Quantitative Precipitation Forecasting Component of the 2011 Hazardous Weather Testbed Spring Experiment

1. INTRODUCTION

The 2011 Spring Experiment was conducted at the National Weather Center in Norman, Oklahoma from May 9-June 10, 2011. This year's experiment featured components on severe weather, led by the Storm Prediction Center (SPC), convective initiation, led by the National Severe Storms Laboratory (NSSL), and quantitative precipitation forecasting (QPF), led by the Hydrometeorological Prediction Center (HPC). Over 80 forecasters, researchers, and academics participated during the five week experiment.

The overall goals of the QPF component were to document the strengths and limitations of the experimental model guidance for QPF and determine how to best use the experimental and operational data in a complementary manner. In particular, the QPF component aimed to answer 5 specific questions:
- Is the pending 4 km NMMB nest better than the operational 12 km NAM?
- Is the HRRR QPF as skillful as the QPF from the 4 km NMMB nest?
- Is the HRMOS an improvement over the SREF?
- Is the SSEO a feasible "poor-man's" approach to storm scale ensembles?
- Is the bias-corrected SSEF mean an improvement over the raw SSEF mean?

This report summarizes the QPF component activities, results, and operational impacts.

2. QPF COMPONENT DESCRIPTION

Data

The experimental datasets used in the QPF component of the Spring Experiment are summarized in Table 1. The deterministic high resolution guidance featured two Weather Research and Forecasting (WRF) models with both Advanced Research WRF (ARW) and Non-hydrostatic Mesoscale Model (NMM) dynamic cores, the High Resolution Rapid Refresh (HRRR), and the Non-hydrostatic Multiscale Model on a rotated Arakawa B-grid (NMMB). The WRF and HRRR are both convection-allowing. The NMMB is a nested model in which the 12 km parent uses the full BMJ convective parameterization scheme (Janjic 1994) while the 4 km nest uses a light version of the BMJ scheme prior to becoming convection-allowing. These models were contributed by NSSL, the Environmental Modeling Center (EMC), and the Global Systems Division (GSD) of NOAA's Earth System Research Laboratory (ESRL). Data from the operational 12 km North American Mesoscale model (NAM) was used for comparison to the high resolution deterministic models.

The experiment also featured two high resolution ensembles. The University of Oklahoma's Center for Analysis and Prediction of Storms (CAPS) provided a 4 km 50-member multi-model,

Table 1. Experimental high resolution models used during the QPF component of the 2011 HWT Spring Experiment. Each model was initialized at 00Z and had forecasts to 15 hrs (HRRR), 36 hrs (SSEF, SSEO, NSSL WRF-ARW), 48 hrs (NCEP HRW-NMM, NCEP HRW-ARW), 60 hrs (NMMB), or 156 hrs (HRMOS). 12Z initializations were also available for the SSEO, NMMB, NCEP HRW-NMM, NCEP HRW-ARW, and HRMOS. The HRRR is initialized hourly.

Provider	Model	Delta X	Notes	Label
CAPS	WRF/ARPS 50 member ensemble	4 km	Multi-model, multi-physics, multi-initial condition ensemble system with radar assimilation	SSEF
SPC	WRF/NMMB 7 member ensemble	4 km	Combination of available high resolution deterministic runs	SSEO
NCEP/EMC	NMMB	12 km and 4 km	Pre-implementation version of the NAM	NMMB
NSSL	WRF-ARW	4 km	NAM initial conditions NAM lateral boundary conditions	NSSL WRF-ARW
NCEP/EMC	WRF-NMM	4 km	NAM initial conditions NAM lateral boundary conditions	NCEP HRW-NMM
NCEP/EMC	WRF-ARW	5.1 km	NAM initial conditions NAM lateral boundary conditions	NCEP HRW-ARW
GSD	HRRR	3 km	Hourly updating with radar assimilation	HRRR
MDL	HRMOS	4 km	GFS-based statistical regression	HRMOS

multi-physics, and multi-initial condition Storm Scale Ensemble Forecast system (SSEF) that assimilated radar and other observational data. SPC provided a 7-member Storm Scale Ensemble of Opportunity (SSEO) consisting of readily available 4 km convection-allowing model runs which included two time lagged members (Table 2). Data from the operational 32 km Short Range Ensemble Forecast System (SREF) was used for comparison to the high resolution ensembles.

A variety of experimental ensemble products were available from both the SSEF and the SSEO. In addition to the ensemble mean, both high resolution ensembles provided point probabilities, neighborhood probabilities, and spaghetti plots. The neighborhood probabilities were calculated at a point using a smoothing parameter of sigma = 30. The SSEF ensemble output also included the probability matched mean, which combines the spatial pattern of the ensemble mean QPF with the frequency distribution of the rainfall rates from the individual ensemble members (Ebert 2001), a bias corrected mean, and the maximum from any ensemble

Table 2. SSEO membership.

Model	Ensemble Members	Forecast Hours
SSEO	00Z NSSL WRF-ARW 00Z SPCWRF4 00Z NCEP HRW-NMM 00Z NCEP HRW-ARW 00Z NCEP NMMB nest 12Z NCEP HRW-NMM (time lagged) 12Z NCEP HRW-ARW (time lagged)	36

member. All forecast products from the SSEF were calculated based on the 24 ensemble members with mixed initial conditions and physics perturbations.

In addition to these numerical models, the Meteorological Development Laboratory (MDL) also contributed a Global Forecast System (GFS)-based high resolution Model Output Statistics product (HRMOS, Charba and Samplatsky 2011). Data from this statistically based model include both probabilities and a derived continuous QPF field. HRMOS output was also compared to the operational SREF.

Daily Activities

Each morning, QPF component participants used 00Z model guidance to issue experimental probabilistic QPFs for the 18-00Z (18-24 hr forecast) and 00-06Z (24-30 hr forecast) periods over a selected forecast domain. These forecasts outlined areas that had a slight (25%), moderate (50%), or high (75%) probability of exceeding 0.50 in and 1.0 in of precipitation during the 6 hr period (Fig. 1). When a probability of exceeding the 1.0 in threshold was identified, participants were also asked to predict the most likely maximum areal average precipitation amount within that area to give an indication of the magnitude of the expected event. After completing the graphical forecasts, participants wrote a brief forecast discussion focused on the uncertainties in the model guidance and their forecast rationale.

A briefing with the Experimental Warning Program (EWP) about the morning forecast activities as held in the early afternoon. Participants then used both 00Z model guidance and any recently available 12Z guidance to update the forecast for the 00-06Z period and issue a new experimental forecast for the 06-12Z period.

In addition to issuing experimental forecasts, participants also subjectively evaluated both their forecasts from the previous day and the available model guidance. The subjective model evaluations consisted of a series of survey questions designed to evaluate whether the experimental high resolution models had provided better forecast guidance than the operational 12 km NAM (deterministic) or the operational SREF (ensemble). There were 23 days available for evaluation over the course of the 5 week experiment.
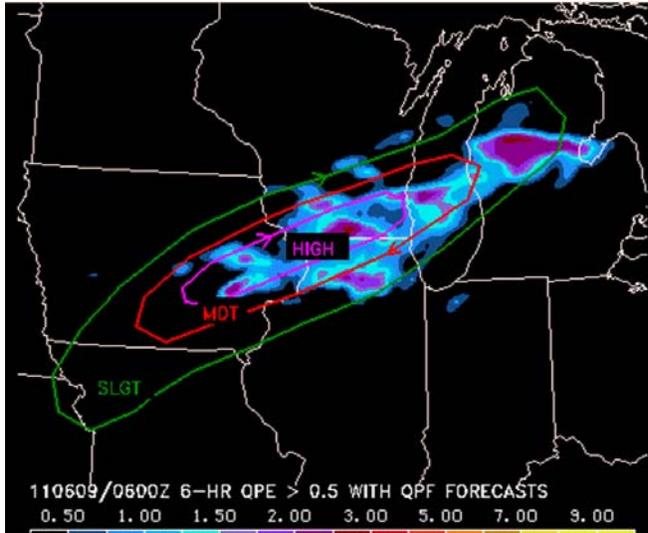
Figure 1. Experimental forecast (contours) from 8 June 2011 valid 06Z 9 June 2011 indicating the probability of exceeding 0.50 in of precipitation during the 00-06Z period. Shading indicates the observed 6 hr precipitation ≥ 0.50 in from NSSL Q2 QPE.

3. DETERMINISTIC HIGH RESOLUTION MODEL PERFORMANCE

As part of the subjective model evaluations, participants were asked to rate the high resolution deterministic guidance as much better, better, about the same, worse, or much worse than the operational 12 km NAM based on the observed precipitation. All results shown below are based on these subjective responses. The Development Testbed Center (DTC) may be providing additional objective verification results at a later date.

NCEP NMMB

The 4 km NMMB nest consistently provided better forecast guidance than the operational 12 km NAM and was the best performing deterministic model evaluated for QPF (Fig. 2). Participants found that the 4 km NMMB nest produced both reasonable precipitation amounts and realistic convective details. Although the results were somewhat mixed during the 18-00Z period, the 4 km NMMB nest showed significant improvement relative to the operational NAM during the later forecast periods.

In contrast, the 12 km NMMB provided worse guidance than the current operational NAM and overall was the worst performing deterministic model evaluated for QPF. Unlike the 4 km nest, the 12 km NMMB had a pronounced dry bias, often producing less precipitation than the operational NAM, which itself was too dry on many occasions. EMC is aware of the dry bias in the parent model and has identified a change to the convective parameterization that will improve the bias, but the change cannot be applied until after the scheduled model implementation in mid-September. Like the 4 km NMMB nest, the 12 km NMMB showed greater improvement at longer forecast lead times, with the 06-12Z forecast comparing more favorably with the operational 12 km NAM than the 18-00Z forecast.

2011 HWT Spring Experiment
High Resolution Model Performance Compared to the
NAM/SREF
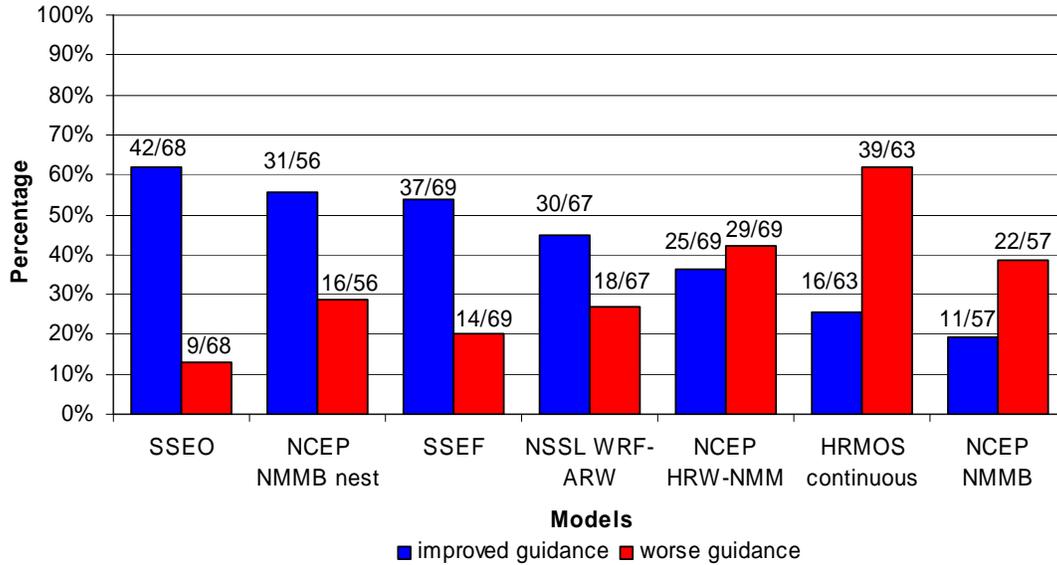
Figure 2. Experimental model performance across all three forecast periods based on participant feedback from subjective evaluation surveys conducted during the QPF component of the 2011 HWT Spring Experiment. Experimental ensembles were compared to the operational SREF mean while experimental deterministic models were compared to the 12 km operational NAM.

Figure 3 shows an example of the differences observed between the 12 km NMMB, 4 km NMMB nest, and the current operational NAM for forecasts valid 00Z 14 May 2011. Heavy precipitation was observed from central Missouri south and east through western Tennessee into northern Mississippi. The operational 12 km NAM indicated generally light precipitation across the area with isolated higher amounts mainly across western Tennessee. The 12 km NMMB was even lighter across the area of interest. The 4 km NMMB nest correctly indicated the potential for heavier precipitation extending from Missouri south into western Tennessee and northern Mississippi with amounts that were similar to the observations. The 4 km NMMB nest also correctly indicated the potential for heavier precipitation across southern Alabama.
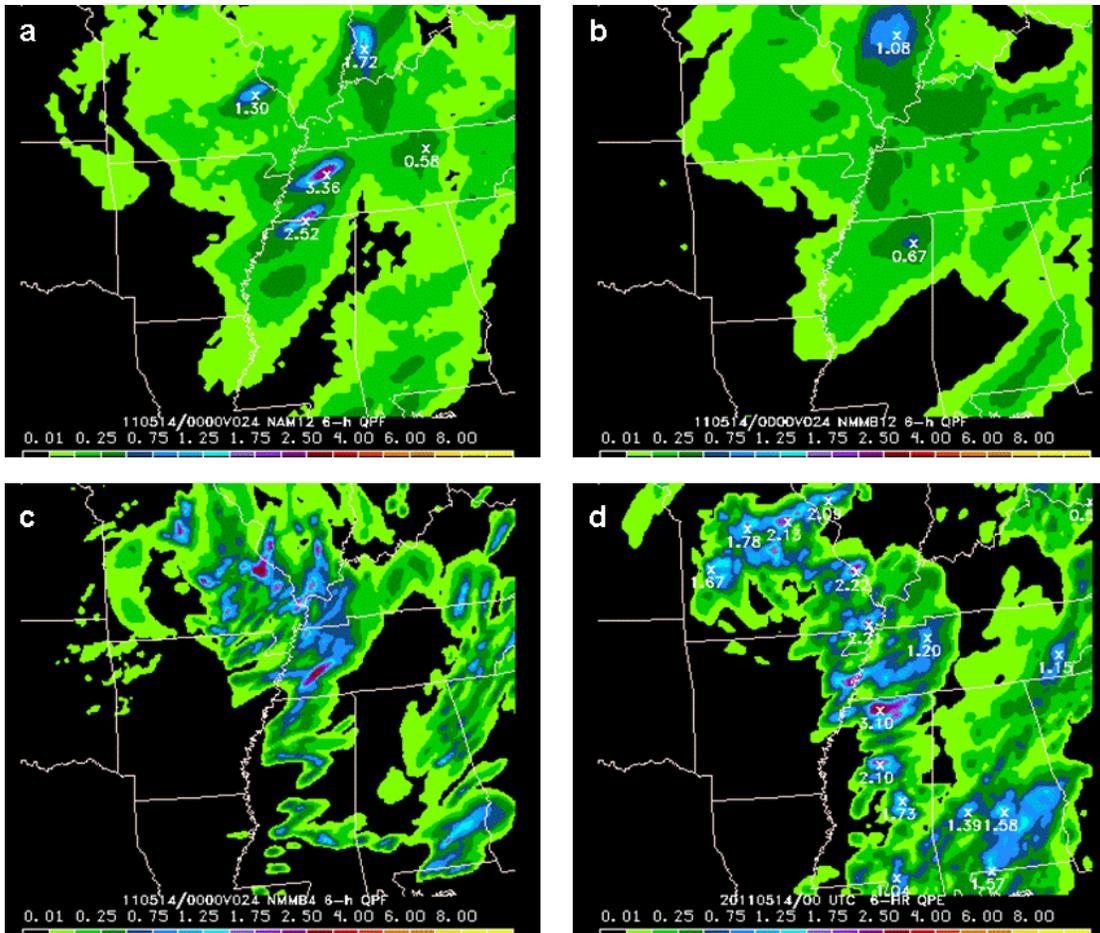
Figure 3. 6 hr precipitation forecasts valid 00Z 14 May 2011 from the (a) 12 km operational NAM, (b) 12 km NMMB, (c) 4 km NMMB nest, and (d) the corresponding 6 hr observations from NSSL Q2 QPE.

GSD HRRR

The QPF component also explored the utility of HRRR QPF guidance relative to the 4 km NMMB nest during the 18-00Z forecast period. The goal of this analysis was to determine if the HRRR provides additional value for short term forecasts relative to the upcoming operational guidance. The HRRR is an hourly updating convection-allowing model run with forecasts to 15 hours (Alexander et al. 2011). Because of the short forecast lead time, the model evaluations compared the 12Z HRRR to the 12Z 4 km NMMB nest. Figure 4 summarizes the performance of the HRRR and the other available high resolution models during the 18-00Z forecast period. Unfortunately, problems with the HRRR verification displays during the first week of the experiment and the 12Z NMMB runs during the last two weeks of the experiment limited the number of days available for evaluation to nine. Over this small sample, the HRRR typically provided worse guidance than the 4 km NMMB nest and was one of the poorer performing models overall.

2011 HWT Spring Experiment
High Resolution Model Performance Compared to the
NAM/SREF During the 18-00Z Forecast Period
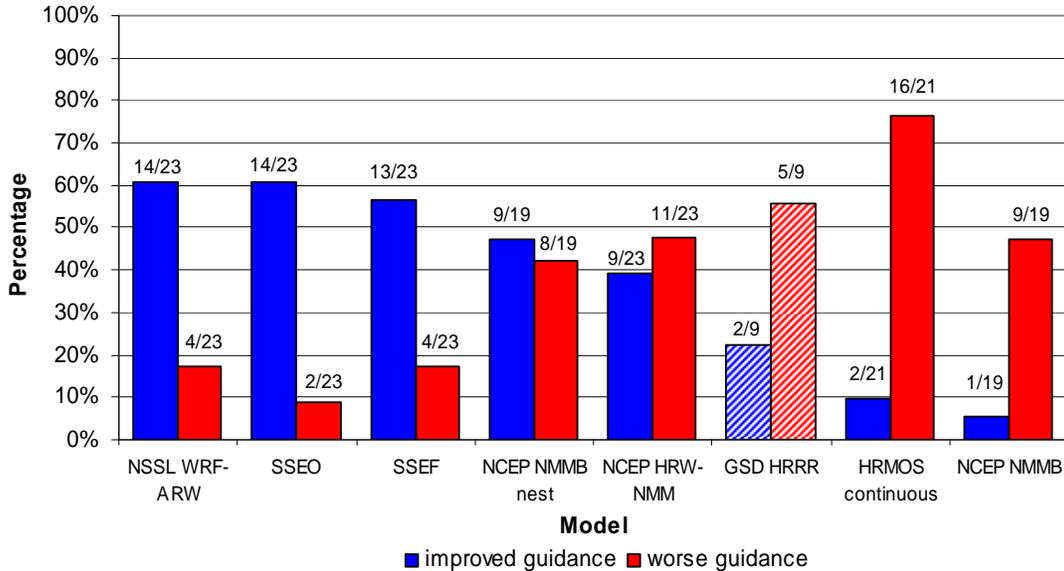
Figure 4. Experimental model performance during the 18-00Z period based on participant feedback from subjective model evaluation surveys conducted during the 2011 HWT Spring Experiment. The 12Z HRRR was compared to the 12Z 4 km NMMB nest while all other experimental models were initialized at 00Z and compared to the corresponding 00Z 12 km operational NAM (deterministic) or 21Z operational SREF (ensemble).

Since the feedback gathered during the experiment was quite limited, a supplemental evaluation survey was conducted in which the HRRR forecast during the 18-00Z period was rated on a scale of 1-5, where 1 is very poor and 5 is very good. This evaluation was available for 17 days. Of the available data, approximately 40% of the HRRR forecasts were rated as either "poor" or "very poor," with an equal percentage rated as "fair." On several occasions the precipitation amounts forecast by the HRRR were good, but the areal extent of those amounts was too limited. Figure 5 provides an example of this issue. Focusing on the observed area of heavier precipitation in northern Virginia, the HRRR underpredicts the areal extent of the heavier precipitation in this region relative to both the observations and the 4 km NMMB nest, although the forecast amount itself looks good. One of the biggest challenges to using the HRRR operationally appears to be the combination of the short forecast lead time and the rapidly updating nature of the model, as there can be significant differences between forecasts from consecutive model runs.
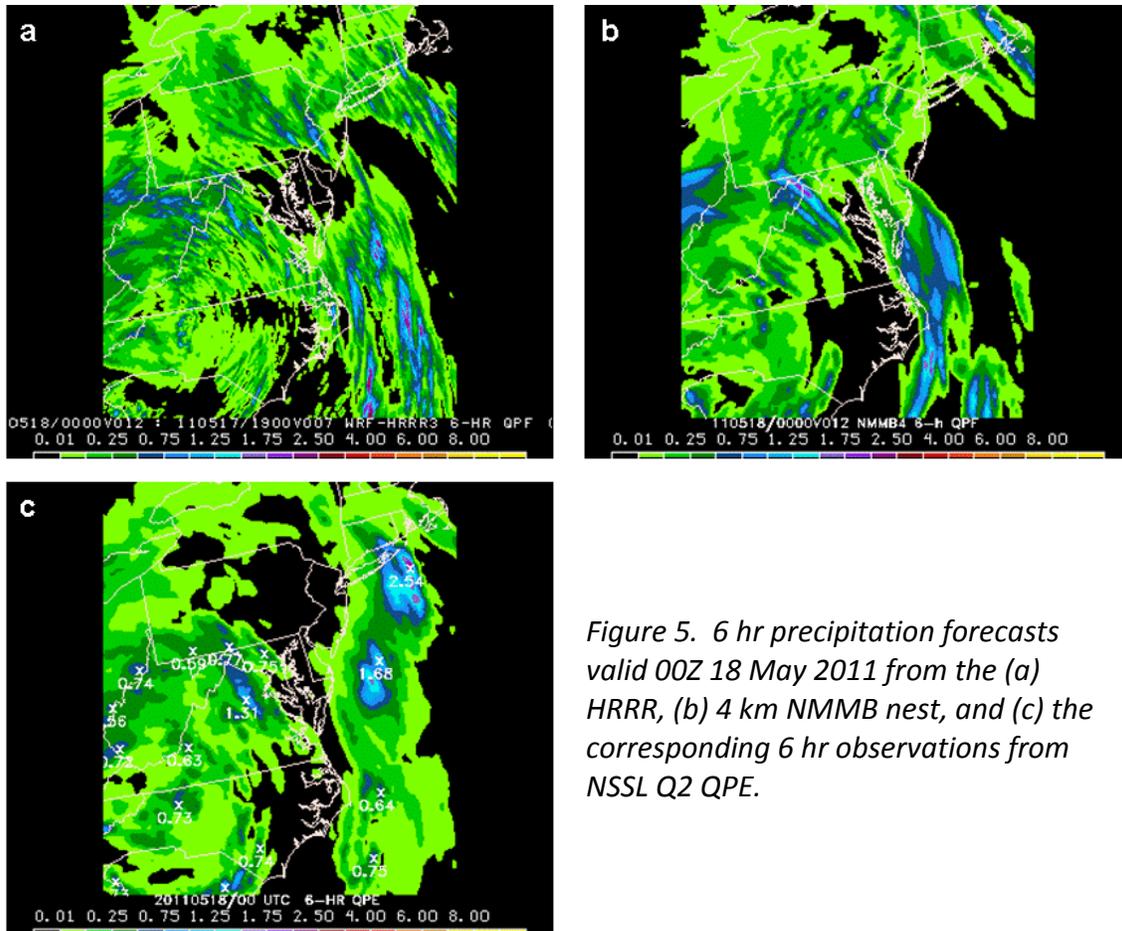
Figure 5. 6 hr precipitation forecasts valid 00Z 18 May 2011 from the (a) HRRR, (b) 4 km NMMB nest, and (c) the corresponding 6 hr observations from NSSL Q2 QPE.

Other Deterministic Guidance

In addition to the NMMB and HRRR, participants also evaluated the NSSL WRF-ARW and NCEP HRW-NMM. The NSSL WRF-ARW, which was the best deterministic high resolution model for QPF in the 2010 Spring Experiment, again provided forecast guidance that consistently improved on the guidance from the operational 12 km NAM (Fig. 2). Like the 4 km NMMB nest, the NSSL WRF-ARW provided useful information about both precipitation placement and amounts. Interestingly, while the NSSL WRF-ARW provided good forecast guidance overall, it tended to struggle during the 06-12Z period (36 hr forecast), receiving a higher percentage of "worse" and "much worse" ratings (41%) than "better" or "much better" ratings (32%). One possible explanation is that the solution is beginning to be influenced by the NAM lateral boundary conditions. The performance of the NSSL WRF-ARW compares favorably between 2010 and 2011 across the 18-00Z and 00-06Z periods common to both experiments (Fig. 6).

The NCEP HRW-NMM tended to provide worse guidance than the operational 12 km NAM more often than it provided improved guidance and was the poorest performer among the 4 km deterministic guidance (Fig. 2). The poor evaluation results may be due in part to an observed high bias in the precipitation amounts, which has been noted before in the NCEP HRW-NMM.

8

HWT Spring Experiment
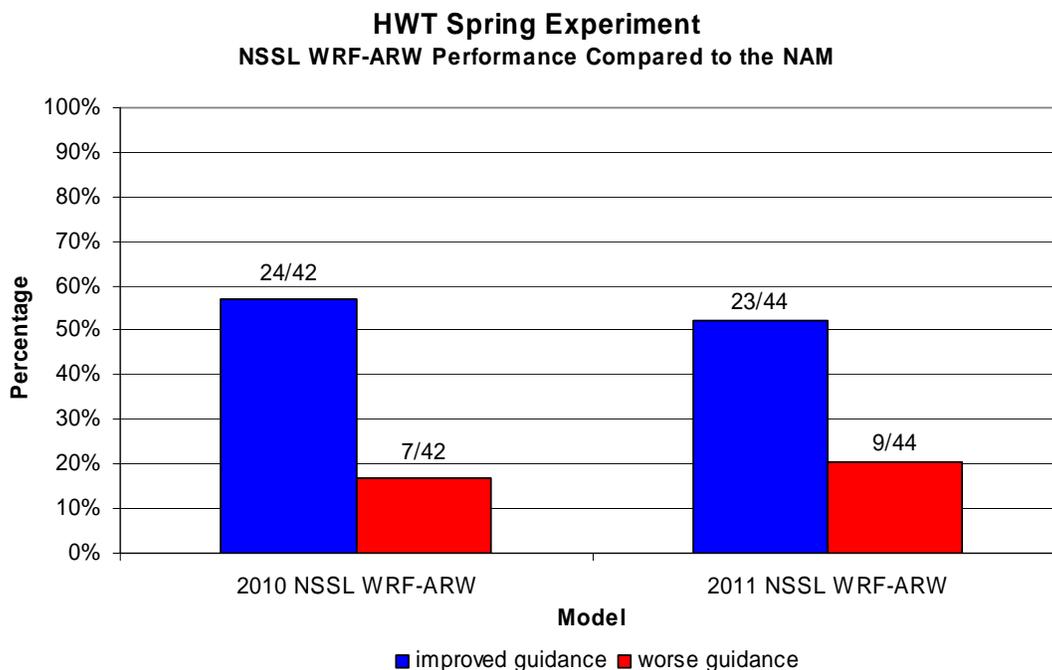**NSSL WRF-ARW Performance Compared to the NAM**

Figure 6. Comparison of the subjective evaluation results for the NSSL WRF-ARW between the 2010 and 2011 HWT Spring Experiments during the 18-00Z and 00-06Z periods.

4. HIGH RESOLUTION ENSEMBLE PERFORMANCE

Like the high resolution deterministic guidance, participants were asked to subjectively evaluate the ensemble means from the high resolution ensembles as much better, better, about the same, worse, or much worse than the ensemble mean from the operational 21Z SREF based on the observed precipitation. The results of this evaluation are summarized in Figure 2. High resolution ensemble data was among some of the most useful guidance for QPF in this year's experiment, highlighting the potential for this type guidance to significantly improve warm season QPF in the future.

SSEF and SSEO

a. Means
The SSEO mean consistently provided improved forecast guidance compared to the operational SREF and received the highest subjective evaluations of any model system tested for QPF during this year's Spring Experiment. In particular, participants noted that the SSEO mean typically provided useful information about the location of the axis of the heaviest precipitation. The SSEF mean also provided substantially improved forecast guidance relative to the operational SREF, although it struggled somewhat during the 06-12Z forecast period (36 hr forecast). The positive feedback about the SSEO mean in this year's experiment is particularly significant because the SSEO relies on currently available

9

high resolution guidance and can be run in real time at an operational center, whereas an operationally viable SSEF-like ensemble is likely many years away.

While the SSEF provided useful forecast guidance overall, participants were less impressed with its performance this year compared to the 2010 Spring Experiment. Compared to 2010, the relative improvement of the SSEF mean over the SREF mean was slightly worse for the two common forecast periods (18-00Z and 00-06Z), with an increase in the number of responses in the "worse" and "much worse" categories (not shown). Numerous changes were made to the SSEF between the 2010 and 2011 experiments, the most significant of those being an increase in the total number of ensemble members from 26 to 50 and an increase in the number of members contributing to the ensemble mean from 15 to 24. Whereas the 2010 version of the SSEF mean routinely provided forecast guidance that helped identify both the location and the amount of the heaviest precipitation, the 2011 version was often characterized by widespread lighter precipitation that made it difficult to use to identify the location of the heaviest precipitation.

The difference in participant feedback about this year's SSEF and the impressive performance of the SSEO leads to the question of how many members are required to produce a useful high resolution ensemble mean. To address this, a supplemental evaluation was conducted which asked whether the ensemble means from 5-member and 15-member subsets of the SSEF were much better, better, about the same, worse, or much worse than the full 24-member SSEF mean. The 15-member subset used in this analysis was configured identically to the 2010 SSEF mean. The results of this evaluation show that the small-membership ensembles typically provide forecast guidance that is at least as good as, if not better than, the full SSEF mean (Fig. 7). While some of this improvement is undoubtedly related to the higher mean precipitation amounts that result from using fewer ensemble members, this is still an important finding.

b. Probabilities
While the SSEF ensemble mean may have been degraded by the increased number of ensemble members, the additional members likely improved the probabilistic output. Figure 8 shows the results of a probability analysis conducted during the experiment in which participants were asked whether the observed 0.50 in precipitation area fell entirely within the 1% probability contour during each 6 hr forecast period. The goal of this analysis was to estimate the reliability of the probabilities on a daily basis. Based on this analysis, the SSEF probabilities were able to capture the observed precipitation almost three times as often as the probabilities from the SSEO. While the SSEO mean provided more improvement over the operational SREF than the SSEF mean, the SSEF provided more reliable probabilistic guidance than the SSEO.

2011 HWT Spring Experiment
Small-Membership SSEF Ensemble Performance Compared to the Full
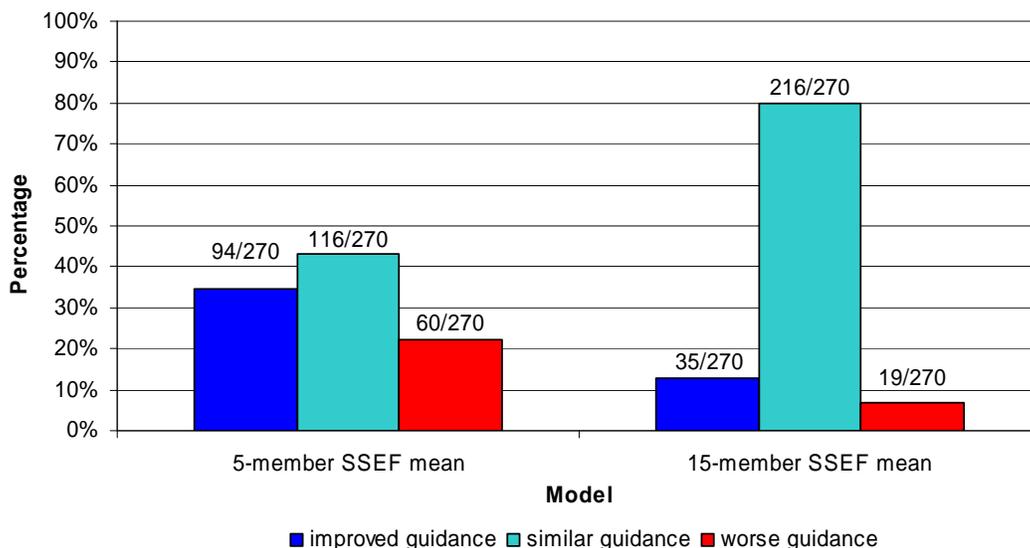**SSEF Ensemble**

Figure 7. Performance of 5-member and 15-member subsets of the SSEF ensemble mean relative to the full 24-member SSEF mean across all three forecast periods based on feedback from subjective evaluation surveys.

2011 HWT Spring Experiment
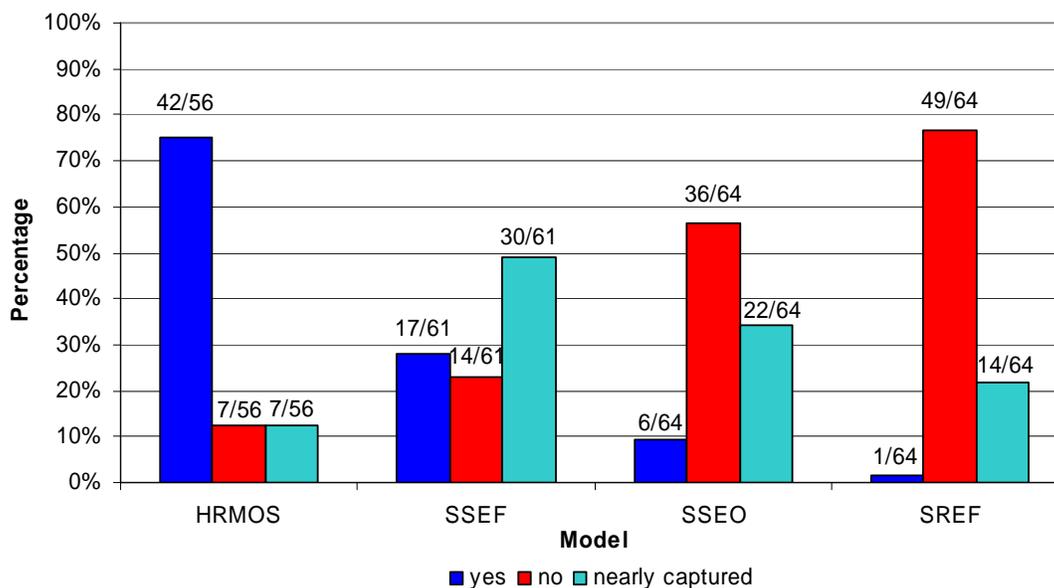**Ability to Capture 0.50"/6hr Rainfall Events**

Figure 8. Performance of the experimental probabilistic guidance in capturing 0.50 in/6 hr precipitation events using the respective 1% probability contours based on feedback from subjective evaluation surveys. The "nearly captured" category represents cases in which there were only very small observed areas of 0.50 in precipitation outside of the 1% probability contour.

11

HRMOS

The HRMOS continuous QPF product was found to provide worse forecast guidance than the operational SREF mean and was one of the poorer performing datasets overall for QPF (Fig. 2). Participants noted that in many cases, HRMOS produced large areas of generally light precipitation and seemed to struggle to identify the precipitation axis. Figure 9 shows an example of the HRMOS continuous QPF, ensemble means from the SSEF, SSEO, and SREF, and the observed precipitation valid 00Z 12 May 2011. While the observed precipitation was much heavier than any of the forecasts, the SSEF mean, SSEO mean, and to a lesser extent the SREF mean all provide a better indication of the potential for heavier precipitation in eastern Texas and Oklahoma than the HRMOS continuous QPF. In cases where the HRMOS continuous QPF did indicate heavy precipitation amounts, the areal coverage of those amounts was often too large. Figure 10 shows an example of 6 hr precipitation valid 12Z 10 May 2011 in which the area encompassed by the 0.50 in contour in the HRMOS continuous QPF product is much larger than either the SREF mean forecast or the corresponding observations. Additionally, the heaviest QPF in the HRMOS forecast actually lies in between the two observed areas of heavy precipitation. The combination of these two characteristics made it difficult to use the HRMOS to determine where to focus the highest precipitation amounts in the forecast.

Like the probabilities from the high resolution ensembles, HRMOS probabilities were evaluated by determining whether the observed 0.50 in precipitation area fell entirely within the 1% probability contour during each of the three 6 hr forecast periods. Compared to the available ensemble guidance, the HRMOS probabilities were most likely to encompass the range of possible solutions (Fig. 8). Despite being statistically reliable, participants thought that the probability values themselves tended to be quite low (typically less than 30%) and wondered what situation would be required to produce a high probability. Additionally, at times the probability values and the continuous QPF output did not appear to be consistent with each other. The HRMOS continuous QPF product for 6 hr precipitation valid 12Z 10 May 2011 indicated a broad 0.50 in area with a maximum of just over 2.0 in in western North Dakota. The corresponding probabilistic output for this time, however, indicates only a 10-30% probability of exceeding 0.50 in over this entire area (Fig. 10).
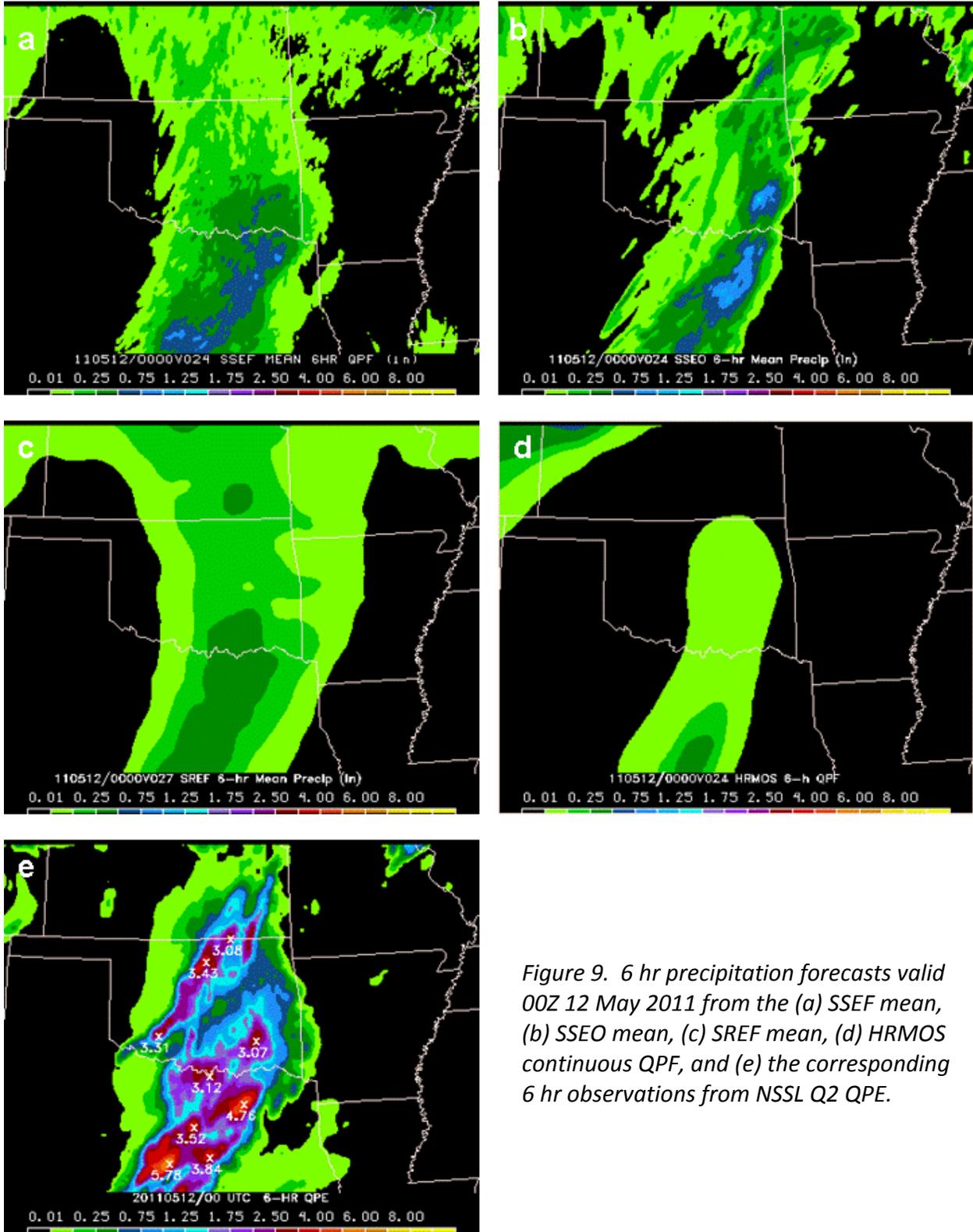
Figure 9. 6 hr precipitation forecasts valid 00Z 12 May 2011 from the (a) SSEF mean, (b) SSEO mean, (c) SREF mean, (d) HRMOS continuous QPF, and (e) the corresponding 6 hr observations from NSSL Q2 QPE.
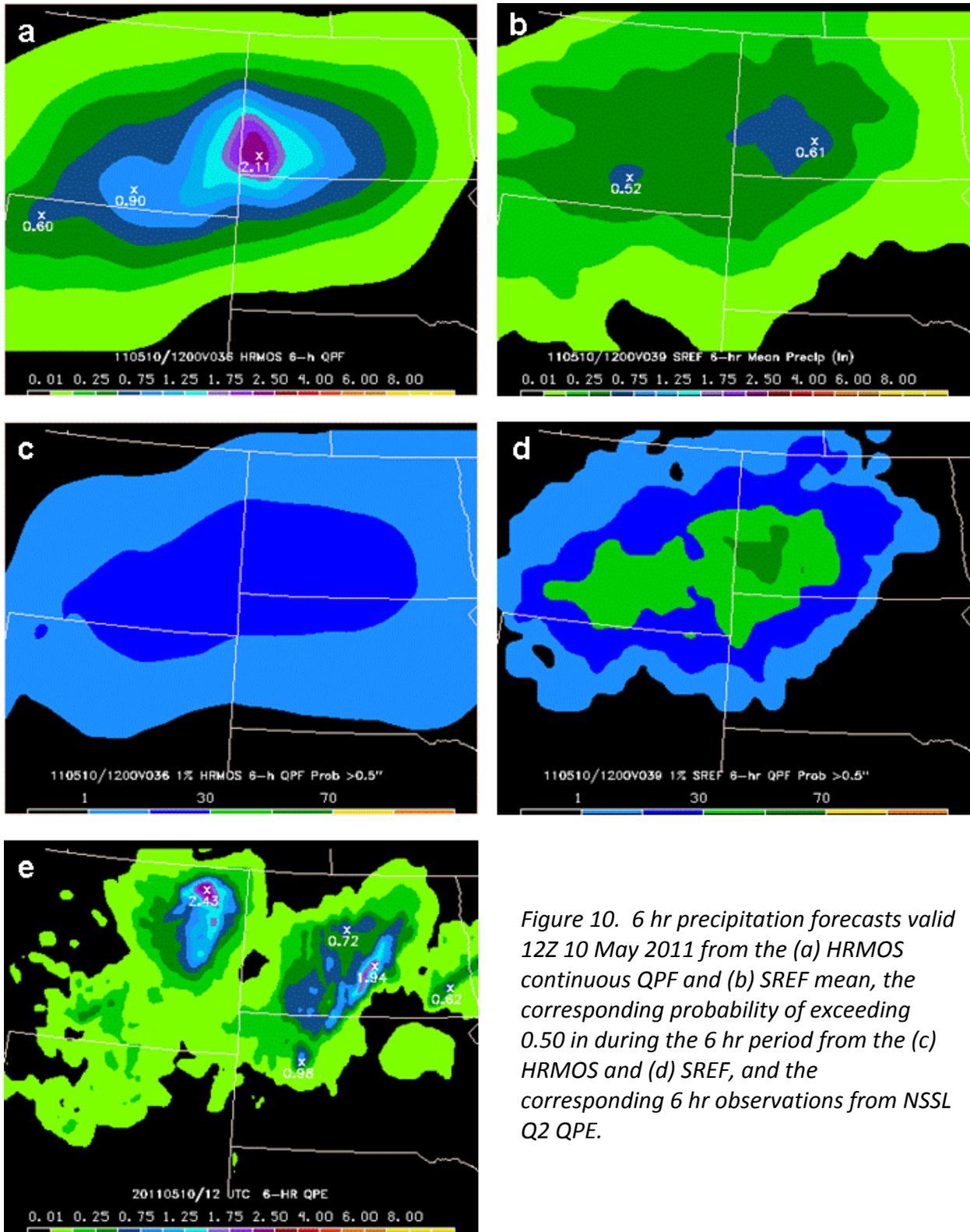
Figure 10. 6 hr precipitation forecasts valid 12Z 10 May 2011 from the (a) HRMOS continuous QPF and (b) SREF mean, the corresponding probability of exceeding 0.50 in during the 6 hr period from the (c) HRMOS and (d) SREF, and the corresponding 6 hr observations from NSSL Q2 QPE.

Post-Processing Techniques

In addition to the raw ensemble mean, the SSEF also featured bias-corrected and probability matched ensemble means. Unfortunately, the bias corrected data wasn't available until the last two weeks of the experiment, limiting the feedback to seven days. Based on this limited sample, the bias corrected guidance was often found to be quite similar to the standard ensemble mean, and did not provide substantial improvement. The probability matched mean was available throughout the experiment and provided better guidance than the standard SSEF mean about as often as it provided worse guidance (Fig. 11).

2011 HWT Spring Experiment
SSEF Post-Processing Technique Performance Compared to the Standard SSEF
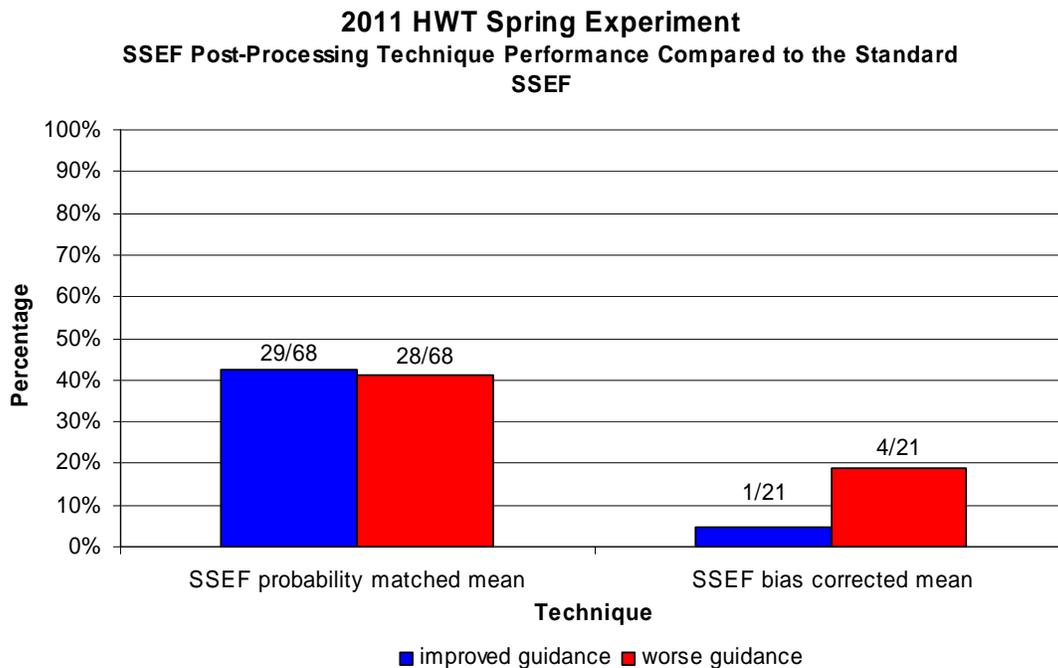
Figure 11. Performance of the probability matched and bias corrected SSEF means compared to the standard SSEF mean across all three forecast periods based on feedback from subjective evaluation surveys conducted during the 2011 HWT Spring Experiment.

Finally, participants were asked whether the available post-processing techniques were ready to be implemented operationally. While participants thought that point probabilities from both the SSEF and SSEO provided valuable operational guidance, probabilities from the HRMOS were not considered ready for operationally use (Fig. 12). Although participants didn't find neighborhood probabilities from either the SSEF or SSEO particularly useful for QPF, they have been found to provide useful forecast guidance for severe weather parameters, and different ways of calculating these fields such that they are relevant to precipitation forecasting should be explored.

Spaghetti plots from both the SSEF and SSEO were found to be operationally useful, with those from the SSEO being particularly valuable since they can be used to provide a quick summary of all of the available high resolution guidance. Participants provided a mixed assessment of the

potential use of the ensemble maximum field in operations. The ensemble maximum provides the maximum precipitation from any ensemble member at each grid point. This often results in precipitation amounts that are unrealistically high, although in some cases this information might be useful for getting a sense of the worst case scenario.
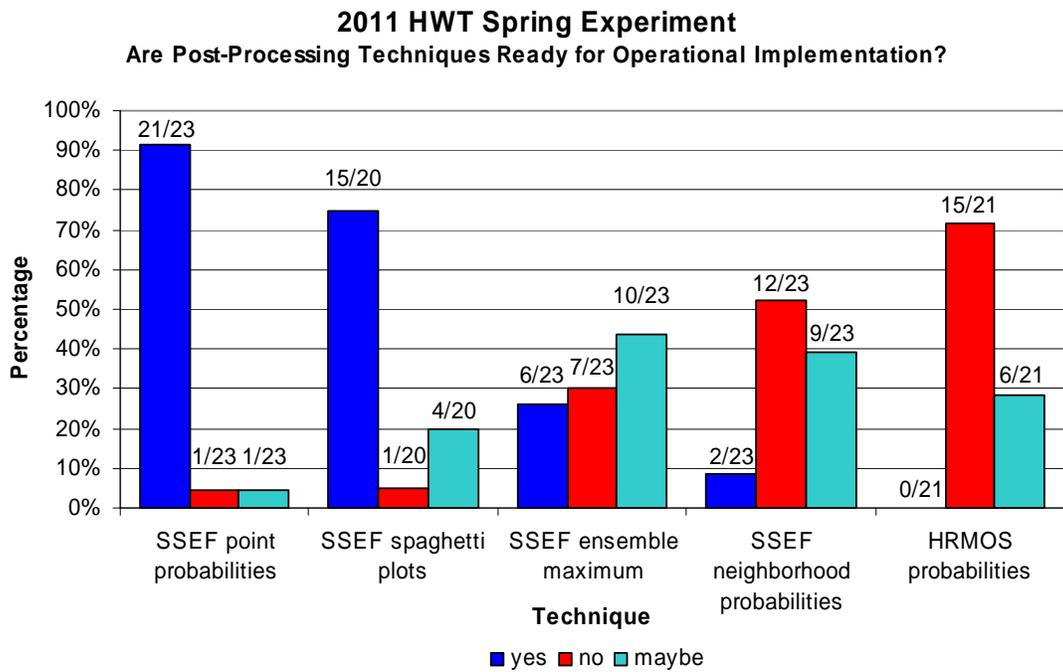
2011 HWT Spring Experiment
Are Post-Processing Techniques Ready for Operational Implementation?

Figure 12. Evaluation of post-processing techniques.

5. LIMITATIONS OF RESULTS

Although the experiment results represent an incredible collaborative effort focused on warm season QPF, they are subject to important limitations. Namely, the results are based on subjective survey responses from a variety of individuals. The Development Testbed Center (DTC) may be providing additional objective verification results at a later date. Analysis of the 2010 Spring Experiment results revealed that the subjective survey results were generally supported by objective verification. Another key limitation is that the sample size is relative small (23 days) and may not be fully representative. Finally, the focus of the experiment was on warm-season QPF. Thus the relative performance of the respective datasets may not necessarily be similar for other phenomenon or seasons.

6. SUMMARY AND OPERATIONAL IMPACTS

The QPF component of the HWT Spring Experiment was conducted May 9-June 10, 2011 with the goal of exploring the use of high resolution convection-allowing model guidance for precipitation forecasting. Each day during the experiment, participants issued experimental probabilistic QPFs for the 0.50 in and 1.0 in precipitation thresholds for the 18-00Z, 00-06Z, and

16

06-12Z forecast periods. In addition to the experimental forecasting activities, participants subjectively evaluated both their forecasts and the available experimental model guidance.

The subjective model evaluations revealed that high resolution guidance is useful and can improve warm season QPFs. While some of the data used during the Spring Experiment was purely experimental, much of the data is currently available in HPC's operational workstations (Table 3). Highlights from this year's experiment included:

- The SSEO mean provided valuable information about the axis of the heaviest precipitation. The performance of the SSEO is particularly significant because it shows that a high resolution "poor-man's" ensemble can provide useful forecast guidance. **Based on the positive experiment results, the SSEO is now available to HPC forecasters.**
- The 4 km NMMB nest provided both realistic precipitation amounts and convective details. **Based on the experiment results forecasters now have confidence in using the NMMB nest as it becomes operational with the NAM upgrade (currently scheduled for mid-September).**
- Spaghetti plots were found to be a useful way to display high resolution ensemble data and can be used to provide a quick summary of the available high resolution guidance. **Based on the experiment results DTB/HMT-HPC is investigating making spaghetti plots available from the SSEO.**

Table 3. Summary of high resolution guidance available to HPC forecasters in NMAP.

Model	NMAP Location	Available Fields
SSEO	sseo_mean	QPF
NCEP NMMB	namp	Same as operational NAM
NCEP NMMB nest	namp_conest	QPF, standard fields (500mb heights, etc.)
NSSL WRF-ARW	wrf4nssl	QPF, standard fields (500mb heights, etc.)
NCEP HRW-NMM	hrw_east_nmm	QPF, standard fields (500mb heights, etc.)
NCEP HRW-ARW	hrw_east_arw	QPF, standard fields (500mb heights, etc.)
HRMOS	gfsmos_qpf	QPF, exceedance probabilities

The Hazardous Weather Testbed provides a unique collaborative environment that brings the research and operational forecasting communities together to explore new forecasting techniques and evaluate emerging model guidance. The experimental forecast and model evaluation process has helped raise awareness of the challenges associated with warm season QPF and fostered discussion about the strengths and limitations of high resolution model data. In addition, the experimental forecasting and evaluation activities have influenced model development by identifying guidance that adds significant value compared to the current operational models.

ACKNOWLEDGEMENTS

REFERENCES

Alexander, C. R., S. S. Weygandt, S. G. Benjamin, T. G. Smirnova, J. M. Brown, P. Hofmann, and E. P. James, 2011: The High Resolution Rapid Refresh (HRRR): Recent and future enhancements, time-lagged ensembling, and 2010 forecast evaluation activities. Preprints, *24th Conference on Wea. and Forecasting/20th Conference on NWP*, Amer. Meteor. Soc., Seattle, WA, 12B.2.

Charba, J. P., and F. G. Samplatsky, 2011: High-resolution GFS-based MOS quantitative precipitation forecasts on a 4-km grid. *Mon. Wea. Rev.*, **139**, 39-68.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480.

Janjic, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev*., **122**, 927-945.