Chaos and Ensemble Forecasting

Josh Hacker

National Center for Atmospheric Research
Boulder, CO, USA
A chaotic system:

- *Is* deterministic
  If we know exactly the initial conditions and how the system evolves, we can forecast it perfectly

- *Looks* random
  Allows us to use probability and statistics
A chaotic system:

- More than 2 degrees of freedom
  *Just x and y are not enough*
- Nonlinearity
  \[xx, yy, xy \text{ when } y = f(x)\]

*A system with these properties is very sensitive to initial conditions*
The Butterfly Effect

Does the flap of a butterfly’s wings in Brazil set off a tornado in Texas?

- Edward Lorenz, 1972
The Lorenz Equations

\[
\begin{align*}
\frac{dx}{dt} &= a(y-x) \\
\frac{dy}{dt} &= bx-y-xz \\
\frac{dz}{dt} &= xy-cz
\end{align*}
\]

3 variables: \( x, y, z \)
3 parameters: \( a, b, c \)
2 nonlinearities

Represents simple convection
Lorenz demo

http://www.cmp.caltech.edu/~mcc/chaos_new/Lorenz.html
Atmospheric models

• $10^6$ to $10^{10}$ variables
  – All variables in the discretized Navier-Stokes equations

• Highly nonlinear

• Side note: subgrid-scale processes may appear stochastic rather than deterministic
The real atmosphere

• Infinite degrees of freedom

• Different from our models, making ensemble forecasting difficult and probabilistic forecasts (perhaps) unattainable
NWP example: IC sensitivity
NWP example: IC sensitivity

ECMWF

ECMWF ensemble forecast - Air temperature
Date: 26/06/1995 London Lat: 51.5 Long: 0

Degree C

Forecast day

UK

ECMWF ensemble forecast - Air temperature
Date: 26/06/1994 London Lat: 51.5 Long: 0

Degree C

Forecast day

UK
Ensemble Forecasting: Goals

• To forecast the uncertainty
  – Extremes (tails of the distribution)
  – Most likely outcome

• Produce an ensemble mean that filters out the unpredictable components of the flow

• To produce a probabilistic forecast
  – Requires calibration
Disadvantages

• Ensemble-mean loses the temporal and spatial variability of any one ensemble member

• Spread-skill relationships only hold on average

• Typical approaches (medium range) depend on nonlinear error growth
Advantages

• Gives forecaster an idea of the range of possibilities
  – At least the sensitivity of the forecast to initial errors

• Can produce more than one likely scenario
  – Ensemble members might clump around two outcomes

• Can be calibrated to produce useful probabilistic forecasts
Very common ways to use ensembles

- Ensemble-mean and spread
- Postage stamps
- Spaghetti diagrams
- Time series at forecast sites
- Anomalies
- Pseudo-probabilistic forecasts
Ensemble mean forecasts

Gray curves are the control forecast
Ensemble spread forecasts
Spaghetti diagrams

NCEP ENSEMBLE 500mb Z
000H Forecast from: 00Z Mon JUL 25 2005
Valid time: 00Z Mon JUL 25 2005

NCEP ENSEMBLE 500mb Z
120H Forecast from: 00Z Mon JUL 25 2005
Valid time: 00Z Sat JUL 30 2005
Spaghetti diagrams

Model Guidance for Hurricane Debby’s Track
8/23/00
Time series

ECMWF ensemble forecast - Air temperature
Date: 26/06/1995 London Lat: 51.5 Long: 0

ECMWF ensemble forecast - Air temperature
Date: 26/06/1994 London Lat: 51.5 Long: 0
Anomalies

NCEP ENS MEAN ANOM – 850mb TEMP
120H Forecast from: 00Z Mon JUL,25 2005
Valid time: 00Z Sat JUL,30 2005

NCEP ENS. MEAN NORM ANOM – 850mb TEMP
120H Forecast from: 00Z Mon JUL,25 2005
Valid time: 00Z Sat JUL,30 2005
Pseudo-probabilities

NCEP ENS ANOM PROB (1sigma) - 850mb TEMP
120H Forecast from: 00Z Sat JUL 25 2005
Valid time: 00Z Sat JUL 30 2005

NCEP ENS ANOM PROB (2sigma) - 850mb TEMP
120H Forecast from: 00Z Sat JUL 25 2005
Valid time: 00Z Sat JUL 30 2005
Methods for Ensemble Generation

• **NCEP:** Breeding of Growing Modes
  – Average directions of error growth

• **ECMWF, FNMOC:** Singular Vectors
  – Most rapidly growing errors (extremes)

• **CMC:** Monte Carlo
  – Perturbations from ensemble data assimilation system
Breeding (NCEP)

From Toth and Kalnay (1993)
Singular vectors (ECMWF, FNMOC)

1. 48-h Forecast with Tangent Linear Model (TLM)

2. Choose a sensitivity metric that may be relevant to the forecast users (e.g. total energy over Europe)

3. Run the adjoint model backwards to t=0 to get gradients of metric with respect to initial conditions

4. Solve an eigen-decomposition problem to get the first N singular vectors

5. Add and subtract each of the leading SVs to populate ensemble
Singular vectors and breeding

- Both designed for medium-range forecasts
- They rely on a high-quality model
  - Models are good at larger scales
  - Perturbations are “balanced”
- They rely on nonlinear growth in spread
  - This does not always occur before $t = 48$ h

*As resolution and mesoscale models improve, the effectiveness of these approaches will have to be tested*
Mesoscale ensembles: a different animal

• Lateral boundary conditions
  – Sweeping
  – Errors in LBCs

• Forecasts may be dominated by model error of many forms
  – Large biases
  – Lack of variability and nonlinear error growth

• An active area of research

• Smoothing when taking the ensemble mean can hurt you more than when large-scales only are considered
Example: convective cells

Four individual ensemble members

Ensemble mean

Source: NOAA HPC
Are multiple models the answer?

- Maybe for now, but probably not in the long run
- Distributions of parameters (e.g. land-surface properties) may prove more fruitful
- Research on stochastic approaches to parameterization is suited to ensemble approaches
Calibration

• Just like single model runs, the ensemble needs some correction for model error.

• The needs for an ensemble go one step further because the distributions need to be corrected for probability.
The model climate is different from reality
- mean
- spread
Summary

• Chaos and nonlinearity (the butterfly effect)

• How ensembles are used in the medium-range and for large scales

• How ensembles are generated in the medium-range and for large scales

• Remaining issues
  – Calibration
  – Mesoscale ensembles