FINAL REPORT FOR UCAR COMET NWS PARTNERS PROJECT

Peirong Lin and Zong-Liang Yang (PI), University of Texas at Austin (UT-Austin)
Larry Hopper and Jon Zeitler (Co-PIs), NOAA/NWS Austin-San Antonio Forecast Office (EWX)

WRF-Hydro-RAPID Performance in Predicting Recent Texas Hill Country Flash Floods

UCAR Award No.: No. Z16-23463
Date: 31 August 2017

---

1. Summary of Project Objectives

The hydrological prediction capability of the United States has been enhanced with the recent implementation of the National Oceanographic and Atmospheric Administration's (NOAA) high resolution National Water Model (NWM) that became operational in August 2016. The NWM essentially runs the 1-km WRF-Hydro modeling framework with the NHDPlus hydrologic data frame, to predict streamflow for 2.67 million river reaches. In order to help understand and calibrate the model's ability to predict hydrological conditions, one important first step is to evaluate this large-scale framework with historical flood events, particularly in areas with complex terrain and extreme rain rates. In addition, the abilities and limitations of the model configuration remain unclear, especially at local scales, which is of great concern for the National Weather Service (NWS) forecasters and emergency management officials.

In order to improve operational hydrological forecasting, this COMET NWS Partners Project used a similar framework (WRF-Hydro-RAPID) to perform the following objectives:

- Evaluate model hindcast skill in several recent major floods over the Texas Hill Country, one of the most flood-prone regions of the nation with challenging hydrological features
- Assess the uncertainties of two seamless radar-based quantitative precipitation estimation (QPE) products [i.e. National Stage IV QPE (4-km, 1-hourly) and Multi-Radar/Multi-Sensor (MRMS, 1-km, 2-min)], and quantify their influence on the streamflow prediction skill
- Understand current model limitations and seek potential solutions to improve the model's predictive skill for streamflow and stage height.

In the following section, we summarize our key findings on these three aspects. Some of these materials are being incorporated into a manuscript we are preparing for submission to *Journal of Hydrometeorology* within the next few months.

2. Project Accomplishments and Findings

2.1 Assessment of uncertainties with the Stage IV and MRMS QPE products

We obtained all of the available gauged precipitation data within our study domain (Fig. 1, both 24-h accumulated and hourly rainfall) to assess the Stage IV (ST4) and MRMS QPE uncertainty. We focused on the May 23-24, 2015 and subsequent May 25, 2015 flash floods that caused numerous fatalities. In addition, we also studied a heavy rainfall event associated with the remnants of Hurricane Patricia on October 24-25, 2015 that was followed by a more extreme rainfall event on October 30, 2015. Hereafter, the May events are combined together, whereas the two October events are referred to as the Oct-I and Oct-II events, respectively, for simplicity.

Figure 2 shows that all QPE products better captured the 24-h accumulated precipitation for May event (first row) than the October events (second and third row). Overall, ***MRMS QPE tends to have lower 24-h accumulated precipitation than the ST4 QPE*** for all three events, with the May event having the best QPEs relative to that observed based on its CC closest to 1 and its lowest RMSE. ***For the May event, the ST4 QPE has the best skill whereas the MRMS gauge-corrected (Q3GC) and radar-only products (Q3RAD) slightly underestimate the rainfall*** (i.e. bias ratio ~0.9). For Oct-I, the QPE behaviors are similar as in May, albeit with slightly lower CC and higher RMSE. However, for the much more extreme Oct-II event, all QPE products tend to overestimate the rainfall for several gauges with rainfall less than 50 mm day^{-1}.

Maximum precipitation intensity (MPI) estimation by the QPE is shown in Fig. 3, an important factor for the peak discharge prediction. ***Q3RAD massively underestimated the MPI for the May event*** shown in Fig. 3c, which may be related to instantaneous (i.e., 2 min) MPI values being truncated at ~46 mm hr^{-1} (~1.9 in hr^{-1}) for non-tropical stratiform rain echoes noted by Zhang et al. (2016). Q3GC and ST4 both underestimate (overestimate) rainfall above (below) ~33 mm hr^{-1} (~1.3 in hr^{-1}), but in general these sources are much closer to observed than Q3RAD. Therefore***, incorrect application of the convective-stratiform classification on 25 May 2015 likely led to the underestimated precipitation for the MRMS QPEs.*** This finding points to the critical importance that NWS forecasters carefully scrutinize QPE uncertainties when diagnosing potential errors in river forecasts when MRMS QPE forcing is used.

2.2 Model performance for the 23-25 May 2015 flash floods over the Texas Hill Country

These QPE products were processed to the same resolution as the model runs (1-km) and used together with bi-linearly-interpolated NLDAS-2 forcing to drive the WRF-Hydro-RAPID model. Fig. 4 shows the spatial distribution of the Nash-Sutcliffe Efficiency (NSE) for these experiments. Although calibration tests were not performed, model parameters for these experiments were carefully chosen based on the best-available knowledge on the land surface physical quantities, and remained constant through all experiments. Figure 4 shows that the ***lowest NSEs are observed over flatter areas with much slower rates of river stage rise (or locations without much rainfall and/or river rises), whereas some of the higher NSEs are observed over regions with greater terrain and more extreme rain rates.***

In Fig. 5, the NSE statistics for the May event are summarized in boxplot for all gauges (Fig. 5a), forecasting vs. non-forecasting points (Figs. 5b & 5c, separated based on whether or not a point is a river forecasting location used by NOAA/NWS), and reference vs. non-reference gauges (Figs. 5d & 5e, separated based on the degree of human disturbance). This type of analysis helps us better understand model performance from different perspectives. Overall***, the***

range of NSEs for ST4 is almost always larger than Q3GC and Q3RAD, suggesting MRMS leads to more stable statistics relative to ST4 estimates regardless of their accuracy. ST4 QPE ingests Hydrometeorological Analysis and Support (HAS) forecaster experience from the River Forecast Center (RFC), whereas MRMS does not (Zhang et al. 2016). Although the median of the Q3GC and Q3RAD experiments is slightly lower (not significant) than that of ST4 due to the slightly underestimated 24-h rainfall for the May event (Fig. 2), we see the automatic algorithmic feature of the MRMS products results in fewer stations with lower NSEs (smaller NSE range). This possibly implies that ST4 could display both good and bad performance depending on how much the human forecaster improves or degrades the analysis. Fig. 5c shows that the ***model exhibits reasonable simulation skills for the non-forecast points,*** suggesting the potential extensibility of this framework to locations without existing forecasting information by NOAA. We also found that the ***simulation skill is better for reference gauges (i.e., gauges with least-disturbed watersheds) than non-reference gauges***, indicating that the model performance may still suffer from inaccurately representing human interference (e.g., river diversions) despite eliminating gauges downstream of major reservoirs in our analysis.

Fig. 6 clearly shows ***experiments with the MRMS QPE forcing (Q3RAD and Q3GC) outperform ST4 at gauge locations with smaller drainage areas, especially for those with drainage areas less than 150 km^2*** as shown in Figs. 6a and 6c. Due to the finer spatiotemporal resolution of the MRMS QPE, Q3GC, and Q3RAD are better able to capture the localized fine-scale rain storms that contribute to these gauges than ST4. Thus, the advantages of the MRMS QPE are mainly manifested for gauges with smaller drainage areas that tend to experience rapidly rising flash floods. ***Q3GC and Q3RAD are no better than ST4 in this particular event for gauges with drainage areas larger than 800 km^2.*** These results generally indicate that in future regional- to-continental-scale flood prediction studies, MRMS may be a favored QPE product, because it does not require human quality control, it exhibits more stable simulation skills at smaller range, and it has higher spatiotemporal resolution. However, our results might be event-specific and more case study simulations using the MRMS QPE should be conducted to help inform its future algorithm development (Qi et al., 2016; Zhang et al., 2016).

Fig. 7 shows the flood hydrograph at a few selected gauges. The model's predictive skill appears to be reasonably good at these gauges of high strategic concern and impact considering the model is not calibrated. The ***timing of the record flood peak along the Blanco River at Wimberley (WMBT2) on the morning of 24 May 2015 is very well predicted with all three QPE datasets, but the magnitude of the forecast peak is still too low.*** Potential model errors contributing to this under-prediction may include the model not accounting for possible spring contributions from upstream, the soil saturation degree of upstream drainage areas being under-predicted, and the river flow velocity being under-predicted. Determining which factor dominates the magnitude of error in the under-predicted flow peak is difficult because no direct observations for these three factors exist. However, the fact that the model predicted the peak flow timing well shows great promise for the general applicability of the modeling framework.

Finally, we also sought to understand the limitations that the current NWM has by testing potential model deficiencies for future development. The current modeling framework uses the trapezoidal river geometry assumption for all 2.67 million river reaches, with several categories of detailed characterization determined based on the stream order. These hydraulic properties are not only key for accurate inundation mapping, but are also for hydrologic routing performances.

In our research, we experimented with using the actual river geometry extracted from 10-m DEMs developed by Zheng *et al.* (2017, *in revision*) and Liu et al. (2017, *in revision*). A preliminary forecast flood hydrograph at the Blanco River at Wimberley (WMBT2; Fig. 8a) and Shoal Creek at West 12th Street in Austin (AHOT2; Fig. 8b) shows that the peak streamflow forecast is increased and occurs faster than the standard model runs made in this study. The peak at WMBT2 increases from 2400 m^3/s (Fig. 7) to 3600 m^3/s (Fig. 8a) closer to the official crest of 4955 m^3/s, resulting in peak stage height rising over 2 m to 11.6 m closer to its official peak stage height of 13.7 m despite the peak being too early. AHOT2's peak streamflow increase from 360 m^3/s (Fig. 7) to 420 m^3/s (Fig. 8b) actually degrades the forecast relative to its observed crest at 312 m^3/s. However, practical impact to the forecast is low given the crest is only about 0.5 m higher than its 6.3 m observed crest and because forecast skill still increased due to capturing the flood wave's earlier timing. These results suggest that future work using more accurate river geometries may help be beneficial in helping advance existing model capabilities

3. Benefits and Lessons Learned: Operational Partner Perspective

Findings from this project have allowed forecasters at the Austin/San Antonio WFO to better understand the issues that make hydrologic prediction during heavy rainfall events over our region so challenging. The errors associated with ST4 and MRMS QPE's differing spatial and temporal resolutions and processing methods are critical due to the role QPE plays with flash flood guidance (FFG) in providing our scientific basis for issuing flash flood warnings in addition to surface hydrology and socioeconomic considerations. Working with UT-Austin has increased our awareness of how differences in river channel geometry, antecedent soil moisture conditions, and subsurface groundwater contributions affect streamflow prediction. Finally, seeing how the Muskingum-Cunge routing method with variable routing parameters improved hydrological prediction along the Escarpment relative to Muskingum methods helped our office better understand why the SAC-SMA model struggles along the Escarpment in heavy rainfall events. Future proposals investigating whether the Muskingum method performs better for near steady-state moderate to heavy rainfall events over flatter terrain as observed in WRF-Hydro-RAPID simulations for the October 2015 Hurricane Patricia event and SAC-SMA operational forecasts generated during Hurricane Harvey may be beneficial.

This project has also improved forecaster awareness of advances in hydrologic prediction that may be utilized operationally. Heavy rainfall distributions (Figs. 9a and 9b) have been used in forecasting QPF on several occasions before and during heavy rain events, including for the City of Austin's Flood Early Warning System (FEWS) group to run their contingency forecasts a few hours ahead of moderate flooding along Onion Creek in southeast Austin on 27 May 2016. Co-PI Hopper and another lead forecaster plan to submit a publication on this work and educate forecasters on how to formulate their QPF expectations based on these distributions during flash flood events. Several forecasters and two summer NOAA Hollings Scholars are also working on synergistic operationally-based projects including flood extent mapping, archiving streamflow and river stage crest times, developing verification indices for impact-based seasonal flash flood forecasts, and ingesting FLASH data into real-time operations. Forecasters are also more aware of the NWM's potential strengths and limitations. Finally, this project's partial funding of the inaugural Texas Weather Conference in February 2016 at UT-Austin has allowed several Texas NWS offices to network with academic, broadcast, and emergency management partners across the entire to share cutting-edge research and operational experiences to minimize the detrimental impacts of weather and climate on society. We held a second conference at UT-Austin in March 2017 with partial support from the UT-Austin Jackson School and have a third planned for fall 2018 in the Dallas-Fort Worth Metroplex.

4. Benefits and Lessons Learned: University Partner Perspective

We, as university partners (Prof. Zong-Liang Yang and graduate student Lin), benefitted from working with the WFO scientists to better understand the local storm events, rainfall rates, and streamflow gauge sites. Specifically, we have had an eye-opening experience by visiting the WFO office and it's helpful to gain the first-hand knowledge of the NWS operational capacities. Such an experience is particularly useful to us in terms of better understanding the operational needs while evaluating and developing the model capabilities. In addition, through this project, we have extended our collaborations with Dr. David Maidment's research group at UT-Austin, who seeks to improve the hydraulic representations of river channel elements used by the NWM. We have done some preliminary investigations on using their datasets to improve the model, which turns out to be quite promising (the last part in Section 2.2).

5. Publications and Presentations

- Hopper, L. J., Jr. and N. L. Hampshire, 2016: "Rainfall distributions for varying environmental conditions along the Balcones Escarpment, Texas." *30th Conf. on Hydrology,* New Orleans, LA, Amer. Meteor. Soc., 1.3.
    - [https://ams.confex.com/ams/96Annual/webprogram/Paper281700.html]
- Lin, P., Z.-L. Yang, L.J. Hopper, Jr., and J.W. Zeitler, 2016: "WRF-Hydro-RAPID performance in predicting recent Texas Hill Country flash floods." *1st Texas Weather Conference,* Austin, TX.
    - [https://drive.google.com/file/d/0B1_TYLU2DdMFR2VobkJrUmxMZVk/view]
- Lin, P., Z.-L. Yang, L.J. Hopper, Jr., and J.W. Zeitler, 2017: "WRF-Hydro-RAPID performance in predicting recent Texas Hill Country flash floods." 3*1st Conf. on Hydrology,* Seattle, WA, Amer. Meteor. Soc., P474.
- Lin, P., L.J. Hopper, Jr., Z.-L. Yang, M. Lenz, and J.W. Zeitler: Quantifying local and regional uncertainties with radar-based quantitative rainfall estimation (QPE) in predicting recent Texas Hill Country floods. (in preparation for *J. Hydromet.)*
- Hopper, L. J., Jr., and N. L. Hampshire: "Observed heavy rainfall distributions for varying environmental conditions along the Balcones Escarpment, Texas. (*in preparation for Wea. Forecasting or J. Operational* Meteorology)

6. Summary of University/Operational Partner Interactions and Roles

Peirong Lin performed the majority of the work for this project with constant guidance and feedback from her Ph.D. advisor, Dr. Liang Yang, and NWS co-PIs Larry Hopper and Jon Zeitler at the Austin/San Antonio NWS WFO. Hopper developed rainfall distributions for over 58 flash flood events from 2009-2015 to help select the cases utilized in this study and place these cases within context of other flood events in the area. Lin and Yang performed offline

WRF-Hydro-RAPID simulations of cases selected by co-PIs Hopper and Zeitler using ST4 and MRMS gauge-corrected and radar-only QPE products. Several face-to-face meetings were held between all four collaborators, with WFO Austin/San Antonio's Service Hydrologist (Mark Lenz) joining some of them. Data interpretation and analysis was primarily coordinated between Lin and Hopper with advisement from Yang, Zeitler, and Lenz. Finally, the inaugural Texas Weather Conference was co-chaired by Hopper and Zeitler with Yang serving on the committee. Yang provided partial support to the second Texas Weather Conference using his Jackson School funds.

7. References

Liu, Y., D.R. Maidment, D. Tarboton, X. Zheng, and S. Wang, 2017. A CyberGIS integration and computational framework for high-resolution continental-scale flood inundation mapping. *J. Am. Water Res. Assoc.*, in revision.

Qi, Y., S. Martinaitis, J. Zhang, and S. Cocks, 2016: A Real-Time Automated Quality Control of Hourly Rain Gauge Data Based on Multiple Sensors in MRMS System. *J. Hydrometeor.*, DOI: http://dx.doi.org/10.1175/JHM-D-15-0188.1.

Zhang, J., K. Howard, C. Langston, B. Kaney, Y. Qi, L. Tang, H. Grams, Y. Wang, S. Cocks, S. Martinaitis, A. Arthur, K. Cooper, J. Brogden, and D. Kitzmiller, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, 97, 621–638, doi:10.1175/BAMS-D-14-00174.1.

Zheng, X., D. Tarboton, D.R. Maidment, Y. Liu, and P. Passalacqua, 2017: River channel geometry and rating curve estimation using the Height Above Nearest Drainage. *J. Am. Water Res. Assoc.*, in revision.

APPENDIX: FIGURES

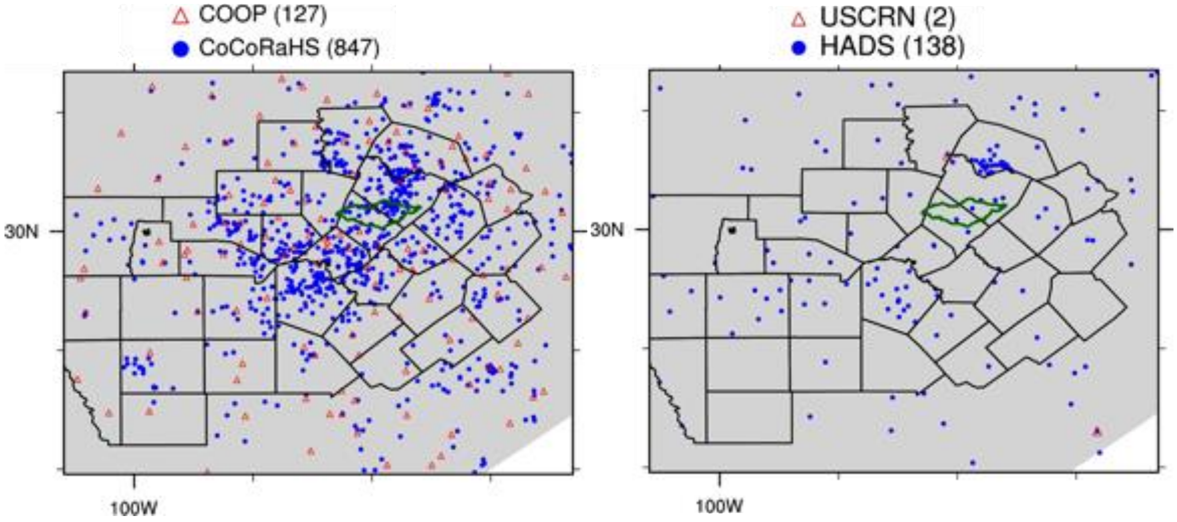


Figure 1. Precipitation gauge network within our study domain used in assessing radar-based QPE uncertainty. **(left)** COOP and CoCoRaHS gauges that report 24-hour accumulated rainfall **(mm)** on a daily basis; (right) USCRN and HADS gauges that report hourly rainfall intensity (mm hour^{-1}). The numbers in the parenthesis show the number of available gauges.

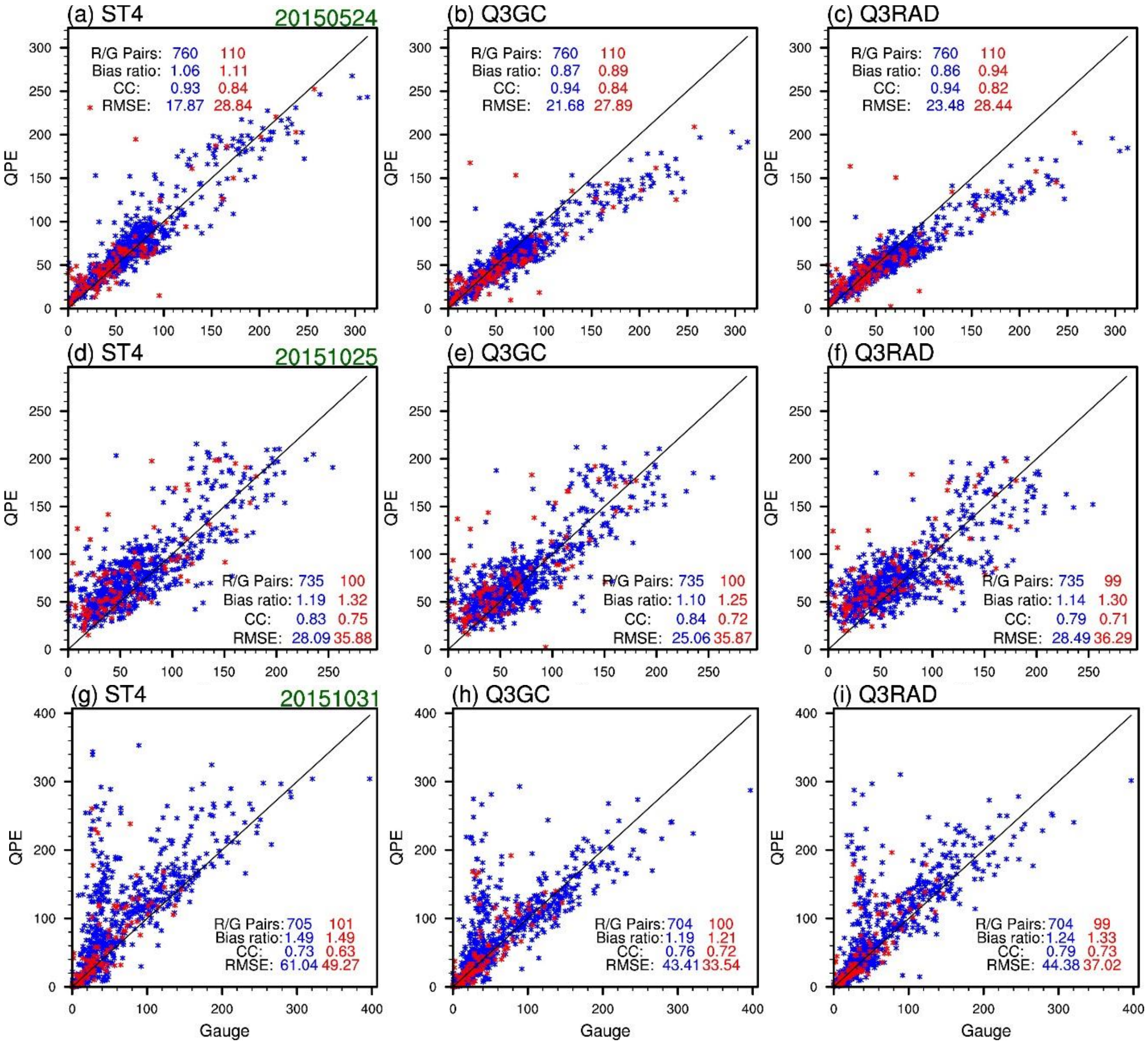


Figure 2. Scatter plot between 24-hour accumulated precipitation from radar-based QPE (y-axis) and that from gauged data (x-axis). The texts show the statistics including radar-gauge ratio (R/G ratio), correlation coefficient (CC), and root-mean-square-error (RMSE).

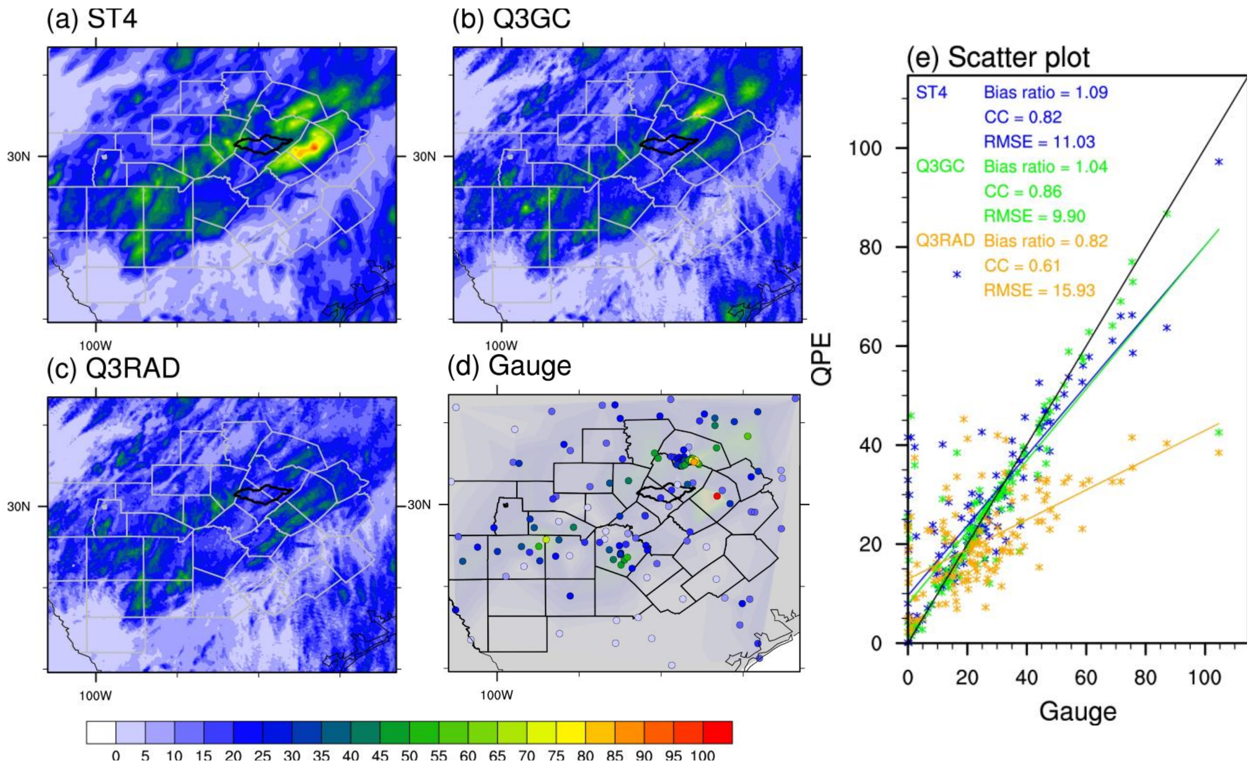


Figure 3. (a–d) Spatial map of the maximum precipitation intensity (MPI, mm hr-1) from 0000 UTC 23 May to 0000 UTC 26 May 2015 for **(a)** ST4, **(b)** Q3GC, **(c)** QCRAD, and **(d)** Individual gauge points. The scatter plot with the regression line and the statistics is shown in **(e)**. The black line is the 1:1 line.

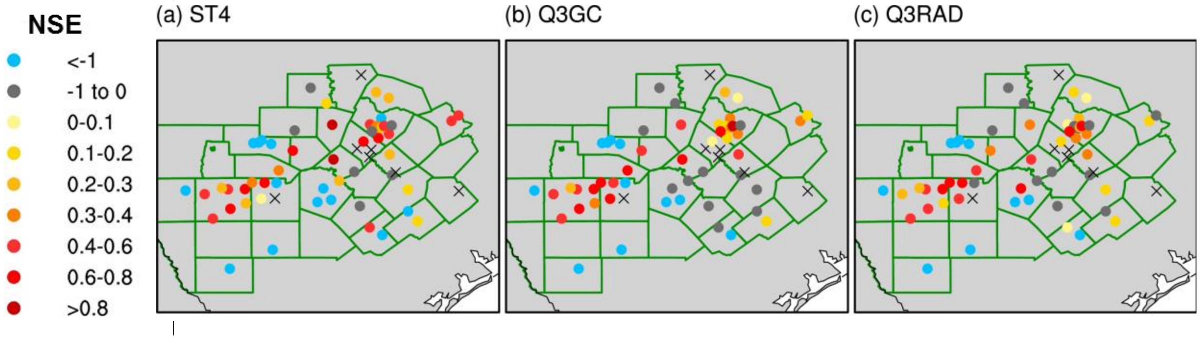


Figure 4. Spatial map of the Nash-Sutcliffe Efficiency (NSE) for different modeling experiments. The statistics are calculated based on hourly streamflow simulation from 0000 UTC 23 May 2017 to 2300 UTC 27 May 2017. Locations with >33% of missing data in the gauged records are represented by a cross.

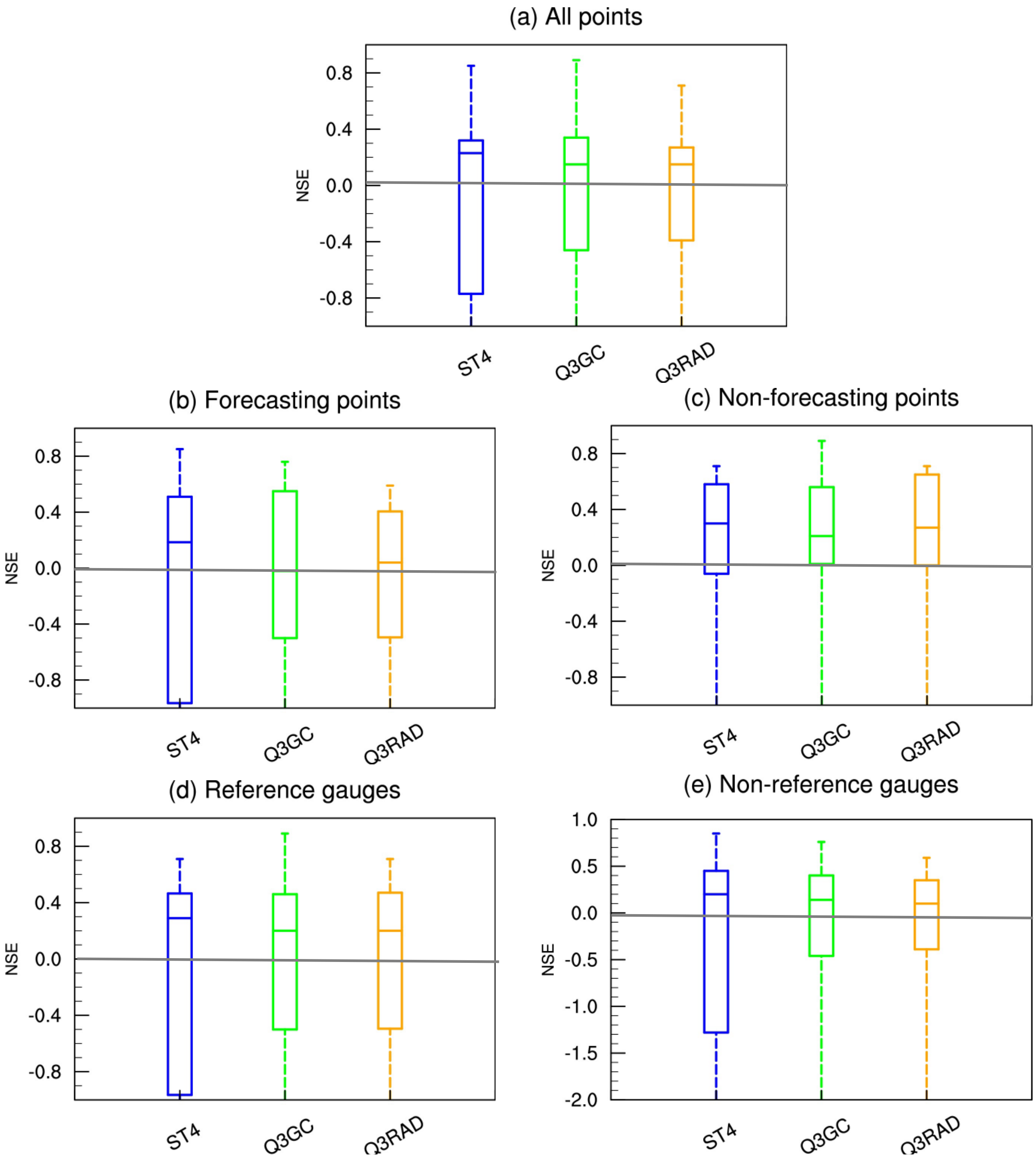


Figure 5. Boxplot of the NSE values of Fig. 4 for different modeling experiments. The whiskers show the maximum, 75th, median, 25th, and minimum NSE values. (a) shows the statistics for all 60 gauges; (b) and (c) separate the data into 33 forecasting points and 27 non-forecasting points; (d) and (e) separate the data into 26 reference gauges and 34 non-reference gauges. Horizontal lines show values when NSE = 0.

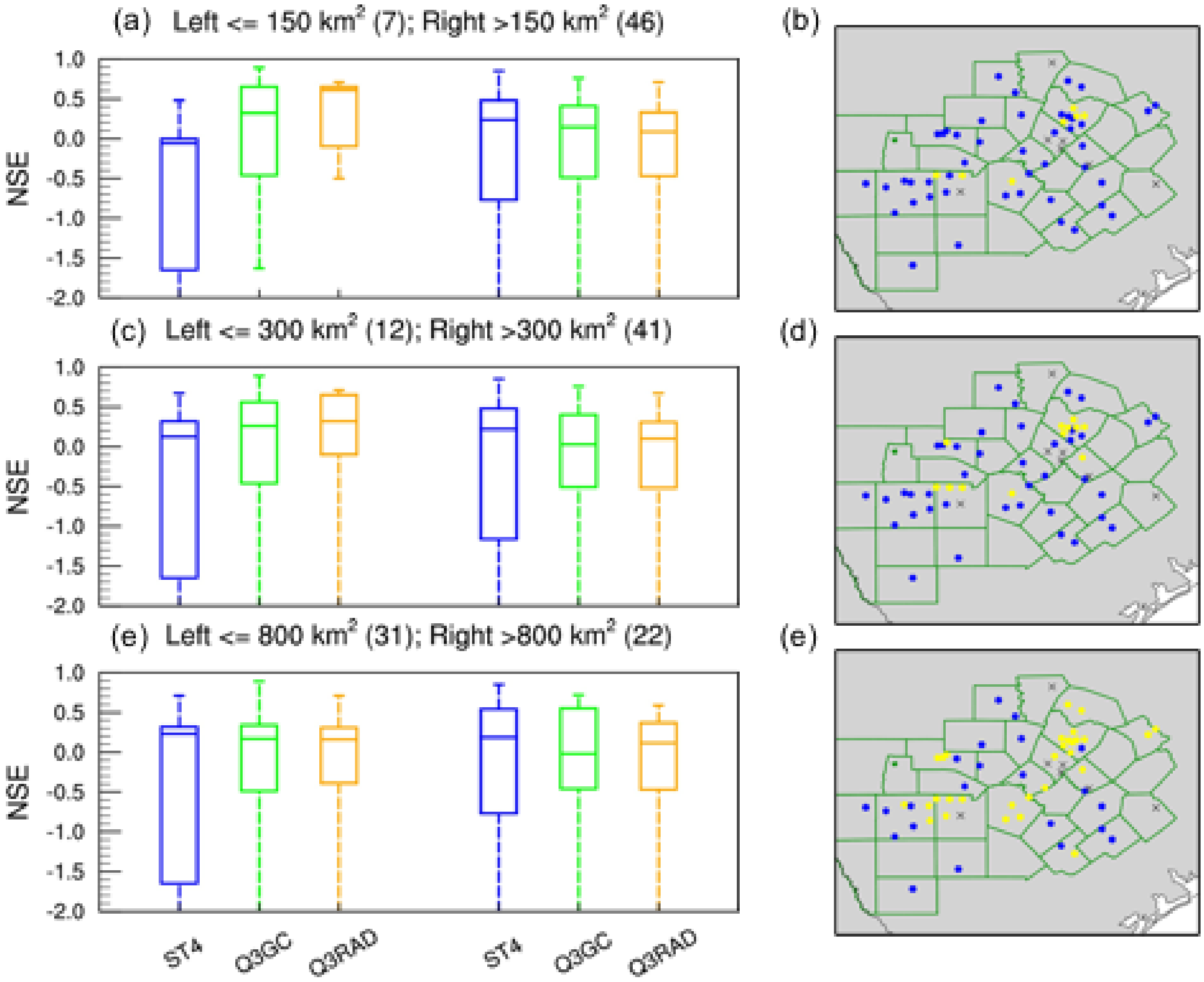
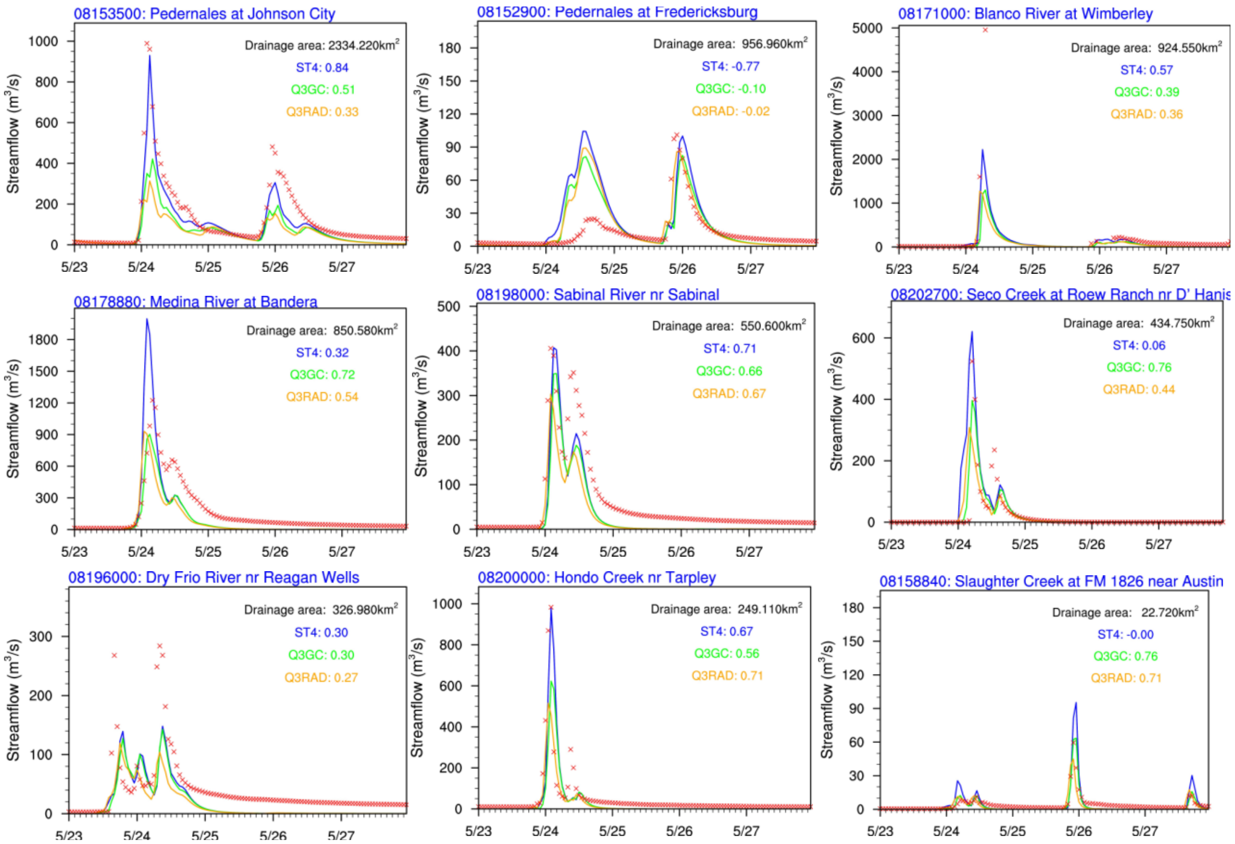


Figure 6. The NSE statistics as categorized by the drainage areas. Left panels (a, c, e) show the boxplot (with maximum, 75th percentile, median, 25th percentile, and minimum as the whiskers), and right panels (b, d, f) show the spatial distribution of gauges with smaller drainage areas (yellow) vs. those with larger drainage areas (blue). Cross indicates that the location has more than 33% of missing data in the gauged records. Enclosed in the parenthesis is the number of gauges within the category.



08153500: Pedernales at Johnson City

08152900: Pedernales at Fredericksburg

08171000: Blanco River at Wimberley

08178880: Medina River at Bandera

08198000: Sabinal River nr Sabinal

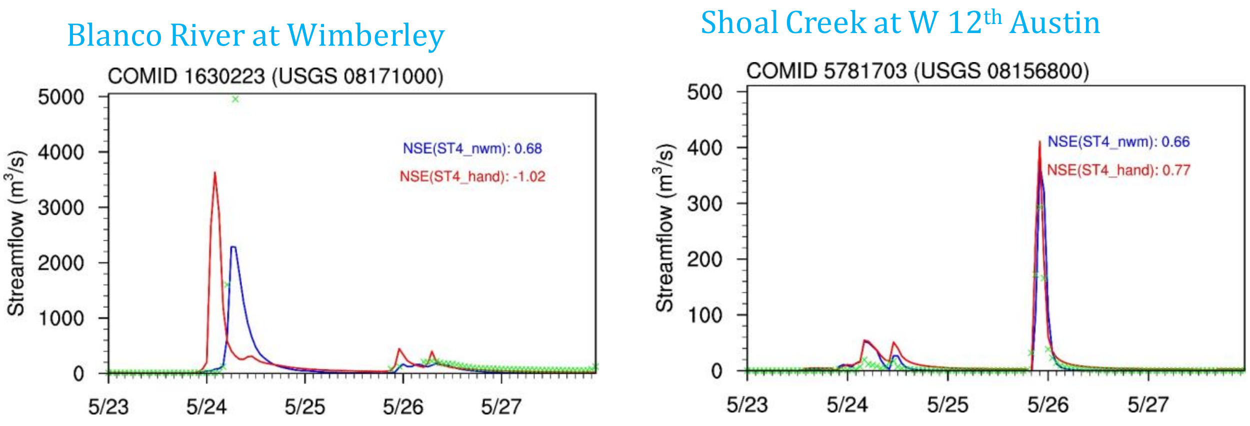
08202700: Seco Creek at Roew Ranch nr D' Hanis

08196000: Dry Frio River nr Reagan Wells

08200000: Hondo Creek nr Tarpley

08158840: Slaughter Creek at FM 1826 near Austin

Figure 7. Flood hydrographs at a few selected gauges-of-interest. The hourly streamflow simulation from 0000 UTC 23 May 2015 to 2300 UTC 27 May 2015 are shown here.



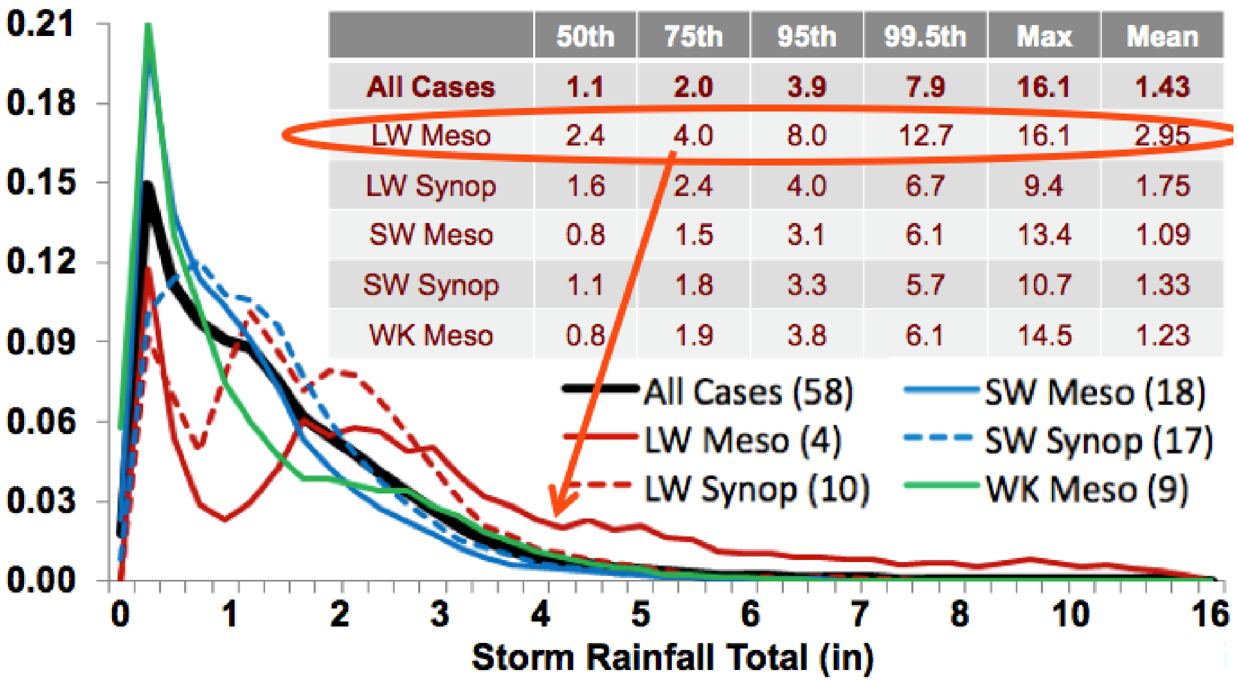
Blanco River at Wimberley

COMID 1630223 (USGS 08171000)

Shoal Creek at W 12th Austin

COMID 5781703 (USGS 08156800)

Figure 8. Flood hydrograph at two stations using different river channel geometry. This figure demonstrates that the model flood prediction skill is very sensitive to the river channel geometry representation as segmented by the NHDPlus dataset.



	50th	75th	95th	99.5th	Max	Mean
All Cases	**1.1**	**2.0**	**3.9**	**7.9**	**16.1**	**1.43**
LW Meso	2.4	4.0	8.0	12.7	16.1	2.95
LW Synop	1.6	2.4	4.0	6.7	9.4	1.75
SW Meso	0.8	1.5	3.1	6.1	13.4	1.09
SW Synop	1.1	1.8	3.3	5.7	10.7	1.33
WK Meso	0.8	1.9	3.8	6.1	14.5	1.23

Figure 9a. Stage IV rainfall probability distributions simulated for 58 heavy rainfall events along the Balcones Escarpment, Texas from 2009–2015 separated by their midlevel flow pattern (longwave, shortwave, or weak) and surface boundary type (synoptic or mesoscale).

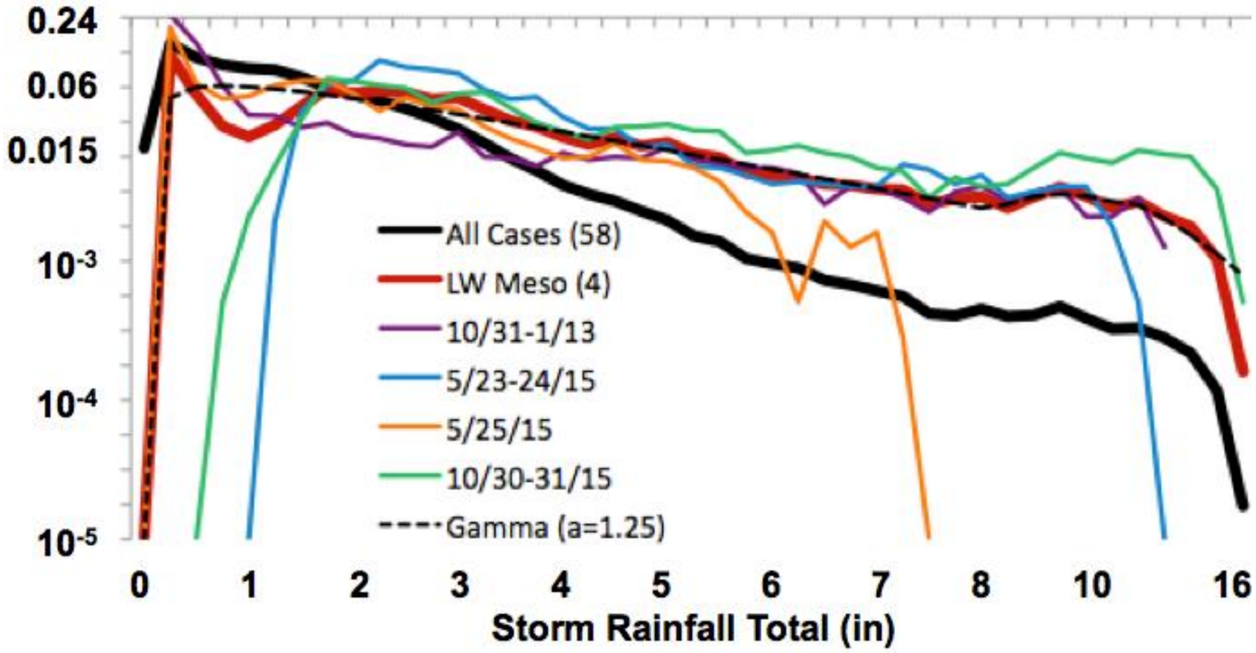


Figure 9b. Stage IV rainfall probability distributions simulated for the four cases associated with longwave midlevel flow patterns and mesoscale surface boundaries. Three of these four cases have been simulated in WRF-Hydro-RAPID for model-observational comparisons in this study.