A rapid multivariate method for estimating regional forecast uncertainty

Albert J. Hermann
UW-JISAO/NOAA-PMEL
hermann@uw.edu

MANY others involved with models shown here!

General notions of predictability

- The ability to predict is what distinguishes science from non-science
  - no testable predictions = no science!
- We cannot set an upper limit on predictability, only a lower one (how well can we do now?)
- Detailed predictions are usually more difficult than general ones:
  - "it will be rainy next summer in WA"

    vs.
  - "it will rain at 2pm July 14 next summer in Seattle"

Methods of Prediction

- Methods:
  - Direct numerical modeling
  - Analogues (find a similar situation from the past)
  - Eulerian persistence (inertia)
  - Lagrangian persistence (advection or waveguide)
- Tools:
  - hydrodynamic/biological models
  - simple correlations
  - EOFs/PCAs – correlations of patterns

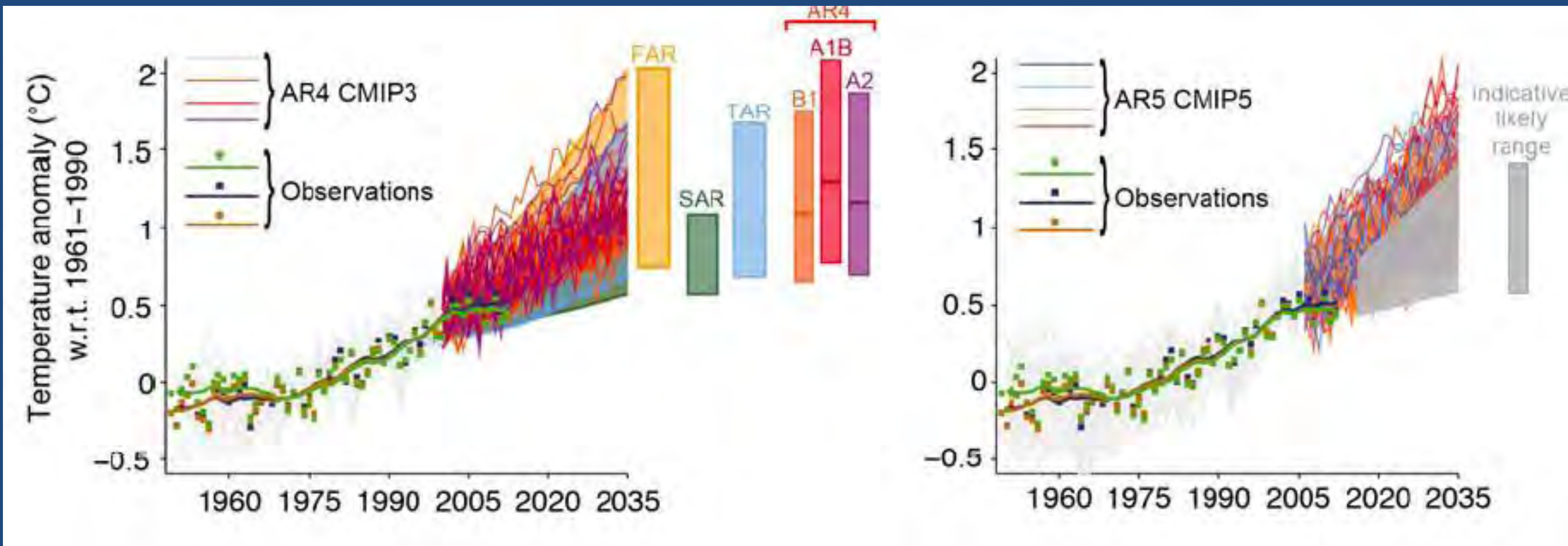
Should predictability get better at larger space and time scales?

- Yes ->
  - because there is always small-scale signal (turbulence, eddies) which cannot be predicted
  - Expect more success with *recurring patterns* (e.g. EOFs)
- No ->
  - spatial average may integrate over a local region with high predictability
  - temporal average may integrate over strongly forced oscillations (annual cycle, tides)

Model-based downscaling

- is a popular approach used to forecast the effects of global climate change on regional seas and their ecosystems over multiple decades.

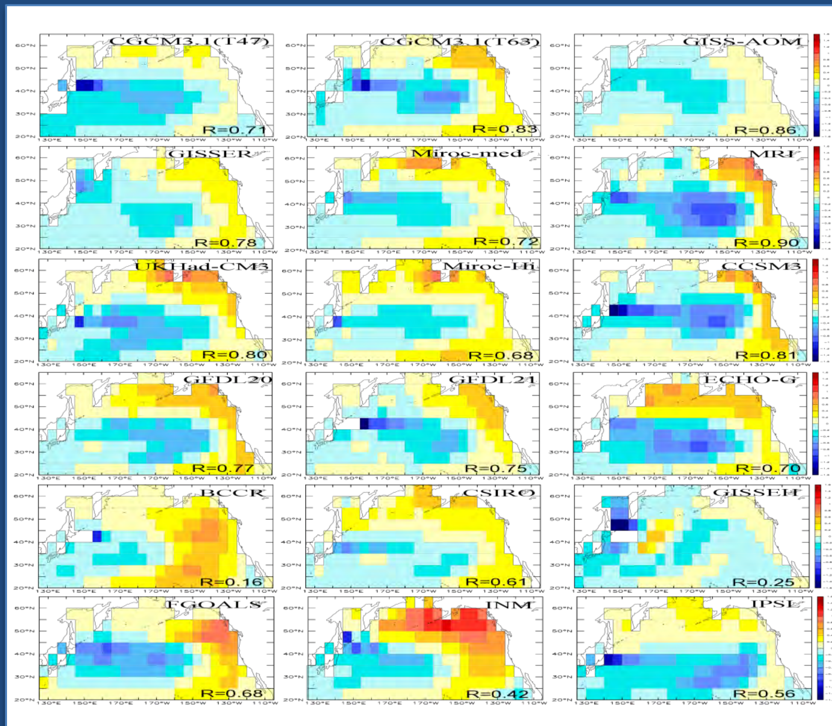
There is an inherent uncertainty in the global projections themselves!



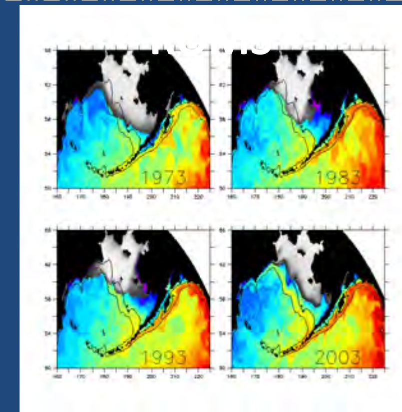
Individual realizations of a global forecast model can be applied one-by-one as forcing to a regional model.

- this generates a regional forecast ensemble which retains all of the nonlinear processes in both global and regional simulations.
- However, *computational and human resources* presently limit the size of this ensemble, since the regional models typically utilize a finer spatial grid, a smaller time step, and more biological detail than their global counterparts

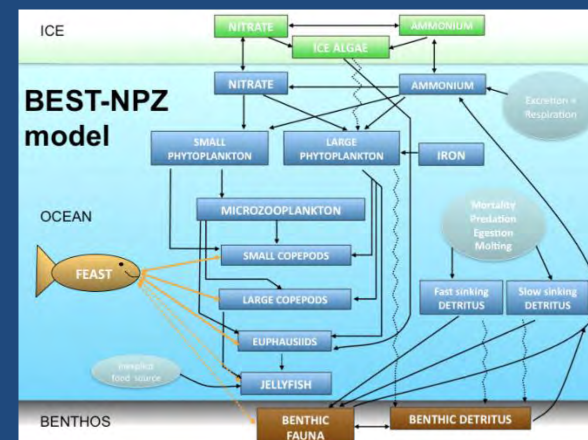
Climate models provide BCs/ICs to *regional coupled models*



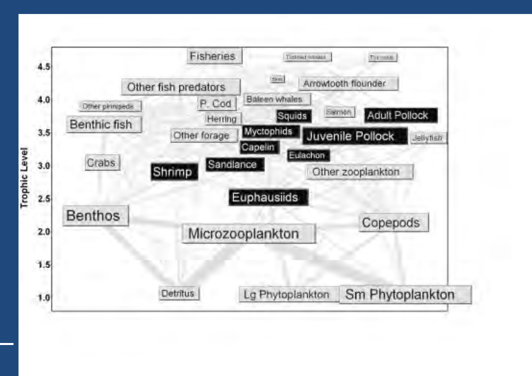
ensemble of runs



NPZ



FOOD WEB (FEAST)



GOAL:
mutidecadal
projections of physics and biology in the Bering Sea

ensemble of projected futures

Reduce Dimensionality with Multivariate Statistics

- Summarize covarying modes of behavior in large- and small-scale models
- Convolve the large-scale patterns with multiple realizations of the large-scale forecast
- Interpret resulting time series as realizations of the covarying small-scale phenomena. Get mean trajectory and spread of small-scale "forecast"

Idealized examples

- Hit a bell (forcing) -> it rings (response)
  - Space/time pattern of impact is the forcing: normal modes of the bell are the response
  - Look for coupled mode of hitting and ringing
  - Convolve the hitting pattern with a forecast of hits to get a forecast of rings
- Large-scale winds (forcing) -> upwelling (response)
  - Look for coupled mode which has large-scale pattern associated with fine-scale coastal conveyor belt response (upwelling -> nutrients -> phytoplankton -> zooplankton)
  - Convolve that large-scale pattern with a single global forecast realization to infer upwelling "realization"
  - Repeat for many forecasts to get many "realizations" of the future upwelling response

How to get useful coupled modes?

- Coupled Correlation Analysis (CCA) and related methods can be used to link covarying spatial and temporal patterns
  - But usually requires being on the same spatial grid
- Coupled PCA is more flexible – can have an arbitrary collection of time series
  - As in classical EOFs may still need space/time average to avoid quadrature
  - Beware of domination by a single variable
  - May need to rotate factors for proper interpretation

“PRINCIPAL COMPONENTS” use multivariate timeseries at a single location:

$N(x_1,t_1)$, $N(x_1,t_2)$, $N(x_1,t_3)$,……$N(x_1,t_N)$
$P(x_1,t_1)$, $P(x_1,t_2)$, $P(x_1,t_3)$,……$P(x_1,t_N)$
$Z(x_1,t_1)$, $Z(x_1,t_2)$, $Z(x_1,t_3)$,……$Z(x_1,t_N)$

OR a set of multivariate observations scattered in space and time

$N(x_1,t_1)$, $N(x_2,t_2)$, $N(x_3,t_3)$,……$N(x_M,t_N)$
$P(x_1,t_1)$, $P(x_2,t_2)$, $P(x_3,t_3)$,……$P(x_M,t_N)$
$Z(x_1,t_1)$, $Z(x_2,t_2)$, $Z(x_3,t_3)$,……$Z(x_M,t_N)$

“EMPIRICAL ORTHOGONAL FUNCTIONS” use timeseries of a single variable at multiple locations:

$T(x_1,t_1)$, $T(x_1,t_2)$, $T(x_1,t_3)$,……$T(x_1,t_N)$
$T(x_2,t_1)$, $T(x_2,t_2)$, $T(x_2,t_3)$,……$T(x_2,t_N)$
$T(x_3,t_1)$, $T(x_3,t_2)$, $T(x_3,t_3)$,……$T(x_3,t_N)$

COUPLED PCA ("MULTIVARIATE EOFS") use spatially gridded multivariate timeseries:

$N(x_1,t_1)$, $N(x_1,t_2)$, $N(x_1,t_3)$,……$N(x_1,t_N)$
$P(x_1,t_1)$, $P(x_1,t_2)$, $P(x_1,t_3)$,……$P(x_1,t_N)$
$Z(x_1,t_1)$, $Z(x_1,t_2)$, $Z(x_1,t_3)$,……$Z(x_1,t_N)$
$T(x_1,t_1)$, $T(x_1,t_2)$, $T(x_1,t_3)$,……$T(x_1,t_N)$

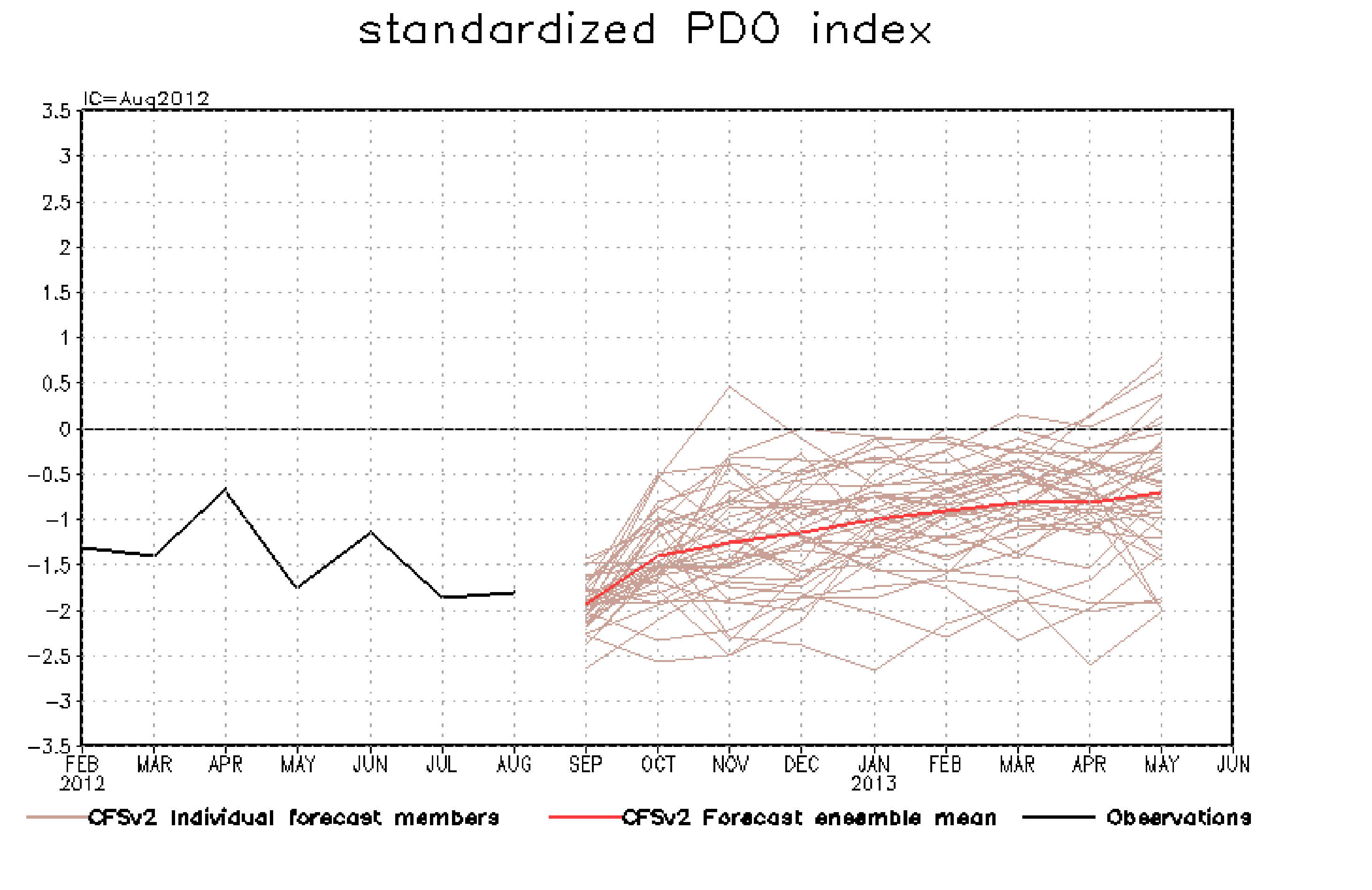
$N(x_2,t_1)$, $N(x_2,t_2)$, $N(x_2,t_3)$,……$N(x_2,t_N)$
$P(x_2,t_1)$, $P(x_2,t_2)$, $P(x_2,t_3)$,……$P(x_2,t_N)$
$Z(x_2,t_1)$, $Z(x_2,t_2)$, $Z(x_2,t_3)$,……$Z(x_2,t_N)$
$T(x_2,t_1)$, $T(x_2,t_2)$, $T(x_2,t_3)$,……$T(x_2,t_N)$

.
.
.
$N(x_M,t_1)$, $N(x_M,t_2)$, $N(x_M,t_3)$,……$N(x_M,t_N)$
$P(x_M,t_1)$, $P(x_M,t_2)$, $P(x_M,t_3)$,……$P(x_M,t_N)$
$Z(x_M,t_1)$, $Z(x_M,t_2)$, $Z(x_M,t_3)$,……$Z(x_M,t_N)$
$T(x_M,t_1)$, $T(x_M,t_2)$, $T(x_M,t_3)$,……$T(x_M,t_N)$

Summary of Technique

- See if we can get significant variance in a coupled (CPCA) mode which includes BOTH large-scale (global) and small-scale (downscaled) patterns
- IF have such a powerful mode, *convolve the large-scale part with large-scale forecast output* to get an inferred time series of small-scale output.

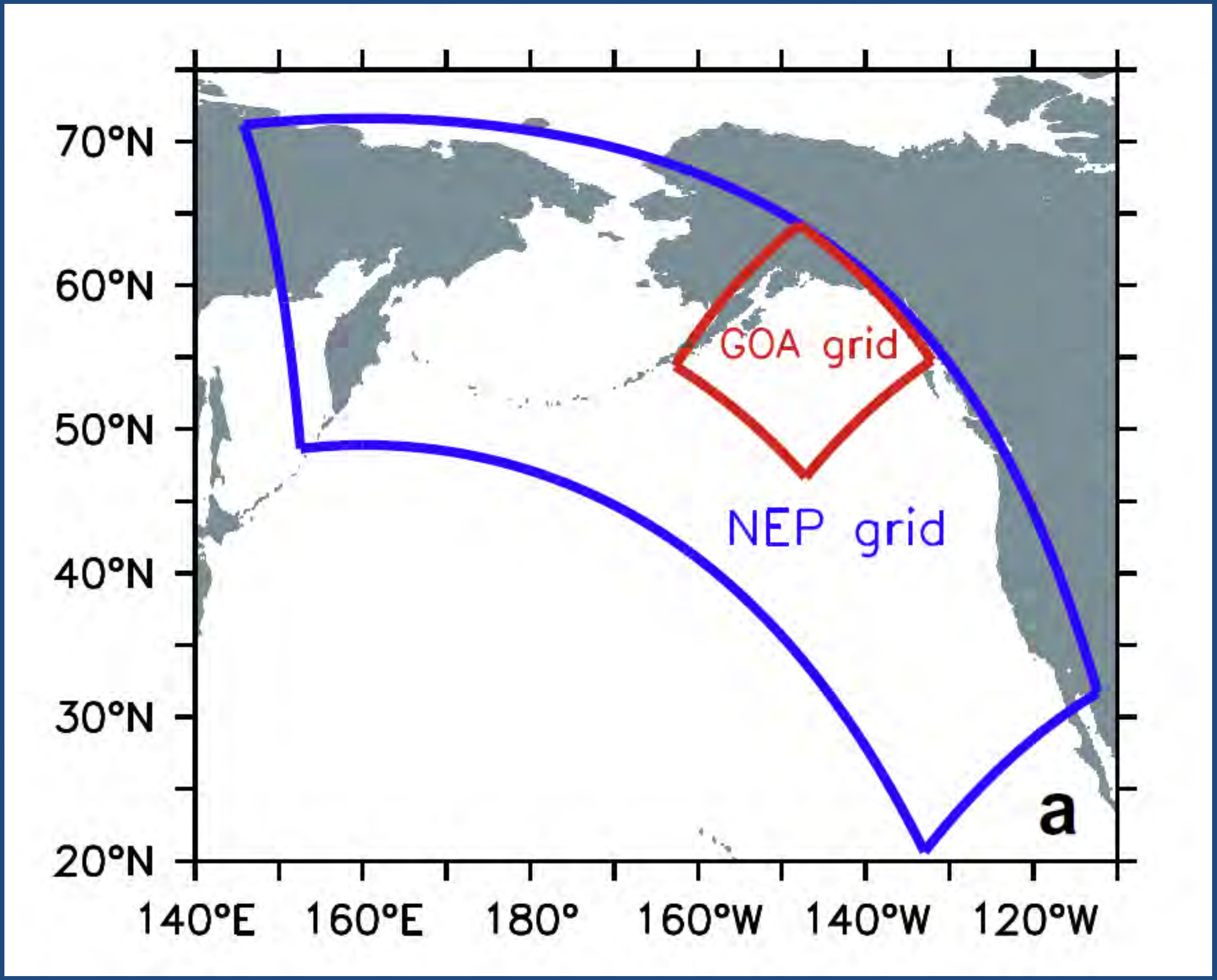
Note: big modeling centers already convolve popular EOFs, e.g. NCEP Climate Forecast System predictions of PDO



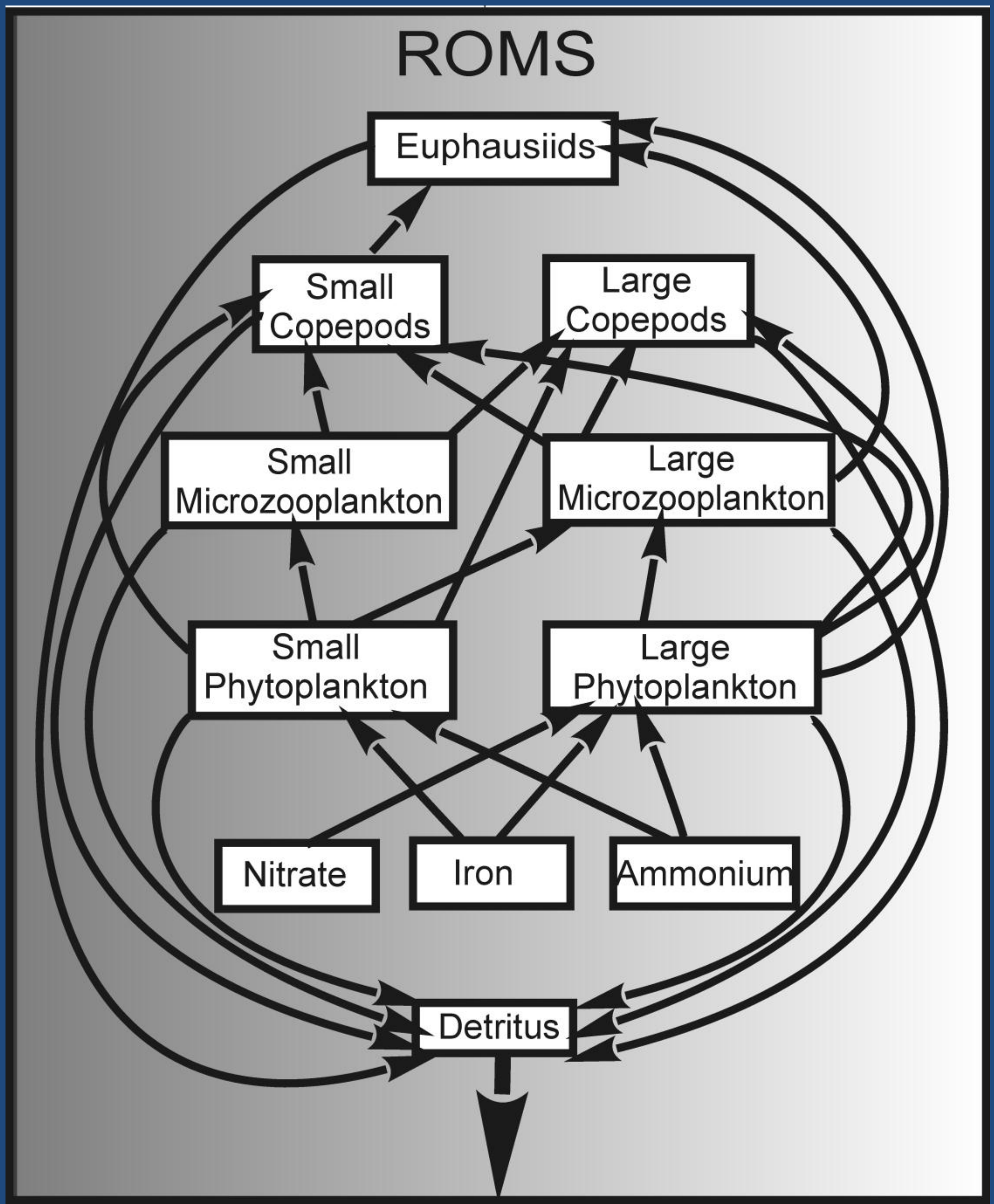
A few regional models (w/ resolution) which may benefit from this approach

1. GOA physical (10 km) and biophysical (3 km)
2. Bering Sea biophysical (10 km)
3. Pacific Northwest biophysical (1.5 km)

Example 1: GOA regional models (NEP=10km resolution, GOA = 3 km resolution)



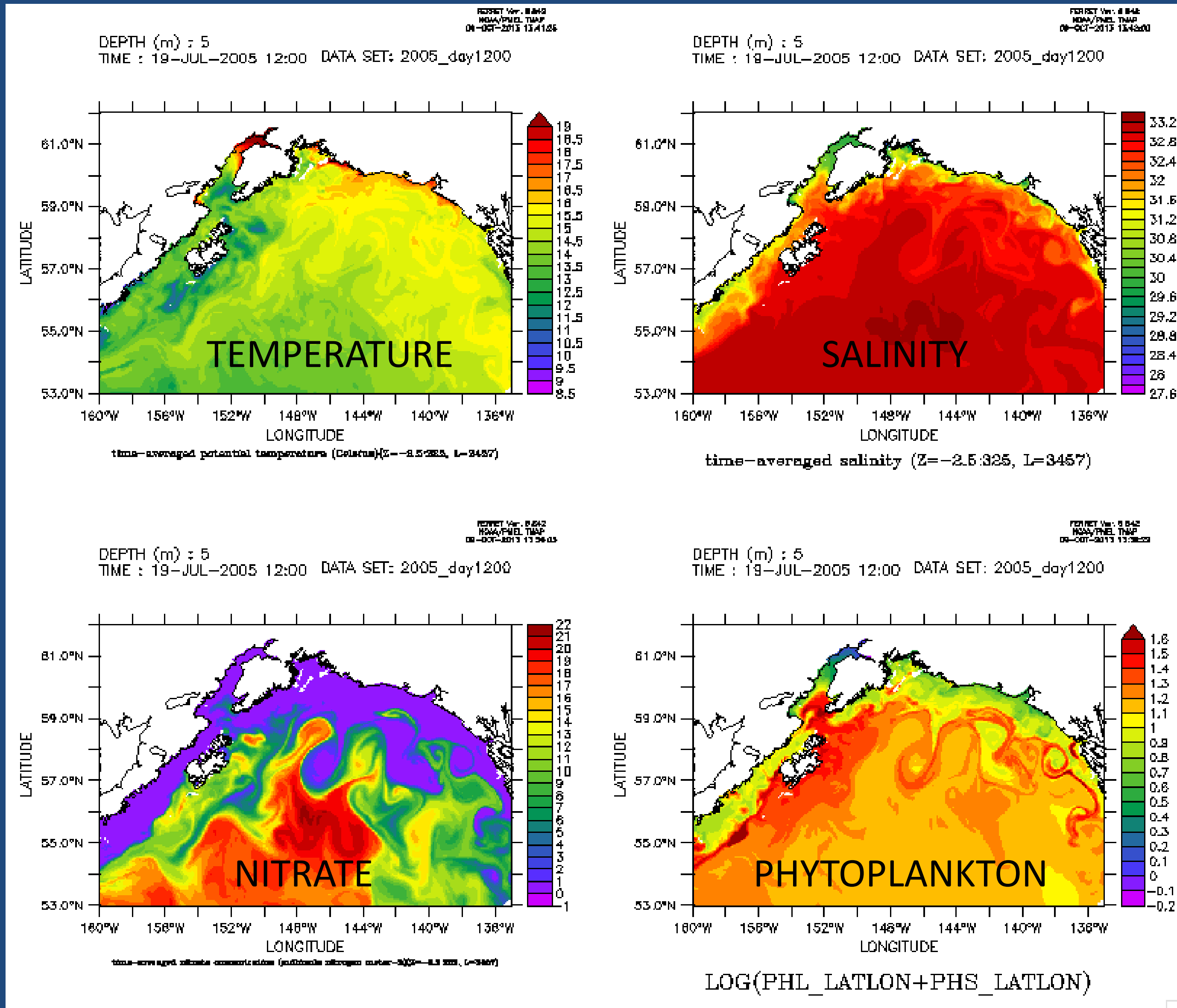
GOA-IERP regional biological model



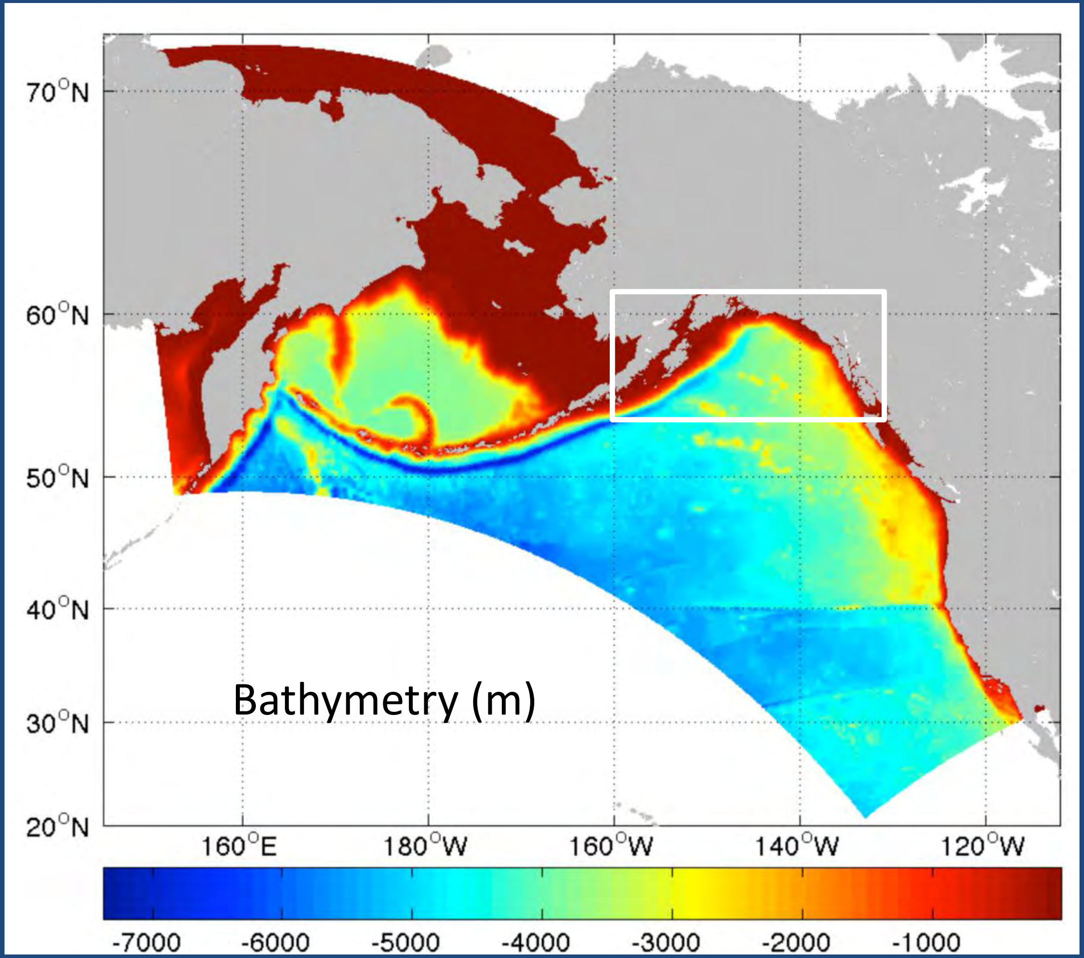
Hindcast run 1995-2012

New pilot project will be adding carbonate and pH!

Example of GOA biophysical output



CPCA has been performed on GOA subsection of NEP domain (NEP5 model of Curchiter, Hedstrom, Danielson)



GOA MODEL VARIABLES ANALYZED

SURFACE SCALARS

- Surface Temperature (SST)
- Surface Height (SSH)
- Surface Salinity (SSS)
- Mixed Layer Depth (SBLD)

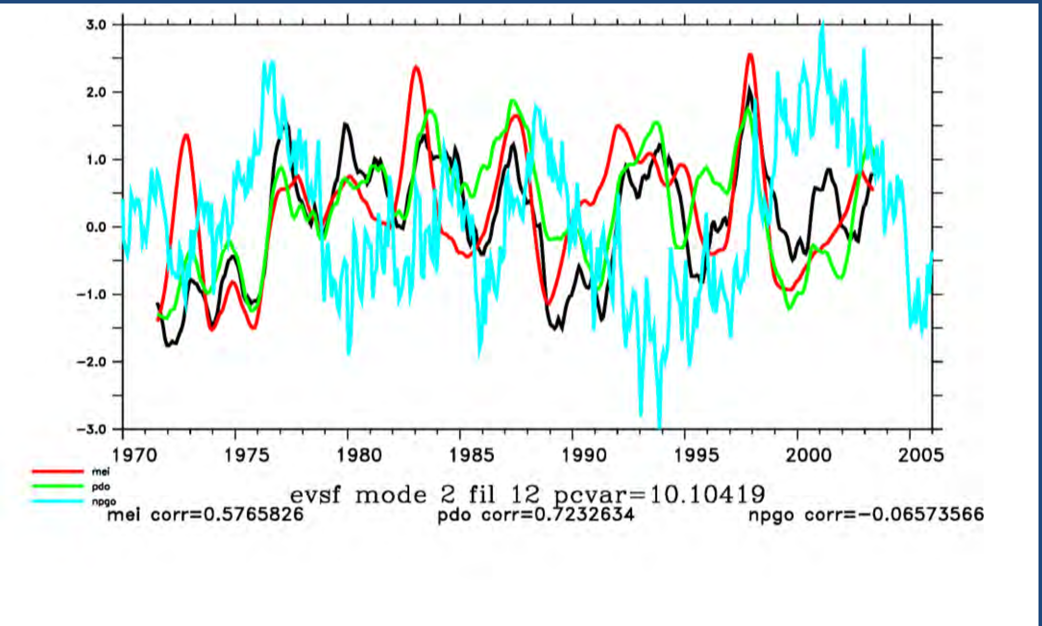
HEAT FLUX (into the ocean)

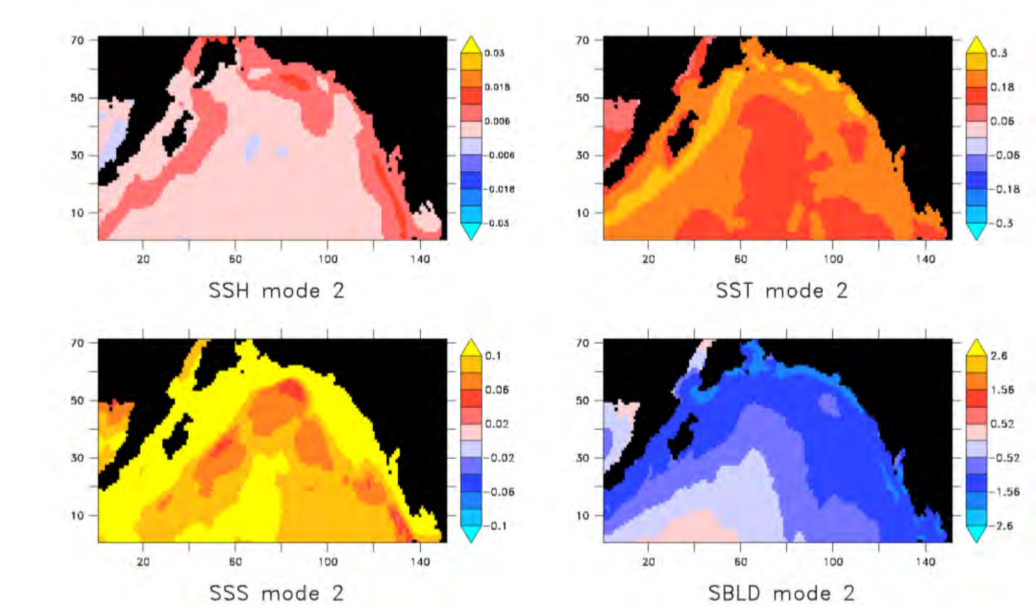
- Latent Heat Flux (LATEN)
- Sensible Heat Flux (SENSI)
- Net Longwave (LW)
- Shortwave (SW)

STRESS and VELOCITY

- E and N Velocity (UV)
- E and N Windstress (UVSTRESS)

Multivariate EOF results based on NEP HINDCAST





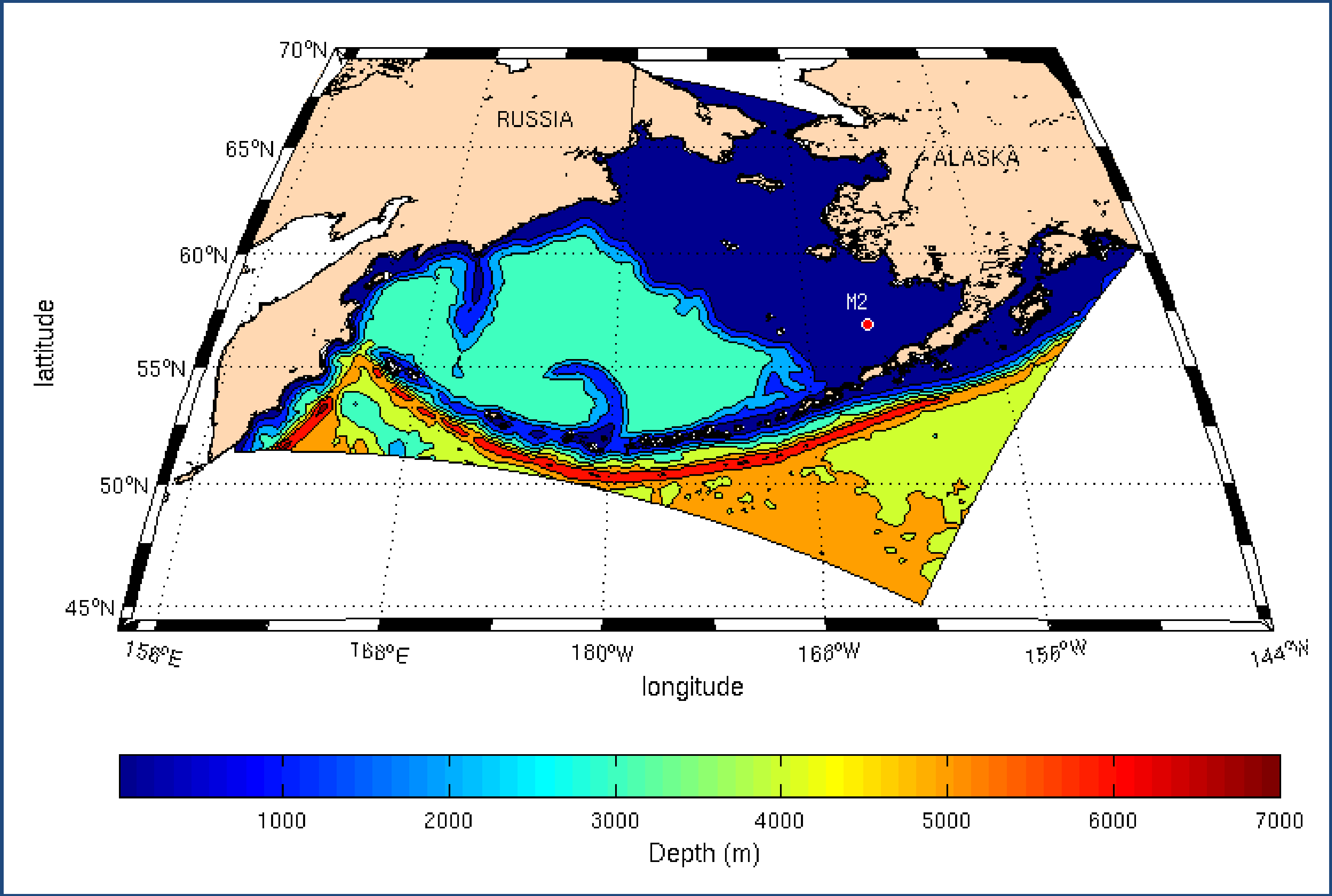
ENSO/PDO are strongly correlated to variability of the Gulf of Alaska

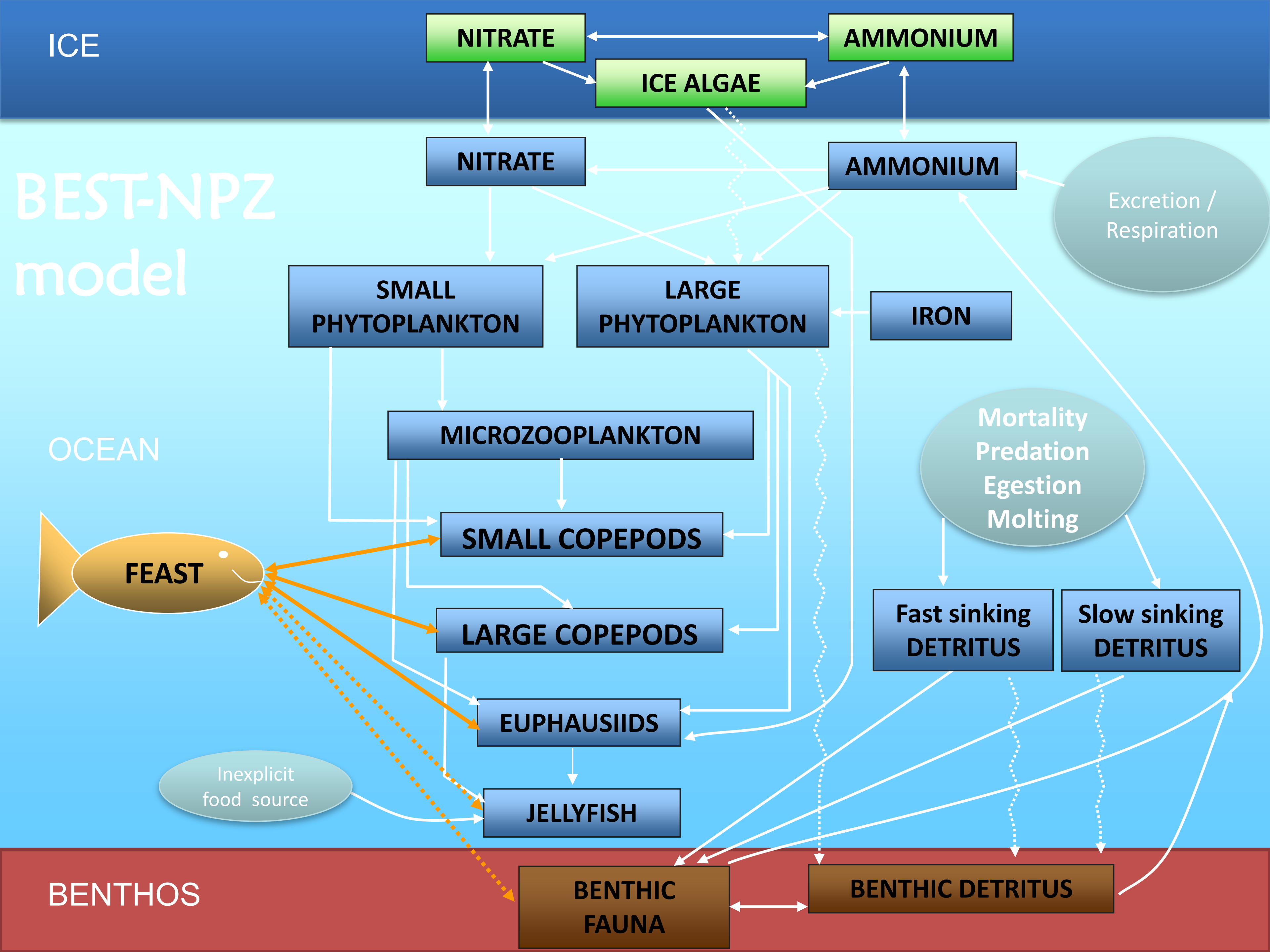
Positive ENSO/PDO is associated with:

- stronger northeastward wind stress
- stronger Alaska Gyre circulation
- higher coastal sea surface height
- warmer sea surface temperatures
- deeper coastal mixed layer depths
- enhanced shortwave radiation
- enhanced sensible heat losses

Example 2: Bering Sea regional model

10 layers, 10-km grid
Includes ice and tides
CCSM heat flux algorithms




ICE
NITRATE
AMMONIUM
ICE ALGAE
NITRATE
AMMONIUM
BEST-NPZ model
Excretion / Respiration
SMALL PHYTOPLANKTON
LARGE PHYTOPLANKTON
IRON
MICROZOOPLANKTON
OCEAN
Mortality
Predation
Egestion
Molting
SMALL COPEPODS
FEAST
LARGE COPEPODS
Fast sinking DETRITUS
Slow sinking DETRITUS
EUPHAUSIIDS
Inexplicit food source
JELLYFISH
BENTHOS
BENTHIC FAUNA
BENTHIC DETRITUS

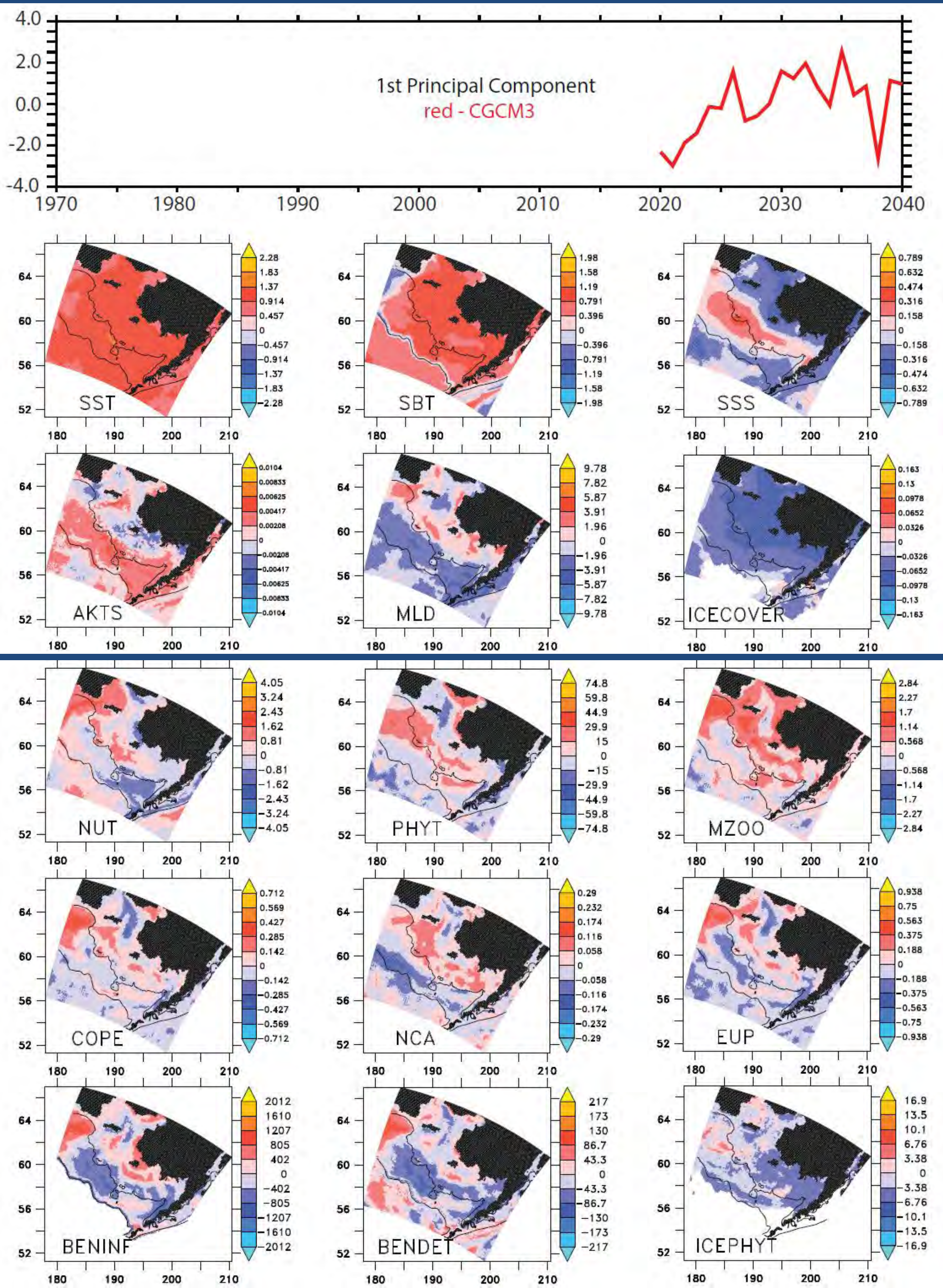
BERING SEA MODEL VARIABLES ANALYZED

PHYSICAL (ROMS)

- Surface Temperature (SST)
- Bottom Temperature (SBT)
- Surface Salinity (SSS)
- Ice (ICECOVER)
- Mixed Layer Depth (MLD)
- Vertical Mixing (AKTS)

BIOLOGICAL (BEST-NPZ)

- Nitrate + Ammonium (NUT)
- Ice Phytoplankton (ICEPHYT)
- Phytoplankton (PHYT)
- Microzooplankton (MZOO)
- Small Copepods (COPE)
- Neocalanus (NCA)
- Euphausiids (EUP)
- Benthic detritus (BENDET)
- Benthic infauna (BENTHIC)

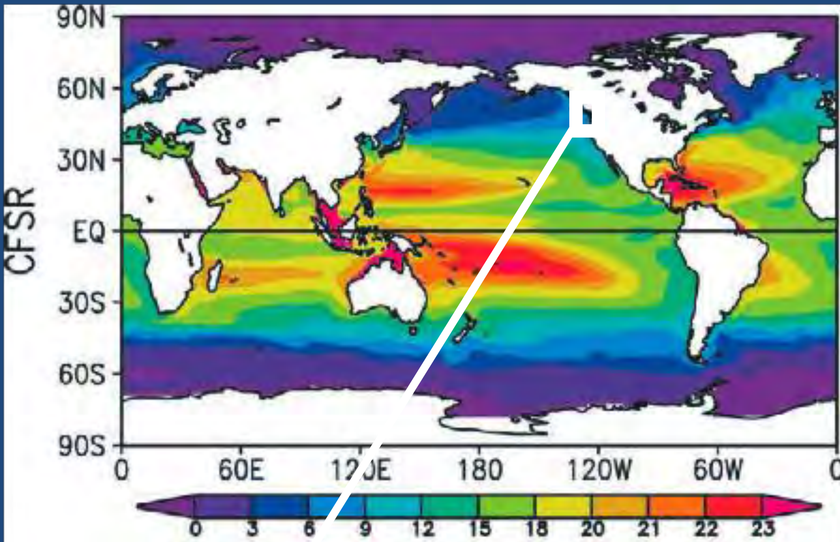


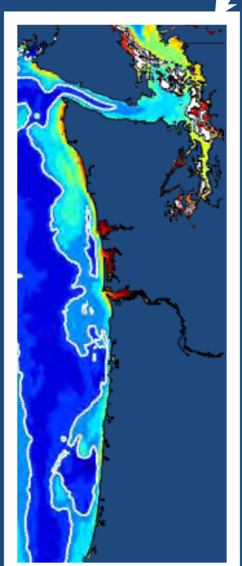
Multivariate EOF results from Bering Sea FORECAST driven by IPCC forecast (Hermann et al. DSR2 2013)

Higher SST associated with:

- Reduced ICECOVER
- Band of high SSS
- Increased inner-shelf MZOO
- Decreased outer-shelf EUP and NCA
- Reduced outer-shelf BENINF and BENDET

Example 3: J-SCOPE project seeks to predict physics and biology of the coastal Pacific NW

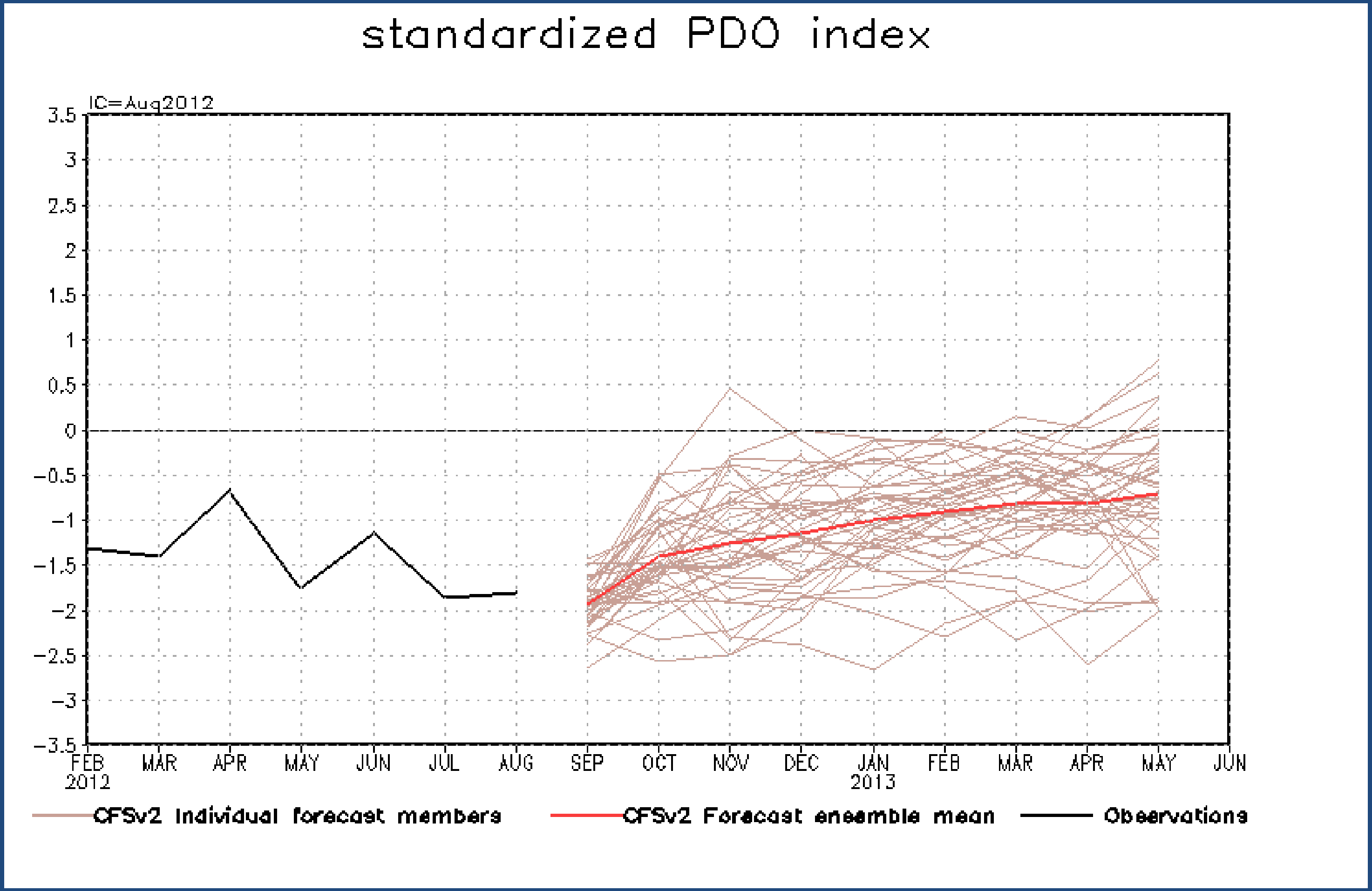




- The Climate Forecast System (CFS) – a global coupled air/sea/land model –is used for boundary conditions and atmospheric forcing of an established ROMS-based regional model with biogeochemistry (Cascadia domain, ~1.5 km resolution)
- Empirical relationships from observations are applied to the modeled fields to predict additional quantities (e.g. pH and fish)

CMG
UW Coastal Modeling Group

CFS individual realizations of the future PDO:



ideally would run regional model for each of these realizations but very expensive

CFS monthly predictions of ENSO and PDO (based on 15-member ensemble means)

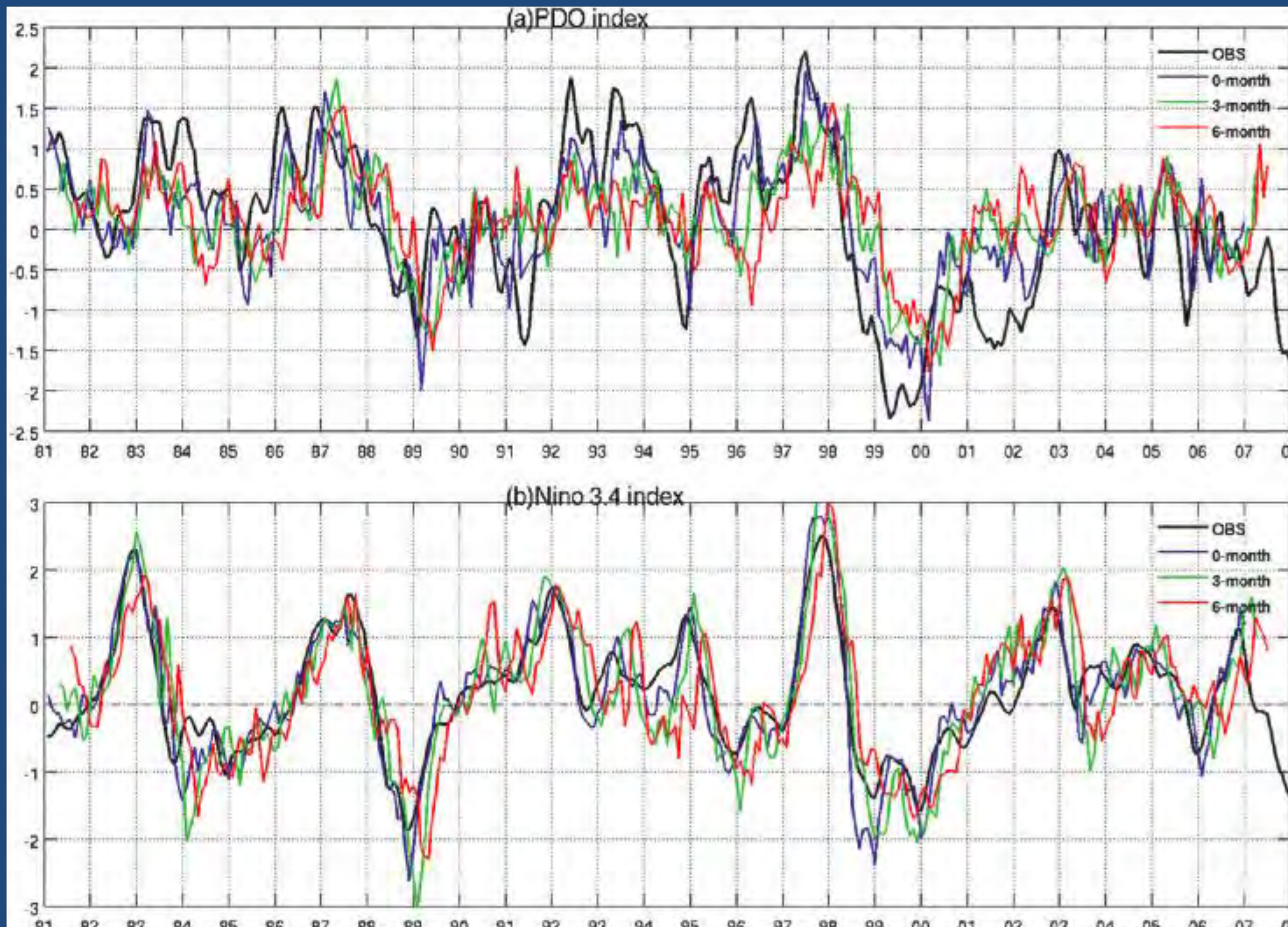
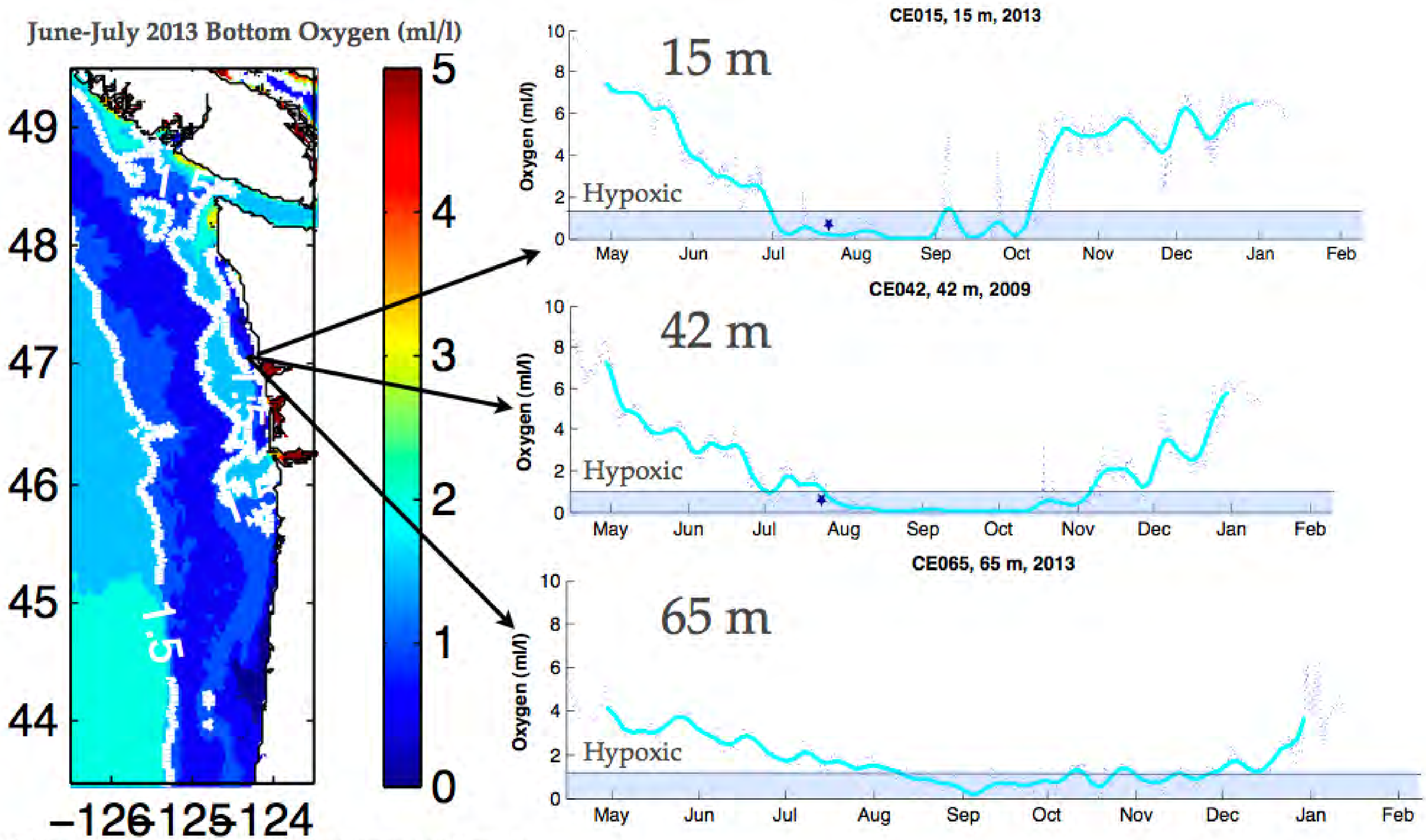


FIG. 5. Time series of (a) the PDO index and (b) the Niño-3.4 SST anomaly during 1981–2006. Black line denotes observation, and blue, green, and red lines denote CFS predictions at 0-, 3-, and 6-month lead, respectively.

Wen et al, 2012

A successful forecast was made from an individual realization



White contour outlines Hypoxic Zones (<1.5 ml/l)

Forecast: Hypoxia begins in July, 2013
for Cape Elizabeth region of WA coast

J-SCOPE

Oxygen model - Siedlecki et al, in prep


DEAD CRABS OBSERVED THIS YEAR OFF WA COAST
Photo credit:
Ellen Starr

Summary of Technique

- See if we can get significant variance in a coupled (CPCA) mode which includes BOTH large-scale (global) and small-scale (downscaled) patterns
- IF have such a powerful mode, *convolve the large-scale part with large-scale forecast output* to get an inferred time series of small-scale output.