Numerical Weather Prediction

Thomas Desautels

Gatsby Computational Neuroscience Unit

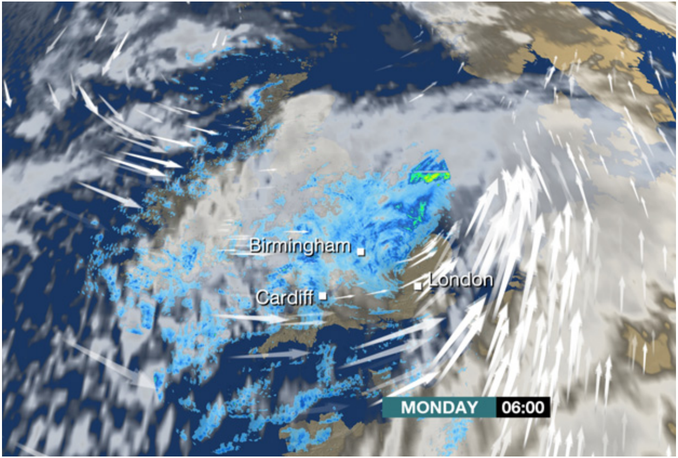
1 November, 2013

Outline

Early Predictions of Monday’s Storm

Figure : This ‘Thumb’ Was Identified as Significant on Saturday
http://www.bbc.co.uk/news/uk-24674537

Monday's Results



Monday's Results

Monday’s Results

Motivating Question

- So how can a weather system near Boston be reliably identified as important to weather conditions in the UK two days later?

Outline

Numerical Weather Prediction

- Method of simulating weather as a complex, non-linear dynamical system.
  - This requires sophisticated models.
- Challenges:
  - Weather phenomena are inherently multi-scale
    - Discretizations will miss sub-grid phenomena
  - These systems have chaotic dynamics
    - Lorentz system (butterfly attractor) is a simplified model of atmospheric convection

Outline

Ensemble Forecasting

- Major problems to be addressed: sensitivity to unmodeled dynamics and measurement errors.
- Monte Carlo methods which sample over model parameters, initial conditions (ICs), maybe even models, are thus appealing.
- However, running the model is very expensive: too costly to compute more than a few samples from the predictive distribution (typically a few dozen).
- Methods for creating and working with such *ensembles* of simulations are vital.

Ensemble Creation

- Creating an ensemble which allows both robust prediction and reliable measure of confidence is challenging.
- Naïvely allocating samples according to *measurement* uncertainty in the ICs is wasteful or may underestimate *predictive* uncertainty.
- Choice of samples according to the sensitivity of the dynamics to perturbation of the initial conditions is much more representative.
  - *Singular vectors* method: finds those perturbations in the ICs which have the largest effect on the forecast and allocates samples to these directions (see Palmer & Zanna, 2013).

Met Office Unified Model

- The UK Met Office uses several systems in parallel, with a new cycle run every few hours.
  - Global, regional, and UK scales: large-scale models give boundary conditions to finer local models.
  - At each scale, systems include high-resolution, single-run models (high precision) and coarser resolution ensembles (quantify uncertainty).
  - Several of these systems use the Ensemble Transform Kalman Filter: choose the ICs for new cycle's ensemble using the results from previous cycle.
- On the web: `http://www.metoffice.gov.uk/research/modelling-systems/unified-model`
- For a more technical description (of one component): see Bowler et. al, 2008, "The MOGREPS short-range ensemble prediction system."

So that question's (sort of) answered...

What's under the hood (besides messy non-linear dynamics)?

Outline

Typical Kalman Filter (KF)

- Assume linear dynamics, Gaussian innovations and observation error.
- Calculate posterior mean $\hat{x}_t^{t-1} \in \mathbb{R}^d$ and variance $V_t^{t-1} \in \mathbb{R}^{d \times d}$ over the state x_t, given observations $y_1, \ldots, y_{t-1}$.
- Computationally easy updates of mean $\hat{x}_t^{t-1} \to \hat{x}_t^t$ and variance $V_t^{t-1} \to V_t^t$, given observation y_t.
- Yields nice recursion relationships which consist of a propagation step and an update step.

Ensemble Kalman Filter (En KF)

- If d is very large, the variance V_t^{t-1} may be very cumbersome to keep in data.
- Approximate V_t^{t-1} as the ensemble variance:

$$V_t^{t-1} \approx P_t = \frac{X_t^{t-1}(X_t^{t-1})^T}{n-1},$$

where $X_t^{t-1} \in \mathbb{R}^{d \times n}$ is the matrix of ensemble perturbations, i.e., the ith column is the deviation of the ith particle from the ensemble mean, $x_{t,i}^{t-1} - \bar{x}_t^{t-1}$.
- P_t is low-rank and only depends on the values of the individual ensemble "particles:" possible to avoid storage entirely, do computations cheaply.
- Now treat propagation step via non-linear dynamics.

Ensemble Transform Kalman Filter (ET KF)

(Bishop, Etherton, & Majumdar 2001)

- To incorporate observations, we usually need to calculate the expected observation and the Kalman gain matrix.
- If we assume that the dynamics are piecewise linear[1], then $X_\tau^{\tau-1} = X_t^t T_{t,\tau} \Pi_t$ where $T_{t,\tau} \in \mathbb{R}^{n \times n}$ and $\Pi_t \in \mathbb{R}$.
- Reduces our computational problem to calculating T_t, and performing a low-dimensional matrix multiplication.
- Importantly, $T_{t,\tau}$ captures *which* observations are available at time t.
- Key calculations to produce $T_{t,\tau}$ are cheap (partial eigenvector decomposition, see Bishop et. al 2001).

[1] This is still quite fuzzy for me.

Outline

1. Motivation
2. Numerical Weather Prediction
   - Problem Sketch
   - Ensemble Forecasting
3. An Example: Met Office Unified Model
4. Kalman Filtering in Weather Prediction
   - Kalman Filter Variants
   - Experimental Design Using the ET KF

Experimental Design Using the ET KF

- Given a decent ensemble model, how should we use the ensemble to select possible new observations (airplane measurements, etc.) to reduce our uncertainty in the verification region?
    - This is an experimental design problem.
- In the ET KF, this reduces to determining which T_t matrices are the best allocation of a budget of observations.
- There are (possibly) combinatorially many allocations of observations, making direct optimization intractable.
    - Greedy observation selection, where the measure of success is reduction in the predictive variance in the verification region.

Summary

- Numerical weather prediction is typically done using hierarchical non-linear dynamical system models.
- Ensembles of such models are used to quantify uncertainty.
- Simulation techniques used in numerical weather prediction are similar to ML and control techniques, though buried under a hefty layer of field-specific jargon.