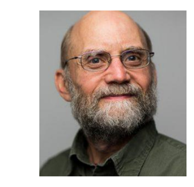
Emulation of MOM6-based downscaling results in the Northeast Pacific using Machine Learning methods

Albert Hermann (albert.j.Hermann@noaa.gov)[1,2], Wei Cheng[2] , Vivek Seelanki [1,2], Phyllis Stabeno[2]

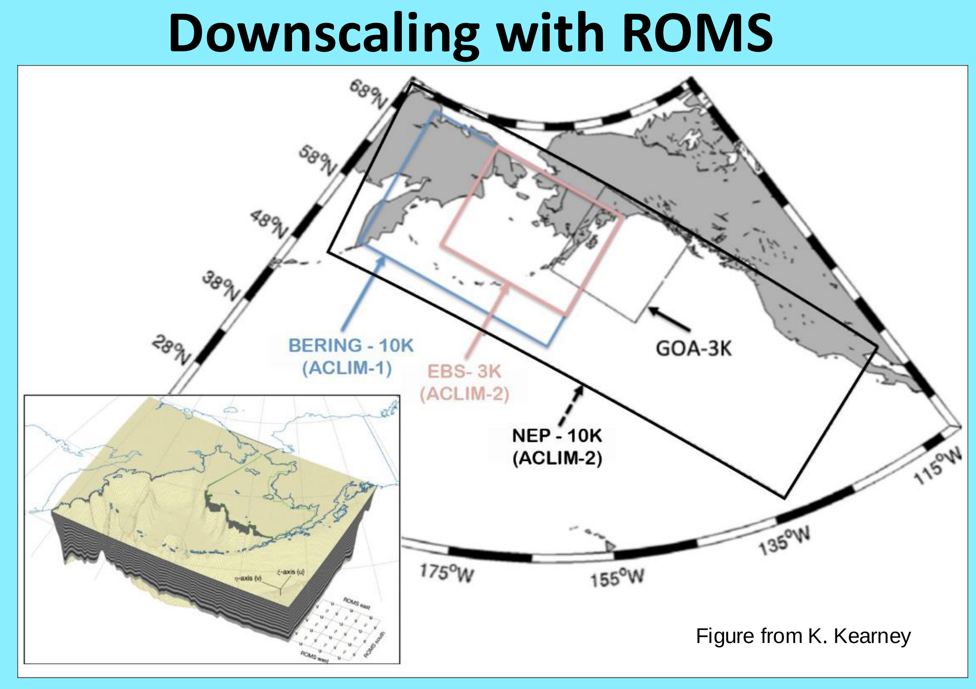
1. University of Washington/Cooperative Institute for Climate Ocean and Ecosystem Studies (CICOES)
2. NOAA/Pacific Marine Environmental Laboratory



ABSTRACT

As part of the new Climate and Ecosystems and Fisheries Initiative (CEFI), we have been comparing results from a dynamical downscaling regional model, based on the Modular Ocean Model version 6 (MOM6), with hydrographic and current meter data from the Bering Sea and surrounding waters. Ultimately under CEFI this regional model is to be used for hindcasts, seasonal to multiyear forecasts, and multidecadal projections of conditions in the Northeast Pacific. Early results have shown an impressive correspondence between the new model and observations in the Bering Sea, including metrics which are central to the management of fisheries. Here we present preliminary results from an application of Machine Learning methods, used to create a compact emulator of the full MOM6-based dynamical model. The method is based on dimensional reduction via 3D EOFs, followed by an application of a Recurrent Neural Network method (specifically, the Long Short-Term Memory technique). We explore whether the emulator can capture major features of the Bering Sea such as the summertime "cold pool". By virtue of its computational efficiency, a validated emulator of this type could be used to: 1) expand dynamically-produced ensembles of regional forecasts and projections of conditions in the Bering Sea; 2) expand our sensitivity analyses of the regional downscaling model.

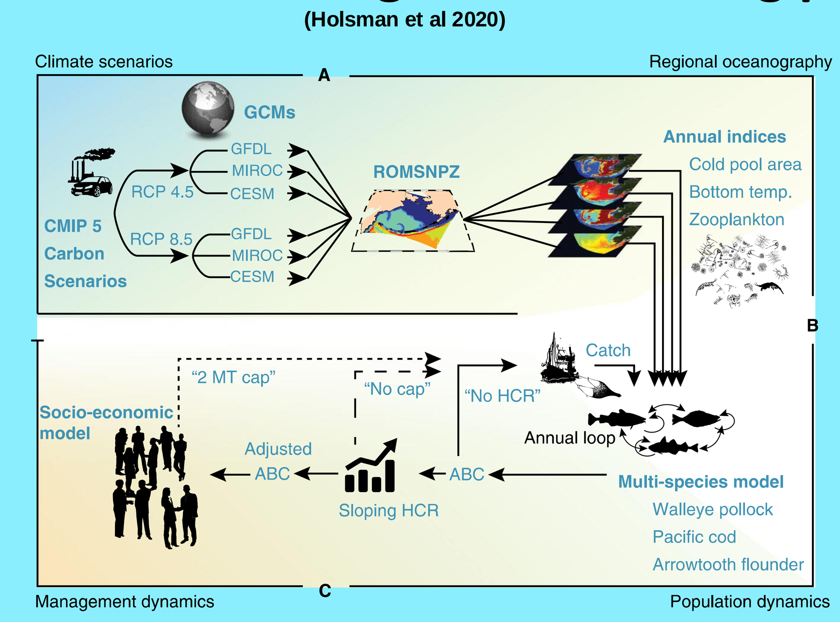
Downscaling with ROMS



Over many years, we have used the Regional Ocean Modeling System (ROMS) for hindcasts and projections of the Bering Sea and Gulf of Alaska.

Alaska Climate Integrated Modeling program

(Holsman et al 2020)



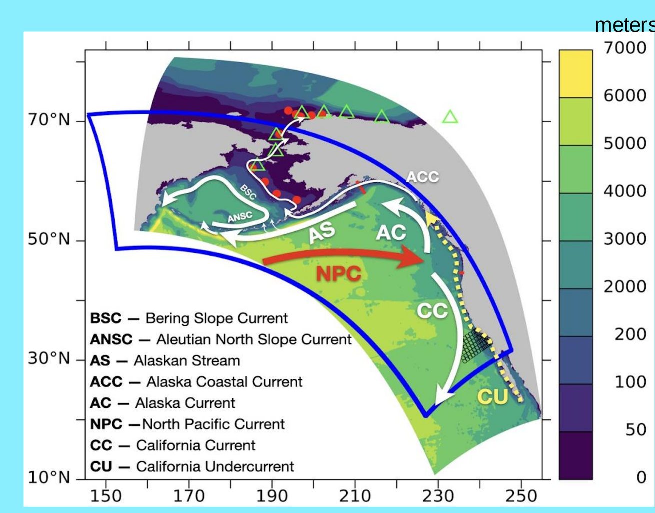
Projection results have been used in Management Strategy Evaluation. Ideally want BIG ensembles for management applications. These can be very costly, hence need for emulators.

The Climate Ecosystems and Fisheries Initiative (CEFI)

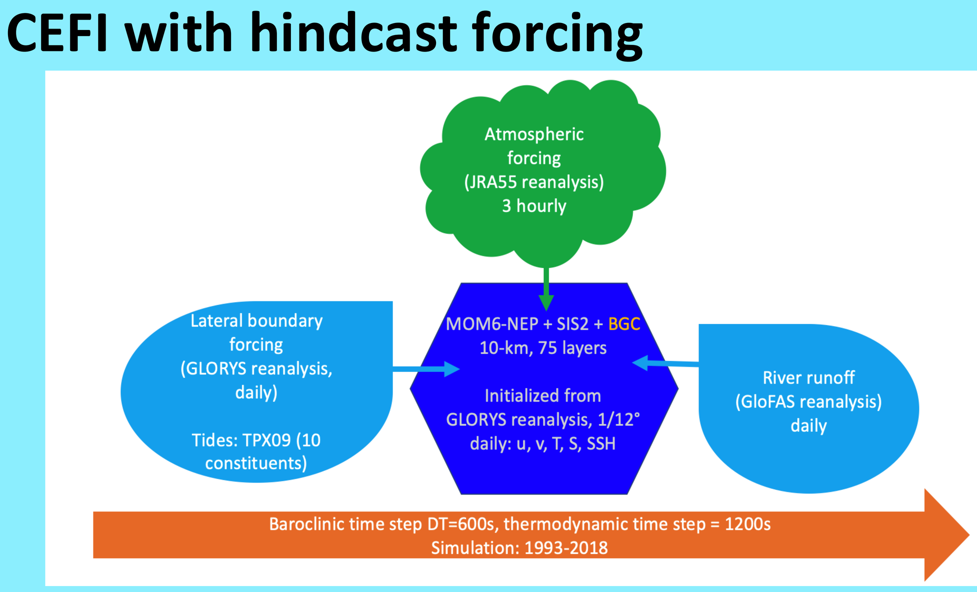
The CEFI program aims to *operationalize* regional ocean hindcasts/forecasts/projections within NOAA

MOM6-NEP Domain for CEFI with hindcast forcing

As part of CEFI, we have been calibrating and validating the new MOM6-NEP model (developed by Drenkard et al., 2024) with data from the Bering Sea.



Bathymetry, grid domains and EcoFOCI moorings

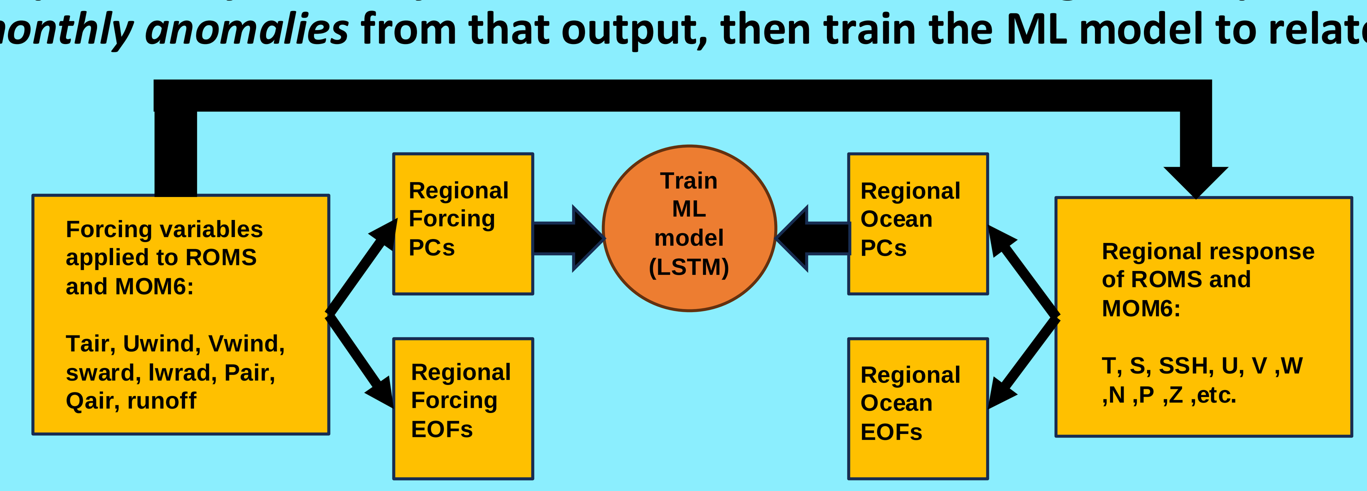


Hindcast forcing and boundary conditions

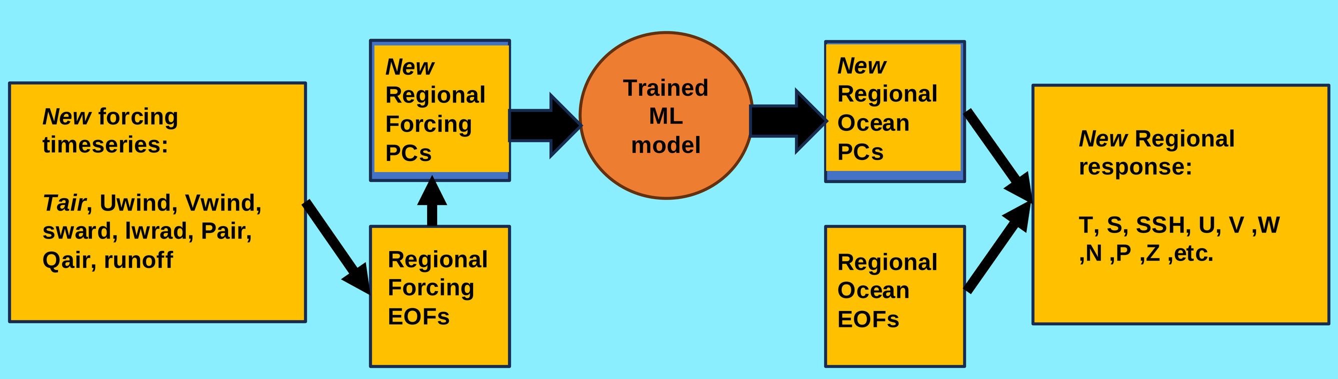
Construction of a Machine Learning (ML) based emulator

- Output from a complex dynamical model can be used to train a surrogate (aka "*emulator*") which **compactly** approximates the behavior of the full system (and is **~10^6 times faster** than dynamical version). The use of a compact emulator allows a broader range of model experiments, e.g.:
  - Quantify sensitivities to forcing and parameters
  - Broaden ensemble of predictions
- Here, we explore the use of Machine Learning to construct emulators for regional NEP models based on ROMS and MOM6

Step 1: We dynamically downscale, calculate the forcing and response EOFs of *monthly anomalies* from that output, then train the ML model to relate the PCs

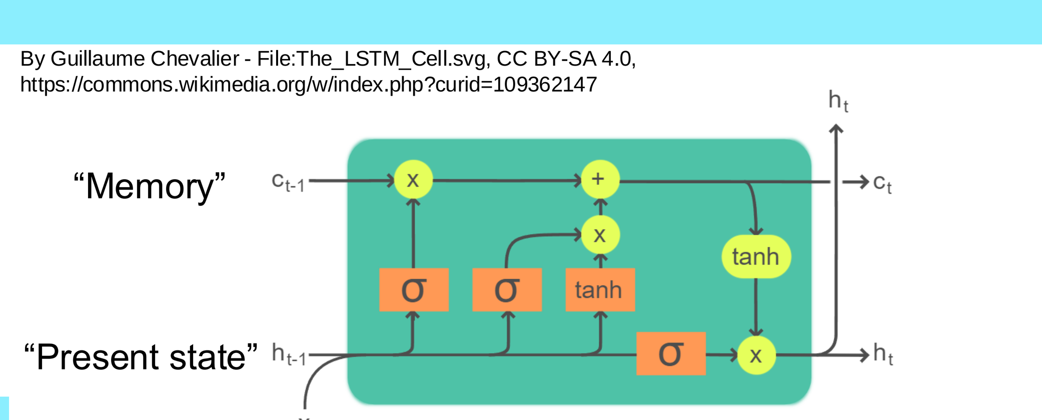


Step 2: We then project new forcing sets onto the regional forcing EOFs and use the trained ML model to emulate the regional response to that *new* forcing



Details of the method

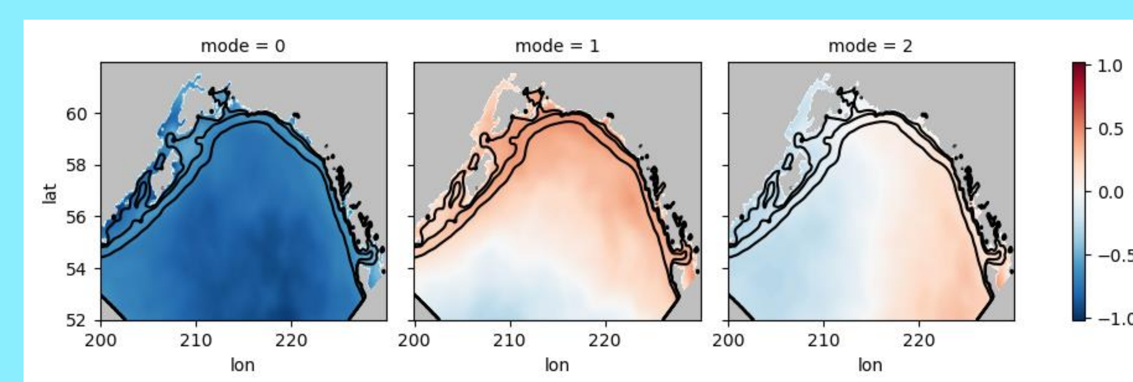
The ML model is based on the Long Short-Term Memory (LSTM) algorithm, which is a type of Recurrent Neural Network:



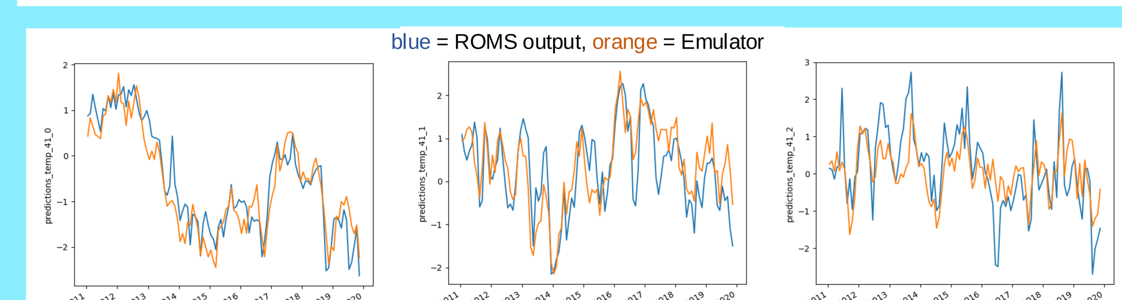
- Include 12 months past forcing for training and emulation
- Include top 10 PCs of each forcing and response variable (these can be either 2D or 3D fields)
- Use 200-400 "neurons" in the LSTM
- The optimization target for each "training session" can be a single PC of a single regional response variable ***or can train many variables/modes simultaneously –*** which yields superior results in many cases!
- Here we ***train*** the ML models using only dynamical hindcast results prior to 2010 and ***test*** them using dynamical hindcast results from 2010 onwards (which were not used in training)

A ROMS-NEP emulator trained using the top 10 EOFs of 8 forcing variables and 1 oceanic response variable (SST) is used to hindcast *SST in the Gulf of Alaska*

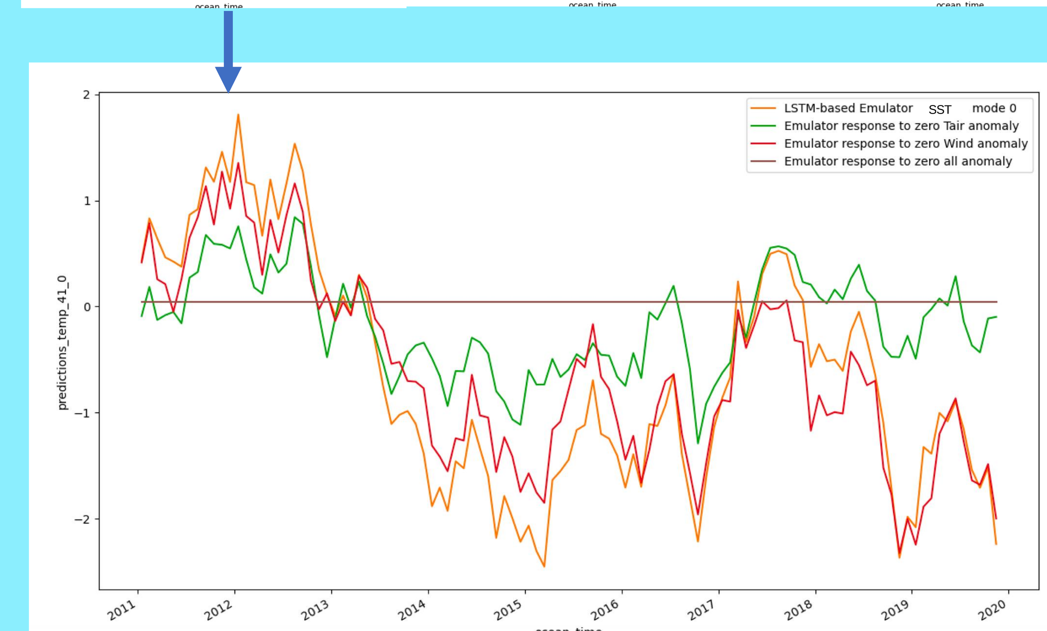
Top 3 EOFs of SST from ROMS output (deg C)



Top 3 PCs of SST from ROMS output vs. ROMS emulator (testing years 2011-2020)

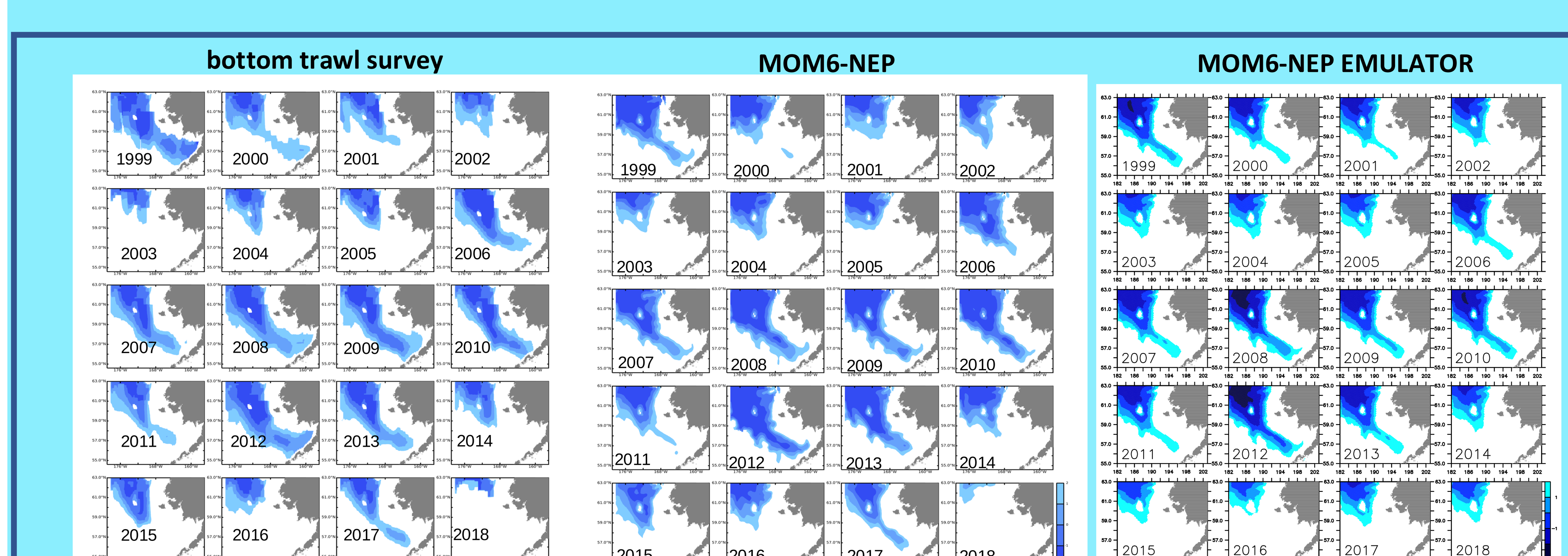


Top PC of SST from ROMS Emulator: sensitivity to forcing elements



The ROMS ML Emulator (orange line) does very well at replicating the top 3 PCs of SST monthly anomalies from the original ROMS output (blue line). Selective removal of atmospheric forcing anomalies reveals a greater sensitivity of mode 0 to air temperature (Tair, green line) than winds (Wind, red line).

A MOM6-NEP emulator trained using the top 10 EOFs of 8 forcing variables and 3 oceanic response variables (3D Temperature, 3D Salinity, SSH) is used to hindcast *bottom temperatures on the Bering Sea shelf*



The MOM6-NEP Emulator is used to reconstruct the full signal (monthly anomalies plus monthly means) of bottom temperatures (deg C) on the Bering Sea Shelf. It does well at replicating the original MOM6 output (and the measured values) in most (but not all) years, including the testing years of 2011-2018. Note a slightly different colorbar is used in the display of emulator results.