Propagating and quantifying uncertainties in numerical ocean simulations using Polynomial Chaos

An example problem in the Gulf of Mexico

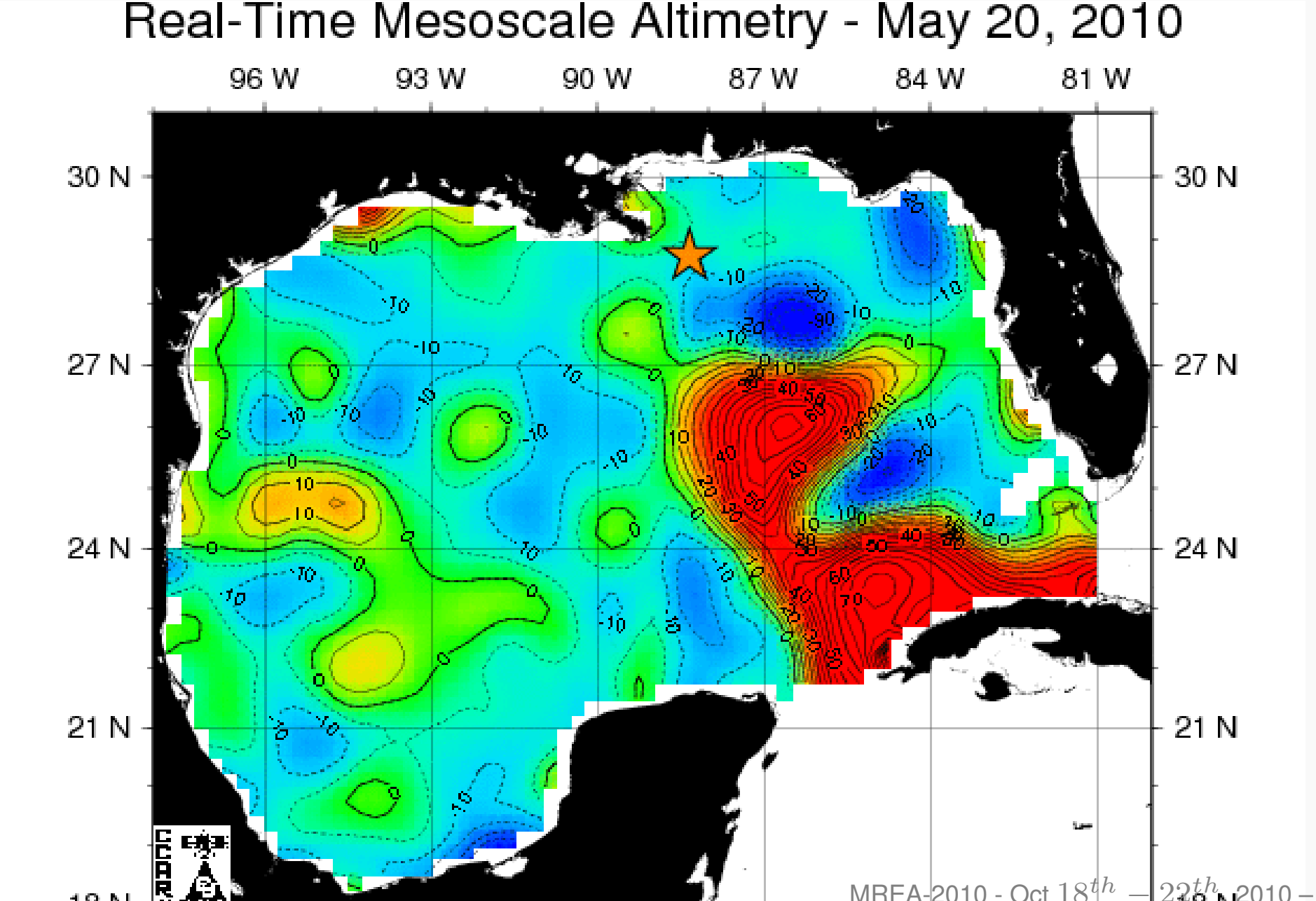
Ashwanth Srinivasan, RSMAS, University of Miami, Miami, FL, USA

Carlisle Thacker, CIMAS, University of Miami, Miami, FL, FL, USA

Mohamed Iskandarani, RSMAS, University of Miami, Miami, FL, USA

Omar Knio, Johns Hopkins University, Baltimore, MD, USA

Gulf of Mexico: Circulation and Loop Current



Polynomial Chaos

- Polynomial Chaos originally introduced by Norbert Wiener (1938)
- Addressed the question of efficiently estimating uncertainties of a dynamical system when given probability distributions for the system's parameters.
- Expanding the solution in a polynomial series in the uncertain parameters reduces the problem to that of determining the expansion coefficients.
- The name "polynomial chaos", which has become popular in the engineering literature, stems from Wiener's referring to uncertainty as "chaos" and his use of a polynomial expansion.
- Introduced by Ghanem and Spanos (1991) in the engineering community
- Manner et al. 1997 and Chen et al. (2010) used this in the Atmospheric and Oceanographic context.

Polynomial Chaos Expansions

Parameterize uncertain model inputs as functions of random variables $\xi_1, \xi_2, \ldots, \xi_N$

The model solution can be represented as a spectral expansion in terms of suitable orthogonal polynomial basis functions associated with random variables of a given probability density as:

$$\mathbf{u}(x, t, \xi_1, \xi_2, \ldots, \xi_N) = \sum_{k=0}^{\infty} \mathbf{u}(x, t)_k \mathbf{\Psi}_k(\xi_1, \xi_2, \ldots, \xi_N)$$

where x and t denote space and time and ξ sample a random event space. The $\mathbf{u}(x, t)_k' s$ are deterministic spectral coefficients, reffered to as PC coefficients. The $\mathbf{\Psi}_k' s$ are multi-dimensional orthogonal polynomials of uncorrelated random variables $\{\xi\}_1^N$ with the following properties:

$$\langle \mathbf{\Psi}_0 \rangle \equiv \mathbf{E}[\mathbf{\Psi_0}] = \mathbf{1}, \quad \langle \mathbf{\Psi_k} \rangle = \mathbf{0}, \mathbf{k} > \mathbf{0},$$
$$\langle \mathbf{\Psi}_i \mathbf{\Psi}_j \rangle = h_i \delta_{ij}$$

where h_i are normalization constants and δ_{ij} is the Kronecker delta. These spectral expansions are generally referred to as polynomial-chaos (PC) expansions, following Wiener(1938).

Truncating the Polynomial Chaos Expansions

In practice the PC expansion is truncated to P terms as:

$$\mathbf{u}(x,t,\xi_1,\xi_2,\ldots,\xi_N) = \sum_{k=0}^{P} \mathbf{u}(x,t)_k \mathbf{\Psi}_k(\xi_1,\xi_2,\ldots,\xi_{Ndim})$$

The value of P is determined by the number N of random variables and the highest degree of polynomial used to represent the solution $\mathbf{u}$ with the formula:

$$P = \frac{(N+K)!}{(N!K!)} - 1$$

- Each uncertain parameter adds a dimension to the probability space that must be explored
- Leads to a practical limitation that Bellmann termed "the curse of dimensionality"

Non-Intrusive Spectral Projection

We apply the Non-Intrusive Spectral Projection (NISP) approach to obtain the PC coefficients of the spectral expansion. In this case the orthogonal modes are obtained as:

$$\mathbf{u}(x,t)_k = \langle \frac{\mathbf{u}(x,t,\xi_1,\xi_2,\ldots,\xi_N)\mathbf{\Psi}_k}{\mathbf{\Psi}_k^2} \rangle \quad k = 0, 1....P$$

where the expectations are found by evaluating the equivalent stochastic integrals over ξ-space, e.g.,

$$\langle \mathbf{u}(x,t,\xi_1,\xi_2,\ldots,\xi_N)\psi_k \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{u}(x,t,\xi_1,\xi_2,\ldots,\xi_N)\mathbf{\Psi}_k e^{\frac{-\xi^2}{2}} d\xi$$

using quadrature rules: e. g., Gauss-Hermite quadrature

$$\int_{-\infty}^{\infty} f(x) e^{\frac{-x^2}{2}} dx = \sum_{i=0}^{n} w_i f(x_i)$$

Computing Statistics of the Solution

Once we have the PC coefficients we can compute approximate statistics of the solution with the following formulas:

$$E[u] = E\left[\sum_{k=0}^{P} \mathbf{u}(x,t)_k \mathbf{\Psi}_k\right] \quad = u_0 E[\mathbf{\Psi}_0] + \textstyle\sum_{k=0}^{P} u_0 E[\mathbf{\Psi}_0] \quad = u_0$$

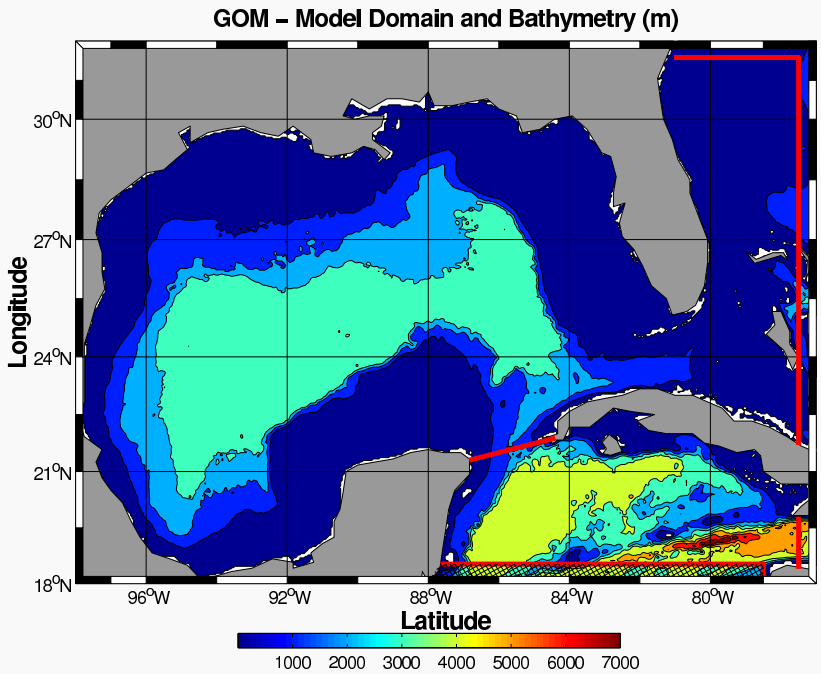
$$Var[u] = E[(u - E[u])^2] = E[((\sum_{k=0}^{P} \mathbf{u}(x,t)_k \mathbf{\Psi}_k) - u_0)^2]$$

$$= E[(\sum_{k=1}^{P} \mathbf{u}(x,t)_k \mathbf{\Psi}_k) - u_0)^2] = \sum_{k=1}^{P} \mathbf{u}(x,t)_k{}^2 E[\mathbf{\Psi}_k{}^2]$$

and $Std[u] = \sqrt{var[u]}$. We can also approximate the PDF of U by sampling from the distribution of ξ and plugging them into the PCE.

A Gulf of Mexico HYCOM example

- 1/25° Gulf of Mexico HYCOM (Configuration similar to one used at NRL- Pat Hogan & Ole Martin Smedstad)
- We use two Gaussian distributed random variables sampled to parameterize the uncertain southern boundary inflow



- 5th Gauss Hermite polynomial expansions to obtain solutions – *requires 36 model runs*

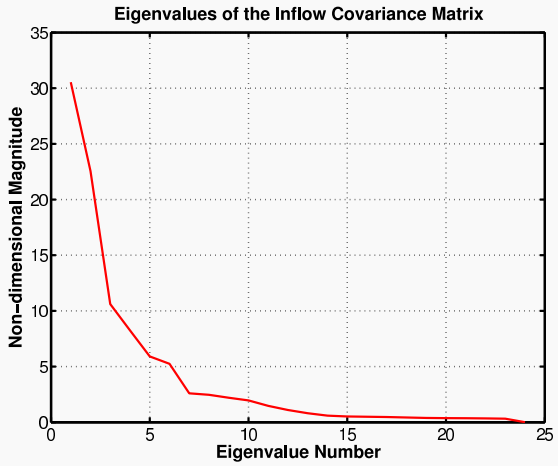
Specifying Inflow Uncertainties

We represent the open boundary conditions in the south as spatially varying random field with a Gaussian probability density, using the Karhunen-Loeve expansions (KLE), $\alpha(x,t,\omega)$ with a mean $\bar{\alpha}(x,t)$ and covariance $C_\alpha(x1,x2,t1,t2)$. Then stochastic boundary condition $\alpha(x,t,\omega)$ is:

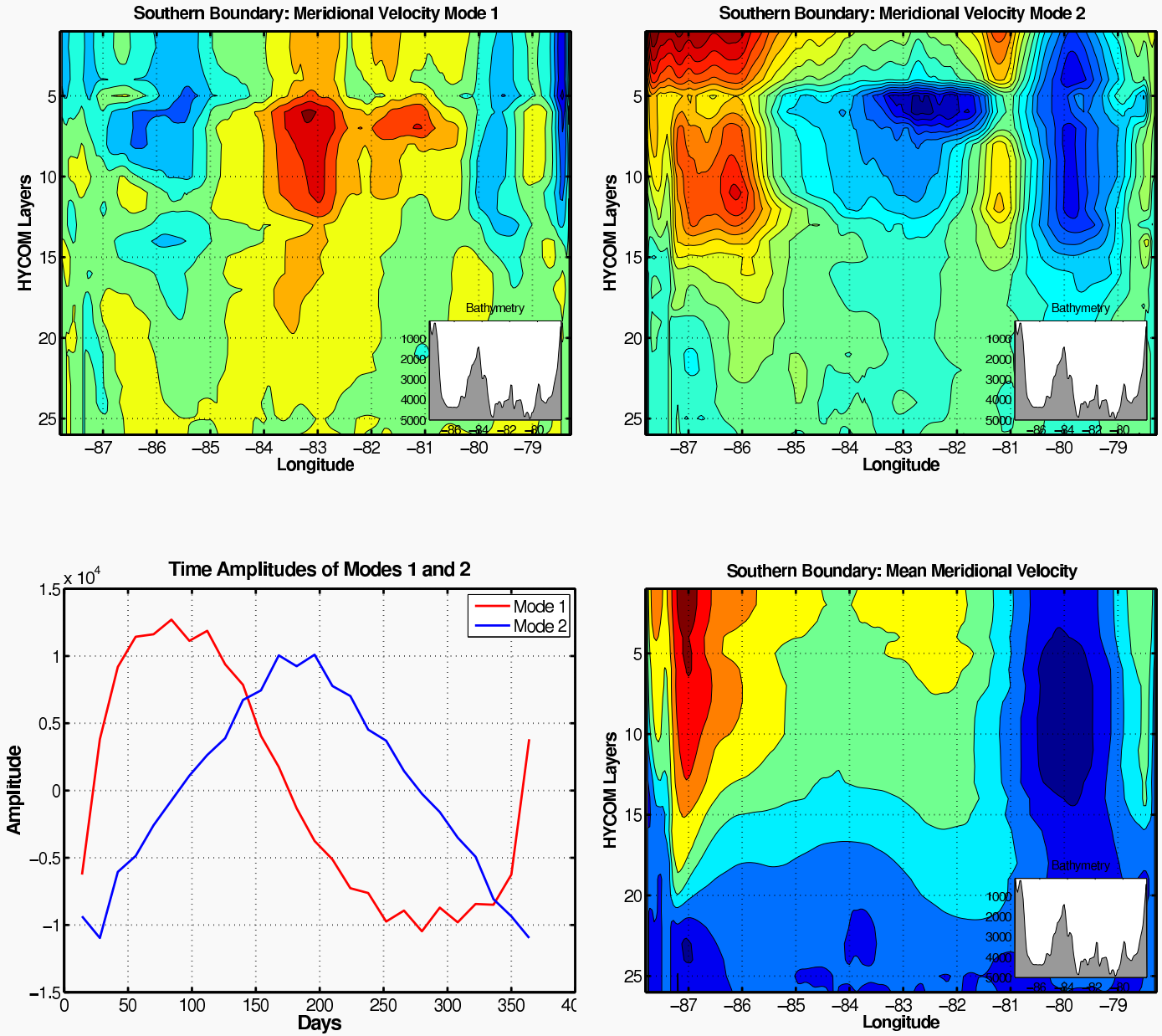
$$\alpha(x,t,\omega) = \bar{\alpha}(x,t) + \sum_{k=1}^{2} \sqrt{(\lambda_k)}\phi_k \xi_\omega$$

In the above series, λ_k and the ϕ_k are the eigenvalues and eigenfunction of the covariance matrix. They are be approximated numerically for our purposes. We truncate the KLE expansion to the first two terms - problem has two random dimensions

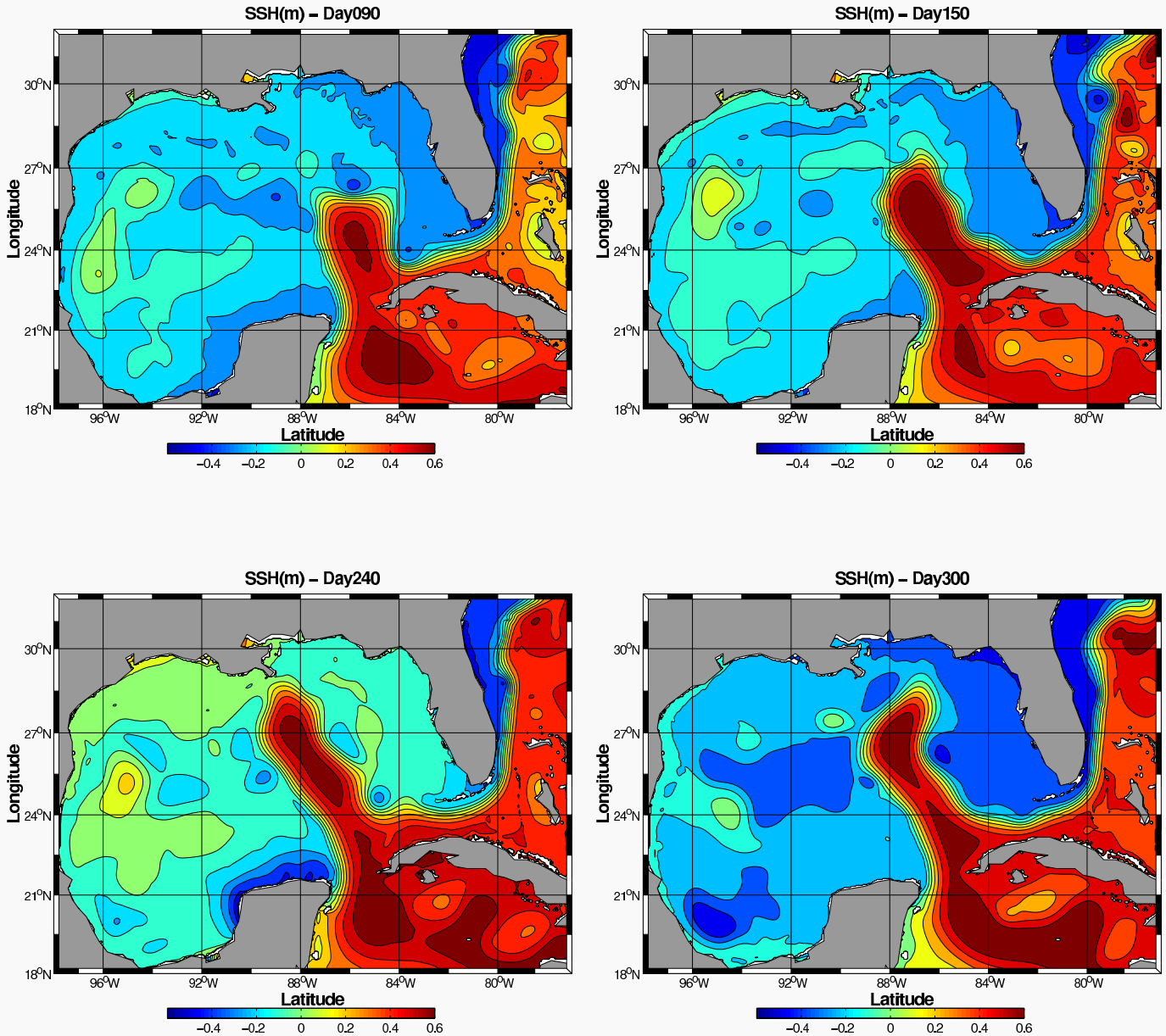
The first two modes capture about 60 % energy.



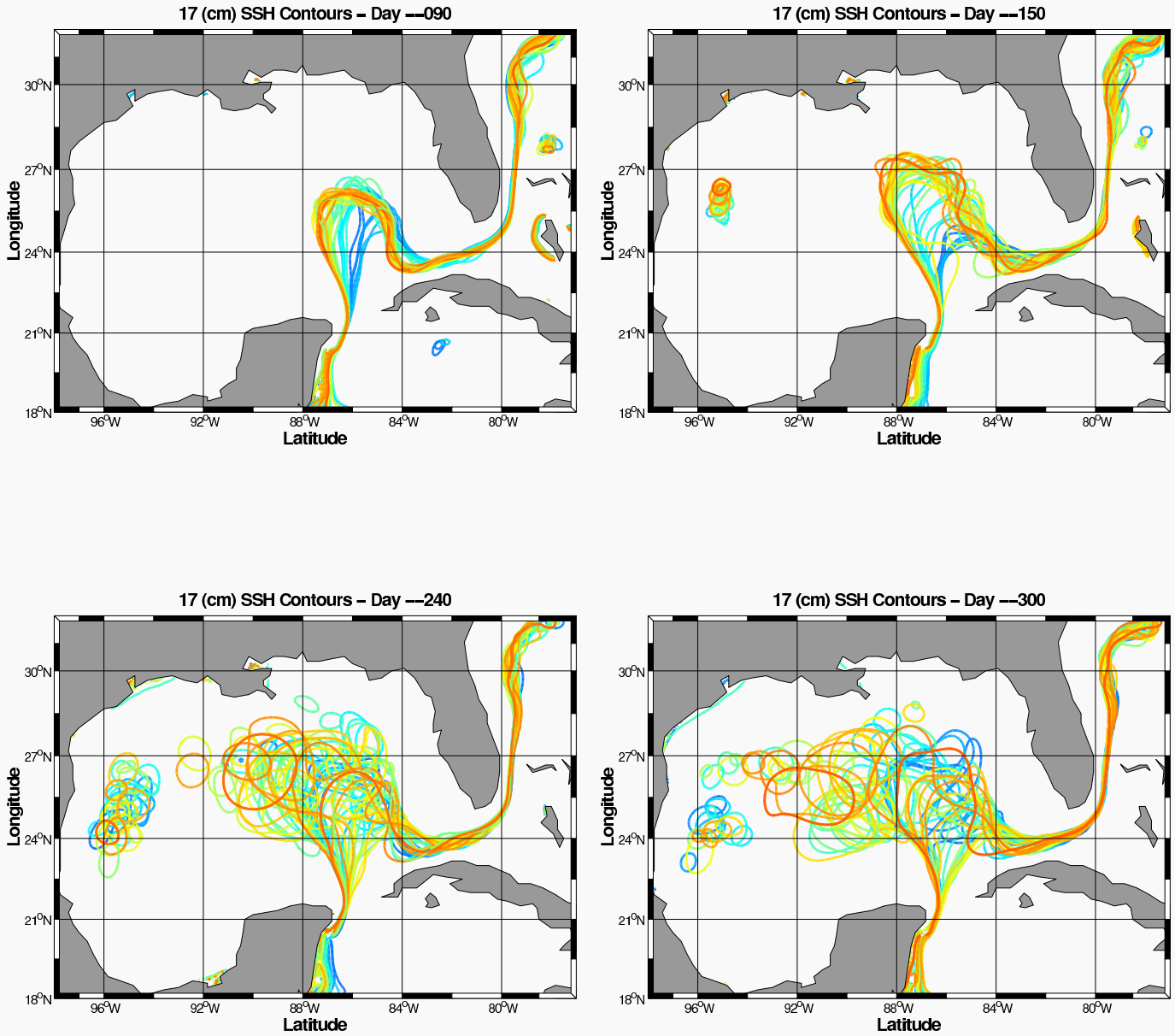
The two spatial modes at the Southern Boundary



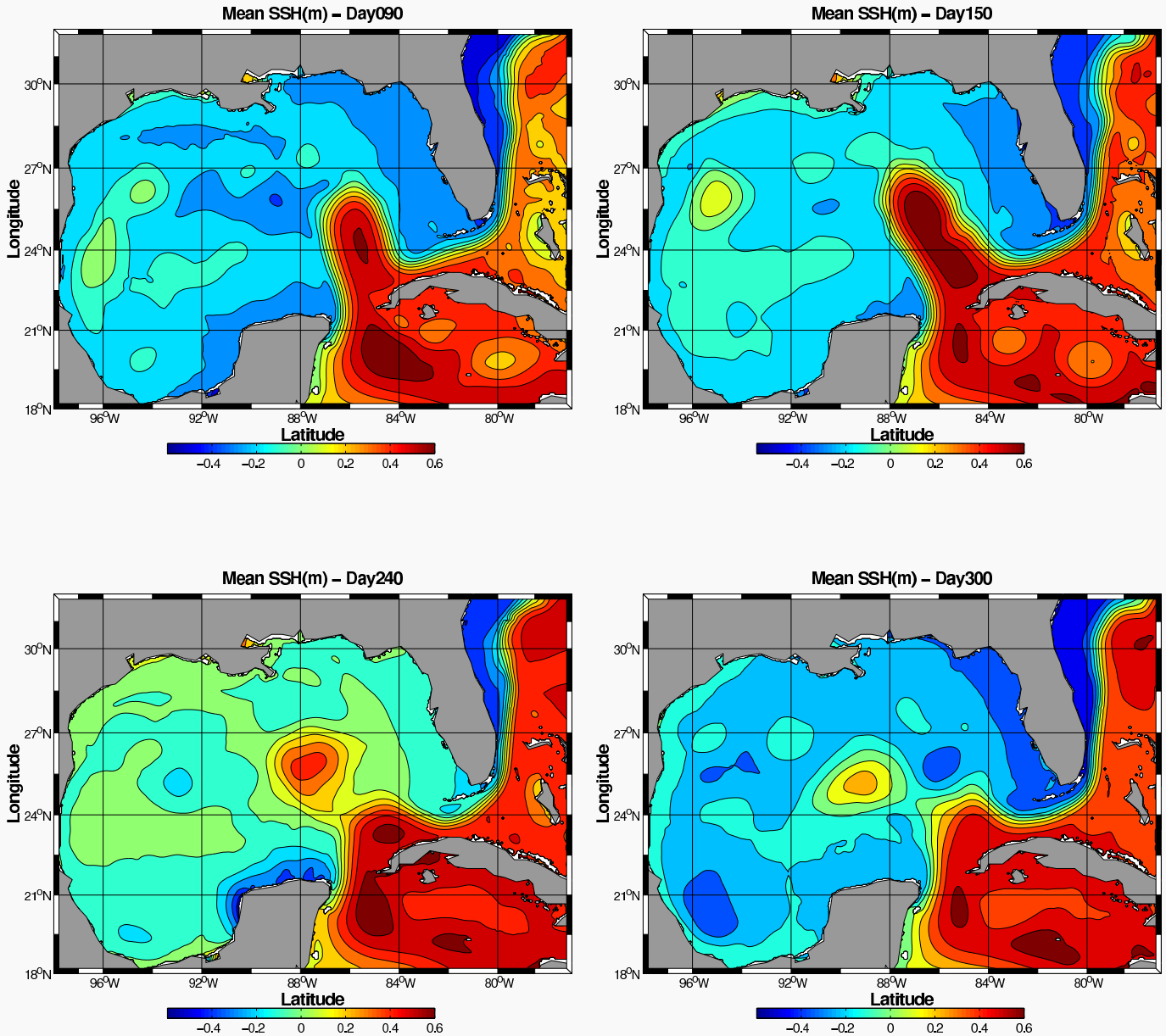
Snapshots of the Unperturbed Reference Run



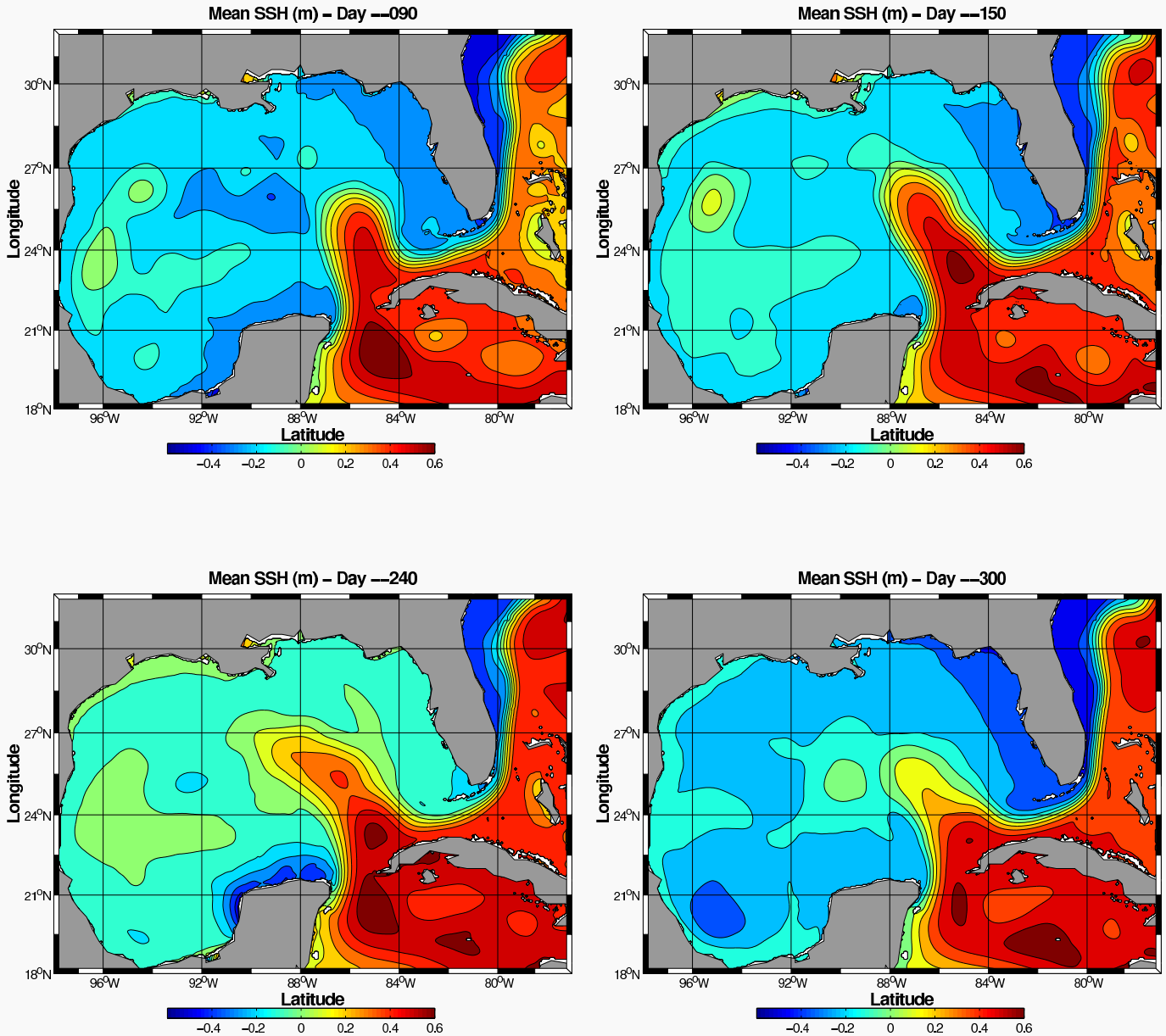
Snapshots of the 36 Realizations



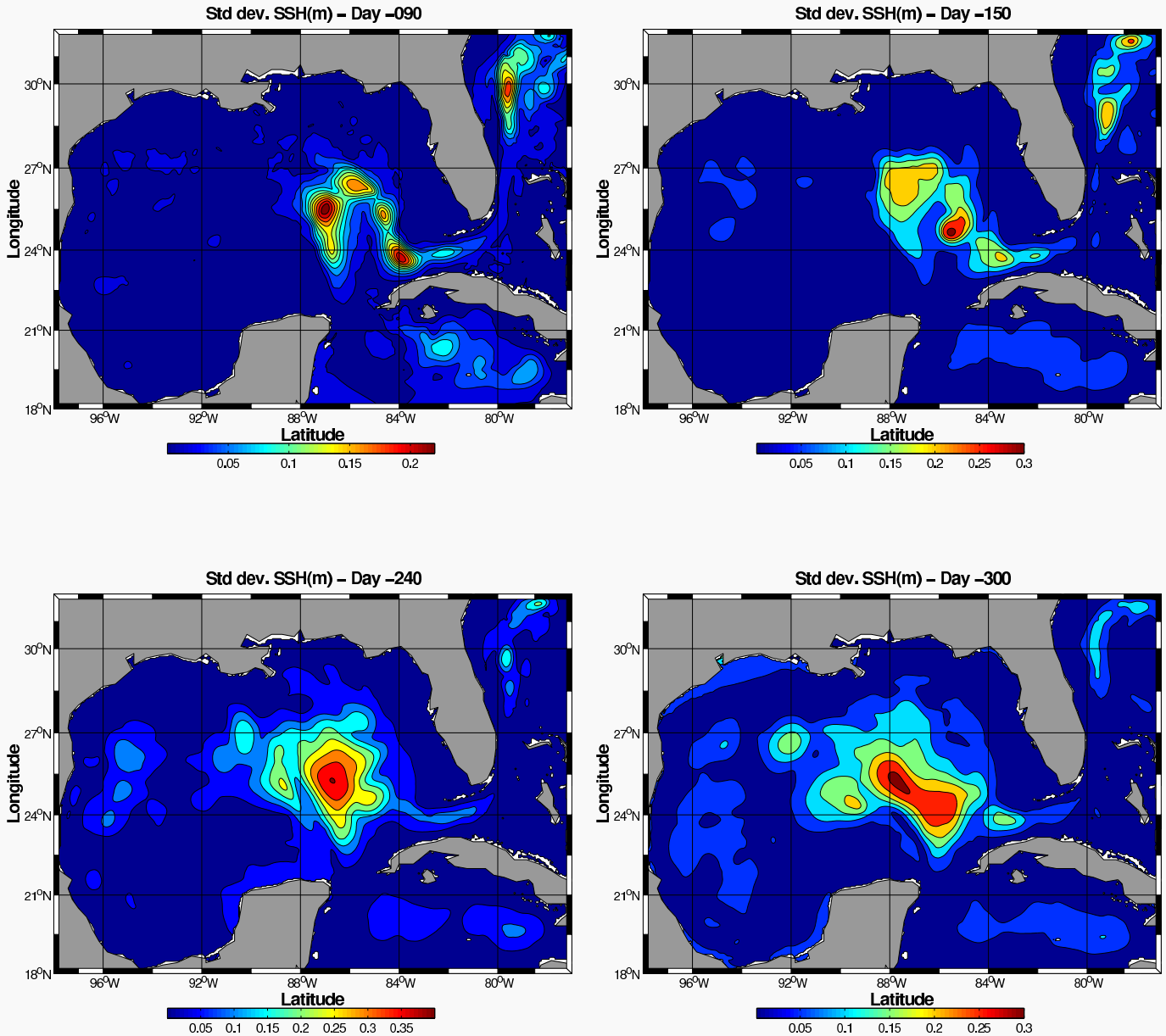
Mean of the Solution obtained from the PC expansion



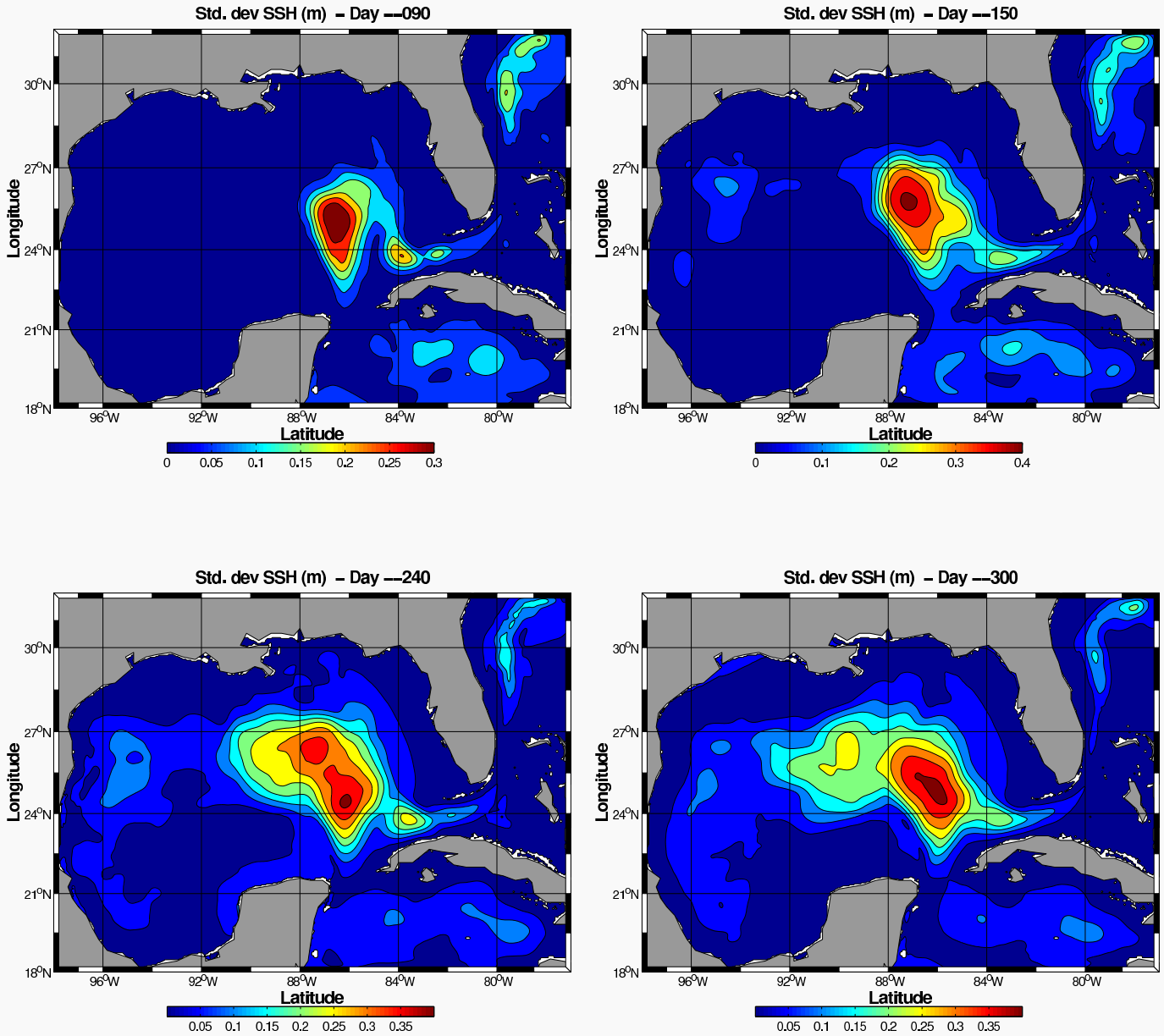
Simple Ensemble Mean of the 36 Realizations



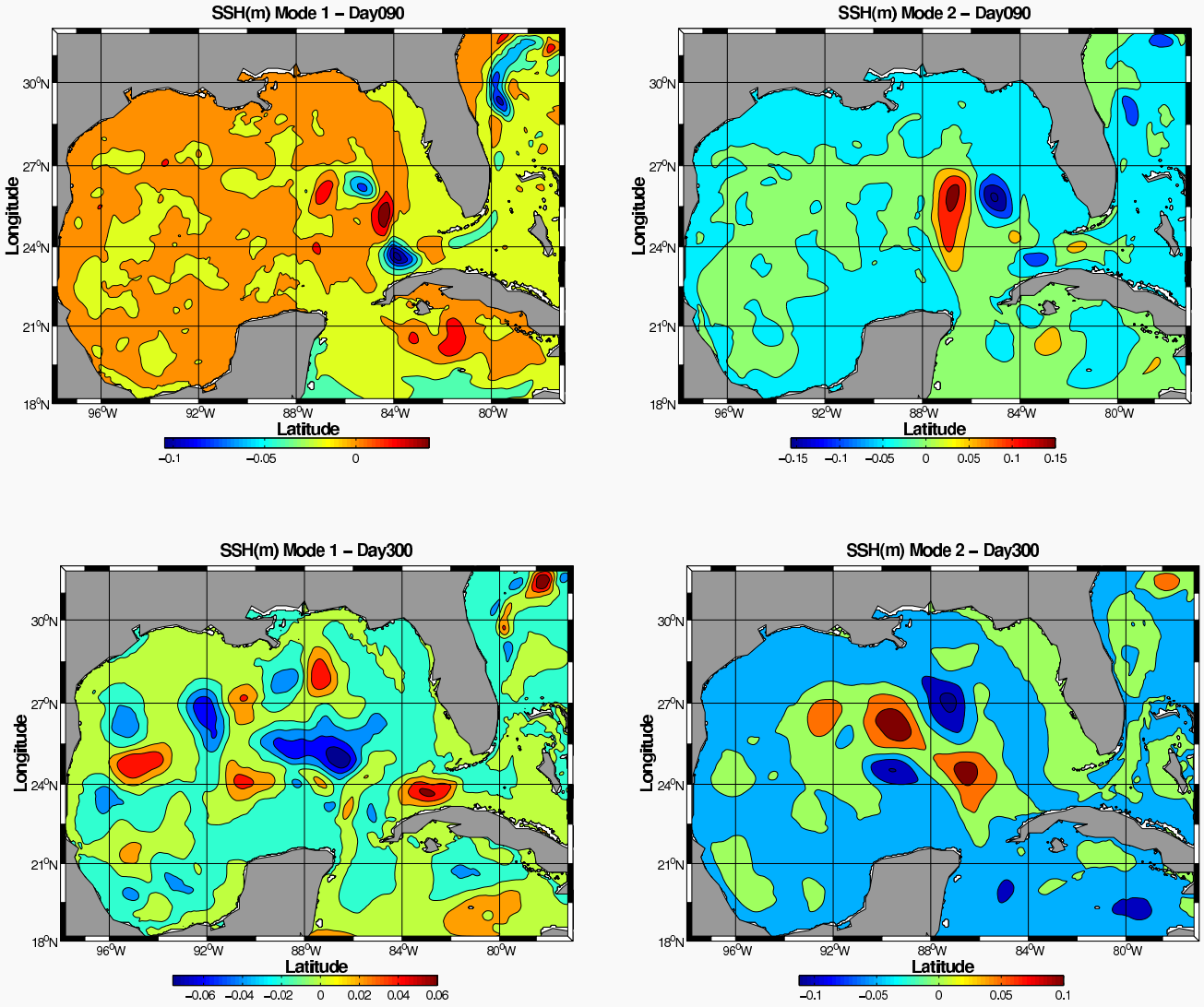
Std dev. of the Solution obtained from the PC expansion



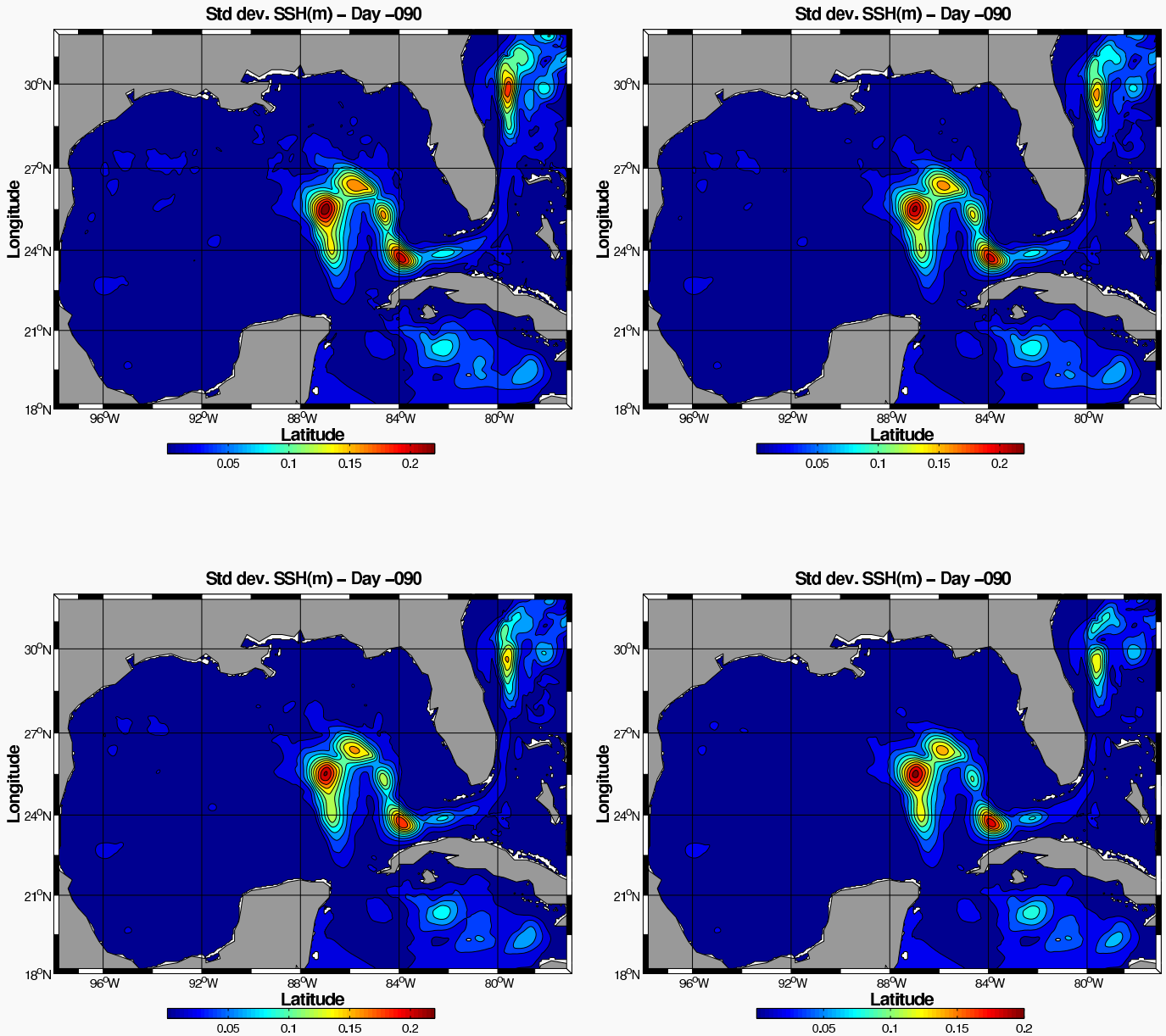
Std dev. of the 36 Realizations



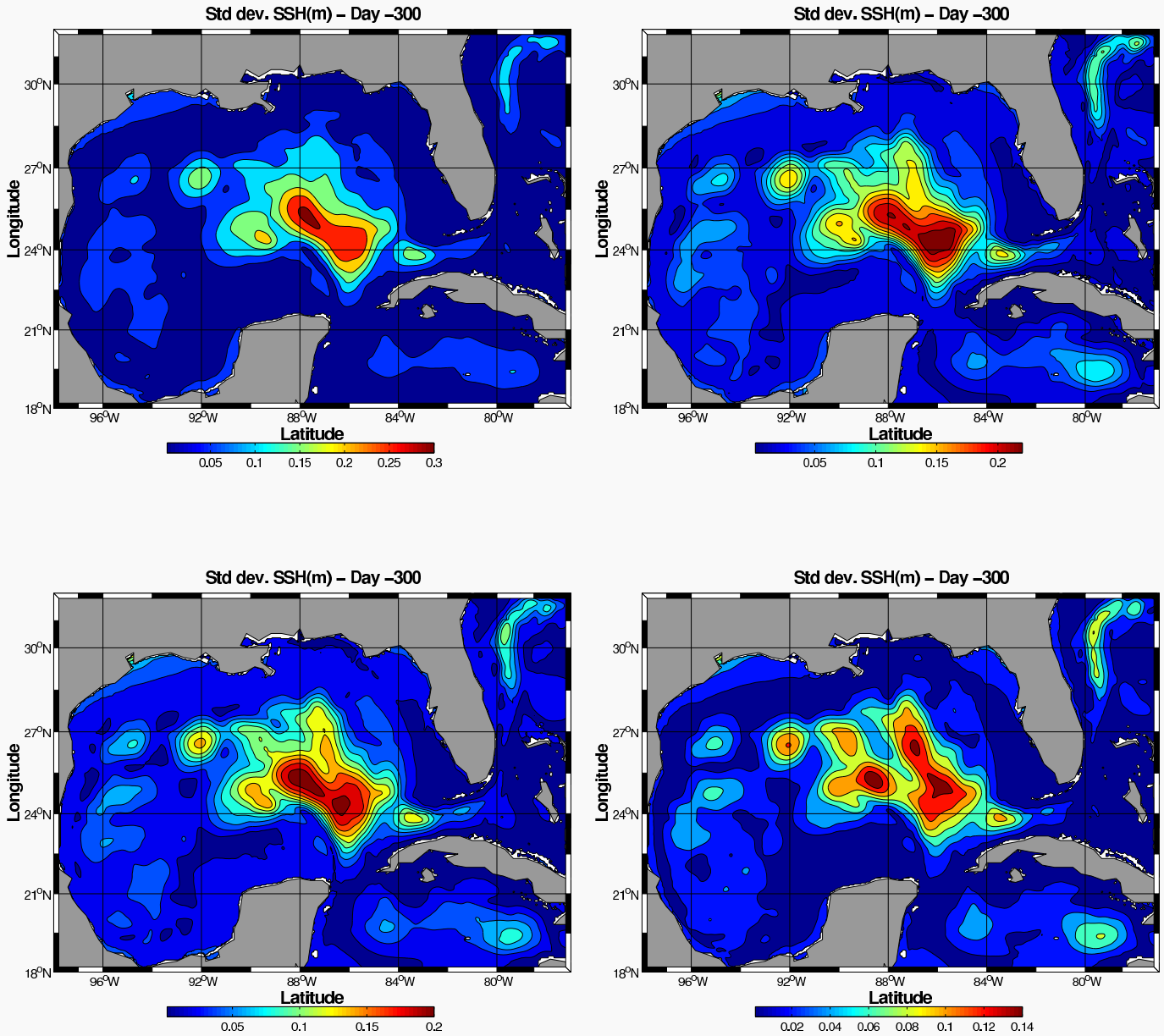
PC Modes


SSH(m) Mode 1 – Day090
SSH(m) Mode 2 – Day090
SSH(m) Mode 1 – Day300
SSH(m) Mode 2 – Day300
Longitude
Latitude
30°N
27°N
24°N
21°N
18°N
96°W
92°W
88°W
84°W
80°W
−0.1 −0.05 0
−0.15 −0.1 −0.05 0 0.05 0.1 0.15
−0.06 −0.04 −0.02 0 0.02 0.04 0.06
−0.1 −0.05 0 0.05 0.1

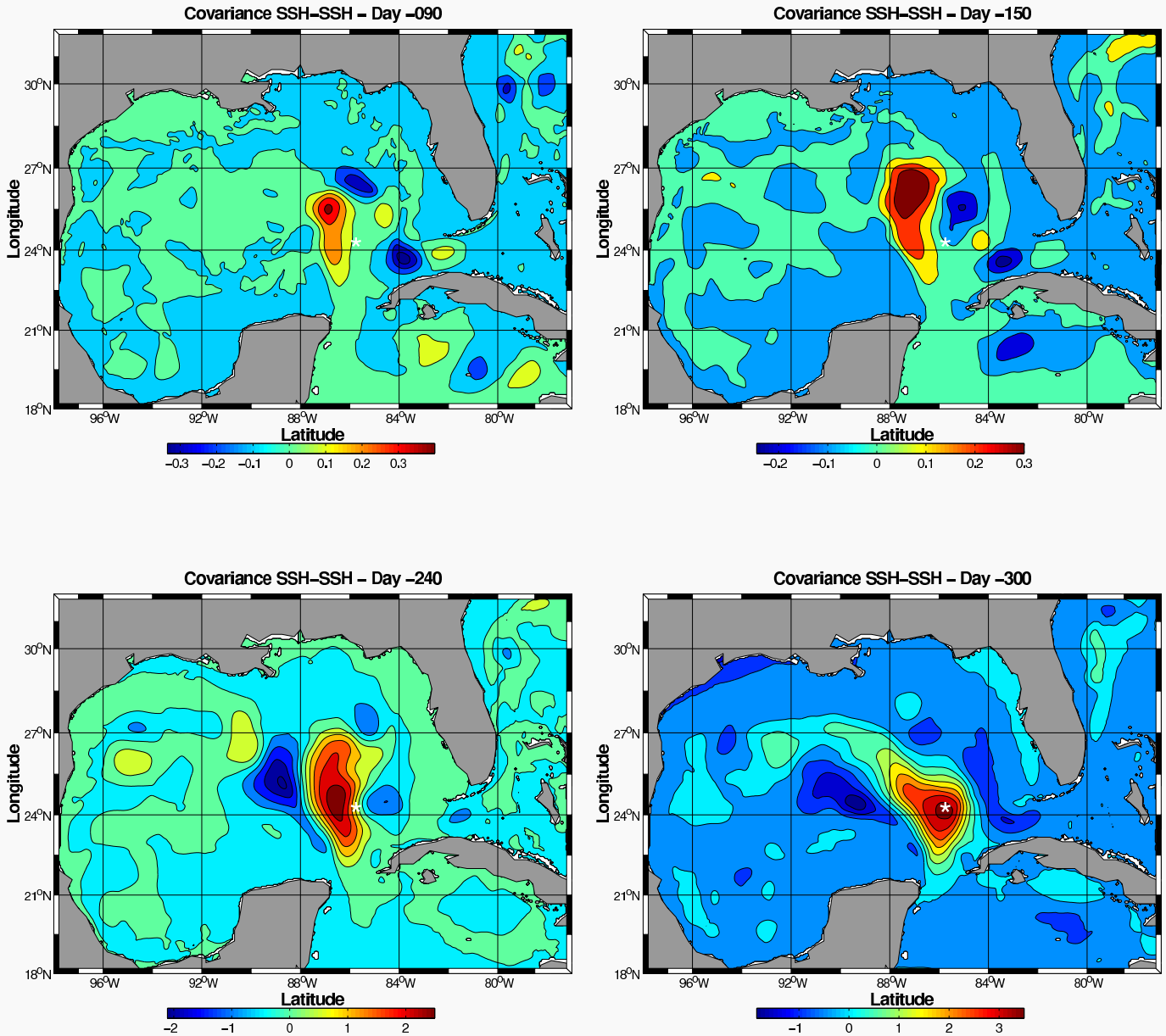
Convergence Day 90



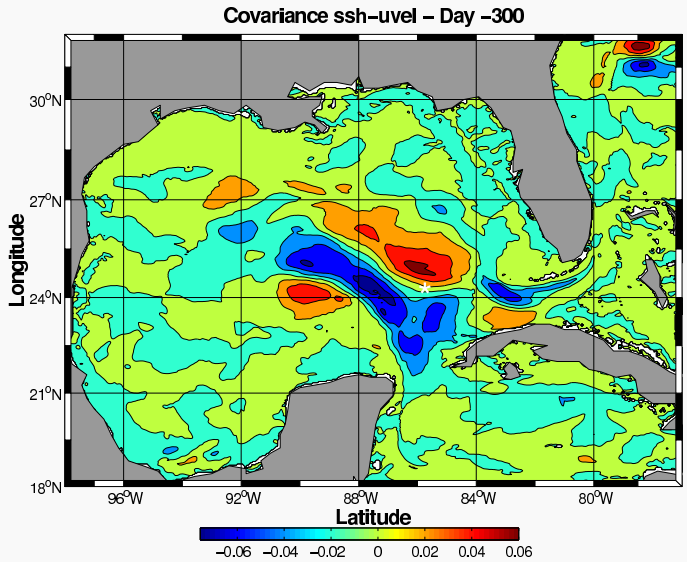
Convergence Day 300

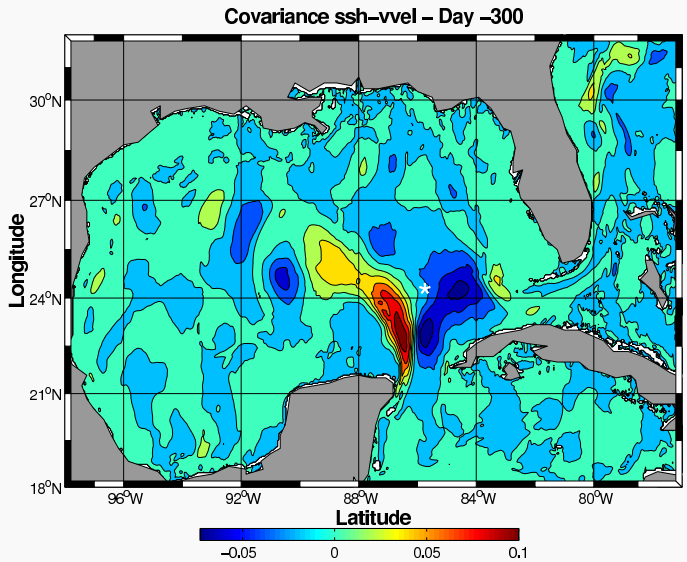


SSH - Auto Covariance from the PC expansion



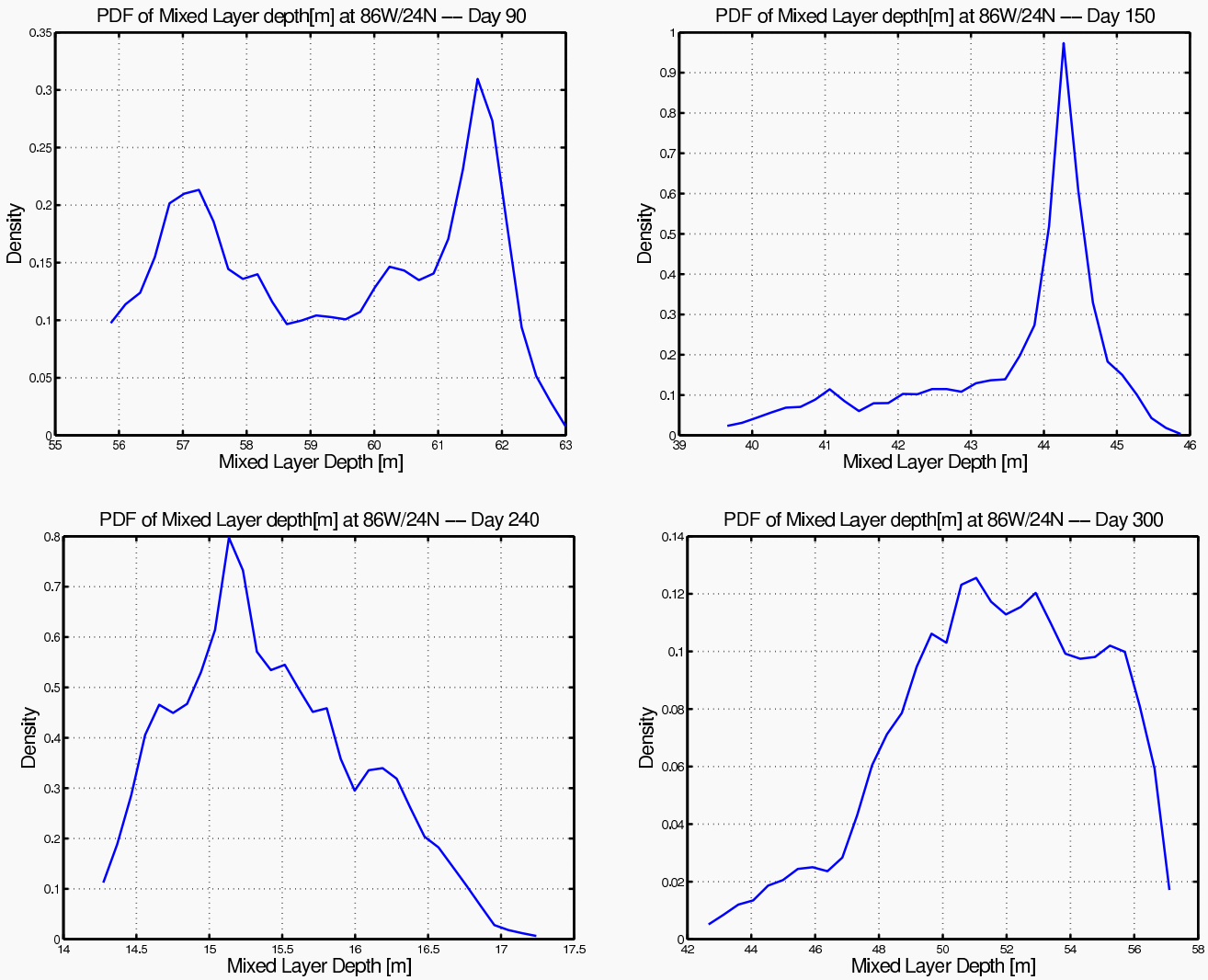
SSH - U/V Cross Covariance from the PC expansion





PDFs of Mixed Layer Depth at 24N/86W

Obtained by sampling the random variables without re-running HYCOM.



Summary

- PC methodology is a viable alternative to other methods for exploring the consequences of a small number of important but uncertain inputs.
- PC coefficients can be used to reduce uncertainty through data assimilation
- Once we have the PC coefficients we can generate PDF's to quantify uncertainty without re-running the model