Ensemble Weather Forecasts

(a lecture in ATSC 507 - Num. Weather Prediction)

Prof. Roland Stull, PMet., CCM, CFII
Earth, Ocean & Atmospheric Sciences Dept.
University of British Columbia
rstull @ eoas.ubc.ca

...with lots of help from
Thomas Nipen, Dominique Bourdin, Katelyn Wells,
Henryk Modzelewski & Luca Delle Monache

Outline

• Role of ensembles in improving forecast skill.
• Operational ensemble forecast methods.
• Deterministic ensemble forecasts (DEF).
• Probabilistic forecasts from ensembles.
• Analog ensembles
• Ensemble-to-ensemble (E2E) models.
• Ways to display ensemble forecasts.

Ways to Correct Deterministic NWP Forecasts

• Systematic errors (bias) - remove with statistical post-processing (e.g., regression, neural networks, genetic programming...)

• Random errors (non-linear dynamics/chaos) - reduce by combining an ensemble of many weather forecasts into an ensemble average.

Lorenz & the Birth of Chaos Theory (1963)

Deterministic Nonperiodic Flow

Edward N. Lorenz
Massachusetts Institute of Technology

(Manuscript received 25 November 1962, in revised form 7 January 1963)

Abstract

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions. A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.
Sensitive Dependence on Initial Conditions

- Defines the limits of weather predictability (Unless you make a new discovery.)
- Even if NWP model was a perfect description of the weather, then if the model starts with a slightly different initial condition than the real weather, then the forecast diverges from truth.
Need to Fill the Pacific Data Void

- Paucity of upstream in-situ data over the Pacific causes errors in the initial conditions of NWP.
- This is currently the weakest link in making more accurate NWP forecasts for W. Canada, because of the sensitive dependence of fcsts on initial conditions.

As McCollor mentioned:

Forecasts are Less Accurate in W. Can. because of Bad ICs Upstream

- The potential economic loss in different regions of Canada is proportional to the area between the solid curve and the dashed line.
- W. Canada has the largest losses, starting at shorter range forecasts.

Data Denial Experiment

Kelly et al, 2007, QJRMS

Relative RMS errors in the 50 kPa geopotential heights, when all observations over the Pacific are excluded from the ECMWF data assimilation for Day 0, vs. those normally retained by ECMWF.

Green, blue, dark purple show worse forecasts, while yellow and red show positive impact.

The Problem gives the Solution

- Intentionally start with many slightly wrong initial conditions (ICs), and use the spread of the resulting forecasts to hopefully “bracket” the true weather.
Ensemble Mean & Spread

- The average of the ensemble forecasts is the best “deterministic” forecast (i.e., better than any individual run when averaged over many days of forecasts).
- The spread indicates the (a) uncertainty or skill, & (b) the probability of alternate outcomes.

Sample from ECMWF

(a) Temperature (°C)

Tight cluster => high confidence
Spread cluster => low confidence

Sampling from: North American Ensemble Fcst System (NAEFS) for Wolf River station, as modified at UBC

Verification of ensemble over past month at Wolf Riv.

Findings:
- Ensemble-average forecasts are better than any individual deterministic forecast, on average.

UBC Ensemble Hub-height Wind-speed Verification at Wind Farms

Accumulated Absolute Error for Day 2 forecast for each individual SEIF ensemble member from the start of the ensemble. Less is better.
Error Growth

For IC perturbations, ensemble average gives worse forecast for first 1 to 2 days. Thus NOT useful for short-range forecasting.

Outline

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.

Many ways to perturb a reference fcst to generate ensembles

1. ICs
2. BCs (for limited-area model)
3. physics
4. numerics
5. grid resolutions
6. lagged in time
7. models (= multi-model ensemble)
8. terrain
9. compilers & compiler optimizations

At UBC, we use (1-5), (7) & (soon 8) for Short-range Ensemble Fcsts (SREF)

History of Operational Ensembles

- 1988 - UK Met. Office (UKMO)
- 1992 - ECMWF - multi-IC (singular vectors)
- 1992 - NCEP - multi-IC (bred vectors)
- 1996 - UBC - multi-model, multi-resolution
- 1996 - CMC - multi-IC (via data assim; perturbed obs.)
- 2004 - CMC & NCEP & NMSM => NAEFS (super-ensemble)
We make operational, daily ensemble forecasts using multi models, multi initial & boundary conditions, and multi resolutions. As of 2020, it is a 45 member ensemble.

**Outline**
- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- **Deterministic ensemble forecasts (DEF).**
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.

**Ways to Create the Ensemble-based Deterministic Fcsts**
- Use average
- Use weighted average (weighted inversely by error variance)
- Use median
- Use nonlinear combinations of ensemble members.
- Use Gene-Expression Programming (GEP), see Bakhshaii & Stull 2009 WAF.
- Use analogs (Julia’s PhD research)
Deterministic Ensemble Forecasts (DET)

- Simple (linear) ensemble mean:
  \[ T_{\text{ens}} = (1/N) [T_{\text{model1}} + T_{\text{model2}} + T_{\text{model3}} + \ldots + T_{\text{model n}}] \]

- Example of nonlinear DET from GEP (see next pages.)

Examples of non-unique fits by GEP to the synthetic sigmoid data

In this illustration, GEP was allowed to use only:

- \((+, -, L)\) where \(L\) is the logistic function \(L(x) = \{1 + \exp(-x)\}^{-1}\) (thick line)
- \((+, -, \text{Power})\) (thin dashed line, mostly hidden behind the thick line)
- \((+, -, \text{mod})\) where "mod" is the floating-point remainder (thin solid line)

These illustrations show how GEP inexorably approaches a best fit, even when forced to use non-optimum functions.
Comments on DEFs (e.g., ensem. averages)

- Ensemble average is physical unrealizable. (but Rachel will investigate its utility)
  (E.g., winds don’t agree with pressures)
- For ensembles sharp features (e.g., fronts) that progress at different speeds, the ensemble average causes unphysical smoothing.
- Ensemble average loses skill during regime changes.
- Must Bias-correct each member BEFORE combining into an ensemble.
- Ensembles with more (well chosen) members are more accurate than small ensembles. But don’t include known bad members.
- Data voids (such as NE Pacific) cause all ensemble members to have similar errors (i.e., small spread NOT due to high confidence of an accurate forecast).
- Imposed lateral boundary conditions can dominate limited-area models (LAMs), resulting in poor ensemble spread.

But why stop with deterministic fcsts ...

- ...when you have so much more info from the ensemble spread?

Outline

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.
Deterministic Forecasts

after postprocessing & ensemble-average

Wind Speed (m/s)

Threshold

0  25  50

Temperature (°C)

Threshold

0  25  50

...but the forecast can have errors.

Probabilistic Forecasts

Wind Speed (m/s)

Threshold

0  25  50

Temperature (°C)

Threshold

0  25  50

$p$ = probability of threshold exceedance

---

Decision-making with Probabilistic Forecasts

Let: $C = $$$ cost to try to mitigate/avoid an event
    $L = $$$ lost if event happens (without mitigation)

Then: $r = C/L$ cost loss ratio

Let: $p = $forecasted probability of exceeding a threshold

Take action whenever: $p > r$

Namely, take action when: $C < p \cdot L$

E.g.: Loss is $$ needed to replace overheated transmission line.
     Cost is $$ not earned by reducing the amperage, or fines.

Solution:

Cost to protect the blades (postpone the replacement) = $165k.
Loss if blades damaged during attempt = $970k.

Cost/Loss ratio $R \approx 0.17$

$p > R$, Therefore do not replace today.
Sample Probability Fcst for Dokie wind farm in NE BC

Comments on Probability Forecasts

• Needs of meteorologists are different from needs of end users.
• Many industries could use prob. fcsts. to very good economic advantage, but are unable to realize it. (The industry needs a meteorologist on staff to use prob. fcsts.)
• Laypeople are totally clueless, and think that prob. fcsts. are a joke.
Outline

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.

Analog Ensemble

- Method: Use past weather forecasts that best match today’s synoptic regime as ensemble members.
- Advantages: handles regime changes better, and automatically eliminates NWP model biases.

Delle Monache, Nipen, Liu, Roux, Stull, 2011, MWR.
Analog Ensembles

Finally, use only the closest analogs, but take their verifying observations, not their NWP fcsts, as the ensemble members.

How skillful is AnEn?

- AnEn generated with Environ. Canada GEM (15 km), 0-48 hours
- Comparison with Environment Canada Regional Ensemble Prediction System (REPS, 20 members, 33 km grid spacing)
- Period of 15 months (verification over the last 3 months)
- 10-m wind speed, 2-m temperature
- 550 surface stations over CONUS
- Probabilistic prediction attributes: reliability & sharpness, statistical consistency, utility/value

Ground truth dataset

- 550 hourly METAR Surface Observations
  - 1 May 2010 – 31 July 2011, for a total of 457 days
  - 10-m wind speed

AnEn Results

Delle Monache et al 2011, MWR
Analog Ensembles (AnEn)

Results:
AnEn (from only one NWP run) had comparable verification scores (reliable, sharp, consistent, valuable) as the 20-run GEM ensemble.

AnEn is best at capturing changes in synoptic regime.

THIS COULD BE A GAME CHANGER !!

Julia at UBC is researching AnEn.

Outline

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.
(E2E) Ensemble outputs as inputs to other ensemble models

- \( \text{wx} \) -- CMAQ (air pollution dispersion)

- \( \text{wx} \) -- hydrologic (for reservoirs)
  See papers by Bourdin & Stull, 2013.

**Outline**

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.
CMC contour of 50 kPa sfc
Spaghetti Map - 96 h

CMC contour of 50 kPa sfc
Spaghetti Map - 144 h

CMC contour of 50 kPa sfc
Spaghetti Map - 192 h
Separate zero from nonzero precip. members at UBC:

Average of ONLY those ens. members that have nonzero forecasts exceeding a threshold.

Out of all members, fraction of ens. members that have nonzero forecasts exceeding a threshold.

CMC: contours of ensemble average SLP, with red and blue showing center locations for Lows & Highs.

UBC Ensemble Cloud Cover: version for meteorologists.

UBC Ensemble Cloud Cover: version for public.
Each model was bias-corrected using Kalman filter (KF) before combining into the ensemble.

...for future lecture by Dr. Thomas Nipen

- ensemble calibration / calibrated probabilistic forecasts
- other postprocessing of ensembles

Summary: Ensembles

- Role of ensembles in improving forecast skill.
- Operational ensemble forecast methods.
- Deterministic ensemble forecasts (DEF).
- Probabilistic forecasts from ensembles.
- Analog ensembles
- Ensemble-to-ensemble (E2E) models.
- Ways to display ensemble forecasts.