Special Session 20: Stochastic-Statistical Modeling of Climate

Dimitris Giannakis, New York University, USA
John Harlim, North Carolina State University, USA
Andrew Majda, New York University, USA

Fundamental barriers to advancing weather and climate prediction on time scales from several days to years are limited by the capability of contemporary operational and research prediction systems (GCMs) to represent coupled processes in the climate involving precipitating convection, atmospheric teleconnection patterns, low-frequency modes in the ocean, and variability in the cryosphere. A grand challenge of contemporary applied science is to understand these patterns and their changes in a globally warming world, as well as the impact of these changes on long range forecasting. These problems are beyond the regime of traditional weather forecast models, and instead both effects of the initial state and the change in mean forcing are important. In contrast to using comprehensive GCMs, it is very natural to develop stochastic-statistical models for these patterns, as well as their interaction, for use in long-range forecasting, sensitivity, and attribution studies. This special session aims to bring together researchers from across the spectrum of disciplines related to statistical-stochastic modeling of climate to discuss the development and application of emerging ideas and techniques for these important and difficult practical issues.

Quantifying uncertainty for predictions with model error in non-Gaussian systems with intermittency

Michal Branicki
Courant Institute, NYU, USA
branicki@cims.nyu.edu
Andrew J. Majda

Synergy between empirical information theory and fluctuation-dissipation theorem provides a systematic framework for improving sensitivity and predictive skill for imperfect models of complex natural systems. We utilize a suite of increasingly complex nonlinear models with intermittent hidden instabilities and time-periodic features to illustrate the advantages of such an approach, as well as the role of model errors due to coarse-graining, moment closure approximations, and the memory of initial conditions in imperfect prediction.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Using stochastic models to diagnose the origins of leading atmospheric zonal modes

Grant Branstator
NCAR, USA
branst@ucar.edu

In terms of the meridional wind, the leading structures of interannual variability have a simple structure: to a first approximation each pattern is concentrated in a narrow band of latitudes and is dominated by a single zonal wavenumber. For example zonal wavenumber three variability is prominent at high latitudes while zonal wavenumber five variability is prominent in midlatitudes. Many general circulation models (GCMs) are not able to reproduce this behavior, which can affect their ability to respond properly to an external forcing, including increasing greenhouse gases. We have used mechanistic models generated by linearizing the governing equations about climate mean states and driven by noise to determine aspects of the mean state that may be responsible for the prominent wavenumbers of variability in nature and GCMs. For example we find that whether the mean state has a strong projection onto zonal wavenumber three at high latitudes is a factor in determining whether a climate system will have prominent variability in that wavenumber in that region. We have also used response operators based on the fluctuation-dissipation theorem to determine regions from which interannual variability in tropical rainfall may force these prominent patterns, thus identifying another factor that may differentiate model behavior.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Stochastic subgrid-scale parameterization designed for a finite-difference model discretization

Stamen Dolaptchiev
Goethe-University Frankfurt, Germany
dolaptchiev@iau.uni-frankfurt.de
U. Achatz, I. Timofeyev

We present a new approach for the construction of stochastic subgrid-scale parameterizations. Starting from a high-resolution finite-difference discretization of some model equations, the approach is based on splitting the model variables into fast, small-scale and slow, large-scale modes by averaging the model discretization over neighboring grid cells. After that, a closed form effective stochastic model for the slow modes is derived applying a stochastic mode reduction procedure. An advantage over heretofore applications of stochastic mode reduction to spectrally discretized models is that the resulting closure is local and thus remains applicable even if the number of slow variables is large. The new approach is implemented for the discretized Burgers equation and compared with other benchmark parameterizations.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Quantifying long-range predictability and model error through data clustering and information theory

Dimitrios Giannakis
New York University, USA
dimitris@cims.nyu.edu
Andrew J. Majda

We present a framework blending data clustering and information theory to quantify long-range initial-value predictability and forecast error with imperfect models in complex dynamical systems. With reference to wind-driven ocean circulation, we demonstrate that the pertinent information for long-range forecasting can be represented via a coarse-grained partition of the set of initial data available to a model. A related formalism is applied to assess the forecast skill of Markov models of ocean circulation regimes.

⟶ ∞ ⋄ ∞ ⟵

Stochastic homogenization for an energy conserving multi-scale toy model of the atmosphere

Georg Gottwald
University of Sydney, Australia
georg.gottwald@sydney.edu.au
Jason Frank

We study a simple Hamiltonian toy model for a Lagrangian fluid parcel in the semi-geostrophic limit which exhibits slow and fast dynamics. We first reinject unresolved fast dynamics into the deterministic equation through stochastic parametrization which respects the conservation of the energy of the deterministic system. In a second step we use stochastic singular perturbation theory to derive an effective reduced stochastic differential equation for the slow dynamics. We verify the results in numerical simulations.

⟶ ∞ ⋄ ∞ ⟵

Optimal filtering of complex turbulent systems with memory depth through consistency constraints

John Harlim
North Carolina State University, USA
jharlim@ncsu.edu

We present the AR(p)-filter for assimilating weakly chaotic dynamical systems with long memory depth. In particular, a data-driven autoregressive stochastic models with non-Markovian nature is used as surrogate filter prior models. We will show that the autoregressive filter is not as sensitive as standard ensemble filtering strategies to additional intrinsic model errors. Secondly, we will also discuss offline mathematical conditions for optimal autoregressive filters. In particular, we will rigorously and numerically show that if the autoregressive model parameters are chosen to satisfy a certain subset of the consistency conditions and absolute stability of multistep numerical discretization scheme, then the optimal autoregressive filtering is guaranteed. We will also demonstrate how to apply this result to improve signals with long memory depth: the first Fourier coefficient of the truncated Burgers-Hopf model and the Lorenz-96 model in weakly chaotic regime.

⟶ ∞ ⋄ ∞ ⟵

A Bayesian approach to parameter estimation and model error quantification of stochastic models for turbulent signals

Radu Herbei
The Ohio State University, USA
herbei@stat.osu.edu
Dimitris Giannakis, Andrew J. Majda

We study the performance of Markov Chain Monte Carlo (MCMC) algorithms for parameter estimation and model error quantification in the setting of a model with time-dependent stochastic parameters, which has high skill in reproducing intermittency, transient instability, and other important features of turbulent signals. The test parameters of the model are chosen to simulate two distinct regimes: (1) frequent, short-lasting transient instabilities, and (2) large amplitude transient instabilities. We implement several state-of-art MCMC algorithms and study their ability to correctly recover these parameters from partial observations. Our main focus is to compare an adaptive MCMC approach to a particle MCMC approach under two scenarios: (a) correctly specified model and (b) with model error. We discuss the benefits and drawbacks of a Bayesian approach to this setting.

⟶ ∞ ⋄ ∞ ⟵

Stochastic data assimilation methods for estimating ocean eddy heat transport

Shane Keating
Cournat Institute of Mathematical Science, USA
skeating@cims.nyu.edu
Andrew J. Majda, K. Shafer Smith

The role of ocean eddies in redistributing heat from the tropics to the poles remains a poorly constrained feature of the global energy balance. Attempts to diagnose eddy transport are limited by the sparseness of available observations and the nonlinearity of the underlying dynamics. In this study, a suite of stochastic data assimilation methods are tested in idealized two-layer simulations of oceanic turbulence at high and low latitudes under a range of observation scenarios. A novel feature is the use of

inexpensive stochastic models to forecast the eddy dynamics. The stochastic model parameters can be estimated by regression fitting to climatological energy spectra and correlation times or by adaptively learning these parameters "on-the-fly" from the observations themselves. We show that, by extracting high-wavenumber information that has been aliased into the low wavenumber band, one can derive "stochastically superresolved" velocity fields with a nominal resolution increase of a factor of two or more. The filtered estimates of the upper and lower layer streamfunctions produce time-mean poleward eddy heat transports that are significantly closer to the true value when compared with standard estimates based upon optimal interpolation. Implications for estimating poleward eddy heat transport using current and next-generation altimeters are discussed.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Sparse adaptive polynomial chaos representations for ocean general circulation models

Omar Knio
Duke University, USA
omar.knio@duke.edu
P. Conrad, J. Winokur, I. Sraj, A. Alexanderian, M. Iskandarani, A. Srinivasan, Y. Marzouk, O. Knio

A database of high-resolution HYCOM simulations of the oceanic circulation in the Gulf of Mexico is used to conduct an a priori analysis of the performance of adaptive refinement schemes for uncertainty quantification. The database includes realizations corresponding to isotropic sparse sampling of the uncertain model inputs, namely parameterizations of subgrid mixing and wind drag. The analysis is used to determine performance gains due to a sparse, adaptive, pseudospectral projection approach, and to study the impact of different refinement criteria. Predictions of adaptive refinement are validated against results obtained using a Latin hypercube sampling approach. Finally, the analysis is used to explore the potential of stochastic preconditioning in constructing sparse representations of model outputs as well as improving the performance of adaptive refinement schemes.

$\longrightarrow \infty \diamond \infty \longleftarrow$

A digital filtering framework for the local ensemble transform Kalman filter

Eric Kostelich
Arizona State University, USA
kostelich@asu.edu
Istvan Szunyogh

Data assimilation refers to the process by which initial conditions for geophysical models are determined from noisy observations, typically with maximum likelihood methods. In the classical Kalman filter, one seeks to minimize an appropriate sum of squares that is weighted according to one's relative confidence in the observations and the model forecasts. This talk will explore alternative formulations to include an internal digital filter as a weak constraint within the Local Ensemble Transform Kalman Filter. Preliminary results of its application to the Global Forecast System atmospheric model will be described.

$\longrightarrow \infty \diamond \infty \longleftarrow$

How do you determine whether the earth is warming up?

Juan Restrepo
University of Arizona, USA
restrepo@math.arizona.edu
Darin Comeau, Hermann Flaschka

How does one determine whether the extreme summer temperatures in Moscow in 2010 was an extreme climatic fluctuation or the result of a systematic global warming trend? It is only under exceptional circumstances that one can determine whether a climate signal belongs to a particular statistical distribution. In fact, climate signals are rarely "statistical," other than measurement errors, there is usually no way to obtain enough field data to produce a trend or a tendency, based upon data alone. We propose a trend or tendency methodology that does not make use of a parametric or statistical assumption. The most important feature of this trend strategy is that it is defined in very precise mathematical terms.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Blended reduced subspace algorithms for uncertainty quantification

Themistoklis Sapsis
New York University, USA
themis.sapsis@gmail.com
Andrew J. Majda

We study uncertainty-quantification (UQ) properties of the Quasi-linear Gaussian (QG) closure method and we compare it with the results from order-reduction based on dynamical orthogonality (DO). We find that each of these approaches suffer from disadvantages that can be overcome by combining them. Specifically, the QG method is incapable to capture strong energy transfers among linearized modes. On the other hand, due to the reduced order character the DO approach is incapable to capture the full-order effect of the linearized operator which in many cases (e.g. skew systems) can be critical for the correct evolution of the statistics. We formulate a blended approach based on these two methods which can be further improved by adding empirical information.

$\longrightarrow \infty \diamond \infty \longleftarrow$

A stochastic model for tropical rainfall and extreme events

Samuel Stechmann
University of Wisconsin-Madison, USA
stechmann@wisc.edu
J. David Neelin

Recently it has been discovered that tropical rainfall patterns, on scales of 20-200 km or larger, have statistics that resemble critical phenomena from statistical physics. Through, for instance, the power-law distributions and long-range correlations in these statistics, the characteristics of extreme rainfall events can be quantified. To gain further insight into these statistics and extreme events, a stochastic model is designed and analyzed to reproduce the statistics that are local in space (and evolving in time). The model includes the interaction of a stochastic jump process and Gaussian processes to represent different aspects of tropical convection, a highly complex system that, if fully resolved, involves nonlinear turbulent interactions of fluid dynamics and moist thermodynamics. The stochastic model can be thought of as a simplified subgrid-scale parameterization of moist convection for atmospheric models with grid spacings of 20-200 km.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Sub-sampling in parametric estimation of effective stochastic models

Ilya Timofeyev
University of Houston, USA
ilya@math.uh.edu
R. Azencott, A Beri

It is often desirable to derive an effective stochastic model for the physical process from observational and/or numerical data. Various techniques exist for performing estimation of drift and diffusion in stochastic differential equations from discrete datasets. In this talk we discuss the question of sub-sampling of the data when it is desirable to approximate statistical features of a smooth trajectory by a stochastic differential equation. In this case estimation of stochastic differential equations would yield incorrect results if the dataset is too dense in time. Therefore, the dataset has to sub-sampled (i.e. rarefied) to ensure estimators' consistency. Favorable sub-sampling regime is identified from the asymptotic consistency of the estimators. Nevertheless, we show that estimators are biased for any finite sub-sampling time-step and construct new bias-corrected estimators.

$\longrightarrow \infty \diamond \infty \longleftarrow$

Sampling in and out of equilibrium when the tails matter

Jonathan Weare
University of Chicago, USA
weare@math.uchicago.edu

Analyzing and simulation rare events in stochastic process is important in many areas. For example, in the context of weather and climate, prediction can be dramatically hampered by unlikely, but important transitions in the underlying system. The past decade or so has seen dramatic improvements in our ability to simulate and analyze these events. So far this progress has come mostly in the contexts of Chemistry and Computer Science but rare event ideas seem ripe for application in geophysical contexts. I will survey a bit of my work in rare event simulation generally as well as show some preliminary work toward geophysical applications.

$\longrightarrow \infty \diamond \infty \longleftarrow$