Submitted to the Annals of Applied Statistics

INFERRING LIKELIHOODS AND CLIMATE SYSTEM CHARACTERISTICS FROM CLIMATE MODELS AND MULTIPLE TRACERS *

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

The Pennsylvania State University

An important potential outcome of anthropogenic climate change is a possible collapse of the Atlantic meridional overturning circulation (AMOC). Assessing the risk of an AMOC collapse is of great importance since it may result in major temperature and precipitation changes and a shift in terrestrial ecosystems. One key source of uncertainty in AMOC predictions is uncertainty about background ocean vertical diffusivity (K_v), a key model parameter. K_v cannot be directly observed but can be inferred by combining climate model output with observed data on ocean tracers. In this work, we combine information from multiple tracers, each observed on a spatial grid. Our two stage approach emulates the 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. We utilize kernel mixing and matrix identities in our Gaussian process model in order to make computations tractable for our 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 make probabilistic projections of the AMOC.

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 acceptable strategies for climate change mitigation and carries considerable expected economic value of information (Keller et al., 2004). One key source of uncertainty in

*This work was supported by the NSF.

AMOC predictions is uncertainty about background ocean vertical diffusivity (K_v) (Dijkstra, 2008; Schmittner et al., 2009), a key model parameter.

K_v is a model parameter which quantifies the intensity of vertical mixing in the ocean, specifically the speed of transfer (e.g. heat transfer) from the surface to the deep ocean. K_v parameterizes a range of processes occurring below the model resolution. K_v is hence model dependent and cannot be directly observed in the ocean (Schmittner et al., 2009). 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 often a large amount of information about K_v in the form of oceanic tracers, which are observations that provide information about ocean patterns. In the climate model, these tracers are strongly affected by the value of K_v. We can therefore use observations of these tracers to infer K_v. For example, larger observed values of Δ^{14}C 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). 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. Earth system models 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). 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. Here we analyze a previously published ensemble of runs from an Earth System Model of Intermediate Complexity (EMIC). Specifically, we analyze an ensemble of the University of Victoria (UVic) Earth System Climate Model described in Schmittner et al. (2009). 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.

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), and Forest et al. (2008), 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. (2007) 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 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. We use kernel mixing (Higdon, 1998) and matrix rank reduction identities (c.f. Cressie and Johannesson, 2008; Stein, 2008) to enable the Gaussian process modeling of large 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 reduce parameter identifiability issues. Our approach improves on previous work in three main ways by (i) enabling the analysis of larger data sets, (ii) allowing for a flexible model to combine information from multiple tracers, and (iii) reducing identifiability problems.

The rest of our 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 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. 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 ques-

tionable (Rougier, 2008a). By modeling the emulator separately, we obtain improved diagnostics regarding the accuracy of the emulator. This approach has much in common with Bayarri et al. (2007). 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:

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

where we assume a linear mean function, $\mu_{\boldsymbol{\beta}}(\theta) = \mathbf{X}\boldsymbol{\beta}$, with $\mathbf{X}$ 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. $\boldsymbol{\xi}$ is a vector of covariance parameters that specify the covariance matrix $\Sigma(\boldsymbol{\xi})$ (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 $(\boldsymbol{\xi}, \boldsymbol{\beta})$ be $(\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\beta}})$. 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 $(\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\beta}})$ in place of $(\boldsymbol{\xi}, \boldsymbol{\beta})$ 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.* In order 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

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

where $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ is as described in Section 2.1, $\boldsymbol{\epsilon} \sim N(0, \psi I)$, where $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_N)^T$ is the observation error with $\psi > 0$ as the observation error variance. 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* value of θ. We can now perform inference on θ and ψ, by specifying a prior for these parameters. Using Markov Chain Monte Carlo (MCMC), we can estimate a posterior distribution for θ. We will discuss prior selection for θ and ψ in Section 4. It should be noted 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 reduce the complexity and nonstationarity of the covariance (Royle and Berliner, 1999).

3.1. *Stage 1: Emulation for Multivariate Data.* The climate model output for the tracers are modeled using a hierarchical framework as described below in equation (1):

$$\begin{aligned} \mathbf{Y}_1 \mid \mathbf{Y}_2, \boldsymbol{\beta}_1, \boldsymbol{\xi}_1, \gamma &\sim N(\mu_{\boldsymbol{\beta}_1}(\theta) + \mathbf{B}(\gamma)\mathbf{Y}_2, \Sigma_{1.2}(\boldsymbol{\xi}_1)) \\ \mathbf{Y}_2 \mid \boldsymbol{\beta}_2, \boldsymbol{\xi}_2 &\sim N(\mu_{\boldsymbol{\beta}_2}(\theta), \Sigma_2(\boldsymbol{\xi}_2)) \end{aligned} \tag{1}$$

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 $\mu_{\boldsymbol{\beta}_i}(\theta) = \mathbf{X}\boldsymbol{\beta}_i$ (for i=1,2) where $\mathbf{X}$ is the covariate matrix of dimension $M \times b$, with covariates as specified in Section 2.1. $\mathbf{B}(\boldsymbol{\gamma})$ is an $M \times M$ matrix that specifies the relationship between $\mathbf{Y}_1$ and $\mathbf{Y}_2$, and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \cdots, \gamma_5)$. 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:

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

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}).

$\Sigma_{1.2}(\boldsymbol{\xi}_1)$ is the conditional covariance matrix of $\mathbf{Y}_1|\mathbf{Y}_2$, and $\Sigma_2(\boldsymbol{\xi}_2)$ is the marginal covariance matrix of $\mathbf{Y}_2$, both have dimensions $M \times M$. $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ are vectors of covariance parameters for $\mathbf{Y}_1 \mid \mathbf{Y}_2$, $\mathbf{Y}_2$ respectively. As in Section 2.1, we find MLEs for $\boldsymbol{\xi}_1, \boldsymbol{\beta}_1, \boldsymbol{\gamma}, \boldsymbol{\xi}_2, \boldsymbol{\beta}_2$ and obtain a predictive distribution at $\mathbf{S}$ by plugging in the MLEs and conditioning on $\mathbf{Y}_1, \mathbf{Y}_2$. We note that the log-likelihood for $\ell(\boldsymbol{\beta}_1, \boldsymbol{\xi}_1, \boldsymbol{\beta}_2, \boldsymbol{\xi}_2, \boldsymbol{\gamma} \mid \mathbf{Y}_1, \mathbf{Y}_2)$ can be broken up into the sum of two functions, $\ell_1(\boldsymbol{\beta}_1, \boldsymbol{\xi}_1, \boldsymbol{\gamma} \mid \mathbf{Y}_1, \mathbf{Y}_2)$ and $\ell_2(\boldsymbol{\beta}_2, \boldsymbol{\xi}_2 \mid \mathbf{Y}_2)$, 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. $\Sigma_{\mathbf{Y}_{1.2}}$ and $\Sigma_{\mathbf{Y}_2}$ are the covariance matrices for $\mathbf{Y}_1|\mathbf{Y}_2$ and $\mathbf{Y}_2$ using the MLE for $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ respectively. We can in a similar manner derive the predictive distribution $\boldsymbol{\eta}(\mathbf{Y}, \theta)$ from the conditional and marginal distributions described in the previous paragraph.

$$\text{(2)} \qquad \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{X}_Y \hat{\boldsymbol{\beta}}_1 + \mathbf{B}(\hat{\boldsymbol{\gamma}}) \mathbf{X}_Y \hat{\boldsymbol{\beta}}_2 \\ \mathbf{X}_Y \hat{\boldsymbol{\beta}}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{\mathbf{Y}_{1.2}} + \mathbf{B}(\hat{\boldsymbol{\gamma}}) \Sigma_{\mathbf{Y}_2} \mathbf{B}^T(\hat{\boldsymbol{\gamma}}) & \mathbf{B}(\hat{\boldsymbol{\gamma}}) \Sigma_{\mathbf{Y}_2} \\ \Sigma_{\mathbf{Y}_2} \mathbf{B}^T(\hat{\boldsymbol{\gamma}}) & \Sigma_{\mathbf{Y}_2} \end{bmatrix} \right)$$

3.2. *Stage 2: Inference for Multiple Tracers.* We model the observed data $\mathbf{Z} = (\mathbf{Z}_1 \ \mathbf{Z}_2)$ as follows:

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

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

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

where $\psi_1, \psi_2 > 0$ are the observation error variances for the two tracers. We perform inference on θ, ψ_1, and ψ_2 as discussed in Section 2.2 by writing the log-likelihood $\ell(\theta, \psi_1, \psi_2 \mid \mathbf{Z})$ as the sum of

two multivariate normal functions by using a standard result for partitioned matrices (Anderson, 2003). See Appendix B for details. This reduces the matrix computations from dimension $2N \times 2N$ to $N \times N$. Using the observations $\mathbf{Z}_1, \mathbf{Z}_2$, we estimate the posterior distributions of θ, ψ_1, and ψ_2 given $\mathbf{Z}_1, \mathbf{Z}_2$ 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 without computational tractability issues. However, for larger data sets such as the ocean tracer data set that we study here, computing becomes infeasible. 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 (c.f. Golub and Van Loan (1996)); used in spatial modeling contexts by Cressie and Johannesson (2008) and Stein (2008)).

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 observation 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

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

Usually 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 $\mu(\mathbf{s})$:

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

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. *Applying Kernel Mixing to Our Approach.* In this section, we describe how we apply kernel mixing 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 write 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:

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

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

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

with $\boldsymbol{\phi} = (\phi_s, \phi_d, \phi_c)$, $\kappa, \phi_s, \phi_d, \phi_c > 0$. The kernel is separable over location, depth, and climate parameters, although a nonseparable kernel could be chosen if appropriate. Note that this function can be easily adapted to models for multiple climate parameters.

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 $\mathbf{K}_{\mathbf{Y}_i}(\boldsymbol{\phi}_i, \kappa_i)$ is defined in equation (3). Hence, the model is as described below, where $\mathbf{X}_Y$ is a covariate matrix of dimension $M \times b$ for the model output:

$$\mathbf{Y}_i \sim N(\mathbf{X}_Y \boldsymbol{\beta}_i, \zeta_i I_N + \mathbf{K}_{\mathbf{Y}_i} \mathbf{K}_{\mathbf{Y}_i}^T)$$

where $\mathbf{K}_{\mathbf{Y}_i}$ is a $N \times J$ dimensional kernel matrix. We estimate MLEs of the regression and covariance parameters $\boldsymbol{\beta}_i, \boldsymbol{\xi}_i$ respectively. To do so, we must invert an $M \times M$ matrix, which is computationally expensive since $M = 5926$ in our dataset. However, by using kernel mixing to obtain a covariance matrix with a specific structure, $\Sigma_{\mathbf{Y}_i} = \zeta_i I + \mathbf{K}_{\mathbf{Y}_i} \mathbf{K}_{\mathbf{Y}_i}^T$, where $\zeta_i > 0$, and rewriting the inverse of this matrix by using the Sherman-Woodbury-Morrison identity (see Appendix A), the matrix inversions can be reduced 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. Also it can be seen as accounting for microscale variation (c.f. 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 $\mathbf{Y}_1, \mathbf{Y}_2$: $\zeta_1, \kappa_1, \boldsymbol{\phi}_1, \boldsymbol{\gamma}, \boldsymbol{\beta}_1, \zeta_2, \kappa_2, \boldsymbol{\phi}_2, \boldsymbol{\beta}_2$.

If 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 (4) for the covariance matrices in equation (2). The inverses of these matrices can be rewritten in a manner that requires inversions

only of J by J matrices (see Appendix A).

$$\Sigma_{\mathbf{Y}_{1.2}} = \zeta_1 I + \mathbf{K}_{\mathbf{Y}_1}\mathbf{K}_{\mathbf{Y}_1}^T \quad \Sigma_{\mathbf{Y}_2} = \zeta_2 I + \mathbf{K}_{\mathbf{Y}_2}\mathbf{K}_{\mathbf{Y}_2}^T \tag{4}$$

When N is large, computation in the second stage becomes intractable, and we need to 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 distributions of the predictions, $\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta)$ and $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$. We write the distribution of $\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta)$ as follows:

$$\boldsymbol{\eta}_2(\mathbf{Y}_2, \theta) \sim N(\mu_2^*, \Sigma_2^*)$$

where the mean vector μ_2^* and covariance matrix Σ_2^* are written below in (5) and (6) using kriging equations and matrix simplification:

$$\mu_2^* = X_{\mathbf{Z}}(\theta)\boldsymbol{\beta}_2 + \mathbf{K}_{\mathbf{Z}_2}\mathbf{K}_{\mathbf{Y}_2}^T\Sigma_{\mathbf{Y}_2}^{-1}(\mathbf{Y}_2 - X_{\mathbf{Y}}\boldsymbol{\beta}_1) \tag{5}$$

$$\Sigma_2^* = \zeta_2\mathbf{I} + \mathbf{K}_{\mathbf{Z}_2}(\mathbf{I}_J - \mathbf{K}_{\mathbf{Y}_2}^T\Sigma_{\mathbf{Y}_2}^{-1}\mathbf{K}_{\mathbf{Y}_2})\mathbf{K}_{\mathbf{Z}_2}^T \tag{6}$$

We write the distribution of $\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta)$, in a similar manner in equation (7) with mean vector and covariance matrix in equations (8) and (9):

$$\boldsymbol{\eta}_1(\mathbf{Y}_1, \mathbf{Y}_2, \theta) \sim N(\mu_{1.2}^*, \Sigma_{1.2}^*) \tag{7}$$

$$\mu_{1.2}^* = X_{\mathbf{Z}}(\theta)\boldsymbol{\beta}_1 + \mathbf{K}_{\mathbf{Z}_1}\mathbf{K}_{\mathbf{Y}_1}^T\Sigma_{\mathbf{Y}_1}^{-1}(\mathbf{Y}_1 - X_{\mathbf{Y}}\boldsymbol{\beta}_1) \tag{8}$$

$$\Sigma_{1.2}^* = \zeta_1\mathbf{I} + \mathbf{K}_{\mathbf{Z}_1}(\mathbf{I}_J - \mathbf{K}_{\mathbf{Y}_1}^T\Sigma_{\mathbf{Y}_1}^{-1}\mathbf{K}_{\mathbf{Y}_1})\mathbf{K}_{\mathbf{Z}_1}^T \tag{9}$$

where $\mathbf{K}_{\mathbf{Z}_1}$ and $\mathbf{K}_{\mathbf{Z}_2}$ are the kernel matrices for the locations of the observations $\mathbf{S}$ and parameters $\boldsymbol{\beta}_1, \boldsymbol{\xi}_1$ and $\boldsymbol{\beta}_2, \boldsymbol{\xi}_2$ substituted respectively. $X_{\mathbf{Z}}(\theta)$ is a covariate matrix of dimension $N \times b$ with $\mathbf{S}$ and climate parameter θ as covariates. In addition, the covariance matrices $\Sigma_{1.2}^*$ and Σ_2^* can be rewritten in a manner that requires matrix inversions only of J by 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:

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

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

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

where $\psi_1, \psi_2 > 0$ are the observation error variances for the two tracers. We perform inference on θ, ψ_1, and ψ_2 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):

$$\mu_{1.2} = \mu^*_{1.2},\ \mu_2 = \mu^*_2,\ \Sigma_{1.2} = \psi_1 \mathbf{I}_N + \Sigma^*_{1.2},\ \Sigma_2 = \psi_2 \mathbf{I}_N + \Sigma^*_2$$

Matrix inversions can be reduced to dimension $J \times J$ (see Appendix A). Using the observations $\mathbf{Z}_1, \mathbf{Z}_2$, we estimate the posterior distributions of θ, ψ_1, and ψ_2 given $\mathbf{Z}_1, \mathbf{Z}_2$ 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 are 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 compare and constrain climate models (England, 1995), and K_v in particular (Schmittner et al., 2009).

Δ^{14}C and CFC11 measurements were collected for all oceanic basins in the 1990s, with locations denoted by a latitude, longitude, and depth. The data have been controlled for quality and gridded by Key et al. (2004). We use the observations from the data synthesis project by Key et al. (2004). We then aggregate the observations zonally (i.e. aggregated over longitudes), resulting in a data set with $N = 3,706$ locations of latitude and depth. In addition, 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 were evaluated from University of Victoria (UVic) Earth System Climate Model as described in Schmittner et al. (2009). The model output were also aggregated zonally, 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. Observations from latitudes above 60° N and depths below 3000 m are excluded 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 and the locations do not coincide. 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.

The model runs are set up in a manner such that ^{14}C model output can be compared to the observations of Δ^{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

based on work of Bryden (1973) and Fofonoff (1977) and the seawater density using 1980 UNESCO International Equation of State (IES80) (UNESCO, 1981). In addition, for density and temperature calculations, the model ocean pressure field was obtained from depth and latitude using simplified equations (Lovett, 1978).

Next we discuss some of the details of the application of our approach to the ocean tracer data. Due to the large amount of data, we use the 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 in two ways, 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. These knots were found to produce a reasonable model based on cross-validation. Also in preliminary analyses using more knots (say 360), we found virtually no difference in results. We verify the emulator using a leave one out cross-validation approach, where we leave out the model output for one climate parameter (Rougier, 2008b) and predicted at all locations for that climate parameter setting. We also used a spatial cross-validation approach, where we removed approximately one-ninth of the locations, evenly across space, for all climate parameter settings, and used the emulator to predict at those locations. Both approaches of cross-validation appear to result in predictions for the removed locations similar to the original model output.

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) and a Uniform (0, 0.60). The Lognormal prior reflects the geoscientists' (KK and RT) prior uncertainty about K_v based on previous research. We use an inverse gamma prior for the observation error variances, specifically $\psi_1 \sim IG(2, 100)$ for Δ^{14}C and $\psi_2 \sim IG(2, 5)$ for CFC11. These priors were obtained after an exploratory analysis of the data suggested the approximate scale of these parameters.

5. 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 2). 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 0.206 and 0.332 cm^2/s, the 90% credible region for K_v using the single tracer Δ^{14}C is between 0.153 and 0.256 cm^2/s, and the 90% credible region for K_v using the tracers jointly is between 0.183 and 0.260 cm^2/s. 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 bivariate approach appear to be more similar to the original observations than the univariate approach, although we obtain reasonable predictions from both approaches (Figure 3). 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 (c.f. Dijkstra, 2008; Forest et al., 2002). This reduced uncertainty about K_v results, for example, in reduced uncertainty about AMOC strength projections. As is apparent from our results, there is a substantial reduction in the uncertainty of AMOC strength 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 4.5 (Figure 5(b)). We note that AMOC strength increases proportionally with K_v (Figure 5(a)); AMOC strength is dependent on the transfer of cold, dense water from the surface to the deep ocean. Stronger intensity of vertical mixing results in the faster movement of dense water to the deep ocean and a stronger AMOC (Bryden et al., 2005).

To ensure convergence of our MCMC based estimates in the second stage, we obtained Monte Carlo standard errors for the posterior mean estimates of θ, ψ_1, and ψ_2 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 ψ_1 and ψ_2 were less than 10^{-3} and 10^{-4}, respectively, 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 half a day (12 hours) and about 5 days (65 hours) for the univariate and the bivariate approaches, respectively. To obtain the posterior density of K_v using the Uniform prior, we used importance resampling (c.f. 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.

6. Discussion. We develop and apply a novel approach for inferring climate parameters by combining information from observations and climate model output. Our methodology allows for a flexible model for relating multiple tracers and accounts for spatial dependence among the observations. To improve the computational tractability for large data sets, we model dependence using special covariance structures and take advantage of matrix identities that allow for fast matrix

computations.

Our approach offers several attractive advantages over current approaches. These advantages include improved computational efficiency for large data sets, a richer modeling approach to express the relationship between multiple tracers, and the potential of a reduction in identifiability issues. We obtain computational gains by splitting the model into two stages; resulting in covariance matrices of dimension $M \times M$ and $N \times N$ rather than a single covariance matrix of dimension $(N + M) \times (N + M)$. By using kernel mixing and a specific structure for the covariance matrix, our approach can be easily extended to both larger spatial datasets and to spatiotemporal data. Our hierarchical approach allows us to model the relationship between tracers as a piecewise linear mean function in the mean function, rather than as a single correlation in the covariance function. Further advantages of this modeling approach include computational gains due to lower dimension covariance matrices. We can, in principle, extend our approach to modeling hierarchies of three or more tracers, but this may prove challenging as the number of tracers gets large, especially if there is relatively little domain expertise to guide the development of the hierarchy.

Note that the kernel mixing approach has the advantage of additional flexibility, for example, the ability to allow the process to be non-stationary (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. It may be of interest to investigate other approaches for constructing computable nonstationary models for large data sets (c.f. Cressie and Johannesson, 2008; Fuentes, 2007; Gramacy and Lee, 2008; Jun and Stein, 2008) in the context of inference for climate parameters.

Previous work has shown that attempting to separate observation error and model error can impose nontrivial computation and conceptual problems (Kennedy and O'Hagan, 2001; Sansò et al., 2008). Sansò et al. (2008), Keller and McInerney (2008), and Ricciuto et al. (2008) combined observation error and model error into a single term, rather than estimate them separately. Even this approach requires substantial compromises in computing techniques in order to fit the model. A further difficulty in including model error in our statistical model is the high dependence between the climate parameter and model error term (Liu et al., 2009). While our goal here has been inference for the climate parameters, we may be interested in learning more about the individual error terms in the future. With enough data, our two stage approach may allow us to separately identify information about both observation error and model error.

APPENDIX A: COMPUTATIONAL DETAILS

The Sherman-Woodbury-Morrison identity states that the inverse of a matrix of the form $\boldsymbol{A} + \mathbf{UCV}$, where $\boldsymbol{A}$ is of dimension $N \times N$, $\mathbf{U}$ is dimension $N \times J$, $\mathbf{V}$ is dimension $J \times N$, and $\mathbf{C}$ is

dimension $J \times J$ can be written as

$$(\boldsymbol{A} + \mathbf{UCV})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\boldsymbol{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\boldsymbol{A}^{-1}$$

In addition the determinant of a matrix $\boldsymbol{A} + \mathbf{UCV}$ can be written as

$$|\boldsymbol{A} + \mathbf{UCV}| = |\mathbf{C}^{-1} + \mathbf{V}\boldsymbol{A}^{-1}\mathbf{U}| \times |\mathbf{C}| \times |\boldsymbol{A}|$$

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

The matrix form $(\psi I_N + \mathbf{K}\mathbf{K}^T)$ 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 (10), which only require computations of matrices of dimension $J \times J$.

$$\begin{aligned}(\psi I_N + \mathbf{K}(I_J)^{-1}\mathbf{K}^T)^{-1} &= \frac{I_N}{\psi} - \frac{\mathbf{K}}{\psi}\left(I_J - \frac{\mathbf{K}^T\mathbf{K}}{\psi}\right)^{-1}\frac{\mathbf{K}^T}{\psi} \\ |\psi I_N + \mathbf{K}(I_J)^{-1}\mathbf{K}^T| &= \left|I_J - \frac{\mathbf{K}^T\mathbf{K}}{\psi}\right| \cdot \psi^N\end{aligned} \tag{10}$$

To compute the likelihoods in this paper, we compute the Cholesky decomposition of the matrix $\left(I_J - \frac{\mathbf{K}^T\mathbf{K}}{\psi}\right)$ 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:

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

Then denoting the mean vector and covariance matrix of $\mathbf{Z}_1 \mid \mathbf{Z}_2$ by $\mu_{11.2} = \mu_{\mathbf{Z}_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{Z}_2 - \mu_{\mathbf{Z}_2})$ 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:

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

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

ACKNOWLEDGEMENTS

We thank Andreas Schmittner for providing us the output of the published runs in Schmittner et al. (2009). The authors thank Rui Paulo, Andreas Schmittner, and Nathan Urban for helpful discussions. We acknowledge support from the National Science Foundation. Opinions, findings, and conclusions expressed in this work are those of the authors alone, and do not necessarily reflect the views of the NSF.

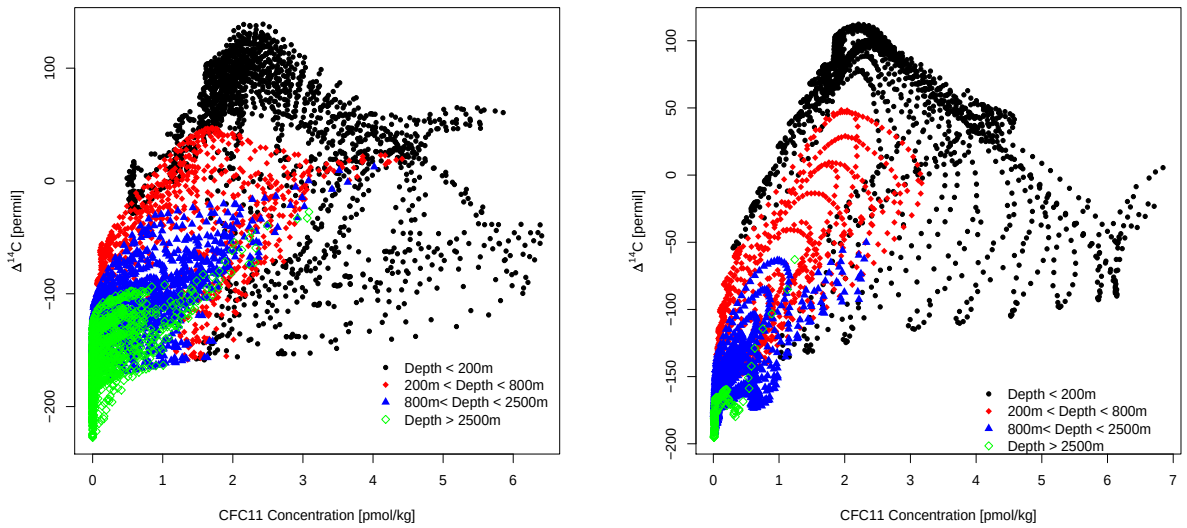

FIG 1. *Relationship between $\Delta^{14}C$ and CFC11 model output for all K_v settings at different depths (less than 200 m, 200-800 m, 800-2500 m, and greater than 2500 m) for model output (left), and observations (right).*

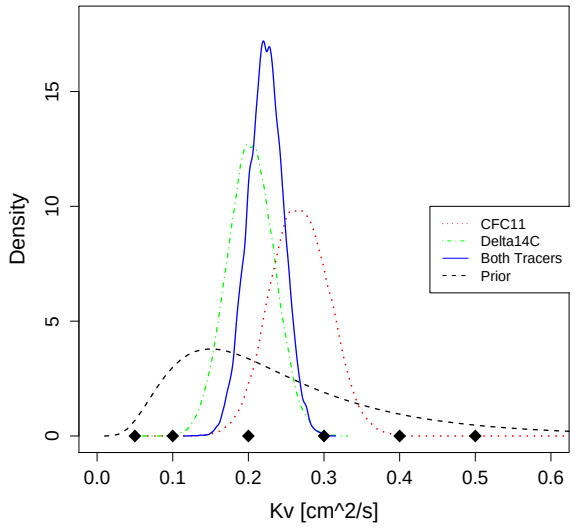

FIG 2. *Log Normal Prior (dotted black line) and posterior distributions of vertical diffusivity (K_v) using (i) CFC11 tracer (dotted red line), (ii) $\Delta^{14}C$ tracer (dotted-dashed green line), (iii) CFC11 and $\Delta^{14}C$ tracers jointly (solid blue line).*

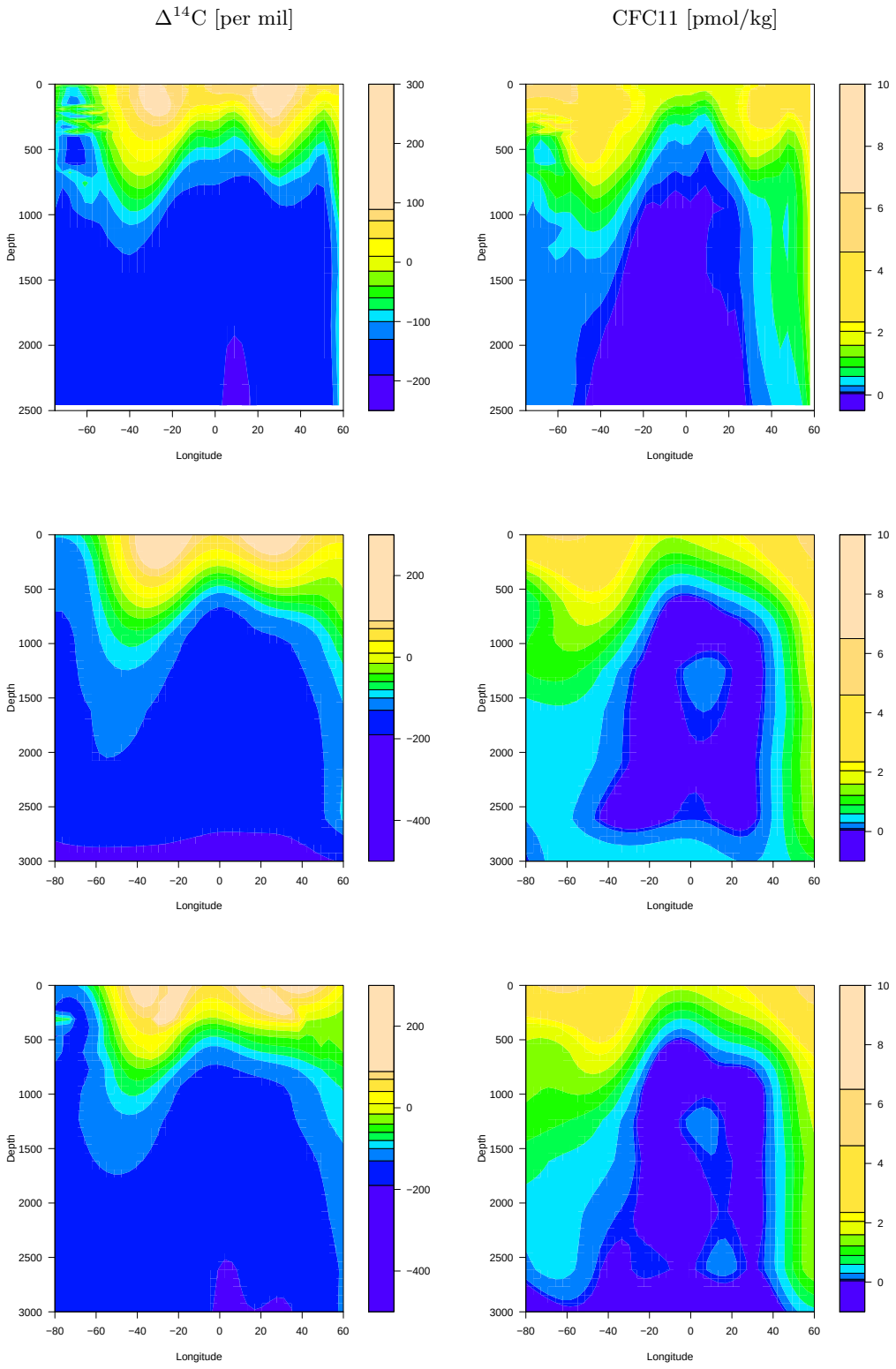

FIG 3. *Contour plot of observations (Key et al., 2004) and the posterior mean predictions using our approach. Left: $\Delta^{14}C$. Right: CFC11, Top : Observations, Middle: Univariate approach for single tracer, Bottom: Bivariate approach for both tracers.*

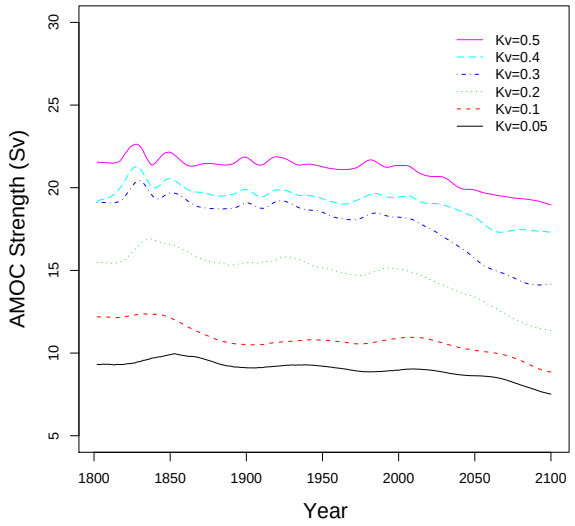

FIG 4. *AMOC strength projections in Sv (Sverdrups) for several K_v values between 0.05 and 0.5.*

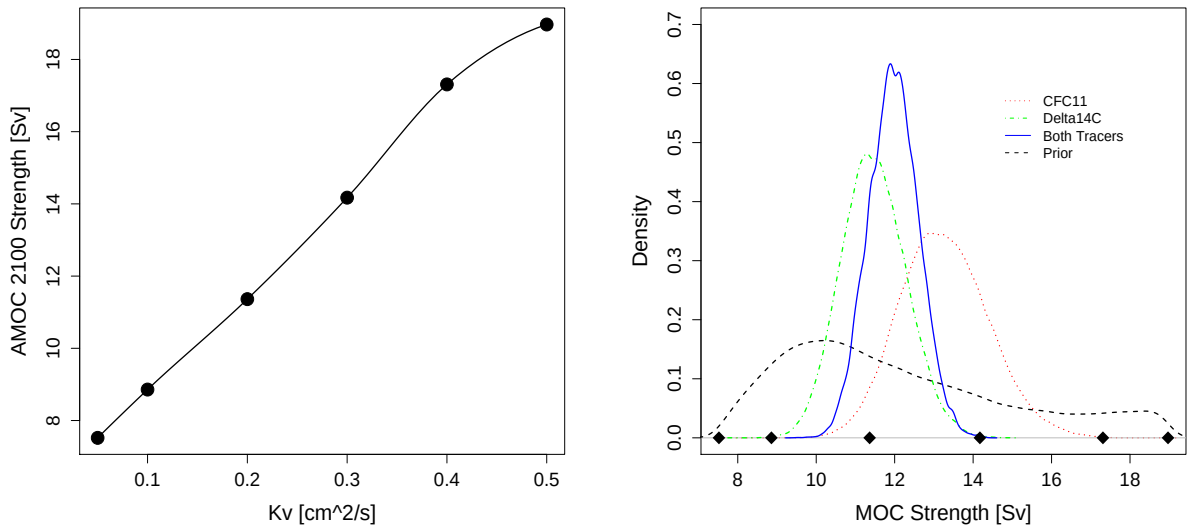

FIG 5. *Relationship between projected AMOC in 2100 and K_v (part a), and the resulting AMOC projection in 2100 (part b) using prior, univariate and bivariate posteriors for K_v.*

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., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C., et al. (2007). 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.

Conti, S. and OHagan, A. (2007). *Bayesian emulation of complex multi-output and dynamic computer models*. Dept. of Probability & Statistics, University of Sheffield.

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.

Dijkstra, H. (2008). Scaling of the Atlantic meridional overturning circulation in a global ocean model. *Tellus A*, 60(4):749–760.

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.

Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*, 102(477):321–331.

Fuentes, M. and Smith, R. (2001). A new class of nonstationary spatial models. *North Carolina State University Institute of Statistics Mimeo Series*, 2534.

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

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

Gramacy, R. and Lee, H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association*, 103(483):1119–1130.

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.

Jun, M. and Stein, M. (2008). Nonstationary covariance models for global data. *Ann. Appl. Statist*, 2:1271–1289.

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. and McInerney, D. (2008). The dynamics of learning about a climate threshold. *Climate Dynamics*, 30(2):321–332.

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).

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.

McCarthy, R., Bower, F., and Jesson, J. (1977). The fluorocarbon-ozone theoryI. Production and releaseworld 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. *Working 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.

Qian, P., Wu, H., and Wu, C. (2008). Gaussian Process Models for Computer Experiments With Qualitative and Quantitative Factors. *Technometrics*, 50(3):383–396.

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.

Ricciuto, D., Davis, K., and Keller, K. (2008). A Bayesian synthesis inversion of carbon cycle observations: How can

observations reduce uncertainties about future sinks. *Global Biogeochemical Cycles*, 22.

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*.

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.

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.

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.

Department of Statistics
Penn State University
University Park, PA 16802
E-mail: kgb130@psu.edu
E-mail: mharan@stat.psu.edu

Department of Geosciences
The Pennsylvania State University
University Park, PA 16802
E-mail: romantonk@psu.edu
E-mail: klaus@psu.edu