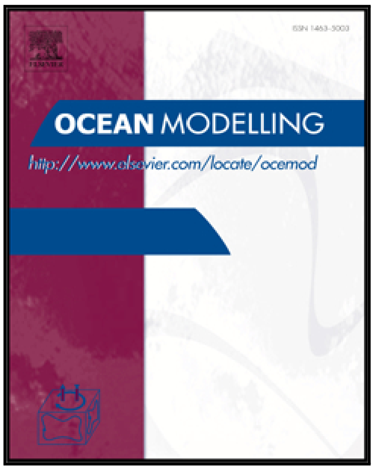

Verifying and assessing the performance of the perturbation strategy in polynomial chaos ensemble forecasts of the circulation in the Gulf of Mexico

Shitao Wang, Guotu Li, Mohamed Iskandarani, Matthieu Le Hénaff, Omar M. Knio

PII: S1463-5003(18)30311-1
DOI: https://doi.org/10.1016/j.ocemod.2018.09.002
Reference: OCEMOD 1339

To appear in: *Ocean Modelling*

Received date: 9 August 2017
Revised date: 10 July 2018
Accepted date: 5 September 2018

Please cite this article as: Shitao Wang, Guotu Li, Mohamed Iskandarani, Matthieu Le Hénaff, Omar M. Knio, Verifying and assessing the performance of the perturbation strategy in polynomial chaos ensemble forecasts of the circulation in the Gulf of Mexico, *Ocean Modelling* (2018), doi: https://doi.org/10.1016/j.ocemod.2018.09.002

Highlights

- Polynomial Chaos methods are used to quantify uncertainties in ocean forecasting.
- The EOF-based perturbations lead to realistic uncertainty representation.
- Two EOFs for initial condition perturbations captures Loop Current uncertainty.
- The analysis of the SSH PDFs shows a strong non-Gaussian signal.

Verifying and assessing the performance of the perturbation strategy in polynomial chaos ensemble forecasts of the circulation in the Gulf of Mexico

Shitao Wang[a,*], Guotu Li[d], Mohamed Iskandarani[a], Matthieu Le Hénaff[b,c], Omar M. Knio[d,e]

[a]*Rosenstiel School for Marine and Atmospheric Sciences, University of Miami, Miami, FL*
[b]*NOAA Atlantic Oceanographic and Meteorological Laboratory, Miami, FL*
[c]*Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL*
[d]*Duke University, Durham, NC*
[e]*King Abdullah University of Science and Technology, Saudi Arabia*

Abstract

We present an analysis of two recent efforts aimed at quantifying the uncertainties in a 30-day HYbrid Coordinate Ocean Model forecast of the circulation in the Gulf of Mexico, with particular emphasis on the separation of Loop Current Eddy Franklin, using Polynomial Chaos methods. The analysis herein explores whether the model perturbations lead to realistic representation of the uncertainty in the Gulf Circulation. Comparisons of model output with Sea Surface Height and current mooring data show that the observational data generally falls within the envelope of the ensemble and that the modal decomposition delivers "realistic" perturbations in the Loop Current

*Corresponding author
Email addresses: `swang@rsmas.miami.edu` (Shitao Wang), `guotu.li@duke.edu` (Guotu Li), `miskandarani@rsmas.miami.edu` (Mohamed Iskandarani), `mlehenaff@rsmas.miami.edu` (Matthieu Le Hénaff), `omar.knio@duke.edu` (Omar M. Knio)

region. We use information theory metrics to quantify the information gain and the computational trade-offs between different wind and initial conditions perturbation modes. The relative entropy measures indicate that two modes for initial condition perturbations are enough, in our model configuration, to represent the uncertainty in the Loop Current region; while two modes for wind forcing perturbations are necessary in order to estimate the uncertainty in the coastal zone. The ensemble statistics are then explored using the Polynomial Chaos surrogate and the newly developed contour boxplot methods.

Keywords: Uncertainty quantification, polynomial chaos, relative entropy, ocean modeling, data depth

---

1. Introduction

The 2010 Deepwater Horizon oil spill underscored the need for reliable oceanic and atmospheric forecasts in order to predict the trajectory and evolution of the oil spill. Forecasting systems are, however, inherently uncertain because of uncertainties in, among other things, the input data used to produce these forecasts such as initial conditions, boundary conditions, and subgrid parametrization. Useful forecasts need to quantify the uncertainties in their predictions so that the reliability of the forecast can be assessed.

The present article analyzes the performance of two recent efforts (Iskandarani et al., 2016a; Li et al., 2016) that have relied on Polynomial Chaos (PC) methods (Ghanem and Spanos, 1991; Xiu and Karniadakis, 2002; Le Maître and Knio, 2010; Iskandarani et al., 2016b) to quantify the uncertainties in forecasting the circulation in the Gulf of Mexico stemming from uncertainties in the initial conditions alone (Iskandarani et al., 2016a) or in combination with wind forcing uncertainties (Li et al., 2016). The forecast timeline covers the oil spill period from May 1–30 2010 and coincides with an extended Loop Current (LC) that threatened to spread the oil along the south Florida coast and, eventually, the Eastern Seaboard of the United States. Fortunately, a LC detachment (LC eddy Franklin) occured and confined the oil to the northern and central parts of the Gulf of Mexico. The uncertainty analysis explores primarily whether the uncertainty in the LC location can be quantified given the uncertainties in the forecast model's data.

The studies in Iskandarani et al. (2016a) and Li et al. (2016) were based on perturbing the model fields (initial conditions and wind forcing) with space-time patterns obtained from an Empirical Orthogonal Function (EOF) decomposition where the amplitudes of these patterns were considered uncertain parameters. The PC formalism was then applied to propagate the uncertainties forward efficiently by: first, running an ensemble of simulations using HYbrid Coordinate Ocean Model (HYCOM) to sample the uncertain parameter space; second, constructing polynomial-based model-surrogates that accurately represent the changes in model outputs caused by changes in model inputs; and third, using these surrogates to perform a reliable and efficient statistical analysis once the validity of the surrogates was established.

The choices made during the course of the uncertainty analysis in those two articles, and which will be detailed in later sections, have raised a number of issues that we wish to address here concerning the "realism" of the uncertainty analysis, the computational and information trade-offs in choosing different uncertain inputs, and the exploration of the statistical information conveyed by the PC approach. Specifically, in the present study, we 1) assess the performance of the EOF-perturbed PC-ensemble by comparing it to observational data, both at the surface and at depth, to verify whether the measurement data falls within the envelope of the PC ensemble; 2) leverage the ability of PC methods to deliver output Probability Density Function (PDF) to quantify, using information theoretical measures, the uncertainty lost by omitting some uncertain inputs or by limiting their variability. A second aim of this paper is to explore the statistics of the ensemble. In order to obtain the most representative ensemble member and to identify the outliers, contour boxplot (Whitaker et al., 2013), a generalization of the conventional boxplot, is applied to the ensemble. Furthermore, the output PDFs delivered by the PC method are used to explore the non-Gaussian statistics in the vicinity of the LC region.

In summary, the present article is a follow on to Iskandarani et al. (2016a) and Li et al. (2016). Iskandarani et al. (2016a) identified the two leading EOF modes whose amplitudes represented the uncertainties in the strength of an LC Frontal Eddy; these modes were subsequently used to perturb the initial conditions of a control forecast. Iskandarani et al. (2016a) relied on a 49 member ensemble to build surrogates of model outputs, validated their accuracy and used them for the statistical analysis. Li et al. (2016) expanded the previous study by including additional EOFs modes in the initial conditions perturbations as well as perturbations to the surface wind forcing. Their parameter space was eight-dimensional and required a compressed-sensing based procedure to construct model surrogates using a 798-member ensemble. A variance-based sensitivity analysis showed that uncertainties in the initial conditions dominated the forecast uncertainties in the deep parts of the Gulf of Mexico while wind forcing uncertainties were the dominant contributors on the continental shelves. The present study compares the ensembles simulations to observations to assess whether the EOF perturbations were adequate at representing the uncertainties in the forecast, performs a cost-benefit analysis regarding the enlargement of the uncertain parameter space and perform additional analysis regarding the statistical distribution of sea surface height at the end of the forecast. No additional experiments were performed in the present study.

The layout of this paper is as follows. Section 2 provides a quick overview of the LC dynamics in the Gulf of Mexico, summarizes the experimental setup of the two uncertainty experiments, provides a brief description of the PC methodology and describes the specification of the input uncertainties. Section 3 compares the ensemble results against observational data. The information trade-offs between the different choices of the sources and variability of the input uncertainties are shown in section 4. Section 5 presents the contour boxplot of the LC edge and the sea surface height (SSH) PDFs. Finally, we conclude with a summary section.

2. Model and ensemble prediction

The Gulf of Mexico, where the Deepwater Horizon oil spill took place, is a suitable test bed for uncertainty studies. It is a well-observed regional sea that presents many dynamical features typical of the deep ocean such as currents and eddying jets. As shown in Figure 1, the LC is a particularly dominant feature of the circulation in the Gulf of Mexico as it flows from the Yucatan Channel between Mexico and Cuba, to the Straits of Florida between Cuba and the Southeastern U.S. The LC presents a time varying extension, from a retracted path at the south of the basin, to an extended one reaching the edge of the continental shelf in the northeastern Gulf. When it is extended, the LC sheds a large, anticyclonic eddy, called LC Eddy (indicated by the anticyclonic arrow, in black, in the western Gulf), which then drifts westward, and the LC retracts to the south. This shedding sequence often implies temporary detachments of the LC Eddy from the current, before final separation. Small, cyclonic eddies, also called LC Frontal Eddies (shown in white arrows), at the edge of the LC play an active role in necking down and chopping the extended LC, leading to the LC Eddy detachments or separation (Zavala-Hidalgo et al., 2003; Schmitz, 2005; Athié et al., 2012; Le Hénaff et al., 2012a, 2014). The Deepwater Horizon oil spill took place during such a LC Eddy shedding sequence, and the fate of the spilled oil was partly influenced by the LC evolution and its frontal dynamics (Walker et al., 2011). The model setup described below was configured primarily to investigate the uncertainties in this eddy shedding scenario.

2.1. HYCOM setup

The forecast model is the Hybrid Coordinate Ocean Model (HYCOM). The model configuration is the same as GOMl0.04 expt_20.1 run by the Navy Research Laboratory (NRL) for the near-real time system in the period 2003-2010. The details of this configuration can be found at HYCOM website[1].

[1]https://hycom.org/data/goml0pt04/expt-20pt1 (last access on July 3$^{\text{rd}}$, 2018)

The model has a horizontal grid resolution of 1/25 degree and 20 vertical layers. Since the vertical layers in HYCOM are hybrid, their thickness changes at each time step. In this configuration, there are more vertical layers toward the surface, with their depth, in the Eastern Gulf, ranging from 1.5m to about 2700 m in the Eastern Gulf. The computational domain is open along portions of its southern, eastern and northern boundaries, where values are provided by a lower resolution 1/12 degree North Atlantic HYCOM simulation (Chassignet et al., 2007). This model configuration has been used extensively in the literature, especially in studies of the Deepwater Horizon oil spill (e.g. Mezić et al. (2010); Valentine et al. (2012); Le Hénaff et al. (2012b); Paris et al. (2012)). The model is forced by the 27 km resolution Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) atmospheric products. The initial condition for the model is from the expt_20.1 (McDonald, 2006) near-real time simulation run at NRL, which includes data assimilation. The model assimilates available satellite altimeter observations (along track data altimetry obtained via the NAVOCEANO Altimeter Data Fusion Center), satellite and in situ sea surface temperature (SST) as well as available in situ vertical temperature and salinity profiles from XBTs, ARGO floats and moored buoys. The model is then integrated forward in time without data assimilation, in forecast mode, for 30 days from May 1, 2010 to May 30, 2010.

2.2. PC surrogates

We give a brief overview of PC methods in order to set the stage for the subsequent analysis; more background information can be found in Le Maître and Knio (2010); Iskandarani et al. (2016b) and references therein. The PC paradigm is based on describing the dependence of a specific model output, say $M(\boldsymbol{x}, t, \boldsymbol{\xi})$ where $\boldsymbol{\xi}$ represents the vector of uncertain inputs and $\boldsymbol{x}$ and t refer to space and time, by a spectral series M_P of the form:

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

where the $\Psi_m(\boldsymbol{\xi})$ are the user specified multi-dimensional basis functions (usually tensorized orthogonal polynomials from the Askey family,(Xiu and Karniadakis, 2002)), and the $\widehat{M}_k(\boldsymbol{x}, t)$ are $(P+1)$ coefficients. These coefficients are determined by sampling the parameter space $\boldsymbol{\xi}$ and minimizing the error $\|M - M_P\|$. Different versions of PC methods can be derived by choosing different error norms and sampling strategies. For example, the traditional Galerkin approach uses the so-called $\mathcal{L}_2$ norm:

$$\|M - M_P\|_2^2 = \int (M - M_P)^2\ \rho(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \tag{2}$$

where $\rho(\boldsymbol{\xi})$ is the probability density function of the uncertain inputs $\boldsymbol{\xi}$. This approach takes advantage of the orthogonal basis and uses quadrature rules to calculate the coefficients as:

$$\widehat{M}_k = \frac{\sum_{q=1}^{Q} \Psi_k(\boldsymbol{\xi}_q) M(\boldsymbol{\xi}_q)\omega_q}{\sum_{q=1}^{Q} \Psi_k(\boldsymbol{\xi}_q)\Psi_k(\boldsymbol{\xi}_q)\omega_q} \tag{3}$$

where $\boldsymbol{\xi}_q$ and ω_q are multi-dimensional quadrature points and weights (Le Maître and Knio, 2010; Iskandarani et al., 2016b). Other approaches to determining the coefficients include spectral collocation and regression (useful when the model output M is noisy, see Iskandarani et al. (2016b) for a comparison of these different techniques). Regardless of the specific approach to calculate the coefficients, the PC series requires sampling of the parameter space to compute $M(\boldsymbol{\xi}_q)$, and this constitutes the most expensive portion of the calculation as each sample requires a model run with the uncertain input set to $\boldsymbol{\xi_q}$. In general the cost increases exponentially with the dimension of the uncertain input $\boldsymbol{\xi}$ and must be mitigated by resorting to either sparse quadrature rules or sparse series construction. Once the coefficients $\widehat{M}_k$ are available (and the series approximation errors have been verified to be small), the spectral series in equation (1), often referred to as a surrogate or emulator, can be used in lieu of the model to estimate the response of the model M_P to changes in the uncertain input data $\boldsymbol{\xi}$. The PC approach provides an efficient way to propagate model uncertainties, quantify principal contributors to the model output uncertainties and infer the posterior distributions of uncertain inputs given observational data. PC methods have been successfully applied to many different uncertainty quantification tasks for oceanic and atmospheric simulations (Thacker et al., 2012; Li et al., 2016; Iskandarani et al., 2016a; Winokur et al., 2013; Wang et al., 2015; Alexanderian et al., 2012; Sraj et al., 2013).

2.3. PC input uncertainties

The two HYCOM ensembles in Iskandarani et al. (2016a) and Li et al. (2016) relied on reduced state space methods (Kleeman, 2011) to characterize the input uncertainties, so that the variance of the uncertain inputs was maximized while retaining as few uncertain inputs as possible. More specifically, EOF decompositions were used to identify modes of variability in the initial conditions and wind forcing. The spatial patterns of the perturbation were thus provided by the EOFs while the time series were associated with their principal components.

The EOF modes used to perturb the initial conditions were obtained from a multivariate EOF analysis of two weeks of daily outputs of the operational, and data-assimilated, simulation prior to our experiment. LC processes and their frontal instabilities are the dominant contributors to variability during this 14-day time period, and contamination of the EOF modes by other, longer time scale processes is thus minimized. Iskandarani et al. (2016a) analyzed the first two of these EOF modes and showed that their addition to the initial condition of the unperturbed run led to a stronger frontal eddy in the northeast corner of the extended LC and an early separation of a LC Eddy, whereas their subtraction had the opposite effect. The wind forcing EOF modes were calculated from a 2 month time-series (May and June,2010) of a 27 km resolution COAMPS simulation. The EOF analysis was performed on the wind velocity vectors (u wind and v wind) and then projected onto the wind speed and the wind stress vectors, which are the actual components of the HYCOM wind forcing inputs. Figure 2 shows the cumulative variance of each mode identified in the 2-week EOF decomposition of daily operational HYCOM outputs (left panel), and in the 60-day EOF decomposition of the COAMPS winds.

The initial conditions and wind-forcing fields can now be constructed as the sum of products of spatial patterns and time series as follows:

$$u(\boldsymbol{x}, t=0, \boldsymbol{\xi}) = u_0(\boldsymbol{x}) + \alpha^{ic} \sum_{k=1}^{K^{ic}} \xi_k^{ic} \mathcal{U}_k(\boldsymbol{x}) \tag{4}$$

$$f(\boldsymbol{x}, t, \boldsymbol{\xi}) = f_0(\boldsymbol{x}, t) + \alpha^{w} \sum_{k=1}^{K^{w}} \xi_k^{w} \lambda_k^{w}(t) \mathcal{F}_k(\boldsymbol{x}) \tag{5}$$

where $u(\boldsymbol{x}, t = 0, \boldsymbol{\xi})$ is the perturbed initial condition field, u_0 is the unperturbed initial field, α^{ic} is a coefficient that controls the size of the perturbation, $\mathcal{U}_k(\boldsymbol{x})$ are the EOFs obtained from the decomposition of the two-week HYCOM daily output, $-1 \leq \xi_k^{ic} \leq 1$ are standardized uncertain input random variables controlling the amplitude of the EOFs modes, and K^{ic} refers to the number of EOF modes retained. The terms f, f_0, α^w, $\mathcal{F}_k$ and K^w are the analogous quantities for the wind-forcing field. The λ_k^w in the wind forcing perturbations refer to the principal components of the two-month wind time series.

The exploratory study in Iskandarani et al. (2016a) included only the first two initial condition EOF modes in order to keep the computational cost tractable; the vector of uncertain input was $\boldsymbol{\xi}^\top = (\xi_1^{ic}, \xi_2^{ic})$ with $K^{ic} = 2$ and $K^w = 0$. Li et al. (2016) increased the number of initial conditions modes to four to explore the impact of additional variability on the forecast uncertainty, and included four wind forcing modes to account for additional sources of uncertainties. The uncertain input vector consisted thus of $\boldsymbol{\xi}^\top = (\xi_1^{ic}, \xi_2^{ic}, \xi_3^{ic}, \xi_4^{ic}, \xi_1^{w}, \xi_2^{w}, \xi_3^{w}, \xi_4^{w})$ with $K^{ic} = 4$ and $K^w = 4$. Note that Li et al. (2016) decreased the size of their perturbation by setting $\alpha^{ic} = 0.8$ and $\alpha^w = 0.04$ in order to avoid repeated crashes of the forward model when the full perturbation was applied. The PC approach treats the perturbation amplitudes as independent and continuous random variables characterized by their PDFs which were assumed to be uniform[2].

The focus on initial conditions and wind forcing uncertainties in the studies of Iskandarani et al. (2016a) and Li et al. (2016) is largely a compromise between computational cost[3], and the desire to account for most of the uncertain processes influencing LC dynamics. For example, uncertainty in boundary conditions was omitted since the domain boundaries were too remote to influence the LC within a 30 day time frame[4]. Uncertainty in the wind field was deemed the second most important contributor to LC dynamics which was then included in Li et al. (2016). Additional sources of uncertainty, such as surface heat-flux, could be included at the expense of increasing the dimension of the uncertain space and the sampling cost.

The PC surrogate in Iskandarani et al. (2016a) relied on a Galerkin projection with sampling on the Gauss-Legendre quadrature points to determine the coefficients, and was shown to be valid for a period of 40 days when validated against independent model simulations. The eight-dimensional space of Li et al. (2016) required a different surrogate construction approach and the latter was built using a Basis Pursuit Denoising[5] algorithm. The va-

[2]Specifying these PDFs can be difficult in practice due to the scarcity of observational data. The availability of the surrogate, however, allows the user to explore the effect of different input PDFs at little extra cost.

[3]The cost increases quickly with the number of uncertain inputs.

[4]Roughly 60 days are needed for a perturbation in the boundary condition to reach the LC region (Thacker et al., 2012).

[5] The basis pursuit denoising algorithm seeks to find the shortest series possible whose coefficients minimize the square of the surrogate error; it mitigates the cost of sampling an 8-dimensional space to compute the $\widehat{M}_k$'s (a Gauss quadrature procedure as in Iskandarani et al. (2016a) would have required $7^8 = 5,764,801$ samples).

	Iskandarani et al. (2016a)	Li et al. (2016)
uncertain inputs	Initial Conditions	Initial Conditions & Wind Forcing
# of IC EOF modes K^{ic}	2	4
# of WF EOF modes K^{w}	0	4
dimension of $\boldsymbol{\xi}$-space	2	8
perturbation scale α^{ic}	1	0.8
perturbation scale α^{w}	0	0.04
input pdf $\rho(\boldsymbol{\xi})$	uniform: 2^{-2}	uniform: 2^{-8}
surrogate basis	Legendre polynomials	Legendre polynomials
Coefficient	Galerkin Projection	Basis Pursuit Denoising
Sampling	Gauss Quadrature	Latin Hypercube
Ensemble size	49	798

Table 1: Summary of the two uncertainty quantification experiments.

lidity of the surrogates was also established by comparing their estimates to those of an independent validation ensemble. The analyses herein focus on the first 30 days of the simulation when both surrogates delivered accurate representation of the model output. Table 1 lists the settings for the two uncertainty quantification analyses in Iskandarani et al. (2016a) and Li et al. (2016); the reader is referred to these articles for more details on the surrogate construction and their validation.

The goals of the present study is to evaluate the perturbation strategy used to generate these ensembles and to compare the evolution of the individual members in relation to the evolution of the observed LC system, and to assess whether the increase in the number of uncertain parameters, and the associated increase in the sampling requirements, yield to a better estimate of the output uncertainties. Furthermore, the output PDFs of the ensembles are computed and compared to those of the climatological observations.

3. Model-data comparison

For a "good" ensemble prediction, the forecast uncertainties should be properly represented (Slingo and Palmer, 2011) such that the true evolution appears to be a plausible realization in the ensemble. In order to assess whether the EOF perturbations satisfy this requirement, we compare the envelop of the model ensembles against observational data both at the surface and at depth.

3.1. Comparison against satellite SSH

For the surface model-data comparison, we use the AVISO gridded satellite SSH data optimally interpolated to a $1/4° \times 1/4°$ grid (Le Traon et al., 1998). Specifically, we compare the edges of the LC and the LC eddies, which are defined by the contours of the 17cm SSH anomaly with respect to the basin mean value. Leben (2005) introduced this 17 cm anomaly as a reliable indicator of the Loop Current edge's position by comparing it to other criteria that were traditionally used before. It is now a commonly used metric to identify the edge of the LC (e.g., Le Hénaff et al. (2012a); Dukhovskoy et al. (2015)). Figure 3 shows time snapshots of the LC edge every 10 days starting from May 1, 2010 for both HYCOM ensembles. The background color is the gridded satellite SSH data. The LC edge from the HYCOM ensemble members (black contours) are compared against that from the gridded satellite SSH data (white contours).

A basic observation for both ensembles is that the LC contour derived from satellite data (shown in white contours in Figure 3) generally fall within the envelope of the HYCOM ensemble contours (shown in black) near the

LC region. This indicates a “good” ensemble since the observational data appears to be a plausible realization of the model ensemble. On day 30, it is clear that the LC eddy shedding process is affected by the ensemble perturbation, in which several ensemble members have already shed an eddy while others have not. The difference between these two HYCOM ensembles is visually small, which is a first indication that the additional uncertainty in the initial conditions and in the wind forcing do not contribute much to the uncertainty in the LC edge position.

Both ensembles, however, deviate from the observational data in the vicinity of a detached LC Eddy in the western part of the Gulf of Mexico; moreover, the deviation increases with time (this is not too surprising as each ensemble member is a “free” run without data assimilation). The EOF perturbations in both the initial conditions and the wind forcing seem to have missed the local uncertainty in the vicinity of the western LC Eddy. We speculate that the reason for this is, first, mainly concerned with initial conditions uncertainties; and second, that the EOF decomposition of the initial conditions picked up the largest variability in the eastern side of the basin, namely the one associated with LC dynamics. As a result local variability away from the most dynamic region might not have been captured in the first four EOF modes. The PC paradigm, with its focus on establishing a functional relationship between the uncertain inputs and the model output, requires that modelers be careful and deliberate in selecting their input perturbations. It also suggests that a more tailored decomposition (or model reduction method) would be useful if the user is interested in quantifying the uncertainties in multiple regions simultaneously.

3.2. Comparison against mooring data

In order to compare the model with observational data at depth, we compare both HYCOM ensembles against 9 full-depth mooring observations deployed by the Bureau of Ocean Energy Management (BOEM)/Science Applications International Corporation (SAIC) (Hamilton et al., 2016) during the mutual HYCOM simulation period (May 1, 2010 - May 30, 2010). Here, we compare the mean and standard deviation ellipses of the point velocity for the entire 30 days, like in Xu et al. (2013). Figure 4 shows the comparison of the mean and standard deviation ellipses between the HYCOM ensemble members and the mooring data at different depths. In each subfigure, the mean velocity vectors of the ensemble members are shown in black and the standard deviation ellipses of the ensemble members are represented in red. The mean and standard deviation ellipses of the mooring observations are in blue. The mooring data falls in general within the envelope of the HYCOM ensembles at different depths; this indicates that the two ensembles at depth capture reasonably well the observations. The 49-member PC ensemble, with only two initial conditions EOF perturbations, underestimates the variability observed in the two most northeastern moorings at the 100m and 300m depths, whereas it seems to be well-represented in the 798 member ensemble.

Both at the surface and at depth, the ensembles are "realistic" as evidenced by the model-data comparison: the ensembles captured the observed evolution, especially in the targeted LC region. Regarding the LC edge, the small ensemble can capture similar amount of variability compared with the large ensemble at the surface and at depth.

4. Relative Entropy

The uncertainty experiment in Li et al. (2016) was largely motivated by two considerations: first to explore the impact of including additional EOF modes, and, second, to include other sources of input uncertainties, specifically wind forcing uncertainties. Other sources of uncertainty, such as surface heat flux, river-runoff or open boundary conditions, were deemed to be less important or too remote for short term forecasting, and were thus not considered. The enlargement of the uncertain parameter space increases the sampling requirement substantially, and the natural question is whether the variability "gained" in the output, alternatively the missed uncertainties, justifies the increased sampling cost. To address this question, we attempt to quantify this variability loss by considering various scenarios where the dimension of the uncertain input space is reduced. Note that the reduction can be achieved either by discarding high order EOF modes and retaining only the leading order ones, or by supressing independent (initial conditions or wind forcing) sources of input uncertainty.

We leverage the ability of PC methods to deliver output PDFs to quantify the variability loss using an information theoretic measure, relative entropy (Kullback and Leibler, 1951), which measures the "distance" between two probability density functions p and q. Relative entropy can be defined in discrete form as follows (Kleeman, 2002):

$$D(p,q) = \sum_{i=1}^{K} p_i \log\left(\frac{p_i}{q_i}\right) \tag{6}$$

where p and q denote the PDFs of two distributions, i is the discrete bin index, and K is the total number of bins. The relative entropy is a measure of "distance" between the reference PDF, q, and the PDF p; it is zero when p and q are identical and increases as they grow apart. In what follows q refers to the output PDF that is obtained by including all uncertain parameters whereas p refers to the output PDF that is obtained by restricting the number of uncertain inputs. The relative entropy $D(p, q)$ thus quantifies the amount of variability lost by restricting the input uncertainty space.

The discrete PDFs are calculated as follows: a large number of samples is drawn from a PC surrogate, the range of a model variable (SSH for example) is then divided into bins and the probability of a specific bin is set to the number of samples in this bin divided by the total number of samples, i.e.

$$p_i = \frac{N_i}{\sum_{i=1}^{K} N_i} \tag{7}$$

where N_i represents the number of samples in bin i, and K is the number of bins. We use the PC surrogate to generate a large number of samples (100,000) and set $K = 20$ for all relative entropy calculations[6]. The PC surrogate used here is the one constructed from the large ensemble since it has already been built and validated (Li et al., 2016), and since it encompasses the largest uncertain input space. Table (2) displays several scenarios of restricting the uncertain input space.

[6]Experimentation has shown that our results are not sensitive to K when $K \geq 20$.

Scenarios	Initial Condition	Wind Forcing
0) Reference calculation	$(\xi_1^{ic}, \xi_2^{ic}, \xi_3^{ic}, \xi_4^{ic})$	$(\xi_1^{w}, \xi_2^{w}, \xi_3^{w}, \xi_4^{w})$
1) omitting high order EOF modes	$(\xi_1^{ic}, \xi_2^{ic}, 0, 0)$	$(\xi_1^{w}, \xi_2^{w}, 0, 0)$
2) only leading IC EOF modes	$(\xi_1^{ic}, \xi_2^{ic}, 0, 0)$	$(0, 0, 0, 0)$
3) only leading wind EOF modes	$(0, 0, 0, 0)$	$(\xi_1^{w}, \xi_2^{w}, 0, 0)$
4) only 1st IC EOF modes	$(\xi_1^{ic}, 0, 0, 0)$	$(0, 0, 0, 0)$
5) only 2nd IC EOF modes	$(0, \xi_2^{ic}, 0, 0)$	$(0, 0, 0, 0)$

Table 2: Left column: Different uncertainty perturbation scenarios. Middle colum: The range of the initial condition uncertain inputs variables $|\xi_i^{ic}| \leq 1$ where $\xi_i^{ic} = 0$ means no perturbation. Right column: The range of the wind forcing uncertain inputs variables $|\xi_i^{w}| \leq 1$ where $\xi_i^{w} = 0$ means no perturbation.

4.1. Results for sea surface height

We first investigate the variability loss caused by retaining two EOF modes only instead of four (scenario 1 in table 2), for both sources of uncertainty. Figure 5.a shows the time evolution of the relative entropy between the reference scenario that perturbs all eight EOF modes and the scenario that perturbs only the first two leading EOF modes of each uncertain source. The results show that, for the time span considered, little variability is lost by ignoring the uncertainty due to higher EOF modes (modes three and four of each uncertain source), and that there is very little gain in expanding the uncertainty space to include these higher order modes.

Next, we investigate the variability loss caused by omitting uncertainty in the wind forcing. Since Figure 5.a shows that the high order EOF modes contribute little to the information content of the SSH PDF, we focus only on the low order EOF modes here. Figure 5.b shows the variability loss by omitting uncertainty in the wind forcing (scenario 2 in table 2), in which the ensemble that perturbs only the first two initial condition EOF modes is compared with the fully perturbed ensemble. The impact of omitting the wind forcing EOF modes is mainly in the coastal region and little influence can be found in the LC region and the western Gulf of Mexico at day 30.

The variability loss by omitting the initial condition EOF modes is shown in Figure 5.c (scenario 3 in table 2). The coastal signal disappears in this case, which indicates that the first 2 wind forcing EOF modes dominate SSH signal in the coastal region. Instead, a strong signal is observed in the LC region and its magnitude grows as time evolves, which indicates that the initial condition uncertainty is the dominant contributor to uncertainty in the LC region. These results are consistent with the analysis of variance in Li et al. (2016), in which the SSH in the LC region is more sensitive to the initial condition EOF perturbation modes while in the coastal region the SSH is more sensitive to the wind forcing EOF modes.

Since the first two initial condition EOF modes contribute the most to the SSH variability in the deep Gulf of Mexico, we further investigate the influence of each individual initial condition EOF mode. Figure 5.d shows the SSH variability loss by retaining only the 1st initial condition EOF mode and Figure 5.e shows the SSH variability loss by retaining only the 2nd initial condition EOF mode. Without the 2nd initial condition EOF mode (shown in Figure 5.d), the SSH variability loss is localized in the north side of the LC , especially at day 30. Without the 1st initial condition EOF mode (shown in Figure 5.e), the SSH variability loss is substantial in a localized region along the LC and its amplitude grows with time. In the deep Gulf of Mexico, two initial condition EOF modes are necessary in order to capture most of the SSH variability.

In summary, for SSH, the variability loss caused by retaining only the two leading EOF modes for initial condition and wind forcing uncertainties is rather negligible. Additionally, the variability loss by omitting wind forcing input uncertainty is small in the LC region but is quite large in the coastal zone. Thus, this posterior analysis suggests that a more optimal ensemble could have been designed by retaining only the first two leading EOF modes of initial conditions and wind forcing uncertainties. These results also indicate that the small HYCOM ensemble used in Iskandarani et al. (2016a), that perturbs only the first two initial condition EOF modes, is a suitable choice for studying the ensemble statistics in the LC region as we do in the following section. It should be noted that the relative entropy approach here could be generalized to any ensemble simulations to investigate the balance between the size of the ensemble and the information content contained in the ensemble.

5. Ensemble statistics

5.1. Ensemble visualization

Ensemble simulations provide us with a way to derive statistics of the model, which leads to an estimate of the confidence of the model prediction. However, mining useful information from the ensemble can be challenging since ensemble simulations usually involve a large number of single model runs. The recently developed contour boxplot (Whitaker et al., 2013), built on the notion of data depth, enables us to extract valuable information from an ensemble. Here, we are particularly interested in answering: 1) what is the most representative ensemble member? 2) what are the outliers of the ensemble?

Simply speaking, the contour boxplot can be considered as a generalization of the conventional boxplot. Both of these methods are designed for order statistics which requires an ordering of the data. The difference between conventional boxplot and contour boxplot is that the former can only be applied to scalar quantities while the latter can be applied to functions and contours. At the heart of the contour boxplot, a measure of centrality is defined by the concept of data depth proposed from the statistics community (Whitaker et al., 2013). The gist of data depth concept is to quantify the centrality or depth of a data sample with respect to an ensemble of data samples. In practice, the centrality or depth can be measured by how many times a function or contour falls within the band formed by an arbitrary set of other functions or contours. Figure 6, adopted from Whitaker et al. (2013), shows an example of how to measure the centrality or depth. In the left subfigure, three different blue curves form a grey band and three red curves are tested against the band. The solid red curve falls completely within the grey band, while the two dashed red curves partially fall within the grey band. Therefore the data depth of the solid red curve is larger than that of the dashed red curves. For contours in the right subfigure, the same logic applies. We refer the reader to Whitaker et al. (2013); Mirzargar et al. (2014) for more descriptions of the methodology as well as an application in Meteorology. We apply the contour boxplot concept to the LC edges obtained from the 49-member HYCOM ensemble. The key effort of the contour boxplot is to sort the 49 ensemble contours by their data depth defined in Whitaker et al. (2013). The "deeper" data sample is considered to be more representative than the "shallower" data sample. The "deepest" data sample can be considered as the most representative ensemble member and the "shallowest" data samples can be considered as outliers.

Figure 7 shows the contour boxplot for the edges of the LC and LC Eddies, as defined by the 17 cm SSH contours, at day 30. The satellite SSH observation is shown in black for reference. The satellite observations show that a LC eddy has separated from the LC at that date, while in the mean of the ensemble simulations, indicated by the green line, the LC is still in its extended position. On the other hand, the most representative ensemble member identified by the median of the ensemble (in yellow) shows a similar LC eddy shedding stage compared with the satellite observations. We mark the "shallowest" three ensemble members (shown in red) as outliers. It is clear that these outliers are still in their early stage for LC eddy shedding process, which is somewhat "slow" compared with other ensemble members. In the uncertain parameter space, these outliers are the ensemble members with extreme negative perturbations in both modes as seen in the inserted box in Figure 7.

Next, we connect the normalized uncertain perturbation to the estimated initial condition perturbation pattern. Figure 8 shows the initial perturbation approximated by the SSH difference between the perturbed and unperturbed runs one day after the start of the simulation. The initial perturbations are shown according to their normalized random variables shown in the middle of the figure. The bottom left subfigure represents the most negative perturbations for both EOF modes; the top right subfigure represents the most positive perturbations for both EOF modes. The dynamical process associated with EOF mode1 (the ξ_1 direction) can be explained by the strengthening or weakening of the LC frontal eddies. The signature of EOF mode2 (the ξ_2 direction) is associated with variability along the edge of the LC, as well as in the western Gulf of Mexico. These signals strongly affect the intensity of the LC frontal eddies, which play an important role in the LC eddy shedding process (Le Hénaff et al., 2012a, 2014). The outliers identified by the data depth concept in Figure 7 are located at the most negative perturbation of EOF mode1, which is consistent with the related dynamical processes: negative perturbation on EOF mode1 is associated with the weakening of the LC frontal eddies, which delays the LC eddy shedding event.

5.2. An exploration of SSH PDF

The estimate of the full PDF usually requires a large sample size. Sometimes, only the low order statistical moments are calculated on the assumption that the underlying PDF can be approximated by a normal distribution. PC method provides us with an efficient way to estimate the full PDF of a model output. We thus investigate whether the underlying distribution of the SSH field in the Gulf of Mexico is normal using the PC surrogate constructed by the 49-member HYCOM ensemble. We first explore the pointwise SSH PDF at different locations in the Gulf of Mexico. In Figure 9, the location of four selected points (A1-A4) and the LC edges are shown on the right subfigure. On the left subfigure, the SSH PDF from the four different locations are presented. Clearly, the SSH PDF is not always normally distributed. For example, the SSH PDF at point A3 along the LC edge shows a bimodal distribution whereas the PDF at A2 shows a bias towards higher values. Next, we investigate where in the Gulf of Mexico the SSH are normally distributed. We reuse the relative entropy metric to quantify the distance between the

PC-surrogate PDF and its Gaussian counterpart; this Gaussian counterpart is obtained by specifying the mean and variance as calculated from the PC surrogate.

Figure 10 shows the relative entropy map between the SSH PDF in each grid cell and its Gaussian counterpart. We only plot the regions where the associated relative entropy is larger than or equal to 0.4 and the LC edges are also shown in the figure for reference. It is clear that the strongest non-Gaussian signal appears in the LC region (indicated by red color), especially in the place where the variation in the LC edge is high. This is not surprising, as the LC is a highly nonlinear dynamical feature (Oey et al., 2005), so the LC region is expected to show non-Gaussian statistic. The relative entropy is an efficient tool to identify the locations of highly nonlinear features in an ensemble of simulations.

6. Summary and Discussion

This paper analyzes two ensembles designed to quantify, using a polynomial chaos approach, the uncertainties in forecasting the circulation in the Gulf of Mexico during the Deep Water Horizon period of May 1 to May 30, 2010. The two ensembles differed in the sources of uncertainty (initial conditions and wind forcing uncertainties) and in the amount of variability (number of perturbation modes) accounted for. Both ensembles relied on EOF decomposition to perturb the initial conditions and wind forcing fields, and considered the amplitude of these modes as the uncertain input parameters. The EOF-based perturbations served to maximize the "amount" of input uncertainty using the smallest number of uncertain inputs, so that the size of the ensemble required to sample the uncertain space remains manageable.

The model data comparison reveals that, in general, the observations fall within the envelope of uncertainty generated by the EOF perturbations during the 30-day forecast period. This is particularly true in the target area of the LC region where two EOF modes used to perturb the initial conditions are enough to capture the variability in the system. On the other hand, the altimetry data shows the remnant of a LC Eddy in the western side of the basin evolving outside the envelope predicted by the two ensembles. This could be explained by an inability of our basin-wide EOF modes to capture simultaneously localized uncertainty on the western side of the basin and in the LC region, and that the local uncertainty in the western side did not project on the first four EOF modes. One remedy is to compute separate EOF modes in the Eastern and Western sides of the basin so that the variability in the former would not dominate the variability in the latter. Other approaches would involve abandoning the EOF decomposition and resorting to perturbing the system using either singular modes (Buizza and Palmer, 1995) or bred vectors (Toth and Kalnay, 1997). However, the computations of the singular vectors would require the availability of the tangent linear model and its adjoint, and would inccur additional computational costs. Likewise, identifying the bred vectors would require running an ensemble of simulations prior to obtaining the PC ensemble itself.

The relative entropy metric was used to quantify the variability loss/gain caused by accounting for different uncertainty sources and by including different "amounts" of input variability. It shows that the variability loss caused by omitting higher EOF modes in the input uncertainty is small, at least for the 30 day period considered. This result suggests that adding additional input uncertainty sources is more useful than adding high order EOF modes of the same uncertainty source. The uncertainty in forecasting the SSH field in the shelf regions, for example, is primarily caused by uncertainties in the wind forcing while the initial conditions uncertainty plays a secondary role. The wind forcing uncertainty adds little to the SSH forecast uncertainty in the deep parts of the Gulf where two EOF modes used to perturb the initial conditions are enough to account for the forecast uncertainty in these regions. The conclusions obtained here using the relative entropy metric are consistent with the variance-based sensitivity analysis in Li et al. (2016).

The analysis of the SSH PDFs shows a strong non-Gaussian signal in the LC region, which is reflective of the bifurcation in the state of the LC caused by the eddy detachment. Furthermore, the contour boxplot allowed us to identify the most representative ensemble member and the ensemble outliers.

The application of uncertainty quantification techniques in ocean modeling is in its early stages. The most challenging part is to reduce the dimensionality of the problem to minimize the sampling cost of the uncertain input space while capturing the largest amount of input uncertainty. The PC paradigm emphasizes the link between the input and output uncertainties by explicit construction of a surrogate, and allows forecasters to identify the dominant contributors to the output uncertainties. The efficiency of this uncertainty quantification implementation can be applied for model calibration to guide the selection of model parameters. The availability of the full output PDF will open the door for data assimilation and predictability study in a non-Gaussian paradigm.

Acknowledgements

This research was made possible in part by a grant from The Gulf of Mexico Research Initiative, and in part by the US Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research, under Award Number DE- SC0008789. This research was conducted in collaboration with and using the resources of the University of Miami Center for Computational Science, as well as the resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. M. Le Hénaff received partial support for this work from the NOAA Quantitative Observing System Assessment Program (QOSAP, grant NA15OAR4320064) and the base funds of the NOAA Atlantic Oceanographic and Meteorological Laboratory. Data are publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC) at https://data.gulfresearchinitiative.org (doi:10.7266/N77H1GNF, doi:10.7266/N7QZ2813). This study has been conducted using E.U. Copernicus Marine Service Information. The mapped altimetry products were produced by Ssalto/Duacs and distributed by the Copernicus Marine Environment Marine Service (http://marine.copernicus.eu).

References

Alexanderian, A., Winokur, J., Sraj, I., Srinivasan, A., Iskandarani, M., Thacker, W. C., Knio, O. M., 2012. Global sensitivity analysis in an ocean general circulation model: a sparse spectral projection approach. Computational Geosciences 16 (3), 757–778.

Athié, G., Candela, J., Ochoa, J., Sheinbaum, J., 2012. Impact of caribbean cyclones on the detachment of loop current anticyclones. Journal of Geophysical Research: Oceans 117 (C3).

Buizza, R., Palmer, T., 1995. The singular-vector structure of the atmospheric global circulation. Journal of the Atmospheric Sciences 52 (9), 1434–1456.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., Bleck, R., 2007. The hycom (hybrid coordinate ocean model) data assimilative system. Journal of Marine Systems 65 (1), 60–83.

Dukhovskoy, D. S., Leben, R. R., Chassignet, E. P., Hall, C. A., Morey, S. L., Nedbor-Gross, R., 2015. Characterization of the uncertainty of loop current metrics using a multidecadal numerical simulation and altimeter observations. Deep Sea Research Part I: Oceanographic Research Papers 100, 140–158.

Ghanem, R. G., Spanos, P. D., 1991. Stochastic finite elements: a spectral approach. Vol. 41. Springer.

Hamilton, P., Lugo-Fernández, A., Sheinbaum, J., 2016. A loop current experiment: Field and remote measurements. Dynamics of Atmospheres and Oceans 76, 156–173.

Iskandarani, M., Le Hénaff, M., Thacker, W. C., Srinivasan, A., Knio, O. M., 2016a. Quantifying uncertainty in gulf of mexico forecasts stemming from uncertain initial conditions. Journal of Geophysical Research: Oceans 121 (7), 4819–4832.

Iskandarani, M., Wang, S., Srinivasan, A., Carlisle Thacker, W., Winokur, J., Knio, O. M., 2016b. An overview of uncertainty quantification techniques with application to oceanic and oil-spill simulations. Journal of Geophysical Research: Oceans 121 (4), 2789–2808.

Kleeman, R., 2002. Measuring dynamical prediction utility using relative entropy. Journal of the atmospheric sciences 59 (13), 2057–2072.

Kleeman, R., 2011. Information theory and dynamical system predictability. Entropy 13 (3), 612–649.

Kullback, S., Leibler, R. A., 3 1951. On information and sufficiency. The Annals of Mathematical Statistics 22 (1), 79–86.

Le Hénaff, M., Kourafalou, V. H., Dussurget, R., Lumpkin, R., 2014. Cyclonic activity in the eastern gulf of mexico: Characterization from along-track altimetry and in situ drifter trajectories. Progress in Oceanography 120, 120–138.

Le Hénaff, M., Kourafalou, V. H., Morel, Y., Srinivasan, A., 2012a. Simulating the dynamics and intensification of cyclonic loop current frontal eddies in the gulf of mexico. Journal of Geophysical Research: Oceans 117 (C2).

Le Hénaff, M., Kourafalou, V. H., Paris, C. B., Helgers, J., Aman, Z. M., Hogan, P. J., Srinivasan, A., 2012b. Surface evolution of the deepwater horizon oil spill patch: Combined effects of circulation and wind-induced drift. Environmental science & technology 46 (13), 7267–7273.

Le Maître, O., Knio, O. M., 2010. Spectral methods for uncertainty quantification: with applications to computational fluid dynamics. Springer Science & Business Media.

Le Traon, P.-Y., Nadal, F., Ducet, N., 1998. An improved mapping method of multisatellite altimeter data. Journal of Atmospheric and Oceanic Technology 15 (2), 522–534.

Leben, R. R., 2005. Altimeter-derived loop current metrics. Circulation in the Gulf of Mexico: Observations and Models, 181–201.

Li, G., Iskandarani, M., Le Hénaff, M., Winokur, J., Le Maître, O. P., Knio, O. M., 2016. Quantifying initial and wind forcing uncertainties in the gulf of mexico. Computational Geosciences, 1–21.

McDonald, M., 2006. HYCOM + NCODA Gulf of Mexico 1/25° analysis (GOMl0.04/expt_20.1).
URL https://hycom.org/data/goml0pt04/expt-20pt1

Mezić, I., Loire, S., Fonoberov, V. A., Hogan, P., 2010. A new mixing diagnostic and gulf oil spill movement. Science 330 (6003), 486–489.

Mirzargar, M., Whitaker, R. T., Kirby, R. M., 2014. Curve boxplot: Generalization of boxplot for ensembles of curves. IEEE transactions on visualization and computer graphics 20 (12), 2654–2663.

Oey, L.-Y., Ezer, T., Lee, H.-C., 2005. Loop current, rings and related circulation in the gulf of mexico: A review of numerical models and future challenges. Circulation in the Gulf of Mexico: Observations and models, 31–56.

Paris, C. B., Le Hénaff, M., Aman, Z. M., Subramaniam, A., Helgers, J., Wang, D.-P., Kourafalou, V. H., Srinivasan, A., 2012. Evolution of the macondo well blowout: Simulating the effects of the circulation and synthetic dispersants on the subsea oil transport. Environmental Science & Technology 46 (24), 13293–13302.

Schmitz, W. J., 2005. Cyclones and westward propagation in the shedding of anticyclonic rings from the loop current. Circulation in the Gulf of Mexico: observations and models, 241–261.

Slingo, J., Palmer, T., 2011. Uncertainty in weather and climate prediction. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 369 (1956), 4751–4767.

Sraj, I., Iskandarani, M., Srinivasan, A., Thacker, W. C., Winokur, J., Alexanderian, A., Lee, C.-Y., Chen, S. S., Knio, O. M., Jul. 2013. Bayesian Inference of drag parameters using AXBT data from typhoon Fanapi. Monthly Weather Review 141 (7), 2347–2367.

Thacker, W. C., Srinivasan, A., Iskandarani, M., Knio, O. M., Le Hénaff, M., 2012. Propagating boundary uncertainties using polynomial expansions. Ocean Modelling 43, 52–63.

Toth, Z., Kalnay, E., 1997. Ensemble forecasting at ncep and the breeding method. Monthly Weather Review 125 (12), 3297–3319.

Valentine, D. L., Mezić, I., Maćešić, S., Črnjarić-Žic, N., Ivić, S., Hogan, P. J., Fonoberov, V. A., Loire, S., 2012. Dynamic autoinoculation and the microbial ecology of a deep water hydrocarbon irruption. Proceedings of the National Academy of Sciences 109 (50), 20286–20291.

Walker, N. D., Pilley, C. T., Raghunathan, V. V., D'Sa, E. J., Leben, R. R., Hoffmann, N. G., Brickley, P. J., Coholan, P. D., Sharma, N. N., Graber, H. C., et al., 2011. Impacts of loop current frontal cyclonic eddies and wind forcing on the 2010 gulf of mexico oil spill. Monitoring and modeling the Deepwater Horizon oil spill: a record-breaking enterprise, 103–116.

Wang, S., Iskandarani, M., Srinivasan, A., Thacker, W. C., Winokur, J., Knio, O. M., 2015. Propagation of uncertainty and sensitivity analysis in an integral oil-gas plume model. Journal of Geophysical Research: Oceans.

Whitaker, R. T., Mirzargar, M., Kirby, R. M., 2013. Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. IEEE Transactions on Visualization and Computer Graphics 19 (12), 2713–2722.

Winokur, J., Conrad, P., Sraj, I., Knio, O., Srinivasan, A., Thacker, W. C., Marzouk, Y., Iskandarani, M., 2013. A priori testing of sparse adaptive polynomial chaos expansions using an ocean general circulation model database. Computational Geosciences 17 (6), 899–911.

Xiu, D., Karniadakis, G. E., 2002. The wiener–askey polynomial chaos for stochastic differential equations. SIAM journal on scientific computing 24 (2), 619–644.

Xu, F.-H., Oey, L.-Y., Miyazawa, Y., Hamilton, P., 2013. Hindcasts and forecasts of loop current and eddies in the gulf of mexico using local ensemble transform kalman filter and optimum-interpolation assimilation schemes. Ocean Modelling 69, 22–38.

Zavala-Hidalgo, J., Morey, S. L., O'Brien, J. J., 2003. Cyclonic eddies northeast of the campeche bank from altimetry data. Journal of physical oceanography 33 (3), 623–629.

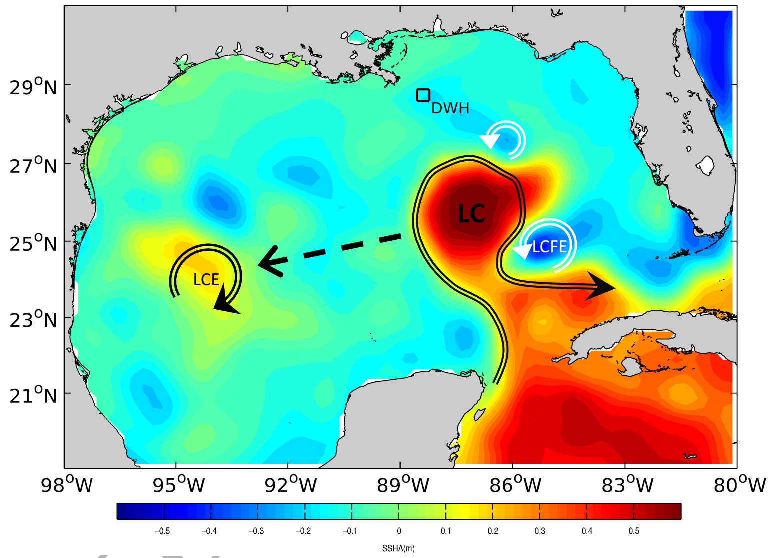

Figure 1: Illustration of LC, LC Eddy indicated by the black arrows and cyclonic LC Frontal Eddy indicated by the white arrows. The background is AVISO SSH Anomaly data at day 20 of the forecast time window. The DWH oil spill location is marked with a black square.

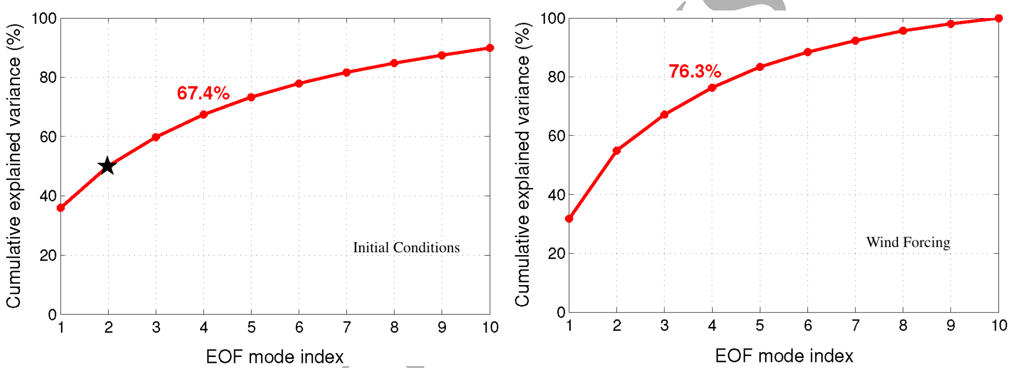

Figure 2: The cumulative variance explained by the first 10 EOF modes. The red labels specify the cumulative variance of the first 4 EOF modes retained in the large ensemble. The black star indicates the cumulative variance associated with the 2 EOF modes of the small ensemble, and which is about 50% of the total variance.

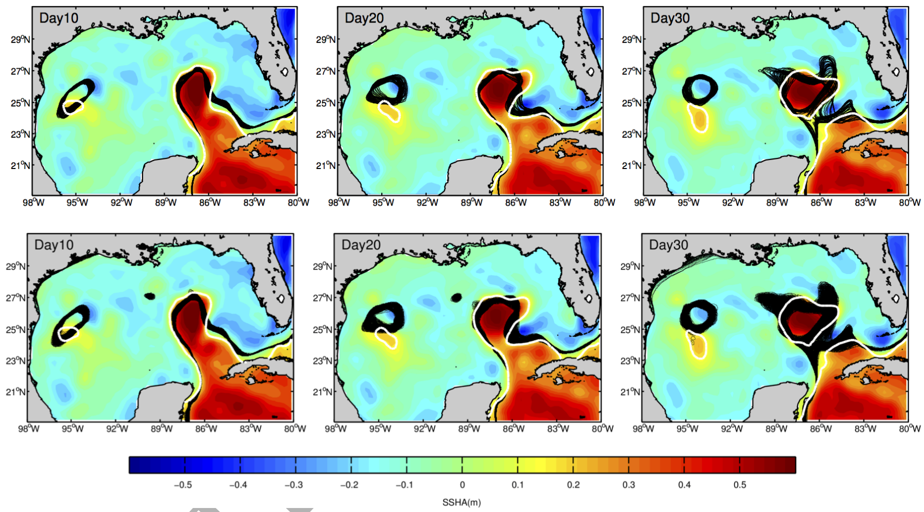

Figure 3: SSH 17 cm contour from HYCOM ensemble (black) and AVISO SSH (white) at indicated forecast time. The background is AVISO SSH Anomaly data. Top: 49-member HYCOM ensemble; Bottom: 798-member HYCOM ensemble.

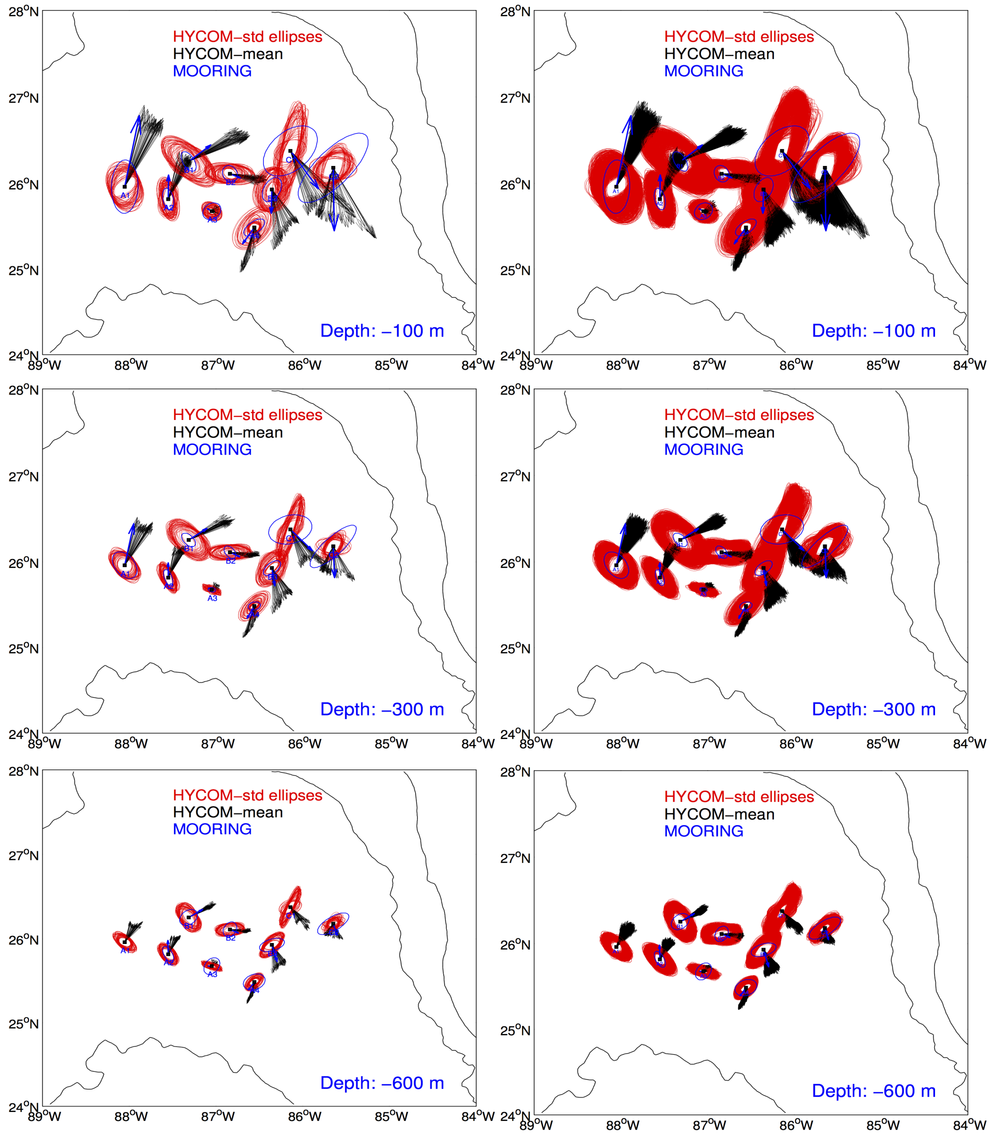

Figure 4: Temporal mean velocity and standard deviation ellipses at different depth built from 9 Bureau of Ocean Energy Management (BOEM)/Science Applications International Corporation (SAIC) full-depth mooring data (blue) and HYCOM ensemble (black and red). The period is from 05/01/2010 to 05/30/2010. Left: 49-member HYCOM ensemble; Right: 798-member HYCOM ensemble.

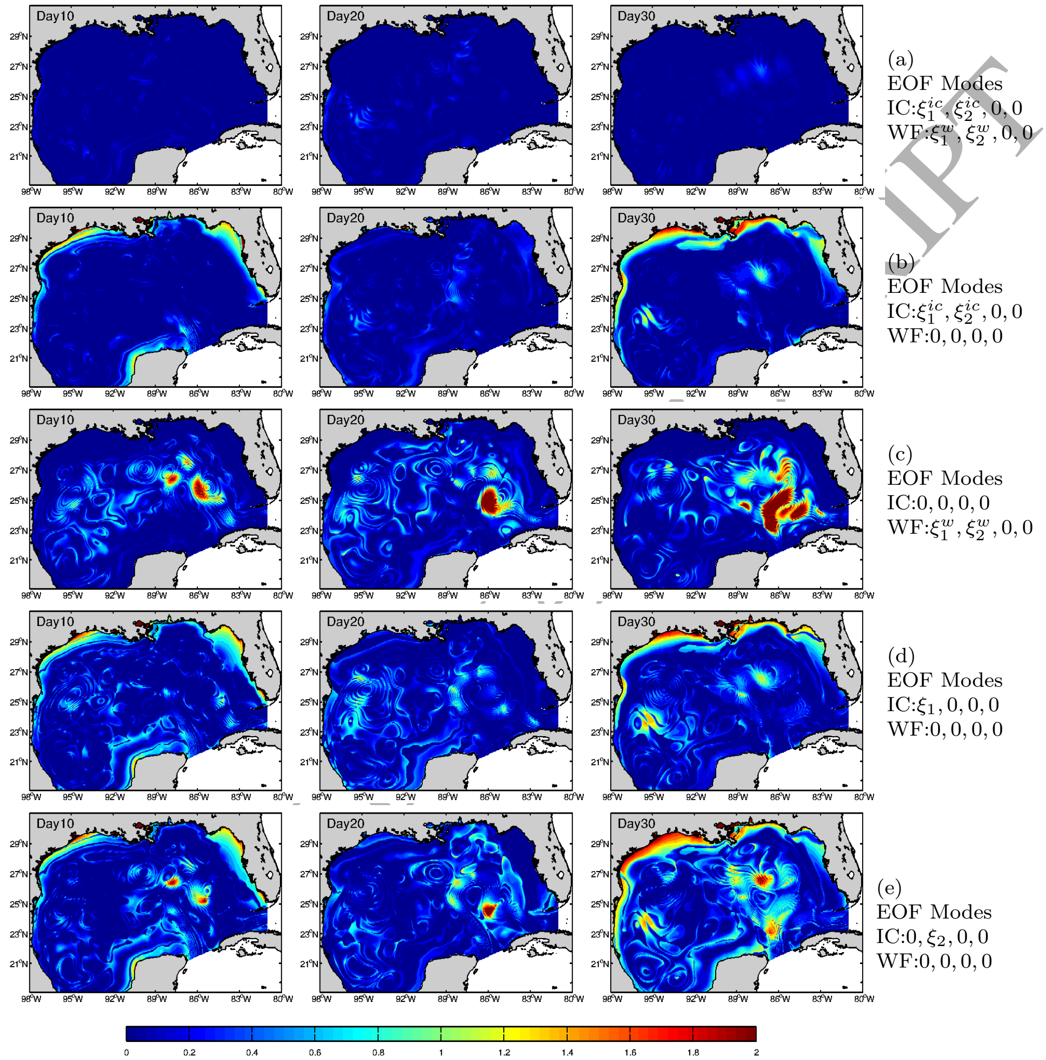

Figure 5: The relative entropy measures the SSH variability loss when variability/uncertainty in the input data is reduced. In all cases shown the reference pdf q refers to perturbing all 8 EOF modes while the pdf p refers to perturbing only a subset of these modes. The unperturbed modes have their amplitudes set to 0 and correspond, from top to bottom, to the 5 scenarios shown in table 2. Areas with no variability loss appear in blue. The abbreviation IC and WF refer to Initial Condition and Wind Forcing modes, respectively.

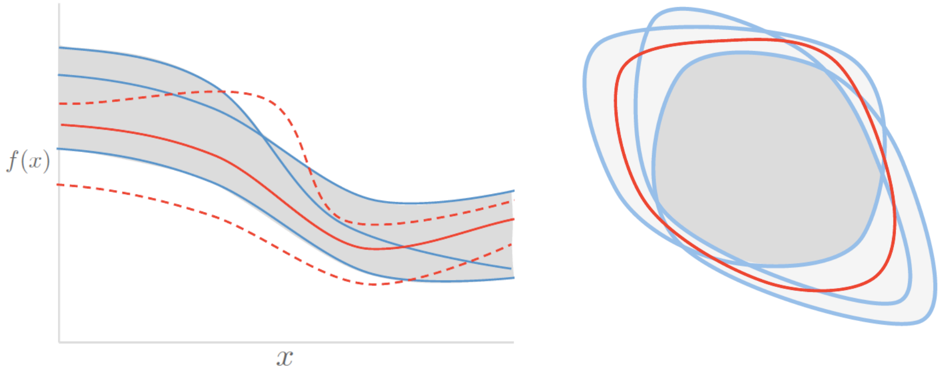

Figure 6: Examples of the generalization of boxplot to curves and contours, adopted from Whitaker et al. (2013). Left: a grey band is formed by three different blue curves and three red curves are test against the band, with only the solid red curve falls completely within the grey band. Right: a red curve falls completely within the light grey band formed by three blue contours.

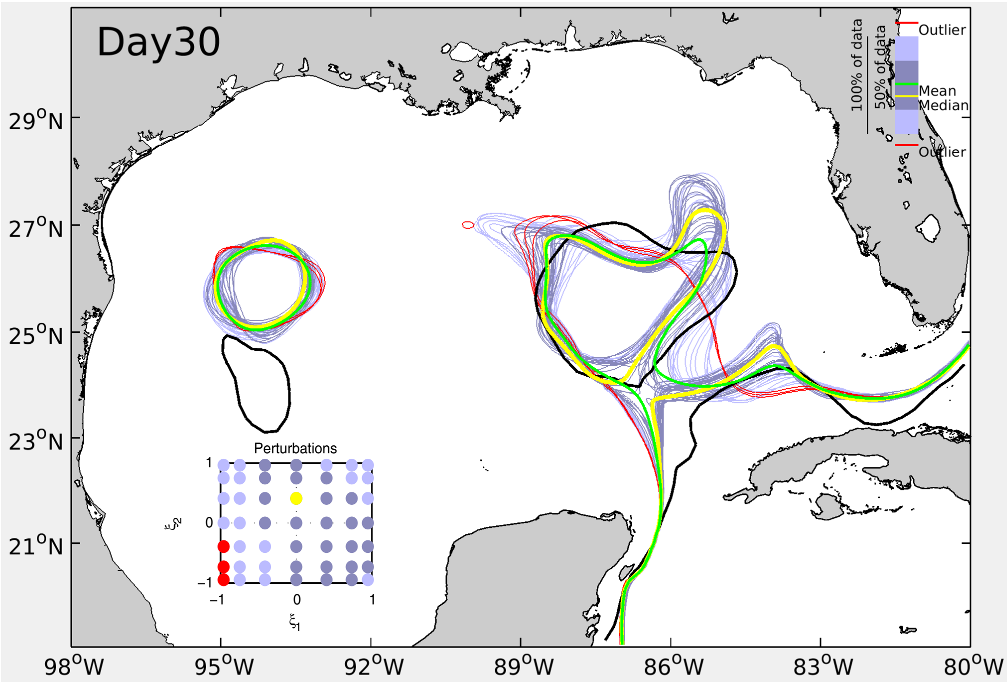

Figure 7: Contour boxplot of the edges of the LC and LC Eddy of the 49-member HYCOM ensemble. All HYCOM ensemble contours are color coded according to the legend. The inserted box shows the location, in the parameter space, of the various simulations of the ensemble with the corresponding color code. The satellite SSH LC edge is shown in black for reference. The edges of the LC and LC Eddy from the mean of the ensemble is in green.

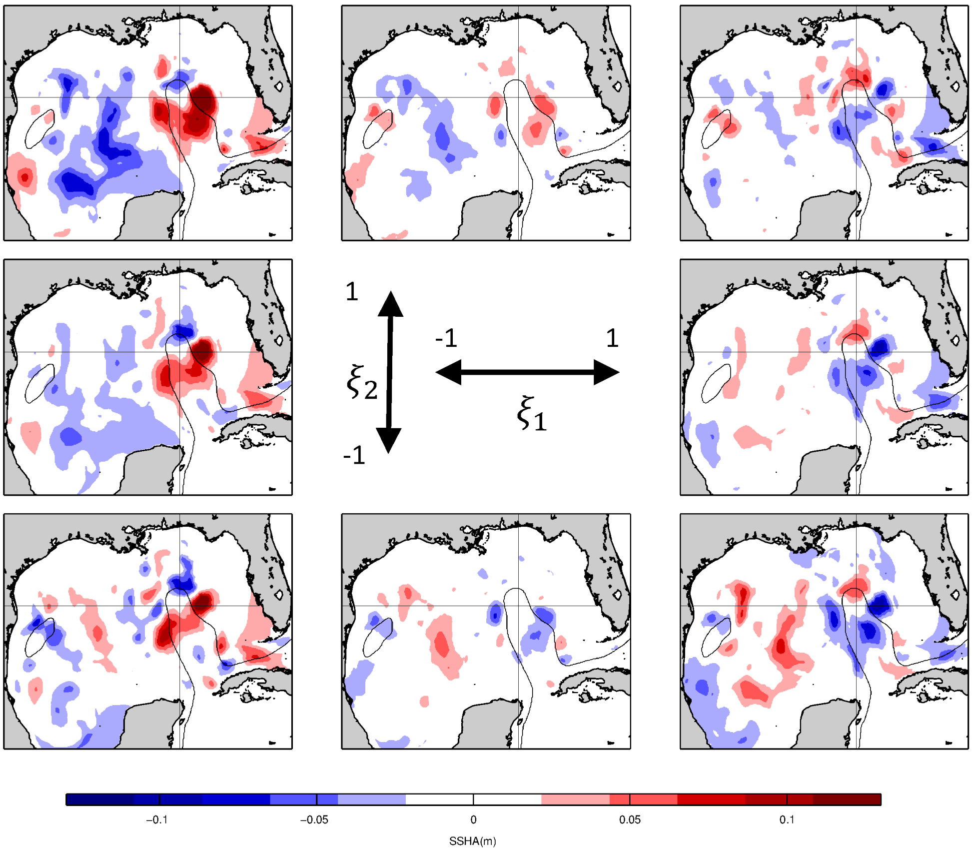

Figure 8: Initial Condition perturbation in SSH according to their normalized random variables (shown in the center). The perturbation is represented by the difference in SSH between the perturbed run and the unperturbed run 1 day after the start of the simulation. The LC contour of the unperturbed simulation is added for reference. The bottom left subfigure represents the most negative perturbations for both EOF modes; the top right subfigure is the most positive perturbations for both EOF modes.

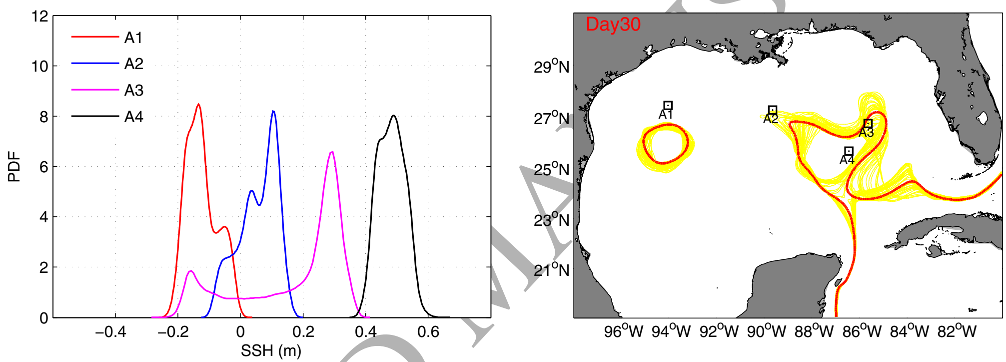

Figure 9: Left: SSH PDF from the ensemble simulations at day 30, sampled at four different locations indicated by the map on the right. Right: The contours on the map shows the LC edge from ensemble simulations represented by the 17cm SSH Anomaly contour. The red contour is the unperturbed simulation.

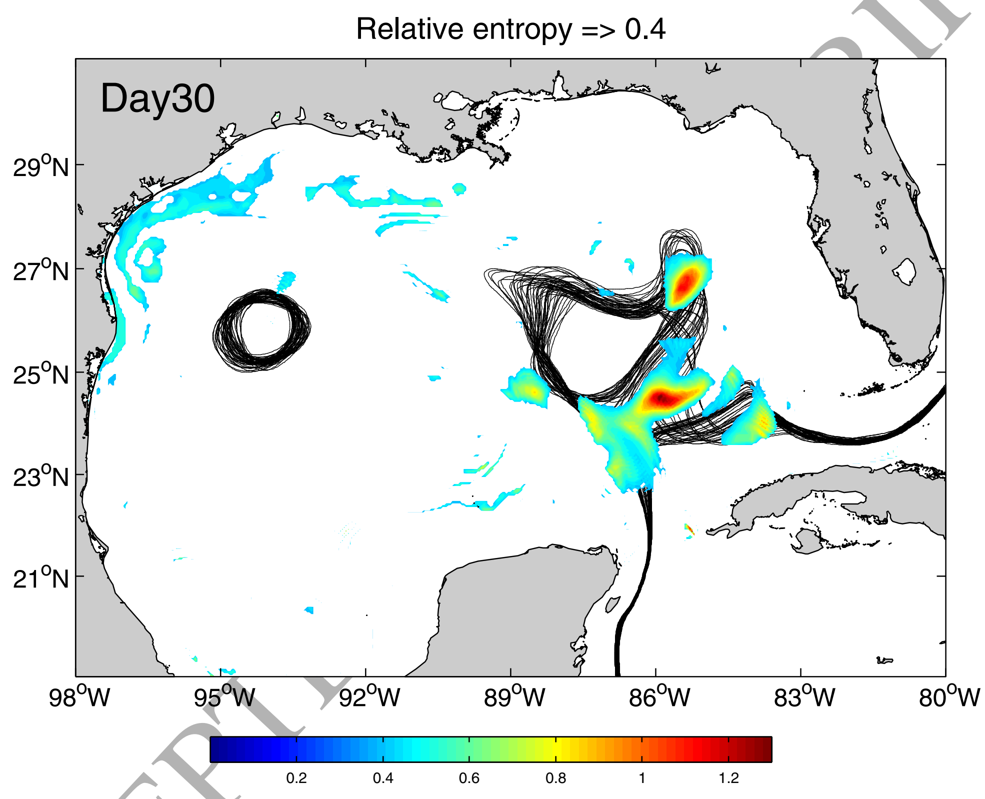

Figure 10: Relative entropy, at day 30, between the ensemble SSH PDF and the Gaussian PDF calculated with the ensemble mean and variance estimated from the PC analysis. The color shows only the region where the relative entropy is greater than 0.4. The black contours are ensemble LC contours for reference.