Deep-Sea Research II 134 (2016) 30–47

Contents lists available at ScienceDirect

Deep-Sea Research II

journal homepage: www.elsevier.com/locate/dsr2

Regular article

Projected future biophysical states of the Bering Sea

Albert J. Hermann [a,e,*], Georgina A. Gibson [b], Nicholas A. Bond [a,e], Enrique N. Curchitser [c], Kate Hedstrom [d], Wei Cheng [a,e], Muyin Wang [a,e], Edward D. Cokelet [e], Phyllis J. Stabeno [e], Kerim Aydin [f]

[a] Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98195, USA
[b] International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[c] Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901, USA
[d] Arctic Region Supercomputing Center, Fairbanks, AK 99775, USA
[e] Ocean Environment Research Division, NOAA/PMEL, Seattle, WA 98195, USA
[f] Alaska Fisheries Science Center, NOAA/PMEL, Seattle, WA 98195, USA

ARTICLE INFO

Available online 10 November 2015

Keywords:
USA
Alaska
Bering Sea
Modelling

ABSTRACT

Three global climate simulations from the Intergovernmental Panel on Climate Change Fourth Assessment (AR4) were used as physical forcing to drive a regional model that includes both physical and biological elements of the Bering Sea. Although each downscaled projection indicates a warming of 1–2 °C between 2010 and 2040 on the Bering Sea shelf, the interannual and interdecadal details of this trend vary considerably among the three realizations. In each case, the magnitude of presently observed interannual variability of bottom temperatures and ice cover is found in the models to be maintained out to at least 2040, but with a steadily increasing probability of warm years with less ice on the southern shelf. The overall trends indicate warmer temperatures and the retreat of ice in the southeastern Bering Sea, but continued ice cover in the northeastern Bering Sea. Sensitivity analyses suggest both increasing air temperature and northward wind stress as primary drivers of higher water-column temperatures. Based on currently available models, changes in shortwave radiation are not likely to have a significant role in this warming. Warming trends on the outer shelf may lead to decreased production of large crustacean zooplankton at that location, but could increase such production on the inner shelf.

1. Introduction

1.1. Overview of the Bering Sea ecosystem

Prominent physical features of the Bering Sea include seasonal ice cover, strong advection of ice, and tidally generated biophysical domains. Ice formed each winter in the northern Bering Sea is advected to the southeast, where it gradually melts as it encounters warmer water and air temperatures. This southward advection contributes to the latitudinal salinity gradient of the Bering Sea and its interannual variability. A cross-shelf gradient in the vertical penetration of tidal mixing sets up distinct biophysical regimes with associated biota. Classically, the southeastern shelf is classified as having three biophysical domains: a vertically well-mixed inner region (~0–50 m), a middle region which is well-mixed in the winter and has two distinct layers separated by a sharp thermocline in the summer (~50–100 m), and an outer region which is more gradually stratified (~100–200 m) (Kinder and Schumacher, 1981; Coachman, 1986; Kachel et al., 2002).

Distinct biological features of the Bering Sea ecosystem include ice algae as a potential food source to secondary producers, and strong benthic–pelagic coupling. Within the different biophysical regimes, the relative magnitude of pelagic vs. benthic pathways of carbon flux varies interannually, and is believed to be strongly influenced by the extent of seasonal ice through its effects on stratification (Hunt et al., 2002, 2011). The relative importance of pelagic vs. benthic pathways is likely to shift under the influence of global warming, partially through its impact on seasonal ice extent in the Bering Sea. Field data suggest that recent cold temperatures in the Bering Sea have led to an increase in large crustacean zooplankton, favored as food items by juvenile pollock in the fall season (Coyle et al., 2011).

The present hydrography and climatology of the Bering Sea result in a highly productive ecosystem, with plankton biomass ultimately supporting large populations of shellfish and finfish (and major fisheries), marine birds and marine mammals. Such intense production derives, in part, from a broad shelf with strong tidally

* Correspondence to: Pacific Marine Environmental Laboratory, 7600 Sand Point Way NE, Seattle, WA 98115, USA. Tel.: +1 206 526 6495; fax: +1 206 526 6485.
E-mail address: albert.j.hermann@noaa.gov (A.J. Hermann).

http://dx.doi.org/10.1016/j.dsr2.2015.11.001
0967-0645/

induced mixing, a plentiful supply of the micro-nutrient iron, and seasonal stratification which maintains the phytoplankton in the euphotic zone, adjacent to a deep, macronutrient-rich basin. Cooling trends in the Bering Sea from 2006 to 2011 (Stabeno et al., 2012a) have been documented by the Bering Sea Ecosystem Program (BEST), the Bering Sea Integrated Ecosystem Research Program (BSIERP), the U.S Bering-Aleutian Salmon International Survey (BASIS), and the North Pacific Climate Regimes and Ecosystem Productivity Program (NPCREP). Measurements since late 2013 indicate a return to warmer conditions, with reduced ice (Stabeno et al., 2016). The response of Bering Sea production to changes in temperature is not yet completely clear and may depend on the timescale under consideration. While ocean color observations over the Bering Sea suggest that primary production during warm years may be enhanced by 40–50% compared to cold years (Brown and Arrigo, 2013) it has also been suggested that very warm temperatures suppress summer production, because intense water column stratification (Coyle et al., 2008) reduces the re-supply of nutrients to the upper mixed layer.

A model-based multivariate analysis was used to help explore the relationships between physical and biological factors on the Bering Sea shelf (Hermann et al., 2013, henceforth referred to as H2013). The analysis suggested that the Bering Sea shelf may not respond uniformly to changes in climate forcing. For example, large crustacean zooplankton (LCZ) are negatively correlated with temperature on the outer, southwestern shelf, and positively correlated to temperature on the inner, northeastern shelf. Areas of positive correlation tend to correspond with those areas with greatest change in ice cover. As in the revised Oscillating Control Hypothesis of Hunt et al. (2011), the ratio of large to total zooplankton is enhanced at lower temperatures. On the outer shelf, higher temperatures may be leading to reduced LCZ production either through effects on stratification (and hence nutrient limitation), or through direct effects of temperature on growth, respiration, predation and vertical migration. Changes on the northern shelf may involve a complex interplay of light and nutrient limitation effects, as modulated by a reduction in the duration of seasonal ice cover.

Ice dynamics of the Bering Sea have been explored in both observational and modeling studies (Stabeno et al., 2010; Danielson et al., 2011a, 2011b, Cheng et al., 2014; Li et al., 2014a, 2014b; Sullivan et al., 2014). Ice is formed seasonally in the northern Bering Sea and is advected southward, resulting in a net transfer of freshwater from north to south. Heat budgets from these studies have underscored the importance of sensible heat flux between the atmosphere and the ice in the northern Bering, and between the ocean surface and the ice in the southern Bering, where the ice edge retreats each spring.

1.2. Organization of this paper

We begin with a description of the global and regional models used for the biophysical modeling of this region, and the physical data used for comparison with the models. This is followed by a comparison of the model output with some of the moored and gridded data collected during 1971–2012. We next consider our three downscaled projections of future conditions in the Bering Sea, and highlight the most significant changes from present conditions. Finally, we consider which elements of the physical forcing appear most likely to govern the projected changes, using the coherence among relevant pairs of physical and biological features.

2. Methods

2.1. The physical models

Both global and regional models were used in our analysis. The method of coupling is described in H2013. Briefly, the global ocean model output is interpolated in time and space, and applied as initial and boundary conditions for the finer grid regional model (one-way nesting). Similarly, the global atmospheric output is interpolated and applied as surface forcing on our regional ocean model. A description of both global and regional models follows.

2.1.1. Global reanalyses used for regional hindcasts

As described in H2013, there are two global hindcasts and one larger regional model used for our regional downscaling to the Bering Sea. The relevant model products are: (1) the Common Ocean Reference Experiment atmospheric reanalysis (CORE; Large and Yeager, 2008); the Simple Ocean Data Assimilation oceanic reanalysis (SODA, Carton and Giese, 2008); the regional northeast Pacific (NEP-5) simulation of Danielson et al. (2011a); and the Climate Forecast System Reanalysis [a combined atmospheric and oceanic reanalysis] (CFSR; Saha et al., 2010). CORE was utilized as atmospheric forcing for our hindcast of years 1969–2004, with oceanic boundary conditions interpolated from NEP-5. NEP-5 itself utilized CORE atmospheric forcing and oceanic boundary conditions derived from SODA. CFSR was utilized for both atmospheric forcing and oceanic boundary conditions for the hindcast of years 2003–2012. Note that CORE is a global atmospheric reanalysis spanning 1950–2004, while CFSR is a coupled atmospheric and oceanic reanalysis spanning 1979-present. Use of the CORE product for most of our hindcast is based on availability of CORE forcing variables for the earlier decades, as well as its broad acceptance within the oceanic community. CORE products were not available for years beyond 2005 at the time our analyses were begun, while CFSR products were not available for years before 1979; hence both were employed for different portions of our hindcast. Overlapping simulations of 2003 and 2004 allowed a comparison of results using the two reanalyses; these were used to adjust CFSR for compatibility with CORE. Ultimately our “continuous” (i.e. concatenated) hindcast series was composed of the CORE results for 1970–2004, and CFSR results for 2005–2012.

2.1.2. Global models used for regional forecasts

Global forecasts used in the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4) were archived by the Coupled Model Inter-comparison Project (CMIP3; Meehl et al., 2007). These models and a subsequent generation of models (CMIP5) have been used by the IPCC to explore the coupled global atmospheric and oceanic response to anticipated changes in atmospheric CO_2. Three of these global forecasts, under AR4 emissions scenario A1B, were selected for application to regional downscaling based on their fidelity to present mean conditions in the Bering Sea and the northeastern Pacific (Wang et al., 2010). The A1B scenario was chosen as it is in the middle of the range of AR4 scenarios; it assumes rapid economic growth and continued use of fossil fuels along with other, non-fossil energy sources. The three models chosen were CGCM3-t47, ECHO-G and MIROC3.2 (each described below). A single realization (projection) from each of these three models was used for atmospheric forcing and oceanic boundary conditions for our regional, downscaling projections of the Bering Sea, spanning years 2003–2040. Among these three, the ECHO-G model realization exhibited the least warming, and the CGCM3-t47 model realization exhibited the greatest warming. The seasonality of these changes also differed among the three models.

The Coupled Global Climate Model, t47 grid (CGCM3-t47), from the Canadian Centre for Climate Modelling and Analysis (CCCMA),

Table 1
Spatial and temporal resolution of the three IPCC models.

MODEL	CGCM3.1-t47	MIROC	ECHOG
OCEAN	1.85° Latitude 1.85° Longitude Monthly	~1.0° Latitude ~0.5° Longitude Monthly	~2.8° Latitude[a] ~2.8° Longitude Monthly
ATMOSPHERE	3.75° Latitude 3.75° Longitude Daily	~2.5° Latitude ~1° Longitude Daily	~3.7° Latitude ~3.75° Longitude Daily

[a] Finer near equator.

makes use of the same ocean component as that used in the earlier CGCM2 (Flato et al., 2000), but employs an updated atmospheric component (AGCM3; McFarlane et al., 2005; Scinocca et al., 2008). The ocean component is described in Flato and Boer (2001), Kim et al. (2002, 2003), and references therein. A two-category sea-ice component (mean thickness and concentration) is based on Flato and Hibler (1992) and includes a prognostic equation for ice concentration (Hibler, 1979).

The Hamburg Atmosphere-Ocean Coupled Circulation Model (ECHO-G; Legutke and Voss, 1999), from the Max Planck Institute in Germany, is based on the atmospheric model ECHAM4 (Roeckner et al., 1996) and the ocean model HOPE (Wolff et al., 1997), and includes a thermodynamic-dynamic sea ice model (Hibler, 1979).

The Model for Interdisciplinary Research on Climate, medium-resolution version (MIROC3.2-Medres) was developed by a consortium of agencies in Japan. This coupled model includes the AGCM5.7b atmospheric model and the COCO3.3 ocean and sea ice model (K-1 Model Developers et al., 2004).

The resolution of output from these models is shown in Table 1. In each case, the monthly ocean values were used as initial conditions for the regional ocean simulations. The daily atmospheric and monthly oceanic outputs were interpolated in space and time for use in the surface forcing and boundary conditions of the regional ocean model, as described in H2013.

2.1.3. The regional model

The regional model used for downscaling (also referred to as the "Bering10K" model) has been described extensively in H2013. Briefly, this regional model is based on the Regional Ocean Modeling System (ROMS, Haidvogel et al., 2008; Shchepetkin and McWilliams, 2005), and is a direct descendant of the 60-layer model used by Danielson et al. (2011a) in their analysis of the Bering Sea. Our regional model is implemented at ~10-km resolution, with 10 vertical layers. This limited vertical and horizontal resolution allowed a larger number of tuning runs, a multidecadal hindcast, and a modest "ensemble" of three future downscaling projections. As is frequently the case in modeling, it reflects a compromise between realism and feasibility. Vertical mixing is based on the algorithms of Large et al. (1994). Both ice (Budgell, 2005) and tidal dynamics are included in this model; the explicit inclusion of tidal flows allows tidally-generated mixing and tidal residual flows to develop. Freshwater runoff was applied by freshening of the salinity field within a few gridpoints of the coastline, using the monthly runoff values of Dai et al. (2009). Bulk forcing, based on algorithms of Large and Yeager (2008), were used to relate winds, air temperature, specific humidity, and downwelling shortwave and longwave radiation to surface stress and the net transfers of sensible heat, latent heat, net shortwave and net longwave radiation through the sea surface. The full domain of the regional model spans the entire Bering Sea; model bathymetry for the domain relevant to the present analysis is shown in Fig. 1.

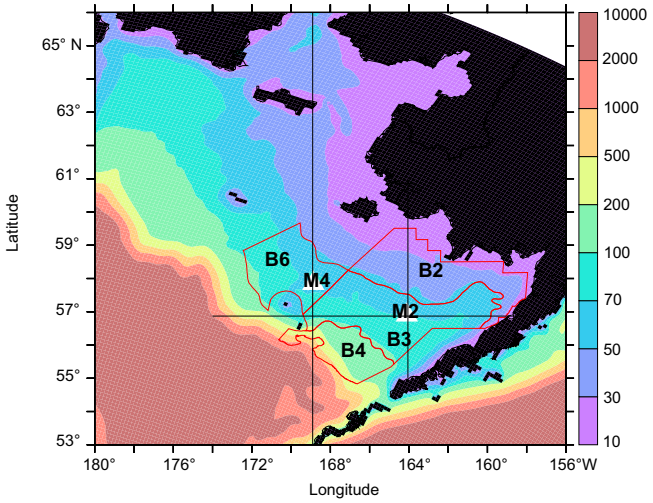

Fig. 1. A subregion of the ROMS model domain with shaded bathymetry (m) (for full domain see H2013). The locations of biophysical moorings M2 and M4 are shown as white triangles. Biophysical domains B2 (south inner shelf), B3 (south middle shelf), B4 (south outer shelf) and B6 (midnorth middle shelf) used in BEST-BSIERP studies (Ortiz et al., 2012), and in the present analysis, are outlined in red (thick black in print version). Thin black lines indicate the location of north–south and east–west sections used for model/data comparisons. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

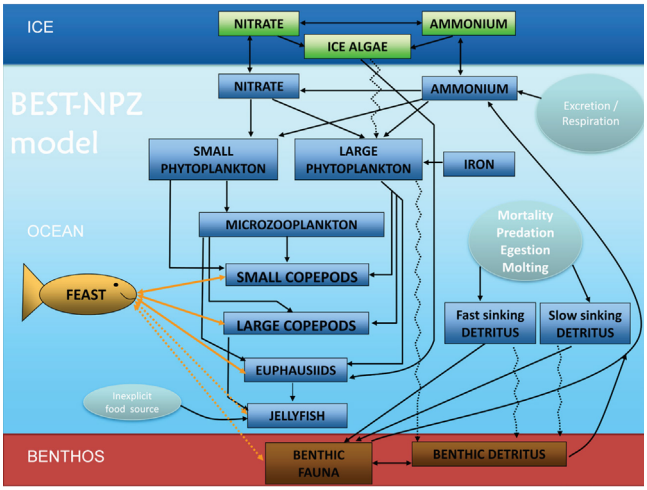

Fig. 2. Structure of the NPZ model used in this study (from Gibson and Spitz, 2011). Arrows indicate the direction of material flow. "FEAST" refers to a size- and age-structured fish model (not used in the present study), described by Ortiz et al. (submitted).

The biological model used in our simulations is based on the NPZ model of Gibson and Spitz (2011). The pelagic components of this model are: nitrate, ammonium, iron, small phytoplankton, large phytoplankton, microzooplankton, small copepods, large copepods, krill (euphausiids), jellyfish, and slow and fast sinking detritus. The model's pelagic core is coupled to an ice biology module comprising ice algae, ice nitrate, and ice ammonium, and to a benthic module comprising benthic infauna and benthic detritus. The NPZ model components and material flows are illustrated in Fig. 2.

Two new features distinguish the present model from the version described in H2013:

1) Phytoplankton and bottom depth are now used to calculate the local depth profile of absorption of shortwave radiation in the water column, for use in the heat budget. The formula used for this absorption is similar to that used for PAR in Gibson and Spitz (2011), as follows:

$$I = I_o e^{-kz} \quad (1)$$

$$k = k_w + 2.0e^{-0.05H} + k_c\left(<c_s>/r_s + <c_l>/r_l\right)^{.428} \quad (2)$$

where

I/I_o = fraction of incident shortwave I_o at (x,y) penetrating to depth (z) (m)

$<c_s>$ = depth-average small phytoplankton above (z) at (x,y) (mg C m^{-3})

$<c_l>$ = depth-average large phytoplankton above (z) at (x,y) (mg C m^{-3})

r_s = carbon-to-chlorophyll ratio for small phytoplankton = 65.0

r_l = carbon-to-chlorophyll ratio for large phytoplankton = 25.0

k_w = background extinction coefficient = .046 (m^{-1})

k_c = extinction coefficient due to chlorophyll = .121 (m^{-1})

H = bottom depth at (x,y) (m)

This formula produced greater near-surface absorption of heat in near-shore areas and in summer, resulting in a better fit to observed temperature patterns on the southeastern shelf (see Section 3) The use of bottom depth H in (2) is a rough proxy for an assumed higher sediment concentration in the nearshore environment.

2) the factor Φ_f governing the increase of ice thickness versus horizontal ice coverage during freezing conditions (Mellor and Kantha, 1989) - that is, the ratio of horizontal to vertical growth of ice - has been changed from its original value of 4.0 (as suggested in Mellor and Kantha, 1989) to a value of 8.0. This yielded a better fit with the observed seasonal formation of ice on the southeastern shelf. This improvement and remaining seasonal biases (e.g. late melting of ice) will be discussed in Section 3.

Subsequent to model runs, output was re-gridded to a regular latitude-longitude-depth grid for comparison with data. The destination grid included horizontal spacing of 0.1 degrees latitude and 0.1 degrees longitude with the following depth levels in meters: 0, 5, 10, 15, 20, 30, 40, 50, 60, 75, 100, 125, 150, 200, 250, 300.

Spatial averages of model output were computed, using the standard BEST/BSIERP bio-regions, which were chosen to minimize within-region variance and to maximize variance across regions (Ortiz et al., 2012). Specifically we used the south inner shelf, south middle shelf, south outer shelf, and midnorth middle shelf domains, which are BEST-BSIERP bio-region numbers 2,3,4 and 6, respectively (Fig. 1).

2.2. Hydrographic data used for model validation

Several features of the previous version of this model were compared with observations in H2013. We will not attempt a complete replication of these comparisons with the updated version. We nevertheless focus on a few of the most important features, and utilize newly available datasets for an extended validation through 2012. As in H2013, we did not utilize any forward-based (e.g. 3DVAR) data assimilation for the hindcast runs, as the ultimate purpose of the runs was simulation of future conditions, for which all data is unavailable. Rather, model refinement and tuning of free parameters was conducted using available hydrographic data through 2012.

Multiple types of hydrographic data were used for this purpose. First, as in H2013, we utilized temperature profiles from long-term biophysical moorings maintained by the EcoFOCI group at stations M2 (56.87°N, 164.05°W) and M4 (57.85°N, 168.87°W) along the 70 m isobath (mid-shelf, Fig. 1). These data have been described extensively in previous publications (Stabeno et al., 2001, 2007, 2010, 2012a, 2012b); temperatures were measured approximately every 3 m in the upper 30 m and every 5–7 m below 30 m. As in H2013, vertical averages from these profiles (and some individual profiles as well) are compared with the equivalent time series from the updated Bering10K model hindcast.

Second, a long interannual time series of bottom temperatures from the Alaska Fisheries Science Center Bottom Trawl Survey (BTS; Stauffer, 2004), spanning 1971-present and centered on July each year, was used for model-data comparison. There are slight interannual variations in the spatial coverage of the BTS, which typically spans the EBS shelf from the Alaska Peninsula to ~62° N. For each sampled year, we binned the measured (BTS) and modeled (July average) values onto a $0.25 \times 0.25°$ latitude–longitude grid whose northern boundary is at 62° N. Equivalent spatial averages were computed for data and model, using only those bins containing BTS data.

Third, we compared modeled vs. observed areal ice coverage data (local fraction of the sea surface which is covered by ice) using data from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA; Donlon et al., 2012). Values were compared at the location of mooring M4, and along a north–south transect passing through M4 (Fig. 1).

Finally, we utilized newly available temperatures and salinities from regular yearly hydrographic (CTD) surveys of the Bering Sea, conducted in concert with a recent BTS (Cokelet, 2016). The BTS sampled the shelf between the Alaska Peninsula and 62°N in 2008–2009, and up to 66.67°N in 2010, from the 30-m isobath to the shelf break (~180 m), with a 37×37 km grid for a total of ~300 CTD casts per year. Cokelet (2016) quality-checked and interpolated these observations onto a regular latitude–longitude-depth grid ($1/3° \times 2/3° \times 1$ m), which provided an excellent resource for comparison with the model hindcast. These surveys took place from early June through mid-August of each year, and are hence not "synoptic" in the typical oceanographic sense of that term. Hence for detailed model-data comparisons, we extracted temperature and salinity data from the model at the precise dates and locations of the regridded CTD data. We focus on the data from 2010, which had the most spatially extensive survey (Fig. 3).

In addition to simple difference maps and line plots, we quantify the degree of fit/misfit using four metrics: the standard deviation of the model vs. that of the data, the correlation (Pearson's r) between model and data, the mean bias (where positive bias indicates the model had higher values than the data), and the RMS difference between model and data. For the 2010 CTD data, difference maps and metrics are calculated: (1) for the entire domain at all depths, (2) for the entire domain at 5 m depth, (3) for the entire domain at the sea bottom, (4) along an east–west section passing through mooring M2, and (5) along a north–south section passing through mooring M2. Locations of these sections are noted in Fig. 1.

2.3. Coherence analysis

Coherence analysis was used to examine relationships within bio-regions among atmospheric and oceanic variables, and between pairs of oceanic variables. For this purpose, spatially averaged time series of vertically averaged ocean temperature,

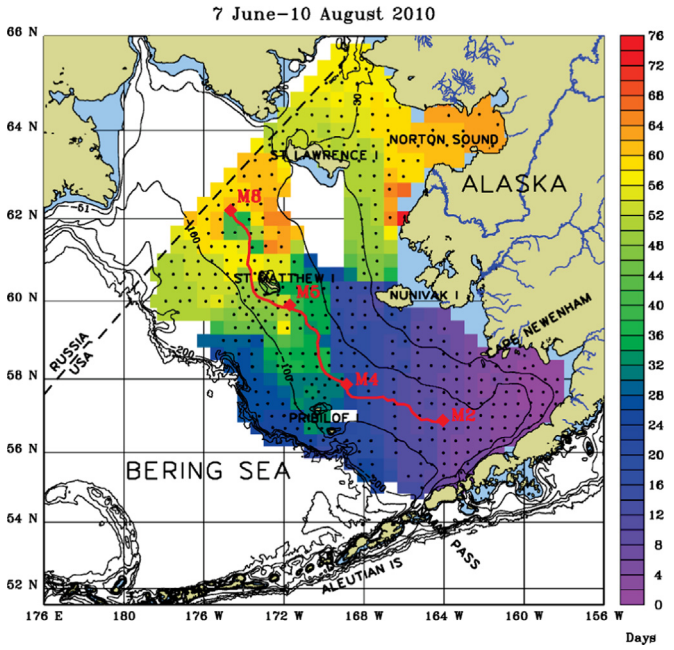

Fig. 3. Dates of hydrographic casts performed as part of the groundfish survey in 2010 (from Cokelet, 2016). Shading represents time since June 1, 2010 (days). Place names and the locations of biophysical moorings M2 and M4 are shown. Contours indicate actual bottom topography in m; 70 m isobath is highlighted in red.

vertically averaged euphausiid biomass, atmospheric temperature at 10 m above the sea surface, alongshore (northwestward) wind stress, and downwelling shortwave radiation were binned into 3-month (seasonal) averages, and anomalies from the seasonal climatology were calculated. The resulting anomaly series were used to compute spectra of individual series, as well as coherence and phase lag between series at seasonal to decadal periods. A shorter time series was used for shortwave radiation during the CORE reanalysis, as that product had only climatological monthly averages for shortwave (due to a lack of data for assimilation) prior to 1985.

3. Results

3.1. Hindcasts and comparison with data

3.1.1. Results at fixed moorings

In H2013, we compared measured temperatures with their model equivalent at moorings M2 and M4. Here, we focus our attention on the integrated heat content (mean temperature) at these locations (Fig. 4). As in H2013, the model time series illustrate strong variability at interannual to interdecadal time scales, such as the strong cold-to-warm "regime shift" of 1976 (Ebbsmeyer et al., 1991; Hare and Mantua, 2000). In recent decades, the model captures the sequence of relatively warm (2000–2005) and cold (2006–2012) years. To compare model versions, we focus on the period 2003–2009, which was simulated for both. The new algorithms for shortwave absorption and ice growth yield a similar overall fit at M2 and M4 (RMS model-data misfit of $\sim$1 °C for both M2 and M4 using either old or new model version); however, the updated ("new") version of the model yields a smaller overall bias at both locations, especially at M4, where the previous bias of $+1$ °C is now only -0.2 °C. As in H2013, the model captures the seasonal formation of the thermocline and its destruction in the fall at both locations (not shown here) and the north-to-south gradient in temperature (compare M2 and M4). Due partially to

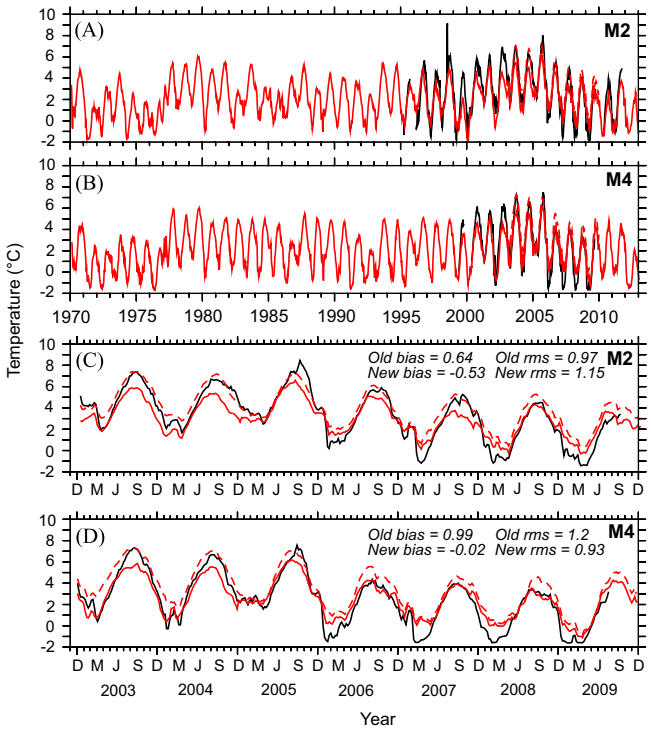

Fig. 4. Comparison of weekly depth-averaged temperatures (°C) at M2 (a) and M4 (b) mooring locations from measurements and model. Black line indicates measured values; dashed red line (grey in print version) indicates result from the model used in H2013 ("old" model); solid red line (grey in print version) indicates result from the updated model used in the present study ("new" model). An expanded view of years 2003–2009 is presented for M2 (c) and M4 (d). Also shown are the mean bias (*bias*) and the root-mean-square difference between model and data (*rms*) for each model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

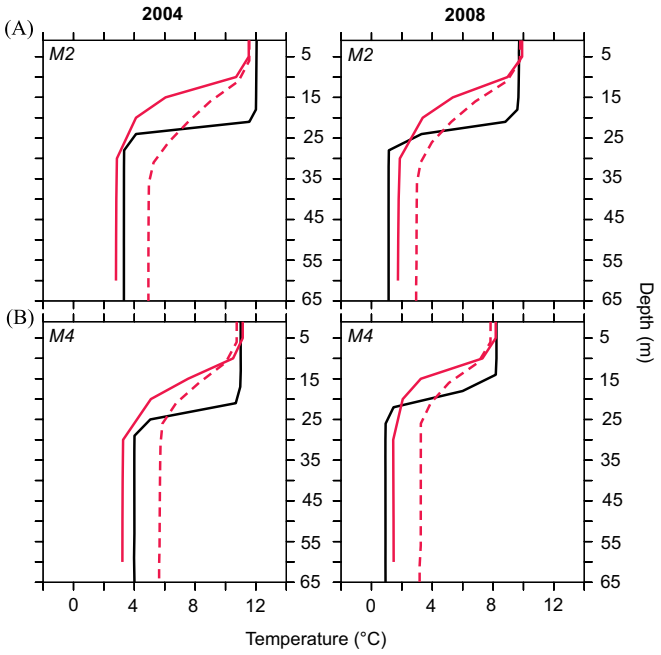

Fig. 5. Comparison of August 15 temperature (°C) at M2 (a) and M4 (b) mooring locations from measurements and model during a warm year (2004, left) and a cold year (2008, right). Black line indicates measured values; dashed red line (grey in print version) indicates result from the model used in H2013; solid red line (grey in print version) indicates result from the updated model used in the present study. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the limited (10-layer) model resolution, vertical profiles differ between the model and the data. As an example, we compare measured vs. modeled profiles at M2 and M4 on August 15 for a "warm" year (2004) and a "cold" year (2008) (Fig. 5). In both years, the M2 and M4 data exhibit a thermocline at ~25 m whereas the model thermocline is ~10 m shallower. The older version of the model failed to resolve the slope of the thermocline; the updated version is closer to the observed slope, with vertical gradients nearly twice as strong as the older version. The enhanced gradient in late summer apparently derives from the combined effects of the new, phytoplankton- and bathymetry- based heat absorption algorithm.

3.1.2. Comparison with bottom temperatures

As in the Fig. 4, periods of warm and cold conditions are clearly evident in the spatially averaged bottom temperatures from the BTS data (Fig. 6). Using the updated model, spatially averaged bottom temperatures are strongly correlated with the data (r=0.83), biased by only 0.02 °C, and yield an RMS difference from data of only 0.47 °C. Note how the old model has a 1.25 °C warm bias relative to data, and how this has been rectified using the new algorithms.

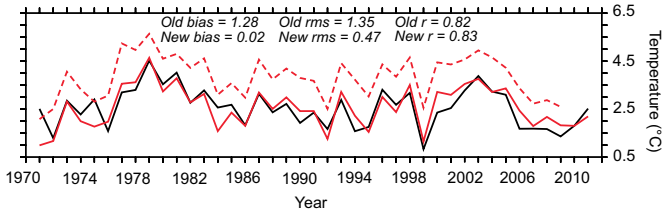

Fig. 6. Comparison of spatially averaged measured summer bottom temperatures (°C) from the Bering Trawl Survey with equivalent model results (July average for each year). Black line indicates measured values; dashed red line (grey in print version) indicates results from the model used in H2013 ("old" model); dark red line (grey in print version) indicates results from the updated model used in the present study ("new" model). Also shown are the Pearson's r-value (r), the mean bias (*bias*) and the root-mean-square difference between model and data (*rms*) for each model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1.3. Comparison with ice data

We restrict our attention to the areal ice cover, as little data are available for ice thickness. In both model and data, the onset of ice begins in late November 2009 in the northern Bering and gradually spreads to the south, reaching maximum extent in March 2010 (Fig. 7a and b). Ice cover is slightly higher than data in the model during December–April, but the largest discrepancy occurs in May and June, where the model is retaining a partial ice cover which is no longer observed in the satellite data. The modified ice formation algorithm (increased horizontal growth vs. thickening) yields a closer match to observed (OSTIA) ice coverage at M4 in recent years (and especially in 2009), but a stubborn bias towards late ice retreat was not affected by this change (Fig. 7a). The spatial extent of this bias can be seen by comparing a north–south transect of ice coverage for the new model vs. data for year 2010 (Fig. 7b). Note, however, that the OSTIA data may themselves be slight underestimates of total ice cover. During the melt season, SSM/I microwave radiometers (on which OSTIA estimates are partially based) cannot perfectly distinguish melt ponds from open water (Rösel et al., 2012). Such melt ponds may be seasonally extensive in the Bering Sea (indeed, Fig. 4 of Rösel et al., 2012 shows evidence of extensive melt ponds in the northern Bering for May 2008, a period of misfit shown in our Fig. 7a).

3.1.4. Comparison with gridded CTD data

For the remainder of this paper we utilize only the updated model results. The CTD survey for June–August 2010 allows a broad spatial comparison of temperatures and salinities. As noted in the Methods, this is not a synoptic survey of hydrography, as the survey spans approximately 75 days proceeding roughly from south to north. Indeed, the measurements indicate cooler near-surface temperatures

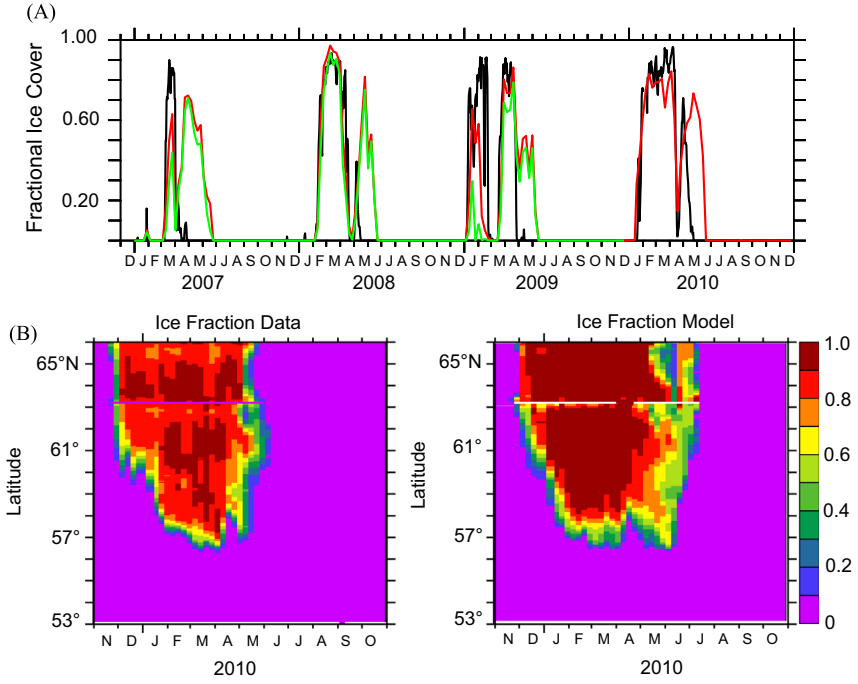

Fig. 7. (a) Comparison of weekly averaged fractional ice cover at mooring M4 from measurements and model. Black line indicates measured values; red line indicates result from model used in H2013; green line indicates result from model used in the present study. (b) measured (OSTIA) fractional ice cover along a north–south line passing through M4 during 2010 (left) vs. modeled fractional ice cover from the present study (right).

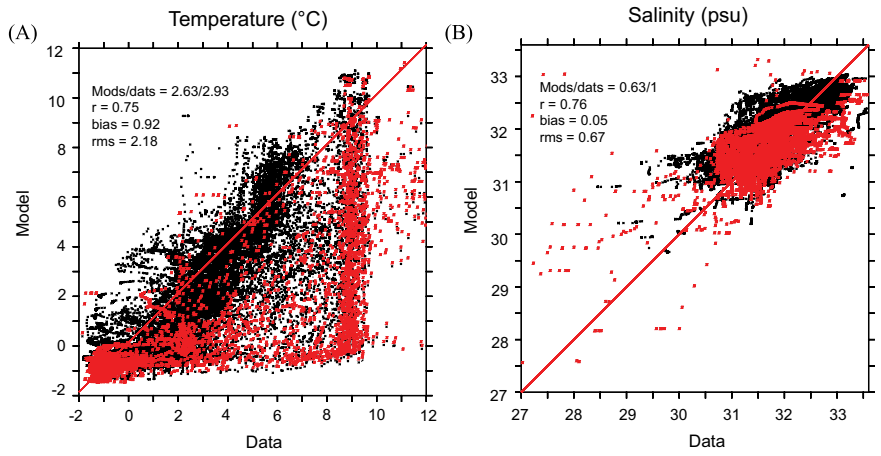

Fig. 8. Scatterplot comparison of all measured temperatures (a) (°C) and salinities (b) (psu) from the trawl survey (*x*-axes) with corresponding values from the model (*y*-axes), sampled at the same locations and times as the measurements. Values highlighted in red (grey in print version) indicate values from DOY 202 or later (i.e. more than 50 days since the start of the trawl survey, as shown in Fig. 3). Standard deviation of model (*mods*) data (*dats*) are shown, along with the Pearson's *r*-value (*r*), the mean bias (*bias*) and the root-mean-square difference between model and data (*rms*). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in the south, contrary to expectation for a synoptic survey in this region.

We begin with an overall comparison of all temperature data with their model equivalent (Fig. 8a). Standard deviations of model and data are similar at ~ 2.75 °C, and the *r*-value is 0.75, with a mean cold bias of ~ −0.9 °C. The RMS difference is ~2.2 °C. The best fit is obtained for measurements made prior to late July; the biggest misfit is for data above 8 °C, which were obtained after late July.

When all salinity data are considered (Fig. 8b), the model exhibits smaller standard deviation than the data (0.63 vs. 1.0 psu, respectively). The *r*-value is 0.76, and the mean bias is small (0.05 psu). The RMS difference is 0.67 psu. Lowest measured and modeled salinity values appear after late July.

Near-surface (5 m) temperature data from this non-synoptic trawl survey exhibited coldest values in the middle shelf south of 60 N, and highest values in Norton Sound (Fig. 9a). The model exhibits the best skill at representing the measured temperature values on the south middle shelf, and has a cold bias in excess of 2 °C in the northern areas. This finding is broadly consistent with the late ice retreat of the model relative to data. For the entire horizontal domain at 5 m depth, the mean model bias is 0.75 °C colder than the data and the RMS difference is 2.48 °C.

Observed near-surface salinities were highest along the shelf break, directly south of Nunivak Island, and near Bering Strait (Fig. 9b). Fresher waters were found near St. Matthew Island and Bristol Bay, with freshest values in Norton Sound. The model reproduces some of these features, but has a bias towards ~1 psu fresher waters than are observed on the south middle shelf (50–100 m bottom depths), and is strongly biased (> 2 psu too salty) in Norton Sound. The horizontally averaged mean model bias is 0.28 psu saltier than the data, and the RMS difference is 1.19 psu.

Near-bottom temperature measurements reveal the contour of the cold pool ($T < 2$ °C), which approximately follows the 70 m isobaths (Fig. 9c). Warmer temperatures were found seaward and shoreward of this feature. Highest temperatures were observed in Norton Sound. The model near-bottom temperatures exhibit a general correspondence with measurements, however, the model exhibits a tendency to extend the cold pool farther inshore (e.g. along the 30 m isobath) than is observed. The mean bias is −0.72 °C (note this includes the bias in Norton Sound, which is outside the standard 62°N of the BTS data shown in Fig. 6). The *r* value is 0.78 and the RMS difference is 1.75 °C.

Spatial patterns of the measured near-bottom salinities are similar to those of the near-surface measurements, and the model exhibits a general correspondence with those patterns (Fig. 9d). The model has only a slight (−0.03 psu) fresh bias, with an *r*-value of 0.67 and an RMS difference of 0.9 psu. Generally, the fit (skill) of the model is superior in the case of near-bottom measurements. This is likely due in part to the slower evolution of near-bottom properties; higher frequency events which occur near the surface are less likely to be captured well by a model which does not explicitly assimilate data.

Zonal and meridional cross-sections of temperature and salinity, passing through station M2, are likewise compared with data (Fig. 10). As in Fig. 5, the model tends to have a shallower mixed layer than is observed. A cold pool signature is clearly evident; as in Fig. 9 the model places its center slightly inshore of the observed location. Temperatures and salinities rise seaward toward the shelf break in both the model and the observations. The halocline is stronger in the model, and the model tends to be fresher. As noted for the near-surface and bottom plots, a superior fit with temperature and salinity data is found at depth, where measured seasonal changes are smaller in amplitude than at the surface. For temperature, the mean model bias for the zonal (meridional) section is −0.4 (−1.08) °C, the *r*-value is 0.88 (0.79) and RMS difference is 1.23 (1.54) °C. For salinity, the mean model bias for the zonal (meridional) section is −0.12 (−0.1) psu, the *r*-value is 0.83 (0.05) and the RMS difference is 0.34 (0.46) psu. The poor fit of the model to the meridional salinity data is primarily due to an excess of surface freshwater in the model near the latitude of M2, where data are from early June (Fig. 3). This freshwater excess is apparent in the zonal section as well.

3.2. Comparison of projected futures

3.2.1. Effects of downscaling

Following the official standards of the latest IPCC report (IPCC et al., 2013), here we use the term “projections” to denote our three possible realizations of future conditions. Before analyzing the downscaled projections, it is useful to compare IPCC model output with downscaled results on specific dates (Fig. 11). This illustrates the different spatial resolution among the IPCC models themselves, and gives us a sense of the greater spatial resolution obtained through model downscaling. We compare temperature output at 20 m depth from each of the three chosen IPCC projections with their corresponding regional model result on two different dates (July 15 of 2010 and 2039). Note in particular how the Bering10K model resolves the cold pool, whereas none of the IPCC

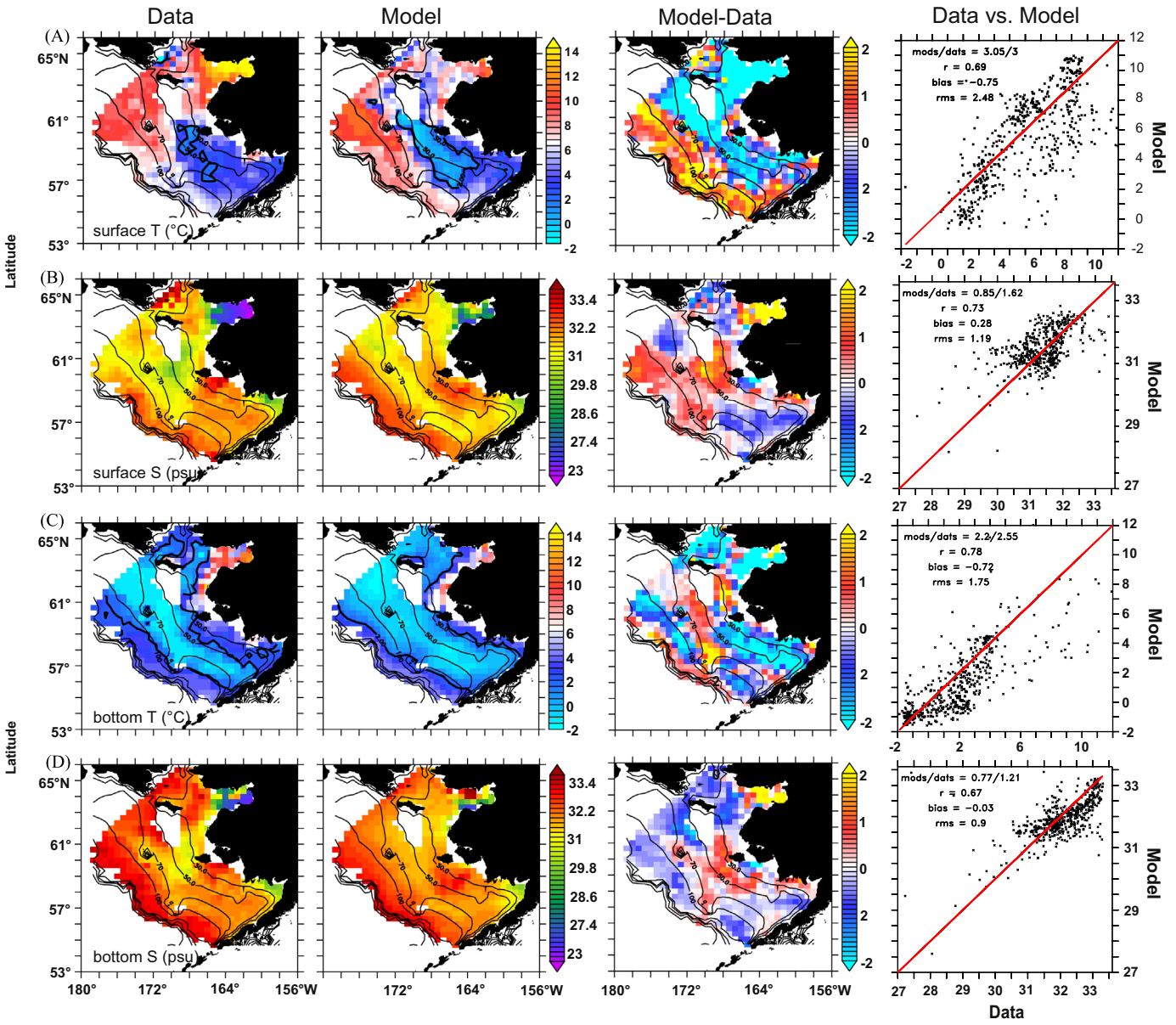

Fig. 9. Comparison of measured vs. modeled temperatures (T, °C) and salinities (S, psu) during June–August 2010. Measurements are from the trawl survey of Cokelet (2016). Shown in each row are: the data; corresponding values from the model, sampled at the same locations and times as the measurements; model minus data values; scatterplot of data vs. model, with statistics as in Fig. 8. (a) near-surface (5 m) temperatures; (b) near-surface (5 m) salinities; (c) bottom temperatures; (d) bottom salinities. Isobaths (m) are shown as thin black lines. The 2 °C isotherm (boundary of the "cold pool") is highlighted as a thick black line in a and c.

models are able to replicate this feature, given their limited horizontal and vertical resolution. More generally, note that downscaling results in different spatial means – that is, a spatially averaged version of the Bering10K output is not identical to the IPCC model output which was used as forcing. This is expected due to the additional physical processes (including tidal mixing) resolved by the finer-scale model.

3.2.2. *Projected spatial averages for biophysical domains*

Here we focus on two of the standard BEST-BSIERP bio-regions (Fig. 1): (1) the south middle shelf (bio-region 3) and (2) the midnorth middle shelf (bio-region 6). These encompass the biophysical moorings M2 and M4, respectively. Annual means of the vertically averaged temperature exhibit considerable interannual variability in these bio-regions over both hindcast and forecast simulations (Fig. 12). As in Figs. 5 and 6, the strong cold-to-warm "regime shift" of 1976, and the transition from warm to cold years around 2005, are both evident in the hindcast.

In the time range where hindcast and forecast simulations overlap (2003–2010), we see how the free-running (without data assimilation) IPCC models differ from the actual hindcast yearly values (due to intrinsic variability), and from each other (due to both intrinsic variability and model biases). This underscores the need to examine long-term trends within a single model projection. The 10-year running means of the temperature series indicate a range of 0.5–1.5 °C increase over present conditions by 2040 for both the south middle shelf and midnorth middle shelf domains. A slightly greater envelope of uncertainty is present in the midnorth case, possible reflecting ice-albedo feedback effects.

Annual mean ice coverage in the south middle shelf exhibits substantial interannual variability during 2010–2030, but reduced variability thereafter as the mean ice cover disappears (Fig. 12). The ECHO-G model, which exhibits the coldest projection, also

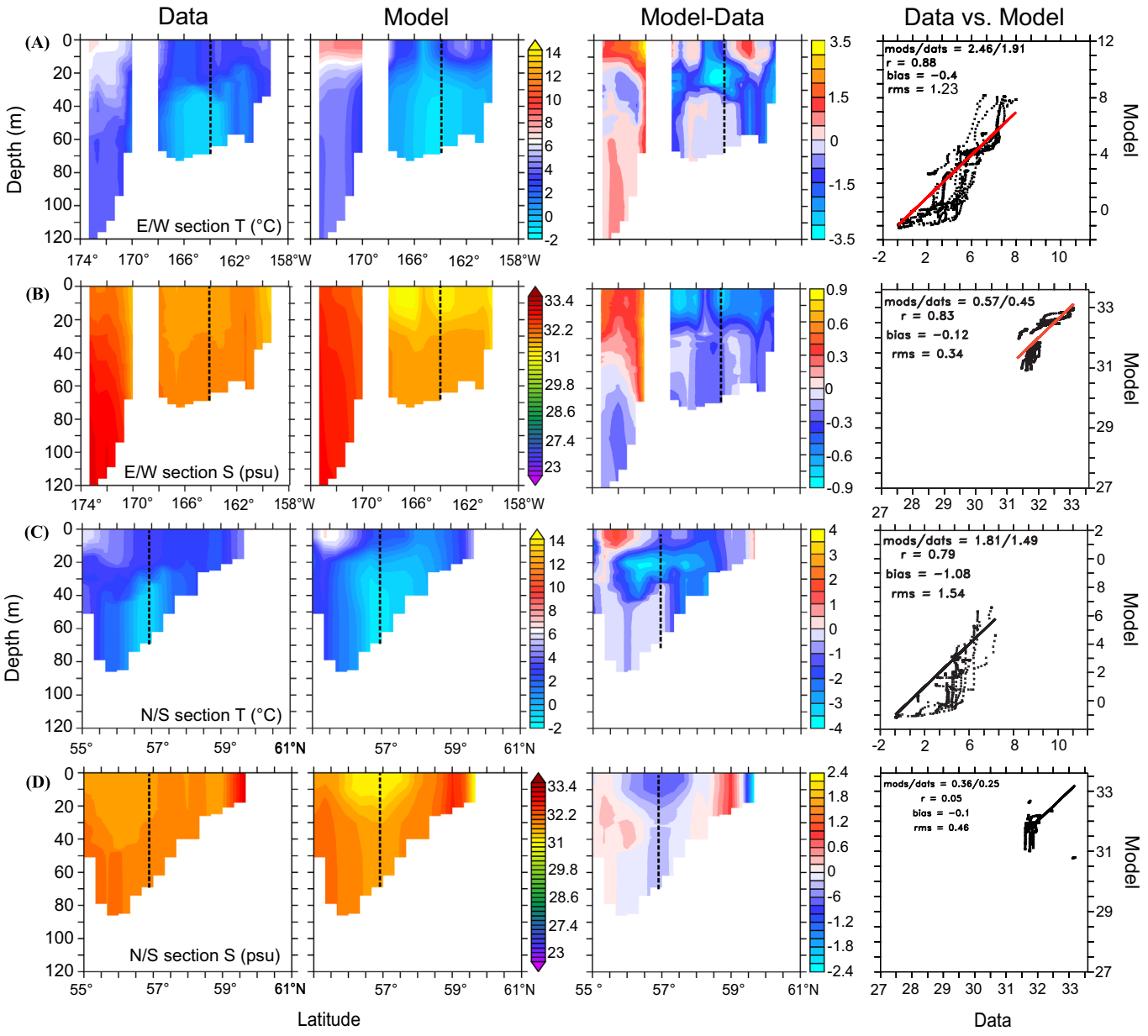

Fig. 10. As in Fig. 9 for an east–west section (a,b) and a north–south section (c,d), both through station M2 (56.87°N, 164.05°W; see Fig. 1). Dashed lines indicate the location of the M2 mooring.

exhibits the largest ice cover. Ice coverage on the midnorth middle shelf exhibits even greater interannual variability than the south middle shelf; again, the ECHO-G model exhibits the largest ice cover. Despite gradual warming of the region, there is substantial decadal average ice coverage out to 2040; however, note the increasing incidence of individual years with no ice in the south, and a few years with no ice in the midnorth, especially in the MIROC3.2 results.

3.2.3. *Projected spatially averaged temperatures for the groundfish survey domain*

Fisheries managers use the spatially averaged bottom temperature of the BTS (south of 62°N, between 20 m and 200 m isobaths on the Bering Sea shelf) as a metric of average yearly conditions. Specifically, they have found it useful to classify past conditions on the shelf as "cold" ($T < 1.75$ °C), "neutral" ($1.75 < T < 2.5$ °C) and "warm" ($T > 2.5$ °C), based on that metric (S. Kotwicki, NOAA/AFSC, pers. comm.). It is instructive to calculate the areal mean of bottom temperatures in this area for July (the temporal midpoint of that survey) of each projected year, to get a sense of how this metric might change in the future. Fig. 13 illustrates this metric for each of the IPCC-driven downscaling projections. As shown, there is continued variability at interannual-to-interdecadal scales all the way through 2040. The average trend (regression line) clearly suggests warming in each case, with the CGCM3-t47 projection yielding the greatest interdecadal warming overall. By this metric, there will be an increase in the number of warm years in the future, and a decrease in the number of cold years. Significance testing (*F*-test) for the regression lines, with degrees of freedom based on the autocorrelation in each series, indicates significance at the 99, 96 and 89 percent levels for the CGCM3-t47, MIROC3.2, and ECHO-G cases, respectively.

Further insight regarding these projected futures is obtained by averaging the spatial patterns separately for all years with projected warm, neutral, and cold July bottom temperatures (Fig. 14). In "cold"

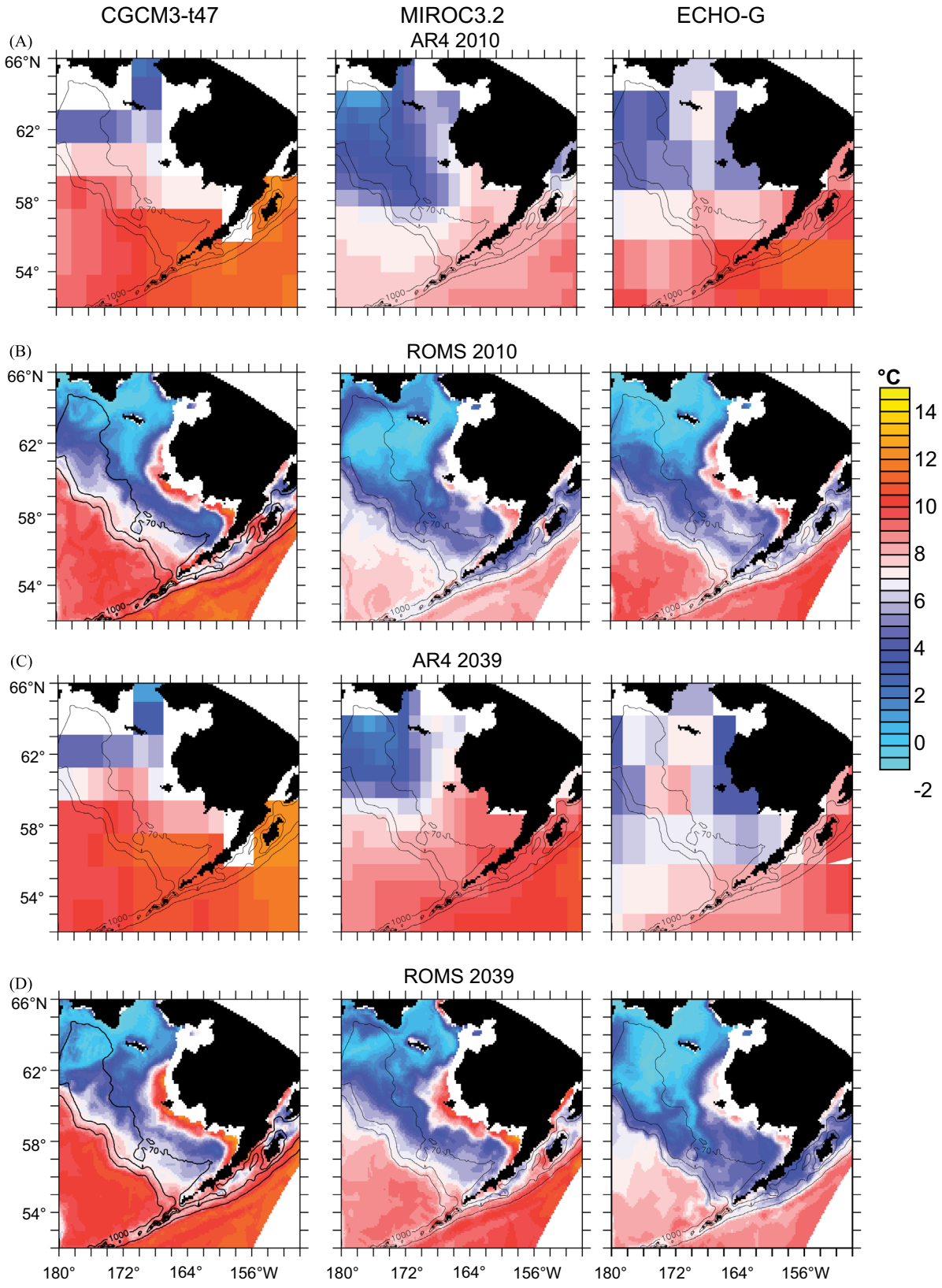

Fig. 11. Comparison of AR4 global models (CGCM3-t47, MIROC3.2, ECHO-G) (a,c) with the companion, downscaling regional Bering10K model (b,d), showing their resolution of temperature (°C) at 20 m depth for July 15, 2010 (a,b) and July 15, 2039 (c,d).

and "neutral" years, the cold pool extends to ~56 °N, whereas in "warm" years it extends only to ~58 °N. Note in particular how, despite higher temperatures in the south during warm years, a cold environment is on average retained in the north during all years. These patterns are in fact very similar to warm–neutral–cold averages derived from the hindcast run (not shown).

Decadal histograms of the spatially averaged bottom temperatures for the BTS domain (Fig. 15) summarize the values plotted in

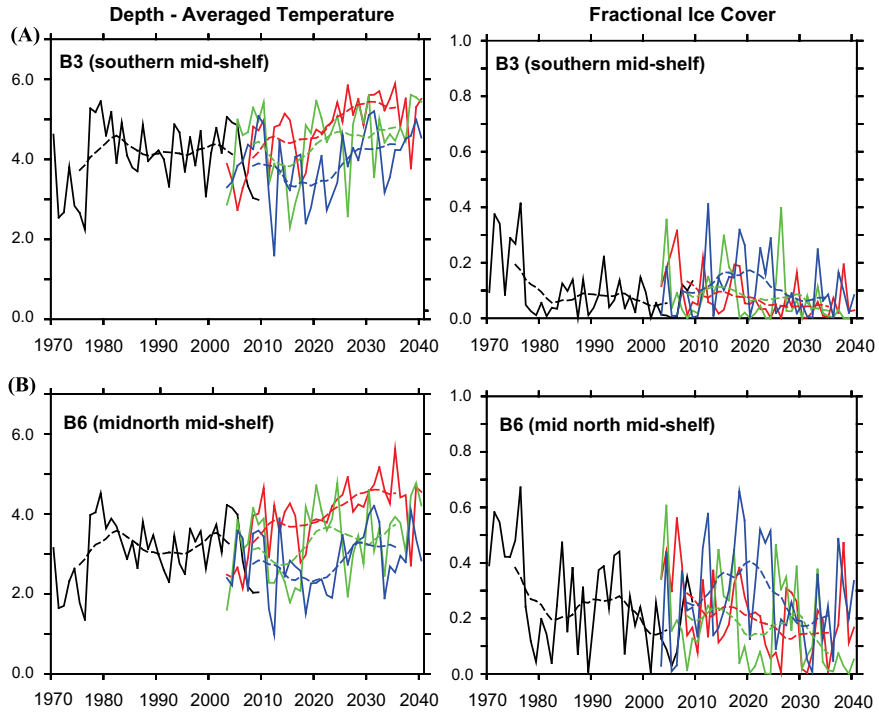

Fig. 12. Depth-averaged temperature (°C) (left) and areal (fractional) ice cover (right) from the Bering10K model for: (a) the south middle shelf domain (bio-region 3); (b) the midnorth middle shelf domain (bio-region 6). Solid lines are annual averaged values downscaled from the CORE/CFSR hindcast (black), CGCM3-t47 projection (red), MIROC3.2 projection (green) and ECHO-G projection (blue). Dashed lines show decadal running means.

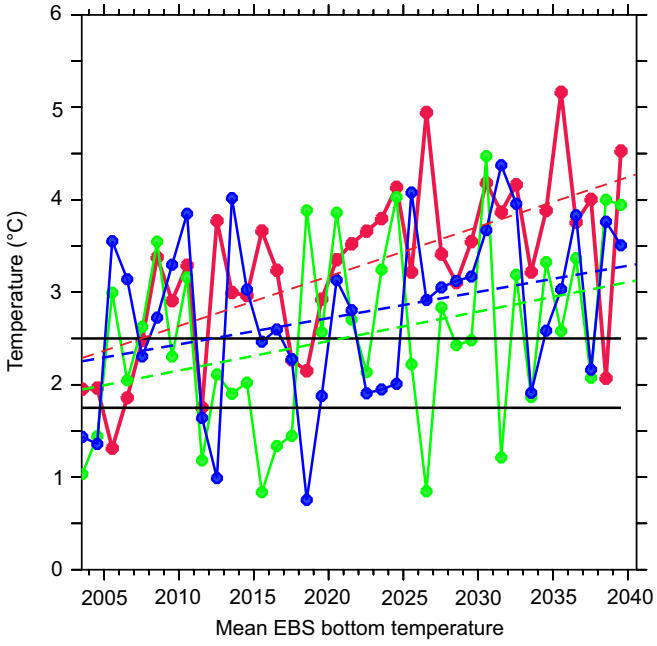

Fig. 13. Mean bottom temperature (°C) in July on the Southeastern Bering Sea shelf over the area of the standard groundfish survey (between 20 and 200 m isobaths, south of 62 N) for each of the three downscaled projections of future climate. Red = CGCM3-t47, Green = MIROC3.2, Blue = ECHO-G. Dashed lines indicate mean trend for each case. Solid lines show standard definitions presently used by fisheries managers for "cold" (< 1.75), "neutral" (1.75–2.5) and "warm" (> 2.5) years.

Fig. 13 for the first (2003–2012) vs. the last (2030–2039) decade of the downscaled projections. Observed conditions for 2001–2010 (from values shown in Fig. 6) are plotted for comparison. By this measure, the CGCM3-t47 run exhibits the closest distribution to present conditions, whereas the MIROC3.2 and ECHO-G distributions exhibit more variability than is presently observed. All three of the projections exhibit a substantial positive (greater than 1 °C) shift of the most likely temperature by 2030–2039. Using these histograms are rough measures of probability, the data indicate a 50% chance of a "warm" year during 2001–2010, and each of the projections suggests a 30–50% increase in the likelihood of "warm" years from 2003–2012 to 2030–2039 (CGCM3-t47: 40 to 90%; MIROC3.2: 40 to 70%; ECHO-G: 50 to 80%).

3.2.4. *Projected decadal averages*

Decadal averaging of the model domain affords a useful method to compare average present with average future conditions. As in 3.2.3, we have chosen a 10-year time interval, sufficient to eliminate much of the interannual variability (owing to low-frequency variability – or more precisely, a red spectrum – there is no specific averaging window one may choose to eliminate *all* of the interannual variability). In Fig. 16 we compare the means under "recent" (2003–2012) conditions, produced using the three IPCC models. Specifically we compare SST, areal ice cover, the "alongshore" (northwestward) component of wind stress, and large crustacean zooplankton (LCZ). The SST patterns are similar among the three models, with the CGCM3-t47 model yielding the warmest conditions overall. Areal ice cover is likewise similar in all three cases, and each exhibits southeastward wind stress in the northern half of the domain. While spatial patterns are similar, the MIROC3.2 model produces significantly higher biomass of LCZ along the outer shelf.

To obtain a sense of future trends, we calculate a ten year average during 2030–2039, and subtract the ten year average during 2003–2012 (Fig. 17). CGCM3-t47 and MIROC3.2 each exhibit a spatially broad rise in SST, exceeding 1.5 °C in some areas (Fig. 17a), with a corresponding loss of sea ice cover (Fig. 17b). In CGCM3-t47, the rise in SST is greatest along the middle shelf and shelf break. ECHO-G exhibits only modest changes in SST or ice. All three of the models exhibit substantial changes in the

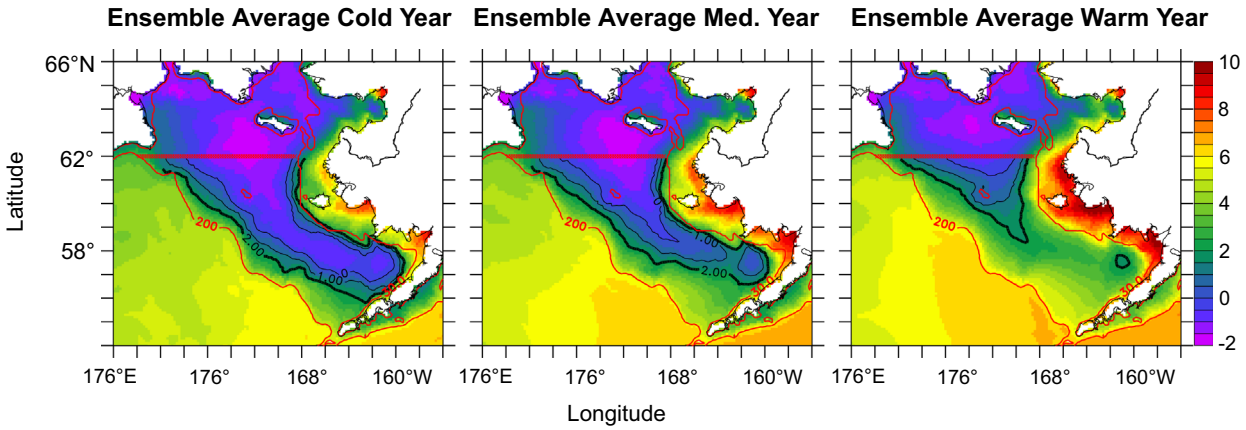

Fig. 14. Ensemble average spatial patterns for: (a) "cold" ($T < 1.75$) , (b) "neutral" ($1.75 < T < 2.5$) and (c) "warm" ($T > 2.5$) years, as defined by the spatially averaged bottom temperature (°C). Thin red lines indicate the eastern and western boundaries of the areal average (i.e. the 30 m and 200 m isobaths); thick red line indicates the northern boundary of the areal average (62°N).

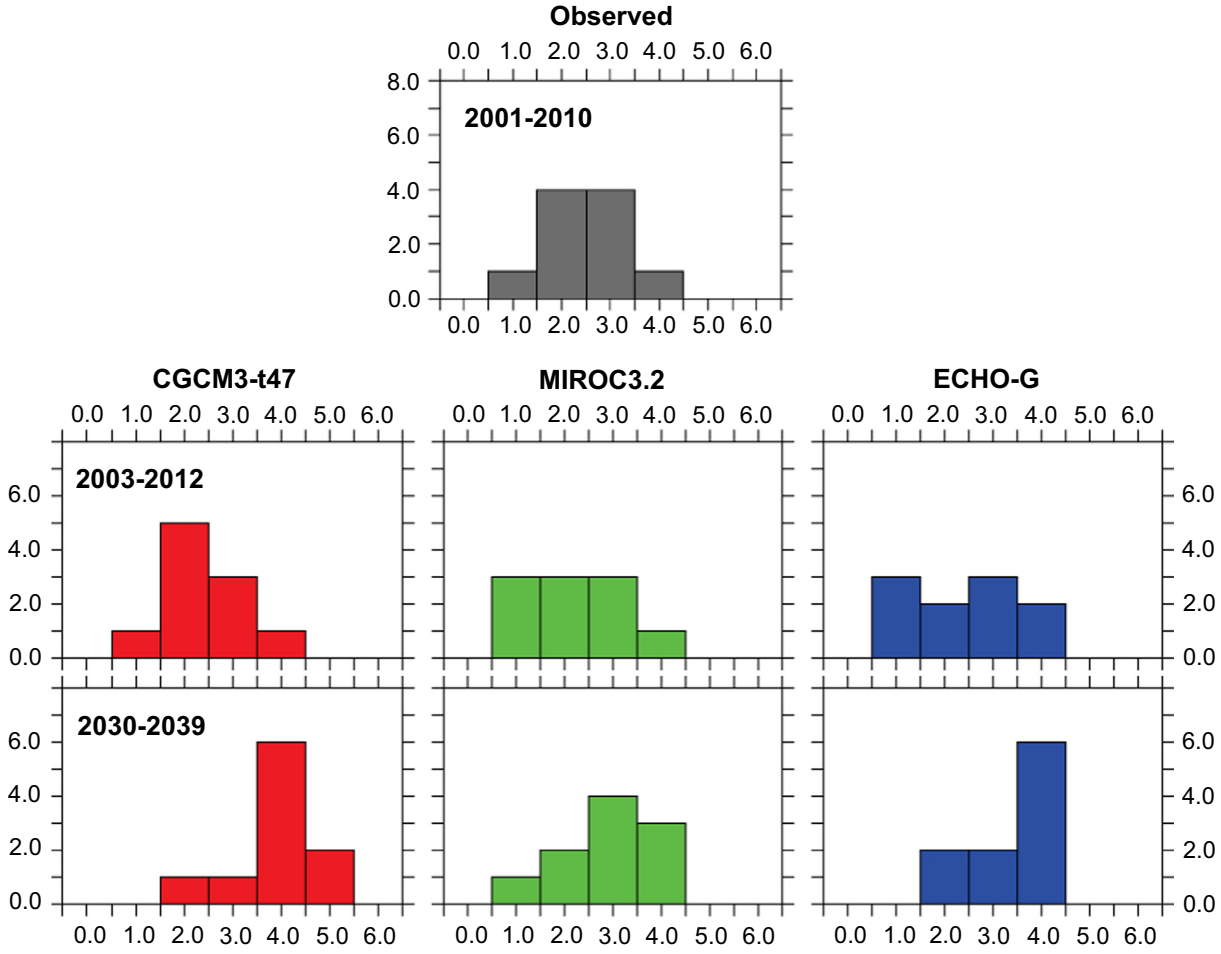

Fig. 15. Histograms of observed (from Fig. 6) and projected (from Fig. 13) mean bottom temperatures. Upper figure displays observed values for 2001–2011. Middle row displays downscaled AR4 model values for 2003–2012; bottom row displays downscaled AR4 values for 2030–2039.

northwestward component of wind stress (Fig. 17c). The spatial patterns are different, but each entails a trend toward more northwestward stress over most of the shelf. CGCM3-t47 also includes a trend to more southeastward stress in the northwestern Bering Sea. LCZ decrease on the outer shelf and increase on the inner shelf in CGCM3-t47 and MIROC3.2, whereas they increase on the outer shelf and Gulf of Alaska in ECHO-G (Fig. 17d).

4. Discussion

4.1. Sources of regional ocean variability at multiple time scales

Here we consider the sources of observed variability from short (seasonal) to long (decadal) timescales in the model. We explore this through a coherence analysis of seasonal anomaly time series

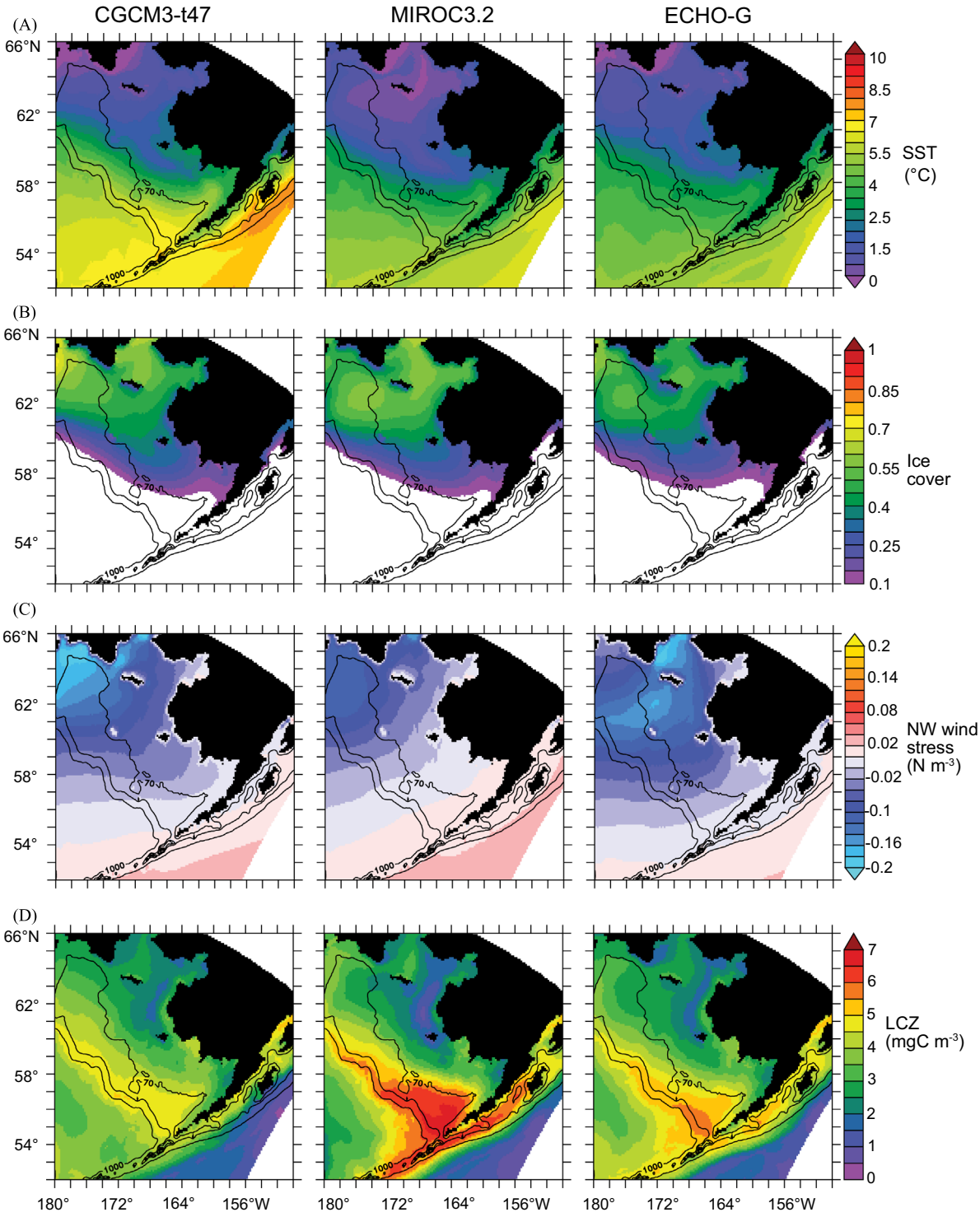

Fig. 16. Ten-year means from the three AR4-driven downscaling simulations of recent (2003–2012) conditions. Results from CGCM3-t47 (left), MIROC3.2 (center) and ECHO-G (right) are shown. Properties shown are: (a) SST (°C); (b) fractional ice cover; (c) northwestward wind stress (N m^{-2}); and (d) large crustacean zooplankton (LCZ) (mg C m^{-3}).

of: (1) forcing terms vs. oceanic response in the regional model; and (2) pairs of biophysical variables in the regional model. In each case these time series represent spatial averages over the biophysical domains shown in Fig. 1. Rather than including all possible pairs (a task more suited to multivariate techniques, as in H2013), we focus a few relevant combinations which yielded strongest and weakest coherence: (1) depth-averaged ocean temperature vs. air temperature, northwestward winds, and shortwave radiation; (2) depth-averaged ocean temperature vs. ocean euphausiids. In each case similar results were obtained using CORE/SODA-driven hindcast simulations as were found using forecast/projection simulations. Here we focus on the CORE/SODA-driven hindcast, as it is a multidecadal reanalysis which included both atmospheric and oceanic data.

4.1.1. Ocean temperature coherence with physical forcing

Danielson et al. (2011a, 2011b) suggested that net surface heat flux and southeast–northwest advection contribute in equal measure

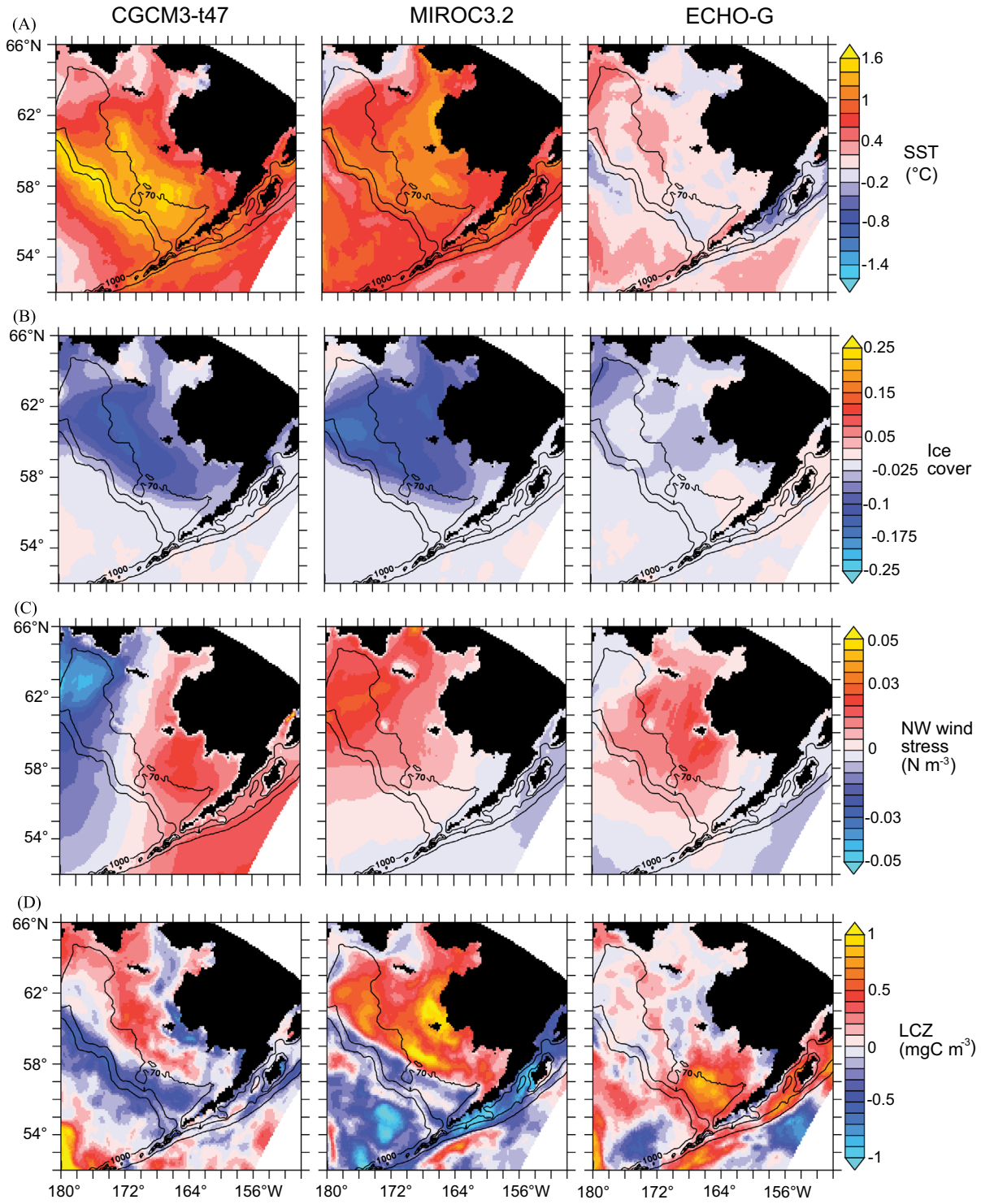

Fig. 17. Differences between ten-year means from the three IPCC-driven downscaling simulations of future (2030–2039) minus recent (2003–2012) conditions. Results from CGCM3-t47 (left), MIROC3.2 (center) and ECHO-G (right) are shown. Properties shown are: (a) SST (°C); (b) fractional ice cover; (c) northwestward wind stress (N m^{-2}); and (d) large crustacean zooplankton (LCZ) (mg C m^{-3}).

to the year-to-year variability of heat content in the central shelf waters of the Bering Sea. Our coherence analysis supports the idea that both of these terms are strong contributors to the interannual heat budget. We find that mid-shelf ocean temperature is coherent with air temperature (and by inference, sensible heat flux) on annual to decadal scales (Fig. 18a). Note the corresponding rise of both series during the large regime shift of 1976, which contributes to the low-frequency coherence. A similar coherence at slightly reduced amplitude (not shown) was found between mid-shelf temperature and specific humidity of the atmosphere (which is strongly correlated with the air temperature in our forcing data).

Shortwave radiation is a dominant surface forcing term in the annual heat budget of the Bering Sea shelf (Reed, 2003). Given this fact, we initially expected to find a strong coherence between shortwave and ocean temperature anomalies. The obvious lack of coherence between these two series (Fig. 18b) suggests that the

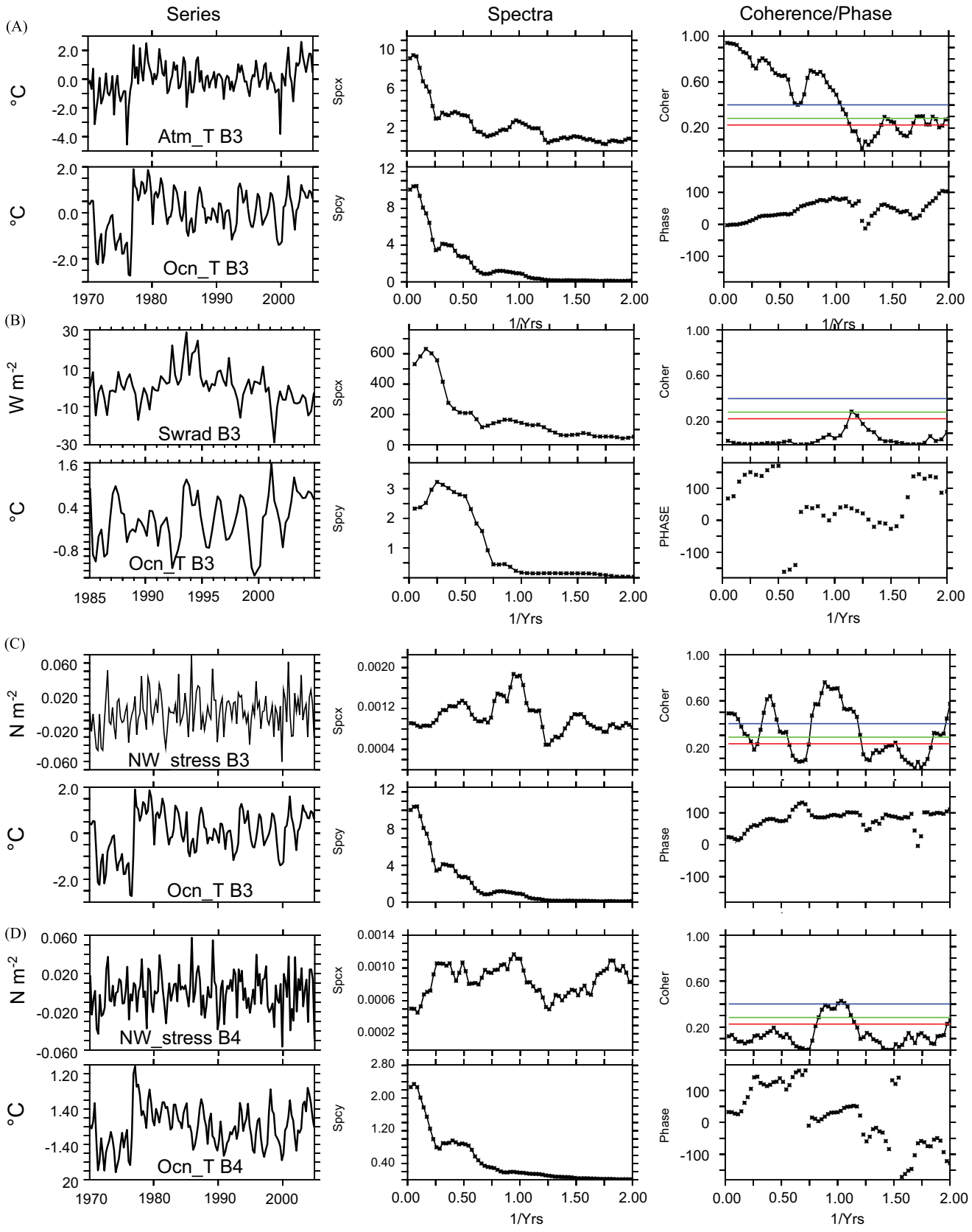

Fig. 18. Coherence of seasonal anomalies from the CORE/SODA-driven regional simulation. For each chosen pair of variables we present: the seasonal anomaly series for each variable (left column); the frequency spectrum for each variable (middle column); the coherence and phase for that pair of variables (right column; upper panel is coherence and lower panel is phase lag in degrees). Positive phase lag indicates that upper variable leads the lower variable. Red, green, blue lines (black dashed, grey dashed, solid black in print version) indicate 90%, 95%, 99% confidence levels, respectively, for coherence significantly different than zero. Pairs shown here are as follows: (a) atmospheric temperature at 10 m above the sea surface (Atm_T, °C) compared with the vertically averaged (top 100 m) ocean temperature (Ocn_T, °C) in bio-region 3 (south middle shelf); (b) shortwave radiation (sward, W m^{-2}) compared with Ocn_T in bio-region 3 (note the different time axes in this figure, relative to a,c,d); (c) alongshelf (northwest/southeast) windstress (NW_stress, N m^{-2}) compared with Ocn_T in bio-region 3; and (d) NW_stress compared with Ocn_T, bio-region 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

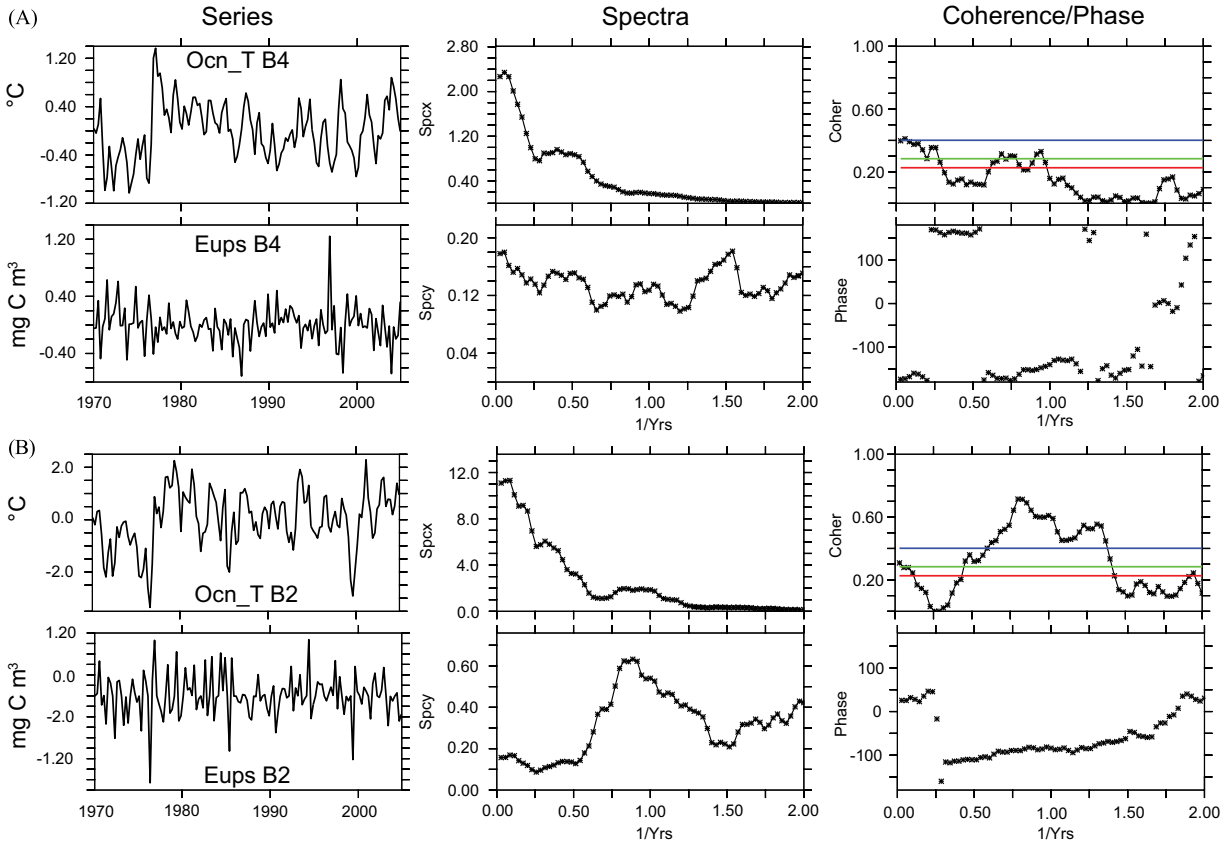

Fig. 19. As in Fig. 18, comparing seasonal anomalies for vertically averaged (top 100 m) ocean temperature (Ocn T, °C) with seasonal anomalies for euphausiid biomass (Eups, mg C m^{-3}) in: (a) bio-region 4 (south outer shelf); and (b) bio-region 2 (south inner shelf).

shortwave anomalies are too small to be significant, relative to other forcing terms. While peak anomalies have magnitudes as high as 30 W m^{-2}, the RMS value is closer to ~10 W m^{-2}, which is small compared to the peak summer input of ~200 W m^{-2}. Similarly small variations are found in the three forecast/projection time series (not shown). Bond (unpublished) similarly found only a few percent interannual variation in cloudiness for the Bering Sea during years 2003–2008. It is recognized that the clouds of the region (particularly low stratus and fog) are a challenge to simulate properly, and so this aspect of our results is tentative. Nevertheless, our findings suggest a minimal role for shortwave radiation in climate change on the Bering Sea shelf over the next few decades.

Alongshelf (southeast-to-northwest) wind stress is found to be coherent with ocean temperature at annual and most interannual frequencies at the southern mid-shelf (bio-region 3, Fig. 18c); a similar result was found for the south inner shelf (bio-region 2, not shown). Note how this coherence with air temperature and wind stress falls off at higher frequencies; as noted in Danielson et al. (2011b), the ocean response on the central shelf is best considered a seasonal integral of atmospheric forcing and advection. On the outer shelf there is less coherence with the wind stress (bio-region 4, Fig. 18d); this is expected due to the greater impact of local wind forcing in shallower seas and shelves. While not specifically analyzed here, a strong coherence between wind stress and broad-scale flow patterns is expected over the Bering shelf (Danielson et al., 2012a, 2012b).

4.1.2. *Euphausiids coherence with ocean temperature*

In observational studies such as Coyle et al. (2011), it has been noted that large crustacean zooplankton production (euphausiids and neocalanus) on the outer shelf tends to decrease at higher temperatures. This effect has been noted in past analyses of the NPZ model results (Gibson et al., unpublished; H2013), and is suggested in the 10-year summaries of forecasts shown in Fig. 17. Here we explicitly look for coherence between euphausiid biomass and ocean temperature on the south outer shelf (bio-region 4, Fig. 19a) and the south inner shelf (bio-region 2, Fig. 19b). Note how the spectra for euphausiids are whiter (more evenly distributed across frequency bands) than the spectra for temperature on the outer shelf (Fig. 19a). Significant coherence appears at sub-annual bands, with the ±180° phase lag indicating a negative correlation at those frequencies, consistent with large crustacean zooplankton production decreasing as temperature increases. By contrast, on the inner shelf, euphausiids exhibit a spectral peak near annual frequencies (Fig. 19b). Since we have used seasonal anomalies for the analysis, this likely reflects year-to-year changes in both: (1) their total annual production; and (2) the seasonal timing (phenology) of that total. The negative phase lag values (euphausiids leading temperature by ~90° [3 months] at annual scales) are curious; however, given that we are comparing specific frequency bands, it may be suggesting that inner shelf euphausiids are responding positively to higher temperatures from much earlier in the year (that is, a 270° [9-month] lag, as one might find with higher temperatures in early spring promoting greater euphausiid biomass in late fall). The contrasting response of these large crustacean zooplankton to temperatures on the inner vs. outer shelves was noted in H2013; it was inferred that the greater interannual variability of ice cover on the inner shelf may contribute to this contrast.

4.2. *Relative significance of these results and potential improvements*

While two out of three of our regional projections indicate significantly warmer conditions in the Bering Sea by the 2030s,

there remains considerable uncertainty regarding future states. We must emphasize that the three IPCC realizations chosen are only a small sampling of the variability seen in the larger set of all such realizations carried out under AR4. The mean trajectory of the AR4 models is towards a warmer future summer Bering Sea SST (~1.0–1.5° rise between 2010 and 2040; Mueter et al., 2011), and our simulations thus far suggest that a regionally downscaled version of each would be likely to follow this trend, while better resolving the cold pool (Fig. 11). A larger set of direct downscaling runs is now underway; also, multivariate methods of summarizing regional model output (discussed in H2013) might offer a way to compactly downscale the larger set IPCC projections. A future publication will explore this possibility for "hybrid" statistical/model downscaling.

While observed bottom temperatures (and especially their areal average, Fig. 6) were relatively well replicated by our model, and stratification was more realistic than in H2013 (Fig. 5), the upper ocean still has significant bias (Fig. 9). Our relatively shallow mixed layer (Fig. 5) did not significantly bias the onset of seasonal ice (Fig. 7), but presumably could still lead to overestimates of future SST. Further refinements to the ice and heat flux algorithms, as well as greater vertical resolution, could reduce this shallow bias. Improved estimates/projections of recent/future coastal runoff are also needed; presently a monthly climatology is used for years beyond 2004.

The NPZ model is undergoing updates to better reflect ecosystem understanding gained during the BEST-BSIERP study. These include improved rates and algorithms for light attenuation by chlorophyll and for euphausiid dynamics, as new evidence suggests that euphausiids in the Bering Sea live multiple years and consume appreciable detritus (e.g. Ressler et al., 2014). The work reported here includes only the lower trophic level dynamics of the system; other simulations couple these with a size- and age-structured fish model (Ortiz et al., submitted) to examine both top-down and bottom-up control of zooplankton. These modifications could ultimately change the details of coherence relationships between modeled temperature and modeled large crustacean zooplankton; however, Coyle et al.'s (2011) observed pattern on the outer shelf (more abundant large crustacean zooplankton under cold conditions) has thus far persisted in the revised models.

5. Summary and conclusions

Three IPCC models from AR4 were used as physical forcing to drive a regional model which includes both physical and biological elements of the Bering Sea. This set was chosen based on their fidelity to represent historical conditions in the Bering Sea and the northeastern Pacific. Although each of the downscaled projections indicates warming of the Bering Sea, the interannual and interdecadal details of this trend vary considerably among the three realizations. In each case, the magnitude of presently observed interannual variability of bottom temperatures and ice cover is found in the models to be maintained out to at least 2040, with a steadily increasing probability of warm years with less ice on the southern shelf. The overall trends indicate warmer temperatures and a retreat of ice in the southeastern Bering Sea, but continued ice cover in the northeastern Bering Sea. Sensitivity analyses suggest both increasing air temperature and northward wind stress as primary drivers of higher water column temperatures. Based on currently available models, changes in shortwave radiation are not likely to have a significant role in this warming. Warming trends on the outer shelf may lead to decreased production of large crustacean zooplankton at that location, but may increase such production on the inner shelf.

Acknowledgements

We thank the reviewers and S. Danielson for useful discussions which strengthened the manuscript. This research is contribution number 4183 from NOAA/Pacific Marine Environmental Laboratory, and contribution ecoFOCI-0851 to NOAA's Ecosystems Fisheries Oceanography Coordinated Investigations. This publication is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA cooperative agreement NA10OAR4320148, Contribution no. 2435. The research was generously supported by two grants from NSF: "Bering Sea Ecosystem Study" (NSF-0732534, NSF-0732538) and "Impact of sea-ice on bottom-up and top-down controls of crustacean zooplankton and the mediation of carbon and energy flow in the eastern Bering Sea" (NSF-1107250, NSF-1107203), as well as the North Pacific Research Board (NPRB) sponsored Bering Sea Integrated Ecosystem Research Program (BSIERP projects B52 and B70) and NOAA's North Pacific Climate Regimes and Ecosystem Productivity programs. The Arctic Region Supercomputing Center provided additional computational resources and technical support to this project. This is BEST-BSIERP Bering Sea Project publication number 172 and NPRB publication number 557.

References

Budgell, W.P., 2005. Numerical simulation of ice-ocean variability in the Barents Sea region: towards dynamical downscaling. Ocean Dyn. 55, 370–387.
Brown, Z.W., Arrigo, K.R., 2013. Sea ice impacts on spring bloom dynamics and net primary production in the Eastern Bering Sea. J. Geophys. Res. Oceans 118, 43–62. http://dx.doi.org/10.1029/2012JC008034.
Carton, J.A., Giese, B.S., 2008. A reanalysis of ocean climate using simple ocean data assimilation (SODA). Mon. Weather Rev. 136, 2999–3017.
Cheng, W., Curchitser, E., Ladd, C., Stabeno, P.J., Wang, M., 2014. Influences of sea ice on the eastern Bering Sea: NCAR CESM simulations and comparison with observations. Deep-Sea Res. II 109, 27–38. http://dx.doi.org/10.1016/j.dsr2.2014.03.002.
Coachman, L.K., 1986. Circulation, water masses, and fluxes on the southeastern Bering Sea shelf. Cont. Shelf Res. 5 (1–2), 23–108.
Cokelet, E.D., 2016. 3-D water properties and geostrophic circulation on the eastern Bering Sea shelf. Deep-Sea Res. II, Topical Studies 134, 65–85.
Coyle, K.O., Pinchuk, A.I., Eisner, L.B., Napp, J.M., 2008. Zooplankton species composition, abundance and biomass on the eastern Bering Sea shelf during summer: the potential role of water-column stability and nutrients in structuring the zooplankton community. Deep-Sea Res. II 55, 1775–1791.
Coyle, K.O., Eisner, L.B., Mueter, F.J., Pinchuk, A.I., Janout, M.A., Cieciel, K.D., Farley, E.V., Andrews, A.G., 2011. Climate change in the southeastern Bering Sea: impacts on pollock stocks and implications for the oscillating control hypothesis. Fish. Ocean. 20 (2), 139–156.
Dai, A., Qian, T., Trenberth, K.E., Milliman, J.D., 2009. Changes in continental freshwater discharge from 1948–2004. J. Clim. 22, 2773–2791.
Danielson, S., Curchitser, E., Hedstrom, K., Weingartner, T., Stabeno, P., 2011a. On ocean and sea ice modes of variability in the Bering Sea. J. Geophys. Res. 116, C12034. http://dx.doi.org/10.1029/2011JC007389.
Danielson, S., Eisner, L., Weingartner, T., Aagaard, K., 2011b. Thermal and haline variability over the central Bering Sea shelf: seasonal and inter-annual perspectives. Cont. Shelf Res. . http://dx.doi.org/10.1016/j.csr.2010.12.010
Danielson, S., Weingartner, T., Aagaard, K., Zhang, J., Woodgate, R., 2012a. Circulation on the central Bering Sea shelf, July 2008 to July 2010. J. Geophys. Res. 117, C10003. http://dx.doi.org/10.1029/2012JC008303.
Danielson, S., Hedstrom, K., Aagaard, K., Weingartner, T., Curchitser, E., 2012b. Wind-induced reorganization of the Bering shelf circulation. Geophys. Res. Lett. 39, L08601. http://dx.doi.org/10.1029/2012GL051231.
Donlon, C.J., Martin, M., Stark, J.D., Roberts-Jones, J., Fiedler, E., Wimmer, W., 2012. The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sens. Environ. 116. http://dx.doi.org/10.1016/j.rse.2010.10.017.
Ebbesmeyer, C.C., Cayan, D.R., McLain, D.R., Nichols, F.H., Peterson, D.H., Redmond, K.T., 1991. 1976 step in the Pacific climate: forty environmental changes between 1968–1975 and 1977–1984. In: Betancourt J.L., Tharp V.L. (Eds.). Proceedings of the Seventh Annual Climate (PACLIM) Workshop, April 1990. California Department of Water Resources. Interagency Ecological Studies Program Technical Report 26, pp. 115–126.
Flato, G.M., Boer, G.J., Lee, W.G., McFarlane, N.A., Ramsden, D., Reader, M.C., Weaver, A.J., 2000. The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. Clim. Dyn. 16, 451–467.
Flato, G.M., Boer, G.J., 2001. Warming asymmetry in climate change simulations. Geophys. Res. Lett. 28, 195–198.

Flato, G.M., Hibler III, W.D., 1992. Modelling pack ice as a cavitating fluid. J. Phys. Oceanogr. 22, 626–651.
Gibson, G.A., Spitz, Y.H., 2011. Impacts of biological parameterisation, initial conditions, and environmental forcing on parameter sensitivity and uncertainty in a marine ecosystem model for the Bering Sea. J. Mar. Syst. 88, 214–231.
Gibson, G.A., Hermann, A.J., Hedstrom, K. Curchitser, E.N., Response of euphausiid production to 'cold' and 'warm' years in the Bering Sea, unpublished results.
Hare, S.R., Mantua, N.J., 2000. Empirical evidence for North Pacific regime shifts in 1977 and 1989. Prog Ocean 47 (2–4), 103–145.
Haidvogel, D.B., Arango, H., Budgell, W.P., Cornuelle, B.D., Curchitser, E., Di Lorenzo, E., Fennel, K., Geyer, W.R., Hermann, A.J., Lanerolle, L., Levin, J., McWilliams, J.C., Miller, A.J., Moore, A.M., Powell, T.M., Shchepetkin, A.F., Sherwood, C.R., Signell, R.P., Warner, J.C., Wilkin, J., 2008. Regional ocean forecasting in terrain-following coordinates: model formulation and skill assessment. J. Comput. Phys. 227, 3595–3624.
Hermann, A.J., Gibson, G.A., Bond, N.A., Curchitser, E.N., Hedstrom, K., Cheng, W., Wang, M., Stabeno, P.J., Eisner, L., Cieciel, K.D., 2013. A multivariate analysis of observed and modeled biophysical variability on the Bering Sea shelf: multi-decadal hindcasts (1970–2009) and forecasts (2010–2040). Deep Sea Res. II 94, 121–139. http://dx.doi.org/10.1016/j.dsr2.2013.04.007.
Hibler, W., 1979. A dynamic-thermodynamic sea-ice model. J. Phys. Ocean. 9, 815–846.
Hunt Jr., G.L., Stabeno, P., Walters, G., Sinclair, E., Brodeur, R.D., Napp, J.M., Bond, N. A., 2002. Climate change and control of the southeastern Bering Sea pelagic ecosystem. Deep-Sea Res. II 49 (26), 5821–5853. http://dx.doi.org/10.1016/S0967-0645(02)00321-1.
Hunt Jr., G.L., Coyle, K.O., Eisner, L., Farley, E.V., Heintz, R., Mueter, F., Napp, J.M., Overland, J.E., Ressler, P.H., Salo, S., Stabeno, P.J., 2011. Climate impacts on eastern Bering Sea foodwebs: a synthesis of new data and an assessment of the oscillating control hypothesis. ICES J. Mar. Sci. 68 (6), 1230–1243. http://dx.doi.org/10.1093/icesjms/fsr036.
IPCC, 2013. Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 1535.
K-1 model developers, 2004. K-1 coupled model (MIROC) description, K-1 technical report, 1. In: Hasumi, H., Emori, S. (Eds.), K-1 Coupled Model (MIROC) Description, K-1 Technical Report, 1. Center for Climate System Research, University of Tokyo, p. 34.
Kachel, N.B., Hunt Jr., G.L., Salo, S.A., Schumacher, J.D., Stabeno, P.J., Whitledge, T.E., 2002. Characteristics and variability of the inner front of the southeastern Bering Sea. Deep-Sea Res. Pt. II 49 (26), 5889–5909. http://dx.doi.org/10.1016/S0967-0645(02)00324-7.
Kim, S.-J., Flato, G.M., Boer, G.J., McFarlane, N.A., 2002. A coupled climate model simulation of the Last Glacial Maximum, Part 1: transient multi-decadal response. Clim. Dyn. 19, 515–537.
Kim, S.-J., Flato, G.M., Boer, G.J., 2003. A coupled climate model simulation of the Last Glacial Maximum, Part 2: approach to equilibrium. Clim. Dyn. 20, 635–661.
Kinder, T.H., Schumacher, J.D., 1981. Hydrographic structure over the continental shelf of the southeastern Bering Sea. In: Hood, D.W., Calder, J.A. (Eds.), Eastern Bering Sea Shelf: Oceanography and Resources, vol. 1. USDOC/NOAA/OMPA, pp. 31–51 (Chapter 4).
Large, W.G., Yeager, S.G., 2008. The global climatology of an interannually varying air–sea flux data set. Clim. Dyn. 33, 341–364.
Large, W.G., McWilliams, J.C., Doney, S.C., 1994. Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. Rev. Geophys. 32, 363–403.
Legutke, S. and Voss, R., 1999. The Hamburg Atmosphere-ocean Coupled Model ECHO-G. Technical Report #18, German Climate Computer Center (DKRZ).
Li, L., McClean, J., Miller, A., Eisenman, I., Hendershott, M., Papadopoulos, C., 2014a. Processes driving sea ice variability in the Bering Sea in an eddying ocean/sea ice model: mean seasonal cycle. Ocean. Model. 84, 51–66. http://dx.doi.org/10.1016/j.ocemod.2014.09.006.
Li, L., Miller, A., McClean, J., Eisenman, I., Hendershott, M., 2014b. Processes driving sea ice variability in the Bering Sea in an eddying ocean/sea ice model: anomalies from the mean seasonal cycle. Ocean Dyn. 64, 1693–1717. http://dx.doi.org/10.1007/s10236-014-0769-7.
McFarlane, N.A., Scinocca, J.F., Lazare, M., Harvey, R., Verseghy, D. and Li, J., 2005. The CCCma Third Generation Atmospheric General Circulation Model. CCCma Intern. Rep., 25pp.
Meehl, G.A., Arblaster, J.M., Tebaldi, C., 2007. Contributions of natural and anthropogenic forcing to changes in temperature extremes over the U.S. Geophys. Res. Lett. 34, L19709.
Mellor, G.L., Kantha, L., 1989. An ice-ocean coupled model. J. Geophys. Res. 94 (C8), 10937–10954.
Mueter, F.J., Bond, N.A., Ianelli, J.N., Hollowed, A.B., 2011. Expected declines in recruitment of walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea under future climate change. ICES J. Mar. Sci. . http://dx.doi.org/10.1093/icesjms/fsr022
Ortiz, I., Wiese, F.K., Grieg, A., 2012. Marine Regions Boundary Data for the Bering Sea Shelf and Slope. UCAR/NCAR-Earth Observing Laboratory/Computing, Data, and Software Facility. Dataset. ⟨http://dx.doi.org/10.5065/D6DF6P6⟩.
Ortiz, I., Aydin, K., Hermann, A.J., Gibson, G. Climate to fisheries – exploring processes in the eastern Bering Sea based on a 40 year hindcast Deep-Sea Res. II, submitted.
Reed, R.K., 2003. A surface heat flux climatology over a region of the eastern Bering Sea. Cont. Shelf Res. 23 (14–15), 1255–1263.
Ressler, P.R., DeRobertis, A., Kotwiki, S., 2014. The spatial distribution of euphausiids and walleye pollock in the eastern Bering Sea does not imply top-down control by predation. Mar. Ecol. Prog. Ser. 503, 111–122. http://dx.doi.org/10.3354/meps10736.
Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., Schulzweida, U., 1996. The Atmospheric General Circulation Model ECHAM4: Model Description and Simulation of Present-day Climate. Technical Report 218, Max Planck Institut für Meteorologie.
Rösel, A., Kaleschke, L., Birnbaum, G., 2012. Melt ponds on Arctic sea ice determined from MODIS satellite data using an artificial neuronal network. Cryosphere 6, 1–19. http://dx.doi.org/10.5194/tc-6-1-2012.
Saha, S., et al., 2010. The NCEP climate forecast system reanalysis. Bull. Am. Meteorol. Soc. 91, 1015–1057. http://dx.doi.org/10.1175/2010BAMS3001.1.
Scinocca, J.F., McFarlane, N.A., Lazare, M., Li, J., Plummer, D., 2008. The CCCma third generation AGCM and its extension into the middle atmosphere. Atmos. Chem. Phys. 8, 7055–7074.
Shchepetkin, A.F., McWilliams, J.C., 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Model. 9 (4), 347–404.
Stabeno, P.J., Bond, N.A., Kachel, N.B., Salo, S.A., Schumacher, J.D., 2001. On the temporal variability of the physical environment over the south-eastern Bering Sea. Fish. Oceanogr. 10 (1), 81–98. http://dx.doi.org/10.1046/j.1365-2419.2001.00157.x.
Stabeno, P.J., Bond, N.A., Salo, S.A., 2007. On the recent warming of the southeastern Bering Sea shelf. Deep-Sea Res. II 54 (23–26), 2599–2618. http://dx.doi.org/10.1016/j.dsr2.2007.08.023.
Stabeno, P.J., Napp, J., Mordy, C., Whitledge, T., 2010. Factors influencing physical structure and lower trophic levels of the eastern Bering Sea shelf in 2005: sea ice, tides and winds. Prog Ocean 85 (3–4), 180–196. http://dx.doi.org/10.1016/j.pocean.2010.02.010.
Stabeno, P.J., Kachel, N.B., Moore, S.E., Napp, J.M., Sigler, M., Yamaguchi, A., Zerbini, A.N., 2012a. Comparison of warm and cold years on the southeastern Bering Sea shelf and some implications for the ecosystem. Deep-Sea Res. II 65–70, 14–30. http://dx.doi.org/10.1016/j.dsr2.2012.02.019.
Stabeno, P.J., Farley, E., Kachel, N., Moore, S., Mordy, C., Napp, J.M., Overland, J.E., Pinchuk, A.I., Sigler, M., 2012b. A comparison of the physics of the northern and southern shelves of the eastern Bering Sea and some implications for the ecosystem. Deep-Sea Res. II 65–70, 31–45. http://dx.doi.org/10.1016/j.dsr2.2012.02.020.
Stabeno, P.J., Danielson, S., Kachel, D., Kachel, N.B., Mordy, C.W., 2016. Currents and transport on the eastern Bering Sea shelf. Deep-Sea Res. II 134, 13–29.
Stauffer, G., 2004. NOAA Protocols for Groundfish Bottom-trawl Surveys of the Nation's Fishery Resources. United States Department of Commerce, NOAA Technical Memorandum. NMFS-F/SPO-65, 205 pp.
Sullivan, M.E., Kachel, N.B., Mordy, C.W., Salo, S.A., Stabeno, P.J., 2014. Sea ice and water column structure on the eastern Bering sea shelf. Deep-Sea Res. II 109, 39–56. http://dx.doi.org/10.1016/j.dsr2.2014.05.009.
Wang, M., Overland, J.E., Bond, N.A., 2010. Climate projections for selected large marine ecosystems. J. Mar. Syst. 79 (3–4), 258–266. http://dx.doi.org/10.1016/j.jmarsys.2008.11.028.
Wolff, J., Maier-Reimer, E., Legutke, S., 1997. The Hamburg Primitive Equation Model HOPE. Technical Report 18, German Climate Computer Center (DKRZ).