An Introduction to Probability and Stochastic Processes for Ocean, Atmosphere, and Climate Dynamics

2: Stochastic Processes

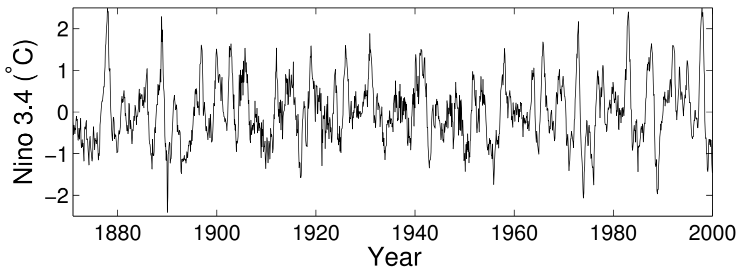
Adam Monahan

monahana@uvic.ca

School of Earth and Ocean Sciences, University of Victoria

Climate variability: “signal vs. noise”, “fast vs. slow”

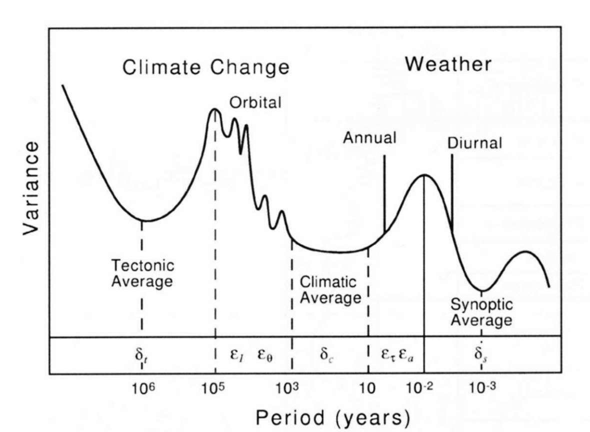
- Atmosphere and ocean flows generally unsteady; often turbulent
- Some aspects of variability predictable; others not
- Niño3.4 time series (has both regular & irregular variability)



- Irregular, unpredictable variability often well-modelled as random; need to think about evolution of random variable through time

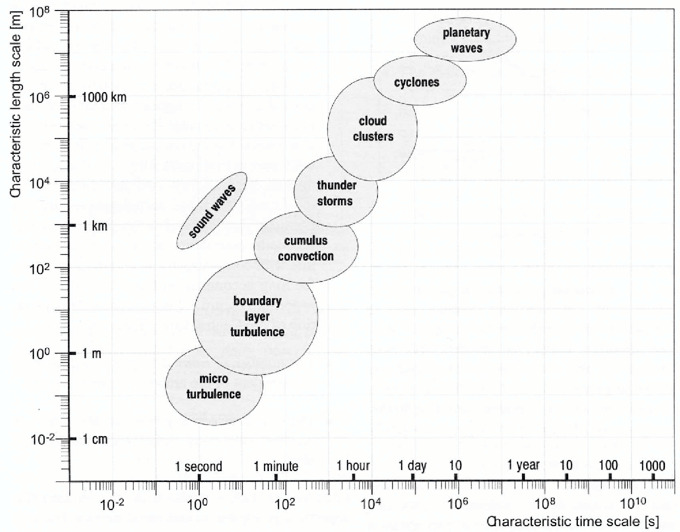
Climate variability: multiple temporal scales

- Climate system displays variability over broad range of space and time scales, and involves different interacting components



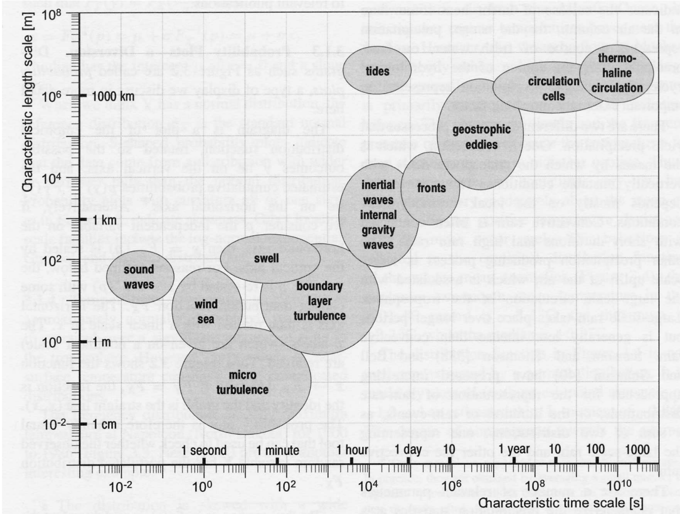
From Saltzman, 2002

Climate variability: multiple spatial scales



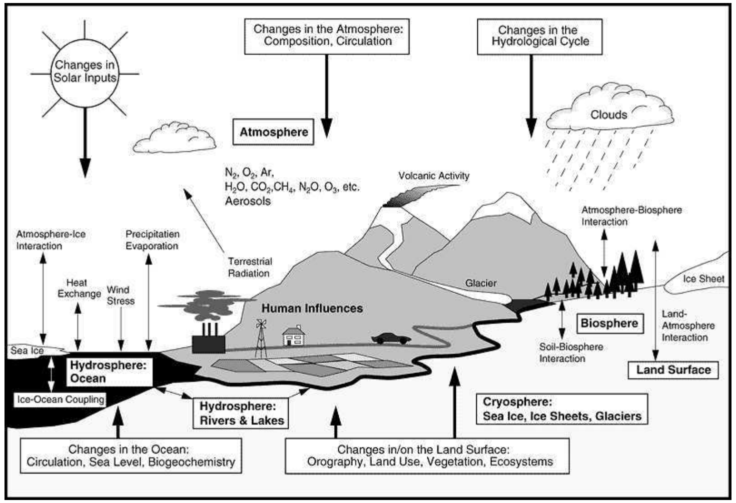
From von Storch and Zwiers, 1999

Climate variability: multiple spatial scales



From von Storch and Zwiers, 1999

Climate variability: multiple systems



From IPCC Third Assessment Report

Where might “randomness” come from?

- Common strategy in atmosphere/ocean dynamics is to split variability across scales, e.g.
  - mean vs. eddy (Reynolds averaging)
  - climate vs. weather
  - ocean vs. atmosphere
  - resolved vs. subgrid-scale
- One scale is modelled explicitly, the other isn’t
- A classical example: Reynolds averaging

$$\overline{f}(t) = \int_{t-\tau/2}^{t+\tau/2} f(t')dt' \quad , \quad f'(t) = f(t) - \overline{f}(t)$$

Where might “randomness” come from?

- Reynolds averaged 2d horizontal momentum budget:

$$\frac{\overline{d\overline{u_i}}}{dt} = -\frac{1}{\rho}\partial_i \overline{p} + f\epsilon_{3ik}\overline{u}_k + \nu\nabla^2\overline{u}_i - \partial_j \overline{u_i' u_j'}$$

- Evolution of “slow” variable depends on statistics of “fast” variable: classic closure problem
- Can try to represent Reynolds stress in terms of mean state, but in absence of very large scale separation between “mean” and “eddy” flows, why should a configuration of resolved flow be associated with a (statistically) unique configuration of unresolved flow?
- Turbulence is dissipative & bursty; perhaps a better model is

$$\partial_j \overline{u_i' u_j'} = -K\nabla^2\overline{u}_i + \text{“noise”} \sim P(\overline{u_i' u_j'}|\overline{u}_i)$$

Where might "randomness" come from?

- A statistical mechanical analogy: consider a box of N identical molecules of temperature T
- Each individual molecule will have random kinetic energy E_i, with pdf from Maxwell-Boltzmann distribution; total energy of molecules sum of individual energies $E = \sum_{i=1}^{N} E_i$
- std$(E) \sim N^{-1/2}$

⇒ as N becomes large, $E \to$ mean(E) with very small fluctuations ("thermodynamic limit"), and can talk about "large scale" without worrying about fluctuations on "small scale"

- **However**, if N is not small (if there is not a "scale separation"), then these fluctuations may not be negligible

Stochastic processes

- A sequence of random variables x_t (where t may be a discrete or continuous index) is called a **stochastic process**
- Describes the evolution of a random variable in time (or space)
- Stochastic processes will generally have **memory**; that is, the pdf of x_t will generally depend on the values taken by the process at an earlier time (or times)
- Important quantities are the joint and conditional pdfs across time, e.g. $p(x_{t_1}, x_{t_2})$, $p(x_{t_2}|x_{t_1})$, $p(x_{t_2}, x_{t_3}|x_{t_1})$, ...
- Joint pdfs of **stationary processes** are time translation invariant, e.g.

$$p(x(t_1), x(t_2)) = p(x(t_1 + \tau), x(t_2 + \tau))$$

so joint pdfs depend only on time differences $t_1 - t_2$

Markov processes

- A very important class are **Markov processes** for which

$$p(x_{t_l}|x_{t_k}, x_{t_j}, x_{t_i}, ...) = p(x_{t_l}|x_{t_k}) \text{ for } t_l > t_k > t_j, t_i, ...$$

- For these processes, knowledge of the state of the system at time t_k provides all possible information about the state at future time; information about state at times before t_k is irrelevant
- A deterministic analogue:
  - future evolution of ODE $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ for $\mathbf{x} \in \mathrm{R}^N$ fully specified if full state $\mathbf{x} = \mathbf{x}_0$ known at $t = t_0$; state at earlier times irrelevant
  - If only some subspace of $\mathbf{x}$ known at $t = t_0$, information about state at earlier times needed to specify trajectory

An example: discrete random walk

- Motion with position x_n on a lattice such that each time there is a 50% chance of moving one step either right or left

$$p(x_n = k|x_{n-1} = j) = \begin{cases} 0.5 & \text{if } k = j+1 \text{ or } j-1 \\ 0 & \text{otherwise} \end{cases}$$

- mean$(x_n) = 0$ (by symmetry)
- std$(x_n) = \sqrt{n}$

⇒ variability of x_n grows without bound; size of fluctuations grows as square root of time

- By central limit theorem, $p(x_n)$ becomes Gaussian as $n \to \infty$