Estimation and Prediction of Complex Systems: Progress in Weather and Climate

Gregory J. Hakim

University of Washington

Math Across Campus Lecture—3 December 2009

Thanks: David Battisti, Karin Bumbaco, Seb Dirren, Helga Huntley, Rahul Mahajan, Cliff Mass, Guillaume Mauger, Phil Mote, Angie Pendergrass, Gerard Roe, Chris Snyder, & Ryan Torn
Goals

- Opportunities for fusing models & observations
- Reach both non-specialists & practitioners in this area
- Meet people from other fields—opportunity to organize?
Fusing Observations & Models

- **Motivation**: Challenges in complex systems
- **Estimation primer**: least squares & Bayesian methods
- **Prediction I**: Weather
  - methods to deal with complexity
  - success: decreasing forecast uncertainty (obs & estimation)
- **Prediction II**: Climate
  - methods compound complexity
  - failure: forecast uncertainty not decreasing (feedbacks)
- **Rethinking models**: complexification vs. simplification
Motivation

Estimation & prediction are ubiquitous activities

Randomly chosen examples

- **social**: social trends; agents (Sugarscape); elections
- **medicine**: diagnosis; disease spread; flu vaccines
- **biology**: protein structure; neuron models; populations
- **finance**: portfolios; markets; inventory management
- **networks**: web traffic; highways; electrical grid
- **engineering**: autonomous vehicles; robotics; tracking
- **environment**: weather; hydrology; ecosystems

Lots of field-specific jargon. Is there a *lingua franca*?

Yes: **Math!**
WARNING

MATH AHEAD!
Simple scalar example: $x$

- Estimate $x$ (e.g. true temperature at a point)
- Single observation: $y_1$
- Uncertainty = error variance ($\sigma^2$) of the observation
Simple scalar example: $x$

- Two observations: $y_1$ and $y_2$; error variance $\sigma_1^2$ and $\sigma_2^2$
- Least squares: minimize, e.g., error variance in $x$

\[ \hat{x} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} y_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} y_2 \]

\[ \sigma_\hat{x}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} < \sigma_1^2, \sigma_2^2 \]
Multivariate probability density

marginal probability density

joint probability density & likelihood

conditional probability density
Conditional probability

- $P(x|y)$ “probability of $x$, given $y$”
- **Bayes rule:** $P(x|y) = \frac{P(y|x)P(x)}{P(y)}$
  - $P(y|x)$: “likelihood function” for measurement errors.
  - $P(x)$: “prior distribution”; e.g. the first observation.
  - $P(y)$: “marginal distribution”—a normalizing constant.

- **If** $P(y|x) \sim N(x, \sigma_y^2)$ and $P(x) \sim N(\bar{x}, \sigma_x^2)$
  - Then $P(x|y) \sim N \left( \frac{\sigma_y^2}{\sigma_x^2 + \sigma_y^2} \bar{x} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} y, \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2} \right)$
  - Same as least squares if $\bar{x}$ is the first observation!
  - $\hat{x} = \bar{x} + K(y - \bar{x})$  \hspace{0.5cm} $\sigma_{\hat{x}}^2 = (1 - K)\sigma_x^2$  \hspace{0.5cm} $K = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}$

Gregory J. Hakim
Estimation and Prediction of Complex Systems
Estimation primer: Bayesian example #1
Estimation primer: Bayesian example #2

[Graph showing probability density functions for prior, observation, and posterior distributions with x-axis from -1.5 to 3 and y-axis from 0.005 to 0.025]
Estimation primer: Bayesian example #3

Gregory J. Hakim

Estimation and Prediction of Complex Systems
Large numbers of state elements and observations: vectors

- \( \mathbf{x} = [x_1, x_2, \cdots, x_n]^T \) “state vector”
- \( P(\mathbf{x}_t|\mathbf{Y}_t) \) “\( P(\mathbf{x}) \) at time \( t \), given all current and past obs”
- \( \mathbf{Y}_t = [\mathbf{y}_t, \mathbf{Y}_{t-1}] \)

If observation errors are uncorrelated in time, then

\[
P(\mathbf{x}_t|\mathbf{Y}_t) \propto P(\mathbf{y}_t|\mathbf{x}_t) \ P(\mathbf{x}_t|\mathbf{Y}_{t-1})
\]

- \( P(\mathbf{y}_t|\mathbf{x}_t) \): observation likelihood
- \( P(\mathbf{x}_t|\mathbf{Y}_{t-1}) \): prior, given all past observations: model

Recursive if we can update \( P(\mathbf{x}_t|\mathbf{Y}_{t-1}) \)!
State dynamics & conservation of probability

- **State dynamics:** \( \frac{dx}{dt} = F(x) \)
- **Liouville equation:** \( \frac{dP}{dt} = -P \nabla \cdot F(x) \)

Solution:

\[
P(x_{t+1}|Y_t) = P(x_t|Y_t) e^{-\int_{t}^{t+1} \nabla \cdot F(x) dt'}
\]

- Probability density decreases exponentially when state-space trajectories diverge. **chaos!**
- Completely impractical for complex systems: weather models \( \sim (10^8)^n \) for \( n \) moments.
Example for the Lorenz attractor

Now I don’t care what the weatherman says
When the weatherman says it’s raining
You’ll never hear me complaining, I’m certain the sun will shine.

—Louis Armstrong (“Jeepers Creepers,” 1938)

“Numerical” weather prediction began in the 1950s.
- application of laws of physics to atmosphere.
- mathematics made the problem tractable.
Weather Prediction

The numbers (European Centre for Medium Range Weather Forecasting)

- analyses and forecasts every 6-12 hours.
- model: \( \sim 128 \text{ million} \) degrees of freedom
- observations: \( \sim 9 \text{ million} \) per cycle
- computers: 2 x 156 TFlops (US: 2 x 94 TFlops)
Satellites dominate the observation data stream
Lines show the forecast day at which skill is lost.
from: ECMWF (2009)
Increasingly realistic forecasts from: ECMWF (2009)

New product: model simulated satellite images

Sunday 31 May 2009 12UTC © ECMWF t+48 VT:Tuesday 2 June 2009 12 UTC
Model simulated METEOSAT 9 SEVIRI (Channel 5 WV6.2) Brightness Temperature and Mean sea level pressure

Forecast Products Users Meeting 10-12 June 2009

from: ECMWF (2009)
Assumption #1: **Gaussian statistics**

- \( P(x_t|Y_t) \propto P(y_t|x_t) \ P(x_t|Y_{t-1}) \)
- \(-\ln P(x_t|Y_t) \propto (x_t - \bar{x})^T \ B^{-1} (x_t - \bar{x}) + [y - \mathcal{H}(\bar{x})]^T \ R^{-1} \ [y - \mathcal{H}(\bar{x})] \)
- \( J_b \) and \( J_o \)
- \( \bar{x} \) prior. \( B \) and \( R \) error covariance matrices.
- Note: \(-\ln P \propto \) Shannon’s information metric.

**Two solution paths**

- (1) let \( J = -\ln P \) and search for the minimum. **Variational**.
- (2) direct solve: \( \frac{\partial J}{\partial x} = 0 \). **Kalman filter**.
(1) Variational approach

**3DVAR**
- descent algorithm based on $\frac{\partial J}{\partial x}$
- assumes fixed $B$ (empirical)
- this is a sequential filter

**4DVAR**
- extension into time domain
- $J_0 = \sum J_0(t)$
- this is a smoother
(2) Kalman Filter approach

- analytical derivation of posterior distribution
- \(x_t^a = x_t^b + K[y_t - \mathcal{H}(x_t^b)]\)
- \(K = f(B, R)\)

Need to advance \(B\) in time; too big \((10^8 \times 10^8)\)
Approximate with an **ensemble** of forecasts

**Ensemble filters**

- Estimate \(\bar{x}\) and \(B\) with \(O(100)\) forecasts
- Why does this work? Strong, state-dependent, correlations
  - I.e., problem is not as high dimensional as it seems

**Ensemble prediction**

- Approximation to Liouville equation
- Advance analysis ensemble in time
- Note: *variational approach is deterministic*
Ensemble forecast spread: $t = 0$
Ensemble forecast spread: $t = 3$ days
Ensemble forecast spread: $t = 5$ days
Applications

- Sensitivity analysis
  - What structure most affects the forecast?

- Observation impact
  - Which observations contain the most information?
  - Filter redundant observations.
  - Optimally design observing networks.

- Adaptive sampling
  - Target observations to reduce forecast errors.

- System dynamics
  - Use ensemble statistics for hypothesis testing.
red = greater forecast sensitivity.
Forecast sensitivity example: medium-time forecast

Gregory J. Hakim
Estimation and Prediction of Complex Systems

red = greater forecast sensitivity.
Forecast sensitivity example: long-time forecast

red = greater forecast sensitivity.
Real weather example
With Ryan Torn (SUNY Albany)
Typhoon Tokage (2004): Large forecast errors

Sea-level Pressure

0 hour fest.

24 hour fest.

48 hour fest.

500 hPa Height

b)

c)

d)

e)

f)
Typhoon Tokage (2004): Observation Impact

greatest impact well removed from the storm!
Perturbed forecast for a single observation

Sea-level Pressure

500 hPa Height

Gregory J. Hakim
Estimation and Prediction of Complex Systems
Climate Prediction
IPCC Projected global temperature change

- This large range of uncertainty has changed little with time.
Earth’s Energy Balance

$$T_s = \left[ \frac{S(1-A)}{2\sigma(2-\varepsilon)} \right]^{1/4}$$

- No atmosphere: 0, 0.3, −18°C
- 20th century: 0.75, 0.3, 14°C
- Global warming #1: 0.85, 0.3, 20°C
- Global warming #2: 0.85, 0.35, 15°C

Use climate models to be more precise...
2100–1990 global-mean surface temperature change (A1B scenario)

<table>
<thead>
<tr>
<th>Report</th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAS (1979)</td>
<td>N/A</td>
<td>1.6+°C</td>
</tr>
<tr>
<td>AR1 (1992)</td>
<td>2.9°C</td>
<td>1.9–4.4°C</td>
</tr>
<tr>
<td>AR3 (2001)</td>
<td>3.0°C</td>
<td>2.0–4.0°C</td>
</tr>
<tr>
<td>AR4 (2007)</td>
<td>2.8°C</td>
<td>1.7–4.4°C</td>
</tr>
</tbody>
</table>

Why hasn’t the uncertainty (range) changed?
- **climate sensitivity**: equilibrium change in surface temperature due to “forcing.”
- **forcing**: anything that changes the net energy (e.g. absorbing gases)
- **feedbacks**: changes that reinforce forcing.
  - e.g. “water-vapor” feedback: \( T \uparrow \Rightarrow H_2O \text{ vapor} \uparrow \Rightarrow T \uparrow \)

**Roe & Baker (2007):**
- linear approximation: \( \Delta T = \lambda \Delta R \)
  - without feedbacks: \( \lambda \equiv \lambda_0 \approx 0.3 \, \text{K/ (W/m}^2\) \)
  \( \Delta T_0 = 1.2^\circ \text{C for 2 x CO}_2 \).
- feedback: \( \Delta T = \lambda_0 \Delta R + \lambda_0 C \Delta T \)
  \( \Rightarrow \Delta T = \frac{\Delta T_0}{1-f} \quad f = \lambda_0 C \text{ feedback factor} \)
“first principles” models

- “faithful” representation of basic processes
- Moore’s Law enables better representation; more processes (“complexification”)
- parameterized processes often do not go away
- complexification + parameterization + feedbacks = trouble

“simplified” models

- objective model formulation; e.g., information theory
- let complexity emerge from simple building blocks (e.g. agents)
- careful calibration before increasing complexity
Summary

**estimation & prediction**
- powerful “fusion” of models with observations
- equivalent Bayesian & least-squares approaches
- variational & ensemble methods for complex systems

**application to weather & climate**
- **weather**: success due to better estimation & observations
  - huge increase in observations
  - better estimation & faster computers
  - sensitivity theory: observation targeting & selection
- **climate**: failure due to model “complexification” & feedbacks
  - No argument: world is warming, and we know why
  - Uncertainty in projections resistant to CPU and $\$
  - Uncertainty in feedbacks; complexification does not help
Thank You!