internal climate variability (cf. Meehl et al., 2007). We hope that accounting for structural model errors in our approach through the model discrepancy term mitigates this issue to some degree.

Appendix A: Matrix Identities

The Sherman-Woodbury-Morrison identity states that the inverse of a matrix of the form $A + UCV$, where A is of dimension $N \times N$, U is dimension $N \times J$, V is dimension $J \times N$, and C is dimension $J \times J$ can be expressed as the following,

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}.$$

The determinant of a matrix $A + UCV$ can be expressed as the following,

$$|A + UCV| = |C^{-1} + VA^{-1}U| \times |C| \times |A|.$$

using the matrix determinant lemma (Harville, 2008). This identity reduces matrix inversions and determinant computations to dimension J rather than N (cf. p. 50 Golub and Van Loan, 1996).

The matrix form $(\psi I_N + KK^T)$ comes up regularly in our computations, for which we obtain the inverse and determinant (using Sylvester's Theorem, see Golub and Van Loan (1996)) below in equation (11), which only require computations of matrices of dimension $J \times J$.

$$\begin{aligned} (\psi I_N + K(I_J)^{-1}K^T)^{-1} &= \frac{I_N}{\psi} - \frac{K}{\psi}\left(I_J - \frac{K^T K}{\psi}\right)^{-1}\frac{K^T}{\psi} \\ |\psi I_N + K(I_J)^{-1}K^T| &= \left|I_J - \frac{K^T K}{\psi}\right| \cdot \psi^N \end{aligned} \tag{11}$$

To compute the likelihoods in this paper, we compute the Cholesky decomposition of the matrix $\left(I_J - \frac{K^T K}{\psi}\right)$ rather than the inverse directly, which also reduces the computation time of the determinant (Golub and Van Loan, 1996).

Appendix B: Likelihood Computation Using Partitioned Matrices

In this section, we describe the simplification of the computation of the likelihood using the decomposition of a partitioned matrix, which reduces the dimensionality of matrix inversions.

The distribution of $\mathbf{Z}$ is as follows:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\mu}_{z_1} \\ \boldsymbol{\mu}_{z_2} \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right),$$

Then denoting the mean vector and covariance matrix of $\mathbf{Z}_1 \mid \mathbf{Z}_2$ by $\mu_{11.2} = \boldsymbol{\mu}_{z_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{Z}_2 - \boldsymbol{\mu}_{z_2})$ and $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ respectively, the log-likelihood of $\mathbf{Z}$ can be written as follows:

$$-\frac{1}{2}\log(|\Sigma_{11.2}|) - \frac{1}{2}(\mathbf{Z}_1 - \mu_{11.2})^T\Sigma_{11.2}^{-1}(\mathbf{Z}_1 - \mu_{11.2}) - \frac{1}{2}\log(|\Sigma_2|) - \frac{1}{2}(\mathbf{Z}_2 - \boldsymbol{\mu}_{z_2})^T\Sigma_2^{-1}(\mathbf{Z}_2 - \boldsymbol{\mu}_{z_2})$$

All computations here involve matrices of dimension $N \times N$. By using the matrix structure in Appendix A, these computations are reduced to dimension $J \times J$.

Appendix C: Details About Computation of Inverse and Determinant

The matrices in Section 3.3.2 are expressed as

$$\psi_i I_N + K_{di}K_{di}^T + \Sigma_{z_i}^* = (\psi_i + \zeta_i)I_N + K_{di}K_{di}^T + K_{zi}(I_J - K_{yi}^T\Sigma_{y_i}^{-1}K_{yi})K_{zi}^T \tag{12}$$

Using the matrix form for the Sherman-Woodbury-Morrison Theorem from Appendix A, we let $A = (\psi_i + \zeta_i)I_N + K_{di}K_{di}^T$, $U = K_{zi}$ $V = K_{zi}^T$, and $C = I_J - K_{yi}^T\Sigma_{y_i}^{-1}K_{yi}$. We can clearly compute the inverse and determinant of the matrix in (12) using matrix computations with dimension $J \times J$ (see Appendix A) if we can perform the computations of the matrices A and C. We use the identities in (11) to rewrite the A into a form that requires matrix computations of dimension $J \times J$. For C, which is of dimension $J \times J$, we need only compute $\Sigma_{y_i} = \zeta_i I + K_{yi}K_{yi}^T$, which can also be computed using (11).

References

Alley, R., Berntsen, T., Bindoff, N. L., Chen, Z., Chidthaisong, A., Friedlingstein, P., Gregory, J., Hegerl, G., Heimann, M., Hewitson, B., Hoskins, B., Joos, F., Jouzel, J., Kattsov, V., Lohmann, U., Manning, M., Matsuno, T., Molina, M., Nicholls, N., Overpeck, J., Qin, D., Raga, G., Ramaswamy, V., Ren, J., Rusticucci, M., Solomon, S., Somerville, R., Stocker, T. F., Stott, P., Stouffer, R. J., Whetton, P., Wood, R. A., and Wratt, D. (2007). *Climate Change 2007: The Physical Science Basis: Summary for Policymakers: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* Wiley Series in Probability and Statistics.

Ardia, D. (2007). The DEoptim package: Differential Evolution Optimization. *R Foundation for Statistical Computing*.

Banerjee, S. (2005). On Geodetic Distance Computations in Spatial Modeling. *Biometrics*, 61(2):617–625.

Bayarri, M., Berger, J., Cafeo, J., Garcia-Donato, G., Liu, F., Palomo, J., Parthasarathy, R., Paulo, R., Sacks, J., and Walsh, D. (2007a). Computer Model Validation with Functional Output. *The Annals of Statistics*, 35(5):1874–1906.

Bayarri, M., Berger, J., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C., et al. (2007b). A Framework for Validation of Computer Models. *Technometrics*, 49(2):138–154.

Bryden, H. (1973). New Polynomials for Thermal Expansion, Adiabatic Temperature Gradient and Potential Temperature of Sea Water. *Deep Sea Res*, 20:655–657.

Bryden, H., Longworth, H., and Cunningham, S. (2005). Slowing of the Atlantic Meridional Overturning Circulation at 25 N. *Nature*, 438:401–408.

Calder, C. A., Holloman, C. H., and Higdon, D. M. (2002). Exploring space time structure in ozone concentration using a dynamic process convolution model. In *Case Studies in Bayesian Statistics Volume VI*, pages 165–177. Springer-Verlag Inc.

Chen, M., Shao, Q., and Ibrahim, J. (2000). *Monte Carlo Methods in Bayesian Computation*. Springer.

Cressie, N. and Johannesson, G. (2008). Fixed Rank Kriging for Very Large Spatial Data Sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):209–226.

Cressie, N. A. (1993). *Statistics for Spatial Data*. John Wiley & Sons, New York, 2nd. edition.

Cubasch, U., Meehl, G., Boer, G., Stouffer, R., and Dix, M. (2001). Coauthors, 2001: Projections of Future Climate Change. *Climate Change 2001: The Scientific Basis*, pages 525–582.

Dalan, F., Stone, P., Kamenkovich, I., and Scott, J. (2005a). Sensitivity of the Oceans Climate to Diapycnal Diffusivity in an EMIC. Part I: Equilibrium State. *Journal of Climate*, 18(13):2460–2481.

Dalan, F., Stone, P., and Sokolov, A. (2005b). Sensitivity of the Oceans Climate to Diapycnal Diffusivity in an EMIC. Part II: Global Warming Scenario. *Journal of Climate*, 18(13):2482–2496.

Dijkstra, H. (2008). Scaling of the Atlantic Meridional Overturning Circulation in a Global Ocean Model. *Tellus A*, 60(4):749–760.

Doney, S., Lindsay, K., Caldeira, K., Campin, J., Drange, H., Dutay, J., Follows, M., Gao, Y., Gnanadesikan, A., Gruber, N., et al. (2004). Evaluating Global Ocean Carbon Models: The Importance of Realistic Physics. *Global Biogeochem. Cycles*, 18(3).

Drignei, D., Forest, C., and Nychka, D. (2008). Parameter Estimation for Computationally Intensive Nonlinear Regression with an Application to Climate Modeling. *Ann. Appl. Statist*, 2:1217–1230.

England, M. (1995). Using Chlorofluorocarbons to Assess Ocean Climate Models. *Geophysical Research Letters*, 22(22):3051–3054.

Flegal, J., Haran, M., and Jones, G. (2008). Markov Chain Monte Carlo: Can We Trust the Third Significant Figure? *Statist. Sci*, 23(2):250–260.

Fofonoff, N. (1977). Computation of Potential Temperature of Seawater for an Arbitrary Reference Pressure. *Deep-Sea Res*, 24:489–491.

Forest, C., Stone, P., and Sokolov, A. (2008). Constraining Climate Model Parameters from Observed 20th Century Changes. *Tellus A*, 60(5):911–920.

Forest, C., Stone, P., Sokolov, A., Allen, M., and Webster, M. (2002). Quantifying Uncertainties in Climate System Properties with the use of Recent Climate Observations. *Science*, 295(5552):113–117.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis*. CRC Press.

Goes, M., Urban, N., Tonkonojenkov, R., Haran, M., and Keller, K. (2010). The Skill of Different Ocean Tracers in Reducing Uncertainties About Projections of the Atlantic Meridional Overturning Circulation. *Journal of Geophysical Research-Oceans*, in press.

Golub, G. and Van Loan, C. (1996). *Matrix Computations*. Johns Hopkins University Press.

Han, G., Santner, T. J., Notz, W. I., and Bartel, D. L. (2009). Prediction for Computer Experiments Having Quantitative and Qualitative Input Variables. *Technometrics, to appear*.

Harville, D. (2008). *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag New York Inc.

Higdon, D. (1998). A process-convolution approach to modelling temperatures in the North Atlantic Ocean (Disc: p191-192). *Environmental and Ecological Statistics*, 5:173–190.

Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008). Computer Model Calibration Using High-Dimensional Output. *Journal of the American Statistical Association*, 103(482):570–583.

Higdon, D., Swall, J., and Kern, J. (1999). Non-stationary spatial modeling. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A., editors, *Bayesian Statistics 6 – Proceedings of the Sixth Valencia International Meeting*, pages 761–768. Clarendon Press [Oxford University Press].

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5:299–314.

Jones, G. L., Haran, M., Caffo, B. S., and Neath, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. *Journal of the American Statistical Association*, 101:1537–1547.

Keller, K., Bolker, B., and Bradford, D. (2004). Uncertain climate Thresholds and Optimal Economic Growth. *Journal of Environmental Economics and Management*, 48(1):723–741.

Keller, K., Deutsch, C., Hall, M., and Bradford, D. (2007). Early Detection of Changes in the North Atlantic Meridional Overturning Circulation: Implications for the Design of Ocean Observation Systems. *Journal of Climate*, 20(2):145–157.

Kennedy, M. and O'Hagan, A. (2001). Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(3):425–464.

Key, R., Kozyr, A., Sabine, C., Lee, K., Wanninkhof, R., Bullister, J., Feely, R., Millero, F., Mordy, C., and Peng, T. (2004). A Global Ocean Carbon Climatology: Results from Global Data Analysis Project (GLODAP). *Global Biogeochem. Cycles*, 18(4).

Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S. (2007). On the Driving Processes of the Atlantic Meridional Overturning Circulation. *Rev. Geophys*, 45.

Levitus, S. (1998). *World Ocean Database 1998*. US Dept. of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service.

Liu, F., Bayarri, M., and Berger, J. (2009). Modularization in Bayesian Analysis, with Emphasis on Analysis of Computer Models. *Bayesian Analysis*, 4(1):119–150.

Lovett, J. (1978). Merged Seawater Sound-speed Equations. *The Journal of the Acoustical Society of America*, 63:1713.

Matear, R. and Wong, C. (1997). Estimation of Vertical Mixing in the Upper Ocean at Station P from Chlorofluorcarbons. *Journal of Marine Research*, 55(3):507–521.

McCarthy, R., Bower, F., and Jesson, J. (1977). The Fluorocarbon-ozone theory, I. Production and Release: World Production and Release of CCl3F and CCl2F2 (fluorocarbons 11 and 12) Through 1975. *Atmospheric Environment (1967)*, 11(6):491–497.

Meehl, G., Stocker, T., Collins, W., Friedlingstein, P., Gaye, A., Gregory, J., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S., Watterson, I., Weaver, A., and Zhao, Z.-C. (2007). *Climate Change 2007: The Physical Science Basis. Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*.

Meissner, K. J. (2007). Younger Dryas: A Data to Model Comparison to Constrain the Strength of the Overturning Circulation. *Geophysical Research Letters*, 34, L21705.

Paciorek, C. and Schervish, M. (2006). Spatial Modelling Using a new Class of Nonstationary Covariance Functions. *Environmetrics (London, Ont.)*, 17(5):483.

Rappold, A., Lavine, M., and Lozier, S. (2007). Subjective Likelihood for the Assessment of Trends in the Ocean's Mixed-Layer Depth. *Journal of the American Statistical Association*, 102(479):771.

Roemmich, D. (1983). Optimal Estimation of Hydrographic Station Data and Derived Fields. *Journal of Physical Oceanography*, 13(8):1544–1549.

Rougier, J. (2008a). Comment on Article by Sanso et al. *Bayesian Analysis*, 3(1):45–56.

Rougier, J. (2008b). Efficient Emulators for Multivariate Deterministic Functions. *Journal of Computational and Graphical Statistics*, 17(4):827–843.

Royle, J. and Berliner, L. (1999). A Hierarchical Approach to Multivariate Spatial Modeling and Prediction. *Journal of Agricultural, Biological, and Environmental Statistics*, 4(1):29–56.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments (C/R: P423-435). *Statistical Science*, 4:409–423.

Sansò, B., Forest, C., and Zantedeschi, D. (2008). Inferring Climate System Properties Using a Computer Model. *Bayesian Analysis*, 3(1):1–38.

Schmittner, A., Urban, N., Keller, K., and Matthews, D. (2009). Using tracer observations to reduce the uncertainty of ocean diapycnal mixing and climate carbon-cycle projections. *Global Biogeochemical Cycles*, in press, doi:10.1029/2008GB003421.

Schmittner, A. and Weaver, A. (2001). Dependence of Multiple Climate States on Ocean Mixing Parameters. *Geophys. Res. Lett*, 28(6):1027–1030.

Schneider, S. (2001). What is 'Dangerous' Climate Change? *Nature*, 411(6833):17–19.

Schneider, S., Semenov, S., Patwardhan, A., Burton, I., Magadza, C., Oppenheimer, M., Pittock, A., Rahman, A., Smith, J., Suarez, A., Yamin, F., Corfee-Morlot, J., Finkel, A., Fssel, H.-M., Keller, K., MacMynowski, D., Mastrandrea, M. D., Todorov, A., Sukumar, R., van Ypersele, J.-P., and Zillman, J. (2007). Assessing key vulnerabilities and the risk from climate change. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds.*, pages 779–810.

Simmons, H., Jayne, S., Laurent, L., and Weaver, A. (2004). Tidally Driven Mixing in a Numerical Model of the Ocean General Circulation. *Ocean Modelling*, 6(3-4):245–263.

Sriver, R., Goes, M., Mann, M., and Keller, K. (2010). Climate Response to Tropical Cyclone-Induced Ocean Mixing in an Earth System Model of Intermediate Complexity. Technical report.

Stein, M. (2008). A Modeling Approach for Large Spatial Datasets. *Journal of the Korean Statistical Society*, 37:3–10.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer-Verlag Inc.

Storn, R. and Price, K. (1997). Differential Evolution–A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11(4):341–359.

Toole, J., Schmitt, R., and Polzin, K. (1994). Estimates of Diapycnal Mixing in the Abyssal Ocean. *Science*, 264(5162):1120–1123.

UNESCO (1981). Tenth Report of the Joint Panel on Oceanographic Tables and Standards. *Unesco Technical Papers in Marine Science*, 36.

Vellinga, M. and Wood, R. (2008). Impacts of Thermohaline Circulation Shutdown in the Twenty-first Century. *Climatic Change*, 91(1):43–63.

Weaver, A., Eby, M., Wiebe, E., Bitz, C., Duffy, P., Ewen, T., Fanning, A., Holland, M., MacFadyen, A., Matthews, H., et al. (2001). The UVic Earth System Climate Model: Model Description, Climatology, and Applications to Past, Present and Future Climates. *Atmosphere-Ocean*, 39(4):361–428.

Wunsch, C. and Ferrari, R. (2004). Vertical Mixing, Energy, and the General Circulation of the Oceans. *Annu. Rev. Fluid Mech*, (36):281–314.

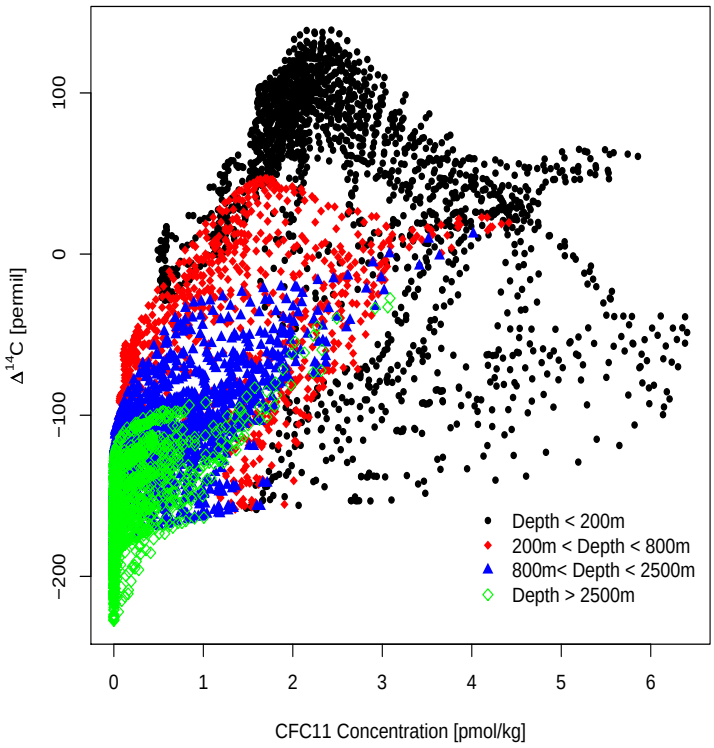

Figure 1: Relationship between Δ^{14}C and CFC11 model output for all K_v settings at different depths (less than 200 m, 200-800 m, 800-2500 m, and greater than 2500 m) for the model output

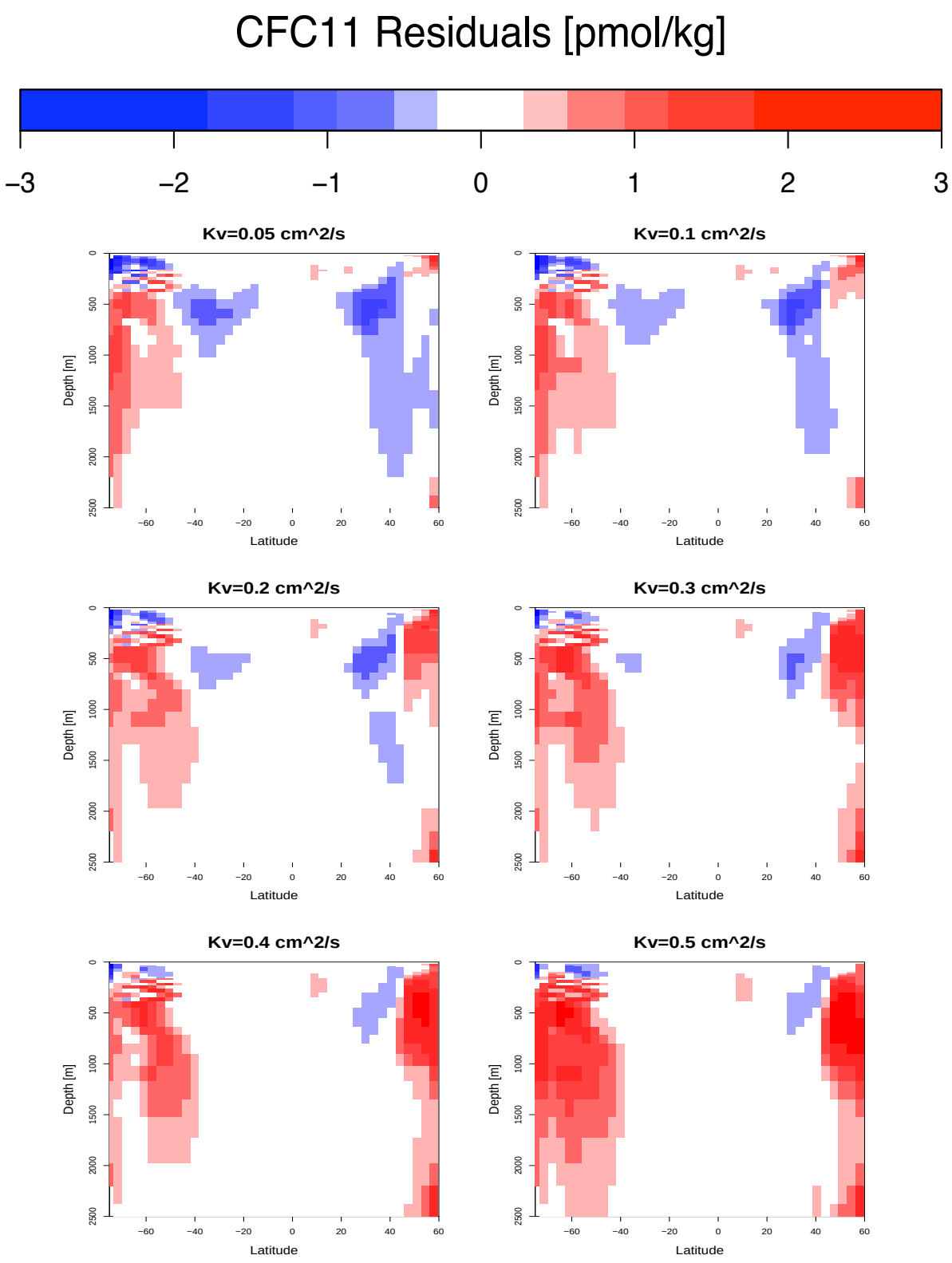

Figure 2: Difference between model output and observations (model output-observation) for the six values of K_v for CFC11. Top left: $K_v = 0.05$ cm^2/s, Top right: $K_v = 0.10$ cm^2/s , Middle left: $K_v = 0.20$ cm^2/s, Middle right: $K_v = 0.30$ cm^2/s, Bottom left: $K_v = 0.40$ cm^2/s, Bottom right: $K_v = 0.50$ cm^2/s.

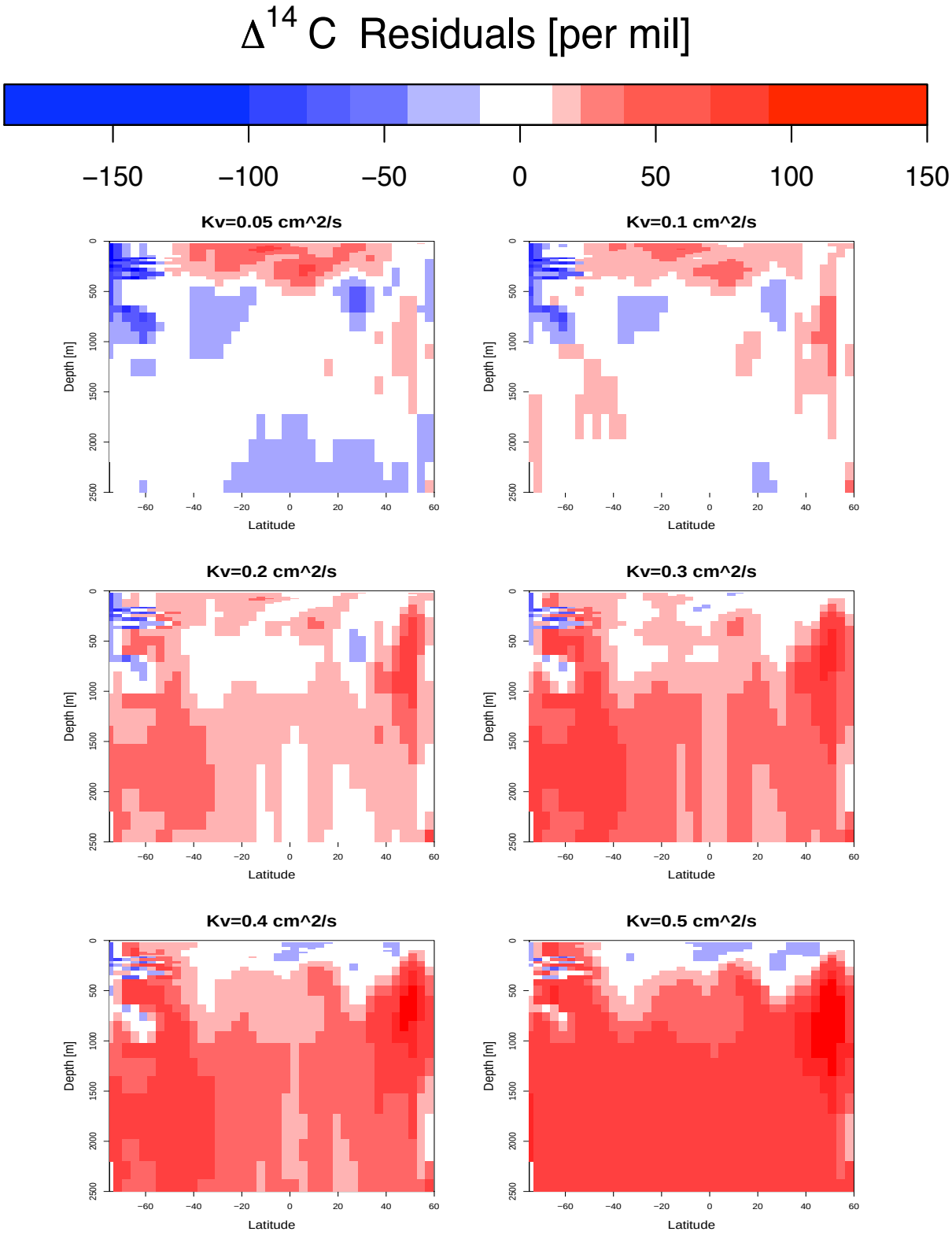

Figure 3: Difference between model output and observations (model output-observation) for the six values of K_v for Δ^{14}C. Top left: $K_v = 0.05$ cm^2/s, Top right: $K_v = 0.10$ cm^2/s , Middle left: $K_v = 0.20$ cm^2/s, Middle right: $K_v = 0.30$ cm^2/s, Bottom left: $K_v = 0.40$ cm^2/s, Bottom right: $K_v = 0.50$ cm^2/s.

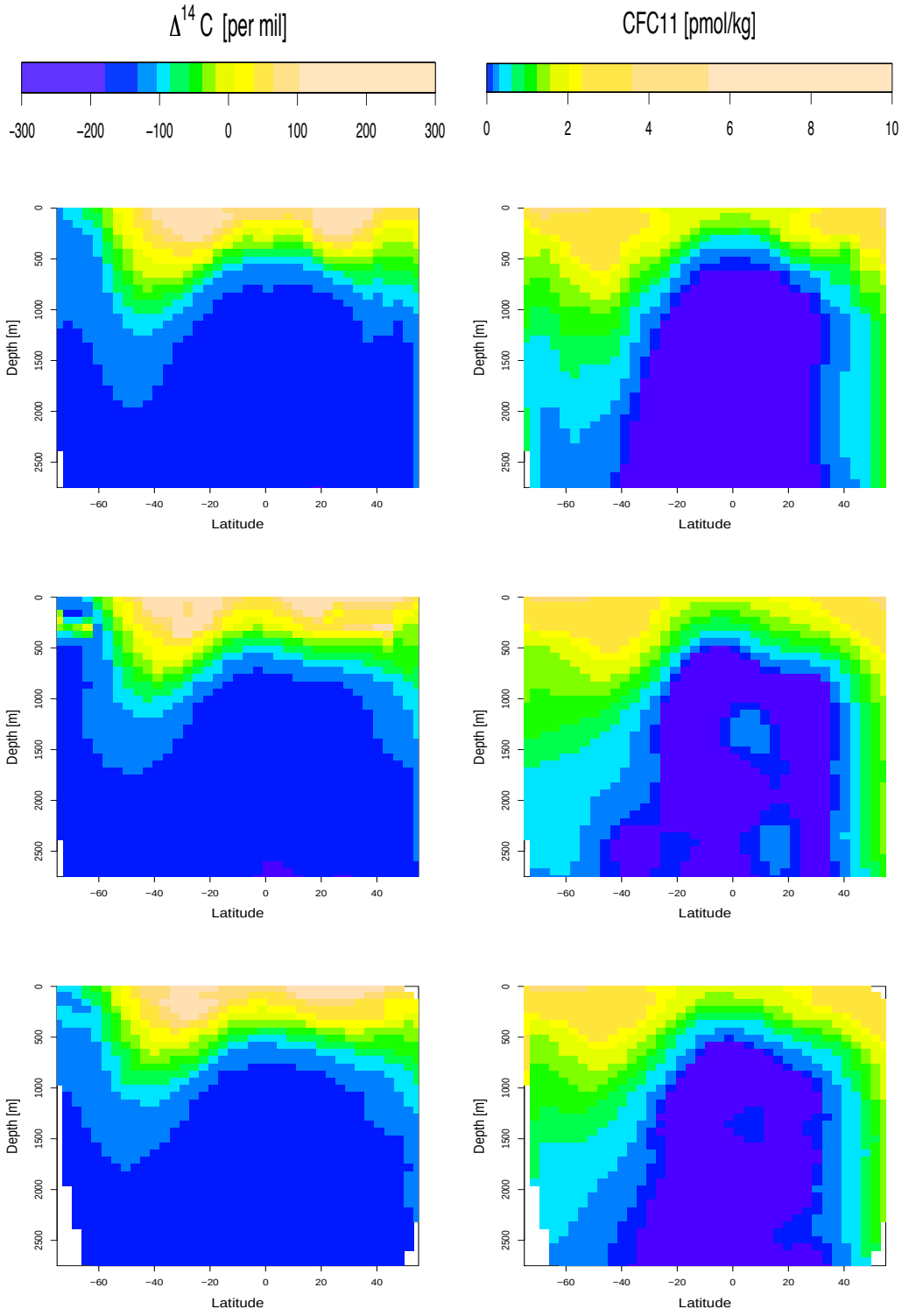

Figure 4: Plots of model output at $K_v = 0.2$ cm^2/s (top row), Cross-validation plots of predictions at K_v=0.2 cm^2/s with model output at K_v held out (middle row), Cross-validation plots of predictions at K_v=0.2 cm^2/s with one-ninth of spatial locations held out (bottom row). Left: Δ^{14}C, Right: CFC11.

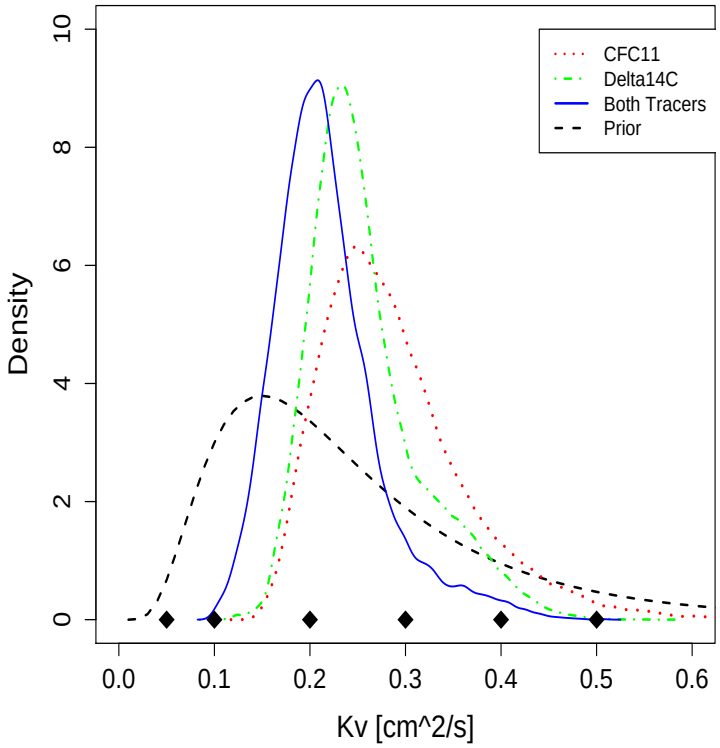

Figure 5: Log Normal Prior (dashed line) and posterior distributions of vertical diffusivity (K_v) using (i) CFC11 tracer (dotted line), (ii) Δ^{14}C tracer (dotted-dashed line), (iii) CFC11 and Δ^{14}C tracers jointly (solid line).

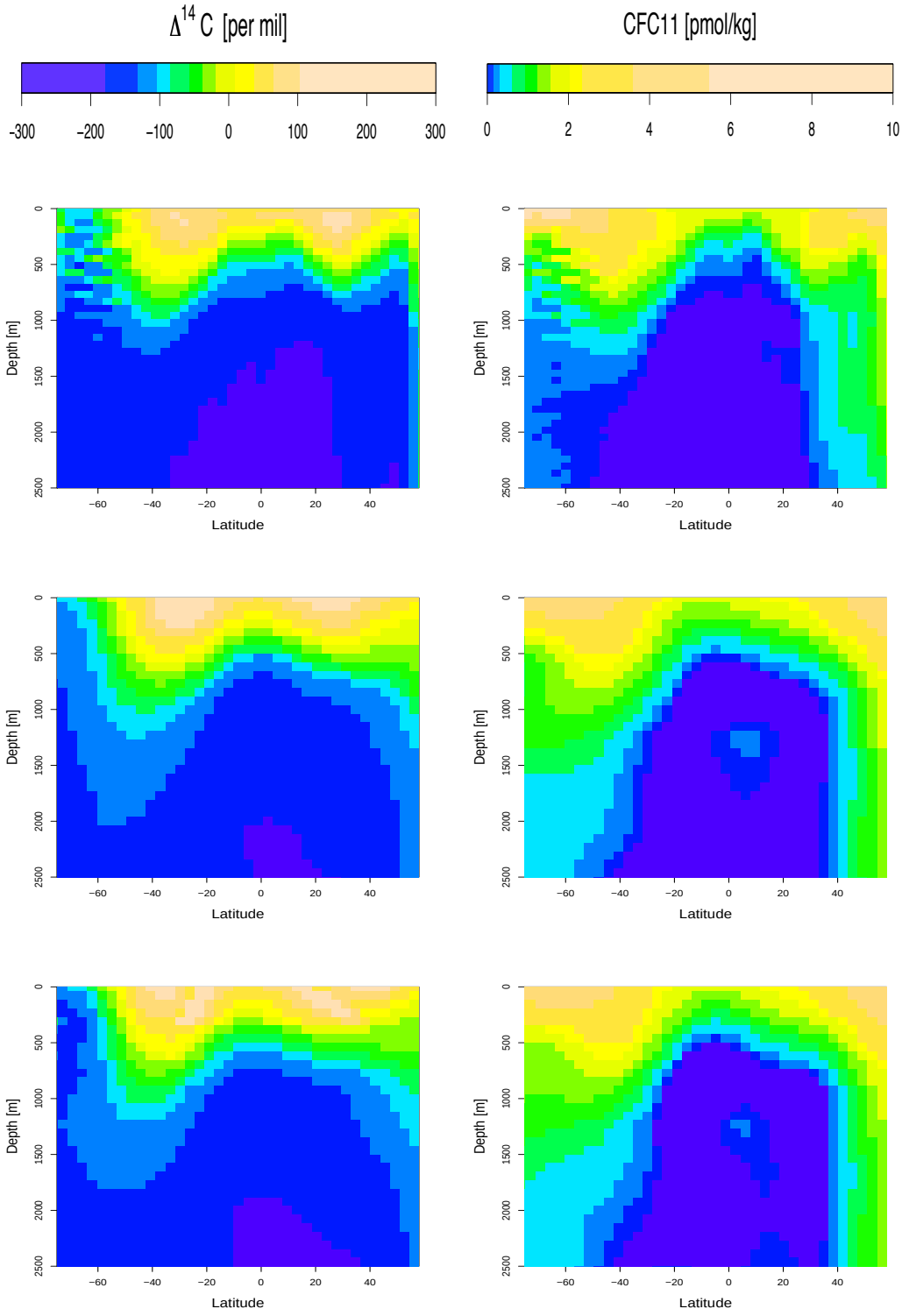

Figure 6: Plot of observations (Key et al., 2004) and the posterior mean predictions using our approach. Left: Δ^{14}C . Right: CFC11, Top : Observations, Middle: Univariate approach for single tracer, Bottom: Bivariate approach for both tracers.

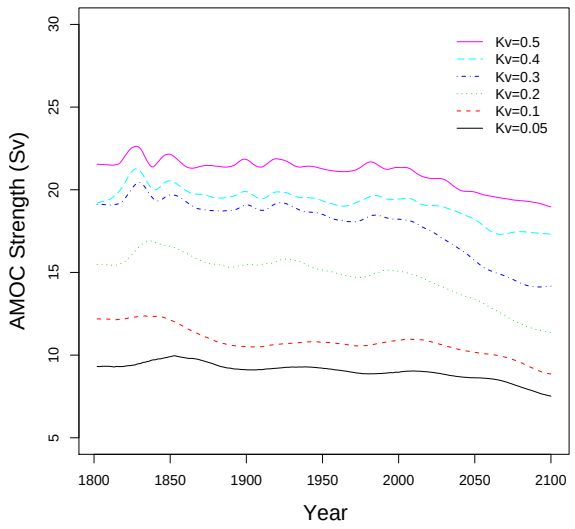

Figure 7: AMOC strength projections in Sv (Sverdrups) for the K_v values considered between 0.05 and 0.5.

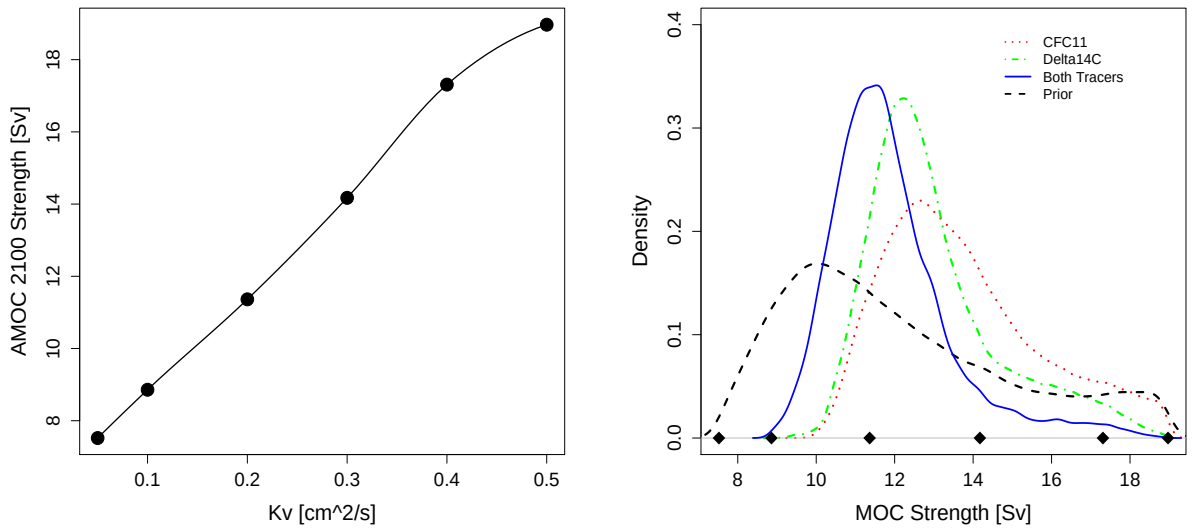

Figure 8: Interpolated relationship between projected AMOC in 2100 and K_v (part a), and the resulting AMOC projection in 2100 (part b) using prior, univariate and bivariate posteriors for K_v.