Regional Reanalyses

Christopher A. Edwards (UC Santa Cruz)

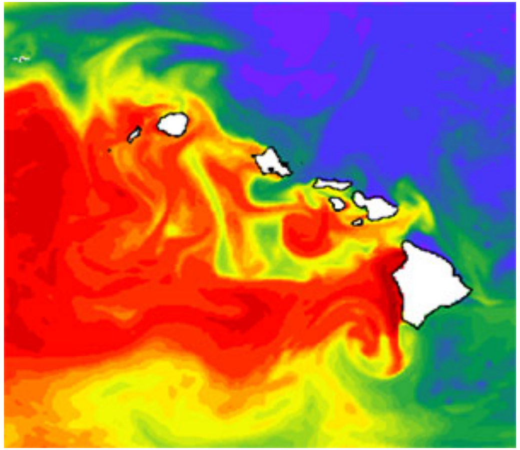
CLIVAR Daily to Decadal Ecological Forecasting along North American Coastlines Workshop

Woods Hole, MA

April 12-14, 2022

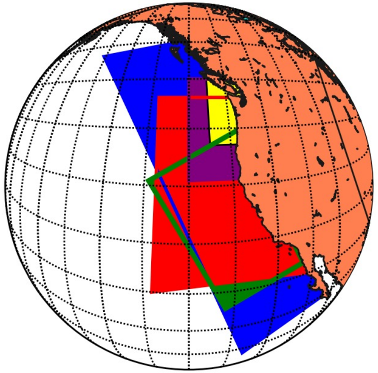
Credits to: **Andy Moore**, Paul Mattern, Hajoon Song

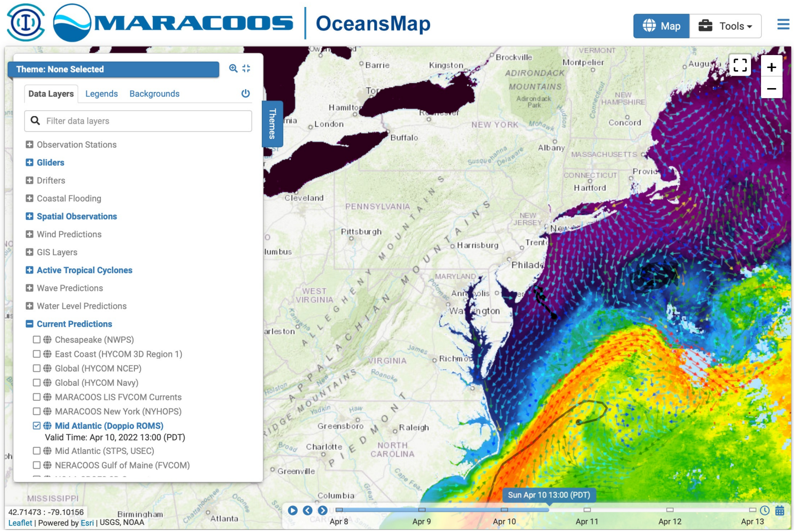
Some Regional Reanalysis Products



PacIOOS
PACIFIC ISLANDS OCEAN OBSERVING SYSTEM

These physical products use ROMS 4D-Var





MARACOOS | OceansMap

Theme: None Selected

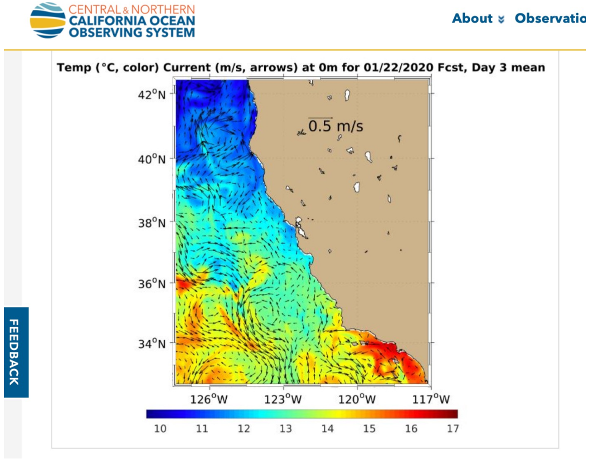
Data Layers | Legends | Backgrounds

Filter data layers

- Observation Stations
- **Gliders**
- Drifters
- Coastal Flooding
- **Spatial Observations**
- Wind Predictions
- GIS Layers
- **Active Tropical Cyclones**
- Wave Predictions
- Water Level Predictions
- **Current Predictions**
  - ☐ Chesapeake (NWPS)
  - ☐ East Coast (HYCOM 3D Region 1)
  - ☐ Global (HYCOM NCEP)
  - ☐ Global (HYCOM Navy)
  - ☐ MARACOOS LIS FVCOM Currents
  - ☐ MARACOOS New York (NYHOPS)
  - ☑ **Mid Atlantic (Doppio ROMS)**
    Valid Time: Apr 10, 2022 13:00 (PDT)
  - ☐ Mid Atlantic (STPS, USEC)
  - ☐ NERACOOS Gulf of Maine (FVCOM)

42.71473 : -79.10156
Leaflet | Powered by Esri | USGS, NOAA

Sun Apr 10 13:00 (PDT)

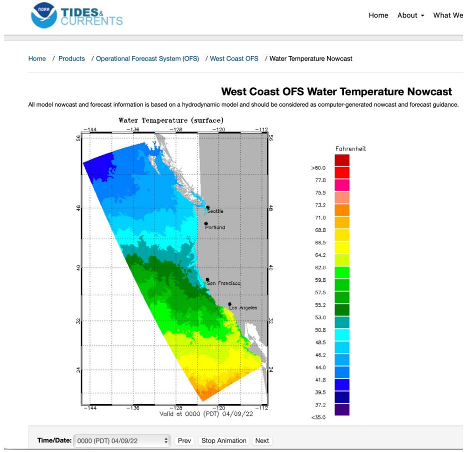


CENTRAL & NORTHERN CALIFORNIA OCEAN OBSERVING SYSTEM

About ⌄ Observatio

FEEDBACK

Temp (°C, color) Current (m/s, arrows) at 0m for 01/22/2020 Fcst, Day 3 mean



TIDES & CURRENTS

Home | About ▾ | What We

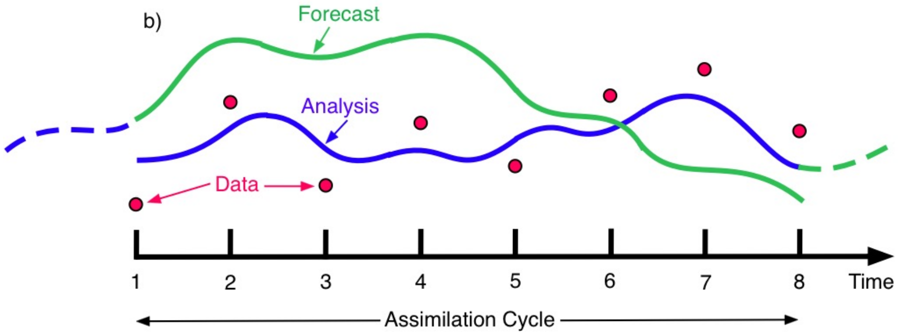
Home / Products / Operational Forecast System (OFS) / West Coast OFS / Water Temperature Nowcast

West Coast OFS Water Temperature Nowcast

All model nowcast and forecast information is based on a hydrodynamic model and should be considered as computer-generated nowcast and forecast guidance.

Time/Date: 0000 (PDT) 04/09/22 | Prev | Stop Animation | Next

ROMS 4D-Var

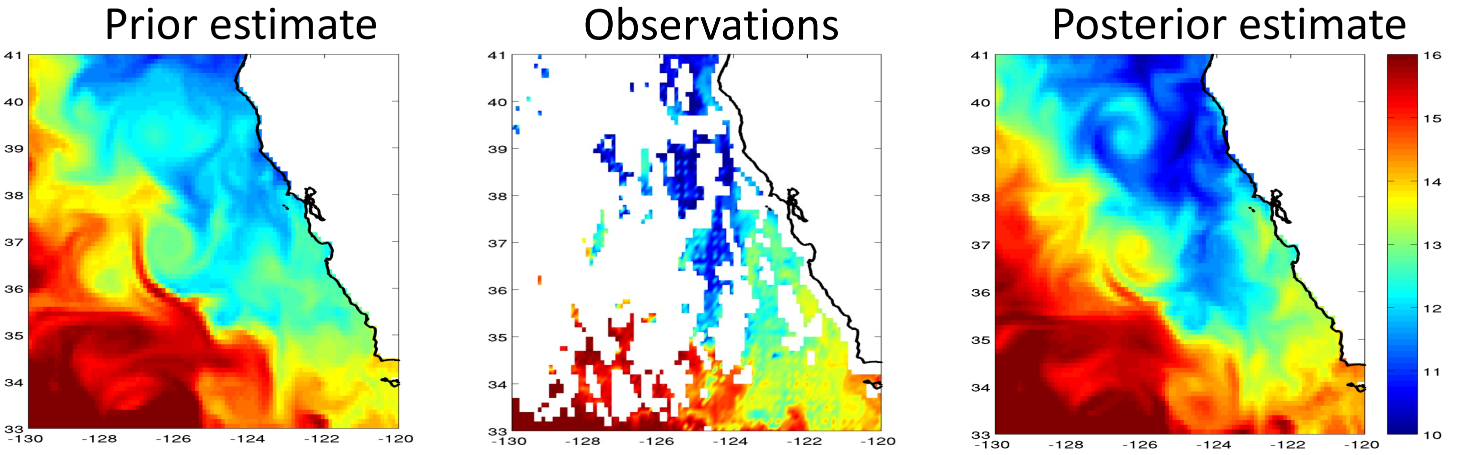


- Regional Ocean Modeling System (ROMS)
- 4-Dimensional Variational Assimilation
- Linearized model dynamics connect observations at different times
- Data can be continuous in time
- Long cycles (days-week)

Minimize:

$$J = \left(\mathbf{z} - \underset{\substack{\uparrow \\ \textit{Prior}}}{\mathbf{z}_\mathrm{b}}\right)^T \underset{\substack{\uparrow \\ \textit{Prior} \\ \textit{error cov.}}}{\mathbf{B}}^{-1} \left(\mathbf{z} - \mathbf{z}_\mathrm{b}\right) + \left(\underset{\substack{\uparrow \\ \textit{Obs}}}{\mathbf{y}} - \underset{\substack{\uparrow \\ \textit{Obs} \\ \textit{operator}}}{H}(\mathbf{z}_\mathrm{b})\right)^T \underset{\substack{\uparrow \\ \textit{Obs} \\ \textit{error cov.}}}{\mathbf{R}}^{-1} \left(\mathbf{y} - H(\mathbf{z}_\mathrm{b})\right)$$

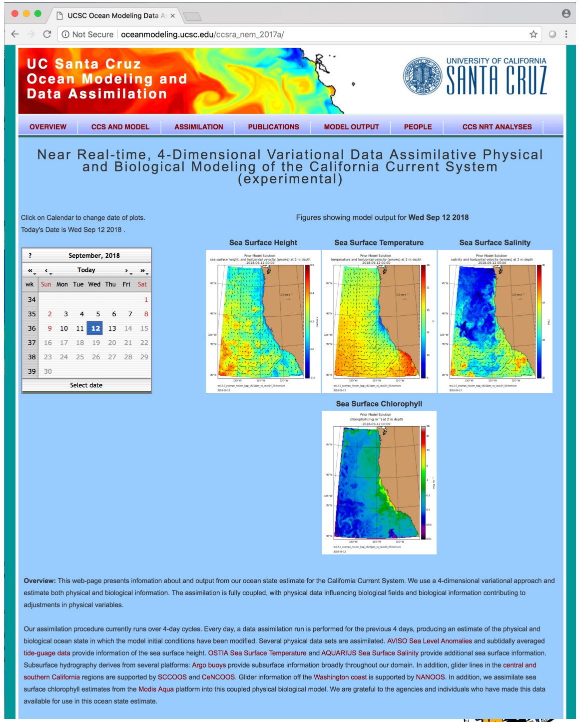
One example cycle showing SST
Control variables are model initial conditions


Prior estimate
Observations
Posterior estimate

- Though SST is shown, all variables are adjusted in ways consistent with background and observation error covariances

UCSC ROMS 4D-Var **Historical Reanalyses** 1980-2010 (ERA) and 1999-2012 (COAMPS) and **near real-time** system (2011-present)

- 1/10º CCS ROMS configuration
- Reanalyses 8-day assimilation cycles
- NRT: 4-day assimilation cycles
- Assimilates SST, SSH, SCHL, glider T/S, Argo T/S, HF RADAR velocities
- Model output available on a TDS
- Calendar searchable with figures
- Focus on nowcast and potentially short-term prediction (~ 1 week).
- Why do this?
    - Marine resource management (HABS, Fisheries, Sanctuaries)
    - Industry: aquaculture, shipping
    - CGSAR (in principle)



Near Real-time, 4-Dimensional Variational Data Assimilative Physical and Biological Modeling of the California Current System (experimental)

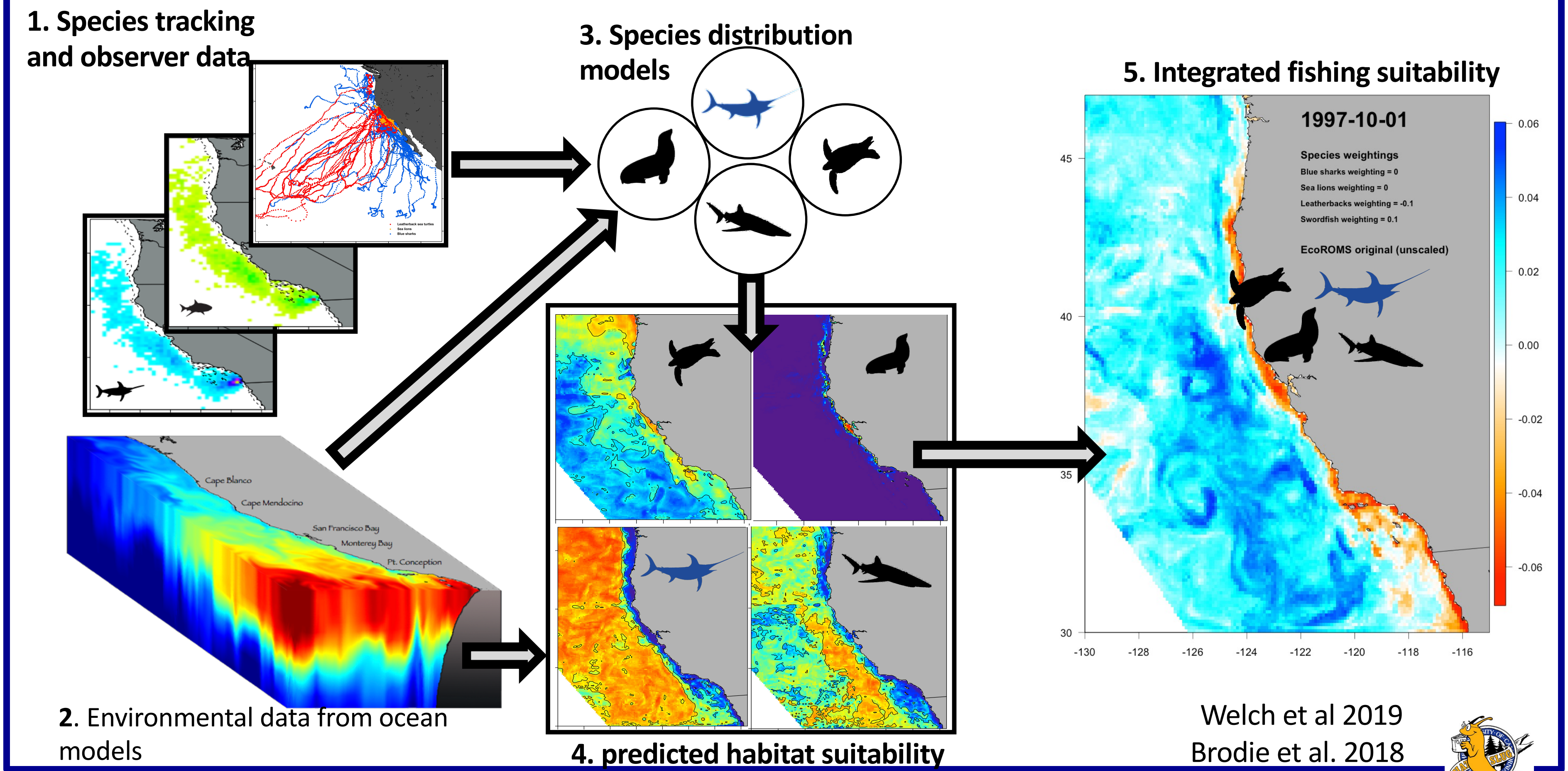
Click on Calendar to change date of plots.
Today's Date is Wed Sep 12 2018 .

Figures showing model output for **Wed Sep 12 2018**

Overview: This web-page presents infomation about and output from our ocean state estimate for the California Current System. We use a 4-dimensional variational approach and estimate both physical and biological information. The assimilation is fully coupled, with physical data influencing biological fields and biological information contributing to adjustments in physical variables.

Our assimilation procedure currently runs over 4-day cycles. Every day, a data assimilation run is performed for the previous 4 days, producing an estimate of the physical and biological ocean state in which the model initial conditions have been modified. Several physical data sets are assimilated. AVISO Sea Level Anomalies and subtidally averaged tide-guage data provide information of the sea surface height. OSTIA Sea Surface Temperature and AQUARIUS Sea Surface Salinity provide additional sea surface information. Subsurface hydrography derives from several platforms: Argo buoys provide subsurface information broadly throughout our domain. In addition, glider lines in the central and southern California regions are supported by SCCOOS and CeNCOOS. Glider information off the Washington coast is supported by NANOOS. In addition, we assimilate sea surface chlorophyll estimates from the Modis Aqua platform into this coupled physical biological model. We are grateful to the agencies and individuals who have made this data available for use in this ocean state estimate.

One motivation for state estimation



1. Species tracking and observer data

2. Environmental data from ocean models

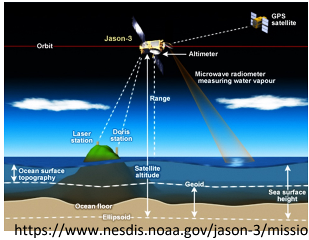
3. Species distribution models

4. predicted habitat suitability

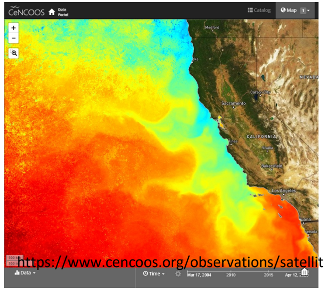
5. Integrated fishing suitability

Welch et al 2019
Brodie et al. 2018

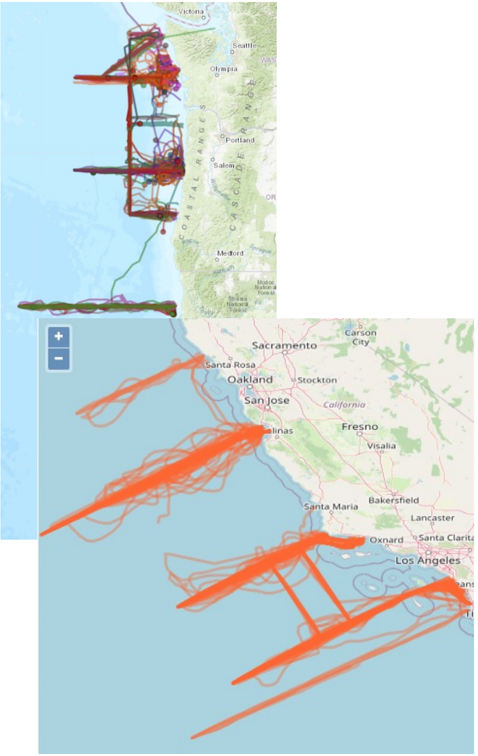
Physical Data Available for regional NRT assimilation

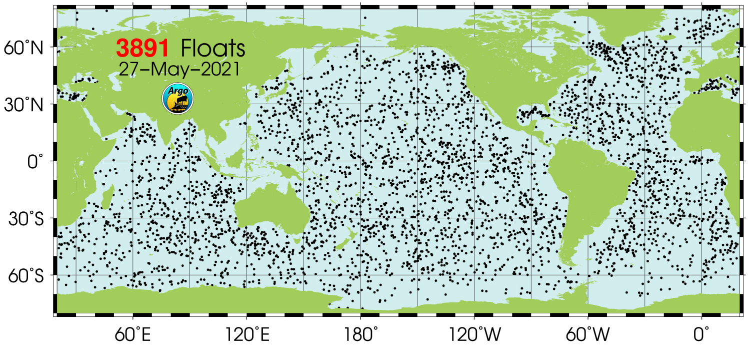


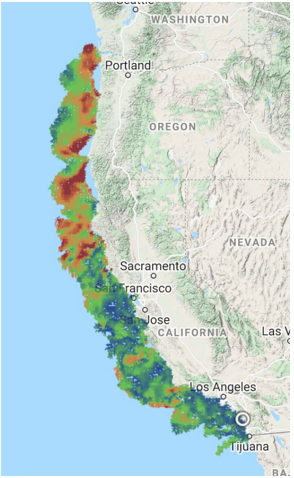
https://www.nesdis.noaa.gov/jason-3/mission.html



https://www.cencoos.org/observations/satellites/



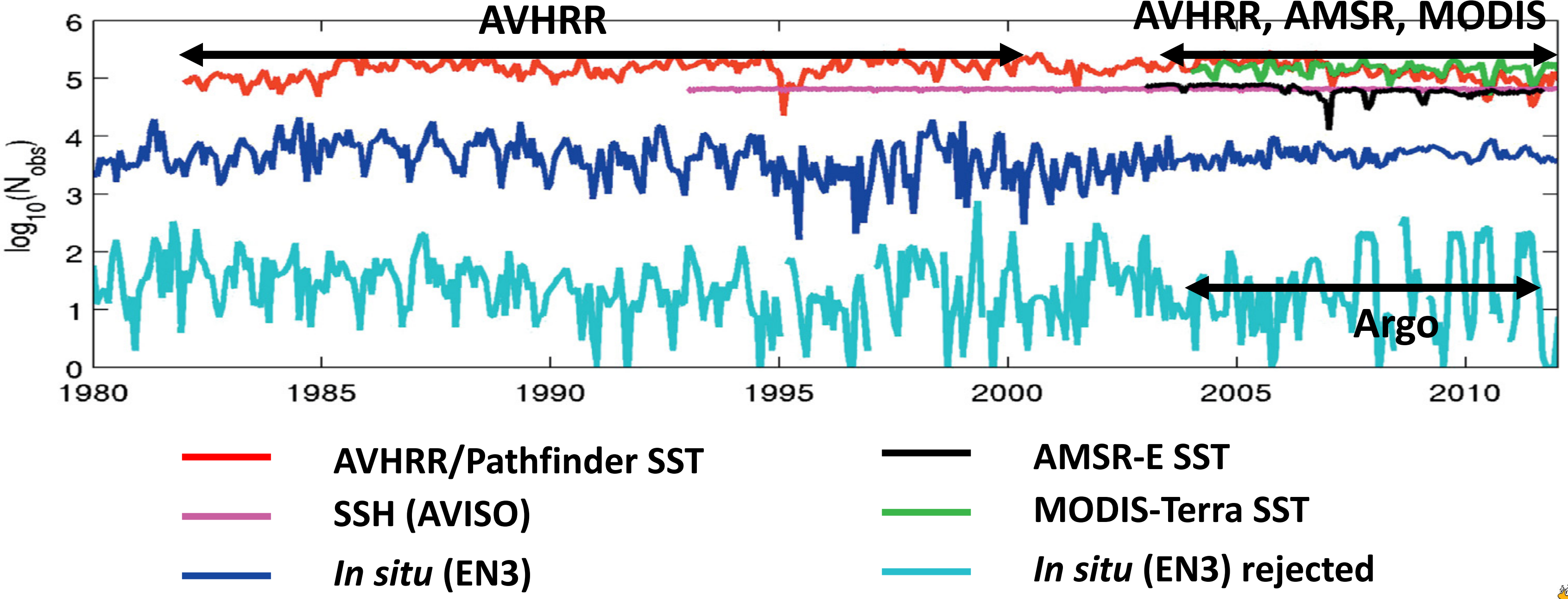




CCS Historical Observation Summary

(1980-2010)

Number of Obs:



Long reanalyses allow interpretive, predictive studies (e.g., of chlorophyll response during 2015-2016 El Nino)

- Estimates of historical 26.0 kg/m3 density surface put 2015-16 El Nino in context

- Along with EOF analysis of chl, allowed a couple month prediction of muted ecosystem response

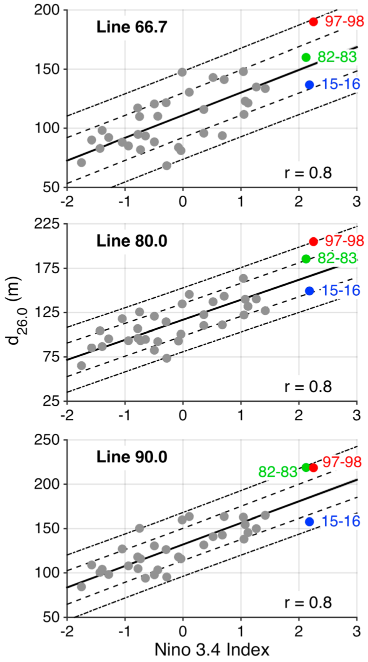
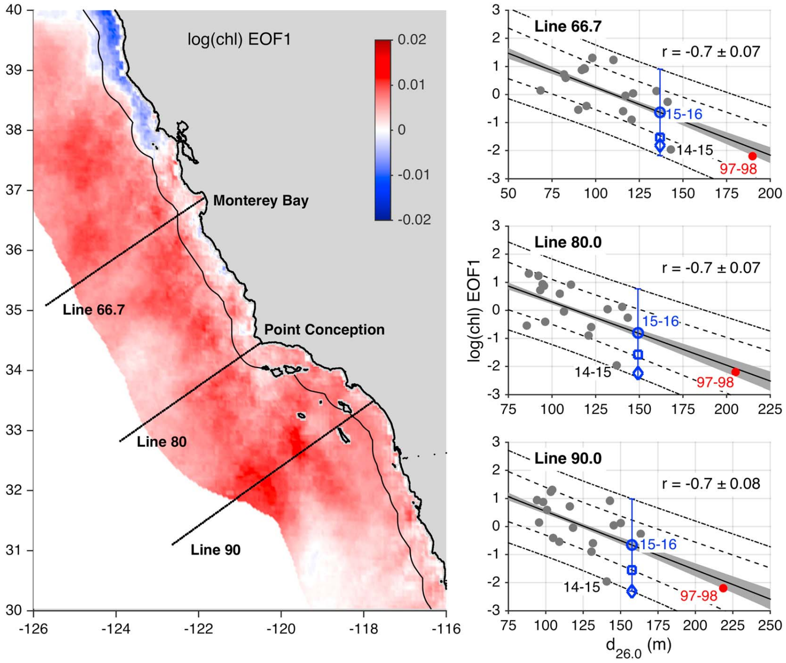
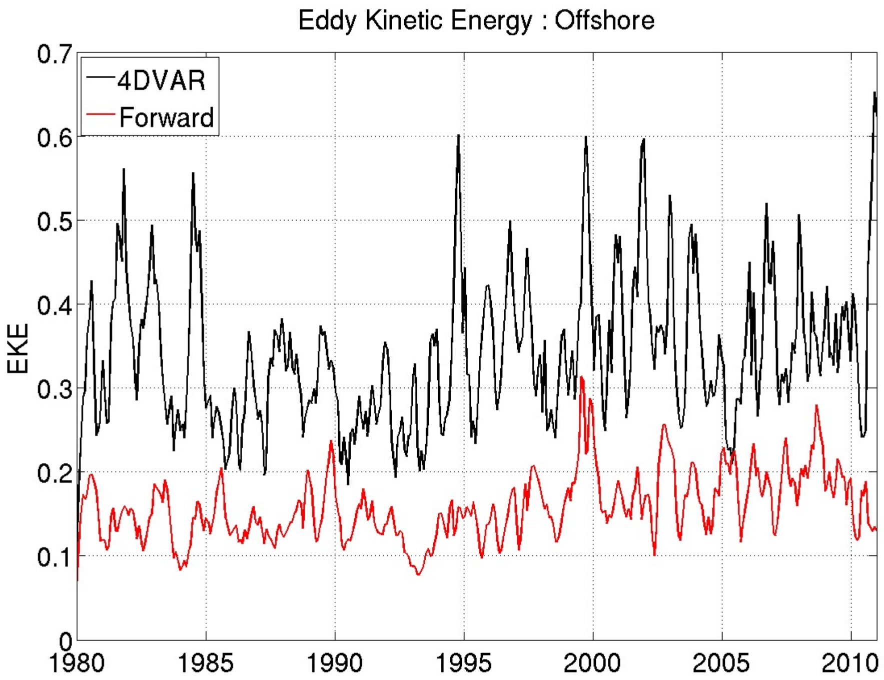


Figure 5. December–February mean $d_{26.0}$ (from merged model-glider data) plotted against November–January mean Niño 3.4 anomalies for each winter from 1981–1982 to 2015–2016. Isopycnal depths are averaged within 50 km of shore. The solid lines are linear fits to the data; the dashed and dash-dotted lines are ±1 and ±2 standard deviations from the linear fit.

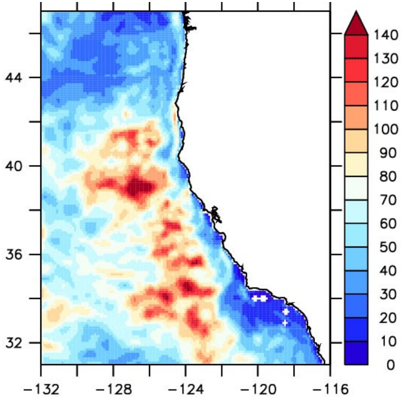


Jacox et al. (2016)

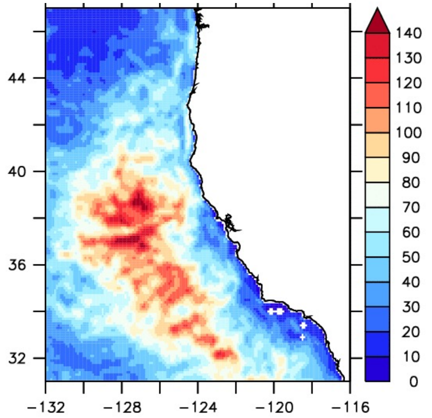
Impact of assimilation on Eddy Kinetic Energy


Eddy Kinetic Energy : Offshore

No Assimilation EKE



4D-Var EKE



Impact of Observations on Circulation Estimates (like OSSE)

4D-Var circulation estimate: $\mathbf{x_a} = \mathbf{x_b} + \mathbf{K}(\mathbf{y} - H(\mathbf{x_b}))$

analysis **gain** **obs operator**

background **obs**

Consider a scalar function $I(\mathbf{x})$ (e.g. transport)

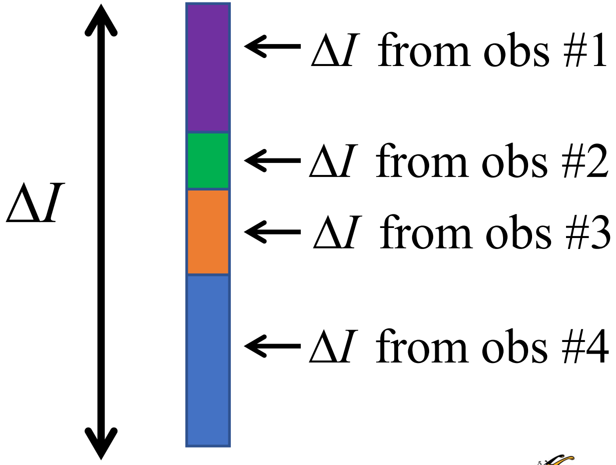
Change in $I(\mathbf{x})$ due to 4D-Var: $\Delta I = I(\mathbf{x_a}) - I(\mathbf{x_b})$

Impact of the observations on ΔI :

$$\Delta I \xrightarrow{\mathbf{K}^T} \Delta I_{obs1} + \Delta I_{obs2} + \Delta I_{obs3} + \ldots$$

Impact of controls variables on ΔI :

$$\Delta I \xrightarrow{\mathbf{K}^T} \Delta I_{ic} + \Delta I_{fc} + \Delta I_{obc}$$



ΔI

$\leftarrow \Delta I$ from obs #1

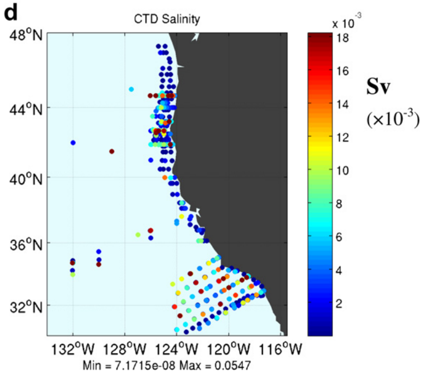
$\leftarrow \Delta I$ from obs #2

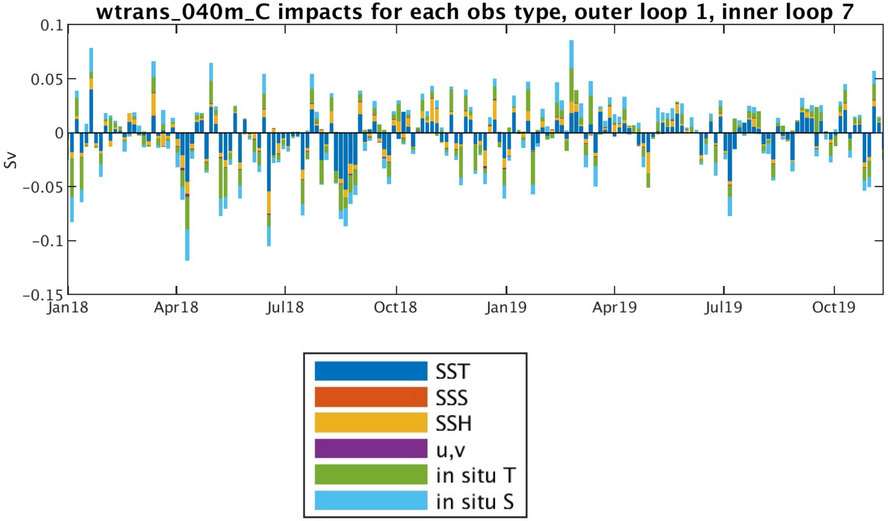
$\leftarrow \Delta I$ from obs #3

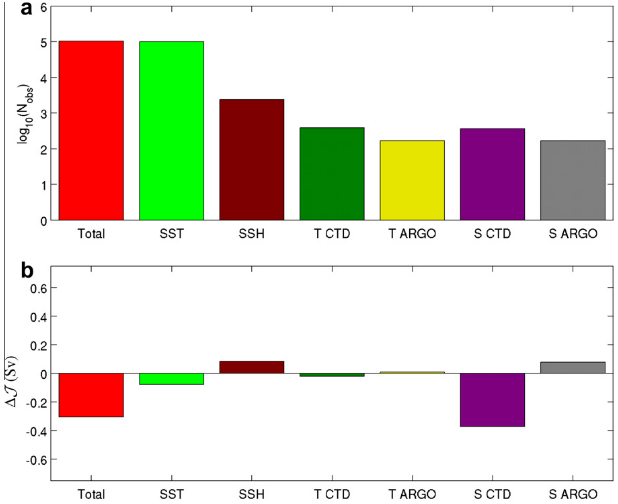
$\leftarrow \Delta I$ from obs #4

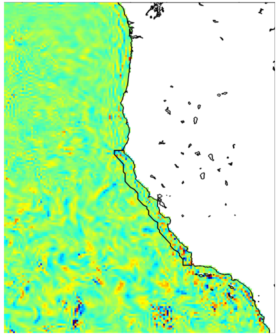
Quantifying the impact of observations and platforms on model estimates (like OSSE)

(e.g., impact on nearshore upwelling transport across 40 m and alongshore transport)





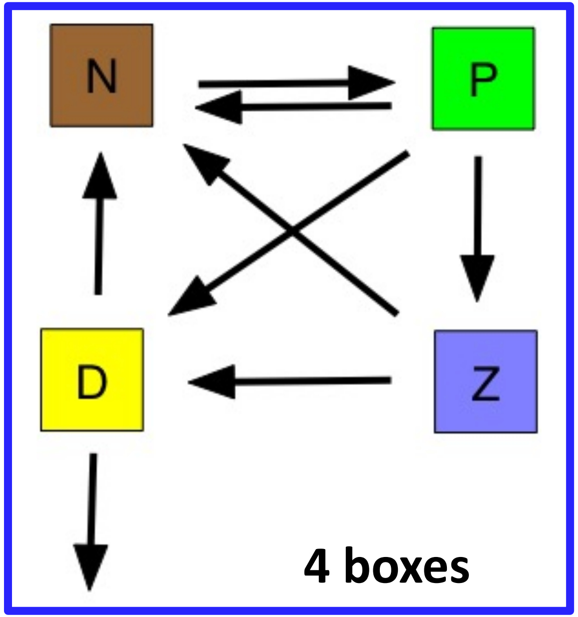




We have developed version of 4D-Var for use with ROMS coupled with biogeochemistry (for two ecosystem models)

NPZD

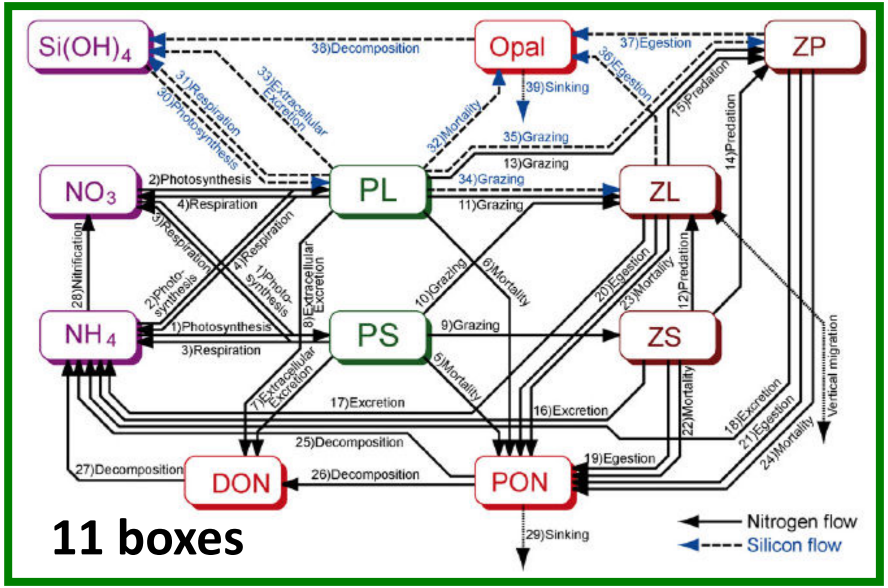
(Powell et al. 2006)



4 boxes

NEMURO

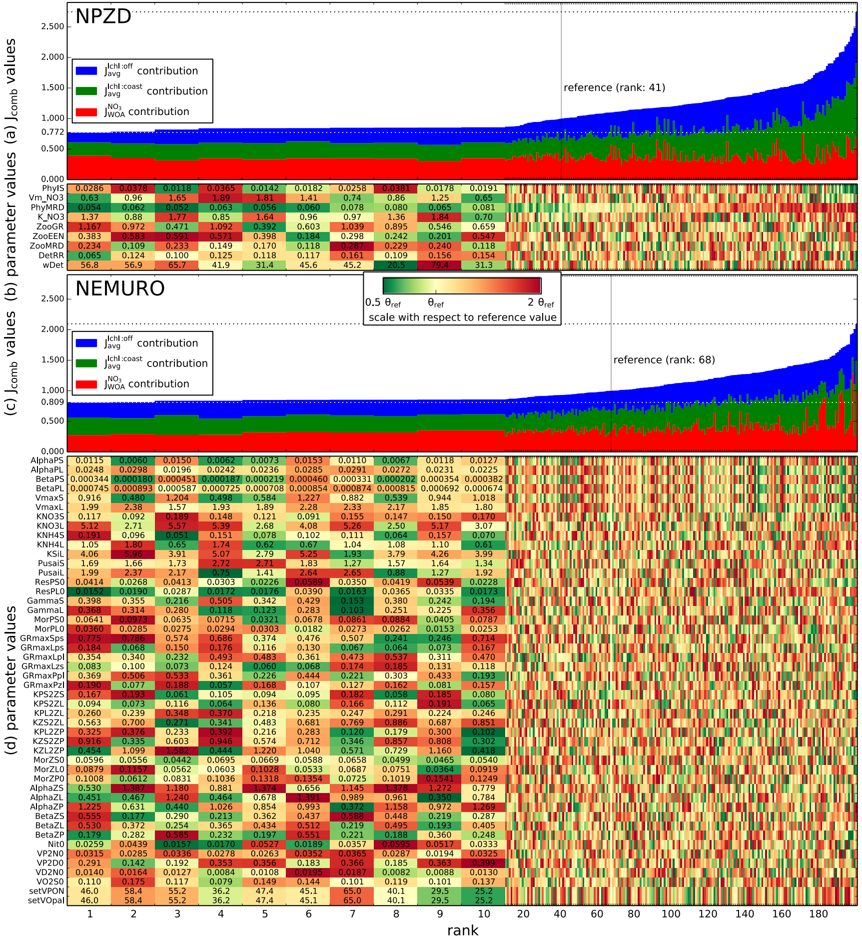
(Kishi et al. 2011)



11 boxes

One challenge: Parameter Sensitivity
Monte Carlo optimization

- 9 parameters (NPZD)
- 43 parameters (NEMURO)
- Multiple minima
- No clear parameter bias



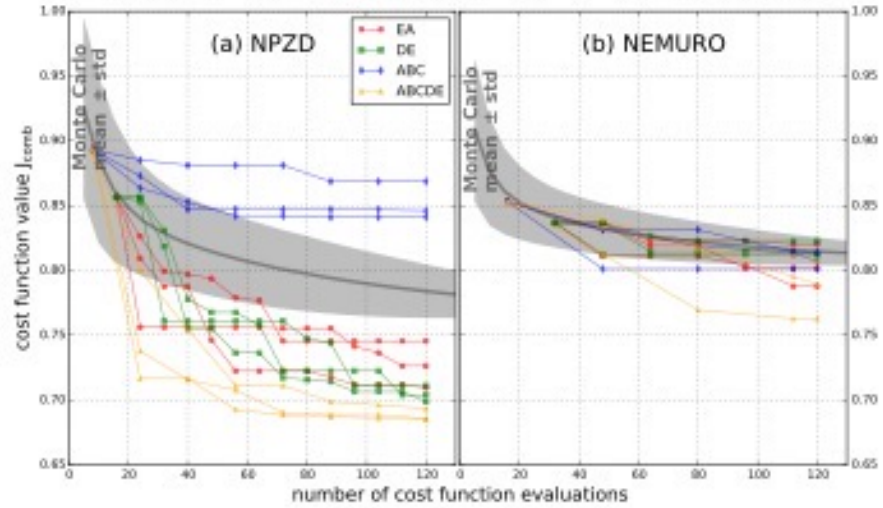
	1	2	3	4	5	6	7	8	9	10
PhyIS	0.0286	0.0378	0.0118	0.0365	0.0142	0.0182	0.0258	0.0381	0.0178	0.0191
Vm_NO3	0.63	0.96	1.65	1.89	1.81	1.41	0.74	0.86	1.25	0.65
PhyMRD	0.054	0.062	0.052	0.063	0.056	0.060	0.078	0.080	0.065	0.081
K_NO3	1.37	0.88	1.77	0.85	1.64	0.96	0.97	1.36	1.84	0.70
ZooGR	1.167	0.972	0.471	1.092	0.392	0.603	1.039	0.895	0.546	0.659
ZooEEN	0.383	0.583	0.591	0.571	0.398	0.184	0.298	0.242	0.201	0.547
ZooMRD	0.234	0.109	0.233	0.149	0.170	0.118	0.287	0.229	0.240	0.118
DetRR	0.065	0.124	0.100	0.125	0.118	0.117	0.161	0.109	0.156	0.154
wDet	56.8	56.9	65.7	41.9	31.4	45.6	45.2	20.5	79.4	31.3

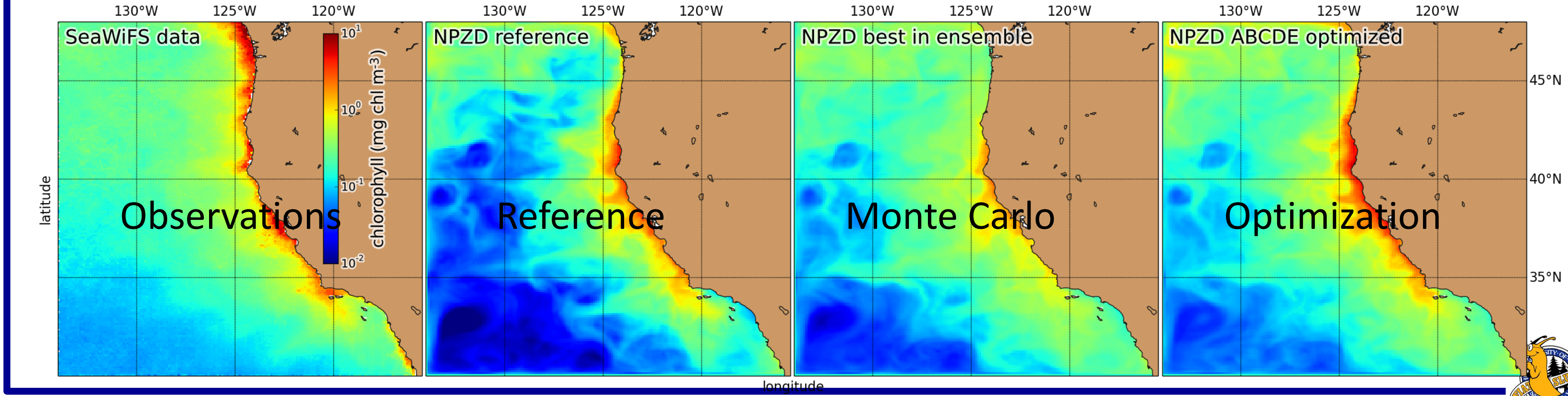
	1	2	3	4	5	6	7	8	9	10
AlphaPS	0.0115	0.0060	0.0150	0.0062	0.0073	0.0153	0.0110	0.0067	0.0118	0.0127
AlphaPL	0.0248	0.0298	0.0196	0.0242	0.0236	0.0285	0.0291	0.0272	0.0231	0.0225
BetaPS	0.000344	0.000180	0.000451	0.000187	0.000219	0.000460	0.000331	0.000202	0.000354	0.000382
BetaPL	0.000745	0.000893	0.000587	0.000725	0.000708	0.000854	0.000874	0.000815	0.000692	0.000674
VmaxS	0.916	0.480	1.204	0.498	0.584	1.227	0.882	0.539	0.944	1.018
VmaxL	1.99	2.38	1.57	1.93	1.89	2.28	2.33	2.17	1.85	1.80
KNO3S	0.117	0.092	0.189	0.148	0.121	0.091	0.155	0.147	0.150	0.170
KNO3L	5.12	2.71	5.57	5.39	2.68	4.08	5.26	2.50	5.17	3.07
KNH4S	0.191	0.096	0.051	0.151	0.078	0.102	0.111	0.064	0.157	0.070
KNH4L	1.05	1.80	0.65	1.74	0.62	0.67	1.04	1.08	1.10	0.61
KSiL	4.06	5.96	3.91	5.07	2.79	5.25	1.93	3.79	4.26	3.99
PusaiS	1.69	1.66	1.73	2.72	2.71	1.83	1.27	1.57	1.64	1.34
PusaiL	1.99	2.37	2.17	0.75	1.41	2.64	2.65	0.88	1.27	1.92
ResPS0	0.0414	0.0268	0.0413	0.0303	0.0226	0.0589	0.0350	0.0419	0.0539	0.0228
ResPL0	0.0152	0.0190	0.0287	0.0172	0.0176	0.0390	0.0163	0.0365	0.0335	0.0173
GammaS	0.398	0.355	0.216	0.505	0.342	0.429	0.153	0.380	0.242	0.194
GammaL	0.368	0.314	0.280	0.118	0.123	0.283	0.103	0.251	0.225	0.356
MorPS0	0.0641	0.0973	0.0635	0.0715	0.0321	0.0678	0.0861	0.0884	0.0405	0.0787
MorPL0	0.0360	0.0285	0.0275	0.0294	0.0303	0.0182	0.0273	0.0262	0.0153	0.0253
GRmaxSps	0.775	0.786	0.574	0.686	0.374	0.476	0.507	0.241	0.246	0.714
GRmaxLps	0.184	0.068	0.150	0.176	0.116	0.130	0.067	0.064	0.073	0.167
GRmaxLpl	0.354	0.340	0.232	0.493	0.483	0.361	0.473	0.537	0.311	0.470
GRmaxLzs	0.083	0.100	0.073	0.124	0.060	0.068	0.174	0.185	0.131	0.118
GRmaxPpl	0.369	0.506	0.533	0.361	0.226	0.444	0.221	0.303	0.433	0.193
GRmaxPzl	0.190	0.077	0.188	0.057	0.168	0.107	0.127	0.162	0.081	0.157
KPS2ZS	0.167	0.193	0.061	0.105	0.094	0.095	0.182	0.058	0.185	0.080
KPS2ZL	0.094	0.073	0.116	0.064	0.136	0.080	0.166	0.112	0.191	0.065
KPL2ZL	0.260	0.239	0.348	0.370	0.218	0.235	0.247	0.291	0.224	0.246
KZS2ZL	0.563	0.700	0.271	0.341	0.483	0.681	0.769	0.886	0.687	0.851
KPL2ZP	0.325	0.376	0.233	0.392	0.216	0.283	0.120	0.179	0.300	0.102
KZS2ZP	0.916	0.335	0.603	0.946	0.574	0.712	0.346	0.857	0.808	0.302
KZL2ZP	0.454	1.099	1.552	0.444	1.220	1.040	0.571	0.729	1.160	0.418
MorZS0	0.0596	0.0556	0.0442	0.0695	0.0669	0.0588	0.0658	0.0499	0.0465	0.0540
MorZL0	0.0879	0.1157	0.0562	0.0603	0.1028	0.0533	0.0687	0.0751	0.0364	0.0919
MorZP0	0.1008	0.0612	0.0831	0.1036	0.1318	0.1354	0.0725	0.1019	0.1541	0.1249
AlphaZS	0.530	1.387	1.180	0.881	1.374	0.656	1.145	1.378	1.272	0.779
AlphaZL	0.451	0.467	1.240	0.464	0.678	1.391	0.989	0.961	0.350	0.784
AlphaZP	1.225	0.631	0.440	1.026	0.854	0.993	0.372	1.158	0.972	1.269
BetaZS	0.555	0.177	0.290	0.213	0.362	0.437	0.588	0.448	0.219	0.287
BetaZL	0.530	0.372	0.254	0.365	0.434	0.512	0.219	0.495	0.193	0.405
BetaZP	0.179	0.282	0.585	0.232	0.197	0.551	0.221	0.188	0.360	0.248
Nit0	0.0259	0.0439	0.0157	0.0170	0.0527	0.0189	0.0357	0.0595	0.0517	0.0333
VP2N0	0.0315	0.0285	0.0336	0.0278	0.0263	0.0352	0.0365	0.0287	0.0194	0.0325
VP2D0	0.291	0.142	0.192	0.353	0.356	0.183	0.366	0.185	0.363	0.399
VD2N0	0.0140	0.0164	0.0127	0.0084	0.0108	0.0195	0.0187	0.0082	0.0088	0.0130
VO2S0	0.110	0.175	0.117	0.079	0.149	0.144	0.101	0.119	0.101	0.137
setVPON	46.0	58.4	55.2	36.2	47.4	45.1	65.0	40.1	29.5	25.2
setVOpal	46.0	58.4	55.2	36.2	47.4	45.1	65.0	40.1	29.5	25.2

Mattern and Edwards (2017)

Directed search can improve parameter values over Monte Carlo methods

- EA = Evolutionary Algorithm
- DE = Differential Evolution
- ABC = Artificial Bee Colony

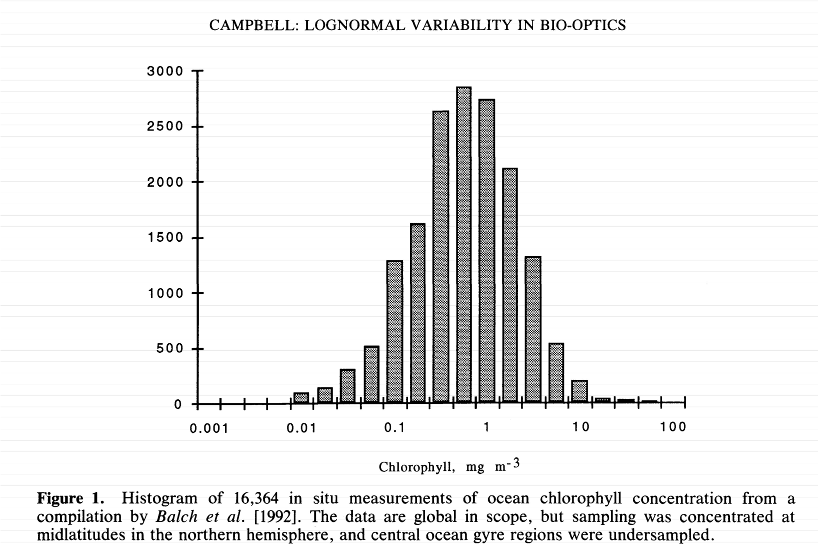




Logarithmic 4D-Var

Logarithm transformation
Surface chl-a

- Gaussian data vs skewed data
- Positive and negative variables vs positive definite concentrations
- We assume lognormal variables
- For 4D-Var, requires additional linearizations



CAMPBELL: LOGNORMAL VARIABILITY IN BIO-OPTICS

Figure 1. Histogram of 16,364 in situ measurements of ocean chlorophyll concentration from a compilation by *Balch et al.* [1992]. The data are global in scope, but sampling was concentrated at midlatitudes in the northern hemisphere, and central ocean gyre regions were undersampled.

Campbell (1995)

Fully Coupled G4DVar and L4DVar using augmented state vector

Gaussian Cost function

$$
\begin{aligned}
J_G(\delta\mathbf{x}_0) &= \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 \\
&\quad + \frac{1}{2}\sum_{i=1}^{N_o}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_{i,0}\delta\mathbf{x}_0)^T\mathbf{R}_i^{-1}(\mathbf{d}_i - \mathbf{H}_i\mathbf{M}_{i,0}\delta\mathbf{x}_0),
\end{aligned}
$$

Lognormal Cost function

$$
\begin{aligned}
&J_L(\delta\mathbf{g}_0) \\
&\quad = \frac{1}{2}\delta\mathbf{g}_0^T\mathbf{B}_L^{-1}\delta\mathbf{g}_0 \\
&\quad + \frac{1}{2}\sum_{i=1}^{N_o}\left(\mathbf{p}_i - \mathbf{L}_i\mathbf{H}_i\mathbf{M}_{i,0}\mathbf{X}_{b,0}\delta\mathbf{g}_0\right)^T\mathbf{R}_{L,i}^{-1}\left(\mathbf{p}_i - \mathbf{L}_i\mathbf{H}_i\mathbf{M}_{i,0}\mathbf{X}_{b,0}\delta\mathbf{g}_0\right),
\end{aligned}
$$

Cost functions can be combined in terms of augmented state vector and error covariances

$$
\delta\mathbf{z} = \begin{bmatrix} \delta\mathbf{x}_G \\ \delta\mathbf{g}_L \end{bmatrix}
$$

$$
\mathbf{B} = \begin{bmatrix} \mathbf{B}_G & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_L \end{bmatrix}
$$

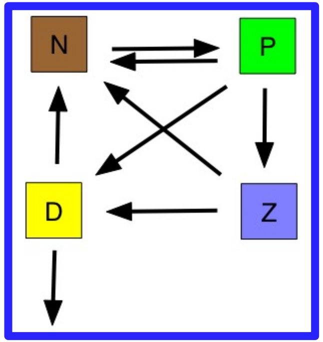
$$
\mathbf{R} = \begin{bmatrix} \mathbf{R}_G & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_L \end{bmatrix}
$$

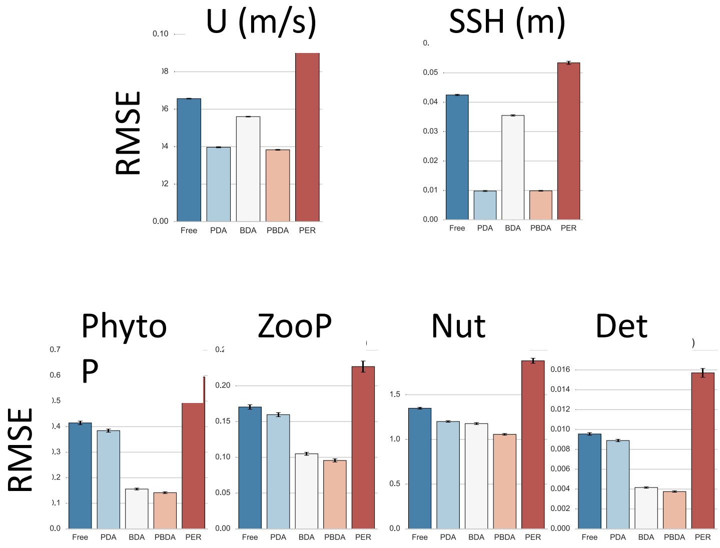
Song et al. (2012, 2016b)

Fully coupled 4DVar
Gaussian (physical) lognormal (biogeochemical)
A ROMS model twin experiment

- Statistics from 30 1-month runs.
- Assimilating physical data and surface Phytoplankton

- Lowest error from combined PBDA





Free Run

Physical DA

Biological DA

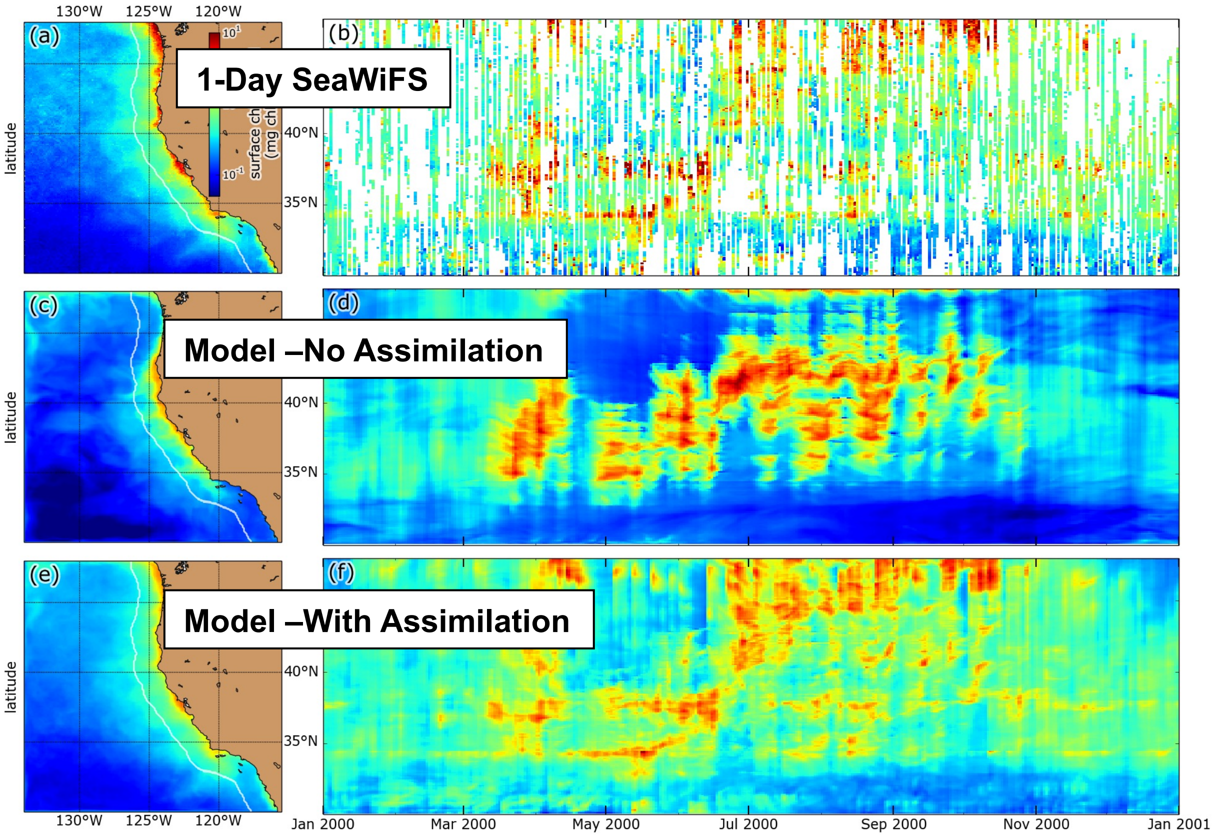
Physical and

Persistence (1 month)

Song et al. (2016b)

Demonstration: fully coupled 4D-Var using NEMURO

- Surface chl-a
- Year 2000



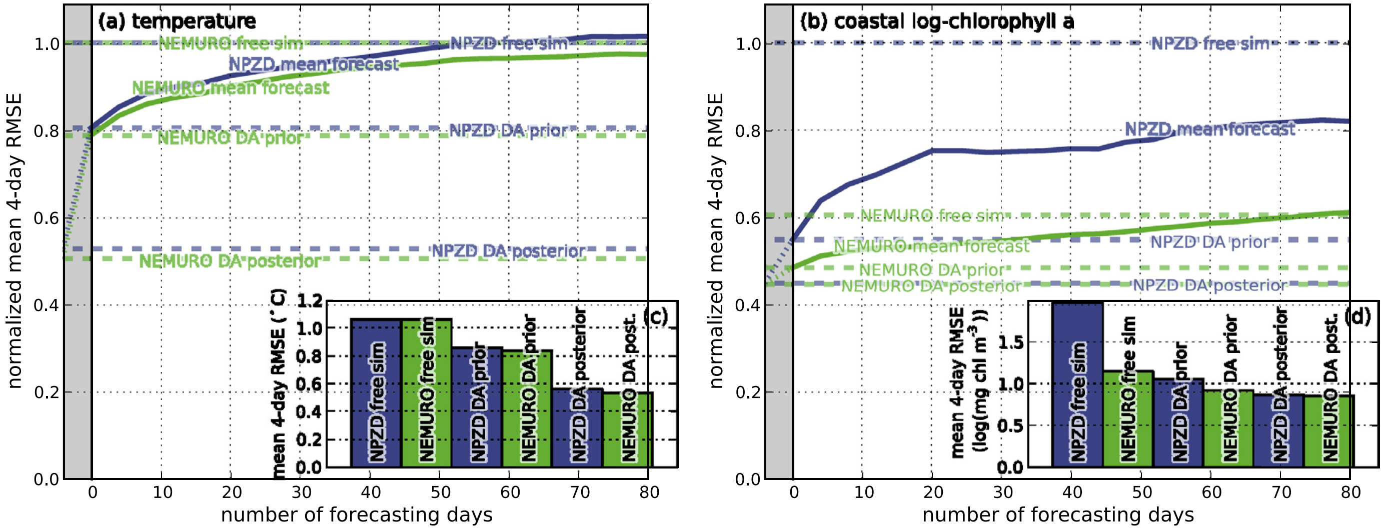
1-Day SeaWiFS

Model –No Assimilation

Model –With Assimilation

Mattern et al. (2017)

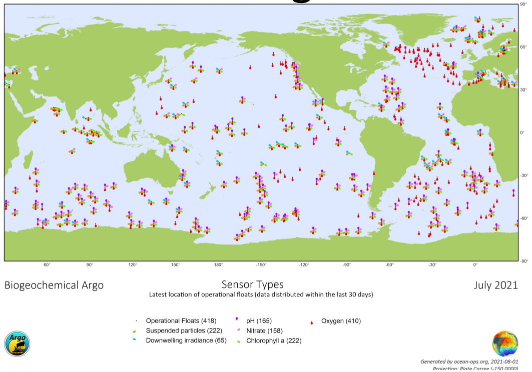
Forecast skill following assimilation is longer for BGC than for physics



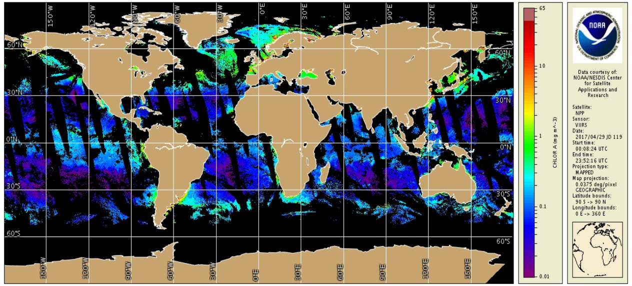
Mattern et al. (2017)

Biogeochemical Data Available for assimilation

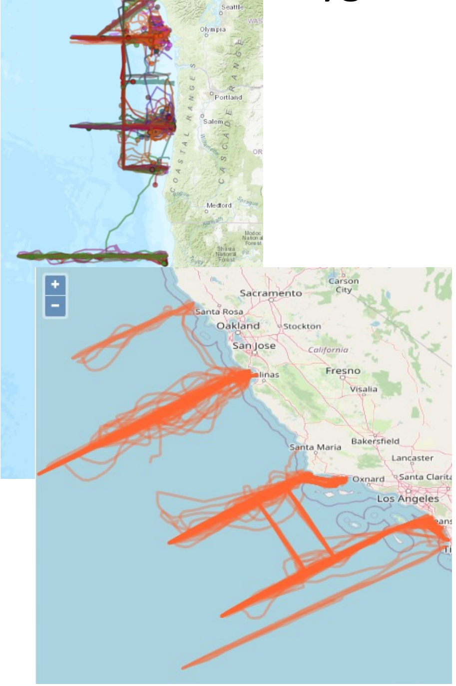
BioArgo



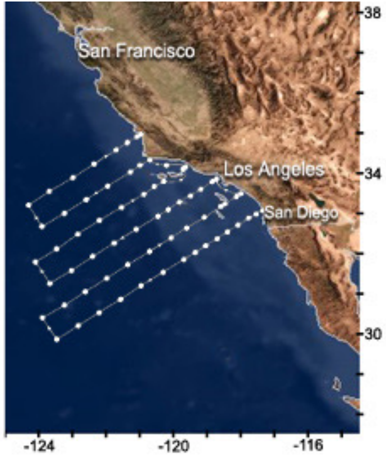
Satellite Chlorophyll



Gliders increasingly have fluorescence, oxygen, nitrate, pH



Shipboard

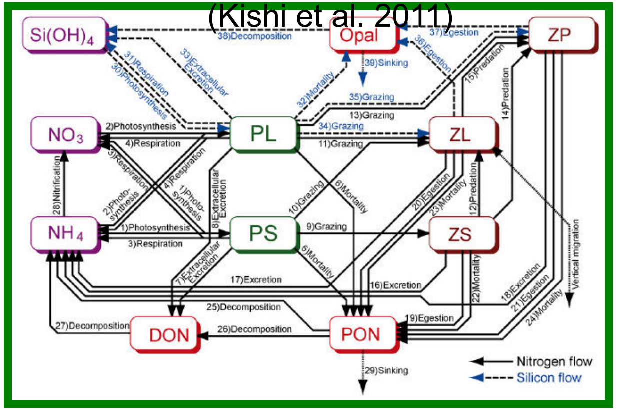


The observational challenge for biogeochemical assimilation

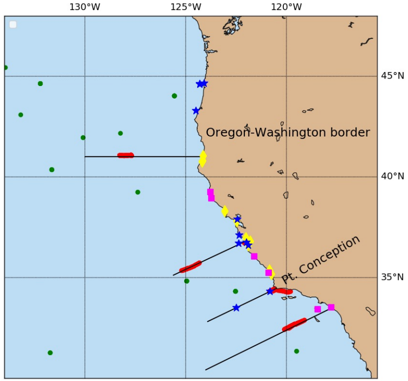
- Mismatch between state variables and observations
- Available (SCHL, in situ chl, nitrate, oxygen)
- Needed
  - Better spatial coverage (true of Physics too)
  - More state variables observed
    - Phytoplankton type (starting to be product at CCI)
    - Zooplankton (obs in counts, hard to convert to biomass)
    - PON, DON
    - Carbonate system requires pH (starting to become available) and one other component (e.g., TIC, pCO_2).

NEMURO

(Kishi et al. 2011)



In situ assets for one cycle



Summary

- Regional physical data assimilation using 4D-Var are quite mature
- Routinely used in multiple IOOS Regional Associations
- NOAA WCOFS product is operational since March 2021
- Biogeochemical data assimilation using 4D-Var and a logarithm transform well-developed
- Routinely used in CCS
- Multiple studies possible with long reanalyses
- Limited data is a real challenge
  - Physics would benefit from increased subsurface T&S.
  - BGC would benefit from both spatial coverage and new types of observations