Tutorial on Uncertainty Quantification with Emphasis on Polynomial Chaos Methods

Mohamed Iskandarani Ashwanth Srinivasan Carlisle Thacker,
Shuyi Chen Chiaying Lee,
University of Miami
Omar Knio Ihab Sraj Alen Alexandrian Justin Winokur,
Duke University
Youssef Marzouk Patrick Conrad,
MIT

October 3, 2013

Outline

What is UQ

Uncertainty Quantification revolves around:

- **Identification**: What are the uncertainty sources?,
- **Characterization**: aleatoric (intrisically random) or epistemic (fixed but have unknown values) Characterization may be scale dependent
- **Forward Propagation**: Propagate input uncertainty through numerical model to calculate output uncertainty
- **Inverse Propagation**: Use observations/experiments to correct *input* uncertainties
- **Sensitivity Analysis**: Which uncertainties contribute the most to output uncertainties
- **Reduction**: Improve forecast by assimilating observations

UQ assesses confidence in model predictions and allows resource allocation for fidelity improvements

Quantifying Ocean Model Uncertainties

- Model equations
- Initial Conditions: Observation sparse in space-time
- Boundary Conditions
  - Momentum, heat and fresh water fluxes
  - Lateral Boundary Conditions in Regional Models
  - Bottom boundary conditions
- Parameterization of small scale processes
  - mixed layer and bottom boundary layer parameters
  - bulk formula for air-sea fluxes

Predictive simulation requires careful assessment of all sources of error and uncertainty

UQ Approaches

- Many UQ approaches exist fulfilling specific needs.
- Emphasis here will be on representation of uncertain variables
- Emphasis on Forward Propagation which enables analysis and inverse propagation
- Topics centered on Generalized Polynomial Chaos methods (reflecting the presentor biases and experience).

What is Polynomial Chaos (PC)

PC combines probabilistic and approximation frameworks to express dependency of model outputs on uncertain model inputs

- Series representation:

$$M(\boldsymbol{x}, t, \xi) \approx M_P \doteq \sum_{k=0}^{P} \widehat{M}_k(\boldsymbol{x}, t)\psi_k(\xi) \tag{1}$$

  - ξ: uncertain input characterized by its PDF $\rho(\xi)$
  - $M(\boldsymbol{x}, t, \xi)$: model output aka observable
  - $\widehat{M}_k(\boldsymbol{x}, t)$: series coefficients
  - $\psi_k(\xi)$: basis (shape) functions in ξ-space
- Basic Questions
  - How to choose ψ_k?
  - How to determine the coefficients $\widehat{M}_k$?
  - Where to truncate the series?

Benefit of functional representation

What can you do with a series?

- **Sum** series to interpolate in ξ-space
  - series is computationally (much) cheaper than a complex model
  - can sum it millions of time to build histogram or effect Monte Carlo sampling
- **Integrate** in ξ-space for statistical moments
  - Mean: $E[M] = \int M\rho(\xi)\mathrm{d}\xi = \sum_k \widehat{M}_k \int \rho(\xi)\psi_k(\xi)\mathrm{d}\xi$
  - Variance: $var[M] = \int \left(\sum_k \widehat{M}_k \psi_k(\xi) - E[M] \right)^2 \rho(\xi)\mathrm{d}\xi$
- **Differentiate** in ξ-space (no adjoint code!)

$$\frac{\partial M}{\partial \xi} = \sum_k \widehat{M}_k \frac{\partial \psi_k}{\partial \xi}$$

Series must be reliable to reap benefits

Example 1 of input uncertainties and $\rho(\xi)$

- Drag Coefficient is uncertain: $C_D = \alpha C_D^{ref}$
  - α is a multiplicative factor, with $\alpha \in [\alpha_{\min}, \alpha_{\max}]$
  - Map it to standard interval $-1 \leq \xi \leq 1$
    $\alpha = (\alpha_{\max} - \alpha_{\min})\frac{\xi+1}{2} + \alpha_{\min}$
  - If all values are equally likely than $\rho(\xi) = \frac{1}{2}$.
  - To weigh an area more than others choose a beta distribution:

$$\begin{aligned} \rho(\xi) &= \frac{(1+\xi)^{\alpha}(1-\xi)^{\beta}}{2^{\alpha+\beta+1}B(\alpha+1,\beta+1} \\ E[\xi] &= \frac{\alpha+1}{\alpha+\beta+2} \\ var[\xi] &= \frac{(\alpha+1)(\beta+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)} \end{aligned}$$

Example 2 of input uncertainties and $\rho(\xi)$

Uncertainty in Initial Boundary Conditions via Empirical Orthogonal Functions perturbations:

$$u(\boldsymbol{x}, 0, \xi_1, \xi_2) = \overline{u}(\boldsymbol{x}, 0) + \left[\sqrt{\lambda_1}\mathcal{U}_1\xi_1 + \sqrt{\lambda_2}\mathcal{U}_2\xi_2\right] \qquad (2)$$

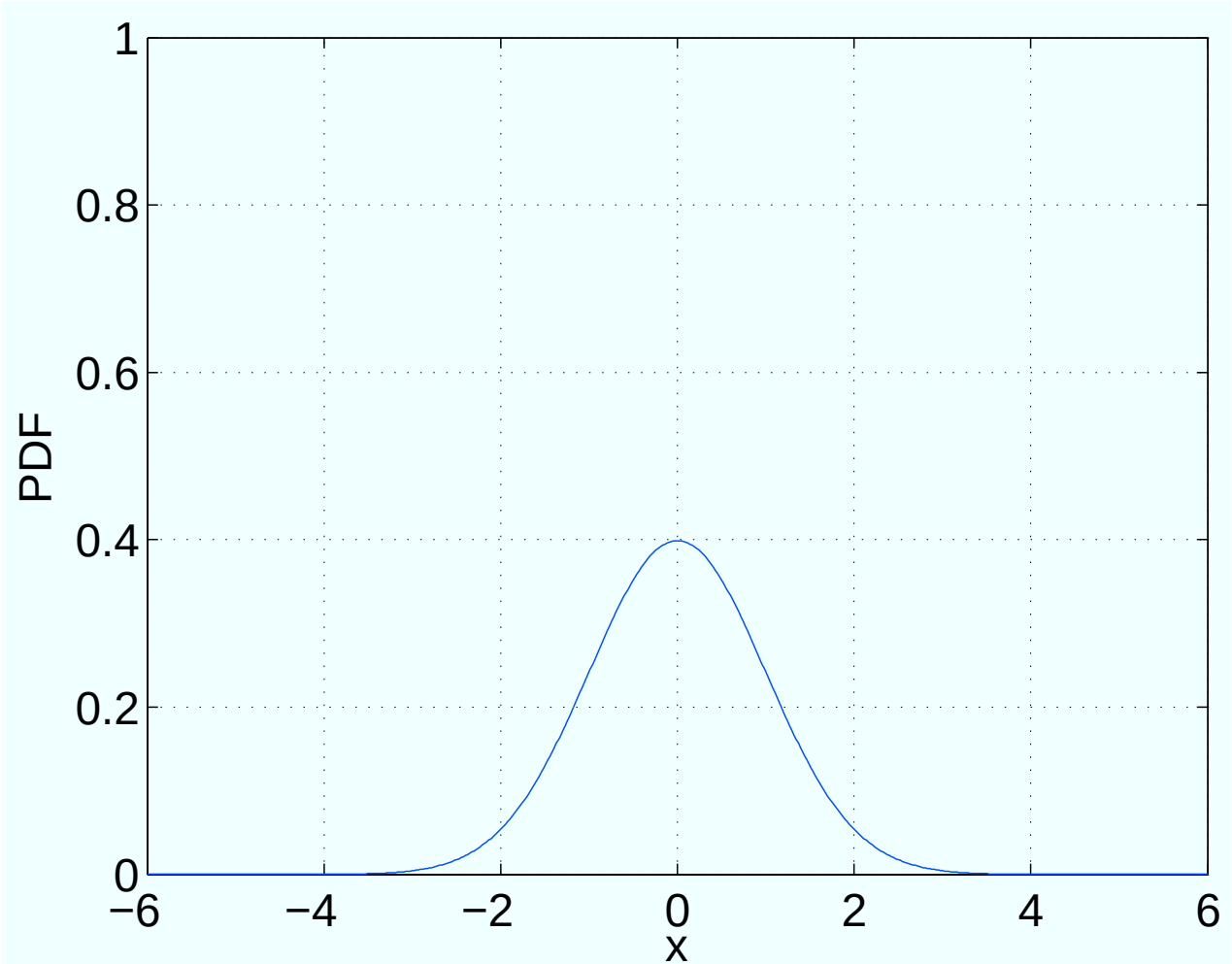
- $(\lambda_k, \mathcal{U}_k)$: are eigenvalues/eigenvectors of covariance matrix obtained from free-run simulation
- $\overline{u}$: unperturbed initial condition
- $u(\boldsymbol{x}, 0, \xi_1, \xi_2)$: Stochastic initial condition input
- The two *independent* uncertain variables are the modes amplitudes: $\xi_{1,2}$
- Uniform distributions: $\rho(\xi_{1,2}) = \frac{1}{2}$
- Gaussian distributions: $\rho(\xi_{1,2}) = \frac{e^{-\frac{\xi_{1,2}^2}{2}}}{\sqrt{2\pi}}$

Polynomial Chaos Basis

ξ-distribution	Domain	weight $\rho(\xi)$	basis $\psi_k(\xi)$	parameter
Gauss	$(-\infty, \infty)$	$\frac{e^{-\frac{\xi^2}{2}}}{\sqrt{2\pi}}$	Hermite	none
Gamma	$(0, \infty)$	$\frac{\xi^\alpha e^{-\xi}}{\Gamma(\alpha+1)}$	Laguerre	$\alpha > 1$
Beta	$[-1, 1]$	$\frac{(1+\xi)^\alpha (1-\xi)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}$	Jacobi	$\alpha, \beta > 1$
Uniform	$[-1, 1]$	$\frac{1}{2}$	Legendre	none

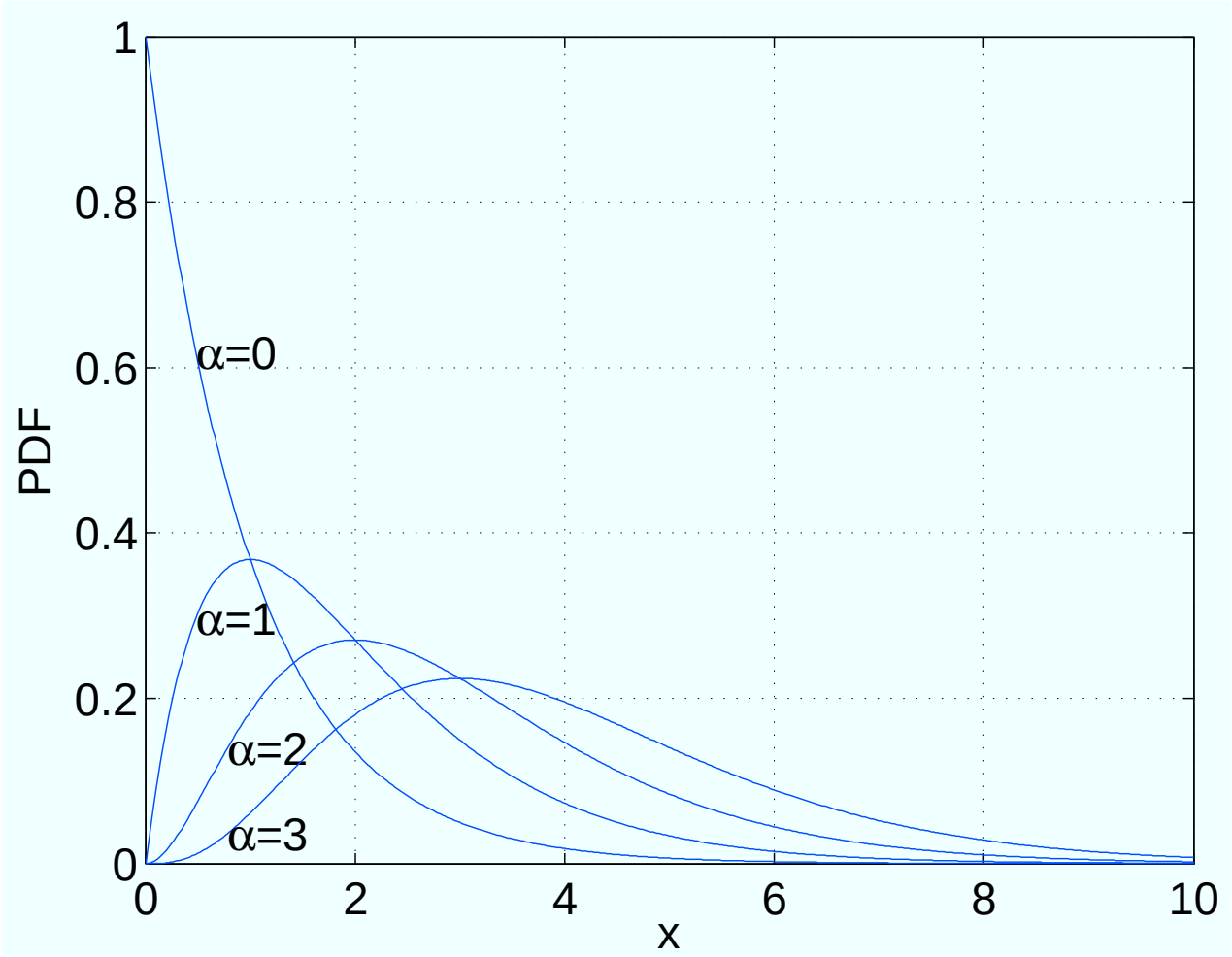
- Inner Product in ξ-space: $\langle \psi_j, \psi_k \rangle = \int \psi_k(\xi)\, \psi_j(\xi)\, \rho(\xi) \mathrm{d}\xi$
- Polynomial basis is orthonormal w.r.t. $\rho(\xi)$: $\langle \psi_j, \psi_k \rangle = \delta_{i,j}$
- Input parameter domain and distribution often dictate the most convenient basis. $\langle \psi_j, \psi_k \rangle = \delta_{i,j}$
- Wiener-Askey scheme provides a hierarchy of possible continuous PC bases

Normal Distribution



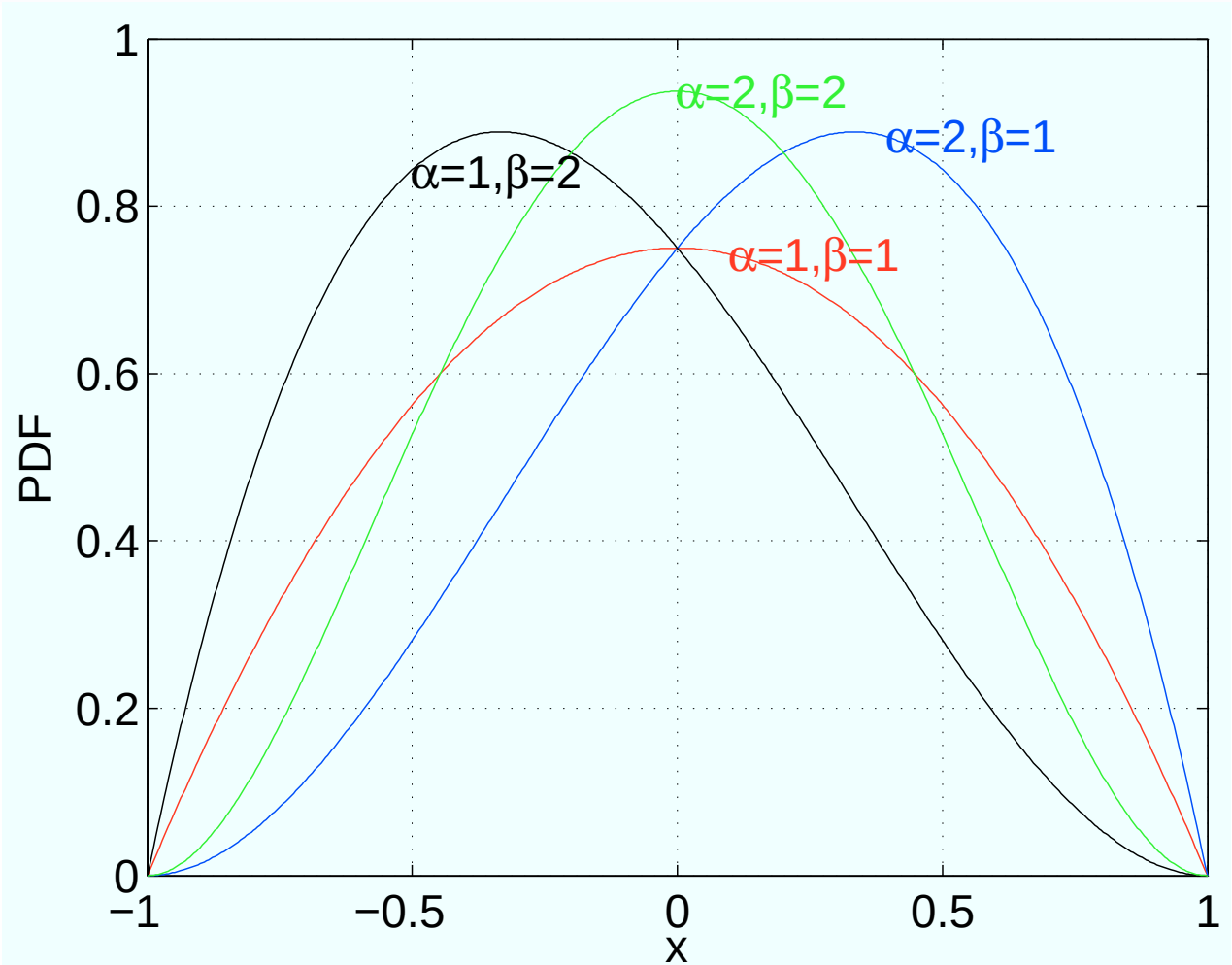
- Most commonly used input distribution
- Support on $(-\infty, \infty)$

Gamma Distribution



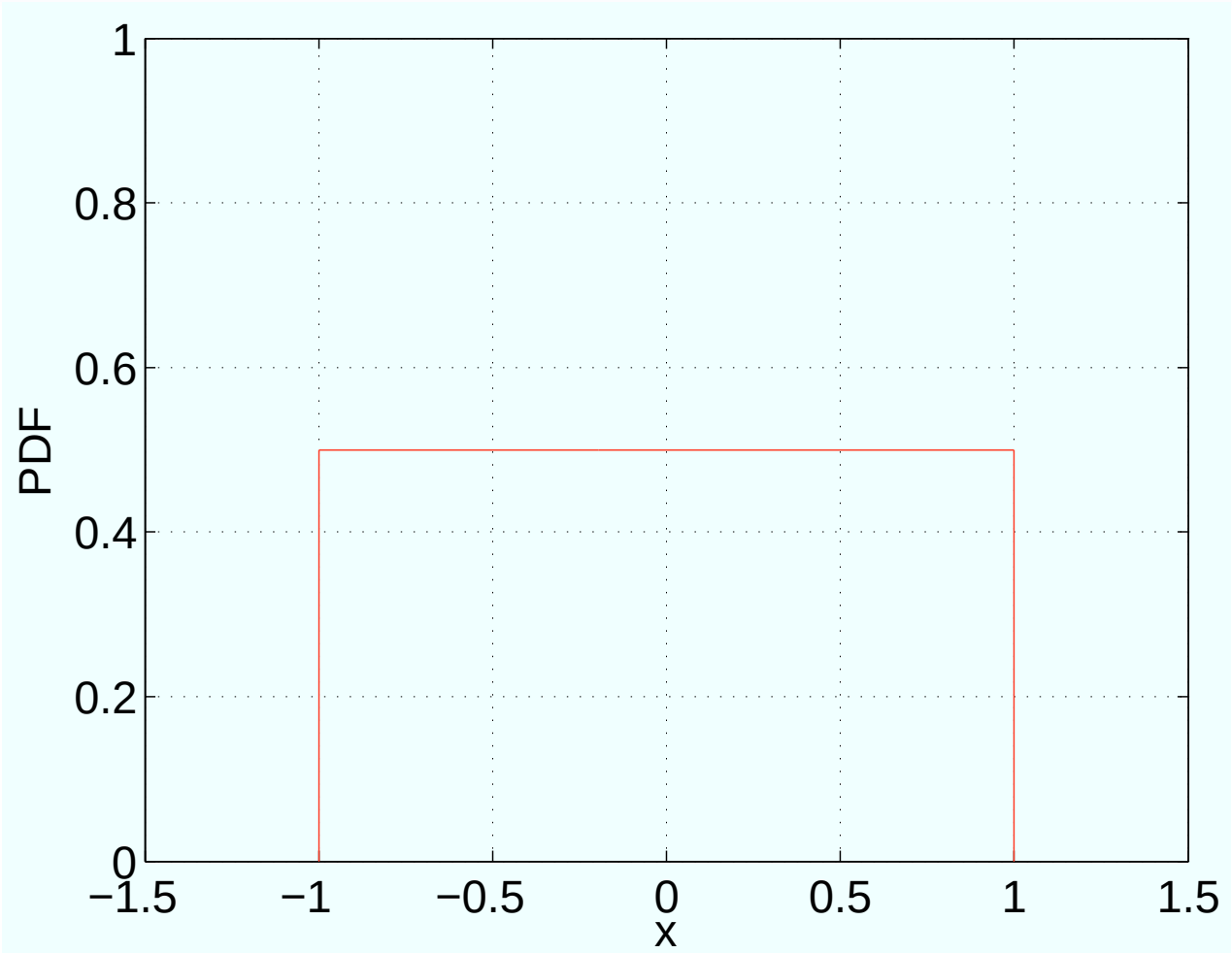
- Useful to represent uncertainties in positive quantities.
- Support on $(0, \infty)$

Beta Distribution



- Useful for uncertainties that varies between set quantities.
- Can be tailored to weigh some values more than others
- Support on $[-1, 1]$

Uniform Distribution



- Useful for uncertainties with sharp bounds
- or not much is known about input distribution
- Support on $[-1, 1]$

Polynomial Chaos Basis

- Series: $M(\boldsymbol{x}, t, \xi) = \sum_{k=0}^{P} \widehat{M}_k(\boldsymbol{x}, t)\psi_k(\xi)$
- Expectation:

$$E[\psi_k] = \int \psi_k(\xi)\rho(\xi)\mathrm{d}\xi = \langle \psi_k, \psi_0 \rangle = \delta_{k,0}$$

- mean:

$$E[M] = \sum_{k=0}^{P} u_k(\boldsymbol{x}, t)E[\psi_k(\xi)] = u_0(\boldsymbol{x}, t)$$

- Variance:

$$E\left[(M - E[M])^2\right] = \sum_{k=1}^{P} \widehat{M}_k^2(\boldsymbol{x}, t)$$

- Covariance:

$$E[\ (u - E[u])\ (v - E[v])\ ] = \sum_{k=1}^{P} u_k(\boldsymbol{x})v_k(\boldsymbol{x}, t)$$

Multidimensional basis

Multi-dimensional basis functions $\Psi_k(\xi_1, \xi_2, \ldots, \xi_n)$ are tensor products of 1D basis functions:

$$\Psi_k(\xi_1, \xi_2, \ldots, \xi_n) = \psi_{\alpha_1^k}(\xi_1)\psi_{\alpha_2^k}(\xi_2)\ldots\psi_{\alpha_3^k}(\xi_n)$$

- 1D Legendre basis: $L_0(\xi) = 1$, $L_1(\xi) = \xi$, $L_2(\xi) = \frac{3\xi^2-1}{2}$
- 2D Example

  $\psi_0 = L_0(\xi_1)L_0(\xi_2)$ $\psi_2 = L_0(\xi_1)L_1(\xi_2)$ $\psi_5 = L_0(\xi_1)L_2(\xi_2)$ $\psi_9 = L_0(\xi_1)L_3(\xi_2)$
  $\psi_1 = L_1(\xi_1)L_0(\xi_2)$ $\psi_4 = L_1(\xi_1)L_1(\xi_2)$ $\psi_8 = L_1(\xi_1)L_2(\xi_2)$
  $\psi_3 = L_2(\xi_1)L_0(\xi_2)$ $\psi_7 = L_2(\xi_1)L_1(\xi_2)$
  $\psi_6 = L_3(\xi_1)L_0(\xi_2)$
- Triangular truncation is common, max order=3
- number of coefficient is $P + 1 = \frac{(N+p)!}{N!p!}$
  N is the number of stochastic variables
  p is the max polynomial degree in 1D

How do we determine PC coefficients?

- Series: $M(\boldsymbol{x}, t, \xi) = \sum_{k=0}^{P} \widehat{M}_k(\boldsymbol{x}, t)\psi_k(\xi)$
- Galerkin Projection on ψ_k basis (minimizes L_2-error norm)

$$\widehat{M}_k(\boldsymbol{x}, t) = \langle M, \psi_k \rangle = \int M(\boldsymbol{x}, t, \xi)\psi_k(\xi)\rho(\xi)\mathrm{d}\xi$$

- Non Intrusive Spectral Projection: Approximate integral numerically via **quadrature**

$$\widehat{M}_k(\boldsymbol{x}, t) \approx \sum_{q=1}^{Q} M(\boldsymbol{x}, t, \xi_q)\psi_k(\xi_q)\omega_q$$

- ξ_q/ω_q quadrature points/weights
- Quadrature requires an ensemble run at ξ_q.
- No code modification is necessary

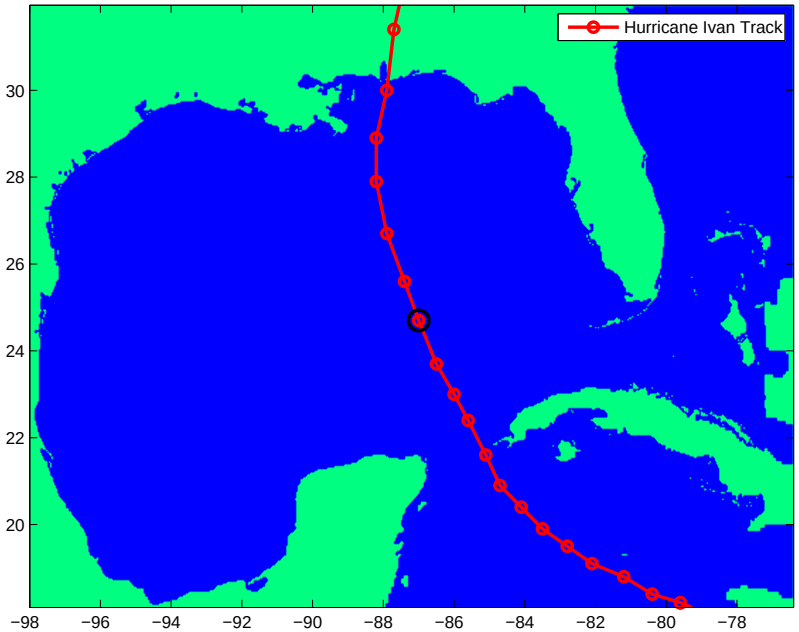
Choice of Quadrature

- Gauss quadrature most accurate/point ($\psi_{p+1}(\xi_q) = 0$) but Naive tensorization cost grows exponentially: p^N.
- Rely on Nested Sparse Smolyak Quadrature Tempers the curse of dimensionality
- Adaptive Quadrature

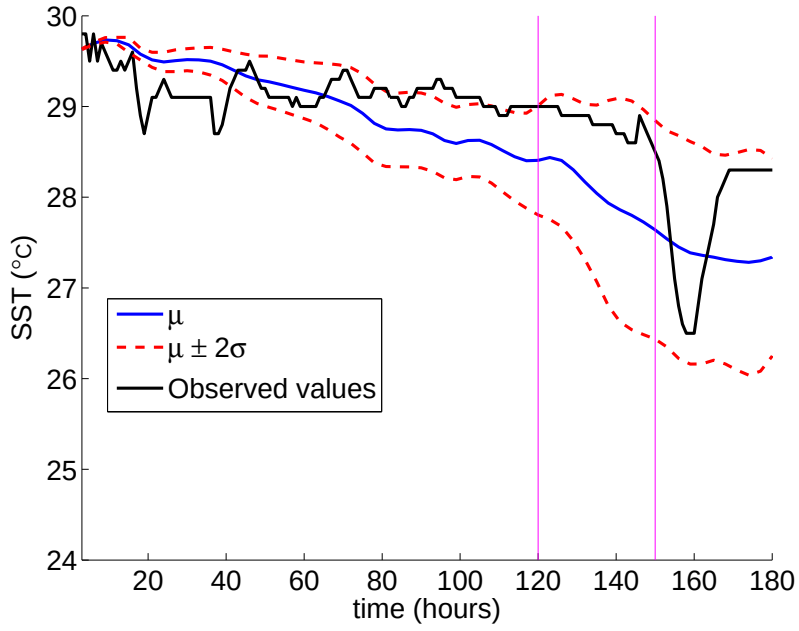
Polynomial Chaos Expansions Summary

- Paradigm shift from statistical to combined probabilistic/approximation view
- Can quantify approximation error and "convergence" to solution
- No a-priori restriction/assumption on output statistics
- Approach robust to model non-linearity and model differentiability
- Can be done non-intrusively via ensembles.
- Multiple independent stochastic variables can be handled by multi-dimensonal tensorization of 1D basis functions and quadratures.
- Sampling Challenges for high N or p

Forward Problem: Parametric Sensitivity



Hurricane Ivan track



Comparing mean & observed SST. Vertical lines show when Ivan enters GoM and when it is nearest buoy.

Description	Range
critical Richardson #	$p_1 \in [0.25, 0.7]$
background viscosity	$p_2 \in [10^{-4}, 10^{-3}]$
background diffusivity	$p_3 \in [10^{-5}, 10^{-4}]$
drag coefficient factor	$p_4 \in [0.2, 1.0]$

Table: HYCOM uncertain inputs.

- Legendre basis with $p = 5$
- 210 unknown coefficients
- Nested sparse Smolyack Ensemble size 385 ($\ll 6^4 = 1,296$ Gauss quadrature)

Variance Analysis

$$T_i = \frac{\text{Variance due to parameter} p_i}{\text{Total variance}}$$

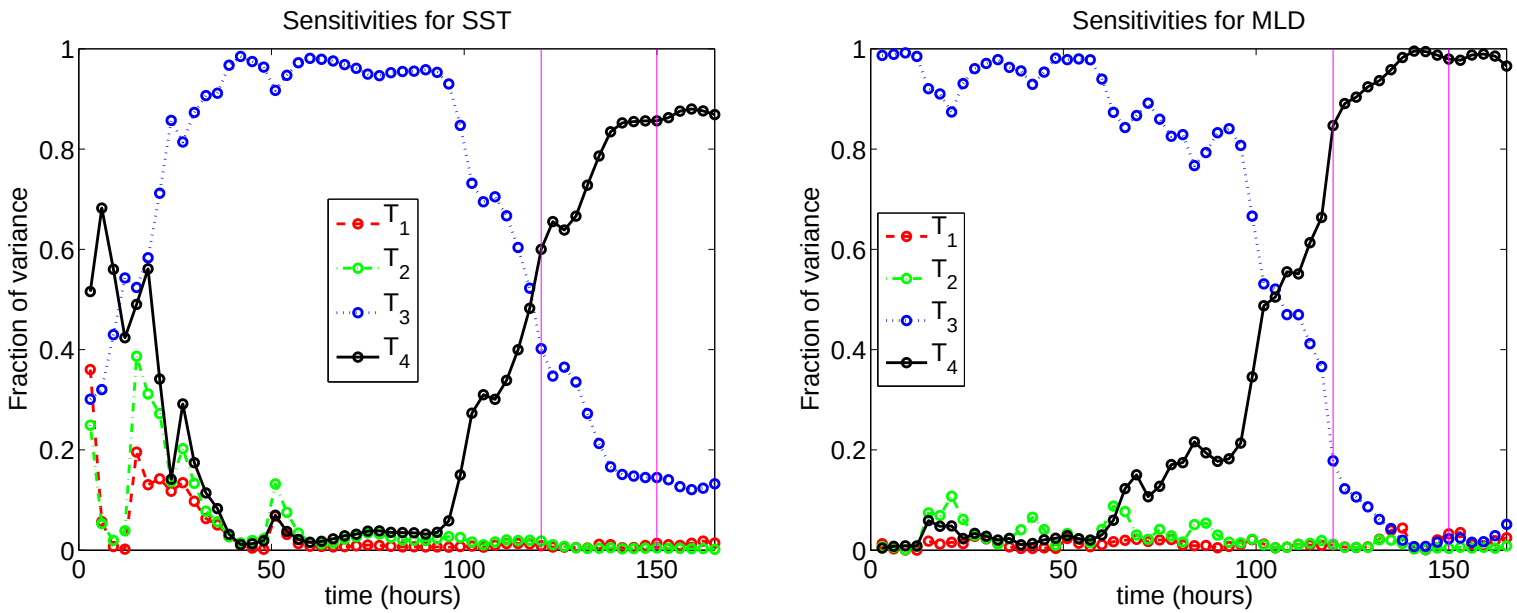


Figure: Evolution of the global sensitivity indices $T_1, \ldots, T_4$ for SST and MLD (bottom). The first vertical line indicates the time the hurricane enters the GOM whereas the second indicates a time at which the hurricane is close to the buoy.

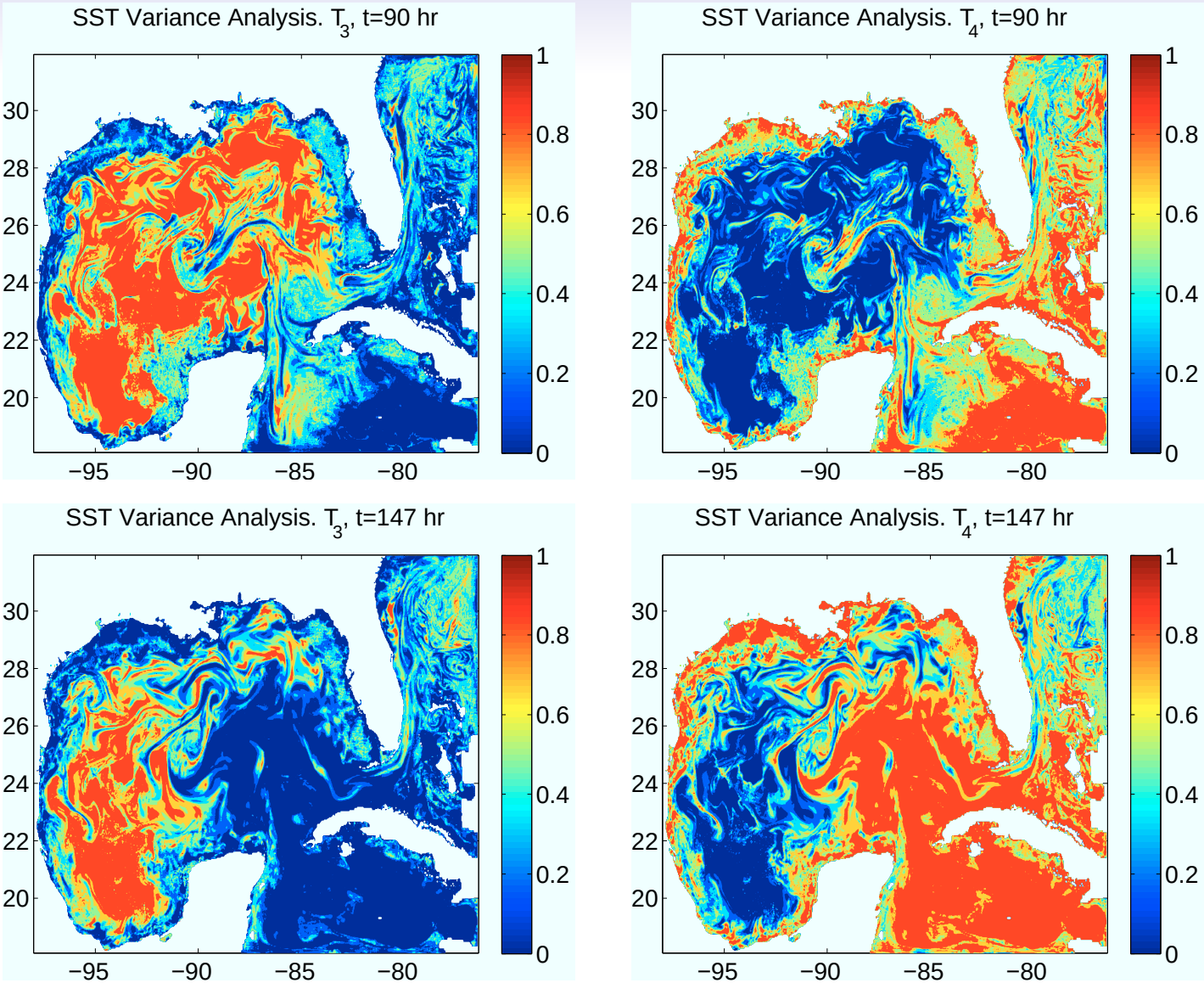
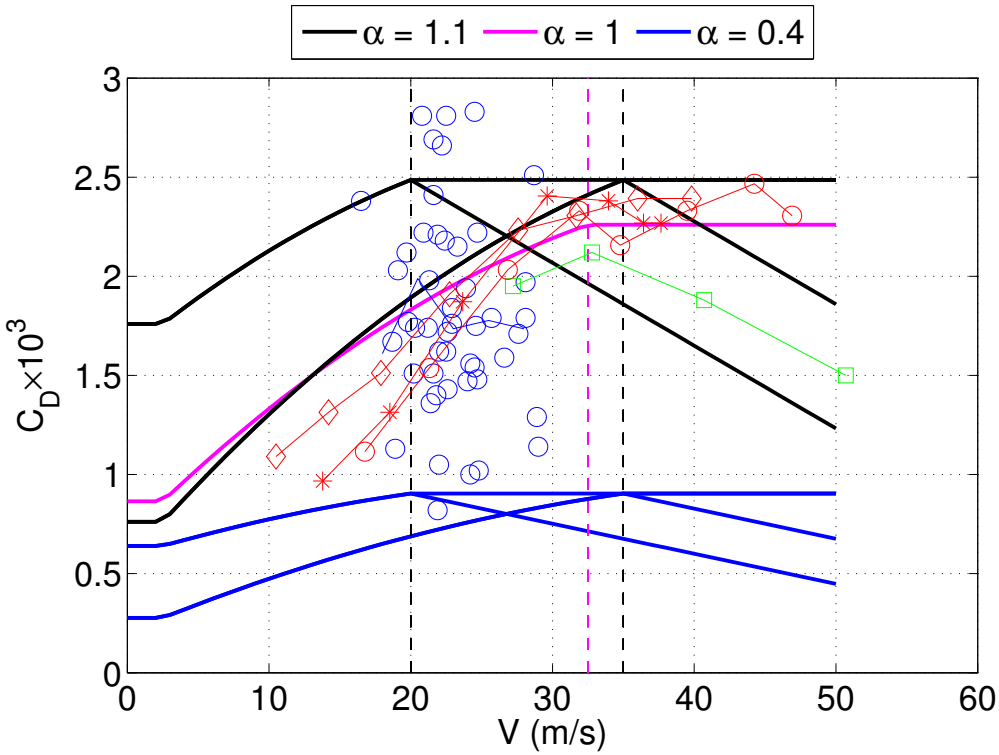


Figure: T_3 (left) and T_4 (right) sensitivity contours for SST. Drag dominates uncertainty during high winds, otherwise it is background diffusivity.



$$\vec{\tau} = \rho_a C_D V \vec{V}$$

$$C_D = C_{D0} + C_{D1}(T_s - T_a)$$

$$C_{D0} = a_0 + a_1 \tilde{V} + a_2 \tilde{V}^2$$

$$C_{D1} = b_0 + b_1 \tilde{V} + b_2 \tilde{V}^2$$

$$\tilde{V} = \max[\, V_{\min},\ \min(V_{\max}, V)$$

C_D is drag coefficient
V is wind speed at 10 m.
C_D saturates for $V > V_{\max}$

- Blue circles: aircraft observations
- red: wind tunnel
- green: drop sondes
- magenta: HYCOM fit to COARE 2.5,
- Problem: $V_{\max}$ and $C_D^{\max}$ are not well-known and does C_D decrease for $V > V_{\max}$ as drop sondes suggest?

Inverse Modeling Problem

- Perturb C_D by introducing 3 control variables (α, $V_{\max}$, m)

$$C_D{}' = \alpha C_D \text{ for } V < V_{\max} \quad (3)$$

$$C_D{}' = \alpha[C_D + m(V - V_{\max})] \text{ for } V > V_{\max} \quad (4)$$

- multiplicative factor $0.4 \leq \alpha \leq 1.1$
- vary $V_{\max}$ between 20 and 35 m/s
- m is a linear slope modeling decrease for $V > V_{\max}$ with $-3.8 \times 10^{-5} \leq m \leq 0$
- Use ITOP data to learn about likely distribution of α, $V_{\max}$ and m.

Bayes Theorem: $p(\boldsymbol{\theta} \mid \boldsymbol{T}) \propto p(\boldsymbol{T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})$

- Likelihood: $\epsilon = \boldsymbol{T} - \boldsymbol{M}$ is normally distributed

$$p(\boldsymbol{T} \mid \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right) \tag{5}$$

Bayes Theorem: $p(\boldsymbol{\theta} \mid \boldsymbol{T}) \propto p(\boldsymbol{T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})$

- Likelihood: $\epsilon = \boldsymbol{T} - \boldsymbol{M}$ is normally distributed

$$p(\boldsymbol{T} \mid \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right) \tag{5}$$

- σ^2 unknown, treated as hyper-parameter. Assume a Jeffreys prior

$$p(\sigma^2) = \begin{cases} \frac{1}{\sigma^2} & \text{for } \sigma^2 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Bayes Theorem: $p(\boldsymbol{\theta} \mid \boldsymbol{T}) \propto p(\boldsymbol{T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})$

- Likelihood: $\epsilon = \boldsymbol{T} - \boldsymbol{M}$ is normally distributed

$$p(\boldsymbol{T} \mid \theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right) \tag{5}$$

- σ^2 unknown, treated as hyper-parameter. Assume a Jeffreys prior

$$p(\sigma^2) = \begin{cases} \frac{1}{\sigma^2} & \text{for } \sigma^2 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

- Uninformed priors for α, $V_{\max}$ and m:

$$p(\{\alpha, V_{\max}, m\}) = \begin{cases} \frac{1}{b_i - a_i} & \text{for } a_i \leq \{\alpha, V_{\max}, m\} \leq b_i, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $[a_i, b_i]$ denote the parameter ranges.

Final Form of Bayes theorem

$$p(\{\alpha, V_{\max}, m\}, \sigma^2 | \boldsymbol{T}) \propto \left[\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right)\right] p(\sigma^2)\, p(\alpha)\, p(V_{\max})\, p(m)$$

Final Form of Bayes theorem

$$p(\{\alpha, V_{\max}, m\}, \sigma^2 | \boldsymbol{T}) \propto \left[\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right)\right] p(\sigma^2)\, p(\alpha)\, p(V_{\max})\, p(m)$$

- Build full posterior with Markov Chain Monte Carlo (MCMC)
  MCMC requires $O(10^5)$ estimates of M_i: prohibitive

Final Form of Bayes theorem

$$p(\{\alpha, V_{\max}, m\}, \sigma^2 | \boldsymbol{T}) \propto \left[\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right)\right] p(\sigma^2)\, p(\alpha)\, p(V_{\max})\, p(m)$$

- Build full posterior with Markov Chain Monte Carlo (MCMC)
  MCMC requires $O(10^5)$ estimates of M_i: prohibitive
- Solve for center and spread of posterior
  minimization problem requiring access to cost function gradient and Hessian: Needs an adjoint model

Final Form of Bayes theorem

$$p(\{\alpha, V_{\max}, m\}, \sigma^2 | \boldsymbol{T}) \propto \left[\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(T_i - M_i)^2}{2\sigma^2}\right)\right] p(\sigma^2)\, p(\alpha)\, p(V_{\max})\, p(m)$$

- Build full posterior with Markov Chain Monte Carlo (MCMC) MCMC requires $O(10^5)$ estimates of M_i: prohibitive
- Solve for center and spread of posterior minimization problem requiring access to cost function gradient and Hessian: Needs an adjoint model
- Rely on Polynomial Chaos expansions to replace HYCOM by a polynomial series that could be either summed for MCMC or differentiated for the gradients.

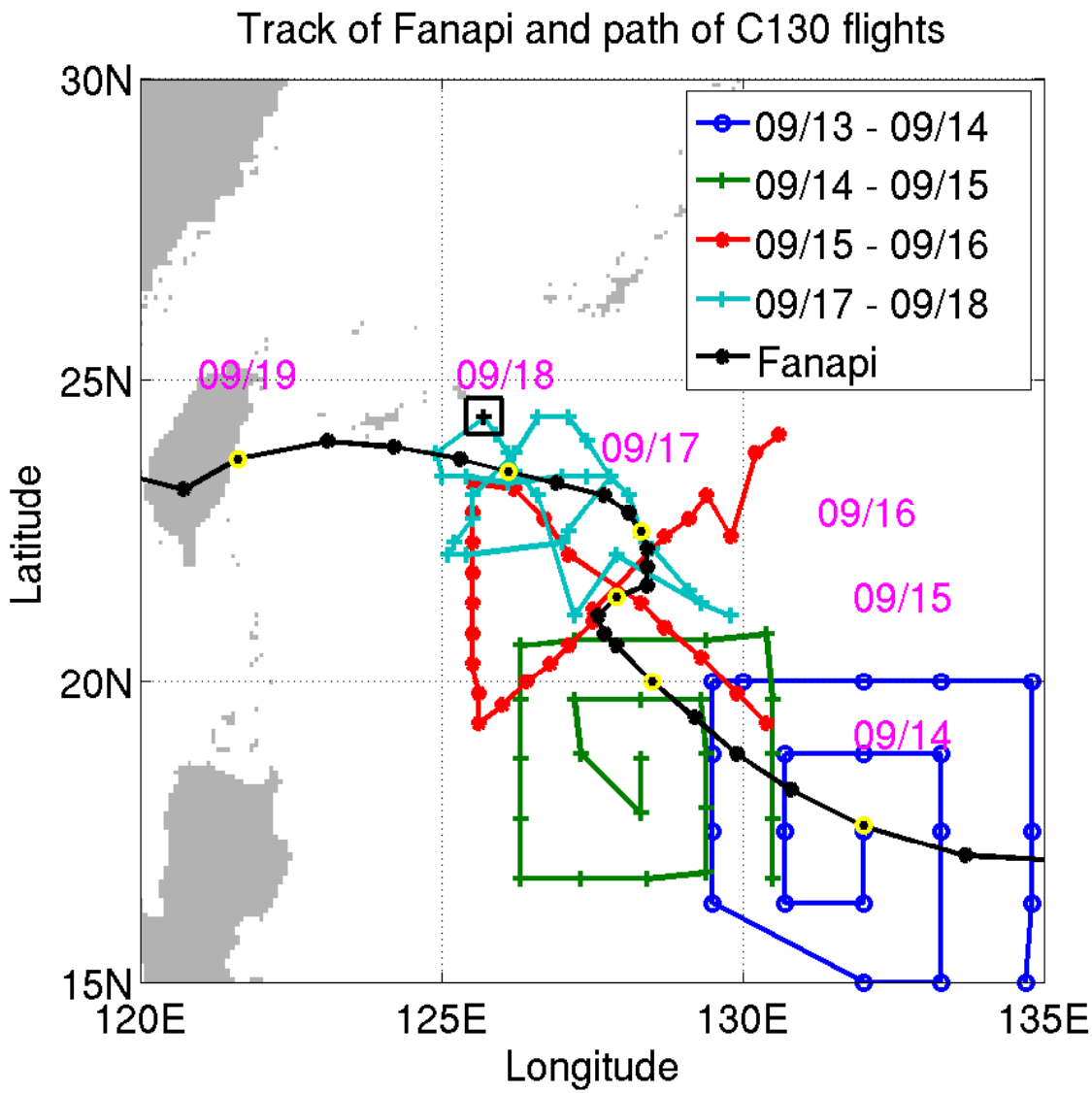


Figure: Fanapi's JTWC track (black curve) and paths of C-130 flights. The yellow circles on the track represent the typhoon center at 00:00 UTC. The circles on the flight paths mark the 119 AXBT drops. The 42×42 km^2 analysis box is also shown.

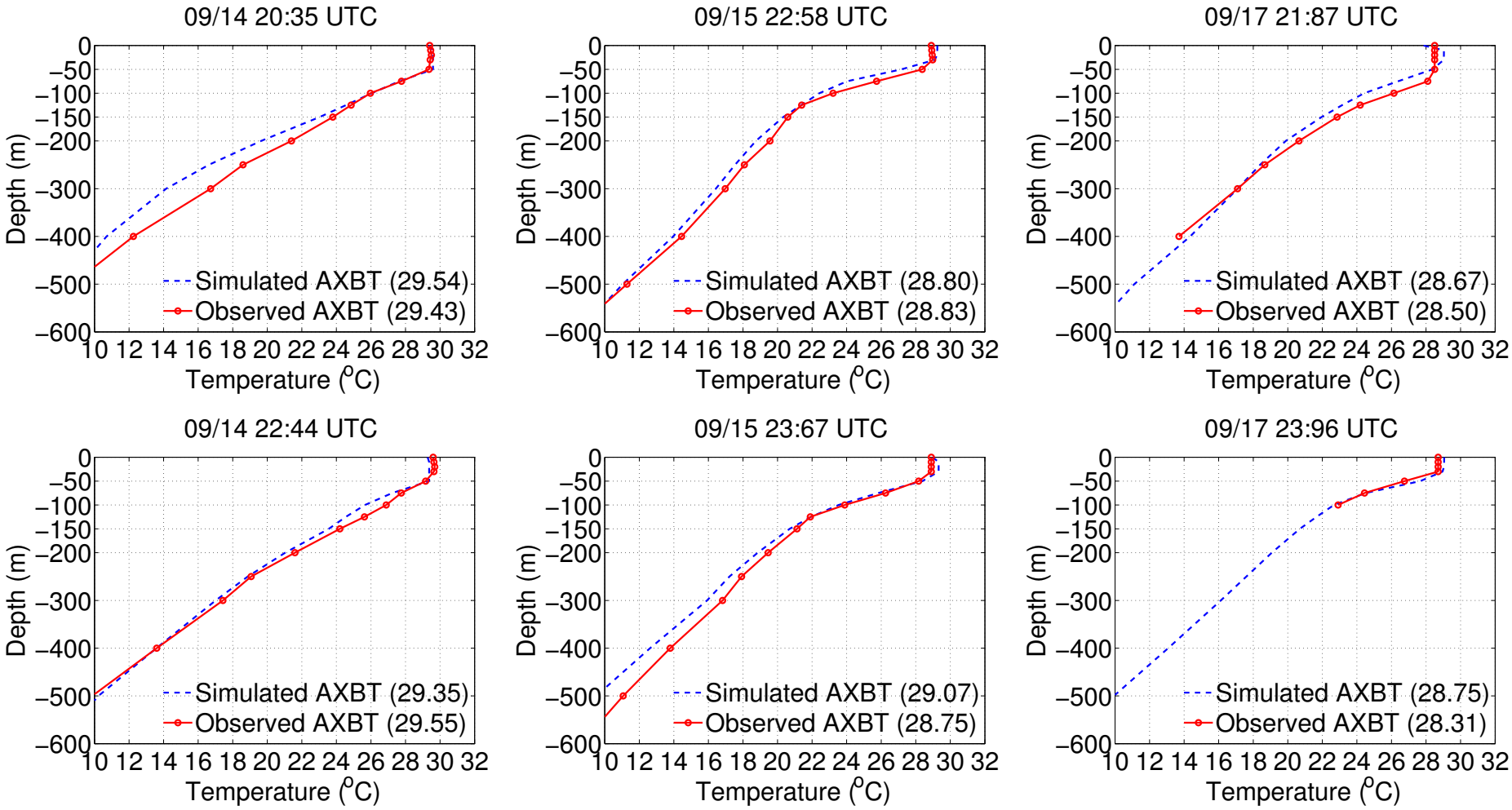
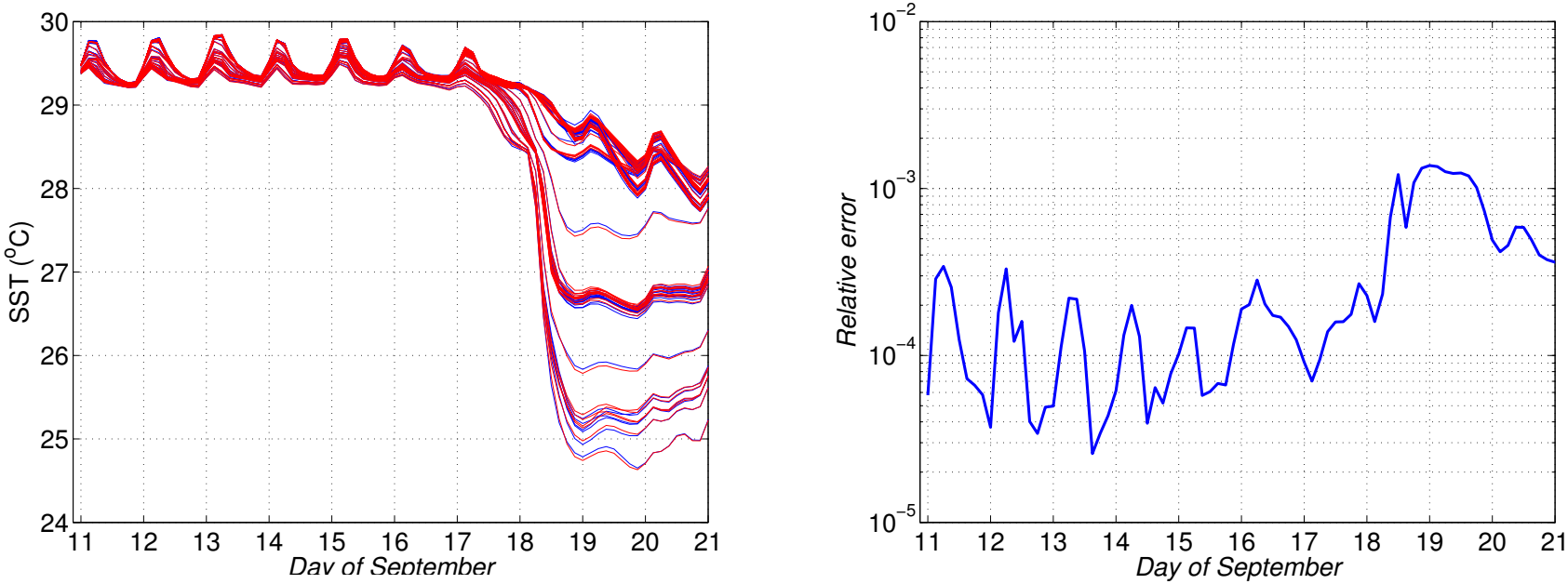


Figure: Comparison of HYCOM vertical temperature profiles with AXBT observations on Sep 14 (left), 15 (center) and 17 (right). Temperature averages over the first 50 m are shown in the legend.

PC Representation Errors



Evolution of the area-averaged SST realizations (blue) and of the corresponding PC estimates (red). The normalized rms error (right panel) remains below 0.1% for the duration of the simulation.

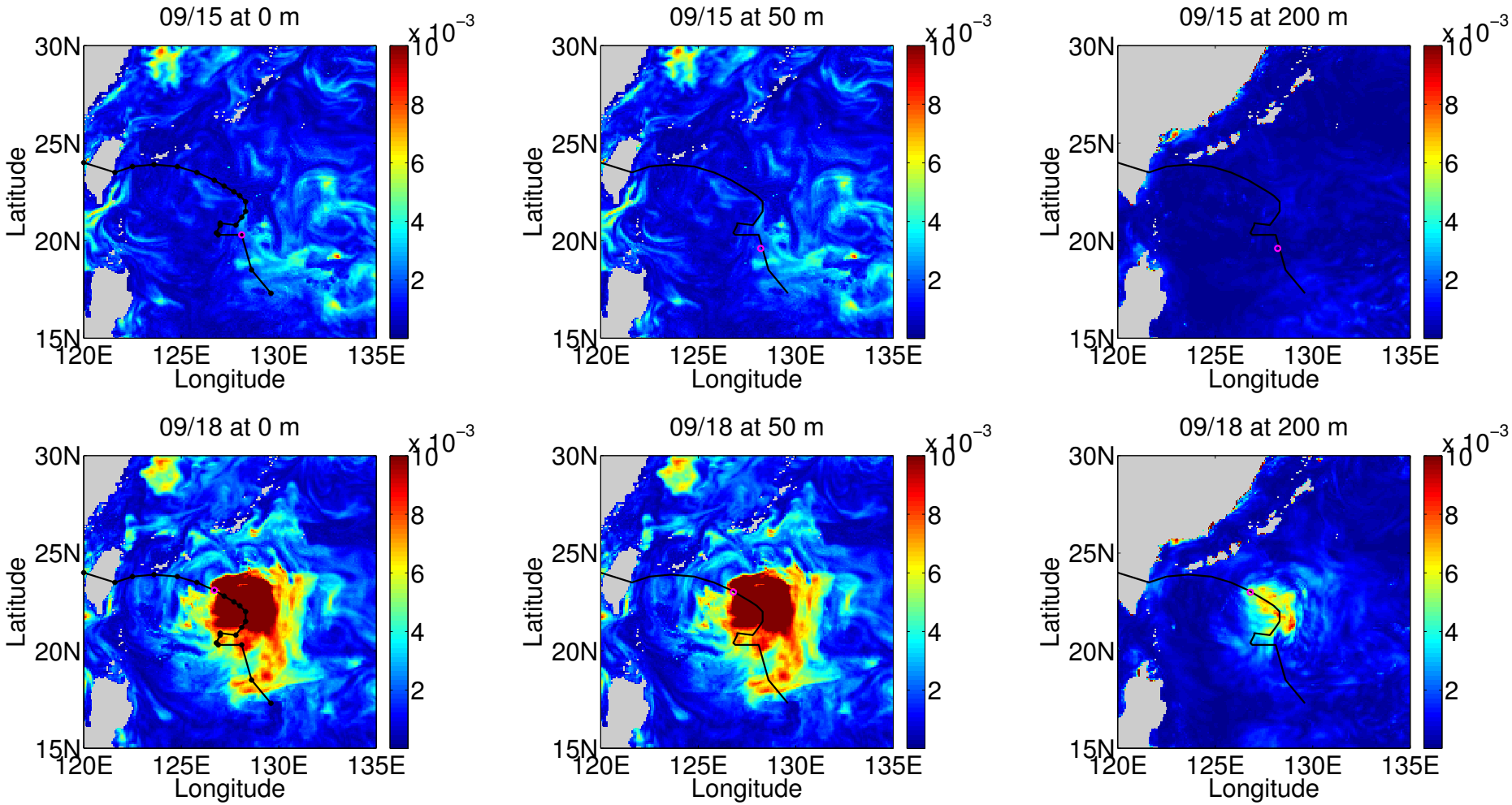
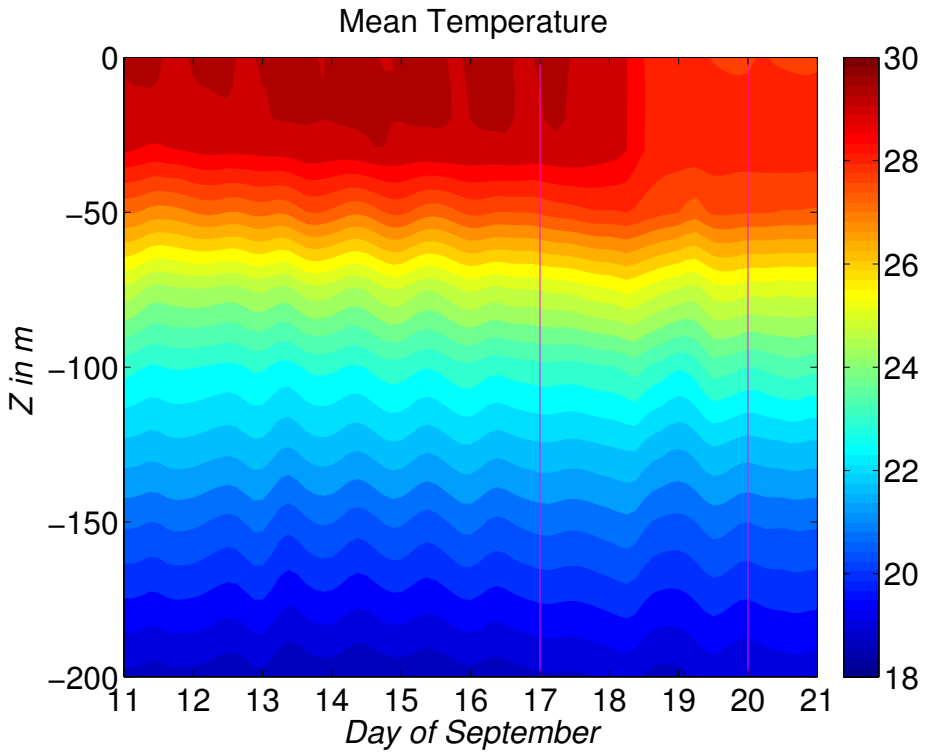
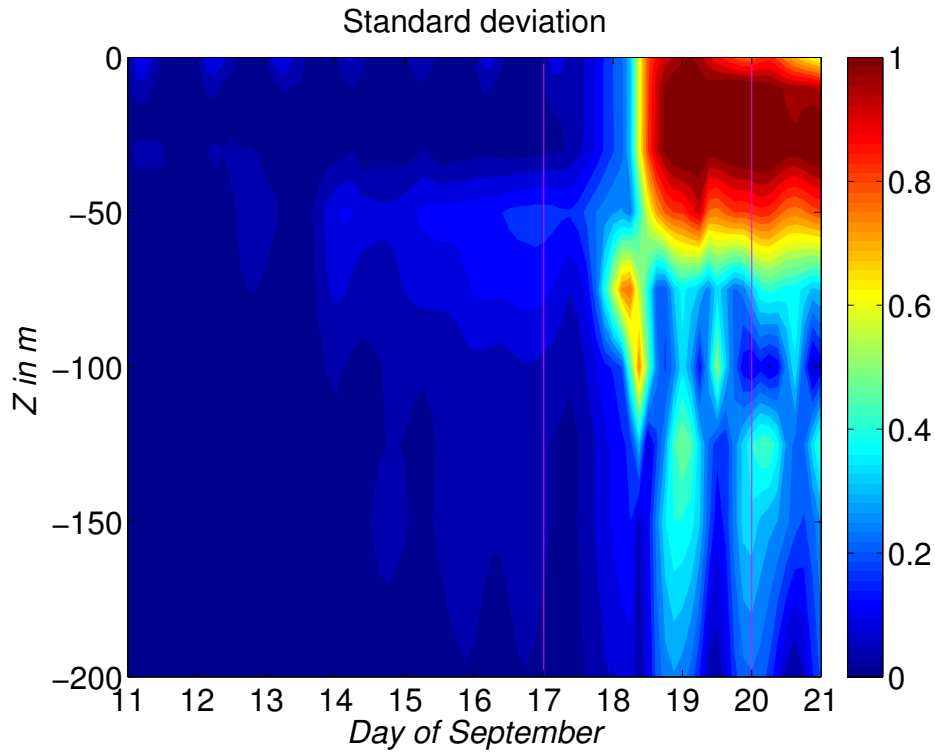


Figure: Normalized error between realizations and the corresponding PC surrogates at different depths; Top row: 00:00 UTC Sep 15; bottom row: 00:00 UTC Sep 18.

Depth Profile of Temperature Statistics





50m-deep mixed layer

2°C cooling after Fanapi arrives

Uncertainties confined to top 50 m.

SST Response Surface

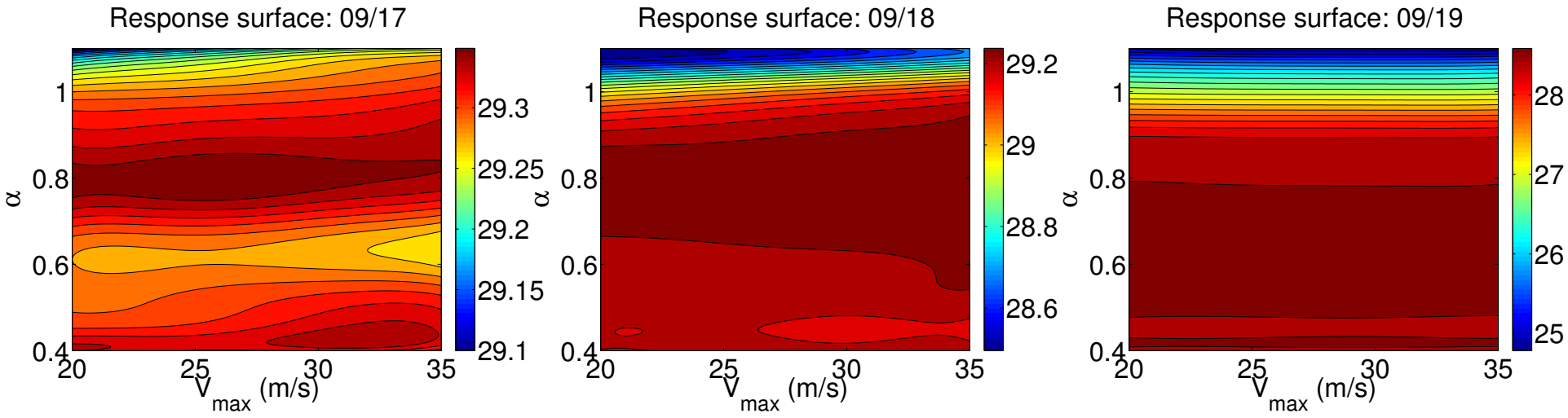


Figure: SST response surface as function of α and V_{max}, with fixed $m = 0$. Plots are generated on different days, as indicated. SST's dependence on $V_{\max}$ decreases after 09/17.

Markov Chain Monte Carlo

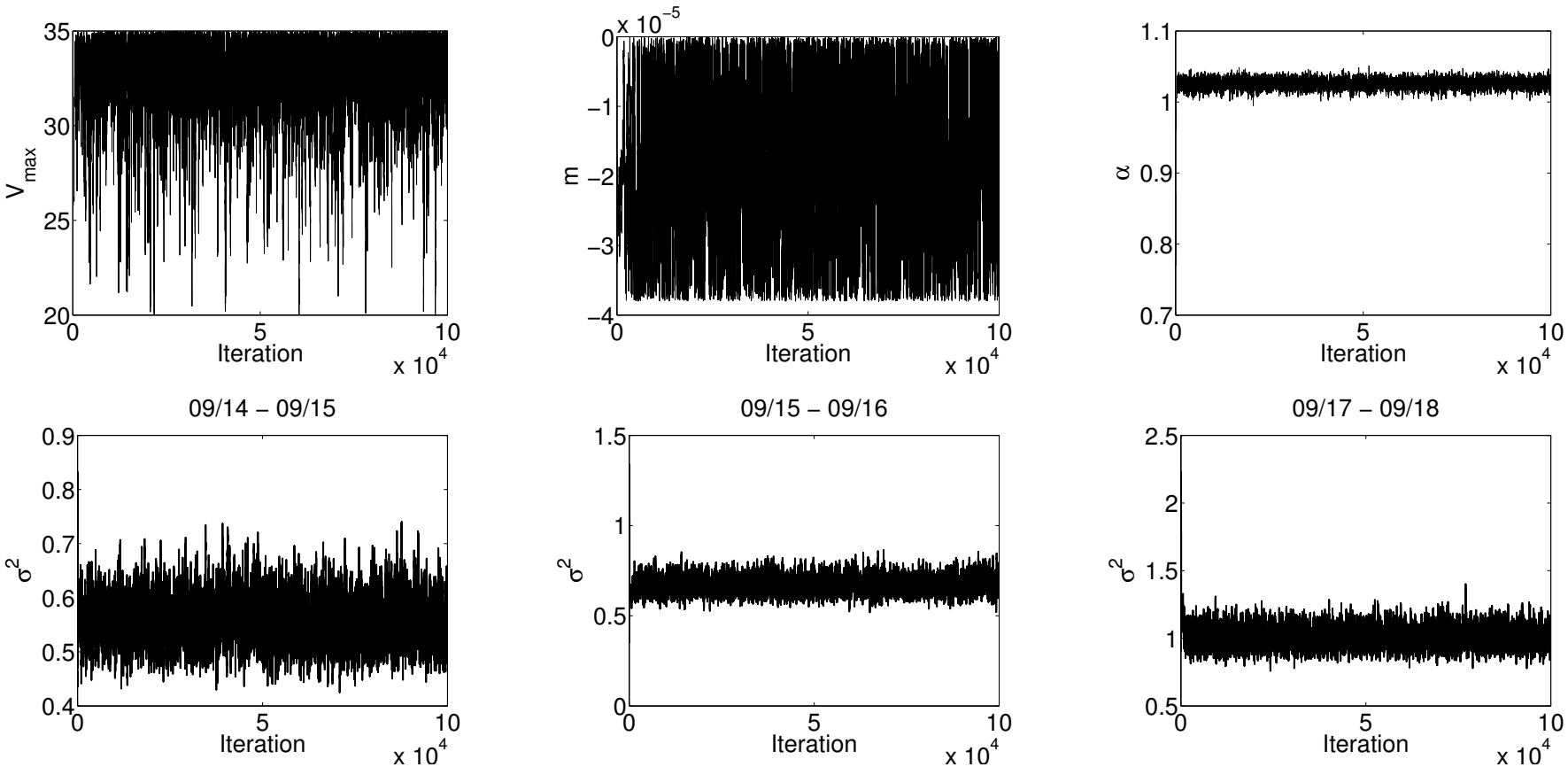


Figure: Top row: chain samples for V_{max}, m and α. Bottom row: chain samples for σ^2 generated for different days, as indicated.

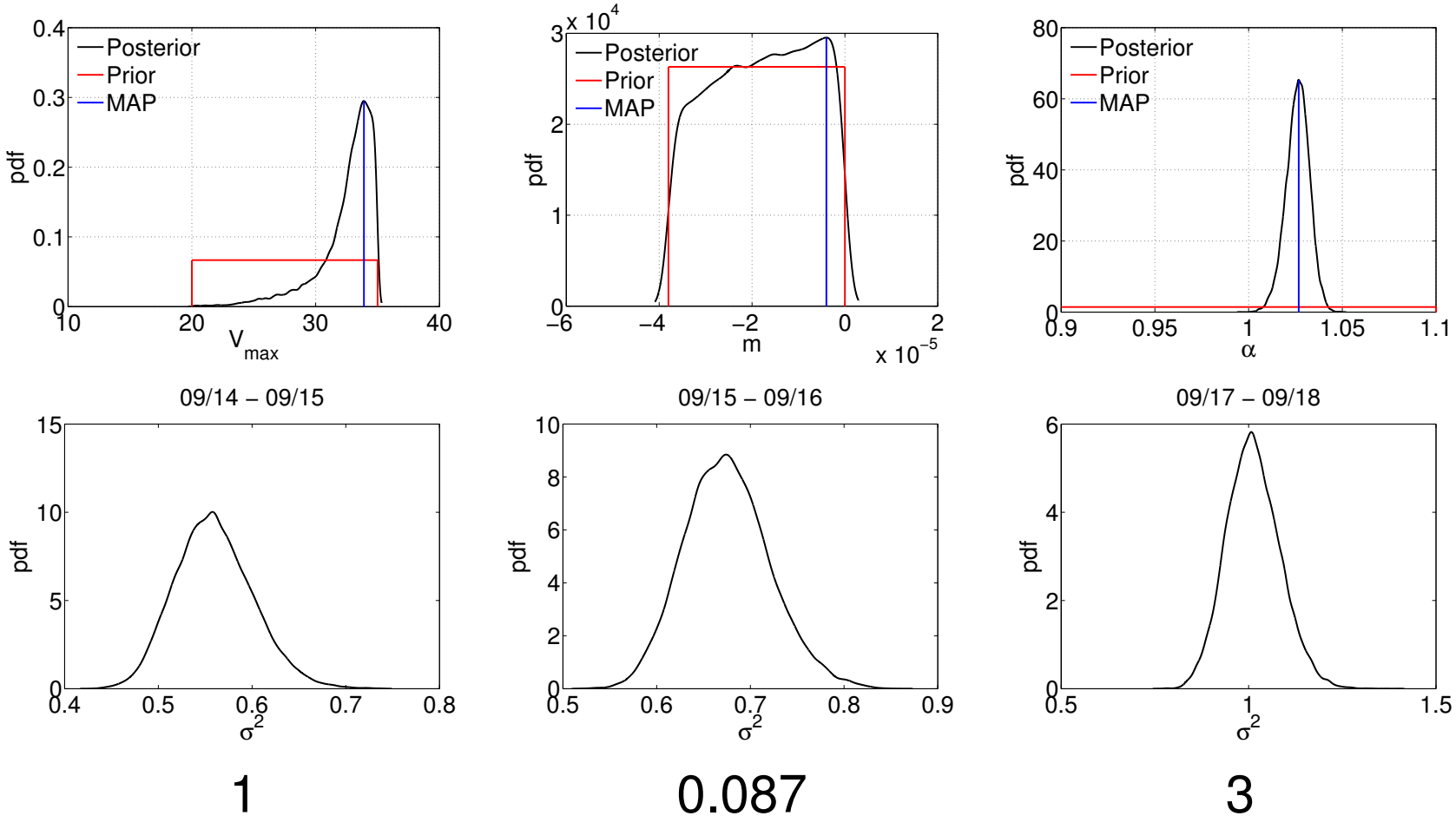


Figure: Posterior distributions for the drag parameters (top) and the variance between simulations and observations (bottom). The numbers show the Kullback-Liebler divergence quantifying the distance between 2 prior and posterior pdfs, i.e. the information gain.

Remarks on posteriors

- $V_{\max}$ exhibits a well-defined peak at 34 m/s.
- Posterior of m resembles prior. Data added little to our knowledge of m.
- α shows a definite peak at 1.03 with a Gaussian like-distribution.
- $\sqrt{\sigma^2}$ is a measure of the temperature error expected. This error grows with time from about 0.75° to 1°C.

Joint posterior PDFs

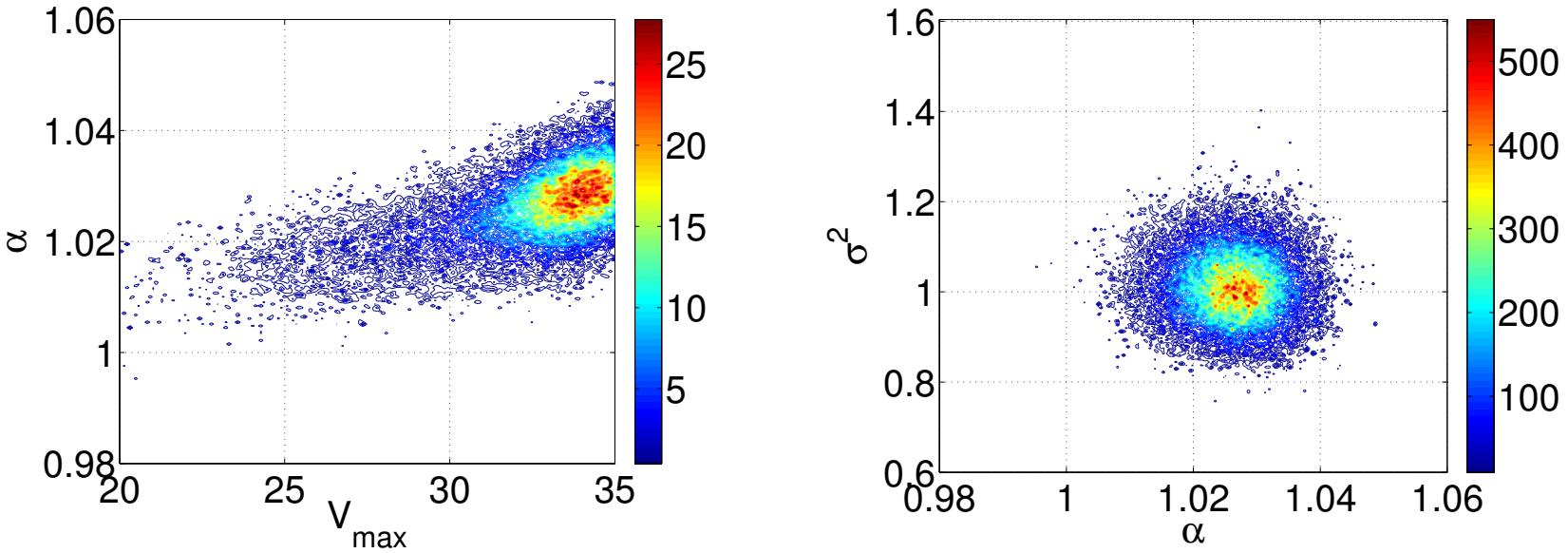


Figure: Left: joint posterior distribution of α (left) and V_{max}; right: joint posterior of α and σ^2, generated for Sep 17-Sep 18. Single peak located at $V_{\text{max}} = 34$ m/s and $\alpha = 1.03$. The posterior shows a tight estimate for α with little spread around it.

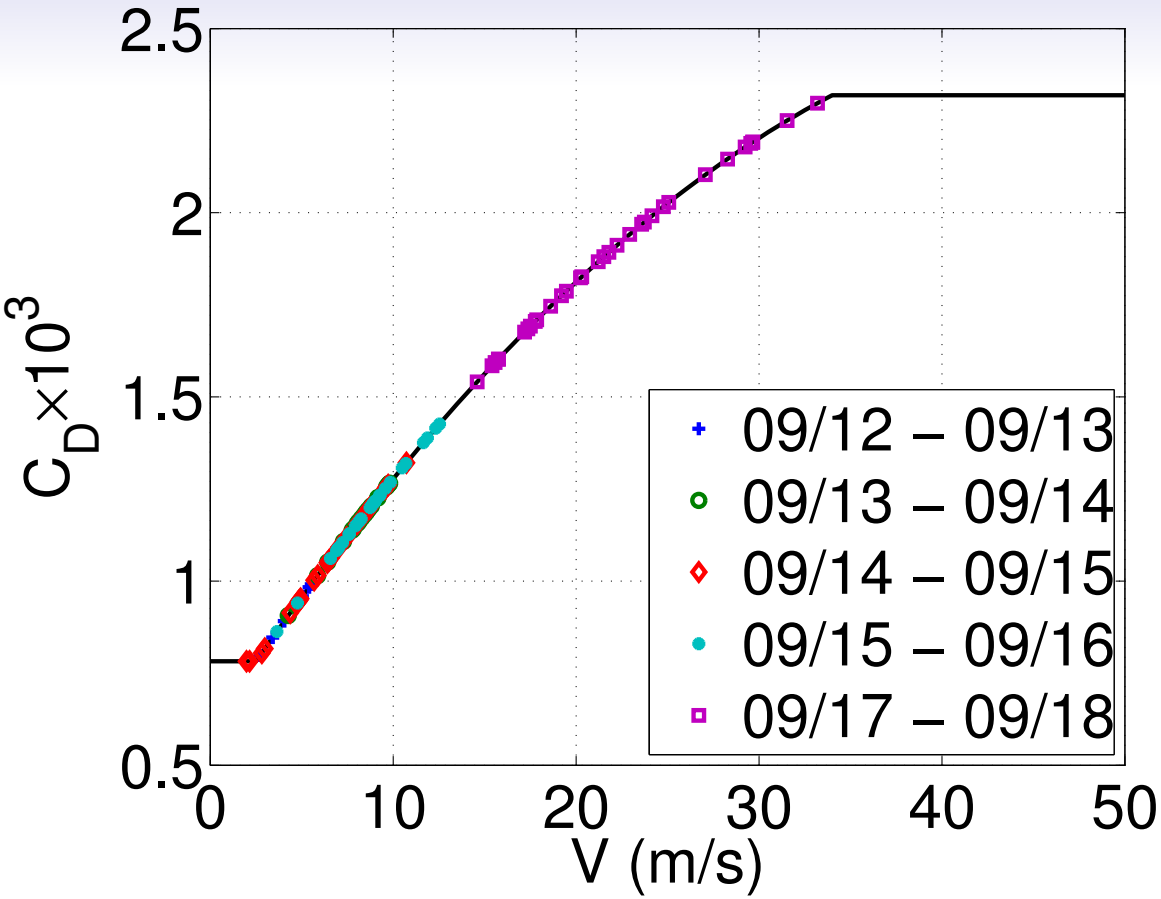


Figure: Optimal wind drag coefficient C_D using MAP estimate of the three drag parameters. The symbols refer to AXBT data used in the Bayesian inference.

Variational Form

- maximize the posterior density, or equivalently, minimize the negative of its logarithm

$$\mathcal{J}(\alpha, V_{max}, m, \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2) = \sum_{d=1}^{5} \left[J_d + \left(\frac{n_d}{2} + 1 \right) \ln(\sigma_d^2) \right] , \tag{8}$$

where J_d is the misfit cost for day d, the $\ln(\sigma_d^2)$ terms come from the normalization factors of the Gaussian likelihood functions and from the Jeffreys priors.

- The expression for J_d is:

$$J_d(\alpha, V_{max}, m, \sigma_d^2) = \frac{1}{2\sigma_d^2} \sum_{i \in \mathcal{I}_d} [M_i - T_i]^2 , \tag{9}$$

where $\mathcal{I}_d$ is the set of n_d indices of the observations from day d.

Adjoint-Free Gradients

Minimization requires cost function gradients

$$\left[\frac{\partial \mathcal{J}}{\partial \alpha}, \frac{\partial \mathcal{J}}{\partial V_{\max}}, \frac{\partial \mathcal{J}}{\partial m}\right] = \sum_{d=1}^{5} \frac{1}{\sigma_d^2} \left(\sum_{i \in \mathcal{I}_d} (M_i - T_i) \left[\frac{\partial M_i}{\partial \alpha}, \frac{\partial M_i}{\partial V_{max}}, \frac{\partial M_i}{\partial m}\right] \right)$$

Compute them from PC expansion

$$\left[\frac{\partial M}{\partial \alpha}, \frac{\partial M}{\partial V_{\max}}, \frac{\partial M}{\partial m}\right] = \sum_{k=0}^{P} \hat{M}_k(\boldsymbol{x}, t) \left[\frac{\partial \psi_k}{\partial \alpha}, \frac{\partial \psi_k}{\partial V_{\max}}, \frac{\partial \psi_k}{\partial m}\right].$$

- $\frac{\partial \psi_k}{\partial \alpha}$ easy to compute
- No adjoint model needed
- For Hessian just differentiate above again.

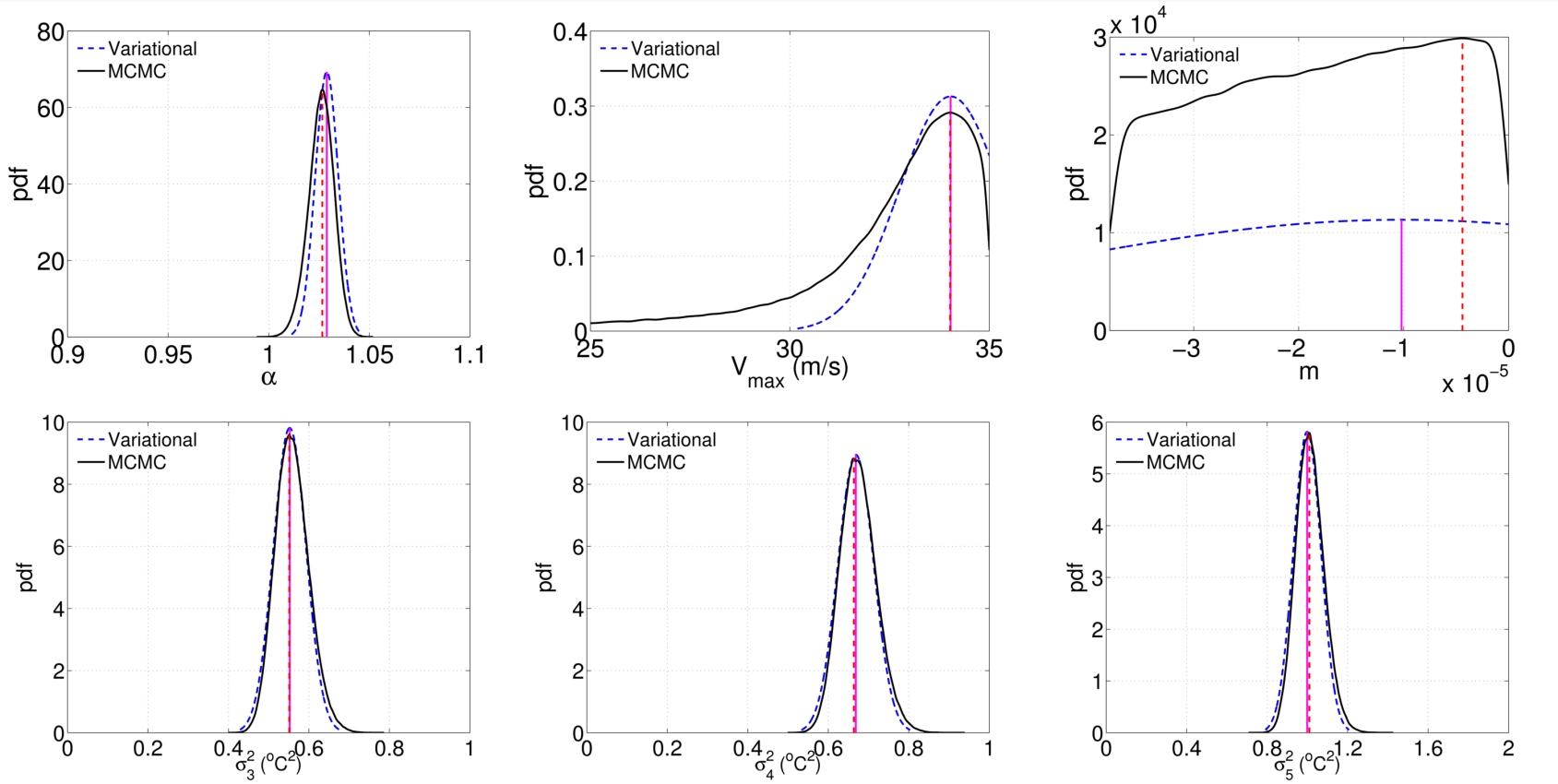


Figure: Posterior probability distributions for (top) drag parameters and (bottom) variances σ_d^2 at selected days using variational method and MCMC. The vertical lines correspond to the MAP values from MCMC and optimal parameters found using the variational method.

Uncertainty in Initial Boundary Conditions

Rely on EOFs to characterize uncertainty and reduce the number of stochastic variables. For 2 EOFs mode we have:

$$u(\vec{x}, 0, \xi_1, \xi_2) = \overline{u}(\vec{x}, 0) + \alpha \left[\sqrt{\lambda_1}\mathcal{U}_1\xi_1 + \sqrt{\lambda_2}\mathcal{U}_2\xi_2\right] \quad (10)$$

- $(\lambda_k, \mathcal{U}_k)$: are eigenvalues/eigenvectors of covariance matrix obtained from free-run simulation
- $\overline{u}$: unperturbed initial condition
- $u(\vec{x}, 0, \xi)$: Stochastic initial condition input
- α: multiplicative factor to control size of “kick”

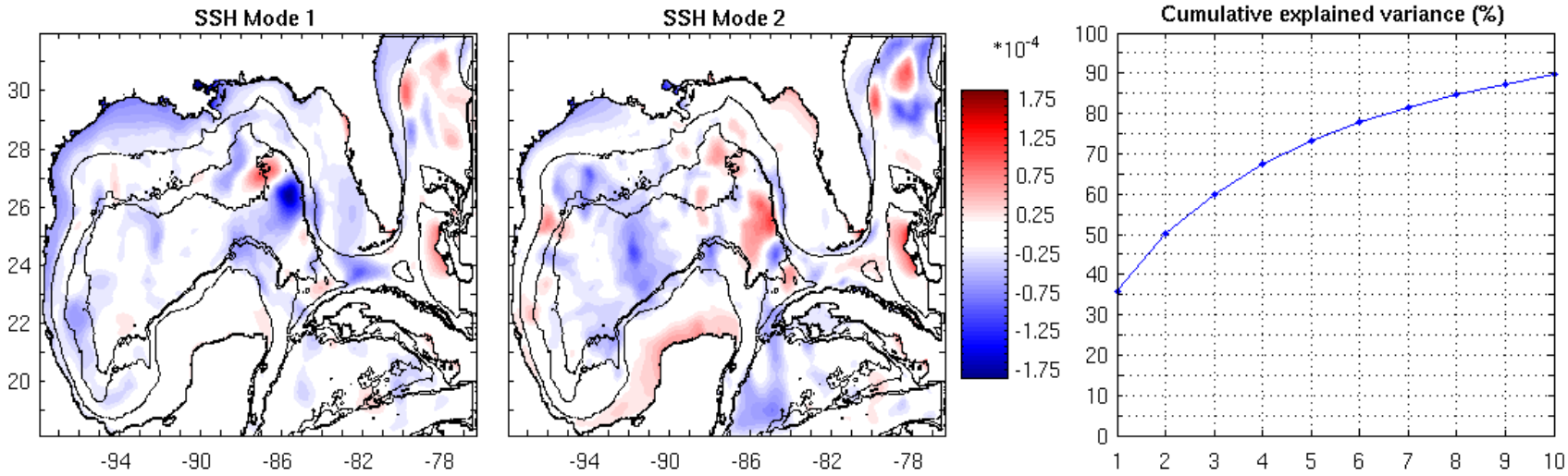


Figure: First and Second SSH modes from a 14-day series. The 2 modes account for 50% of variance during these 14 days.

- Characterize local uncertainty: get perturbation from short, 14-day, simulation.
- Uncertainty dominated by Loop Current (LC) dynamics
- Mode 1 seems associated with a frontal eddy

PC representation

- (ξ_1, ξ_2) independent and uniformly distributed random variables
- PC basis: Legendre polynomials of max degree 6, $P = 28$
- Ensemble of 49 realizations for Hermite quadrature

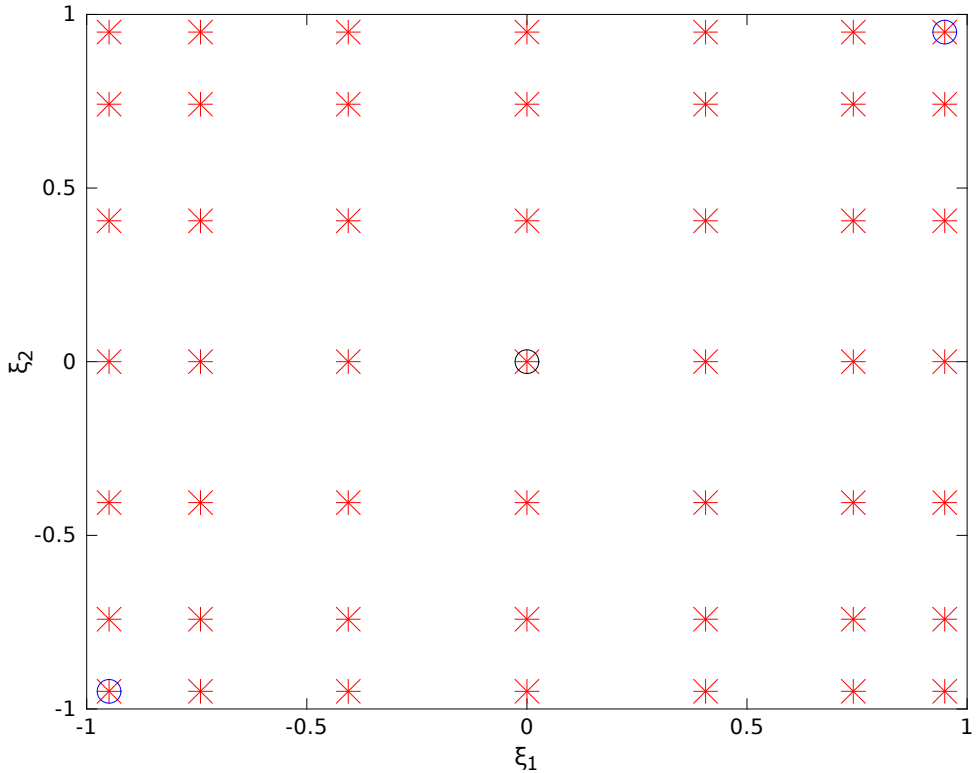
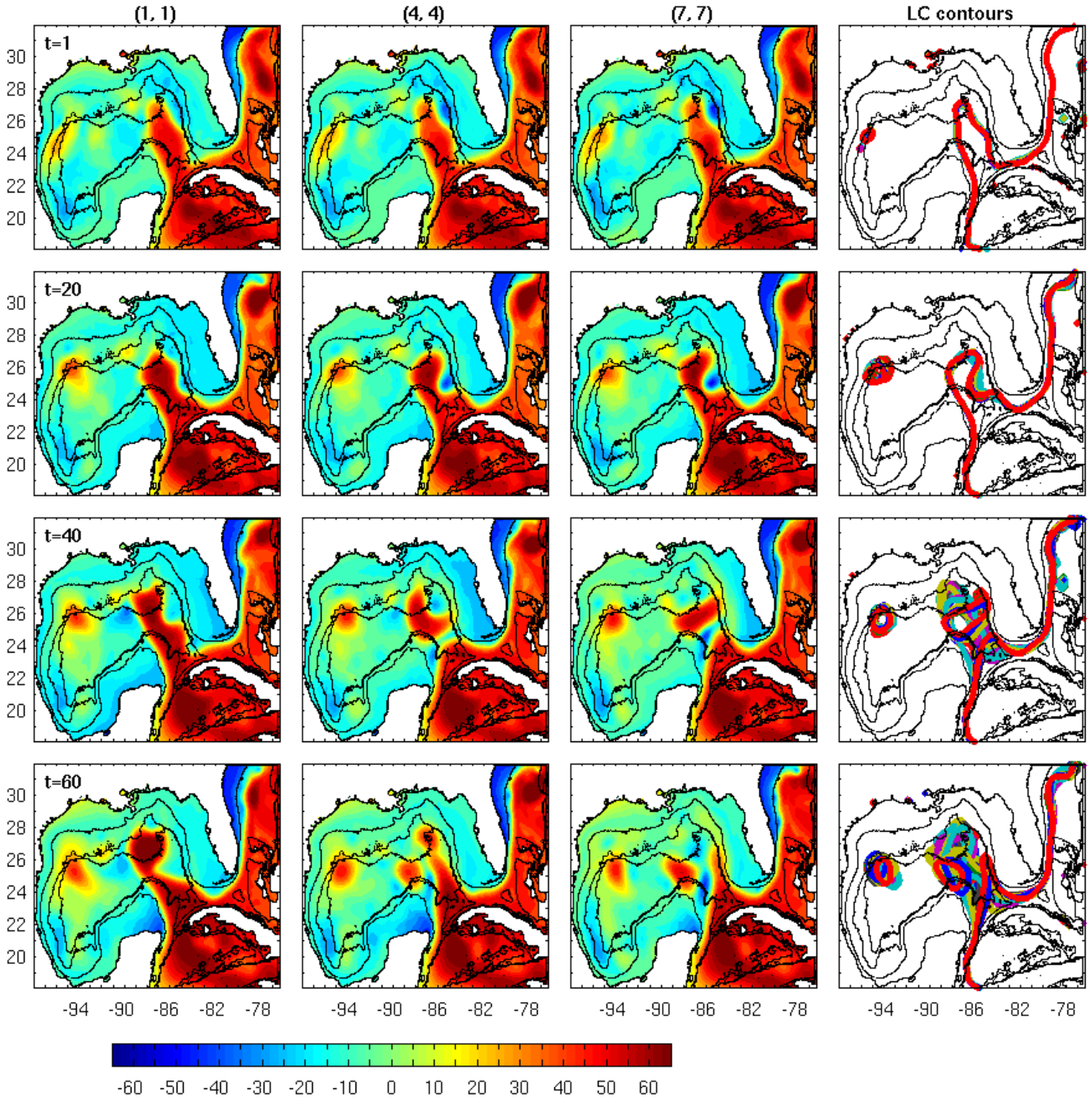


Figure: Quadrature/Sample points in ξ_1, ξ_2 space. Center black circle corresponds to unperturbed run, while blue circles correspond to largest negative and positive perturbations.

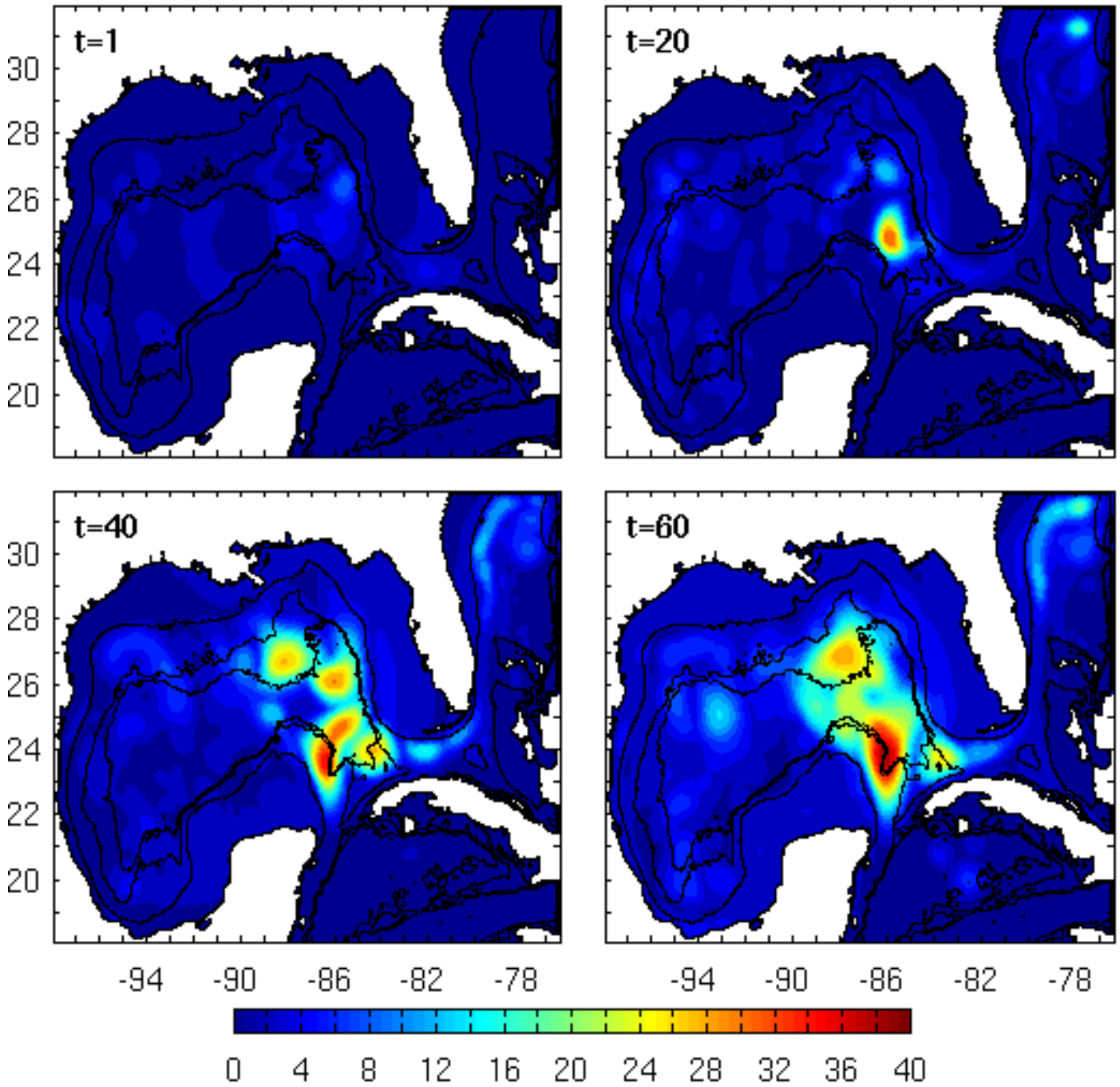


Col 1: SSH of realization (1,1) with weakest frontal eddy

Col 2: SSH of unperturbed realization (4,4) has medium strength frontal eddy

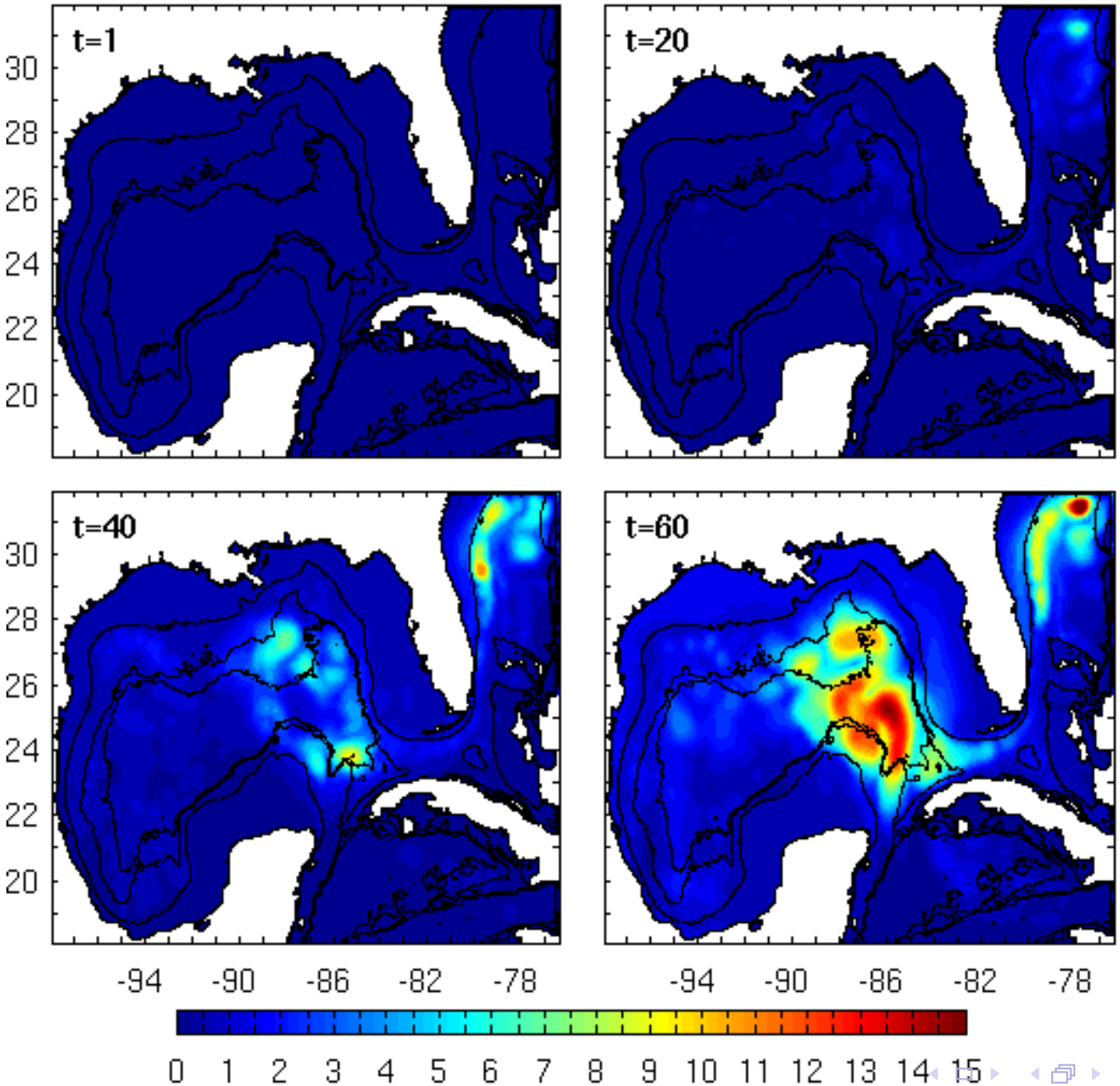
Col 3: SSH of realization (7,7) has strogest frontal eddy and earliest LC separation

Col 4: Loop current edge in ensemble



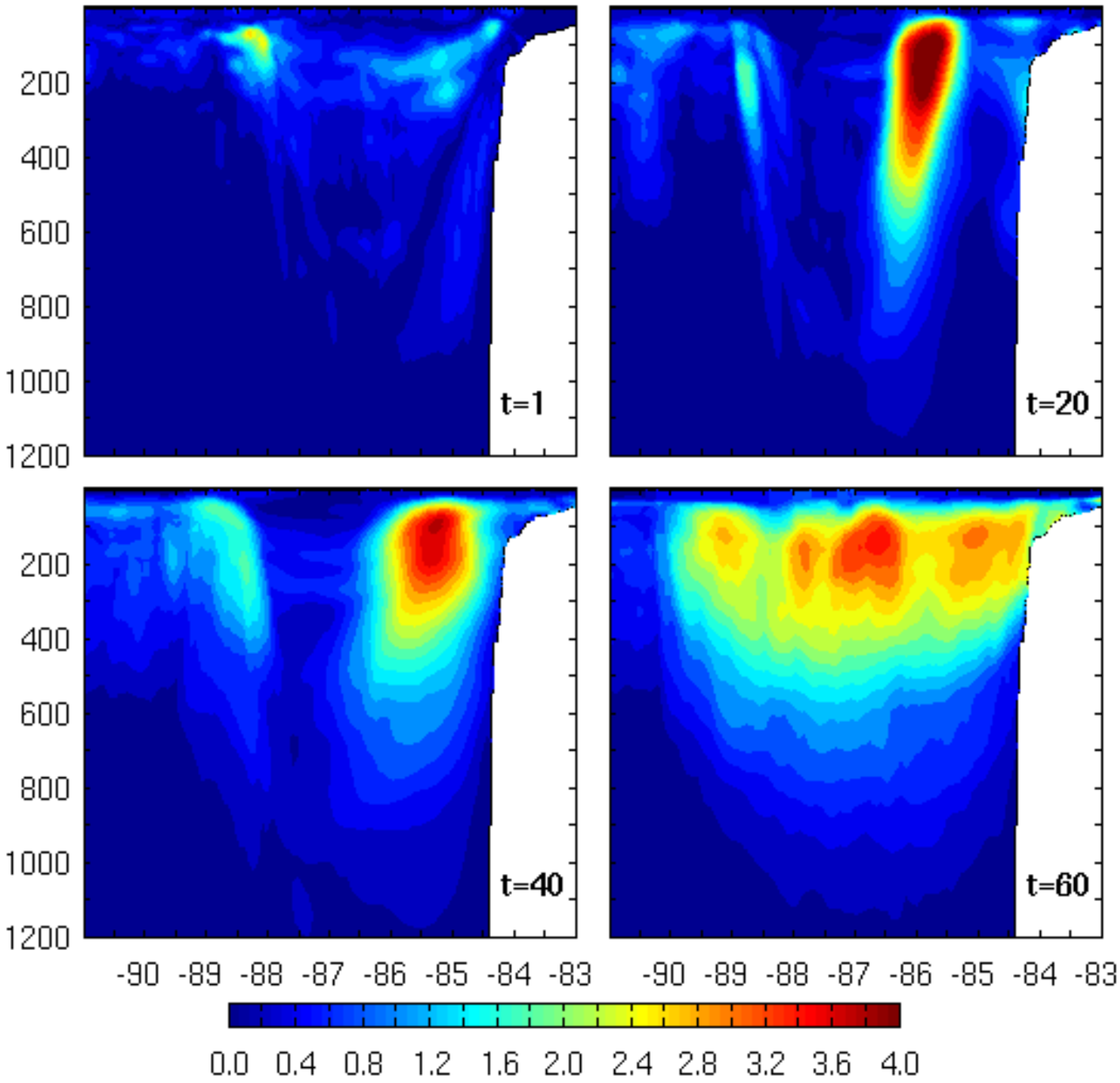
SSH stddev (cm) grows in time with maximum in LC region

PC-error: $\|\epsilon\|_2^2 = \sum_q \left[\eta(\vec{x}, t, \xi_q) - \eta_{PC}(\vec{x}, t, \xi_q)\right]^2 \omega_q$



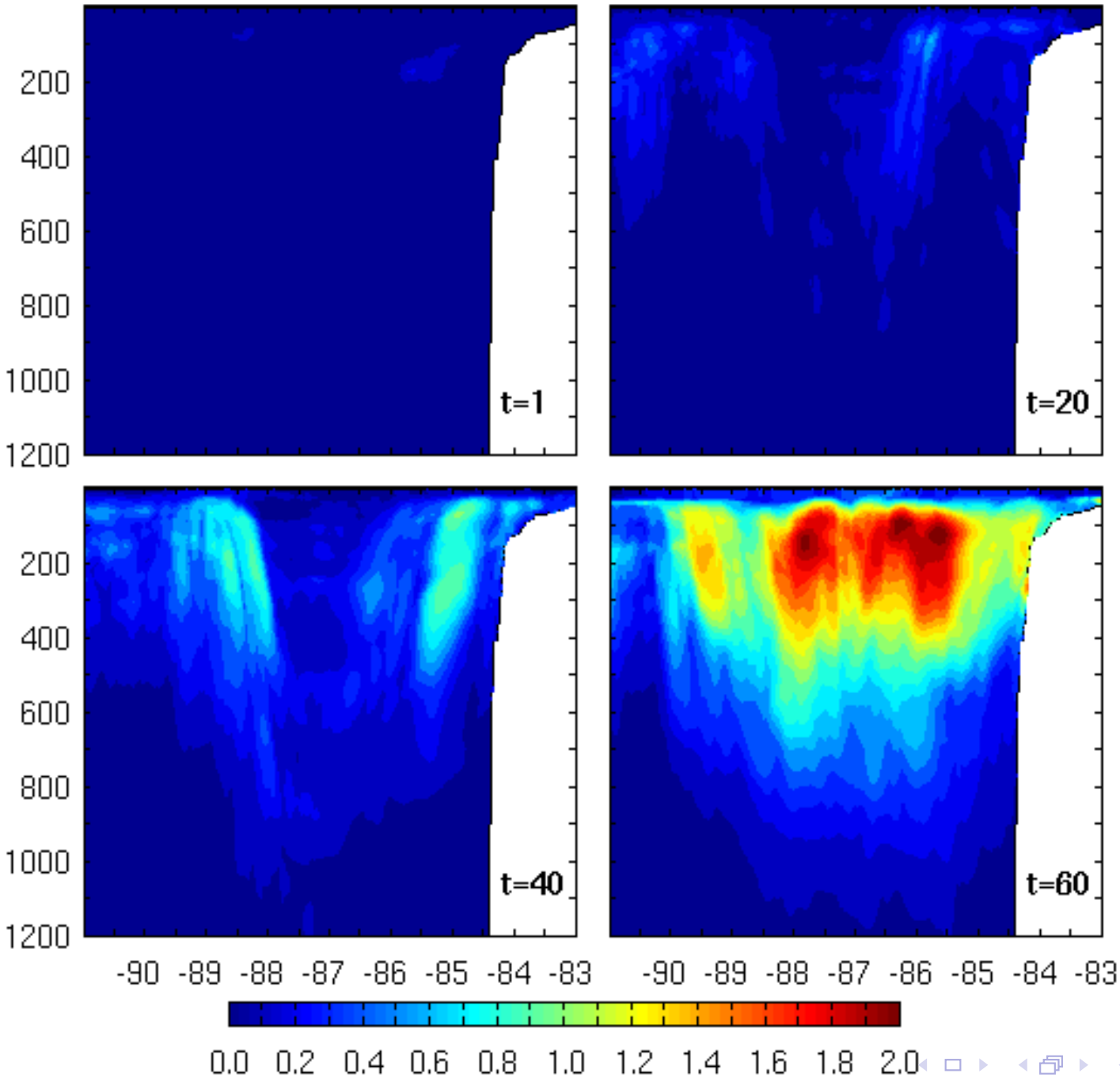
SSH PC-errors (cm) grow in time with maxima in LC region

On day 60 PC-error is about 38% of stddev



T-section along 25N, stddev grows in time with maxima coinciding with Frontal Eddy during days 20–40.

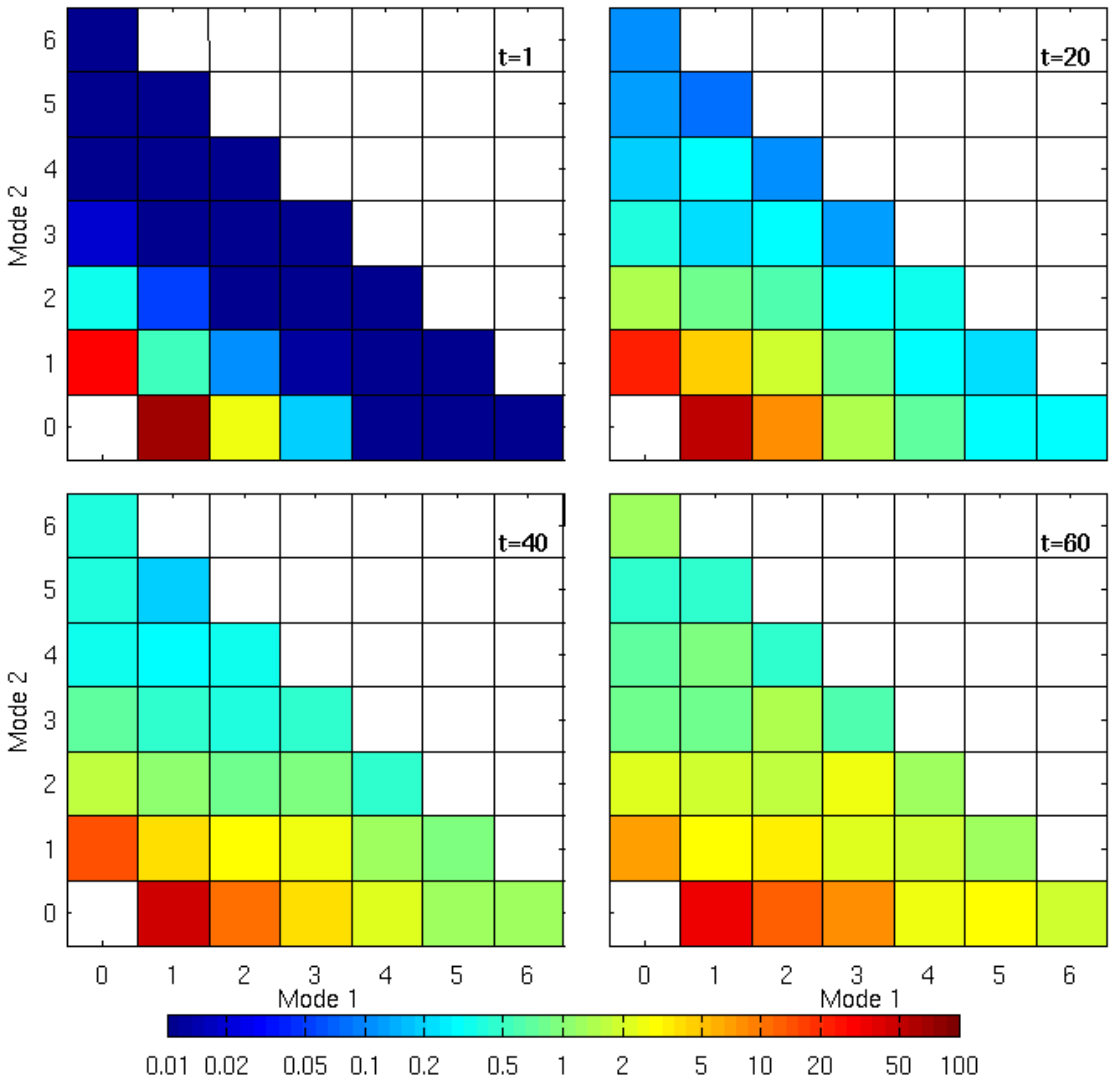
PC-error: $\|\epsilon\|_2^2 = \sum_q \left[T(\vec{x}, t, \xi_q) - T_{PC}(\vec{x}, t, \xi_q)\right]^2 \omega_q$



T PC-errors (cm) grow in time with maxima in LC region

On day 60 PC-error is about 50% of stddev

Distribution of SSH PC coefficients



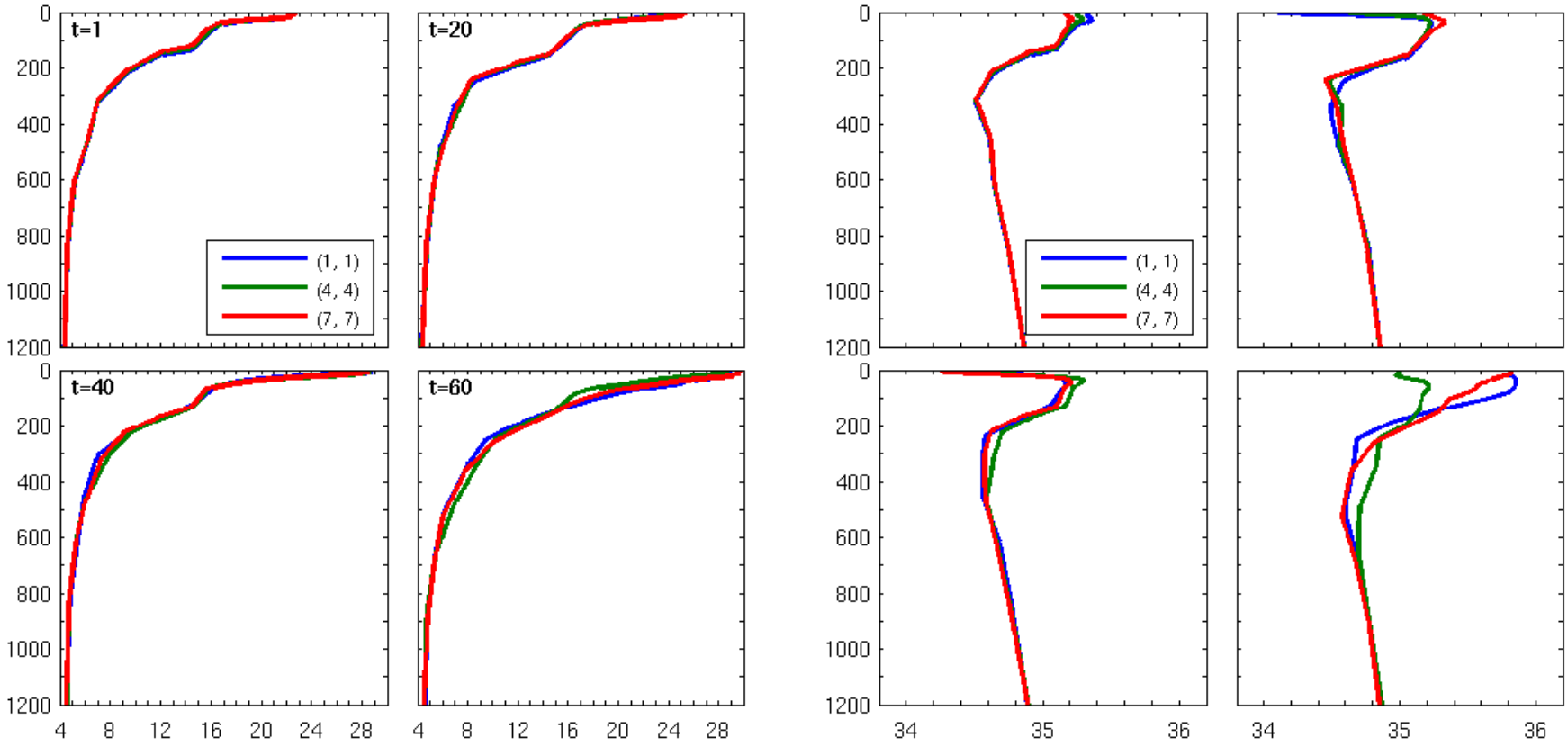
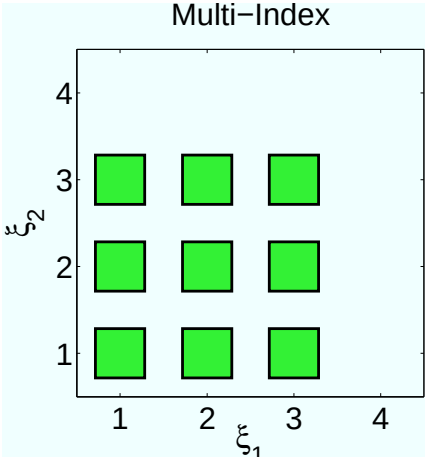
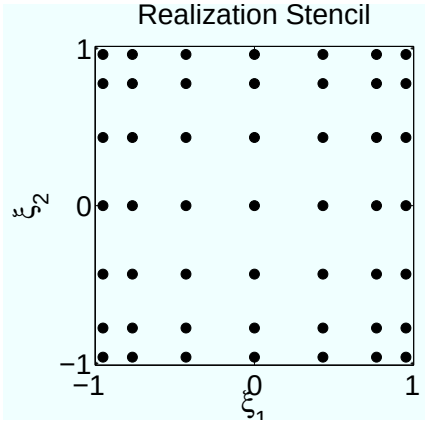
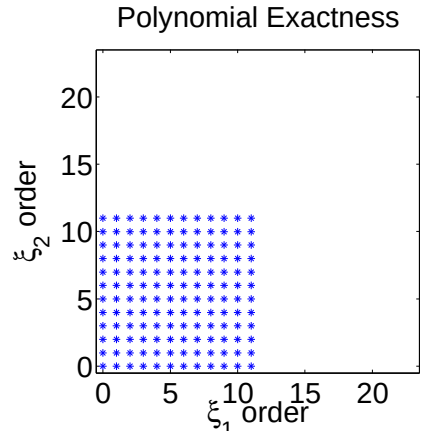


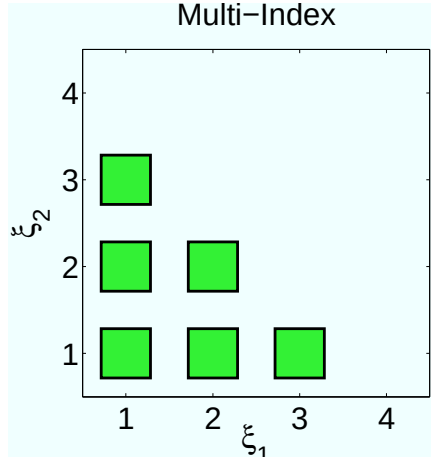
Figure: Temperature (left) and Salt (right) profiles for extreme realizations at DWH

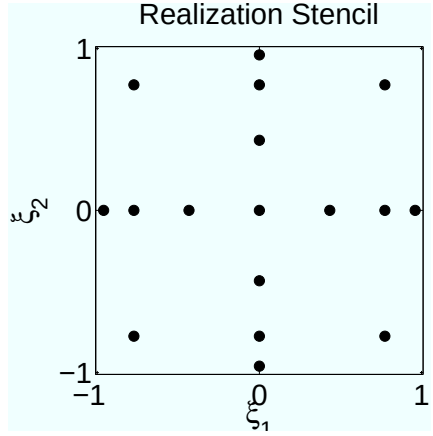


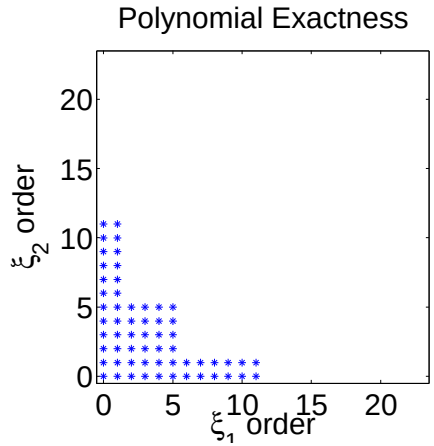




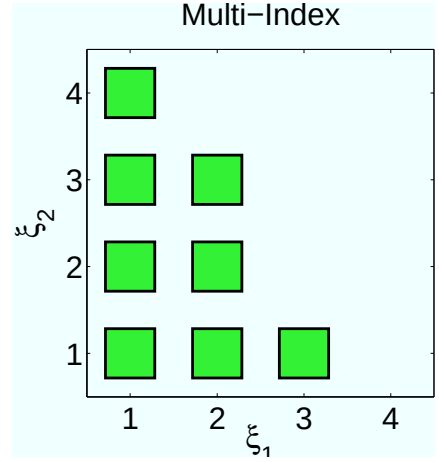
Full-Tensor (49)

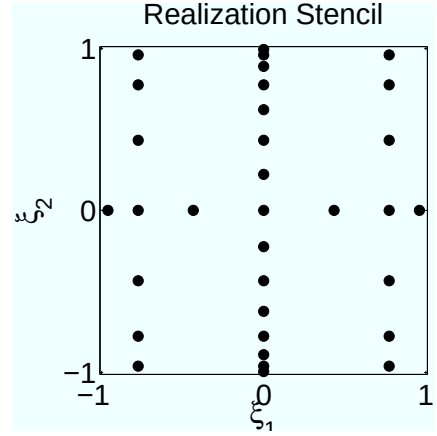


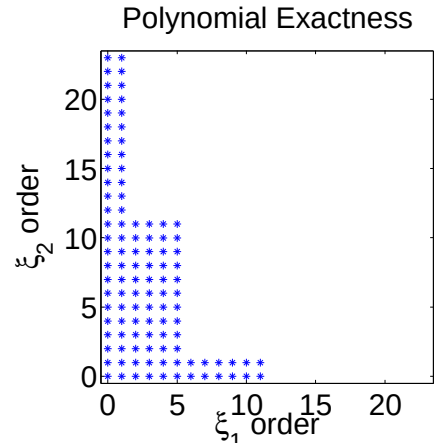




Classic Smolyak (17)







Arbitrary Multi-Index (33)

Examples of 2D tensorizations

Varying polynomial order

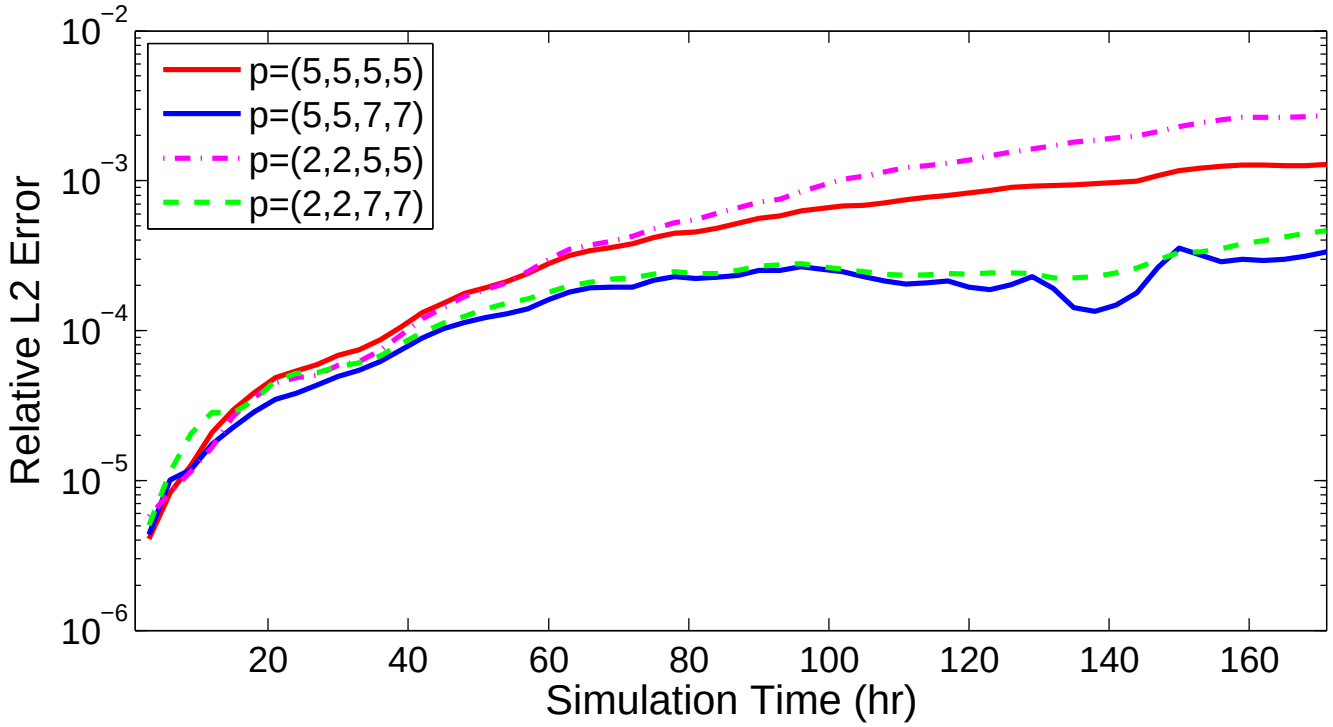


Figure: Relative L2 error between the area-averaged SST and the Latin Hypercube Samples.

Simple Truncation	P	# of realizations
$\mathbf{p} = (\mathbf{5}, \mathbf{5}, \mathbf{5}, \mathbf{5})$	126	385
$\mathbf{p} = (\mathbf{5}, \mathbf{5}, \mathbf{7}, \mathbf{7})$	168	513
$\mathbf{p} = (\mathbf{2}, \mathbf{2}, \mathbf{5}, \mathbf{5})$	36	73
$\mathbf{p} = (\mathbf{2}, \mathbf{2}, \mathbf{7}, \mathbf{7})$	59	169

Smolyak *Projections*

- Apply Smolyak's algorithm directly to construct the PCE instead of purely generating the quadrature. Thus, the final projection becomes a weighted sum of aliasing-free sub-projections. This is an extension of the Smolyak tensor construction from quadrature operators to projection operators.
- Smolyak projection allows a refinement approach based on successive inclusion of any admissible multi-index, $\mathcal{F}$, of quadrature rules while maintaining the representation free of internal aliasing.
- A larger number of polynomials can be integrated than is possible with a classical dimensional truncation / quadrature using the same ensemble, The 513 HYCOM realizations yields 402 coefficient with Smolyak *projection* compared to 168 using Smolyak *quadrature*.

Adaptive *Projections*

- Rewrite projection as tensor products of projection *differences*:

$$\left(\Delta_{k_1} \otimes \ldots \otimes \Delta_{k_d}\right) U,$$

- The L_2 norm of this difference can be readily used to define an error indicator for multi-index $\mathbf{k}$,

$$\epsilon(\boldsymbol{k}) = || \left(\Delta_{k_1} \otimes \ldots \otimes \Delta_{k_d}\right) U||$$

  The indicator represents the variance surplus due to the $\mathbf{k}$ sub-projection.

- The surplus is computed for each $\mathbf{k} \in \mathcal{F}$ and the sub-projection with the highest $\epsilon(\mathbf{k})$ is selected for subsequent refinement, which generally consists of inclusion of admissible forward neighbors.

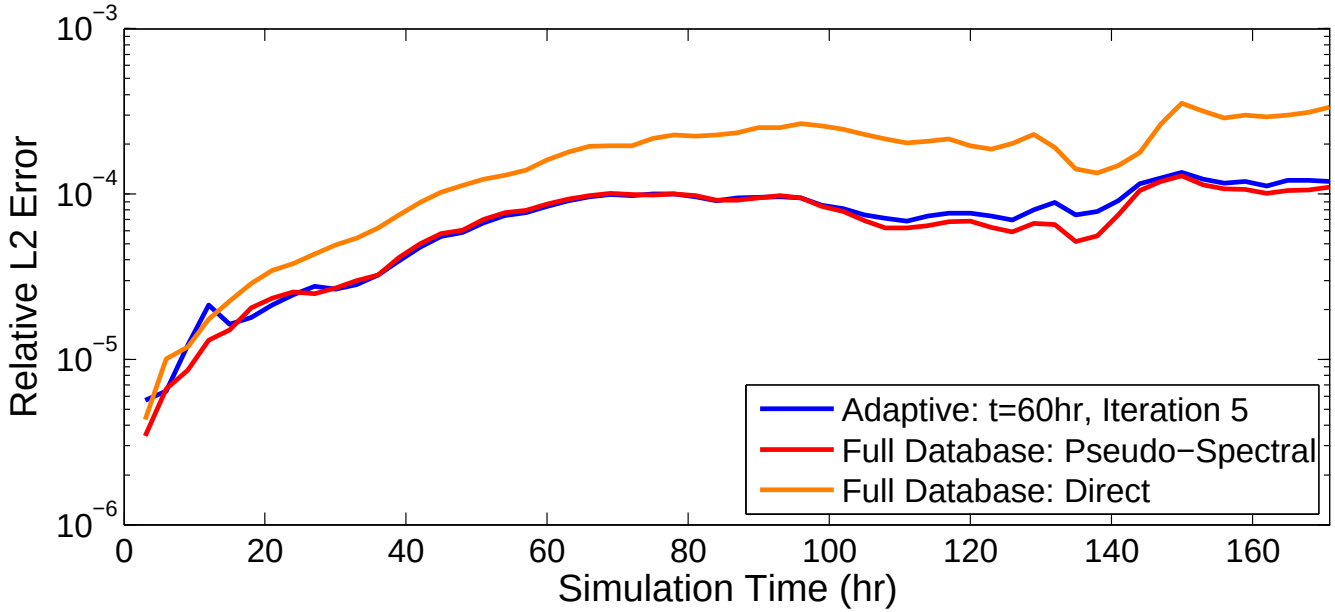


Figure: Relative L2 difference between the PCE of the averaged SST and the LHS sample. Plotted are curves generated with (i) the adaptive Smolyak projection adapted at $t = 60$ hr, (ii) the Smolyak projection with the full database, and (iii) Smolyak classical quadrature using the full database. For the adapted solution, the refinement is stopped after iteration 5, leading to 69 realizations and a PCE with 59 polynomials. The full 513 database curves have 402 polynomials for the pseudo-spectral construction and 168 polynomials for the Smolyak quadrature.

Publications

- A. Srinivasan, J. Helgers, C. B. Paris, M. LeHenaff, H. Kang, V. Kourafalou, M. Iskandarani, W. C. Thacker, J. P. Zysman, N. F. Tsinoremas, and O. M. Knio. Many task computing for modeling the fate of oil discharged from the deep water horizon well blowout. In *Many-Task Computing on Grids and Supercomputers (MTAGS), 2010 IEEE Workshop on*, pages 1–7, November, 2010. IEEE.
- W. C. Thacker, A. Srinivasan, M. Iskandarani, O. M. Knio, and M. Le Henaff. Propagating oceanographic uncertainties using the method of polynomial chaos expansion. *Ocean Modelling*, **43–44**, pp 52–63, 2012.
- A. Alexanderian, J. Winokur, I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker, and O. M. Knio, Global sensitivity analysis in an ocean general circulation model: a sparse spectral projection approach, *Computational Geosciences*, **16**, 757–778, 2012.
- I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker, A. Alexanderian, C. Lee and S. S. Chen and O. M. Knio, Bayesian Inference of Drag Parameters Using AXBT Data from Typhoon Fanapi, *Monthly Weather Review*, **141**, no 7, pp 2347–2366, 2013.
- J. Winokur, P. Conrad, I. Sraj, O. M. Knio, A. Srinivasan, W. Carlisle Thacker, Y. Marzouk, and M. Iskandarani, A Priori Testing of Sparse Adaptive Polynomial Chaos Expansions Using an Ocean General Circulation Model Database, *Computational Geosciences*, 2013.
- I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker and O. M. Knio, Computing Model Gradients from a Polynomial Chaos based Surrogate for an Inverse Modeling Problem, *Monthly Weather Review*, in review.