Propagating boundary uncertainties using polynomial expansions

W. C. Thacker — UM/CIMAS & NOAA/AOML
A. Srinivasan — University of Miami, CCS
M. Iskandarani — University of Miami, MPO
O. M. Knio — Duke University
M. Le Henaff — University of Miami, CIMAS

1. Introduction

This study explores the uncertainty of the circulation within the Gulf of Mexico resulting from the uncertainty of the inflow through the Yucatan Straits. It requires first to characterize the uncertainties of the inflow and then to propagate them dynamically so that they manifest in the circulation at later times. Here, the nature of the uncertainty of the inflow is assumed to be similar to its simulated climatological variability, and polynomial expansions for each simulated variable are used to propagate this inflow uncertainty.

2. Characterizing the uncertainty of the Yucatan inflow

As our high-resolution HYCOM simulations of the Gulf's circulation requires open boundary conditions from a lower-resolution simulation of a larger region, the flow specified at the southern open boundary provide a convenient proxy for the Yucatan inflow. Quantifying its uncertainty requires deciding the nature and likelihood of possible deviations from these boundary conditions.

- Assume possible deviations from default boundary conditions are similar to deviations from a long-term mean of simulated boundary flow.
- Decompose the deviations from long-term mean into spatiotemporal modes using SVD,
- Scale modes by factor α to get boundary variability. ($\alpha = 1$ below.)
- Assume first two modes, which account for 42% of variability are sufficient to model boundary uncertainty.
- Add modes with Gaussian random amplitudes ξ_1 and ξ_2 to default boundary conditions to quantify uncertain boundary conditions.

3. Propagating uncertainty with polynomial expansions

The uncertainty of any simulated output y is dynamically linked to the boundary uncertainties. In principal its uncertainty can be described quantatively by running the model for a large variety of boundary conditions chosen randomly from the distribution of values for ξ_1 and ξ_2. Here we explore a more efficient approach that uses a smaller number of model runs supplemented by inexpensive estimates of y at all values of ξ_1 and ξ_2.

- Assume y can be well approximated by a sum of polynomials:

$$y(\xi_1,\xi_2) \approx \sum_{k_1,k_2}^{k_1+k_2\le K} y_{k_1,k_2} P_{k_1}(\xi_1) P_{k_2}(\xi_2) \ .$$

- Use Hermite polynomials, because they are orthogonal with Gaussian weights.
- Orthogonality provides an expression for the expansion coefficients:

$$y_{k_1,k_2} = \frac{1}{N_{k_1}N_{k_2}} \int\int y(\xi_1,\xi_2) P_{k_1}(\xi_1) p(\xi_1) d\xi_1 P_{k_2}(\xi_2) p(\xi_2) \ d\xi_2 \ .$$

- N_k account for the normalization convention and $p(\xi)$ is the standard Gaussian probability density.

4. Quadrature ensemble

The expansion coefficients for any quantity simulated by HYCOM can be evaluated by Gauss-Hermite quadrature.

$$\int\int y(\xi_1,\xi_2) P_{k_1}(\xi_1) p(\xi_1) d\xi_1 P_{k_2}(\xi_2) p(\xi_2) d\xi_2 \approx \sum_{q_1}\sum_{q_2} y(\xi_{q_1},\xi_{q_2}) P_{k_1}(\xi_{q_1}) w_{q_1} P_{k_2}(\xi_{q_2}) w_{q_2} \ .$$

- (ξ_{q_1},ξ_{q_2}) are quadrature points where values of y must be known.
- w_{q_1} and w_{q_2} are corresponding weights.
- The principal cost of evaluating the coefficients is that of the HYCOM simulations needed to get values y at the quadrature points.

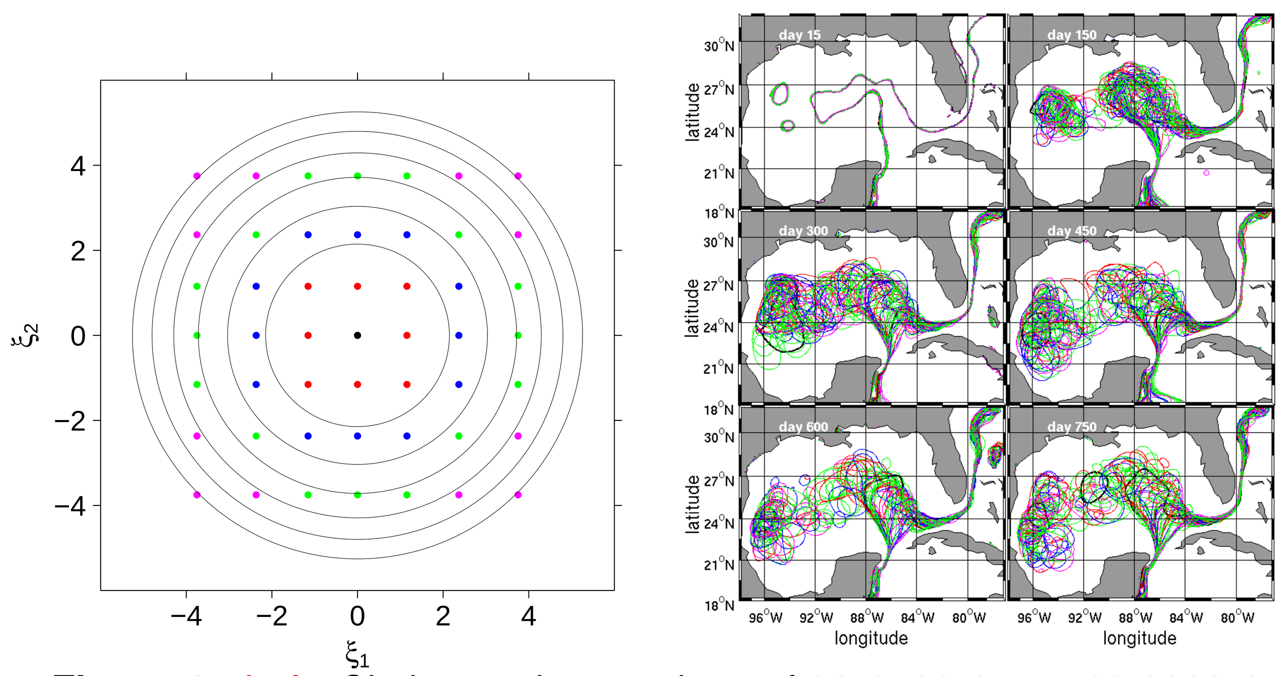

Figure 1: *Left: Circles enclose regions of 90%, 99%, ..., 99.9999% probability. Dots mark locations of 49 Gauss-Hermite quadrature points, with red dots corresponding to relatively likely, blue less likely, green unlikely, and magenta highly unlikely boundary conditions. Right: Locations of the Loop Current and its eddies from 49 HYCOM quadrature runs as indicated by 17 cm sea-surface-height contours. The panels, from upper left to lower right, show the contours at 15, 150, 300, 450, 600, and 750 days after the boundary uncertainties were initiated. The colors of the contours correspond to the colors of the dots .*

5. Uncertainty of the surface elevation field

After the ensemble of quadrature simulations has been run, it is simple to evaluate the uncertainty of any variables that have been saved. In particular, the orthogonality of the polynomials simplifies expectation integrals:

- The mean of y is given by the constant term $y_{0,0}$.
- The variance is given by the sum of squares of the other coefficients.
- The covariance with another variable z is given by the sum of products of their coefficients.

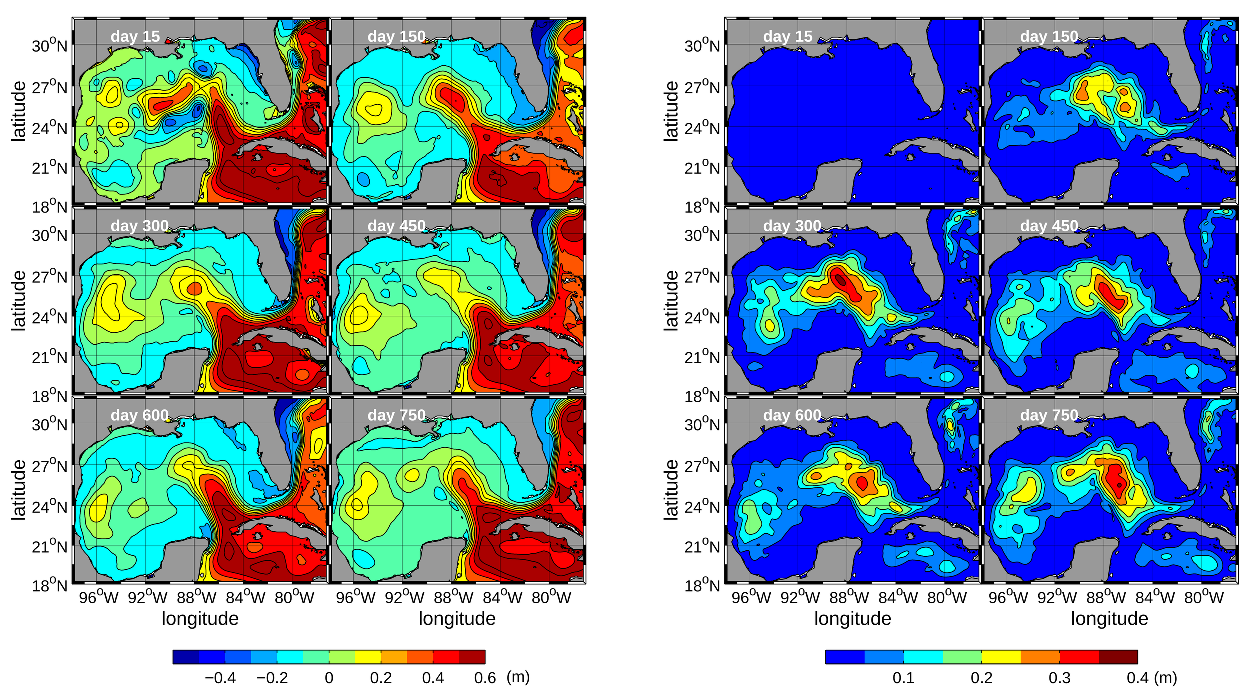

Figure 2: *Left: Mean (m) of the sea-surface-height field from the polynomial chaos expansion at 15, 150, 300, 450, 600, and 750 days after the boundary uncertainties were initiated. Right: Standard deviation (m) of the sea-surface-height field from the polynomial chaos expansion at same days.*

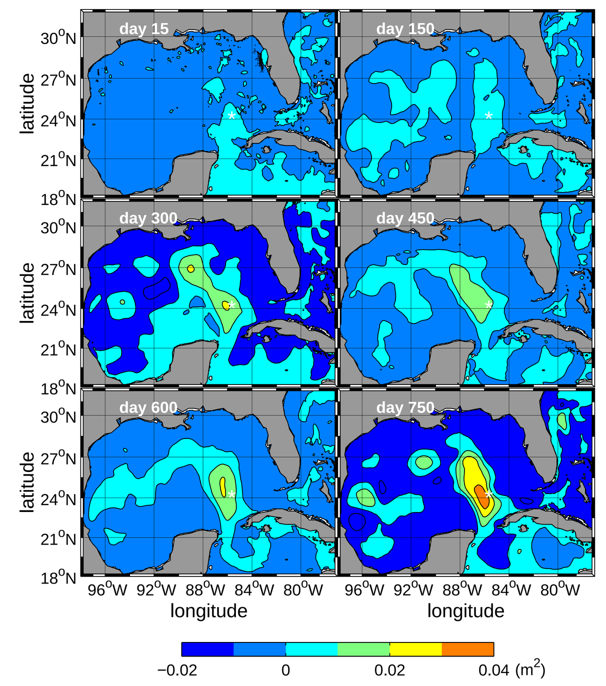

Figure 3: *Estimates of covariance (m^2) of the surface elevation for each grid cell with the surface elevation for the cell at the point (86°E, 24.1°N) marked by the white star.*

6. Convergence of the polynomial expansion

As variance and standard deviation increase as more terms are retained in the polynomial expansion, convergence can be examined by comparing different levels of truncation. Figure 4 shows that going from 5th to 6th degree polynomials contributes significantly less to the standard deviation of surface elevation than does going from 4th to 5th.

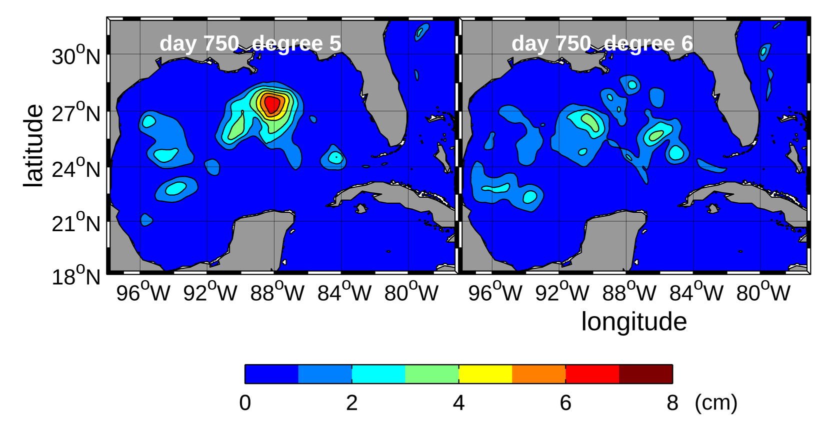

Figure 4: *Incremental contribution to standard deviation (cm) of surface elevation at day 750. Left: contribution of the 6 5th-degree terms relative to the total contributed by the 21 terms of degree less than 6. Right: contribution of the 7 6th-degree terms relative to the total contributed by the 28 terms of degree less than 7.*

7. Emulation

The polynomial expansions provide inexpensive alternatives to the simulations for boundary inflows that were not in the quadrature ensemble. For any value of the amplitudes ξ_1 and ξ_2, there is a polynomial approximation for y. This approximation can be called *emulation* in contrast to values for expensive model simulations.

- The emulated values for y can be regarded as lying on a polynomial response surface.
- The more likely values of ξ_1 and ξ_2 determine the part of the surface with the more likey values for y.
- Because the dynamics are nonlinear, the distribution of y is not Gaussian.
- Sampling ξ_1 and ξ_2 to generate a histogram reveals its non-Gaussian nature.

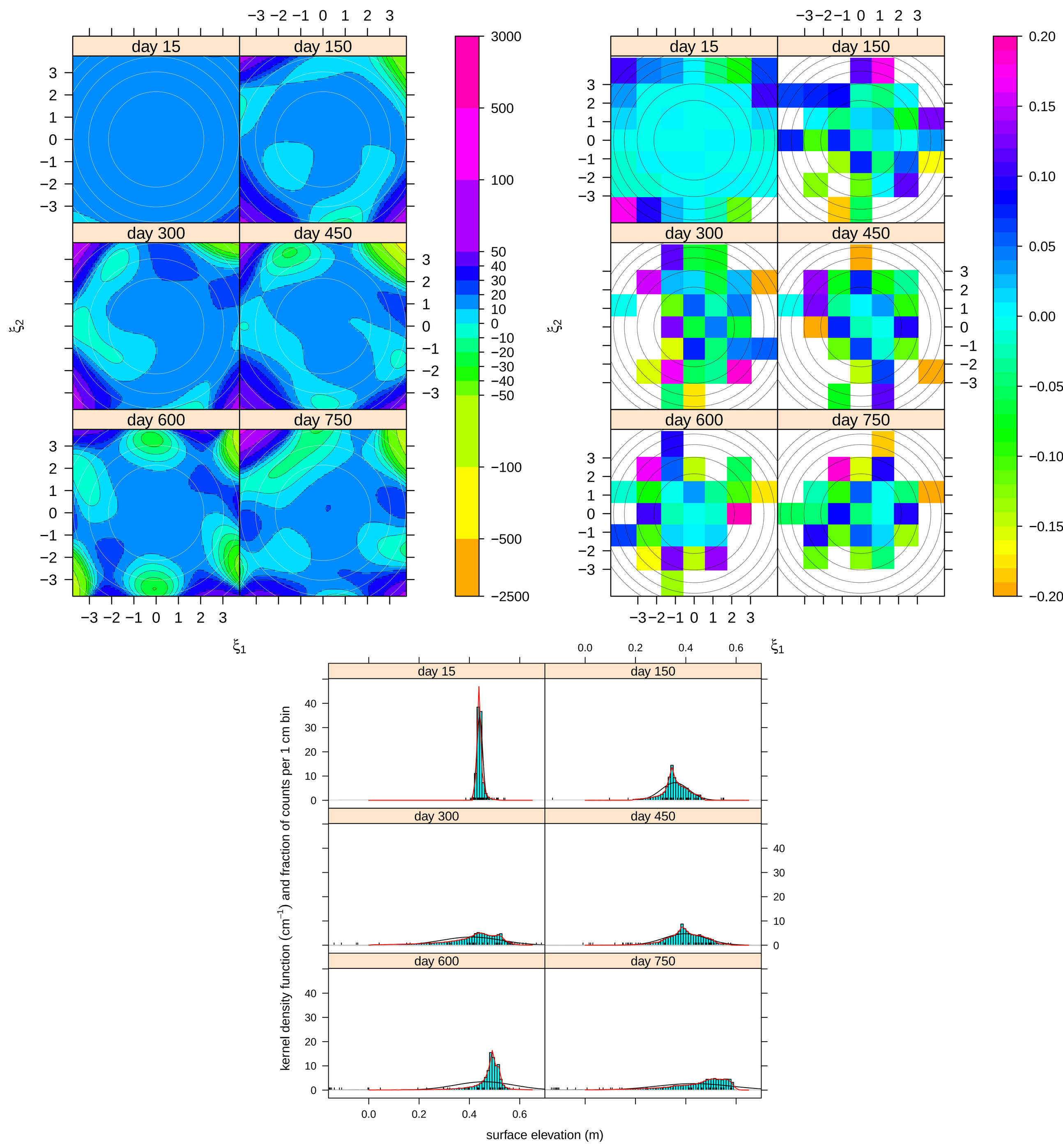

Figure 5: *Left: Surface elevation (cm) at (86°E,24.1°N) as a function of random variables ξ_1 and ξ_2. Note the compressed color scale used to distinguish more likely from less likely responses: Contours are at 10 cm intervals from -50 to 50 cm, and more extreme values are represented with a compressed scale. The circles indicate that the extreme values are highly unlikely. Right: Errors (m) of the polynomial chaos expansion for sea-surface-height at (86°E,24.1°N). The color of each rectangle indicates the difference between the HYCOM simulation and its approximation by the polynomial chaos expansion at each quadrature point, with white indicating errors larger than 20 cm. Bottom: Kernel density estimates for surface elevation (m) at the point (86°E, 24.1°N) derived from histograms generated using polynomial chaos expansion corresponding to 50,000 random boundary conditions. Ticks along the bottom indicate values for the 49 HYCOM simulations. Red curves are kernel density estimates, and black curves are Gaussian densities with means and standard deviations from the polynomial expansions.*

8. Conclusion

Polynomial expansion provide a relatively inexpensive way to explore the consequences of uncertainties in a model's inputs. The approach taken here illustrates the need for quantifying the uncertainties of the inputs. In particular, because the expense increases geometrically with the number of uncertain inputs to be examined, it is important to focus on only the few that are the most important. In this regard, the method has similarities to Kalman filtering as applied to oceanographic and meteorological models. And like Kalman filtering, polynomial expansions offer the possibility of updating prior estimates of input uncertainties by exploiting observational data, although that aspects has not been explored here.

Carlisle.Thacker@noaa.gov