AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2015JC011311

Special Section:
Physical Processes Responsible for Material Transport in the Gulf of Mexico for Oil Spill Applications

A framework to quantify uncertainty in simulations of oil transport in the ocean

Rafael C. Gonçalves[1], Mohamed Iskandarani[1], Ashwanth Srinivasan[2], W. Carlisle Thacker[3], Eric Chassignet[4], and Omar M. Knio[5,6]

[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, USA, [2]Tendral LLC, Miami, Florida, USA, [3]independent scholar, [4]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida, USA, [5]Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina, USA, [6]King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

Abstract An uncertainty quantification framework is developed for the DeepC Oil Model based on a nonintrusive polynomial chaos method. This allows the model's output to be presented in a probabilistic framework so that the model's predictions reflect the uncertainty in the model's input data. The new capability is illustrated by simulating the far-field dispersal of oil in a Deepwater Horizon blowout scenario. The uncertain input consisted of ocean current and oil droplet size data and the main model output analyzed is the ensuing oil concentration in the Gulf of Mexico. A 1331 member ensemble was used to construct a surrogate for the model which was then mined for statistical information. The mean and standard deviations in the oil concentration were calculated for up to 30 days, and the total contribution of each input parameter to the model's uncertainty was quantified at different depths. Also, probability density functions of oil concentration were constructed by sampling the surrogate and used to elaborate probabilistic hazard maps of oil impact. The performance of the surrogate was constantly monitored in order to demarcate the space-time zones where its estimates are reliable.

Key Points:
- Uncertainties in currents and oil buoyancy were characterized using three random input variables
- The dominant contributor to oil concentration uncertainties is identified via variance analysis
- Probabilistic hazard maps of oil impact are constructed based on the input uncertainty

Supporting Information:
- Supporting Information S1

Correspondence to:
R. Gonçalves,
rgoncalves@rsmas.miami.edu

Citation:
Gonçalves, R. C., M. Iskandarani, A. Srinivasan, W. C. Thacker, E. Chassignet, and O. M. Knio (2016), A framework to quantify uncertainty in simulations of oil transport in the ocean, *J. Geophys. Res. Oceans*, *121*, 2058–2077, doi:10.1002/2015JC011311.

Received 14 SEP 2015
Accepted 25 FEB 2016
Accepted article online 1 MAR 2016
Published online 1 APR 2016

1. Introduction

Oil spilled in the ocean poses a threat to the environment as its consequences might be catastrophic to the biota and to the communities of the impacted areas. Oil fate models are important instruments that can provide the scientific background needed for oil spill risk assessment and for decision-making during an oil spill incident. In an emergency response, forecasting the pathways and the oil fate is essential for awareness, preparation, and mitigation actions of the probable impacts.

Several oil fate models of increasing skill have been developed over the decades to represent the complex behavior of the oil in the ocean. They evolved from simple transport models, with limited fate algorithms, to current state-of-the-art models, which simulate the transport of the oil in three dimensions, including processes of advection, diffusion, stranding, emulsification, and oil weathering [*Cekirge et al.*, 1995; *Chao et al.*, 2003; *Berry et al.*, 2012]. In general, oil fate models treat the spill as a collection of Lagrangian particles that are advected by an Eulerian velocity field. This velocity field, which is set as an input for the oil model, includes the effect of ocean currents, waves, and winds, and is usually provided by the outputs of numerical models.

Regardless of how sophisticated the oil fate models are, and the amount of processes they can represent, errors in the input data will inevitably lead the results to diverge from observations [*Sebastião and Guedes Soares*, 2007]. Oil fate predictions have to rely on data from oceanic and atmospheric forecasts, which have limited predictability and inherent uncertainty. In many cases, little is known about the incident itself, and important information, that is essential for setting up the initial conditions and other input parameters, is missing. Information like the exact location where the spill occurred, the amount of oil that was spilled, the flow rate of the oil, in case of a blowout, will likely influence the outcome of an oil fate simulation. With that in mind, the application of uncertainty quantification (UQ) techniques in oil fate forecasts is essential to provide meaningful information for the spill response efforts [*Galt*, 1997a, 1997b].

The goal of UQ is to estimate uncertainty in the model's output given uncertainties in its inputs. It provides means of estimating the most likely output, along with confidence levels and error estimates. Although it

should be the next step for the advancement of operational oil spill forecast [*Hodges et al*., 2015], applications of UQ in oil fate models are still scarce in the literature. The current article presents a framework for quantifying uncertainty in the transport of oil during oil fate simulations based on a polynomial chaos (PC) approach. The method is based on constructing a computationally cheap surrogate for the model, where an output is defined as a polynomial expansion of the uncertain input parameters. The surrogate is built through an ensemble of model realizations, and once constructed, it should be able to approximate the model's output for any given value of the input parameters inside a preestablished range. As obtaining outputs from the surrogate is computationally cheap, the task of sampling the model hundreds of thousands of times is feasible. With that, probability density functions (PDF) of the model's output can be estimated and used to obtain relevant statistical information, as it would be done in a Monte-Carlo approach. The ability of estimating PDFs at low computational cost is particularly important for emergency response efforts, when decisions have to be made fast. The PC approach also provides an easy and fast way to quantify the contribution of each uncertain input parameter in the output uncertainty (sensitivity analysis). This property has applications in oil spill risk assessments [*Neves et al*., 2015].

The capabilities of the PC approach are explored in a simulation of the Deepwater Horizon (DWH) blowout. In this accident, about 4.9 million barrels of crude oil along with a large amount of gas were released in the Gulf of Mexico [*McNutt et al*., 2011] following the explosion of a drilling platform on 20 April 2010. The spill lasted for 87 days, with oil being continuously released in the ocean at a depth of 1500 m. The fact that the accident occurred in deep water increases drastically the complexity of the problem of simulating the oil spill. High pressure along with other ambient variables, such as temperature, salinity, and ocean currents, affect the upward movement of the plume formed by oil and gas [*Johansen*, 2003; *Yapa et al*., 2012]. The dissolution of gas and oil components along with the entrainment of seawater causes the plume to lose buoyancy and momentum; eventually, the plume reaches a neutrally buoyant level. At this level, oil droplets and gas bubbles separate from the plume, that now is composed of dissolved gas and oil mixed with seawater, and start moving upward [*Socolofsky and Adams*, 2002]. The initial plume stage of a deep water blowout is referred to as the near field, and is usually simulated through integral models [*Socolofsky et al*., 2011]. The stage when oil droplets separate from the plume and move upward, while being advected by the ocean currents, is the far field, which is simulated through Lagrangian transport models mentioned above [*Mariano et al*., 2011; *North et al*., 2011].

Here the newly developed DeepC Oil Model (DCOM) is applied to simulate the far field of the DWH spill. Two sources of uncertainty are considered. The advecting ocean currents, which are provided by the outputs of a general ocean circulation model (OGCM), and the oil droplet size distribution, which affects the buoyancy of the spilled oil. The forward propagation of uncertainty is carried out in a hindcast simulation of the first 30 days of the oil spill. We present a framework to estimate the uncertainty in the simulated oil distribution, to quantify the contribution of each uncertain input parameter considered and to construct probabilistic hazard maps of oil impact. Also, we test the skill of the PC surrogate in representing the model's output by comparing it to a set of actual realizations of the model.

The paper is organized as follows. Section 2 describes the models used and the setup of the simulations. Section 3 discusses the sources of uncertainties and the choice of the input parameters. Section 4 introduces the PC methodology. Section 5 displays the results, which include the mean and standard deviation of the oil distribution, sensitivity analysis, and probabilistic hazard maps. Summary and discussion are presented in section 6.

2. Model Description

2.1. Lagrangian Oil Fate Model

The DeepC Oil model (DCOM) is a 3-D advection-diffusion and weathering code that simulates the transport and evolution of oil accidentally released from a subsea well, pipeline, or a surface source such as a leaking ship. The algorithms for transport and transformations are largely derived from the works of *Proctor et al*. [1994], *Al-Rabeh et al*. [2000], *Spaulding et al*. [2000], *Zheng and Yapa* [2000], *Lehr et al*. [2002], *Zheng et al*. [2003], *Tkalich et al*. [2003], *Korotenko et al*. [2010], and *Socolofsky et al*. [2011]. The intended functionality of the model is to predict where, when, and in what state the oil is most likely to be found based on information about the ocean currents, winds, waves, oil properties, and other environmental conditions. The model

is based on a Lagrangian formulation where the oil spill is represented by a large number of particles with time-varying position, mass, volume, and composition. The model takes as input the release conditions, initial position, and properties of the oil along with time-varying environmental conditions including winds, waves, water currents, and water temperature and salinity. The position and composition of the particles are then updated based on environmental conditions. The present article focuses solely on the evolution of the oil caused by transport, the effect of other processes have been turned off. We thus focus on describing in some details the transport processes including the buoyancy calculation.

The positions of the particles ($\boldsymbol{x}$) are evolved in time (t) according to the following equation

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{u}_a + \boldsymbol{u}_s + \boldsymbol{k}w_b, \tag{1}$$

Here $\boldsymbol{u}_a$ represents the sum of advective processes, namely ocean currents, Stokes drift, and wind drag. These velocities must be obtained from operational forecasting systems and account for the transport due to dynamical processes that can be resolved on the forecasting systems computational grids. The unresolved processes are represented by $\boldsymbol{u}_s$ and account for subgrid-scale turbulence; DCOM uses a random walk model [*Griffa*, 1996] which adds a random displacement to the particles at each time step. The last term on the right-hand side of equation (1) is the buoyant velocity of the oil droplets, w_b. The buoyancy of the droplets is a function of their size, and w_b is computed following a two equations approach. For droplet sizes below a critical diameter (d_c), so that the Reynolds number is low, w_b is given by Stokes law. For diameters greater than d_c (higher Reynolds number), w_b follows a formulation proposed by *Zheng and Yapa* [2000]. In DCOM, equation (1) is solved by a fourth-order Runge-Kutta scheme.

DCOM uses a probabilistic approach to represent the range of droplet sizes found in a deep water oil spill: each model particle is assigned an oil droplet diameter sampled from a distribution. DCOM uses the Rosin-Rammler distribution, as that was found to be a good fit based on different laboratory and field experiments [*Yapa and Chen*, 2004; *Brandvik et al*., 2013; *Johansen et al*., 2013; *Aman et al*., 2015]. Its cumulative probability function, $P(d)$, is

$$P(d) = 1 - \exp\left[\ln\left(\frac{1}{2}\right)\left(\frac{d}{d_{50}}\right)^{\alpha}\right], \tag{2}$$

where d is the droplet size, d_{50} is the fiftieth percentile of the distribution, and α is the spreading parameter.

2.2. Hydrodynamic Model

To force DCOM, we used ocean currents, temperature, and salinity data from a Gulf of Mexico simulation performed with the Hybrid Coordinate Ocean Model (HYCOM, http://hycom.org). HYCOM is a generalized vertical coordinate model, which is widely used for research and operational tasks [*Bleck*, 2002; *Chassignet et al*., 2003, 2006, 2009]. Typically, the vertical coordinate is isopycnic in the stratified interior, z-level or pressure coordinates in the mixed layer, and terrain following in coastal regions. Here outputs of a data assimilative 2003–2012 hindcast were used [*Chassignet and Srinivasan*, 2015]. The hindcast is based on the 1/25° GoM-HYCOM system, which is used in near real time by the Naval Research Lab (http://www7320.nrlssc.navy.mil/hycomGOM/). The model is configured with 20 vertical layers, and its domain extends from approximately 98°W to 76.5°W and from 18.9°N to 31.5°N. It includes a realistic bottom topography, derived from a recent high-resolution bathymetry available from Florida State University. The hindcast is forced at the surface with the North American Regional Reanalysis (NARR) atmospheric product [*Mesinger et al*., 2006]. Surface fluxes of heat and momentum are calculated through bulk formula using fields of wind stress, 10 m wind speed, 2 m air temperature, 2 m water vapor mixing ratio, precipitation, and incoming solar radiation. A correction to the NARR-derived radiation fluxes was derived by comparing model surface temperature with Advanced Very High Resolution Radiometer (AVHRR) estimates of surface temperature fields. The NARR wind speed was corrected by regression to QuikSCAT wind speeds. Monthly climatological Gulf of Mexico river data from USGS database (http://waterdata.usgs.gov/nwis/sw) and the RIVDIS database (http://www.RivDis.sr.unh.edu) were specified for the freshwater discharge (treated as a virtual salt flux). The hindcast simulation uses biweekly climatological lateral open boundary conditions derived from a North Atlantic climatological run.

The data assimilation is implemented with the Tendral-Statistical Interpolation System [*Halliwell et al.*, 2014]. The primary data set for constraining the model is daily along-track sea level anomalies. The data were obtained from multiple operational satellite altimeters, available from Collecte Localisation Satellites in a delayed time quality controlled mode. Daily satellite sea surface temperature, obtained from the National Oceanic and Atmospheric Administration (NOAA), in situ profiles of temperature and salinity from the World Ocean Atlas 2009, and data from cruises during the DWH incident were also used. For an evaluation of the hindcast, the reader is referred to *Chassignet and Srinivasan* [2015].

2.3. Model Setup

The numerical experiments were set to represent the first 30 days of the DWH oil spill. Despite the fact that the blowout occurred at 1500 m depth, studies have indicated that the transition between the near field and far field in the DWH spill occurred around 1200 m [*Socolofsky et al.*, 2011; *North et al.*, 2011]. So, a thousand particles were released every 30 min from a source located at 88.3°W and 28.7°N at a depth of 1200 m. The velocities used to transport the particles ($\boldsymbol{u}_a$ in equation (1)) were given by the outputs of the HYCOM hindcast, from 20 April 2010 to 19 May 2010, and displacements of the particles were computed with a time step of 30 min. The oil density was set to 858 kg/m^3, while the seawater density was computed using temperature and salinity from the HYCOM hindcast. In our experiments, only ocean currents were considered.

We define our quantity of interest as the oil concentration at some specific location after a period of time. Despite the difference in droplet sizes between the particles, it was considered that each particle carried the same amount of oil. This way, the simulated oil spill had the same flow rate for the different realizations of the model, as the number of particles released in each realization was the same. The oil concentration was computed by binning the particles in a grid, where the number of particles at each grid cell was multiplied by a constant oil mass, and divided by the volume of the cell. Oil concentrations were obtained for three different depths: between 0 and 5, 300 and 600, and 900 and 1200 m depth, which will be referred to as upper layer, middepth layer, and deep layer, respectively. In all of these layers, the particles were binned in a grid with 0.1° resolution in both zonal and meridional directions.

3. Input Parameters Uncertainty

The choice of the input parameters to be perturbed has to fall on those whose uncertainty is more likely to have a higher influence on the quantity of interest. On the other hand, the number of parameters to be perturbed is limited by the "curse of dimensionality," i.e., the number of simulations required to build the PC surrogate grows exponentially with the number of uncertain parameters considered (or dimensions in the uncertain space). Here we are interested in the sources of uncertainty that controls the transport of oil in a deep water blowout. For that matter, we focused solely on the oil droplet size distribution, which controls the droplets rise velocity, and the advecting currents, which is the main factor responsible for the horizontal spread of the oil throughout the water column. Other important sources of uncertainty, which would not be considered in this study, include the subgrid-scale parametrization, wind drag at the surface, Stokes drift, and the oil degradation.

A key aspect of UQ is the choice of the range that an input parameter is to be perturbed. We would like this range to be wide enough to account for all possible values of the parameter. On the other hand, the response of the quantity of interest to the perturbations inside the chosen range should be fairly smooth, so that the PC expansion can approximate the behavior of the model. If the range of the input values is too broad, the response of the model might be more complicated and would require a longer series to represent accurately. Also, the parameter space might present different regimes, with each regime being characterized by a local uncertainty range. If such local ranges are known, they should be accounted for separately.

In the PC approach, it is convenient to describe the input uncertainty by standardized random variables $|\xi| \leq 1$. Considering the vector of uncertain input parameters $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_d,\ldots,\theta_D)^\top$, where D is the number of uncertain inputs, each parameter is related to a standardized random variable by:

$$\theta_d = \frac{\theta_d^{\max}+\theta_d^{\min}}{2} + \xi_d \frac{\theta_d^{\max}-\theta_d^{\min}}{2}, \tag{3}$$

where $\theta_d^{\min} \leq \theta_d \leq \theta_d^{\max}$ is the range of the dth uncertain input. An important and difficult step in UQ is the definition of the PDFs of the uncertain inputs. Due to the lack of information about the input parameters

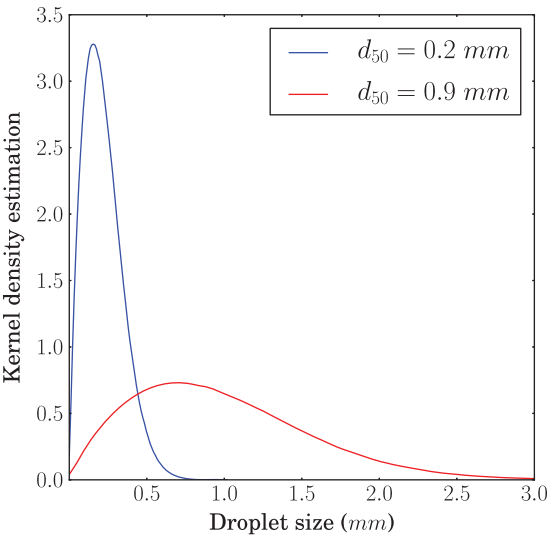

Figure 1. Rosin-Rammler distributions considering $d_{50} = 200\ \mu$m (blue) and $d_{50} = 900\ \mu$m (red). For both distributions, the spreading parameter is $\alpha = 1.8$.

considered, we defined the PDFs of all input parameters with an uniform distribution. A description of the input parameters and their perturbation range is presented in the following subsections.

3.1. Droplet Size Distribution Uncertainty

In situ measurements of the droplet size distribution during accidents like the DWH blowout are usually difficult and highly uncertain. Several conditions influence the size of the droplets, as they break up/coalesce into smaller/bigger droplets; these include the ambient state (pressure, temperature, ocean currents, etc.), the intensity of the flux at the source, the oil viscosity, and oil-water surface tension [*Yapa et al.*, 2012; *Zhao et al.*, 2014]. The use of dispersants as a measure to prevent the oil from reaching the surface further increases the uncertainty of the droplet size distribution, as they reduce the size of the droplets by decreasing the oil-water interfacial tension.

The Rosin-Rammler distribution, used here to represent the droplet size distribution, is defined by α and d_{50} (equation (2)), which control its shape and range. Due to the high uncertainty of the droplet size distribution and its importance in determining the vertical velocity of the oil, both parameters would be good candidates to be perturbed. In a preliminary experiment, not presented here, both parameters were considered uncertain; d_{50}, however, turned out to have a much bigger impact on the oil fate than α. Henceforth, only d_{50} will be perturbed here in order to reduce the dimensionality of the uncertain space, while α is set to 1.8.

To define a range of values for d_{50}, we considered that the oil plume was under the influence of dispersants, as about 2900000 L. of dispersants were released next to the wellhead during the DWH spill [*Kujawinski et al.*, 2011]. Our choice for the range of d_{50} was based on the results from *Zhao et al.* [2014], where a droplet formation model was applied to estimate the droplet size distribution in the DWH blowout considering different degrees of reduction on the oil-water interfacial tension. For a one thousandfold reduction, their model produced a distribution with $d_{50} = 260\ \mu$m, while for a tenfold reduction, the droplet median diameter was $d_{50} = 890\ \mu$m. We defined the range of d_{50} to be between 200 and 900 μm. Figure 1 shows the droplet size distributions in the extreme cases, when $d_{50} = 200\ \mu$m and $d_{50} = 900\ \mu$m.

3.2. Eulerian Velocity Field

There are several sources of uncertainty that are present in the modeled HYCOM Eulerian velocity fields. In addition to truncation errors associated with the model numerics and grid choices, the solution is quite sensitive to choices in subgrid-scale parameterizations and the specification of initial and boundary conditions. Quantifying the effects of all sources would require thousands of HYCOM realizations with different combinations of values for each parameter and for each collection of initial and boundary conditions. Unfortunately, that task is impractical due to the overwhelming computational cost. Instead, we would like to be able to propagate uncertainty in the velocity field using a reduced number of parameters. Moreover, we would like to perturb the velocity field in every grid cell without running extra realizations of HYCOM, using only the outputs of a single run, and with the perturbations evolving in time. Our approach to this problem was to characterize the perturbations as a sum of products of spatial patterns and time series using empirical orthogonal function (EOF) decomposition. The spatial patterns are given by the multivariate EOFs of the zonal and meridional components of the velocity field, and the time series are given by the respective principal components. A similar approach was applied by *Thacker et al.* [2012] to propagate uncertainty through boundary conditions in a HYCOM simulation of the Gulf of Mexico circulation.

The perturbations on the Eulerian velocity field were carried out in the following manner:

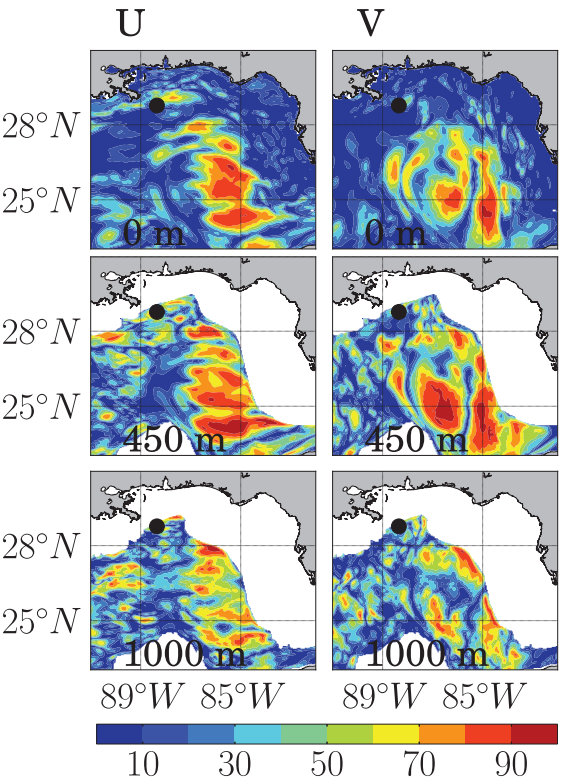

Figure 2. Percentage of the total variance explained by the first and second EOFs combined for the (left) zonal (U) and (right) meridional (V) velocity components, in three selected depths. The black dots depict the position of the Deepwater Horizon wellhead.

$$\boldsymbol{u_a} = \boldsymbol{u_0} + \sum_{k=1}^{K} \xi_k \boldsymbol{u_k}, \tag{4}$$

where $\boldsymbol{u_0}$ is the unperturbed horizontal velocity field, $\boldsymbol{u_k}$ is the kth perturbation of the horizontal velocity field, ξ_k is the standardized random variable that controls the kth perturbation, and K is the number of modes used. The perturbations are given by:

$$\boldsymbol{u_k} = N_T c_k r_k, \tag{5}$$

where c_k is the kth EOF, r_k is its principal component, and N_T is a normalization factor, which here is the number of time steps considered in the velocity field.

The goal of this approach is to generate a distribution of velocity fields that capture the uncertainty in the HYCOM velocity field $\boldsymbol{u}_0$ used in the DeepC Oil model. This task is similar to that of characterizing an error-covariance matrix for data assimilation and suffers from the same limitation of insufficient information to quantify the error variances of velocity at each grid point and their mutual covariances. For our approach to be practical, these uncertainties must be parameterized by just a few numbers, the coefficients ξ_n used in the expansion, each adjusting the strength of a pattern characterizing covarying uncertainties.

The issue becomes how to design these patterns. Clearly, their design should reflect the uncertainties likely to have the greatest impact on the particular outputs of the DeepC Oil model that are of greatest interest. And even then, there are more ways to design them than can be explored here. We have chosen to base them on an EOF decomposition of the variability of the HYCOM velocity field during the interval of the simulation, as that clearly assigns greater uncertainty in regions of higher variability. To insure that the uncertainty is appropriate to the region of the spill, it was restricted to the box bounded by latitudes 23°N and 31°N and longitudes 82°W and 92°W. The first two EOFs, which represent 56% of the total variance (41% the first mode, and 15% the second), were used to propagate uncertainty in the velocity field. Figure 2 presents maps of the percentage of the variance that is explained by both modes together at different depths. This plot illustrates where the imposed perturbations will mostly impact the velocity field.

The unperturbed surface velocity ($\boldsymbol{u}_0$) and the perturbations ($\boldsymbol{u}_1$ and $\boldsymbol{u}_2$) are presented on Figure 3 for days 10, 20, and 30 of the DWH oil spill (29 April 2010, 9 May 2010, and 19 May 2010, respectively). Both modes are dominated by variability from meanders of the Loop Current and Loop Current frontal eddies, although variability related to coastal currents are also present. These perturbations may intensify, weaken, or deform circulation features. Figure 4 presents three perturbed velocity fields for day 30, when both $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are strong (see Figure 3 and Figure S1 of the supporting information). These velocity fields ($\boldsymbol{u_a}$) were obtained considering $\xi_1 = 1$ and $\xi_2 = 1$ ($\boldsymbol{u_a} = \boldsymbol{u}_0 + \boldsymbol{u}_1 + \boldsymbol{u}_2$), $\xi_1 = 1$ and $\xi_2 = 0$ ($\boldsymbol{u_a} = \boldsymbol{u}_0 + \boldsymbol{u}_1$), and $\xi_1 = 0$ and $\xi_2 = 1$ ($\boldsymbol{u_a} = \boldsymbol{u}_0 + \boldsymbol{u}_2$). The effect of the perturbations on the dynamics is illustrated in the maps of vertical vorticity and lateral strain, which are presented in Figures S2 and S3 of the supporting information. The vorticity maps show the intensification of frontal cyclones and on the Loop Current in the perturbed velocity fields, most notably for $\boldsymbol{u_a} = \boldsymbol{u}_0 + \boldsymbol{u}_1 + \boldsymbol{u}_2$. The strain rate maps show intensification in the Loop Current system and over the continental shelf break of the northern Gulf of Mexico coast, including the Mississippi delta region. The circulation in the latter is dominated by wind-driven currents, and by a cyclonic eddy east of the DWH wellhead. The increased strain rate in the perturbed velocities indicates an enhanced dispersion around the

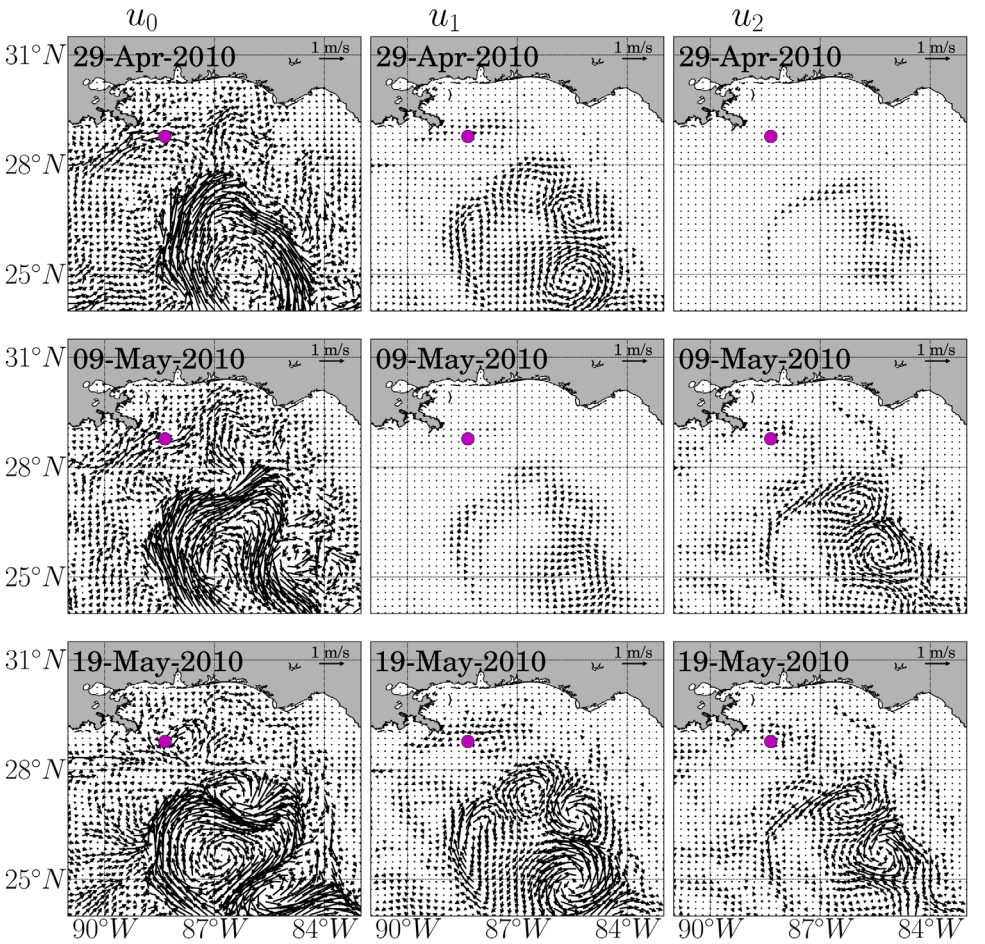

Figure 3. Surface velocity vectors (left) of the unperturbed velocity field (left) and of the perturbations constructed from the (middle) first and (right) second EOFs. The magenta dots depict the position of the Deepwater Horizon wellhead.

DWH site. An opposite effect occurs for negative values of ξ_1 and ξ_2, where for the same time, the perturbations induce a weakening in the vorticity and strain rates of the same features (not shown). Figure S4 of the supporting information shows the maps of lateral divergence of the same fields presented on Figure 5 and supporting information Figures S2 and S3. Qualitatively, there is no apparent difference between the perturbed fields and the unperturbed one.

4. The Polynomial Chaos Approach

The uncertainty quantification is based on a polynomial chaos formulation [M. Iskandarani et al., An overview of uncertainty quantification techniques for oceanic simulations, submitted to Journal of Geophysical Research: Oceans, 2015; *Knio and Le Maître*, 2006; *Najm*, 2009; *Thacker et al.*, 2015] whereby a surrogate for the model is constructed to represent the dependence of the model output on the uncertain data. This surrogate is subsequently used to explore the uncertain parameter space using a large ensemble to tally robust estimates of the output's statistical distribution. A polynomial chaos surrogate uses a spectral series of the form

$$C(\boldsymbol{x},t,\xi)=\sum_{k=0}^{P}\hat{C}_k(\boldsymbol{x},t)\psi_k(\xi), \tag{6}$$

where $C(\boldsymbol{x},t,\xi)$ represents the oil concentration, $\hat{C}_k(\boldsymbol{x},t)$ represents the polynomial chaos coefficients, $\psi_k(\xi)$ are the multidimensional orthogonal-polynomial basis functions representing the variations of C in the uncertain space, and P represents the truncation of the series to $(P+1)$ terms. The notation $C(\boldsymbol{x},t,\xi)$ emphasizes the dependence of the model output not only on space $\boldsymbol{x}$ and time t, but also on the vector of uncertain data

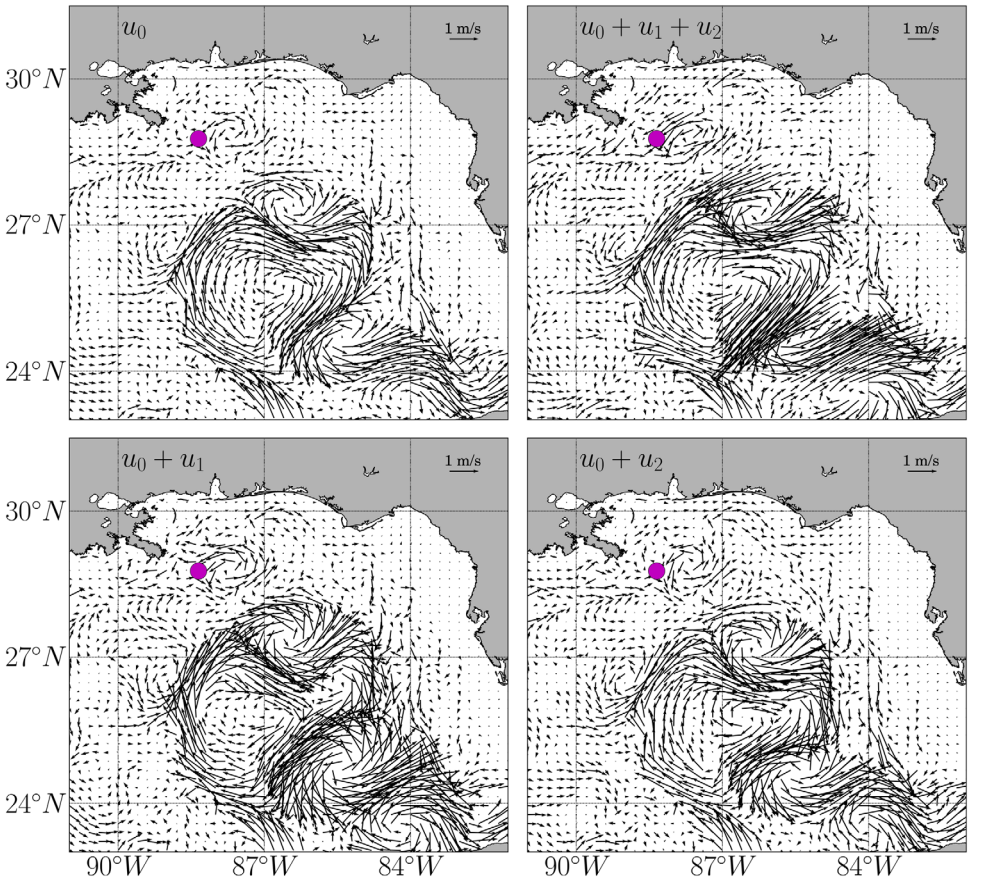

Figure 4. Surface velocity vectors at 19 May 2010 (top left) for the unperturbed velocity field $\boldsymbol{u_a}=\boldsymbol{u_0}$ and for three perturbed fields (top right) $\boldsymbol{u_a}=\boldsymbol{u_0}+\boldsymbol{u_1}+\boldsymbol{u_2}$, (bottom left) $\boldsymbol{u_a}=\boldsymbol{u_0}+\boldsymbol{u_1}$, and (bottom right) $\boldsymbol{u_a}=\boldsymbol{u_0}+\boldsymbol{u_2}$. The magenta dots depict the position of the Deepwater Horizon wellhead.

ξ. The basis functions are chosen to be orthogonal with respect to an inner product where the PDF of the uncertain variables ξ appears as the weight function. The multidimensional basis functions are product of 1-D basis functions. In a 3-D uncertain space, as we have here (three uncertain parameters), they take the form:

$$\psi_k(\boldsymbol{\xi})=P_{m_1}(\xi_1)P_{m_2}(\xi_2)P_{m_3}(\xi_3), \tag{7}$$

where $P_{m_d}(\xi_d)$ represents the 1-D basis function of degree m_d in the variable ξ_d, $d = 1, 2, 3$. As the standardized random variables ξ_d are uniformly distributed, $P_{m_d}(\xi_d)$ are Legendre polynomials of degree m_d. The

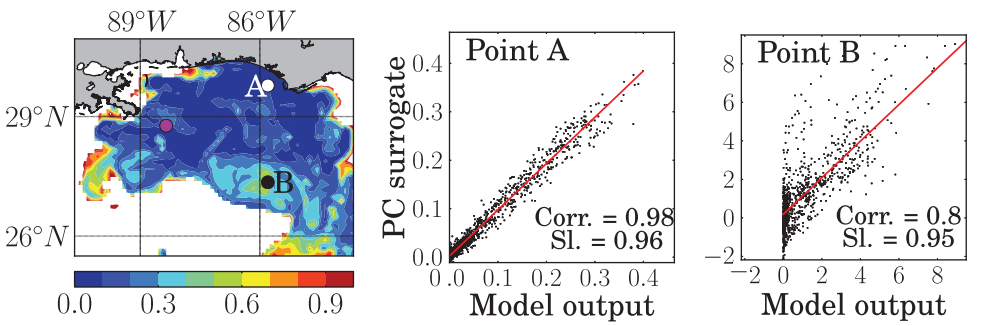

Figure 5. (left) RMSE of the oil concentration in the upper layer after 30 days. The RMSE was computed through a set of 1024 model realizations using randomly chosen sets of the uncertain input parameters and their respective surrogate. The values of oil concentration (in 10^4 kg/m^3) from the model realizations and their surrogates in points (middle) A and (right) B are presented in the scatterplots. The red lines are the linear regression of the scatterplots, and Corr. and Sl. are the correlation coefficients and slopes, respectively. The magenta dot indicates the DWH wellhead location.

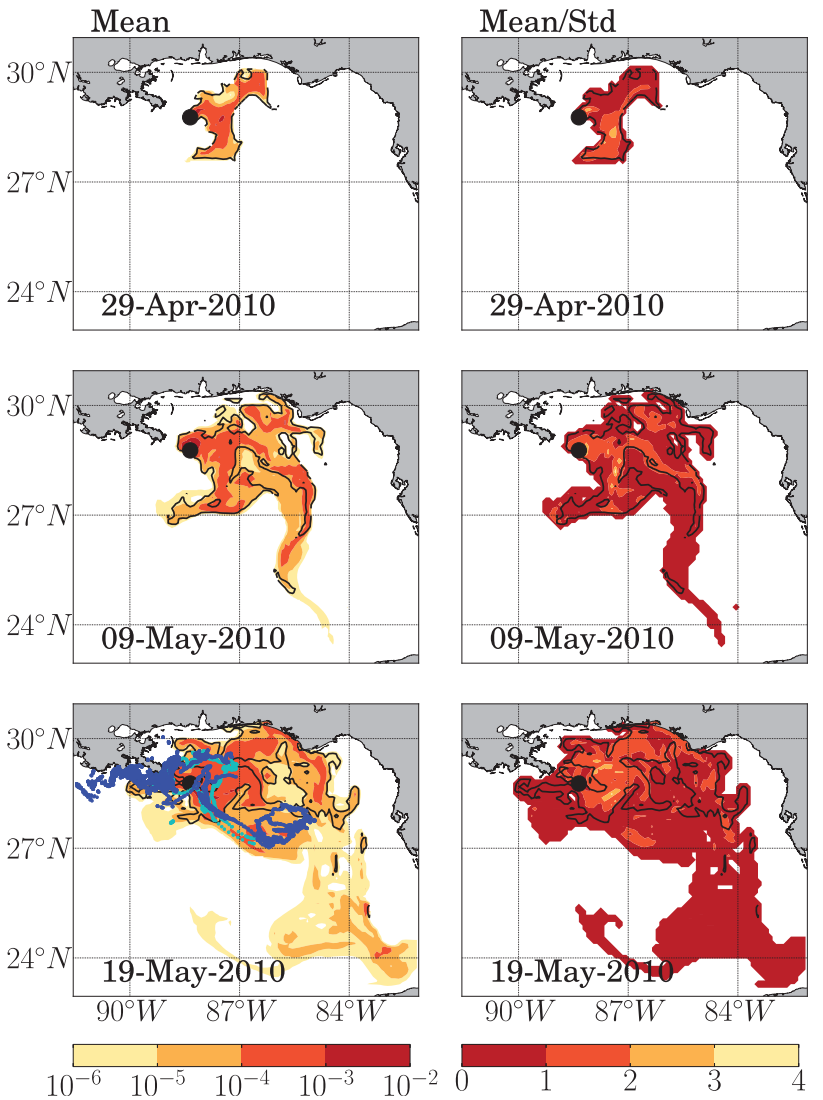

Figure 6. (left) Mean oil concentration between 0 and 5 m depth (in kg/m^3), and (right) mean over standard deviation ratio, on days 10 (29 April 2010), 20 (9 May 2010), and 30 (19 May 2010) of the oil spill. Both mean and standard deviation were extracted from the PC coefficients. The contours on the plots of 19 May 2010 indicate the position of the oil spill on 17 May 2010 (cyan) and 20 May 2010 (blue) as observed from a composite of satellite images. The black contours enclose cells for which the surrogate reproduces the model outputs with a *RMSE* $\leq$ 0.1, and the black dot indicates the DWH wellhead location.

polynomial degree in each dimension varies between $0 \leq m_d \leq M$, and we have used a total degree truncation for the multidimensional basis, i.e., $m_1+m_2+m_3 \leq M$. The total number of basis functions is then $P+1=\frac{(M+3)!}{M!3!}$.

In the present article, we use a nonintrusive Galerkin projection to compute the series coefficients and the associated integrals are approximated with tensorized Gauss quadrature:

$$\hat{C}_k(\boldsymbol{x},t)=\frac{\sum_{i=1}^{Q}\psi_k(\boldsymbol{\xi_i}^Q)C(\boldsymbol{x},t,\xi_i^Q)\Omega_i^Q}{\sum_{i=1}^{Q}\psi_k(\boldsymbol{\xi_i}^Q)\psi_k(\xi_i^Q)\Omega_i^Q}, \tag{8}$$

where Ω_i^Q and ξ_i^Q are the multidimensional quadrature weights and nodes, respectively. The number of unidimensional quadrature points needed to maintain an exact evaluation of the denominator is $q=(M+1)$ and hence the number of 3-D quadrature points needed is $Q=(M+1)^3$. The determination of the

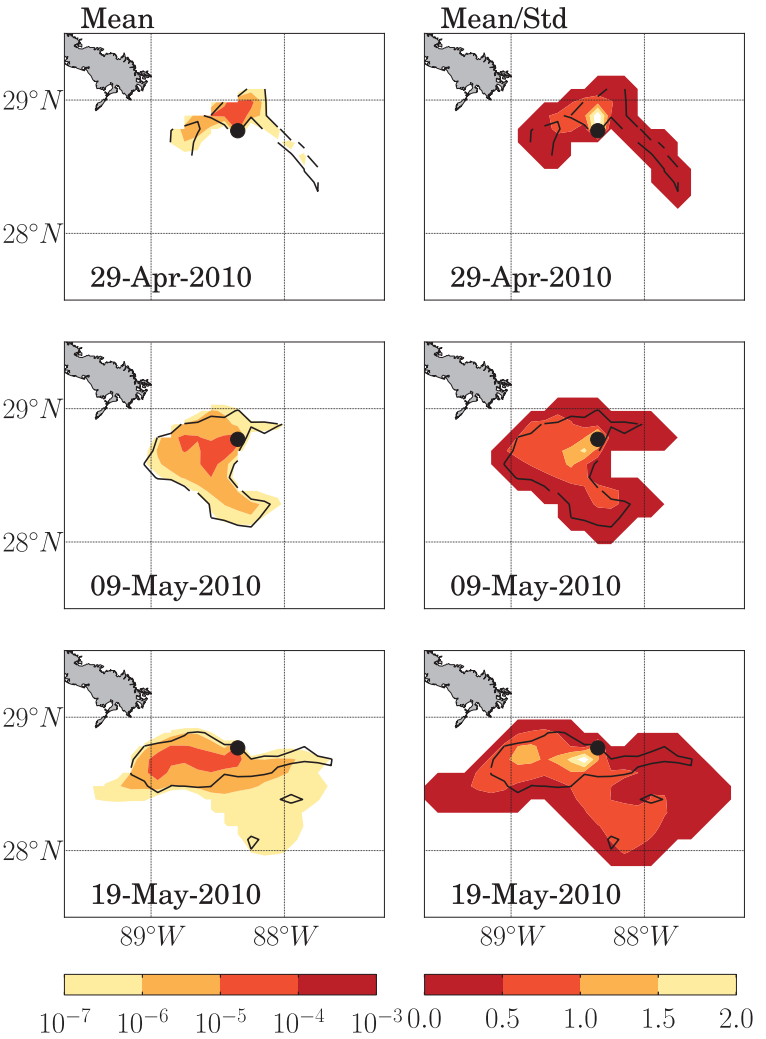

Figure 7. (left) Mean oil concentration between 300 and 600 m depth (in kg/m^3), and (right) mean over standard deviation ratio, on days 10 (29 April 2010), 20 (9 May 2010), and 30 (19 May 2010) of the oil spill. The black contours enclose cells for which the surrogate reproduces the model outputs with a *RMSE* $\leq$ 0.1, and the black dot indicates the DWH wellhead location.

coefficients thus requires an ensemble with Q members and this constitutes the bulk of the computational cost. Here the maximum degree of the polynomial was set to $M = 10$, which required $Q = 1331$ realizations of DCOM to compute the coefficients. Once the coefficients are available, the series can be used instead of the model to compute any required statistical information. Some statistical information can be obtained directly from the coefficients, e.g., the mean concentration over the uncertain data is simply given by $\hat{C}_0(\boldsymbol{x}, t)$ while the standard deviation is $\sum_{k=1}^{P} \hat{C}_k^2 ||\psi_k||^2$.

5. Results

5.1. PC Representation of Oil Concentration

In order to use the PC expansion as a surrogate for the model, it is important to check its consistency against actual model outputs. A metric that approximates the relative root-mean-square error (RMSE) of the PC surrogate was computed by using differences between the surrogate and the model at a large number of different combinations of the random inputs [*Alexanderian et al.*, 2012a]:

$$RMSE = \frac{\left(\sum_{\xi \in S} |C(\boldsymbol{x}, t, \xi) - \tilde{C}(\boldsymbol{x}, t, \xi)|^2 \right)^{1/2}}{\left(\sum_{\xi \in S} |C(\boldsymbol{x}, t, \xi)|^2 \right)^{1/2}}. \quad (9)$$

Here S is an ensemble of 1024 *independent* realizations of the model, where the values of the uncertain input parameters $\boldsymbol{\xi}(\xi_1, \xi_2, \xi_3)$ are randomly chosen from the assumed three-dimensional uniform distribution, C represents the oil concentration computed from the model's outputs, and $\tilde{C}$ is the respective PC surrogate. Figure 5 shows the scatterplot of all 1024 samples of $\tilde{C}$ and C at two grid cells on 19 May 2010. Point A illustrates a grid cell where the PC expansion represents satisfactorily the model outputs, with $RMSE = 0.13$. The scatterplot for this grid cell shows little spread around the linear regression line, which has a slope $s = 0.96$, and the concentrations present a correlation coefficient of 0.98. On the other hand,

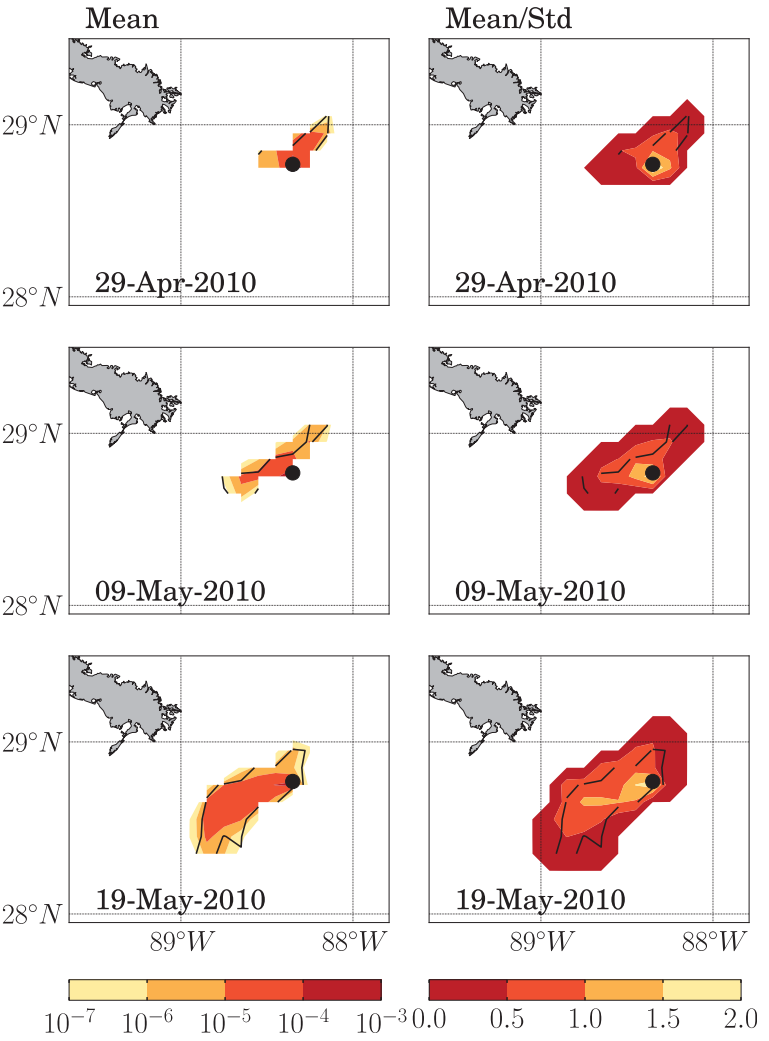

Figure 8. (left) Mean oil concentration between 900 and 1200 m depth (in kg/m^3), and (right) mean over standard deviation ratio, on days 10 (29 April 2010), 20 (9 May 2010), and 30 (19 May 2010) of the oil spill. The black contours enclose cells for which the surrogate reproduces the model outputs with a *RMSE* $\leq$ 0.1, and the black dot indicates the DWH wellhead location.

point B illustrates a grid cell where the PC expansion poorly approximate the model outputs, with *RMSE* = 0.59. In this case, many samples of the surrogate produced negative concentrations. Here PC expansions with *RMSE* $\leq$ 0.1 will be considered as suitable to be used as a surrogate for the model.

As presented on section 4, the mean and standard deviation of the quantity of interest can be computed directly from the coefficients of the PC expansion. They are a proxy for the mean and standard deviation of the distribution of all possible model runs inside the range of the uncertain parameters. Maps of the mean oil concentration in the upper layer (between 0 and 5 m depth) are presented in Figure 6 (left) at days 10, 20, and 30 of the spill (29 April 2010, 9 May 2010, and 19 May 2010). As the values range within several orders of magnitude, the maps are presented in a logarithmic scale. The plots on the right show the ratio of the mean over the standard deviations. Smaller values (in red) indicate relatively higher standard deviation, and therefore, higher uncertainty. In most cases, places with higher mean oil concentration presented relatively smaller uncertainty. On the edges of the oil plume, where concentrations are smaller, the standard deviation reached values which were equal or higher than the mean. Uncertainty is also high in the regions dominated by the mesoscale variability of the Loop Current system.

Although the scope of the paper is mainly to present the PC approach to simulations of oil transport in the ocean, a comparison of our results to the observed oil slick from the DWH spill may bring some insight of what processes were well represented, and what was missed from our simulations. We compared the results of oil concentration in the surface layer to daily composite products from NOAA's Satellite Analysis Branch [*Streett*, 2011], which is composed by synthetic aperture radar and visible/near infra-red multispectral satellite imagery. Only qualitative comparisons can be made, as the daily composites just provide information about the presence of the oil slick, not giving any information that can be related to oil concentration. The detection of the oil slick through SAR imagery presents some limiting factors that may affect the quality of the composites. The most important one is that the slick cannot be detected under strong (greater than 12.8 m/s) or low (less than 2.1 m/s) wind conditions, as these conditions induce false positive and false negative, respectively. Other important factors that affect the quality and availability of

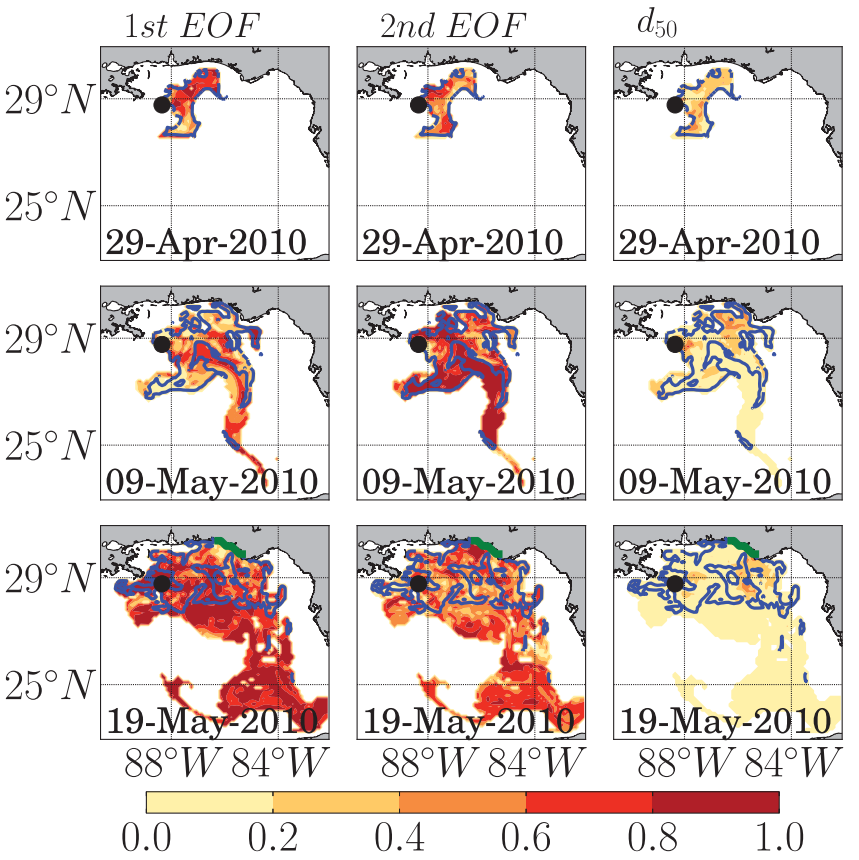

Figure 9. Global sensitivity index at the upper layer for the first and second EOF modes, and the fiftieth percentile of the droplet size distribution (d_{50}), on days 10 (29 April 2010), 20 (9 May 2010), and 30 (19 May 2010) of the oil spill. The blue contour encloses cells for which the surrogate reproduces the model outputs with a $RMSE \leq 0.1$. The green line on the 19 May 2010 plots indicates the grid cells used to compute the total amount of oil that reached the northwestern coast of Florida during the first 30 days of the oil spill (Figure 12).

the composites are cloud cover, and limited satellite image coverage, which may not cover the whole spill area every day [*Streett*, 2011; *Dietrich et al*., 2012]. For the period considered (initial 30 days of the oil spill), only two daily composites are available; from 17 May 2010 and from 19 May 2010, although the quality of the last one is poor due to the lack of imagery. Thereby, we compared the observed oil slick from days 17 May 2010 and 20 May 2010 with the simulated oil concentration at the surface layer on day 19 May 2010. The contours of the observed oil slick are superimposed on the plot of the mean oil concentration from 19 May 2010 (Figure 6). During this period, part of the oil slick was advected southward into a frontal cyclone north of the Loop Current, while another portion was advected westward along the Louisiana coast by wind-driven currents [*Walker et al*., 2011]. The first was well captured by the simulations, while the westward advection of the oil was underestimated. Overall, the observed spread of the oil was mostly inside the uncertainty range of the simulations. On the other hand, the simulated oil transport to the east, which resulted in oil arriving in the northwestern coast of Florida by 9 May 2010, and in the Straits of Florida by 14 May 2010, was not observed during the oil spill. In the simulations, the oil was advected eastward/northeastward during the first 10 days of the spill (see day 29 April 2010 on Figure 6). Part of the oil that was advected eastward was then transported to the south, where it reached the Loop Current by 9 May 2010. The oil, then, got into the Straits of Florida by 14 May 2010. During this period when simulations presented a strong eastward transport, observations have shown that the oil was not transported east of 87°W [*Liu et al*., 2011a; *Walker et al*., 2011]. The discrepancy between simulations and observations could be simply explained by ocean circulation patterns that were not present in the HYCOM velocity field. On the other hand, it could also be caused by the neglected effects of wind and waves. Numerical experiments of *Le Henaff et al*. [2012] showed that the direct effect of the wind on the DWH oil slick might have constrained the spread of the oil closer to the Louisiana coast in the initial stages of the oil spill, which kept the oil from reaching the Loop Current, and thereafter the Straits of Florida. *Huntley et al*. [2011] tested the direct effect

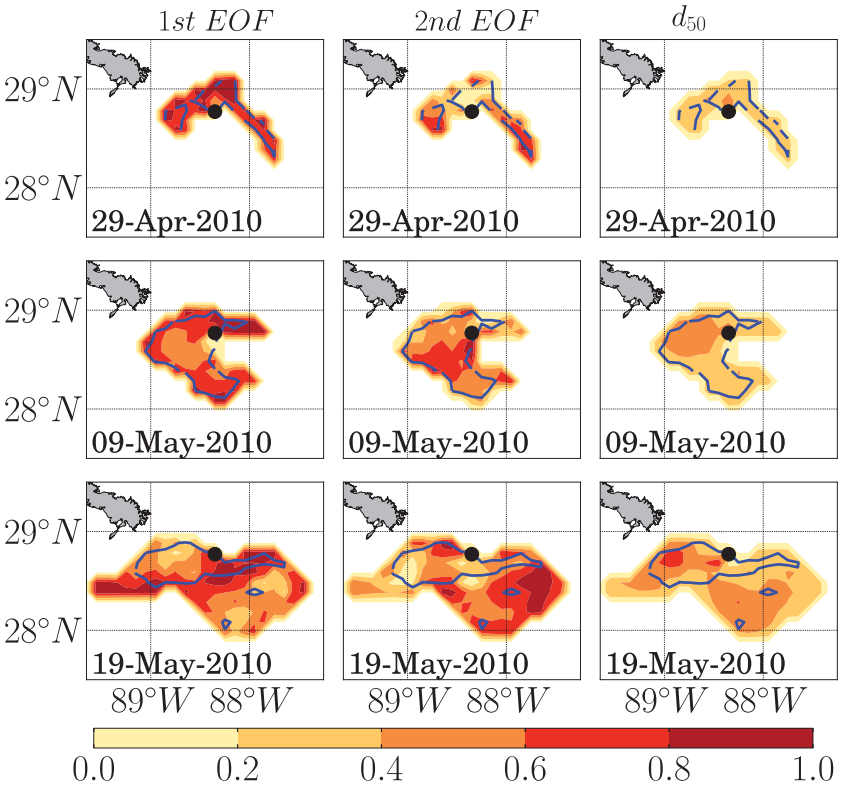

Figure 10. Same as Figure 9 but for the middepth layer.

of the wind in different stages of the DWH spill, and obtained better results with a parameterized wind effect for the oil transport in coastal areas, although this effect was negligible away from the shelf. The effect of waves through stokes drift in the transport of the surface oil slick is very similar to the direct effect of wind, and as *Carratelli et al*. [2011] described, it is hard to distinguish both conceptually and practically. *Dietrich et al*. [2012] used a high-resolution finite element model coupled with a wave model to simulate the transport of oil from the DWH over the continental shelf, and found better results without adding the direct effect of the wind.

The mean and standard deviation for the middepth and deep layers are presented in Figures 7 and 8. The deep layer reached higher concentrations than the middepth layer, as ocean currents are stronger in the latter, and consequently, dispersion is higher. Another aspect which influenced this behavior was that particles with smaller droplet sizes got trapped in the deep layer, while all particles that reached the middepth range kept moving toward the surface. As both of these layers span a depth range of 300 m (300–600 m for the middepth layer and 900–1200 m for the deep layer), oil concentrations cannot be directly compared to the concentrations in the upper layer, which span a 5 m depth range, as the Lagrangian particles are binned

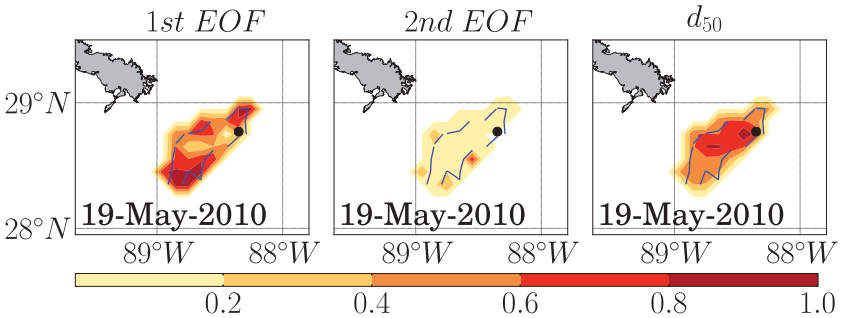

Figure 11. Global sensitivity index for the deep layer for the three input parameters on day 30.

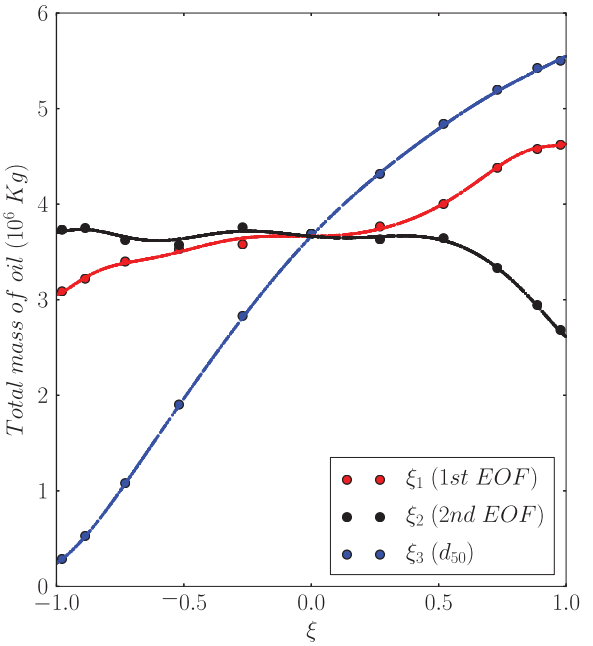

Figure 12. Amount of oil that reached the coast of Florida (green line on Figure 9) after 30 days considering different input parameters. ξ_1, ξ_2, and ξ_3 are the standardized random variables used to propagate uncertainty through the first and second EOFs, and the fiftieth percentile of the droplet size distribution, respectively. The red dots are the outputs of simulations considering $\xi_1 \in [-1, 1], \xi_2 = 0, \xi_3 = 0$, the black dots are from simulations with $\xi_2 \in [-1, 1], \xi_1 = 0, \xi_3 = 0$, and the blue dots are from simulations with $\xi_3 \in [-1, 1], \xi_1 = 0, \xi_2 = 0$. The big dots are outputs from actual realizations of the model, while the smaller ones were obtained from the PC surrogate.

in a larger volume. Uncertainty was high in both layers, with values of standard deviation reaching twice the mean concentration in the edges of the oil plume, where concentrations are low.

In the upper layer, most of the oil is inside the $RMSE \leq 0.1$ margin in the first 10–15 days (Figure 6 shows day 10), indicating that the surrogate is a good representation of the model during this period. After 15 days, the oil spreads toward and around the Loop Current system where the errors of the surrogate increase above 0.1. At all depth ranges analyzed, the areas with a $RMSE > 0.1$ presented high uncertainty, and for the two deeper layers, most of these areas occurred where the standard deviation was at least twice as high as the mean. Also, in the deeper layers, most of the areas with $RMSE > 0.1$ presented an oil concentration $C \leq 10^{-6}\text{kg/m}^3$.

The results in the upper layer show that the quality of the PC surrogate deteriorates in areas dominated by mesoscale variability. This is explained by the fact that the perturbations in the velocity field are larger there. Another problem occurs in areas with a relatively small oil concentration. In these areas, the grid cells used to compute the oil concentration are reached by a small number of Lagrangian particles. With that, a slight variation in the number of particles in these grid cells, among different realizations of the model, results in a relatively abrupt change of oil concentration. In this case, the response of the model to the different values of input parameters is not accurately represented by the 10° PC surrogate. These problems could be attenuated by increasing the order of the PC expansions, although this measure would require a considerably higher number of model realizations to compute the coefficients. For example, increasing the order from 10 to 11 would require another 397 model realizations. Another measure to reduce the errors in the surrogate would be to increase the area of the grid cells in which the Lagrangian particles are binned. In this case, the drawback is the loss of spatial resolution.

5.2. Sensitivity Analysis

In order to quantify the contribution of each input parameter on the uncertainty of our quantity of interest, we took advantage of the PC coefficients, as they provide an efficient and computationally cheap way to perform variance-based sensitivity analysis. A global sensitivity index was computed as the ratio between the fraction of the variance due to a specific input parameter and the total variance considering all parameters (for more information, the reader is referred to *Alexanderian et al.* [2012b]). As the index quantifies the total sensitivity of the model to a specific parameter, it also takes into account the interaction between the different parameters.

The sensitivity indices for the three input parameters in the upper layer for days 10, 20, and 30 of the oil spill are presented in Figure 9. The blue contour indicates the $RMSE = 0.10$, presented previously. Overall, the perturbations on the velocity field dominated the variance, with the first EOF mode being dominant on days 10 and 30, and the second mode being more significant on day 20. This behavior follows the patterns of the principal components, as on day 20, the absolute value of the second mode is higher than the first mode (Figure S1 of the supporting information). Comparing Figures 3 and 9 elucidates how each mode

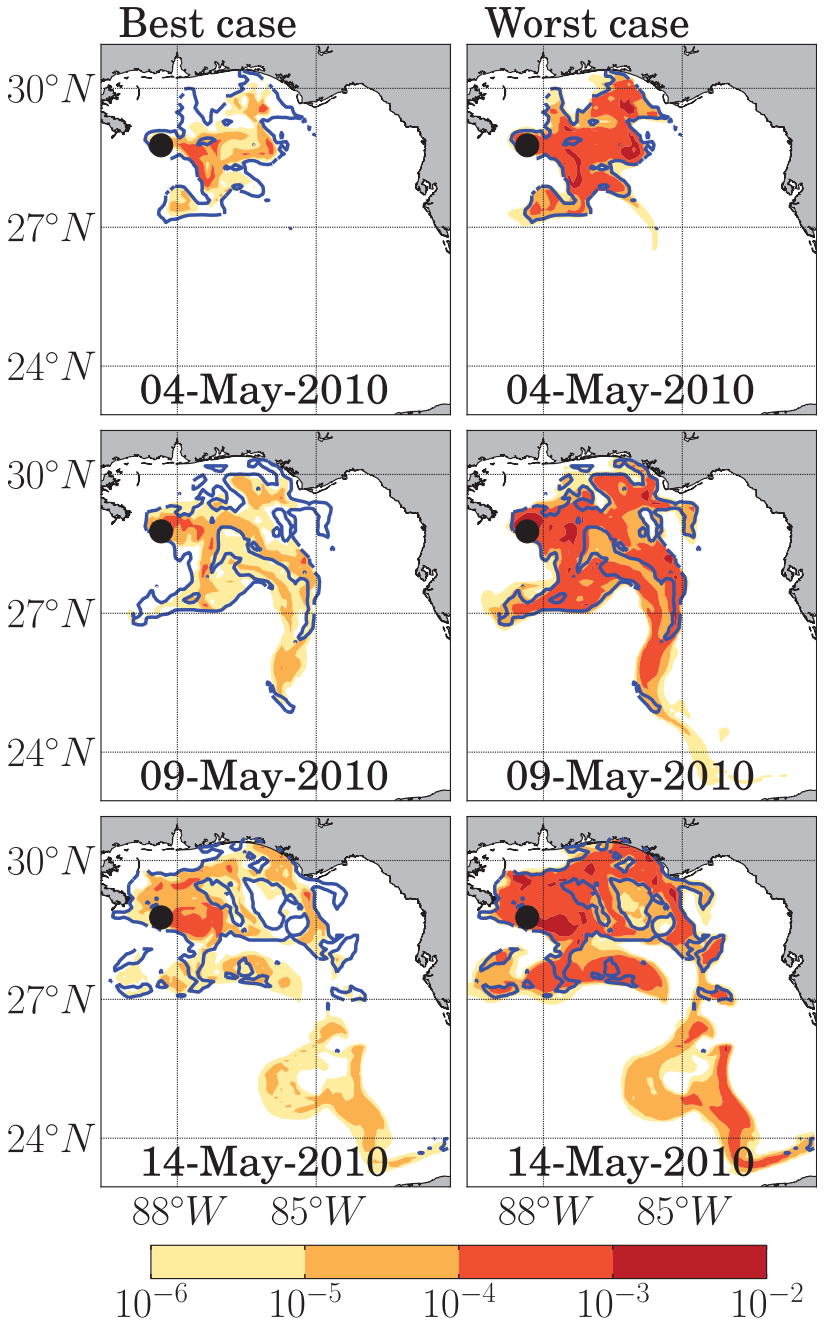

Figure 13. The **(left) ten**th and (right) ninetieth percentiles of the PDF of oil the concentration (kg/m^3) between 0 and 5 m depth for days 15 (4 May 2010), 20 (9 May 2010), and 25 (14 May 2010) of the oil spill, considering 10^5 samples of the PC proxy. The blue contour encloses cells for which the surrogate reproduces the model outputs with a $RMSE \leq 0.1$, and the black dot indicates the DWH wellhead location.

impacts the oil concentration uncertainty. On day 10, the oil is being transported toward the shelf by wind-driven currents, which at this moment are affected by the first EOF. On day 20, the oil is reaching the Loop Current, which is perturbed by the second EOF. On day 30, both modes impact the circulation on the Loop Current system and closer to the coast. The contribution of the droplet size distribution was more prominent on day 10, but it still reached as much as 60% of the total variance in some areas on days 20 and 30. In the middepth layer, the velocity perturbations are still dominant, although the contribution of the droplet size distribution uncertainty is much higher than in the surface layer (Figure 10). In the deep layer, the uncertainty in the droplet size distribution dominated the variance in the core of the deep plume, while the first EOF mode contributed to the uncertainty on its edges. Figure 11 shows the sensitivity indices only for

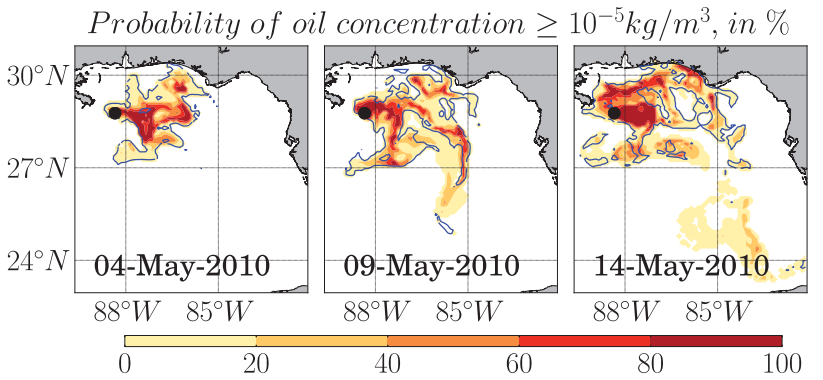

Figure 14. Probability of oil concentration (in %) to reach at least 10^5kg/m^3 between 0 and 5 m depth on days 15 (4 May 2010), 20 (9 May 2010), and 25 (14 May 2010) of the oil spill. The blue contour encloses cells for which the surrogate reproduces the model outputs with a $RMSE \leq 0.1$, and the black dot indicates the DWH wellhead location.

day 30, as on the other days the results followed the same pattern. The effect of the second EOF at this depth range is negligible, which reflects that the second EOF signal around this area is relatively weak.

In the upper layer, in the areas outside the $RMSE = 0.1$ contour, where the PC surrogate did not approximate the output of the model accurately, the variance was largely dominated by the EOF modes. This indicates that the perturbations of the velocity field induced a complex response in the oil concentration in those areas that could not be captured by the tenth-order polynomial expansion. This problem could be attenuated if we increased the size of the grid cells in which we binned the Lagrangian particles to compute the oil concentration. As the Eulerian velocity is what controls the horizontal displacement of the particles, a coarser grid would reduce the impact of the velocity perturbations in the uncertainty of our quantity of interest. Similarly, a bigger depth range of the layer used to compute the oil concentrations would reduce the impact of the droplet size distribution uncertainty. If instead of defining our quantity of interest as oil concentration in a grid, we define it as the amount of oil that reached some sensitive area, the response of the model to perturbations in each parameter will be different. Let us consider that our quantity of interest now is the amount of oil that reaches the shoreline of northwestern Florida (green line on day 19 May 2010 plot of Figure 9) during the first 30 days of oil spill. Figure 12 shows the curves of the total amount of oil that reaches this area as a function of each uncertain parameter. For each curve, the PC surrogate was sampled 10,000 times with the values of one specific parameter being randomly chosen from an uniform distribution, while the other two parameters were kept zero. These results indicate that the uncertainty in this quantify of interest was largely dominated by the droplet size distribution, with the amount of oil varying by an order of magnitude between the simulation considering a distribution with $d_{50} = 0.2$ mm (related to $\xi_3 = -1$), and a simulation considering $d_{50} = 0.9$ mm ($\xi_3 = 1$). The increased impact of d_{50} in this case is not surprising, as the area considered to compute the amount of oil that reached the shoreline is 10 times bigger than the horizontal area of the grid cells used to compute the oil concentrations. On the other hand, these curves do not show the combined effect of the parameters, as the global sensitivity index does.

5.3. Probabilistic Maps

Probability density functions of oil concentration were constructed at each grid point by sampling the PC expansion 100,000 times with randomly chosen values of the variables ξ_1, ξ_2, and ξ_3. These PDFs provide valuable information for impact assessment and for contingency planning during oil spill events. By extracting the tenth percentile of each distribution, we can get a measure of "best case scenario" of the oil spread, where there is a 90% chance of the concentration to reach that value for the uncertain parameters space considered. On the other hand, the ninetieth percentile provides a measure of "worst case scenario," where in only 10% of the uncertain parameters space the oil concentration suppress its value. Figure 13 presents the maps of the tenth and ninetieth percentiles of the PDFs of oil concentration in the upper layer at days 15, 20, and 25. These plots highlight the importance of taking into account the uncertainty of the model in the oil fate forecast. For instance, in the worst case scenario, the oil reaches the coast of Florida, near Panama City, on day 15,

while on the best case scenario, the oil reaches this area after day 20. Also, in the worst case scenario, the oil slick would reach the Mississippi Delta by day 25, which did not happen in the best case scenario.

Another way to use these PDFs is to build probabilistic hazard maps given a threshold value [*Madankan et al*., 2014]. Considering that in our simulations we did not take into account any type of oil degradation, it is hard to define what concentration would represent some risk to the environment, which could be used as threshold. Our choice here was arbitrary, as we picked a concentration that was 2 orders of magnitude smaller than the maximum concentration obtained in the upper layer. Figure 14 shows the probability of the oil concentration to reach at least $10^{-5}\mathrm{kg/m^3}$ in the upper layer on days 15, 20, and 25. According to these maps, it is more likely that the oil would reach the coast of Florida after 20 days. Also, the maps show a high probability for the oil slick to reach the Mississippi Delta by day 25.

6. Summary and Discussion

The purpose of this paper was to present a methodology to account for model uncertainty in oil fate simulation during a deep water spill. The ensemble-based PC approach was used to propagate uncertainty in the Eulerian velocity field and droplet size distribution of a Lagrangian oil fate model. The method aims to represent a model output as a polynomial series expansion of model input parameters, and this representation can then be used as a surrogate for the model. In order to test the capabilities of the PC expansion, we applied the technique on the first 30 days of the DWH oil spill, and checked the quality of the PC representation of the model during the simulation period. A variance-based sensitivity analysis was carried out through the PC coefficients, where the total contribution of each source of uncertainty considered was quantified at different depths. Finally, we used the PC surrogate to construct probability density functions of oil concentrations, and applied them to produce probabilistic hazard maps.

The biggest challenge we faced was to find a reasonable way to perturb the Eulerian velocity field without having to run extra realizations of HYCOM, and prescribe the perturbations with a small number of parameters. By using the EOF modes to accomplish that, we assumed that the uncertainty in the velocity field from HYCOM had patterns of variability that resemble the variability it exhibited during the period analyzed. This way, higher uncertainty was assigned where the flow was more energetic, like in regions dominated by mesoscale variability, which is a reasonable assumption. On the other hand, if HYCOM missed completely some physical process or feature, like a mesoscale eddy, due to misspecification of some parametrization (or boundary/initial conditions), this source of uncertainty was not accounted for by the EOF modes. Our approach targets the effect of relatively small deviations from the Eulerian velocity field, but, as in any other approach to estimate oil transport in the ocean, a reliable representation of the ocean state is paramount.

It is important to point out that, as our focus was to demonstrate the capabilities of the methodology, we were not concerned about accurately reproducing observations of the oil slick from the DWH spill. Actually, several important aspects prevented our results from correctly reproducing the observations of the surface oil slick. First, as discussed on section 5.1, we did not account for the direct impact of wind or Stokes drift in the transport of oil, which are likely to impact the spread of the surface oil slick closer to the shore. Another important factor is that the predictability limit of Lagrangian trajectories in the area affected by the oil spill ranges from about 1.5–6 days [*Mariano et al*., 2011], which makes the task of forecasting the oil transport for long periods impractical without updating the location of the oil with observations. Applications of Lagrangian models to simulate the transport of the oil from the DWH by *Liu et al*. [2011b], *Macfadyen et al*. [2011], *Huntley et al*. [2011], and *Mariano et al*. [2011] highlight the importance of applying observations to correct the location of the oil, in order to accomplish a faithful forecast. Also, we neglected oil weathering processes, like biodegradation and evaporation, which might decrease the horizontal distribution of oil by thousands of kilometers [*Adcroft et al*., 2010; *Mariano et al*., 2011; *North et al*., 2015]. Finally, physical processes which are not resolved by the HYCOM velocity field have been identified as important factors in the dispersion of the DWH oil spill [*Poje et al*., 2014].

Nevertheless, the capabilities of the PC approach presented can improve significantly the oil fate forecast, and support contingency planning during oil spill events. In order to apply the methodology operationally in a 7 days forecast, the order of the polynomial expansion can be reduced to decrease the computational cost of calculating the expansion coefficients, and provide a faster response. In previous experiments within the same scenario, we were able to satisfactorily reproduce our quantity of interest up to 10 days using a

sixth-order PC expansion. In this case, the computation of the PC coefficients required only 343 realizations of the oil fate model, against the 1331 used to evaluate the tenth-order PC expansions presented. Further, other sources of uncertainty should be considered, such as oil weathering, wind drag, and stokes drift. The last two can be added in the Eulerian velocity field, and be perturbed along with the ocean currents through EOF decomposition. This way, the number of uncertain input parameters would not change, and no extra realizations of the model would be necessary. At last, alternative ways to perturb the velocity field are possible, and should be the focus of further investigation. Considering a sixth-order PC expansion, and limiting the number of variables used to perturb the velocity field, we can propagate uncertainty on an ensemble of HYCOM realizations [*Thacker et al*., 2012; M. Iskandarani et al., Quantifying initial conditions uncertainty in the Gulf of Mexico using polynomial chaos expansions, submitted to Journal of Geophysical Research: Oceans, 2016]. If we are able to reduce the number of uncertain parameters to two, as we did in the present study, a maximum of 49 realizations of HYCOM would be necessary. This number can be further reduced by the application of adaptive quadrature techniques to compute the PC coefficients [*Constantine et al*., 2012; *Winokur et al*., 2013; M. Iskandarani et al., An overview of uncertainty quantification techniques for oceanic simulations, submitted to Journal of Geophysical Research: Oceans, 2015].

Acknowledgments

This research was made possible in part by a grant from BP/The Gulf of Mexico Research Initiative to the Deep-C and CARTHE Consortia, by the Office of Naval Research, award N00014-101-0498, and by the US Department of the Interior, Bureau of Ocean Energy Management under the cooperative agreement MC12AC00019.
R. Gonçalves acknowledges support by the Brazilian Ministry of Science, Technology and Innovation (CNPq-Council for Scientific and Technological Development) through a PHD scholarship from the Science Without Borders program, grant 202263/2012-6. O. Knio acknowledges partial support from the US Department of Energy, Office of Advanced Scientific Computing Research, under award DE-SC0008789. This research was conducted in collaboration with and using the resources of the University of Miami Center for Computational Science. The outputs of the DeepC Oil Model used here are publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC). [Available at https://data.gulfresearchinitiative.org/data/R1.x138.077:0026.]

References

Adcroft, A., R. Hallberg, J. P. Dunne, B. L. Samuels, J. A. Galt, C. H. Barker, and D. Payton (2010), Simulations of underwater plumes of dissolved oil in the Gulf of Mexico, *Geophys. Res. Lett.*, *37*, L18605, doi:10.1029/2010GL044689.

Alexanderian, A., O. P. Le Maître, H. N. Najm, M. Iskandarani, and O. M. Knio (2012a), Multiscale stochastic preconditioners in non-intrusive spectral projection, *J. Sci. Comput.*, *50*(2), 306–340, doi:10.1007/s10915-011-9486-2.

Alexanderian, A., J. Winokur, I. Sraj, A. Srinivasan, M. Iskandarani, W. C. Thacker, and O. M. Knio (2012b), Global sensitivity analysis in an ocean general circulation model: A sparse spectral projection approach, *Comput. Geosci.*, *16*(3), 757–778, doi:10.1007/s10596-012-9286-2.

Al-Rabeh, A., R. W. Lardner, and N. Gunay (2000), Gulfspill Version 2.0: A software package for oil spills in the Arabian Gulf, *Environ. Modell. Software*, *15*, 425–442.

Aman, Z. M., C. B. Paris, E. F. May, M. L. Johns, and D. Lindo-Atichati (2015), High-pressure visual experimental studies of oil-in-water dispersion droplet size, *Chem. Eng. Sci.*, *127*, 392–400, doi:10.1016/j.ces.2015.01.058.

Berry, A., T. Dabrowski, and K. Lyons (2012), The oil spill model OILTRANS and its application to the Celtic Sea, *Mar. Pollut. Bull.*, *64*(11), 2489–2501, doi:10.1016/j.marpolbul.2012.07.036.

Bleck, R. (2002), An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modell.*, *4*(1), 55–88, doi:10.1016/S1463-5003(01)00012-9.

Brandvik, P. J., O. Johansen, F. Leirvik, U. Farooq, and P. S. Daling (2013), Droplet breakup in subsurface oil releases—Part 1: Experimental study of droplet breakup and effectiveness of dispersant injection, *Mar. Pollut. Bull.*, *73*(1), 319–326, doi:10.1016/j.marpolbul.2013.05.020.

Carratelli, E. P., F. Dentale, and F. Reale (2011), On the effects of wave-induced drift and dispersion in the Deepwater Horizon oil spill, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 197–204, AGU, Washington, D. C., doi:10.1029/2011gm001109.

Cekirge, H., M. Koch, C. Long, C. Giammona, K. Binkley, R. Engelhardt, and R. Jamail (1995), State-of-the-art techniques in oil spill modeling, in *International Oil Spill Conference*, vol. 1995, no. 1, pp. 67–72, Amer. Petrol. Inst., Washington, D. C.

Chao, X., N. J. Shankar, and S. S. Y. Wang (2003), Development and application of oil spill model for Singapore coastal waters, *J. Hydraul. Eng.*, *129*(7), 495–503, doi:10.1061/(ASCE)0733-9429(2003)129:7(495).

Chassignet, E. P., and A. Srinivasan (2015), Data assimilative hindcast for the Gulf of Mexico, *OCS Study BOEM 2015-035*, 46 pp., U.S. Dep. of the Inter., Bur. of Ocean Energy Manage., Sterling, Va.

Chassignet, E. P., L. T. Smith, G. R. Halliwell, and R. Bleck (2003), North Atlantic Simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity, *J. Phys. Oceanogr.*, *33*(12), 2504–2526, doi:10.1175/1520-0485(2003)033 < 2504:NASWTH>2.0.CO;2.

Chassignet, E. P., et al. (2006), Generalized vertical coordinates for eddy-resolving global and coastal ocean forecasts, *Oceanography*, *19*(1), 118–129, doi:10.5670/oceanog.2006.95.

Chassignet, E. P., et al. (2009), U.S. GODAE: Global ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM), *Oceanography*, *22*(2), 64–75.

Constantine, P. G., M. S. Eldred, and E. T. Phipps (2012), Sparse pseudospectral approximation method, *Comput. Methods Appl. Mech. Eng.*, *229-232*, 1–12, doi:10.1016/j.cma.2012.03.019.

Dietrich, J. C., et al. (2012), Surface trajectories of oil transport along the Northern Coastline of the Gulf of Mexico, *Cont. Shelf Res.*, *41*, 17–47, doi:10.1016/j.csr.2012.03.015.

Galt, J. A. (1997a), The integration of trajectory models and analysis into spill response information systems, *Spill Sci. Technol. Bull.*, *4*(2), 123–129, doi:10.1016/S1353-2561(98)00008-5.

Galt, J. A. (1997b), Uncertainty analysis related to oil spill modeling, *Spill Sci. Technol. Bull.*, *4*(4), 231–238, doi:10.1016/S1353-2561(98)00027-9.

Griffa, A. (1996), Applications of stochastic particle models to oceanographic problems, *Stochastic Modell. Phys. Oceanogr.*, *39*, 113–140, doi:10.1007/978-1-4612-2430-3_5.

Halliwell, G. R., A. Srinivasan, V. Kourafalou, H. Yang, D. Willey, M. Le Hénaff, and R. Atlas (2014), Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mexico, *J. Atmos. Oceanic Technol.*, *31*(1), 105–130, doi:10.1175/JTECH-D-13-00011.1.

Hodges, B. R., A. Orfila, J. M. Sayol, and X. Hou (2015), Operational oil spill modeling: From science to engineering applications in the presence of uncertainty, in *Mathematical Modelling and Numerical Simulation of Oil Pollution Problems*, edited by M. Ehrhardt, chap. 5, pp. 99–126, Springer, doi:10.1007/978-3-319-16459-5.

Huntley, H. S., B. L. Lipphardt Jr., and A. D. Kirwan Jr. (2011), Surface drift predictions of the Deepwater Horizon spill: The Lagrangian perspective, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 179–195, AGU, Washington, D. C., doi:10.1029/2011gm001097.

Johansen, O. (2003), Development and verification of deep-water blowout models, *Mar. Pollut. Bull.*, *47*(9-12), 360–368, doi:10.1016/S0025-326X(03)00202-9.

Johansen, O., P. J. Brandvik, and U. Farooq (2013), Droplet breakup in subsea oil releases—Part 2: Predictions of droplet size distributions with and without injection of chemical dispersants, *Mar. Pollut. Bull.*, *73*(1), 327–335, doi:10.1016/j.marpolbul.2013.04.012.

Knio, O. M., and O. P. Le Maître (2006), Uncertainty propagation in CFD using polynomial chaos decomposition, *Fluid Dynamics Research*, *38*(9), 616–640, doi:10.1016/j.fluiddyn.2005.12.003.

Korotenko, K. A., M. J. Bowman, and D. E. Dietrich (2010), High-resolution numerical model for predicting the transport and dispersal of oil spilled in the black sea, *Terr. Atmos. Oceanic Sci.*, *21*(1), 123–136, doi:10.3319/TAO.2009.04.24.01(IWNOP).

Kujawinski, E. B., M. C. Kido Soule, D. L. Valentine, A. K. Boysen, K. Longnecker, and M. C. Redmond (2011), Fate of dispersants associated with the Deepwater Horizon oil spill, *Environ. Sci. Technol.*, *45*(4), 1298–1306, doi:10.1021/es103838p.

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

Lehr, W., R. Jones, M. Evans, D. Simecek-Beatty, and R. Overstreet (2002), Revisions of the ADIOS oil spill model, *Environ. Modell. Software*, *17*(2), 189–197, doi:10.1016/S1364-8152(01)00064-0.

Liu, Y., R. H. Weisberg, C. Hu, C. Kovach, and R. Riethmüller (2011a), Evolution of the loop current system during the deepwater horizon oil spill event as observed with drifters and satellites, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 91–101, AGU, Washington, D. C., doi:10.1029/2011gm001127.

Liu, Y., R. H. Weisberg, C. Hu, and L. Zheng (2011b), Trajectory forecast as a rapid response to the deepwater horizon oil spill, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 153–165, AGU, Washington, D. C., doi:10.1029/2011GM001121.

Macfadyen, A., G. Y. Watabayashi, C. H. Barker, and C. J. Beegle-Krause (2011), Tactical modeling of surface oil transport during the deepwater horizon spill response, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 167–178, AGU, Washington, D. C., doi:10.1029/2011GM001128.

Madankan, R., S. Pouget, P. Singla, M. Bursik, J. Dehn, M. Jones, A. Patra, M. Pavolonis, E. Pitman, T. Singh, and P. Webley (2014), Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion, *J. Comput. Phys.*, *271*, 39–59.

Mariano, A. J., V. H. Kourafalou, A. Srinivasan, H. Kang, G. R. Halliwell, E. H. Ryan, and M. Roffer (2011), On the modeling of the 2010 Gulf of Mexico Oil Spill, *Dyn. Atmos. Oceans*, *52*(1-2), 322–340, doi:10.1016/j.dynatmoce.2011.06.001.

McNutt, M. K., R. Camilli, T. J. Crone, G. D. Guthrie, P. A. Hsieh, T. B. Ryerson, O. Savas, and F. Shaffer (2011), Science applications in the deepwater horizon oil spill special feature: Review of flow rate estimates of the deepwater horizon oil spill, *Proc. Natl. Acad. Sci. U. S. A.*, *109*(50), 20,260–20,267, doi:10.1073/pnas.1112139108.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360, doi:10.1175/BAMS-87-3-343.

Najm, H. N. (2009), Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics, *Annu. Rev. Fluid Mech.*, *41*(1), 35–52, doi:10.1146/annurev.fluid.010908.165248.

Neves, A. A. S., N. Pinardi, F. Martins, J. Janeiro, A. Samaras, G. Zodiatis, and M. De Dominicis (2015), Towards a common oil spill risk assessment framework Adapting ISO 31000 and addressing uncertainties, *J. Environ. Manage.*, *159*, 158–168, doi:10.1016/j.jenvman.2015.04.044.

North, E. W., E. E. Adams, Z. Schlag, C. R. Sherwood, R. He, K. H. Hyun, and S. A. Socolofsky (2011), Simulating oil droplet dispersal from the deepwater horizon spill with a Lagrangian approach, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 217–226, AGU, Washington, D. C., doi:10.1029/2011GM001102.

North, E. W., E. E. Adams, A. E. Thessen, Z. Schlag, R. He, S. A. Socolofsky, S. M. Masutani, and S. D. Peckham (2015), The influence of droplet size and biodegradation on the transport of subsurface oil droplets during the Deepwater Horizon spill: A model sensitivity study, *Environ. Res. Lett.*, *10*(2), 024016, doi:10.1088/1748-9326/10/2/024016.

Poje, A. C., et al. (2014), Submesoscale dispersion in the vicinity of the Deepwater Horizon spill, *Proc. Natl. Acad. Sci. U. S. A.*, *111*, 12,693–12,698, doi:10.1073/pnas.1402452111.

Proctor, R., A. J. Elliot, and R. A. Fiather (1994), Forecast and hindcast simulations of the Braer oil spill, *Mar. Pollut. Bull.*, *28*(4), 219–229, doi:10.1016/0025-326X(94)90097-3.

Sebastião, P., and C. Guedes Soares (2007), Uncertainty in predictions of oil spill trajectories in open sea, *Ocean Eng.*, *34*(3-4), 576–584, doi:10.1016/j.oceaneng.2006.01.014.

Socolofsky, S. A., and E. E. Adams (2002), Multi-phase plumes in uniform and stratified crossflow, *J. Hydraul. Res.*, *40*(6), 661–672, doi:10.1080/00221680209499913.

Socolofsky, S. A., E. E. Adams, and C. R. Sherwood (2011), Formation dynamics of subsurface hydrocarbon intrusions following the Deepwater Horizon blowout, *Geophys. Res. Lett.*, *38*, L09602, doi:10.1029/2011GL047174.

Spaulding, M., P. Bishnoi, E. Anderson, and T. Isaji (2000), An integrated model for prediction of oil transport from a deep water blowout, in *23rd AMOP Technical Seminar 2000*, pp. 611–635, Vancouver, B. C., Canada.

Streett, D. (2011), NOAA s satellite monitoring of marine oil, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 9–18, AGU, Washington, D. C., doi:10.1029/2011GM001104.

Thacker, W. C., A. Srinivasan, M. Iskandarani, O. M. Knio, and M. Le Hénaff (2012), Propagating boundary uncertainties using polynomial expansions, *Ocean Modell.*, *43-44*, 52–63, doi:10.1016/j.ocemod.2011.11.011.

Thacker, W. C., M. Iskandarani, R. C. Gonçalves, A. Srinivasan, and O. M. Knio (2015), Pragmatic aspects of uncertainty propagation: A conceptual review, *Ocean Modell.*, *95*, 25–36, doi:10.1016/j.ocemod.2015.09.001.

Tkalich, P., K. Huda, and K. Y. Hoong Gin (2003), A multiphase oil spill model, *J. Hydraul. Res.*, *41*(2), 115–125, doi:10.1080/00221680309499955.

Walker, N. D., et al. (2011), Impacts of loop current frontal cyclonic eddies and wind forcing on the 2010 Gulf of Mexico oil spill, in *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophys. Monogr. Ser.*, vol. 195, pp. 103–116, AGU, Washington, D. C., doi:10.1029/2011GM001120.

Winokur, J., P. Conrad, I. Sraj, O. Knio, A. Srinivasan, W. C. Thacker, Y. Marzouk, and M. Iskandarani (2013), A priori testing of sparse adaptive polynomial chaos expansions using an ocean general circulation model database, *Comput. Geosci.*, *17*(6), 899–911, doi:10.1007/s10596-013-9361-3.

Yapa, P. D., and F. Chen (2004), Behavior of oil and gas from deepwater blowouts, *J. Hydraul. Eng.*, *130*(6), 540–553, doi:10.1061/(ASCE)0733-9429(2004)130:6(540).

Yapa, P. D., M. R. Wimalaratne, A. L. Dissanayake, and J. A. DeGraff (2012), How does oil and gas behave when released in deepwater?, *J. Hydro-Environ. Res.*, *6*(4), 275–285, doi:10.1016/j.jher.2012.05.002.

Zhao, L., M. C. Boufadel, S. A. Socolofsky, E. Adams, T. King, and K. Lee (2014), Evolution of droplets in subsea oil and gas blowouts: Development and validation of the numerical model VDROP-J, *Mar. Pollut. Bull.*, *83*(1), 58–69, doi:10.1016/j.marpolbul.2014.04.020.

Zheng, L., and P. D. Yapa (2000), Buoyant velocity of spherical and nonspherical bubbles/droplets, *J. Hydraul. Eng.*, *126*(11), 852–854, doi:10.1061/(ASCE)0733-9429(2000)126:11(852).

Zheng, L., P. D. Yapa, and F. Chen (2003), A model for simulating deepwater oil and gas blowouts—Part I: Theory and model formulation, *J. Hydraul. Res.*, *41*(4), 339–351, doi:10.1080/00221680309499980.