Contribution to the Symposium: 'The Effects of Climate Change on the World's Oceans'
Original Article

Projected biophysical conditions of the Bering Sea to 2100 under multiple emission scenarios

Albert J. Hermann[1,2]*, Georgina A. Gibson[3], Wei Cheng[1,2], Ivonne Ortiz[1,4], Kerim Aydin[4], Muyin Wang[1,2], Anne B. Hollowed[4], and Kirstin K. Holsman[4]

[1]Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98195, USA
[2]Ocean Environment Research Division, NOAA/PMEL, Seattle, WA 98115, USA
[3]International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[4]Alaska Fisheries Science Center, NOAA, Seattle, WA 98115, USA

*Corresponding author: tel: + 206 526 6495; fax: + 206 526 6485; e-mail: albert.j.hermann@noaa.gov.

A regional biophysical model is used to relate projected large-scale changes in atmospheric and oceanic conditions from CMIP5 to the finer-scale changes in the physical and biological structure of the Bering Sea, from the present through the end of the twenty-first century. A multivariate statistical method is used to analyse the results of a small (eight-member) dynamically downscaled ensemble to characterize and quantify dominant modes of variability and covariability among a broad set of biophysical features. This characterization provides a statistical method to rapidly estimate the likely response of the regional system to a much larger (63-member) ensemble of possible future forcing conditions. Under a high-emission [Representative Concentration Pathway 8.5 (RCP8.5)] scenario, results indicate that decadally averaged Bering Sea shelf bottom temperatures may warm by as much as 5°C by 2100, with associated loss of large crustacean zooplankton on the southern shelf. Under a lower emission scenario (RCP4.5), these effects are predicted to be approximately half their calculated change under the high emission scenario.

Keywords: Bering Sea, biophysical modelling, climate change, regional modelling

Introduction

Widespread change is anticipated for the Bering Sea (AK) under climate change, including substantial oceanographic warming that scales with future carbon mitigation scenarios (IPCC, 2013, 2014). Climate-driven changes to oceanographic conditions have the potential to propagate through the food web and impact fish and fisheries in the region (Holsman *et al.*, 2018). The Bering Sea is a highly productive system that supports a wide diversity of species, some critically endangered, as well as multiple small coastal fishing communities that depend on subsistence harvest (Haynie and Huntington, 2016) and large-scale commercial fisheries that annually represent more than 40% of the U.S. commercial catch (Fissel *et al.*, 2017). In this article, we report estimates of anticipated change to the physical and lower trophic level dynamics of the Bering Sea, derived both through application of dynamical model downscaling and through statistical projections based on those results.

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 shelf domain (ocean depth is between ∼0 and 50 m), a middle shelf domain (with ocean depth ∼50–100 m) which is well-mixed in the winter and has two distinct layers separated by a sharp thermocline in the summer, and an outer shelf domain (ocean depth ∼100–200 m) which is more gradually stratified (Kinder and Schumacher, 1981; Coachman, 1986; Kachel *et al.*, 2002). A map of the region with these features identified is provided in Figure 1.

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 cold temperatures in the Bering Sea lead to an increase in large crustacean zooplankton, favoured as food items by juvenile pollock in the fall season (Coyle *et al.*, 2011).

The present hydrography and seasonal ice-driven 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 (Sigler *et al.*, 2016). Such intense production derives, in part, from a broad shelf with strong tidally 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. Interannual variation in winter ice extent over the Bering Sea modulates annual variability in productivity in the system. A cold period in the Bering Sea from 2006 to 2011 (Stabeno *et al.*, 2012) was followed by a return to warmer conditions, with reduced ice (Stabeno *et al.*, 2016, 2017) and attendant changes in primary and secondary productivity (Sigler *et al.*, 2016).

In previous studies, 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). 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 tended 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 was 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
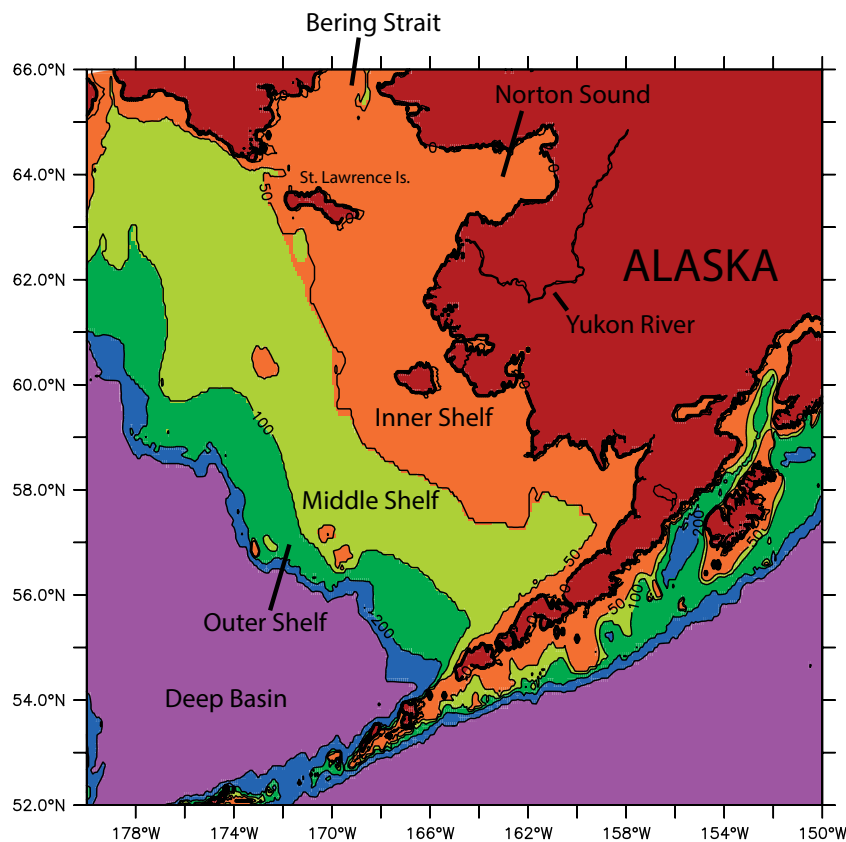

Figure 1. Bathymetry (m) with biophysical domains of the Bering Sea. Shown are inner shelf domain (0–50 m), middle shelf domain (50–100 m), outer shelf domain (100–200 m), shelf break (200–1000 m), and deep basin (>1000 m). Colour version of this figure is available online at *ICESJM* online.

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 modelling studies (Stabeno et al., 2010; Danielson et al., 2011; Cheng et al., 2014; Li et al., 2014a, b; 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.

In previous publications (Hermann et al., 2013, 2016a), we projected future Bering Sea conditions out to 2040 under an intermediate carbon mitigation scenario (i.e. A1B) from phase 3 of the Coupled Model Intercomparison (CMIP3; Meehl et al., 2007). Here, we extend this work to include a larger ensemble of global models under unmitigated (RCP 8.5) and moderate (RCP 4.5) mitigation scenarios from phase 5 of the CMIP (CMIP5) out to 2100 (Taylor et al., 2012). The goals of this contribution are: (i) to summarize the results of dynamical biophysical downscaling of a limited ensemble of CMIP5 models to the Bering Sea; (ii) to describe a multivariate method which allows extrapolation of those dynamically obtained results to a much larger ensemble of CMIP5 output; (iii) to apply that method in order to obtain a more robust estimate of the regional response of the Bering Sea to global change. Output from these simulations are being used in fisheries models, both to project stocks and for management-strategy evaluation, as part of NOAA's Alaska Climate Integrated Modeling Project (ACLIM; Hollowed et al., in prep.).

Methods
Dynamical downscaling regional model
The regional model used for these analyses (Bering10K) is identical to the one described in Hermann et al. (2016a). In that earlier publication, the regional model was driven by results from three global simulations from CMIP3 under emission scenario A1B. Here we take a similar approach, using global simulations from CMIP5 under emission scenarios RCP4.5 and RCP8.5. Collectively these downscaling runs span a wide range of model types and assumed future behaviours of mankind (e.g. RCP 4.5 vs. RCP 8.5 emission scenarios).

Major features of the regional model are as follows. The model is based on the Regional Ocean Modeling System version 3.2. ROMS is a sigma-coordinate model with curvilinear horizontal coordinates; a description of basic features and implementation can be found in Haidvogel et al. (2008) and Shchepetkin and McWilliams (2005). The Bering10K regional grid has approximately 10 km horizontal resolution, with ten vertical levels. Fine-scale bathymetry is

based on ETOPO5 and supplementary datasets as described in Danielson et al. (2011); smoothing of that bathymetry was utilized for numerical stability. Any oceanic regions shallower than 10 m were set to be 10 m deep. 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 surface salinity field within a few grid points of the coastline, using climatological monthly runoff values based on Dai et al. (2009). Bulk forcing, based on algorithms of Large and Yeager (2008), were used to relate winds, air temperature, relative humidity, and downward 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. Further detail of model tuning, implementation, and biases are available in Hermann et al. (2016a).

The lower trophic level dynamics (Nutrient-Phytoplankton-Zooplankton, NPZ) model is described in detail in Gibson and Spitz (2011). Briefly, this model includes two size categories of phytoplankton and ice plankton, and distinguishes among micro-zooplankton, copepods, neocalanus, and euphausiids, as well as jellyfish, benthic detritus, and epibenthos. Limiting nutrients are nitrate, ammonium, and dissolved iron. Metabolic and grazing rates are temperature dependent, which leads to substantially different food web structure under cold vs. warm conditions. Many of the state variables from this model were utilized in our analyses here. Results from hindcasts and forecasts with this model are described in Hermann et al. (2013, 2016a) and Ortiz et al. (2016).

Global models used for downscaling
The three CMIP5 models used as the basis for this study are: GFDL-ESM2M (Dunne et al., 2012), CESM (Kay et al., 2015), and MIROC (Watanabe et al., 2011). These were chosen based on: (i) performance in the Bering Sea under present conditions; (ii) the desire for a representative subset of CMIP5 members, spanning their range of variability; (iii) the ready availability of both physical and biogeochemical output. Spatial and temporal resolution of the output from the three chosen global models is shown in Table 1. Knutti et al. (2013) performed a dendogram analysis of the reference state of CMIP3/CMIP5 models; the GFDL-ESM2M, MIROC, and CESM models in fact span a broad range of global patterns for precipitation and SST. A regional summary of CMIP5 output was obtained through the NOAA climate change web portal (https://www.esrl.noaa.gov/psd/ipcc/cmip5/).

Air temperatures over the eastern Bering Sea from these three models span ∼80% of the range of CMIP5 results, with the CESM model being closest to the ensemble average (Figure 2). These models likewise span a wide range of ice climatologies for

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

Model	GFDL-ESM2M	CESM	MICOC
Ocean	0.33–1.0° latitude 1.0° longitude 50 levels vertically Monthly	nominally 1° latitude nominally 1° longitude 60 levels vertically Monthly	0.56° near equator, 1.71° latitude at the poles 1.4 ° longitude 44 levels vertically Monthly
Atmosphere	2.0° latitude 2.5° longitude 6 hourly	0.94° latitude 1.25° longitude Daily	2.79° latitude 2.81° longitude Daily

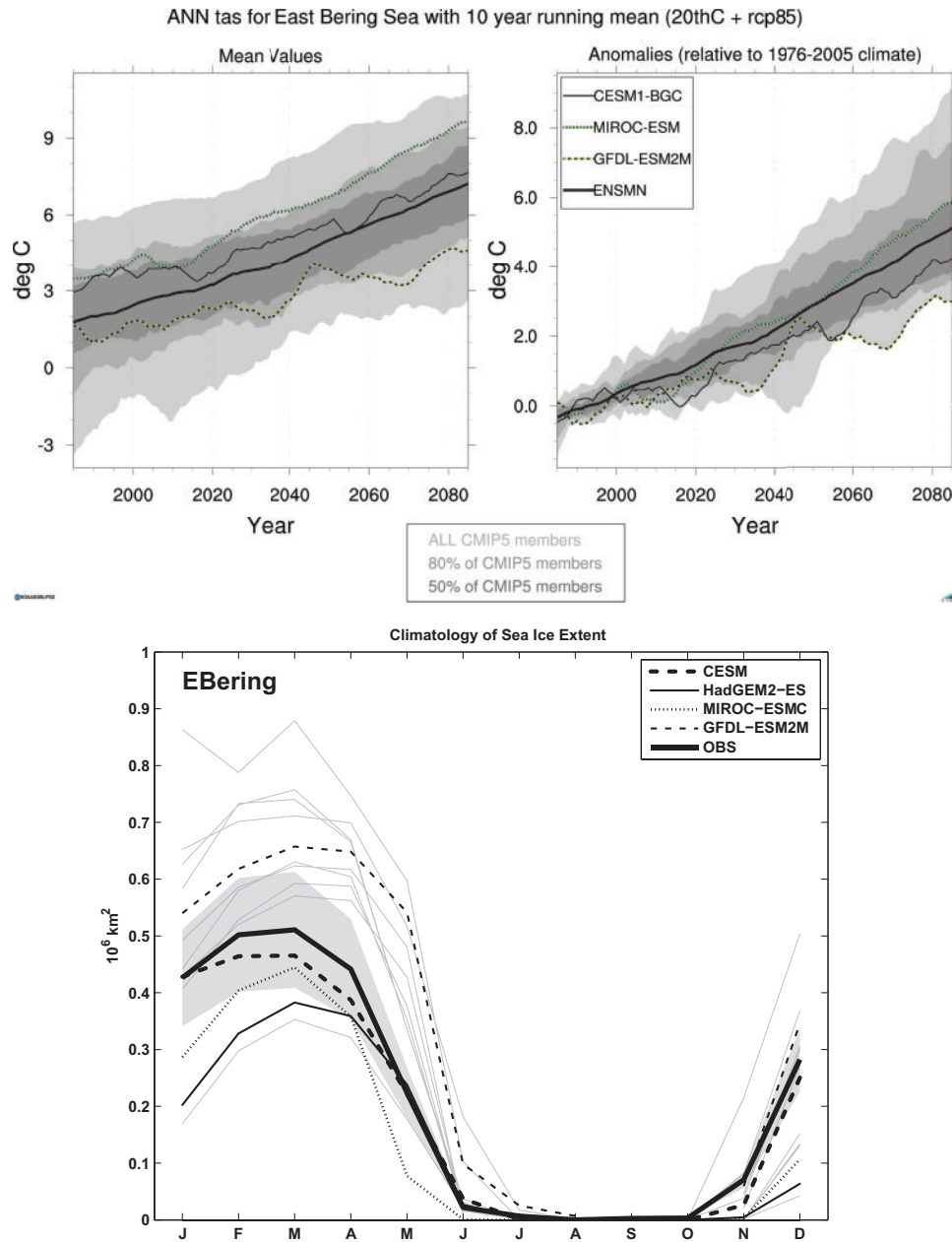ANN tas for East Bering Sea with 10 year running mean (20thC + rcp85)

Figure 2. Characterization of the chosen ensemble members for the Bering Sea, relative to other CMIP5 models. Upper left panel: low-passed (10-year running mean), spatially averaged air temperature for the eastern Bering Sea from CMIP5 members under RCP8.5, from 1976 through 2080, obtained from the NOAA climate change web portal (https://www.esrl.noaa.gov/psd/ipcc/cmip5/). Ensemble mean of all CMIP5 models (ENSMN) is shown along with individual trajectories of CESM, MIROC, and GFDL models. Light grey, medium grey, and dark grey illustrate the range of: 100% of CMIP5 members, 80% of CMIP5 members nearest to ensemble mean, and 50% of CMIP5 members nearest to ensemble mean, respectively. Upper right panel illustrates temperature change relative to individual model climatologies during 1976–2005. Lower panel: seasonal climatology of sea ice extent among different CMIP5 models as compared to observations during 1980–1999. Solid thick black line shows observed climatology (OBS); grey shading indicates ± 1 *SD* of all observations. Thin solid and light grey lines represent other CMIP5 model climatologies.

the Bering Sea, where that climatology was calculated using observations and model output for the period 1980–1999; GFDL has more ice than observed, MIROC has less ice than observed, and CESM has close to observed climatological ice. The models diverge considerably in their estimates of Bering Sea air temperature change after 2040.

For the CMIP5 downscaling simulations, we continuously span years 2006–2100. Oceanic physical boundary conditions for the Bering10K model were derived from the CMIP5 models' monthly averages, while atmospheric forcing is either daily (MIROC, CESM) or 6 hourly (GFDL). For the primary downscaling runs of the CMIP5 models (GFDL, MIROC, and CESM),

biological ICs and BCs were derived from present monthly climatologies (or assumed near-zero) as described in Hermann *et al.* (2013, 2016a).

Yearly (GFDL) and monthly (CESM) nitrate and dissolved iron values were obtained from the global biogeochemical projections. These were interpolated as the ICs and BCs for two additional realizations of Bering10K. These two additional downscaling runs (GFDL_RCP8.5_BIO, CESM_RCP8.5_BIO) utilized the same physical ICs, BCs and forcing as the runs with purely climatological biological conditions (GFDL_RCP8.5,

CESM_RCP8.5). These additional runs sample additional structural uncertainty of the projections, as they utilize projected changes in the large-scale biological conditions (e.g. those described in Rykaczewski and Dunne, 2010).

Variables analysed

As a summary of model behaviour, we examined a broad spectrum of physical and biological variables. A full listing of the variables chosen, with their units, are shown in Table 2. The variables chosen are similar to those used in the multivariate analysis of Hermann *et al.* (2013); a major new element here is the inclusion of atmospheric forcing variables. Ocean variables analysed include relevant surface and bottom values, as well as vertical averages. A derived variable used to summarize zooplankton biomass, here termed "large crustacean zooplankton," is as follows:

$$lcz \; = \; cope \; + \; nca \; + \; eup$$

where lcz, cope, nca, and eup refer to large crustacean zooplankton, copepods, neocalanus, and euphausiids, respectively (Table 2).

Analysis methods

Here we describe both univariate and multivariate analyses of the dynamically downscaled results. A summary of the CMIP5 realizations used in the various calculations is provided in Table 3.

Univariate analyses

To summarize behaviours of the model through time, we calculate yearly averages of the variables shown in Table 2 within 1.0 degree longitude $\times$ 0.5 degree latitude bins spanning the area between 180W–150W and 52N–66N (hence each bin contains approximately 5×5 of the native gridpoints of the Bering10K model). This averaging highlights the interannual/multidecadal variability of the output at $\sim$50 km spatial resolution. These binned values are used to calculate annual/decadal averages and internal vs. model variability, and for derivation of the multivariate modes used in the statistical projections.

For the dynamically downscaled results, mean change for each variable under RCP8.5 was summarized by two methods: (i) averaging over all the RCP8.5 ensemble members (GFDL_RCP8.5, CESM_RCP8.5, MIROC_RCP8.5, GFDL_RCP8.5_BIO, and CESM_RCP8.5_BIO) and over all spatial bins to obtain a single yearly time series; (ii) averaging over all RCP8.5 ensemble

Table 2. Variables used in the multivariate analysis of the simulations, with short names and units.

Surface temperature	sst	°C
Bottom temperature	sbt	°C
Surface salinity	sss	psu
Ice cover	iceco	fractional area
Mixed layer depth	mld	m (positive up coordinates; hence, negative change denotes *deepening* MLD)
Vertical mixing (depth avg.)	**akts**	$m^2\ s^{-1}$
Nitrate + ammonium (depth avg.)	nut	$mgN\ m^{-3}$
Ice phytoplankton	**iceph**	$mgC\ m^{-2}$
Small plus large phytoplankton (depth avg.)	**phyt**	$mgC\ m^{-3}$
Microzooplankton (depth avg.)	**mzoo**	$mgC\ m^{-3}$
Small copepods (depth avg.)	**cope**	$mgC\ m^{-3}$
Neocalanus (depth avg.)	**nca**	$mgC\ m^{-3}$
Euphausiids (depth avg.)	**eup**	$mgC\ m^{-3}$
Benthic detritus	**detbe**	$mgC\ m^{-2}$
Benthic infauna	**benth**	$mgC\ m^{-2}$
Sea surface height	ssh	m
Sea surface cross-shelf velocity	utop	$m\ s^{-1}$
Sea surface along-shelf velocity	vtop	$m\ s^{-1}$
Air temperature	tair	°C
Air pressure	pair	Pa
Specific humidity	qair	$kg\ kg^{-1}$
Zonal wind	uwind	$m\ s^{-1}$
Meridional wind	vwind	$m\ s^{-1}$
Downward longwave radiation	lwrad	$W\ m^{-2}$
Downward shortwave radiation	swrad	$W\ m^{-2}$

Variables shown in **bold** were log-transformed prior to use in the statistical analyses.

Table 3. Summary of the regional downscaling realizations and their use in dynamical-statistical results.

Regional model run acronym	CMIP5 model driver	Emission scenario	Bio-BCs	EOFs	RCP85 dynamical results	RCP45 dynamical results	Hybrid method results
GFDL_RCP85	GFDL	RCP85	clim	x	x		x
GFDL_RCP85_BIO	GFDL	RCP85	interp		x		x
CESM_RCP85	CESM	RCP85	clim	x	x		x
CESM_RCP85_BIO	CESM	RCP85	interp		x		x
MIROC_RCP85	MIROC	RCP85	clim	x	x		x
GFDL_RCP45	GFDL	RCP45	clim	x		x	x
CESM_RCP45	CESM	RCP45	clim	x		x	x
MIROC_RCP45	MIROC	RCP45	clim	x		x	x

CMIP5 model driver, CMIP5 global model output used as forcing and boundary conditions; **Emission scenario**, emission scenario of the driver; **Bio-BCs**, biological boundary conditions used for regional model (**clim**, present-day climatology; **interp**, interpolate from biogeochemical output of the global driver); **EOFs**, models used in derivation of univariate/multivariate EOFs; **dynamical results**, models used in calculation of dynamical method results; **hybrid results**, models used in calculation of hybrid (dynamical-statistical) method results.

members and over individual decades to obtain ensemble mean maps of change between decades.

Estimates of uncertainty

Our calculation of internal vs. model variability (also known as "intrinsic" vs. "structural" uncertainty) is similar to methods used elsewhere (e.g. Hawkins and Sutton, 2009). While it is recognized that oceanic and atmospheric time series have significant energy over a broad range of frequencies (red spectra), it is useful for our present purposes to divide these into sub-decadal and intra-decadal frequencies. Here, internal variability within a particular decade was calculated using the variance of each model's yearly averages during that decade, followed by averaging that statistic over all members of the ensemble under a particular emission scenario—that is, our working definition of "internal variability" refers to fluctuations of the annual averages over periods shorter than a decade (intra-decadal frequencies). Model variability was calculated as the variance of the decadal average change across ensemble members at each horizontal location—hence it is a measure of differences in decadal means (sub-decadal frequencies) due to model structure, where the "internal" variability has been filtered out. Using a similar distinction, Figure 2 displayed the model variability of spatially averaged, low-pass-filtered air temperature from the CMIP5 models. Here we will calculate internal and model variability for all of the regional and forcing variables at each spatial bin during the period 2090–2099. These definitions and methods for internal vs. model variability are separately applied to the dynamically and statistically derived ensembles of this study, and are subsequently combined into weighted estimates of uncertainty (see "Weighted estimates of change and uncertainty" section).

Multivariate analyses

Multivariate statistical methods can be useful in summarizing the behaviour of oceanographic models and how they respond to large-scale atmospheric forcing. Here we employ a variant of the canonical correlation analysis (CCA) technique to identify the dominant spatial patterns of multiple variables which rise and fall together through time. Similar analyses (Hermann *et al.*, 2013, 2016b) utilized coupled principal component analysis (CPCA) to achieve this aim. Related techniques have been described in Preisendorfer (1988) and Bretherton *et al.* (1992).

As an example of CCA, the time amplitude of a dominant regional spatial pattern of sea surface temperature, identified through univariate EOF analysis, may well be correlated with the time amplitudes of dominant spatial patterns of air temperature and primary production. These spatial patterns need not be similar across the variables; indeed, the primary production response, due to advection and bathymetry, will typically look very different than any wind, air temperature or SST pattern with which it is correlated through time.

Details of the statistical procedure are provided in the Appendix; here we summarize the basic steps. To find the most significant modes of *multivariate* behaviour—those which efficiently reproduce most of the covariance in the original time series of N variables—we begin with the univariate EOFs, which decompose the original fields into pairs of space and time functions. To obtain the most universal patterns, encompassing the widest range of scenario, model, and internal variabilities in a balanced fashion, we

include the following six dynamically downscaled ensemble members: GFDL_RCP4.5, GFDL_RCP8.5, CESM_RCP4.5, CESM_RCP8.5, MIROC_RCP4.5, and MIROC_RCP8.5. We demean the binned annual time series for each realization, then concatenate all members together along an abstract time axis. Note that the EOF analysis calculates the eigenvectors of the covariance matrix, and this matrix is unaffected by the values of time per se or their ordering; here we are preparing a collection of independent yearly realizations of each field, which happens to include both multiple years and multiple ensemble members.

The top few (say M) EOFs and their modulating "time series" for each of the N variables provide a new, dimensionally reduced, set of functions summarizing the patterns in the original data. A subsequent principal component analysis of the set of $N \times M$ multivariate time series will then identify how the spatial modes of these different variables tend to covary in our yearly realizations. This yields a final set of modulating series, with associated spatial patterns for each variable, which efficiently describes much of the correlated variability across both space and variables:

$$V_{ilt} = \sum_j X_{jil} T_{jit} = \sum_k C_{kil} \Gamma_{kt} \qquad (1)$$

where X and T are the traditional EOF space and time functions, whereas C and Γ are the new modes emphasizing covariance among the different variables. The indices i, j, k, l, t represent variable type i, spatial location l, univariate mode number j, multivariate mode number k, and time index t. We map the spatial structures of the most significant univariate and multivariate modes, and note the spatially averaged variance explained by each.

The value of this decomposition lies not only in its illumination of multivariate modes of variability, but also in its potential use in predicting any subset of "response" variables (here, the regional quantities not yet generated through dynamical downscaling) from a collection of "forcing" variables (readily available from coarse-scale CMIP5 output). This is achieved by projecting those forcing variables onto the previously derived multivariate modes—that is, by estimating how much of each multivariate mode C (which include the response variables) is contained in those readily available fields, for each year where downscaling has not been performed, to obtain the corresponding Γ value for that year. The result is an estimate of likely regional biophysical response, given only the large-scale or coarse-scale atmospheric forcing from some new global model (e.g. other CMIP5 or CMIP6 models), not yet dynamically downscaled to the Bering Sea:

$$Vest_{ilt} = \sum_k C_{kil} \Gamma est_{kt} \qquad (2)$$

Effectively this bootstraps the results from a few, computationally intensive, dynamically downscaling runs to an estimate of what downscaling results would have been obtained from a much larger ensemble. The success of this approach depends on finding multivariate modes which explain non-trivial variance in *both* the forcing variables (for which we have a large ensemble of realizations) and the regional response (for which we have a more limited ensemble).

In the present case, we choose air temperature, air pressure, zonal winds, and meridional winds as our "forcing" variables, since they are readily available online for most of the CMIP5

models, subjected to RCP4.5 (35 different models) and RCP8.5 (28 different models) emission scenarios (http://apdrc.soest. hawaii.edu/datadoc/cmip5.php). We compare the estimates from statistical projection of our six-member ensemble with the corresponding dynamically downscaled output, to quantify how much of the signal in each variable is captured by the technique. Specifically, we calculate the mean square difference between the original and estimated fields, normalized by the variance of the original field (*Rfit*, henceforth termed "fractional variance"):

$$Rfit_{il} = 1.0 - \left[\sum_t \left(Vest_{ilt} - V_{ilt} \right)^2 \middle/ \sum_t \left(V_{ilt} \right)^2 \right] \quad (3)$$

Hence a value of 1 indicates a perfect fit between original and estimated fields (and values less than 0 are possible for an exceptionally bad fit, though in practice all values fell between 0 and 1).

Finally, we use the technique to generate a 63-member ensemble of regional projections under RCP4.5 and RCP8.5, by estimating the Γ_{kt} for each of those members in each projected year, and summing the modes to obtain *Vest* [Equation (2)]. Details regarding the derivation of Γest (essentially, projection of four atmospheric forcing variables onto the multivariate modes) are shown in the Appendix.

Weighted estimates of change and uncertainty

The goal of our hybrid dynamical-statistical downscaling method is an improved estimate of the expected change, and the uncertainty of that estimate, due to the internal variability of the real system across a broad range of time scales and the structural uncertainty (many possible formulations of equations and parameters) of both the global and regional models. Dynamical downscaling methods are more likely to capture large, nonlinear changes in the regional system under global climate change, compared to simple statistical downscaling methods based on correlations between present-day global and regional data. However, there is no guarantee that the (affordably) small ensemble used for dynamical downscaling will faithfully capture the mean and variability which would have been obtained had the larger ensemble of CMIP5 global results all been used. The application of multivariate statistical rules, derived from the smaller ensemble, to the larger ensemble of global results, serves to correct for these discrepancies. This provides improved estimates of mean expected change and its uncertainty, when a full dynamical downscaling of all CMIP5 members is not feasible.

Note, however, that the statistical method is still just an approximation, and will fail to capture some of the changes observed under full dynamical downscaling. Its efficacy as a replacement will be both space and variable dependent. Hence, a final "best" estimate of change and uncertainty combines both the dynamical and statistical results. A reasonable approach here is to utilize the fractional variance (goodness of fit) metric to derive a weighted average of the two:

$$Vwgt_{ilt} = Rfit_{il} < Vest_{ilt} > + (1 - Rfit_{il}) < V_{ilt} > \quad (4)$$

where $<.>$ denotes the ensemble average over all available dynamical (*V*) or statistical (*Vest*) members. That is, for each variable, in areas where the statistical method is least successful, we preferentially utilize our small ensemble, dynamically downscaled results, while in areas where the statistical method is most successful, we preferentially utilize our large ensemble, statistically derived results. An analogous weighted mean is derived from the internal and model variability estimates, derived separately from the dynamical and statistical members.

Results

We begin by comparing the seasonal means for 10-year periods of the "present" (2010–2019), the "near-future" (2050–2059), and the "far-future" (2090–2099) under RCP8.5, as derived from direct dynamical downscaling. We then present results from the multivariate analysis, and statistically expand our original ensemble to include a much larger set of CMIP5 forcings. Finally we combine the dynamical and statistical estimates, present the anticipated changes by 2100 under RCP4.5 and RCP8.5, and present the magnitude of internal and model variability for these combined estimates.

Mean spatial patterns under present forcing

Decadal ensemble mean spatial patterns of "present" (2010–2019) conditions from the five-member RCP8.5 ensemble are displayed in Figure 3 using the raw (untransformed) variable values (for all subsequent plots and analyses, some of the variables were log transformed as noted in Table 2). Prominent spatial features include the following: (i) warmer sea surface temperatures in the Gulf of Alaska and cooler temperatures to the northwest; (ii) lowest sea bottom temperatures at mid shelf (the "cold pool"); (iii) lowest salinity near the coast; (iv) elevated phytoplankton on the middle shelf and near the coast; (v) depleted nitrate plus ammonium on the shelf; (vi) elevated microzooplankton, copepods, and euphausiids on the southeastern and middle shelf; (vii) elevated neocalanus on the southeastern shelf; (viii) elevated mixing on the northwestern shelf; (ix) elevated benthic detritus and epi-benthos along the 50 m isobath, with maxima to the north and south; (x) elevated ice cover and ice phytoplankton on the northwestern shelf; (xi) deepest mixed layer depth at the outer shelf; (xii) highest sea surface height in the Gulf of Alaska and on the southeastern shelf; (xiii) highest air temperature and absolute humidity in the southeast; (xiv) northeastward winds in the southeast and southeastward winds in the northwest; (xv) strongest downward longwave and shortwave radiation in the southeast.

Many of these model-generated features correspond to observed properties of the present-day Bering Sea. The mid-shelf "cold pool" of low bottom temperatures is a prominent feature of the region, with strong impacts on fish production. Correspondence of the modelled cold pool and ice cover with observations under past and present-day forcing, as well as extensive comparisons of temperature and salinity throughout the water column, have been described in Hermann *et al.* (2013, 2016a) and Ortiz *et al.* (2016). Similar patterns were observed here under "present" (2010–2019) CMIP5 forcing. The spatial pattern of yearly averaged phytoplankton—higher in the southeastern Bering (e.g. Bristol Bay and along the Aleutian chain), the northwestern Bering (e.g. the Bering Strait), and along the middle shelf domain—resembles the satellite-derived annual chlorophyll climatology reported in Brown *et al.* (2011). The southeasterly, mid-shelf maxima of euphausiids resemble the acoustically derived euphausiid biomass patterns reported in Ressler *et al.* (2012, 2014). A similar (southeasterly and mid-shelf) distribution of copepod biomass generally corresponds to the (spatially limited) copepod abundance estimates (Coyle *et al.,*
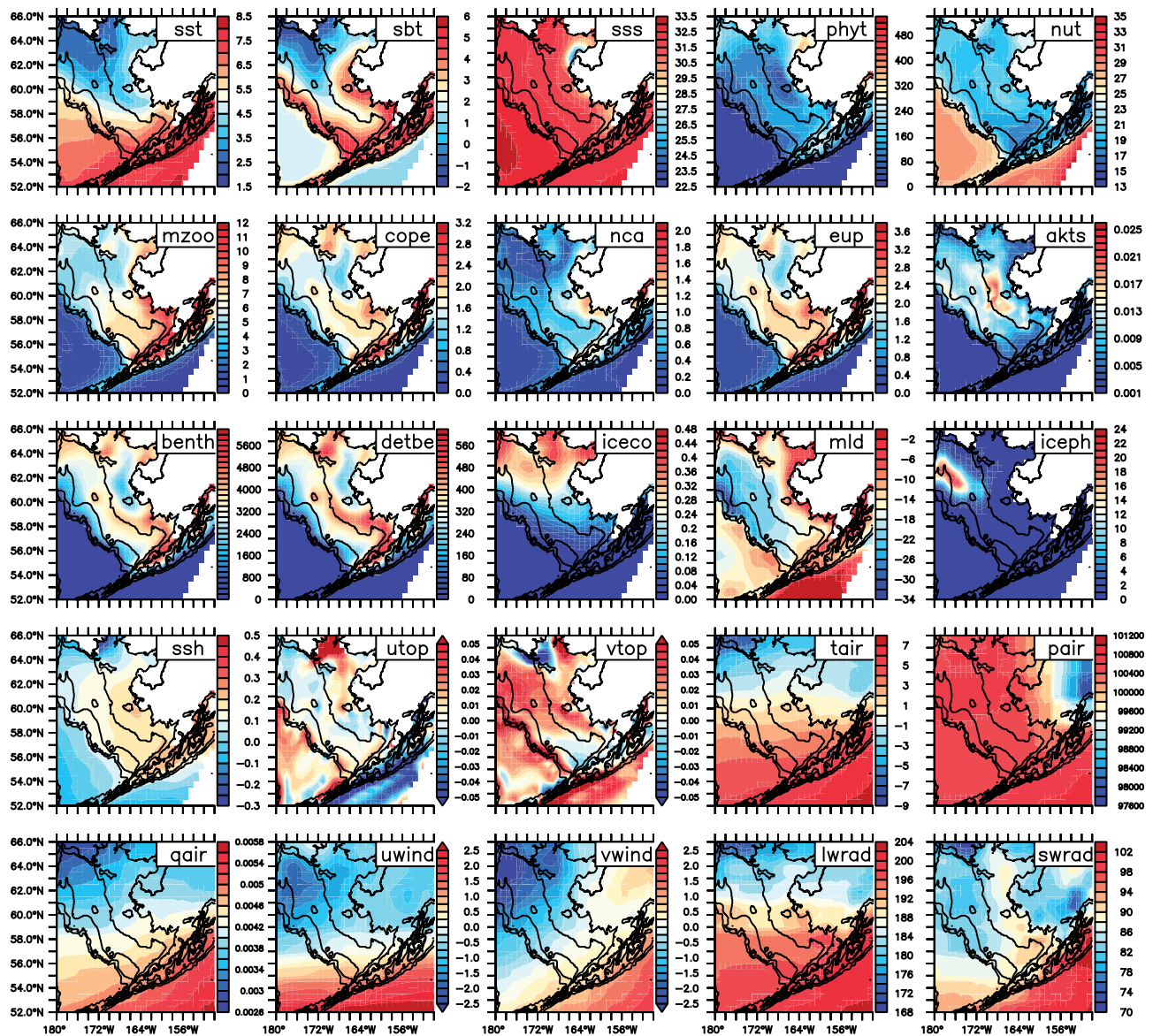
Figure 3. Decadal average spatial patterns for the years 2010–2019, downscaled from the five-member ensemble of RCP8.5 runs. Descriptions of each variable and its units are listed in Table 2. In this subsequent areal plots, isobaths are shown at depths of 50, 100, 200, and 1000 m. In these plots, the raw (untransformed) values are displayed; for all subsequent plots and analyses, some of the variables are log transformed as shown in Table 2. Colour version of this figure is available online at *ICESJM* online.

2008; Hunt *et al.*, 2011; D. Kimmel, pers. comm.). The model is less successful at replicating observed patterns of neocalanus, where the available data suggest maximal biomass near the shelf break (D. Kimmel, pers. comm.). Some of these pelagic biomass comparisons are complicated by the fact that we are using vertical averages from the model, as compared to near-surface vertical integrals or surface values reported in the observational studies. Spatial patterns of modelled benthic detritus and infaunal biomass in the northern Bering (e.g. local maxima near Bering Strait) roughly correspond to the infaunal patterns reported for the northern shelf in Grebmeier (2012) and Grebmeier *et al.* (2015). Limited information is available for total benthic infaunal biomass in the southeastern Bering; however, bottom trawl surveys typically indicate maxima of red king crabs along the Aleutian peninsula and the inner shelf domain near Bristol Bay, and maxima of snow crabs in

the northern portion of the middle shelf domain (e.g. Siddon and Zador, 2017). Our model indicates relative maxima of total benthic infaunal biomass in these regions.

Spatially averaged trends under RCP8.5

As a simplest index of projected change, we now examine the time series of the spatial average of each quantity, averaged over the five-member RCP8.5 ensemble (Figure 4; see Table 3 for models used). Quantities exhibiting consistent *upward* trends include sea surface temperature, sea bottom temperature, sea surface height, cross-shelf transport, air temperature, absolute humidity, and downward longwave radiation. Quantities exhibiting consistent *downward* trends include sea surface salinity, phytoplankton, nutrients, copepods, euphausiids, benthic detritus
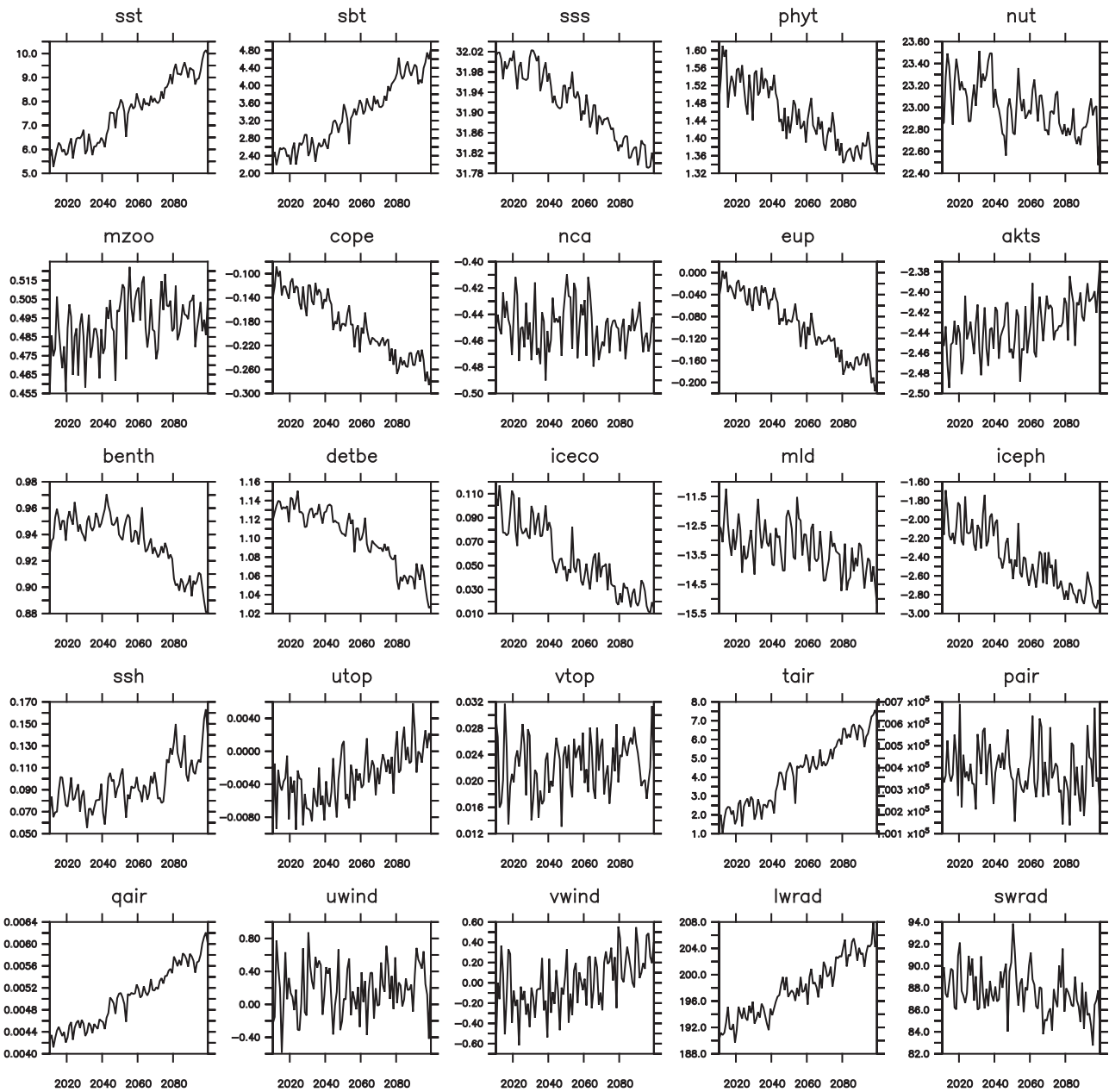
Figure 4. Time series of spatial averaged properties, averaged over the five-member RCP8.5 ensemble.

and epibenthos, ice cover, mixed layer depth (more negative values indicate mixed layer deepening), and ice phytoplankton. For most of these variables, the largest changes appear after 2040. Note how even after averaging across space and ensemble members, there is still substantial interannual as well as interdecadal variability, reflecting the intrinsic variability within each simulation.

Projected change in spatial patterns under RCP8.5

To examine the spatial patterns of these changes in more detail, and to correct for systematic bias in each of the models, we plot the average decadal change between the 2010s and the 2050s for the five-member RCP8.5 ensemble in Figure 5 (same models used

as in Figure 4). These patterns differ substantially from those of the mean state shown in Figure 3. Sea surface temperature (*sst*) increases uniformly across the model domain. Changes to sea bottom temperature (*sbt*) are focused on the inner and middle shelf, with maximum increase to the northwest. Sea surface salinity (*sss*) exhibits greatest decrease near the coast. Depth-averaged phytoplankton, copepods, and euphausiids (*phyt, cope, eup*) exhibit greatest losses along the outer shelf, while neocalanus (*nca*) exhibits a slight increase on the northwestern shelf, and microzooplankton (*mzoo*) exhibits a substantial increase all across the shelf. Ice cover (*iceco*) and an ice phytoplankton (*iceph*) decrease in the northwest. Sea surface height (*ssh*) increases along the coast, consistent with enhanced shoreward flow across the shelf (*utop*). Air temperature (*tair*) and absolute humidity (*qair*)
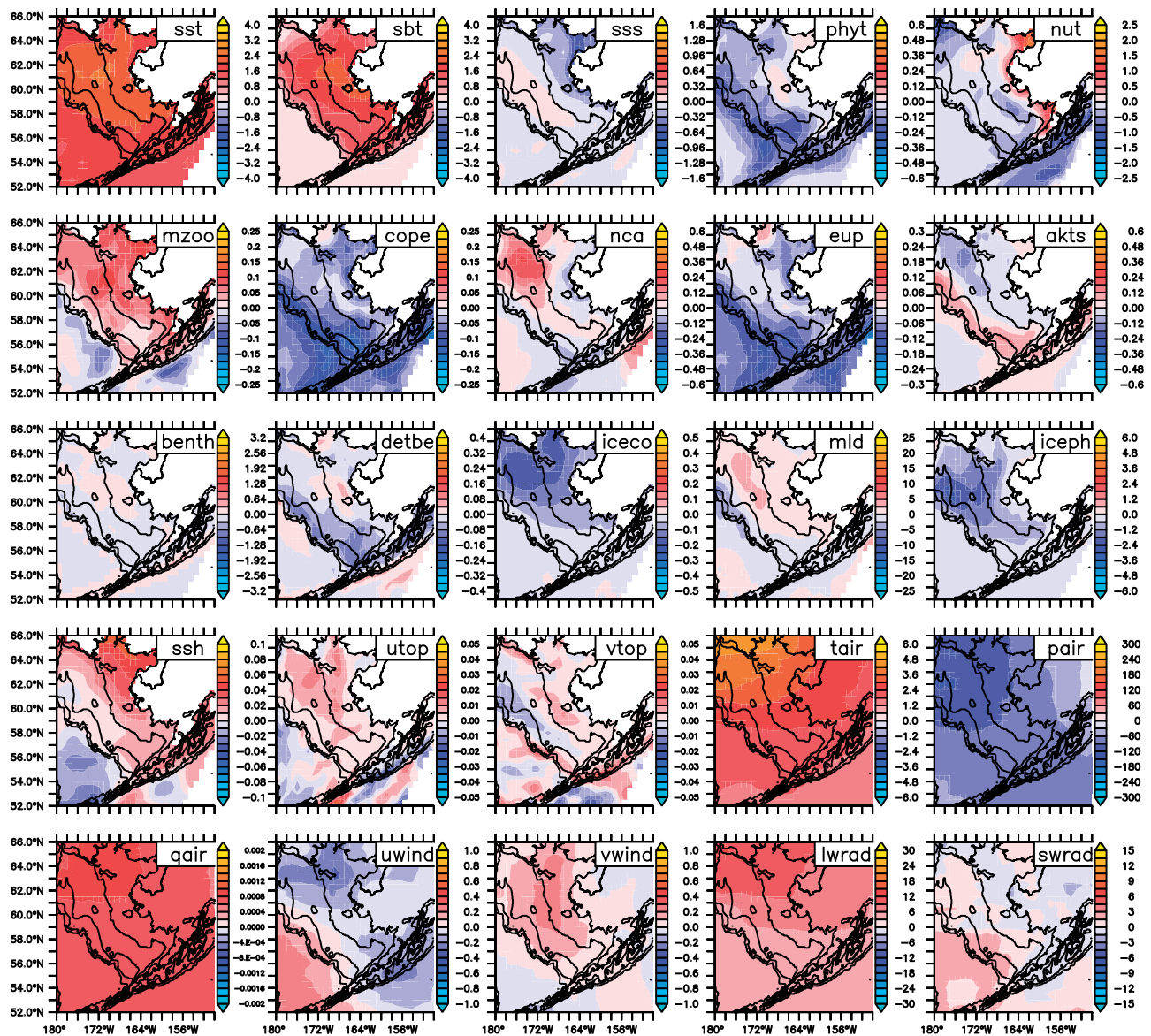
Figure 5. Decadal average change between 2010–2019 and 2050–2059 from downscaled projections based on GFDL-ESM2M, CESM, and MIROC RCP 8.5 global projections. Colour version of this figure is available online at *ICESJM* online.

increase everywhere, especially in the northwest, while air pressure (*pair*) decreases, especially in the northwest. Changes to winds (*uwind, vwind*) are geostrophically consistent with the changes in air pressure (*pair*), i.e. enhanced northeastward winds along the enhanced atmospheric pressure gradient. Increased downward longwave radiation (*lwrad*) is especially prominent in the northwest.

These broad spatial trends under RCP8.5 continue through the remainder of the twenty-first century (Figure 6). In addition, a substantial increase in vertical mixing (*akt*) and deepening of the mixed layer (*mld*) are produced along the shelf break. Sea surface and sea bottom temperatures increase by as much as 5°C, while copepods, euphausiids, and benthos (all $\log_{10}$ transformed in Figures 5 and 6) decrease by as much as half their initial values ($\log_{10}$ decreased by 0.3 or greater). Changes to shortwave radiation lack a clear trend or

spatial pattern by mid-century, but are reduced by ∼10% of their present-day mean by the end of the century.

To summarize: the CMIP5 projections of the small ensemble (based on GFDL, CESM, and MIROC models) anticipate a shift to warmer air temperatures (especially in the northern Bering Sea) and a shift to more northward winds under the RCP8.5 scenario. When dynamically downscaled through the Bering10K model, these and related changes in the forcing lead to substantially warmer surface and bottom temperatures (hence a smaller "cold pool"), reduced ice cover, enhanced cross-shelf surface flux, enhanced growth of small zooplankton and neocalanus in the (increasingly ice-poor) northern Bering, and reduced biomass of phytoplankton, copepods, and euphausiids on the outer shelf of the (increasingly warm) southern Bering.
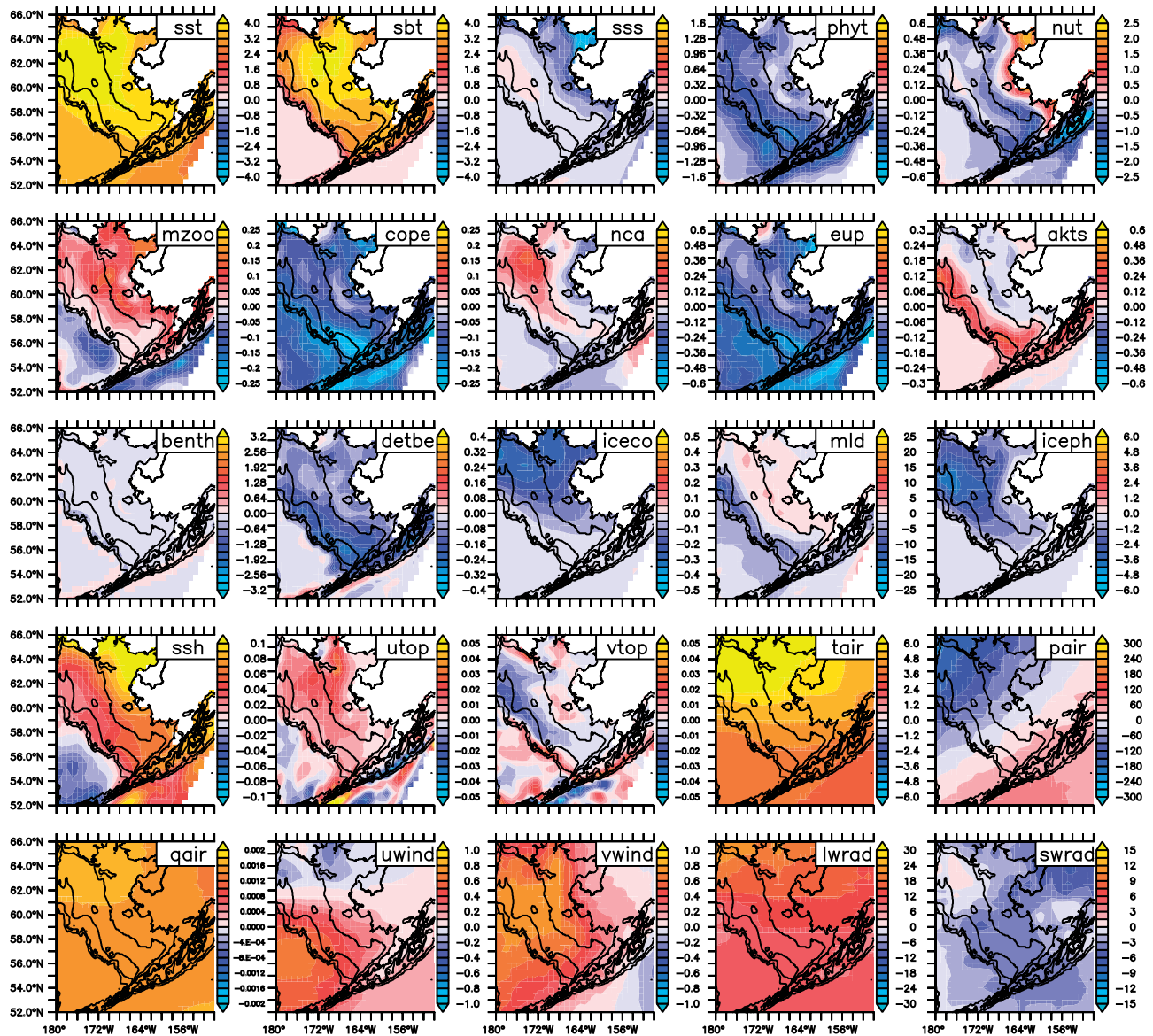
Figure 6. Decadal average change between 2010–2019 and 2090–2099 from downscaled projections based on GFDL-ESM2M, CESM, and MIROC RCP 8.5 global projections. Colour version of this figure is available online at *ICESJM* online.

Multivariate analysis

Univariate EOFs

These analyses use the six-member ensemble (see Table 3), which includes both RCP4.5 and RCP8.5 members. We begin by presenting the univariate EOFs of each analysed variable, based on the annually averaged fields (Figure 7). For many but not for all cases, these spatial patterns are similar to those of the long-term decadal trends under RCP8.5 in Figures 5 and 6. Note that this EOF analysis includes interannual through interdecadal scales of variability, as well as the considerable structural and scenario variability across the CMIP5 models driving the Bering10K realizations. This suggests a consistent, spatially dependent response in those realizations to recurring patterns in the forcing. In some cases (e.g. *sst*, *tair*, *qair*), the first univariate mode captures nearly all of the variance in the original time series, whereas in others

(e.g. surface velocities *utop* and *vtop*) less than 25% is explained by the first univariate mode.

Multivariate modes

When the multivariate modes are calculated (Figure 8), we find many spatial patterns similar to those of the first univariate EOFs (compare Figures 7 and 8, bearing in mind the sign of each pattern in Figure 7 is arbitrary). The spatial patterns of the first multivariate mode clearly indicate many of the patterns rise and fall in synchrony across variables; for example, note how different spatial patterns of sea bottom temperature and air temperature co-occur, and how these co-occurring patterns explain over 70% of each variable's original signal. Variables most strongly explained by the multivariate mode are: sea surface and bottom temperature, ice cover, air temperature, and absolute humidity. Variables most
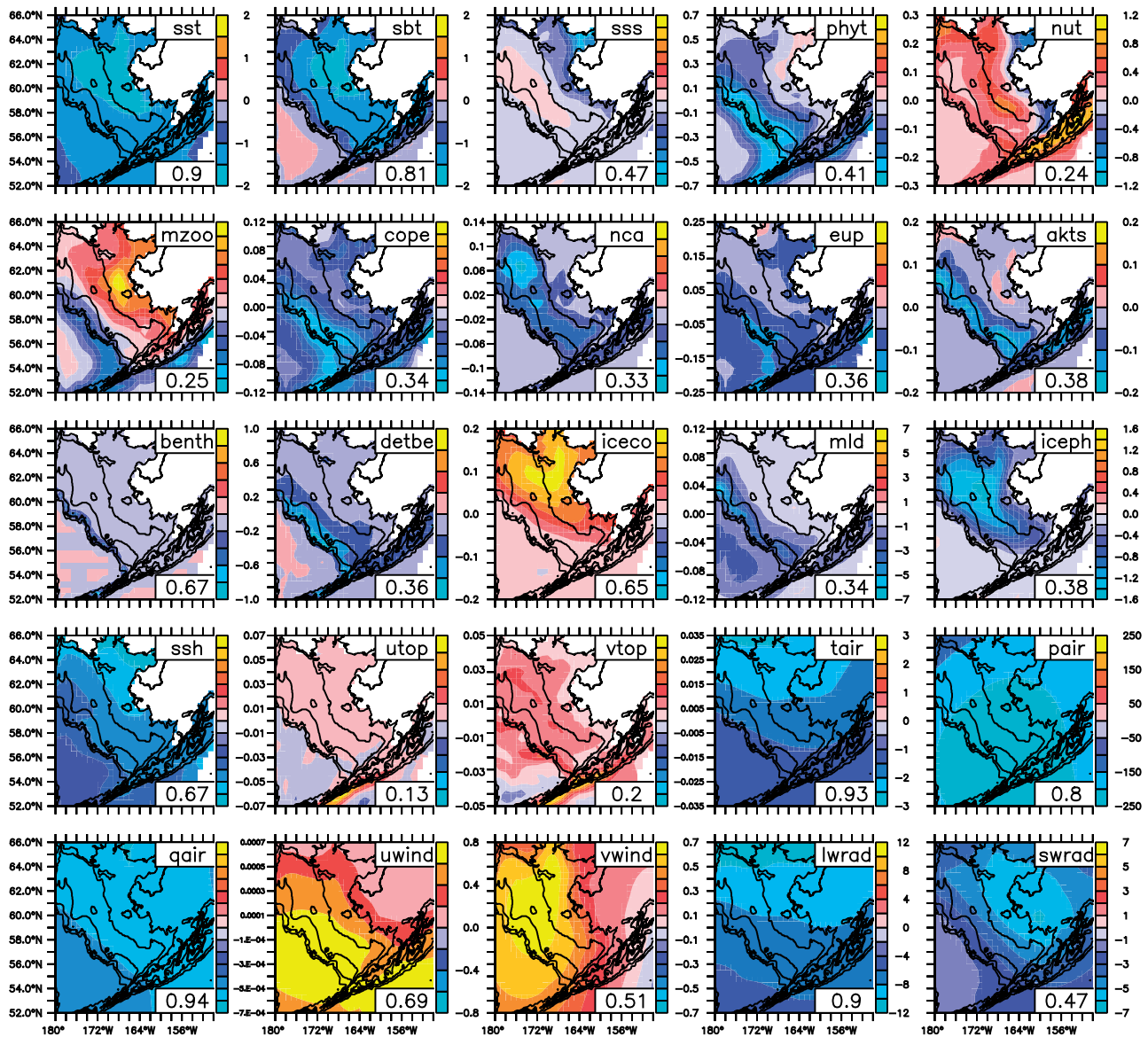
Figure 7. First-mode univariate EOFs for forcing and regional response variables, calculated using a six-member ensemble which includes both RCP4.5 and RCP8.5. Fractional variance explained by the univariate EOF is listed for each variable. Colour version of this figure is available online at *ICESJM* online.

weakly explained by this mode are: air pressure, surface velocities, winds, and shortwave radiation. A significant fraction of the biological signals are explained. In many locations, this "warm mode" is associated with reduced biomass of phytoplankton and large crustacean zooplankton, especially along the outer shelf. In some locations, the higher temperatures are associated with increased biomass, for example for microzooplankton and neocalanus.

The second multivariate mode (Figure 9) explains substantially less of the original signals; however, averaged over the entire domain, the following variables have at least 10% of their original variance explained: air pressure, zonal and meridional winds, along-shelf surface velocity, downward longwave radiation, shortwave radiation, vertical mixing (greater on the outer shelf), and benthic infauna. None of the biological variables have more than

10% of their spatially averaged variance explained by this mode, although in some cases there are spatially localized effects where much of the original signal is captured (e.g. note the near-shore reduction in many of the biological groups, and the near-shore increase in nutrients).

The differences between the two modes are further illustrated by plotting the variance explained for each variable (shown on Figures 8 and 9) on a 2D scatter plot (Figure 10). The two multivariate modes appear as fundamentally different modes of variability (a "heat" mode and a "wind" mode), with nearly all variables falling into one or the other "factor" group. The notable exceptions are benthic detritus and downward longwave radiation; modes 1 and 2 both explain a significant percentage of the original signal for those variables. Note also that both modes
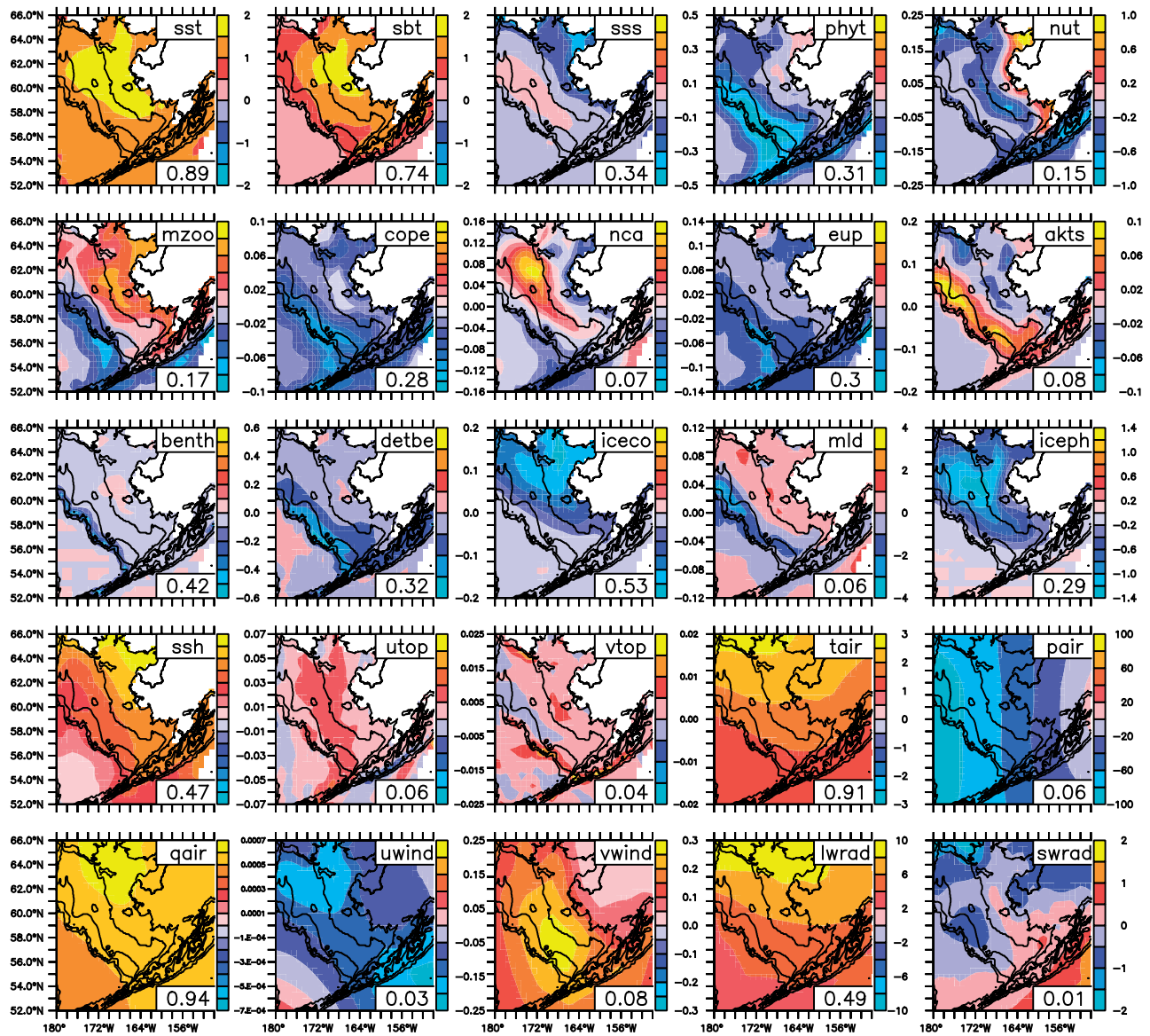
Figure 8. First-mode multivariate EOFs for the collection of forcing and regional response variables, with fractional variance of the original time series explained by that mode. Colour version of this figure is available online at *ICESJM* online.

include increased vertical mixing and deepening of the mixed layer on the outer shelf (Figures 8 and 9).

We now calculate how much of the variance in our "training" set *V* can be replicated by *Vest*, with multivariate mode amplitudes calculated exclusively by projecting only four of atmospheric forcing variables used in our dynamical downscaling: air temperature, air pressure, zonal winds, and meridional winds (Figure 11). In general, greater variance is explained within those areas where large change was calculated between 2010s and 2090s (Figure 6); these areas naturally dominate the univariate and multivariate modes as well (Figures 7–9). In many areas over 50% of the variance is replicated by the method, that is, we capture over 50% of the dynamically downscaled results using our statistical estimator alone (Figure 11). As described in "Weighted estimates of change and uncertainty" section, this fractional variance at each location is used in our final weighted estimates of change and uncertainty for each variable.

Statistical expansion of the ensemble using the multivariate modes

We now estimate the regional response to an expanded set of CMIP5 models under emission scenarios RCP4.5 and RCP8.5. As described in the Methods, this proceeds by projecting each member's available atmospheric "forcing" variables (here, air temperature, air pressure, zonal winds, and meridional winds) onto the multivariate modes calculated from the dynamical downscaling results, and subsequently taking a weighted average of the statistically and dynamically downscaled results [Equation (4)] The statistically calculated change in decadal sea bottom temperature (Figure 12) shows considerable variability among the different members, reflecting substantial structural variability among CMIP5 models. An ensemble mean is calculated from those members; as with the dynamically downscaled results, it exhibits an increase of up to 5°C on the northern Bering Sea shelf. The
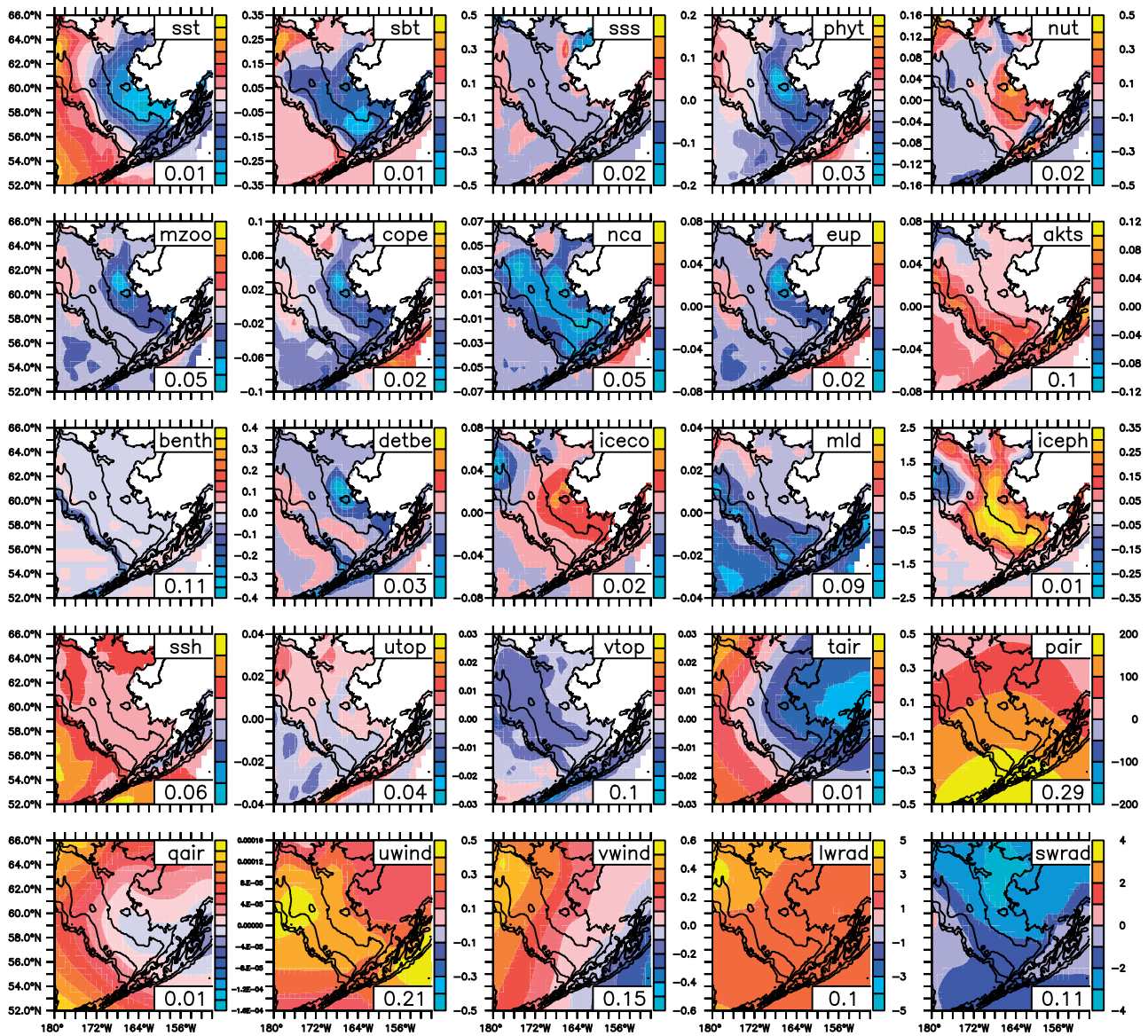
Figure 9. Second-mode multivariate EOFs for the collection of forcing and regional response variables, with fractional variance of the original time series explained by that mode. Colour version of this figure is available online at *ICESJM* online.

same procedure applied to total large crustacean zooplankton (*lcz*), using the logarithm of those values, predicts a decrease of −0.2 log units on the southeastern shelf by 2100—that is, a reduction nearly by half from the values earlier in the century (Figure 13). Conversely, there is mean increase in *lcz* over the northern shelf.

The weighted means of the (5-member) dynamical ensemble and (28-member) statistical ensemble result for decadal change under RCP8.5 for all variables (Figure 14) bear a very strong resemblance to the means calculated from the dynamical results alone under that emission scenario (compare Figures 6 and 14). Relative to the small ensemble results, the large ensemble appears to have greater northward winds (*vwind*), leading to greater onshore flow (*utop*) and slightly reduced ice cover (*iceco*). Other differences include slightly reduced changes for air temperature (*tair*), sea surface temperature (*sst*), and sea bottom temperatures

(*sbt*). Some of these differences between Figures 6 and 14 are hard to discern, as the spatial patterns are so similar, but emerge in difference maps (not shown). The strong resemblance of the spatial patterns by the two methods results from two factors: (i) the mean change in air temperature of the five-member dynamical ensemble is very similar to the mean change of the full CMIP5 ensemble (e.g. see Figure 2), and that particular forcing variable strongly correlates with the regional responses (the "heat" mode revealed by the analysis); (ii) not all areas were well reproduced by the statistical modes (low fractional variance explained, see Figure 11), and hence the original, dynamical estimate was preferentially used in those areas [Equation (4)].

The analogous weighted maps based on the (3-member) dynamical and (35-member) statistical ensembles under RCP4.5 (Figure 15) exhibit marked differences from the results under RCP8.5. Of particular note, they predict only half the increase in
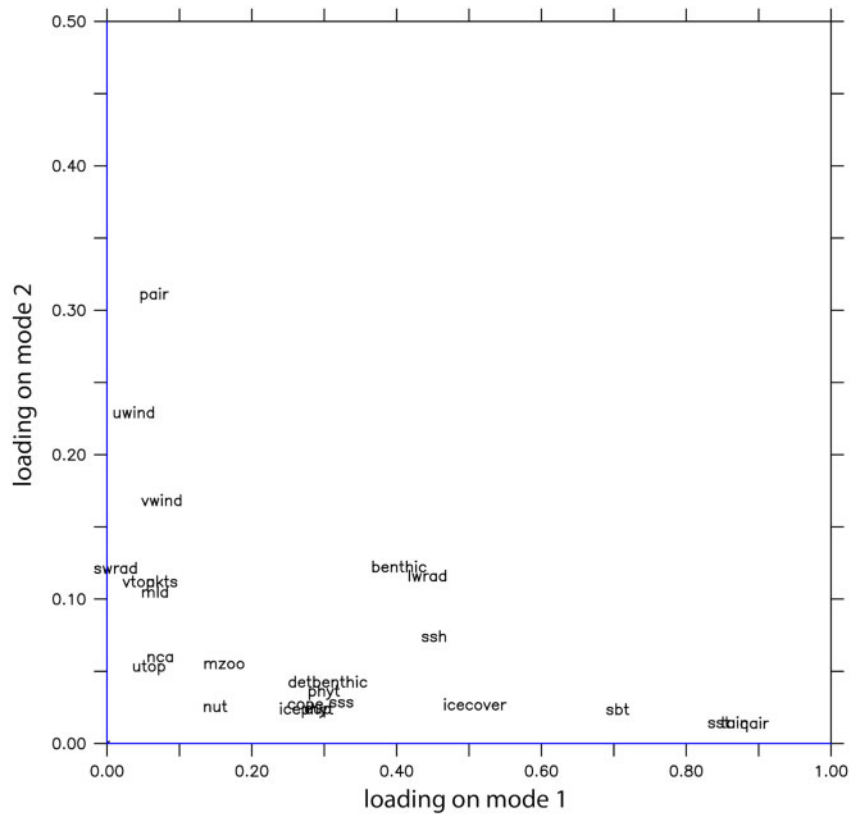
Figure 10. Variable loadings on the first 2 multivariate modes; a scatter plot of total variance explained for each variable by multivariate mode 1 (*x*-axis) vs. multivariate mode 2 (*y*-axis). Variances plotted here are also listed in Figures 8 and 9, respectively

surface and bottom temperatures anticipated under RCP8.5. The general sense of the spatial patterns is retained, but at much smaller amplitude as compared to RCP8.5.

Internal vs. model variability of the projections

Variabilities ("uncertainties") of projections computed from the dynamical method alone (not shown here) were in fact very different from those computed by the statistical method, a reflection of the small sample size of the dynamical set. The much larger statistical set presumably gives us a superior estimate of the true uncertainty among projections, at least in those areas where the statistical method is able to capture a substantial fraction of the dynamical variance (see Figure 11). The weighted dynamical-statistical means for internal and model variability among the RCP8.5 projections are shown in Figures 16 and 17, respectively. The model variability is larger than the internal variability for nearly all variables and locations; this is especially true for sea surface height. The exceptions are the atmospheric pressure and associated winds, which exhibit far greater internal than model variability. For many of the variables, model variability is greatest in areas where the largest mean interdecadal change was calculated (Figure 6). In particular, sea bottom temperature exhibits greatest such uncertainty (greater than 2°C) in the northern Bering Sea.

For a majority of the variables, both the model and internal variability are smaller than the calculated mean change in most areas by 2100. Winds and air pressure exhibit some of the greatest internal variability relative to mean change, whereas sea surface height, longwave radiation, and shortwave radiation exhibit some of the greatest model variability relative to mean change.

Discussion

Causal pathways and limits of the dynamical model

The basic patterns of change from the dynamically downscaled results suggest large changes to the future Bering Sea, with considerable spatial detail—not all areas warm to the same extent, and large zooplankton are reduced most severely on the outer shelf. These results are consistent with our previous downscaling runs of the model through 2040 (Hermann *et al.*, 2016a), as well as recent observations of warm vs. cold years in the Bering Sea (Hunt *et al.*, 2011). The warmer surface and bottom temperatures are largely a response to warmer air temperatures and reduced ice (hence greater penetration of shortwave radiation). Reduced southward advection of ice (as winds become more northward) may contribute secondarily to warming in the south. Reduced formation of ice also manifests as reduced salinity in Norton Sound, as brine rejection is reduced there. One probable causal pathway reducing the large zooplankton is heightened grazing pressure on those groups (represented as a temperature-dependent quadratic loss term in the model). The northern increase in neocalanus is difficult to interpret, especially as the present-day pattern produced by the model may not reflect the true observed distribution. The broad increase in microzooplankton is likely driven in part by an increase in temperature-dependent growth rates. The decreased phytoplankton on the outer shelf may be partly
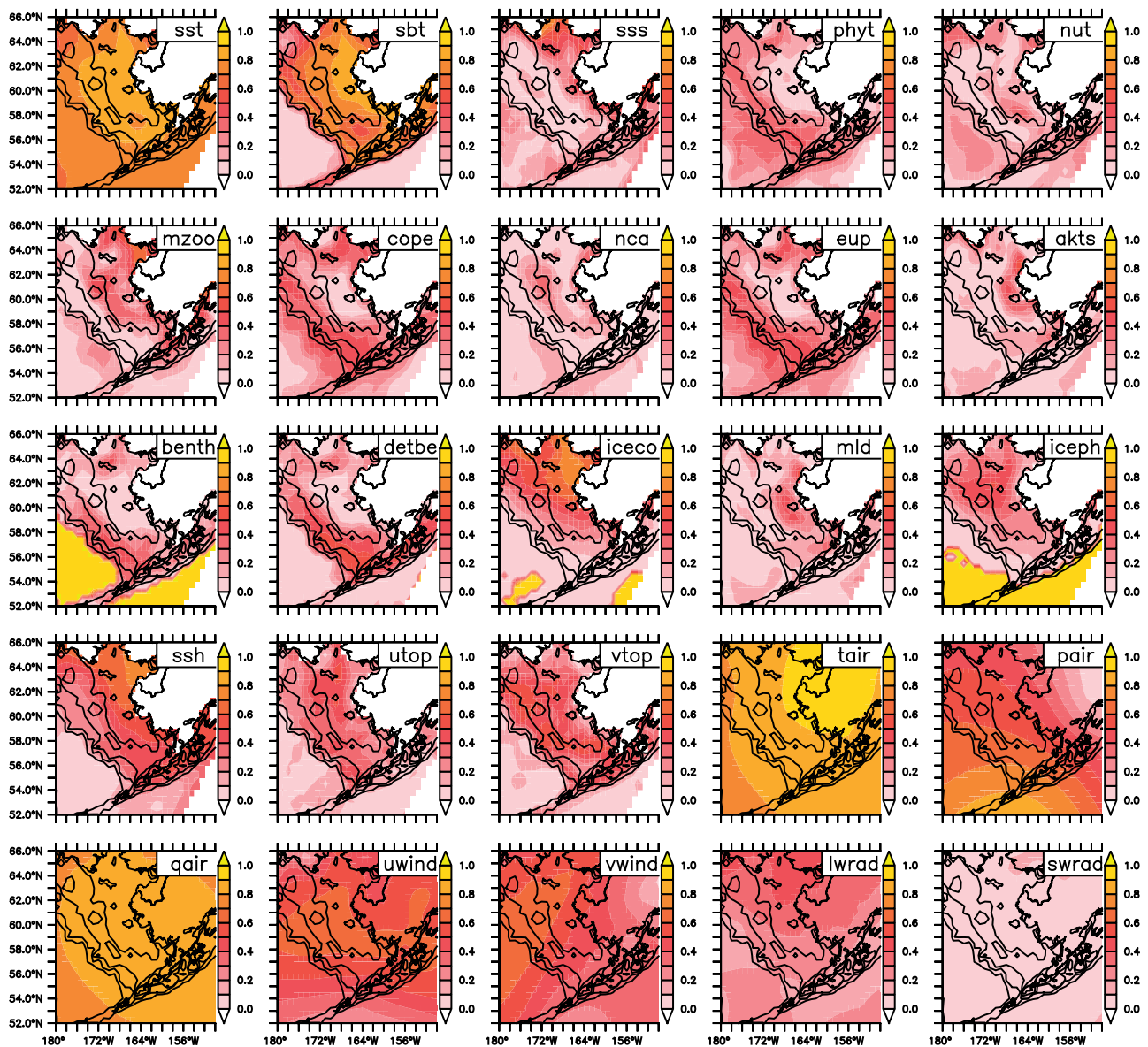
Figure 11. Fractional variance of the dynamically downscaled output reproduced by projecting four atmospheric forcing variables (tair, pair, uwind, and vwind) onto the top three multivariate modes. Colour version of this figure is available online at *ICESJM* online.

due to the deepened mixed layer, for example through light limitation. Further analysis of the model output may help to clarify these mechanisms.

Known biases and imperfections of the regional dynamical model used here include a limited number (10) of vertical layers, a fixed Bering Sea through flow, and a simple monthly climatology (no interannual variability) for coastal runoff. Several of these issues have been addressed in more recent simulations, which include more (30) layers and improved algorithms for light extinction, primary production, and zooplankton diapause (K. Kearney, pers. comm.). Previous work has indicated that the use of only 10 layers does not strongly limit our ability to reproduce the fundamental seasonal and interannual temperature patterns, although it may contribute to a persistent shallow bias in mixed layer depth (Hermann *et al.*, 2016a). An observed bias toward late

melting of ice in the hindcasts (Ortiz *et al.*, 2016), while undesirable, should not strongly affect the annual averages employed for our analyses. Recent improvements to the ice dynamics and thermodynamics code (K. Hedstrom, pers. comm.) have significantly reduced this bias.

More generally, it is recognized that for the biological elements in particular: (i) there is considerable uncertainty about the present structure of the food web and appropriate rates for each variable, and (ii) we are implicitly assuming that the present major groups (state variables) and their rates will endure over the twenty-first century. As in global models, we have no guarantee that this will be the case; new species types, with different rates and strategies, may ultimately come to dominate the Bering Sea as it warms. Caution is of course warranted in the use of such projections.
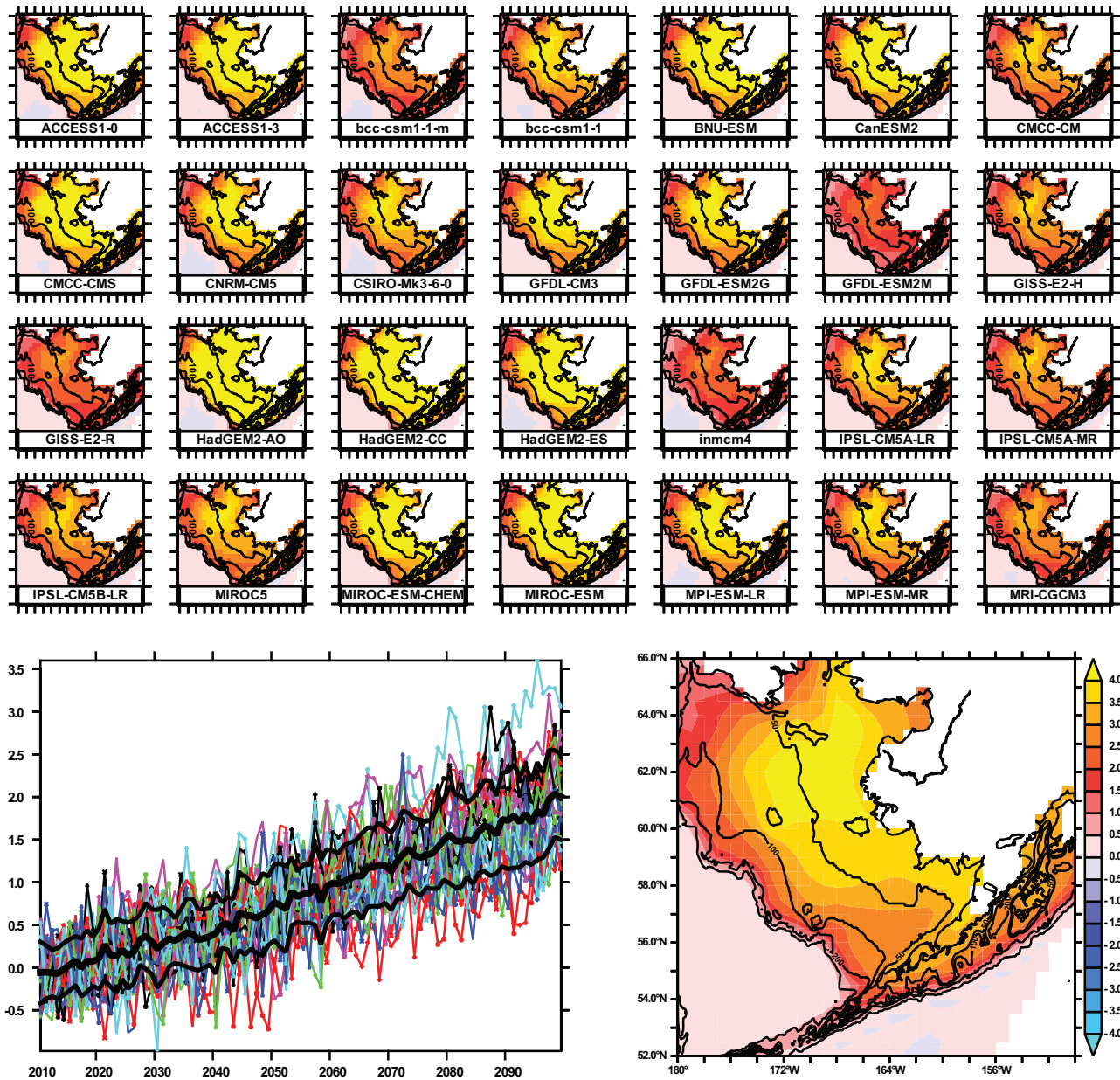
Figure 12. Ensemble results for sea bottom temperature (sbt), obtained by projecting atmospheric forcing variables (from 28 different CMIP5 models under emission scenario RCP8.5) onto the multivariate modes. Upper panel: calculated change in 10-year average sbt between 2010–2019 and 2090–2099 for each CMIP5 realization. The forcing CMIP5 model is listed at the bottom of each panel. Isobaths are shown at depths of 50, 100, 200, and 1000 m. Lower left panel: yearly areal average for each CMIP5 realization, relative to the 2010–2019 mean. Dark black line shows ensemble mean; light black lines indicate $\pm$ 1 SD for that year. Lower right panel: ensemble mean change based on the 28 CMIP5 models. Colour version of this figure is available online at ICESJM online.

Limits of the hybrid dynamical-statistical methodology

The differences in temporal change calculated from the smaller (dynamically downscaled) and larger (statistically downscaled) ensembles reflect the added model (structural) uncertainty of CMIP5 encompassed by the latter. The five-member RCP8.5 ensemble (actually only three different models, with two additional runs using biological boundary conditions from GFDL and CESM models), while deliberately chosen to span a range of "warmer" and "colder" models from CMIP5, is still a small subsample of that set. Indeed, our initial choice of three models

(GFDL, CESM, and MIROC) is a small sample from a population with large variance (e.g. see Figure 2), hence subject to significant error as an estimate of the true mean and variance (uncertainty) of that population. This fact—and the high cost of dynamical downscaling—is the primary motivator for seeking an efficient hybrid (dynamical-statistical) method to downscale a larger population of CMIP5 output.

The fact that the statistically generated ensemble results are so similar in pattern to the dynamically generated results is to some extent a consequence of so much of the variance and covariance
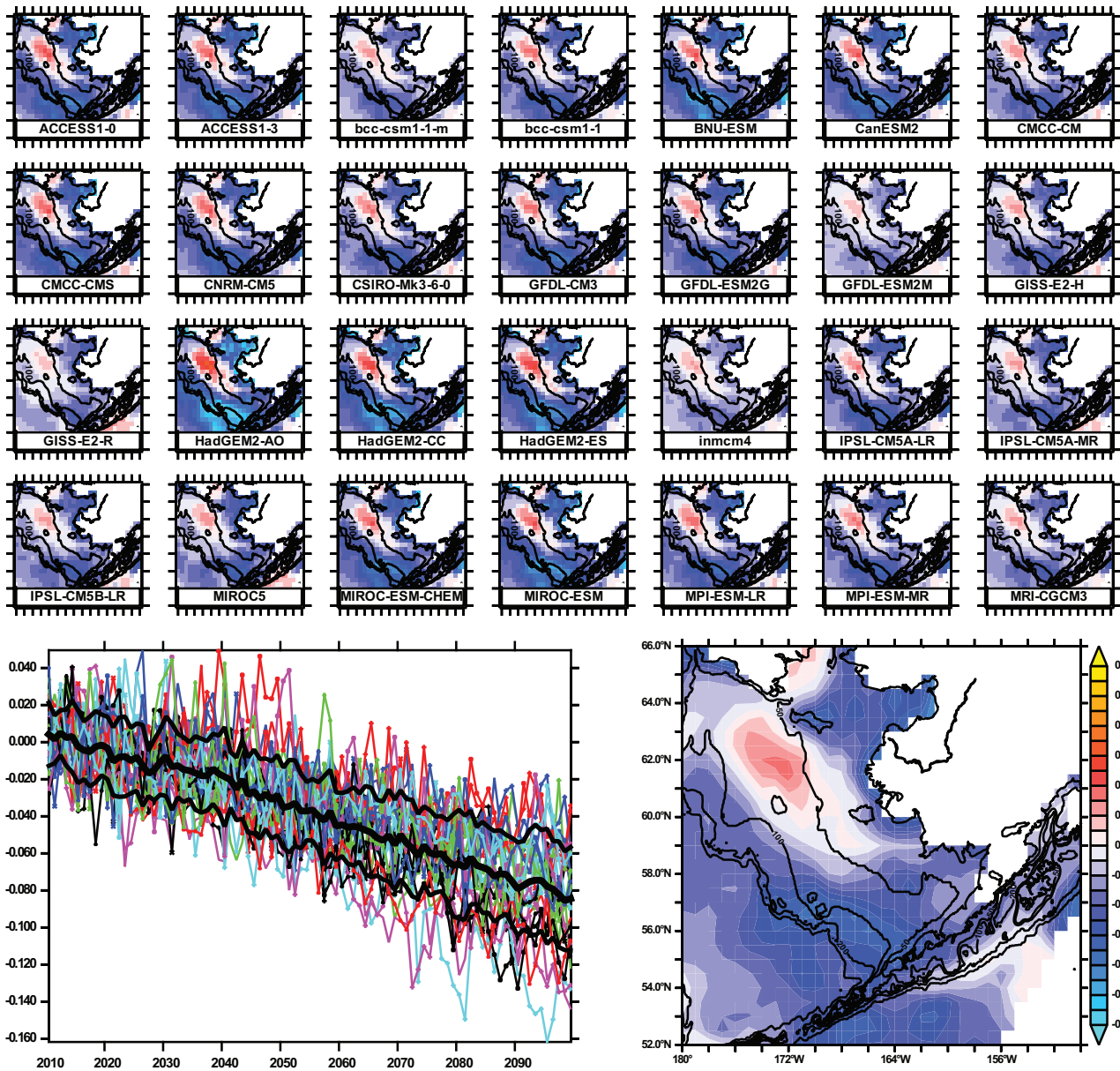
Figure 13. Ensemble results as in Figure 12, for $\log_{10}$ (large crustacean zooplankton). Colour version of this figure is available online at *ICESJM* online.

being contained in the first, "heat," multivariate mode. The forcing data from the larger set of CMIP5 output projects strongly onto this mode, which generates a final pattern similar to that of the dynamically generated results. However, we now have the added benefit of the much larger ensemble, to inform us regarding the expected mean amplitudes of those patterns in any future year, and their uncertainty. We achieve this benefit without dynamically downscaling all of the CMIP5 members. Obviously we will not capture all of the signal which would emerge from a full dynamical downscaling through this hybrid method. To compensate, we use an objective measure of our statistical method (fractional variance explained, Figure 11) in our final weighted sums [Equation (4)], to extract the greatest value from both methods.

Simpler statistical methods (e.g. linear regression between observed air temperature and regional ocean temperature, applied to predict the future ocean temperature using projected air temperature from CMIP5) have been used in other studies to quickly predict the regional response to a broad range of possible futures. The fact that our multivariate technique emphasizes covariance across space and across variables (e.g. air temperature is correlated with absolute humidity, and the sea level pressure field is correlated with the winds) reduces the possibility of spurious correlations being obtained for any single location and/or single pair of predictor/predictand variables. Instead, we have assimilated the covariance structure of the entire set of multivariate fields into our unified statistical rule.
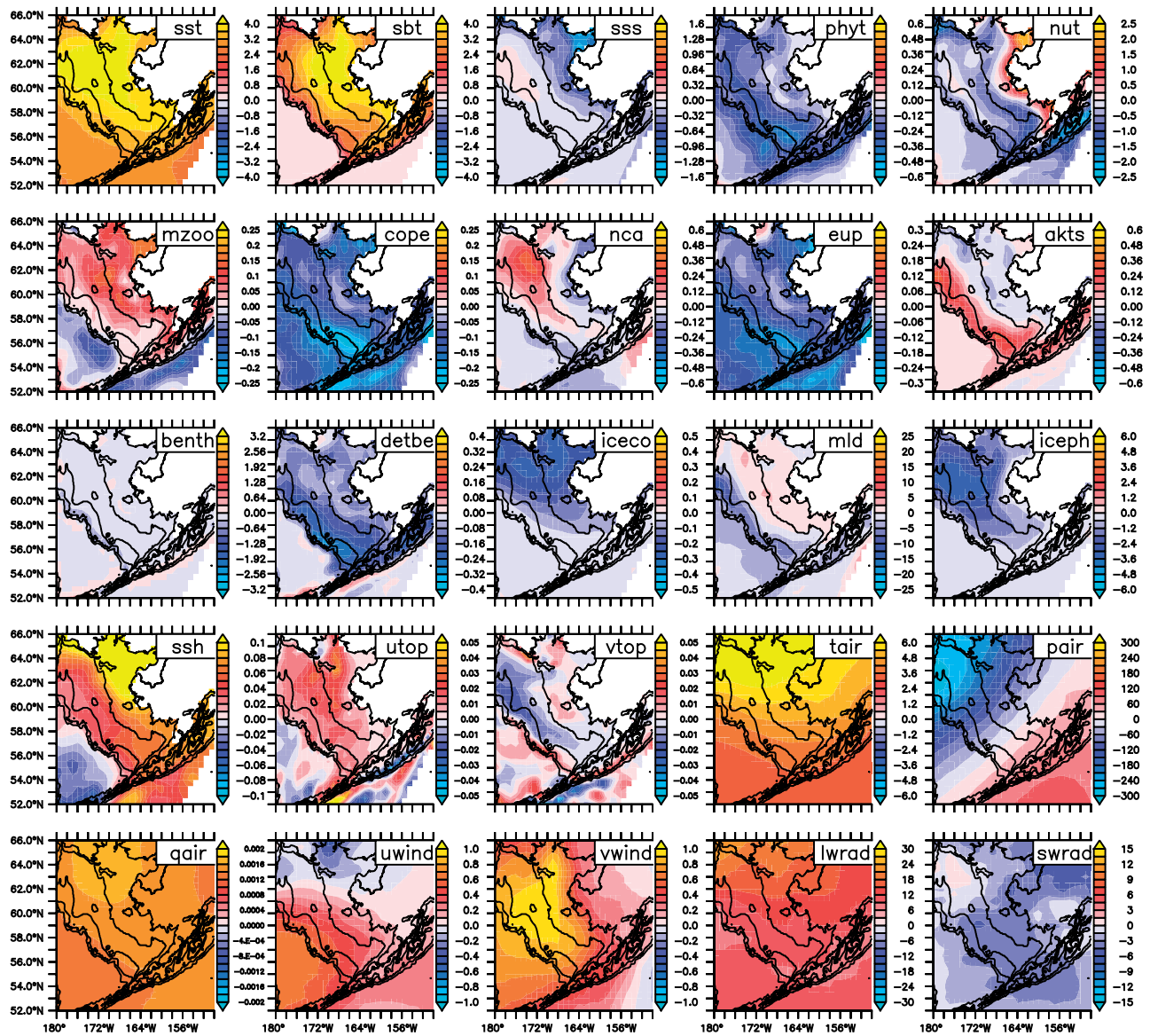
Figure 14. Predicted decadal average change between 2010–2019 and 2090–2099 under RCP8.5, using statistical extrapolation to include 28 ensemble members (weighted average of dynamical and statistical results). Colour version of this figure is available online at *ICESJM* online.

Further testing and alternate methods

In the present work, we used the same set of data both for: (i) generation of the multivariate modes and (ii) testing how well a subset of the variables could replicate, through projection onto those multivariate modes, the original set. More rigorous tests of the method will include sequentially withholding individual ensemble members from the original training set, and calculating how well the methods can replicate the withheld member. We will also explore the relative performance of simpler multivariate regression methods at each spatial location. Finally, this and alternate methods will be tested for application to seasonal patterns. Seasonal evolution entails strong autocorrelation of signals, and hence alternate methods such as linear inverse modelling (LIM) (Newman *et al.,* 2003; Alexander *et al.,* 2008) may be more appropriate for that case. Such LIM methods entail the calculation of a space-time correlation matrix, sometimes using a dimensionally reduced dataset (e.g. Capatondi and Sardeshmukh, 2015), as was used in our analysis.

Conclusions

As part of an interdisciplinary project to explore management strategy in a future Bering Sea (ACLIM), an eight-member ensemble of global CMIP5 output was dynamically downscaled using a 10-km resolution regional biophysical model. These downscaled projections indicate that shelf bottom temperatures may increase by as much as 5°C by the end of the twenty-first century, given a continuation of present global trends in the emission of greenhouse gases (i.e. IPCC scenario RCP8.5). These changes may be accompanied by a significant reduction in large
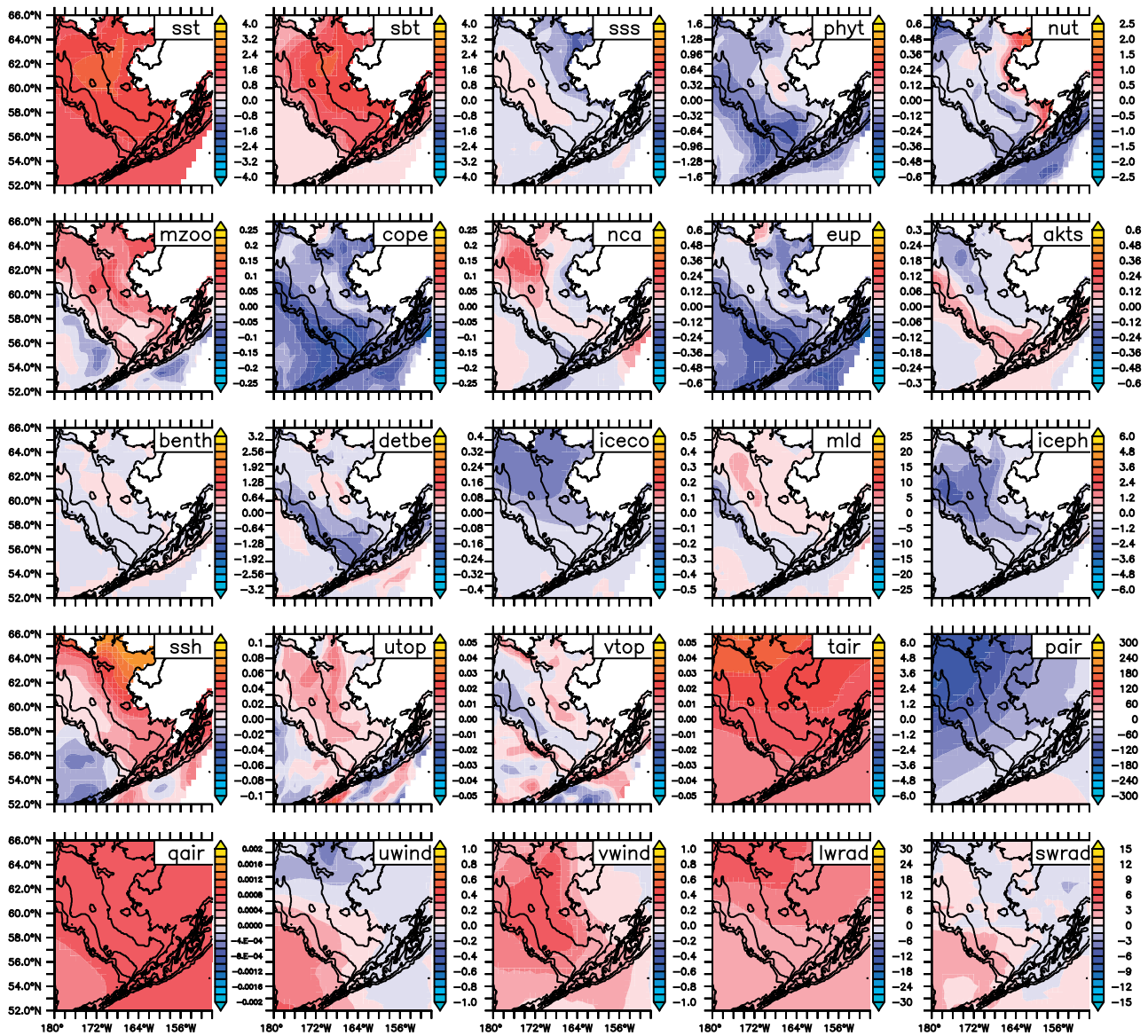
Figure 15. Predicted decadal averages changes between 2010–2019 and 2090–2099 under RCP4.5, using statistical extrapolation to include 35 ensemble members (weighted average of dynamical and statistical results). Colour version of this figure is available online at *ICESJM* online.

crustacean zooplankton over the outer shelf of the southeastern Bering Sea. Such impacts are substantially reduced (approximately by half) by 2100 under a moderate emission mitigation scenario (RCP4.5). These results suggests that future anthropogenic atmospheric carbon emissions will have a strong impact on Bering Sea physics and lower trophic level biology by the end of the twenty-first century. These are in turn expected to have strong consequences on the fisheries of the region; such impacts are being explored in fishery and socioeconomic models under ACLIM.

A hybrid multivariate method, based on a six-member dynamically downscaled ensemble, was used to statistically expand our ensemble to over 60 members (35 for RCP4.5 and 28 for RCP8.5). The multivariate analysis suggests two primary independent modes of variability, the primary one involving temperature and zooplankton productivity (among other covariates), and a secondary one involving air pressure and winds with fewer associated biological effects.

The expanded results from the hybrid dynamical-statistical method are similar in spatial pattern to those from dynamical downscaling alone, but provide an estimate of the magnitude of change we would have obtained if all 63 members were to be dynamically downscaled, and the associated uncertainty of those estimates due to both internal and cross-model (CMIP5) variability. As such it yields improved estimates of future conditions in the Bering Sea, which are useful in management strategy evaluation for fisheries of the region under high vs. low emission scenarios. Results thus far suggest that the projected mean changes of many biophysical attributes, from the present through 2100, exceed the internal variability or model-generated uncertainty of those estimates under RCP8.5.
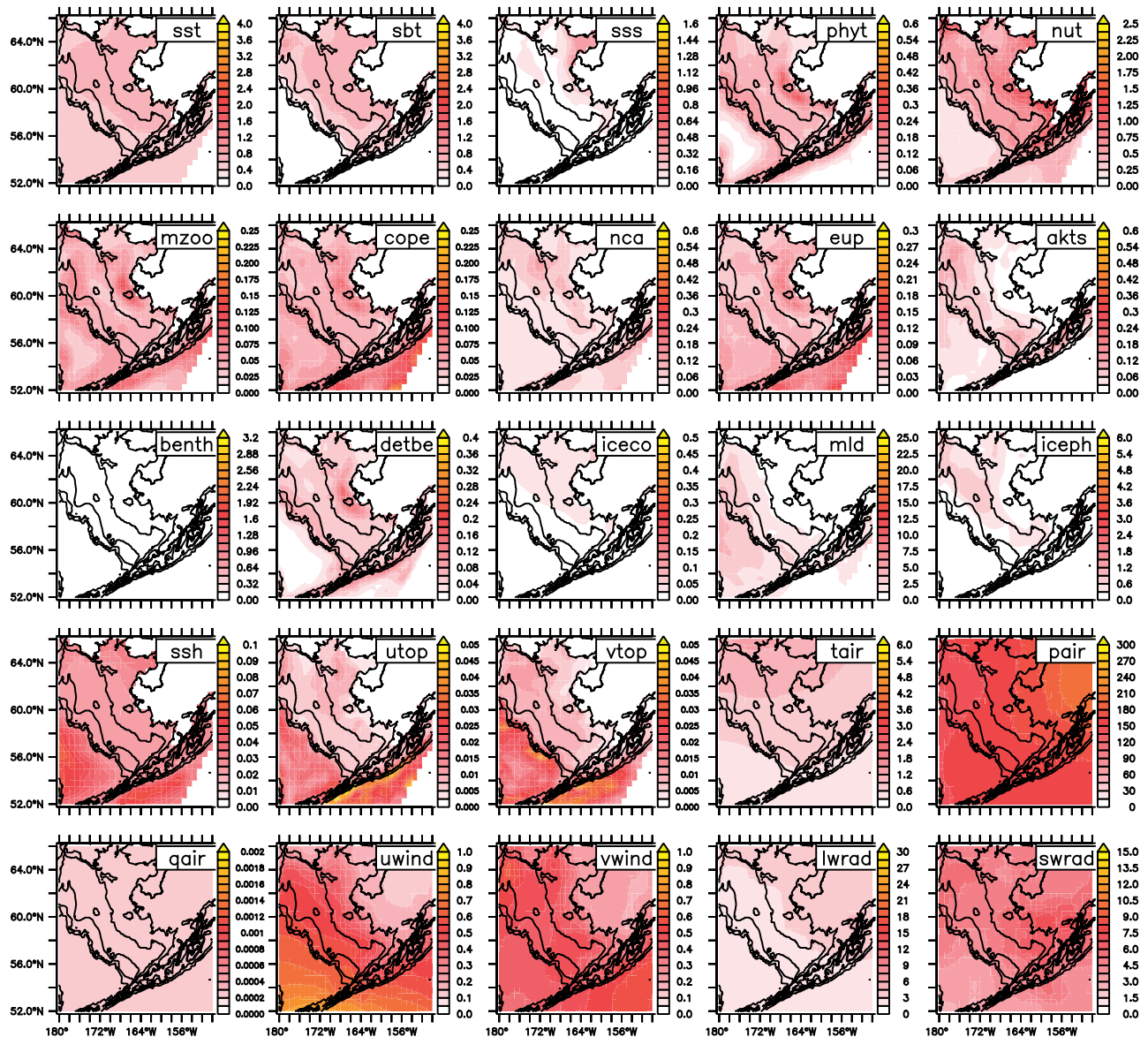
Acknowledgements

Figure 16. Internal (interannual) variability of each variable during 2090–2099, using statistical extrapolation to include 28 ensemble members from RCP 8.5 global projections (weighted average of dynamical and statistical results). Levels and colour key for each variable correspond to those used in plots of decadal change. Colour version of this figure is available online at *ICESJM* online.

Author contributions

A.J.H. and W.C. performed the dynamical downscaling, which includes NPZ model elements developed by G.A.G. A.J.H. developed and applied the hybrid statistical method and drafted the manuscript. All of the co-authors were involved in selecting the appropriate CMIP5 models for dynamical downscaling, and the interpretation of results. A.B.H. and K.K.H. are co-leads of the ACLIM project, and as such provided guidance on the necessary scope of the climate projections and their application to the suite of biological models.
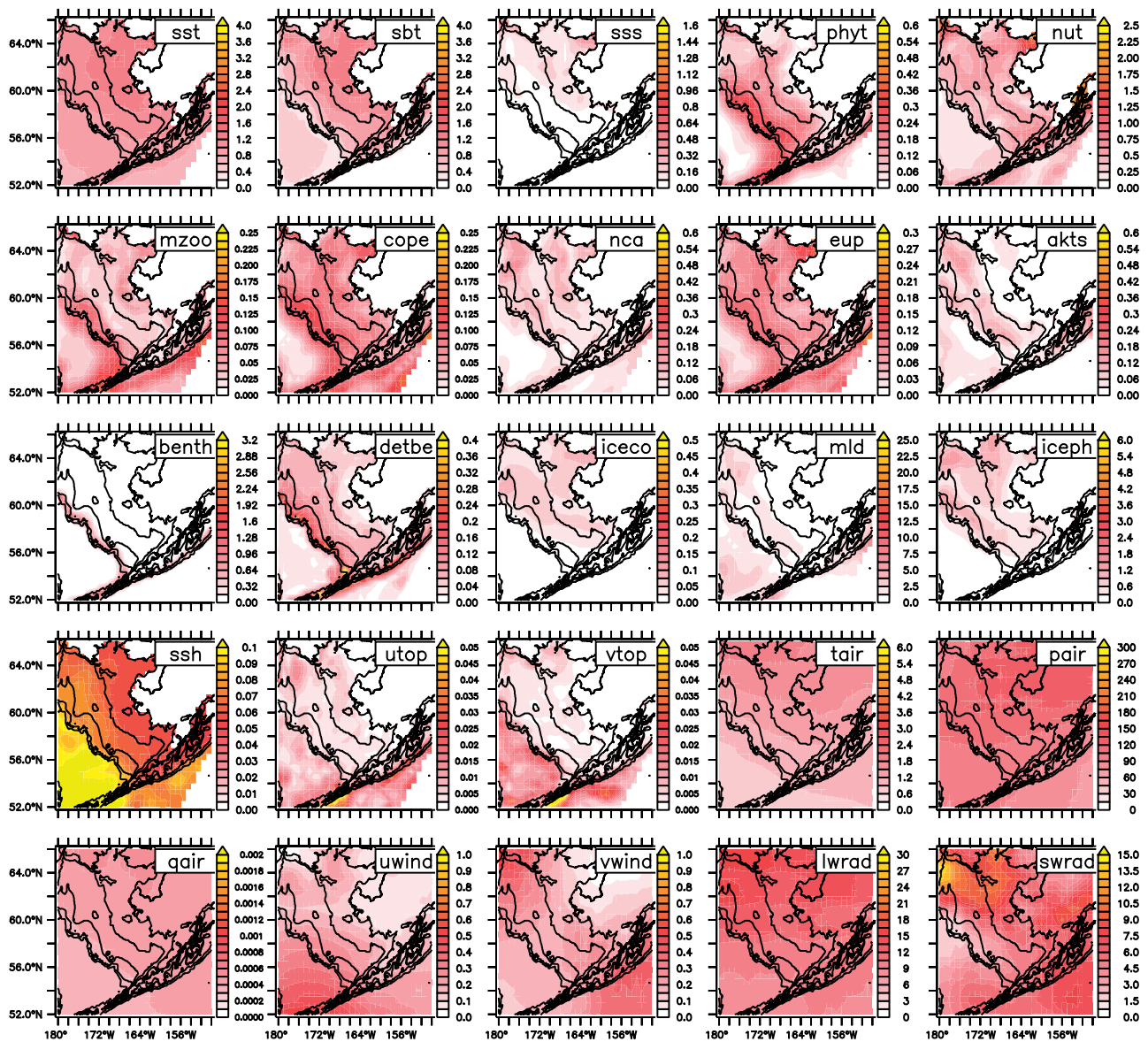
Figure 17. Cross-model variability of each variable during 2090–2099, using statistical extrapolation to include 28 ensemble members from RCP 8.5 global projections (weighted average of dynamical and statistical results). Levels and colour key for each variable correspond to those used in plots of decadal change. Colour version of this figure is available online at *ICESJM* online.

References

Alexander, M. A., Matrosova, L., Penland, C., Scott, J. D., and Chang, P. 2008. Forecasting pacific SSTs: linear inverse model predictions of the PDO. Journal of Climate, 21: 385–402.

Budgell, W. P. 2005. Numerical simulation of ice-ocean variability in the Barents Sea region: towards dynamical downscaling. Ocean Dynamics, 55: 370–387.

Bretherton, C. S., Smith, C., and Wallace, J. M. 1992. An intercomparison of methods for finding coupled patterns in climate data. Journal of Climate, 5: 541–560.

Brown, Z. W. v., Dijken, G. L., and Arrigo, K. R. 2011. A reassessment of primary production and environmental change in the Bering Sea. Journal of Geophysical Research, 116: C08014.

Capotondi, A., and Sardeshmukh, P. D. 2015. Optimal precursors of different types of ENSO events. Geophysical Research Letters, 42: 9952–9960.

Cheng, W., Curchitser, E., Ladd, C., Stabeno, P. J., and Wang, M. 2014. Influences of sea ice on the eastern Bering Sea: NCAR CESM simulations and comparison with observations. Deep Sea Research Part II: Topical Studies in Oceanography, 109: 27–38.

Coachman, L. K. 1986. Circulation, water masses, and fluxes on the southeastern Bering Sea shelf. Continental Shelf Research, 5: 23–108.

Coyle, K. O., Eisner, L. B., Mueter, F. J., Pinchuk, A. I., Janout, M. A., Cieciel, K. D., Farley, E. V. et al. 2011. Climate change in the southeastern Bering Sea: impacts on pollock stocks and implications for the Oscillating Control Hypothesis. Fisheries Oceanography, 20: 139–156.

Coyle, K. O., Pinchuk, A. I., Eisner, L. B., and 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 Research Part II: Topical Studies in Oceanography, 55: 1775–1791.

Dai, A., Qian, T., Trenberth, K. E., and Milliman, J. D. 2009. Changes in continental freshwater discharge from 1948–2004. Journal of Climate, 22: 2773–2791.

Danielson, S., Curchitser, E., Hedstrom, K., Weingartner, T., and Stabeno, P. 2011. On ocean and sea ice modes of variability in the Bering Sea. Journal of Geophysical Research, 116: C12034.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J. et al. 2012. GFDL's ESM2 global coupled climate-carbon earth system models. Part I: physical formulation and baseline simulation characteristics. Journal of Climate, 25: 6646–6665.

Fissel, B., Dalton, M., Garber-Yonts, B., Haynie, A., Kasperski, S., Lee, J., Lew, D. et al. 2017. Stock assessment and fishery evaluation report for the groundfish fisheries of the Gulf of Alaska and Bering Sea/Aleutian Islands area: Economic status of the groundfish fisheries of Alaska, 2016. NPFMC Bering Sea, Aleutian Islands and Gulf of Alaska SAFE, Seattle, WA. Retrieved from http://www.afsc.noaa.gov/refm/docs/2017/economic.pdf (last accessed 9 January 2019).

Gibson, G. A., and 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. Journal of Marine Systems, 88: 214–231.

Grebmeier, J. M. 2012. Shifting patterns of life in the Pacific Arctic and sub-Arctic seas. Annual Review of Marine Science, 4: 63–78.

Grebmeier, J. M., Bluhm, B. A., Cooper, L. W., Danielson, S. L., Arrigo, K. R., Blanchard, A. L., Clarke, J. T. et al. 2015. Ecosystem characteristics and processes facilitating persistent macrobenthic biomass hotspots and associated benthivory in the Pacific Arctic. Progress in Oceanography, 136: 92–114.

Haidvogel, D. B., Arango, H., Budgell, W. P., Cornuelle, B. D., Curchitser, E., Di Lorenzo, E., Fennel, K. et al. 2008. Regional ocean forecasting in terrain-following coordinates: model formulation and skill assessment. Journal of Computational Physics, 227: 3595–3624.

Hawkins, E., and Sutton, R. 2009. The potential to narrow uncertainty in regional climate predictions. Bulletin of the American Meteorological Society, 90: 1095–1108.

Haynie, A. C., and Huntington, H. P. 2016. Strong connections, loose coupling: the influence of the Bering Sea ecosystem on commercial fisheries and subsistence harvests in Alaska. Ecology and Society, 21: 6.

Hermann, A. J., Gibson, G. A., Bond, N. A., Curchitser, E. N., Hedstrom, K., Cheng, W., Wang, M. et al. 2013. A multivariate analysis of observed and modeled biophysical variability on the Bering Sea shelf: multidecadal hindcasts (1970–2009) and forecasts (2010–2040). Deep Sea Research Part II: Topical Studies in Oceanography, 94: 121–139.

Hermann, A. J., Gibson, G. A., Bond, N. A., Curchitser, E. N., Hedstrom, K., Cheng, W., Wang, M. et al. 2016a. Projected future biophysical states of the Bering Sea. Deep-Sea Research Part II: Topical Studies in Oceanography, 134: 30–47.

Hermann, A. J., Ladd, C., Cheng, W., Curchitser, E. N., and Hedstrom, K. 2016b. A model-based examination of multivariate physical modes in the eastern and western Gulf of Alaska. Deep Sea Research Part II: Topical Studies in Oceanography, 132: 68–89.

Holsman, K., Hollowed, A., Ito, S.-I., Bograd, S., Hazen, E., King, J., Mueter, F. et al. 2018. Climate change impacts, vulnerabilities and adaptations: North Pacific and Pacific Arctic marine fisheries. In Impacts of Climate Change on Fisheries and Aquaculture: Synthesis of Current Knowledge, Adaptation and Mitigation Options, Fisheries and Aquaculture Technical Paper 627, pp. 113–138. Ed. by M. Barange, T. Bahri, M. C. M. Beveridge, K. L. Cochrane, S. Funge-Smith, and F. Poulain. FAO, Rome.

Hunt, Jr, G. L. , Coyle, K. O., Eisner, L., Farley, E. V., Heintz, R., Mueter, F., Napp, J. M. et al. 2011. Climate impacts on eastern

Bering Sea foodwebs: a synthesis of new data and an assessment of the Oscillating Control Hypothesis. ICES Journal of Marine Science, 68: 1230–1243.

Hunt, Jr, G. L. , Stabeno, P., Walters, G., Sinclair, E., Brodeur, R. D., Napp, J. M., and Bond, N. A. 2002. Climate change and control of the southeastern Bering Sea pelagic ecosystem. Deep Sea Research Part II: Topical Studies in Oceanography, 49: 5821–5853.

IPCC. 2013. Climate change 2013: the physical science basis. In Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Ed. by Stocker T. F., Qin D., Plattner G.-K., Tignor M. M. B., Allen S. K., Boschung J., Nauels A. et al. Cambridge University Press, Cambridge, UK and New York. 1535 pp. https://doi.org/10.1017/CBO9781107415324 (last accessed 9 January 2019).

IPCC. 2014. Climate change 2014: impacts, adaptation, and vulnerability, Part. B: regional aspects. In Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Ed. by Barros V. R., Field C. B., Dokken D. J., Mastrandrea M. D., Mach K. J., Bilir T. E., Chatterjee M. et al. Cambridge University Press, Cambridge, UK and New York. 688 pp.

Kachel, N. B., Hunt, Jr, G. L. , Salo, S. A., Schumacher, J. D., Stabeno, P. J., and Whitledge, T. E. 2002. Characteristics and variability of the inner front of the southeastern Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 49: 5889–5909.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M. et al. 2015. The Community Earth System Model (CESM) large ensemble project: a community resource for studying climate change in the presence of internal climate variability. Bulletin of the American Meteorological Society, 96: 1333–1349.

Kinder, T. H., and Schumacher, J. D. 1981. Hydrographic structure over the continental shelf of the southeastern Bering Sea. In Eastern Bering Sea Shelf: Oceanography and Resources, Vol. 1, pp. 31–51. Ed. by D.W. Hood, and J. A. Calder. USDOC/NOAA/OMPA, Washington, D.C.

Knutti, R., Masson, D., and Gettelman, A. 2013. Climate model genealogy: generation CMIP5 and how we got there. Geophysical Research Letters, 40: 1194–1199.

Large, W. G., McWilliams, J. C., and Doney, S. C. 1994. Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. Reviews of Geophysics, 32: 363–403.

Large, W. G., and Yeager, S. G. 2009. The global climatology of an interannually varying air-sea flux data set. Climate Dynamics, 33: 341–364.

Li, L., McClean, J., Miller, A., Eisenman, I., Hendershott, M., and Papadopoulos, C. 2014. Processes driving sea ice variability in the Bering Sea in an eddying ocean/sea ice model: mean seasonal cycle. Ocean Modelling, 84: 51–66.

Li, L., Miller, A., McClean, J., Eisenman, I., and Hendershott, M. 2014. Processes driving sea ice variability in the Bering Sea in an eddying ocean/sea ice model: anomalies from the mean seasonal cycle. Ocean Dynamics, 64: 1693–1717.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J. et al. 2007. The WCRP CMIP3 multimodel dataset: a new era in climate change research. Bulletin of the American Meteorological Society, 88: 1383–1394.

Newman, M., Sardeshmukh, P. D., Winkler, C. R., and Whitaker, J. S. 2003. A study of subseasonal predictability. Monthly Weather Review, 131: 1715.

Ortiz, I., Aydin, K., Hermann, A. J., Gibson, G., Punt, A. E., Wiese, F., Eisner, L. B. et al. 2016. Climate to fish: synthesizing field work, data and models in a 39-year retrospective analysis of seasonal processes on the eastern Bering Sea shelf and slope. Deep Sea Research Part II: Topical Studies in Oceanography, 134: 390–412.

Preisendorfer, R. W. 1988. Principal Component Analysis in Meteorology and Oceanography, Elsevier, Amsterdam. 425 pp.

Ressler, P. H., De Robertis, A., and Kotwicki, S. 2014. The spatial distribution of euphausiids and walleye pollock in the eastern Bering

Sea does not imply top-down control by predation. Marine Ecology Progress Series, 503: 111–122.

Ressler, P. H., De Robertis, A., Warren, J. D., Smith, J. N., and Kotwicki, S. 2012. Developing an acoustic index of euphausiid abundance to understand trophic interactions in the Bering Sea ecosystem. Deep Sea Research Part II: Topical Studies in Oceanography, 65–70: 184–195.

Rykaczewski, R. R., and Dunne, J. P.2010. Enhanced nutrient supply to the California Current Ecosystem with global warming and increased stratification in an earth system model. Geophysical Research Letters, 37: L21606.

Shchepetkin, A. F., and McWilliams, J. C. 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling, 9: 347–404.

Siddon, E., and Zador, S. 2017. Ecosystem Considerations 2017: Status of the Eastern Bering Sea Marine Ecosystem, Stock Assessment and Fishery Evaluation Report. North Pacific Fishery Management Council, Anchorage, AK.

Sigler, M. F., Napp, J. M., Stabeno, P. J., Heintz, R. A., Lomas, M. W., and Hunt, G. L. 2016. Variation in annual production of copepods, euphausiids, and juvenile walleye pollock in the southeastern Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 134: 223–234.

Stabeno, P. J., Danielson, S., Kachel, D., Kachel, N. B., and Mordy, C. W. 2016. Currents and transport on the eastern Bering Sea shelf: an integration of over 20 years of data. Deep Sea Research Part II: Topical Studies in Oceanography, 134: 13–29.

Stabeno, P. J., Duffy-Anderson, J. T., Eisner, L. B., Farley, E. V., Heintz, R. A., and Mordy, C. W. 2017. Return of warm conditions in the southeastern Bering Sea: physics to fluorescence. PLoS One, 12: e0185464.

Stabeno, P. J., Kachel, N. B., Moore, S. E., Napp, J. M., Sigler, M., Yamaguchi, A., and Zerbini, A. N. 2012. Comparison of warm and cold years on the southeastern Bering Sea shelf and some implications for the ecosystem. Deep Sea Research Part II: Topical Studies in Oceanography, 65–70: 14–30.

Stabeno, P. J., Napp, J., Mordy, C., and Whitledge, T. 2010. Factors influencing physical structure and lower trophic levels of the eastern Bering Sea shelf in 2005: sea ice, tides and winds. Progress in Oceanography, 85: 180–196.

Sullivan, M. E., Kachel, N. B., Mordy, C. W., Salo, S. A., and Stabeno, P. J. 2014. Sea ice and water column structure on the eastern Bering sea shelf. Deep Sea Research Part II: Topical Studies in Oceanography, 109: 39–56.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. 2012. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93: 485–498.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T. *et al.* 2011. MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. Geoscientific Model Development, 4: 845–872.

Appendix

The steps of our hybrid statistical procedure, based on dynamically downscaled output, may be summarized as follows:

1. Calculate individual spatial modes (EOFs) of each variable with associated time series (PCs). For this analysis, we use the series of annual averages at each bin prepared as described in "Univariate analyses" section, with each of the ensemble members concatenated together after removal of the time mean.

$$V_{ilt} = \sum_j X_{jil} T_{jit}$$

where V represents the demeaned and concatenated data, X represents the spatial EOFs, T represents the modulating time series, and the indices i, j, l, t represent variable type i, spatial location l, mode number j, and time index t. In this decomposition, T has zero mean and unit variance.

2. Calculate "multivariate PCs" using the univariate Ts as the input set of time series. (In practice the best results are obtained when Ts are each scaled by the percent variance of V explained by that univariate mode). We are now calculating EOFs across variables, rather than across space, decomposing the univariate Ts into a single multivariate set of time functions Γ and variable loadings M:

$$T_{jit} = \sum_k M_{kji} \Gamma_{kt}$$

where k represents the multivariate mode number. This ultimately yields a new set of multivariate PCs (multivariate time series) with associated loadings (multivariate spatial modes) for each variable:

$$V_{ilt} = \sum_j X_{jil} \sum_k M_{kji} \Gamma_{kt} = \sum_k \Gamma_{kt} \sum_j M_{kji} X_{jil} = \sum_k C_{kil} \Gamma_{kt}$$

Note how the new multivariate spatial modes C are simply a linear combination of the univariate spatial modes X, weighted by

M. This new basis set retains the orthogonality property of the original EOF basis set for each variable, as it entails the multiplication of one orthogonal matrix by another. Hence it can in theory be used to fully reconstruct the original data, in a manner which compactly emphasizes the covariance among the dominant spatial patterns of different variables.

3. For convenience, we convolve each of the multivariate time series with the original time series for each variable at each spatial bin to obtain the new multivariate spatial modes:

$$C_{kil} = \sum_t \Gamma_{kt} V_{ilt}$$

In general, these will look different than the original univariate EOFs, as they represent the dominant spatial patterns of variables rising and falling together through time.

4. Now use the multivariate modes derived from the training set as predictors of the regional response, given only a new set of coarse, large-scale atmospheric/oceanic forcing. We project the new set of available forcing variables (in our application these are atmospheric variables from other CMIP5 members) onto the multivariate basis set at each time step, to obtain the amplitude of each mode through time. Ideally, this allows us to estimate the covarying regional variables not contained in that new set, as a proxy for what would have appeared if we had conducted a full dynamical downscaling of that forcing through the regional model. For each forcing variable i at future time t we construct the estimate as follows:

$$\Gamma_{ikt}^* = \left[\sum_l \left(V_{ilt} C_{kil} \right) \right] / \left[\sum_l \left(C_{kil} C_{kil} \right) \right]$$

Consider each forcing variable as an independent estimate of the modal amplitude. Some of these estimates are likely to be more accurate, as some of the variables are captured more fully by that mode. Hence for our final estimate of the modal amplitudes at each future time, we weight the individual estimates by the percent variance of that variable explained by that multivariate mode:

$$R_{ik} = \left[\sum_{lt} \left(V_{ilt} C_{kil} \right) \right] / \left[\sum_{lt} \left(V_{ilt} V_{ilt} \right) \sum_{lt} \left(C_{kil} C_{kil} \right) \right]^{1/2}$$

$$\Gamma_{kt}^{**} = \left[\sum_{i} \left(\Gamma_{ikt}^{*} R_{ik} \right) \right] / \sum_{i} R_{ik}$$

So our estimate of the "unobserved" (regional biophysical response) variables at some future time, based on the projection of "observed" (large-scale atmospheric forcing) variables onto the multivariate modes, becomes:

$$Vest_{ilt} = \sum_{k} C_{kil} \Gamma_{kt}^{**}$$

This estimate can be further tuned by linearly regressing each $Vest_{ilt}$ against the original V_{ilt} over the original time sequence t; the slopes b_{il} of those regressions may be used in our final estimates:

$$Vfinal_{ilt} = b_{il} Vest_{ilt}$$

In practice, it was found that this tuning made little difference to the results.

Handling editor: Shubha Sathyendranath