of step size variation. It is also worth noting that (22) generates a smooth step size sequence, as $h_n - h_{n-1} = O(h_n h_{n-1})$.

This type of control does not work with an error estimate, but rather tracks a prescribed target function; it corresponds to keeping $hQ(y) = \text{const.}$, where Q is a given functional reflecting the geometric structure of (1). One can then take $G(y) = \text{grad}\, Q(y) \cdot f(y)/Q(y)$. For example, in celestial mechanics, $Q(y)$ could be selected as total centripetal acceleration; then the step size is small when centripetal acceleration is large and vice versa, concentrating the computational effort to those intervals where the solution of the problem changes rapidly and is more sensitive to perturbations.

Literature

Step size control has a long history, starting with the first initial value problem solvers around 1960, often using a simple step doubling/halving strategy. The controller (10) was soon introduced, and further developments quickly followed. Although the schemes were largely heuristic, performance tests and practical experience developed working standards. Monographs such as [1, 2, 6, 7, 10] all offer detailed descriptions.

The first full control theoretic analysis is found in [3, 4], explaining and overcoming some previously noted difficulties, developing proportional-integral (PI) and autoregressive (AR) controllers. Synchronization with Newton iteration is discussed in [5]. A complete framework for using digital filters and signal processing is developed in [11], focusing on moving average (MA) controllers. Further developments on how to obtain improved computational stability are discussed in [12].

The special needs of geometric integration are discussed in [8], although the symmetric controllers are not based on error control. Error control in implicit, symmetric methods is analyzed in [13].

References

1. Butcher, J.C.: Numerical Methods for Ordinary Differential Equations. Wiley, Chichester (2008)
2. Gear, C.W.: Numerical Initial Value Problems in Ordinary Differential Equations. Prentice Hall, Englewood Cliffs (1971)
3. Gustafsson, K.: Control theoretic techniques for stepsize selection in explicit Runge–Kutta methods. ACM TOMS **17**, 533–554 (1991)
4. Gustafsson, K.: Control theoretic techniques for stepsize selection in implicit Runge–Kutta methods. ACM TOMS **20**, 496–517 (1994)
5. Gustafsson, K., Söderlind, G.: Control strategies for the iterative solution of nonlinear equations in ODE solvers. SIAM J. Sci. Comp. **18**, 23–40 (1997)
6. Hairer, E., Nϕrsett, S.P., Wanner, G.: Solving Ordinary Differential Equations I: Nonstiff Problems, 2nd edn. Springer, Berlin (1993)
7. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, 2nd edn. Springer, Berlin (1996)
8. Hairer, E., Söderlind, G.: Explicit, time reversible, adaptive step size control. SIAM J. Sci. Comp. **26**, 1838–1851 (2005)
9. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer, Berlin (2006)
10. Shampine, L., Gordon, M.: Computer Solution of Ordinary Differential Equations: The Initial Value Problem. Freeman, San Francisco (1975)
11. Söderlind, G.: Digital filters in adaptive time-stepping. ACM Trans. Math. Softw. **29**, 1–26 (2003)
12. Söderlind, G., Wang, L.: Adaptive time-stepping and computational stability. J. Comp. Methods Sci. Eng. **185**, 225–243 (2006)
13. Stoffer, D.: Variable steps for reversible integration methods. Computing **55**, 1–22 (1995)

Stochastic and Statistical Methods in Climate, Atmosphere, and Ocean Science

Daan Crommelin[1,2] and Boualem Khouider[3]
[1]Scientific Computing Group, Centrum Wiskunde and Informatica (CWI), Amsterdam, The Netherlands
[2]Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands
[3]Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada

Introduction

The behavior of the atmosphere, oceans, and climate is intrinsically uncertain. The basic physical principles that govern atmospheric and oceanic flows are well known, for example, the Navier-Stokes equations for fluid flow, thermodynamic properties of moist air, and the effects of density stratification and Coriolis force.

Notwithstanding, there are major sources of randomness and uncertainty that prevent perfect prediction and complete understanding of these flows.

The climate system involves a wide spectrum of space and time scales due to processes occurring on the order of microns and milliseconds such as the formation of cloud and rain droplets to global phenomena involving annual and decadal oscillations such as the EL Nio-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) [5]. Moreover, climate records display a spectral variability ranging from 1 cycle per month to 1 cycle per 100,000 years [23]. The complexity of the climate system stems in large part from the inherent nonlinearities of fluid mechanics and the phase changes of water substances. The atmosphere and oceans are turbulent, nonlinear systems that display chaotic behavior (e.g., [39]). The time evolutions of the same chaotic system starting from two slightly different initial states diverge exponentially fast, so that chaotic systems are marked by limited predictability. Beyond the so-called predictability horizon (on the order of 10 days for the atmosphere), initial state uncertainties (e.g., due to imperfect observations) have grown to the point that straightforward forecasts are no longer useful.

Another major source of uncertainty stems from the fact that numerical models for atmospheric and oceanic flows cannot describe all relevant physical processes at once. These models are in essence discretized partial differential equations (PDEs), and the derivation of suitable PDEs (e.g., the so-called primitive equations) from more general ones that are less convenient for computation (e.g., the full Navier-Stokes equations) involves approximations and simplifications that introduce errors in the equations. Furthermore, as a result of spatial discretization of the PDEs, numerical models have finite resolution so that small-scale processes with length scales below the model grid scale are not resolved. These limitations are unavoidable, leading to model error and uncertainty.

The uncertainties due to chaotic behavior and unresolved processes motivate the use of stochastic and statistical methods for modeling and understanding climate, atmosphere, and oceans. Models can be augmented with random elements in order to represent time-evolving uncertainties, leading to stochastic models. Weather forecasts and climate predictions are increasingly expressed in probabilistic terms, making explicit the margins of uncertainty inherent to any prediction.

Statistical Methods

For assessment and validation of models, a comparison of individual model trajectories is typically not suitable, because of the uncertainties described earlier. Rather, the statistical properties of models are used to summarize model behavior and to compare against other models and against observations. Examples are the mean and variance of spatial patterns of rainfall or sea surface temperature, the time evolution of global mean temperature, and the statistics of extreme events (e.g., hurricanes or heat waves). Part of the statistical methods used in this context is fairly general, not specifically tied to climate-atmosphere-ocean science (CAOS). However, other methods are rather specific for CAOS applications, and we will highlight some of these here. General references on statistical methods in CAOS are [61, 62].

EOFs

A technique that is used widely in CAOS is Principal Component Analysis (PCA), also known as Empirical Orthogonal Function (EOF) analysis in CAOS. Consider a multivariate dataset $\Phi \in \mathbf{R}^{M\times N}$. In CAOS this will typically be a time series $\phi(t_1), \phi(t_2), \ldots, \phi(t_N)$ where each $\phi(t_n) \in \mathbf{R}^M$ is a spatial field (of, e.g., temperature or pressure). For simplicity we assume that the time mean has been subtracted from the dataset, so $\sum_{n=1}^{N} \Phi_{mn} = 0 \ \forall m$. Let C be the $M \times M$ (sample) covariance matrix for this dataset:

$$C = \frac{1}{N-1}\, \Phi\, \Phi^T .$$

We denote by (λ_m, v^m), $m = 1, \ldots, M$ the ordered eigenpairs of C:

$$C\, v^m = \lambda_m\, v^m , \quad \lambda_m \geq \lambda_{m+1} \ \forall m.$$

The ordering of the (positive) eigenvalues implies that the projection of the dataset onto the leading eigenvector v^1 gives the maximum variance among all projections. The next eigenvector v^2 gives the maximum variance among all projections orthogonal to v^1, v^3 gives maximum variance among all projections

orthogonal to v^1 and v^2, etc. The fraction $\lambda_m / \sum_l \lambda_l$ equals the fraction of the total variance of the data captured by projection onto the m-th eigenvector v^m.

The eigenvectors v^m are called the Empirical Orthogonal Functions (EOFs) or Principal Components (PCs). Projecting the original dataset Φ onto the leading EOFs, i.e., the projection/reduction

$$\phi^r(t_n) = \sum_{m=1}^{M'} \alpha_m(t_n) v^m \,, \qquad M' \ll M,$$

can result in a substantial data reduction while retaining most of the variance of the original data.

PCA is discussed in great detail in [27] and [59]. Over the years, various generalizations and alternatives for PCA have been formulated, for example, Principal Interaction and Oscillation Patterns [24], Nonlinear Principal Component Analysis (NLPCA) [49], and Nonlinear Laplacian Spectral Analysis (NLSA) [22]. These more advanced methods are designed to overcome limitations of PCA relating to the nonlinear or dynamical structure of datasets.

In CAOS, the EOFs v^m often correspond to spatial patterns. The shape of the patterns of leading EOFs can give insight in the physical-dynamical processes underlying the dataset Φ. However, this must be done with caution, as the EOFs are statistical constructions and cannot always be interpreted as having physical or dynamical meaning in themselves (see [50] for a discussion).

The temporal properties of the (time-dependent) coefficients $\alpha_m(t)$ can be analyzed by calculating, e.g., autocorrelation functions. Also, models for these coefficients can be formulated (in terms of ordinary differential equations (ODEs), stochastic differential equations (SDEs), etc.) that aim to capture the main dynamical properties of the original dataset or model variables $\phi(t)$. For such reduced models, the emphasis is usually on the dynamics on large spatial scales and long time scales. These are embodied by the leading EOFs v^m, $m = 1, \ldots, M'$, and their corresponding coefficients $\alpha_m(t)$, so that a reduced model ($M' \ll M$) can be well capable of capturing the main large-scale dynamical properties of the original dataset.

Inverse Modeling

One way of arriving at reduced models is inverse modeling, i.e., the dynamical model is obtained through statistical inference from time series data. The data can be the result of, e.g., projecting the dataset Φ onto the EOFs (in which case the data are time series of $\alpha(t)$). These models are often cast as SDEs whose parameters must be estimated from the available time series. If the SDEs are restricted to have linear drift and additive noise (i.e., restricted to be those of a multivariate Ornstein-Uhlenbeck (OU) process), the estimation can be carried out for high-dimensional SDEs rather easily. That is, assume the SDEs have the form

$$d\alpha(t) = B\,\alpha(t)\,dt + \sigma\,dW(t)\,, \qquad (1)$$

in which B and σ are both a constant real $M' \times M'$ matrix and $W(t)$ is an M'-dimensional vector of independent Wiener processes (for simplicity we assume that α has zero mean). The parameters of this model are the matrix elements of B an σ. They can be estimated from two (lagged) covariance matrices of the time series. If we define

$$R^0_{ij} = \mathbf{E}\alpha_i(t)\alpha_j(t)\,, \qquad R^\tau_{ij} = \mathbf{E}\alpha_i(t)\alpha_j(t+\tau)\,,$$

with $\mathbf{E}$ denoting expectation, then for the OU process (1), we have the relations

$$R^\tau = \exp(B\,\tau)\,R^0$$

and

$$BR^0 + R^0 B^T + \sigma\sigma^T = 0$$

The latter of these is the fluctuation-dissipation relation for the OU process. By estimating R^0 and R^τ (with some $\tau > 0$) from time series of α, estimates for B and $A := \sigma\sigma^T$ can be easily computed using these relations. This procedure is sometimes referred to as linear inverse modeling (LIM) in CAOS [55]. The matrix σ cannot be uniquely determined from A; however, any σ for which $A = \sigma\sigma^T$ (e.g., obtained by Cholesky decomposition of A) will result in an OU process with the desired covariances R^0 and R^τ.

As mentioned, LIM can be carried out rather easily for multivariate processes. This is a major advantage of LIM. A drawback is that the OU process (1) cannot capture non-Gaussian properties, so that LIM can only be used for data with Gaussian distributions. Also, the estimated B and A are sensitive to the choice of τ, unless the available time series is a an exact sampling of (1).

Estimating diffusion processes with non-Gaussian properties is much more complicated. There are various estimation procedures available for SDEs with nonlinear drift and/or multiplicative noise; see, e.g., [30, 58] for an overview. However, the practical use of these procedures is often limited to SDEs with very low dimensions, due to curse of dimension or to computational feasibility. For an example application in CAOS, see, e.g., [4].

The dynamics of given time series can also be captured by reduced models that have discrete state spaces, rather than continuous ones as in the case of SDEs. There are a number of studies in CAOS that employ finite-state Markov chains for this purpose (e.g., [8,48,53]). It usually requires discretization of the state space; this can be achieved with, e.g., clustering methods. A more advanced methodology, building on the concept of Markov chains yet resulting in continuous state spaces, is that of hidden Markov models. These have been used, e.g., to model rainfall data (e.g., [3, 63]) and to study regime behavior in large-scale atmospheric dynamics [41]. Yet a more sophisticated methodology that combines the clustering and Markov chain concepts, specifically designed for nonstationary processes, can be found in [25].

Extreme Events

The occurrence of extreme meteorological events, such as hurricanes, extreme rainfall, and heat waves, is of great importance because of their societal impact. Statistical methods to study extreme events are therefore used extensively in CAOS. The key question for studying extremes with statistical methods is to be able to assess the probability of certain events, having only a dataset available that is too short to contain more than a few of these events (and occasionally, too short to contain even a single event of interest). For example, how can one assess the probability of sea water level at some coastal location being more than 5 m above average if only 100 years of observational data for that location is available, with a maximum of 4 m above average? Such questions can be made accessible using extreme value theory. General introductions to extreme value theory are, e.g., [7] and [11]. For recent research on extremes in the context of climate science, see, e.g., [29] and the collection [1].

The classical theory deals with sequences or observations of N independent and identically distributed (iid) random variables, denoted here by $r_1, \ldots, r_N$. Let M_N be the maximum of this sequence, $M_N = \max\{r_1, \ldots, r_N\}$. If the probability distribution for M_N can be rescaled so that it converges in the limit of increasingly long sequences (i.e., $N \to \infty$), it converges to a generalized extreme value (GEV) distribution. More precisely, if there are sequences $a_N (> 0)$ and b_N such that $\mathrm{Prob}((M_N - b_N)/a_N \leq z) \to G(z)$ as $N \to \infty$, then

$$G(z) = \exp\left(-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right]\right).$$

$G(z)$ is a GEV distribution, with parameters μ (location), $\sigma > 0$ (scale), and ξ (shape). It combines the Fréchet ($\xi > 0$), Weibull ($\xi < 0$), and Gumbel ($\xi \to 0$) families of extreme value distributions. Note that this result is independent of the precise distribution of the random variables r_n. The parameters μ, σ, ξ can be inferred by dividing the observations $r_1, r_2, \ldots$ in blocks of equal length and considering the maxima on these blocks (the so-called block maxima approach).

An alternative method for characterizing extremes, making more efficient use of available data than the block maxima approach, is known as the peaks-over-threshold (POT) approach. The idea is to set a threshold, say r^*, and study the distribution of all observations r_n that exceed this threshold. Thus, the object of interest is the conditional probability distribution $\mathrm{Prob}(r_n - r^* > z \,|\, r_n > r^*)$, with $z > 0$. Under fairly general conditions, this distribution converges to $1 - H(z)$ for high thresholds r^*, where $H(z)$ is the generalized Pareto distribution (GPD):

$$H(z) = 1 - \left(1 + \frac{\xi z}{\tilde{\sigma}}\right)^{-1/\xi}.$$

The parameters of the GPD family of distributions are directly related to those of the GEV distribution: the shape parameter ξ is the same in both, whereas the threshold-dependent scale parameter is $\tilde{\sigma} = \sigma + \xi(r^* - \mu)$ with μ and σ as in the GEV distribution.

By inferring the parameters of the GPD or GEV distributions from a given dataset, one can calculate probabilities of extremes that are not present themselves in that dataset (but have the same underlying distribution as the available data). In principle, this makes it possible to assess risks of events that have not been observed, provided the conditions on convergence to GPD or GEV distributions are met.

As mentioned, classical results on extreme value theory apply to iid random variables. These results have been generalized to time-correlated random variables, both stationary and nonstationary [7]. This is important for weather and climate applications, where datasets considered in the context of extremes are often time series. Another relevant topic is the development of multivariate extreme value theory [11].

Stochastic Methods

Given the sheer complexity of climate-atmosphere-ocean (CAO) dynamics, when studying the global climate system or some parts of global oscillation patterns such ENSO or PDO, it is natural to try to separate the global dynamics occurring on longer time scales from local processes which occur on much shorter scales. Moreover, as mentioned before, climate and weather prediction models are based on a numerical discretization of the equations of motion, and due to limitations in computing resources, it is simply impossible to represent the wide range of space and time scales involved in CAO. Instead, general circulation models (GCMs) rely on parameterization schemes to represent the effect of the small/unresolved scales on the large/resolved scales. Below, we briefly illustrate how stochastic models are used in CAO both to build theoretical models that separate small-scale (noise) and large-scale dynamics and to "parameterize" the effect of small scales on large scales. A good snapshot on the state of the art, during the last two decades or so, in stochastic climate modeling research can be found in [26, 52].

Model Reduction for Noise-Driven Large-Scale Dynamics

In an attempt to explain the observed low-frequency variability of CAO, Hasselmann [23] splits the system into slow climate components (e.g., oceans, biosphere, cryosphere), denoted by the vector x, and fast components representing the weather, i.e., atmospheric variability, denoted by a vector y. The full climate system takes the form

$$\begin{aligned} \frac{dx}{dt} &= u(x,y) \\ \frac{dy}{dt} &= v(x,y), \end{aligned} \tag{2}$$

where t is time and $u(x,v)$ and $v(x,y)$ contain the external forcing and internal dynamics that couple the slow and fast variables.

Hasselmann assumes a large scale-separation between the slow and fast time scales: $\tau_y = O\left(y_j \left(\frac{dy_j}{dt}\right)^{-1}\right) \ll \tau_x = O\left(x_i \left(\frac{dx_i}{dt}\right)^{-1}\right)$, for all components i and j. The time scale separation was used earlier to justify statistical dynamical models (SDM) used then to track the dynamics of the climate system alone under the influence of external forcing. Without the variability due to the internal interactions of CAO, the SDMs failed badly to explain the observed "red" spectrum which characterizes low-frequency variability of CAO.

Hasselmann made the analogy with the Brownian motion (BM), modeling the erratic movements of a few large particles immersed in a fluid that are subject to bombardments by the rapidly moving fluid molecules as a "natural" extension of the SDM models. Moreover, Hasselmann [23] assumes that the variability of x can be divided into a mean tendency $\langle dx/dt\rangle = \langle u(x,y)\rangle$ (Here $\langle . \rangle$ denotes average with respect to the joint distribution of the fast variables.) and a fluctuation tendency $dx'/dt = u(x,y) - \langle u(x,y)\rangle = u'(x,y)$ which, according to the Brownian motion problem, is assumed to be a pure diffusion process or white noise. However, unlike BM, Hasselmann argued that for the weather and climate system, the statistics of y are not in equilibrium but depend on the slowly evolving large-scale dynamics and thus can only be obtained empirically. To avoid linear growth of the covariance matrix $\langle x' \otimes x'\rangle$, Hasselmann assumes a damping term proportional to the divergence of the background frequency $F(0)$ of $\langle x' \otimes x'\rangle$, where $\delta(\omega-\omega')F_{ij}(\omega) = \langle V_i(\omega)V_j(\omega')\rangle$ with $V(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u'(t)e^{-i\omega t}dt$. This leads to the Fokker-Plank equation: [23]

$$\frac{\partial p(x,t)}{\partial t} + \nabla_x \cdot (\hat{u}(x)\, p(x,t)) = \nabla_x \cdot (D \nabla_x\, p(x,t)) \tag{3}$$

for the distribution $p(x,t)$ of $x(t)$ as a stochastic process given that $x(0) = x_0$, where D is the normalized covariance matrix $D = \langle x' \otimes x'\rangle/2t$ and $\hat{u} = \langle u\rangle - \pi \nabla_x \cdot F(0)$. Given the knowledge of the mean statistical forcing $\langle u\rangle$, the evolution equation for p can be determined from the time series of x obtained either from a climate model simulation of from observations. Notice also that for a large number of slow variables x_i,

the PDE in (3) is impractical; instead, one can always resort to Monte Carlo simulations using the associated Langevin equation:

$$dx = \hat{u}(x)dt + \Sigma(x)dW_t \tag{4}$$

where $\Sigma(x)\Sigma(x)^T = D(x)$. However, the functional dependence of $\hat{u}$ and D remains ambiguous, and relying on rather empirical methods to define such terms is unsatisfactory. Nonetheless, Hasselmann introduced a "linear feedback" version of his model where the drift or propagation term is a negative definite linear operator: $\hat{u}(x) = Ux$ and D is constant, independent of x as an approximation for short time excursions of the climate variables. In this case, $p(x,t)$ is simply a Gaussian distribution whose time-dependent mean and variance are determined by the matrices D and U as noted in the inverse modeling section above.

Due to its simplicity, the linear feedback model is widely used to study the low-frequency variability of various climate processes. It is, for instance, used in [17] to reproduce the observed red spectrum of the sea surface temperature in midlatitudes using simulation data from a simplified coupled ocean-atmosphere model. However, this linear model has severe limitations of, for example, not being able to represent deviations from Gaussian distribution of some climate phenomena [13, 14, 17, 36, 51, 54]. It is thus natural to try to reincorporate a nonlinearity of some kind into the model. The most popular idea consisted in making the matrix D or equivalently Σ dependent on x (quadratically for D or linearly for Σ as a next order Taylor correction) to which is tied the notion of multiplicative versus additive (when D is constant) noise [37, 60]. Beside the crude approximation, the apparent advantage of this approach is the maintaining of the stabilizing linear operator U in place although it is not universally justified.

A mathematical justification for Hasselmann's framework is provided by Arnold and his collaborators (see [2] and references therein). It is based on the well-known technique of averaging (the law of large numbers) and the central limit theorem. However, as in Hasselmann's original work, it assumes the existence and knowledge of the invariant measure of the fast variables. Nonetheless, a rigorous mathematical derivation of such Langevin-type models for the slow climate dynamics, using the equations of motion in discrete form, is possible as illustrated by the MTV theory presented next.

The Systematic Mode Reduction MTV Methodology

A systematic mathematical methodology to derive Langevin-type equations (4) à la Hasselmann, for the slow climate dynamics from the coupled atmosphere-ocean-land equations of motion, which yields the propagation (or drift) and diffusion terms $\hat{u}(x)$ and $D(x)$ in closed form, is presented in [44,45] by Majda, Timofeyev, and Vanden-Eijnden (MTV).

Starting from the generalized form of the discretized equations of motion

$$\frac{dz}{dt} = Lz + B(z,z) + f(t)$$

where L and B are a linear and a bilinear operators while $f(t)$ represent external forcing, MTV operate the same dichotomy as Hasselmann did of splitting the vector z into slow and fast variables x and y, respectively. However, they introduced a nondimensional parameter $\epsilon = \tau_y/\tau_x$ which measures the degree of time scale separation between the two sets of variables. This leads to the slow-fast coupled system

$$\begin{aligned} dx =& \epsilon^{-1}\left(L_{11}x + L_{12}y\right)dt + B_{11}^1(x,x)dt \\ &+ \epsilon^{-1}\left(B_{12}^1(x,y) + B_{22}^1(y,y)\right)dt \\ &+ Dxdt + F_1(t)dt + \epsilon^{-1}f_1(\epsilon^{-1}t) \\ dy =& \epsilon^{-1}\left(L_{21}x + L_{22}y + B_{12}^2(x,y) + B_{22}^2(y,y)\right)dt \\ &- \epsilon^{-2}\Gamma y dt + \epsilon^{-1}\sigma dW_t + \epsilon^{-1}f_2(\epsilon^{-1}t) \end{aligned} \tag{5}$$

under a few key assumptions, including (1) the nonlinear self interaction term of the fast variables is "parameterized" by an Ornstein-Uhlenbeck process: $B_{22}^2(y,y)dt := -\epsilon^{-1}\Gamma y dt + \sqrt{\epsilon}^{-1}dW_t$ and (2) a small dissipation term $\epsilon Dxdt$ is added to the slow dynamics while (3) the slow variable forcing term assumes slow and fast contributions $f_1(t) = \epsilon F_1(\epsilon t) + f_1(t)$. Moreover, the system in (5) is written in terms of the slow time $t \longrightarrow \epsilon t$.

MTV used the theory of asymptotic expansion applied to the backward Fokker-Plank equation associated with the stochastic differential system in (5) to obtain an effective reduced Langevin equation (4) for the slow variables x in the limit of large separation of time scales $\epsilon \longrightarrow 0$ [44, 45]. The main advantage

of the MTV theory is that unlike Hasselmann's ad hoc formulation, the functional form of the drift and diffusion coefficients, in terms of the slow variables, are obtained and new physical phenomena can emerge from the large-scale feedback besides the assumed stabilization effect. It turns out that the drift term is not always stabilizing, but there are dynamical regimes where growing modes can be excited and, depending on the dynamical configuration, the Langevin equation (4) can support either additive or multiplicative noise.

Even though MTV assumes strict separation of scales, $\epsilon \ll 1$, it is successfully used for a wide range of examples including cases where $\epsilon = O(1)$ [46]. Also in [47], MTV is successfully extended to fully deterministic systems where the requirement that the fast-fast interaction term $B_{22}(y, y)$ in (5) is parameterized by an Ornstein-Uhlenbeck process is relaxed. Furthermore, MTV is applied to a wide range of climate problems. It is used, for instance, in [19] for a realistic barotropic model and extended in [18] to a three-layer quasi-geostrophic model. The example of midlatitude teleconnection patterns where multiplicative noise plays a crucial role is studied in [42]. MTV is also applied to the triad and dyad normal mode (EOF) interactions for arbitrary time series [40].

Stochastic Parametrization

In a typical GCM, the parametrization of unresolved processes is based on theoretical and/or empirical deterministic equations. Perhaps the area where deterministic parameterizations have failed the most is moist convection. GCMs fail very badly in simulating the planetary and intra-seasonal variability of winds and rainfall in the tropics due to the inadequate representation of the unresolved variability of convection and the associated cross-scale interactions behind the multiscale organization of tropical convection [35]. To overcome this problem, some climate scientists introduced random variables to mimic the variability of such unresolved processes. Unfortunately, as illustrated below, many of the existing stochastic parametrizations were based on the assumptions of statistical equilibrium and/or of a stationary distribution for the unresolved variability, which are only valid to some extent when there is scale separation.

The first use of random variables in CGMs appeared in Buizza et al. [6] as means for improving the skill of the ECMWF ensemble prediction system (EPS). Buizza et al. [6] used uniformly distributed random scalars to rescale the parameterized tendencies in the governing equations. Similarly, Lin and Neelin [38] introduced a random perturbation in the tendency of convective available potential energy (CAPE). In [38], the random noise is assumed to be a Markov process of the form $\xi_{t+\Delta t} = \epsilon_t \xi_t + z_t$ where z_t is a white noise with a fixed standard deviation and ϵ_t is a parameter. Plant and Craig [57] used extensive cloud-permitting numerical simulations to empirically derive the parameters for the PDF of the cloud base mass flux itself whose Poisson shape is determined according to arguments drawn from equilibrium statistical mechanics. Careful simulations conducted by Davoudi et al. [10] revealed that while the Poisson PDF is more or less accurate for isolated deep convective clouds, it fails to extend to cloud clusters where a variety of cloud types interact with each other: a crucial feature of organized tropical convection.

Majda and Khouider [43] borrowed an idea from material science [28] of using the Ising model of ferromagnetization to represent convective inhibition (CIN). An order parameter σ, defined on a rectangular lattice, embedded within each horizontal grid box of the climate model, takes values 1 or 0 at a given site, according to whether there is CIN or there is potential for deep convection (PAC). The lattice model makes transitions at a given site according to intuitive probability rules depending both on the large-scale climate model variables and on local interactions between lattice sites based on a Hamiltonian energy principle. The Hamiltonian is given by

$$H(\sigma, U) = -\frac{1}{2}\sum_{x,y} J(|x-y|)\sigma(x)\sigma(y) + h(U)\sum_{x}\sigma_x$$

where $J(r)$ is the local interaction potential and $h(U)$ is the external potential which depends on the climate variables U and where the summations are taken over all lattice sites x, y. A transition (spin-flip by analogy to the Ising model of magnetization) occurs at a site y if for a small time τ, we have $\sigma_{t+\tau}(y) = 1 - \sigma_t(y)$ and $\sigma_{t+\tau}(x) = \sigma_t(x)$ if $x \neq y$. Transitions occur at a rate $C(y, \sigma, U)$ set by Arrhenius dynamics: $C(x, \sigma, U) = \frac{1}{\tau_I}\exp(-\Delta_x H(\sigma, U))$ if $\sigma_x = 0$ and $C(x, \sigma, U) = \frac{1}{\tau_I}$ if $\sigma_x = 1$ so that the resulting Markov process satisfies detailed balance with respect to the Gibbs distribution $\mu(\sigma, U) \propto \exp(-H(\sigma, U)$. Here

$\Delta_x H(\sigma, U)) = H(\sigma + [1-\sigma(x)]e_x), U) - H(\sigma, U) = -\sum_z J(|x-z|)\sigma(z) + h(U)$ with $e_x(y) = 1$ if $y = x$ and 0 otherwise.

For computational efficiency, a coarse graining of the stochastic CIN model is used in [34] to derive a stochastic birth-death process for the mesoscopic area coverage $\eta_X = \sum_{x \in X} \sigma(x)$ where X represents a generic site of a mesoscopic lattice, which in practice can be considered to be the GCM grid. The stochastic CIN model is coupled to a toy GCM where it is successfully demonstrated how the addition of such a stochastic model could improve the climatology and waves dynamics in a deficient GCM [34, 42].

This Ising-type modeling framework is extended in [33] to represent the variability of organized tropical convection (OTC). A multi-type order parameter is introduced to mimic the multimodal nature of OTC. Based on observations, tropical convective systems (TCS) are characterized by three cloud types, cumulus congestus whose height does not exceed the freezing level develop when the atmosphere is dry, and there is convective instability, positive CAPE. In return congestus clouds moisten the environment for deep convective towers. Stratiform clouds that develop in the upper troposphere lag deep convection as a natural freezing phase in the upper troposphere. Accordingly, the new order parameter σ takes the multiple values 0,1,2,3, on a given lattice site, according to whether the given site is, respectively, clear sky or occupied by a congestus, deep, or stratiform cloud.

Similar Arrhenius-type dynamics are used to build transition rates resulting in an ergodic Markov process with a well-defined equilibrium measure. Unphysical transitions of congestus to stratiform, stratiform to deep, stratiform to congestus, clear to stratiform, and deep to congestus were eliminated by setting the associated rates to zero. When local interactions are ignored, the equilibrium measure and the transition rates depend only on the large-scale climate variables U where CAPE and midlevel moisture are used as triggers and the coarse-graining process is carried with exact statistics. It leads to a multidimensional birth-death process with immigration for the area fractions of the associated three cloud types. The stochastic multicloud model (SMCM) is used very successfully in [20, 21] to capture the unresolved variability of organized convection in a toy GCM. The simulation of convectively coupled gravity waves and mean climatology were improved drastically when compared to their deterministic counterparts. The realistic statistical behavior of the SMCM is successfully assessed against observations in [56]. Local interaction effects are reintroduced in [32] where a coarse-graining approximation based on conditional expectation is used to recover the multidimensional birth-death process dynamics with local interactions. A Bayesian methodology for inferring key parameters for the SMCM is developed and validated in [12]. A review of the basic methodology of the CIN and SMCM models, which is suitable for undergraduates, is found in [31].

A systematic data-based methodology for inferring a suitable stochastic process for unresolved processes conditional on resolved model variables was proposed in [9]. The local feedback from unresolved processes on resolved ones is represented by a small Markov chain whose transition probability matrix is made dependent on the resolved-scale state. The matrix is estimated from time series data that is obtained from highly resolved numerical simulations or observations. This approach was developed and successfully tested on the Lorenz '96 system [39] in [9]. [16] applied it to parameterize shallow cumulus convection, using data from large eddy simulation (LES) of moist atmospheric convection. A two-dimensional lattice, with at each lattice node a Markov chain, was used to mimic (or emulate) the convection as simulated by the high-resolution LES model, at a fraction of the computational cost.

Subsequently, [15] combined the conditional Markov chain methodology with elements from the SMCM [33]. They applied it to deep convection but without making use of the Arrhenius functional forms of the transition rates in terms of the large-scale variables (as was done in [33]). Similar to [16], LES data was used for estimation of the Markov chain transition probabilities. The inferred stochastic model in [15] was well capable of generating cloud fractions very similar to those observed in the LES data. While the main cloud types of the original SMCM were preserved, an important improvement in [15] resides in the addition of a fifth state for shallow cumulus clouds. As an experiment, direct spatial coupling of the Markov chains on the lattice was also considered in [15]. Such coupling amounts to the structure of a stochastic cellular automaton (SCA). Without this direct coupling, the Markov chains are still coupled,

but indirectly, through their interaction with the large-scale variables (see, e.g., [9]).

References

1. AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., Sorooshian, S. (eds.): Extremes in a Changing Climate, p. 423. Springer, Dordrecht (2013)
2. Arnold, L.: Hasselmann's program revisited: the analysis of stochasticity in deterministic climate models. In: Imkeller, P., von Storch, J.-S. (eds) Stochastic Climate Models. Birkhäuser, Basel (2001)
3. Bellone, E., Hughes, J.P., Guttorp, P.: A Hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. Clim. Res. **15**, 1–12 (2000)
4. Berner, J.: Linking nonlinearity and non-Gaussianity of planetary wave behavior by the Fokker-Planck equation. J. Atmos. Sci. **62**, 2098–2117 (2005)
5. Bond, N.A., Harrison, D.E.: The pacific decadal oscillation, air-sea interaction and central north pacific winter atmospheric regimes. Geophys. Res. Lett. **27**, 731–734 (2000)
6. Buizza, R., Milleer, M., Palmer, T.N.: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Quart. J. R. Meteorol. Soc. **125**, 2887–2908 (1999)
7. Coles, S.: An Introduction to Statistical Modeling of Extreme Values, p. 208. Springer, London (2001); Statistical Methods in the Atmospheric Sciences, 3rd edn., p. 704. Academic, Oxford
8. Crommelin, D.T.: Observed non-diffusive dynamics in large-scale atmospheric flow. J. Atmos. Sci. **61**, 2384–239 (2004)
9. Crommelin, D.T., Vanden-Eijnden, E.: Subgrid-scale parameterization with conditional Markov chains. J. Atmos. Sci. **65**, 2661–2675 (2008)
10. Davoudi, J., McFarlane, N.A., Birner, T.: Fluctuation of mass flux in a cloud-resolving simulation with interactive radiation. J. Atmos. Sci. **67**, 400–418 (2010)
11. de Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction, p. 417. Springer, New York (2006)
12. de la Chevrotière, M., Khouider, B., Majda, A.: Calibration of the stochastic multicloud model using Bayesian inference. SIAM J. Sci. Comput. **36**(3), B538–B560 (2014)
13. DelSole, T.: Stochastic models of quasi-geostrophic turbulence. Surv. Geophys. **25**, 107–149 (2004)
14. DelSole, T., Farrel, B.F.: Quasi-linear equilibration of a thermally maintained, stochastically excited jet in a quasi-geostrophic model. J. Atmos. Sci. **53**, 1781–1797 (1996)
15. Dorrestijn, J., Crommelin, D.T., Biello, J.A., Böing, S.J.: A data-driven multicloud model for stochastic parameterization of deep convection. Phil. Trans. R. Soc. A **371**, 20120374 (2013)
16. Dorrestijn, J., Crommelin, D.T., Siebesma, A.P., Jonker, H.J.J.: Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data. Theor. Comput. Fluid Dyn. **27**, 133–148 (2013)
17. Frankignoul, C., Hasselmann, K.: Stochastic climate models, Part II application to sea-surface temperature anomalies and thermocline variability. Tellus **29**, 289–305 (1977)
18. Franzke, C., Majda, A.: Low order stochastic mode reduction for a prototype atmospheric GCM. J. Atmos. Sci. **63**, 457–479 (2006)
19. Franzke, C., Majda, A., Vanden-Eijnden, E.: Low-order stochastic mode reduction for a realistic barotropic model climate. J. Atmos. Sci. **62**, 1722–1745 (2005)
20. Frenkel, Y., Majda, J., Khouider, B.: Using the stochastic multicloud model to improve tropical convective parameterization: a paradigm example. J. Atmos. Sci. **69**, 1080–1105 (2012)
21. Frenkel, Y., Majda, J., Khouider, B.: Stochastic and deterministic multicloud parameterizations for tropical convection. Clim. Dyn. **41**, 1527–1551 (2013)
22. Giannakis, D., Majda, A.J.: Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. Proc. Natl. Acad. Sci. USA **109**, 2222–2227 (2012)
23. Hasselmann, K.: Stochastic climate models. Part I, theory. Tellus **28**, 473–485 (1976)
24. Hasselmann, K.: PIPs and POPs: the reduction of complex dynamical systems using principal interaction and oscillation patterns. J. Geophys. Res. **93**(D9), 11015–11021 (1988)
25. Horenko, I.: Nonstationarity in multifactor models of discrete jump processes, memory, and application to cloud modeling. J. Atmos. Sci. **68**, 1493–1506 (2011)
26. Imkeller P., von Storch J.-S. (eds.): Stochastic Climate Models. Birkhäuser, Basel (2001)
27. Jolliffe, I.T.: Principal Component Analysis, 2nd edn. Springer, New York (2002)
28. Katsoulakis, M., Majda, A.J., Vlachos, D.: Coarse-Grained stochastic processes and Monte-Carlo simulations in lattice systems. J. Comp. Phys. **186**, 250–278 (2003)
29. Katz, R.W., Naveau P.: Editorial: special issue on statistics of extremes in weather and climate. Extremes **13**, 107–108 (2010)
30. Kessler, M., Lindner, A., Sørensen, M. (eds.): Statistical Methods for Stochastic Differential Equations. CRC, Boca Raton (2012)
31. Khouider, B.: Markov-jump stochastic models for organized tropical convection. In: Yang, X.-S. (ed.) Mathematical Modeling With Multidisciplinary Applications. Wiley, (2013). ISBN: 978-1-1182-9441-3
32. Khouider, B.: A stochastic coarse grained multi-type particle interacting model for tropical convection. Commun. Math. Sci. **12**, 1379–1407 (2014)
33. Khouider, B., Biello, J., Majda, A.J.: A stochastic multicloud model for tropical convection. Commun. Math. Sci. **8**, 187–216 (2010)
34. Khouider, B., Majda, A.J., Katsoulakis, M.: Coarse grained stochastic models for tropical convection. Proc. Natl. Acad. Sci. USA **100**, 11941–11946 (2003)
35. Khouider, B., Majda, A.J., Stechmann, S.: Climate science in the tropics: waves, vortices, and PDEs. Nonlinearity **26**, R1-R68 (2013)
36. Kleeman, R., Moore, A.: A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. J. Atmos. Sci. **54**, 753–767 (1997)
37. Kondrasov, D., Krastov, S., Gill, M.: Empirical mode reduction in a model of extra-tropical low-frequency variability. J. Atmos. Sci. **63**, 1859–1877 (2006)

38. Lin, J.B.W., Neelin, D.: Toward stochastic deep convective parameterization in general circulation models. Geophy. Res. Lett. **27**, 3691–3694 (2000)
39. Lorenz, E.N.: Predictability: a problem partly solved. In: Proceedings, Seminar on Predictability ECMWF, Reading, vol. 1, pp. 1–18 (1996)
40. Majda, A., Franzke, C., Crommelin, D.: Normal forms for reduced stochastic climate models. Proc. Natl. Acad. Sci. USA **16**, 3649–3653 (2009)
41. Majda, A.J., Franzke, C.L., Fischer, A., Crommelin, D.T.: Distinct metastable atmospheric regimes despite nearly Gaussian statistics: a paradigm model. Proc. Natl. Acad. Sci. USA **103**, 8309–8314 (2006)
42. Majda, A.J., Franzke, C.L., Khouider, B.: An applied mathematics perspective on stochastic modelling for climate. Phil. Trans. R. Soc. **366**, 2427–2453 (2008)
43. Majda, A.J., Khouider, B.: Stochastic and mesoscopic models for tropical convection. Proc. Natl. Acad. Sci. **99**, 1123–1128 (2002)
44. Majda, A.J., Timofeyev, I., Vanden-Eijnden, E.: Models for stochastic climate prediction. Proc. Natl. Acad. Sci. **96**, 14687–14691 (1999)
45. Majda, A.J., Timofeyev, I., Vanden-Eijnden, E.: A mathematical framework for stochastic climate models. Commun. Pure Appl. Math. **LIV**, 891–974 (2001)
46. Majda, A.J., Timofeyev, I., Vanden-Eijnden, E.: A priori tests of a stochastic mode reduction strategy. Physica D **170**, 206–252 (2002)
47. Majda A., Timofeyev, I., Vanden-Eijnden, E.: Stochastic models for selected slow variables in large deterministic systems. Nonlinearity **19**, 769–794 (2006)
48. Mo, K.C., Ghil, M.: Statistics and dynamics of persistent anomalies. J. Atmos. Sci. **44**, 877–901 (1987)
49. Monahan, A.H.: Nonlinear principal component analysis by neural networks: theory and application to the Lorenz system. J. Clim. **13**, 821–835 (2000)
50. Monahan, A.H., Fyfe, J.C., Ambaum, M.H.P., Stephenson, D.B., North, G.R.: Empirical orthogonal functions: the medium is the message. J. Clim. **22**, 6501–6514 (2009)
51. Newman, M., Sardeshmukh, P., Penland, C.: Stochastic forcing of the wintertime extratropical flow. J. Atmos. Sci. **54**, 435–455 (1997)
52. Palmer, T., Williams, P. (eds.): Stochastic Physics and Climate Modelling. Cambridge University Press, Cambridge (2009)
53. Pasmanter, R.A., Timmermann, A.: Cyclic Markov chains with an application to an intermediate ENSO model. Nonlinear Process. Geophys. **10**, 197–210 (2003)
54. Penland, C., Ghil, M.: Forecasting northern hemisphere 700-mb geopotential height anomalies using empirical normal modes. Mon. Weather Rev. **121**, 2355–2372 (1993)
55. Penland, C., Sardeshmukh, P.D.: The optimal growth of tropical sea surface temperature anomalies. J. Clim. **8**, 1999–2024 (1995)
56. Peters, K., Jakob, C., Davies, L., Khouider, B., Majda, A.: Stochastic behaviour of tropical convection in observations and a multicloud model. J. Atmos. Sci. **70**, 3556–3575 (2013)
57. Plant, R.S., Craig, G.C.: A stochastic parameterization for deep convection based on equilibrium statistics. J. Atmos. Sci. **65**, 87–105 (2008)
58. Prakasa Rao, B.L.S.: Statistical Inference for Diffusion Type Processes. Arnold Publishers, London (1999)
59. Preisendorfer, R.W.: Principal Component Analysis in Meteorology and Oceanography. Elsevier, Amsterdam (1988)
60. Sura, P., Newman, M., Penland, C., Sardeshmukh, P.: Multiplicative noise and non-Gaussianity: a paradigm for atmospheric regimes. J. Atmos. Sci. **62**, 1391–1409 (2005)
61. Von Storch, H., Zwiers, F.W.: Statistical Analysis in Climate Research. Cambridge University Press, Cambridge (1999)
62. Wilks, D.S.: Statistical Methods in the Atmospheric Sciences, 3rd edn., p. 704. Academic, Oxford (2011)
63. Zucchini, W., Guttorp, P.: A Hidden Markov model for space-time precipitation. Water Resour. Res. **27**, 1917–1923 (1991)

Stochastic Eulerian-Lagrangian Methods

Paul J. Atzberger
Department of Mathematics, University of California Santa Barbara (UCSB), Santa Barbara, CA, USA

Synonyms

Fluid-structure interaction; Fluid dynamics; Fluctuating hydrodynamics; Immersed Boundary Method; SELM; Statistical mechanics; Stochastic Eulerian Lagrangian method; Thermal fluctuations

Abstract

We present approaches for the study of fluid-structure interactions subject to thermal fluctuations. A mechanical description is utilized combining Eulerian and Lagrangian reference frames. We establish general conditions for the derivation of operators coupling these descriptions and for the derivation of stochastic driving fields consistent with statistical mechanics. We present stochastic numerical methods for the fluid-structure dynamics and methods to generate efficiently the required stochastic driving fields. To help establish the validity of the proposed approach, we perform analysis of the invariant probability distribution of the stochastic dynamics and relate our results to statistical mechanics. Overall, the presented approaches are expected to be applicable to a wide variety of systems involving fluid-structure interactions subject to thermal fluctuations.