Stochastic Quantification of Missing Mechanisms in Dynamical Systems

Jinqiao Duan
Illinois Institute of Technology
Department of Applied Mathematics
Chicago, IL 60616, USA
www.iit.edu/~duan

Outline

Dynamical systems as mathematical models

- Deterministic dynamical system

$$\frac{du}{dt} = Au + f(u)$$

- Stochastic dynamical system

$$\frac{du}{dt} = Au + f(u) + g(u)\frac{d}{dt}\xi_t$$

with ξ_t a stochastic process

Where does $g(u)\frac{d}{dt}\xi_t$ come from?

Math models are not perfect

Due to lack of knowledge:

- Missing mechanisms (ignored “too small" or “too fast" or “unimportant" processes)
- Unresolved scales (not represented in models)
- **Model uncertainties**

Math Modeling: A procedure of making better and better models

Impact of uncertainty may be certain

- Impact of model uncertainties may **not** be negligible

  For example:
  Noises destroy bifurcation, noises changes invariant sets, noise-induced phenomena, noise-enhanced phenomena

  L. Arnold: **Random Dynamical Systems**
  Imkeller, Blomker, Wanner

- **Risk for ignoring model uncertainties**

How to obtain better math models?

Modeling a scientific phenomenon:

- **One way:** Try to figure out all mechanisms and put them in the math model
- **Another way:** Use observational data to improve the math model

Let data speak! A part of scientific truth is embedded in data!

Observational data are abundant

- Geophysical, biological and other data
- Need: Taking advantage of observational data to validate models
- Need: Appropriate math modeling techniques

More and more data are available!

A deterministic PDE model

A partial differential equation (PDE):

$$u_t = Au + N(u),$$

where A is a linear (unbounded) differential operator, and N is a nonlinear function of $u(x,t)$ with $x \in D$ and $t > 0$, and satisfies a local Lipschitz condition.

$u(x,t)$: Model prediction (“solution of the model")

How to validate a model: **Observational data**

Does u match with observation $\tilde{u}$?

Model uncertainty

Some mechanisms are not represented due to the lack of scientific understanding

“Unresolved" scales (e.g., microscopic, fast, or small scales)

Impact of these unresolved mechanisms on the resolved ones: May be delicate

Quantify model uncertainty: using data

When the prediction u deviates from the observational data $\tilde{u}$, we then need to modify the model:

$$\tilde{u}_t = A\tilde{u} + N(\tilde{u}) + F(\tilde{u})$$

where the model uncertainty $F(\tilde{u})$ may be a fluctuating process, as the observational data $\tilde{u}$ is so (i.e., various realizations/measurements).

Quantifying model uncertainty: using data

More specifically: discretize the PDE model; fitting data; discrepancy gives the model uncertainty

$$u_t = Au + N(u),$$

Objective

$$\tilde{u}_t = A\tilde{u} + N(\tilde{u}) + F(\tilde{u}),$$

Model uncertainty: $F(\tilde{u})$

- Model uncertainty is determined by historical observation data
- How to quantify it, so that we may predict its impact and get an improved model?

How to rigourously quantify missing mechanisms $F(\tilde{u})$?

Depending on:

1. Statistic properties of $F(\tilde{u})$
2. Our ability for stochastic calculus

- Gaussian process: Brownian motion B_t
- Non-Gaussian process: Lévy motion L_t
- Non-Markov process: fractional Brownian motion B_t^H

Brownian motion B_t

- Independent increments: $B_{t_2} - B_{t_1}$ and $B_{t_3} - B_{t_2}$ independent
- Stationary increments with $B_t - B_s \sim N(0, t-s)$
  In particular, $B_t \sim N(0, t)$
- Continuous sample paths, but nowhere differentiable

Reference:
I. Karatzas and S. E. Shreve,
Brownian Motion and Stochastic Calculus

A sample path for Brownian motion $B_t(\omega)$

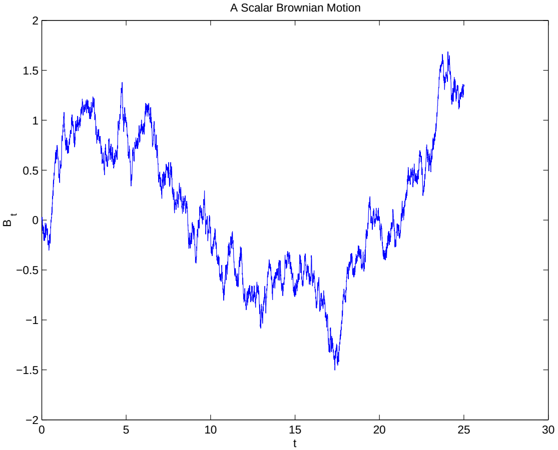


Figure: Continuous path, but nowhere differentiable

Fractional Brownian motion B_t^H

Hurst parameter: $H \in (0, 1)$

- **Dependent** increments: $B_{t_2} - B_{t_1}$ and $B_{t_3} - B_{t_2}$ *dependent*
  Covariance: $\mathbb{E}[B^H(t)B^H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$
- Stationary increments: $\mathbb{E}|B_t^H - B_s^H|^2 = |t-s|^{2H}$
- Continuous sample paths, but nowhere differentiable

$H = \frac{1}{2}$: Usual Brownian motion

Nualart: *Contemp. Math.*, 2003

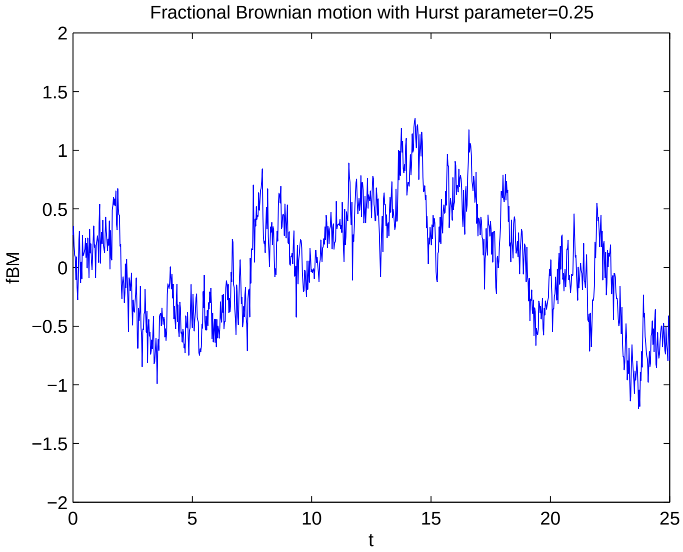


Figure: *Fractional Brownian motion* $B^H(t)$*, with* $H = 0.25$

Stochastic integration theory

- Brownian motion: $H = \frac{1}{2}$
  - Semimartingale, Markov
  - No bounded variation, no Riemann-Stieltjes integral
  - Ito integral
- Fractional Brownian motion: $H \neq \frac{1}{2}$
  - Not semimartingale, not Markov
  - No Ito integral
  - No bounded variation but ok for Riemann-Stieltjes integral (path-wise, $H > \frac{1}{2}$)

A math model for Gaussian white noise $\frac{d}{dt}B_t(\omega)$

Generalized time derivative of Brownian motion

Reason: $B_{t_2} - B_{t_1}$ and $B_{t_3} - B_{t_2}$ are *independent*
$\frac{d}{dt}B_t(\omega)$ and $\frac{d}{dt}B_s(\omega)$ are **uncorrelated**

Fourier transform of correlation of $\frac{d}{dt}B_t(\omega)$ is constant

Spectrum is constant: White light (Therefore: *white noise*)

L. Arnold: **Stochastic Differential Equations with Applications**, 1974

A math model for Gaussian colored noise $\frac{d}{dt}B_t^H(\omega)$

Generalized time derivative of fractional Brownian motion

$B_{t_2} - B_{t_1}$ and $B_{t_3} - B_{t_2}$ are **dependent**

Reason: $\frac{d}{dt}B_t^H(\omega)$ and $\frac{d}{dt}B_s^H(\omega)$ **correlated**

Bender, ***Stoch. Appl. Appl.***, 2003

Quantify model uncertainty using correlated noise

$$\tilde{u}_t = A\tilde{u} + N(\tilde{u}) + F(\tilde{u}),$$

Model uncertainty: $F(\tilde{u})$ (from historical observation data)

$$F = \sigma(x)\, g(u)\, \dot{B}_t^H,$$

$\sigma(x)$: Noise intensity
$g(u)$: Empirical "multiplicative" coefficient

Parameters

For example: $g(u) = u$

$$F = \sigma(x)\, u\, \dot{B}_t^H$$

Parameters: Hurst parameter H; Noise intensity $\sigma(x)$

Estimate Hurst parameter H and noise intensity $\sigma(x)$

$$\tilde{u}_t = A\tilde{u} + N(\tilde{u}) + F(\tilde{u}),$$

Model uncertainty: $F(\tilde{u})$ (from historical observation data)

$$F = \sigma(x)\ u\ \dot{B}_t^H$$

$$\int_0^t F ds = \sigma(x) \int_0^t u(x,s) dB_s^H,\ 0 < t < T.$$

Here the integration is in the sense of Riemann-Stieltjes. In this way, the mean of this integral is not zero, consistent with the mean of the model uncertainty F.

Estimate Hurst parameter H and noise intensity $\sigma(x)$

Let us denote $Z_t = \int_0^t F ds$, then

$$Z_t = \sigma(x) \int_0^t u(x,s) dB_s^H$$

Theorem 1: For a given $p < \frac{1}{1-H}$, as $n \to \infty$, we have the (uniform) convergence in probability:

$$\triangle_n^{1-pH} \; V_p^n(Z)_T \to c_p \; \sigma^p(x) \; \int_0^T |u(x,s)|^p ds$$

where $c_p = \mathbb{E}(|B_1^H|^p) = \frac{2^{p/2}\Gamma(\frac{p+1}{2})}{\Gamma(\frac{1}{2})}$ with Γ the Gamma function, $\triangle_n = \frac{T}{n}$. Variation

$$V_p^n(Z)_T = \sum_{i=1}^{n} |Z_{i/n} - Z_{(i-1)/n}|^p.$$

Corcuera, Nualart and Woerner, *Bernoulli* 2006.

Estimate Hurst parameter H and noise intensity $\sigma(x)$

Thus we have the estimator for the noise intensity for large n:

$$\sigma^p(x) \approx \mathbb{E}\Big[\Big\{ \frac{\triangle_n^{1-pH} \; V_p^n(Z)_T}{c_p \; \int_0^T |u(x,s)|^p ds} \Big\}^{\frac{1}{p}}\Big]. \qquad (1)$$

Estimator $\hat{H}_n$ value is the minimizer for the optimization problem:

$$\min_{0<H\leq 1} \mathbb{E}\big[\triangle_n^{1-pH} \; V_p^n(Z)_T - c_p \, \sigma^p(x) \int_0^T |u(x,s)|^p ds \,\big]. \qquad (2)$$

A SPDE model

$$\tilde{u}_t = A\tilde{u} + N(\tilde{u}) + \sigma(x)\ \tilde{u}\ \dot{B}_t^H$$

Relation with data assimilation

- Quantifying missing mechanisms: Improve the model itself
- Data assimilation: Improve state estimate

Water vapor: An advection-diffusion-condensation model

Relative humidity $q(x, y, t)$: Pierrehumbert et al 2007

$$\frac{\partial q}{\partial t} + (\vec{u} \cdot \nabla)q = \eta \nabla^2 q + S(q, q_s) \tag{3}$$

where

$$S = \begin{cases} -\frac{1}{\tau}[q - q_s], & \text{if } q > q_s(x, y) \\ 0, & \text{if } q \leq q_s \end{cases} \tag{4}$$

with saturated humidity $q_s(x, y)$, $\tau > 0$ being a relaxation time scale, $\vec{u}(x, y)$ the wind velocity field, and $\eta > 0$ the diffusivity.

Model uncertainties: moist convection, etc

Sukhatme and Pierrehumbert 2005

A stochastic model: Taking uncertainty into account

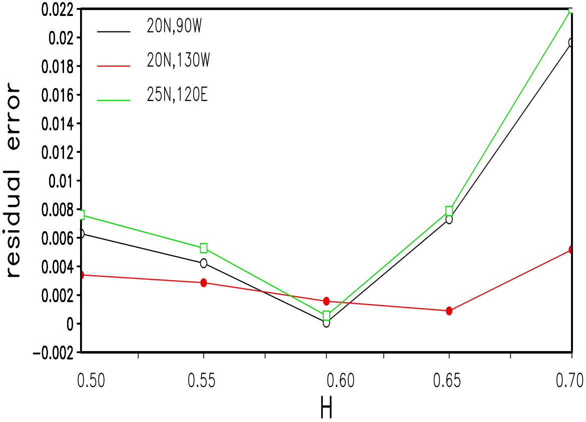
Relative humidity $q(x, y, t)$

$$\frac{\partial q}{\partial t} + (\vec{u} \cdot \nabla) q = \eta \nabla^2 q + S(q, q_s) + \sigma(x, y)\,(q - \frac{q_s}{3})\,\dot{B}_t^H \qquad (5)$$

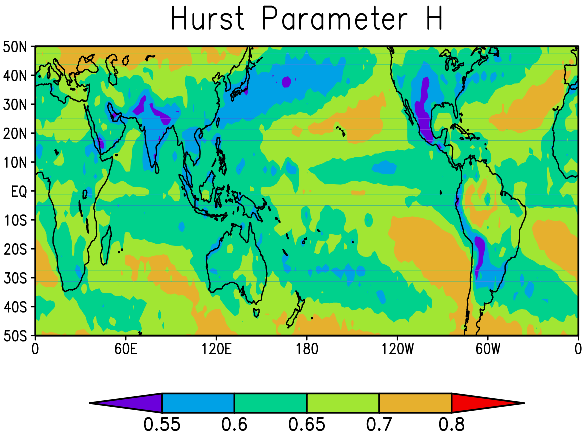
Noise intensity σ and Hurst parameter H determined by observational data

Chen, Duan and Pierrehumbert 2009

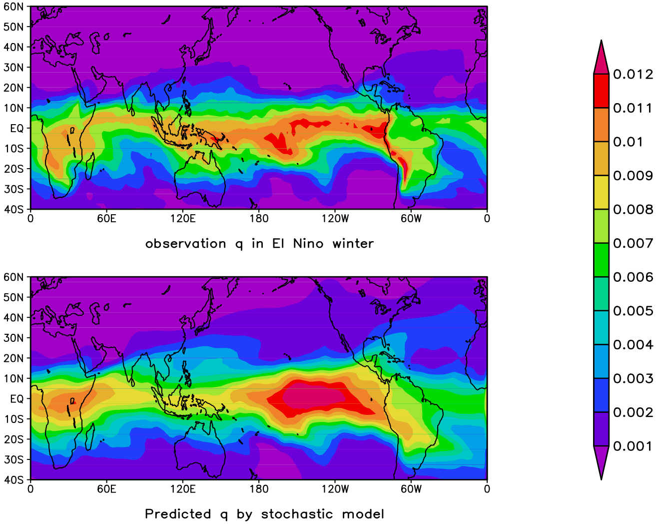
Convergence test



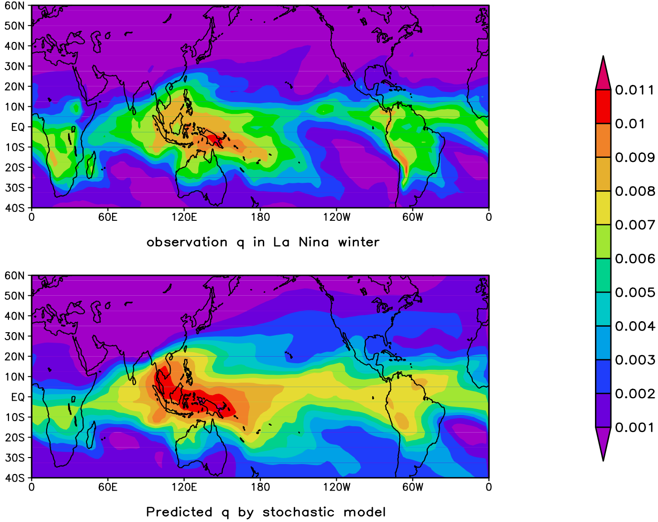
Estimated Hurst parameter H



Numerical Validation 1



Numerical Validation 2



Conclusions

How to quantify model uncertainty?

- **Correlated noises**
- **Parameter estimation**
- **PDE model $\longrightarrow$ SPDE model**
- **Application: Advection-Diffusion-Condensation for water vapor**