Ocean Data Quality Control [and Assimilation]

JCSDA Summer Colloquium on Data Assimilation

July 2009

Kayo Ide
University of Maryland

Basic Ideas for Ocean Data Quality Control

- Quality control (QC) for ocean data from data assimilation perspective
  - Basic concepts are similar to atmospheric systems (theoretical, i.e., Bayesian view).
  - Details may require unique adjustment, development, and implementation, for example:
    - Ocean data assimilation is relatively new and may have different emphases from atmospheric data assimilation.
    - Scales are different in both:
      - » Dynamics
      - » Observations

Outline

- Some examples of ocean data assimilation systems
  - Global
    - ECMWF ORA-S3 [reanalysis and near real-time]
    - SODA [reanalysis]
  - Regional
    - ROMS 3D-Var [near real-time]

- Types of ocean data used in the assimilation
  - Conventional platforms
    - In-situ [mainly T, S]
    - Remote sensing, satellite in particular ⟶ See Bob Miller's lecture on July 15
  - New types of platforms
    - HF (high frequency) radar [surface (u, v)]
    - Lagrangian data [trajectories]

- Quality control
  - Basic concepts ⟶ See Andrew Lorenc's lecture on July 9
  - Examples

ECMWF ORA-S3 (Ocean Re-Analysis System 3)

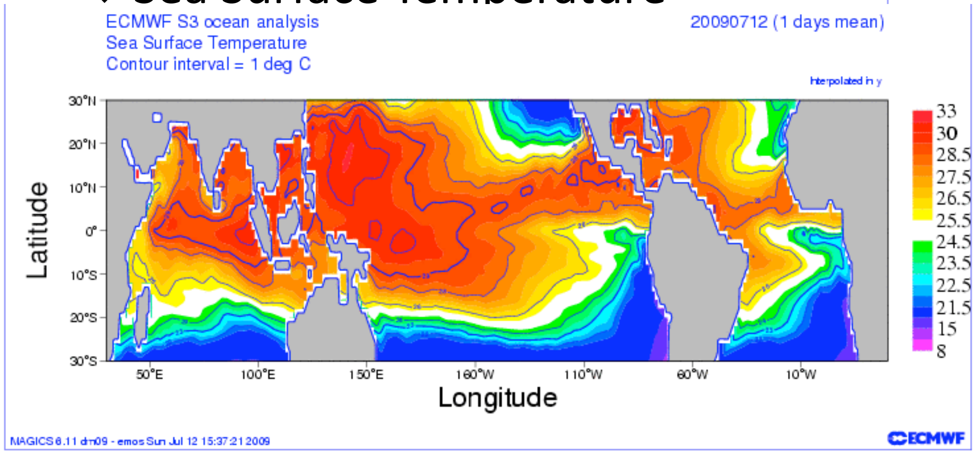
- Global ocean near real-time analysis & reanalysis since 2006
  - Daily starting from January 1, 1959 & continuously maintained up to 11 days behind real time
  - Main purposes: to provide
    - Initial conditions for seasonal/monthly forecasts
    - historical representation of ocean for climate studies
    - Ensemble of ocean analysis (5 total) for uncertainties
  - Featuring
    - Online bias-correction algorithm
    - Assimilation of salinity data
    - Assimilation of altimeter-derived sea level anomalies
  - Designed to reduce spurious climate variability due to observing system change while taking advantages of the new observations
- Previous systems
  - System 1 (ORA-S1) starting 1997. Initial condition for first ECMWF operational seasonal forecasting system
  - System 2 (ORA-S2) introduced in 200.

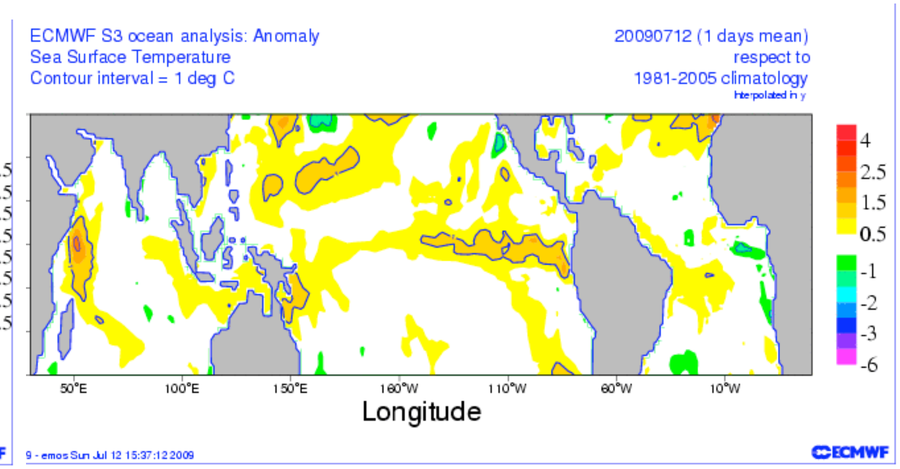
http://www.ecmwf.int/products/forecasts/d/charts/ocean/

Balmaseda, Vidard, Anderson (2008)

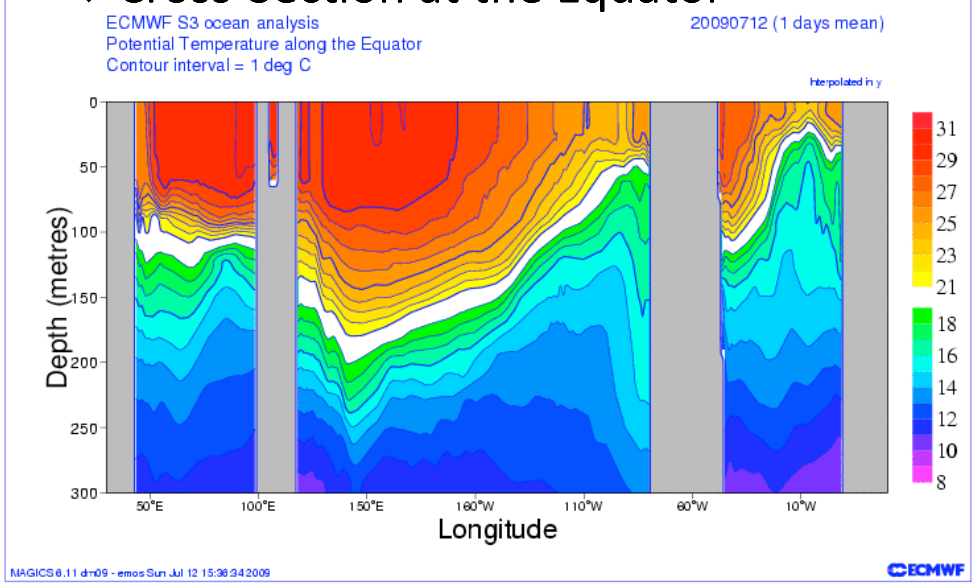
ORA-S3 Real-Time Products: 3D evolution of ocean state

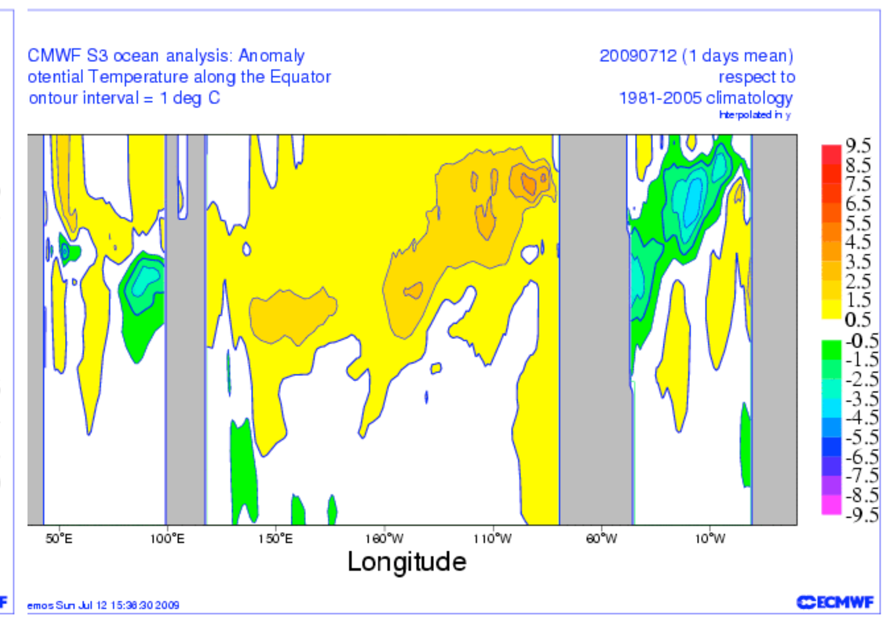
- Sea Surface Temperature





- Cross-Section at the Equator



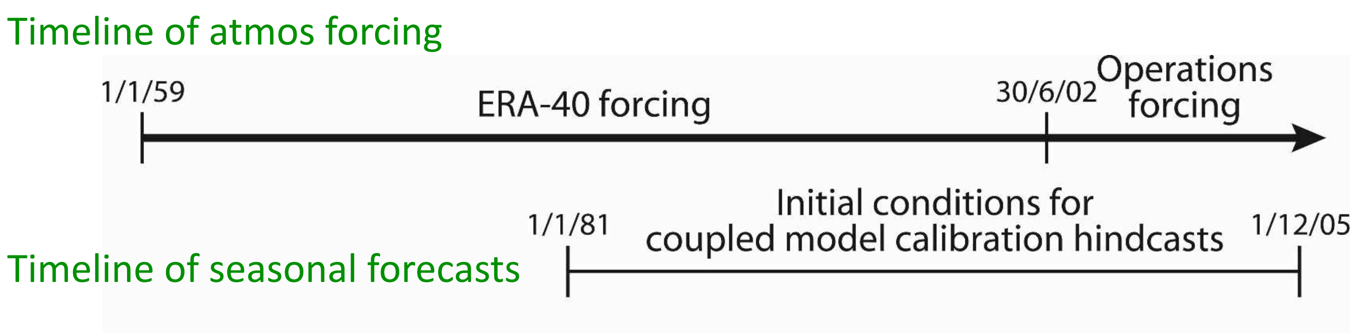


http://www.ecmwf.int/products/forecasts/d/charts/ocean/real_time/xymaps/ & xzmaps

ORA-S3 Model Overview

- Model: ECWMF HOPE ocean model
  - Horizontal resolution 1°x1° with equatorial refinement (0.3° in meridional)
  - Vertical 29 levels, thickness of 10m in upper oceans

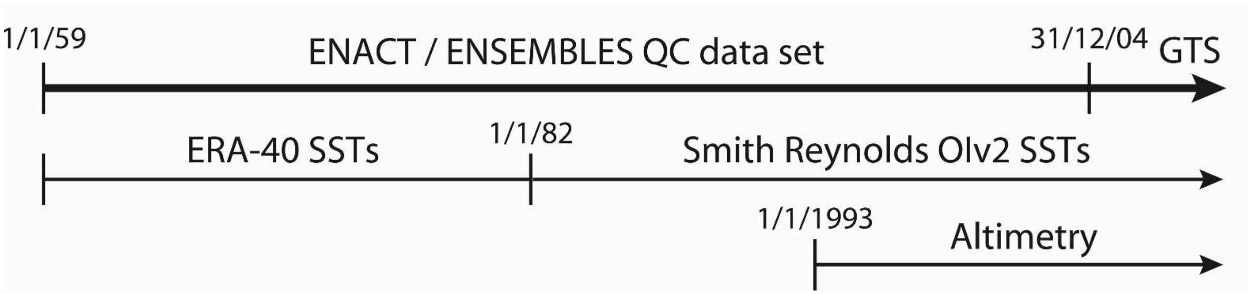
- Forcing
  - ERA-40 from 1959 to June 2002; Operation NWP analysis afterwards



ORA-S3 Data Assimilation System Overview

- Method
  - 3D Optimal Interpolation (OI), simultaneously down to 2000m
- Data types assimilated: all observations in upper 2000m
  - Subsurface temperature (in-situ)
  - Salinity (in-situ)
  - Altimeter-derived sea level anomalies (remote sensing)
- Data sets assimilated:
  - Prepared for ENACT (Enhanced Ocean Data Assimilation and Climate Prediction) & ENSEMBLES (Ensemble-based predictions of climate changes and their impact) until 2004, with quality control (QC: Ingleby and Huddleston, 2007)
  - From GTS (ENACT/Global Telecommunication System) thereafter

Timeline of Observations & Surface Data



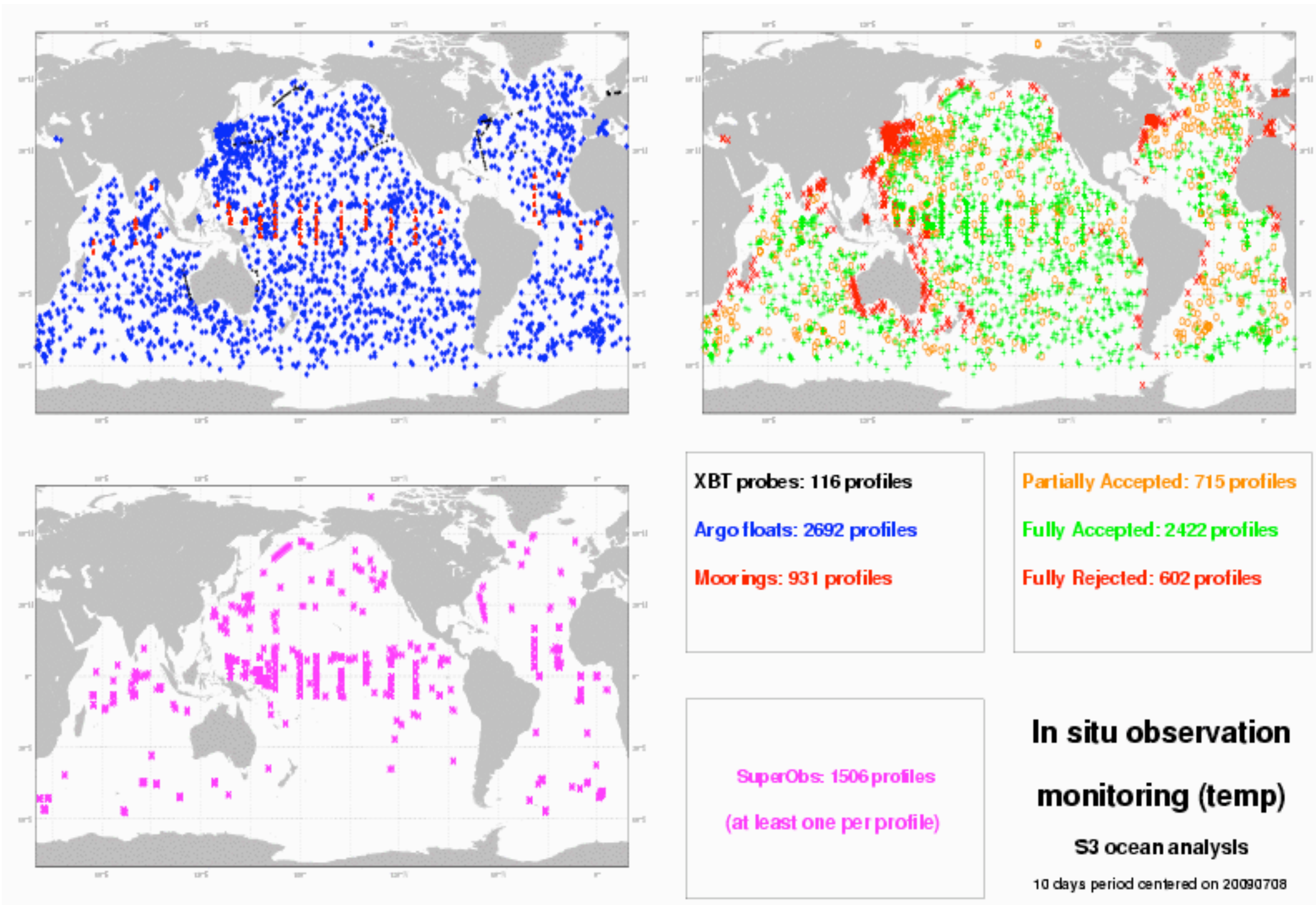
ORA-S3 Considerations for Seasonal Forecasting System

- Data assimilation system needs to take into account of
  - Interannual – decadal variability in the ocean initial conditions are properly represented: strong relaxation to climatology is not desirable
  - Spurious variability due to the change of the observing system network should be reduced as much as possible
  - Large initialization shocks in the coupled model, which may damage the forecast skill, should be avoided.

- In comparison to previous systems
  - the weight to observations has been reduced and
  - the relaxation to climatology is significantly weaker.

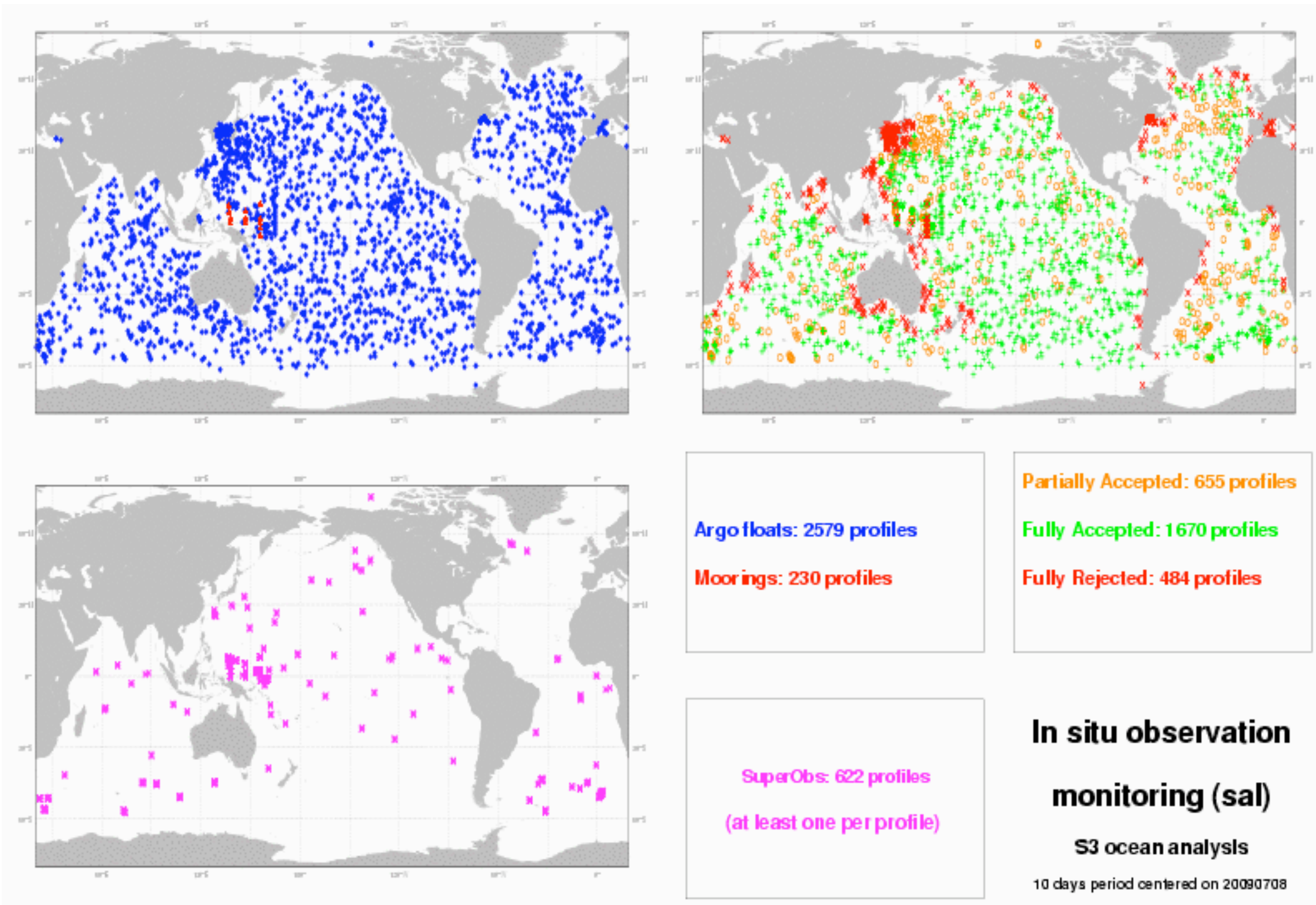
- For ensemble forecasting, five simultaneous analyses are performed.

Data Coverage (Real Time)



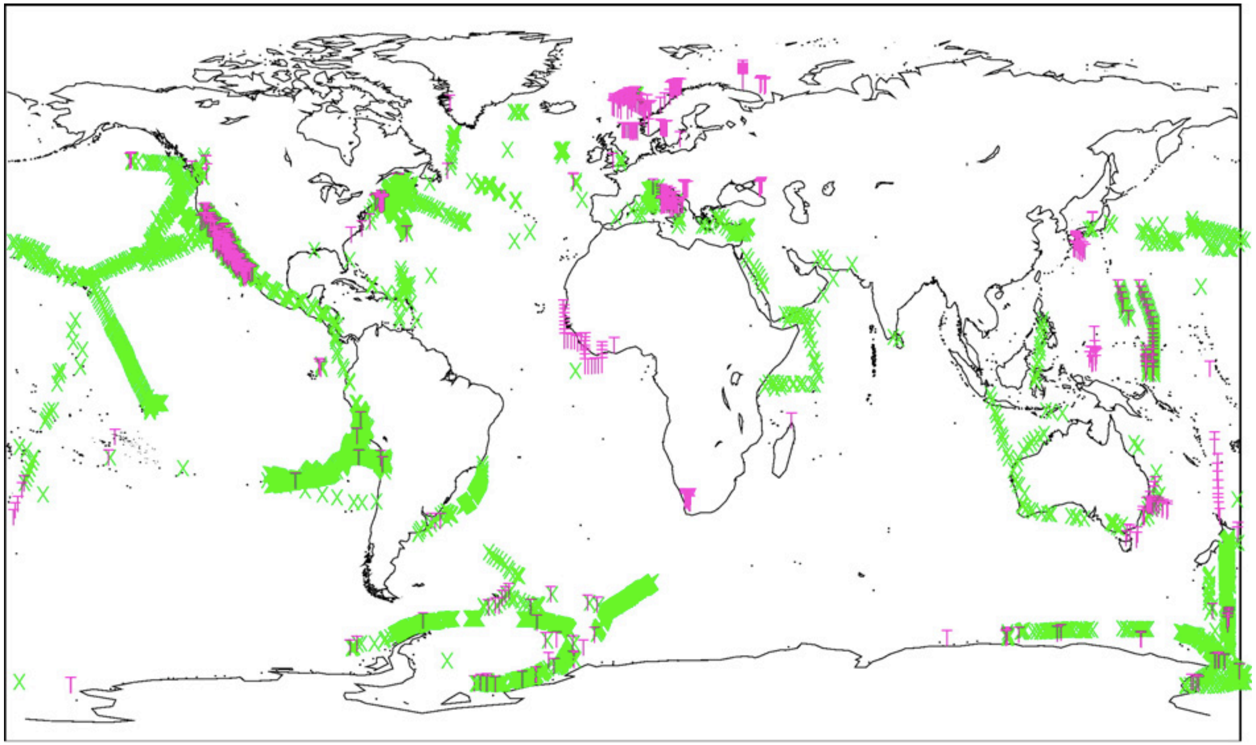
http://www.ecmwf.int/products/forecasts/d/charts/ocean/real_time/obsmap!20090708!Temperature!/

Data Coverage (Real Time)



http://www.ecmwf.int/products/forecasts/d/charts/ocean/real_time/obsmap!20090708!Salinity!/

Data Coverage. January 1958



Reports available from World Ocean Database 2001 (WOD01)

- Green X: Bathy data (3508 reports),
- Purple T: Ocean station data (855 reports).

Actual position is at the bottom left corner of each letter.

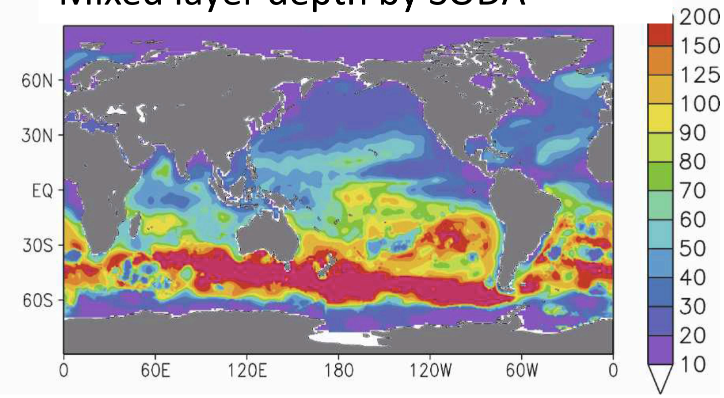
Ingleby, Huddleston(2007)

SODA (Simple Ocean Data Assimilation)

- Global ocean reanalysis system (SODA 1.4.2)
  - Daily starting from 1958. for 44 yrs, span over ECMWF ERA-40 Atmos Reanalysis
  - Used to provide historical representation of ocean for climate studies
  - Featuring assimilation of data:
    - Historical archives of temperature data by XBT, CTD, Mooring, Argo
    - Salinity data by CTD, Mooring, Argo
    - Sea surface temperature by remote-sensing
  - Designed to reduce spurious climate variability due to observing system change while taking advantages of the new observations
- Previous system
  - Previous SODA for global upper ocean (1950-1959)

Carton and Giese (2008)

Mixed layer depth by SODA

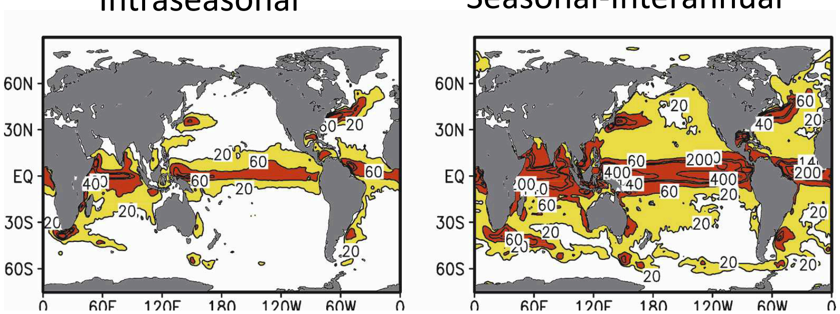


SODA Model Overview & Reanalysis Products

- Model: Princeton MOM
  - Horizontal 0.25°x0.24° on average
  - Vertical 40 levels, thickness of 10m in upper oceans
- Forcing:
  - ERA-40 + Global Precipitation Climatology Project
  - Relaxation to World Ocean Atlas 2001 (WOD01) climatological sea surface salinity

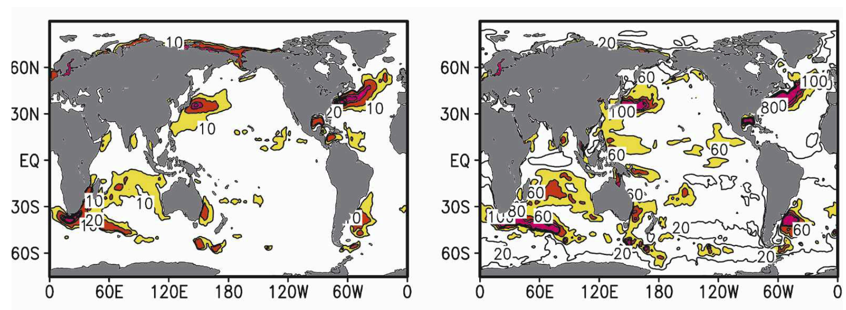
Eddy kinetic energy

Intraseasonal Seasonal-Interannual



Sea level variance

Intraseasonal Seasonal-Interannual

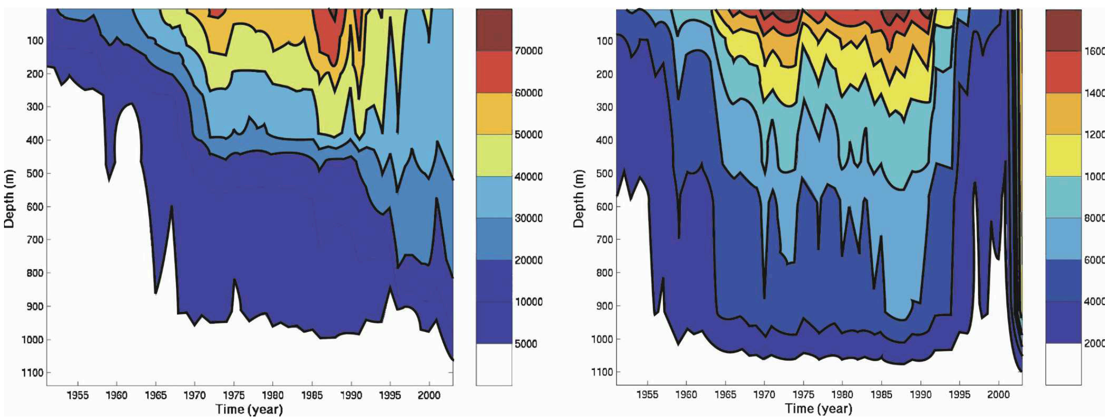


Carton and Giese (2008)

SODA Data Assimilation System Overview

- Data assimilation method:
  - Simple OI-like formula
  - Assimilation performed every 10 days
  - Observations using within the window ±45 days (multiple use of observations) with lesser weights for observations away from the assimilation time
  - Incremental correction for 5 days every time step (Bloom et a; 1996)
- Data and Data QC on in-situ data, similar to ORA-S3
  - WOD01 (upper ocean only up to 1000m): location check, local stability check
  - Buddy check among obs
  - Observation-minus-forecast check

Carton and Giese (2008)



Observations per year for T (left) and S (right) vs depth

ROMS (Regional Ocean Modeling System) 3D-Var

- Regional ocean near real-time analysis and forecast system
  - 4 times daily analysis and forecast (like atmospheric NWP) starting from 2003
  - Regional along the US Western Coast Ocean, currently
    - Southern California Bight
    - Monterey Bay
    - Prince Williams Sound
  - Purpose:
    - Coastal ocean analysis and forecasts
  - Featuring

    Assimilation of
    - Data sampled by movable platforms (possibility for adaptive sampling)
    - Remote sensing observations by satellite and HF radars

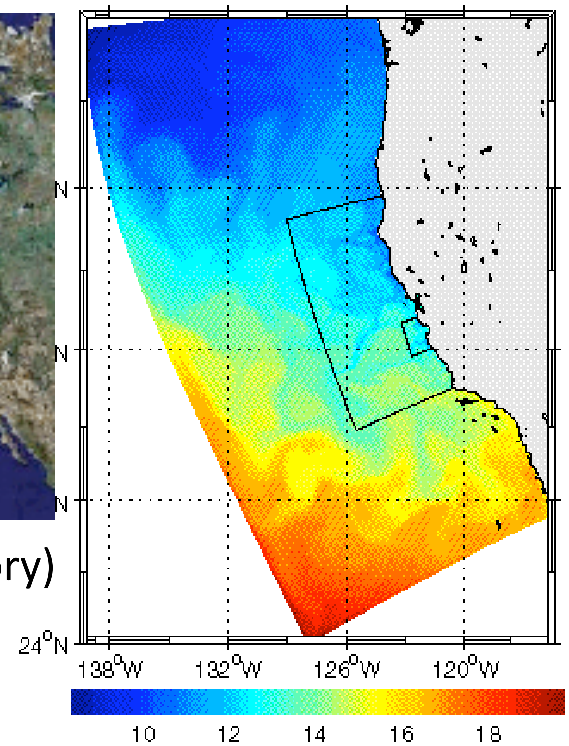
    Evaluation using
    - Independent data sets (mooring data, etc)

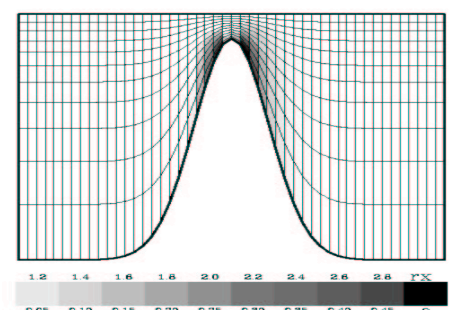
http://ourocean.jpl.nasa.gov/ourocean.html

Li, Chao, McWilliams, Ide (2008a,b)

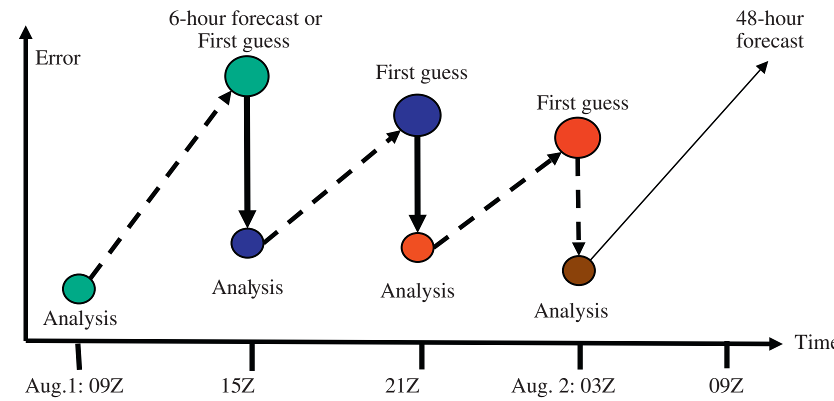
ROMS 3D-Var Model Overview

- Regional Ocean Modeling System (ROMS)
  - One-way nested configuration
    - Pacific basin for largest domain
    - Nested coastal configuration
      - » 15km-5km-1.5km-0.5km
      - » Vertical levels >24
  - Relocatable, in the future
  - Forcing by COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System by Naval Research Laboratory)





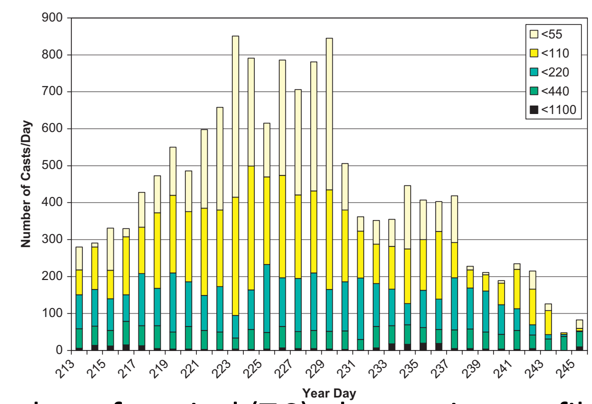
ROMS Data Assimilation System Overview



- ◆ Method
  - – 3D-Var
  - – Assimilation window
    - Analysis: every 6hr
    - 48hr forecast: daily
  - – Weak constraints for
    - Hydrostatic balance
    - Geostrophic balance
  - – Use of
    - Variable transformation to streamfunction-velocity potential, instead of horizontal velocity
    - Kronecker product in the background covariance

    for representing inhomogeneity

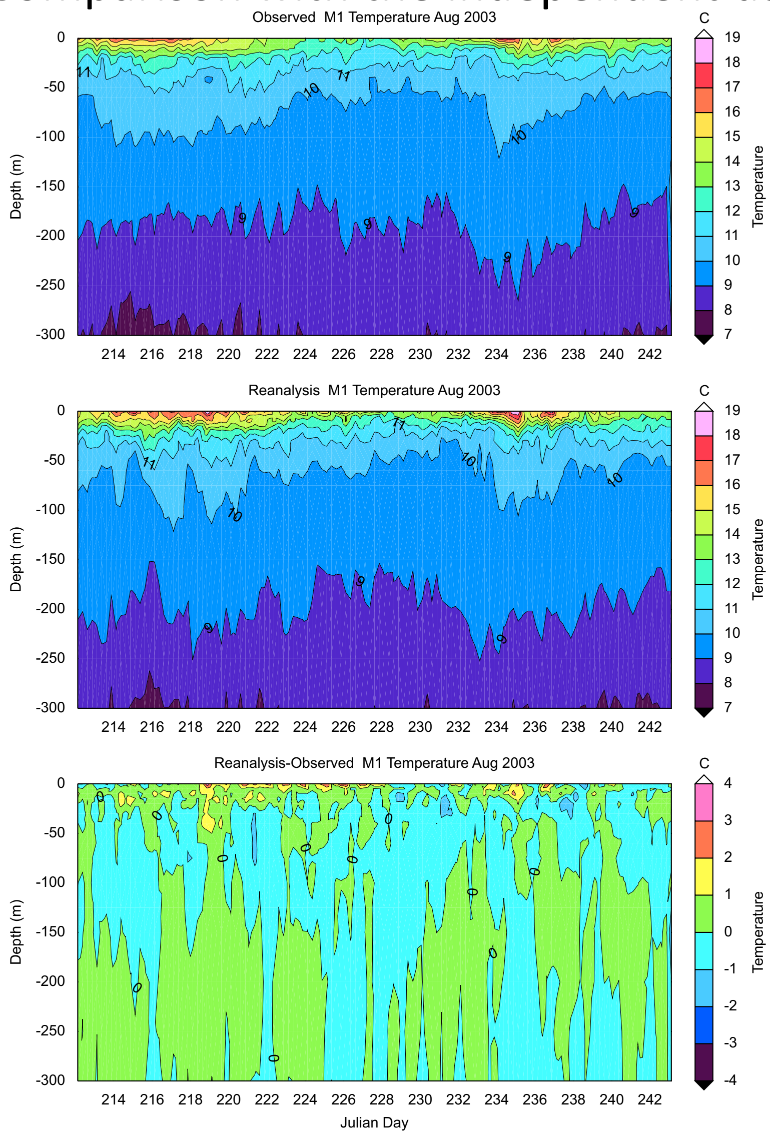
    and anisotropy suitable for the coastal ocean
  - – Assimilation of
    - T, S, SSH (conventional data type)
    - Surface velocity (new type of data)

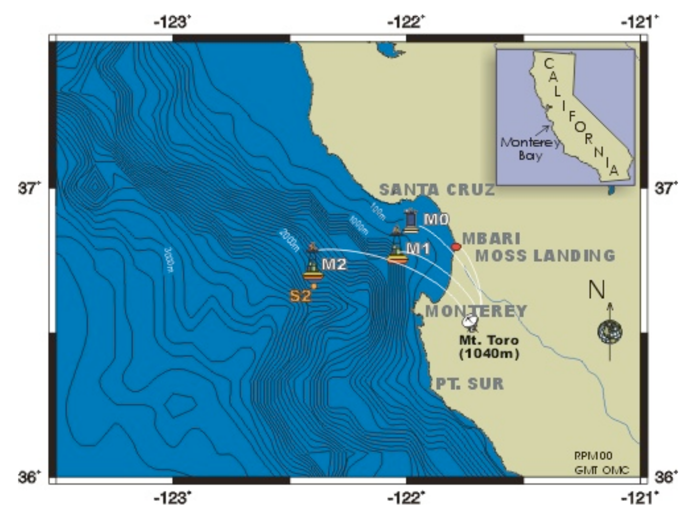


Number of vertical (T,S) observation profiles

ROMS Products Examples

◆ Comparison with the independent data by mooring (vertical profile at M1)



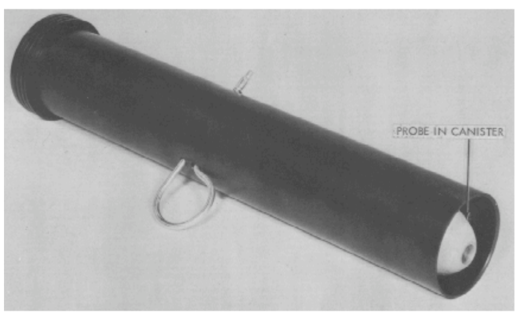


Ocean Observations

- In this presentation, we focus on
  - In situ
  - HF radar (remote-sensing)
- See Bob Miller's lecture for satellite (remote-sensing) observations

In-Situ Instruments for Vertical Profile. XBT

- BathyThermographs: Temperature recorders. $\sigma^{o}_{T} \sim 0.1\text{-}0.2°C$
  - Mechanical (MBTs):
    - Lowered and then winched from the ship down to ~300m.
  - Expendable (XBTs):
    - Dropped from a ship;
    - Designed to fall at a constant rate..
    - Many goes to 460/760m, some goes to 1800m



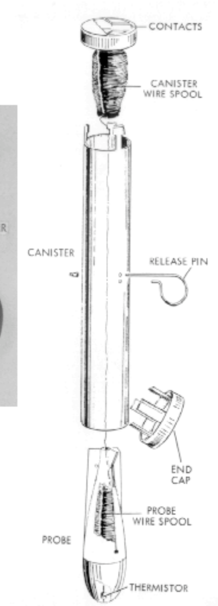
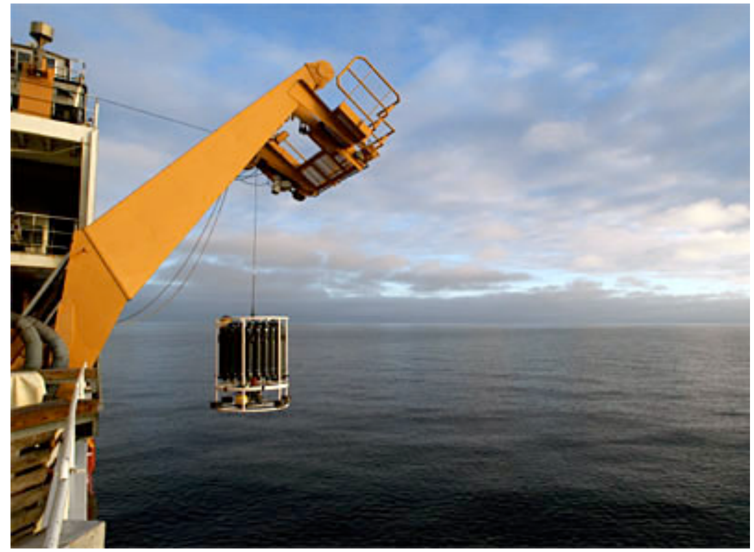
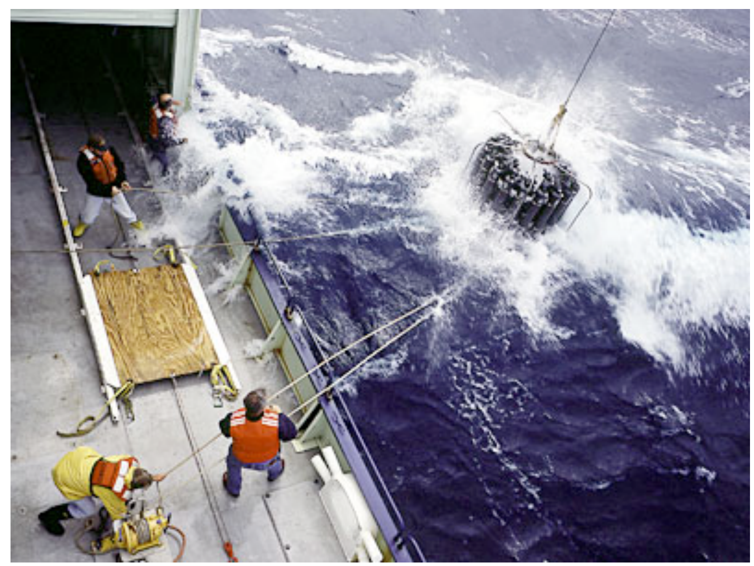


Fig. 1: XBT diagrams: Bathythermograph (probe) and exploded view.

In-Situ Instruments for Vertical Profile. CTD

- Conductivity, Temperature, & Depth (CTD): High-quality T/S profile of 150 levels (σ^o_T, σ^o_S)~(0.002°C, 0.005psu)



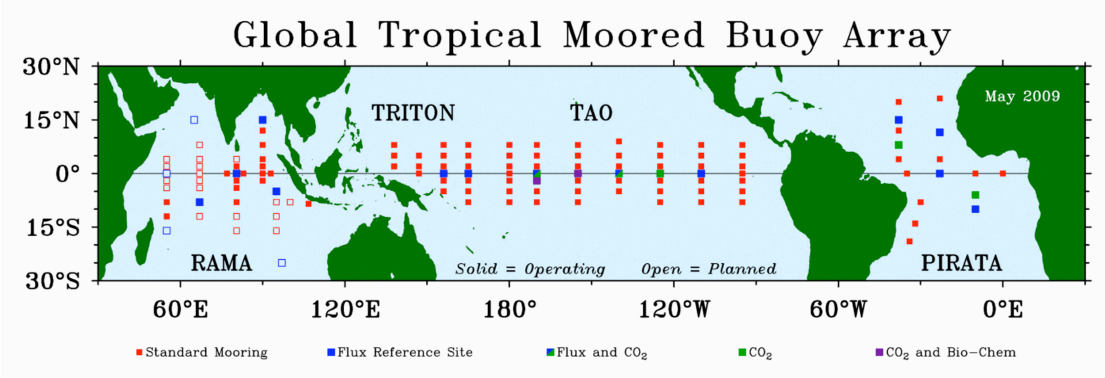


Ship deployed CTD
Woods Hole Oceanographic Institute

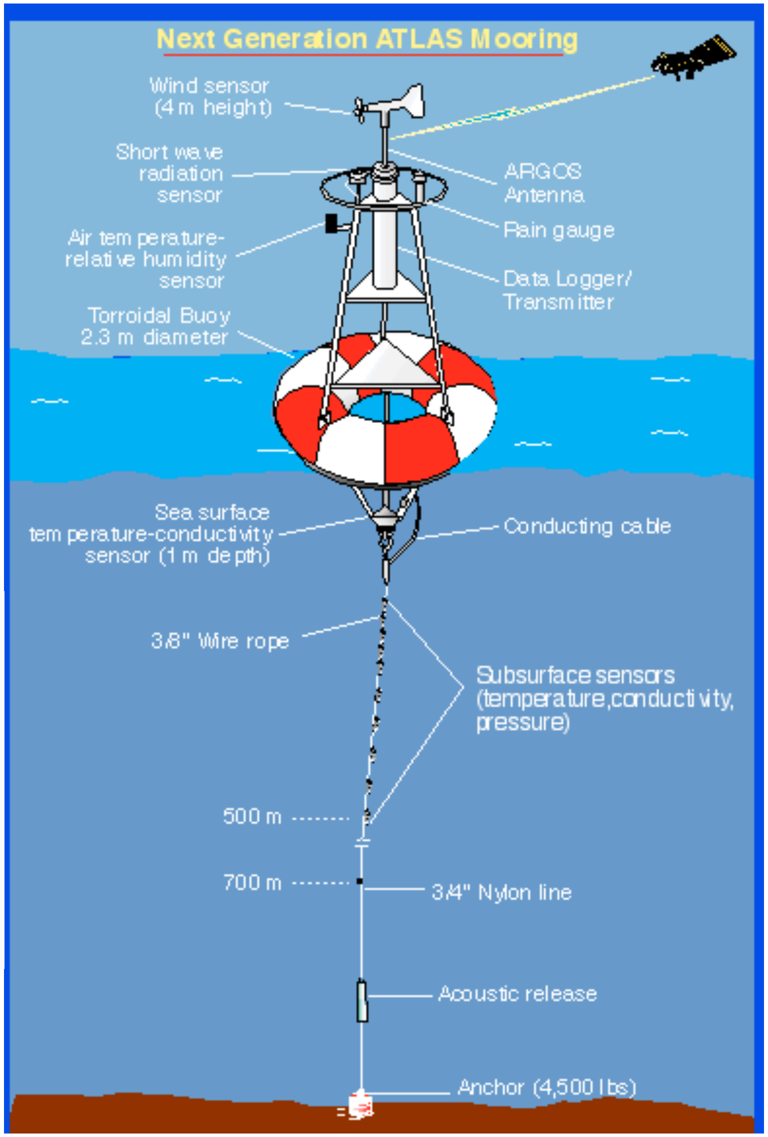
In-Situ Instruments for Vertical Profile. Mooring

- Mooring: CTD, Surface observations
  - TAO (Tropical Atmosphere Ocean):

    For improved detection, understanding and prediction of El Niño and La Niña.
  - TRITON
  - PRITA (small no of deep water buoys)

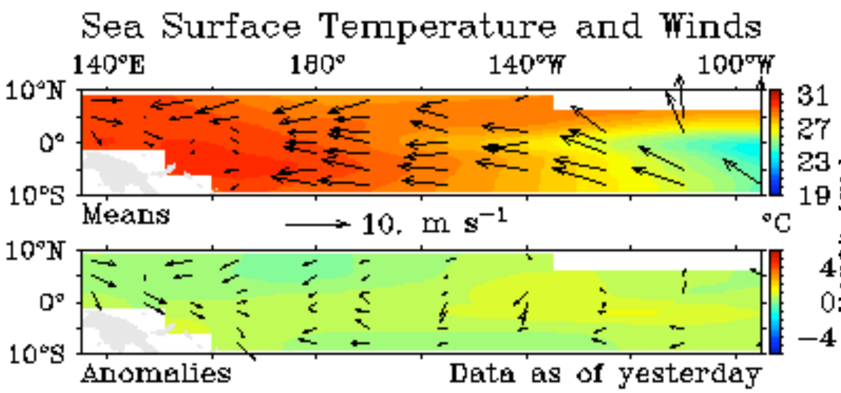


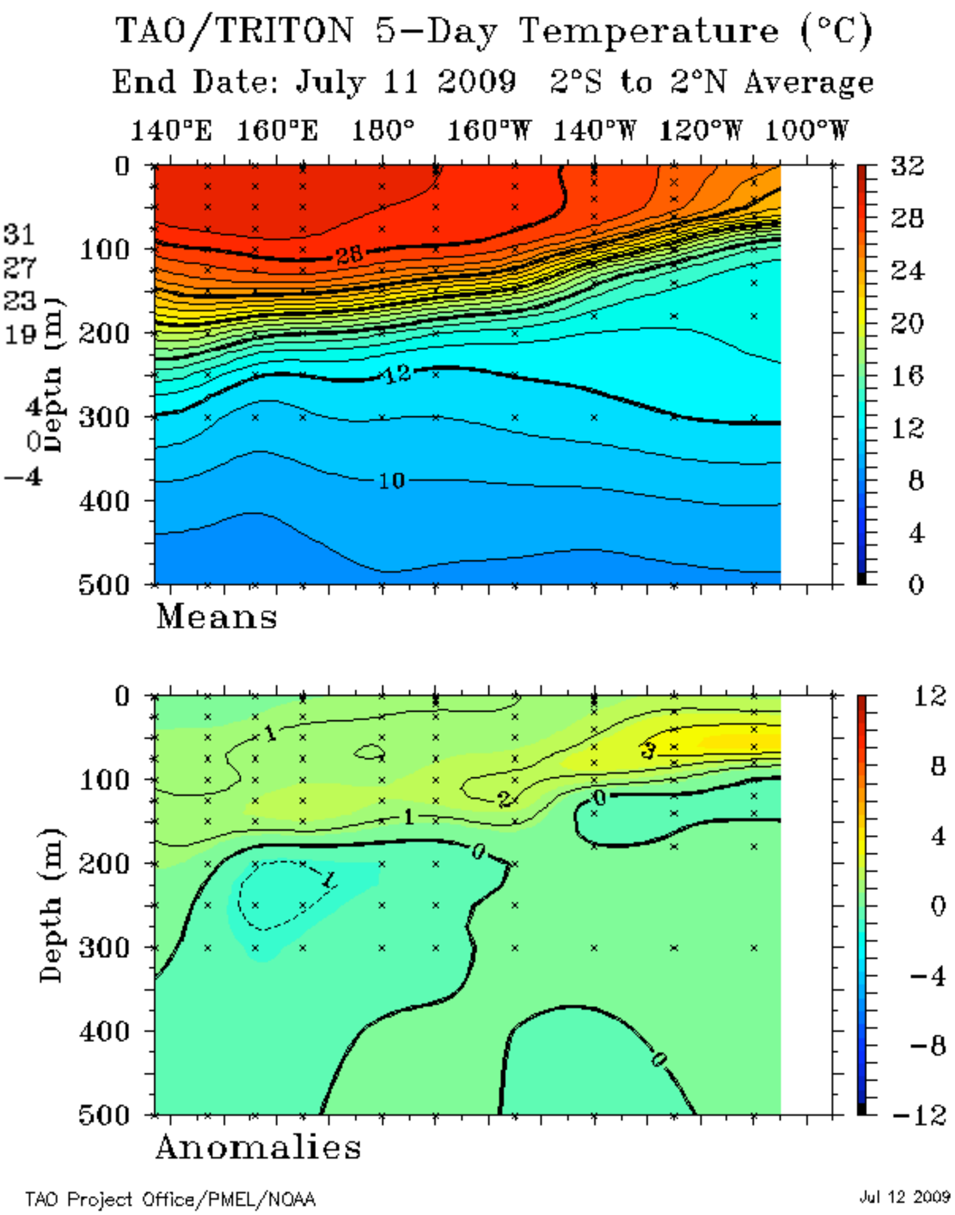
http://www.pmel.noaa.gov/tao/index.shtml



In-Situ Data

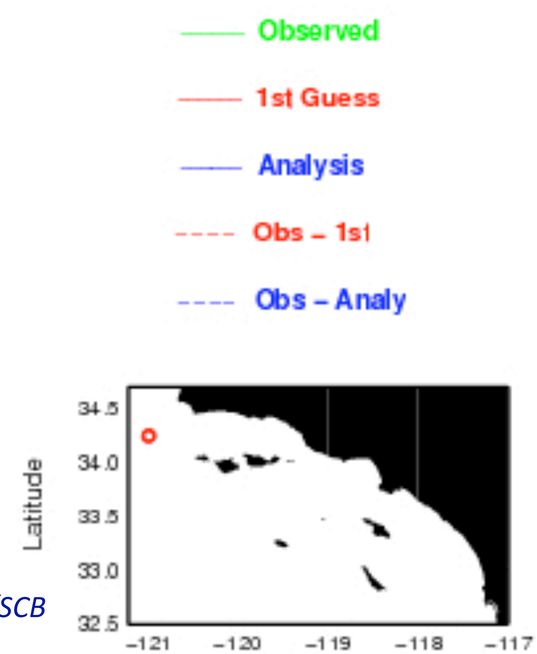
- ◆ Mooring Data



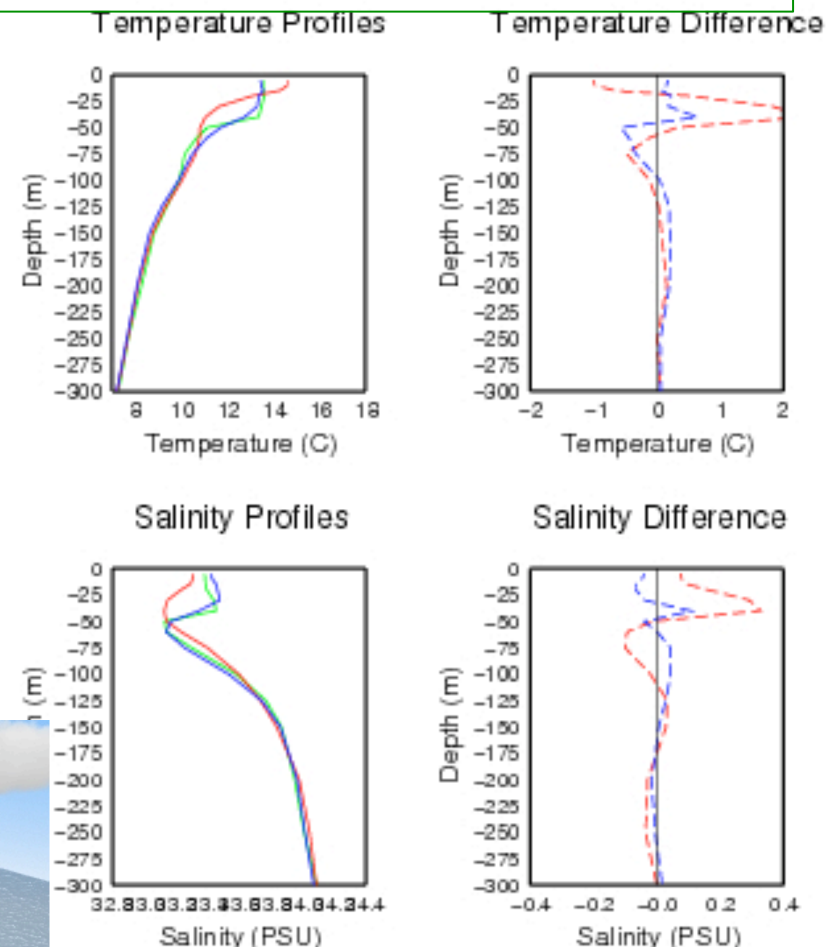


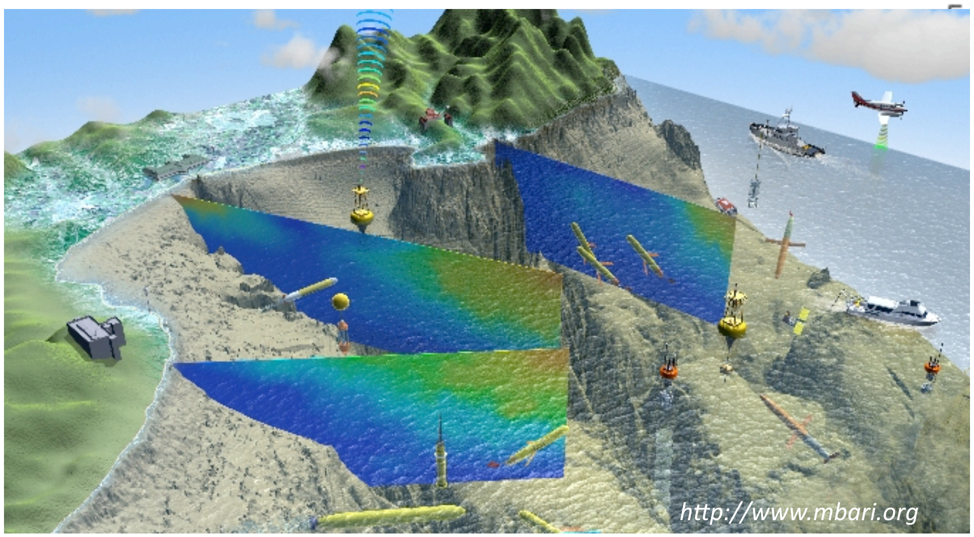
– Mooring data are temporally & vertically ‘continuous’.

In Situ Instruments (Movable): Glider



http://ourocean.jpl.nasa.gov/SCB



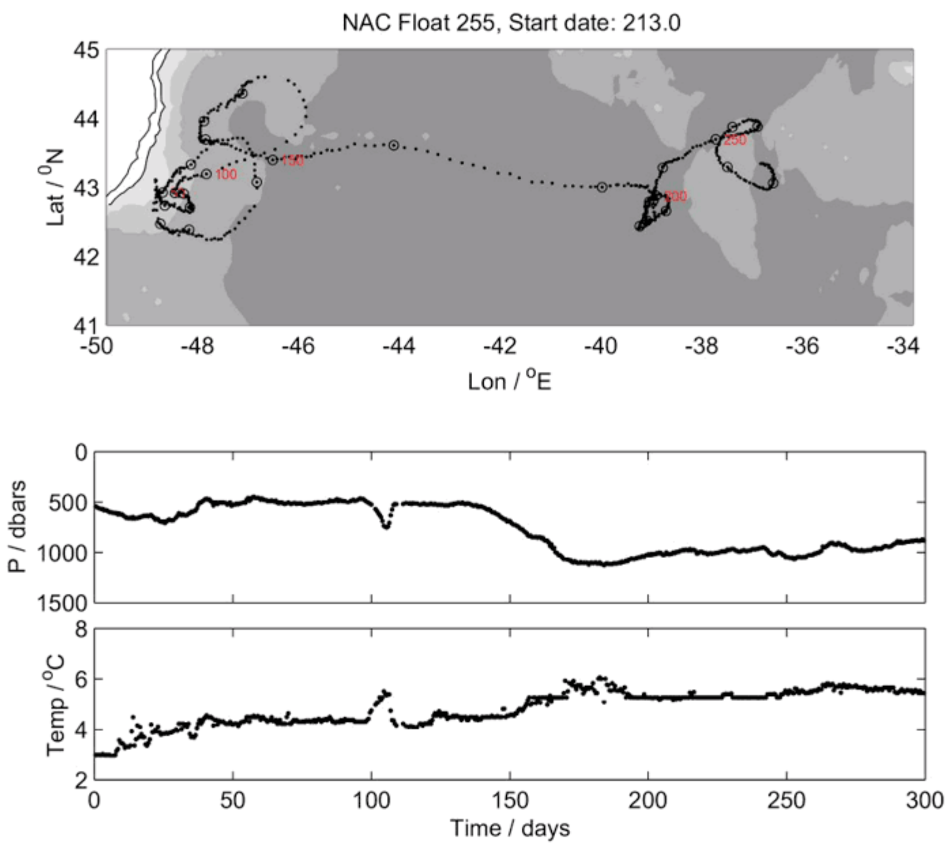


- Glider
  - 2D position, p (depth)
  - T, S

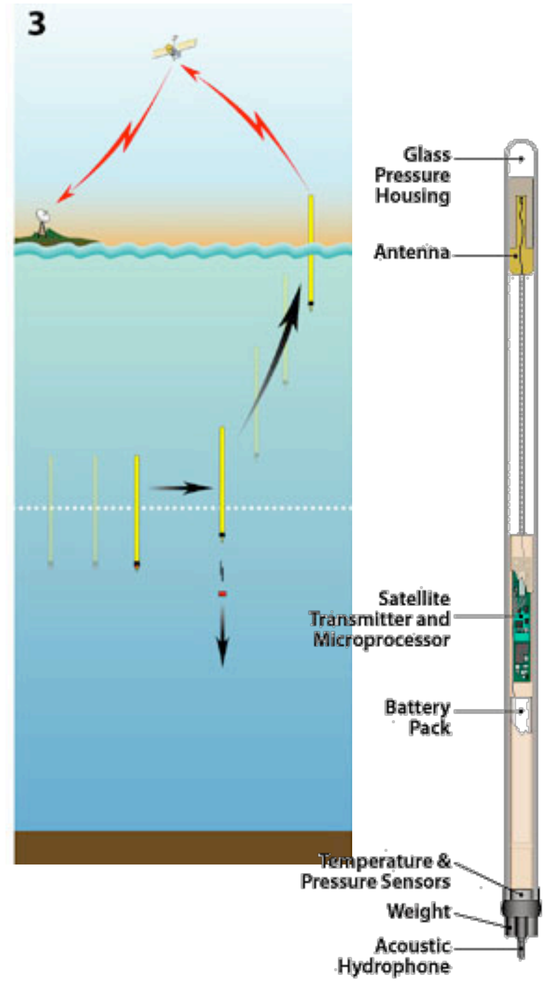
In-Situ Instruments (Movable/Lagrangian). Floats

- Argo Floats. Observation on the isopyncnal surface
  - (T, S) by CTD
  - (u,v) derived from position

  along $(\mathbf{x}^{(2D)})(t_k)$, $p(\mathbf{x}^{(2D)})(t_k)$) in upper 2000m

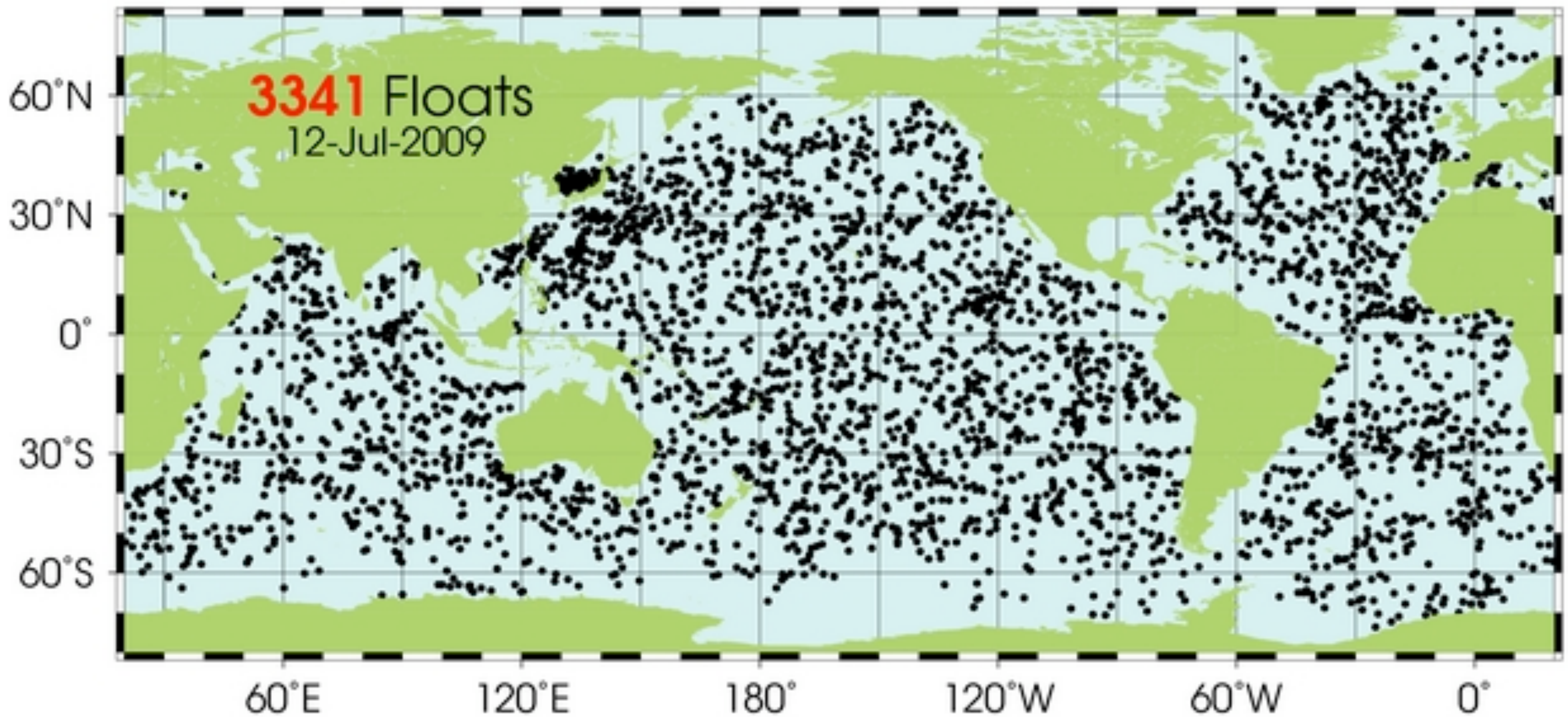


http://www.dosits.org/gallery/tech/ooct/rafos1.htm



http://www.whoi.edu/instruments/

Global Ocean Observing System by ARGO Floats



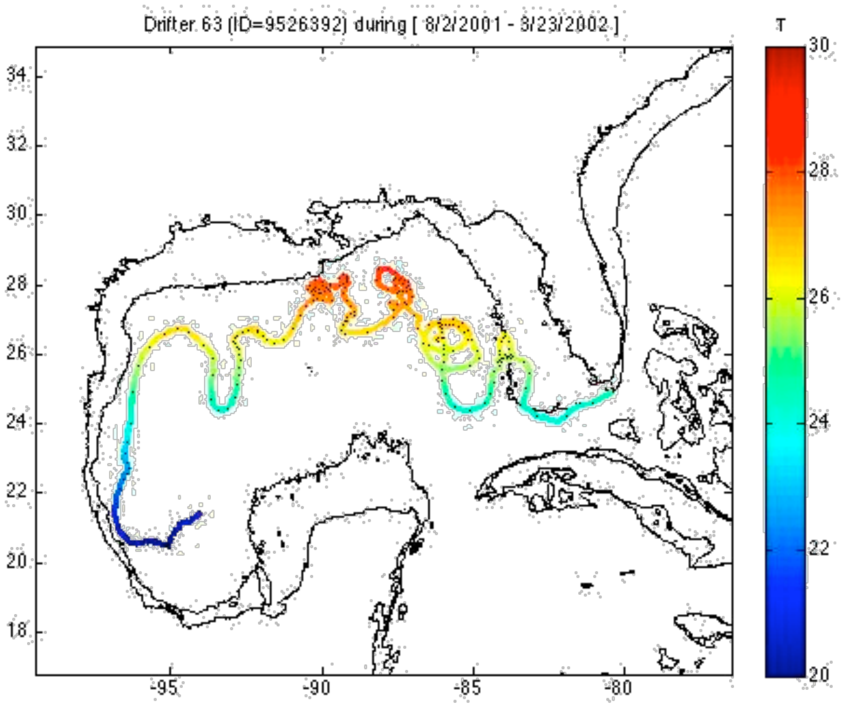
- Argo are used as the platform for continuous observation
  - Eulerian observations of *T, S,* and velocity
- By November 2007, Global observation network by drifters is 100% complete
  - ~3000 profiling at the 5°x5° resolution
  - 800 floats per year to maintain the level

http://www-argo.ucsd.edu

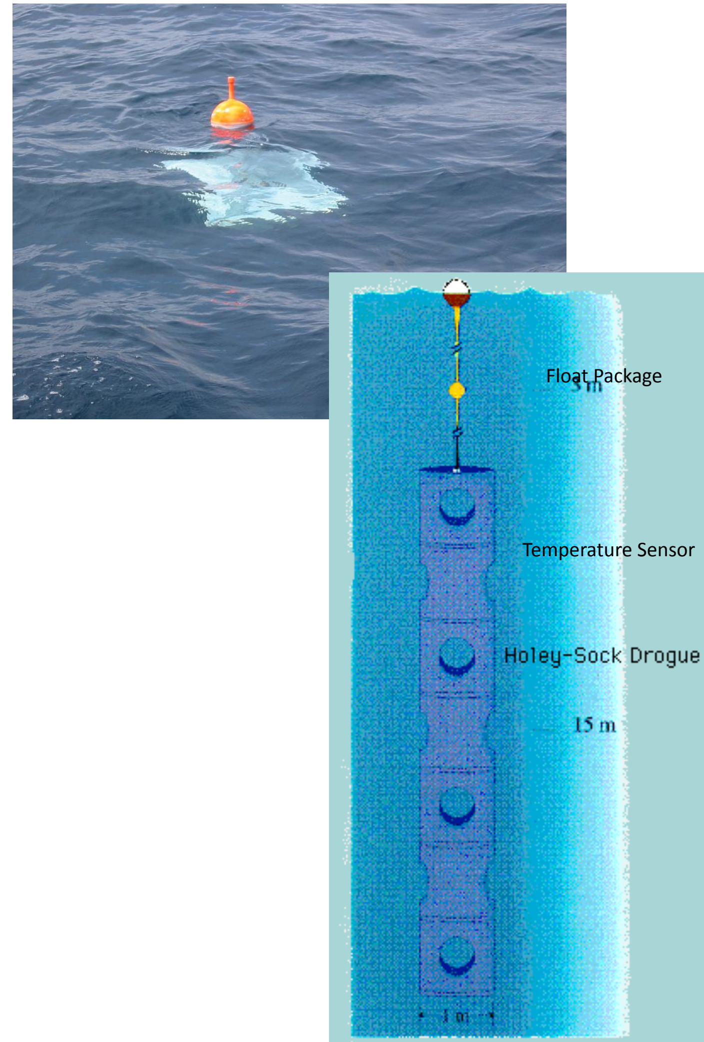
In-Situ Instruments (Movable/Lagrangian). Drifters

- Observations at sea surface
  - T : Temperature

  along $(\mathbf{x}^{(2D)}(t_k))$ at sea surface

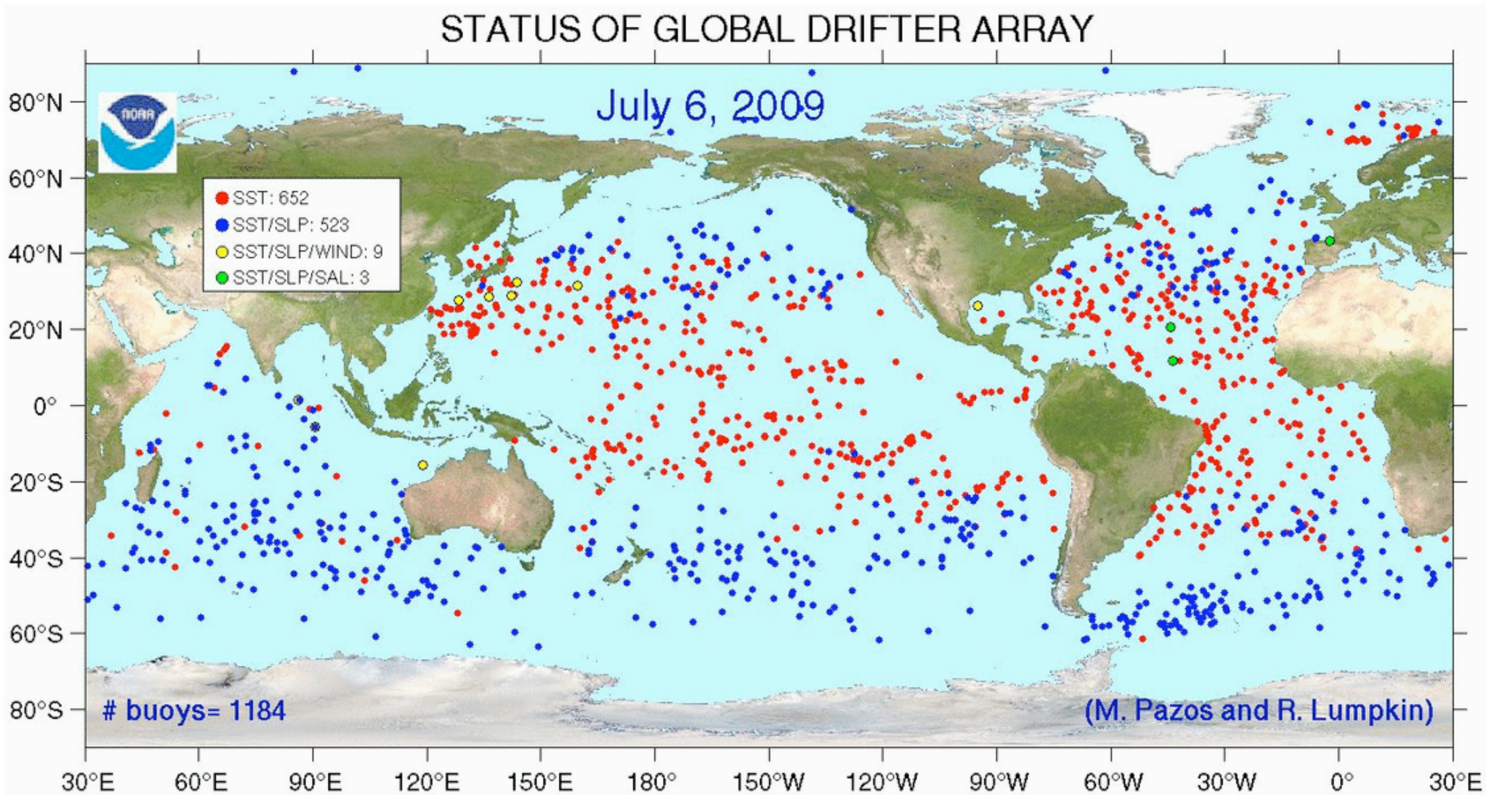


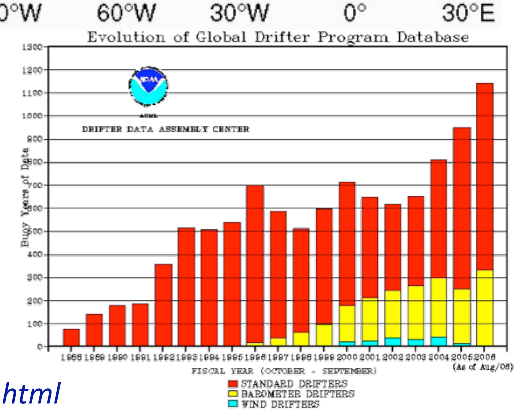
Data available from
http://www.aoml.noaa.gov/phod/dac/dacdata.html



http://www.drifters.doe.gov/design.html

Global Ocean Observing System by Drifters





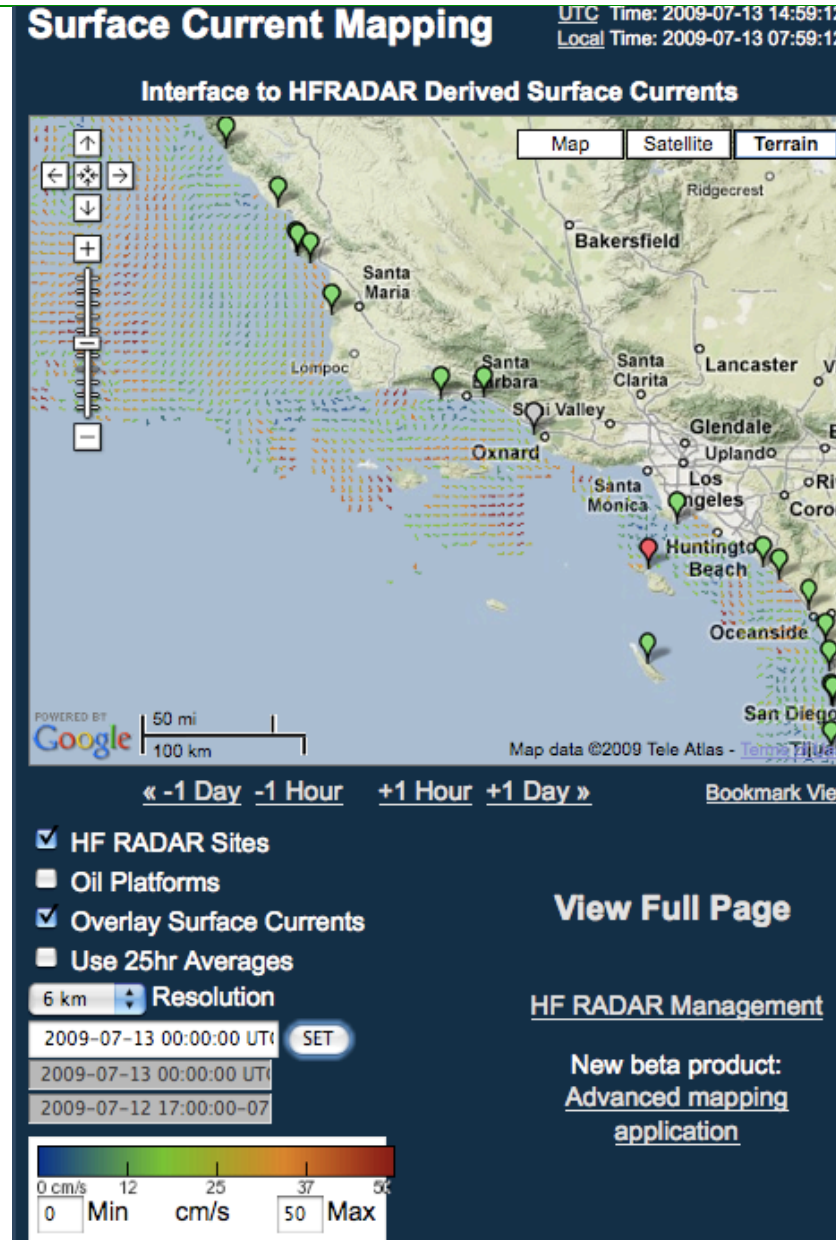
- Global observation network by drifters
  - ~1200 drifters to cover at the $5^o x 5^o$ resolution
- Drifters are used as the platform
  - Eulerian observations of *T* (SLP, Wind)

http://www.aoml.noaa.gov/phod/dac/gdp.html

Remote-Sensing for Surface Current. HF Radar

http://www.cocmp.org/

◆ Actual measurement: radial velocity
  – 2 observation makes 2D surface velocity



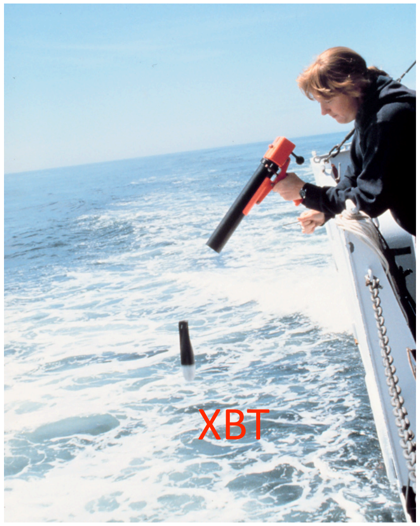
http://www.sccoos.org/data/hfrnet/

Quality Control of In-Situ Profiles

- Developed by Ingleby & Huddleston (2007) at UKMO
- Principles for building the QC system
  - The system has to be automated to cope with the data volume involved
  - Original, reported values should be used as long as possible (flagged, rather than rejected)
  - Any decisions taken by the system should be traceable
  - The system is designed to support data assimilation
  - Tools to monitor system performance and individual cases are available
  - The generic checks and processing use code shared the UKMO atmospheric QC
  - The generic checks have a clear theoretical basis in probability theory
- QC Overview
  - Data specific check
  - Background and buddy check

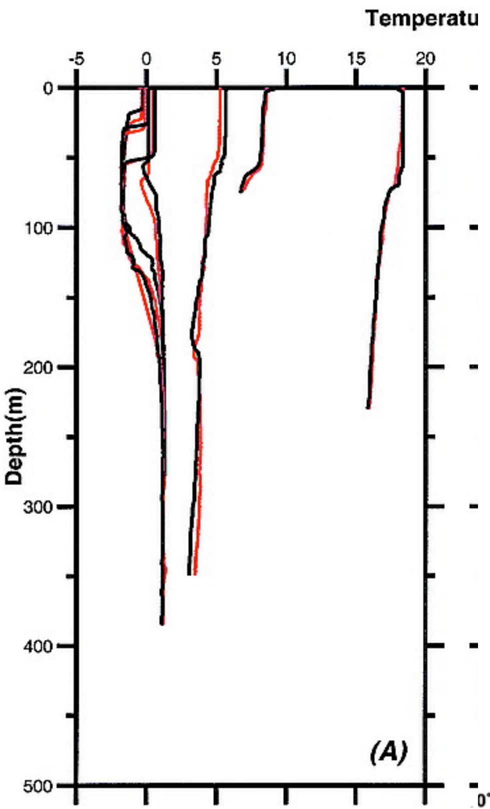
Pre-QC for XBT

- Prior rejects for XBT below 1000m due to inaccuracy of the instruments.
- XBT depth correction: XBT depth is computed based on the time of the release.
  - All XBTs are designed to fall at the known rate, according to the manufacturer's design (formula, or equation for depth vs time from the release).
  - Many won't; revised equation (linear correction) is suggested by Hanawa et al (1995).
  - Depth for the profiles taken before Hanawa et al (1995) were corrected.
  - Additional difficulties:
    - Older data may not have record for type of XBT used
    - In the cold sea, viscosity is higher & drop rate changes.

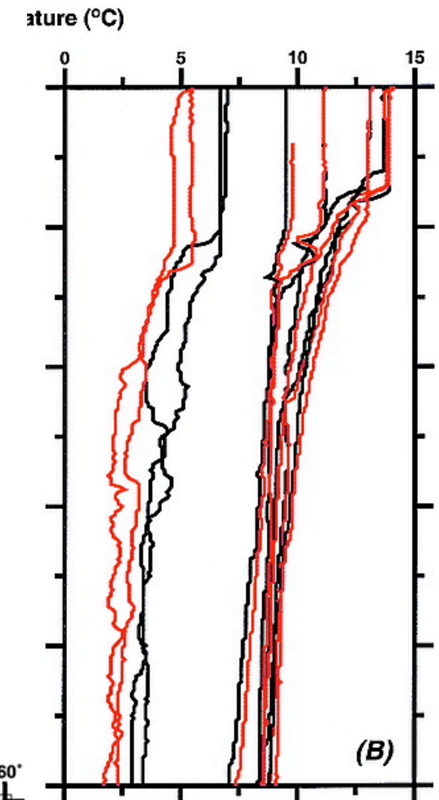


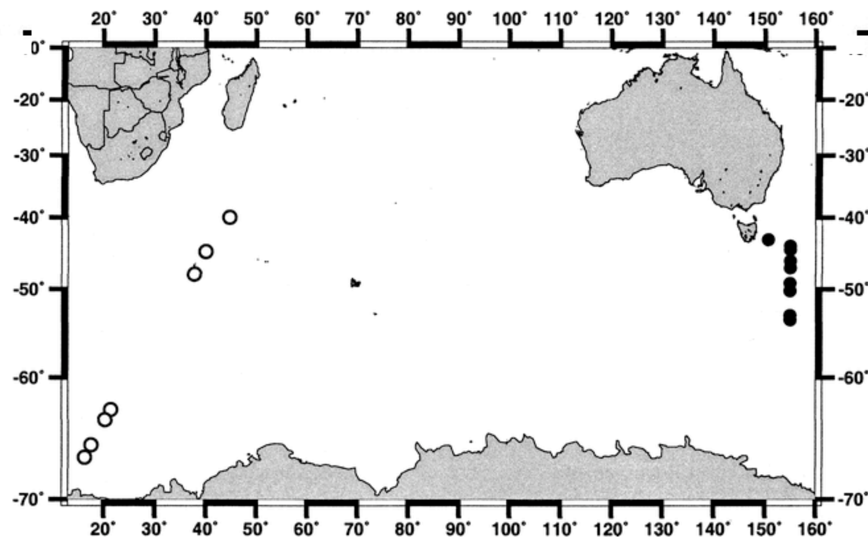
validated against

In Situ Data: XBT vs CTD



Composite profiles of
- XBT (red line) and
- CTD (black line)

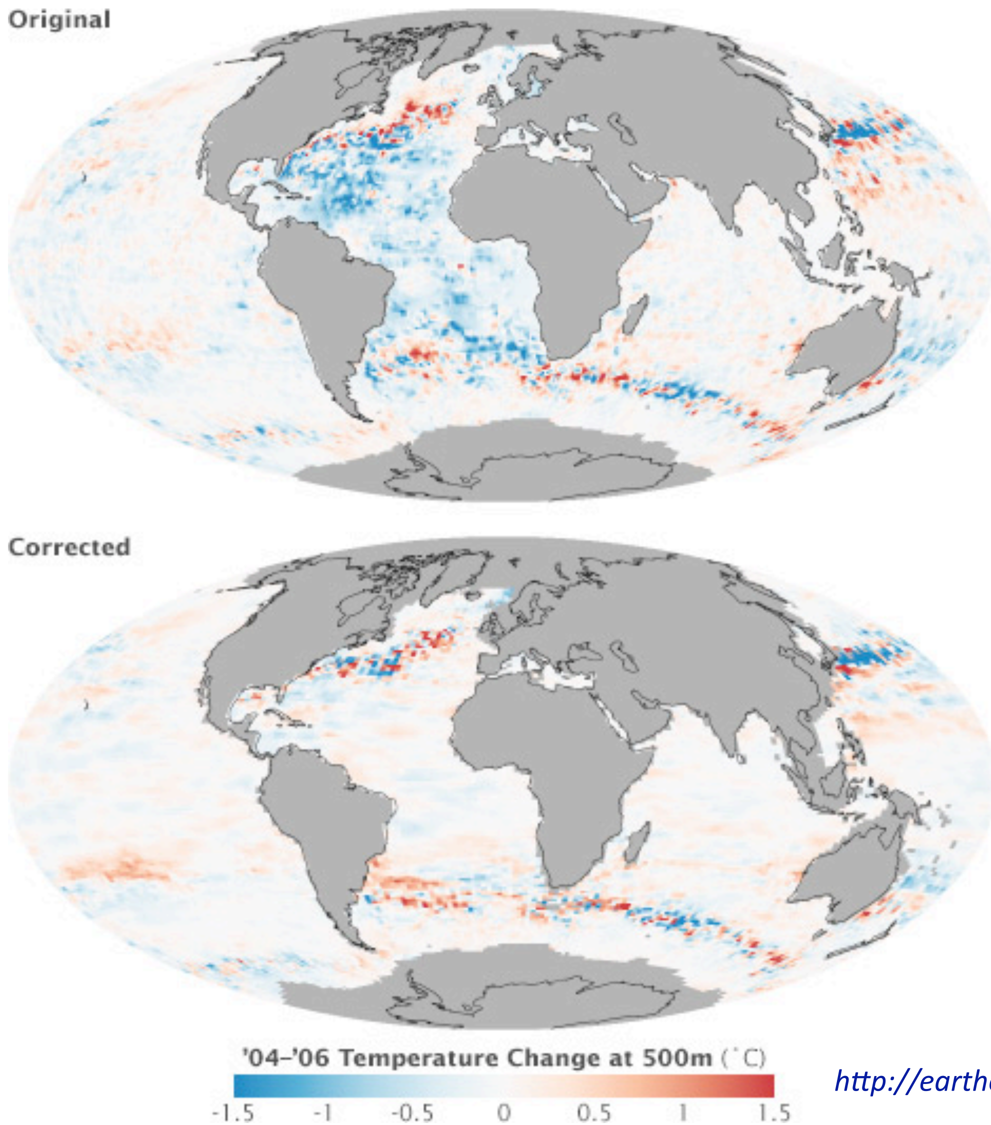




Thadathil et al (2002)

Importance of QC / Bias Correction

- Ocean temperature change from 2004 to 2006



Correction in temperature due to bad data from the Argo floats and XBTs

http://earthobservatory.nasa.gov/Features/OceanCooling/page3.php

QC for Data Specific Check

- Movable instruments
  - Vertical check for the value of data: against constant, spike or step.
  - Horizontal position of the data along the track for each identifier: kinks or jumps.

- Superobs for mooring in time: TRITON are quasi-hourly but formed into daily averages.

- Stability check for T/S based on the density for vertical profiles.

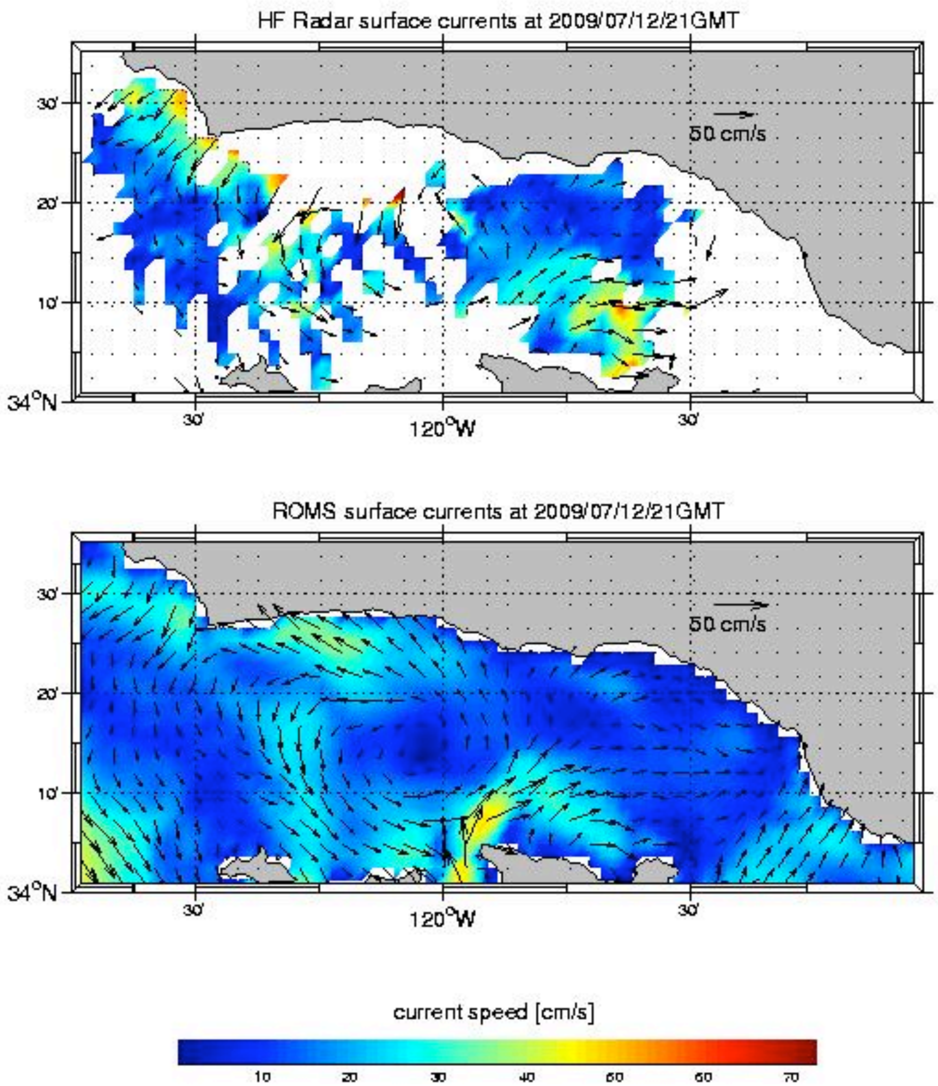
- Duplicate check and thinning in space.

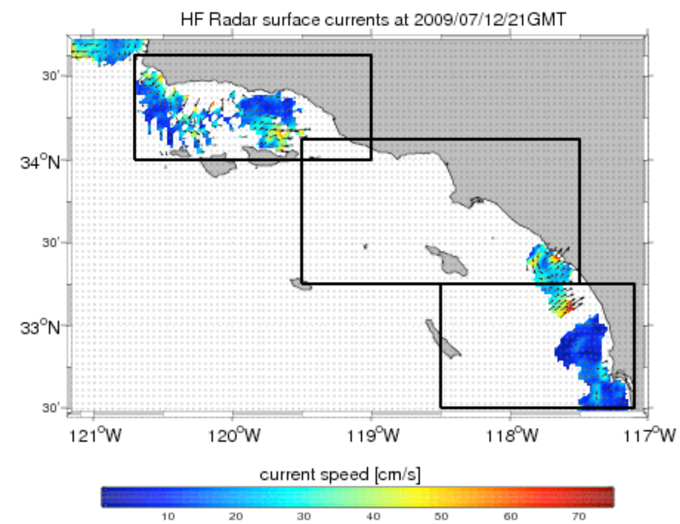
Background and Observation Buddy Checks are Similar to Atmos QC

New Types of Observations

- Ocean observations are (with respected to atmospheric observations)
  - Sporadic
  - Inhomogeneous
  - Limited to upper ocean
- New types of observations that are promising & challenging
  - HF radar observations
  - Lagrangian observations

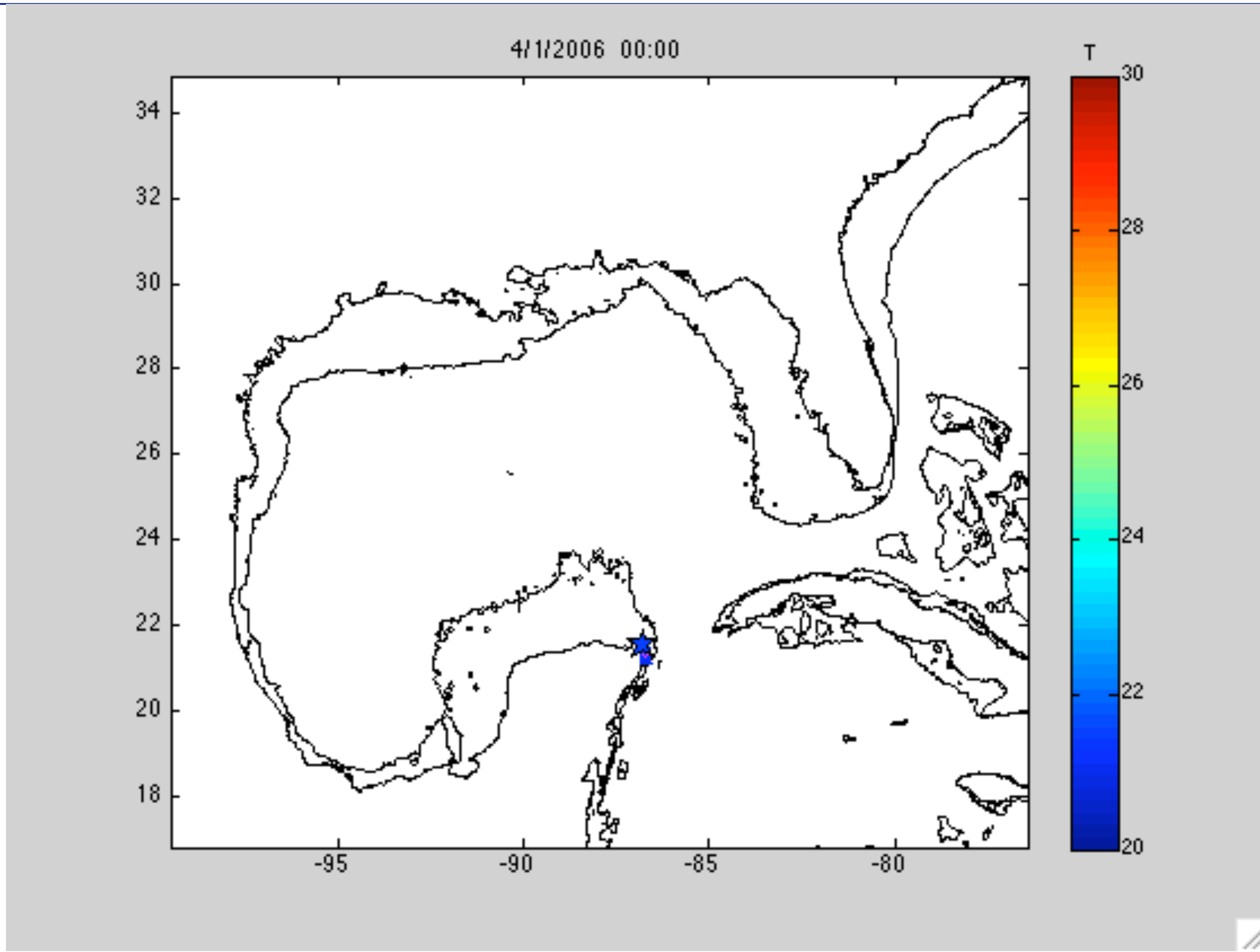
HF Radar: Remote-Sensing Surface Current





- Surface current data
  - Gridded & smoothed (u,v)
- Significant error
  - Representative
  - Instrumental

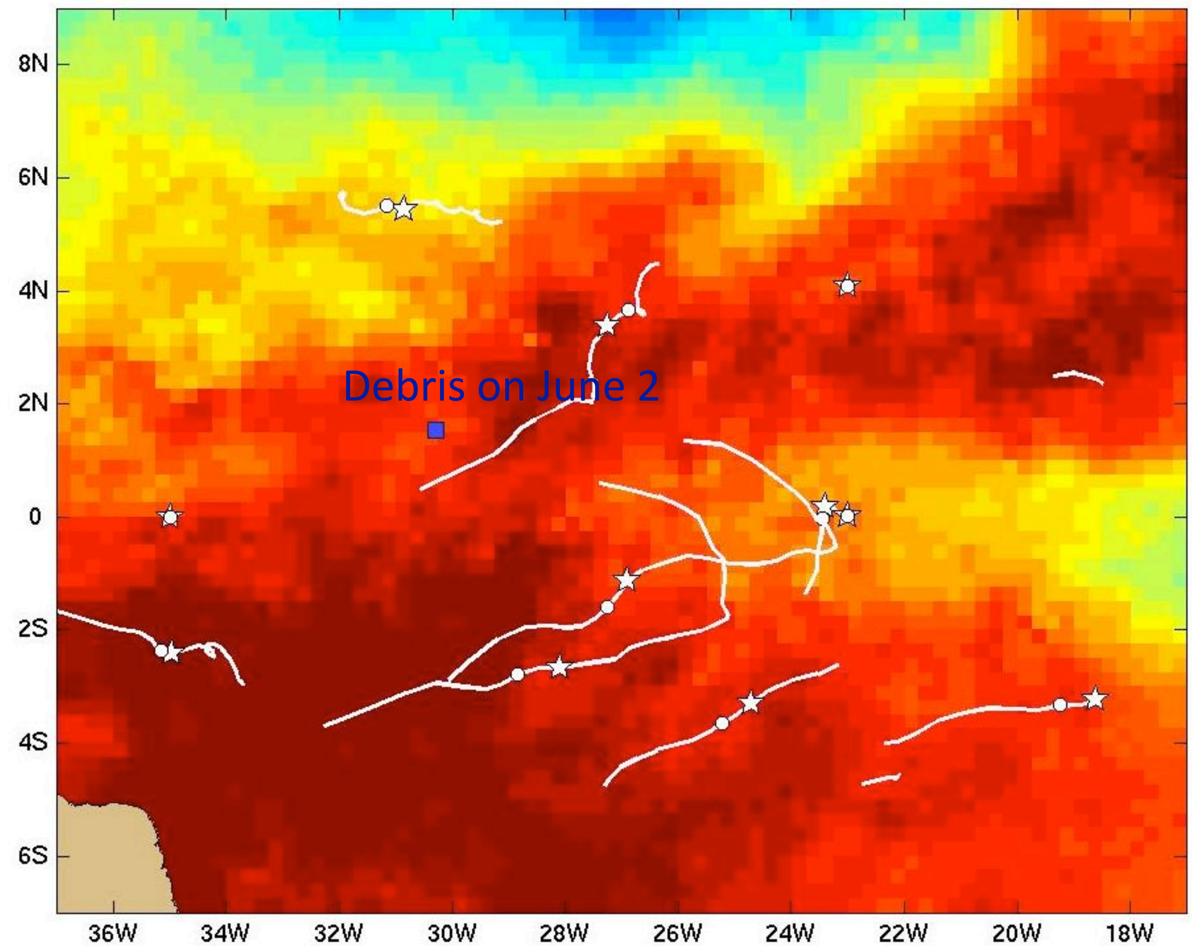
Lagrangian Data



Data available from http://www.aoml.noaa.gov/phod/dac/dacdata.html

Additional Motivation for Lagrangian Data. Rescue Mission

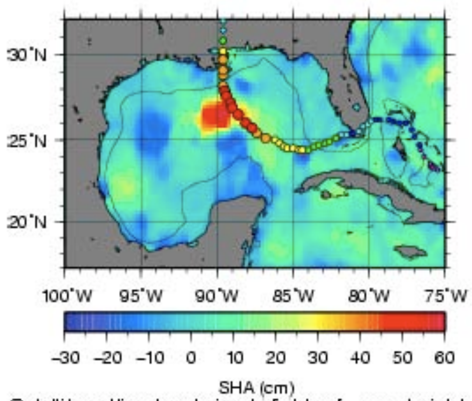
- Air France Flight 447. May 31, 2009
  - Drifters were identified in the area where the plane disappeared
  - Brazilian Navy deployed 5 more drifters on June 14



http://www.aoml.noaa.gov/phod/dac/gdp_information.html

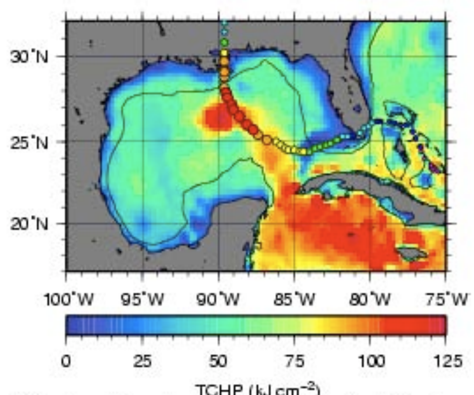
Additional Motivation: Hurricane Prediction

Aug 28



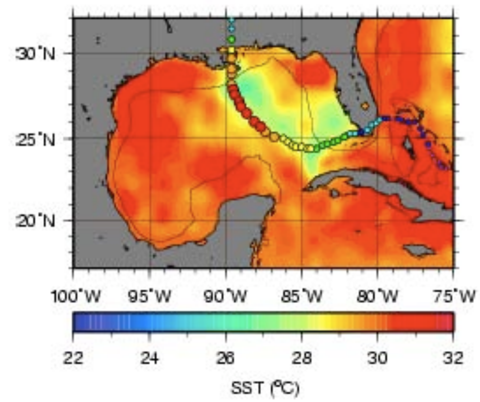
Satellite altimetry-derived field of sea height anomaly (SHA) on August 28, 2005, in the Gulf of Mexico. The large values (red) of SHA in the center of the Gulf are indicative of the presence of a warm anticyclonic ring. The circles of different colors indicate the track and intensity of Hurricane Katrina. The isobath of 200m is superimposed.

Aug 28



Altimeter-derived estimates of Tropical Cyclone Heat Potential (TCHP) for August 28, 2005. The Loop Current and a large warm anticyclonic ring have the largest amount of heat stored in the region. The circles of different colors indicate the track and intensity of Hurricane Katrina. The isobath of 200m is superimposed.

Aug 31

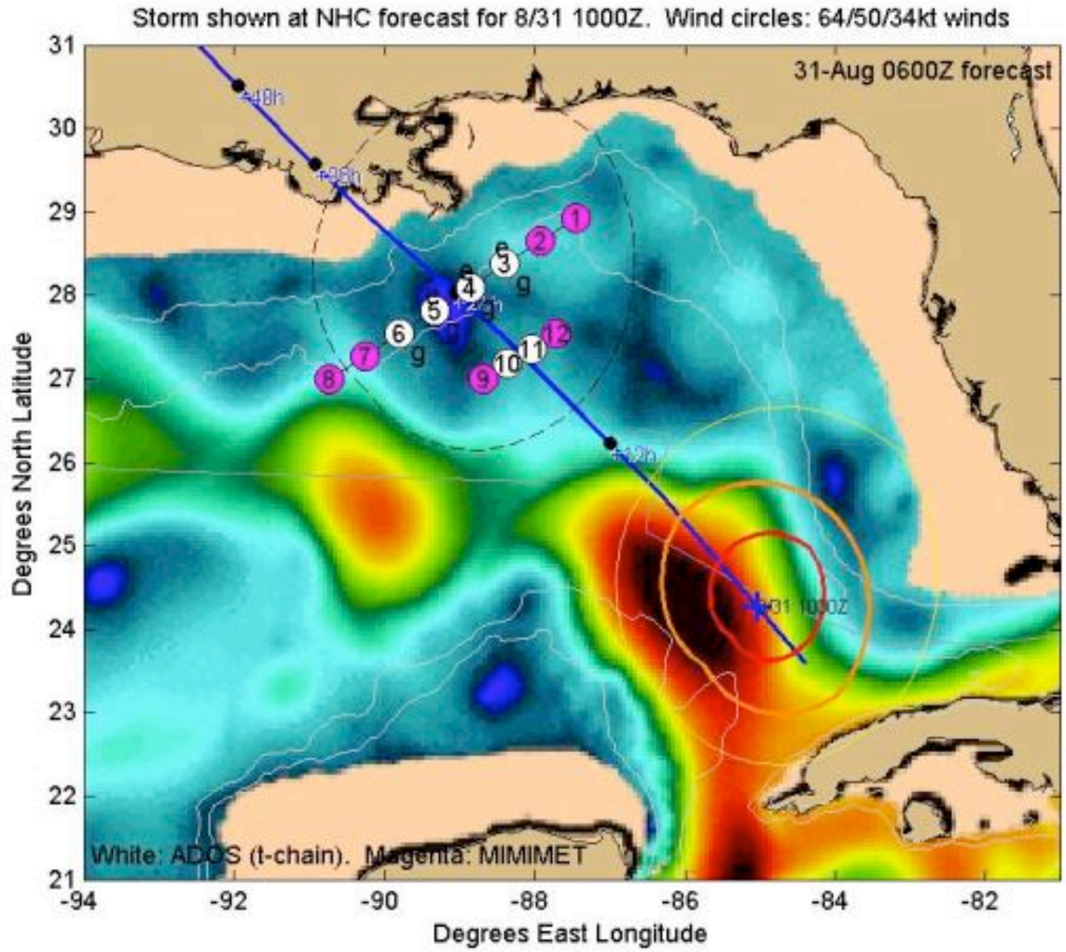


NOAA GOM surface dynamics report for Katrina

http://www.aoml.noaa.gov/phod/altimetry/katrina1.pdf

Adaptive Sampling by Drifters for Better Hurricane Prediction. 1

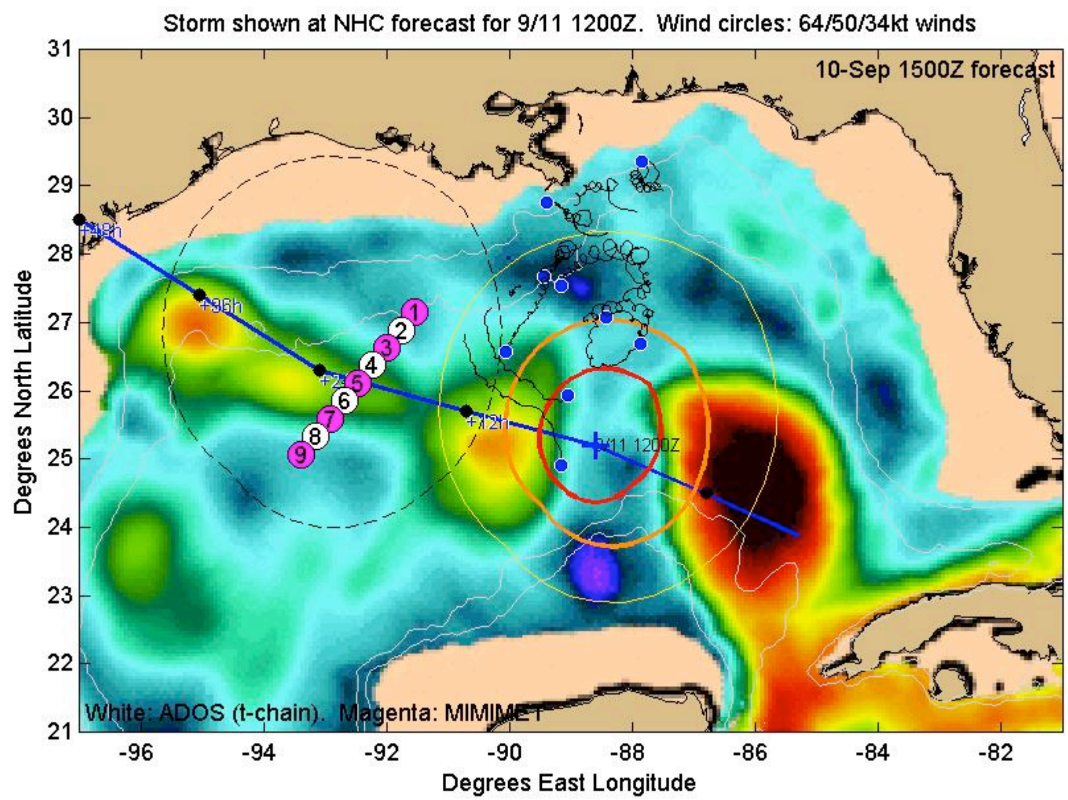
- Hurricane Gustav
  - 12 drifters deployed in the forecast path of Gustav on August 31, 2008
  - All survived and transmitted data



http://www.aoml.noaa.gov/phod/dac/gdp_information.html

Adaptive Sampling by Drifters for Better Hurricane Prediction. 2

- Hurricane Ike
  - 9 drifters deployed in the forecast path of Ike on September 11, 2008
  - Blues are deployed on August 31, 2008



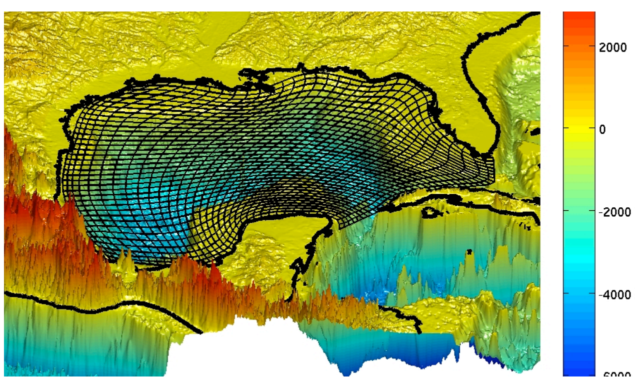
http://www.aoml.noaa.gov/phod/dac/gdp_information.html

Basic Elements of Lagrangian Data Assimilation System

Eulerian Model:
State x_F

$$\mathbf{x}_F(t_k) = \begin{pmatrix} \vdots \\ u_{ijl}(t_k) \\ v_{ijl}(t_k) \\ h_{ijl}(t_k) \\ \vdots \end{pmatrix}$$

$N_F \sim 10^{5\text{-}7}$

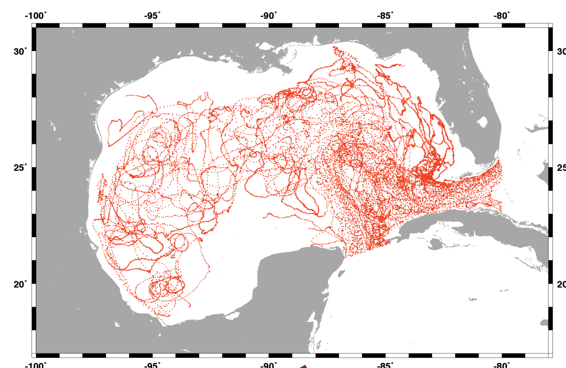


Lagrangian Observation:
Location y_D

$$\mathbf{y}_{D,j}(t_k) = \begin{pmatrix} \vdots \\ r_{D,j}^{(x)}(t_k) \\ r_{D,j}^{(y)}(t_k) \\ \left[r_{D,j}^{(p)}(t_k) \right] \\ \vdots \end{pmatrix}$$

$L_D = 2$ [or 3]

per drifter



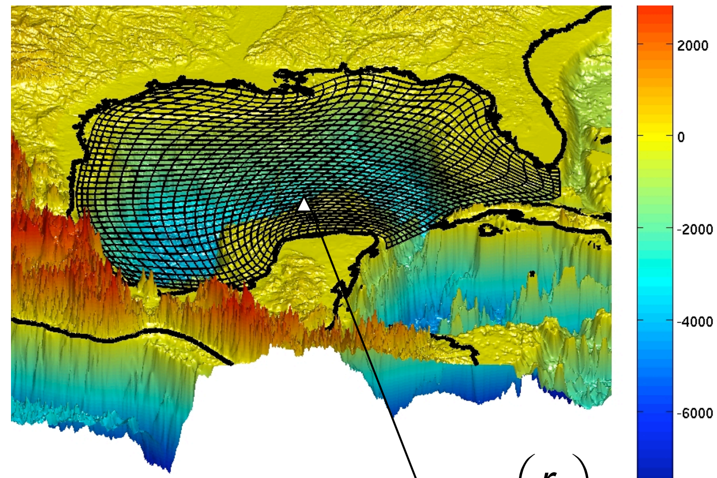
Data Assimilation Method

Ensemble-Based Lagrangian Data Assimilation

- Use of ensemble $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{N_e}\}$ to represent the uncertainty of $\mathbf{x}$ in particular, mean and covariance

  - mean

$$\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_{\mathrm{F}} \\ \bar{\mathbf{x}}_{\mathrm{D}} \end{pmatrix} = \frac{1}{N_e} \sum_{n=1}^{N_e} \begin{pmatrix} \mathbf{x}_{\mathrm{F},n} \\ \mathbf{x}_{\mathrm{D},n} \end{pmatrix}$$



$$\mathbf{x}_{\mathrm{D}} = \begin{pmatrix} r_x \\ r_y \end{pmatrix}$$

  - covariance

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{\mathrm{FF}} & \mathbf{P}_{\mathrm{FD}} \\ \mathbf{P}_{\mathrm{DF}} & \mathbf{P}_{\mathrm{DD}} \end{pmatrix}$$
$$= \frac{1}{N_e - 1} \sum_{n=1}^{N_e} \begin{pmatrix} (\mathbf{x}_{\mathrm{F},n} - \bar{\mathbf{x}}_{\mathrm{F}})(\mathbf{x}_{\mathrm{F},n} - \bar{\mathbf{x}}_{\mathrm{F}})^T & (\mathbf{x}_{\mathrm{F},n} - \bar{\mathbf{x}}_{\mathrm{F}})(\mathbf{x}_{\mathrm{D},n} - \bar{\mathbf{x}}_{\mathrm{D}})^T \\ (\mathbf{x}_{\mathrm{D},n} - \bar{\mathbf{x}}_{\mathrm{D}})(\mathbf{x}_{\mathrm{F},n} - \bar{\mathbf{x}}_{\mathrm{F}})^T & (\mathbf{x}_{\mathrm{D},n} - \bar{\mathbf{x}}_{\mathrm{D}})(\mathbf{x}_{\mathrm{D},n} - \bar{\mathbf{x}}_{\mathrm{D}})^T \end{pmatrix}$$

Kalman filter approach:
- *Extended Kalman filter. Ide, Jones, Kuznetsov (2002)*
- *Ensemble Kalman filter. Salman, Kuznetsov ,Jones, Ide (2006)*

Variational approach ((similar):
- *Ocean 4D-Var. Nodet (2006)*
- *Atmos Operational GEOS5. Meunier et al (2009)*

Mechanisms of Lagrangian Data Assimilation (LaDA)

Forecast from t_{k-1} to t_k:

$$\begin{Bmatrix}\mathbf{x}^{a}_{k-1}\\ \mathbf{P}^{a}_{k-1}\end{Bmatrix} \Rightarrow \begin{Bmatrix}\mathbf{x}^{f}_{k}\\ \mathbf{P}^{f}_{k}\end{Bmatrix}$$

Observation at t_k:

$$\begin{Bmatrix}\mathbf{y}^{o}_{k}\\ \mathbf{R}^{o}_{k}\end{Bmatrix}$$

$$\left.\begin{pmatrix}\mathbf{x}^{f}_{F}\\ \mathbf{x}^{f}_{D}\end{pmatrix}\right|_{k} = \left.\begin{pmatrix} m_{F}\left(\mathbf{x}^{a}_{F}\right)\\ m_{D}\left(\mathbf{x}^{a}_{F},\ \mathbf{x}^{a}_{D}\right)\end{pmatrix}\right|_{k-1}$$

$$\mathbf{y}^{o}_{D}\left(t_{k}\right) \quad \text{with} \quad \mathbf{R}^{o}_{DD}\left(t_{k}\right)$$

$$\left.\begin{pmatrix}\mathbf{P}^{f}_{FF} & \mathbf{P}^{f}_{FD}\\ \mathbf{P}^{f}_{DF} & \mathbf{P}^{f}_{DD}\end{pmatrix}\right|_{k} = M\left.\begin{pmatrix}\mathbf{P}^{a}_{FF} & \mathbf{P}^{a}_{FD}\\ \mathbf{P}^{a}_{DF} & \mathbf{P}^{a}_{DD}\end{pmatrix}\right|_{k-1}$$

Analysis at t_k:

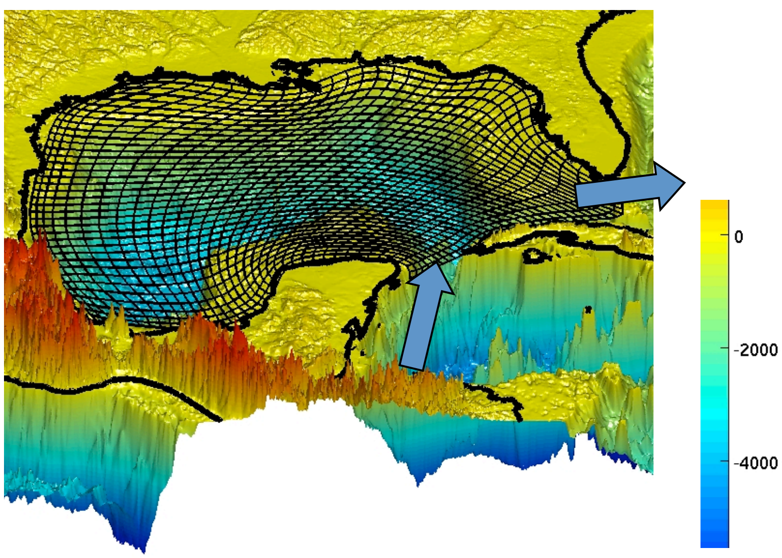
$$\begin{Bmatrix}\mathbf{x}^{f}_{k}\\ \mathbf{P}^{f}_{k}\end{Bmatrix} \Rightarrow \begin{Bmatrix}\mathbf{x}^{a}_{k}\\ \mathbf{P}^{a}_{k}\end{Bmatrix}$$

$$\left.\begin{pmatrix}\mathbf{x}^{a}_{F}\\ \mathbf{x}^{a}_{D}\end{pmatrix}\right|_{k} = \left.\begin{pmatrix}\mathbf{x}^{f}_{F}\\ \mathbf{x}^{f}_{D}\end{pmatrix} + \begin{pmatrix}\mathbf{P}^{f}_{FD}\left(\mathbf{P}^{f}_{DD}+\mathbf{R}^{o}_{DD}\right)^{-1}\\ \mathbf{P}^{f}_{DD}\left(\mathbf{P}^{f}_{DD}+\mathbf{R}^{o}_{DD}\right)^{-1}\end{pmatrix}\left(\mathbf{y}^{o}_{D} - \mathbf{x}^{f}_{D}\right)\right|_{k}$$

$$\left.\begin{pmatrix}\mathbf{P}^{a}_{FF} & \mathbf{P}^{a}_{FD}\\ \mathbf{P}^{a}_{DF} & \mathbf{P}^{a}_{DD}\end{pmatrix}\right|_{k} = F\left.\begin{pmatrix}\mathbf{P}^{f}_{FF} & \mathbf{P}^{f}_{FD}\\ \mathbf{P}^{f}_{DF} & \mathbf{P}^{f}_{DD}\end{pmatrix}\right|_{k}$$

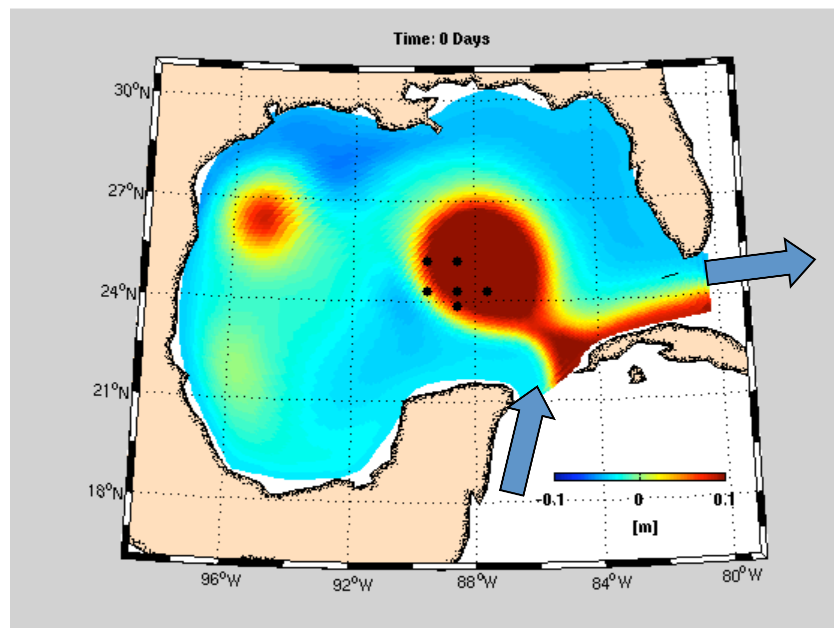
Gulf of Mexico Application

- Ocean circulation:
  Loop-current eddy
  - 3 layer shallow-water model with the structured curvilinear grid
  - Horizontal resolution: 5-13km (average 8.3km)
  - Vertical resolution: 2 layers at 200m, 800m, 2800m
  - Current forcing at 22.4Sv

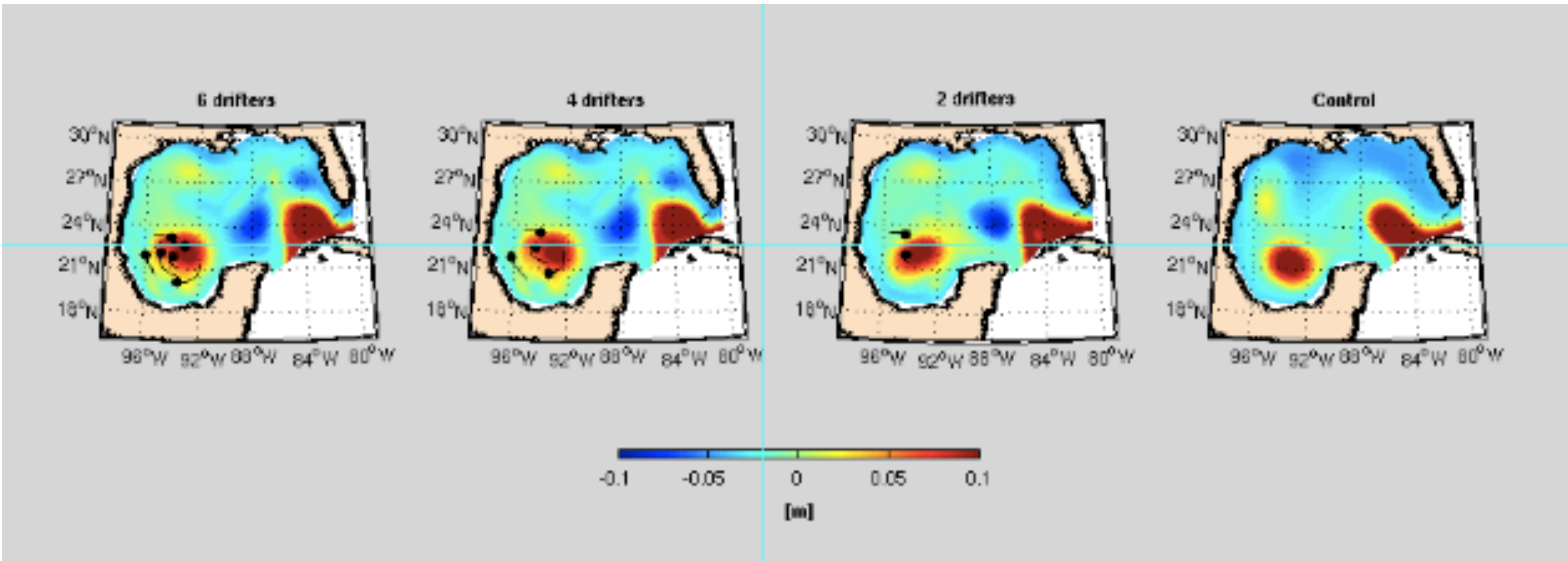


Vernieres, Ide, Jones (2009)

- Data assimilation system
  - Perfect model scenario
    - N_e =32-1028
    - L_D =2-6
  - Initial perturbation in layer depth only (velocity determined by geostrophic balance)



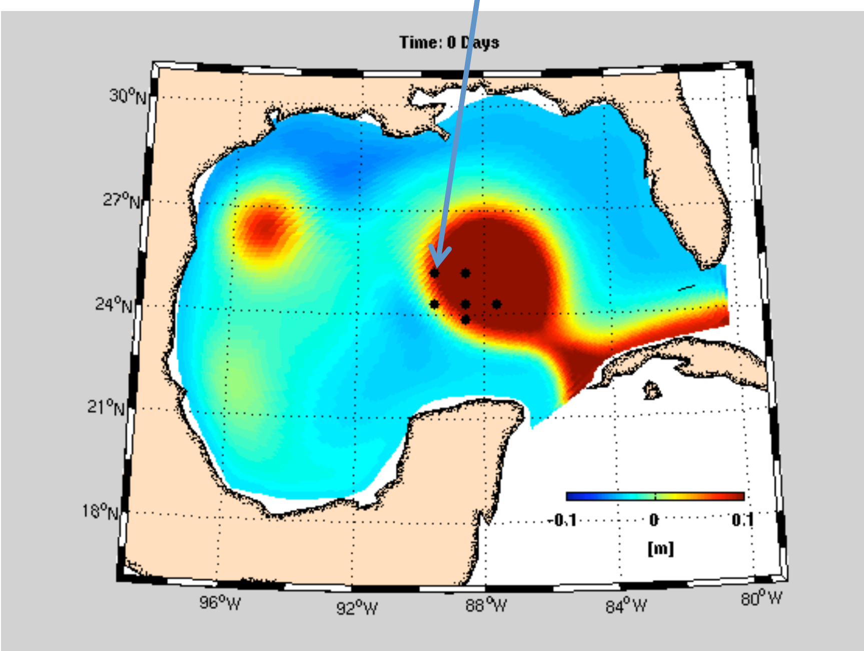
Proof of Concept

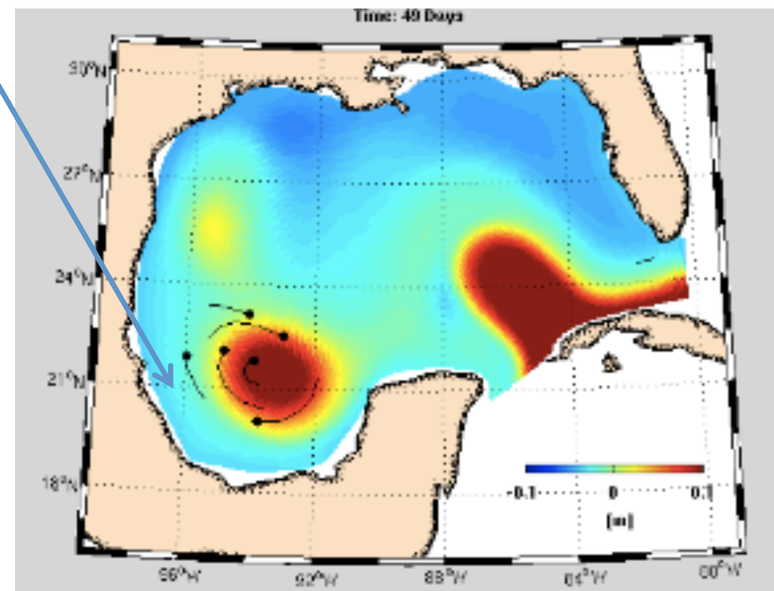


Remarks for Eddy Tracking in the Gulf of Mexico

- LaDA can track the detaching eddy quite effectively.
  - How did I know where to deploy drifters?
- Assimilation of chaotic data???
  - Lagrangian dynamics is chaotic.
  - Necessity for new types of QC can arise from internal/nonlinear dynamics (chaos) of the system

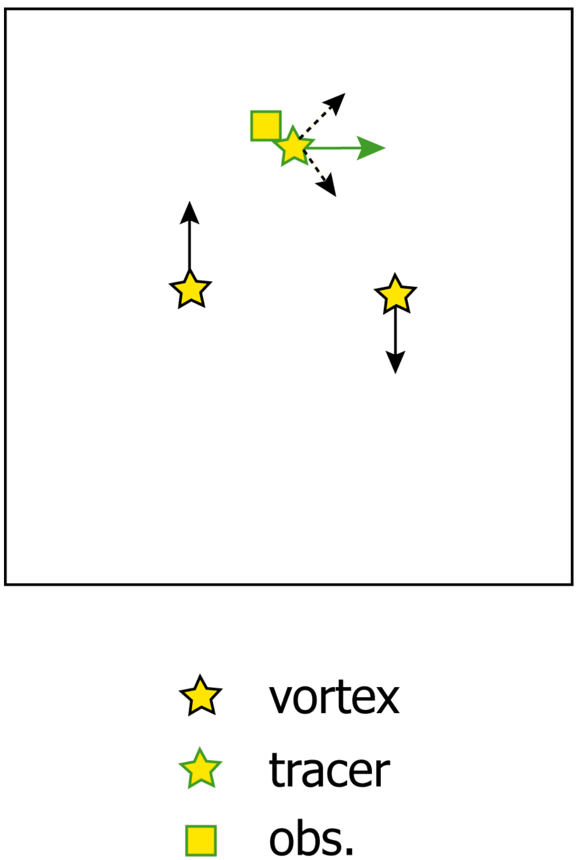
Some drifters look as if they are outside of the eddy
But they will remain in the eddy → How do we know?



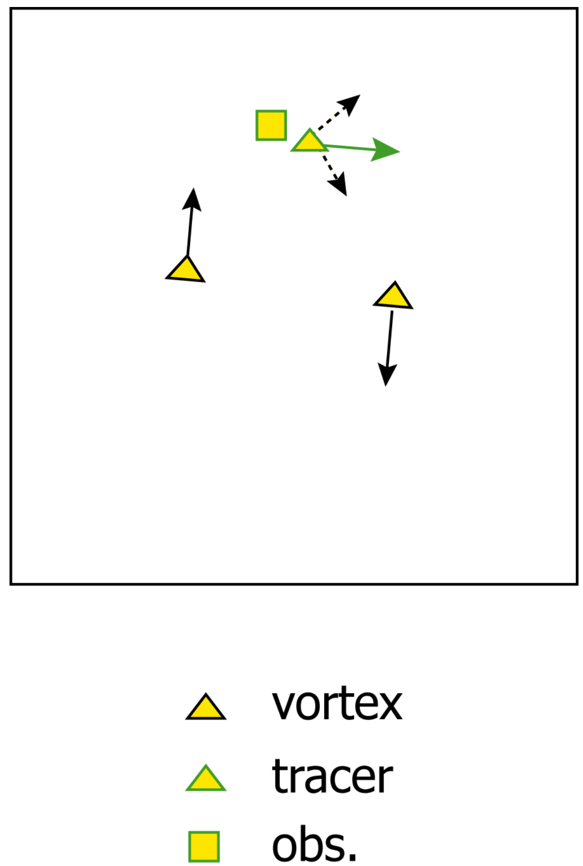


QC. Ocean Dynamics Induced by Vortices (Eddies)

Underlying true system and observation



Data assimilation system

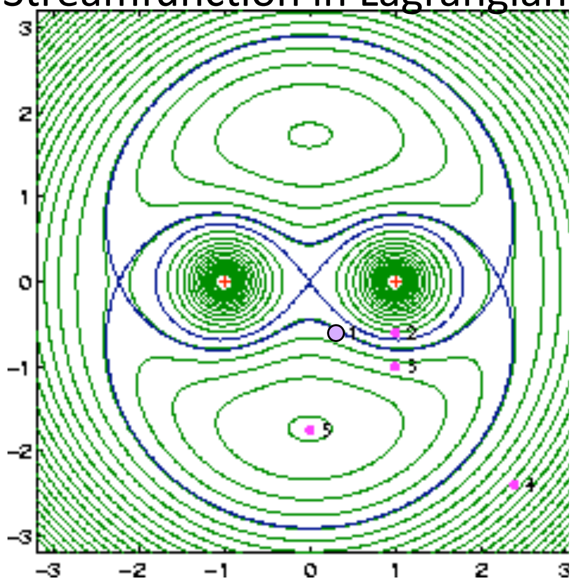


- Ocean
  - Vortex (eddy)
  - Tracer (drifter)
- DA System Parameters
  - σ for $\mathbf{Q}=\sigma^2\mathbf{I}$
  - ρ for $\mathbf{R}=\rho^2\mathbf{I}$
  - ΔT

Assimilation of Chaotic Data

◆ Parameters (σ, ρ, ΔT)=(0.04, 0.02, 1.5)

Streamfunction in Lagrangian Frame



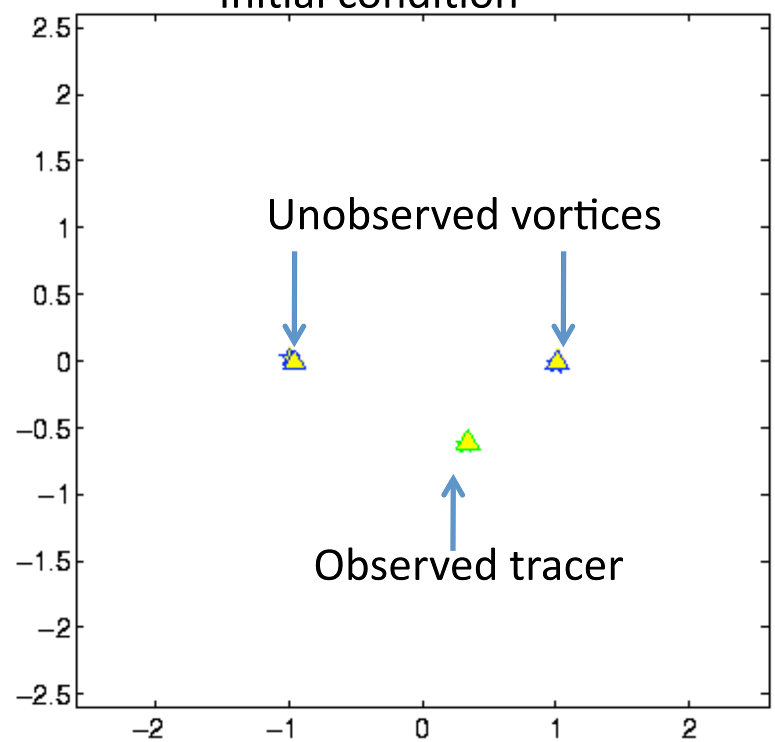
vortex (assimilated)

(true)

tracer (assimilated)

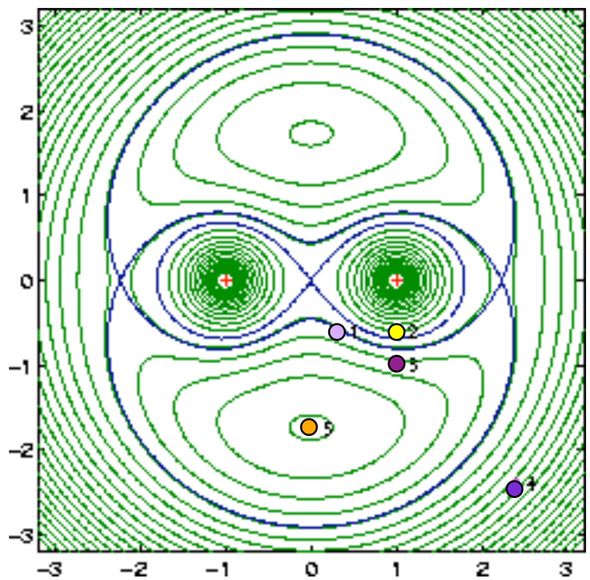
(true)

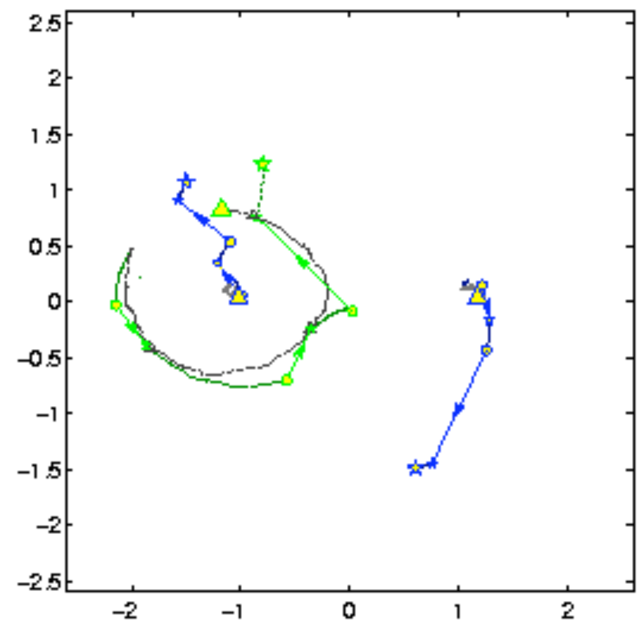
Initial condition



Assimilation of Chaotic Data Without QC

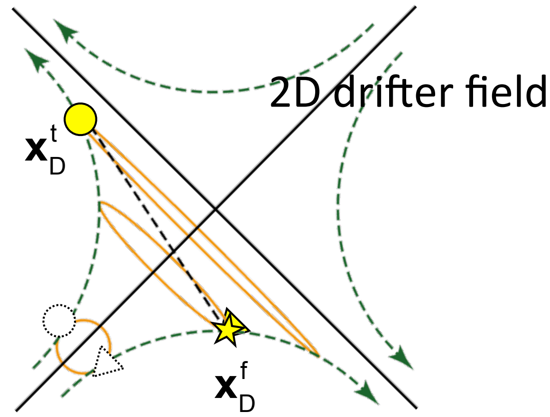
- Parameters (σ, ρ, ΔT)=(0.04, 0.02, 1.5)





(x_T,y_T)	(0.3,-0.6)	(1,-0.6)	(1,-1)	(2.4,-2.4)	(0,-1.75)
Failure %	0	55.5	0.5	15	0
$< \lvert \mathbf{x}^a - \mathbf{x}^t \rvert >$	0.12	1.90	0.11	0.29	0.11

Cause of Sudden Divergence: Hyperbolic Effect



- Hyperbolic effect occurs near the linearly hyperbolic region of velocity
  (λ: hyperbolicity given by the positive local Lyapunov exponent)

$$\frac{d}{dt}\mathbf{x}_D = \mathbf{M}_D(\mathbf{x}_F)\mathbf{x}_D \quad \approx \quad \mathbf{x}_D(t) = \mathbf{F}_D(\mathbf{x}_F)\mathbf{x}_D(t_0)$$

$$\frac{d}{dt}\mathbf{P}_{DD} = \mathbf{M}_D(\mathbf{x}_F)\mathbf{P}_{DD} + \mathbf{P}_{DD}\mathbf{M}_D(\mathbf{x}_F)^T \quad \approx \quad \mathbf{P}_{DD}(t) = \mathbf{F}_D(\mathbf{x}_F)\mathbf{P}_{DD}(t_0)\mathbf{F}_D(\mathbf{x}_F)^T$$

- $\Delta\mathbf{x}^a{}_D$ can be updated correctly: $\mathbf{P}_{DD}$ grows linearly in time with exponent 2λ:

$$\Delta\mathbf{x}_D^a = \mathbf{P}_{DD}^f\left(\mathbf{P}_{DD}^f + \mathbf{R}_D^o\right)^{-1}\left(\mathbf{y}_D^o - \mathbf{x}_D^f\right)$$

- $\Delta\mathbf{x}^a{}_F$ may be unreasonably large because

$$\Delta\mathbf{x}_F^a = \mathbf{P}_{FD}^f\left(\mathbf{P}_{DD}^f + \mathbf{R}_D^o\right)^{-1}\left(\mathbf{y}_D^o - \mathbf{x}_D^f\right)$$

  - $\mathbf{P}_{FD}$ may be approximated by too large with exponent λ:
  - Dragged by a large innovation ($\mathbf{y}^o{}_D$-$\mathbf{x}^o{}_D$)

QC of Chaotic Data: Simple Sanity Check for $\Delta\mathbf{x}^a{}_F$

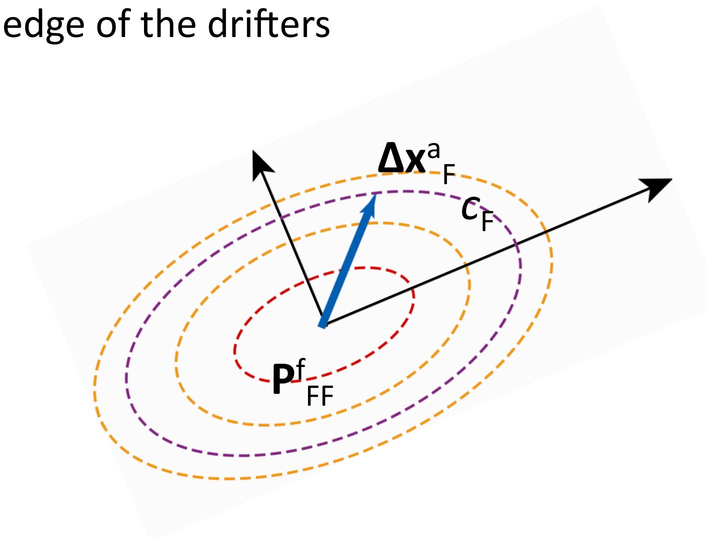
- $C_F \equiv$ Standard deviation of $\Delta\mathbf{x}^a{}_F$ with respect to the expected error
  - $\mathbf{P}^f{}_{FF}$ is independent of $\mathbf{x}^f{}_D$ or $\mathbf{x}^t{}_D$: the flow has no knowledge of the drifters

$$C_F = \left(\Delta\mathbf{x}_F^a\right)^T \left(\mathbf{P}_{FF}^f\right)^{-1} \Delta\mathbf{x}_F^a$$

Doesn't care about $\mathbf{P}^f{}_{FD}$

$$\Delta\mathbf{x}_F^a = \mathbf{P}_{FD}^f \left(\mathbf{P}_{DD}^f + \mathbf{R}_D^o\right)^{-1} \left(\mathbf{y}_D^o - \mathbf{x}_D^f\right)$$

May not be correctly estimated



- Implementation
  - if $C_F < \delta$: Update of $\mathbf{x}_F$
  - if $C_F \geq \delta$: No update of $\mathbf{x}_F$ (but $\mathbf{x}_D$ is updated)
- C_F is computed for $\mathbf{x}_D$ within r_{FD} from $\mathbf{x}_D$ ($r_{FD} < r_{loc}$)
- Control is applied to entire $\mathbf{x}_D$ within r_{loc} from $\mathbf{x}_D$

$\boxed{\text{Method is simple yet works extremely well}}$

Summary

- Ocean data assimilation is relatively new, and may have different
  - Principal goals
  - Scales
  - Observations.
- Conceptually, ocean data QC and atmospheric data QC are similar, but as in the case for atmospheric data QC, we have to handle
  - Individual
  - Data sets
  - Data types
- Observations are extremely important in ocean data assimilation
  - Number is limited
  - Inhomogeneous in time and space
- New types of observations are available
  - New types of QC: for example, QC for chaotic data (e.g., Lagrangian data)
  - Observing system design can be as important as the observed data themselves.