Systematic stochastic modeling of atmospheric variability

Christian Franzke
Meteorological Institute
Center for Earth System Research and Sustainability
University of Hamburg

Daniel Peavoy and Gareth Roberts (Warwick)
Daan Crommelin (CWI) and Andy Majda (NYU)
Outline

1) Motivation

2) Stochastic Mode Reduction

3) Physical Constraints

4) Bayesian Parameter Estimation

5) Results

6) Summary
Scales

Typical Sizes

- Global Scale ~5000 km
- Synoptic Scale ~2000 km
- Mesoscale ~20 km
- Microscale ~2m

Typical Time Scale
- Seconds to minutes
- Minutes to hours
- Hours to days
- Days
- Days to a week and more
  Intraseasonal, interannual
decal and longer

Time and Space Scales in the Climate System

- Coupled Atmosphere–Ocean System
- Teleconnection Pattern (E.g. NAO, PNA)
- Synoptic Weather Systems
- Mesoscale Convective Systems
- Boundary Layer Processes
- Small Turbulent Eddies
Long-Range Forecasts

North Atlantic Oscillation

Positive Mode

Cold & more Sea-Ice
Warm & Wet

Jet Stream

Strong Low

Cool & Dry

Strong High

Negative Mode

Warm & Less Sea-Ice
Cold & Snowy

Jet Stream

“Blocking”

Enhanced Trough

Warm & Wet

Cold & Dry

Associated SST Patterns
(~12-14 yr period)
Reduced Stochastic Climate Models

→ Computationally much cheaper

→ Capture essential dynamics

- Improved extended range forecasting
- Large ensemble forecasting
- Long-term climate studies (e.g. paleo-climate)
- Long control simulations to estimate extremes
- Extreme Event Prediction
Reduced Order Models

Slow Climate Modes: $\frac{d\alpha(t)}{dt} = G(\alpha(t), \beta(t))$

Fast Weather Modes: $\frac{d\beta(t)}{dt} = \frac{1}{\varepsilon} F(\alpha(t), \beta(t))$

Assumption: Time scale separation

Reduced Model: $d\alpha(t) = G_{\text{eff}}(\alpha(t))dt + D_{\text{eff}}(\alpha(t))dW$
Reduced Order Models

Assumption:
Time scale separation

\[ \Delta t_C \]

\[ \Delta t_F \]
Stochastic methods are a crucial area in contemporary climate research and are increasingly being used in comprehensive weather and climate prediction models as well as reduced order climate models. Stochastic methods are used as subgrid-scale parameterizations (SSPs) as well as for model error representation, uncertainty quantification, data assimilation, and ensemble prediction. The need to use stochastic approaches in weather and climate models arises because we still cannot resolve all necessary processes and scales in comprehensive numerical weather and climate prediction models. In many practical applications one is mainly interested in the largest and potentially predictable scales and not necessarily in the small and fast scales. For instance, reduced order models can simulate and predict large-scale modes. Statistical mechanics and dynamical systems theory suggest that in reduced order models the impact of unresolved degrees of freedom can be represented by suitable combinations of deterministic and stochastic components and non-Markovian (memory) terms. Stochastic approaches in numerical weather and climate prediction models also lead to the reduction of model biases. Hence, there is a clear need for systematic stochastic approaches in weather and climate modeling. In this review, we present evidence for stochastic effects in laboratory experiments. Then we provide an overview of stochastic climate theory from an applied mathematics perspective. We also survey the current use of stochastic methods in comprehensive weather and climate prediction models and show that stochastic parameterizations have the potential to remedy many of the current biases in these comprehensive models. © 2014 John Wiley & Sons, Ltd.
Outline

1) Motivation

2) Stochastic Mode Reduction

3) Physical Constraints

4) Bayesian Parameter Estimation

5) Results

6) Summary
Stochastic Mode Reduction

Equations of motions

\[ \frac{du}{dt} = F + Lu + I(u, u) \]

Energy conservation

\[ u \ast I(u, u) = 0 \]

Decompose \( u \) into \((x, y)\)

\( x: \) slow component

\( y: \) fast component
Stochastic Mode Reduction

\[
\frac{dx_i}{dt} = F_i + \sum_j (L_{ij} x_j + I_{ij} x_i x_j)
\]
\[
+ \sum_p (L_{ip} y_p + I_{ip}^M y_p x_i + I_{ip}^A y_p y_p + \ldots)
\]
\[
\frac{dy_p}{dt} = \sum_j (L_{pj} x_j + F_p - I_{pj}^M x_j x_j - I_{pj}^A y_p x_j + \ldots)
\]
\[
- \frac{\gamma_p}{\varepsilon} y_p + \frac{\sigma_p}{\sqrt{\varepsilon}} \dot{W}_p
\]

e.g. Ornstein-Uhlenbeck Process
→ Continuous time version of AR(1)

Fast nonlinear interactions: $I(y,y)$

Stochastic Mode Reduction

Equations of motions

\[ \frac{du}{dt} = F + Lu + I(u, u) \]

Energy conservation

\[ u \ast I(u, u) = 0 \]

Mode Reduction

Reduced Model:

\[ dx = (\tilde{F} + \tilde{L} x + \tilde{I}(x, x) + M(x, x, x)) dt + \sigma_A dW_A + \sigma_M(x) dW_M \]
\[
\frac{d\vec{x}}{dt} = D\vec{x} + (\omega + a_1 x_1 + a_2 x_2)\vec{x}^\perp + \vec{F} + \vec{L} y + \begin{bmatrix} b_1 x_2 y \\ b_2 x_1 y \end{bmatrix},
\]

\[
\frac{dy}{dt} = -\vec{L} \cdot \vec{x} + b_3 x_1 x_2 + F_3 - \frac{\gamma}{\varepsilon} y + \frac{\sigma}{\sqrt{\varepsilon}} \dot{W}(t).
\]
Stochastic Mode Reduction

\[
\frac{d\bar{x}}{dt} = D\bar{x} + (\omega + a_1 x_1 + a_2 x_2)\bar{x}^\perp + \bar{F} + \bar{L}y + \begin{bmatrix} b_1 x_2 y \\ b_2 x_1 y \end{bmatrix},
\]

\[
\frac{dy}{dt} = -\bar{L} \cdot \bar{x} + b_3 x_1 x_2 + F_3 - \frac{\gamma}{\varepsilon}y + \frac{\sigma}{\sqrt{\varepsilon}} \dot{W}(t).
\]

Solve for y:

\[y(t) = e^{-\gamma t/\varepsilon}y(0) + \int_0^t e^{-\gamma(t-s)/\varepsilon} \left[ b_3 x_1(s)x_2(s) - \bar{L} \cdot \bar{x}(s) + F_3(s) \right] ds + g(t)\]

\[g(t) = \frac{\sigma}{\sqrt{\varepsilon}} \int_0^t e^{-\gamma(t-s)/\varepsilon} dW(s).\]
Stochastic Mode Reduction

\[
\frac{d\tilde{x}}{dt} = D\tilde{x} + (\omega + a_1x_1 + a_2x_2)\tilde{x}^\perp + \tilde{F} + \tilde{L}y + \begin{bmatrix} b_1x_2y \\ b_2x_1y \end{bmatrix},
\]

\[
\frac{dy}{dt} = -\tilde{L} \cdot \tilde{x} + b_3x_1x_2 + F_3 - \frac{\gamma}{\varepsilon}y + \frac{\sigma}{\sqrt{\varepsilon}}\dot{W}(t).
\]

For \( \varepsilon \to 0 \),

\[
\int_0^t e^{-\gamma(t-s)/\varepsilon}\left[b_3x_1(s)x_2(s) - \tilde{L} \cdot \tilde{x}(s) + F_3\right] ds \\
\quad \to \frac{\varepsilon}{\gamma}\left[b_3x_1(t)x_2(t) - \tilde{L} \cdot \tilde{x}(t) + F_3\right].
\]

\[
g(t) dt \to \sqrt{\frac{\sigma}{\gamma}} dW(t).
\]
Stochastic Mode Reduction

\[ \frac{d \tilde{x}}{dt} = D \tilde{x} + (\omega + a_1 x_1 + a_2 x_2) \tilde{\mathbf{x}}^\perp + \tilde{F} + \tilde{L} y + \begin{bmatrix} b_1 x_2 y \\ b_2 x_1 y \end{bmatrix}, \]

\[ \frac{dy}{dt} = -\tilde{L} \cdot \tilde{x} + b_3 x_1 x_2 + F_3 - \frac{\gamma}{\varepsilon} y + \frac{\sigma}{\sqrt{\varepsilon}} \dot{W}(t). \]

Plug into equation for \( \mathbf{x} \):

\[
\begin{align*}
    d\tilde{x}(t) &= \tilde{F}_0(\tilde{x}; F) \, dt + \frac{\varepsilon}{\gamma} \tilde{L} [b_3 x_1 x_2 - \tilde{L} \cdot \tilde{x} + F_3] \, dt \\
    &\quad + \frac{\varepsilon}{\gamma} \left[ b_1 x_2(t) \{ b_3 x_1(t) x_2(t) - \tilde{L} \cdot \tilde{x}(t) + F_3 \} \right] \, dt \\
    &\quad + \frac{\sqrt{\varepsilon} \sigma}{\gamma} \left[ \begin{bmatrix} L_1 + b_1 x_2(t) \\ L_2 + b_2 x_1(t) \end{bmatrix} \right] \circ dW(t). 
\end{align*}
\]
Stochastic Mode Reduction

\[
\begin{align*}
\frac{d\bar{x}}{dt} &= D\bar{x} + (\omega + a_1x_1 + a_2x_2)\bar{x}^\perp + \vec{F} + \vec{L}y + \begin{bmatrix} b_1x_2y \\ b_2x_1y \end{bmatrix}, \\
\frac{dy}{dt} &= -\vec{L} \cdot \bar{x} + b_3x_1x_2 + F_3 - \frac{\gamma}{\varepsilon}y + \frac{\sigma}{\sqrt{\varepsilon}}\dot{W}(t).
\end{align*}
\]

Plug into equation for x:

\[
\begin{align*}
d\bar{x}(t) &= \vec{F}_0(\bar{x}; F) dt + \frac{\varepsilon}{\gamma}\vec{L}[b_3x_1x_2 - \vec{L} \cdot \bar{x} + F_3] dt \\
&\quad + \frac{\varepsilon}{\gamma} \left[ b_1x_2(t)\{b_3x_1(t)x_2(t) - \vec{L} \cdot \bar{x}(t) + F_3\} \right] dt \\
&\quad + \frac{\varepsilon}{\gamma} \left[ b_2x_1(t)\{b_3x_1(t)x_2(t) - \vec{L} \cdot \bar{x}(t) + F_3\} \right] dt \\
&\quad + \frac{\varepsilon}{\gamma} \sqrt{\frac{\varepsilon}{\gamma}} \left[ L_1 + b_1x_2(t) \right] \circ dW(t).
\end{align*}
\]

CAM Noise
Stochastic Mode Reduction

\[
\frac{d\tilde{x}}{dt} = D\tilde{x} + (\omega + a_1 x_1 + a_2 x_2)\tilde{x}^\perp + \tilde{F} + \tilde{L}y + \begin{bmatrix} b_1 x_2 y \\ b_2 x_1 y \end{bmatrix},
\]

\[
\frac{dy}{dt} = -\tilde{L} \cdot \tilde{x} + b_3 x_1 x_2 + F_3 - \frac{\gamma}{\varepsilon} y + \frac{\sigma}{\sqrt{\varepsilon}} \dot{W}(t).
\]

Plug into equation for \(x\):

\[
d\tilde{x}(t) = \tilde{F}_0(\tilde{x}; F) \, dt + \frac{\varepsilon}{\gamma} \tilde{L}[b_3 x_1 x_2 - \tilde{L} \cdot \tilde{x} + F_3] \, dt
\]

\[
+ \frac{\varepsilon}{\gamma} \left[ b_1 x_2(t)\{b_3 x_1(t)x_2(t) - \tilde{L} \cdot \tilde{x}(t) + F_3\} \right] \, dt
\]

\[
+ \sqrt{\varepsilon} \frac{\sigma}{\gamma} \left[ L_1 + b_1 x_2(t) \right] \circ dW(t).
\]

Cubic Term

CAM Noise
Outline

1) Motivation

2) Stochastic Mode Reduction

3) Physical Constraints

4) Bayesian Parameter Estimation

5) Results

6) Summary
\[
\frac{dx_i}{dt} = (\text{Bare truncation})_i + A_i \tilde{x} \\
+ F_i + \tilde{B}_i(\tilde{x}, \tilde{x}) + \sum_p \sigma_{ip}^{A} \dot{W}_{ip}^{A} \\
- \sum_p \frac{I_{ip}^{M^2}}{\gamma_p} x_i^3 + \sum_{j \neq i, p} \frac{I_{ip}^{M} I_{pj}^{M}}{\gamma_p} x_i x_j^2 \\
+ \sum_p \frac{\sigma_p}{\gamma_p} (L_{ip} - I_{ip}^M x_i) \circ \dot{W}_p
\]
Constraints on Stochastic Climate Models

\[
\frac{1}{2} \frac{dx_i^2}{dt} = \left( \sum_{j \neq i} \sum_p \frac{A_{ijp} A_{pij}}{\gamma_p} x_j^2 - \sum_p \frac{I_{ip}^2}{\gamma_p} x_i^2 \right) x_i^2
\]

\[
= \sum_i \left( \sum_{j \neq i} \tilde{A}_{ij} x_j^2 - \tilde{I}_i x_i^2 \right) x_i^2
\]

\[
E = |\bar{x}|^2, \quad \tilde{A}_{ij} = \sum_p \frac{I_{ip}^M I_{pj}^M}{\gamma_p} \quad \text{and} \quad \tilde{I}_i = \sum_p \frac{I_{ip}^{M^2}}{\gamma_p}
\]

Only considering cubic terms

Majda et al. 2009; Peavoy et al. 2015
Constraints on Stochastic Climate Models

\[ \frac{1}{2} \frac{dE}{dt} = (\bar{x}^2)^T (\tilde{A} - \tilde{I}E) \bar{x}^2 = (\bar{x}^2)^T Q \bar{x}^2 \]

Stability: \( \frac{dE}{dt} < 0 \)

- Quadratic form \( Q \) is negative-definite
- Allows the system to be linearly unstable

Majda et al. 2009; Peavoy et al. 2015
Physical Constraints

Without constraint about 40% of parameter estimates lead to unstable solutions

Peavoy et al. 2015
Outline

1) Motivation

2) Stochastic Mode Reduction

3) Physical Constraints

4) Bayesian Parameter Estimation

5) Results

6) Summary
Bayesian Parameter Estimation Procedure

\[ d\mathbf{X}_t = \mu(\mathbf{X}_t, \theta)dt + a(\mathbf{X}_t, \theta)dW_t, \quad \mathbf{X}_0 = \mathbf{x}_0, \quad t \in [0, T] \]

Discretisation: Euler-Maruyama scheme

\[ \mathbf{X}_{k+1} = \mathbf{X}_k + \mu(\mathbf{X}_k, \theta)\Delta t_k + a(\mathbf{X}_k, \theta)(W_{k+1} - W_k) \]

Likelihood based parameter estimation (MCMC)

\[ L^\text{Euler}_N(\theta) = -\frac{1}{2\Delta t_k} \sum_{k=0}^{N-1} (\mathbf{x}_{k+1} - \mathbf{x}_k - \Delta t_k \mu(\mathbf{x}_k, \theta))^T \Sigma(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k - \Delta t_k \mu(\mathbf{x}_k, \theta)) \]

Imputing of Data

Modified Linear Bridge

\[ \delta = \Delta / m \]

Peavoy et al. 2015
How to sample negative-definite matrices?

Wishart Distribution

Truncated Normal Algorithm:
A n×n matrix is negative definite if and only if all k≤n leading principal minors obey |M^k|(-1)^k > 0.
The k-th principal minor is the determinant of the upper left k×k sub-matrix.

for i = 1 to n do
    U_i = 0
    for j = i to n do
        x = -\left(\sum_{k \neq i}^{j} (-1)^{i+k} M_{ik} |M^{(j)}_{\{i\}}| \right) / |M^{(j)}_{\{-i\}}|  
    end for
    if x < U_i then
        U_i = x
    end if
end for

Diagonal Elements

for i = 1 to n do
    for j = i + 1 to n do
        u^+ = \infty
        u^- = -\infty
        for k = j to n do
            Calculate a_{ij}^{(k)}, b_{ij}^{(k)} and c_{ij}^{(k)} and solve a_{ij}^{(k)} x^2 + b_{ij}^{(k)} x + c_{ij}^{(k)} = 0.
            Set mn = min(x_1, x_2) and mx = max(x_1, x_2)
            end for
            if mx < u^+ then
                u^+ = mx
            end if
            if mn > u^- then
                u^- = mn
            end if
        end for
        M_{ij} \sim N_+(\mu_{ij}, u^-, u^+, \sigma_{ij}^2)
    end for
end for

Off-Diagonal Elements
Modelling Memory Effects via Latent Variables

\[
dx = (a_1 + a_2 x + a_3 x^2 + a_4 x^3) dt + y dt
\]
\[
dy = -\gamma y dt + \sigma dW
\]

Red Noise
Outline

1) Motivation

2) Stochastic Mode Reduction

3) Physical Constraints

4) Bayesian Parameter Estimation

5) Results

6) Summary
\[
\begin{align*}
\frac{dx_1}{dt} &= \frac{b_1}{\epsilon} x_2 y_1 \\
\frac{dx_2}{dt} &= \frac{b_2}{\epsilon} x_1 y_1 \\
\frac{dy_k}{dt} &= \frac{b_3}{\epsilon} x_1 x_2 \delta_{1,k} - \text{Re} \frac{ik}{2\epsilon^2} \sum_{p+q+k=0} \hat{u}_p^* \hat{u}_q^* \\
\frac{dz_k}{dt} &= -\text{Im} \frac{ik}{2\epsilon^2} \sum_{p+q+k=0} \hat{u}_p^* \hat{u}_q^*
\end{align*}
\]

\[\text{Model Reduction}\]

\[
\begin{align*}
 dx_1(t) &= \frac{b_1}{\gamma} (b_3 x_2^2(t) + \frac{\sigma^2}{2\gamma} b_2) x_1(t)dt + \frac{\sigma}{\gamma} b_1 x_2(t)dW_t \\
 dx_2(t) &= \frac{b_2}{\gamma} (b_3 x_1^2(t) + \frac{\sigma^2}{2\gamma} b_1) x_2(t)dt + \frac{\sigma}{\gamma} b_2 x_1(t)dW_t
\end{align*}
\]

Peavoy et al. 2015
Triad Model Example

(a) Stationary distribution for $x_1$

(b) Autocorrelation function for $x_1$

(c) Stationary distribution for $x_2$

(d) Autocorrelation function for $x_2$
Test: Chaotic Lorenz Model

\[
\begin{align*}
\frac{dx}{dt} &= x - x^3 + \frac{4}{90\epsilon}y_2 \\
\frac{dy_1}{dt} &= \frac{10}{\epsilon^2}(y_2 - y_1) \\
\frac{dy_2}{dt} &= \frac{1}{\epsilon^2}(28y_1 - y_2 - y_1y_3) \\
\frac{dy_3}{dt} &= \frac{1}{\epsilon^2}(y_1y_2 - \frac{8}{3}y_3).
\end{align*}
\]

\[dX_t = (a_1 + a_2X_t + a_3X_t^2 + a_4X_t^3)dt + \sigma dW_t\]

Reduced order model:

Peavoy et al. 2015
Flow over topography on a $\beta$-plane

$$\frac{\partial q}{\partial t} + \nabla^\perp \psi \cdot \nabla q + U \frac{\partial q}{\partial x} + \beta \frac{\partial \psi}{\partial x} = 0$$

$$q = \Delta \psi + h, \quad \frac{dU}{dt} = \frac{1}{4\pi^2} \int h \frac{\partial \psi}{\partial x} \, dxdy.$$  

$$dU = (a_1 + a_2 U + a_3 U^2 + a_4 U^3) \, dt + (\sigma_1 + \sigma_2 U) \, dB_t$$

Stationary Distribution

Peavoy et al. 2015
Arctic Oscillation Index

From NCAR-NCEP reanalysis data covering period 1948-2010

Autocorrelation Function
North Atlantic Jet Stream has three persistent states
Persistent states exhibit variability on interannual and
decadal time scales
Propensity of extreme wind speeds depend on
persistent states

**Scientific Themes**
- Instabilities and Predictability
- Multi-Scale Effects and Stochastic Parameterizations
- Data assimilation and Big Data
- Coupled Data Assimilation
- Complex Network Approaches
- Stochastic and Dynamical Systems Approaches
- Nonlinear Time Series Analysis
- Modeling of Extreme Events
- Environmental Risk Analysis and Sustainability

**Invited Speakers**
- Sonia I. Seneviratne (ETH Zürich, Switzerland)
- Sandra Chapman (University of Warwick, UK)
- Marc Bocquet (CERA, Joint laboratory of École des Ponts ParisTech, France)
- Laure Zanna (Oxford University, UK)
- Erik Chavez (Imperial College London, UK)
- Emilio Hernandez-Garcia (Instituto de Física Interdisciplinar y Sistemas Complejos, Spain)
- Mickael Chekroun (UCLA, USA)
- Judith Berner (NCAR, USA)
- Francisco Doblas-Reyes (ICREA-BSC, Spain)

Information and registration: http://www.clisap.de/dames

Deadline: 15.06.2016
Summary

- Normal form for reduced stochastic climate models predict a **cubic nonlinear drift** and a **correlated additive and multiplicative CAM noise**.
- Bayesian Framework for Physics Constrained Parameter Estimation
- Reduced stochastic climate models perform well

References:
- Franzke, C., T. O'Kane, J. Berner, P. Williams and V. Lucarini, 2015: Stochastic Climate Theory and Modelling. WIREs Climate Change, 6, 63-78.