Research (2013)

Juan M. Restrepo

Contents

1 Brief Description of Research Goals for the Next 5 Years

1.1 Contour Dynamics and Lagrangian Data Assimilation

The primary goal of the proposed study is to develop new feature-based estimation techniques. As opposed to variance minimizers, these are specifically designed to capture such things as clear tracks from a background and where it is crucial to keep these features sharp, even if this is at the expense of physical consistency of *all* dynamic variables. Two such problems are the tracking of hurricanes, or the determination of sharp heat-transport boundaries in the gulf stream, oil spills, and tracking of lost objects and people at sea. Secondly, we add a we focus on the development of new assimilation techniques that are capable of handling buoy/float/drifter data, which at present and in the foreseeable future are sources for significant amounts of data. Their trajectories sample large regions delivering position, velocity and acceleration information about ocean dynamics and the shape of dynamical features. The assimilation techniques must handle this Lagrangian data as well as data that comes from fixed grids (*e.g.*, general circulation model output). The data fusing must be able to handle the Eulerian and Lagrangian information, as well as, highly nonlinear dynamics and non-Gaussian statistics for large state spaces. Practical assimilation schemes will need to exploit dimensional reduction and relaxation of optimality where possible.

For oceanic and atmospheric flows containing coherent features with strong dynamical behavior, such as the Gulf Stream current or hurricanes, a variance-minimizing estimator may be inappropriate since the inherent smoothing deteriorates energetic dynamical features. We propose to develop a new feature-based assimilation algorithm that will do substantially better in fusing data from flows with coherent structures. The key tasks in this proposal are: (1) develop a nonlinear/non-Gaussian estimation strategy, capable of handling moderate-sized dimensions; (2) formulate and implement a feature-based assimilation strategy that uses the estimation algorithm; (3) compute improved estimates of the velocity field and its first few moments, compute the Reynolds stresses and divergence. We will make use of prototypical benchmark flows we are very familiar with, progressing to the use of state of the art data assimilative models (HYCOM), as well as 30 years of near-surface float/drifter data to build a non-Linear/non-Gaussian stochastic model for the Lagrangain uncertainty inherent in the large scale dynamics of the Gulf Stream.

1.2 Calculation of Trends for Non-Statistical Multi-scale Signals

Calculating trends of multi scale signals is a very common task. Methods that rely on statistical stationarity and/or parametric descriptions are commonly used. However, there is plenty of problems where a trend is needed, yet statistical conditions do not exist. How common is this problem? consider calculating the temperature trend in London over the last 100 years. We have developed a notion of a trend and a method for calculating it that does not rely on statistical or deterministic assumptions. In the next few years we want to try it out and refine it, as it presently is not optimal in any sense.

1.3 Breaking Waves, Dissipative Processes, and Stochastic Parametrization

We applied stochastic parametrization in order to capture the important wave-dissipation that is due to white capping, in [1]. The idea is to forgo a physically-based model for the phenomenon (clearly, whitecapping is beyond analytical treatment) in favor for a stochastic paremetrization, in which white-capping is re-interpreted as a stochastic process that creates uncertainty fluctuations of Lagrangian paths of fluid parcels. We then used the Lagrangian-to-Eulerian frame filtering to produce macroscopic expressions for the resulting dissipation. Once a stochastic model is proposed the problem is recast as one of pinning down the parameters of the first few moments of the process, which can be done by using field data as well as models.

We have submitted a proposal to study the Lagrangian motion of water parcels under the action of breaking waves in a laboratory setting. The idea is to develop the correct stochastic parametrization for wave breaking. This is work to be done with Ken Melville (Scripps Institute of Oceanography).

1.4 BP-Sponsored Research: Large-Scale Models for Oil Spills

We are developing a computational framework for oil-spill dynamics and the fate of hydrocarbons in the ocean. The model is being incorporated into a ROMS-style ocean model with data assimilation capabilities. There are a number of novel aspects to the model being developed: in addition to being built numerically as a discontinuous Galerkin, adaptive mesh code, capable of resolutions between 10s of meters to 100s of km, the model is also capable of resolving temporal scales of the order of minutes to seasons. Moreover, smaller scales are parametrized: both waves and wave breaking are properly parametrized into the large scales.

We are also studying the nearshore flow from a large-scale simulation perspective: the challenge is to develop the correct parametrization for nearshore mixing and dispersion. The fundamental issue is that if the velocity at the nearshore is zero, there is no way to get oil to reach the shore (unless you put into the model a rule that drops pollutants there if they reach this neighborhood). The fundamental mechanism that gets oil onto the beach, when finite resolutions are used, is diffusion. Hence models for nearshore dispersion need to be formulated. Our dispersion model includes Langmuir turbulence in the deeper reaches of the should and wave breaking and boundary layer dispersion in the nearshore.

1.5 Sticky Waters

Wind and current-driven flotsam, oil spills, pollutants, and nutrients, approaching the nearshore will often appear to park just beyond the break zone, where waves break. We refer to the parking phenomenon in this environment as nearshore sticky waters. See Figure 1 The portion of these tracers that beach will do so only after a long time. Explaining why these tracers park has important implications on a variety of different nearshore environmental issues and on what subscale processes are important to include in computer models for the simulation of pollutants in the nearshore. We can provide an explanation for the underlying mechanism responsible for the the parking of flotsam and for the physics responsible for the long time it takes a portion of these tracers to reach the shore. Namely, the parking is a natural consequence of the incoming current, via wave residual flows, and a counter current we call the anti-Stokes. The pollutant however, reaches the nearshore because it only floats on a partial layer of the ocean, so even if the velocity is depth-averaged to zero,

Figure 1: *Algal bloom, parked beyond the breakzone.*

there is still shore-ward transport. However, once it reaches a situation in which dissipative effects dominate advective effects, the pollutant reaches the shore via pure diffusion, which is anomalous because of increased dispersion due to breaking and boundary layer effects.

1.6 Oceanic Transport and Its Implications on the Snow Ball Earth

The Budyko/Sellers model [2; 3], and its variants have been used to make sense of terrestrial paleoclimate and the influence of fluctuations of sun radiation. Among other things, this simple balance model is often invoked in order to understand the degree to which the Earth has been covered by a sheet of ice. Recent attention to extensions of this model have also been used to elucidate whether globally-induced warming, which induces a considerable amount of ice melting, is reversible or not (see [4]). A common modeling assumption in these models is that the ocean transport is considered globally balanced. Oceanic transport, as we now know it to be, is highly heterogeneous and does not have to be in a steady state; furthermore, it is one of the most significant mechanisms for Earth's climate balance, after the sun and albedo effects. We intend to revisit the original Budyko/Sellers scenario but will include active oceanic transport.

1.7 Computational Hydrology on Multifractal Subsurface Processes

In collaboration with Prof. Shlomo Neuman (UA) and my PhD student Darin Comeau. The goal of the proposed work is to develop three fundamental and inter-related building blocks crucial to pollutant-source identification simulation strategies, which take as input field measurements of the sub-surface conductivity, the contaminant, and the head, as well as model input for the same. The three aspects are: subsurface characterization of heterogenous media, estimation methodologies suited to nonlinear/non-Gaussian model/data estimation of time dependent problems, the numerical simulation engine that allows one to compute model multi-scaled projections on coarsely-resolved, high gradient field situations efficiently and robustly.

The strategy to be developed with regard to subsurface characterization is natural for a computational setting: it involves a multi-resolution analysis, motivated by the fact that the permeability variogram has incremental stationarity that can be exploited. The result is a nested subspace projection of the original problem into hydrogeological subproblems in which a mean field and a fluctuating field decomposition make mathematical sense. The projection strategy leads to a numerical challenge with regard to simulating the field equations efficiently and robustly. In this regard we are not developing an innovation, but rather, adapting the (nonlinear) discontinuous

Galerkin strategies to the nested assimilation problem. The data assimilation process is designed to deliver predictions/retrodictions of the mean history and the uncertainty of the flow and a contaminant, using available ground permeability measurements, as well as any measurements of the contaminant and flow; moreover, the methodology handles inherent measurement and model errors. Here we will develop new assimilation methodologies appropriate to nonlinear/non-Gaussian time dependent estimation.

In the end we will deliver methodologies that can be applied in source-contaminant identification and contaminant transport which: (1) reduce the uncertainty in subsurface characterization; (2) can handle non-linear, non-Gaussian blending of data and models outcomes, taking into account the uncertainty in the model itself as well as in the data; (3) exploit the robust qualities of the discontinuous nonlinear Galerkin procedure leading to large-scale applications.

1.8 Rip Surges

Randy LeVeque (Washington), Jim McWilliams and I are working on a newly-discovered form of a rip current, which I call the *Rip Surge*. This type of rip surge is new because it has only been possible to see it by the improved efforts of LeVeque's group and ours in creating a numerical solver for shallow water waves that includes a moving shoreface. In this type of rip surge, currents traveling back to the shelf are generated by the action of an excess water mass that accumulates at the shore end of a smooth-sloping beach, subjected to time-periodic waves. It is a residual flow, which when it interacts with the waves, will be guided to follow the stationary nodal lines of the waves, and develop a Kelvin-Helmholtz instability due to excess shearing.

2 Introduction

I am the group leader of the Uncertainty Quantification Group (UQG), at The University of Arizona. I formed this group in 2006 to pursue research on problems in which significant progress will come from the combined statistical and deterministic approaches (rather than from one or the other). The group is comprised of faculty (15), post-docs (3), graduate students (14), and undergraduates (2), from mathematics, statistics, computer science, hydrology, atmospheric sciences, and physics. The three general areas of research are climate/meteorology, hydrology, and computer vision. The very existence of the group, its working dynamics, and its outlook is not accidentally similar to groups in Department of Energy labs: I have a long association with DOE, since my graduate student years. I was an ORISE post-doc and a frequent summer visitor as junior faculty in the Mathematics and Computer Science Division at ANL, where I learned advanced scientific computing, got involved in automatic differentiation and developed wavelet-Galerkin techniques. I also got involved in the CHAMMP effort and this lead to my work in climatology (especially after a short stint at UCLA where a long-lasting collaboration there flourished and still exists today). As a faculty member at the University of Arizona I have spent many summers working with the T7 group (now T5) at LANL, first in biomathematics and fluid mechanics. Presently, I work with the Statistics Division at LANL on sensitivity analysis and estimation theory.

With UQG I brought to an academic setting that interdisciplinary and scientific computing-based approach which has always been a scientific strength of DOE labs. I am presently involved in trying keep alive DOE's connections to academia: I have a proposal for a computational geo-science/energy program that is not only topical and of interest to DOE's mission and longevity, it can only be pulled off if DOE teams up with academia for a mutually beneficial outcome, both in terms of research and training of future DOE scientists.

For more information on UQG, visit our group's web site:

`http://www.physics.arizona.edu/~restrepo/UQ/UQ.html`.

3 Climate Science and Uncertainty Quantification

Climate has very large spatio-temporal scales. To wit, consider the scales glacial/inter-glacial dynamics, El Niño/Southern Oscillation, industrialized-global warming, to name but a few of climatic phenomena. Arguably, the toughest challenge in climate science arises from the lack of high quality, statistically constrained data, with the large spatio-temporal scales of climate itself. More data is being collected but it is clear that significant progress can only be made when data and models are used.

Making improvements on the physics of general circulation models and on coupling oceans, atmosphere, land, biota, and energy models is thus amply justified. So are the efforts on improving the computational efficiency and accuracy of these models. Highly successful groups working on the physics or computation tend to be large enterprises with many years of experience.

Uncertainty quantification as applied to climate and energy issues is now fully recognized by the IPCC as one of the most important aspects of climate change research. Meteorology has been fruitfully using "data assimilation" techniques for many decades now, obtaining considerable improvements in forecasting (so long as the weather does not change considerably).

Climate retrodiction, filtering, and forecasting are borrowing some of these techniques in order to pin down climate variability. Climate is a very high-dimensional nonlinear problem, with data that is sparse and notoriously unconstrained with regard to uncertainties. The statistics are seldom Gaussian. Data assimilation, or the use of models and data with careful handling of inherent errors and uncertainties in the physics and measurement errors, is not a well-developed science: we do not know how to perform high dimensional nonlinear non-Gaussian estimation; we have not fully explored large deviation theory for extreme events, we have just recently found interest in forward and backward sensitivity in order to assess model variability and uncertainties; we have found that variance minimizers are not uniquely appropriate general estimators. We have recognized ensemble statistics as being very useful but have not figured out a way to collect these on high dimensional problems.

4 Summary, Present Research

The three fundamental research questions in UQ have thus far been to develop new and efficient methods for data assimilation, capable of handling nonlinear/non-Gaussian models and data, the derivation of a methodology for the computation of trends that does not make use of statistical assumptions or a deterministic model, and the development of ensemble-based methods for forward sensitivity analysis.

In climate dynamics my focus has been on fundamental *modeling* of transport in oceans and the formulation of multiscale methodologies that circumvent the enormous spatio-temporal scales required to capture the dynamics of waves and currents.

In erosional dynamics, I have calculated *ab-initio* the forces on sand particles, and I have performed experiments on steady shearing flows over lose sedimentary matter. I have also been involved in modeling the process of formation and maintenance of lare sandridges in the shallow reaches of the continental shelf.

If I had to mention specific scientific triumphs, they would be:

- The derivation of a comprehensive and general multiscale model for wave/current interactions in three-space and time at global and shelf scales, with which we clarified important oceanic interactions and showed how and why waves matter at large spatio-temporal scales. The filtering technique yields equations that are set at the coarsest scales, thus circumventing the need for capturing the fine scale dynamics.
- The success of stochastic parametrization as applied to wave breaking, an important source of dissipation for waves and currents.
- The proposal of a variety of data assimilation techniques (3) to handle the nonlinear/non-Gaussian nature of data assimilation for problems not amenable by more traditional means (such as Lagrangian data assimilation). Of these the Diffusion Kernel Filter can be made competitive computationally, compared to the extended Kalman Filter.
- The creation of a new technique for forward sensitivity analysis which has superior computational characteristics as compared to Bred Vectors or finite-time Lyapunov vectors.
- The use of computational techniques to produce state-of-the-art experimental data on lift and drag on particles in flows.
- The formulation of a way to circumvent storage limitations in variational data assimilation.
- Determining whether advective effects could cause the meridional overturning circulation and the thermohaline circulation stability to change; determining whether convective adjustment algorithms are responsible for some of the decadal variability observed in ocean general circulation experiments; determining the extent to which models, rather than physics, determine the extent to which the Southern Hemisphere is critical in the carbon balance during glacial/interglacial times, and the importance of chemical changes in the ocean to consistently account for the observed pCO_2 fluctuations.

An enumeration of present projects:

- A newly proposed method for computing trends that is capable of handling multiscale data without the use of parametric assumptions. It is presently being used to analyse global warming signals.
- A new theoretical basis for the "snow-ball" Earth. The snowball earth is an hypothesis of the state of the Earth some 750 million years ago, which is an important event withing evolutionary biology. Our project seeks to eke out the role played by ocean transport in the evolution of the global climate of this geological era.
- A theoretical description of the phenomenon of "sticky waters" which describes the phenomenon of partial parking of flotsam, oil, nutrients, and pollutants just beyond the break zone.

- An explanation of the role of convergences and divergences of flows that will lead to periodic fluctuations of oil slicks trapped in longshore currents.
- The use of Bayesian estimation methods for the purpose of pinning down critical model parameters for nearshore models, based upon observational data.
- As part of the $112M grant our CARTHE group receives, I am responsible for the creation of a new oil model for use in the Gulf Coast and other near coastal areas. The model is a fully depth-averaged model that includes disperion due to Langmuir turbulence. It properly includes wave, current, and wind forcing. The model will track both floating oil as well as suspended oil. The model is being incorporated into a working wave/current interaction model called ADCIRC. ADCIRC is fully parallelized, adaptive mesh code with data assimilation capabilities. It is presently being transformed into a discontinuous Galerkin formulation.
- Work on the use of Bayesian estimation techniques and polynomial chaos to geo-chronology of sedimentary processes. The goal is to back out an optimal functional dependence of the age and the elevation probability distributions of field data.
- We have just submitted a proposal to work on the experimental determination of the stochastic model for wave breaking. Work on this project depends on funding.

4.1 Uncertainty Quantification

4.1.1 Nonlinear Non-Gaussian Data Assimilation

The aim of data assimilation is to infer the state of a system from a geophysical model and possibly incomplete or non-uniformly distributed spatio-temporal observational data. Used extensively in engineering control theory applications, data assimilation has relatively recently been introduced into meteorological forecasting and climate dynamics. Another example is the problem of recovering the release history of a groundwater contaminant, such as a chemical spill, when the porous medium and the hydrological flow-field are only statistically characterized.

The Mean Field Approach: Our main thrust was to develop data assimilation techniques suited to problems with nonlinear and possibly far from Gaussian statistical distributions. Greg Eyink (Johns Hopkins) and I proposed a variational method to calculate such conditional statistics in a computationally efficient way. There are two main elements of our approach: (i) a new cost function, the effective action, and (ii) a variational Rayleigh-Ritz method to calculate it.

The *effective action* is a nonnegative, globally convex functional of time-histories of the variable to be estimated. Its intuitive significance is that it gives the "cost" for a particular history to occur as a fluctuation in an average over an N-sample ensemble of independently prepared systems. It has been shown (G. Eyink, unpublished) that the minimizer of the effective action subject to constraints approximates the ensemble means conditioned upon those constraints. The approximation involved is of "mean-field type", in that conditioning is on the N-sample averages rather than the individual realizations themselves. The Hessian matrix of the effective action gives furthermore an approximation of the covariance matrix of fluctuations around this optimal history, conditioned upon the same constraints. Higher-order cumulants of fluctuations can also be obtained from higher-order derivatives of the effective action.

What makes the effective action attractive as a means to approximate the required conditional statistics is the existence of a practical variational scheme to calculate it. This *Rayleigh-Ritz*

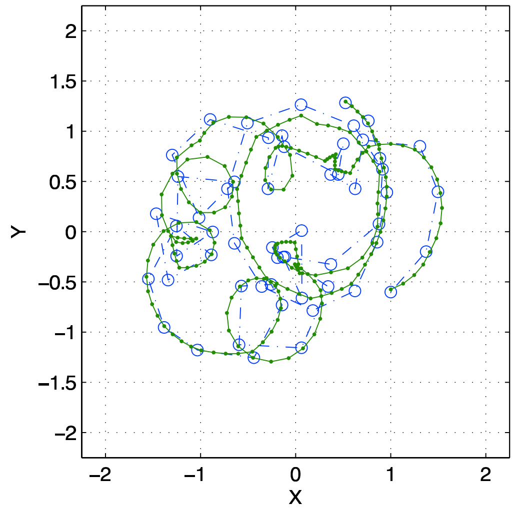


Figure 2: Noisy data (circles); assimilated data with errors in the model of order 20%.

method was proposed earlier by us and its detailed development initiated in a series of works. Its basis is a mathematical characterization of the effective action through a constrained variation of a dynamical action, which is a functional on the joint space of random variables (observables) and probability distributions (states). The Rayleigh-Ritz method is essentially an extension of the standard method of moment-closure approximation, already widely used in the theory of multiscale statistical dynamics, e.g. turbulence and phase-ordering processes. The technique is known as the Mean Field Variational Method [5], [6], [7].

The Path Integral Method: One problem that the Mean Field Approach has (in addition to not being suitable for large problems) is that it requires a great deal of effort in its implementation. It is, nevertheless, useful in benchmarking less optimal methods. We are currently developing Bayesian ensemble approaches: these will readily accept an existing model (legacy code) and will operate as a black box. A promising approach is the Path Integral Method. The basis is to discretize the stochastic differential equations and use these discretizations in the action functional, along with data, in order to propose a general Bayesian-based functional. It is general, fully nonlinear/non-Gaussian. The problem is now shifted, however, to finding fast samplers. The Hybrid Monte Carlo sampling technique is a good basis, but it is clear that it is unsuitable for high dimensional problems. At present we are exploring ways to use the Riemann metric as a way to produce a preconditioner to the molecular dynamics problem in order to achieve a significant speedup.

Applied to a problem in Lagrangian data assimilation Figure 2 conveys the methods' superior assimilation capabilities. In the Figure the data appears as circules, the model for the path of an unobserved tracer (dots) is kept true to within statistical uncertainty by observations of vortices (stars) that generate the flow. This work is reported in [8] and [9].

Particle Filter Resampling: In [10] we propose and evaluate several particle filter resampling parametrizations in order to stave off the particle filter collapse.

The Diffusion Kernel Filter: Techniques such as those proposed by Kushner [11], particle filter methods [12], and particle resampling [10] near-optimal methods such as the Mean Field Variational [5; 6], and Path Integral [9], the Langevin sampler [13], to name but a few, are appropriate to highly

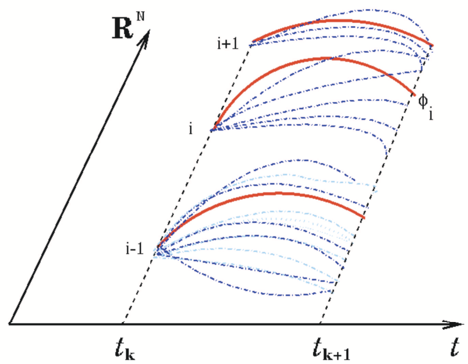


Figure 3: $\mathbf{\Phi}_j(\xi(t_k,i),t) = \phi(\xi(t_k,i),t) + \mathbf{\Phi}'_j(\xi(t_k,i),t),\ j = 1,...,J_{k,i}$. *Branches of prediction labeled* $i = 1,...,I_k$. *Deterministic trajectories* ϕ *in red, the dashed lines represent the samples.*

nonlinear/non-Gaussian problems, but are severely challenged dimensionally.

Recently, we developed a new filtering strategy called the Diffusion Kernel Filter (see [14; 15]). It is well suited for data fusion, and in particular, for Lagrangian data assimilation and trajectory prediction. It is a particle-filter based method (see [16]). It inherits from the particle filter the capability to handle nonlinear/non-Gaussian estimation problems. The parameterization of the noise process at every filtering step leads to sidestepping issues of particle filter sample degeneration. It is a sample-based method and does not rely on a variance-minimizing strategy, which thus allows us to consider a variety of other estimators, *e.g.*, maximum likelihood, average entropy (see [14]), etc. Like the Mori-Zwanzig projection, it converts the system of stochastic ordinary equations into a system of partial differential equations, however, it sidesteps the problematic issue of differentiating noise processes; the system remains nonlinear. We linearize along several *branches of prediction*, unlike the unique global linear update in Extended Kalman Filtering. We can then advance the tangent field using a deterministic time stepping scheme, which affords us several orders of magnitude larger time-stepping and more choices of time integrators. Without making a Gaussian assumption we capture the diffusion process parametrically, along branches of prediction.

Briefly, given a time dependent model for the state vector $\mathbf{x}(t)$ and data $\mathbf{y}(t_k)$, where t refers to continuous time and t_k are discrete times at which data is available the Diffusion Kernel Filter samples the posterior conditional probability of the N-dimension state vector $\mathbf{x}(t)$, subject to N_y-dimensional observations $\mathbf{y}_k$ at discrete times t_k, $k = 1,2,...$. For $t = t_k$, this conditional probability is $p(\mathbf{x}|\mathbf{y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y}_k)}$, which is obtained, sequentially, knowing the conditional probability at t_{k-1}. The fundamental strategy is to linearize along several *branches of prediction*. See Figure 3. Duhamel's principle in turn enables us to write down how the stochastic moment operators are projected along the branches. The second aspect of the procedure is to *parameterize* the diffusion process associated with *each* branch of prediction. Thus, along branch i of prediction the distribution consists of a deterministic trajectory ϕ_i and a fluctuating field ${\phi'}_i$. (Clearly, the overall distribution of the whole field may thus be far from Gaussian). These components are then used within a particle filter. We have found that the bootstrap particle filter (BF) with resampling gives good results. (See [17]. See [18] for a tutorial on particle filters). As can be inferred, the DKF is thus *fully consistent and naturally suited to the Lagrangian to Eulerian estimation problem, i.e., the method is developed using the same expansion of the map.*

Comparison of DKF Estimates to a Benchmark and Current Operational Assimila-

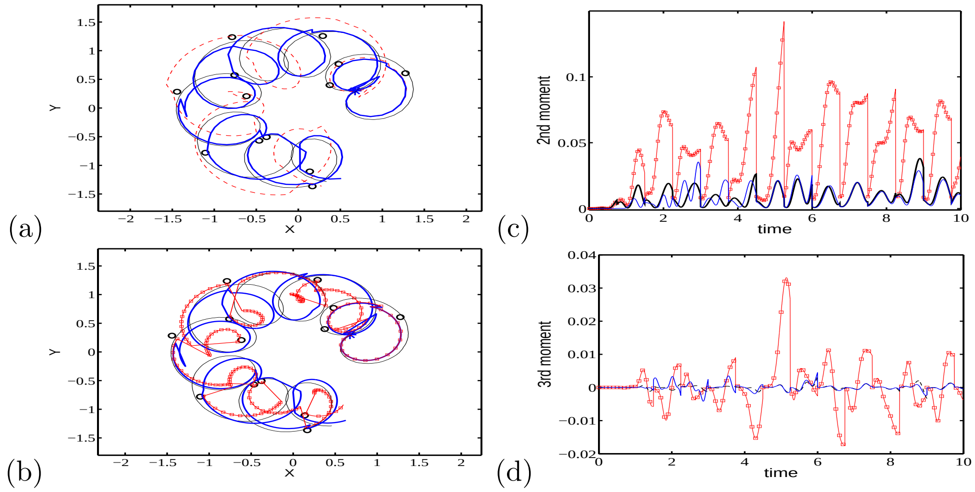


Figure 4: *Mean drifter path* (v_1, w_1) *estimated by DKF and BF (blue), EKF (red, dashed), enKF (red, squares), and the true path (black). (a) DKF, BF, EKF, truth, data (circles) (b) DKF, enKF, data (circles). (c) Second moment, (d) third moment.*

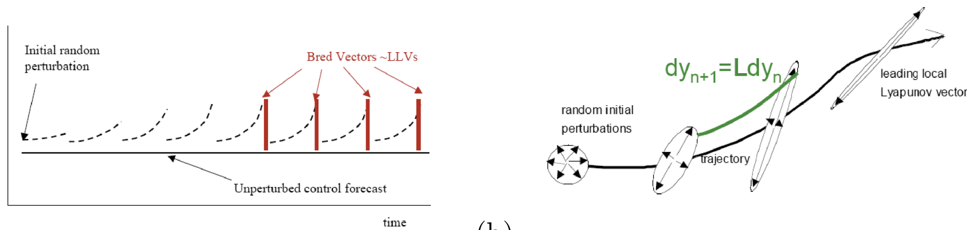
tion State-of-the Art Strategies: Kusnetzov *et al* [19] have proposed a methodology for predicting Lagrangian trajectories and for assimilating Lagrangian data based upon the well-known constrained extended Kalman filter (EKF). Their methodology was first applied to a fully Lagrangian-frame problem: the problem consists of a noisy, two-dimensional flow and passive tracers. Noisy observations are made of the position of one or more of the drifters and of the point vortices. The goal is to produce Lagrangian trajectory estimates, *i.e.* position of the vortices and drifters as a function of time, conditioned on observations. The problem is nonlinear and non-Gaussian. In Figure 4(a) we show that the benchmark particle filter estimate (BF) and the DKF are coincidental. EKF (dashed) is giving a poor estimate. The truth path is shown in black. In Figure 4(b) we highlight the estimate of the enKF (red) and DKF (blue). The enKF is also failing. In Figures 4(c) and (d) we show the second moment and third moment, respectively: BF and DKF blue, enKF (red). The enKF also fails to predict the second and third moment. The codes that produced these results were systematically tested for accuracy on a variety of other prediction problems, see [14] and [15]. This is work with P. Krause (formerly at UA).

4.1.2 Sensitivity Analysis

- **The Ensemble Bred Vector**: We proposed a variant of the Bred Vector (BV) algorithm, orginally introduced by Toth and Kalnay [20] to assess the sensitivity of model outputs to changes in initial conditions for weather forecasting. The new algorithm, which we call the *Ensemble Bred Vector* or EBV, is based on collective dynamics in an essential way. As such, it features distinctive qualities compared to the classical breeding algorithm. By construction, the EBV produces one or more dominant vectors, and is less prone to spurious results than the BV algorithm. It retains the attractive features of the BV with regard to being able to handle legacy codes with minimal additional coding. This work appears in [21]

  We often want to compute the finite-time forward sensitivity of solutions $y(t)$, obeying

$$\frac{dy}{dt} = G(y,t), \quad y(0) = y_0,$$



(a) (b)

Figure 5: (a) Schematic evolution of Bred Vectors; (b) evolution Lyapunov Vectors.

to perturbations of y_0. A computer implementation of such may be

$$Y_{n+1} = M(Y_n, t_n), \quad n = 0, 1, 2, .., \qquad Y_0 = Y(0).$$

where $Y(t_n) \in \mathbb{R}^K$ is an approximation of $y(t)$ at $t = t_n$. K is usually very large in weather and climate problems.

The Bred Vector Algorithm (BV) *BVs computed as follows:*

$$\begin{aligned} \delta Y_{n+1} &:= M(Y_n + \delta\mathcal{Y}_n, t_n) - M(Y_n, t_n), \\ \delta\mathcal{Y}_{n+1} &:= R\,\delta Y_{n+1}. \end{aligned}$$

$$R := \frac{\|\delta Y_0\|}{\|\delta Y_{n+1}\|}, \quad \textit{is the normalization factor.}$$

When the perturbation amplitude is small and the time step small, the BV resemble **the finite-time Lyapunov vectors:**

$$\delta y_{n+1} = L_n \delta y_n, \quad n = 0, 1, 2, ...$$

$$\text{where } L_n := L(Y_n, t_n) = \left.\frac{\partial G(y_n, t_n)}{\partial y}\right| \quad \textit{is the Tangent Linear Model.}$$

The schematic process appears in Figure 4.1.2. The Algorithm:

1. Pick a norm $\|\cdot\|$ and a perturbation magnitude ϵ.
2. Evolve $\mathbb{I}$ perturbations of magnitude ϵ that cover a sphere according to

$$\begin{aligned} \delta Y_{n+1}(\iota) &= M(Y_n + \delta\mathcal{Y}_n(\iota), t_n) - M(Y_n, t_n) \\ \delta\mathcal{Y}_{n+1}(\iota) &= R^{\min}_{n+1}\,\delta Y_{n+1}(\iota), \end{aligned}$$

   for all $\iota \in \mathbb{I}$, where $R^{\min}_{n+1} = \epsilon\,[\max_{\iota\in\mathbb{I}}(\|\delta Y_{n+1}(\iota)\|)]^{-1}$.
3. At n, the EBV is the $\mathbb{I}$ ensemble of $\delta\mathcal{Y}_n(\iota)$.

The crucial difference is that the ensemble is normalized at by the magnitude of the largest ensemble member.

EBV Properties

– EBV is consistent with BV and with finite-time Lyapunov vectors, for small time steps and perturbations

– Will yield a reduced representation of the perturbation field

– Inherently ensemble-based and thus capable of providing probabilistic information

– Not as prone to perturbation amplitude sensitivity as BV

– Not as prone to shape of perturbation dependence as BV

– Less prone to spurious outcomes than BV

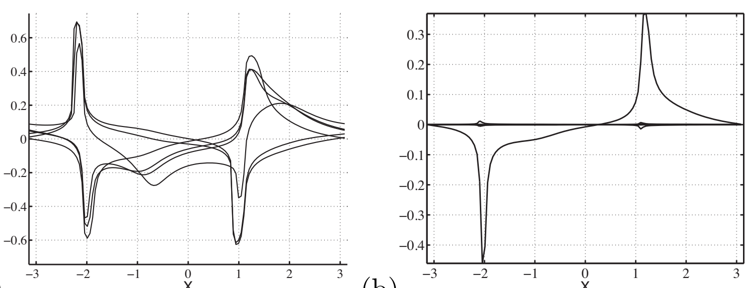
– Only trivially more expensive to compute than an ensemble of BV on moderate to large problems.

– Like BV, the EBV is also sensitive to the choice of norm. This is not the case for finite-time LV.

A comparison of BV and EBV clearly demonstrates the latter's superiority. Consider the forced, dissipative, nonlinear PDE with properties typical of climate models:

$$\frac{\partial S}{\partial t} = \alpha \frac{\partial^2}{\partial x^2}[f(x) + \mu S(S - \sin(x))^2 + S - \gamma \frac{\partial^2 S}{\partial x^2}], \quad S(x,0) = S_0(x).$$

Here $f(x)$ is a fixed forcing. The perturbation field of the initial conditions is compared in Figure 6. It demonstrates the unambiguous nature of the EBV results, compared to those from a BV run.



(a) (b)

Figure 6: (a) At $t = 70$ superposition of 6 BV simulations. (b) At $t = 70$, EBVs, clearly showing a dominant vector. The BV outcome is ambiguous with regard to yielding a dominant and reduced representation of the sensitivity of the initial conditions to perturbations.

– EBVs are superior to BVs and consistent with finite-time Lyapunov vectors.

– EBVs are less prone to sensitivity to size and shape of perturbations and thus significantly easier to interpret than BV (See Figure 6).

– EBVs are, by design, ensemble based and thus deliver probabilistic information.

– EBVs can be used to parameterize the background error covariance using a reduced representation (See Figure 6).

  - EBVs are trivially more expensive than BV computationally on small to moderately large problems.
  - BV and EBV are both sensitive to numerical implementation issues, the EBV significantly less so.

- **Circumventing Storage Limitations in the Exact Calculation of Gradients for Data Assimilation**:

  The problem we addressed is the computation of gradients/Jacobians on finite storage machines. The explosive growth in both on-line computer memory and remote storage requirements of large-scale assimilation studies can be overwhelming and this is a major hurdle in the use of variational methods for data assimilation. The storage problem imposes a severe physical limitation on the size of assimilation studies, even on the largest of computers. Using a recursive strategy, a schedule can be constructed which enables the forward/adjoint model runs to be performed in such a way that storage requirements can be traded for longer computational times. We show that this trade-off is indeed viable and that when the schedule is optimized, the storage and computational times grow at most logarithmically. This work will be described in detail in an extended paper. This work is a collaboration with G. Leaf, ANL, A. Griewank of the University of Dresden. The work is reported in [22]. We also developed the software package for the practical implementation of the recursive algorithm described above.

4.2 Computing Trends

It is hard to believe it, but we do not have a straightforward answer to "if the Earth is warming, by how much?"

Getting a trend correctly in climate can be extremely challenging: you may not have an ensemble of records with which to compute statistics in order to determine an underlying statistical model, nor a way to determine whether the data is stationary; it is not even clear sometimes whether it is statistical in the first place, and if a deterministic model should be proposed, which model to use. *In the absence of a clear reason for choosing one model over another, the trend that is being computed will depend not only on the quality of the data, but on the method itself.*

The aim of this research initiative, which started in 2010, is to create a trend methodology, which is to say, to create a definition for a trend. It will not rely on a statistical assumption or a polynomial ansatz for its shape.

The basic idea is to recursively remove local fluctuations, which consist of consecutive maxima and minima. The signal can thus be represented by a lossless decomposition of the form $x(t) = A(t) + \sum_i B_i(t)$, where the trend is the first term. The B_i are multi scale fluctuations with a locally defined notion of frequency. The decomposition resembles an amplitude equation expansion. Preliminary results appear in Figure 7. At present we have a purely algorithmic definition of the decomposition. Our next goal is to define in precise mathematical terms what are the properties of the decomposition we propose.

4.3 Climate Dynamics, With a Focus on the Oceans

The ocean is the largest heat reservoir on earth and thus of fundamental importance to climate dynamics. The Earth has been exposed to approximately 10^{17} W, for over roughly four billion

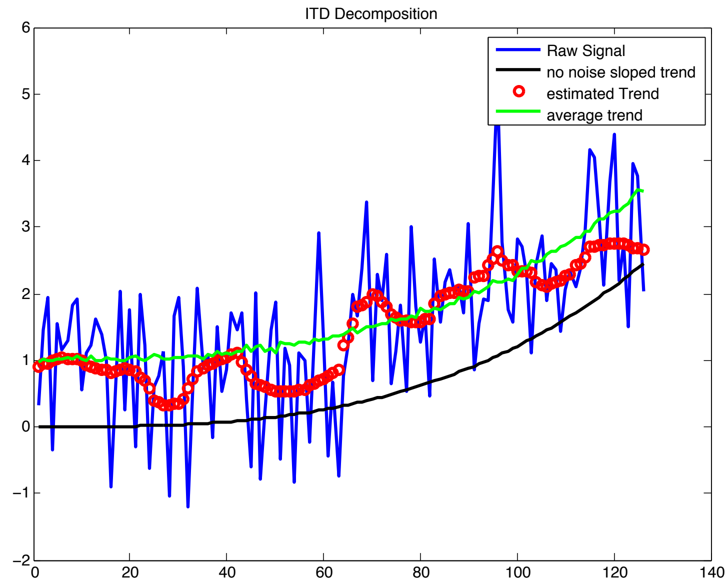


Figure 7: In blue is the raw signal. It has been prepared by superimposing normally-distributed noise and the deterministic black curve plus a bias of 1. The green line represents the traditional regression fit, which relies on assumptions or specific knowledge about what the process that generates the data. The red circles represent our determination of the trend, using no assumptions or knowledge of the process.

years. Explaining why the ocean has not "boiled off", together with explaining the apparent biogeochemical paradox of resources and biota are fundamental to understanding climate; and doing so depends crucially on studying oceanic transport.

Studying how the ocean overturns on transport time scales is essential if we are to understand climatic processes on a global scale. The thermohaline circulation is so called because the two most important components determining the buoyancy of ocean water are its mineral content or salt, and its heat content. Since freshwater inputs from rain, melting ice, river outflow, and precipitation/evaporation, affect the salt concentration in ocean water locally, and since the heat fluxes from the sun and polar ice caps are also spatially nonuniform, the ocean waters tend to sink, move from place to place, and upwell, transporting heat. This transport, of course, is also affected by the Earth's rotation, by oceanic currents, wind forcing, and the blocking and channeling effect of the land masses.

The study of transport using general circulation models (GCM's) is complex and time consuming. More so now, since there is a realization that studies of this sort require coupling multi-physics (atmospheres and oceans, for example). My focus has been instead to use/develop simple yet consistent models to invetigate fundamental conjectures about transport, especially when the problem is subtle yet discernibly difficult to test with GCM's. Below are a few examples of simple models put to good use.

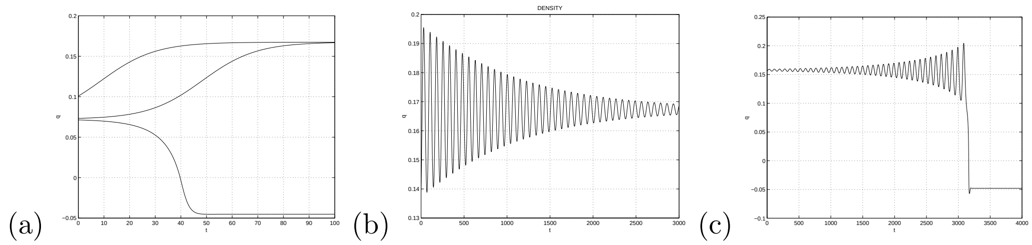


Figure 8: Classic Stommel solution, (b) decaying oscillation due to advection; (c) sustained oscillations and climate transitions due to advection.

4.3.1 Advective Effects in the Thermohaline Circulation

Waves are mostly ignored on climate scales. In [23] we demonstrate, to the contrary, that this is a safe assumption. Until recently, it was thought that waves and other advective effects would have little or no bearing on climatic long scales, as their energy and their coherence over long time scales are small. We revisited Stommel's box model, which was used to establish the non-unique character of the solutions to the climate equations using a low dimensional/low frequency approximation to the full dynamics. The model is modified to include wave effects. We incorporated advective effects into the model by means of delay equations and find that the behavior of the Stommel box-model with advection has self-sustaining oscillatory solutions. This behavior is singular, meaning that the oscillations do not vanish if the delay term is sufficiently large (but physical). The advection has the potential to produce self-sustaining oscillations (See Figure 8). In particular, we show that advective effects can produce climate jumps and decaying oscillatory conditions [24] (See Figure 8). The results may have important implications on climate variability. When these effects are large enough we show that they can cause a transition in the dynamics of the box-model ocean from a thermally dominated solution to a haline-dominated solution. This work is a collaboration with D. Kurtze (St. Joseph University).

4.3.2 Convective Adjustment and Oceanic Climate Outcomes

The process of convective adjustment is also being studied in models ranging from simple box-models to fully 3-D ocean general circulation models. 'Convective adjustment' is the algorithmic approach to incorporating fast mixing in the water column of layer models of the primitive equations. The primitive equations are simplified versions of the equations for the large scale flow of the oceans (and with appropriate modifications, of the atmosphere). The crucial simplification is the assumption of hydrostatic momentum balance between the many layers composing the flow. This simplification has a number of consequences on the resulting flow. Among these is that the flow exhibits unrealistic vertical structure and mixing properties. In order to overcome these shortcomings, enhanced parametrizations, combined with convective adjustments, are invoked along the course of the computation. The result is a nominal improvement by the suppression of certain instabilities and the enhancement of viscous effects.

However, Cessi and Young [25; 26] and Marotzke [27] presented compelling evidence that certain aspects of the climate variability observed in large scale general circulation models (particularly in low resolution numerical experiments) are artifacts of the convective adjustment algorithm. In order to fully understand how convective adjustment affects the solutions in we study a low dimensional approximation to the quasi-geostrophic equations which incorporates convective adjustments. The outcome of our analysis is that Stommel-like exchanges (that is, horizontal and vertical convective exchanges), as well as the heat and salt capacity of the lower regions of the flow, can produce dramatically different solutions. In particular, we show that the finiteness of the lower ocean, with which the upper ocean convectively adjusts, leads to oscillations which at most decay in finite time. The most salient outcome of our investigation is that CA is not capable of producing decadal oscillations, unless the ocean is forced to equilibrium under unrealistic conditions. This work appears in [28] and is a collaboration with an undergraduate student, J. Dittmann (UA), and Douglas Kurtze (St. Joseph University).

4.3.3 Carbon Cycle, Glacial/Interglacial Dynamics

During the last glacial era, the partial pressure of atmospheric carbon dioxide (CO_2) was around 200 ppmv. This is approximately 80 ppmv less than the preindustrial value of atmospheric CO_2. The ocean is the only reservoir large enough to cause such a dramatic change over those timescales. The ocean is able to store large amounts of CO_2 because it reacts with water to form an equilibrium with carbonate and bicarbonate ions. The distribution of total CO_2 (dissolved CO_2, carbonate and bicarbonate) throughout the ocean in controlled by air-sea exchange, the strength of the thermohaline flow, mixing between the surface and the deep water, and by the particulate flux of organic matter sinking to the bottom of the ocean. There is debate over the relative importance of the high and low latitude oceans in the storage of CO_2. Simple box models of the CO_2 cycle suggest that conditions in the high latitude oceans determine the amount of CO_2 the ocean is able to store. However, general circulation models (GCMs), indicate that the low latitude oceans also influence CO_2 storage. We developed a box model which is shown to be consistent with the fundamental fluid and thermodynamic equations governing this flow. In fact, this aspect of consistency is amiss in the existing box models. Using this model, in conjunction with realistic parameter values, we show that both glacial and interglacial solutions may be achieved while conserving total carbon dioxide by changing the forcing of the model to reflect the climate. Our box model thus settles important aspects of the controversy. Since it simple mathematically, it provides a very convenient framework to study CO_2 dynamics in the ocean/atmosphere on grand scales.

In [29] and [30] we derived a reduced complexity model which includes an explicit dependence of the strength of the ocean's overturning flow on density gradients and a phosphate cycle. Using this model the simulated glacial-interglacial variations in atmospheric pCO_2 agree remarkably well with ice-core measurements. Furthermore, the predicted variations in the strength of the overturning between glacial and interglacial periods are also in close agreement with estimates based on paleo-oceanographic data. A fundamental difference between this model and other box models applied to the glacial-interglacial pCO_2 question is that the flow field is not prescribed *a priori*; the model also differs from previous box-models in that it uses relatively few parameters; the parameters used have values in reasonable agreement with data-base estimates.

A sensitivity analysis of the reduced complexity model used in capturing glacial-interglacial variations in atmospheric pCO_2 referred to above shows that the model is not very sensitive to changes in the high latitude ocean (see [29]). This puts into question assertions that box models cannot be faithful to the dynamics; also it puts into question the notion that there is a robust and simple connection between sensitivity of models to high latitude effects and their ability of capturing the carbon cycle. The analysis further demonstrates that several measures for the assessment of the quality of the predictions for the many models for the problem are of questionable utility. This is joint work with S. Peacock (NCAR) and E. Lane (NIWA).

4.4 Wave-Generated Transport and Wave/Current Interactions

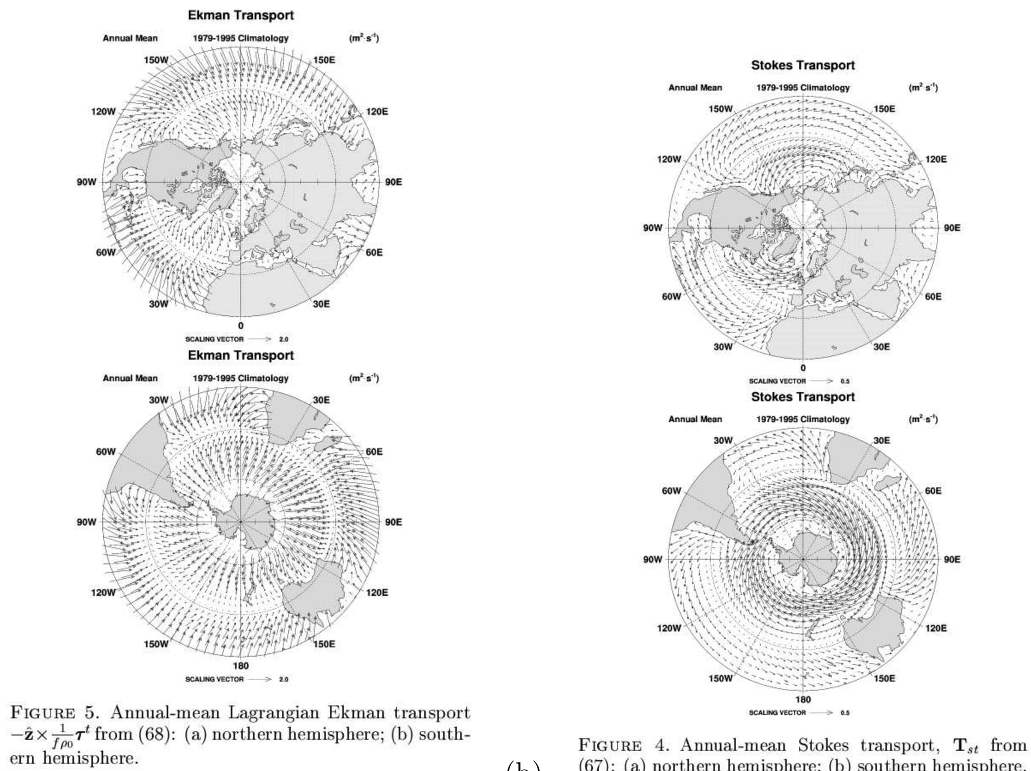
With Jim McWilliams at UCLA, we developed the theory for wave-generated and shows how it is an important transport mechanism in oceanic flows relevant to climate dynamics, vertical mixing and upwelling in the near-shore, and in the transportation of suspended sediment by the interaction of waves and currents. Our goal has been to develop a consistent and complete description of the interactions of wavs and currents.

Wave-generated transport has been shown to modify the dynamics of interacting flows: from nearshore processes, interactions between rivers and oceans. We even showed evidence that gravity waves are shown to affect climate scale oceanic flows. A consequence of momentum and continuity conservation is that short waves are dynamically modified by the mass and momentum fluxes due to finite amplitude long waves. The interaction of currents with the streaming induced by a wave field has been shown to generate circulation cells, which have a striking resemblance to Langmuir cells in the ocean and the atmosphere, and has been shown to modify the Ekman boundary layer in rotating flows.

4.4.1 Multiscale Model for Wave/Current Interactions

Combining asymptotics and time averaging over small scales we have built a hierarchy of multiscale models for short, intermediate and long time scale dynamics. The short scales corresponds to the gravity waves, the intermediate scale to the infra-gravity waves as well as setup and setdown, and the coarsest to currents. The result is a model for the interaction of these flows and a plausible explanation as to how short time scale motions modify the larger scales.

Among our results are: (1) how waves persist at ocean climate scales; (2) by developing a model prescribed specifically for the nearshore environment we also devised how one can come up with a model for the smaller scales that is formulated at the larger resolvable scales; (3) we have rederived the radiation stress under these new assumptions are able to show that the often used expressions are, to lowest order, identically zero; (4) incorporation of dissipative effects due to wave breaking via stochastic parametrization; (5) derivation of a theoretical framework [23] [31], the vortex force framework, which suggests how and why wave generated transport is relevant to climate dynamics, *i.e.*, to oceanic basin-scale motions (Ekman and Stokes flows appear in Figure 9 and mean sea level adjustments predicted by the theory appear in Figure 11b). This is one of the most complete and consistent models for the evolution of wave-current interactions, encompassing gravity waves, long waves, and currents in three space dimensions. The model is also capable of capturing the mechanics of pollutants, heat, and the transport of other tracers. Equally important is that the model is framed exclusively at the largest spatio-temporal scales, thus circumventing the computational hurdle of resolving motions which have extremely large bandwidths. See Figure 10. (6) complete derivation of the radiation stresses in the nearshore and comparison of these to more traditional expressions for these stresses [32]. The aim of this comparison is to settle once and for all the existing misconceptions on what the stress itself yields. Furthermore, our results differ from the more generally accepted expressions, with consequences on our understanding of phenomena on which it has a bearing: rip currents, wave set-up and set-down, undertows, litoral currents.



(a) (b)

Figure 9: (a) Ekman and (b) it Stokes currents as predicted from theory using actual wind data.

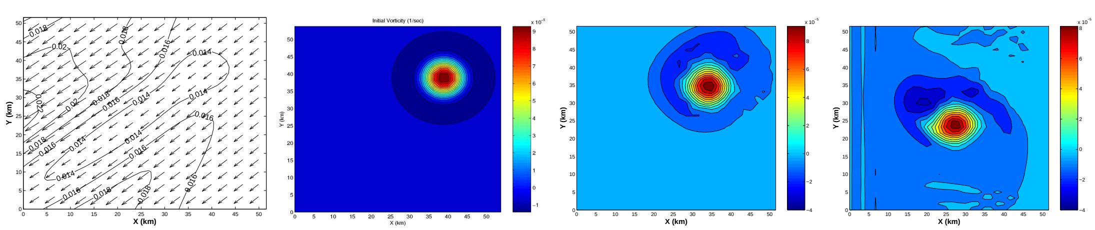
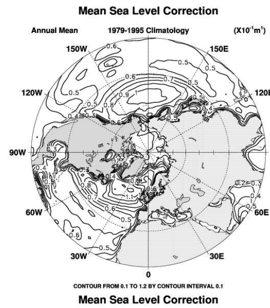


Figure 10: (a) Stokes drift due to waves, (b) initial vortex, (c) final vortex, no waves, (d) final vortex, waves.

(a) 

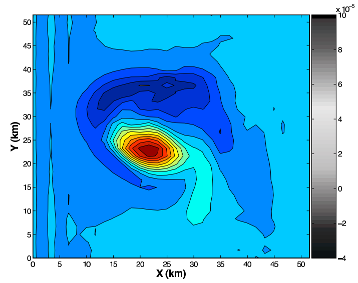
(b) 

Figure 11: (a) Mean sea level correction predicted by theory. (b) Final vortex, dissipation effects included. See Figure 10.

4.4.2 Wave Breaking Stochastic Parametrization in Wave/Current Interactions

One particularly difficult problem in wave/current interactions is to figure out how dissipation enters the dynamics. The major difficulties here are that it is not at all clear what dissipation does to each of these components of the flow and how dissipation distributes itself among the waves and currents.

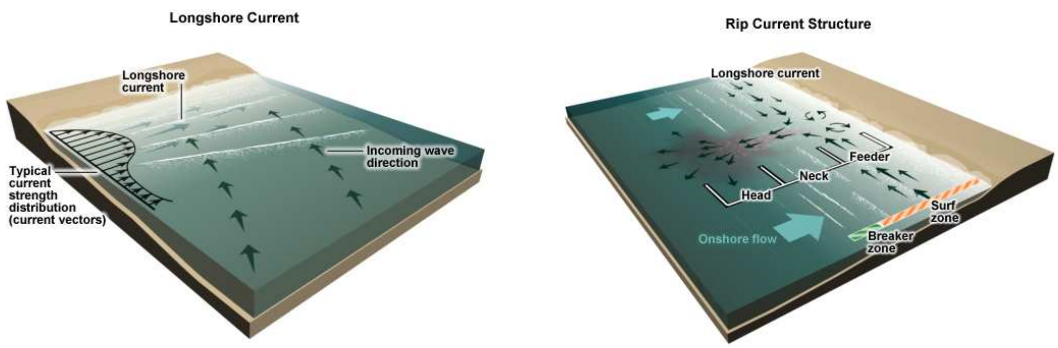
We make use of stochastic differential equation theory. The main idea here is to assert that the dynamics of fluid flow in the Lagrangian frame in the presence of a periodic and spatially-dependent noise source can be modelled by a standard stochastic differential equation. We then apply projection methods and time averaging, as well as averaging over ensembles, to find how effects such as white capping and other sources of dissipation enter the full wave/current model. This is done by recasting the Lagrangian fluid orbit equations as stochastic differential equations and then using projection methods to infer how the stochasticity enters the Eulerian equations at the wave and current scales (see Figure 11). The resulting model explains in what manner white-capping, bottom drag and wind drag appear in the interaction evolution equations. The derivation shows that the dissipative terms appear in the currents in a manner similar to Reynolds stresses, a vast modelling improvement over more ad-hoc methods currently in use. Furthermore it spells out in precise terms how the boundary conditions at the free surface has to be modified due to the presence of dissipation.

The derivation of the dissipation terms in the wave-current interaction model appears in [1; 33]. The ultimate goal is to produce an analogous connection between Lagrangian diffusion and wave/current motions as the one that exists between Brownian motion and molecular diffusion.

Dissipation Effects on Transport in Boundary Layers: Separately, we examined how stochastic effects on the dynamics of wave-generated transport in an oscillatory boundary layer were examined as well (see [34]). We found that noise enhances trapping in progressive wave fields. It has the opposite effect in the standing wave case. This result has a bearing on how wave transport in the noisy environment of natural flows manifests itself and how residual flows interact with an erodible, if present.

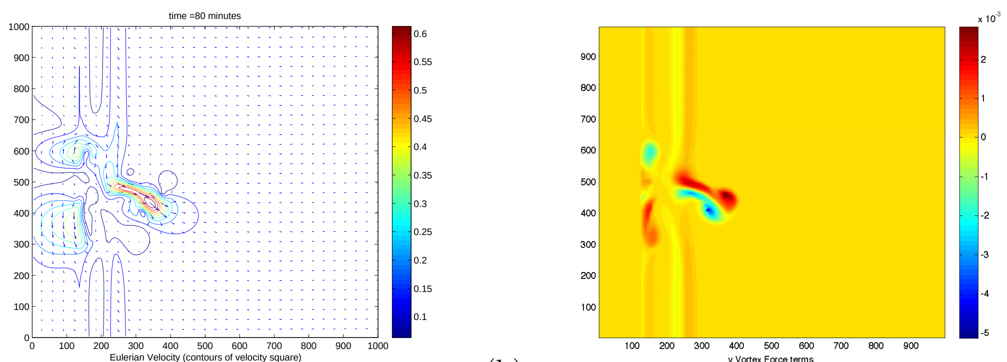
4.4.3 Applications: Longshore and Rip Currents

Having derived in great detail the radiation stresses from our wave/current interaction model it is now possible to revisit these two classical problems with the aim at determining whether significant



(a) (b)

Figure 12: (a) Longshore currents, (b) Rip currents.



(a) (b)

Figure 13: (a) Rip current, predicted by our model, (b) the vortex force component.

qualitative improvements over the traditional view on these problems is possible. Namely, to see if it is possible to determine the position of the runnel in the longshore wave-forced currents (see Figure 12a), and whether the spacing and extent of rip currents (see Figure 12b) is significantly modified by using our model over more traditional means. The centerpiece of our theoretical model are the *Vortex Force*, which couples the total vorticity of the flow and the residual flow due to the waves, as well as the *Bernoulli Head*, which produces an important pressure readjustment as a response to the excess mass due to the waves. The vortex force, in particular, we are finding, can explain a substantial portion of the vortex shearing that occurs in the rip current situation. This is shown in Figure 13a-b.

4.5 Theory and Experiments with Stratified Flows

Stratified flows are ubiquitous in the geophysical setting. Here we consider a series of controlled experiments in which the Boussinesq approximation is known to fail. See the experiment in Figure 14a, and a computational outcome in Figure 14b. In [35] and [36] we consider the flow structure and stability of a planar saline jet descending into a stable, density-stratified fluid. The jet exhibits a rapid acceleration on release, then deceleration, as it encounters the more dense surrounding fluid, yet retains its slender shape due to the low salt diffusion. As the jet descends it entrains fresher

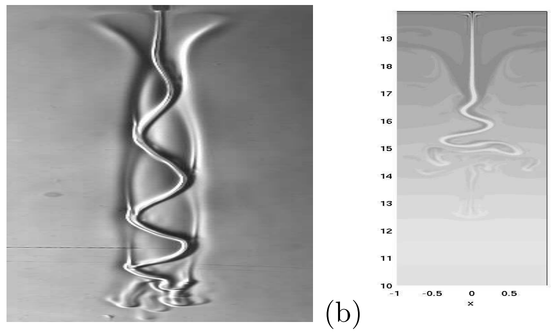
(a) (b)


Figure 14: (a) Actual jet, (b) computational jet. The images are not meant to portray the same dynamic conditions.

water which as it encounters the increasingly dense ambient fluid returns toward the nozzle forming a recirculation zone. Our numerical simulations agree qualitatively with previous experiments and thus serve as a tool to explain the basic kinematics of the jet. We also use numerical means to capture the three instability modes: an antisymmetric instability in the jet core, a symmetric instability in the jet core, and a symmetric instability in the entrained conduit of less saline water. For the dominant antisymmetric instability we determine the range of parameters that demarcate stable and unstable regions.

4.6 Sediment Dynamics

Oceanic erodible beds are very complicated dynamic systems. We do not know fully about the turbulent flow over these, let alone about the dynamics of these interacting granular beds. Even the forces experienced by individual particles are not well known: we thus engaged in a systematic computational campaign, using full 3D Navier-Stokes to find the lift, drag, inertia and buoyancy forces experienced by a particle. We also performed experiment, in which high degree of control not found in nature allows us to pin down some of the mechanisms involved in the formation and maintenance of ripples. We also propose models, focusing mostly on very large scale sand ridges.

4.6.1 Computing Forces on Sedimentary Particles

Calculations are performed because it is impractical to measure these forces in the laboratory, where sensitivity to forces in the micro Newton range would have to be resolved. A primary goal of these experiments is to discover whether lift forces are important to the dynamics of sediment in the oceanic situation. With my colleagues P. Fischer (ANL) and G. Leaf (ANL), we have been using full 3D Navier Stokes calculations of the oscillatory fluid flow in the boundary layer to comprehensively characterize the forces of drag, lift, and torque on particles. Thus far we have fully characterized the lift, drag, and torque as a function of the particles' Reynolds number, the gap, the forcing frequency and the buoyancy. In the future we will address numerically dislogdgement and suspention of particles in the flow. We know how: we will again use a 3D Navier Stokes, however, we will modify it to conform to an accelerated frame. This work appears in [37], [38], [39]. These are 3D Navier Stokes calculations and are considered definitive quantitative and qualitative experimental data. (See Figure 15 for the mesh configuration). Some of the outcomes appear

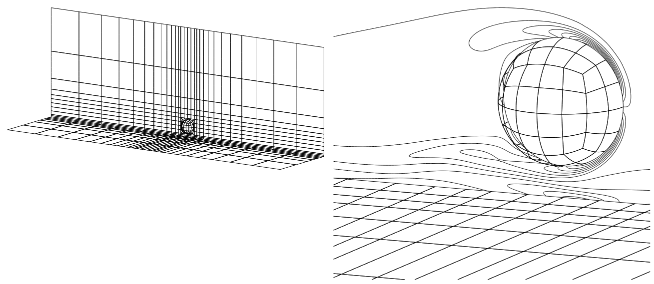
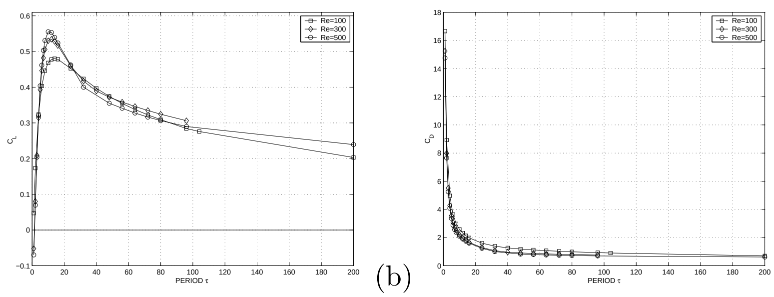


Figure 15: Configurational mesh for 3D Navier Stokes computations of lift and drag on a particle in an oscillatory flow.



(a) (b)

Figure 16: Lift and Drag on a particle as a function of the Reynolds number and the frequency of forcing.

in Figure 16. We established that lift forces play a role in the dynamics of sediment particles under typical oceanic settings [37]. Most every model for sediment dynamics in oceanic and fluvial settings ignores the lift force.

4.6.2 Experiments on Sheared Erodible Beds

Laboratory sediment dynamics experiments were designed and built that give us unprecedented control over the parameters of the problem of the generation and maintenance of bars under the action of shearing flows. The focus here has been primarily with the aspect of modelling fundamental aspects of the sediment dynamics, rather than on the faithful recreation of the oceanic setting in a laboratory. An annular tank has been built in which water and sediment are sheared by a solid rotor. The shearing can be either unidirectional (both steady and accelerated motions are being investigated), or oscillatory. The apparatus includes a video system that enables viewing of the entire tank at once. The shearing rotor, the image capturing system, and the analysis routines are driven by software. At present we are studying the steady shearing case. In this case we have identified two different routes to bed instabilities. These instabilities manifest themselves in bar-like structures that precess around the tank. Transient rearrangements are also possible which lead to disturbances that travel at speeds higher than the precession velocity. We have also found that it is possible for an accumulation of sediment to slowly resolve itself in solitary-wave like structures whose speed of precession is related to their amplitude (and possibly their wavelength). This is but one of many phenomenological outcomes of this experiment. This work is a collaboration with Derek Moulton (U Oxford) and Hermann Uys (NIST). The work appears in [40].

See Figure 17a for an illustration of the experimental arrangement and Figure 17b for some

(a) (b)

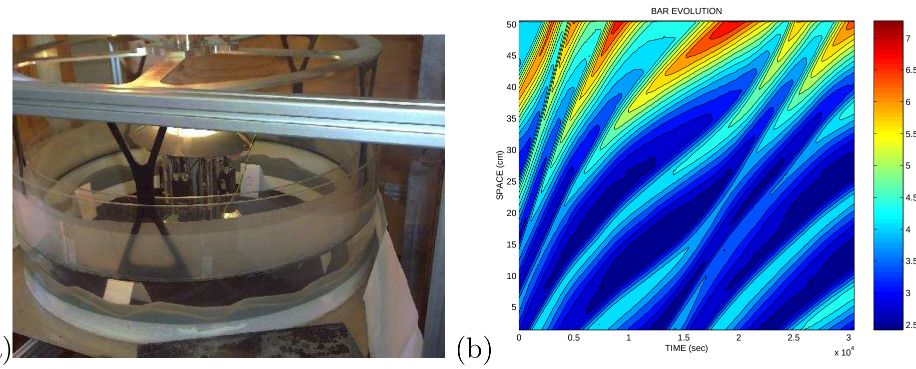


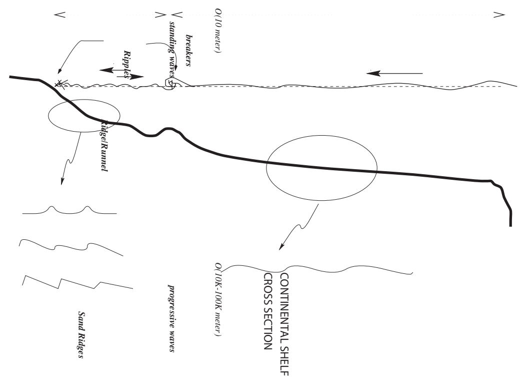
Figure 17: (a) Shearing sand experimental setup. (b) Space-time contours of the sedimentary structures. Contours represent height.

of the space-time data). We have been able to identify at least three types of sedimentary structures: advective, dispersive, and super-advective bars. From a dynamic standpoint we find that the generation of bars is captured by Benjamin's analysis [41]. But once the bars get larger, and boundary layer separation occurs, vortices are produced which modulate the shear stresses at the edge of the bar and naturally scour the lee end of the bars. This is an important new finding because it puts into question the standard view that the fluxes are considered possibly complex functions of the topographical gradient.

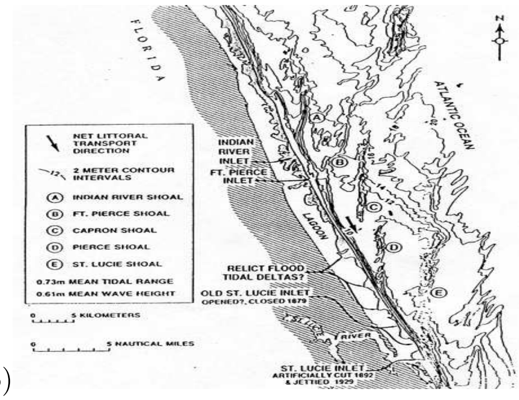
4.6.3 Modeling Sandridges, Shore-Connected Bars

Striking sedimentary structures are found on the continental shelf wherever sand is abundant (See Figure 18 for a schematic view of the continental shelf). Two common arrangements are known as sand ripples and sand ridges. These temporaly evolving sand patterns have piqued the curiosity of researchers for nearly 100 years. Sediment motion has been studied for centuries now. The research is focused mainly on practical aspects of civil engineering, erosion dynamics, petroleum engineering, and industrial processes involving granular flow. Great strides have been made in the formulation of ad-hoc models, which have proven themselves useful in engineering practice. However, due to poor understanding of the dynamics of sediment motion, no evolutionary model has succeeded in accounting for the large variety of observed sedimentary structures.

In collaboration with J. Bona (UIC), a model capturing the evolution of large scale sand ridges in the deeper reaches of the shelf was developed. The results from this model imply that on the time and length scales of sand ridges, sedimentary structures are generated by nonlinear dispersive infra-gravity waves. The energetic interactions (due to nonlinear effects) and the spatial changes (due to dispersion of the spectral components in the infra-gravity range of the waves) have length scales that match those of sand ridges. These nonlinear interactions generate a coherent and structurally organized pattern in the drift velocity. This drift velocity is endowed with a structured pattern by the overlying nonlinear dispersive waves. If these nonlinear interactions are coherent and energetic enough, the drift velocity will scour the ocean bottom in ways that produce the beautiful patterns observed in sandy environments. The work on the development of an evolutionary, parameter-free model, for the formation and evolution of sand-ridges on the continental shelf and detailed analysis of its solutions appears in [42], [43], [44]. A schematic of the domain of the problem appears in



(a)



(b)

Figure 18: (a) Cross section of the continental shelf environment. (b) Shore-connected ridges off the coast of Florida.

Figure 18).

A different type of sand ridge is called "shore-connected" because it is invariably seen to stem from the shore. These are large bars, ubiquitous in the Middle Atlantic Bite and other parts of the world. The burning question was to figure out why they tend to line up roughly at a 12 degree angle with respect to the shore. (see Figure 18b). A model which yields structures that resemble sand ridges on the inner shelf has been formulated in which exponential growth of large features results from the offshore deflection of storm-driven alongshore flows at ridge crests. Our goal was to predict how the inclusion of waves modifies the outcomes of a model with which it is conjectured that current convergences and divergences, driven by instabilities, are capable of explaining the angle, spacing, and size of the ridges. The waves are shown to modify the picture in significant ways, depending on their orientation and their frequency: they modify the celerity of the ridges, they are capable of displacing them, and even suppressing them.

The application of the wave-current model to the study of the formation and evolution of sedimentary dynamics of shore-connected ridges [31], [45], such as those commonly found in the Middle Atlantic Bight (Figure 19), shows that divergences/convergences that occur when longshore currents travel over a critically sloped bottom have the alignment seen in the bars. Furthermore, the waves, rather than always enhancing their appearance can also wipe them out.

5 The Computational Geosciences Program

Just as there are large efforts at many universities to modernize the education of biologists to produce a more quantitatively-skilled class of life scientists, a comparable effort should be made to produce a significant number of quantitatively-oriented geoscientists.

The demand for quantitative geoscientists is most keenly heard in government labs. Talented post-docs come to labs to work on hydrology and geophysics and they spend their first year retooling and learning scientific computing. Engineers, physicists, and mathematicians have little exposure to large-scale computational and theoretical statistics. The problem is also faced by academic departments with strong theoretical geoscience programs: I hear frequent complaints from my

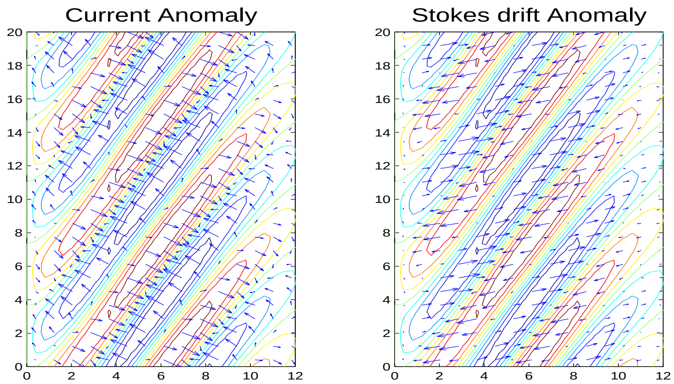


Figure 19: Current and waves anomalies showing covergences and divergences in a patter reminiscent of the shore-connected bars. The bars themselves will have a similar footprint.

geoscience colleagues that they cannot get students who have adequate quantitative skills if their undergraduate education was not math, physics, engineering, chemistry.

The Computational Geosciences Program is an educational and career strategy that addresses these problems, using existing academic and laboratory facilities, and personnel. It is based on building strong and long-term relations between scientists in training and the labs. At the graduate level it would train students who would be recognized by their knowledge of a particular branch of geoscience, but who have do research within scientific computing with a strong large-scale statistics or estimation theory emphasis.

A Appendix: Other Research

A.1 Scientific Computing

- **Preconditioners:** Developed a preconditioner for normal matrices which have gaps in their eigenvalue spectrum [46]. The technique is accurate, numerically stable. This is an alternative to more traditional methods, however, it is a very beautiful method, mathematically.

- **Locating Eigenvalues**: Used of complex polynomial approximation to develop accurately and computationally efficiently the gaps in the spectrum of a Hermitian matrix without the computation of the full spectrum [47].

- **Wavelet-Galerkin**: Development of a wavelet-Galerkin technique for the solution of PDE's. Computational and theoretical studies of wavelet-Galerkin methods for the solution of hyperbolic problems [48], [49]. Analytical estimates of the approximating properties of periodized wavelets. Development of software for the computation of inner products using wavelets for these wavelet-Galerkin methods [50].

- **Computing Software**: Development of software for task farming [51]. When the code was developed it was a novel and simple way to have a heterogeneous network of computers automatically submit sequential jobs in an efficient and fully documented way. This code is now been bypassed by tools with significanly higher capabilities, such as Python, Condor, and GRID.

A.2 Mathematics

- **The gBBM:** In a separate study we are used a pseudo-spectral and time-adaptive code to conduct a large series of numerical experiments which explore the rich behavior of orbitally-stable and unstable solutions to the generalized Benjamin-Bona-Mahony equation $u_t + \alpha u_x + (u^p)_x - u_{xxt} = 0$, where $u(x,t)$ is a real-valued function, $p \geq 2$ is an integer, and $\alpha \geq 0$ is a constant parameter. The equation is important because (for $p = 2$ and $p = 3$) it models waves in plasmas, ion-acoustic media, and water. We confirmed the previously-derived theoretical stability estimates and, more importantly, yielded insights into their behavior. We have found, among other things, that an initially small and hence slow soliton which is unstable will move to the stable regime by growing and speeding up, possibly shedding another soliton and/or a dispersive tail. Perhaps the most important result is that we show that arbitrary initial data will resolve itself into a train of solitary-waves. This is important because the resolution of solitary-waves is occurring in an equation that is clearly non-integrable. This work is in collaboration with W. McKinney, North Carolina State, and Jerry L. Bona, UIC.

- **The Benjamin Equation:** this is an equation that models a new kind of solitary wave. This solitary wave is a solution of an equation that governs, approximately, waves on the interface of a two-fluid system in which surface tension effects cannot be ignored. The equation is $u_t + u_x + 2uu_x - \alpha L u_x - \beta u_{xxx} = 0$, where $\alpha > 0$ and $\beta \geq 0$, and $\alpha \ll \beta$ is the physically relevant range. The operator $L = H\partial_x$ is the composition of the Hilbert transform H and the spatial derivative; L is a Fourier multiplier operator with symbol $|\kappa|$ is essentially the Korteweg-de Vries equation however the fact that $\beta > 0$ endows its solutions with properties that are qualitatively different from the KdV-equation solutions. We established the existence and stability of this new kind of solitary wave. Proved the existence and stability of oscillatory solitary. In [52] we also studied the collisions of the solutions and from them assessed their stability to nonlinear perturbations. This is work in collaboration with T. B. Benjamin, Oxford, J. L. Bona, UIC, and J. Albert, University of Oklahoma, and J. Hyman, now at Tulane.

A.3 Biology

- **Blood Cell Dynamics**:

  We use our existing Navier Stokes solver to solve for the flow inside and outside a blood cell membrane. The membrane has its own mechanical equations. Our novel contribution is to impose the nonlocal surface constraint *on the grid* generator. Thus the determination of cell motions amounts to solving systems of coupled ordinary differential equations and thus easily the 3D motion of the flow and the cells become computationally scalable. This is a practical method because the motion of the grid representing the cell boundary is laminar within the sea of grid points.

  We produced a paper on the distribution of red blood cells in small bifurcating vessels ([53]): while it is approximately true that cells distribute themselves in a bifurcating vessel in a manner proportional to the blood fluxes themselves, this is not the case for very small blood vessels. See Figure 20.

- **Bone Remodelling**:

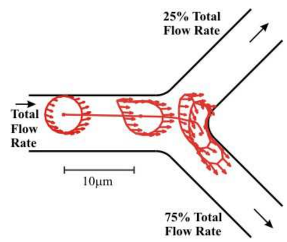


Figure 20: Bifurcating blood vessel and motion of blood cell.

Researchers at SUNY Buffalo and NASA have been able to determine that shaking the body of an animal (a turkey, in fact) tends to preserve already healthy bone in microgravity environments. The shaking is in the order of 10-80 cps and the magnitude is just a a fraction of a g, moreover, the treatment to be effective requires only hours per day. This finding has obvious implications in space travel, the health of bones of bed-ridden patients, etc. Separately, there is a group at U. Washington that has been using high levels of sound radiation to induce bone growth in healthy fractured bones.

Yi Jiang (LANL), Rustum Choksi (McGill) and I developed an alternative model ([54]) and explanation. We examined the mechanistic model for the formation and reforming of bone tissue [54]. Found equilibrium solutions and their stability. The analysis determines very sharply the bounds of applicability of the model in capturing the dynamics of bone reforming. Found that the model from which we generated our own, is not capable of producing sustained oscillating solutions, contrary to the published results.

A.4 Acoustics

Development of regular and high frequency approximations for the covariant acoustic field scattered from randomly-rough surfaces, such as the ocean surface [55] [56].

A.5 Political Theory, The Arts

Studied the influence of exit polls on voting. In [57] we proposed a model with which to study the effect of releasing exit poll data and projections while voting is taking place, in a majority-rule, multi-party election. Will be considering the voting process during the primaries and eventual presidential election as a sequential voting scheme in which information plays a role on voting strategies.

Involvement in the production of the first-ever computer-generated music on a massively parallel distributed supercomputer [58]. Design and supervision of the construction of several concert halls and other complex architectural acoustics structures [59].

References

[1] J. M. Restrepo. Wave breaking dissipation in the wave-driven ocean circulation. *Journal of Physical Oceanography*, 37:1749–1763, 2007.

[2] M. I. Budyko. The effect of solar radiation variations on the climate of the Earth. *Tellus*, 21:611–619, 1969.

[3] W. D. Sellers. A global climate model based on the energy balance of the earth-atmosphere system. *Journal of Applied Meteorology*, 8:392–400, 1969.

[4] I. Eisenman and J. S. Wettlaufer. Nonlinear threshold behavior during the loss of Artic sea ice. *Proceedings of the National Academy of Sciences*, 106:28–32, 2010.

[5] G. L. Eyink and J. R. Restrepo. Most probable histories for nonlinear dynamics: tracking climate transitions. *Journal of Statistical Physics*, 101:459–472, 2000.

[6] G. L. Eyink, J. M. Restrepo, and F. J. Alexander. A mean field approximation in data assimilation for nonlinear dynamics. *Physica D*, 194:347–368, 2004.

[7] G. L. Eyink, J. M. Restrepo, and F. J. Alexander. A statistical-mechanical approach to data assimilation using moment closures, *submitted*, 2002.

[8] F. J. Alexander, G. L. Eyink, and J. M. Restrepo. Accelerated Monte-Carlo for optimal estimation of time series. *Journal of Statistical Physics*, 119:1331–1345, 2005.

[9] J. M. Restrepo. A path integral method for data assimilation. *Physica D*, 237:14–27, 2008.

[10] S. Kim, G. L. Eyink, J. M. Restrepo, F. J. Alexander, and G. Johnson. Ensemble filtering for nonlinear dynamics. *Monthly Weather Review*, 131:2586–2594, 2003.

[11] H. J. Kushner. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *SIAM Journal on Control, Series A*, 2:106–119, 1962.

[12] D. Crisan, P. del Moral, and T.J. Lyons. Discrete filtering using branching and interacting particle systems. *Markov Processes and Related Fields*, 5:293–331, 1999.

[13] a.m. stuart, j. voss, and p. wiberg. Conditional path sampling of sdes and the langevin mcmc method. *Communications in the Mathematical Sciences*, 2:685–697, 2004.

[14] P. Krause and J. M. Restrepo. Lagrangian data assimilation using the Kernel Diffusion Method. *Monthly Weather Review*, 137:4386–4400, 2009.

[15] P. Krause and J. M. Restrepo. Sequential estimation with deterministic models and noisy data using the reduced diffusion kernel filter.

[16] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2002.

[17] N.J. Gordon, D. J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140:107–113, 1993.

[18] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50:174–188, 2002.

[19] L. Kuznetsov, K. Ide, and C.K.R.T. Jones. A method for assimilation of Lagrangian data. *Monthly Weather Review*, 131:2247–2260, 2003.

[20] Z. Toth and E. Kalnay. Ensemble forecasting at NCEP: the generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317–2330, 1993.

[21] N. Balci, A. Mazucatto, J M. Restrepo, and G. Sell. Bred vectors and ensemble dynamics. *Monthly Weather Review*, 2011.

[22] J. M. Restrepo, G. K. Leaf, and A. Griewank. Circumventing storage limitations in variational data assimilation studies. *SIAM Journal on Scientific Computing*, 19:1586–1605, 1998.

[23] J. C. McWilliams and J. M. Restrepo. The wave-driven ocean circulation. *Journal of Physical Oceanography*, 29:2523–2540, 1999.

[24] D. Kurtze and J. M. Restrepo. Advective time lags in box models. *Journal of Physical Oceanography*, 31:1828–1842, 2001.

[25] P. Cessi. Grid-scale instability of convective adjustment schemes. *Journal of Marine Research*, 54:407–420, 1996.

[26] P. Cessi and W. Young. Some unexpected consequences of the interaction between convective adjustment and horizontal diffusion. *Physica D*, 98:287–300, 1996.

[27] J. Marotzke. Influence of convective adjustment on the stability of the thermohaline circulation. *Journal of Physical Oceanography*, 21:903–907, 1991.

[28] D. Kurtze, J. M. Restrepo, and J. Dittmann. Convective adjustment in box models. *Ocean Modelling*, pages 287–300, 2010.

[29] E. Lane, S. Peacock, and J. M. Restrepo. A dynamic-flow carbon-cycle box model and high-latitude sensitivity. *Tellus B*, 58:257–278, 2006.

[30] E. Lane, S. Peacock, and J. M. Restrepo. A possible sequence of events for the generalized Glacial-Interglacial cycle. *Journal of Global Biogeochemical Cycles*, 20:GB2010, 2006.

[31] J. M. Restrepo. Wave-current interactions in shallow waters and shore-connected ridges. *Continental Shelf Research*, 21:1331–1360, 2001.

[32] E. M. Lane, J. M. Restrepo, and J. C. McWilliams. Wave-current interaction: A comparison of radiation-stress and vortex-force representations. *Journal of Physical Oceanography*, 37:1122–1141, 2007.

[33] J. M. Restrepo, J. M. Ramirez, M. Banner, and J. C. McWilliams. Multiscale dissipation and diffusion processes in wave/current interactions. *Journal of Physical Oceanography*, 2010. accepted.

[34] J. M. Restrepo and G. K. Leaf. Wave-generated transport induced by ideal waves. *Journal of Physical Oceanography*, 32:2334–2349, 2002.

[35] S. Schofield and J. M. Restrepo. Stability of planar buoyant jets in stratified fluids. *Physics of Fluids*, 22, 2010.

[36] C. Dombrowski, B. Lewellyn, A. I. Pesci, J. M. Restrepo, J. O. Kessler, and R. E. Goldstein. Coiling, entrianment, and hydrodynamic coupling of decelerated fluid jets. *Physics Review Letters*, 95:184501, 2005.

[37] J. M. Restrepo, P. Fischer, and G. K. Leaf. Forces on particles in oscillatory boundary layers. *Journal of Fluid Mechanics*, 468:327–347, 2002.

[38] P. Fischer, G. K. Leaf, and J. M. Restrepo. Influence of wall proximity on the lift and drag of a particle in an oscillatory flow. *Journal of Fluids Engineering*, 127:583–594, 2005.

[39] P. Fischer, G. Leaf, and J. M. Restrepo. Torque effects on the lift and drag of particles in an oscillatory boundary flow. *Journal of Fluids Engineering*, 2008.

[40] J. M. Restrepo, D. Moulton, and H. Uys. Advective precessive and interacting bars under shear flows, 2010. Accepted.

[41] T.B. Benjamin. Shearing flow over a wavy boundary. *Journal of Fluid Mechanics*, 6:11, 1959.

[42] J. M. Restrepo. Behavior of a sand ridge model. *European Journal of Mechanics B/Fluids*, 6:835–861, 1997.

[43] J. M. Restrepo and J. L. Bona. Three-dimensional model for the formation of longshore sand ridges on the continental shelf. *Nonlinearity*, 8:781–820, 1995.

[44] J.M. Restrepo and J. L. Bona. Discretization of a model for the formation of longshore sand ridges. *Journal of Computational Physics*, 122:129–142, 1995.

[45] E. M. Lane and J. M. Restrepo. Shore-connected ridges under the action of currents and waves. *Journal of Fluid Mechanics*, 582, 2007.

[46] M. Hasson and J.M. Restrepo. Approximating on disjoint intervals and its application to matrix preconditioning. *Accepted, Mathematics of Computation*, 2005.

[47] M. Hasson, J.M. Restrepo, and J. Hyman. A strategy for finding extreme eigenvalues. *Accepted, Computers Mathematics and Applications*, 2005.

[48] J. M. Restrepo and G. K. Leaf. Wavelet-Galerkin discretization of hyperbolic equations. *Journal of Computational Physics*, 122:118–128, 1995.

[49] J. M. Restrepo and G. K. Leaf. Inner product computations using periodized Daubechies wavelets. *International Journal of Numerical Methods in Engineering*, 40:3557–3578, 1997.

[50] J. M. Restrepo and G. K. Leaf. Connection coefficient software. `www.physics.arizona.edu/`. `∼restrepo/myweb/downloads/ntuples.tar.gz`.

[51] C. Chen and J. M. Restrepo. Task farming scheduler for beowulf-class computers. `www.physics.arizona.edu/`. `∼restrepo/myweb/software/TF/index.html`.

[52] J. Albert, J. L. Bona, and J. M. Restrepo. Solitary-wave solutions of the benjamin equation. *SIAM Journal of Applied Mathematics*, 25:2139–2161, 1999.

[53] J. Barber, J. P. Alberding, J. M. Restrepo, and T. Secomb. Two-dimensional computational models of red blood cell motion in microvessel bifurcations and flexibility effects. *Annals of Biomechanics*, 36:1690–1698, 2008.

[54] J. M. Restrepo, R. Choksi, and Y. Jiang. On a mechanistic model for bone remodeling, 2006. Submitted, Journal of Biomechanics.

[55] J. M. Restrepo and S. T. McDaniel. Two models for the spatially covariant field scattered by randomly rough pressure-release surfaces with Gaussian spectra. *Journal of the Acoustical Society of America*, 87:2033–2043, 1990.

[56] J. M. Restrepo and S. T. McDaniel. Spatial coherence in the high frequency limit. *Waves in Random Media*, 2:183–193, 1991.

[57] J. M. Restrepo and R. Rael andJ. Hyman. Modeling the influence of polls on elections: a population dynamics approach. *Journal of Public Choice*, 140:395–420, 2009.

[58] H. Kaper, D. Ralley, J. M. Restrepo, and S. Tipei. Additive synthesis with DIASS on argonne national laboratory's IBM POWERparallel system (sp). *Proceedings, International Computer Music Conference, Banff*, pages 351–352, 1995.

[59] J. M. Restrepo. AEP Acoustics, LLC. www.aepacoustics.com.