Preliminary precipitation verification of some MPEX forecasts

Craig Schwartz and Kevin Manning (or Kevin Manning and Craig Schwartz)

Verification information

- Retrospective evaluation of NCAR-produced MPEX forecasts
 - EnKF-initialized ensemble of WRF model forecasts
 - GFS-initialized deterministic WRF model forecasts
- 30-member ensembles initialized from 1200 UTC (36 cases total)
- All statistics are aggregated hourly over many cases
- NCEP Stage IV precipitation considered the "truth"

WRF domains and verification region



Precipitation "calendar"



Date

Normalized total precipitation over the verification domain, each hour from 1200 UTC May 16 to 1200 UTC June 15

BLACK: Stage IV Colors: Randomly-chosen ensemble members

Average daily precipitation

• 0-36-hr precipitation averaged over 1200 UTC initializations from May 14 to June 15



Average daily precipitation

• 0-36-hr precipitation averaged over 1200 UTC initializations from May 14 to June 15



Average daily precipitation

• 0-36-hr precipitation averaged over 1200 UTC initializations from May 14 to June 15



Total accumulation in verification region accumulated over experiment duration





Total accumulation in verification region accumulated over experiment duration





Average forecast 1-hour accumulations



Average forecast total 1-hour accumulations



Areal coverages

Aggregated over 36 cases (12 UTC forecasts)



Neighborhood approach

- Neighborhood Equitable Threat Score
 - Gives credit for "near misses" within a specified radius
- Fractions Skill Score
 - Similar approach as neighborhood ETS
- How does precipitation forecast skill vary...
 - with scale?
 - with diurnal cycle?
 - with initialization time?
 - with forecast lead time?

24-hour accumulations – 00 Z initializations compared to 12-Z initializations



14

24-hour accumulations – 24- to 48-hour forecasts compared to 12- to 36-hour forecasts



Dash: 00 Z initializations, hours 12-36 Dash: 12 Z initializations, hours 12-36





Threshold = 5.0 mm in 3 hours



12 Z initialization & 00 Z update





Threshold = 5.0 mm in 3 hours



00 Z initialization & 12 Z update

Ensemble members (dash) as compared to GFS initialization (solid) 24-hour accumulations (forecast hours 12 to 36)



00 Z Initializations

12 Z Initializations







12 Z Initializations

Attributes statistics

• Aggregated hourly over first 18-hrs (first diurnal peak) for 36 cases (12 UTC fcsts)



Attributes statistics

• Aggregated hourly between 24-42-hrs (diurnal peak #2) for 36 cases (12 UTC fcsts)



Area under the ROC curve

Aggregated over 36 cases (12 UTC forecasts)



Fractions skill score (FSS)

• Computed with a 50-km radius of influence and aggregated over 36 12 UTC fcsts



Preliminary conclusions

- Forecasts initialized later were better than earlier-initialized forecasts
- Forecasts initialized at 12Z dramatically overforecast the peak precipitation amounts of the first diurnal cycle (12-15 hours after initialization).
- The ensemble did not have enough spread of precipitation, but it had skill at discriminating between events
- The GFS-initialized forecasts were better than any individual ensemble member
- The benefit of the ensemble comes from the probabilistic guidance