How Do Forecasters Utilize Output from a Convection-Permitting Ensemble Forecast System? Case Study of a High-Impact Precipitation Event

Clark Evans (evans36@uwm.edu), Donald F. Van Dyke III, and Todd Lericos

A COMET Partners Project between UW-Milwaukee and National Weather Service WFO Tallahassee, FL

Introduction

We are primarily motivated by two interrelated questions:

- How are ensembles best leveraged to result in improved weather forecasts across scales and forecast elements?
- How does forecasters' utilization of ensemble guidance vary, and what impact does it have upon the forecast?

Background

Previous research has demonstrated that:

- Ensemble usage depends upon the type of forecasts being made: they are used more when ensemble output is matched to specific forecast requirements.
- Ensemble reliability, display methods, and perceived utility all influence how forecasters use of ensemble guidance.

We evaluate how forecasters utilize convection-permitting ensemble guidance when a specific high-impact weather event – Tropical Storm Fay in 2008 (**Figs. 1 and 2**) – is forecast.

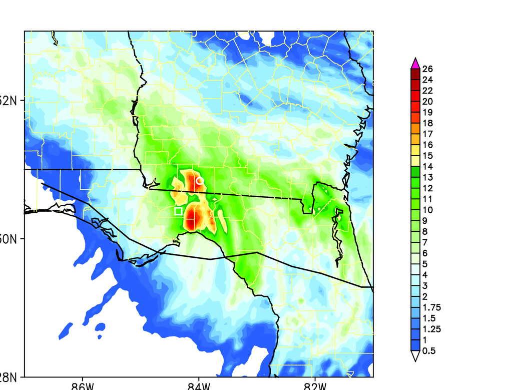
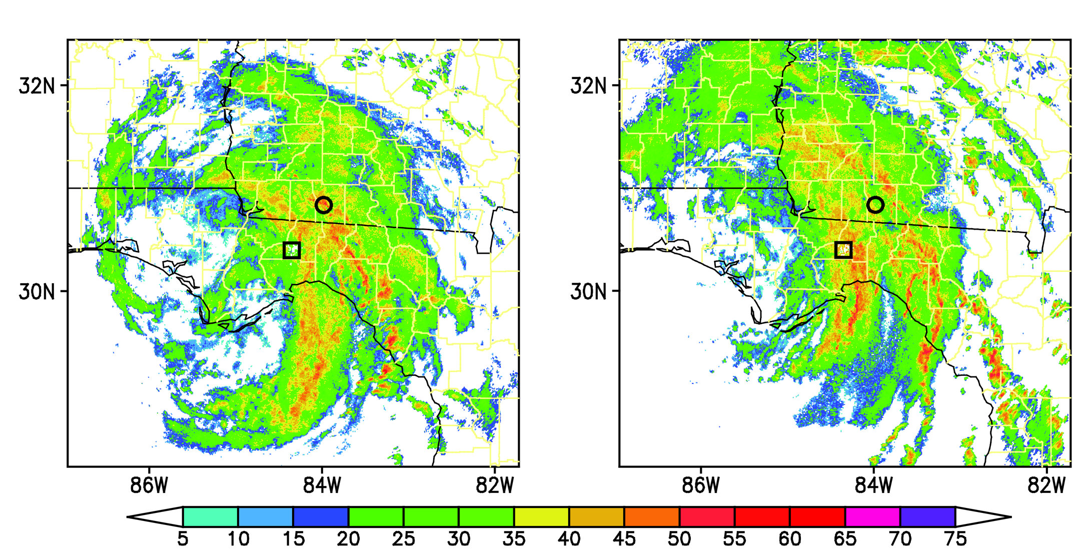


Figure 1 (top): 72-h accumulated precipitation (in) ending 0000 UTC 25 August 2008. The track of Fay is given by the black line.

Figure 2 (bottom): Observed Level II radar (dBZ) from Tallahassee, FL (KTLH) at 1203 UTC (left) and 1801 UTC (right) 23 August 2008.



Ensemble Formulation

A sixteen-member 72-h ensemble forecast of Tropical Storm Fay is conducted using WRF-ARW v3.3 over a SE US domain (Δx = 4 km, 41 σ levels). Forecasts begin 0000 UTC 22 August 2008.

Ensemble diversity is achieved by varying initial and lateral boundary conditions and selected physical parameterizations (PBL, surface layer, cloud microphysics, SKEB perturbations).

Forecasting Exercise

To evaluate how forecasters use ensembles, a forecasting exercise is conducted. All Fay guidance are shifted to the Houston/Galveston, TX area and presented as T.S. "Trixie."

Forecasters create a 72-h QPF forecast using only deterministic and operational guidance, then revise it after being presented with the ensemble forecast guidance in multiple forms.

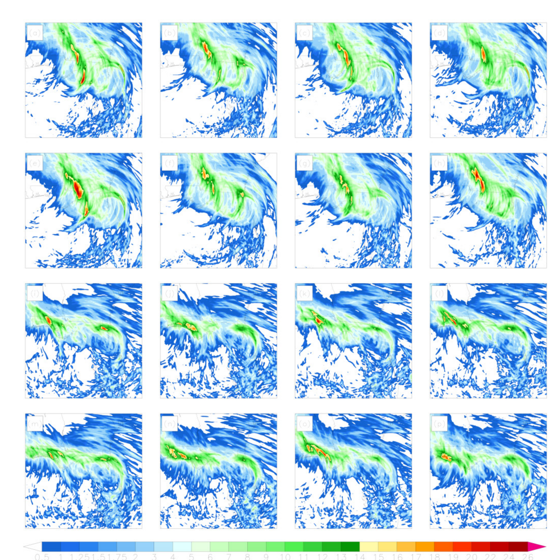
After each phase of the exercise, surveys are conducted to document subjective forecast impressions related to the specific forecast scenario and how the ensemble was used.

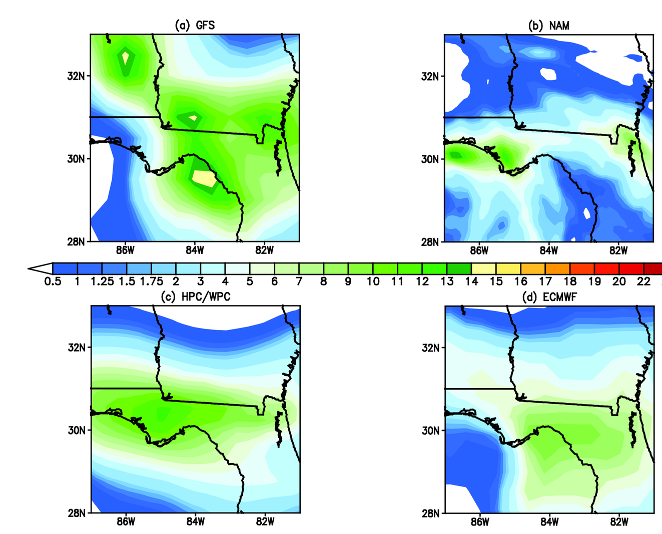
Ensemble Performance and Forecast Verification

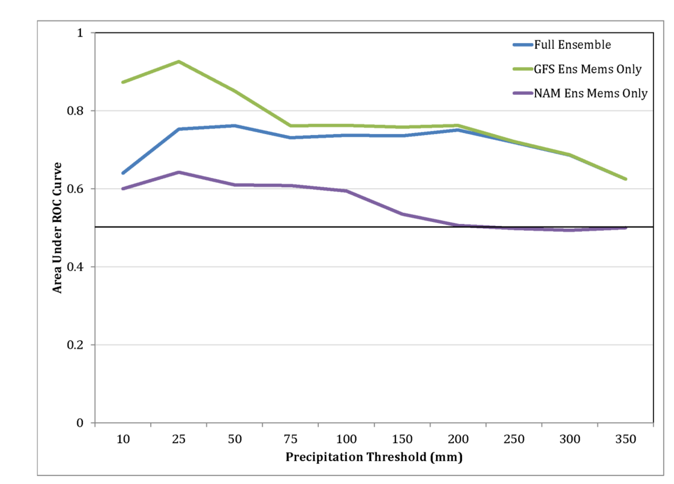
Initial and lateral boundary condition variability (**Fig. 3a-h, i-p**) influenced changes in QPF locations while physical parameterization variability influenced changes in QPF values. Ensemble forecasts reflect downscaled versions of the guidance from which they are initialized (**Figs. 3 and 4**).

The primary forecast challenge, therefore, is pinpointing the location and amount of maximum rainfall within a localized region. To that effect, the ensemble provides a skillful forecast relative to climatology (**Fig. 5**), but one that is underdispersive and biased (**Fig. 6**).

Forecasting exercise participants' forecasts prior to viewing ensemble output are broadly similar (**Fig. 7**). There exist notable differences in how considering ensemble output resulted in changed forecasts (**Figs. 8 and 9**). In the aggregate, forecast skill improved after considering ensemble data. (**Fig. 10**).







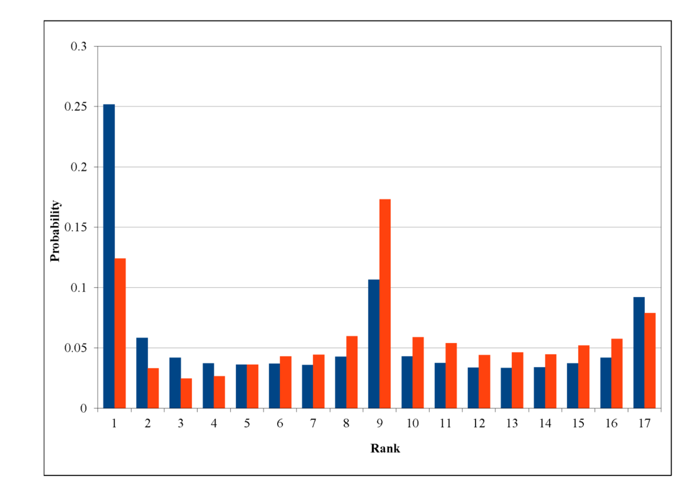


Figure 3 (top): 72-h accumulated precipitation (in) ending 0000 UTC 25 August 2008 from each of the sixteen ensemble members. Members (a-h) utilize GFS initial and lateral boundary conditions; members (i-p) utilize NAM initial and lateral boundary conditions.

Figure 7 (bottom): 72-h QPF (in) forecasts made prior to viewing the ensemble data for (a-g) forecasting exercise participants and (h) the mean of all exercise participants' forecasts.

Figure 4 (top): 72-h accumulated precipitation (in) ending 0000 UTC 25 August 2008 from the (a) GFS, (b) NAM, (c) HPC/WPC, and (d) ECMWF. Panels (a, b, d) are model forecasts, while panel (c) is an operational (human-aided) forecast.

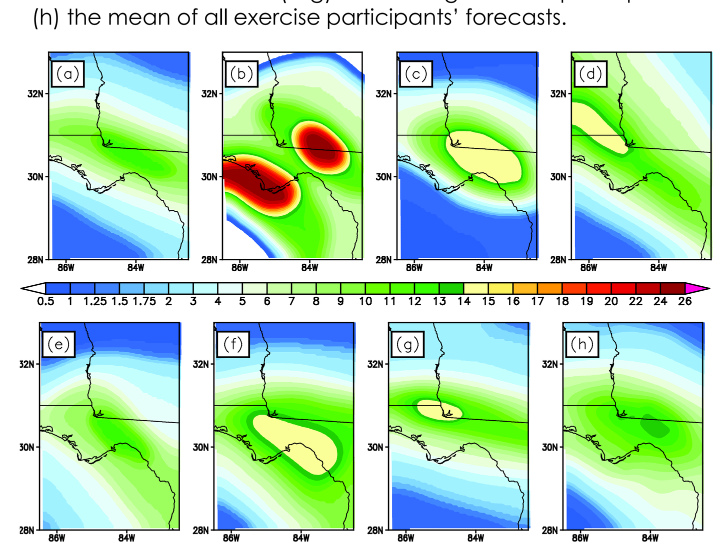
Figure 8 (bottom): 72-h QPF (in) forecasts made after viewing the ensemble data for (a-g) forecasting exercise participants and (h) the mean of all exercise participants' forecasts.

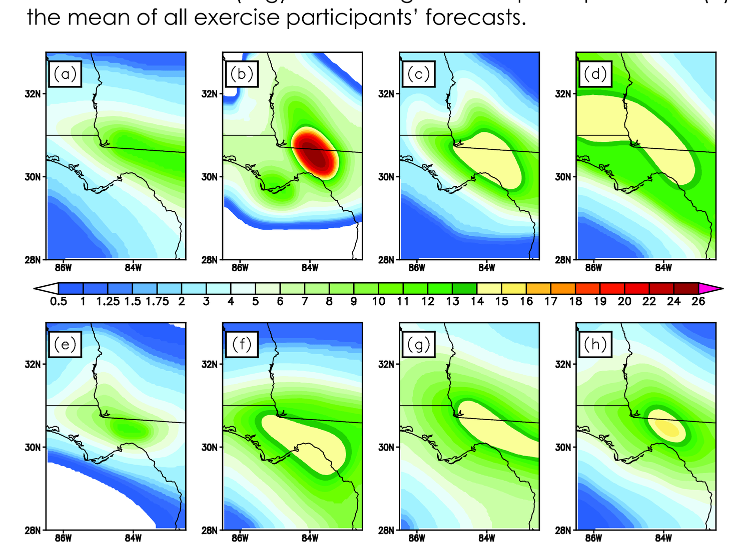
Figure 5 (top): Area under the ROC curve for the full ensemble (blue), GFS-initialized members (green), and NAM-initialized members (purple) at 72-h QPF thresholds of 10 to 350 mm (0.4" to 13.75". Values > 0.5 are said to be skillful relative to climatology.

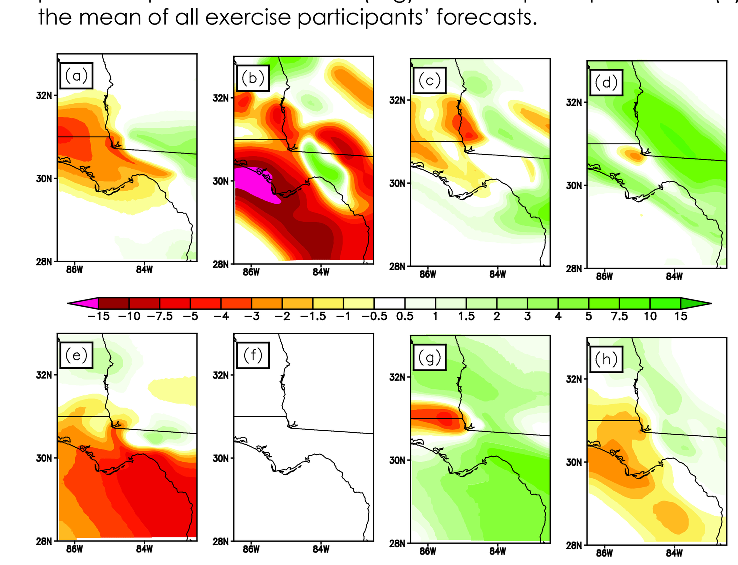
Figure 9 (bottom): Difference in 72-h QPF forecasts (in), defined as pre-minus-post ensemble, for (a-g) exercise participants and (h) the mean of all exercise participants' forecasts.

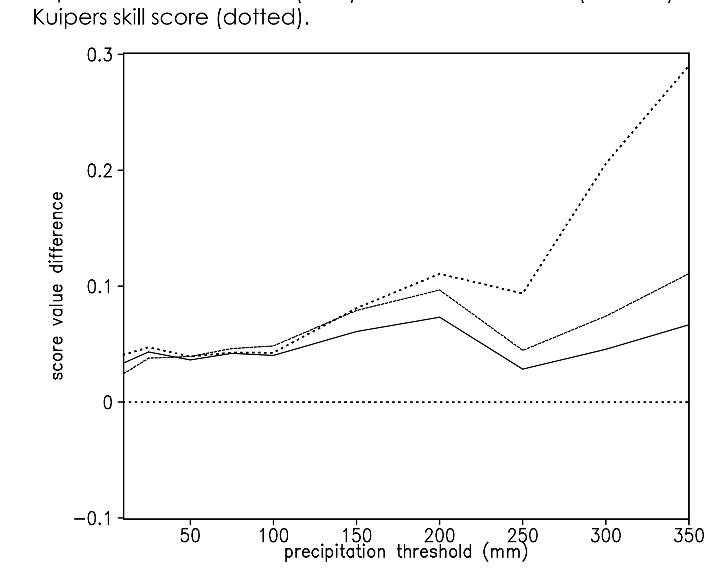
Figure 6 (top): Area under the ROC curve for the full ensemble (blue), GFS-based members (green), and NAM-based members (purple) at 72-h QPF thresholds of 10-350 mm (or 0.4"-13.75").

Figure 10 (bottom): Mean difference in 72-h QPF forecast forecast skill (in), defined as post-minus-pre ensemble, as assessed using the equitable threat score (solid), Heidke skill score (dashed), and Kuipers skill score (dotted).









How *Do* Forecasters Utilize Ensembles?

The forecasting exercise illuminated several key findings regarding how forecasters use ensembles:

- Forecasters typically used ensemble guidance to identify **most likely** forecast outcomes and to **quantify uncertainty**. Post-processed fields that helped them to do so were well-received.
- Approximately half of the forecasters surveyed desired information regarding **only the maximum, minimum, and mean QPF** forecast by the ensemble. The remaining half desired information regarding the **spread of ensemble forecasts**, but only if presented in a concise way (e.g., plumes).
- Most felt that the ensemble guidance added forecast value by **increasing forecast confidence**, whether it agreed with their initial forecast or provided support for forecasting higher QPF amounts.
- Two forecasters **discarded the ensemble** guidance because they perceived it to be **biased and/or underdispersive**; three others noted this but still felt that the ensemble added value to the forecast.

Future Directions

There exist many R2O/O2R challenges regarding ensembles:

- What is the event-to-event, person-to-person, forecast-to-forecast variability in how forecasters use ensembles?
- How are ensembles best constructed, and how are element-specific data best mined from ensembles?
- How are the barriers to ensemble acceptance, perceived or real, best overcome – training, visualization, improved skill?
- How are ensembles *best* used in the forecast process?

Acknowledgments and Further Information

We gratefully acknowledge the assistance of the nine participants in the forecasting exercise. Funding support for this research was provided by a sub-award from UCAR under the sponsorship of NOAA/DOC as part of the COMET Outreach Program. Please scan the QR code at left to access our companion 2014 *Weather and Forecasting* manuscript.