Ocean Modelling 43–44 (2012) 52–63

Contents lists available at SciVerse ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Propagating boundary uncertainties using polynomial expansions

W.C. Thacker [a,b], A. Srinivasan [c,d,*], M. Iskandarani [d], O.M. Knio [e,1], M. Le Hénaff [a]

[a] Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, United States
[b] Atlantic Oceanographic and Meteorological Laboratory, Miami, FL 33149, United States
[c] Center for Computational Science, University of Miami, Miami, FL 33149, United States
[d] MPO Division, Rosensteil School for Marine and Atmospheric Sciences, University of Miami, Miami, FL 33149, United States
[e] Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218, United States

ARTICLE INFO

Article history:
Received 13 June 2011
Received in revised form 25 November 2011
Accepted 27 November 2011
Available online 10 December 2011

Keywords:
Polynomial chaos
Ocean modeling
Error propagation
Uncertainty quantification
Spectral stochastic expansions
Gulf of Mexico
Loop Current
HYCOM

ABSTRACT

The method of polynomial chaos expansions is illustrated by showing how uncertainties in boundary conditions specifying the flow from the Caribbean Sea into the Gulf of Mexico manifest as uncertainties in a model's simulation of the Gulf's surface elevation field. The method, which has been used for a variety of engineering applications, is explained within an oceanographic context and its advantages and disadvantages are discussed. The method's utility requires that the spatially and temporally varying uncertainties of the inflow be characterized by a small number of independent random variables, which here correspond to amplitudes of spatiotemporal modes inferred from an available boundary climatology.

1. Introduction

The object of this paper is to point out how uncertainties of oceanographic simulations might be explored using the method of polynomial chaos expansions. This method was first introduced by Wiener (1938), who addressed the question of efficiently estimating uncertainties of a dynamical simulation stemming from uncertainties in its defining parameters. He realized that, in principle, a probability density describing the uncertainty of the parameters might be propagated dynamically to provide distributional information about any aspects of the simulation, but there was the issue of how to do it in practice. By using polynomial expansions to express the simulation's dependence on the uncertain parameters, he reduced the problem of propagating uncertainties to the task of determining expansion coefficients. The phrase "polynomial chaos", which has become popular in the engineering literature, stems from Wiener's referring to uncertainty as "chaos" and from his use of a polynomial expansion.[2,3] When the outputs of a simulation are well-approximated by polynomials of the inputs, polynomial expansions are appropriate, but when they are not, the expansions may converge slowly or may not converge at all.[4] The "chaos" part of the method relates to the choice of the polynomial basis: as the probability density function describing the uncertainty of the inputs appears in all expectation integrals, it is best to choose polynomials that are orthogonal when weighted by that density.

The method certainly should be of interest, as oceanographic simulations have many uncertain inputs.[5] For example, they depend on initial values of temperature, salinity, and other state variables at each point within the model's domain, on temporally varying values characterizing forcing fluxes everywhere on the

* Corresponding author at: Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, United States.

E-mail addresses: w.thacker@miami.edu, carlisle.thacker@noaa.gov (W.C. Thacker), asrinivasan@rsmas.miami.edu (A. Srinivasan), miskandarani@rsmas.miami.edu (M. Iskandarani), knio@jhu.edu (O.M. Knio), mlehenaff@rsmas.miami.edu (M. Le Hénaff).

[1] Present address: Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, United States.

[2] Chaos within this context should not be confused with its more modern usage to indicate sensitivity to small perturbations (Lorenz, 1963).

[3] For an introduction to the engineering literature see the reviews by Xiu (2009) and Najm (2009).

[4] While the Cameron–Martin theorem (Cameron and Martin, 1948) guarantees convergence for any finite variance process, in practice convergence is tested by checking the impact of retaining more terms in the expansion.

[5] Other approaches to oceanographic uncertainty can be found in the books of Bennett (2002), Evensen (2009), and Wunsch (2006). For discussions of uncertainty in fields other than oceanography, see the article in the special issue of Journal of Computational Physics (Karniadakis and Glimm, 2006) in which Lermusiaux (2006) presents his view of oceanographic uncertainties to a wider audience.

1463-5003/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2011.11.011

air-sea boundary, on values used for a variety of transport coefficients, and when there are open lateral boundaries on the details of their specification. Quantitative information about the impacts of their mis-specification could be quite valuable. Not only would it reveal the limitations of the utility of a simulation, it would also suggest which inputs must be better known to achieve a more useful simulation. It is important to recognize that the method of polynomial chaos expansions, like all methods for dealing with uncertainty, suffers from what Bellman (1957) called the "curse of dimensionality", namely the inescapable fact that computational complexity increases geometrically with increasing numbers of uncertain parameters. Thus, in practice, the method is used to examine the consequences of a limited number of uncertain inputs.

As Kalman filtering (e.g. Evensen, 2009) is better known to oceanographers, especially within the context of data assimilation where its role is to characterize the dynamically evolving uncertainties of the model state, comparing it with the method of polynomial chaos expansions can be instructive. The Kalman filter owes much of its utility to its characterization of the uncertainties using only an evolving mean state and an evolving matrix of covariances characterizing the state's uncertainty. The curse of dimensionality manifests in the size of the error-covariance matrix, which is unmanageably large, so much effort has been devoted to its approximation. For example, the ensemble Kalman filter approximates it using covariances inferred from a manageable number of simulations chosen to sample important aspects of the state's uncertainty. The method of polynomial chaos expansions as illustrated here also uses an ensemble of simulations to characterize the input uncertainties. However, the purpose of the ensemble is to provide quadrature information needed for evaluating the expansion coefficients, so the ensemble members are chosen to optimize the accuracy of the coefficients. The resulting expansions provide not just means and covariances but provide complete distributional information about the model's outputs.

It is also useful to note that Monte Carlo methods (e.g. Gilks et al., 1996), which also seek general distributional information about outputs, generally require a much larger ensemble of simulations to achieve the same accuracy that might be obtained from polynomial chaos expansions with a small quadrature ensemble. Polynomial interpolation between simulations in effect provides additional implicit sampling. While large Monte Carlo ensembles are unachievable for computationally intensive simulations, smaller quadrature ensembles might be affordable using today's computational resources.

If alternative choices for the uncertain parameters are regarded as perturbations of the favorite choice, then this method might be regarded as a perturbation method. However, as there is no requirement that the perturbations be small, the method of polynomial chaos expansions can accommodate information about large but unlikely perturbations. Within the context of automatic differentiation, propagation of infinitesimal perturbations is accomplished using the forward method and tangent-linear codes for accomplishing this can be generated automatically, but unfortunately they have to be run once for each perturbed input (e.g. Griewank and Corliss, 1991). On the other hand, sensitivities of a single output to infinitesimal perturbations of all uncertain inputs can be computed with automatically generated codes that implement the reverse or adjoint method.[6]

To illustrate the method of polynomial chaos expansions, we examine how uncertainties in the inflow through the Yucatan Straits manifest in the Gulf of Mexico's surface-elevation field and in the behavior of the Loop Current. Because of the Gulf's semi-enclosed geography with the Loop Current being the principal dynamical feature, we thought that the consequences of mis-specifying the inflow should be interesting. Our challenge was to find a way to reduce the uncertainties of the spatially and temporally varying inflow to a few parameters, as we could find no published example of a similar problem. As the circulation in the Gulf is simulated using a high-resolution numerical model, the major computational expense is the ensemble of simulations needed to evaluate the coefficients of the polynomial expansions; the cost of evaluating the coefficients and using them to examine the output uncertainties is trivial in comparison.

Section 2 describes the methodology. After describing the numerical model used to simulate the Gulf's circulation, Section 3 explains our approach to reducing the inflow uncertainties to two random parameters. Section 4 discusses how the expansions are truncated and the ensemble of simulations needed for evaluating the coefficients of the polynomials. Then Section 5 presents the mean and standard deviation of the surface elevation field resulting from assumed distribution of possible boundary conditions and discusses surface-elevation covariances. By showing probability densities characterizing the non-Gaussian nature of the model's response, Section 6 illustrates how the polynomial expansions can be used to emulate the numerical model. And section 7 examines the convergence of the polynomial expansion. Finally, Section 8 concludes with comments about what the method might offer for oceanographic applications.

2. The methodology

The objective of the method is to assess how uncertainties of inputs of a dynamical system manifest in its outputs. To see how it works, consider the simple case of only a single uncertain input x, as generalization to two or more is relatively straightforward.[7] To express its uncertainty quantitatively, x can be expressed in terms of a central value x_0, which when not accounting for uncertainty would be used as input, and a spread x_1 characterizing the likely range of values around x_0:

$$x = x_0 + x_1\xi, \tag{1}$$

where ξ is a standardized random variable with probability density function $p(\xi)$.[8] For most problems we might have some idea what values to use for x_0 and x_1, but there may be little empirical basis for our choice of $p(\xi)$. When there are no fixed bounds on the range of x, the probability density might be taken as Gaussian. That was in fact the choice made by Wiener (1938), and that will also be ours, but other, possibly empirical, densities might be used.

Again for simplicity it is useful to focus on a single output $y = y(\xi)$, which might be thought of as the surface elevation at a particular space–time point.[9] The method centers on the assumption that output y can be efficiently described by series of polynomi-

[6] Adjoint codes are typically used to compute the gradient of a cost function for use in algorithms seeking to optimize the choice of a model's uncertain input parameters.

[7] When there is more than one uncertain parameter of interest, x in Eq. (1) becomes a vector, as do x_0 and ξ, while x_1 becomes a matrix.

[8] When constructing software that might be used for a variety of applications, it is useful to standardize ξ so that it has zero for its central value and a spread of unity.

[9] Another approach to polynomial chaos expansion (e.g. Knio and Le Maître, 2006; Le Maître and Knio, 2010) does require that the uncertainty of *all* evolving state variables be computed. In that case the polynomial chaos expansions for all state variables, each similar to Eq. (2), are inserted into the dynamical equations and the condition that the residuals be small in a statistical sense produces a system of equations for the expansion coefficients similar to but more complicated than the original dynamical system. As this would require software at least as demanding to construct as that already existing for the numerical model, this option was not considered for this study. Finette (2006) has proposed this approach for studying uncertainties of underwater acoustics, Ge et al. (2008) for nonlinear shallow-water equations, and Shen et al. (2010) for the Lorenz (1984) model. Somewhat similarly, Sapsis and Lermusiaux (2009) have suggested using a temporal evolving set of basis functions rather than a fixed polynomial basis.