Joint Projections of North Pacific Sea Surface Temperature from Different Global Climate Models.

Francisco Beltrán[*], Bruno Sansó[*], Ricardo T. Lemos[†] and Roy Mendelssohn[†]

January 25, 2011

Abstract

The goal of this work is to develop a general methodology to obtain joint projections of climate indexes, based on ensembles of global climate model (GCM) output and historical records. As a case study, we consider Sea Surface Temperature (SST) in the North Pacific Ocean. We use two ensembles of 17 different GCM results, made available in the 4th Assessment Report of the Intergovernmental Panel on Climate Change: one corresponds to 20th century forcing conditions and the other corresponds to the A1B emissions scenario for the 21st century. Given a representation of the SST spatio-temporal fields based on a common set of empirical orthogonal functions (EOFs), we use a hierarchical Bayesian model for the EOF coefficients to estimate a baseline and a set of model discrepancies. These components are all time-varying. The model enables us to extract relevant temporal patterns of variability from both the observations and simulations and obtain common patterns from all eighteen series. This is used to obtain unified 21st century forecasts of relevant oceanic indexes as well as whole fields of forecast North Pacific SST. We compare the forecast index for different time scales and compare the SST reconstructions to the GCMs for the 21st century. While the coarser time resolution produces clear and faster results, we show that finer time scales produce results with structures that are similar to ones obtained at coarser scales.

1 Introduction

In its fourth assessment report (AR4), the Intergovernmental Panel on Climate Change (IPCC) made data available for 23 Global Climate Models (GCMs), under different emission scenarios. If we inspect the different datasets, we see a substantial disagreement in both hindcasts and future predictions, a fact that hampers the decision making process. The question that thus arises is how to blend the information from ensembles of climate model simulations, in such a way that we can obtain a better depiction of future climate, under the various emissions scenarios. Knutti et al. (2010) summarize the challenges involved in such a task. The simplest approach is to weight all models equally and take an ensemble average. This clearly disregards the fact that some models may be more accurate than others, or accuracy may differ at different time-scales. Alternatively, we can produce a weighted average, assigning weights to models depending on their agreement with observational records. This is known in the climate science literature as Reliability Ensemble

[*]Applied Mathematics and Statistics, University of California, 1156 High St. MS: SOE2, Santa Cruz, CA 95064

[†]NOAA/NMFS Environmental Research Division Southwest Fisheries Science Center 1352 Lighthouse Avenue Pacific Grove, CA 93950-2097

Average (REA; Giorgi and Mearns, 2002). Both approaches are usually applied to very large regions and large periods of time, on the premise that a clear signal can not be found at finer resolution. For example, in Tebaldi et al. (2005), Smith et al. (2008), Tebaldi and Sansó (2008), Smith et al. (2009), and Meehl et al. (2010), decadal temperature and/or precipitation averages are calculated.

A Bayesian hierarchical model (see, for example Gelman et al., 2004) approach is used to obtain unified forecasts, together with their predictive distributions. By assuming that all sources of data provide information about a common unobserved process, the Bayesian hierarchical approach produces a structured weighting scheme. The key assumption is that the differences between the observed climate and the one simulated by the GCM can be explained by an additive discrepancy with respect to a common, underlying, climate process.

An important question remains. Even for large spatial features, is it fair to compare climate model simulations for, say, a given year to the corresponding observational records? A simulation indexed by a given year is not meant to reproduce that year's observations. It is just a sample from the climate that is typical of that year, as estimated by the climate model . Averaging over large areas and time spans to analyze ensembles of climate model simulations is a way to compensate for the fact that such models are not meant to reproduce specific weather conditions that affect individual observations. Nevertheless, the ability to produce forecasts for coastal areas at a seasonal level is key to assess the impact of climate on marine ecosystems. Similarly, the population dynamics of many species can be affected by changes in the phase and amplitude of the seasonal cycles. These are examples of the need to have climate forecasts with high temporal and spatial resolution.

Given the above discussion, it is clear that extracting large scale spatial feature from a spatio-temporal field can be a useful approach to blend information from climate model simulations indexed in space and time. This can be conveniently achieved by representing the field using a set of basis functions. An example is the well known Kahunen-Loève (KL) expansion (see, for example, Yaglom, 1986). This provides a representation of a spatio-temporal random field that is determined by its covariance function. Consider the process $x_t(\boldsymbol{s})$, where t indicates time and $\boldsymbol{s}$ location. Suppose that the covariance function $v(\boldsymbol{s}, \boldsymbol{s}') = \mathrm{cov}(x_t(\boldsymbol{s}), x_t(\boldsymbol{s}'))$ does not depend on t. Then $x_t(s) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\psi_j(s)\alpha_j(t)$ for a set of orthogonal functions $\psi_1, \psi_2, \ldots$, random variables $\alpha_j(t)$ with $\mathrm{cov}(\alpha_j(t), \alpha_i(t)) = \delta_{ij}$, and non-negative λ_j that satisfy the integral equation $\int v(\boldsymbol{s}, \boldsymbol{s}')\psi_j(\boldsymbol{s}')d\boldsymbol{s}' = \lambda_j\psi_j(\boldsymbol{s})$. KL expansions are difficult to obtain in general. In a finite setting, though, they are obtained from the principle component (PC) analysis of the covariance matrix corresponding to the space-time process. The resulting $\psi_j(s)$ are known as the Empirical Orthogonal Functions (EOF) and are used to reduce a dataset's dimensionality by extracting the main modes of variability (Hannachi et al., 2007).

EOFs are very popular in environmental sciences. Unfortunately their estimation often ignores important trends as well as time-varying dependencies between observations at different locations. Additionally, EOF analysis suffers from a number of issues: it depends heavily on data availability; large eigenvalues are likely to be inflated; the absence or addition of a station may alter the eigenvalues, resulting in significant changes in the estimation of spatial and temporal variability; and if the data are non-stationary, the estimation of the covariance matrix becomes problematic. Nevertheless EOFs are simple to calculate and, being a discrete version of a KL expansion, they provide interpretable results. In the analysis of multi-model GCM simulations, using EOFs can filter out some of the noise and allow us to focus on the major patterns of variability. It is possible that there are commonalities between these, even when the different simulated fields are very dissimilar.

Moreover, EOFs have been used in the atmospheric and oceanic sciences to produce environmental indexes. Recent examples of EOF based indexes are the Pacific Decadal Oscillation (PDO; Mantua and Hare, 2002), the North Pacific Gyre Oscillation (NPGO; Di Lorenzo et al., 2008), and the Arctic Oscillation (AO; Thompson and Wallace, 1998). These indexes have been used in studying climate effects on salmon production (Mantua et al., 1997); when describing physical and biological changes in the North Pacific (Di Lorenzo et al., 2009) ; they have been linked to variations in the Kuroshio-Oyashio Extension (Di Lorenzo et al., 2008, 2009); and they have been associated with fish abundance in the San Francisco Bay (Cloern et al., 2010).

In this paper we develop a general methodology to obtain joint projections of climate indexes, using a multi-model ensemble of GCMs. Our approach begins with extracting global spatial features of both observations and simulations. We then model their temporal variability at a seasonal level in a way that allows for smooth temporal changes. We explore three levels of temporal resolution – decadal, seasonal and monthly – to gauge the possibility of blending multi-model ensembles at high temporal frequencies. We apply our methods to indexes that are related to the PDO and the NPGO. We assume that indexes obtained from observational data correspond to noisy versions of the processes of interest. Simulations from the GCMs also produce indexes that are noisy versions of the processes, but with the addition of discrepancy terms. We propose a model that assumes a smooth evolution in time for such discrepancies. We use a Bayesian hierarchical model to weight the simulations and obtain 21st century projections of the underlying process for each oceanic index. The resulting blended time-varying coefficients of the main modes of spatial variability are then used for the reconstruction of the temperature fields. Section 2 contains a description of the data, the simulations, and the methods to obtain the oceanic indexes. Section 3 has a description of the models proposed for the different time scales considered. Section 4 reports the results, and Section 5 discusses the conclusions of our analysis.

2 Data

2.1 Historical Records

We consider 5° gridded SST for the 20th century (1900-1999), in the North Pacific region 22.5°N-62.5°N, 112.5°E-247.5°E. This results in a total of 171 grid cells. Observational data stem from the UK Meteorological Office, Hadley Centre (http://badc.nerc.ac.uk/data/hadisst/; Rayner et al., 2003). We aggregate (to 5° resolution) the available 1° gridded monthly means, to make them compatible with calculations used for the PDO, and we also use three levels of time aggregation: Monthly (M, no aggregation), Quarterly (Q), and Monthly Decadal (D). We use these three levels to see how much information we can extract from each resolution. For the Q data we take the means of December–February, March–May, June–August, and September–November for every year. For D we average the ten months of January of each decade, then the ten months of February, and so on, for each month. Thus, in the M case, we have, for each grid cell, 18 time series with 2400 time steps. In the Q case we have 800 time steps and in the D case we have 240 time steps.

2.2 Global Climate Model data

We consider an ensemble of SST simulations from seventeen different GCMs (see Section A of the Appendix), used to obtain the results in the IPCC AR4. The GCM output, available from `https://esg.llnl.gov:8443/index.jsp`, stemmed from the World Climate Research Programme's

(WCRPs) Coupled Model Intercomparison Project (CMIP3) multi-model data set (Meehl et al., 2007). We obtained model simulations under two types of forcing: Climate of the 20th Century (20C3M), for the years 1900–1999, and Emissions Scenario A1B (Nakicenovic and Swart, 2000), for the years 2000–2099. Under the 20C3M scenario, greenhouse gas forcing is increased as observed in the 20th century. Under A1B, rapid global population and economic growth peak in the mid 21st century, and then start to decline. This scenario assumes that the technological change in energy systems will be balanced between fossil intensive (A1FI) and non-fossil energy sources (A1T). A1B forcing encompasses volcanic aerosols and emissions of sulfur, methane and other greenhouse gases (Randall et al., 2007). Because the spatial resolution of our ensemble of GCMs varies, we aggregate all simulations to a common 5° resolution.

2.3 Data processing

Let $X_t^0(\boldsymbol{s})$ denote the SST observation at time t ($t = 1, \ldots, \frac{T}{2}$) and grid point $\boldsymbol{s}$ ($\boldsymbol{s}_{i,j} = (i, j)$, $i = 1, \ldots, 9$, $j = 1, \ldots, 28$). The value of T corresponds to the end of the 21st century and depends on which level of time aggregation is used. To obtain monthly anomalies in the M dataset case, we consider a given location and average over all years for a given month. This produces a 20th century monthly climatology. For each grid cell, we then subtract the corresponding climatology from the observations. For datasets Q and D we proceed in a similar fashion. We denote the observational anomalies (SSTa) as $\hat{\mathbf{X}}_t^0(\boldsymbol{s})$. Similarly, we obtain the anomalies for each of the climate model simulations.

To obtain the EOF of the observational anomalies we consider the matrix

$$\hat{\mathbf{X}}^0 = \begin{pmatrix} \hat{X}_1^0(1,1) & \hat{X}_2^0(1,2) & \ldots & \hat{X}_{T/2}^0(1,28) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{X}_1^0(9,28) & \hat{X}_2^0(9,28) & \ldots & \hat{X}_{T/2}^0(9,28) \end{pmatrix}$$

and decompose it using the SVD algorithm (Hannachi et al., 2007). Thus

$$\hat{\boldsymbol{X}}^0 = UDV' = \sum_{l=1}^{k} \boldsymbol{u}_l(\boldsymbol{s})\boldsymbol{\psi}_l(t)$$

Letting $p = 9 \times 28$, $\boldsymbol{U}$ is a $p \times p$ orthogonal matrix of left singular vectors, $\boldsymbol{V}$ is a $\frac{T}{2} \times \frac{T}{2}$ orthogonal matrix of right singular vectors, and $\boldsymbol{D}$ is a $p \times T$ diagonal matrix with nonnegative singular values, sorted in decreasing order. $u_l(\boldsymbol{s})$ and $\psi_l(t)$ are $p \times 1$ and $1 \times T$ vectors and are the l-th EOF and the EOF coefficient respectively thus creating a discretized version of the Karhunen-Loéve representation of the spatial surface. As $u_l(\boldsymbol{s})$ form an orthogonal basis, we can obtain a common representation of all climate model simulation SST anomalies by projecting on it. Denote $\hat{\boldsymbol{X}}^j$ as the matrix of anomalies for the j-th climate model simulations. Then

$$\hat{\boldsymbol{X}}^j = UU^T\hat{\boldsymbol{X}}^j = \sum_{l=1}^{k} \boldsymbol{u}_l(\boldsymbol{s})\boldsymbol{\varphi}_{l,j}(t)$$

where $\varphi_{l,j}(t)$ is a $1 \times \frac{T}{2}$ vector and the coefficient for the l-th observational EOF, corresponding to the j-th member of the ensemble. In this paper we focus on the first two largest modes of variability, corresponding to $l = 1$ and 2, for each of the three time indexes. Figure 1 shows the first and second

EOFs of the observational anomalies, for time aggregations M, Q and D. The leading EOF for D is similar to the second EOF for M and Q, suggesting that at the Decadal time level the coastal process is driving the variability. We project the GCM anomalies onto the observational EOF, as opposed to calculating an individual EOF for each model. This is done because some of the GCM EOFs are completely different from the observational EOF. By projecting GCM anomalies onto the observational EOF, we keep the spatial pattern consistent throughout all models, thus giving us a way to compare the temporal disagreements between models.

As described in Mantua and Hare (2002), the PDO is the leading PC from an EOF analysis of North Pacific SST anomalies. To calculate the PDO index, SST anomalies poleward from 20°N are obtained by subtracting the long-term (1900-1993) mean from monthly observations, considering data from November to March only. The global mean anomaly is further subtracted, to remove the effects of "global warming". In our analysis the global mean was not removed. The PDO behaves much like the El Niño-Southern Oscillation but on the time scale of 20-30 years, as opposed to 16-18 months (Mantua and Hare, 2002). Similar to the PDO, the North Pacific Gyre Oscillation (NPGO) corresponds to the second leading PC of Sea Surface Height anomalies (SSHa). The NPGO closely tracks the second PC of SSTs which is referred to as the Victoria Mode (Di Lorenzo et al., 2008).

Figure 2 and 3 show the first and second EOF coefficients for some ensemble members. We note that, while the overall trend of coefficients corresponding to different climate models is similar, there are substantial variabilities between them. For the second half of the 21st century we observe a clear quasi-cyclical pattern for many of the climate model simulations. This is an indication that the simulations are out of phase with the 20th century observational climatology used to produce the anomalies. Monthly decadal series display noticeable jumps, due to the discontinuity that exists, for some decades, between the average SST for January and the average December SST for the previous decade.

3 Models to blend GCM and observational time series

As mentioned in the introduction, our goal is to blend the information from the different GCM simulations using the observations as a reference. To this end we build a hierarchical model that uses the EOF coefficients corresponding to present day observational data to estimate the model discrepancies. These are then propagated into the future to obtain 21st century forecasts.

To perform a comparison between models we use three methods including the Deviance Information Criterion (DIC; Spiegelhalter et al., 2002), Chi-Squared test (West and Harrison, 1997), and a Posterior Predictive P-value technique (PPP; De La Horra and Rodrguez-Bernal, 1999). Using these three tests we toggle different values of the discount factors as well as the seasonal components. We select the values for the discount factors for a specific EOF and time index based on the optimization of these three goodness of fit methods and a visual comparison.

3.1 Models for Monthly and Quarterly Averages

Consider the observational l-th EOF coefficient series $\psi_l(t)$. For simplicity we drop the index l and denote this as ψ_t. We assume that ψ_t follows an underlying level θ_t:

$$\psi_t = \theta_t + \nu_t^0, \tag{1}$$

for some Gaussian errors ν_t^0 uncorrelated in time. Additionally, the j-th model l-th EOF coefficient $\varphi_{l,j}(t)$ has a bias δ_t^j, with respect to θ_t. Again, we drop the index l and simplify the notation to

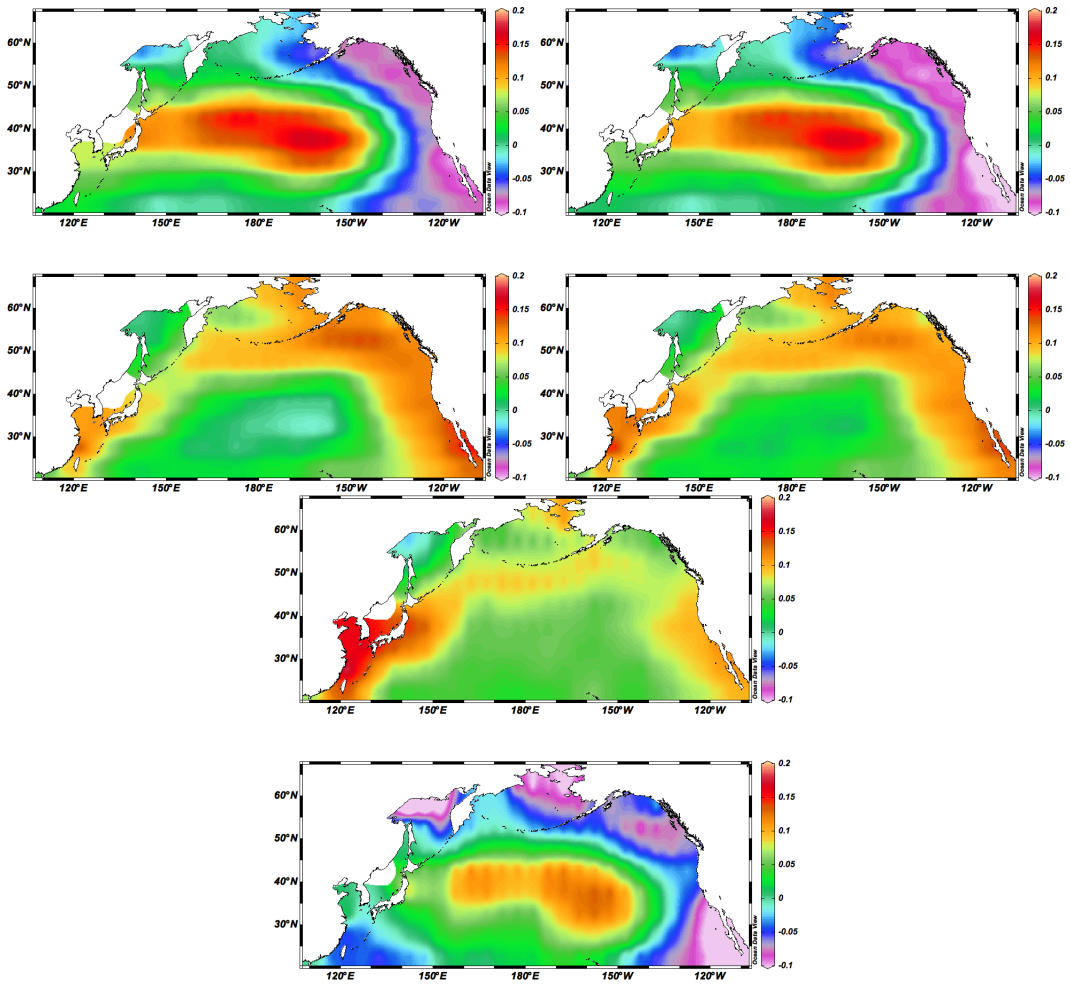

Figure 1: First and Second Monthly EOF(top left), Quarterly EOF(top right), Decadal Monthly EOF (bottom).

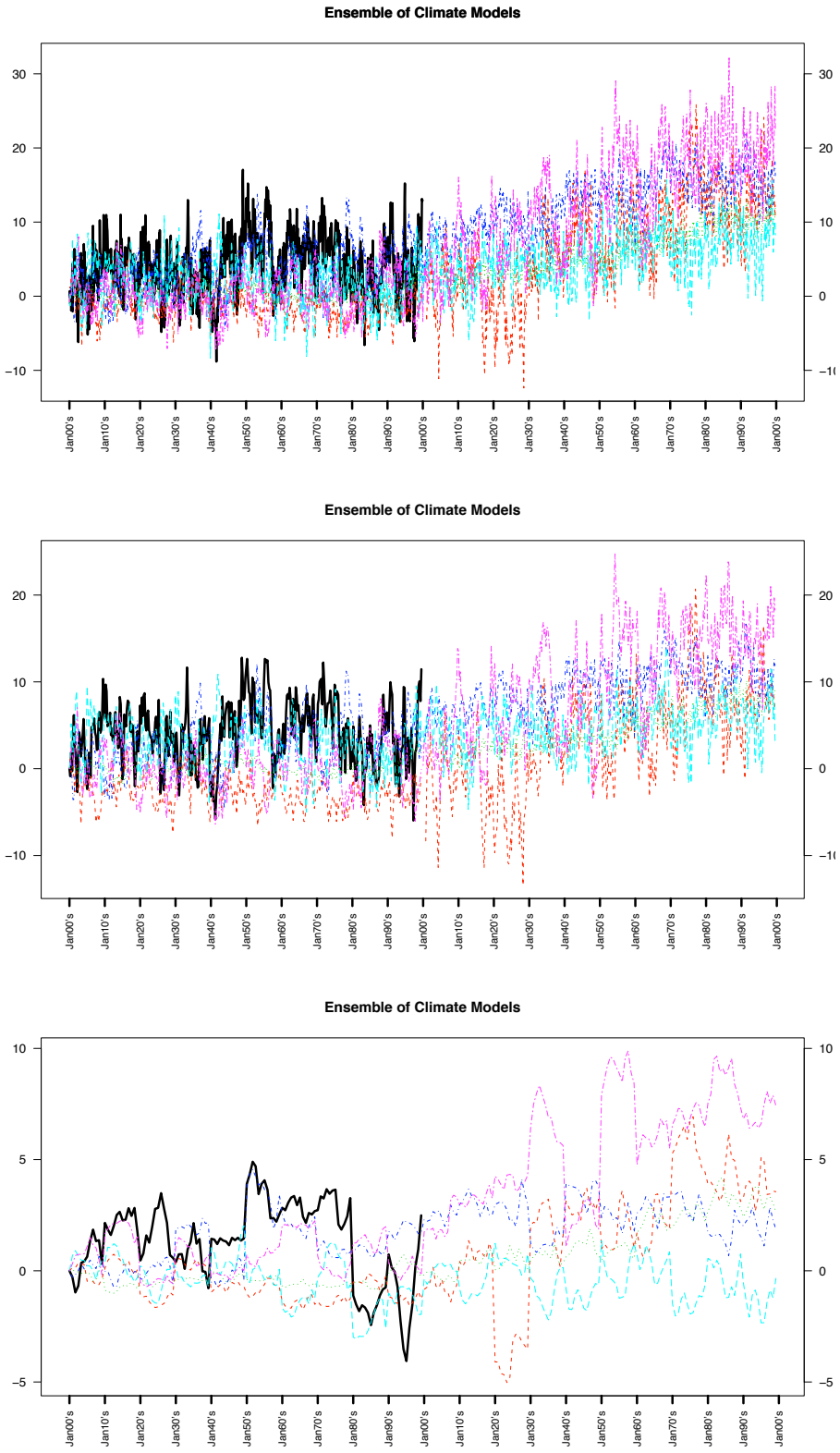

Figure 2: Monthly EOF1 coefficient (top), Quarterly EOF1 coefficient (middle), Decadal Monthly EOF2 coefficient (bottom) . SSTa (Solid Black) with the following model projections: GFDL-CM2.1, GISS-ER, CCSM3, NCAR-PCM, HadCM3.

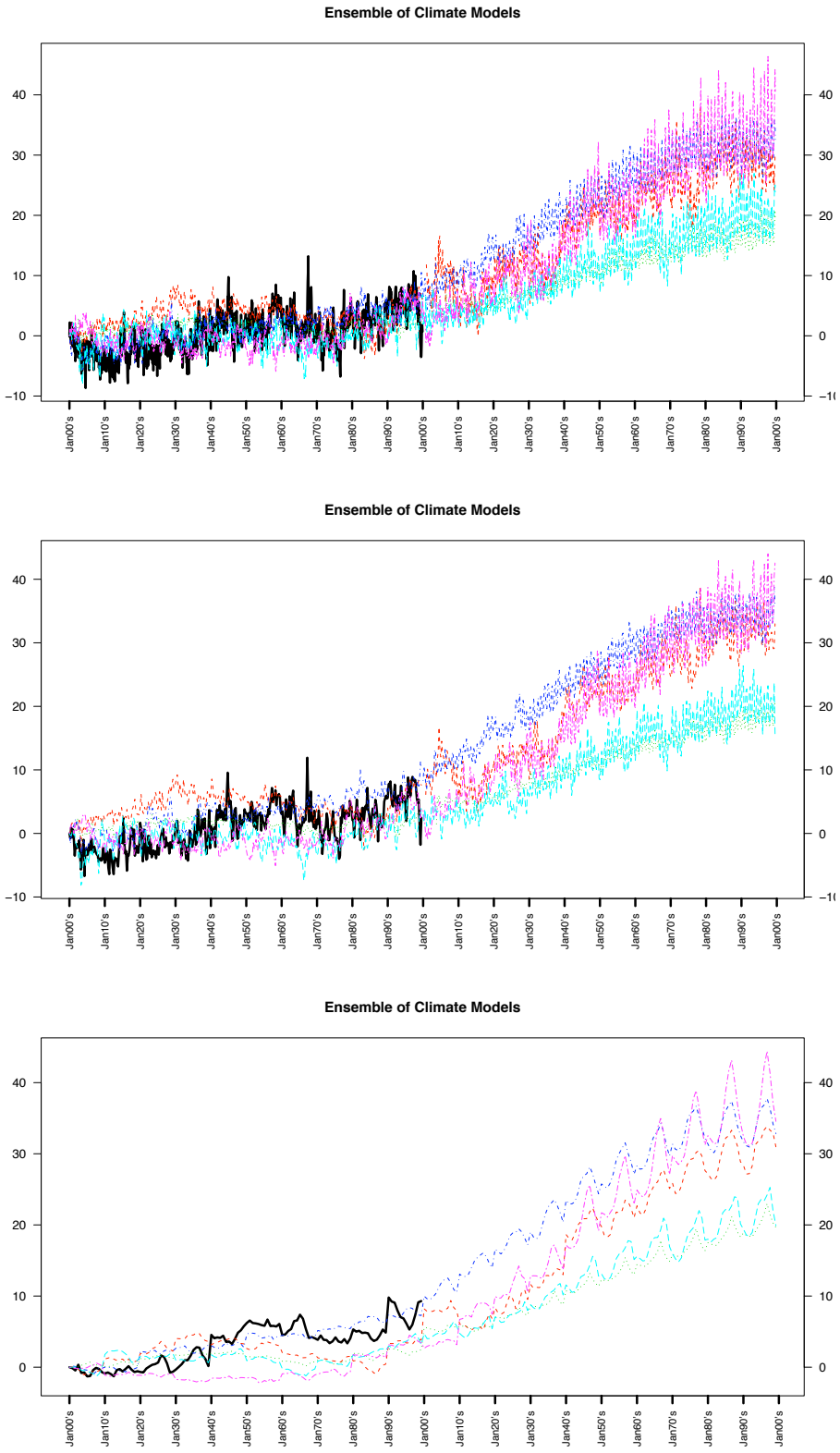

Figure 3: Monthly EOF2 coefficient (top), Quarterly EOF2 coefficient (middle), Decadal Monthly EOF1 coefficient (bottom) . SSTa (Solid Black) with the following model projections: GFDL-CM2.1, GISS-ER, CCSM3, NCAR-PCM, HadCM3.

φ_t^j. Thus

$$\varphi_t^j = \theta_t + \delta_t^j + \alpha_{1,t}^1 + \alpha_{1,t}^2 + \alpha_{1,t}^3 + \nu_t^j, \;\; j = 1, \ldots, 17, \tag{2}$$

where $\alpha_{i,t}^1$ corresponds to a time varying seasonal component having annual ($i = 1$), semestral ($i = 2$) and quarterly ($i = 3$) periods. The Gaussian errors ν_t^j are uncorrelated in time. While we are using anomalies which should eliminate the seasonality, the climate models create an apparent artificial oscillation, which, if not accounted for, will affect the models effectiveness. To complete the model we specify the evolution of the parameters as

$$\begin{array}{rcl} \theta_t &=& \theta_{t-1} + \beta_{t-1} + \omega_t^\theta \\ \beta_t &=& \beta_{t-1} + \omega_t^\beta \\ \delta_t^j &=& \delta_{t-1}^j + \omega_t^{\delta^j} \end{array}, \quad \begin{pmatrix} \alpha_{1,t}^i \\ \alpha_{2,t}^i \end{pmatrix} = \begin{pmatrix} \cos(2\pi\lambda_i) & \sin(2\pi\lambda_i) \\ -\sin(2\pi\lambda_i) & \cos(2\pi\lambda_i) \end{pmatrix} \begin{pmatrix} \alpha_{1,t-1}^i \\ \alpha_{2,t-1}^i \end{pmatrix}, \tag{3}$$

for $i = 1, 2, 3$ and $\lambda_i = i/12$. To specify the distribution of the error terms, denoted as $\boldsymbol{\nu}_t = (\nu_t^1, \ldots, \nu_t^{17})'$, we assume that $(\nu_t^0, \boldsymbol{\nu}_t')' \sim N_{18}(\mathbf{0}, \Sigma)$, for $t = 1, \ldots, \frac{T}{2}$, where $N_J(\cdot, \cdot)$ denotes a J-dimensional normal distribution. For the 21st century there are no observations and so no observational error term. We assume that $\boldsymbol{\nu}_t \sim N_{17}(\mathbf{0}, \Sigma_2)$, $\Sigma \sim W^{-1}(r_\Sigma, S_\Sigma)$, $\Sigma_2|\Sigma \sim W^{-1}(r_{\Sigma_2}, S_{\Sigma_2})$, where $W^{-1}(r, S)$ denotes an inverse-Wishart distribution with r degrees of freedom and scale matrix S. Conditioning the distribution of Σ_2 on Σ establishes a link between the observational variability in the 20th and the 21st centuries. We also assume that $S_{\Sigma_2} = (r_{\Sigma_2} - 17)\Sigma_{2:J,2:J}$, where $\Sigma_{2:J,2:J}$ is the partition of the Σ minus the first row and column. This implies that $E(\Sigma_2|\Sigma) = \Sigma_{2:J,2:J}$. A large value of r_{Σ_2} produces a distribution with small variability around the mean value $\Sigma_{2:J,2:J}$. Let $\boldsymbol{\Theta}_t = (\theta_t, \beta_t, \alpha_{1,t}^1, \ldots, \alpha_{2,t}^3, \delta_t^1, \ldots, \delta^{17})'$. Then, conditional on (Σ, Σ_2), we can write Equations (1) – (3) as a Dynamic Linear Model (DLM; see, for example West and Harrison, 1997) with $\boldsymbol{\Theta}$ as the state-space parameters. Denote as W_t the variance of the evolution equation of $\boldsymbol{\Theta}$. For the 20th century we model W_t using discount factors (West and Harrison, 1997, Chapter 6). We use three blocks, one for θ_t, one for β_t and one for the seasonal components. Each block has a different discount factor $d_i, i = 1, 2, 3$. We set $(d_1, d_2, d_3) = (0.70, 0.95, 0.999)$ and $(d_1, d_2, d_3) = (0.70, 0.95, 0.98)$ for the first and second EOF respectively. For the 21st century we do not discount and we fix W_t to its last values in the 20th century. This is to reflect the loss of observational data.

The Q data set is modeled similarly to the M data set, with two modifications: i) we only include two seasonal components, the annual and the semestral; ii) we place block discount factors on W_t, such that $(d_1, d_2, d_3) = (0.70, 0.95, 0.98)$ and $(d_1, d_2, d_3) = (0.70, 0.80, 0.999)$ for the first and second EOF, respectively.

3.2 Model for Decadal Monthly Averages

Similar to the previous model, we consider the observational EOF coefficient series ψ_t. We assume that ψ_t follows an underlying level θ_t and decadal jump λ_d. The decadal jump parameter corrects for the time discontinuity between decades, an artificial feature caused by the level of averaging. Thus

$$\psi_t = \theta_t + \lambda_d + \nu_t^0 \tag{4}$$

for some Gaussian errors ν_t^0 that have intra-decadal correlation. Additionally the j-th model EOF has a bias δ_t^j with respect to θ_t and decadal jump bias λ_d^j with respect to λ_d . We denote time index $t = 12(d-1) + m = d, m$, where $d = 1, \ldots, \hat{D} = 20$ are decades, $m = 1, \ldots, M = 12$ are the months, so $t = 1, \ldots, T = 240$. Thus,

$$\varphi_t^j = \theta_t + \delta_t^j + \lambda_d + \lambda_d^j + \alpha_{1,t}^1 + \alpha_{1,t}^2 + \alpha_{1,t}^3 + \nu_t^j, \;\; j = 1, \ldots, 17 \tag{5}$$

where $\alpha_{i,t}^1$ has time varying components as in model (3). ν_t^j are Gaussian errors with time correlation extending only within a time span of a decade. We describe the evolution of the parameters as

$$\begin{array}{rcl} \theta_t &=& \theta_{t-1} + \beta_{t-1} + \omega_t^\theta \\ \beta_t &=& \beta_{t-1} + \omega_t^\beta \\ \delta_t &=& \delta_{t-1} + \omega_t^{\delta j} \end{array}, \quad \begin{pmatrix} \alpha_{1,t}^i \\ \alpha_{2,t}^i \end{pmatrix} = \begin{pmatrix} \cos(2\pi\lambda_i) & \sin(2\pi\lambda_i) \\ -\sin(2\pi\lambda_i) & \cos(2\pi\lambda_i) \end{pmatrix} \begin{pmatrix} \alpha_{1,t-1}^i \\ \alpha_{2,t-1}^i \end{pmatrix} \tag{6}$$

for $i = 1, 2, 3$, $\lambda_i = \frac{i}{12}$. To specify the distributions of the error terms we denote $\boldsymbol{\nu_t} = (\nu_t^1, \ldots, \nu_t^{17})$, $\nu_d^j = (\nu_{1,d}^j, \ldots, \nu_{12,d}^j)$ and $\boldsymbol{\nu_d} = (\nu_d^1, \ldots, \nu_d^{17})$, we then assume $(\nu_t^0, \boldsymbol{\nu_t'}) = N_{18}(\mathbf{0}, \tau^2\Sigma)$ and $(\nu_d^0, \boldsymbol{\nu_d}) = N_{18}(\mathbf{0}, \Sigma \otimes \tau^2 I_{12})$, for $t = 1, \ldots, T/2$, $d = 1, \ldots, \hat{D}/2$ where $T/2$ and $\hat{D}/2$ correspond to the end of the 20th century . For the 21st century we lose the observational error term so we assume $\boldsymbol{\nu_t'} = N_{17}(\mathbf{0}, \tau^2\Sigma_2)$ and $\boldsymbol{\nu_d} = N_{17}(\mathbf{0}, \Sigma_2 \otimes \tau^2 I_{12})$. We place prior distributions $\Sigma \sim W^{-1}(r_\Sigma, S_\Sigma)$, $\Sigma_2|\Sigma \sim W^{-1}(r_{\Sigma_2}, S_{\Sigma_2})$ where as stated in the M time index model, S_{Σ_2} is centered at the partition of Σ. Let $\Theta_t = (\theta_t, \beta_t, \alpha_{1,t}^1, \ldots, \alpha_{2,t}^3, \delta_t^1, \ldots, \delta^{17})'$ and $\Lambda_d = (\lambda_d, \lambda_d^1, \ldots, \lambda_d^{17})'$. Conditional on $(\Lambda, \Sigma, \Sigma_2, \tau^2)$ we can write Equations (4) – (6) as a DLM, with state-space parameter Θ. We place block discount factors on the evolution equation similar to the M model, such that $(d_1, d_2, d_3) = (0.50, 0.98, 0.999)$ and $(d_1, d_2, d_3) = (0.50, 0.90, 0.999)$ for the first and second EOF respectively. For the 21st century we fix W_t to its last values of the 20th century to reflect on the loss of observational data. We place prior distributions for $\Lambda_d \sim N(m_{1,d}, C_{1,d})$, $\tau^2 \sim \Gamma^{-1}(r_\tau, S_\tau)$, $m_{1,d} \propto 1$, and $C_{1,d} \sim W^{-1}(220, 50I)$. To explore the posterior distributions of the models above we use the Gibbs sampler, a special case of Markov Chain Monte Carlo methods.

4 Results

Figure 4 shows the first observational EOF and the projected underlying process for the ensemble of climate model projections for the three time indexes. M and Q behave very similarly, in that they capture the same general behavior in the 20th century and have an increase in the 21st century, which ranges from 10 to 30 units and from 5 to 20 units, respectively. The decadal jump term in D corrects for the artificial jumps, and has a less pronounced increase into the 21st century, ranging from 0 to 5 units. The EOF scale is not a temperature scale, but once EOFs are projected onto their spatial counterparts, the outcome is in the original temperature units.

Figure 5 (top) shows the level process and the jump process, together with 95% probability intervals. We can see that the level process increases in the 21st century and that the model is able to describe significant jumps. The model discrepancy terms for time index D in Figure 6 show how the coefficients for some specific GCMs vary from the common level. We observe that two of the projections, NCAR-PCM and HadCM3, over and under estimate at a range of (-3.5,5). We can also see that jumps vary in intensity between models and become more apparent in the 21st century. In figure 6 we can see the model discrepancies for the three different time indexes for the first EOF. The discrepancies for M and Q are very similar in that the model bias follows the same structure, with more noise for M. For D, the weights are similar in the sense that NCAR-PCM and HadCM3 are given large positive and negative weights, suggesting that they are over and under approximating, while the remaining 3 are given approximately zero weight. If we look at the observation variance parameters, $\Sigma_{1,1}$ for the 3 time indexes we see a decrease , from 9.5 to 3.3 to

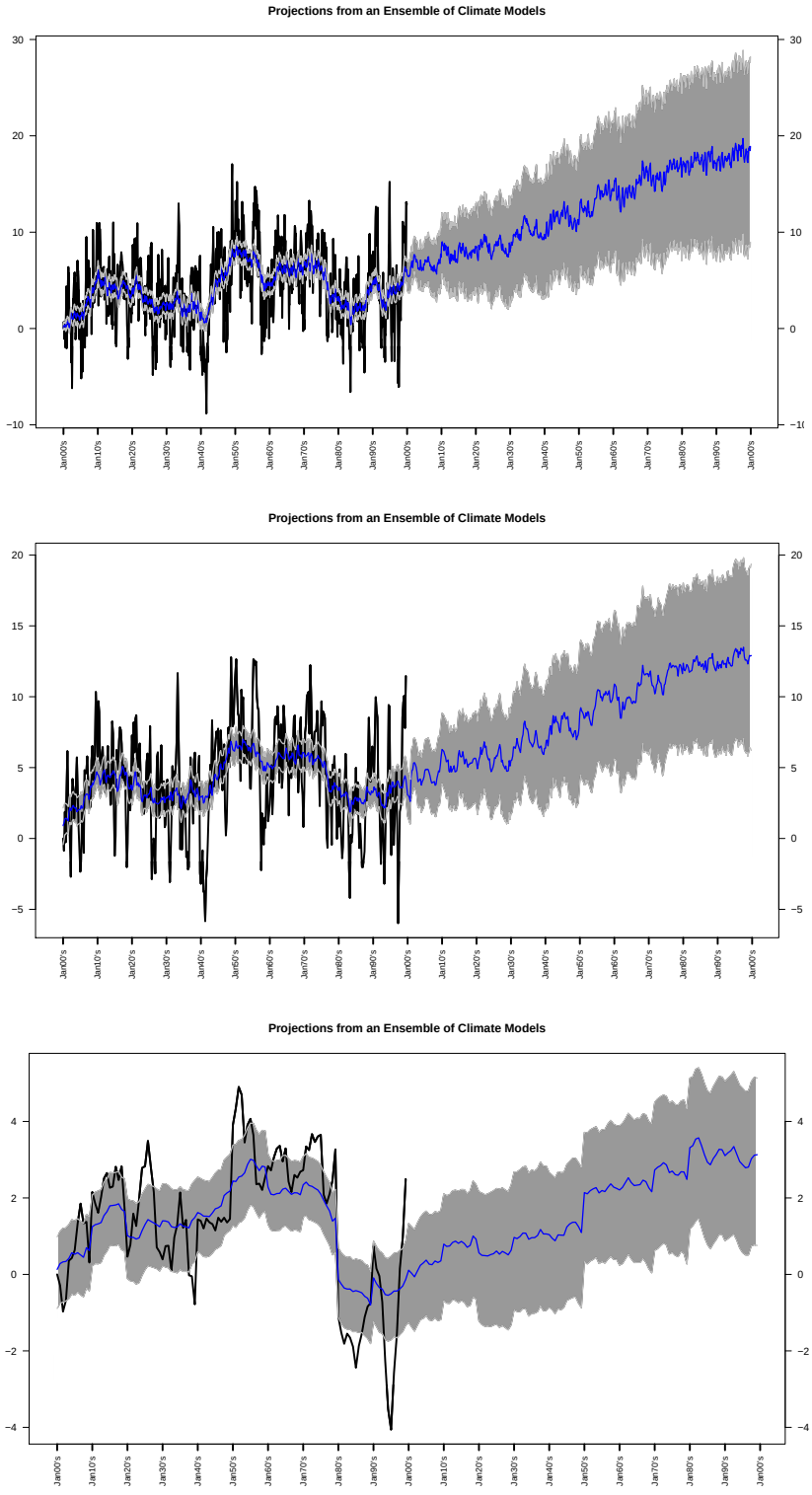

Figure 4: First EOF Projections from an Ensemble of Climate Models from top to bottom: Monthly, Quarterly, and Monthly Decadal.

1.16, when comparing between M, Q, ad D. This is to be expected since we smooth the noise for the monthly data by a third then a tenth for Q and D respectively. The variance parameters do not change for the climate model projections when comparing them over the 3 indexes.

From the ensemble projections, ψ_l, obtained for the EOFs, we perform a reconstruction of the sea surface temperature field. We define the reconstruction as

$$\tilde{X}^r(\mathbf{s},t) = \sum_{l=1}^{r} u_l(\mathbf{s})\psi_l(t) \tag{7}$$

where the climatology is then added to $\tilde{X}^r(\mathbf{s},t)$ to produce SST. We use this reconstruction for D and compare the results to the approximated observational SST using $r = 10$, which gives us most of the explained variability. The methods stated above were used in modeling the other eight indexes.

In Figure 7 we can see the differences between the present observational SST (January 1990's) and the reconstructed SST for the month of January, in the 1960s, 1990s and 2020s, for the model of time index D. Our model suggests that, compared to the present conditions, the past was cooler and the future will be warmer with differences ranging from (-0.45K,-0.10K) for the past and (0.12K,0.44K) for the future for the first and third quantile. The reconstruction for the current SST (middle) provides a visual examination of how well the model is performing, with differences ranging from (-0.24K,0.04K) . The difference is not large , suggesting that our model can describe the process of present day SST. Figure 8 shows the time series of observational SST for the 20th century, as well as the SST reconstruction and GCM for the 20th and 21st century. From this figure we can see reconstructed SST closely follow observational SST, in both slope and amplitude, much closer than any single GCM. In the 21st century, it is important to note that the reconstruction is not just the average of the GCMs. While the amplitude is smaller in some cases, the time series do have a similar increasing slope. In Figure 9 we present the spatial reconstruction for monthly decadal SST, for the month of July in the years 2020, 2040, and 2060. We can see noticeable temperature increase in the far north and along the southeast coastline. This reconstruction suggests that the ocean is getting warmer in certain parts faster than others.

5 Conclusions

We have created a methodology to obtain a unified forecast of oceanic indexes using historical records and GCM simulations. In response to the desideratum of high frequency and high spatial resolution in the projections, we have developed a model that focuses on the main modes of spatial variability. The model blends the coefficients of those modes, using the different sources of information and accounting for seasonal cycles. The resulting index predictions have an interest of their own, as they can be associated with global oceanic dynamics and specific environmental changes. They can also be used to reconstruct the spatio-temporal fields and obtain high spatial resolution SST predictions. The problem of comparing month to month observational data and GCM simulations is tackled by assuming a dynamic model that produces smoothed estimates of the discrepancies.

Predictions show a degeneracy in the GCMs, around the middle of the 21st century. This could be attributed to the strong assumption in our model that the spatial EOF pattern does not change over time, and thus, the GCMs become out of synchrony with the 20th century climatology.

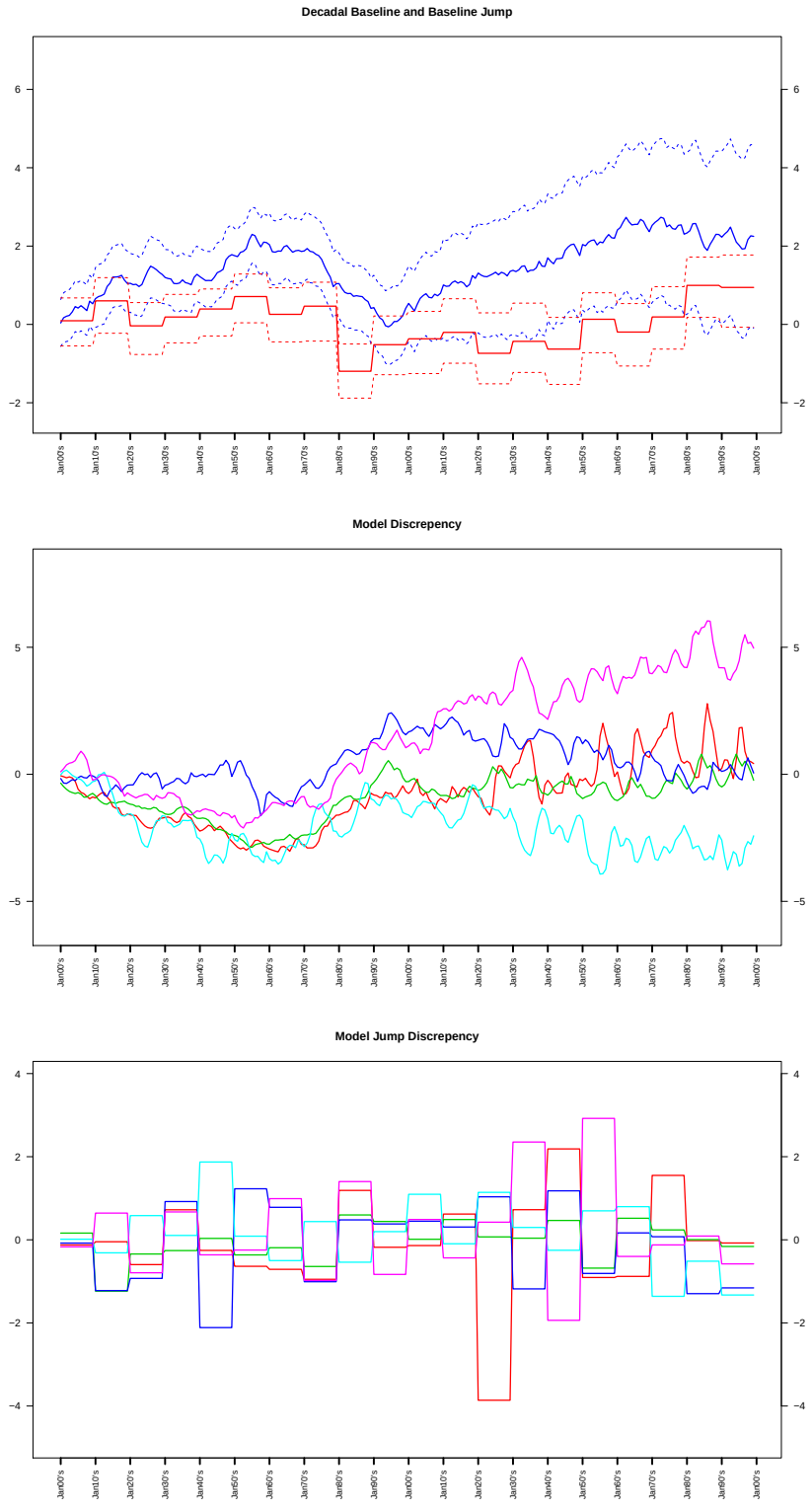

Figure 5: Underlying process and jump (top) with 95% credible intervals for the first EOF of time index D. Model bias terms (middle) and Model Jump bias terms (bottom) for the projections GFDL-CM2.1, GISS-ER, CCSM3, NCAR-PCM, HadCM3.

Figure 6: Model discrepancy terms for M (top), Q (middle) and D (bottom) for the projections GFDL-CM2.1, GISS-ER, CCSM3, NCAR-PCM, HadCM3.

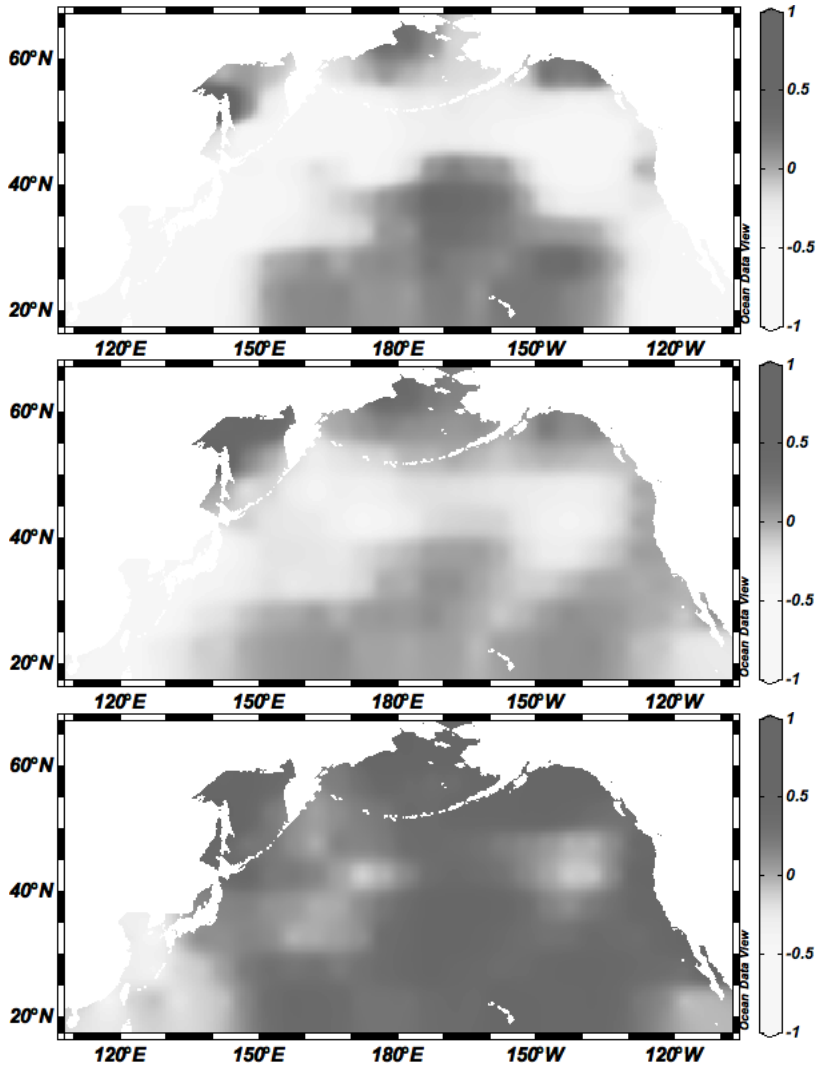

Figure 7: Differences between Monthly Decadal SST for January of 1990's and SST Reconstruction using ten EOFs for the decades of 1960s,1990s,2020s (top to bottom).

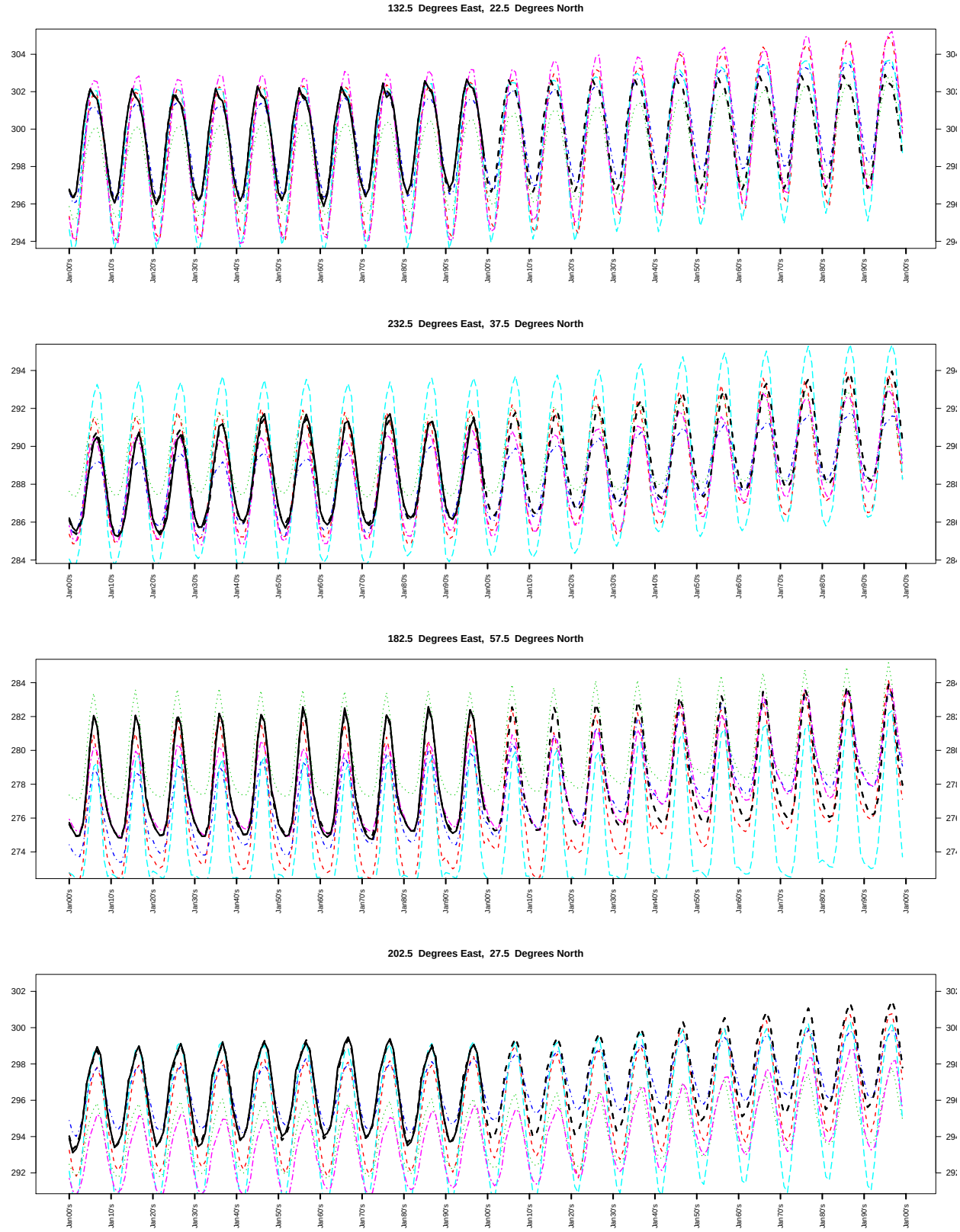

Figure 8: Monthly Decadal Observational SST (Solid Black), Global Climate Models, and SST Reconstruction (Doted Black) for station site (132.5,22.5), (232.5,37.5), (182.5,57.5), (202.5,27.5) where (Degrees East,Degrees North) (Top to Bottom).

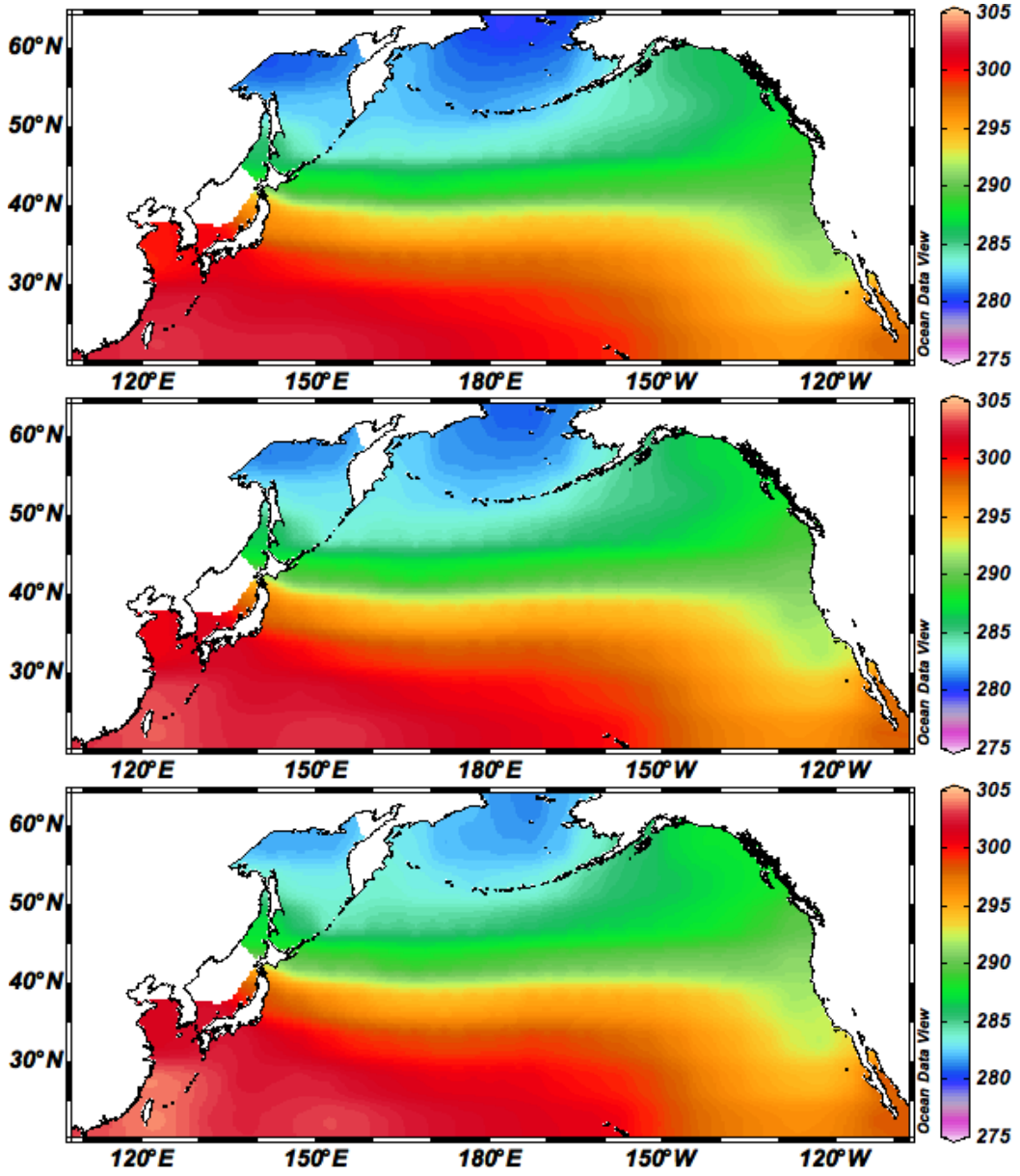

Figure 9: Monthly Decadal SST future reconstruction for July in the year 2020 (top), 2040 (middle), 2060 (bottom).

Alternatively, it could just be an artifact of the GCMs themselves, as they simulate climate into the future. Another possibility is that the climate models are predicting a dynamical shift in seasonal behavior, meaning winter may come earlier or later. Either way, by considering a joint analysis of high frequency data, as opposed to taking each season separately, our statistical model can tackle this issue. Our model is flexible enough to deal with time-varying cycles. The changing seasonality, though, remains an important GCM feature that needs to be considered carefully, and that highlights the difficulties of predicting the climate 100 years into the future.

We performed an analysis that considered three different levels of temporal aggregation. The results show that, at least for the first two EOF coefficients, all three analyses capture similar structural features. This is likely due to the smoothing induced by the dynamical model. We see, in particular, that the results from M are a noisier version of those from Q. We do observe some relevant differences when we compare D to M and Q. The level of increase in the 21st century for D is much smaller than for M and Q, and has less variability. Computationally, D is, of course, much faster to deal with that M, as it involves an order of magnitude less data. D corresponds to the level of time aggregation used most frequently in the literature. By taking decadal averages we create an artificial feature not present in the original data set. This feature is a jump caused by the time discontinuity created at the turn of each decade, which we account for in our modeling structure. The visible jumps in the 21st century projections suggest that the model does not disregard this feature after the 20th century, but incorporates the jumps in the future predictions. The traditional approach to combining multi-model GCM ensembles considers decadal seasonal data separately for each season. In comparison with this approach, our model for D has the advantage of providing continuity between decades and months. This allows a detailed description of the dynamics of future seasonal patterns. We have focused on representation of the spatio-temporal fields in terms of EOFs.

The methodology proposed in this paper can be used for expansions on any set of bases, provided the same set is used for all the GCMs and the observations. The resolution of the reconstructed predictions obtained in this paper is high relative to the large regions commonly used in the literature, but it is still too coarse for most practical ecological studies. We are currently developing extensions that will allow for predictions at much higher level of spatial detail.

Acknowledgments

We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy. The first author was partially funded by a Fisheries and the Environment grant from the National Oceanic and Atmospheric Administration. The second author was partially supported by the National Science Foundation grant DMS-0906765.

A Climate Models

1. Bjerknes Centre for Climate Research, Norway, (BCCR-BCM2.0) 2005
2. Canadian Centre for Climate Modelling and Analysis, Canada, (CGCM3.1.T47) 2005
3. Meteo-France/Centre National de Recherches Meteorologiques, France, (CNRM-CM3) 2004

4. Commonwealth Scientific and Industrial Research Organization Atmospheric Research, Australia, (CSIRO-MK3.0) 2001
5. National Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics/Institue of Atmospheric Physics, China, (FGOALS-g1.0) 2004
6. National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, USA, (GFDL-CM2.1) 2005
7. National Aeronautics and Space Administration/Goddard Institute for Space Studies, USA, (GISS-ER) 2004
8. Institute for Numerical Mathematics, Russia, (INM-CM3.0) 2004
9. National Institute of Geophysics and Volcanology, Italy, (INGV-ECHAM4.6)
10. Institut Pierre Simon Laplace, France, (IPSL-CM4) 2005
11. Center for Climate System Research, National Institute for Environmental Studies, and Frontier Research Center for Global Change, Japan, (MIROC3.2 medres) 2004
12. Meterorological Institude of the University of Bonn, Meteorological Research Institude of the Korea Meteorological Administration and Model and Data Group, German/Korea, (Echo-G) 1999
13. Max Planck Institute for Meteorology, German, (ECHAM5/MPI-OM) 2005
14. Meteorological Research Institute, Japan, (MRI-CGCM2.3.2) 2003
15. National Center for Atmospheric Research, USA, (CCSM3) 2005
16. National Center for Atmospheric Research, USA, (PCM) 1998
17. Hadley Centre for Climate Prediction and Research/Met Office, UK, (UKMO-HadCM3) 1997

References

Cloern, J., Hieb, K., Jacobson, T., Sansó, B., Di Lorenzo, E., Stacey, M., Largier, J., Meiring, W., Peterson, W., Powell, T., Winder, M., and Jassby, A. (2010). Biological communities in San Francisco Bay track a north Pacific climate shift. *Geophysical Research Letters*. To appear.

De La Horra, J. and Rodrguez-Bernal, M. (1999). The posterior predictive p-value for the problem of goodness of fit. *TEST*, 8:117–128. 10.1007/BF02595865.

Di Lorenzo, E., Fiechter, J., Schneider, N., Bracco, A., Miller, A. J., Franks, P. J. S., Bograd, S. J., Moore, A. M., Thomas, A. C., Crawford, W., Pena, A., and Hermann, A. J. (2009). Nutrient and salinity decadal variations in the central and eastern north pacific. *Geophysical Research Letters*, 36(L14601). doi:10.1029/2009GL038261.

Di Lorenzo, E., Schneider, N., Cobb, K. M., Chhak, K., Franks, P. J. S., Miller, A. J., McWilliams, J. C., Bograd, S. J., Arango, H., Curchister, E., Powell, T. M., and Rivere, P. (2008). North pacific gyre oscillation links ocean climate and ecosystem change. *Geophysical Research Letters*, 35(L08607). doi:10.1029/2007GL032838.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis*. Chapman and Hall, New York, second edition.

Giorgi, F. and Mearns, L. O. (2002). Calculation of average, uncertainty range, and reliability of regional climate changes from aogcm simulations via the "reliability endemble averaging" (rea) method. *Journal of Climate*, 15:1141–1158.

Hannachi, A., Jolliffe, I. T., and Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int J Climatolo*, 27:1119–1152.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate*, 23(10):2739–2758.

Mantua, N. and Hare, S. (2002). The pacific decadal oscillation. *Journal of Oceanography*, 58(1):35–44.

Mantua, N., Hare, S., Zhang, Y., Wallace, J. M., and Francis, R. C. (1997). The pacific interdecadal climateoscillation with impacts on salmon production. *American Meteorological Society,*, 78(6):1069–1079.

Meehl, G., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E. (2007). The wcrp cmip3 multi-model dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, 88:1383–1394.

Meehl, G. A., Hu, A., and Tebaldi, C. (2010). Decadal prediction in the pacific region. *American Meteorological Society*, 23:2959–2973.

Nakicenovic, N. and Swart, R. (2000). *IPCC Summary for Policymakers: Emissions Scenarios*. Cambridge University Press, The Edinburgh Building Shaftesbury Road, Cambridge CB2 2RU ENGLAND.

Randall, D., Wood, R., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R., Sumi, A., and Taylor, K. (2007). *Cilmate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, 108(D14). doi:10.1029/2002JD002670.

Smith, R., Tebaldi, C., Nychka, D., and Mearns, L. (2008). Bayesian modeling of uncertainty in ensembles of climate models. *Journal of the American Statistical Association*. (forthcoming).

Smith, R. L., Tebaldi, C., Nychka, D., and Mearns, L. O. (2009). Bayesian modeling of uncertainty in ensembles of climate models. *American Statistical Association*, 104(485):97–116.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, B*, **64**:583–639.

Tebaldi, C. and Sansó, B. (2008). Joint projections of temperature and precipitation change from multiple climate models: A hierarchical Bayes approach. *Journal of the Royal Statistical Society, A*, 172:83–106.

Tebaldi, C., Smith, R., Nychka, D., and Mearns, L. (2005). Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multi-model ensembles. *Journal of Climate*, 18(10):1524–1540.

Thompson, D. W. J. and Wallace, J. M. (1998). The arctic oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters*, 25(9):1297–1300.

West, M. and Harrison, P. (1997). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York, second edition.

Yaglom, A. (1986). *Correlation Theory of Stationary and Related Random Functions I: Basic Results*. Springer Series in Statistics. Springer-Verlag, New York.