NOAA Earth System Research Laboratory

Notes on reforecasting and the computational capacity needed for future SREF systems

Tom Hamill
NOAA Earth System Research Lab

presentation for 2009 National Workshop on Mesoscale Probabilistic Prediction

Part I:
notes on reforecasting and its applicability to SREF

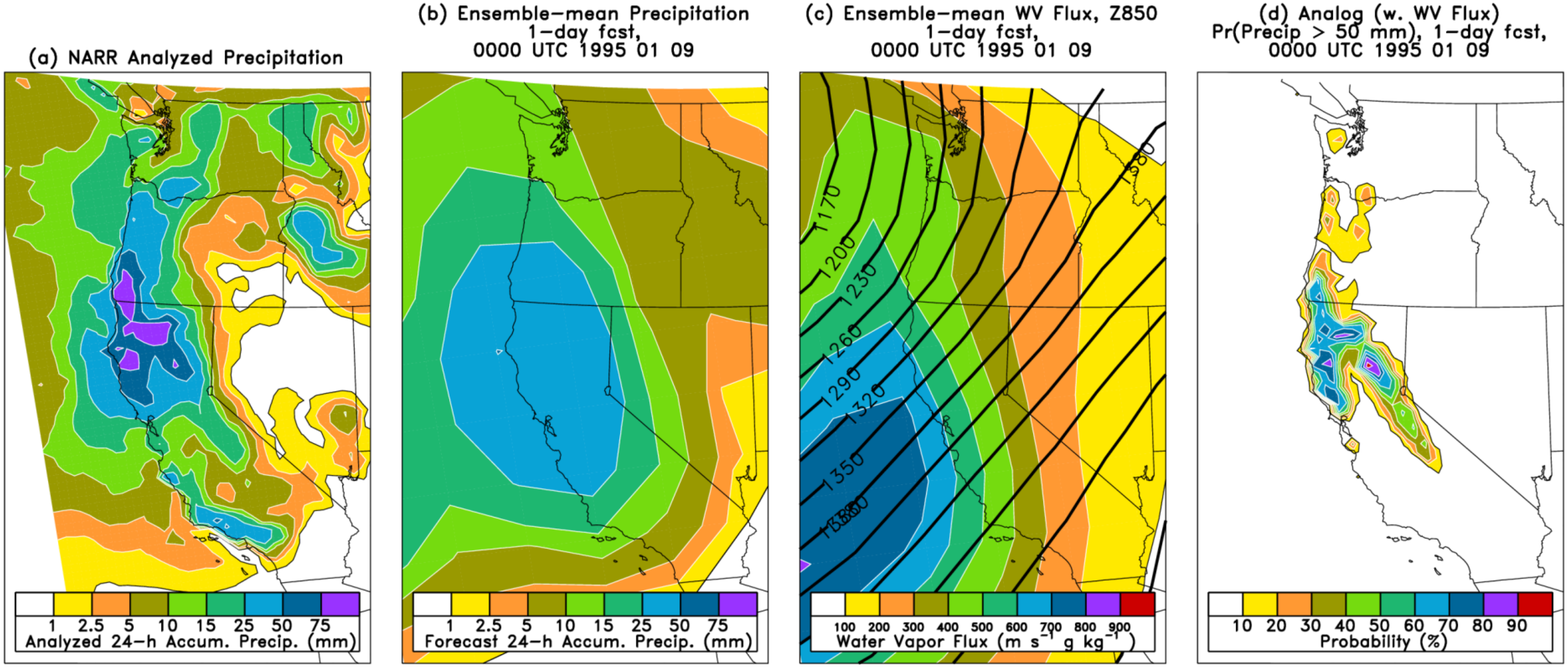
What are reforecasts?

- Forecasts for dates in the past using the same forecast model and data assimilation system as is used operationally.
- Also called “hindcasts” but we prefer “reforecast” to make the connection with reanalyses.
- The (relatively) consistent data set allows effective statistical post-processing algorithms to be developed, deterministic and probabilistic.

Reforecasting, lessons learned

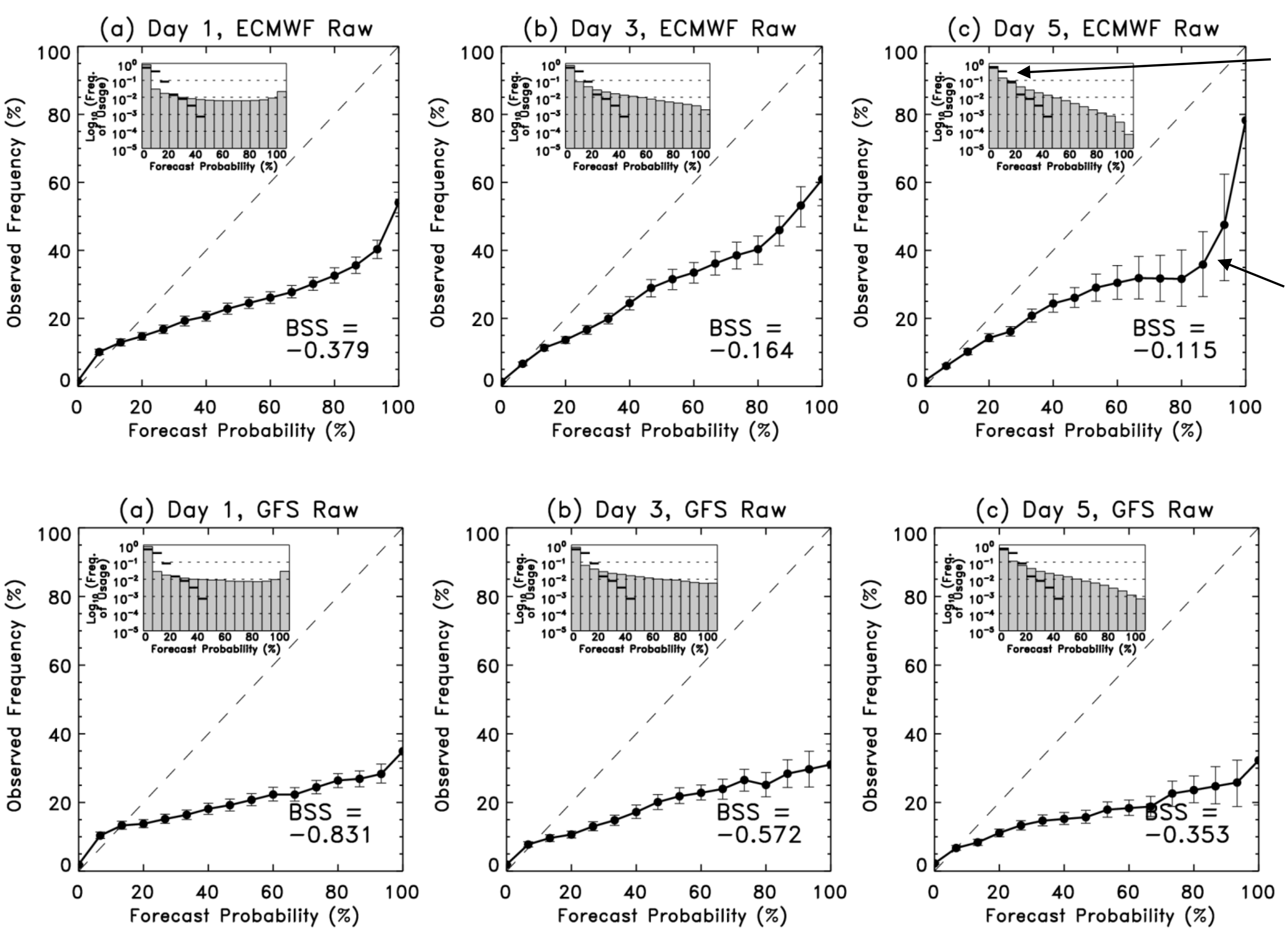
- Common misapprehension: reforecasting is not a calibration technique, but a data set that facilitates a wide array of calibration techniques.
- Reforecasts provide statistical downscaling (as opposed to dynamic, i.e. hi-res models)
- Reforecasts improve calibration of long-lead forecasts, rare-event forecasts. In both cases, need large sample size. Less impact relative to short training data sets for short-term temperature forecasts.
- Reforecasts are important to the hydrologic community, who need to train and test on large samples to cover many flooding events.
- Can't do post-processing without quality reanalysis and observation data sets
  - Up-to date reanalyses, to initialize previous year's forecasts with the same data assimilation system used in real time, achieving consistency between training data set and real-time forecasts.
  - Long time series of observations (for post-processing model development)
- Reforecasts may be more difficult with multi-model SREF, may affect choice of SREF system.
- Reforecasts are no longer an exotic technology; operational at ECMWF, soon at NCEP with Climate Forecast System Reanalysis & Reforecast.

Statistical downscaling using T62 GFS reforecasts



The relative merits of statistical vs. dynamical downscaling is an important research question to answer.

Example of imperfections in 2005 ECMWF, 1998 GFS 5-mm reliability diagrams, raw ensembles

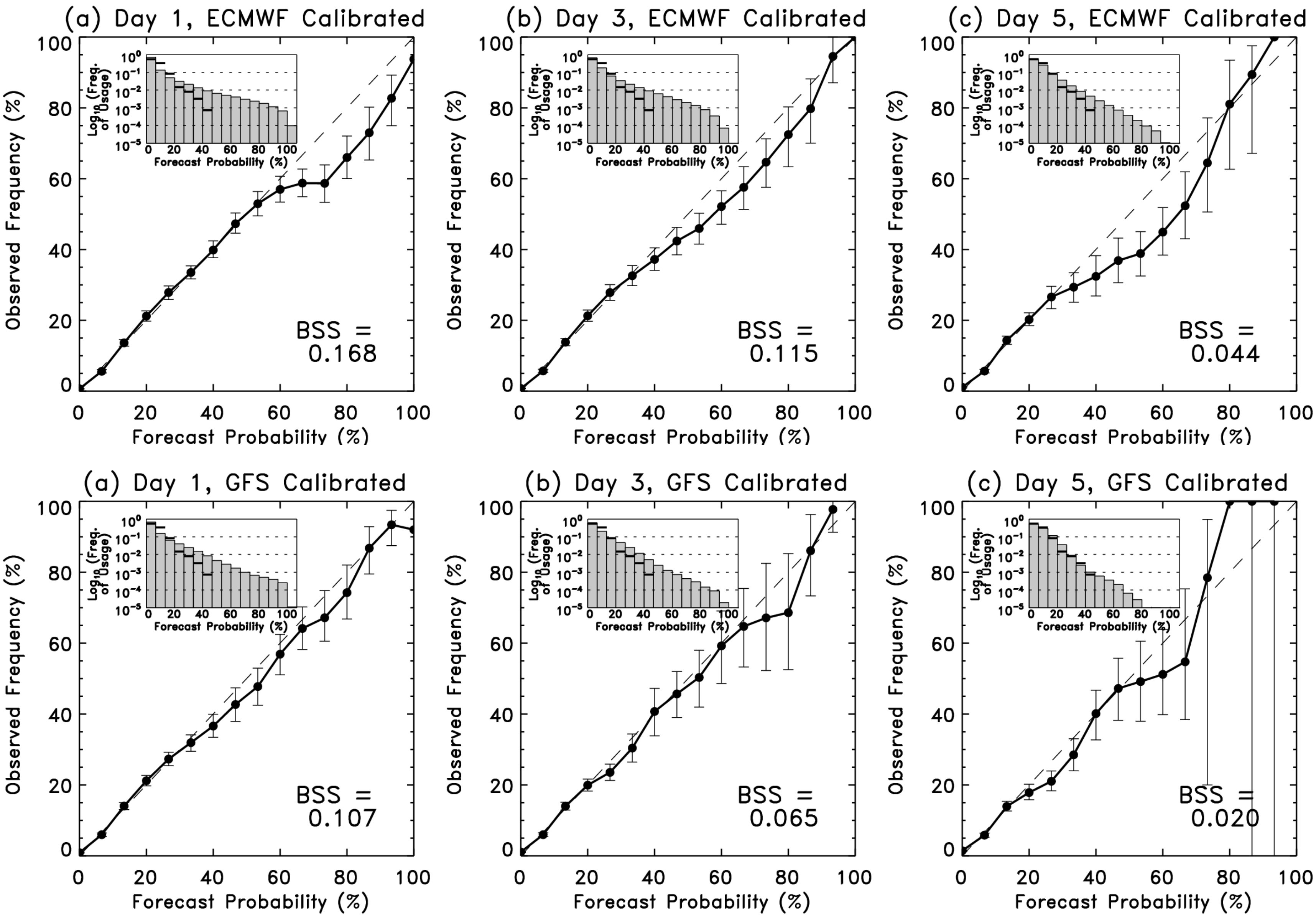


horizontal lines indicate distribution of climatology

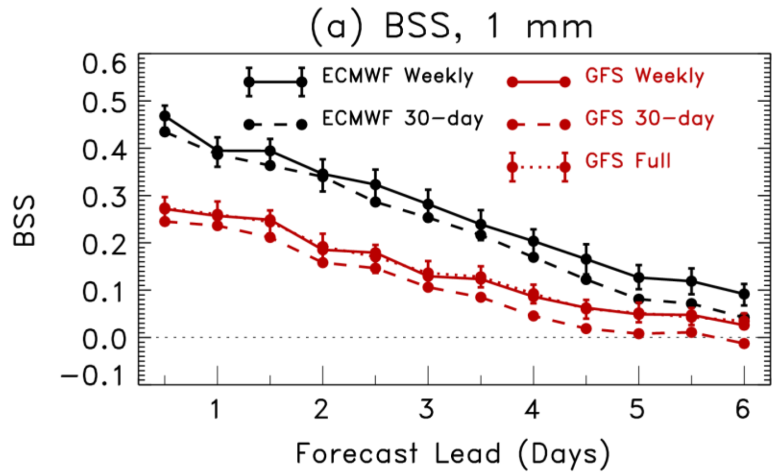
error bars from block bootstrap

Raw forecasts have poor skill in this strict BSS

5-mm reliability diagrams, after reforecast calibration

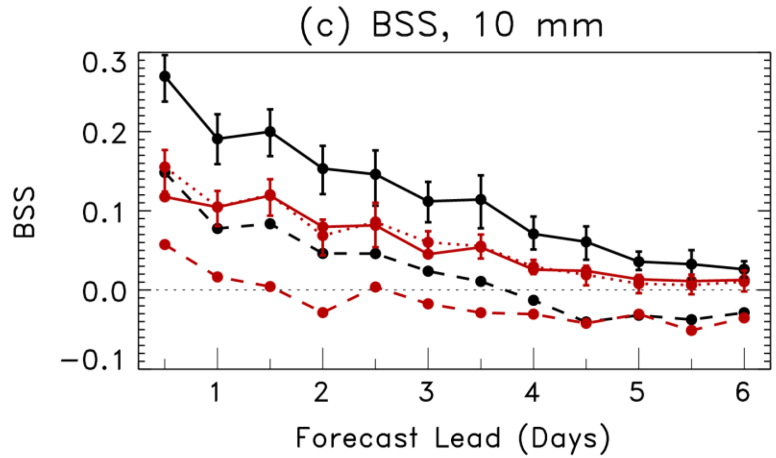


Precipitation skill with 1x weekly (x 20 year), 30-day, and full training data sets



Note the substantial benefit of weekly relative to 30-day training data sets, especially at high thresholds.

Note diminishing returns of increasing sample size beyond 1x weekly.



Part II:
sizing the computers we need.

Current ensemble forecasts at NCEP

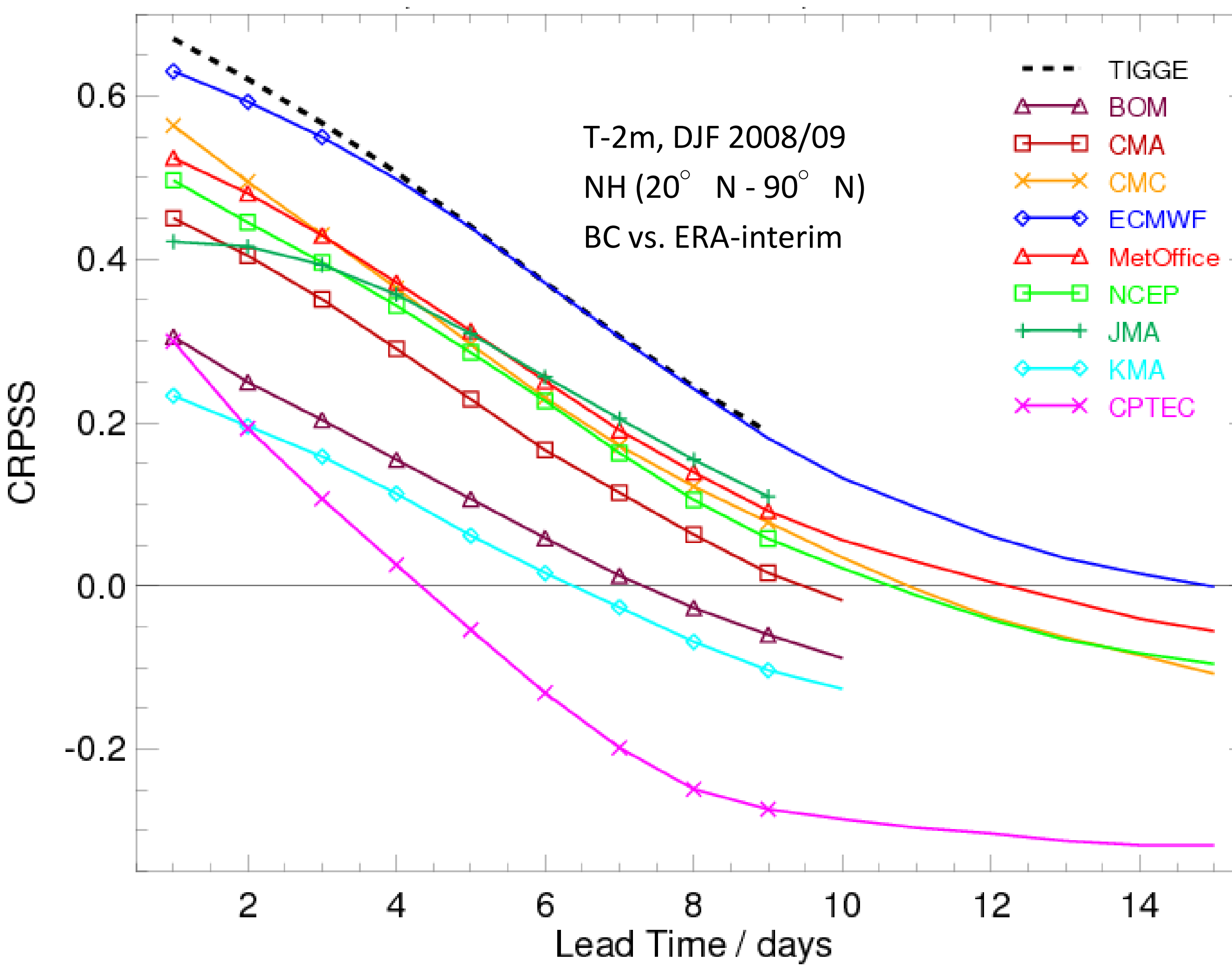
- Current NCEP Global EPS configuration: 4 cycles daily, 20 members/cycle, to 16 days. Uses single-model T128L28 (~150 km). Near-term plans to increase to T190L28, which we will assume as the current baseline. Through North American Ensemble Forecast System (NAEFS), access to Canadian ensemble data (0.9-degree, L58, 20 members/cycle).

- Current NCEP SREF configuration: 4 cycles daily, to 87h lead, 21 members/cycle, horizontal grid spacings ~40 km, !~28 levels. Initial conditions and LBCs interpolated from NCEP global model. Soon resolution upgrade. Assume 32kmL28M20 as baseline for existing SREF system.

What are other weather prediction facilities running currently?

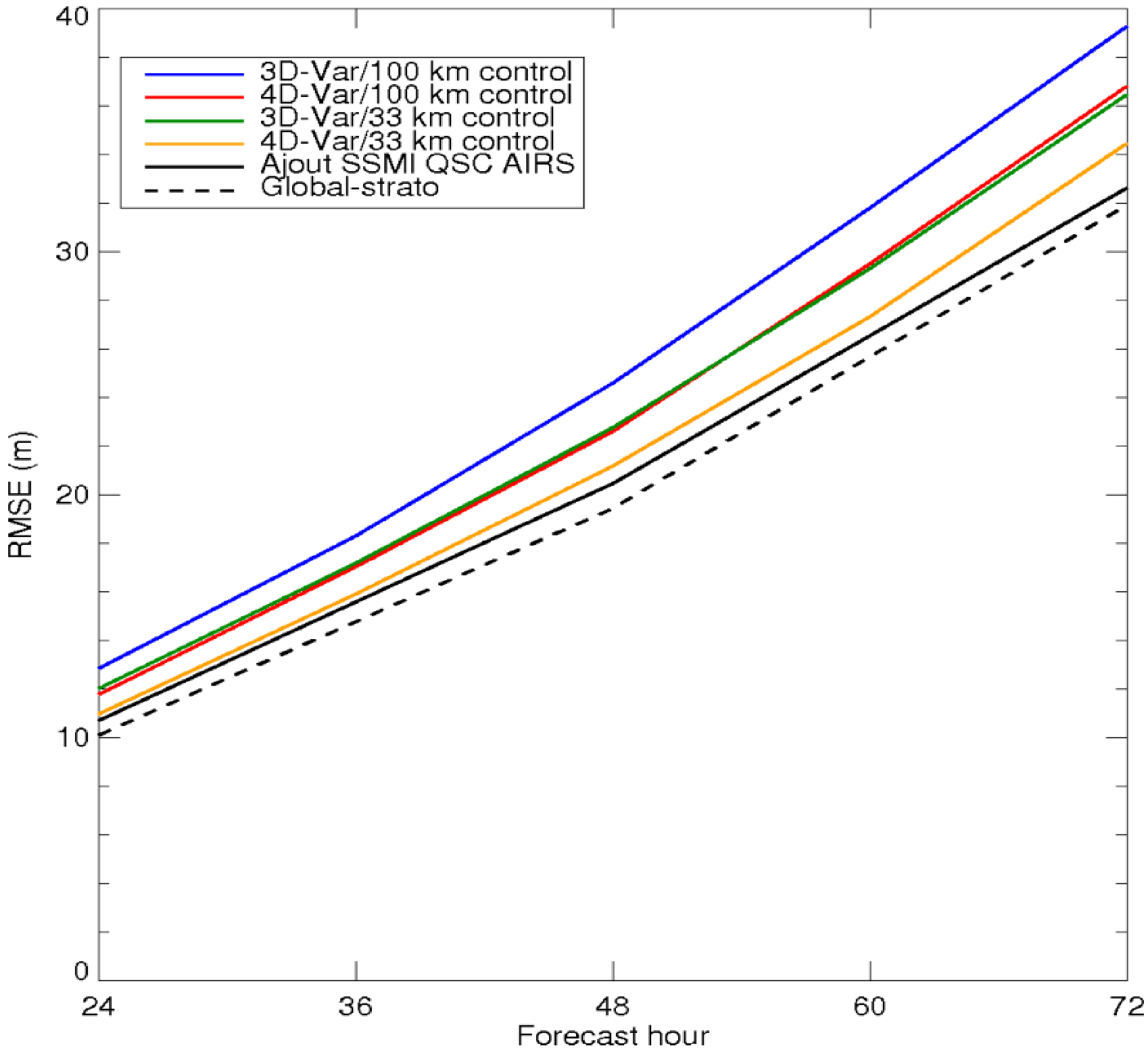
- Global models: current reference standard and world-best is ECMWF modeling system.
  - ECMWF implemented a 2x daily, T255 global ensemble (50 members) in Nov. 2000; a T399 ensemble in Nov. 2006; are planning for T639 ensemble in 2010. Their strategic plan calls for a 10-km deterministic and 20-km global ensemble (~ T1000) by 2015.
- Regional ensemble systems. Future plans for regional ensembles are more vague, and computational requirements depend on domain size, which vary radically. Current regional ensemble prediction systems include:
  - *UK Met Office*: 24 members, 24-km grid spacing to 54 h. Plans for 18-km L70 regional and 1.5-km small regional ensemble around British Isles by 2011.
  - *DWD (Germany)*: 2.8-km, 20 member regional ensemble over Germany, run currently. Plans beyond 2013 TBD.
  - *CMC (Canada)*: Plans for regional ensemble of 33 km/L60, 20 members to 48-h lead by 2010. 22-km L70 to 72h, 20 members in 2013.
  - Several consortia in Europe to share regional ensemble model output and distribute computations, e.g., COSMO consortium; talk. w. Jose Garcia-Moya for more details.

2-m temperature skill from TIGGE data set



ECMWF's investment in better, higher-resolution models, data assimilation provides approximately 2 extra days of forecast lead over the nearest competitor. An investment to improve resolution thus can be expected to provide substantial benefit to NWS.

The impacts of improved data assimilation and higher-resolution models



North America Z500 RMSE for the control experiments and latest upgrades of the MSC global analysis-forecast system (January and February 2007)

Acknowledgements to S. Laroche, Environment Canada

A proposal for the sizing of future NCEP ensemble systems

- Straw(person) : NCEP should be aiming for the following operational capacity by 2016, in order to be competitive with the best in the world (100% capacity):

- Global ensembles: T1000 L100 M20 (~20-km global ensemble) 4x daily to day 3; T500L50M20 to 7 days; T250L50M20 to 15 days lead. Accompanying reforecasts in order to statistically post-process. 255x more CPU than current; 147x more storage than current.
- Regional ensemble:
  - (1) SREF on large regional domain (similar to current domain, covering N America) at 4 km L100 M20, 4x daily to 24h lead, thereafter 12kmL50 M20 to 87-h lead. Accompanying reforecasts (5 members every 3rd day over 20 years) in order to statistically post-process. EnKF initialization.
  - (2) Limited-area, high-impact, on-demand ensemble prediction system run within SREF domain. 2 km, L100 M20 to 24-h lead. Assume 4 regional domain ensembles computed per day. Assume domain will cover 1/4 the size of the existing SREF domain in both E-W and N-S directions (1/16th the area). No reforecasts.
  - 3906x more CPU than current; 1283x more mass storage than current.

- **A laughably high cost estimate**, with NOAA's method of calculating CPU costs (same future CPU cost/cycle as in 2008.

A (more modest) proposal for the sizing of future NCEP ensemble systems

- Global ensembles: T1000 L100 M20 (~20-km global ensemble) 4x daily to day 3; T500L50M20 to 7 days; T250L50M20 to 15 days lead. Accompanying reforecasts in order to statistically post-process. 255x more CPU than current; 147x more storage than current..
- Regional ensemble:
  - (1) 12-km L50M20 4x daily to 87h. Regional 4-km ensemble to 24-h, run ~4x daily.
  - (2) Reforecasts with 12-km ensemble, 10 members, every 3rd day.
  - (3) Total: 265x more CPU than current; 61x more mass storage
- **A more modest cost estimate**, with NOAA's method of calculating CPU costs (same future CPU cost/cycle as in 2008). This didn't cause sticker shock at HQ NOAA

My perception

- Increased CPU for numerical weather prediction is perceived as important, but it’s not yet near the list of the most important upgrades NOAA wants to fund.

- And it darned well ought to be.

A signed statement to the NWS, NOAA, and Dept. of Commerce Leadership?

A Recommendation from the National Workshop on Mesoscale Ensemble Prediction concerning the Procurement of High-Performance Computing by the National Weather Service.

Whereas the United States lags badly in the skill of its numerical weather predictions relative to international competitors;

Whereas the United States must constrain the resolution of its numerical weather prediction models to much less than those of its international peers so that they may run on far smaller computers;

Whereas NOAA's investment in high-performance computing for weather prediction is approximately two orders of magnitude smaller than its investment in weather satellites it is required to process;

Whereas the lack of sufficient high-performance computing is a primary impediment to the development and testing of improved weather prediction models and methods;

Whereas the transition to probabilistic weather prediction involves additional computational burdens, such as the computation of ensembles of forecasts and the statistical post-processing of these using additional past forecasts and observations;

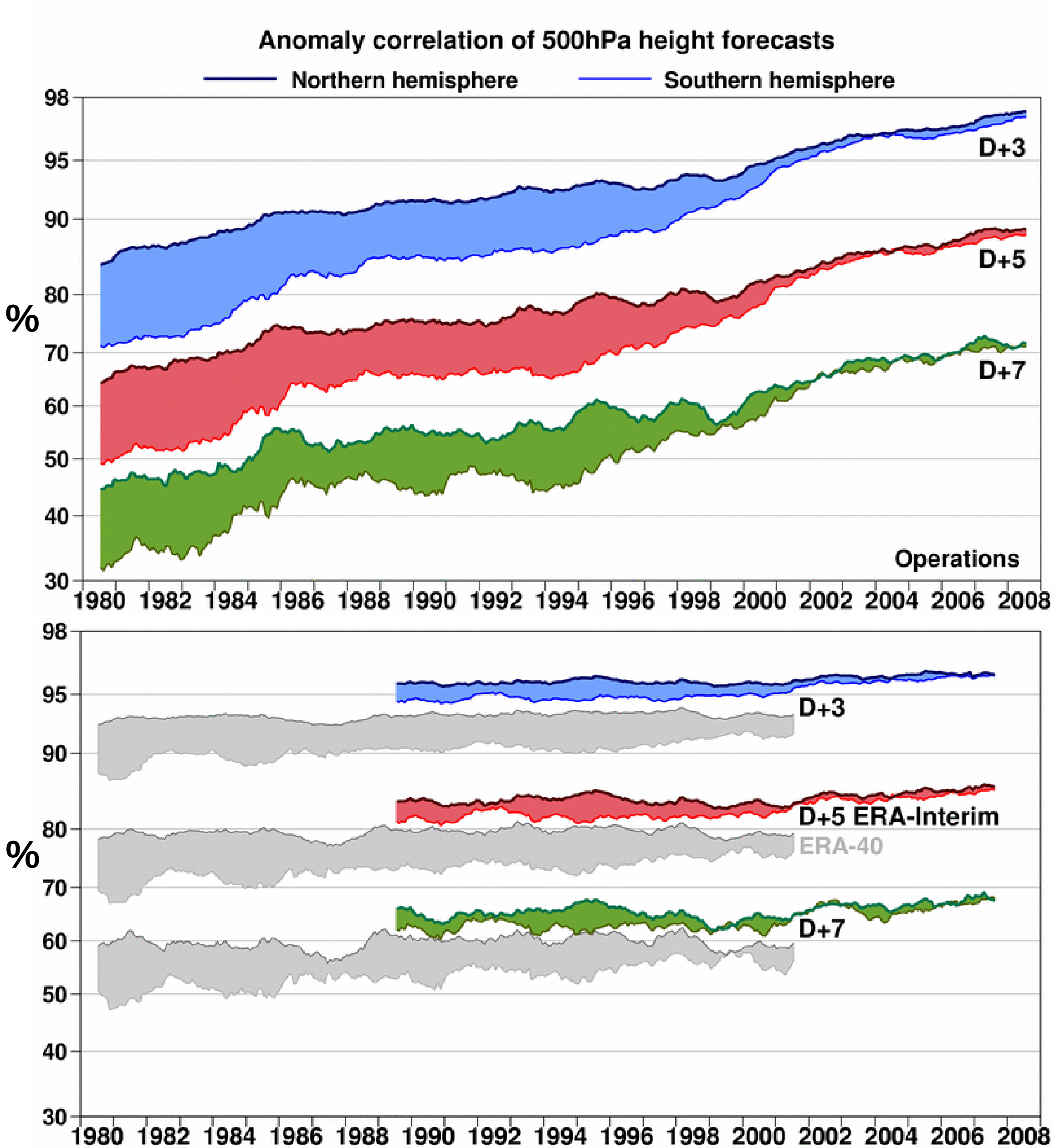
Whereas improving model resolution and implementing ensemble prediction techniques has consistently resulted in improved numerical weather predictions;

Whereas improving the skill of these numerical weather predictions dramatically improves forecasts of high-impact events such as hurricanes, tornadoes, floods, heat waves, and cold spells;

The members of the 2009 National Workshop on Mesoscale Ensemble Prediction unanimously recommend that NOAA procure greatly increased computational capacity for weather prediction to support higher-resolution models and improved ensemble forecast techniques.

Signed, *Lewis F. Richardson* et al.

Forecasting-system improvement at ECMWF



Updated from Simmons & Hollingsworth (2002)
Acknowledgements to A. Simmons

Historical trend

Historical re-forecast project trend using re-analyses

c/o Gilbert Brunet THORPEX presentation