Data Assimilation and Recursive Estimation in Coupled Physical-Biological-Ecosystem Ocean Models

Pierre F.J. Lermusiaux

Division of Engineering and Applied Sciences, Harvard University

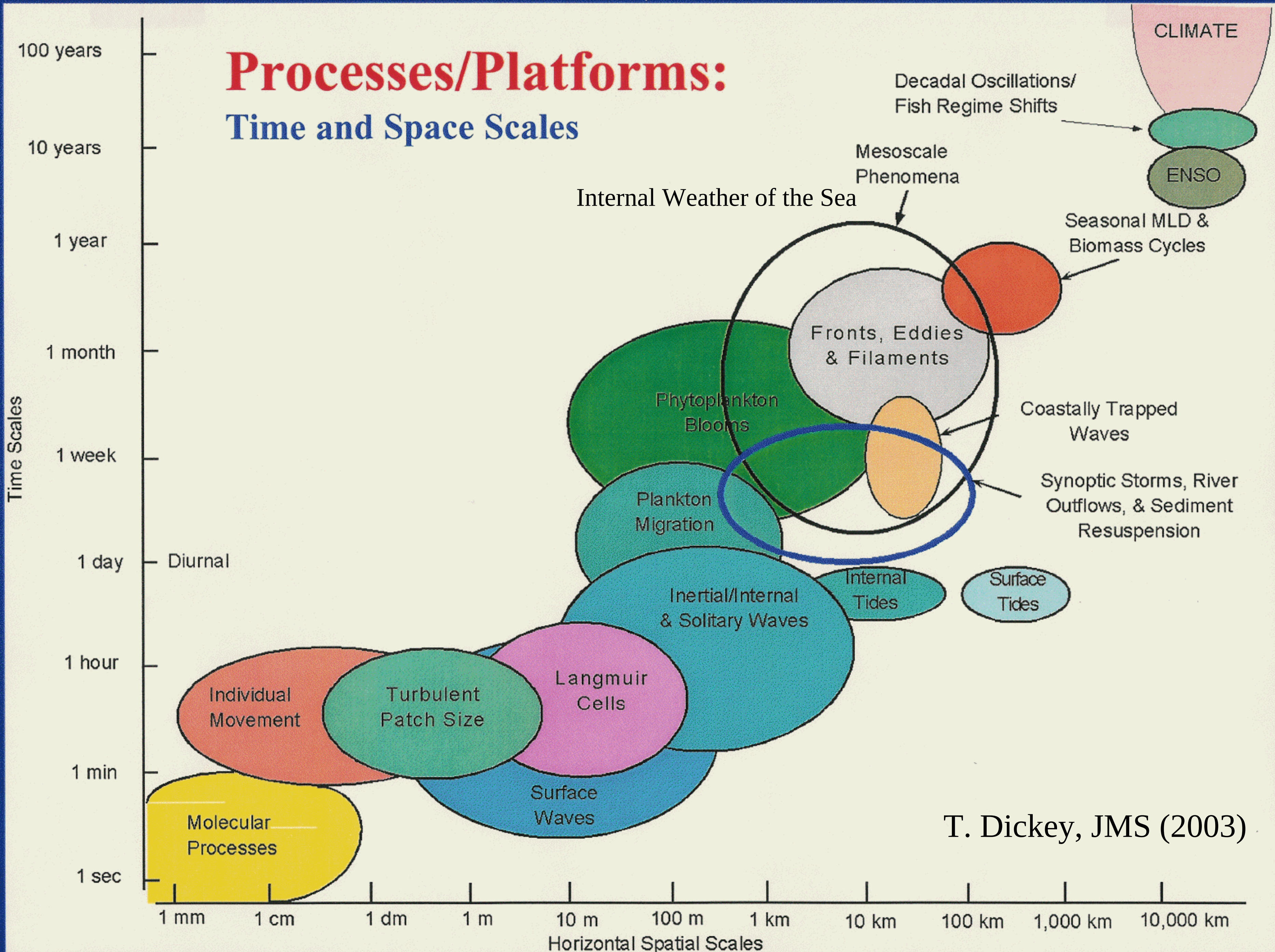
http://www.deas.harvard.edu/~pierrel

---

1. **Interdisciplinary Ocean Science and Data Assimilation**
2. **Different Methods and their Applications**
3. **Error Subspace Statistical Estimation (ESSE)**
   - **Smoothing and Biogeochemical Dominant Dynamical Balances (Mass Bay/ Monterey Bay)**
   - **Error Forecasting, Adaptive Sampling and Adaptive Modeling in Monterey Bay**
4. **Conclusions**

Venice - DARE, September 20, 2004

Processes/Platforms:
Time and Space Scales



T. Dickey, JMS (2003)

OCEANIC FOOD WEB:
Multiple trophic relations

e.g. leading to adult herring (arrows show energy flow)

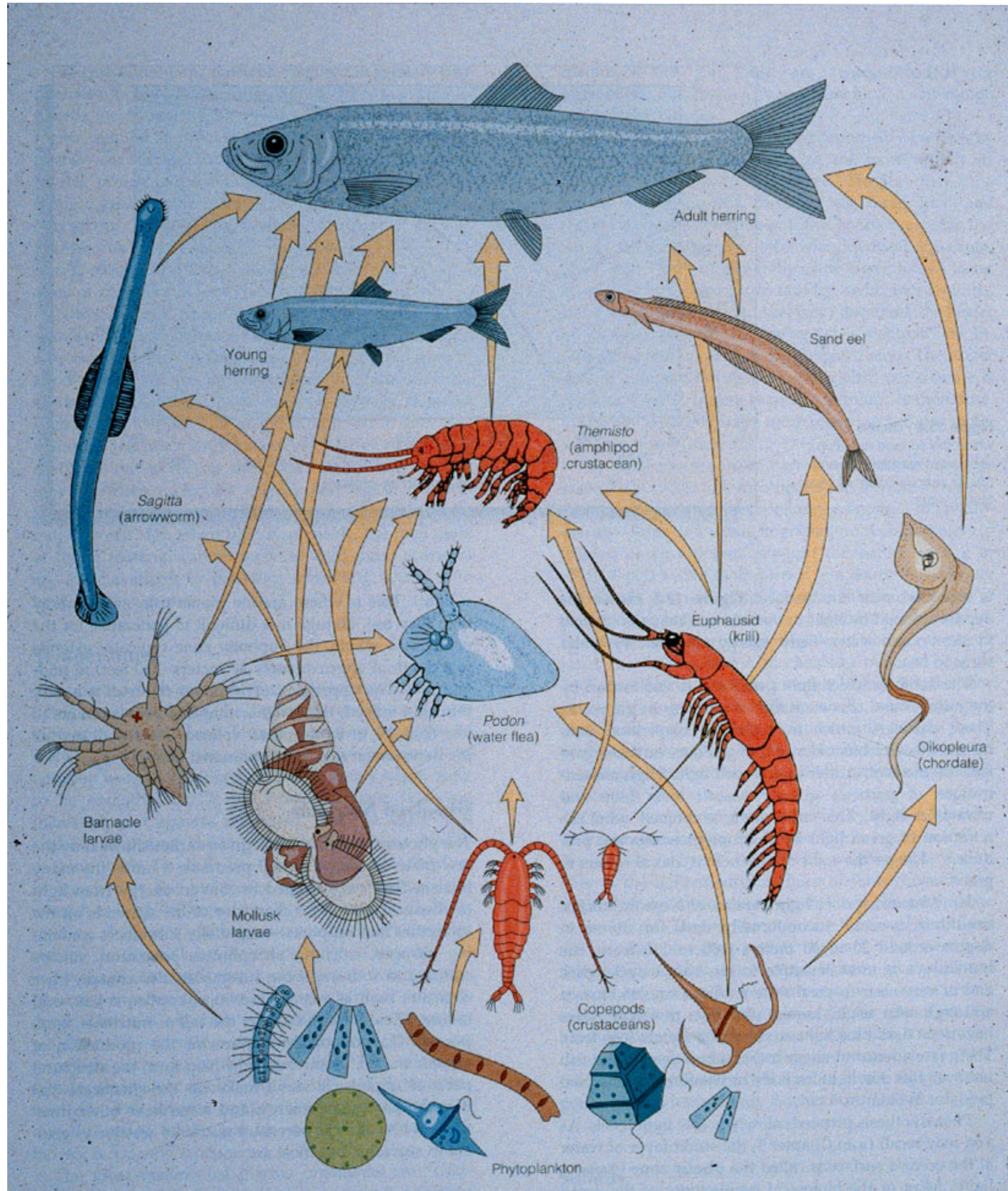


Figure 13.7 A food web, illustrating the major trophic relationships leading to an adult herring. The arrows show the direction of energy flow.

- **Interactions of Physical and Biological/Chemical Dynamical Processes, e.g.**
  - **Primary Productivity**
  - **The Biological Pump and its Role in the Changing Global Carbon Cycle**

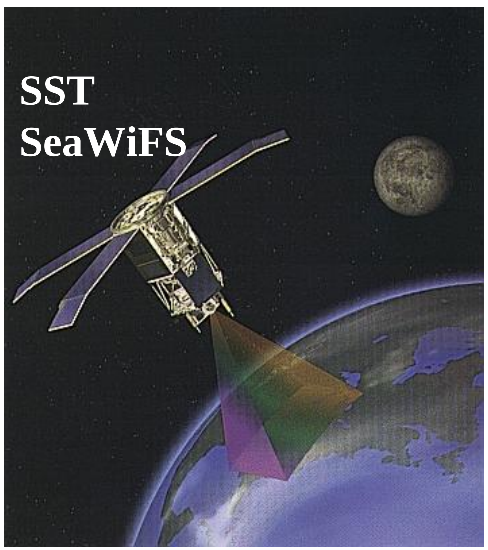
Physical and Multidisciplinary Observations

AUV

Aircraft

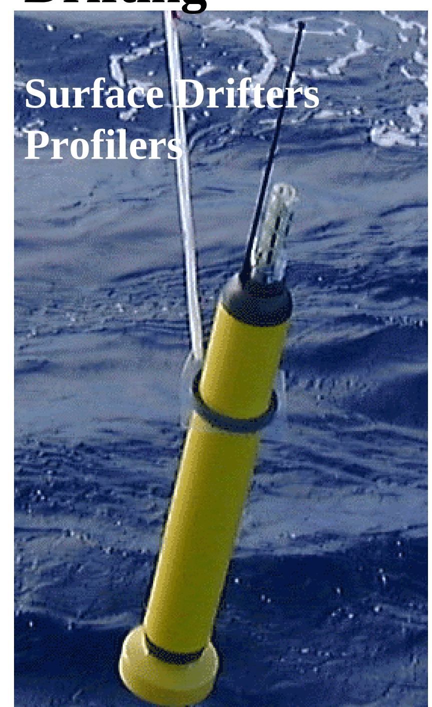
Ships

Satellite



Moored/Fixed

Drifting

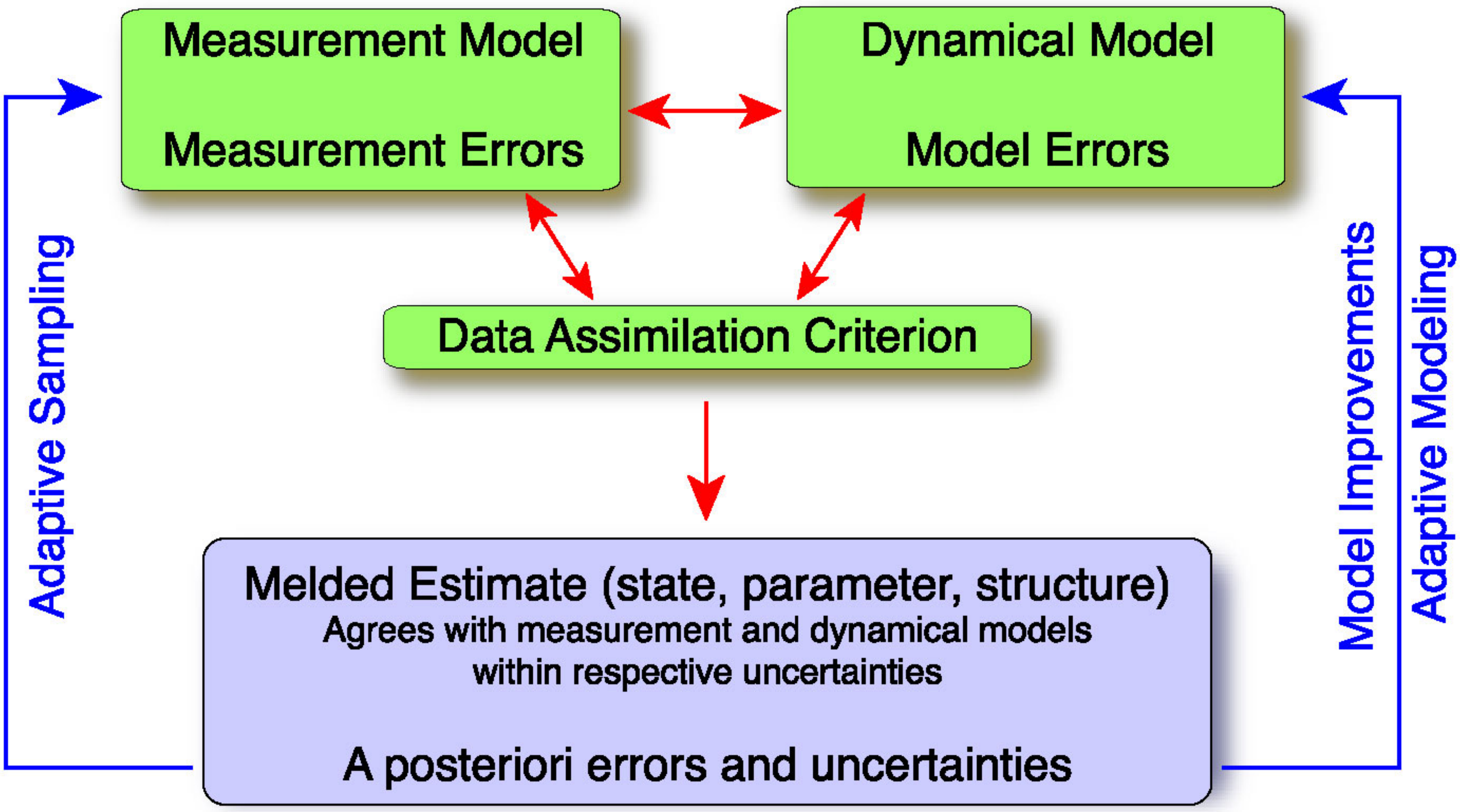


Interdisciplinary Ocean Science and Data Assimilation

- From observations and a priori conservation laws, fundamental ocean science formulates models, usually differential equations, which aim to explain the dynamics of the sea phenomena under study
- Estimation of four-dimensional fields and parameters in the ocean is challenging
  - Multiple interactive scales in space and time (ocean weather: 1-100km, 1-10days)
  - Large domains (e.g. 10-1000km during 10-1000days)
  - Limited ocean data
- Coupled physical-biogeochemical-ecosystem-optical-acoustical modeling and estimations initiated
- Substantial advances require interdisciplinary data assimilation:
  - Quantitative combinations of data and models, in accord with uncertainties
  - Model reductions, simplifications and understanding

WHAT IS DATA ASSIMILATION?

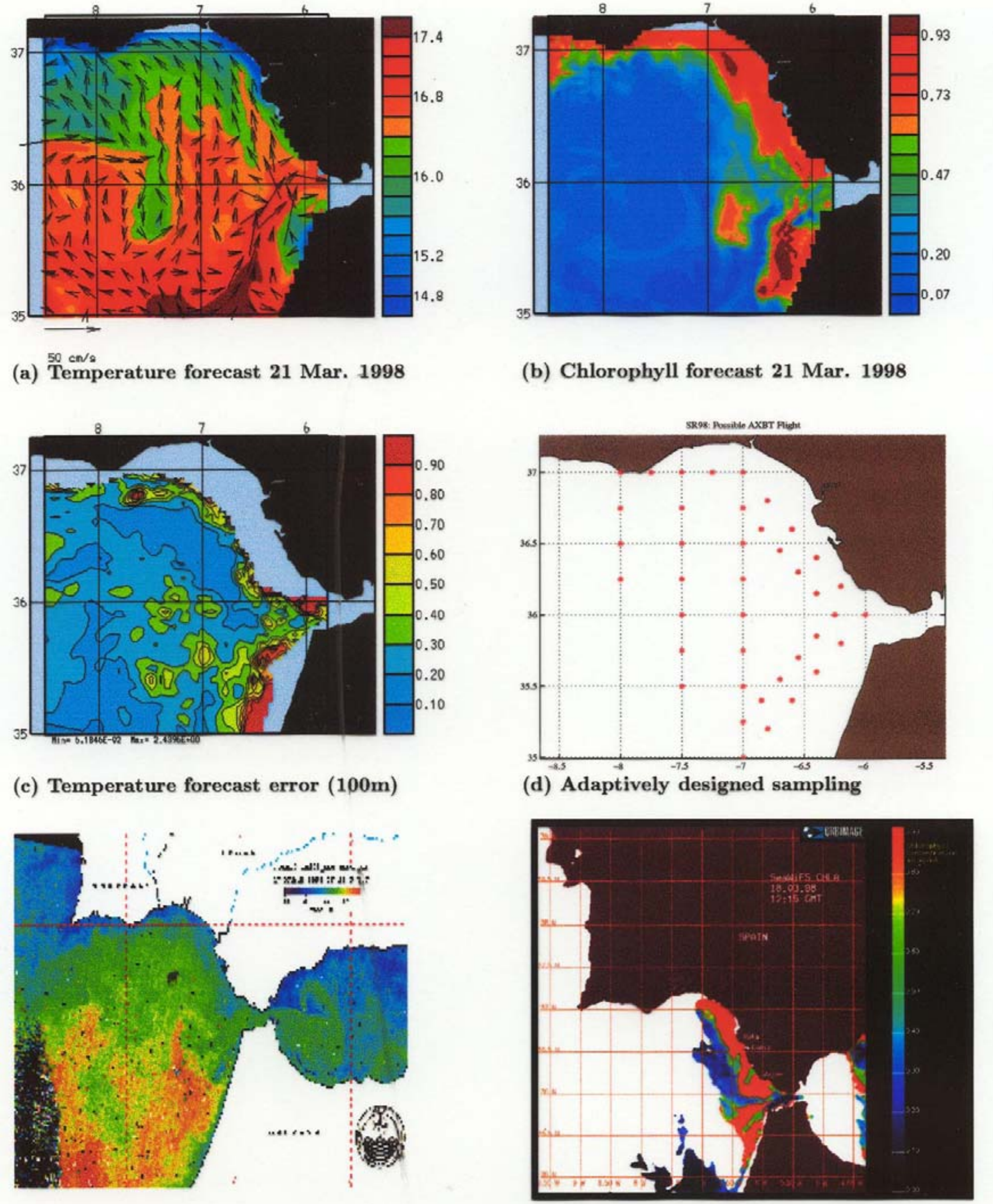
A Melded Estimate of Data and Dynamics



e.g. Robinson A.R., P.F.J. Lermusiaux and N.Q. Sloan, III (1998). *Data Assimilation*. THE SEA, Vol 12.

Gulf of Cadiz, 1998

Real-Time HOPS/ESSE physical-ecosystem predictions



(a) Temperature forecast 21 Mar. 1998

(b) Chlorophyll forecast 21 Mar. 1998

(c) Temperature forecast error (100m)

(d) Adaptively designed sampling

(e) SST from AVHRR 17 Mar. 1998

(f) Chlorophyll from SeaWiFS 18 Mar. 1998

Generic Data Assimilation Problem

Dynamical models:

$$d\phi_i + \mathbf{u} \cdot \nabla \phi_i dt - \nabla (K_i \nabla \phi_i) dt = B_i(\phi_1, \ldots, \phi_i, \ldots, \phi_n) dt + d\eta_i \qquad (i = 1, \cdots, n)$$

e.g. i= u, v, T, $\cdots$, *ZOO*, $\cdots$, p

Parameter equations:

$$dP_\ell = C_\ell(\phi_1, \ldots, \phi_i, \ldots, \phi_n) dt + d\zeta_\ell \qquad (\ell = 1, \cdots, p)$$

e.g. $P_\ell = \{ K_i, R_i, \cdots \}$

Measurement models:

$$y_j = \mathcal{H}_j(\phi_1, \cdots, \phi_i, \cdots, \phi_n) + \epsilon_j \qquad (j = 1, \cdots, m)$$

e.g. $y_j = \{ XBT_j, Fluo_j, SSH_j, CODAR_j \}$

Assimilation criterion:

$$\min_{\phi_i, P_\ell} \quad J(d\eta_i\ ,\ d\zeta_\ell\ ,\ \epsilon_j\ ,\ q_\eta\ ,\ q_\zeta\ ,\ q_\epsilon)$$

CLASSES OF DATA ASSIMILATION SCHEMES

	Error Evol.	Criterion
• **Estimation Theory (Filtering and Smoothing)**		
1. Direct Insertion, Blending, Nudging	- Linear	
2. Optimal interpolation	- Linear	LS
3. Kalman filter/smoother	- Linear	LS
4. Bayesian estimation (Fokker-Plank equations)	- Non-lin.	Non-LS
5. Ensemble/Monte-Carlo methods	- Non-lin.	LS/Non-LS
6. Error-subspace/Reduced-order methods: Square-root filters, e.g. SEEK	- (Non)-Lin.	LS
7. Error Subspace Statistical Estimation (ESSE): 5 and 6	-Non-lin.	LS/Non-LS
• **Control Theory/Calculus of Variations (Smoothing)**		
1. "Adjoint methods" (+ descent)	- Linear	LS
2. Generalized inverse (e.g. Representer method + descent)	- Linear	LS
• **Optimization Theory (Direct local/global smoothing)**		
1. Descent methods (Conjugate gradient, Quasi-Newton, etc)	- Lin	LS/Non-LS
2. Simulated annealing, Genetic algorithms	- Non-lin.	LS/Non-LS

- **Hybrid Schemes**
  - Combinations of the above

Control Theory

(Calculus of Variation Approach, Variational Assimilation)

Adjoint Method

$$\min_{\hat{\boldsymbol{\psi}}_k} J_N = \boldsymbol{\epsilon}_0^T \mathbf{P}_0^{-1} \boldsymbol{\epsilon}_0 + \sum_{k=1}^{N-1} \mathbf{v}_k^T \mathbf{R}_k^{-1} \mathbf{v}_k + \sum_{k=1}^{N} 2\boldsymbol{\lambda}_{k-1}^T \mathbf{w}_{k-1} \quad (19)$$

Dynamical model $\hat{\boldsymbol{\psi}}_k = \mathbf{A}_{k-1} \hat{\boldsymbol{\psi}}_{k-1} \quad k = 1, \ldots, N$ (20)

Initial condition $\hat{\boldsymbol{\psi}}_0 = \boldsymbol{\Psi}_0 + \mathbf{P}_0 \mathbf{A}_0^T \boldsymbol{\lambda}_0$ (21)

Adjoint model $\boldsymbol{\lambda}_{k-1} = \mathbf{A}_k^T \boldsymbol{\lambda}_k + \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \hat{\boldsymbol{\psi}}_k)$ (22)

$k = 1, \ldots, N-1$

Initial condition $\boldsymbol{\lambda}_{N-1} = 0$ (23)

Georges Bank (NW Atlantic)

Estimate full biological source term (RHS) from data (Pseudicalanus spp.)

McGillicuddy et al (1998)

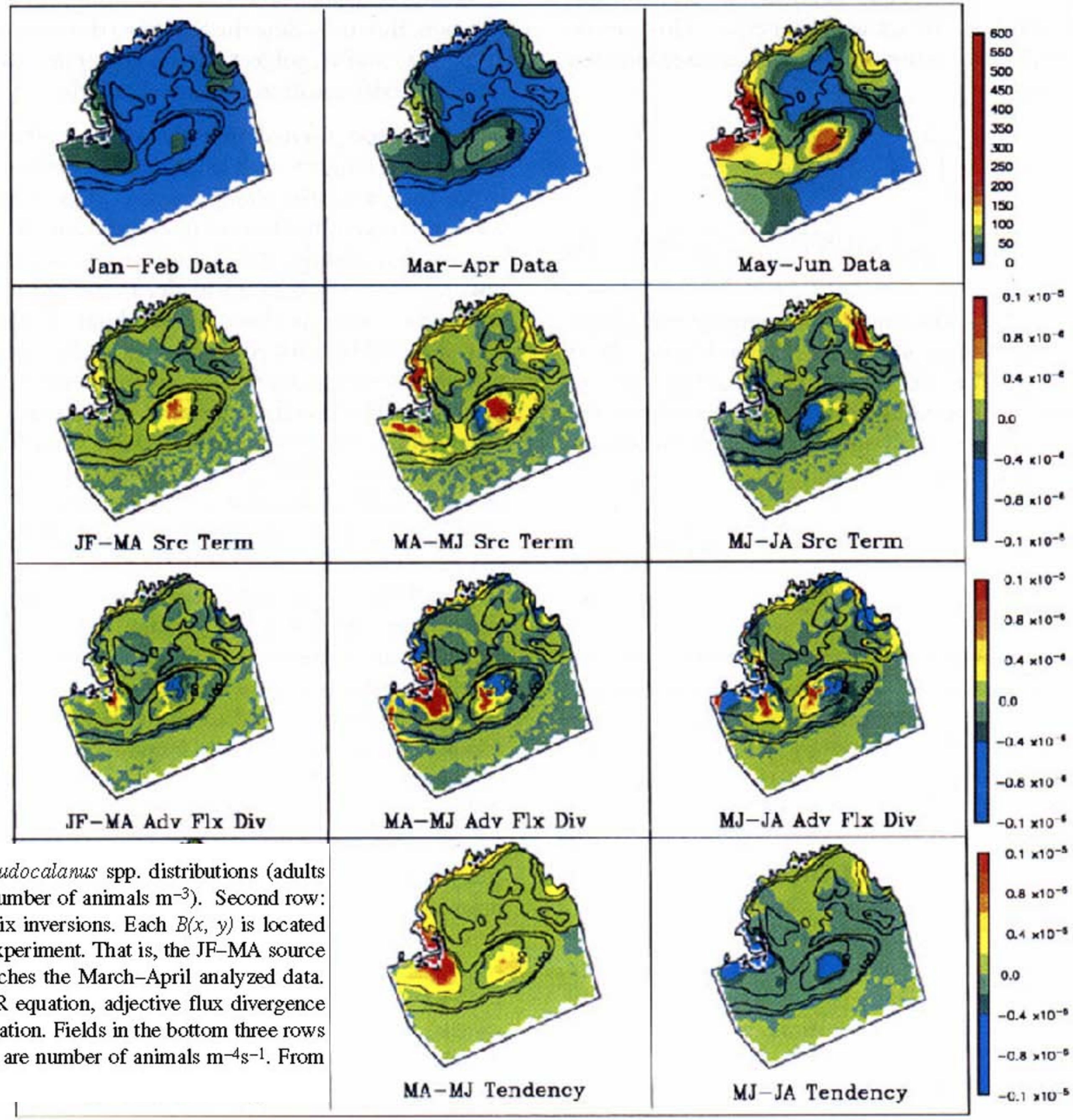


Figure 12.15 Top row: bimonthly climatological *Pseudocalanus* spp. distributions (adults only) objectively analyzed from the MARMAP data (number of animals m^{-3}). Second row: three source terms $B(x, y)$ resulting from three of the six inversions. Each $B(x, y)$ is located directly below the analyzed data used to initialize the experiment. That is, the JF–MA source term results in a forward model integration which matches the March–April analyzed data. Last two rows: two of the remaining terms in the ADR equation, adjective flux divergence and overall tendency, averaged over the period of integration. Fields in the bottom three rows have been normalized to the bottom depth, so the units are number of animals $m^{-4}s^{-1}$. From D. J. McGillicuddy, Jr. et al. (1998)

Direct Minimization Methods

(descent methods, simulated annealing, genetic algorithms, etc)

Comparisons of methods for the estimation of biogeochemical parameters

Vallino, JMR (2000)

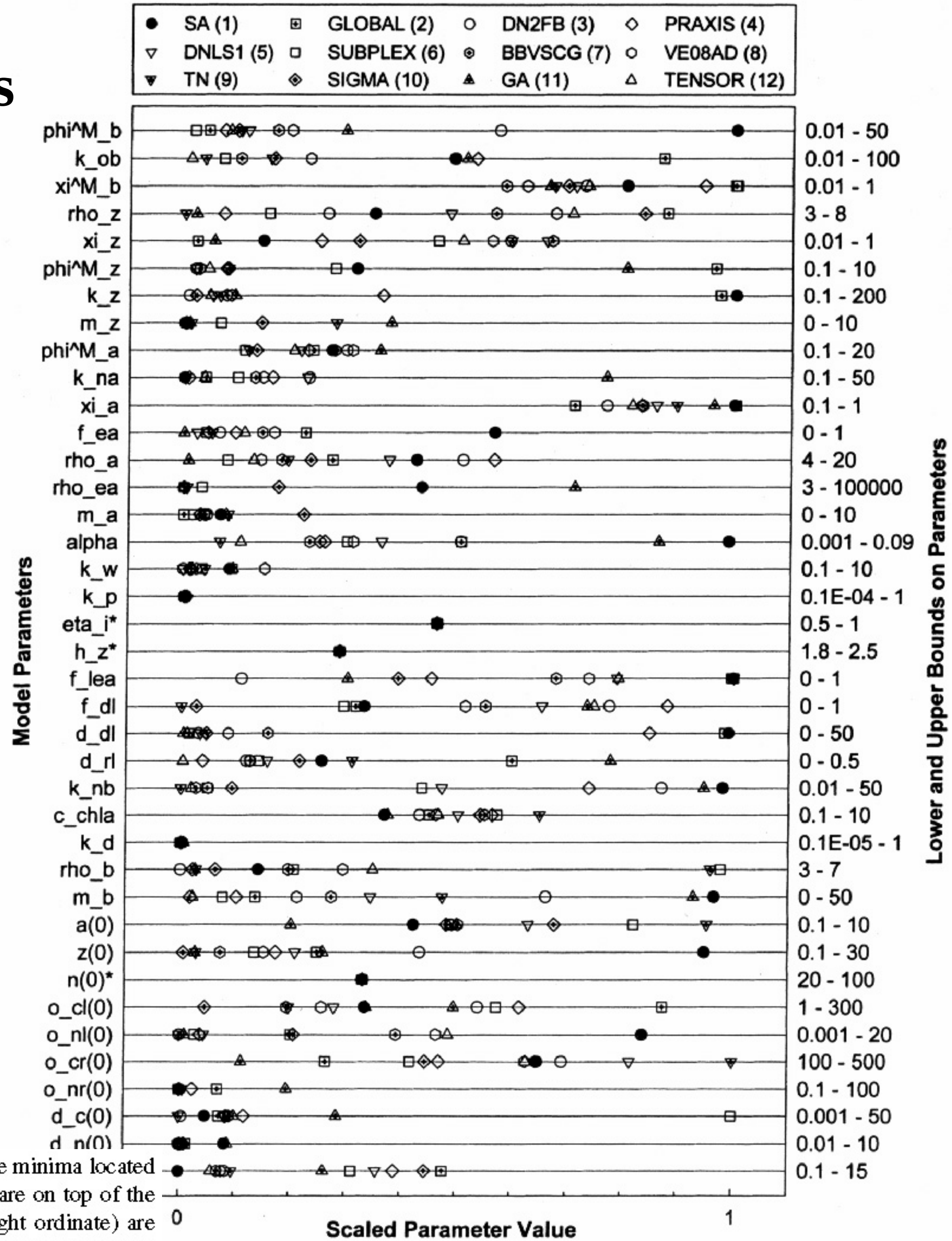


Figure 12.8 Optimized, scaled (0 to 1) parameter values associated with each of the minima located by the twelve optimization routines (abbreviated routine names and their symbols are on top of the figure). Model parameters (left ordinate) and their absolute parameter bounds (right ordinate) are described in Vallino (2000). Parameters marked with an asterisk (left ordinate) were held constant during the data assimilation. From Vallino (2000).

P.F.J. Lermusiaux

Stochastic and Hybrid Methods

Stochastic Methods

Based on (nonlinear) stochastic optimal control

Try to solve the conditional probability density equation (Fokker-Planck)

Minimum error variance, maximum likelihood or minimax estimates

Monte Carlo ensemble calculations

Hybrid methods

Combinations of schemes, for both state and parameter estimation: e.g.,

- Error Subspace Statistical Estimation (ESSE) schemes
- Kalman filter first, then Representer method.

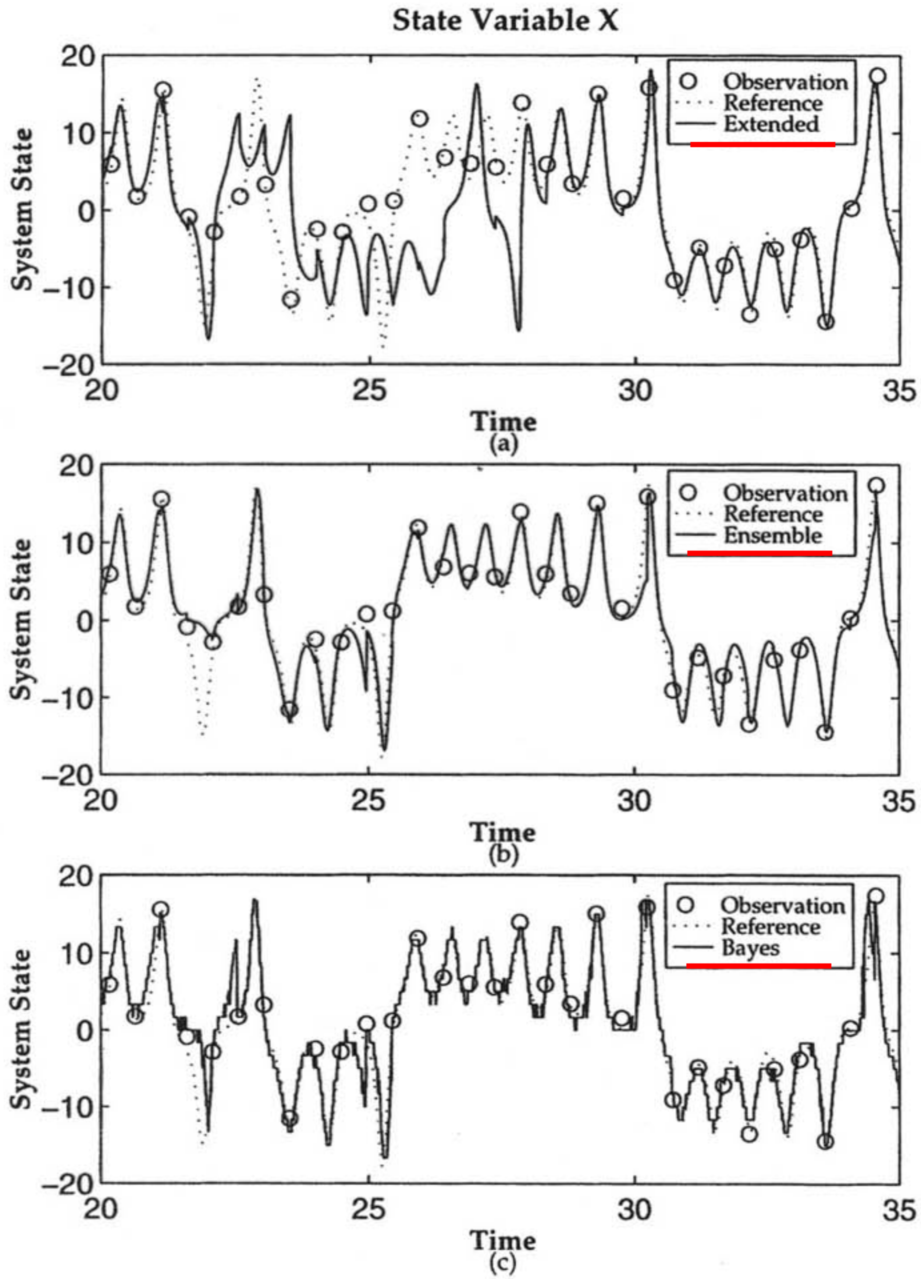


Fig. 8. Results of data assimilation experiments. State variable X, reference solution, observations, and output of data assimilation scheme. (a) The extended Kalman filter (EKF). (b) The ensemble EKF. (c) The mode of the conditional distribution, calculated by Monte-Carlo and Bayes' theorem.

Miller et al, Tellus (1999)
Data assimilation into nonlinear stochastic models

Data Assimilation via ESSE

Table 1. Filtering/Smoothing via ESSE: Continuous-Discrete Problem Statement

Dynamical Model:	$d\widehat{\mathbf{x}} = \mathcal{M}(\widehat{\mathbf{x}})\, dt + d\widehat{\boldsymbol{\eta}}\,, \text{ with } \widehat{\mathbf{x}}(\mathbf{r}_0, t_0) = \widehat{\mathbf{x}}_0 + \widehat{\mathbf{n}}(0)\,.$
Measurement Model:	$\mathbf{y}_k^o = \mathcal{H}(\mathbf{x}_k) + \widehat{\boldsymbol{\epsilon}}_k\,.$
Estimation Criterion:	
Estimate Error Subpace:	$\left\{ \text{Find } \mathbf{P}_k^p = \mathbf{E}_k \boldsymbol{\Pi}_k \mathbf{E}_k^T \text{ with rank}(\mathbf{E}_k) = p \;\middle|\; \min_{\boldsymbol{\Pi}_k, \mathbf{E}_k} \lVert \mathbf{P}_k - \mathbf{P}_k^p \rVert \right\}$
Estimate State by Min. Err. Var. in ES:	$\left\{ \text{Find } \widehat{\mathbf{x}}_k \;\middle|\; \min_{\widehat{\mathbf{x}}_k} J_k = \text{tr}\left[\mathbf{P}_k^p(+)\right] \text{ using } [\mathbf{y}_0^o, \ldots, \mathbf{y}_k^o / \mathbf{y}_N^o] \right\}$

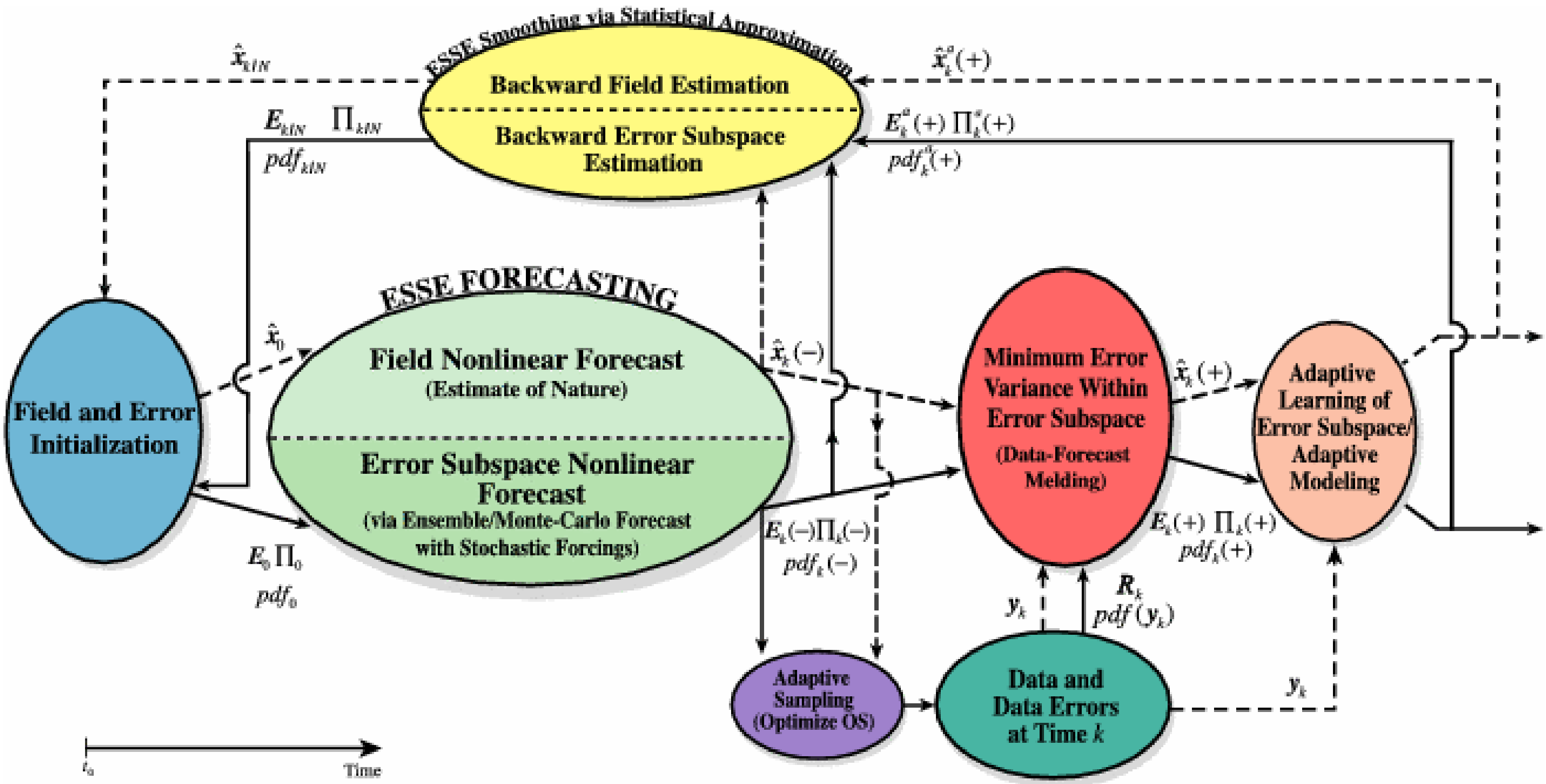
- **Optimal error space reduction and Min. Err. Var. combined:**

"Estimate the ocean evolution by minimizing the largest (most energetic) expected errors, in agreement with the full dynamical model and measurement model (data) constraints, and their respective uncertainties."

- **Linked to POD/Polynomial Chaos, but with time-varying error KL basis:**

$$\mathbf{x}(x,t,\theta) = \bar{\mathbf{x}}(x,t) + \sum_{i=1}^{M} \sqrt{\lambda_i}\;\; \phi_u^s(\mathbf{x},t)\; \zeta_i(\theta)$$

Error Subspace Statistical Estimation (ESSE)



- Uncertainty forecasts (with dynamic error subspace, error learning)
- Ensemble-based (with nonlinear and stochastic model)
- Multivariate, non-homogeneous and non-isotropic DA
- Consistent DA and adaptive sampling schemes
- Software: not tied to any model, but specifics currently tailored to HOPS

Data-Forecast Melding:
Minimum Error Variance within Error Subspace

TRUNCATED Minimum Sample ES Variance, Linear Update (subscript k omitted)

Dynamical State Update:	$\widehat{\mathbf{x}}_{(+)} = \widehat{\mathbf{x}}_{(-)} + \mathbf{K}^p \left(\mathbf{y}^{\mathrm{o}} - \mathcal{H}\big(\widehat{\mathbf{x}}_{(-)}\big)\right)$.
Sample ES Optimal Gain:	$\mathbf{K}^p = \mathbf{E}_- \boldsymbol{\Pi}_{(-)} \mathbf{H}^{p^T} \left(\mathbf{H}^p \boldsymbol{\Pi}_{(-)} \mathbf{H}^{p^T} + \mathbf{R}\right)^{-1}$, where $\mathbf{H}^p \doteq \mathbf{H}\mathbf{E}_-$.
Sample ES Cov. Update:	$\mathbf{L}\boldsymbol{\Pi}_{(+)}\mathbf{L}^T = \boldsymbol{\Pi}_{(-)} - \boldsymbol{\Pi}_{(-)}\mathbf{H}^{p^T}\left(\mathbf{H}^p\boldsymbol{\Pi}_{(-)}\mathbf{H}^{p^T} + \mathbf{R}\right)^{-1}\mathbf{H}^p\boldsymbol{\Pi}_{(-)}$. $\mathbf{E}_+ = \mathbf{E}_- \mathbf{L}$.

ADAPTIVE LEARNING of the Error Subspace (subscript k omitted)

$$\widehat{\mathbf{n}}_{(+)} = \mathbf{K}_{\mathrm{trc}}\big(\mathbf{y}^{\mathrm{o}} - \mathcal{H}\big(\widehat{\mathbf{x}}_{(+)}\big)\big) \ ,$$

$$\mathbf{K}_{\mathrm{trc}} = \mathbf{E}_{\mathrm{trc}(-)}\, \boldsymbol{\Pi}_{\mathrm{trc}(-)}\, \mathbf{H}_{\mathrm{trc}}^T \left(\mathbf{H}_{\mathrm{trc}}\, \boldsymbol{\Pi}_{\mathrm{trc}(-)}\, \mathbf{H}_{\mathrm{trc}}^T + \mathbf{R}\right)^{-1}, \text{ where } \mathbf{H}_{\mathrm{trc}} \doteq \mathbf{H}\mathbf{E}_{\mathrm{trc}(-)} \ .$$

$$\mathbf{E}_+^{\mathrm{a}} \boldsymbol{\Sigma}^{\mathrm{a}}(+) \mathbf{V}_+^{\mathrm{a}\,T} = \mathrm{SVD}_{p+1}\big([\mathbf{E}_+ \boldsymbol{\Sigma}(+)\,,\, \widehat{\mathbf{n}}_{(+)}]\big) \ ,$$

$$\boldsymbol{\Pi}^{\mathrm{a}}(+) = \frac{1}{q+1} \boldsymbol{\Sigma}^{\mathrm{a}\,2}(+) \ .$$

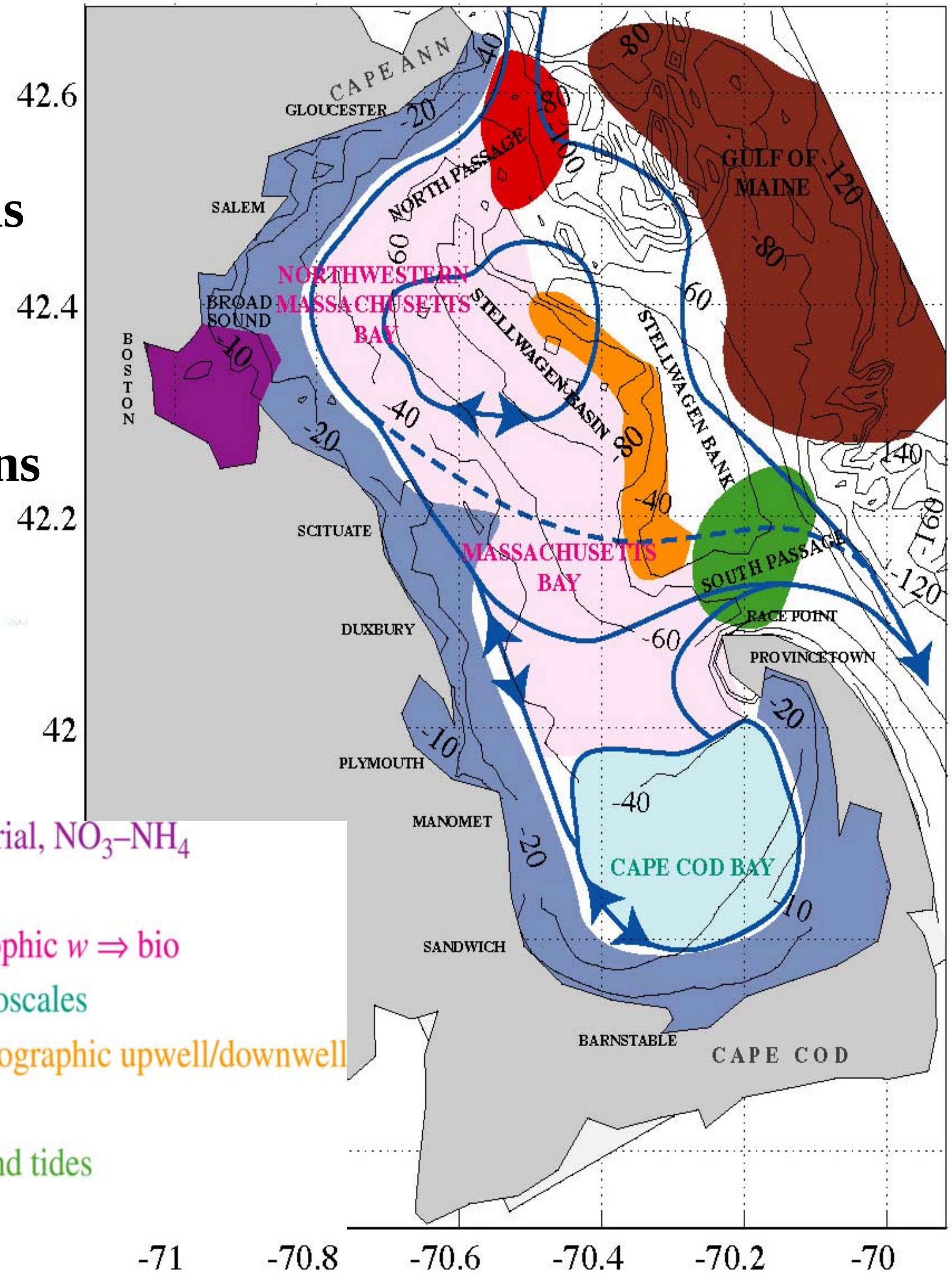
Ocean Regions and Experiments/Operations for which ESSE has been utilized in real-time

- Strait of Sicily (AIS96-RR96), Summer 1996
- Ionian Sea (RR97), Fall 1997
- Gulf of Cadiz (RR98), Spring 1998
- Massachusetts Bay (LOOPS), Fall 1998
- Georges Bank (AFMIS), Spring 2000
- Massachusetts Bay (ASCOT-01), Spring 2001
- Monterey Bay (AOSN-2), Summer 2003

For publications, email me or see http://www.deas.harvard.edu/~pierrel

Massachusetts Bay

Horizontal Circulation Patterns for stratified conditions
(not present at all times)
and
Coupled bio-physical sub-regions in late summer
(Dominant dynamics for trophic enrichment and accumulation)



Boston Harbor: Charles River, sediments, toxic material, NO_3–NH_4

Along Coast: upwelling/downwelling ⇒ bio ↑/↓

Open Bay: submesoscale/mesoscale eddies. Ageostrophic w ⇒ bio

Cape Cod Bay: Horizontal bio advection and submesoscales

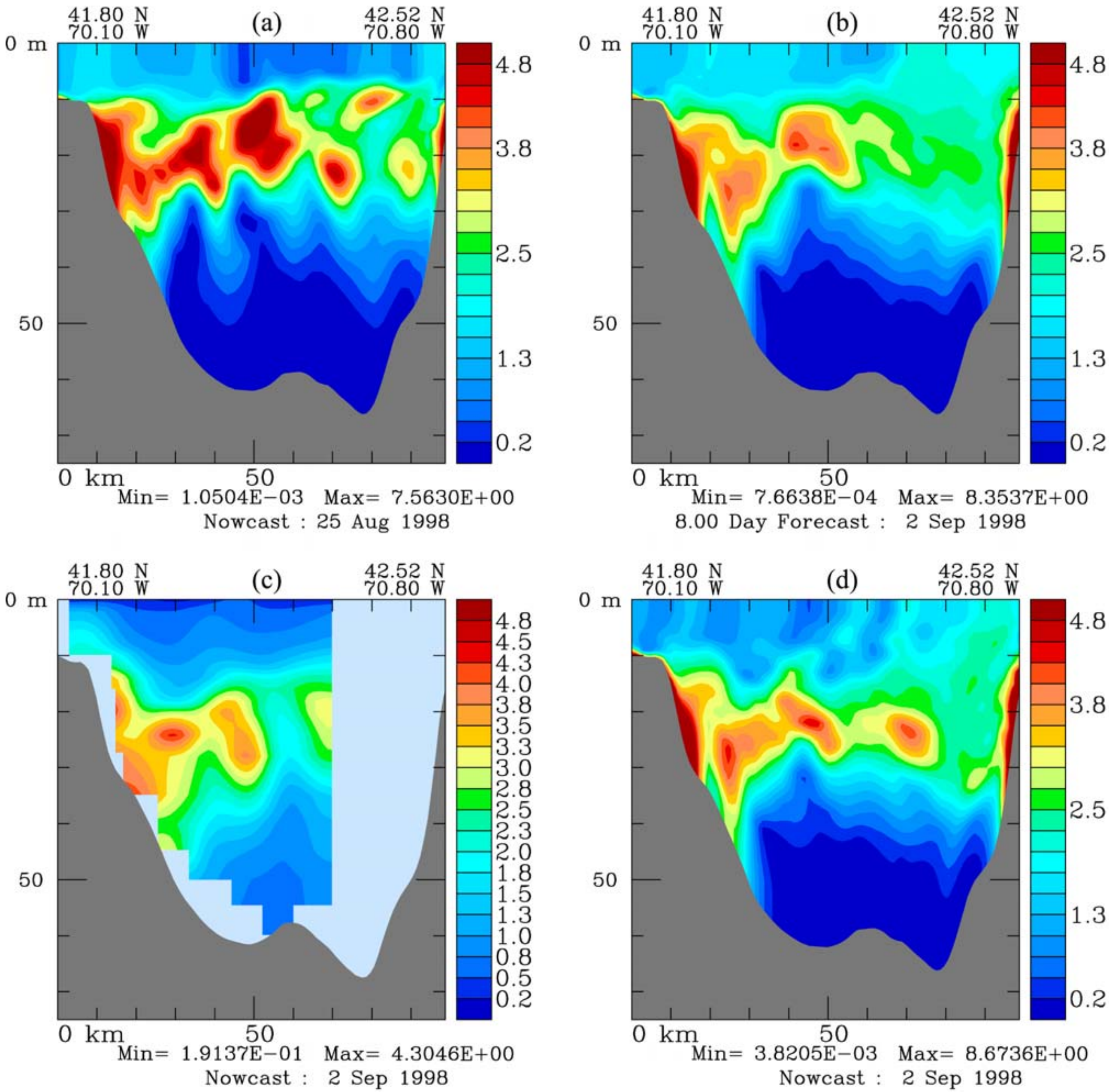
West of Stellwagen Bank: GOM meanders, tides, topographic upwell/downwell

Offshore: GOM meanders

Race Point: Multiple bio advections, accumulation, and tides

Cape Ann: Physical instabilities at GOM inflow

Coupled Physical-Biogeochemical **Smoothing** via ESSE



Cross-sections in Chl-a fields, from south to north along main axis of Massachusetts Bay, with:

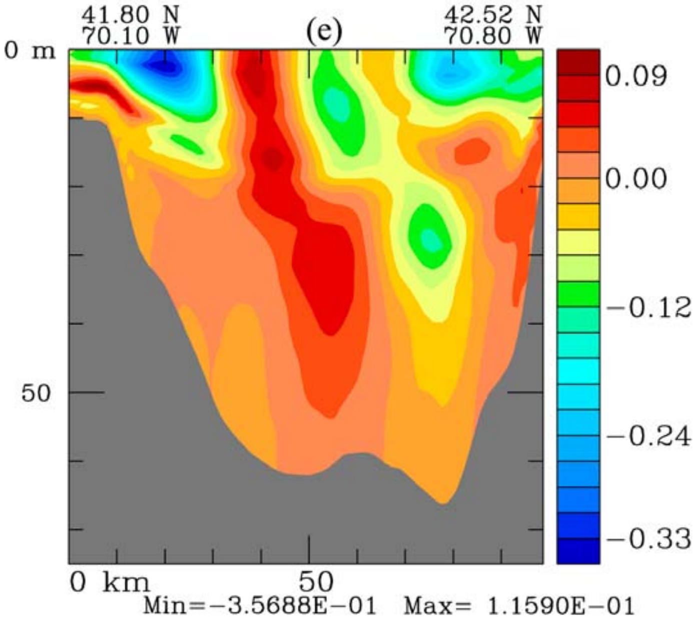
a) Nowcast on Aug. 25

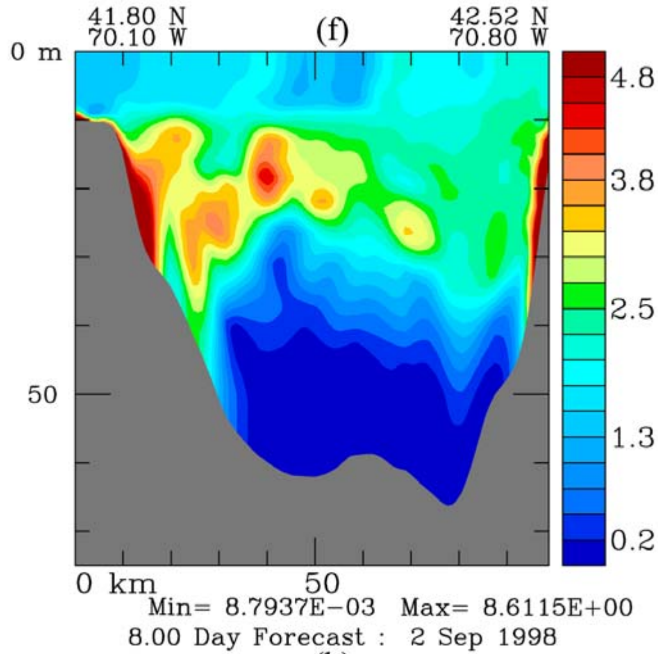
b) Forecast for Sep. 2

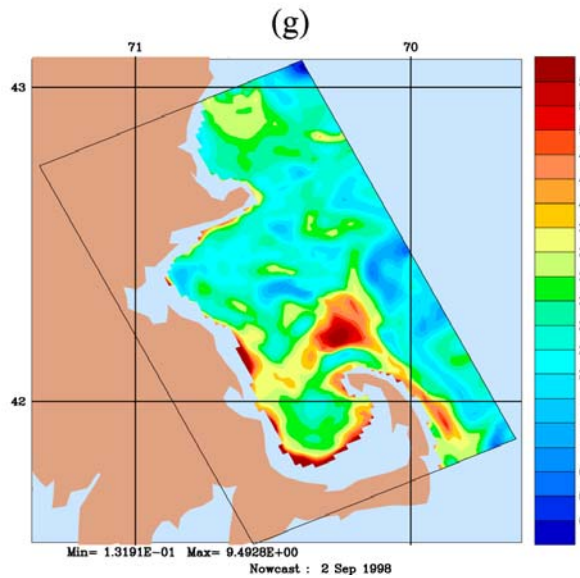
c) 2D objective analysis for Sep. 2 of Chl-a data collected on Sep. 2–3

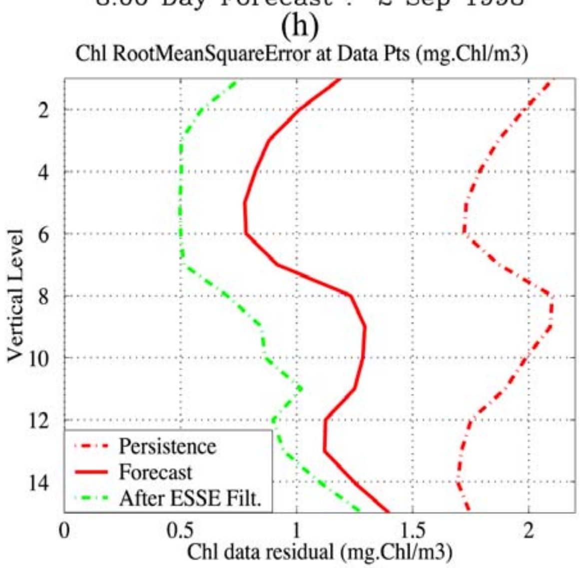
d) ESSE filtering estimate on Sep. 2

Coupled Physical-Biogeochemical DA via ESSE (continued)









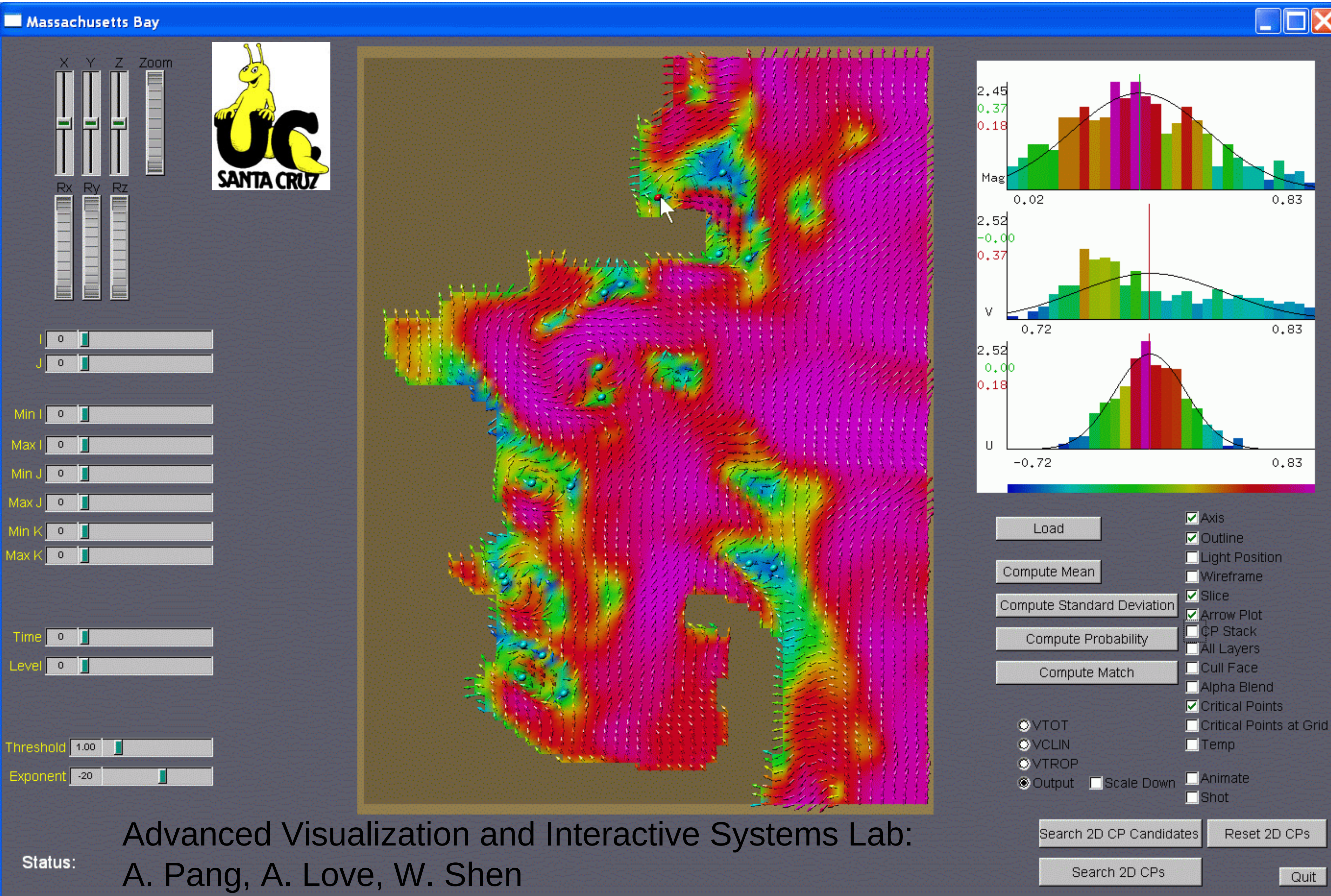
e) Difference between ESSE smoothing estimate on Aug. 25 and nowcast on Aug. 25

f) Forecast for Sep. 2, starting from ESSE smoothing estimate on Aug. 25

(g): as d), but for Chl-a at 20 m depth

(h): RMS differences between Chl-a data on Sep. 2 and the field estimates at these data-points as a function of depth (specifically, "RMS-error" for persistence, dynamical forecast and ESSE filtering estimate) Internal predictability: 2 weeks

Interactive Visualization and Targeting of pdfs



Advanced Visualization and Interactive Systems Lab:
A. Pang, A. Love, W. Shen

A Quest for Dominant Dynamical Balances

- **Ocean dynamics is complex, with multiple scales, processes and features**
- **Ultimate basic understanding is relatively simple but hard to reach**
- **Modern approach:**
  - Combine data and dynamical models quantitatively for realistic studies
  - A road towards understanding and simplified dynamics
- **Many oceanic features can be described by limited number terms, said to be in approximate ``balance'': e.g. geostrophy, Ekman layer**
- **Focus here: explore dominant (dynamical) biogeochemical and biogeochemical-physical balances in coastal ecosystems**
- **Such balances are essential constraints for optimal sampling, biogeochemical initialization and selection of model parameters**

COUPLED PHYSICAL-BIOGEOCHEMICAL MODELS

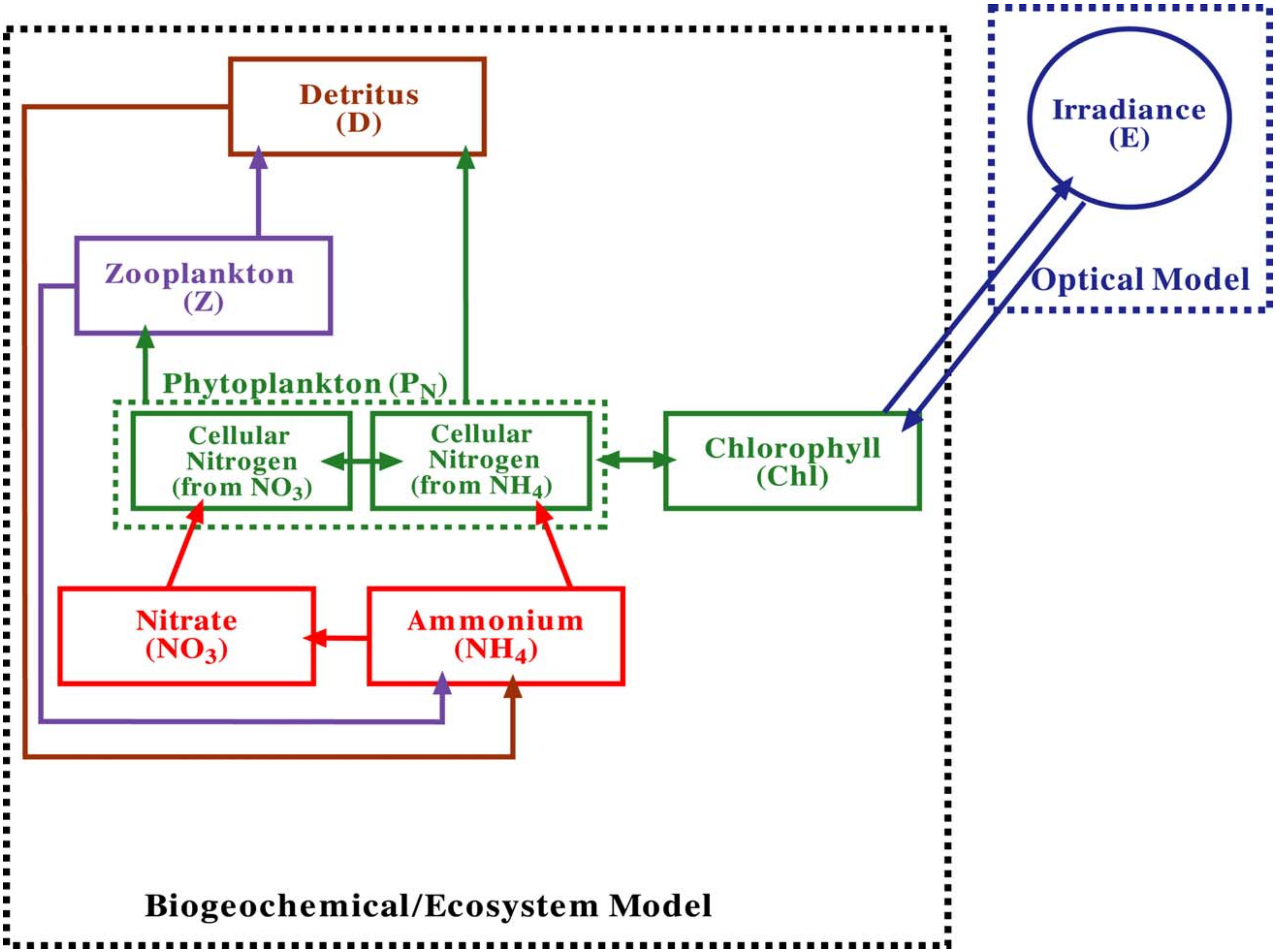
- **Physical model: Primitive-Equation (PDE, x, y, z, t: HOPS)**

Horiz. Mom.	$\frac{D\mathbf{u_h}}{Dt} + f\,\mathbf{e}_3 \wedge \mathbf{u}_h = -\frac{1}{\rho_0}\nabla_h\, p_w + \nabla_h \cdot (A_h \nabla_h \mathbf{u}_h) + \frac{\partial\ A_v\ \partial \mathbf{u}_h/\partial z}{\partial z}$	**(1-2)**
Vert. Mom.	$\rho g + \frac{\partial p_w}{\partial z} = 0$	**(3)**
Thermal en.	$\frac{DT}{Dt} = \nabla_h \cdot (K_h \nabla_h T) + \frac{\partial\ K_v\ \partial T/\partial z}{\partial z}$	**(4)**
Cons. of salt	$\frac{DS}{Dt} = \nabla_h \cdot (K_h \nabla_h S) + \frac{\partial\ K_v\ \partial S/\partial z}{\partial z}$	**(5)**
Cons. of mass	$\nabla \cdot \mathbf{u} = 0$	**(6)**
Eqn. of state	$\rho(\mathbf{r}, z, t) = \rho(T, S, p_w)$	**(7)**

- **Biogeochemical model: Generic ADR equation (PDE, x, y, z, t)**

$$\frac{\partial \phi_i}{\partial t} + \mathbf{u} \cdot \nabla \phi_i - \nabla_h (A_i \nabla_h \phi_i) - \frac{\partial\ K_i\, \partial \phi_i/\partial z}{\partial z} = \mathcal{B}_i(\phi_1, \ldots, \phi_i, \ldots, \phi_7) \qquad (\mathbf{8-14})$$

$i = \mathrm{NO3},\ \ \mathrm{P}_{NO3},\ \ \mathrm{ZOO},\ \ \mathrm{NH4},\ \ \mathrm{DET},\ \ \mathrm{CHL},\ \ \mathrm{P}_{NH4}$



Schematic representation of ecosystem model (seven state variables/compartments)

Dominant dynamical balances for initial biogeochemical fields/parameters

Circadian (daily) 0th order biological balance

Phyto eqn.:

$$\frac{1}{T}\int_0^T V^{NO_3} + V^{NH_4} dt \simeq gZ + n_3 P$$

$$V^{NH_4} = \frac{P_P}{\theta^C_{Chl}} \frac{NH_4}{k_{NH_4} + NH_4} P \;,\quad V^{NO_3} = \frac{P_P}{\theta^C_{Chl}} \frac{NO_3\, e^{-\psi NH_4}}{k_{NO_3} + NO_3} P \;,\quad P_P = P_m(1 - e^{-\alpha E/P_m}) e^{-\beta E/P_m}$$

$$Optical\ model: E(x,y,z,t) = E(x,y,0,t) e^{-(k_w z + k_c \int_0^z Chl\, dz)}$$

Zoo eqn.: *Select non-zero root,*

$$0 \simeq (1 - \gamma_1 - \gamma_2) g - n_1 - n_2 Z \quad , \quad g = R_m(1 - e^{-\Lambda P})$$

Nitrate eqn.:

$$\frac{1}{T}\int_0^T V^{NO_3} dt \simeq k_N [NH_4]$$

Ammonium eqn.:

$$\frac{1}{T}\int_0^T V^{NH_4} dt \simeq -k_N [NH_4] + \gamma_1 g Z + (1 - \epsilon_1) n_1 Z + (1 - \epsilon_2) n_2 Z^2 + k_D D$$

Detritus eqn.: $\int_0^H dz$ *with either null* $D = 0$ *or well-mixed* $\frac{\partial D}{\partial z} = 0$ *surf./bot. BC*

$$\nu_D \frac{\partial D}{\partial z} + k_D D \simeq \gamma_2 g Z + \epsilon_1 n_1 Z + \epsilon_2 n_2 Z^2 + n_3 P + n_4 P^2$$

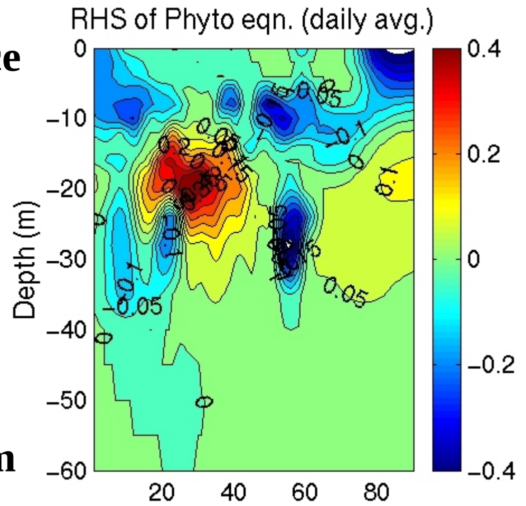
Chlorophyll: *OA-ed from Fluo. For P, assume fully-acclimated C:CHL*

Balance subject to observed variables and parameters constraints

Daily Average of RHS Terms: Residuals vs. Order of Magnitude

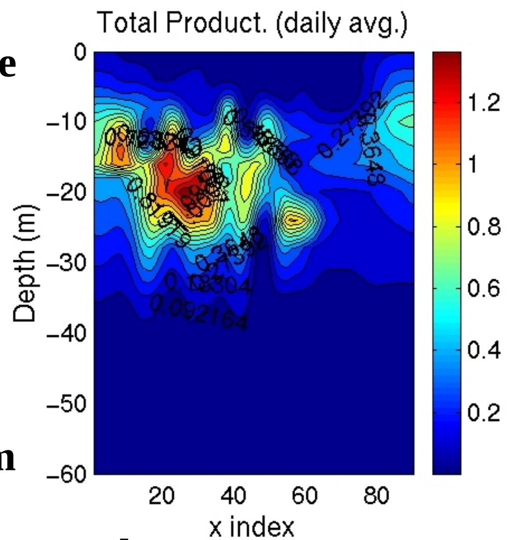
Mass. Bay: Aug 21,

Surface



Bottom

Surface



Bottom

Cape Cod

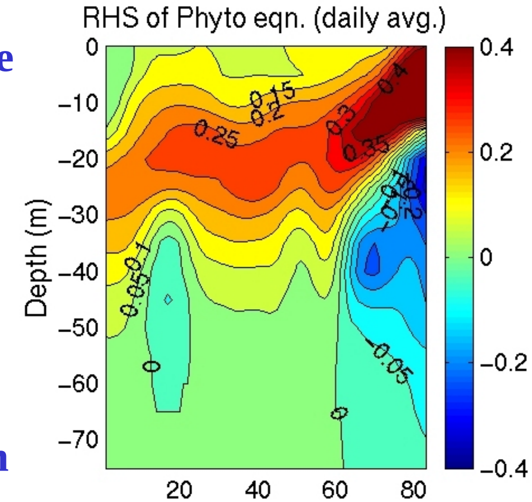
Cape Ann

Residuals

Orders of

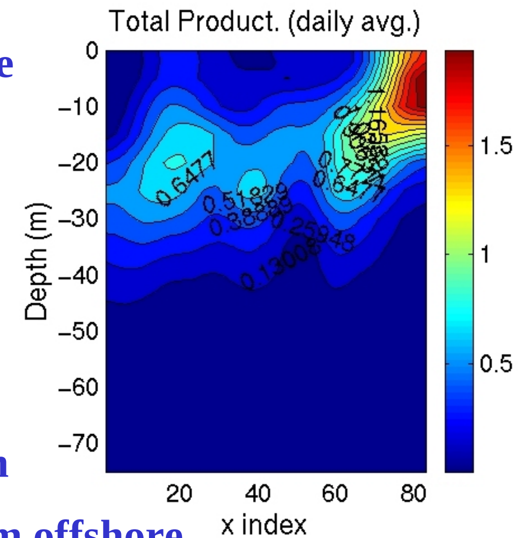
Monterey Bay: Aug 04,

Surface



Bottom

Surface



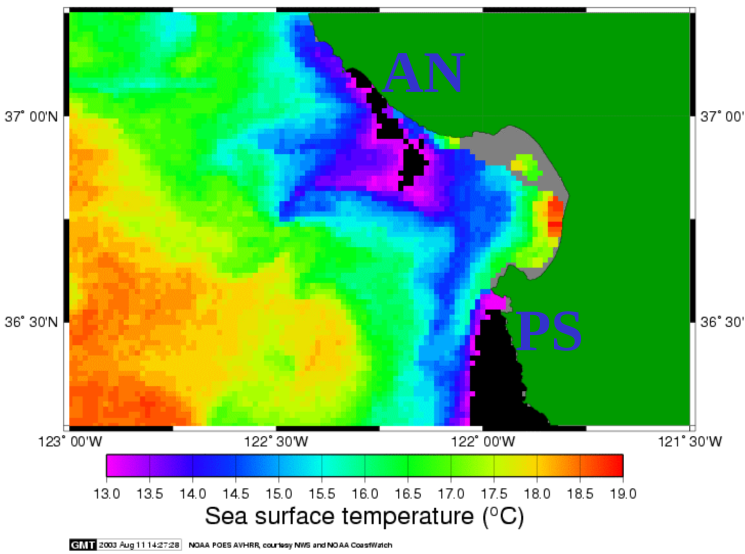
Bottom

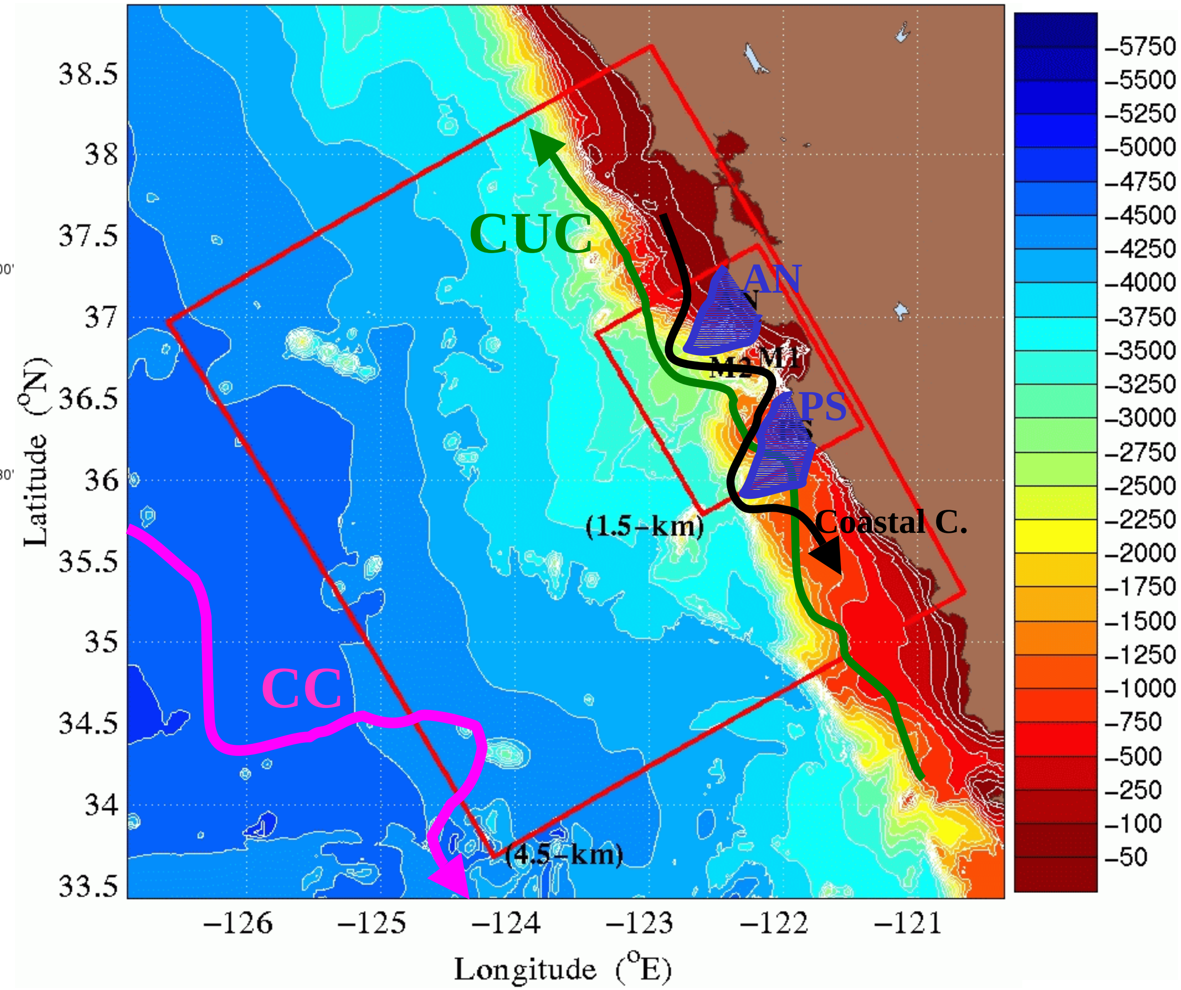
100km offshore

REGIONAL FEATURES of Monterey Bay and California Current System and Real-time Modeling Domains (AOSN2, 4 Aug. – 3 Sep., 2003)

SST on August 11, 2003

Experimental AVHRR HRPT SST August 11, 2003 1850 h UTC

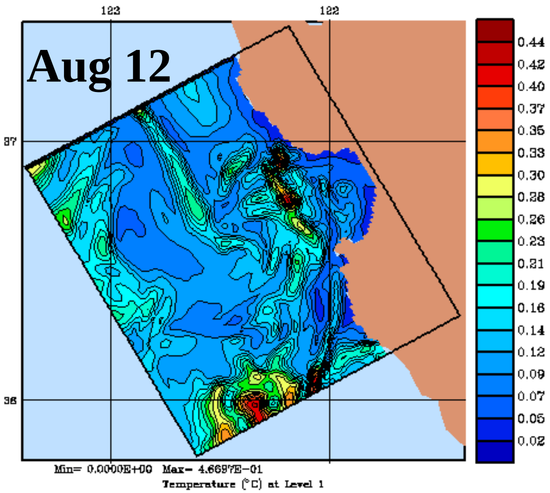




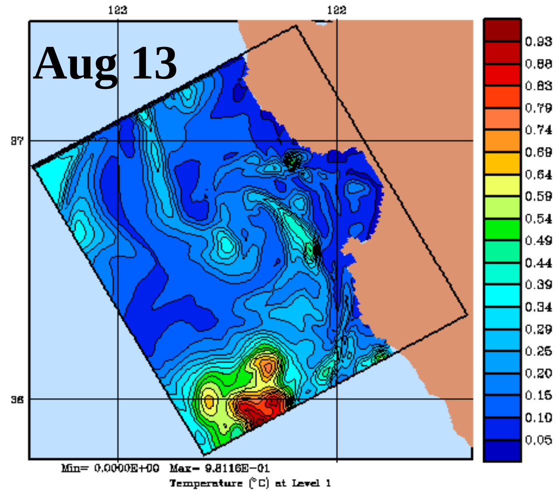
REGIONAL FEATURES

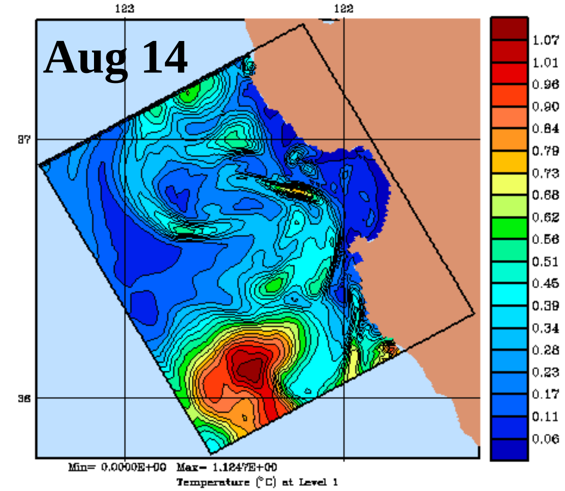
- **Upwelling centers at Pt AN/ Pt Sur**:…………Upwelled water advected equatorward and seaward
- **Coastal current, eddies, squirts, filam., etc:**….Upwelling-induced jets and high (sub)-mesoscale var. in CTZ
- **California Undercurrent (CUC)**:……………..Poleward flow/jet, 10-100km offshore, 50-300m depth
- **California Current (CC)**:……………………Broad southward flow, 100-1350km offshore, 0-500m depth

ESSE Surface Temperature Error Standard Deviation Forecasts

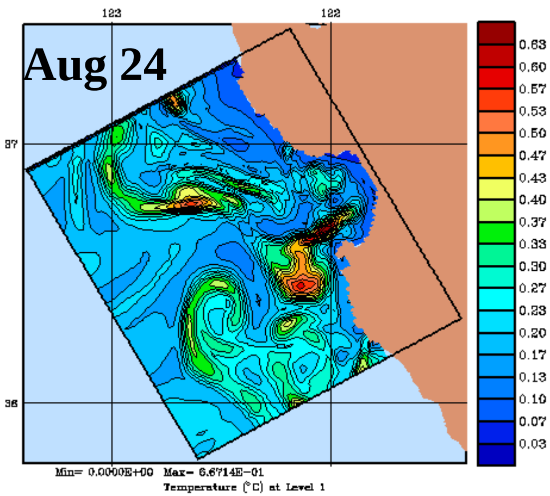


Start of Upwelling

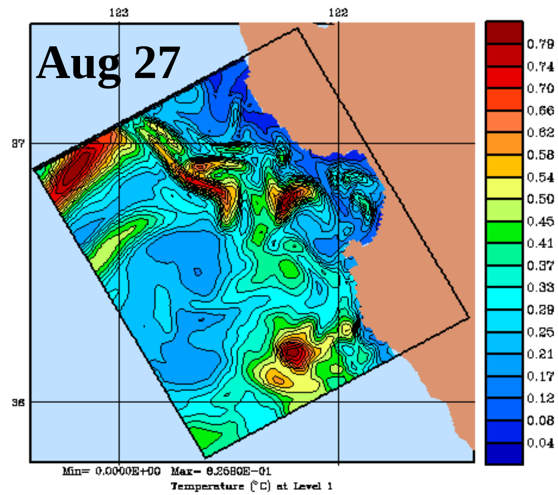


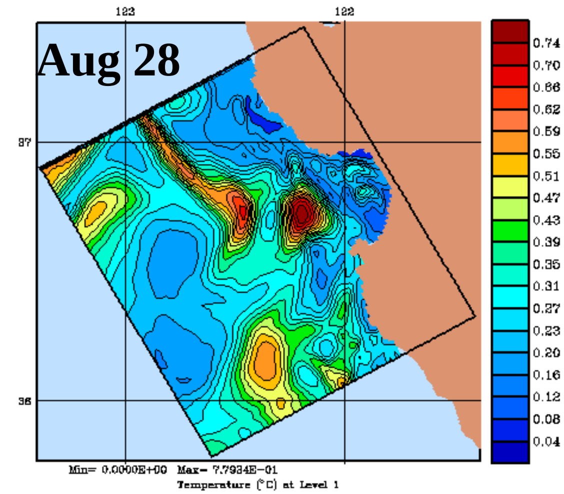


First Upwelling period



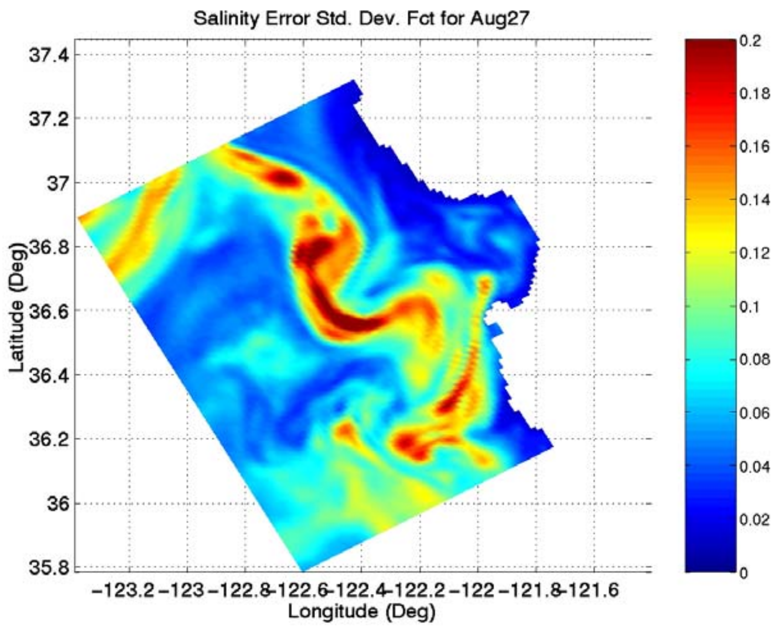
End of Relaxation



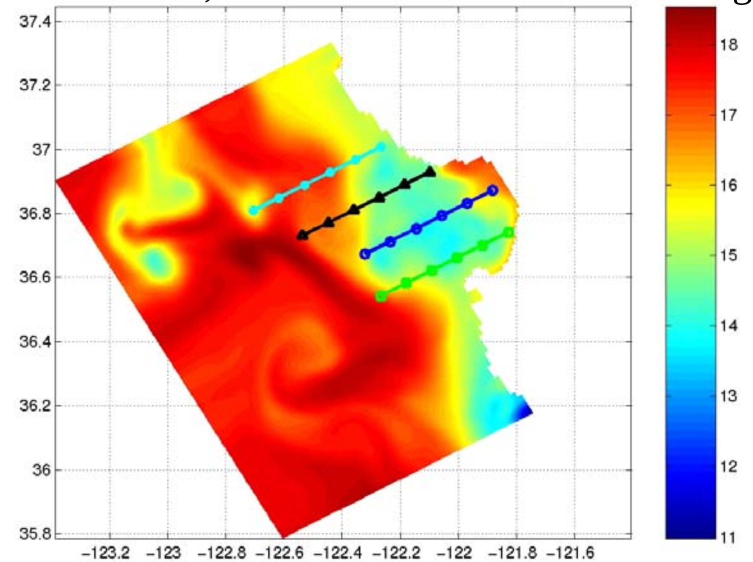


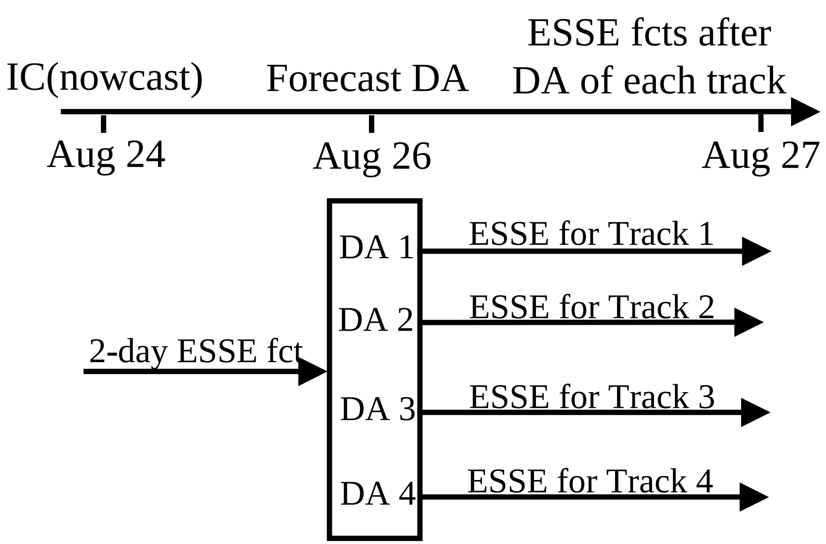
Second Upwelling period

Which sampling on Aug 26 optimally reduces uncertainties on Aug 27?

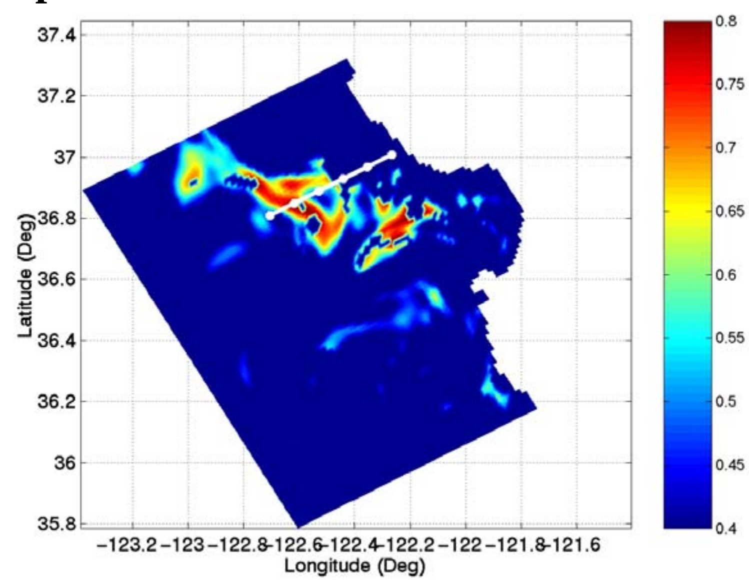


4 candidate tracks, overlaid on surface T fct for Aug 26

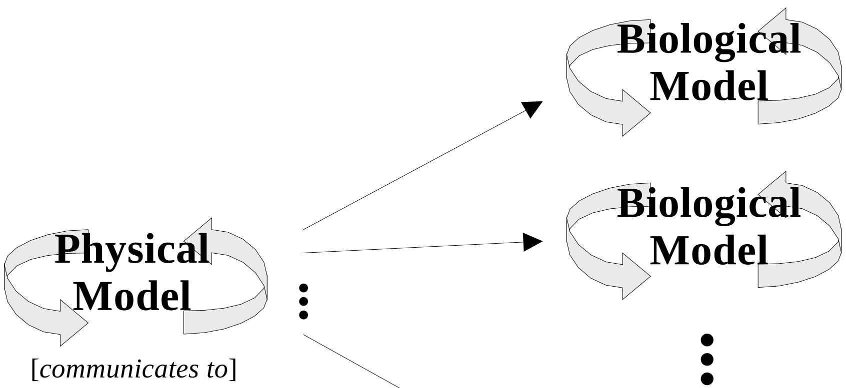


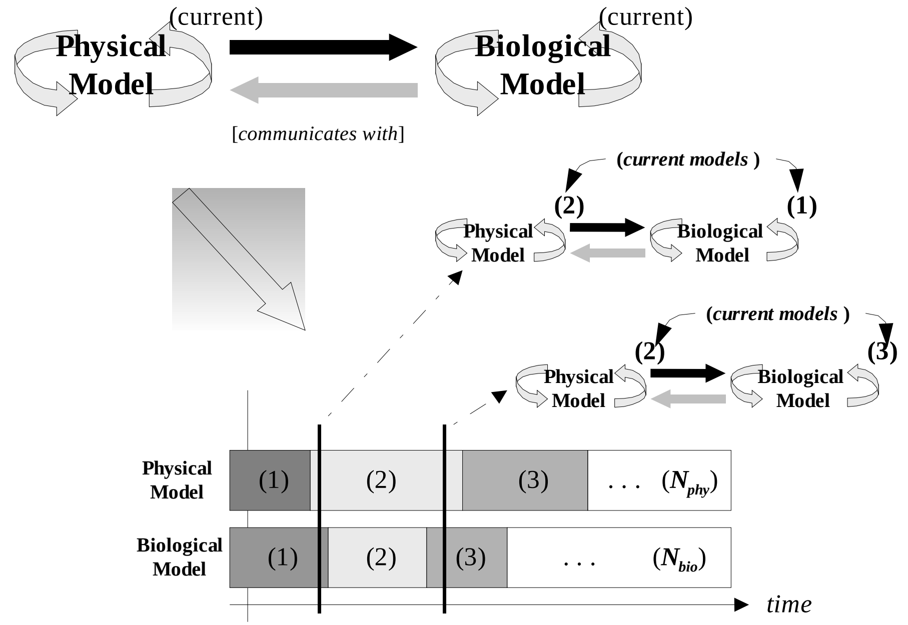


Best predicted relative error reduction: track 1



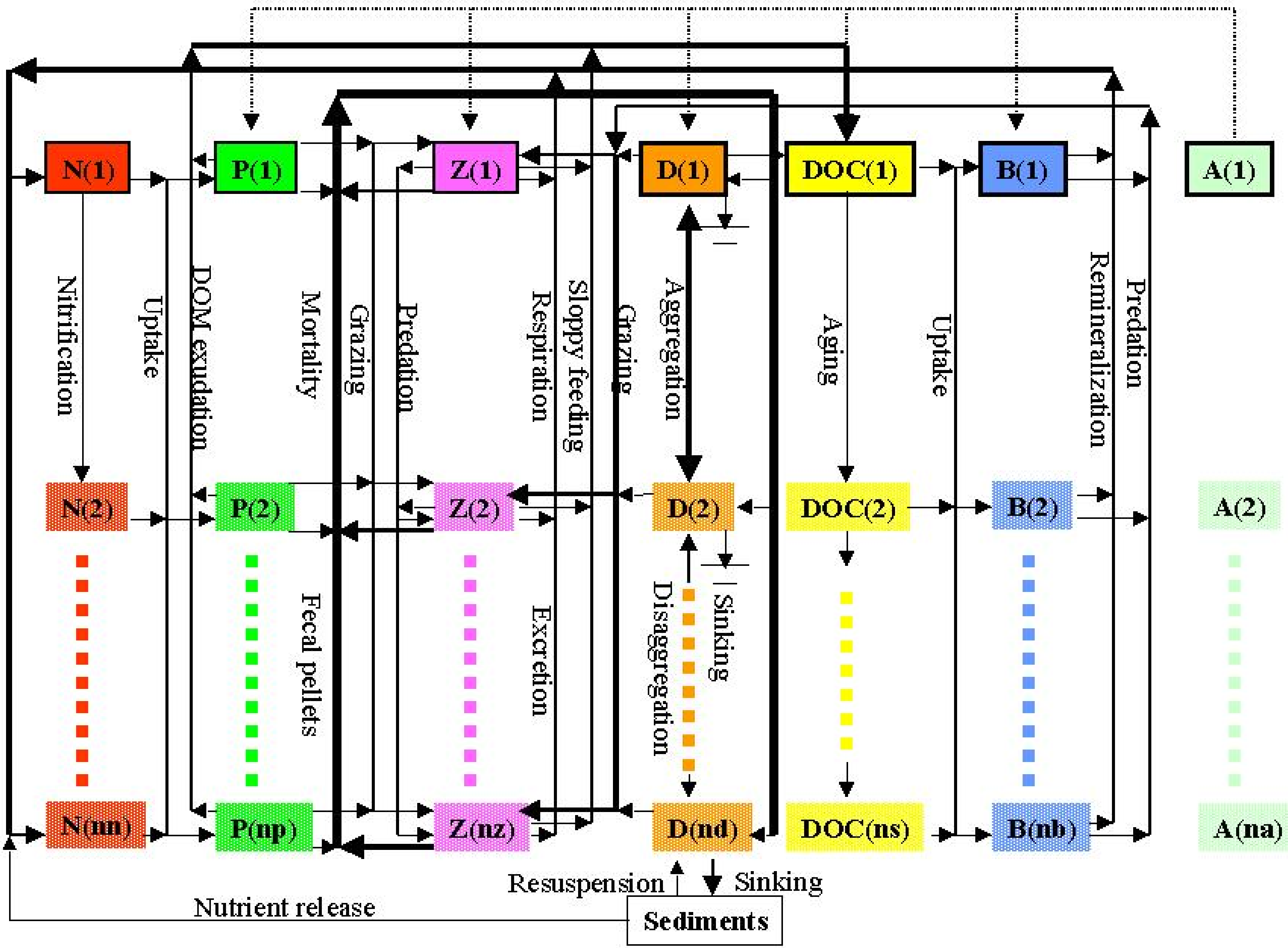
Real-time Adaptive Coupled Models



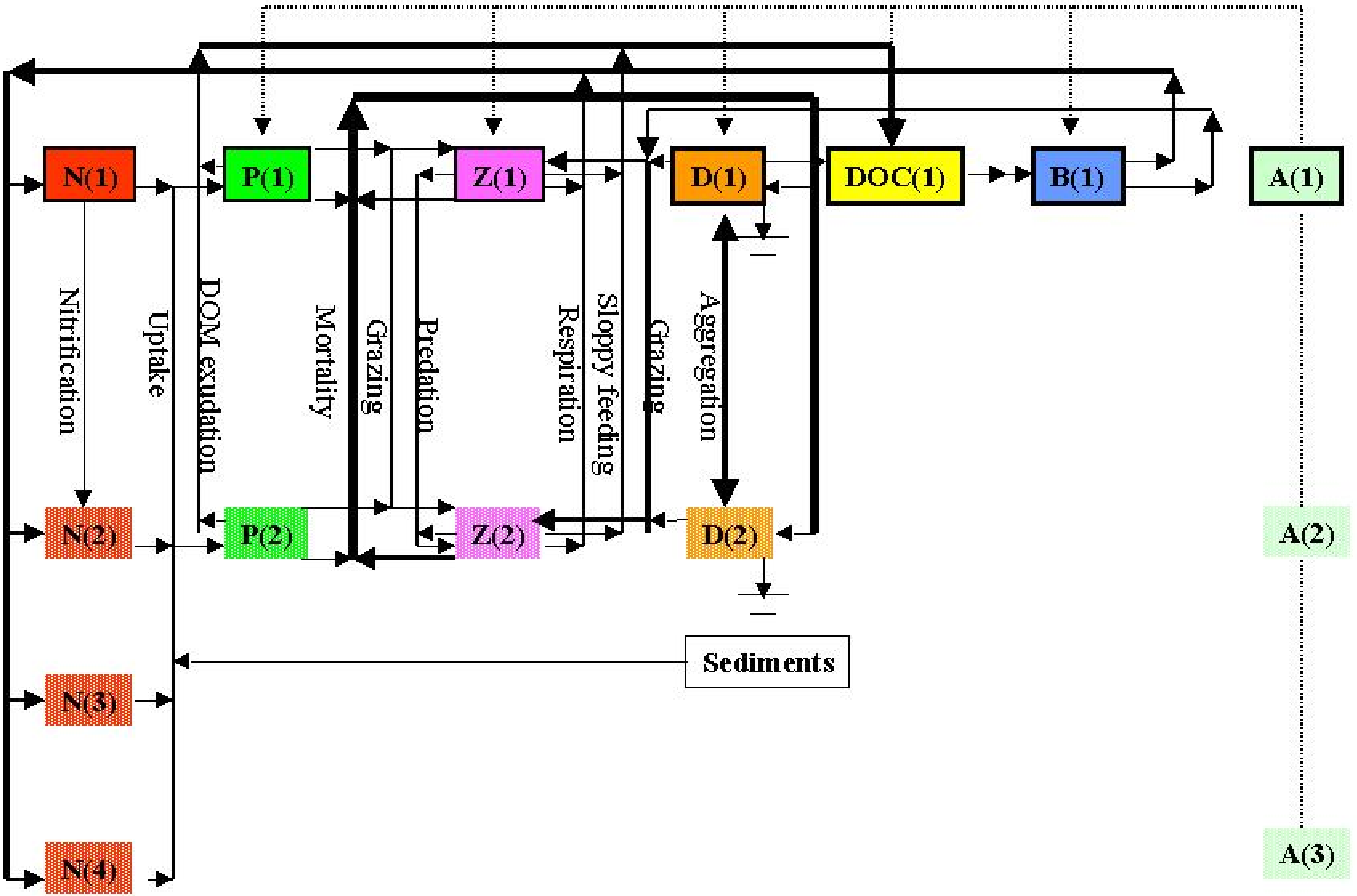


- Different Types of Adaptive Couplings:
  - Adaptive physical model drives multiple biological models (biology hypothesis testing)
  - Adaptive physical model and adaptive biological model proceed in parallel, with some independent adaptation
- Numerical Implementation
  - For performance and scientific reasons, both modes are being implemented using message passing for parallel execution
  - Mixed language programming (using C function pointers and wrappers for functional choices)

Generalized Adaptable Biological Model



A Priori Biological Model



Example: Use P data to select parameterizations of Z grazing

Table 1. Parameterization of grazing on multiple types of prey with passive selection (g_{max}: maximum grazing rate; K: Half-saturation constant (but saturation constant in Eq. 1); P_0 threshold below which grazing is zero; p_i: preference coefficient; ? , *a*, ? : constant).

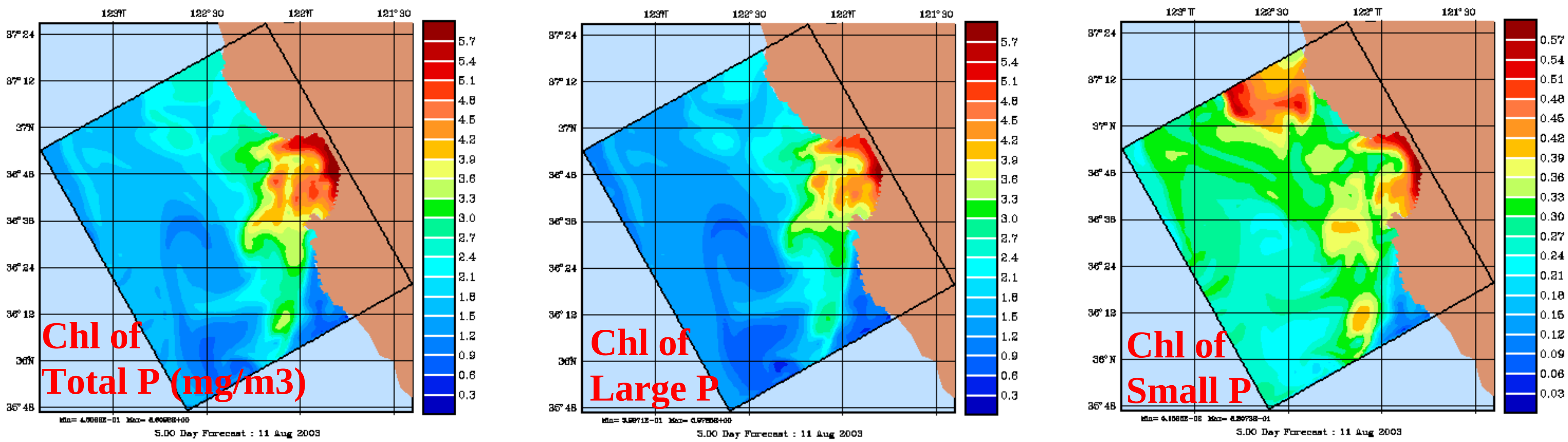
Function	References
(1) Rectilinear $g_i = \begin{cases} g_{\max}\dfrac{p_i P_i}{K}, & for \quad R \le K \\ g_{\max}, & for \quad R > K \end{cases}, \quad R = \sum_{i=1}^{n} p_i P_i$	Armstrong, 1994
(2) Ivlev function for each prey type: $g_i = g_{\max}(1 - e^{-\alpha_i P_i})$	Leonard et al., 1999
(3) Ivlev function with interference between prey types: $g_i = g_{\max}\left(1 - e^{-\alpha R}\right)\dfrac{p_i P_i}{R}$, with $R = \sum_{i=1}^{n} p_i P_i$	Hofmann and Ambler, 1988
(4) Mechanistic disc function: $g_i = g_{\max}\dfrac{a_i N_i}{1 + \sum_{j=1}^{n} a_j \tau_j N_j}$	Murdoch and Oaten, 1975; Holt, 1983; Gismervik and Anderson, 1997; Strom and Loukos, 1998
(5) Michaelis Menten Function: $g_i = g_{\max}\dfrac{p_i P_i}{K + \sum_{j=1}^{n} p_j P_j}$	Murdoch, 1973; Real, 1977; Moloney and Field, 1991; Verity, 1991; Gismervik and Anderson, 1997; Strom and Loukos, 1998
(6) Threshold MM function: $g_i = g_{\max}\left(\dfrac{R - P_0}{K + R - P_0}\right)\dfrac{p_i P_i}{R}$, with $R = \sum_{i=1}^{n} p_i P_i$	Evans, 1988; Lancelot et al., 2000
(7) Modified MM function: $g_i = g_{\max}\dfrac{p_i P_i}{1 + \sum_{j=1}^{n} p_j P_j}$	Verity, 1991; Fasham et al. (1999) and Tian et al. (2001)

Table 2. Parameterization of grazing on multiple types of prey with active switching selection (g_{max}: maximum grazing rate; K: Half-saturation constant; P_0 threshold below which grazing is zero; p_i: preference coefficient; α, a, τ: constant).

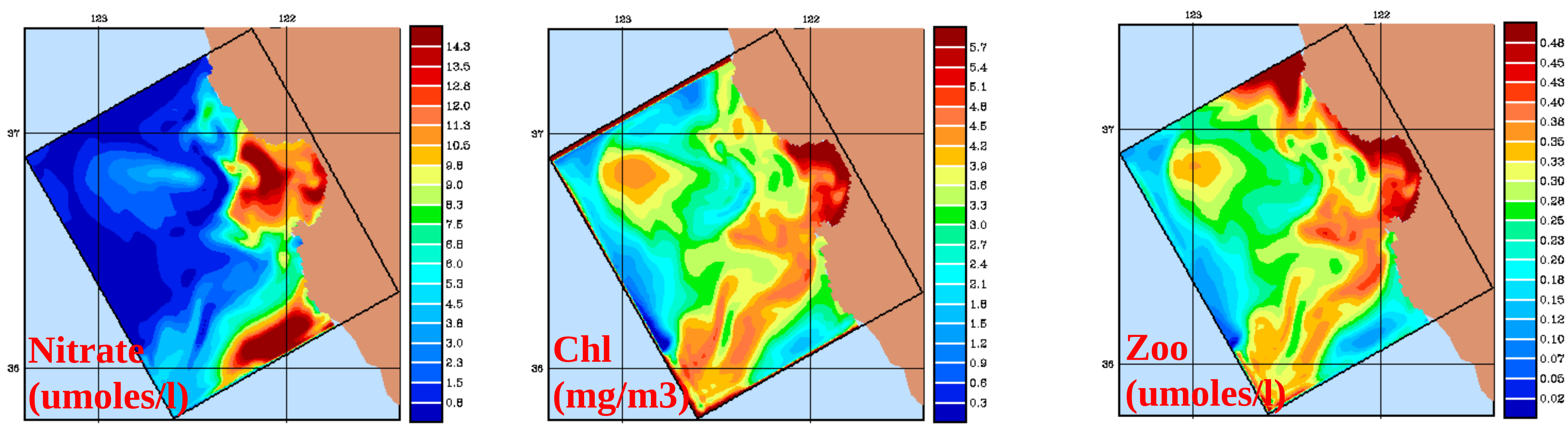
Function	References
(1) Switching MM predation: $g_i = g_{\max}\dfrac{p_i P_i^2}{K\sum_{j=1}^{n} p_j P_j + \sum_{j=1}^{n} p_j P_j^2}$	Fasham et al., 1990; Strom and Loukos, 1998; Pitchford and Brindley, 1999; Spitz et al., 2001
(2) Mechanistic disc switching predation: $g_i = g_{\max}\dfrac{b_i N_i^2}{(1 + c_i N_i)(1 + \sum_{j=1}^{n}\dfrac{b_j h_j N_j^2}{1 + c_j N_j^2})}$	Chesson, 1983
(3) Generalized switching function: $g_i = g_{\max} a_i \dfrac{(p_i P_i)^m}{\sum_{i=1}^{n}(p_i P_i)^m}$	Tansky, 1978; Teramoto, 1979; Matsuda et al., 1986
(4) Generalized switching function: $g_i = g_{\max}\dfrac{(p_i P_i)^m}{\left(\sum_{i=1}^{n}(p_i P_i)\right)^m}$	Vance, 1978
(5) Generalized switching MM function: $g_i = g_{\max}\dfrac{(p_i P_i)^m}{1 + \sum_{i=1}^{n}(p_i P_i)^m}$	Gismervik and Andersen (1997)
(6) Generalized switching MM function: $g_i = g_{\max}\dfrac{(p_i(P_i - P_{0i}))^m}{1 + \sum_{i=1}^{n}(p_i(P_i - P_{0i}))^m}$	This work

Towards automated quantitative model aggregation and simplification

A priori configuration of generalized model on Aug 11 during an upwelling event



NPZ configuration of generalized model on Aug 11 during same upwelling event



Dr. Rucheng Tian

CONCLUSIONS

- ESSE powerful nonlinear scheme for interdisciplinary estimation of oceanic state variables and error fields via multivariate physical-biogeochemical-ecosystem-acoustical data assimilation

- Entering a new era of fully interdisciplinary oceanic dynamical system science, combining models and data

- Multiple novel and challenging opportunities, for example:
  - Quantitative assimilation feedbacks, e.g. via Adaptive (Bayesian) estimation/learning
    - Adaptive modeling/system identification (optimal parameters, structures, state variables and multi-model combinations)
    - Adaptive sampling (optimal data type, quantity and time-space locations)
    - Adaptive model reductions and simplifications
  - Theory and applications of environmental ocean science
    - Dominant dynamical balances for fundamental understanding, and for weak constraints