The future of weather forecasting: high-resolution ensembles

Craig Schwartz
The National Center for Atmospheric Research

schwartz@ucar.edu

NCAR is sponsored by the National Science Foundation
The olden days of weather forecasting

http://www.history.noaa.gov/stories_tales/women6.html
Modern weather forecasting


Supercomputer racks
Weather forecasting has improved!

- Partly due to increases in computing, weather forecasting has greatly improved over the past few decades.

- A happy marriage of computational and scientific progress.
Atlantic Basin Hurricane Track Forecasts

Human forecasts

http://www.nhc.noaa.gov
Atlantic Basin Hurricane Intensity Forecasts

http://www.nhc.noaa.gov

Human forecasts
Components of a numerical weather prediction (NWP) model

• Initialization
  – Data assimilation
• Dynamics
• Physics
• Subjective choices
  – Horizontal resolution
Steady increase in horizontal resolution with time

<table>
<thead>
<tr>
<th>Year</th>
<th>Approximate horizontal grid spacing (km)</th>
<th>Number of vertical levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>375</td>
<td>12</td>
</tr>
<tr>
<td>1983</td>
<td>300</td>
<td>18</td>
</tr>
<tr>
<td>1987</td>
<td>150</td>
<td>18</td>
</tr>
<tr>
<td>1991</td>
<td>105</td>
<td>28</td>
</tr>
<tr>
<td>1998</td>
<td>80</td>
<td>42</td>
</tr>
<tr>
<td>2002</td>
<td>55</td>
<td>64</td>
</tr>
<tr>
<td>2005</td>
<td>35</td>
<td>64</td>
</tr>
<tr>
<td>2010</td>
<td>27</td>
<td>64</td>
</tr>
<tr>
<td>2015</td>
<td>13</td>
<td>64</td>
</tr>
</tbody>
</table>
High-resolution models

- Computer models are pretty good at predicting large-scale systems
- Challenges remain regarding finer-scale details
- To address these challenges, high-resolution models are needed
  - Typically have horizontal grid spacings of 1- to 4-km
Benefit of high-resolution
Cool high-resolution fields

- Maximum 1-km vertical vorticity over 7-hrs
  – Toward tornado prediction
Sensitivity to horizontal grid-spacing

• Within high-resolution model configurations, what resolution is really needed?

• Little dispute that higher-resolution means more realism
  – But does greater realism translate into greater value?
  – If a 4-km model is as useful as 2-km, is it worth the ~8-fold additional cost to have a 2-km model?
Simulated reflectivity snapshots

Schwartz et al. (2009); Monthly Weather Review
Convective evolution

24-hr forecast

27-hr forecast

30-hr forecast

Schwartz et al. (2009); Monthly Weather Review
Colorado Front Range floods of 2013
Operational, coarse-resolution forecasts

- 48-hr accumulated precipitation
  - NAM, GFS, and RR operational models
  - CoCoRaHS gauge measurements overlaid

From Schwartz (2014); *Weather and Forecasting*
4-km WRF model forecasts

- 48-hr accumulated precipitation

- CoCoRaHS gauge measurements overlaid

48-hr accumulated precipitation (mm) between 1200 UTC 11 and 1200 UTC 13 September

From Schwartz (2014); *Weather and Forecasting*
1-km WRF model forecasts

- 48-hr accumulated precipitation

- CoCoRaHS gauge measurements overlaid

From Schwartz (2014); *Weather and Forecasting*
Objective verification of high-resolution models
Traditional objective verification

• Verification at the grid-scale
• Pick an event
  – Precipitation exceeding 1.0 mm/hr
• Compare forecast and observations at each grid point

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Forecast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>$a$</td>
<td>$b$</td>
<td>$a + b$</td>
</tr>
<tr>
<td>No</td>
<td>$c$</td>
<td>$d$</td>
<td>$c + d$</td>
</tr>
<tr>
<td>$a + c$</td>
<td>$b + d$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard 2 x 2 contingency table for dichotomous events
Traditional point-by-point methods

- The event has occurred at the shaded grid points

<table>
<thead>
<tr>
<th>Model output</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Grid point classification

- Green: “hits”
- Red: “false alarms”
- Blue: “misses”
- White: “correct negatives”
It’s a beautiful day in the “neighborhood”

• High-resolution models are not accurate at the grid scale

• To account for spatial displacement errors, specify a radius of influence \((r)\) about each grid point

• Define an event

• Generate a probability at each grid point
- $r = 2.5$ times the grid spacing
- The event has occurred in the shaded boxes
- Event occurs in 8 boxes
- 21 total boxes in neighborhood

Hypothetical model output

\[ P = \frac{8}{21} = 38\% \]
Example Applied to Model and Observations

P = 8/21 = 38%

A perfect forecast using this neighborhood approach
Objective benefit of high-resolution

The fractions skill score compares observed and forecast fractions

WRF2: 2-km
WRF4: 4-km
NAM: 12-km

Schwartz et al. (2009); Monthly Weather Review
Thoughts about horizontal grid spacing

• It appears that 3- or 4-km horizontal grid spacing provides similar value and accuracy as 1- or 2-km horizontal grid spacing over flat terrain
  – Higher-resolution always provides more realism

• In topographically-diverse areas, higher-resolution (~1-km horizontal grid spacing) is usually better
Ensemble prediction systems
Probabilistic predictions

• Probabilistic forecasts are often generated by ensembles of computer models, where variations in model parameters yield different forecast outcomes.

• Different forecasts are called “ensemble members”.

Slightly different realizations of “now” lead to larger differences later.

Now 5-day forecast
Hurricane Joaquin had a very uncertain track.
Why ensemble forecasts are desirable

• Quantification of uncertainty
  – Naturally produces probabilities!
  – Allows forecasters to forecast their “true beliefs”
  – Allows users to make decisions based on expected value and cost-loss scenarios

• Errors of different members cancel when combining forecasts across members
  – Forecasts combining information across all members are better than single deterministic forecasts
Ensembles are better

From Schwartz et al. (2014); *Weather and Forecasting*
Ensemble verification: Calibration

National Weather Service forecasts

Nate Silver, *The Signal and the Noise*
Calibration

Local Media Forecasts

Nate Silver, *The Signal and the Noise*
Challenge with high-resolution ensembles

- One of the forefronts of NWP model research is how to design high-resolution ensembles
  - Vary just initial conditions?
  - Configure different members with different physics or dynamics?

- Each method has advantages and disadvantages

- General goal is to improve calibration
How to initialize high-resolution ensembles?

• Use existing operational ensembles
  – Cheap and easy but potential for mismatches

• Add random noise to a single field
  – A bit ad hoc

• Use ensemble data assimilation
What is data assimilation?

Gridded model forecast...the “background” or “first guess”

Data assimilation algorithm

Real observations

Background error covariances (errors of the background)

Observation errors

Statistically-optimal gridded “analysis”
Two (of many) data assimilation approaches

• Three-dimensional variational (3DVAR)
  – Background error covariances (BECs) typically fixed/time-invariant
  – May yield poor results when actual flow differs from that encapsulated within the fixed “climatology”

• Ensemble Kalman filter (EnKF)
  – Time-evolving, “flow-dependent” BECs estimated from a short-term ensemble forecast
Background errors and observations

• Ensemble spread (standard deviation) of wind speed

More model uncertainty: give observations more weight

Less model uncertainty: give observations less weight

From Schwartz et al. (2013); *Monthly Weather Review*
Continuously cycling data assimilation

- Usually 1- to 6-hrs between each analysis
Continuously cycling EnKF

• Initial conditions for all ensemble members are dynamically consistent
  – No ad hoc assumptions or use of external models
What we’re doing at NCAR/MMM

• Since April 7, 2015, we have been producing real-time, 10-member ensemble forecasts
  – 3-km horizontal grid spacing

• 50-member continuously cycling EnKF
  – 15-km horizontal grid spacing
  – New analysis every 6-hrs
  – Initializes the 10-member, 3-km ensemble forecasts
  – Use of EnKF to initialize high-resolution ensembles is unique

http://www.ensemble.ucar.edu
NCAR ensemble domain
Heavy precipitation probabilities
General precipitation placement

• Average 12-36-hr ensemble mean precipitation between April 7 and July 5, 2015
NCAR ensemble calibration

- Attributes diagrams for 18-36-hr precipitation over ~90 forecasts
Severe weather guidance

- Smoothed probabilities of the union of hail > 1 inch, wind exceeding 25 m/s, and intense mid-level rotation within 25 miles of a point within a 24-hr period

June 23, 2015
Severe weather guidance

- Smoothed probabilities of the *union* of hail > 1 inch, wind exceeding 25 m/s, and intense mid-level rotation within 25 miles of a point within a 24-hr period
Severe weather guidance

- Smoothed probabilities of the *union* of hail > 1 inch, wind exceeding 25 m/s, and intense mid-level rotation within 25 miles of a point within a 24-hr period

June 20, 2015
Closing thoughts

• High-resolution ensembles are the future
• Development at operational centers worldwide
• Challenges
  – Optimal ensemble design?
  – How to get well-calibrated forecasts?
  – How to best use the ensemble output?
  – How many members are necessary?
• Expect much effort on these topics in upcoming years