Joint NPS-NCEI Seminar, NOAA Campus, Silver Spring, Maryland, 22 Sept 2015

A completely gridded, optimally-estimated sea water temperature and salinity data set for ocean model assimilation and assessment

Peter C Chu, Naval Postgraduate School

Charles Sun, NCEI

Collaborators

- Chenwu Fan (NPS)
- Carlos Lozano (NOAA/NCEP)
- Leonid M. Ivanov (NPS)
- Tetyana Margolina (NPS)
- Robin Tokmakian (NPS)
- Oleg Melnichenko (Univ of Hawaii)

References

- Chu, P.C., R.T. Tokmakian, C.W. Fan, and C.L. Sun, 2015: Optimal spectral decomposition (OSD) for ocean data assimilation. Journal of Atmospheric and Oceanic Technology, **32**, 828-841.

- Chu, P.C., 2011: Global upper ocean heat content and climate variability. *Ocean Dynamics*, **61** (8), 1189-1204.

- Chu, P.C., L.M. Ivanov, and O.M. Melnichenko, 2005: Fall-winter current reversals on the Texas-Lousiana continental shelf. Journal of Physical Oceanography, 35, 902-910

- Chu, P.C., L.M. Ivanov, O.M. Melnichenko, and N.C. Wells, 2007: On long baroclinic Rossby Waves in the tropical North Atlantic observed from profiling floats. Journal of Geophysical Research, **112**, C05032, doi:10.1029/2006JC003698..

- Galanis, G., G. Emmanouil, P.C. Chu, and G. Kallos, 2009: A new methodology for the extension of the impact of data assimilation on ocean wave prediction. *Ocean Dynamics*, **59**, 523-535

- http://faculty.nps.edu/pcchu

References

- Chu, P.C., L.M. Ivanov, T.P. Korzhova, T.M. Margolina, and O.M. Melnichenko, 2003a: Analysis of sparse and noisy ocean current data using flow decomposition. Part 1: Theory. Journal of Atmospheric and Oceanic Technology, 20 (4), 478-491.

- Chu, P.C., L.M. Ivanov, T.P. Korzhova, T.M. Margolina, and O.M. Melnichenko, 2003b: Analysis of sparse and noisy ocean current data using flow decomposition. Part 2: Application to Eulerian and Lagrangian data. Journal of Atmospheric and Oceanic Technology, 20 (4), 492-512.

- Chu, P.C., L.M. Ivanov, and T.M. Margolina, 2004: Rotation method for reconstructing process and field from imperfect data. International Journal of Bifurcation and Chaos, 14(8), 2991-2997.

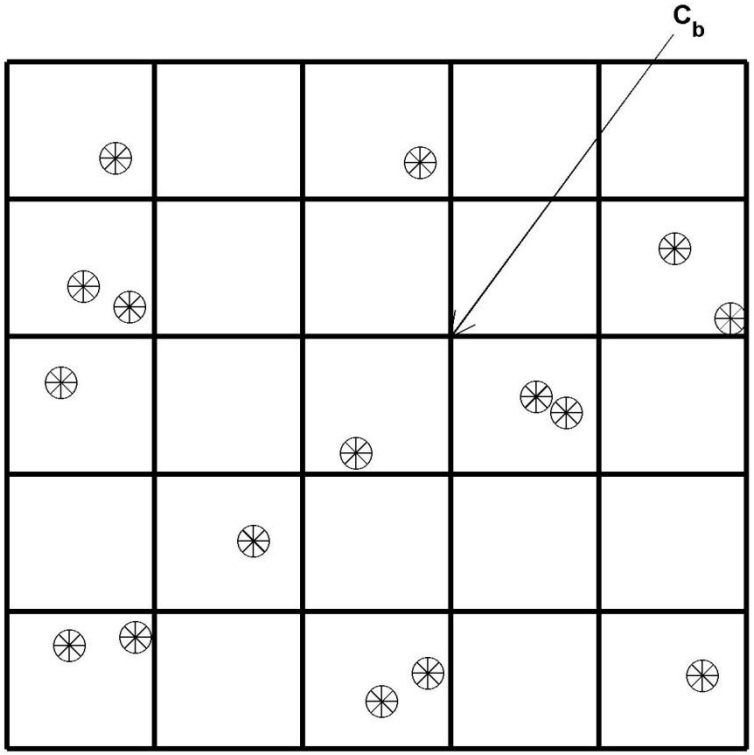
- These papers can be downloaded from:
- http://faculty.nps.edu/pcchu

Part-1

Why do we need a new method for ocean data analysis and assimilation ?

Major Ocean Data Assimilation Schemes

- Optimal Interpolation (OI)
- Kalman Filter
- Variational Methods



Ocean data assimilation with $\boldsymbol{c}_b$ located at the grid points, and $\boldsymbol{c}_o$ located at the points '*'. The ocean data assimilation is to convert the innovation, $\boldsymbol{d} = \boldsymbol{c}_o - \mathbf{H}\boldsymbol{c}_b$, from the observational points to the grid points.

Assimilated Variable $\mathbf{c}(\mathbf{x}, z, t)$

$\mathbf{c}_t$ → "true"

$\mathbf{c}_a$ → analysis (assimilated)

$\mathbf{c}_b$ → background (modeled)

$(\mathbf{c}_t, \mathbf{c}_a, \mathbf{c}_b)$ → grid points → $\mathbf{r}_n$ → total N

$\mathbf{c}_o$ → Observation → Obs points → $\mathbf{r}^{(m)}$

→ total M

Data Assimilation

$$\mathbf{c}_a = \mathbf{c}_b + \mathbf{W}\mathbf{d}$$

Innovation → $\mathbf{d} = \mathbf{c}_o - \mathbf{H}\mathbf{c}_b$

Various ways → **W** – Weight Matrix
→ Different Data Assimilation Schemes

H =$[h_{mn}]$ → the $M \times N$ linear observation operator matrix

W → Depends on **B**, **R**

B → Background Error Covariance Matrix

R → Observational Error Covariance Matrix

Background and Observational Error Covariance Matrices B and R

$$\boldsymbol{\varepsilon}_a = \mathbf{c}_a - \mathbf{c}_t, \quad \boldsymbol{\varepsilon}_o \equiv \mathbf{H}^T \mathbf{c}_o - \mathbf{c}_t$$

$$E^2 = \left\langle \boldsymbol{\varepsilon}_a^T \boldsymbol{\varepsilon}_a \right\rangle \rightarrow \min$$

$$\partial E^2 / \partial w_{nm} = 0$$

- Optimal Interpolation (OI) $\mathbf{W} = \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$
- Kalman Filter (KF) $\mathbf{W}_i = \mathbf{B}^f(t_i)\mathbf{H}^T [\mathbf{R}_i + \mathbf{H}_i \mathbf{B}^f(t_i)\mathbf{H}^T]^{-1}$
- Variational Method $\mathbf{W} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}$
- Data Assimilation Equation (OI/KF) →

$$\mathbf{c}_a = \mathbf{c}_b + \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} \mathbf{d}$$

Gaussian Covariance Function for **B**-Matrix (Bretherton et al. 1976)

Assumption → statistically stationary and homogeneous

$\mathbf{B} = \left[b_{ij} \right]_{N \times N}$ → between two grid points $\mathbf{r}_i$ and $\mathbf{r}_j$

$$b_{ij} = \left(1 - \frac{r_{ij}^2}{r_b^2} \right) \exp\left(1 - \frac{r_{ij}^2}{r_a^2} \right), \quad r_{ij}^2 = \left| \mathbf{r}_i - \mathbf{r}_j \right|^2$$

Two user-input parameters → r_a, r_b

Observational Error Variance Matrix **R**

$$\mathbf{R} = E_o^2 \mathbf{I}, \qquad E_o^2 \equiv \left\langle \boldsymbol{\varepsilon}_o^T \boldsymbol{\varepsilon}_o \right\rangle$$

$\mathbf{I}$ → identity matrix

The third user-input parameter → E_o

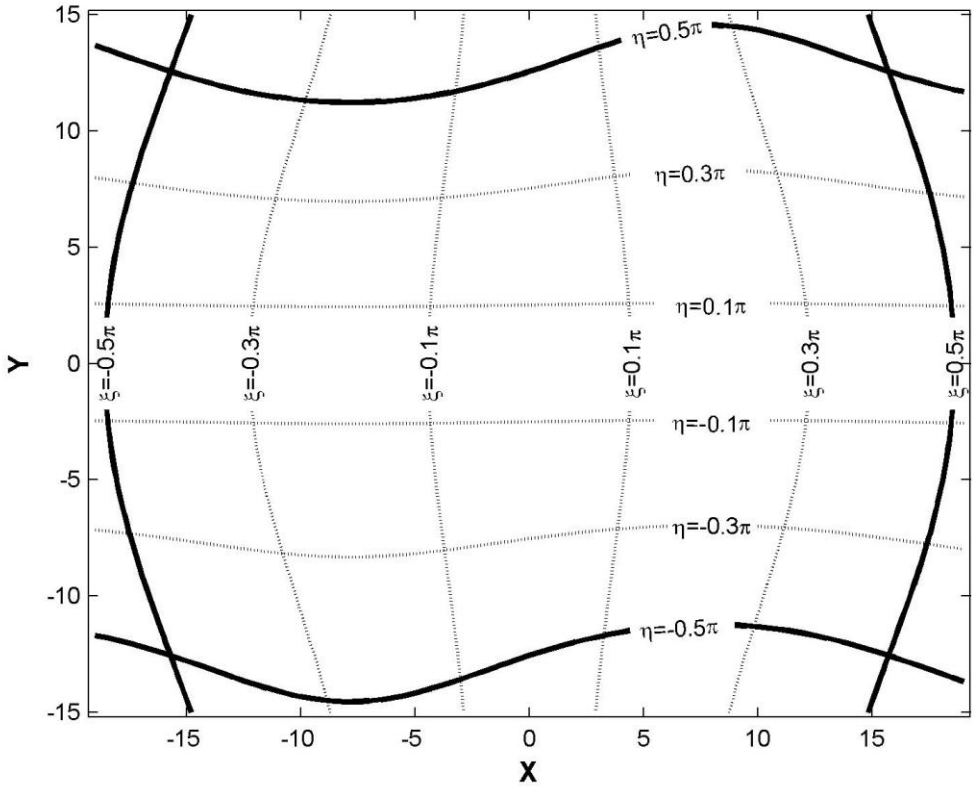
→ Instrument accuracy

OI → three user-input parameters (r_a, r_b , E_o)

Test Case – A Curved Region

$$\frac{x}{10}-0.3\cos\left(\frac{y}{8}\right)\sin\left(\frac{x}{10}\right)=\xi=\begin{cases}-\pi/2 & \text{(west)}\\ \pi/2 & \text{(east)}\end{cases}$$

$$\frac{y}{8}-0.2\sin\left(\frac{x}{5}\right)\left[1-\cos\left(\frac{y}{8}\right)\right]=\eta=\begin{cases}-\pi/2 & \text{(south)}\\ \pi/2 & \text{(north)}\end{cases}$$

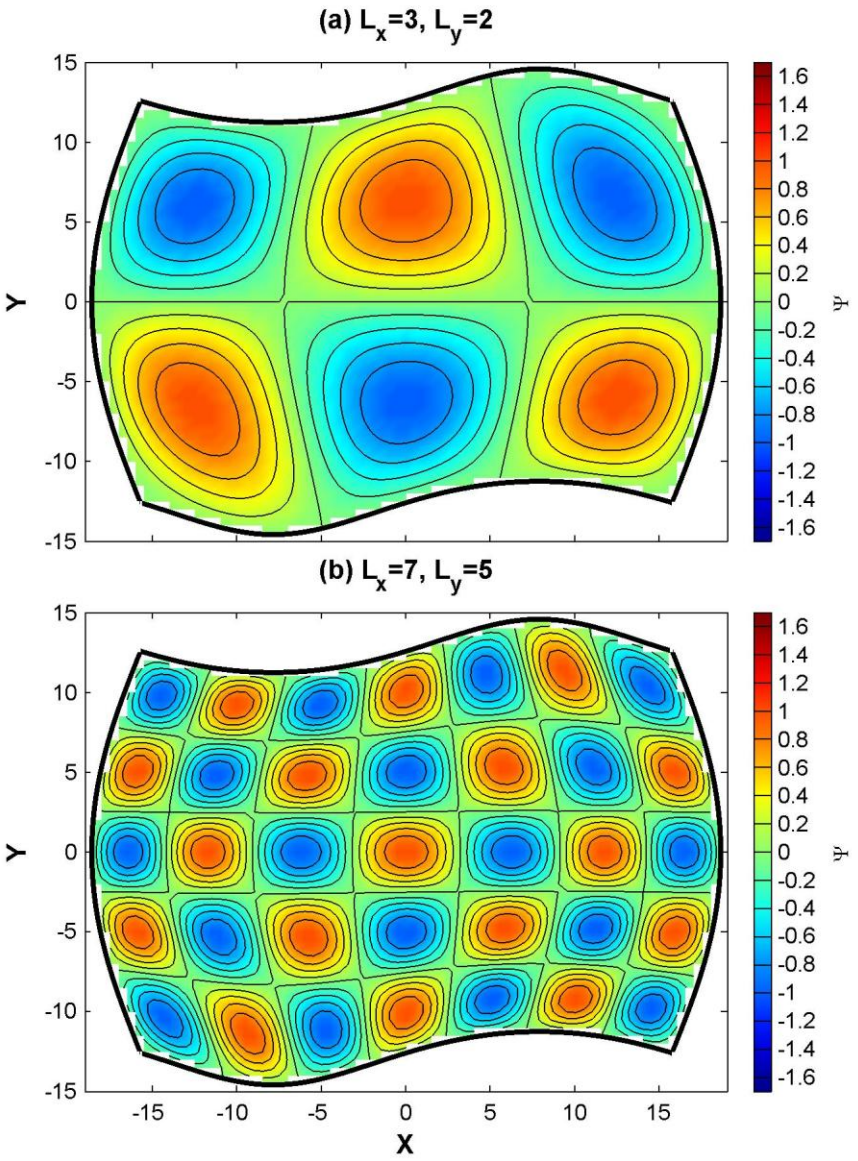


“Truth” – Analytical Streamfunction

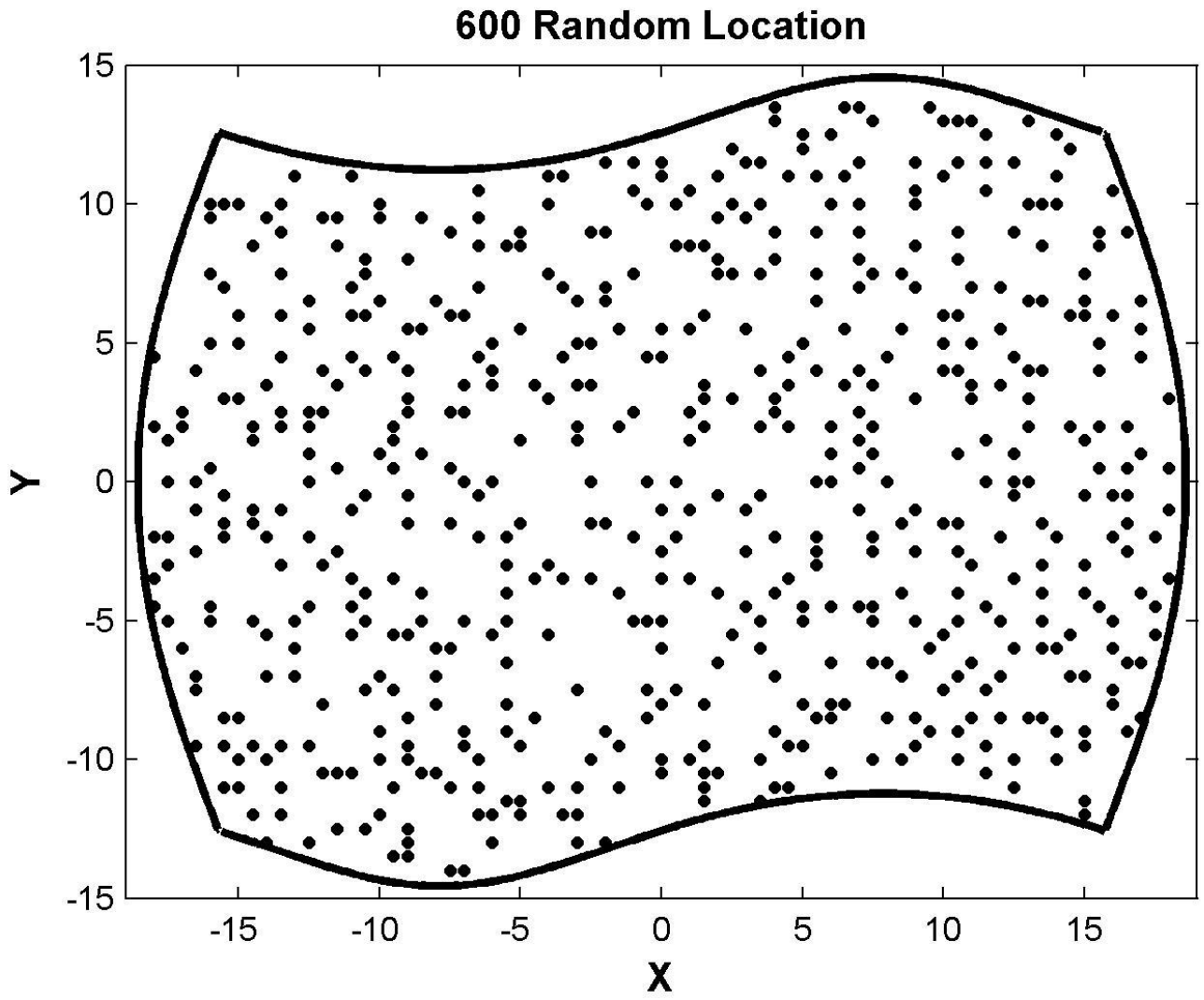
Large-Scale Eddies →

$$\begin{cases} \psi_t(x,y) = \cos\left[L_x\xi(x,y)\right]\cos\left[L_y\eta(x,y)+\beta\right] \\ \xi = \dfrac{x}{10} - 0.3\cos\left(\dfrac{y}{8}\right)\sin\left(\dfrac{x}{10}\right), \\ \eta = \dfrac{y}{8} - 0.2\sin\left(\dfrac{x}{5}\right)\left[1-\cos\left(\dfrac{y}{8}\right)\right] \end{cases}$$

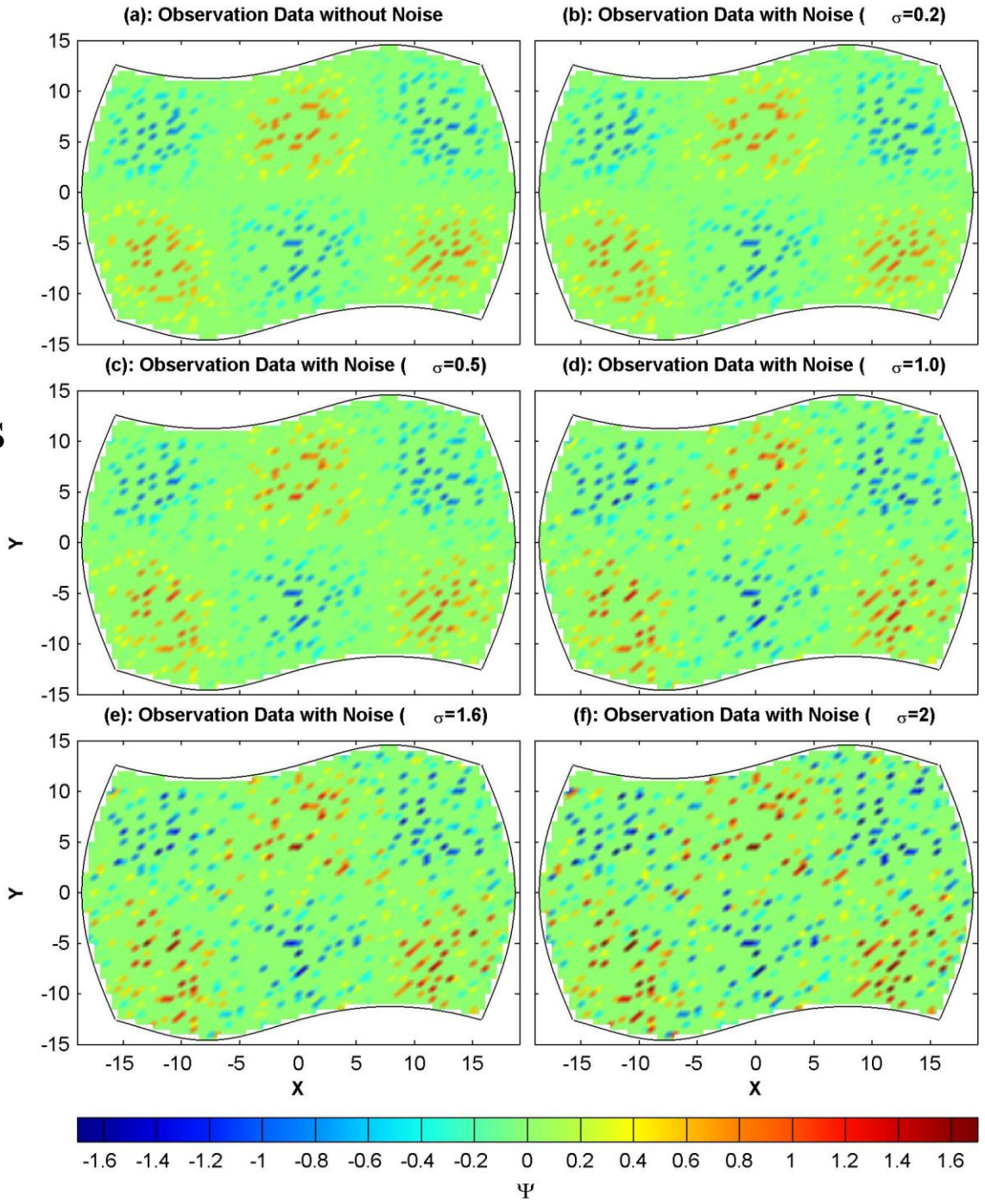
Small-Scale Eddies →



Randomly selected locations (total: 600) inside domain as “observational” points



"Observational" data (ψ_o) from "truth" (large-scale) plus white Gaussian noises of zero mean and various standard deviations: (a) 0 (i.e., no noise) (b) 0.2, (c) 0.5, (d) 1.0, (e) 1.6, and (f) 2.0.



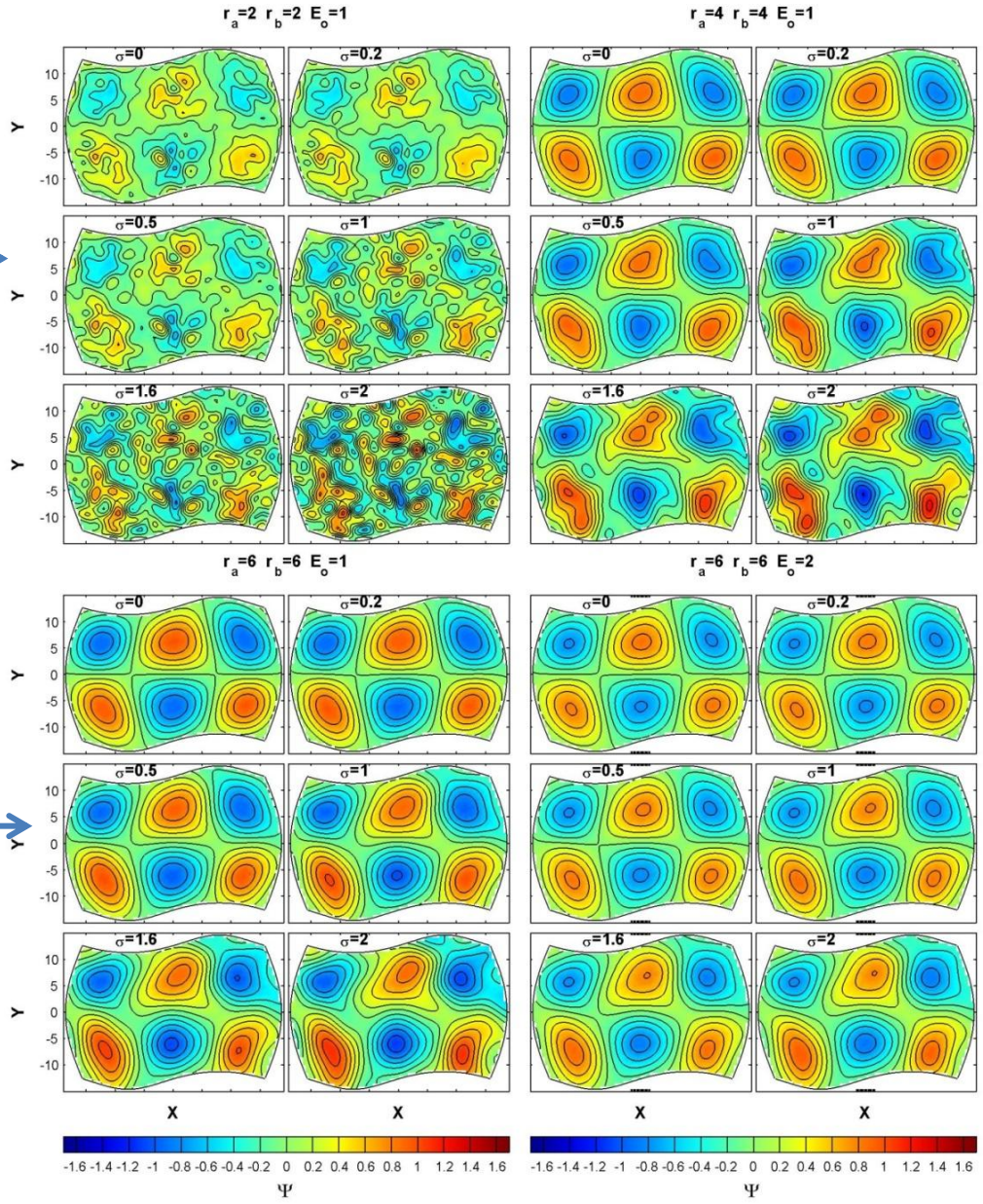
Analysis field ψ_a obtained by the OI – large eddies

$r_a = 2$
$r_b = 2$
$E_o = 1$

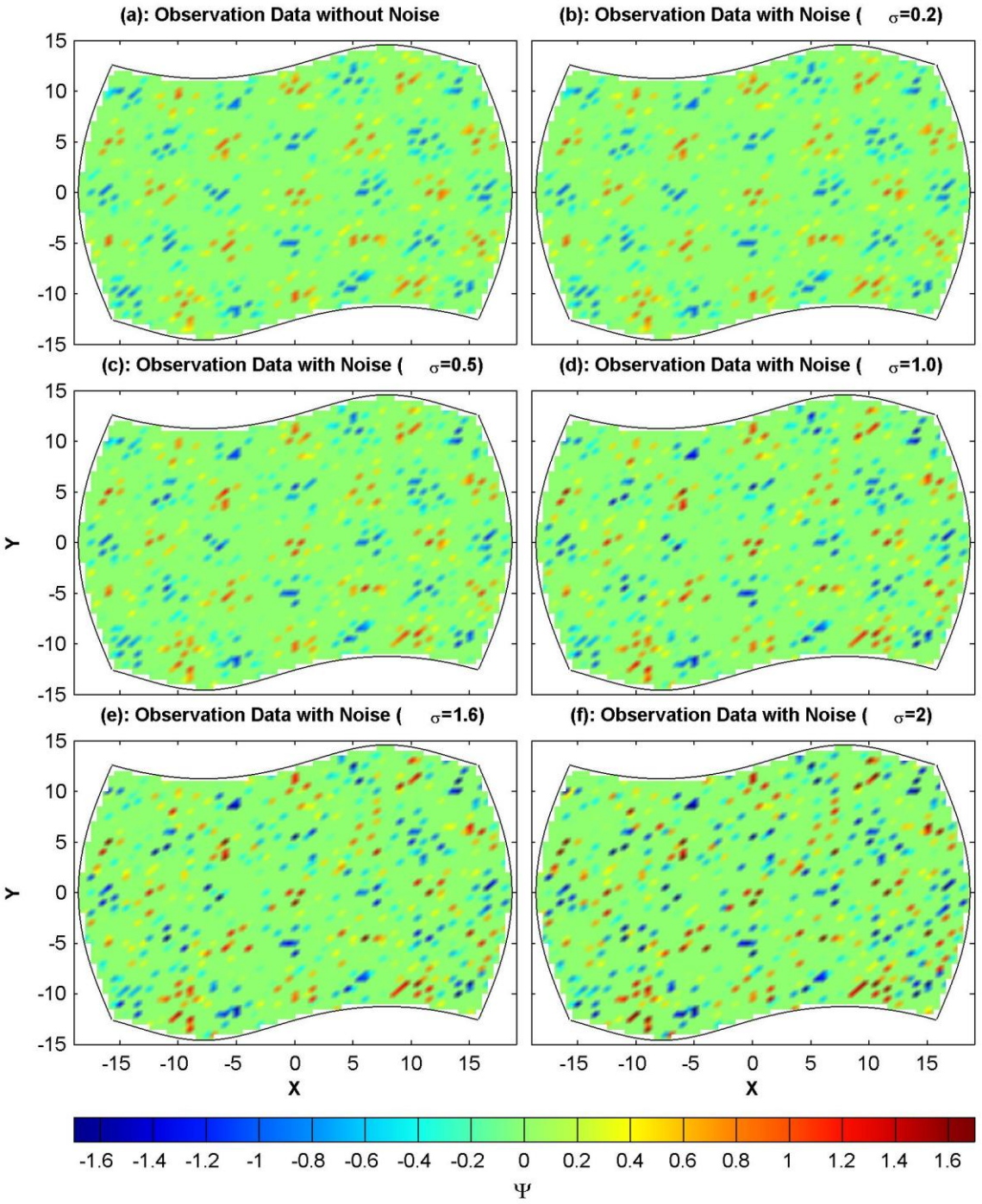
$r_a = 4$
$r_b = 4$
$E_o = 1$

$r_a = 6$
$r_b = 6$
$E_o = 1$

$r_a = 6$
$r_b = 6$
$E_o = 2$

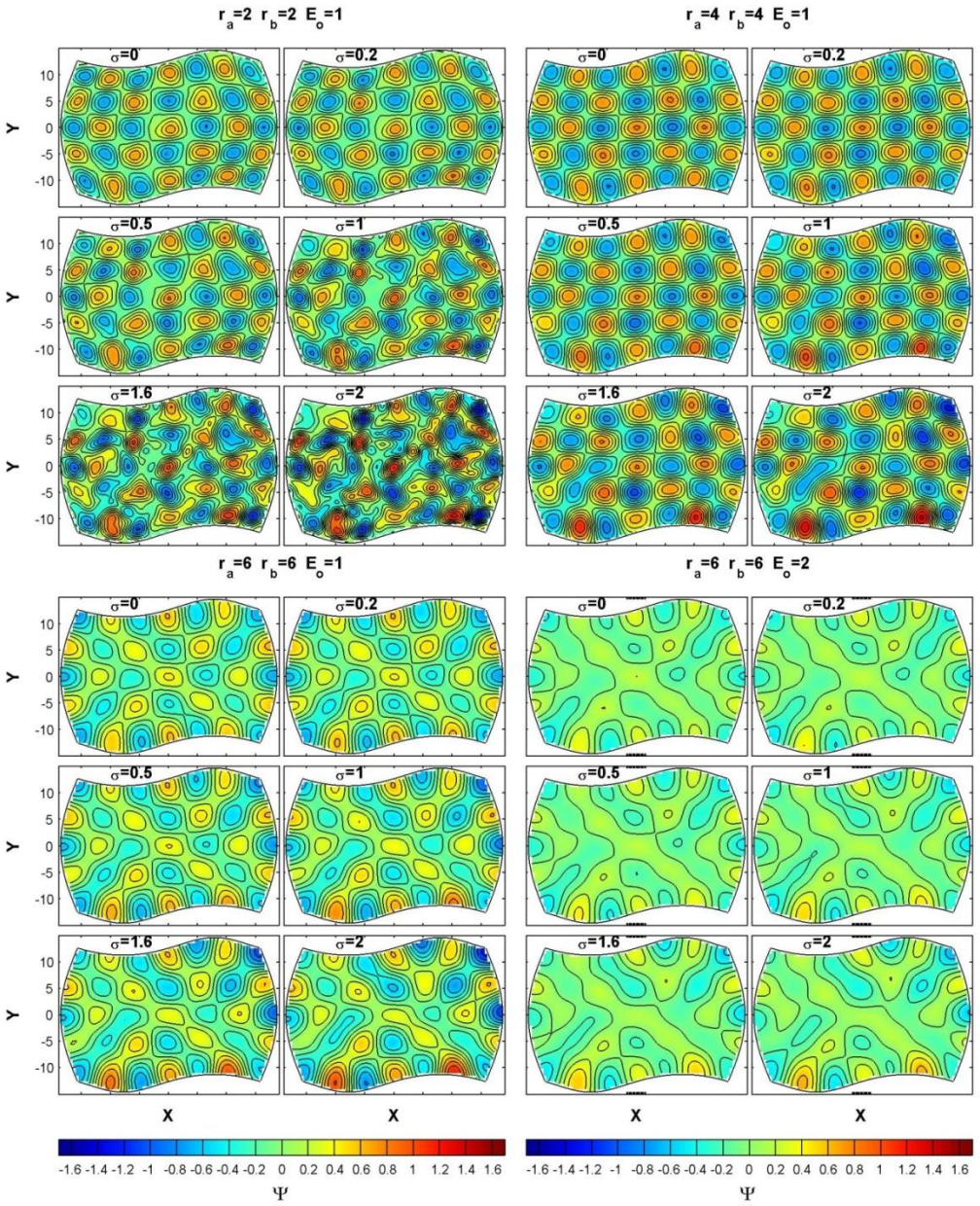


"Observational" data (ψ_o) from "truth" (small-scale) plus white Gaussian noises of zero mean and various standard deviations: (a) 0 (i.e., no noise) (b) 0.2, (c) 0.5, (d) 1.0, (e) 1.6, and (f) 2.0.



Analysis field ψ_a obtained by the OI – small eddies

$r_a = 2$
$r_b = 2$
$E_o = 1$



$r_a = 4$
$r_b = 4$
$E_o = 1$

$r_a = 6$
$r_b = 6$
$E_o = 1$

$r_a = 6$
$r_b = 6$
$E_o = 2$

Effectively using the ocean topographic characteristics



A new spectral ocean data assimilation method without requiring
a priori knowledge of matrices (B, R)

Optimal Spectral Decomposition (OSD) Method

Basis Functions

$$\nabla^2\phi_k = -\lambda_k\phi_{k,} \quad \left[b_1(\tau)\mathbf{n}\bullet\nabla\phi_k + b_2(\tau)\phi_k\right]|_\Gamma = 0, \quad k = 1,\ldots,\infty$$

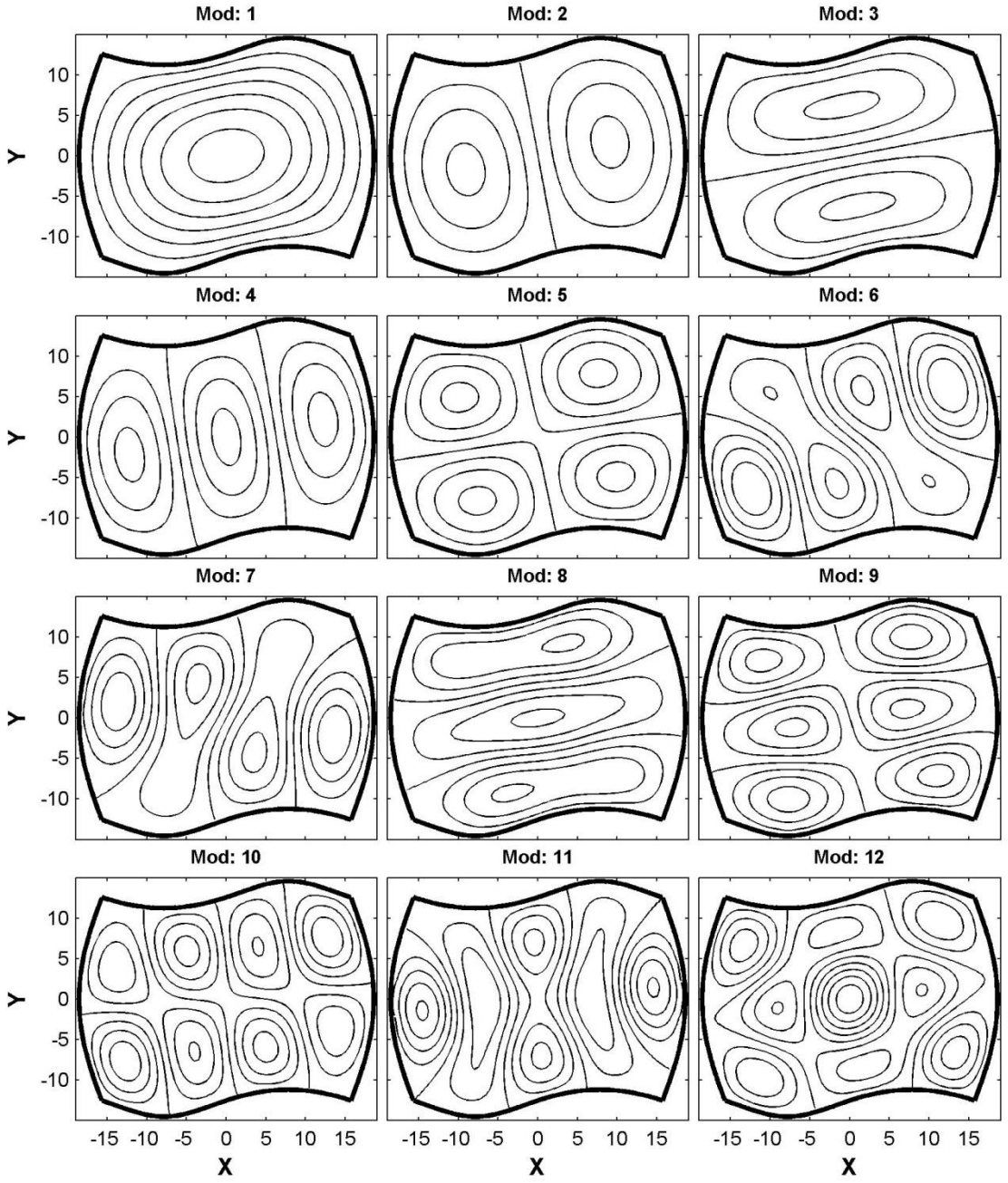
ϕ_k → The eigen functions of the 2D Laplacian Operator

satisfaction of the same homogeneous boundary condition of the assimilated variable anomaly

Φ Matrix →

$$\mathbf{\Phi} = \{\phi_{kn}\} = \begin{bmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \ldots & \phi_K(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \ldots & \phi_K(\mathbf{r}_2) \\ \ldots & \ldots & \ldots & \ldots \\ \phi(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \ldots & \phi_K(\mathbf{r}_N) \end{bmatrix}$$

First 12 Basis Functions



Spectral Ocean Data Assimilation

$$\mathbf{c}_a = \mathbf{c}_b + f_n \mathbf{s}^{(K)}, \quad s_K(\mathbf{r}_n) \equiv \sum_{k=1}^{K} a_k \, \phi_k(\mathbf{r}_n), \qquad f_n \equiv \sum_{m=1}^{M} h_{nm}$$

$\mathbf{H} = [h_{mn}]$ → the $M \times N$ linear observation operator matrix

$$\boldsymbol{\varepsilon}_a \equiv \mathbf{c}_a - \mathbf{c}_t = (\mathbf{c}_a - \mathbf{c}_b) + (\mathbf{c}_b - \mathbf{c}_t) = \boldsymbol{\varepsilon}_K + \boldsymbol{\varepsilon}_o$$

$$\boldsymbol{\varepsilon}_K \equiv \left[f_n \mathbf{s}^{(K)} - \mathbf{H}^T (\mathbf{c}_o - \mathbf{H}\mathbf{c}_b) \right], \qquad \boldsymbol{\varepsilon}_o \equiv \mathbf{H}^T \mathbf{c}_o - \mathbf{c}_t$$

$$\left\langle \boldsymbol{\varepsilon}_o^T \boldsymbol{\varepsilon}_K \right\rangle = 0$$

$$E^2 = \left\langle \boldsymbol{\varepsilon}_a^T \boldsymbol{\varepsilon}_a \right\rangle = E_K^2 + E_o^2, \quad E_K^2 \equiv \left\langle \boldsymbol{\varepsilon}_K^T \boldsymbol{\varepsilon}_K \right\rangle, \; E_o^2 \equiv \left\langle \boldsymbol{\varepsilon}_o^T \boldsymbol{\varepsilon}_o \right\rangle$$

OSD/OI (KF) Data Assimilation Equations

$$E^2 \rightarrow \min, \quad \partial E^2 / \partial a_k = \partial E_K^2 / \partial a_k = 0, \quad k = 1, \ldots, K$$

$$E_K^2 = \sum_{n=1}^{N} f_n \left[\left(\sum_{k=1}^{K} a_k \phi_{kn} - D_n \right)^2 \right] \rightarrow \min$$

$$\sum_{k'=1}^{K} \sum_{n=1}^{N} \left(\phi_{kn} f_n \phi_{nk'} \right) a_{k'} = \sum_{n=1}^{N} \phi_{kn} f_n D_n, \qquad k = 1, 2, \ldots, K$$

$$\mathbf{\Phi F \Phi}^T \mathbf{A} = \mathbf{\Phi F D}, \qquad \mathbf{A} = \left[\mathbf{\Phi F \Phi}^T \right]^{-1} \mathbf{\Phi F D}$$

OSD → $$\mathbf{c}_a = \mathbf{c}_b + \mathbf{F \Phi}^T \left[\mathbf{\Phi F \Phi}^T \right]^{-1} \mathbf{\Phi H}^T \mathbf{d}$$

OI/KF → $$\mathbf{c}_a = \mathbf{c}_b + \mathbf{B H}^T (\mathbf{H B H}^T + \mathbf{R})^{-1} \mathbf{d}$$

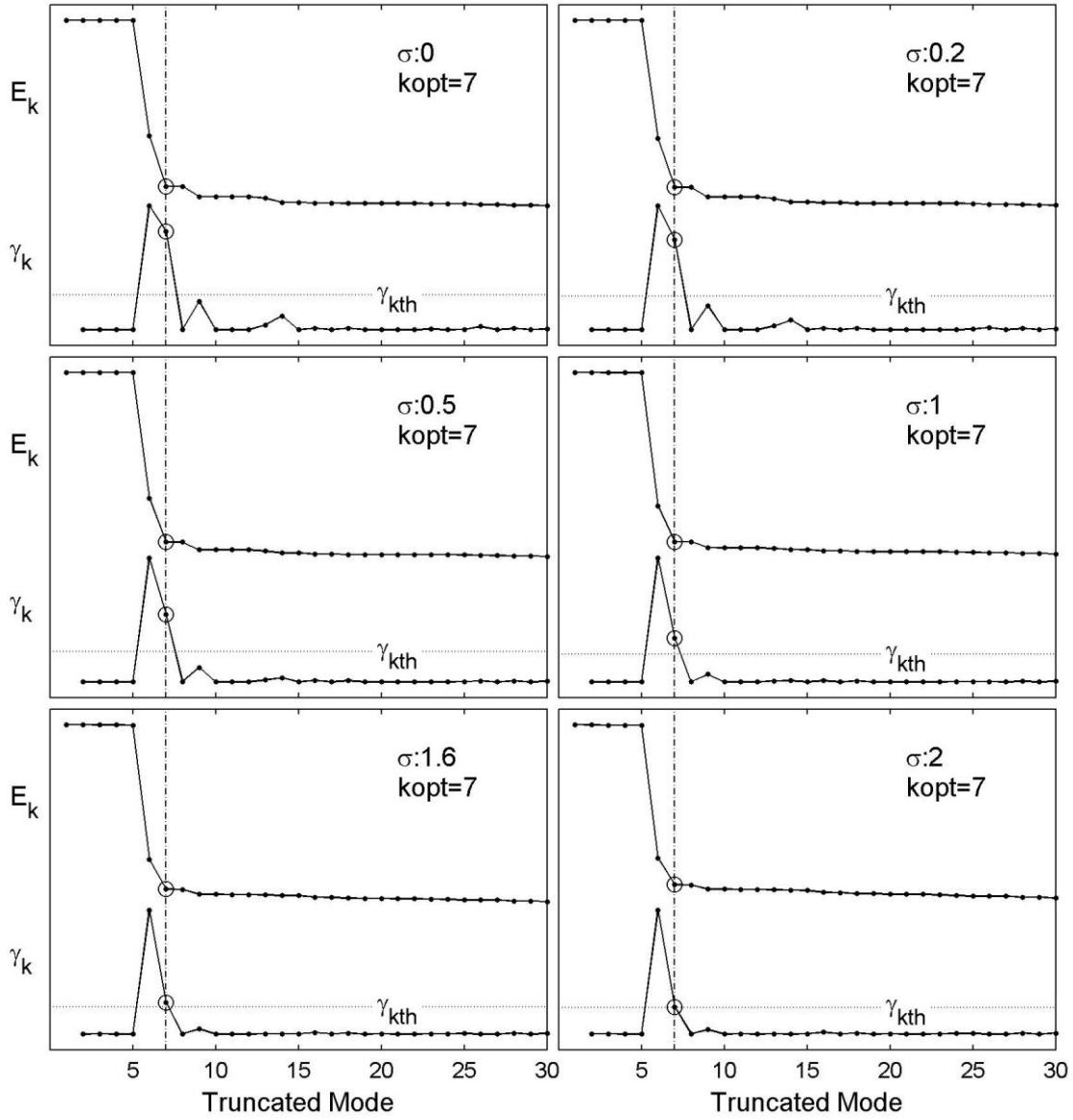
Two Issues in OSD

(1) Determine the optimal truncation using the steep-descending method (Chu and Fan 2015)

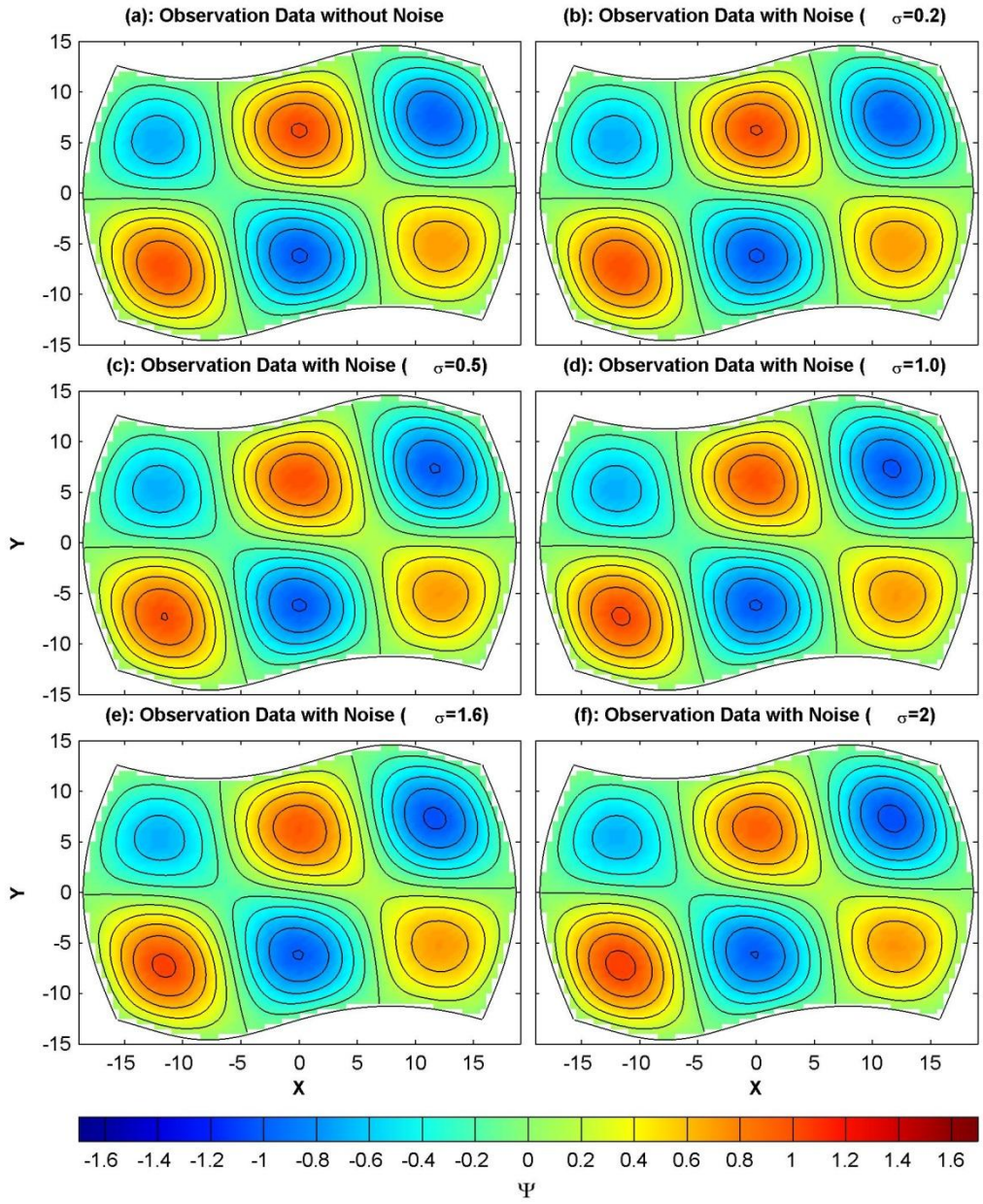
K = Kopt

(2) Obtaining the spectral coefficients at given K from observational data

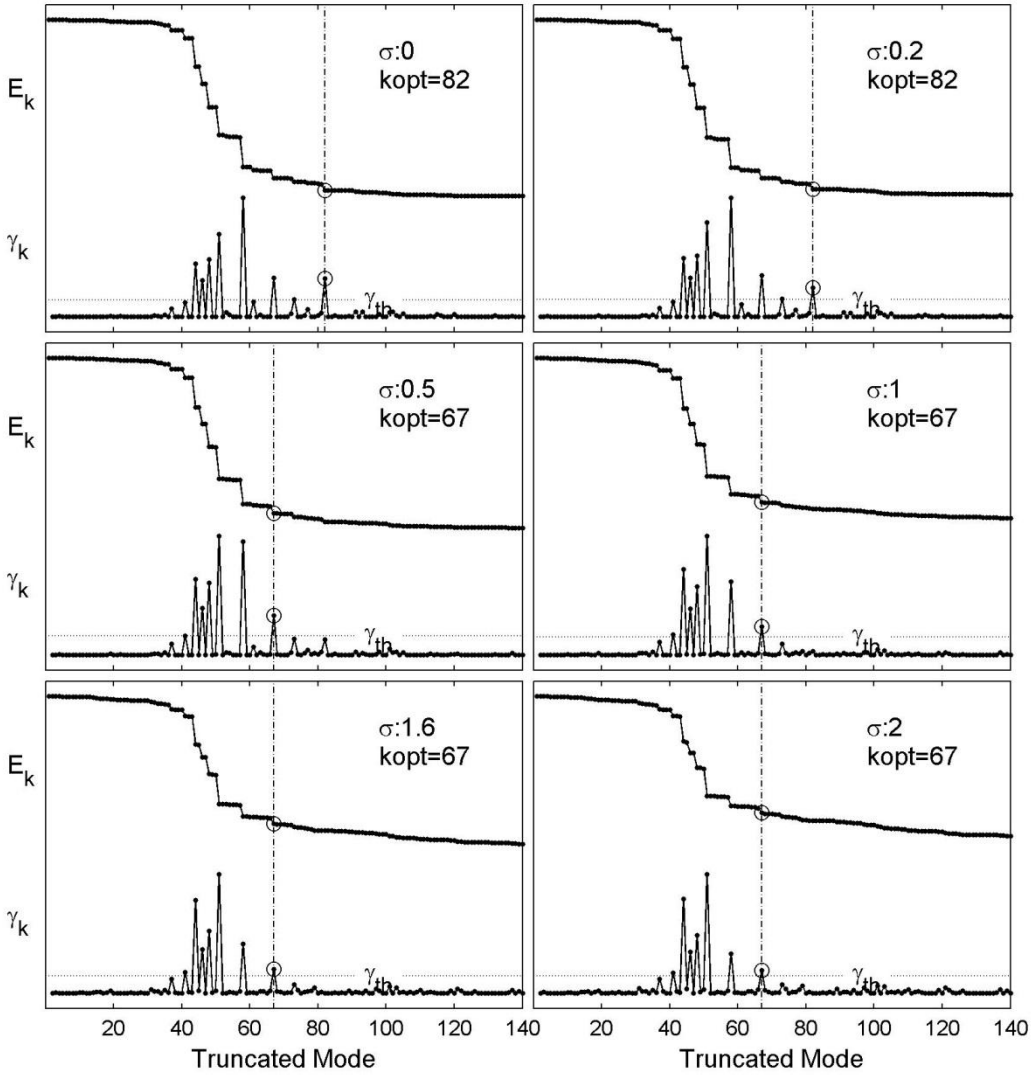
Mode truncation for large-scale “observational data”



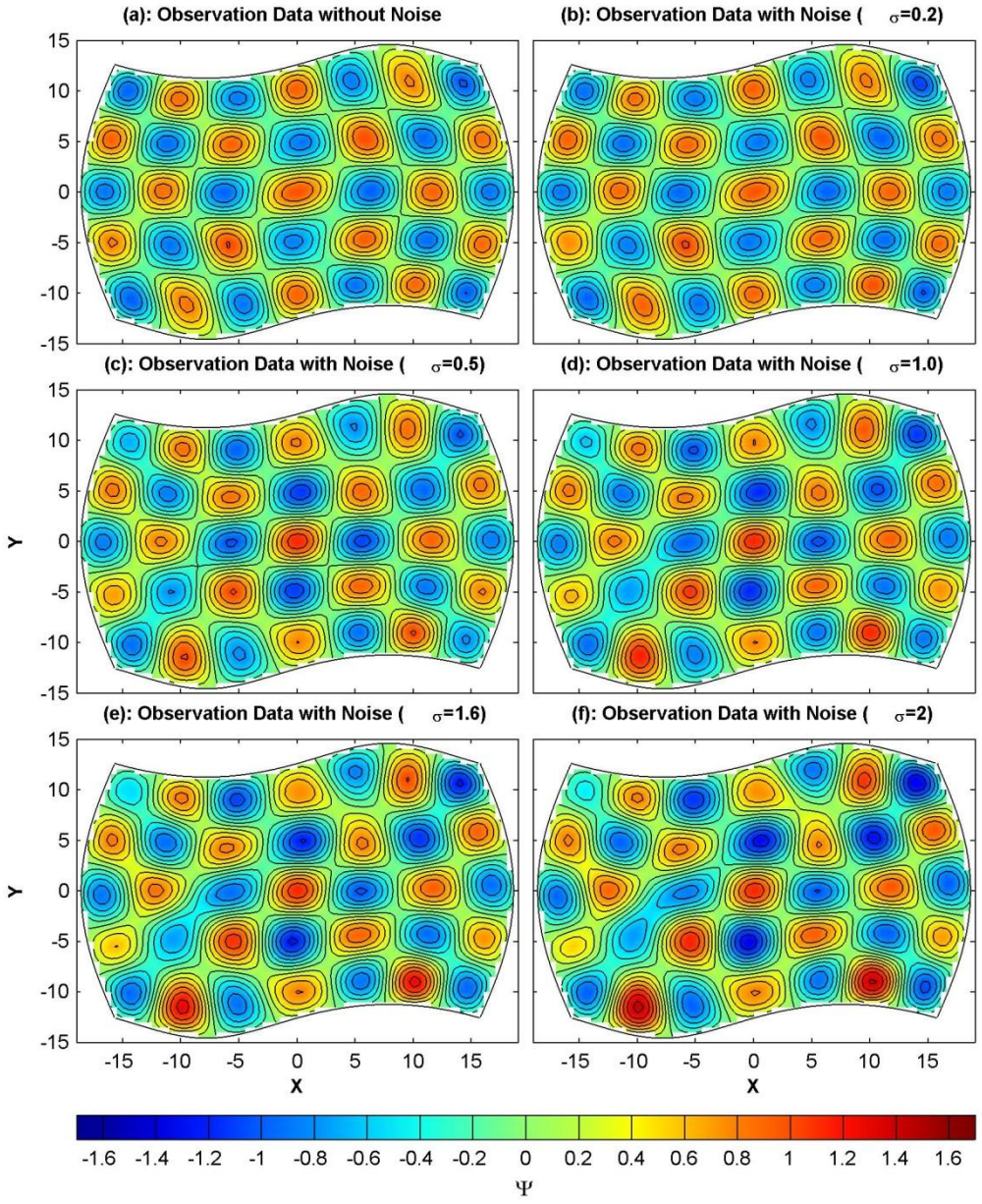
Analysis field ψ_a obtained by the OSD Method



Mode truncation for small-scale "observational data"



Analysis field ψ_a obtained by the OSD



Error Comparison Between OI and OSD

$$R^{OSD}(\sigma) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left[\psi_a^{OSD}(\mathbf{r}_n,\sigma) - \psi_t(\mathbf{r}_n)\right]^2}$$

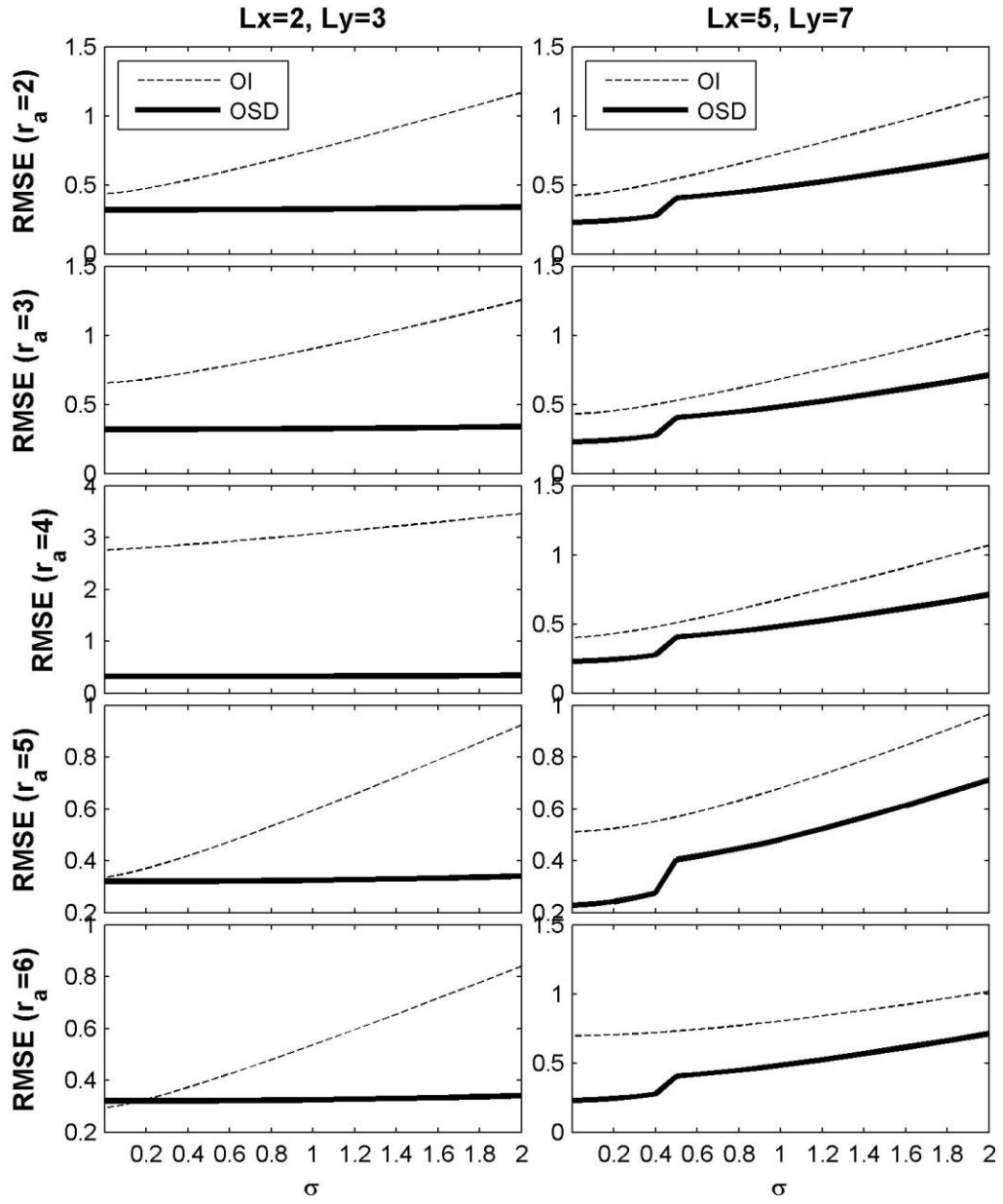
$$R^{OI}(\sigma, r_a) = \sqrt{\frac{1}{NP_bP_E}\sum_{r_b}\sum_{E_o}\sum_{n=1}^{N}\left[\psi_a^{OI}(\mathbf{r}_n,\sigma,r_a,r_b,E_o) - \psi_t(\mathbf{r}_n)\right]^2}$$

$$R^{OI}(\sigma, r_b) = \sqrt{\frac{1}{NP_aP_E}\sum_{r_a}\sum_{E_o}\sum_{n=1}^{N}\left[\psi_a^{OI}(\mathbf{r}_n,\sigma,r_a,r_b,E_o) - \psi_t(\mathbf{r}_n)\right]^2}$$

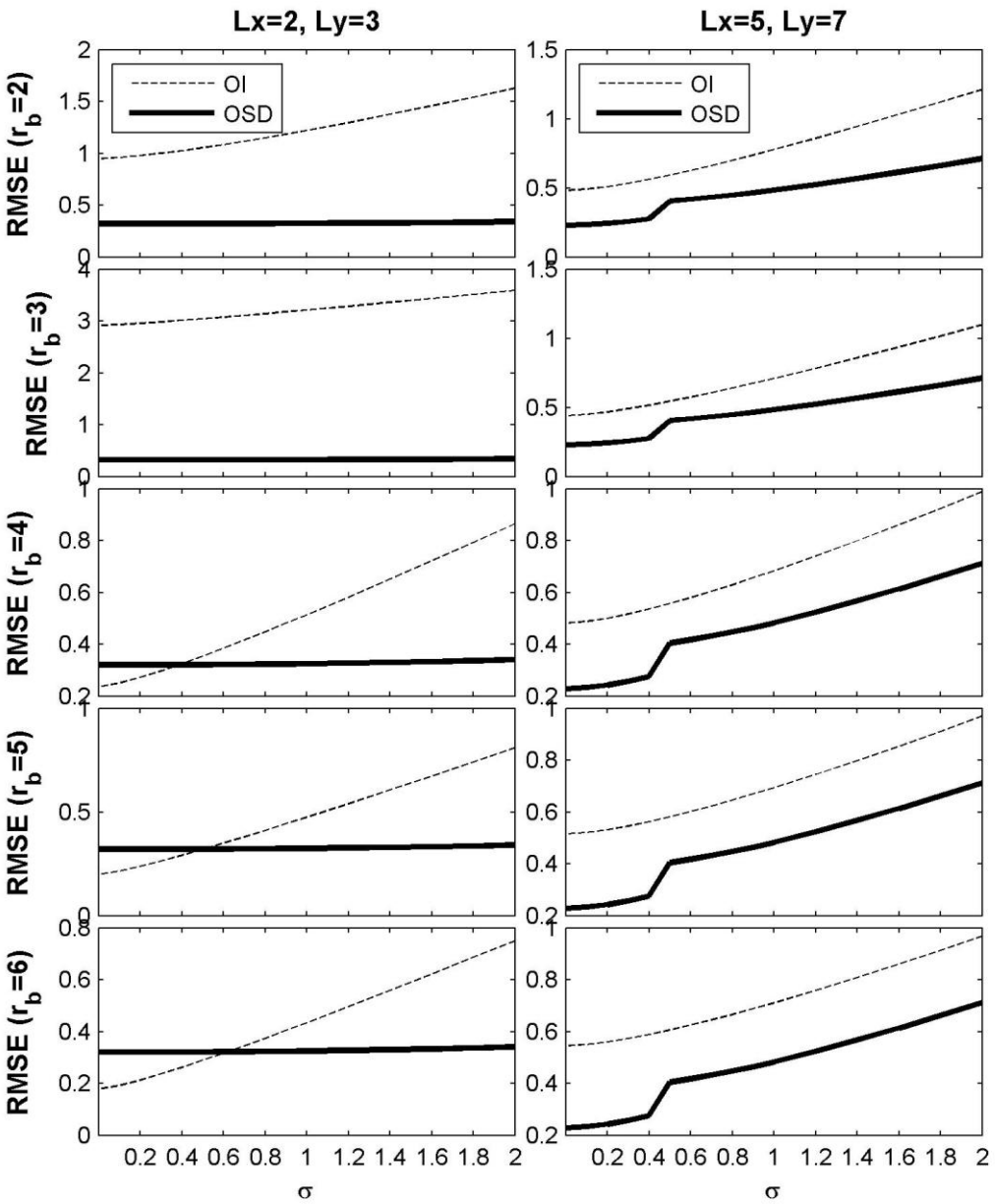
$$R^{OI}(\sigma, E_o) = \sqrt{\frac{1}{NP_aP_b}\sum_{r_b}\sum_{r_b}\sum_{n=1}^{N}\left[\psi_a^{OI}(\mathbf{r}_n,\sigma,r_a,r_b,E_o) - \psi_t(\mathbf{r}_n)\right]^2}$$

$(P_a, P_b, P_E) \rightarrow$ total numbers of (r_a, r_b, E_o)

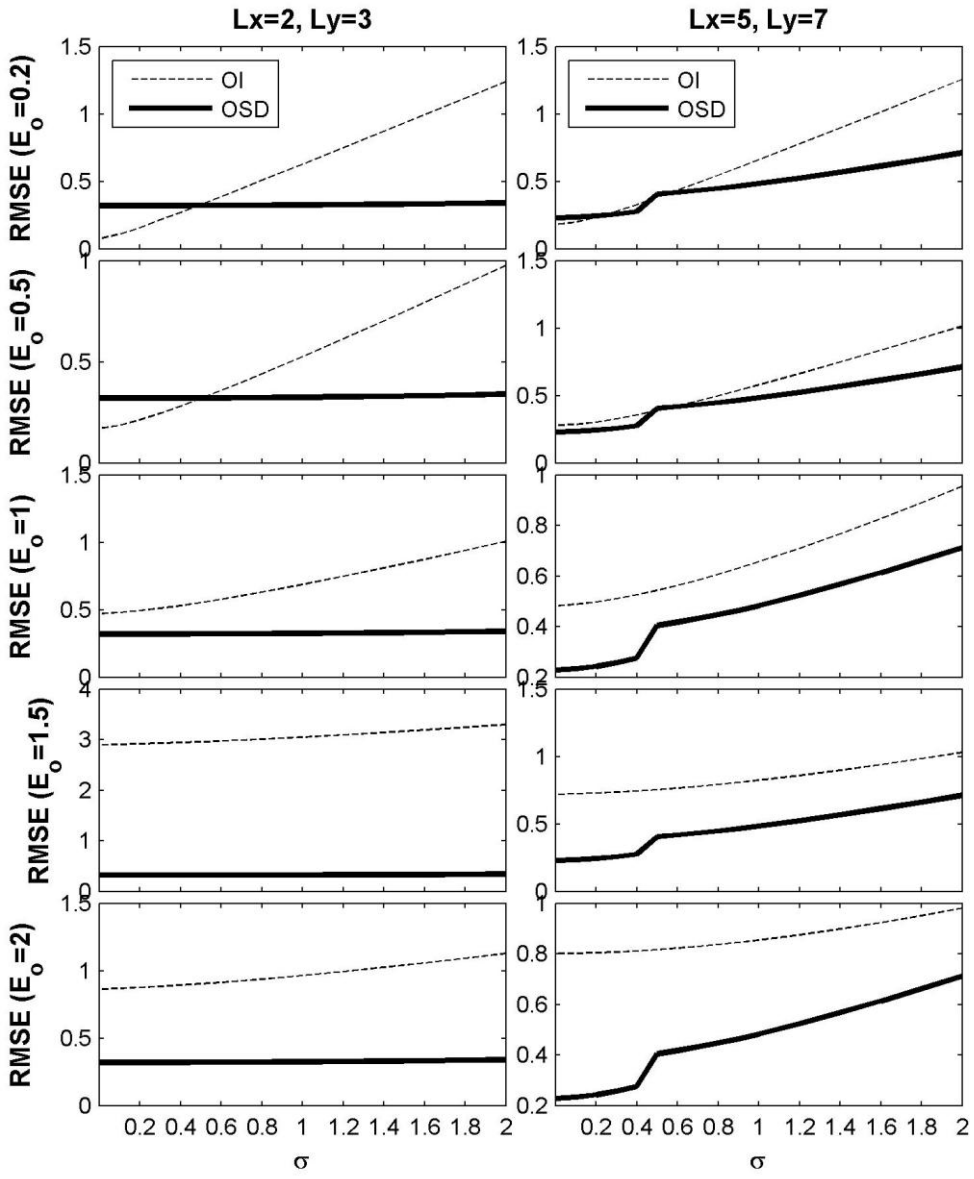
Comparison between $R^{OI}(\sigma, r_a)$ and $R^{OSD}(\sigma)$



Comparison between $R^{OI}(\sigma, r_b)$ and $R^{OSD}(\sigma)$



Comparison between $R^{OI}(\sigma, E_o)$ and $R^{OSD}(\sigma)$



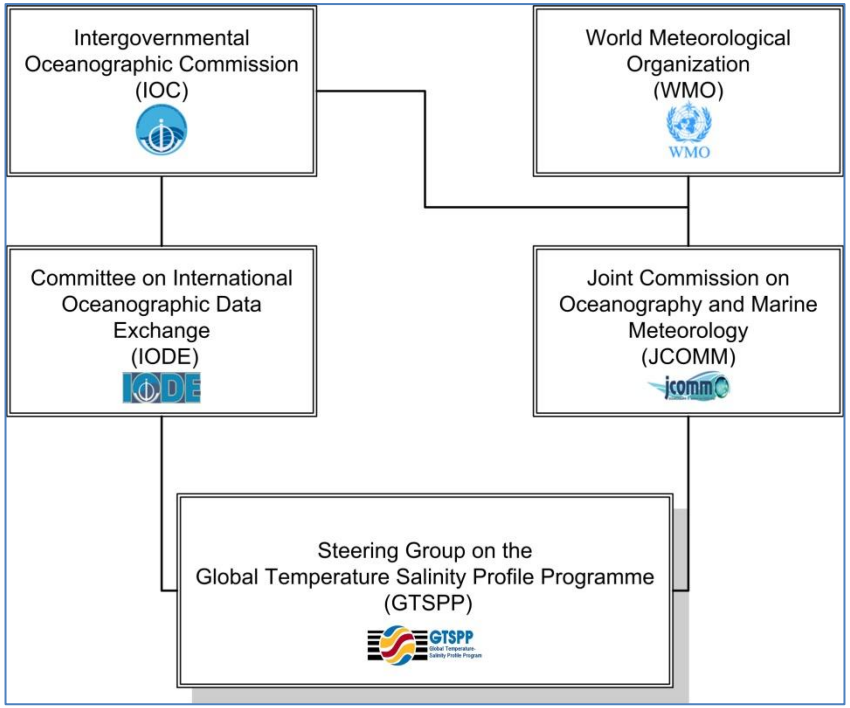
Part 2

Monthly Varying Gridded (T, S) Fields from GTSPP

What's GTSPP

GTSPP stands for Global Temperature and Salinity Profile Programme

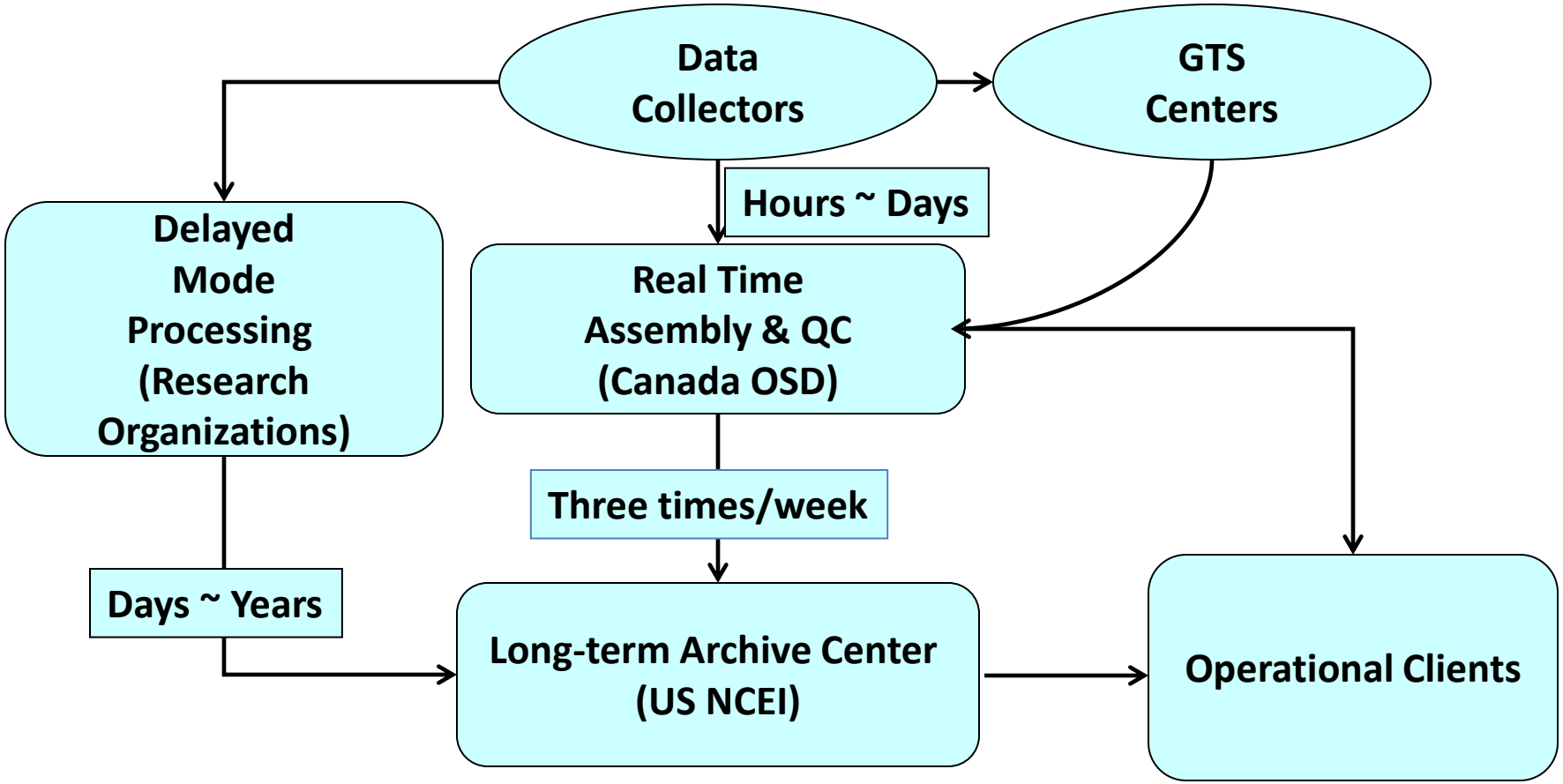
GTSPP is a joint WMO-IOC program designed to provide improved access to the highest resolution, highest quality data as quickly as possible.



GTSPP Infrastructure

1) Global Telecommunication Systems
2) IODE Data Centers
3) Long-Term Archive Center
4) Data Product Centers
5) The Users

GTSPP Data Flows



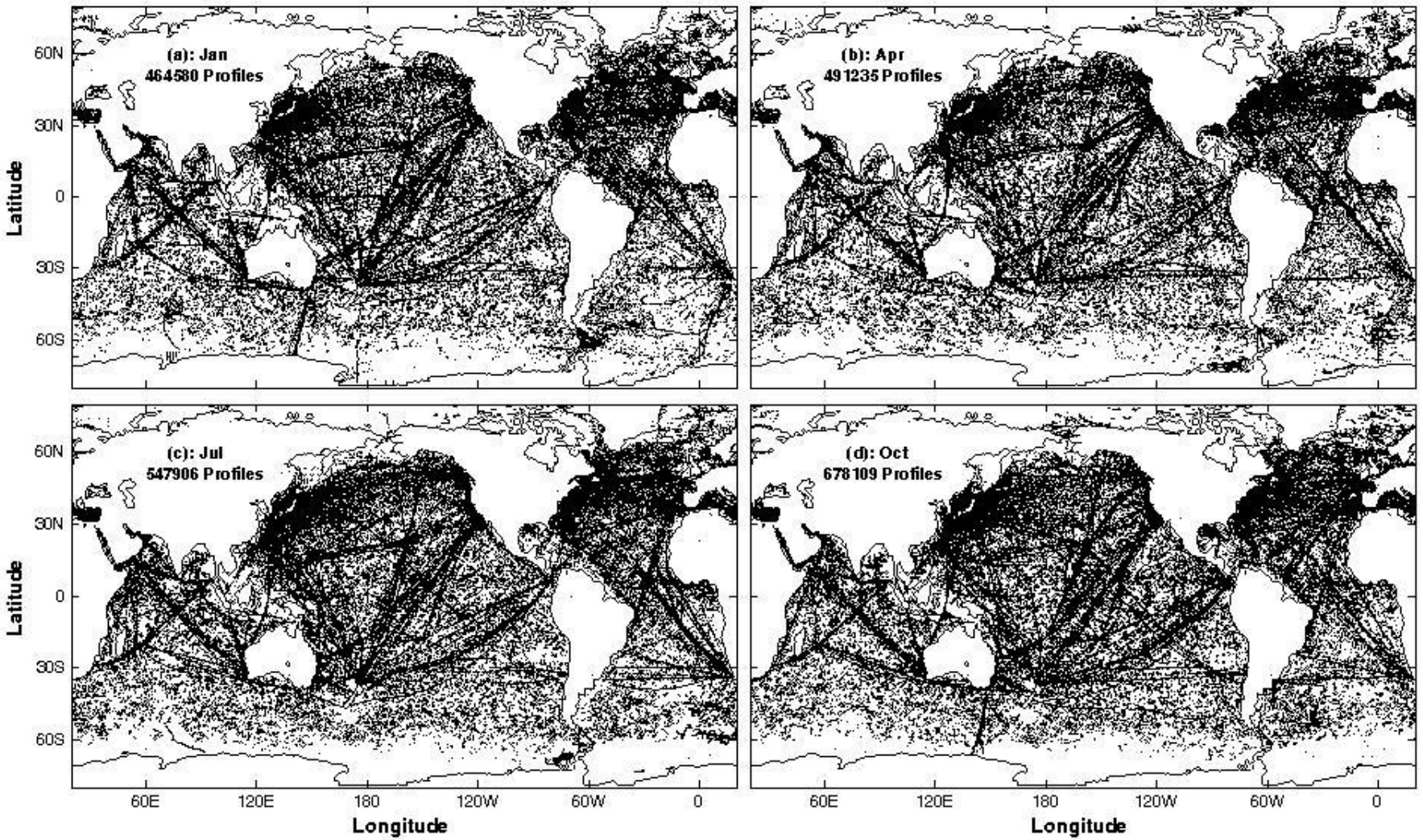
GTSPP Main Web Site (http://www.nodc.noaa.gov/GTSPP/)

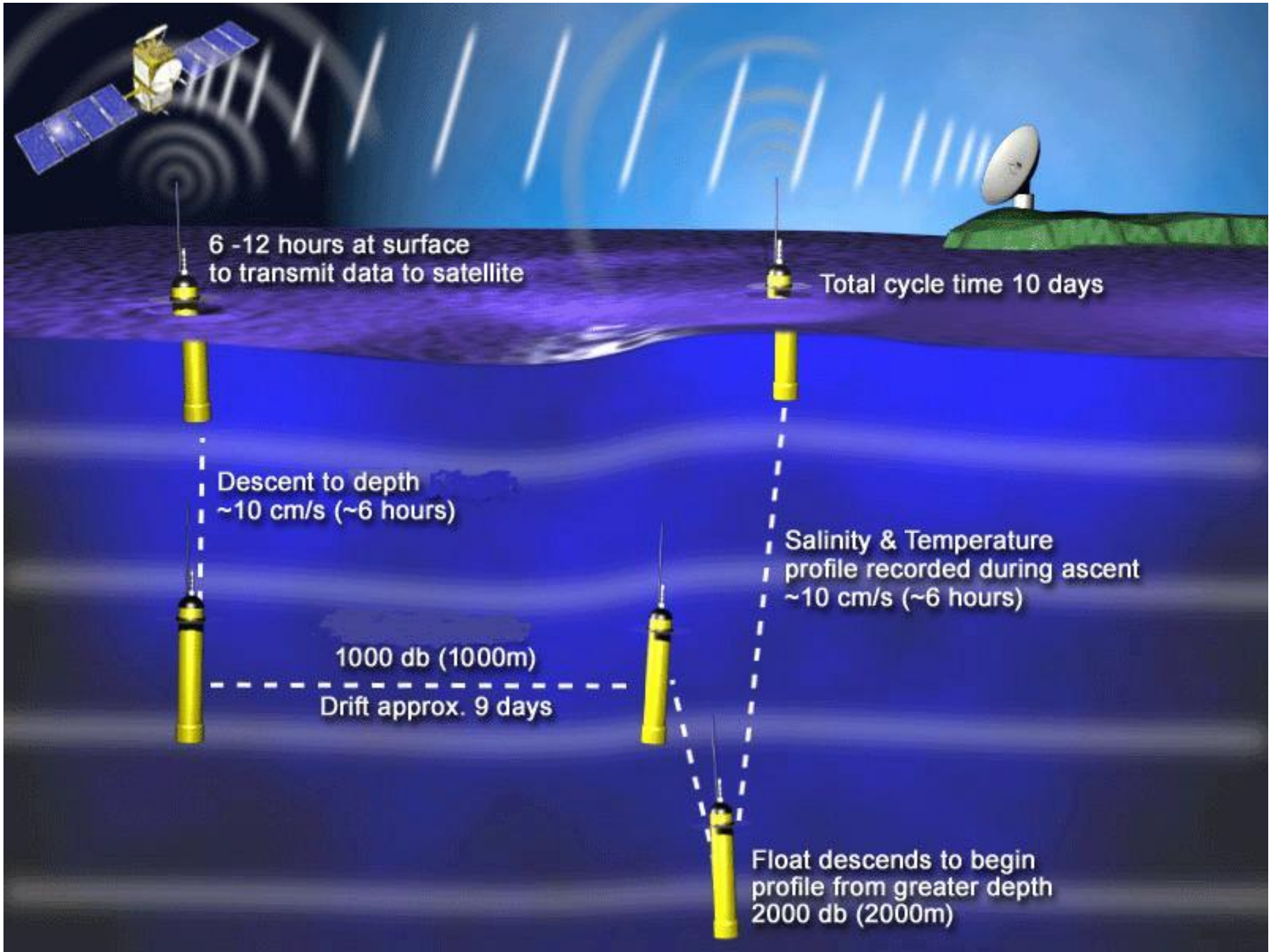


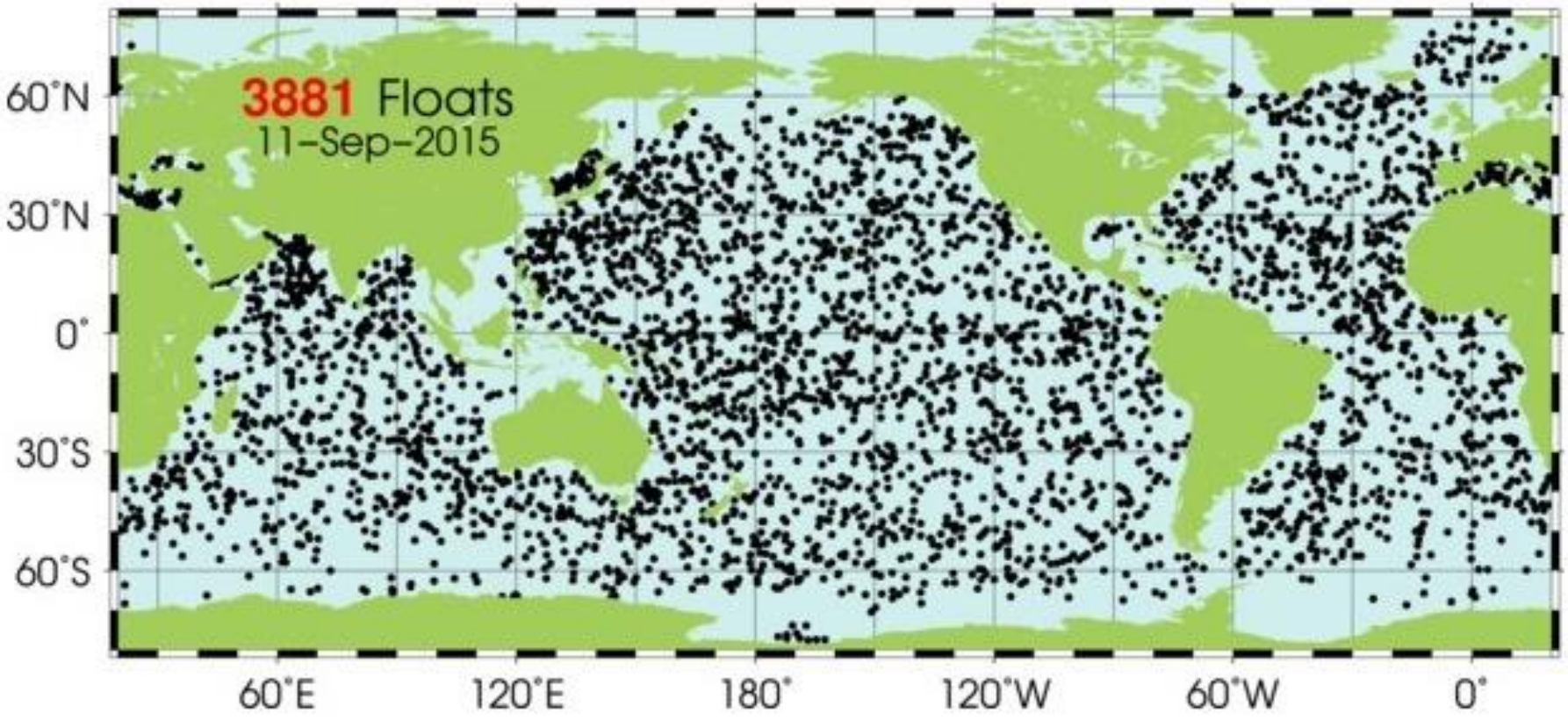
GTSPP Online Data Delivery Servers

1) HTTP (Web) Server:
   http://data.nodc.noaa.gov/gtspp/
2) FTP Server:
   ftp://ftp.nodc.noaa.gov/pub/gtspp/
3) OPeNDAP Server:
   http://data.nodc.noaa.gov/opendap/gtspp/
4) THREDDS Server:
   http://data.nodc.noaa.gov/thredds/catclog/gtspp/

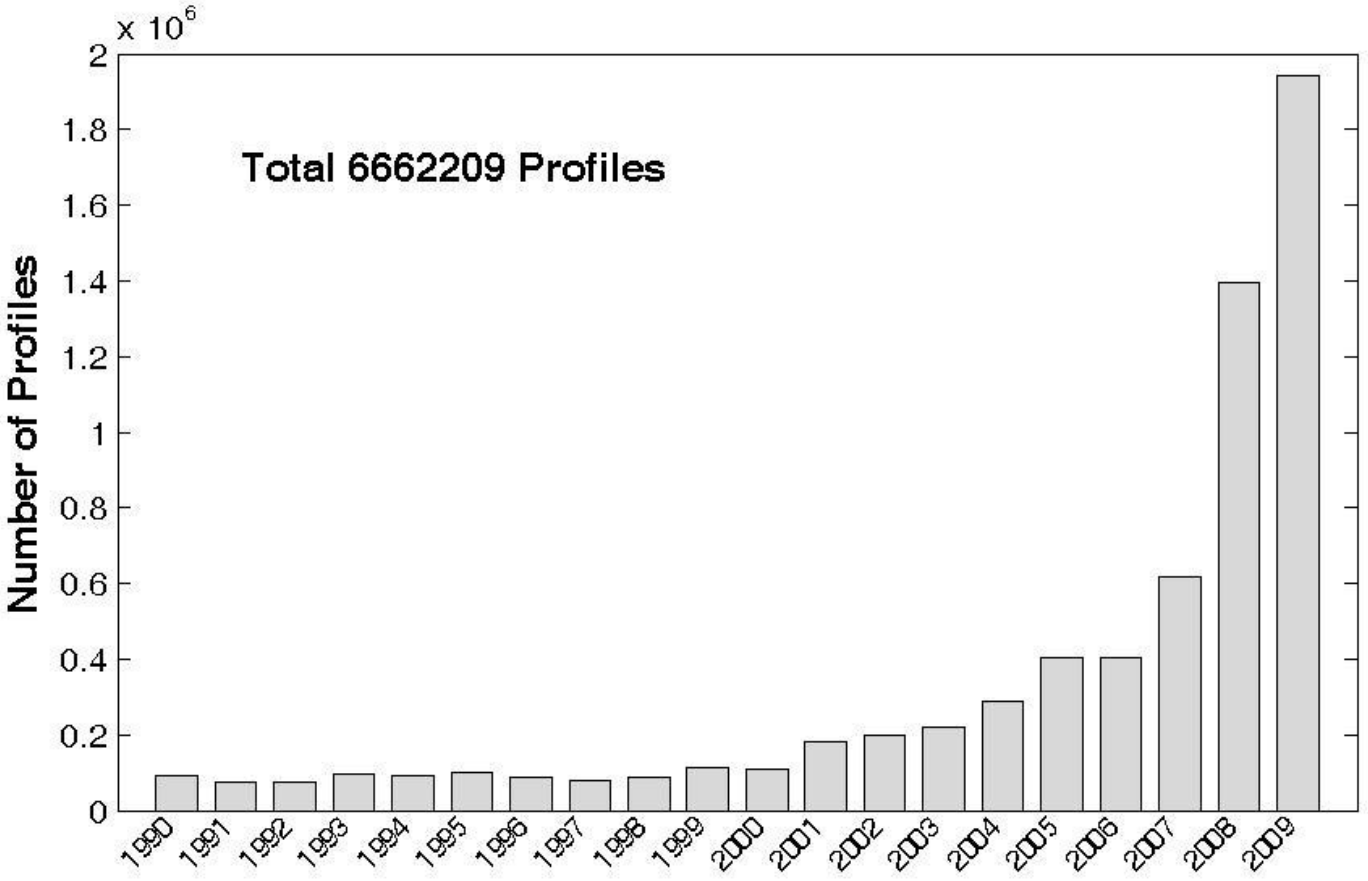
Spatial Distribution of GTSPP Data




6 -12 hours at surface
to transmit data to satellite
Total cycle time 10 days
Descent to depth
~10 cm/s (~6 hours)
Salinity & Temperature
profile recorded during ascent
~10 cm/s (~6 hours)
1000 db (1000m)
Drift approx. 9 days
Float descends to begin
profile from greater depth
2000 db (2000m)


3881 Floats
11-Sep-2015
60°N
30°N
0°
30°S
60°S
60°E
120°E
180°
120°W
60°W
0°

Temporal Distribution of GTSPP Data



Pre-Processing Data → File Merging

One Profile → One File

Total 6,662,209 files



Global Profiles in a Month within a Year → One File

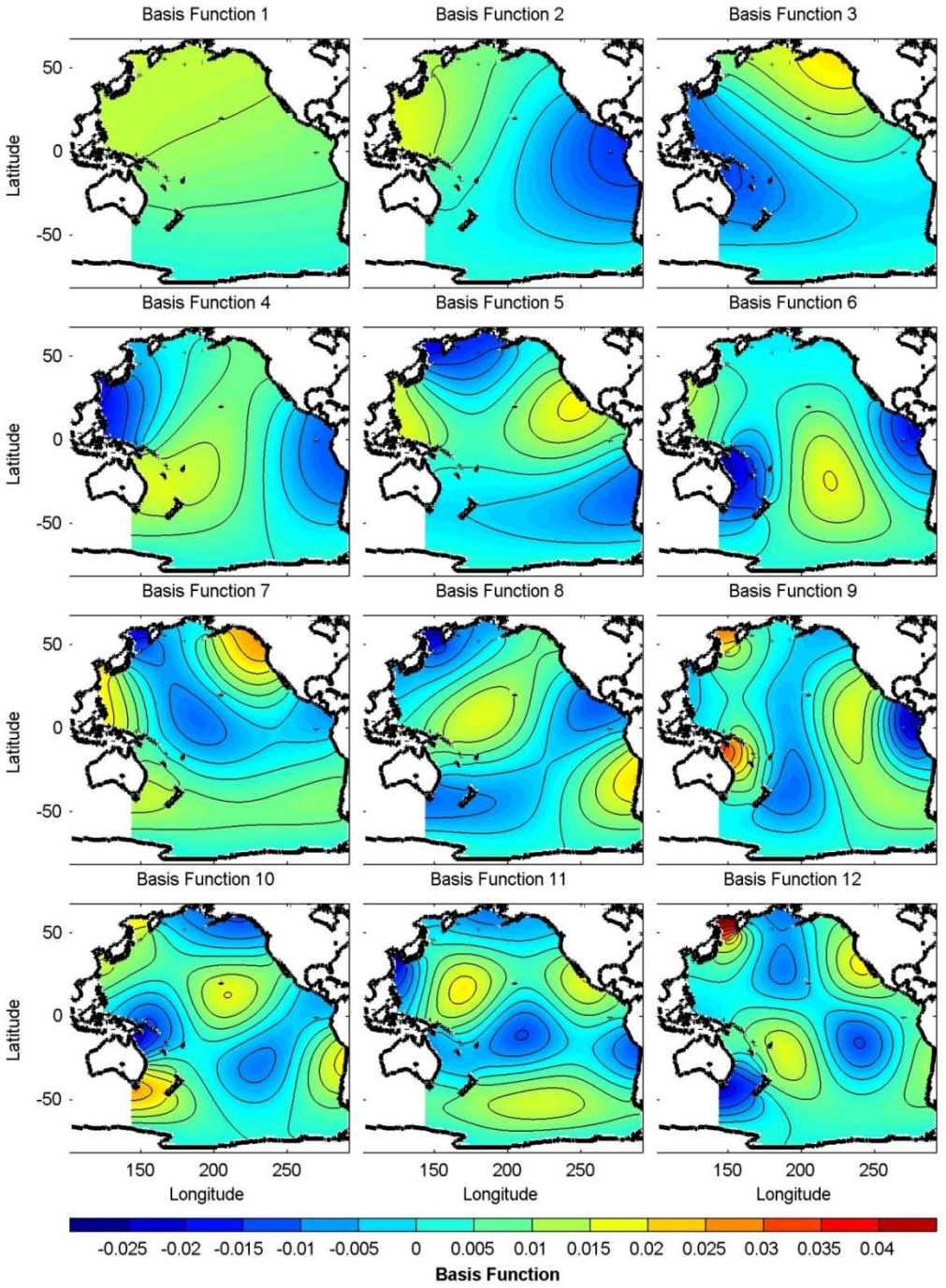
Total 240 (20X12) Files

First 12 basis functions for the Pacific Ocean at the surface.

DBDB5

Dirichlet boundary condition at the southern boundary (Antarctic),

Newmann boundary condition elsewhere



Monthly Gridded Temperature at 10 m in the Atlantic Ocean



Monthly Gridded Temperature at 200 m in the Atlantic Ocean



Monthly Gridded Temperature at 1000 m in the Atlantic Ocean



Monthly Gridded Temperature at 10 m in the Pacific Ocean



Monthly Gridded Temperature at 200 m in the Pacific Ocean



Monthly Gridded Temperature at 1000 m in the Pacific Ocean



Part 3
Upper Ocean Heat Content and Climate Change

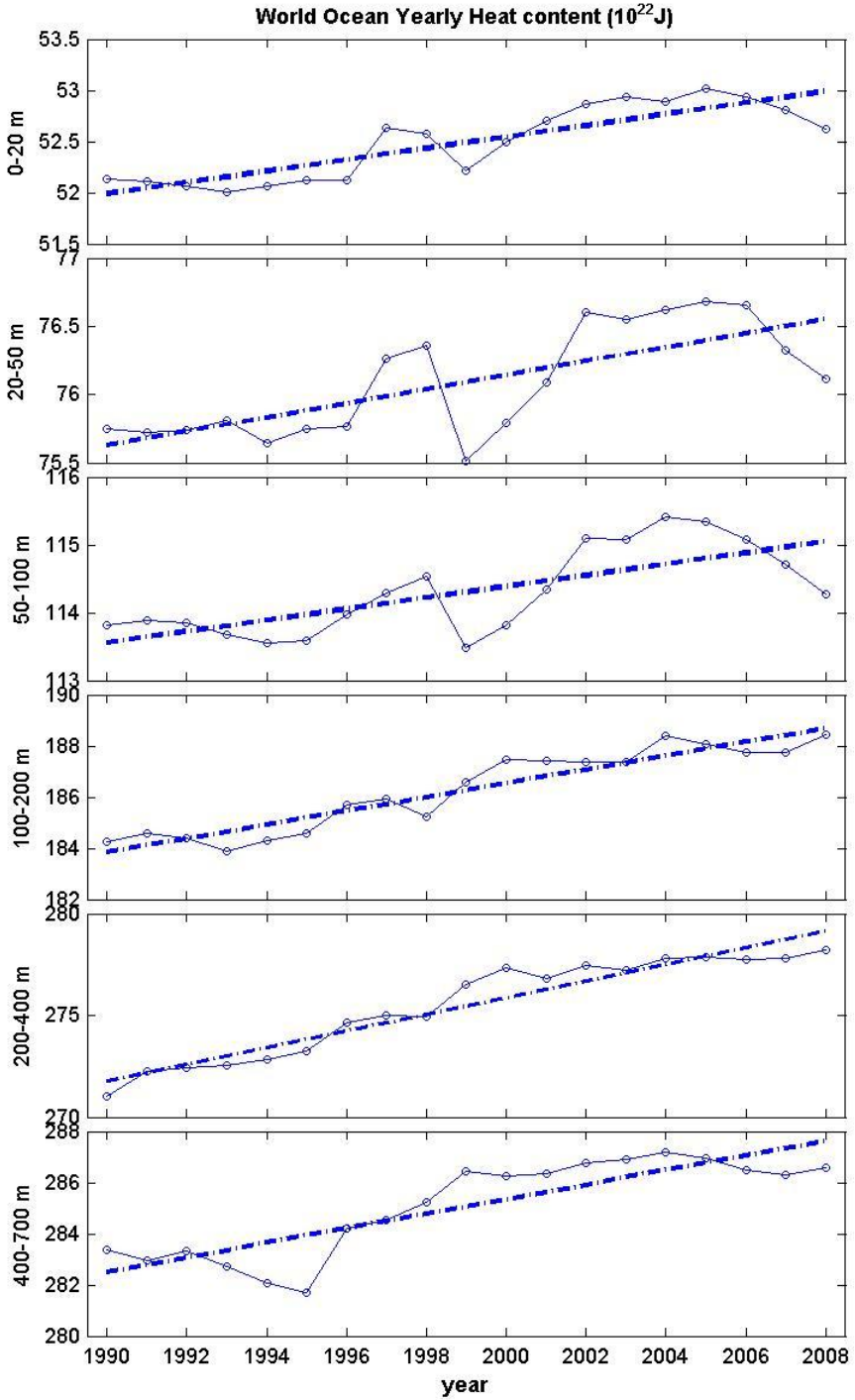
Upper Ocean (0-300 m) Heat Content

$$HC = \int_{-h}^{0} \rho c T dz$$

$HC = HC_{\text{mean}} + HC_{\text{seasonal}} + HC_{\text{anomaly}}$

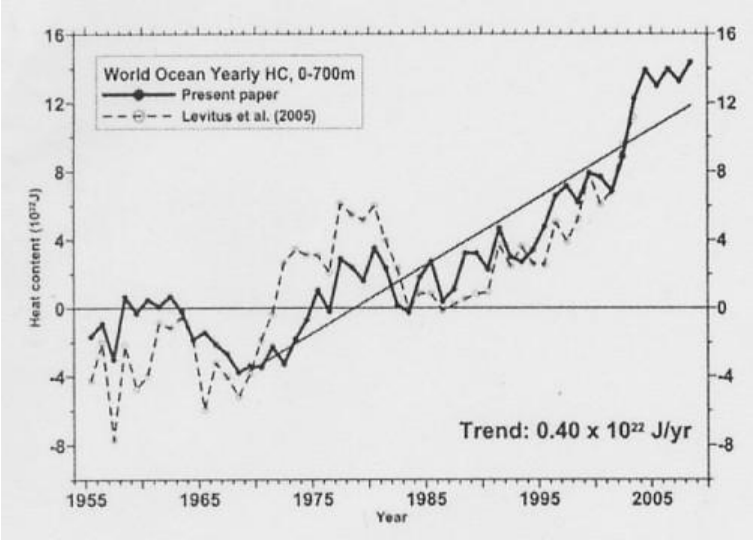
EOF Analysis → HC_{anomaly}

→ Global Ocean Dipole Modes


World Ocean Yearly Heat content (10^{22}J)
0-20 m
20-50 m
50-100 m
100-200 m
200-400 m
400-700 m
53.5
53
52.5
52
51.5
77
76.5
76
75.5
116
115
114
113
190
188
186
184
182
280
275
270
288
286
284
282
280
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
year

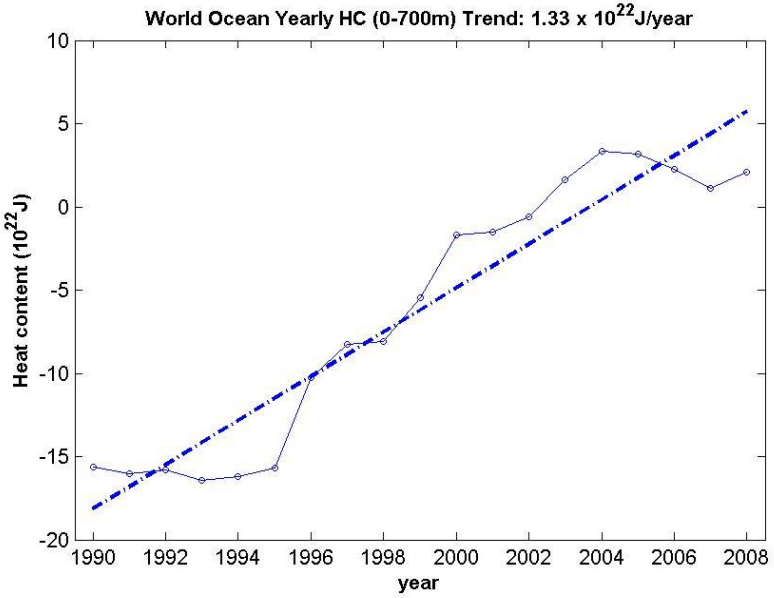
Trend of Upper Ocean (0-700 m) Heat Content

0.4 X 10^{22} J/yr
(1958-2008)
(Levitus et al.,GRL, 2009)
Without Argo data

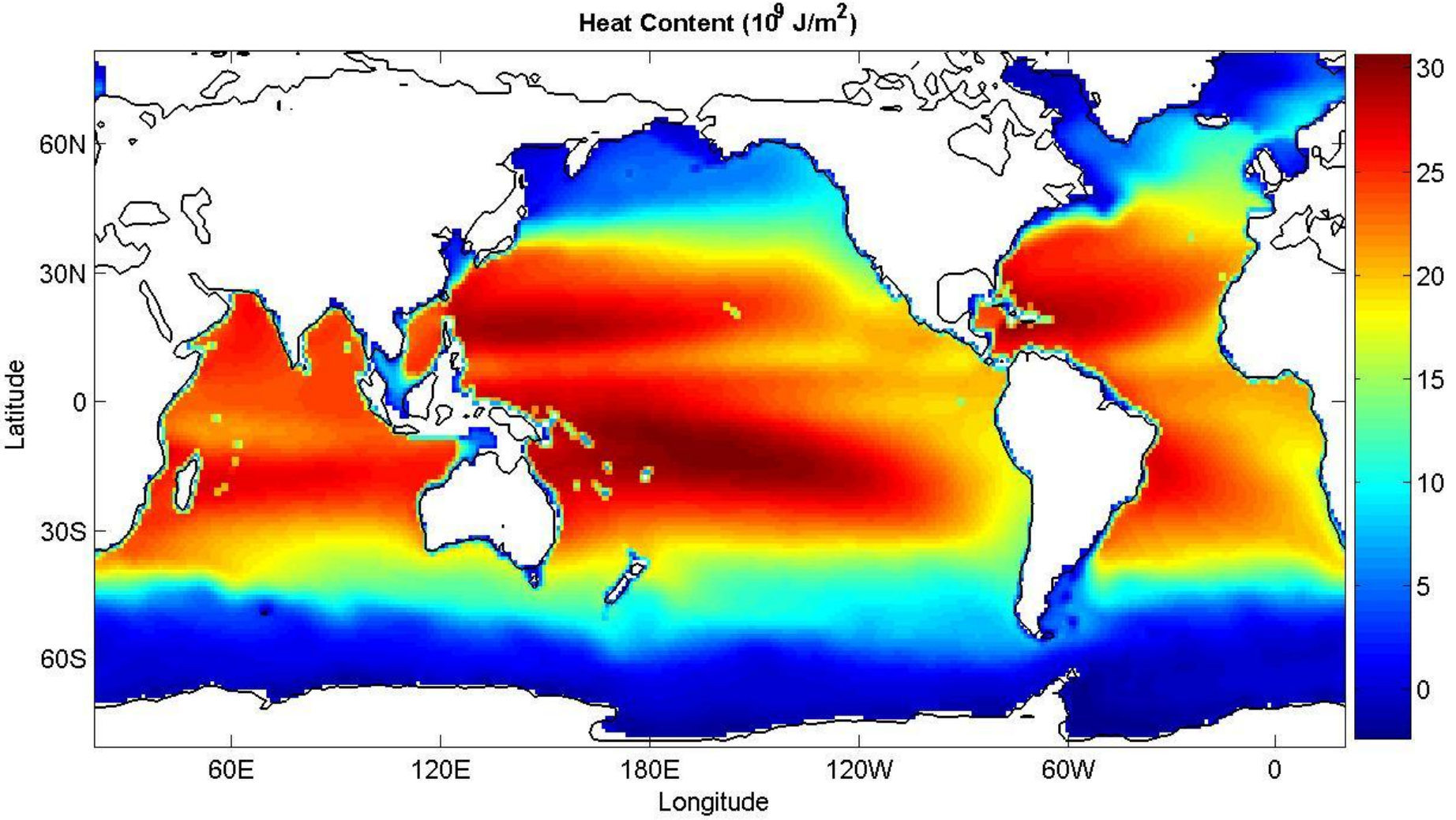


1.3 X 10^{22} J/yr
(1990-2008)

With Argo data

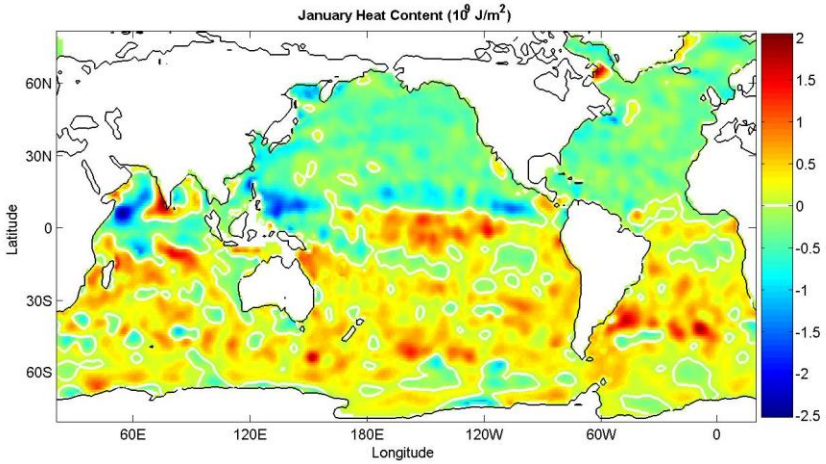


Upper Ocean (0-300 m) Mean Heat Content (J/m^2) (1990-2009)

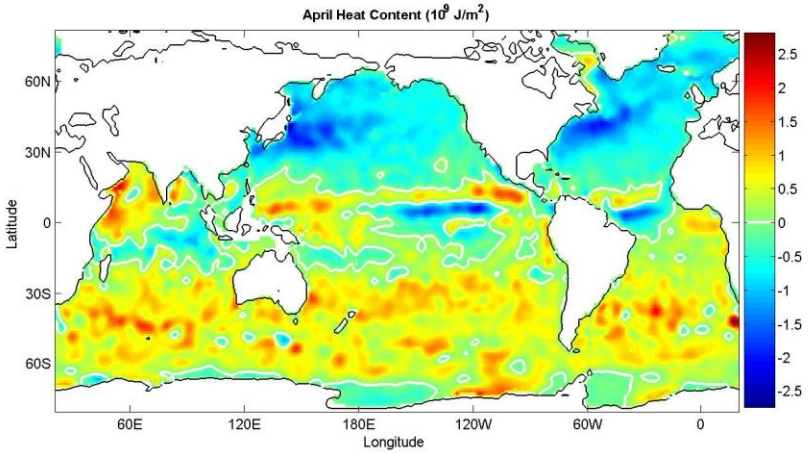


Seasonal Variability of Upper Ocean (0-300 m) Heat Content (J/m^2) (1990-2009)

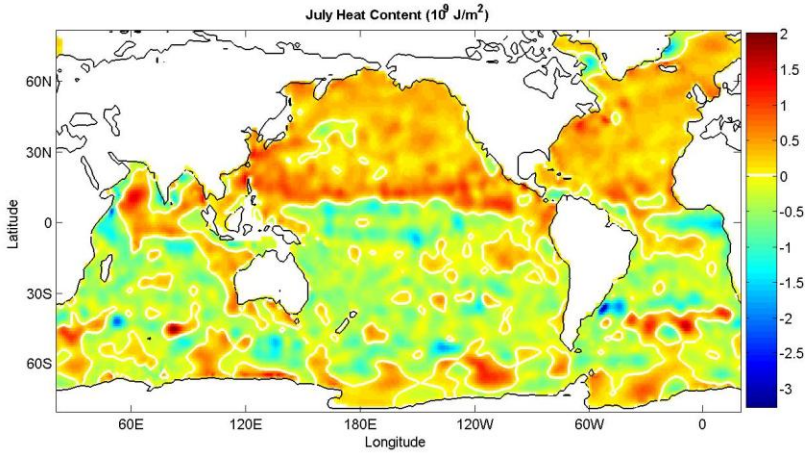
January



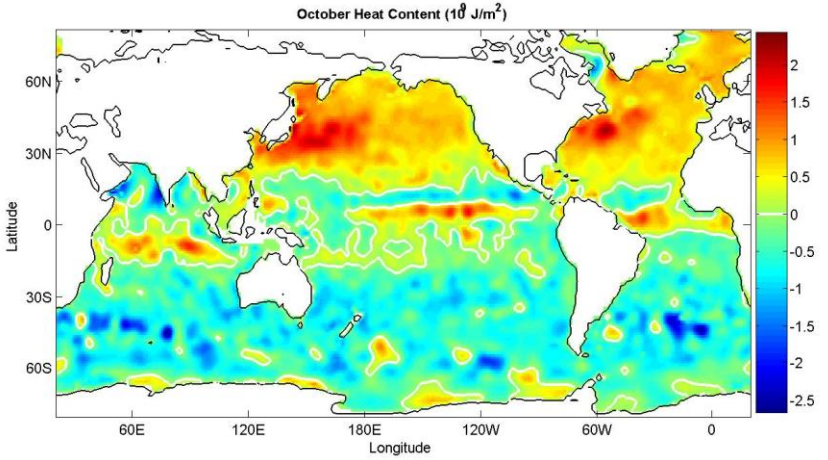
April



July



October



EOF Analysis →

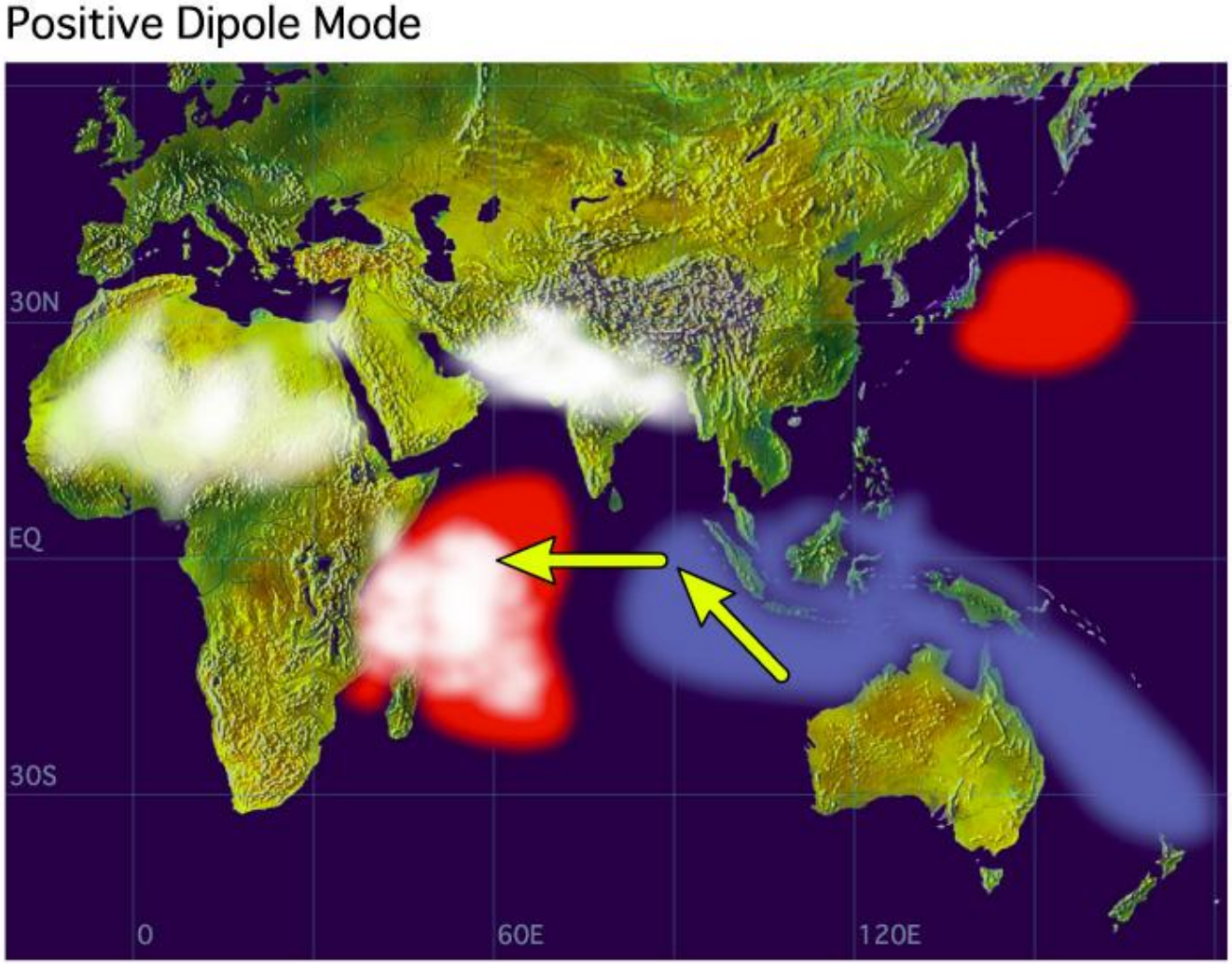
Heat Content Anomaly Relative to Seasonal Variation

→ Inter-annual Variability

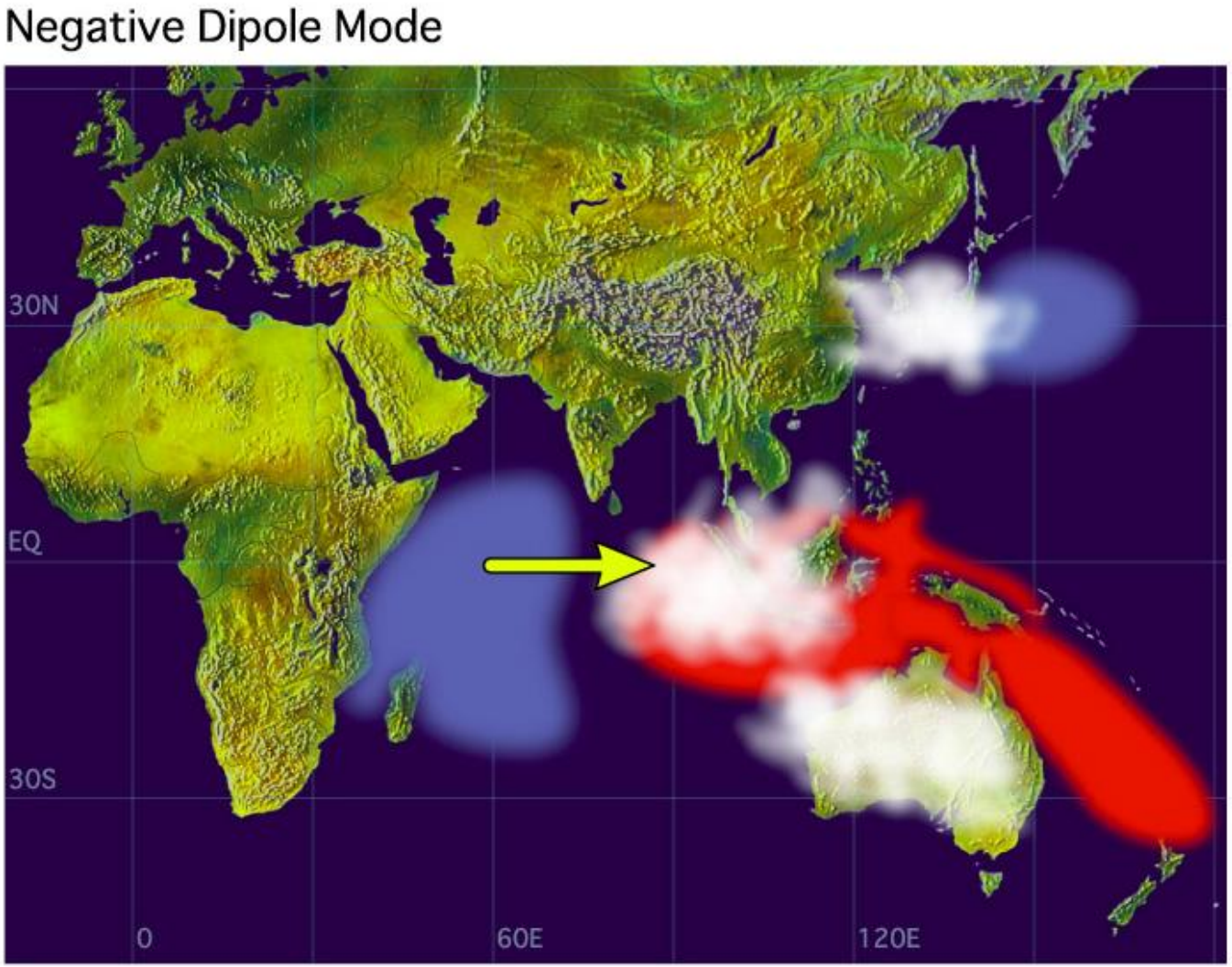
(1) Development in Climate Variability

Indian Ocean Dipole
Canonical (Eastern Pacific) El Nino
El Nino Modoki (Central Pacific El Nino)

Indian Ocean Dipole (Saji et al. 1999)



Indian Ocean Dipole (Saji et al. 1999)



Dipole Mode Index (DMI)

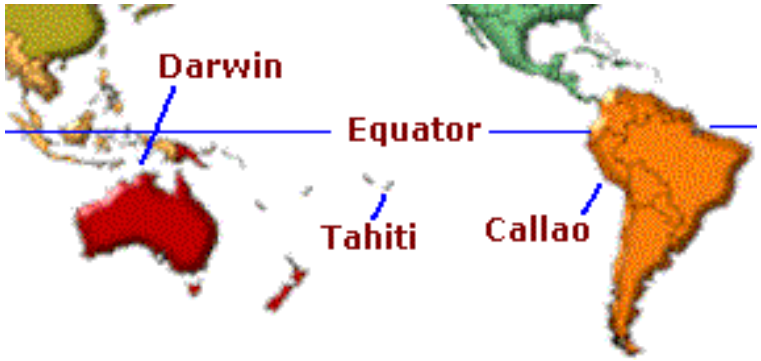
$$DMI = SSTA_{West} - SSTA_{East}$$

West = 50°E-70°E;10°S-10°N

East = 90°E-110°E;10° S-0°

Southern Oscillation Index (SOI)

$$s(\tilde{t}_p) = 10 \times \frac{p_{diff}(\tilde{t}_p) - \langle p_{diff} \rangle}{SD(p_{diff})}$$



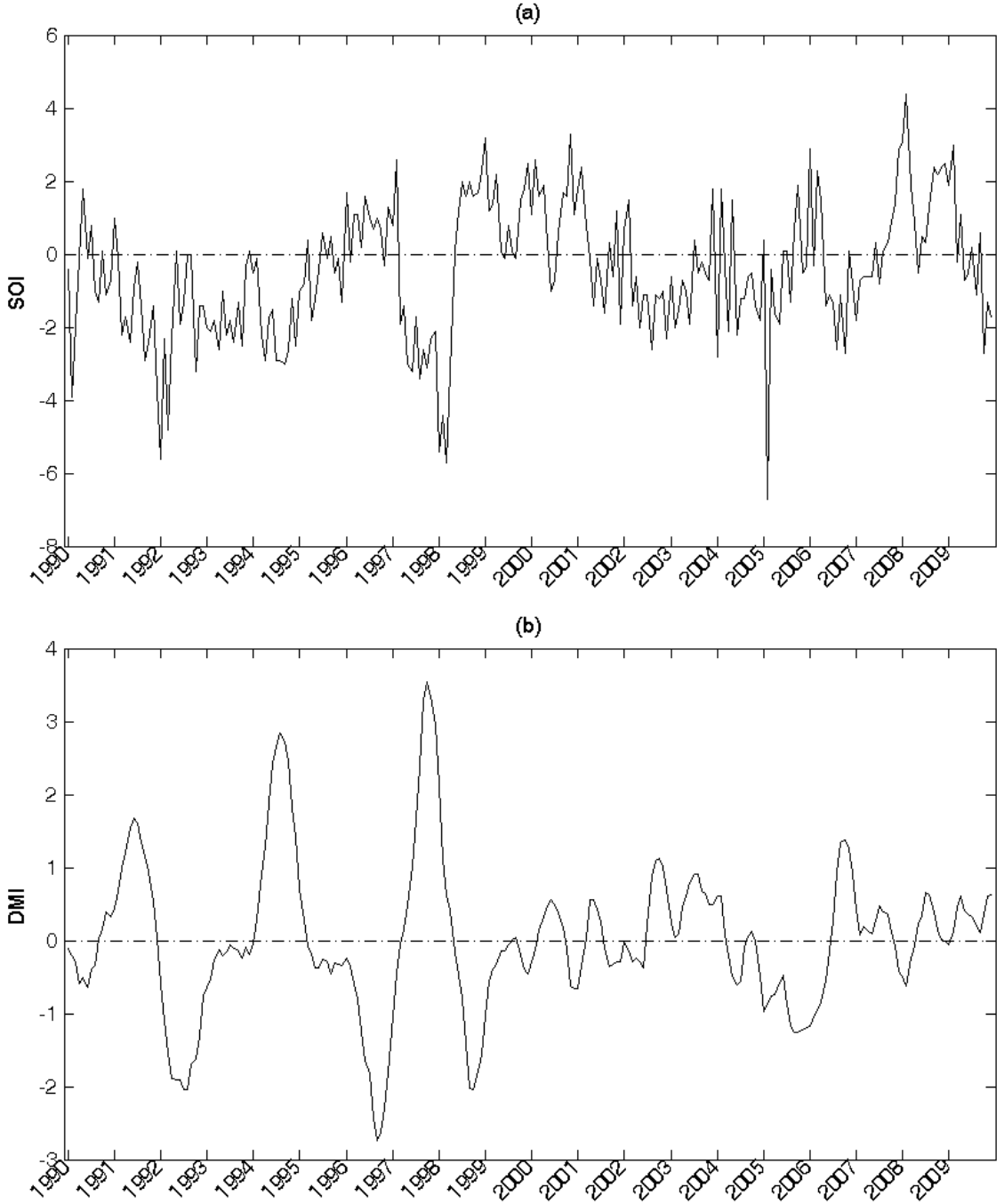
p_{diff} is the mean sea level pressure of Tahiti minus that of Darwin for that month;

$\langle p_{diff} \rangle$ is the long-term average of p_{diff} for the month

SD(p_{diff}) is the long-term standard deviation of p_{diff}

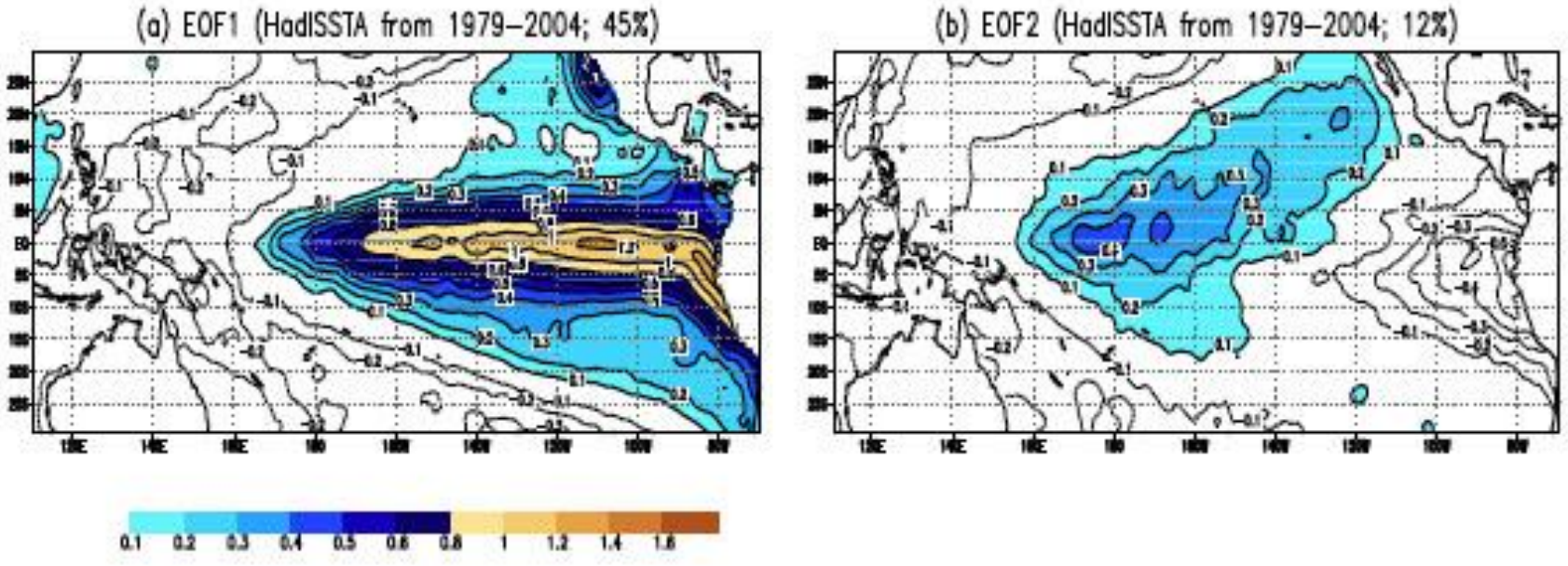
Southern Oscillation Index (SOI)

Dipole Mode Index (DMI)



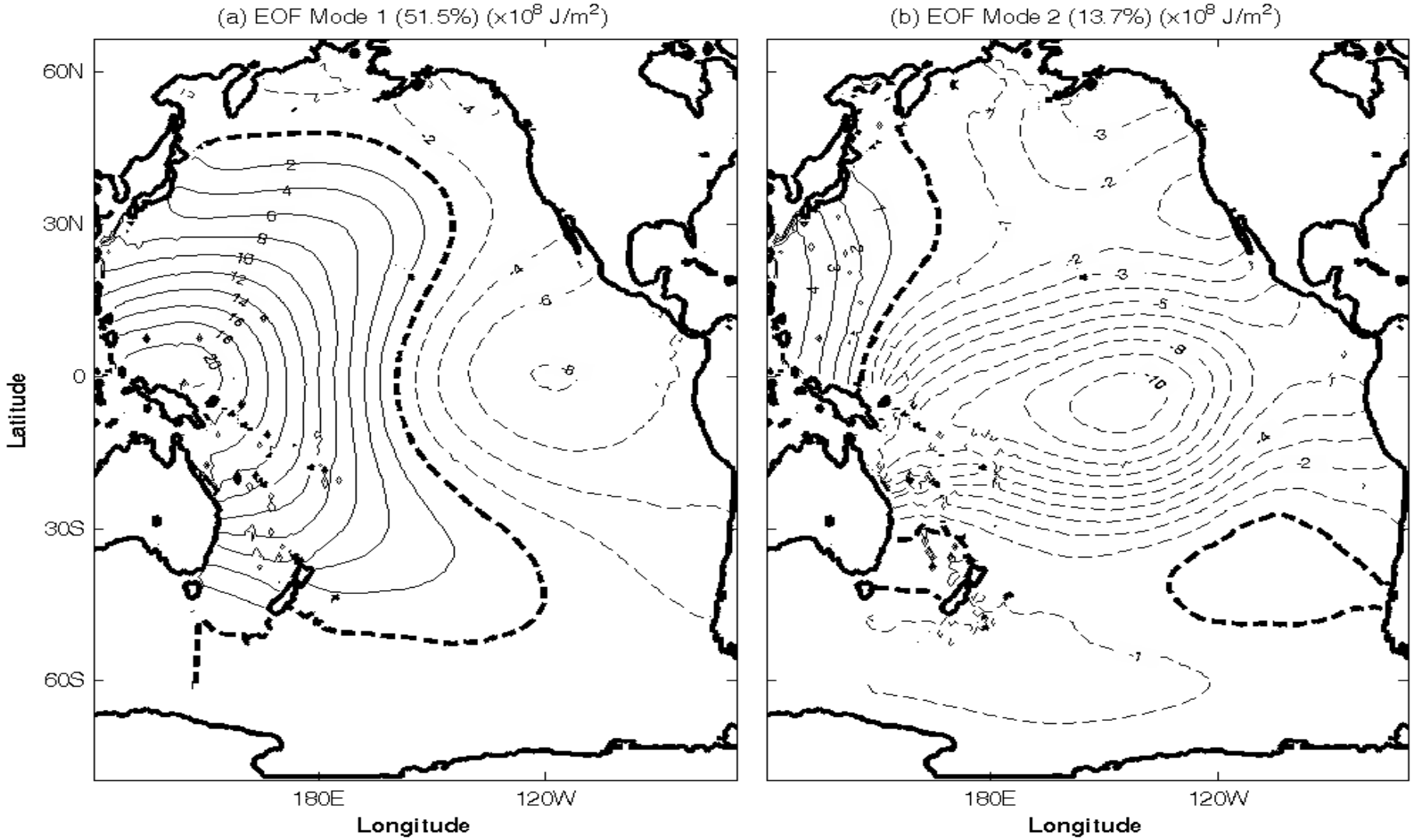
El Nino and El Nino Modoki

(Weng et al. 2007, Ashok et al. 2007)



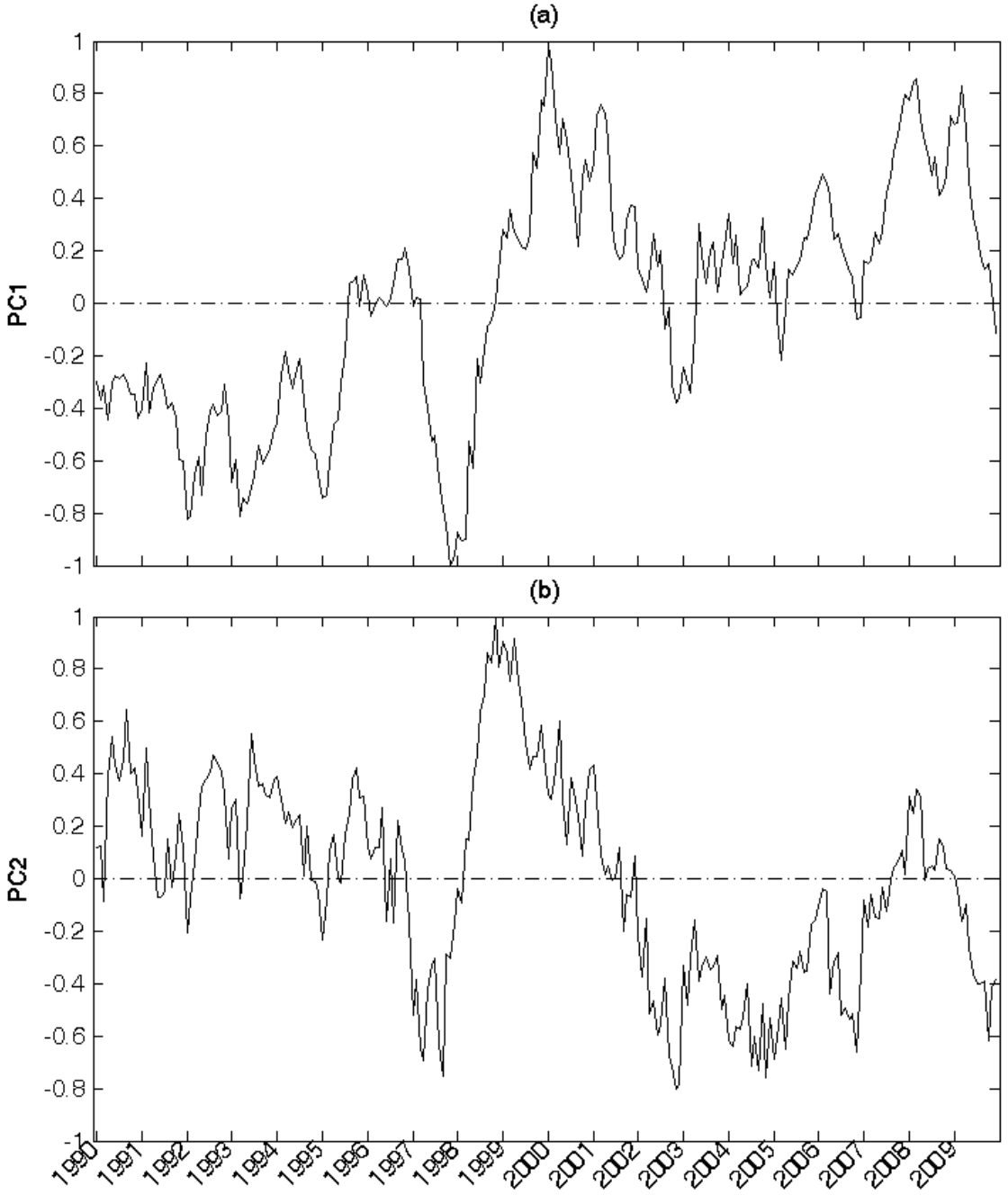
(Images from Karumuri Ashok, APEC Climate Center)

EOF-1 (51.5%) → Canonical El Nino
EOF-2 (13.7%) → El Nino Modoki
Pacific Ocean (in 10^8 J/m^2)



PC1

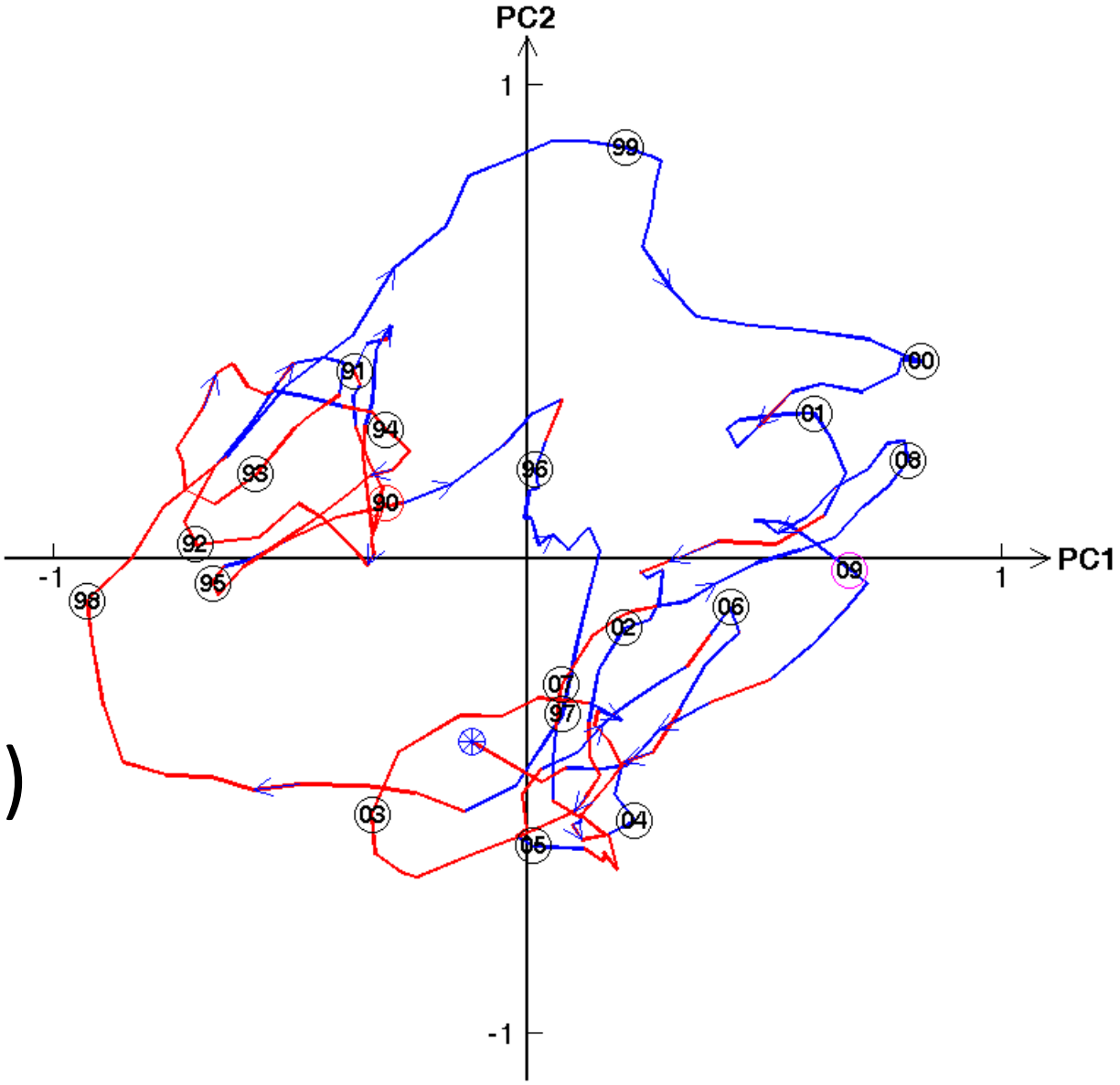
PC2


(a)
PC1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
(b)
PC2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009

Phase Diagram (Pacific)

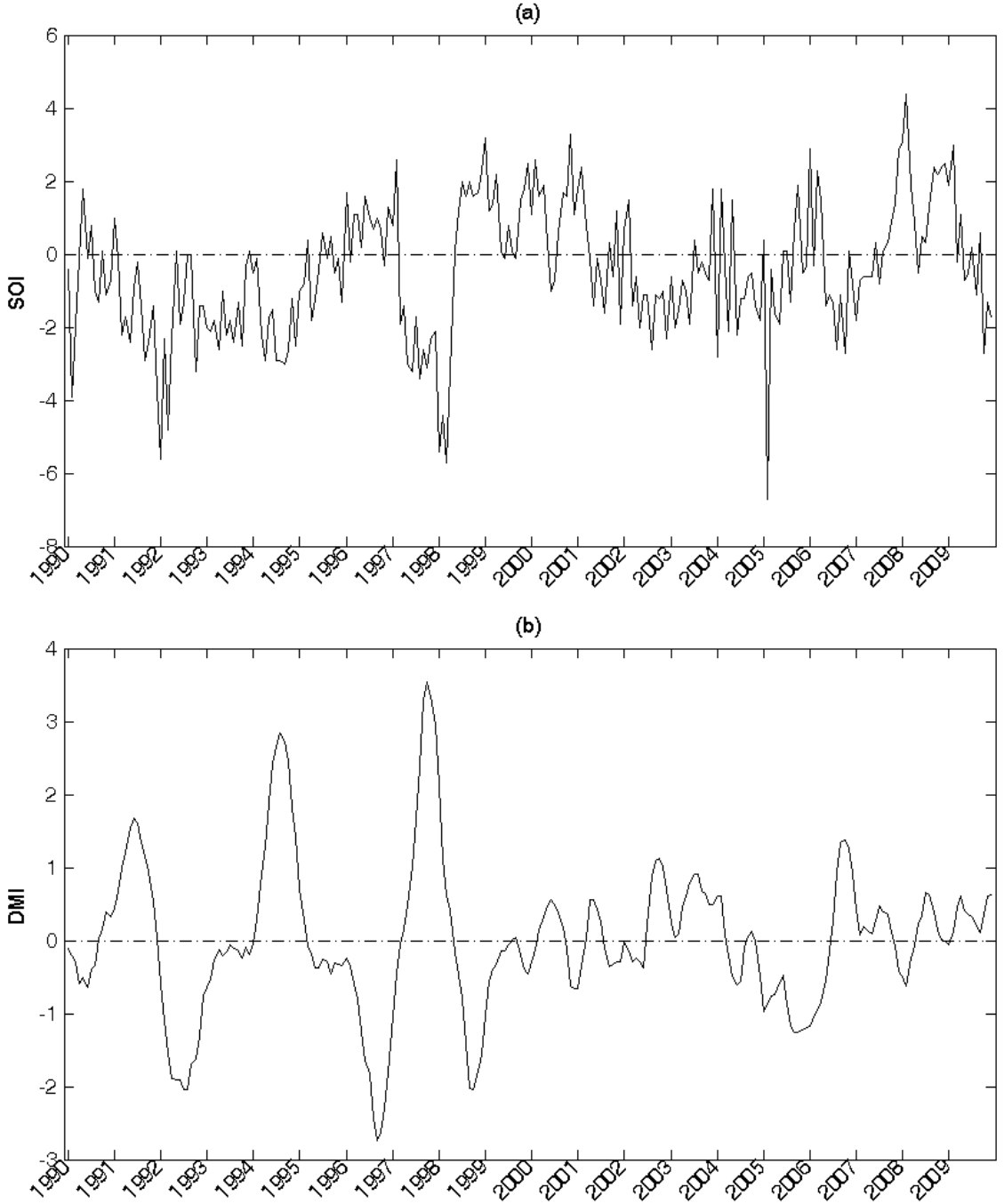
Red → SOI (-)

Blue → SOI (+)



Southern Oscillation Index (SOI)

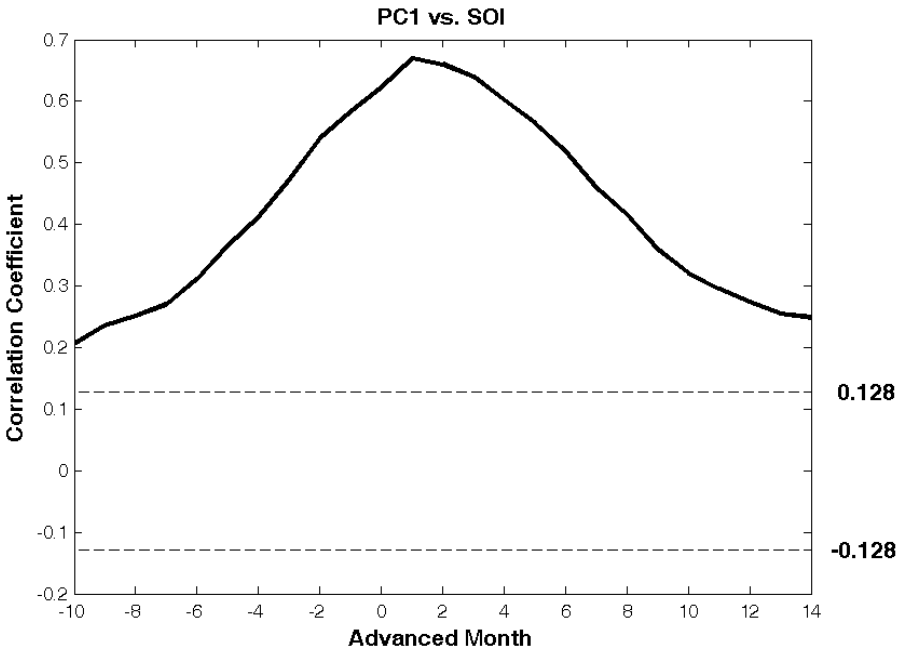
Dipole Mode Index (DMI)

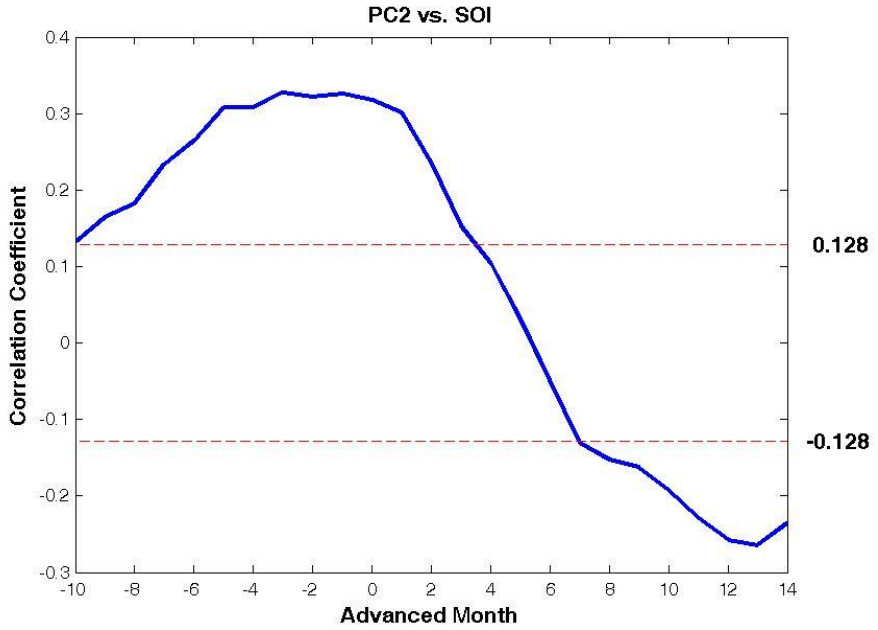


Lag-Correlation Coefficients

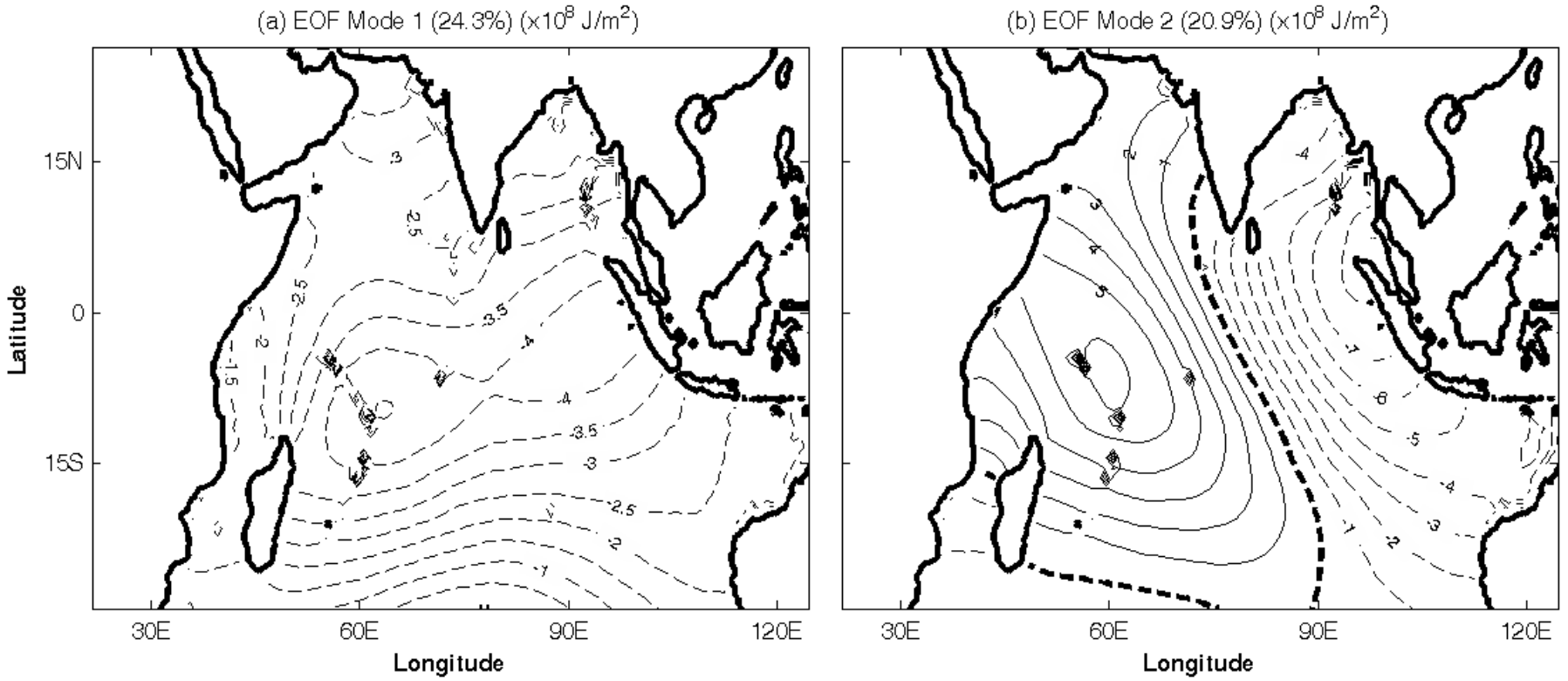
PC1 ←→ SOI

PC2 ←→ SOI





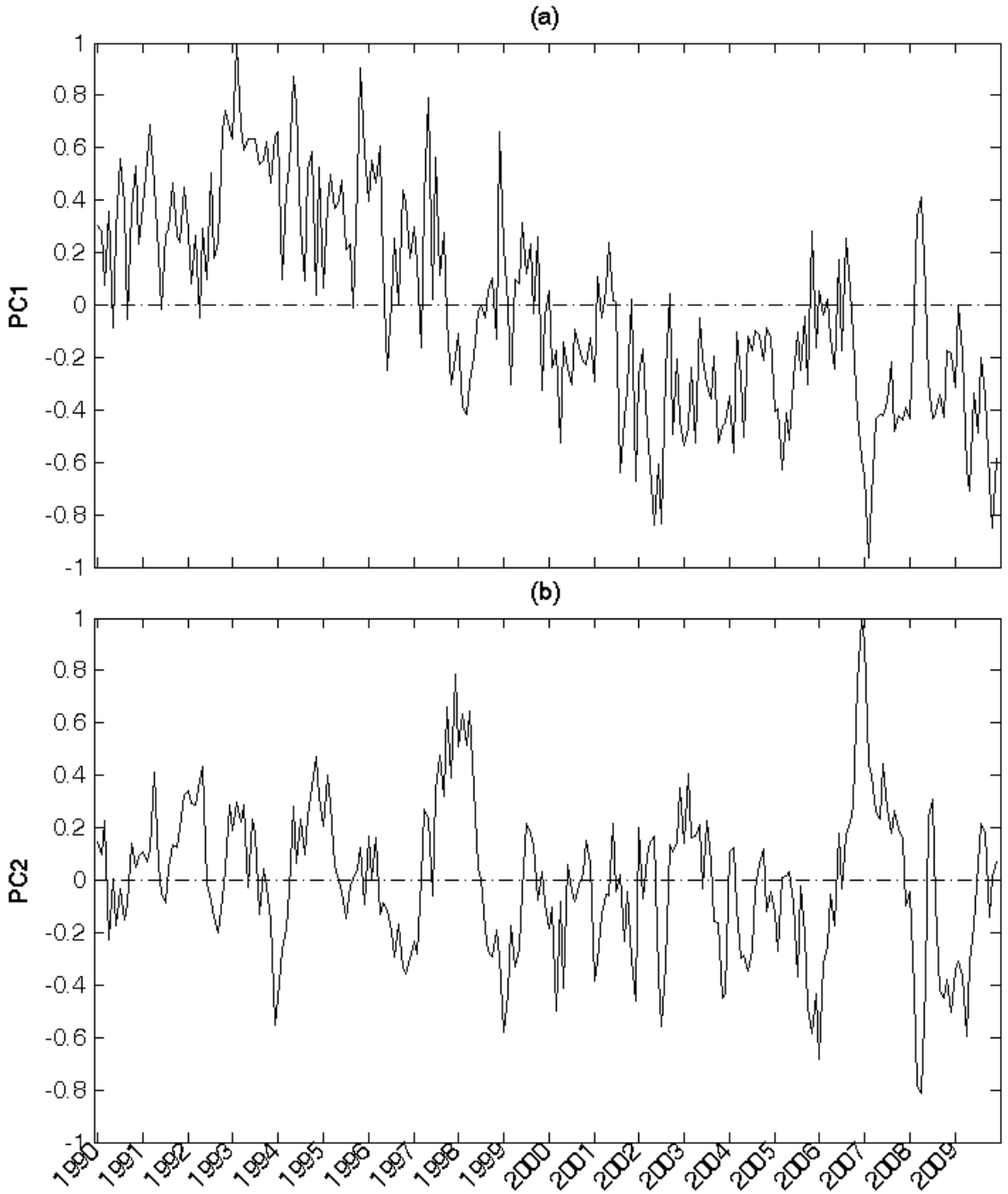
EOF-1 (24.3%) → Basin-scale Oscillation
EOF-2 (20.9%) → Dipole Mode
Indian Ocean (in 10^8 J/m^2)



PC1

→ Decadal

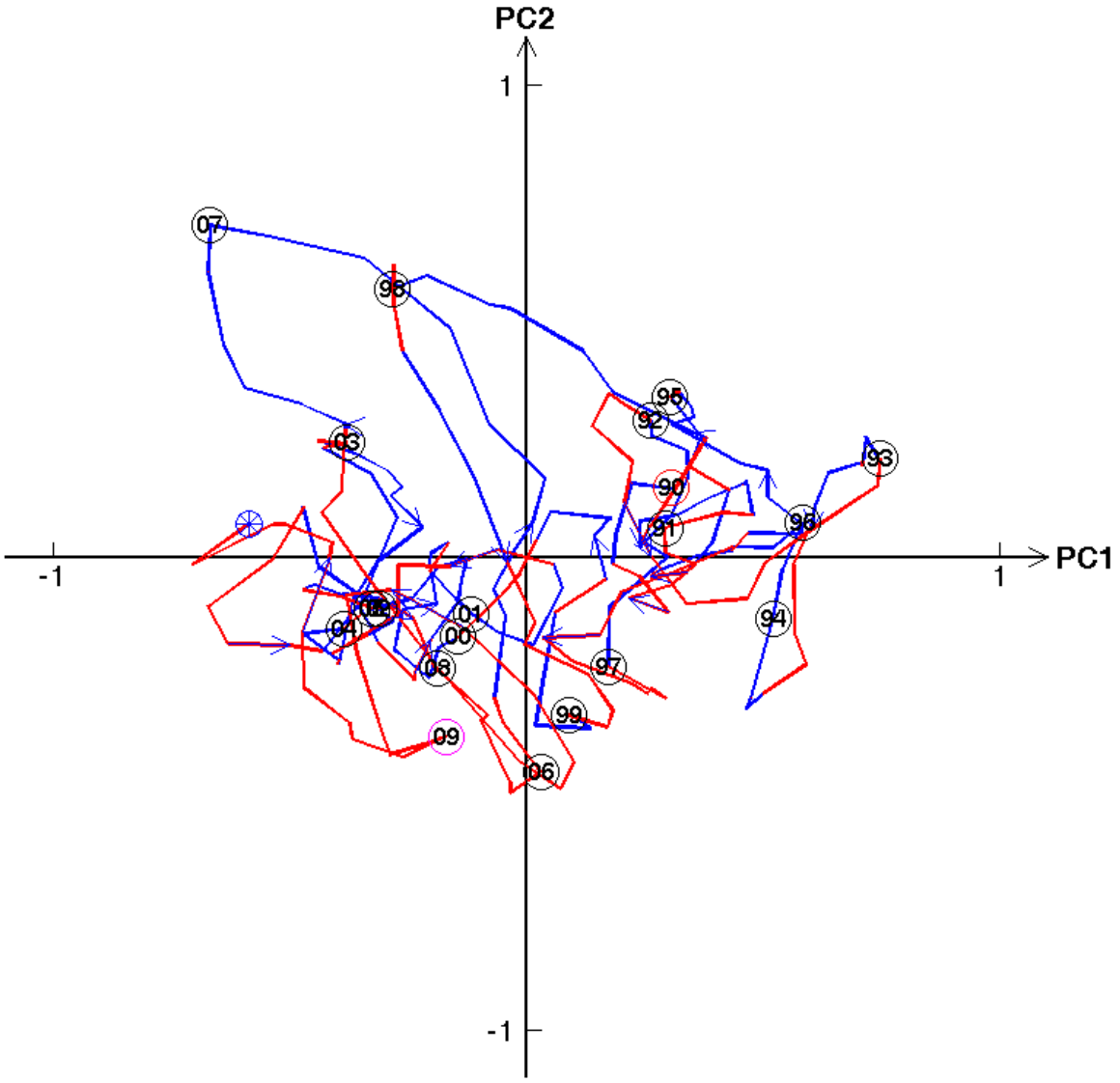
PC2



Phase Diagram (Indian)

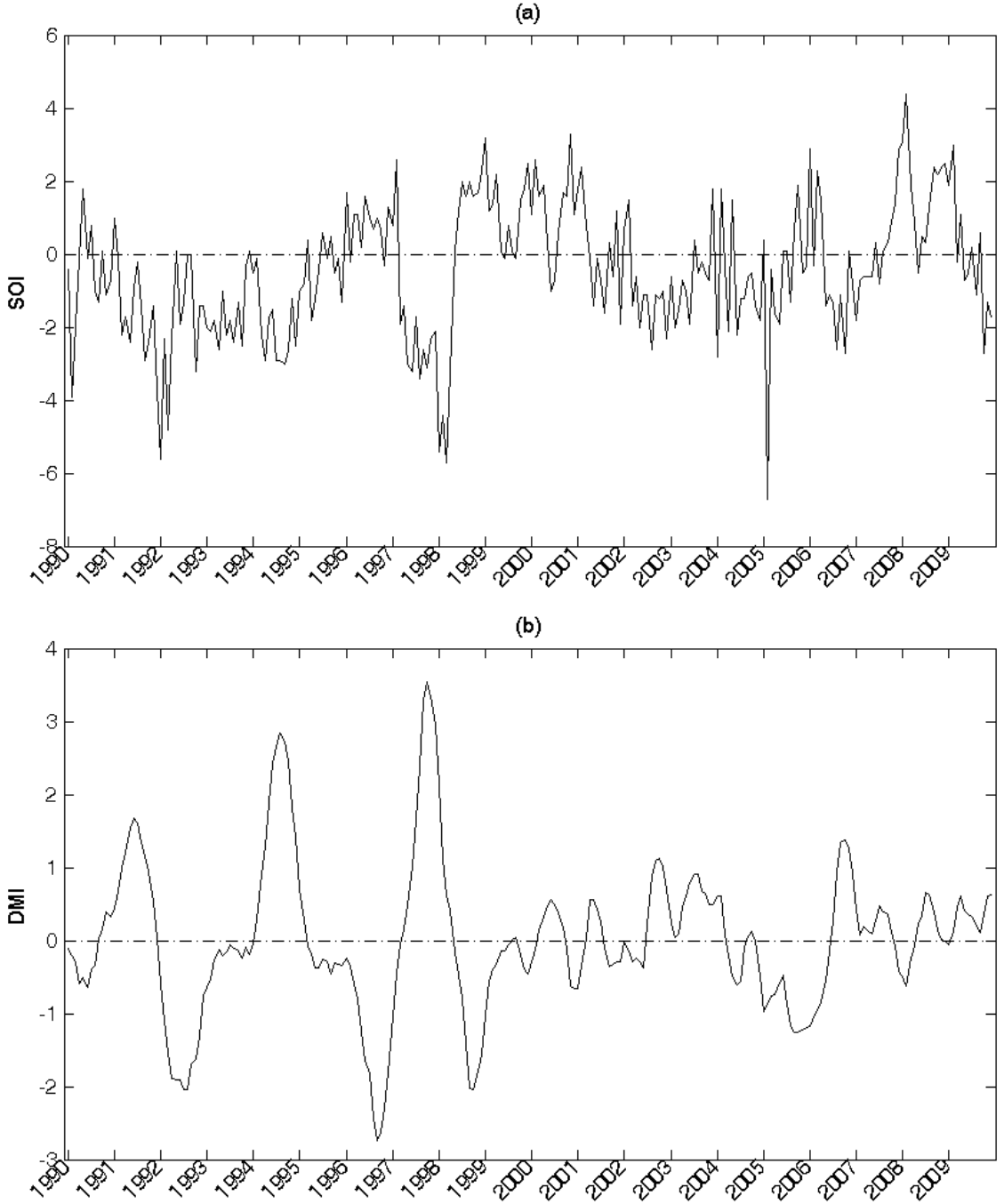
Red → DMI (-)

Blue → DMI (+)

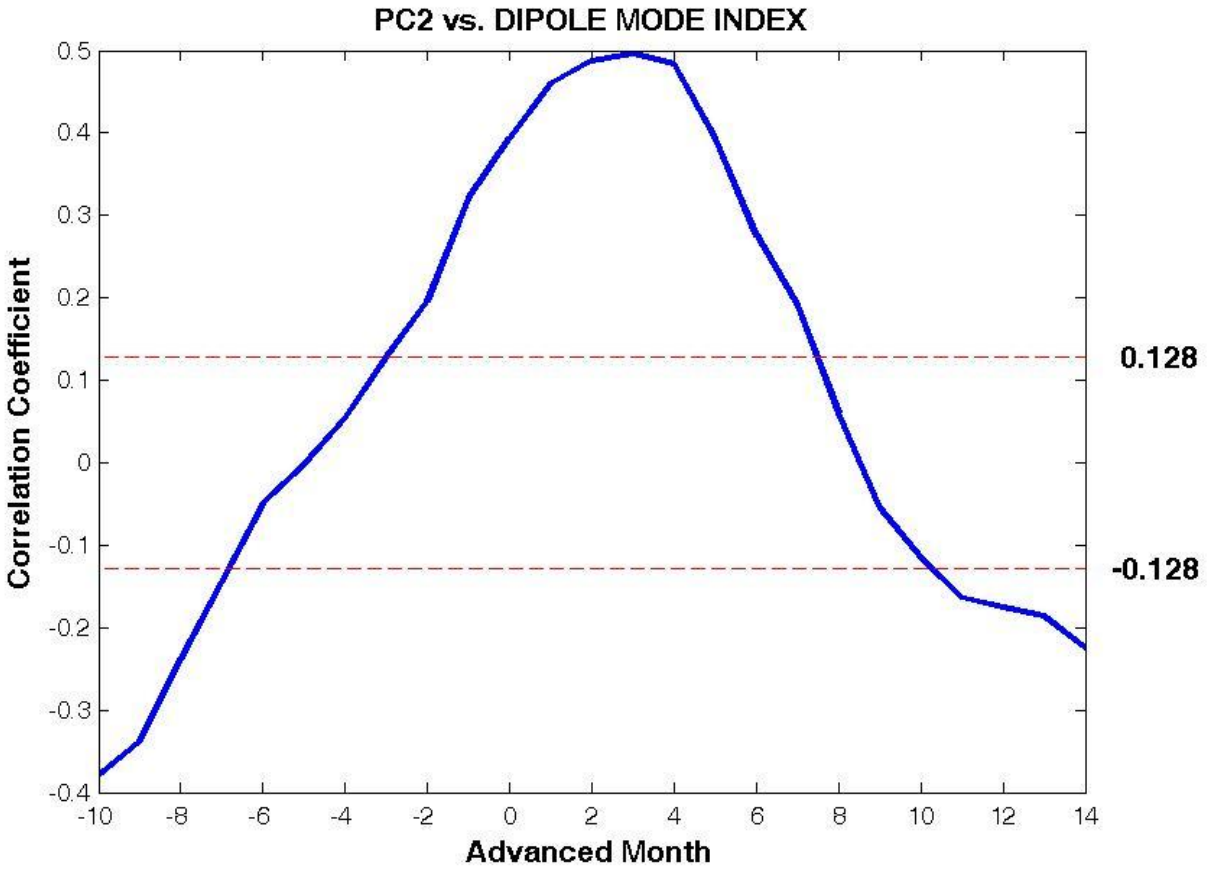


Southern Oscillation Index (SOI)

Dipole Mode Index (DMI)

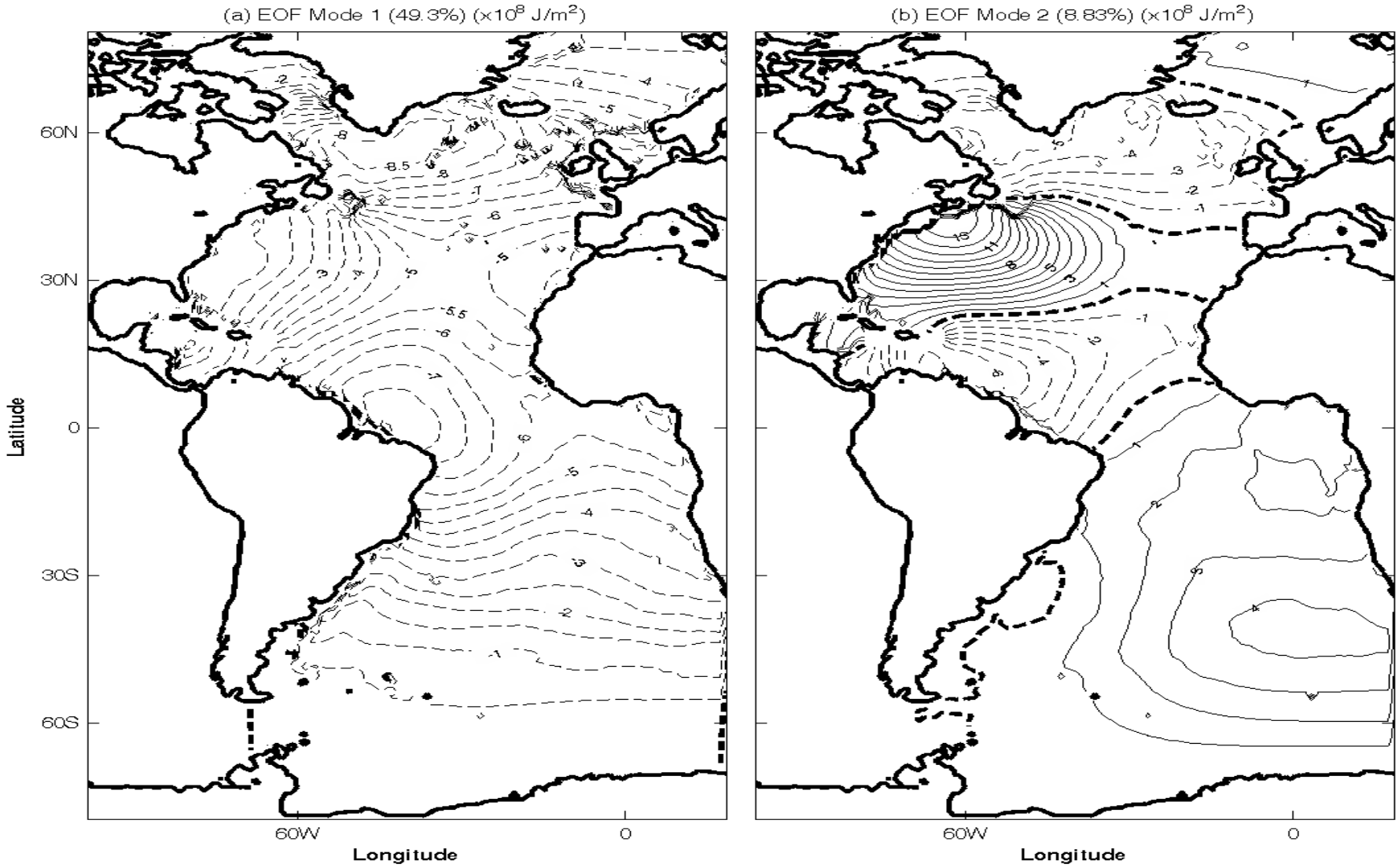


Lag Correlation between PC2 and DMI



Positive Month → PC2 advancing DMI

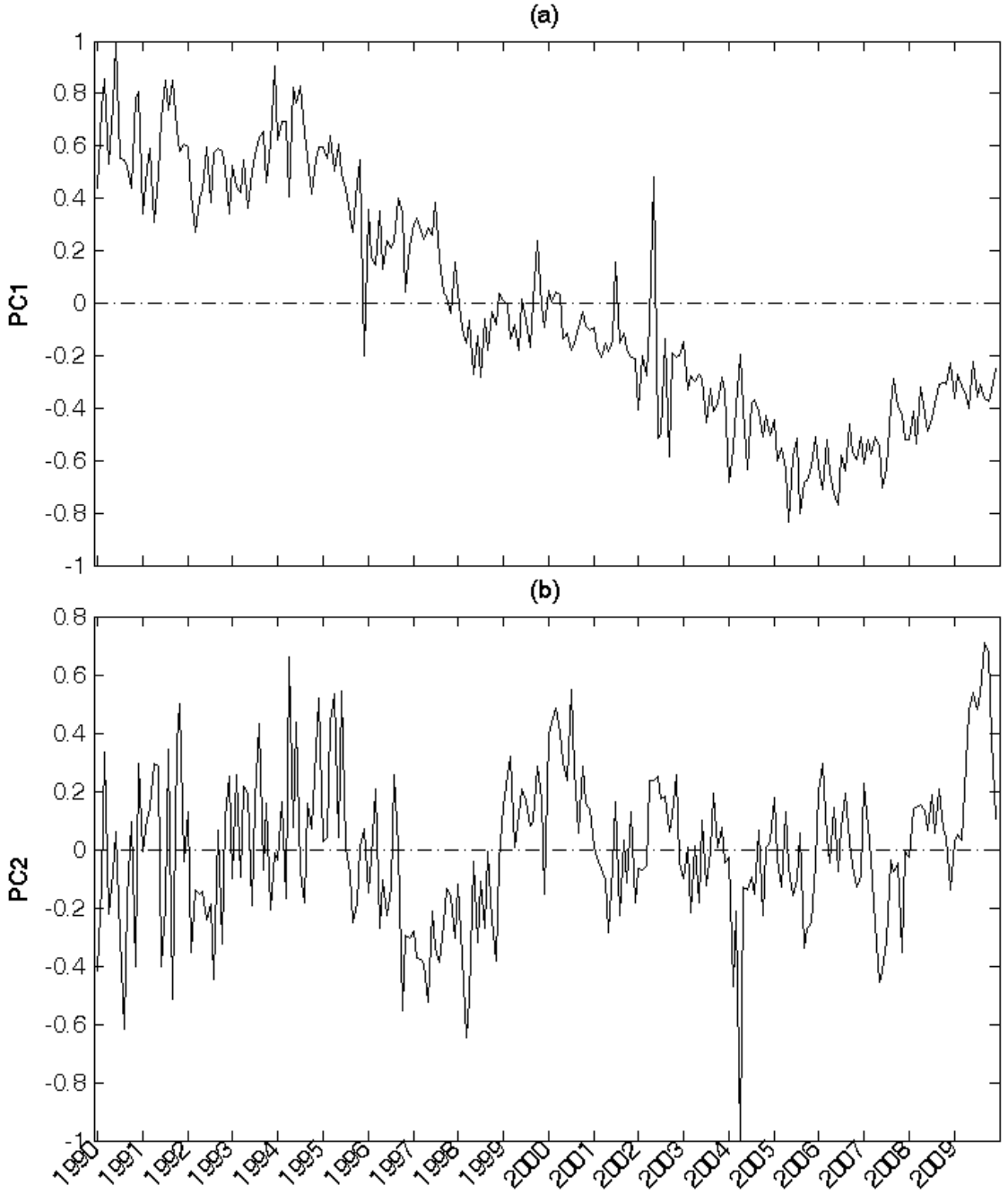
EOF-1 (49.3%) → Basin-Scale Oscillation
EOF-2 (8.8%) → Zonal Variability
Atlantic Ocean (in 10^8 J/m^2)



PC1

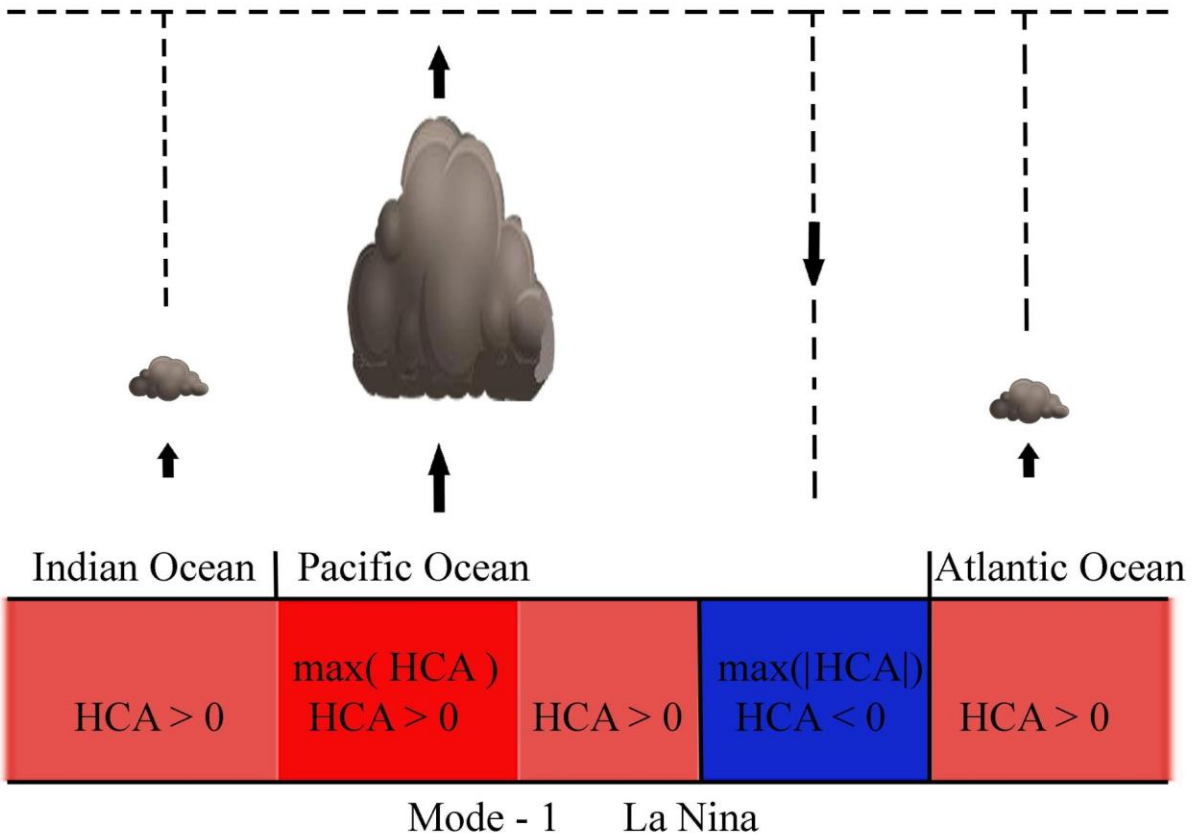
→ Decadal

PC2

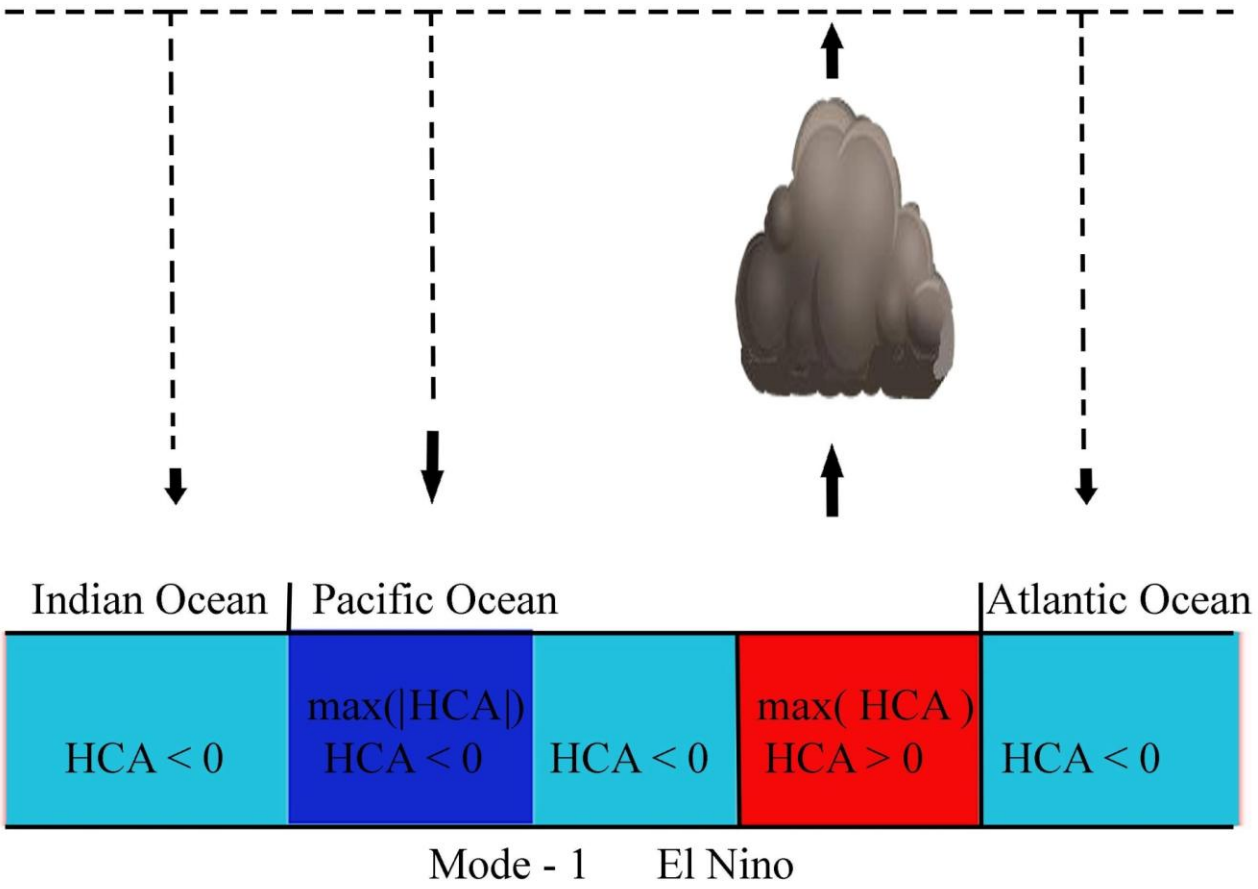


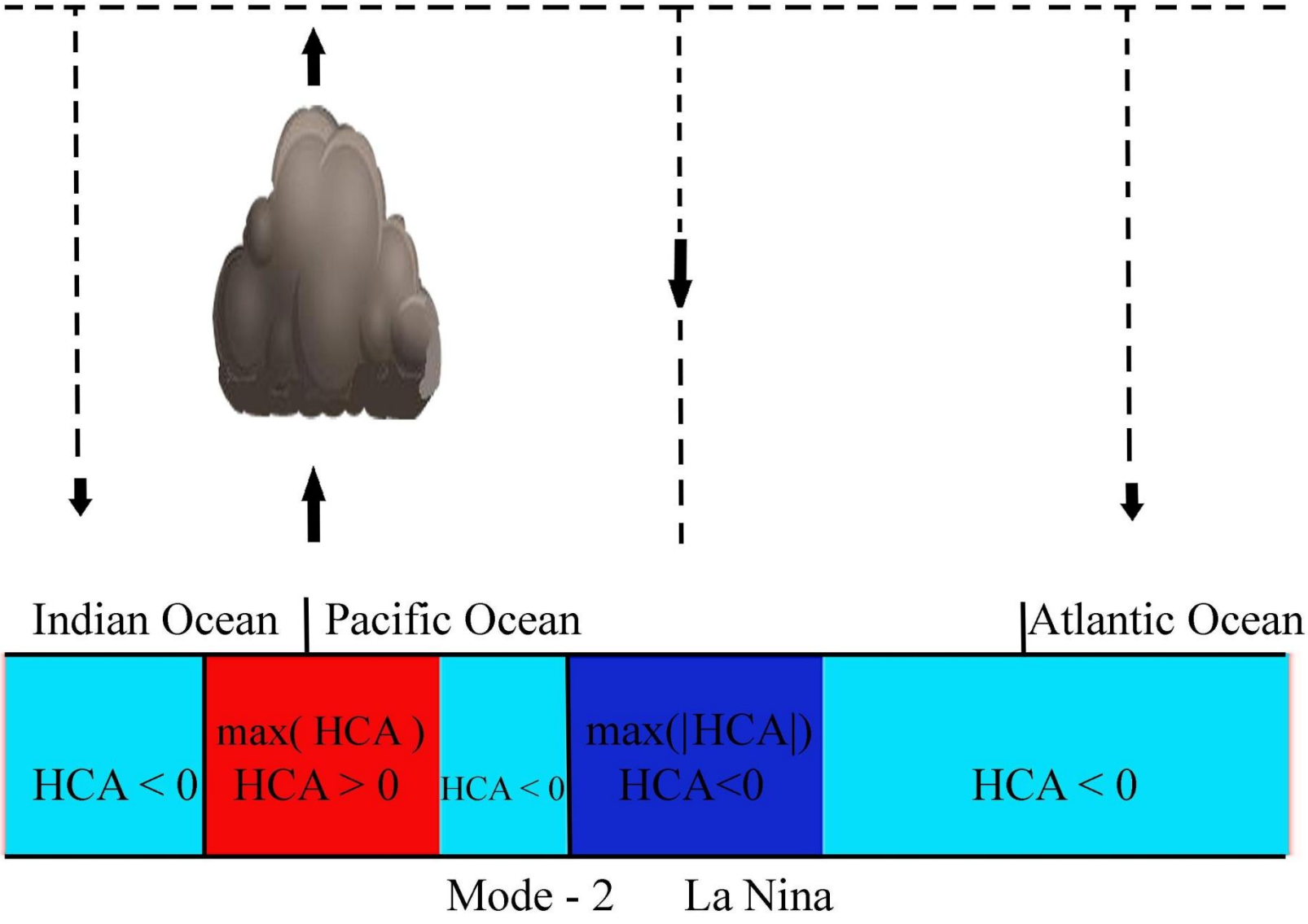
Part 4
Global Ocean Tripole

Canonical La Nina →
More and Stronger Hurricanes in Atlantic (Pielke and Landsea, 1999 BAMS)



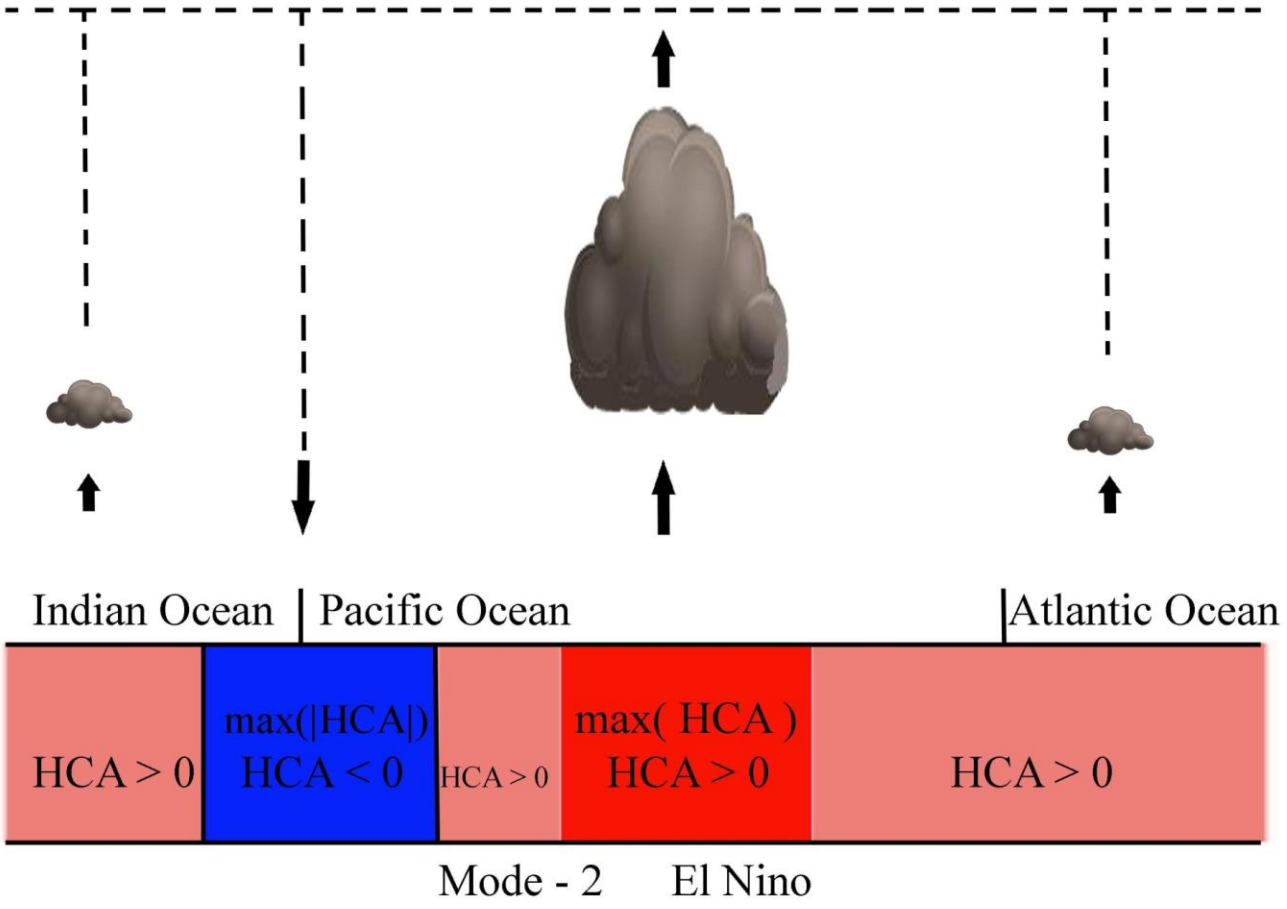
Canonical El Nino




Indian Ocean
Pacific Ocean
Atlantic Ocean
HCA < 0
max(HCA)
HCA > 0
HCA < 0
max(|HCA|)
HCA<0
HCA < 0
Mode - 2
La Nina

El Nino Modoki

More and Stronger Hurricanes in Atlantic



Conclusions

- OSD is a useful method especially for ocean data assimilation. No background error covariance matrix B is needed.
- OSD method effectively uses the ocean topography and lateral boundary conditions.
- The basis functions are orthonormal
- The basis functions are the eigen functions of the Laplacian operator (pre-determined).
- Global gridded monthly varying (T, S) data sets are useful for ocean data assimilation and model verification as well as climate studies.