This is the final report summarizing the research accomplishments conducted for NOAA award NA17NWS4680005

PI: Kristie J. Franz, Department of Geological and Atmospheric Sciences, Iowa State University
Co-PI: William Gallus, Department of Geological and Atmospheric Sciences, Iowa State University

Project overview

The primary goal of this project is to advance the application of ensemble precipitation forecasts for short-term probabilistic streamflow prediction using the High Resolution Rapid Refresh Ensemble (HRRRE) outputs. Under this goal, we addressed these five supporting tasks.

1. Evaluate the use of state-of-the-art PQPF based on HRRR ensembles to create ensemble streamflow information by using different probability thresholds of POP from HRRRE (e.g., every 10%) to input into the hydrologic models.
2. Compare the hydrologic forecast skill when using POP to hydrologic forecast skill when directly using HRRRE members' as input into the models to create probabilistic streamflow forecasts.
3. Examine spatial shifting of HRRRE member QPF to generate streamflow ensembles to improve streamflow forecast errors associated with displacement error in the QPF when forecasting for individual basins.
4. Investigate the mesoscale and synoptic scale environment under which various spatial shifts in QPF occur to identify relationships that can be used to improve hydrologic forecasting by shifting the QPF appropriately based on weather conditions.
5. Quantify the sensitivity of the streamflow ensemble to hydrologic model parameter and state perturbations.

This project covered the forecasting region of the North Central River Forecasting Center (NCRFC), and included forecast basins in Minnesota, Iowa, Wisconsin and Illinois. Throughout this study, we utilized three different NWS hydrologic forecast models: the operational Sacramento Soil Moisture Accounting model (SACSMA), the Hydrology Lab Research Distributed Hydrologic Model (HLRDHM), and the WRF-Hydro version 5.1.1 component of the National Water Model (NWM). When possible, we used the model parameters and configurations used operationally. *Through our work, we have demonstrated that the PQPF could be used with each modeling system, although running ensembles with the spatially distributed models is less practical because of the needed runtime.* In all cases, headwater basins were used and the application of PQPF for forecasting downstream watersheds remain unexplored.

We collected and tested HRRRE data for heavy rain-producing, warm season storm events that occurred throughout our study region for the period of 2017-2019. We held several meetings with personnel from NCRFC to discuss research plans, collect operational data, and present results. This study produced four publications with two more manuscripts still in preparation or review, six national presentations, two M.S. theses, one PhD dissertation, and two undergraduate senior theses.

Our study primarily focused on the precipitation forecasts from the HRRRE system. However, shortly after we started the project, we added forecasts from the High Resolution Ensemble Forecast version 2.0 (HREF) because HREF was considered to be more accurate than the HRRRE among the professional community. To do this, we collaborated with Adam Clark and Brett Roberts from the NOAA National Severe Storms Laboratory (NSSL). A comparison of HRRRE and HREF for flash flood forecasting using the operational SAC-SMA hydrologic model for three Iowa watersheds revealed that the HREF does outperformed the HRRRE for prediction of peak discharge. The average bias in predicted peak discharge was 1459 cubic feet per second (cfs) for the HRRRE-based forecasts and 632 cfs for HREF-based forecasts. Furthermore, the HREF produced forecasts with a better ranked probability score. However, it should be noted that this was a limited comparison and the intent of our work was not to identify the best precipitation ensemble forecast.

Spatially shifted PQPF for ensemble streamflow prediction (Tasks 3 and 4)
Carlberg et al. (2020) presents an initial analysis of an ensemble forecasting method that accounts for spatial errors in QPF ensembles by moving the original precipitation forecast members systematically in space prior to inputting them into a hydrologic forecast model. In this initial study, each of the nine HRRRE members were shifted in eight directions (the four cardinal directions and the four intermediate directions) resulting in an 81-member ensemble. A shifting distance of 0.5° and 1.0° was applied based on our past work of typical displacements for warm season precipitation events.

Compared to the streamflow forecasts created using the original QPF ensemble (raw), the forecasts from the shifted ensemble show an improved probability of detecting a flood (Table 1). The only metrics that did not improve with shifting were the false alarm rate and the ranked probability score. The decrease in skill indicated by the ranked probability score was linked to the large 81 member ensemble having less confident predictions because the probability is calculated based on a sample of 81 rather than 9. We tested a simple weighting scheme to adjust ensemble member weights based on the average displacement direction (i.e., shifts to the east were weighted higher than shifts to the west). The ranked probability scores were improved for these conditioned forecasts (Carlberg et al., 2020).

Table 1 Forecast verification metrics for the streamflow ensemble forecasts created using the 9 raw HRRRE QPF members, the 81-member ensemble using the a 0.5° shift of QPF, the 81-member ensemble using the 1.0° shift of QPF, and the streamflow forecast using the NCRFC Mean Precipitation Estimate (MPE). All forecasts were created using the HLRDHM.

Metric	**Raw**	**0.5° shift**	**1.0° shift**	**NCRFC MPE**	**Perfect score**
Frequency of non-exceedance (FNE)	0.43	0.72	0.78	N/A	1
Probability of detection (POD)	0.33	0.60	0.80	0.73	1
False alarm rate (FAR)	0.29	0.44	0.45	0	0
Critical success index (CSI)	0.29	0.41	0.48	0.73	1
Equitable threat score (ETS)	0.18	0.23	0.27	0.65	1
Ranked probability score(RPS)	0.75	0.74	0.82	2.21	0
Weighted RPS	0.75	0.70	0.78	N/A	0

Results in Table 1 show that the 81-member ensemble tends to have better forecast verification metrics in many cases, but the ensemble contains many non-events resulting in low probabilities that potentially reduce the usefulness of the forecasts. Therefore, the next step in this analysis was to refine the shifting technique to both improve the accuracy and confidence of the forecast and reduce the number of ensembles that are run.

To develop an improved shifting scheme, we first characterized, quantified and compared the spatial displacement errors in the HRRRE and the HREF (Kiel et al., in review). We found that *for total accumulated precipitation, the HRRRE QPF contained a slight westward bias (QPF was displaced to the west), whereas HREF QPF displayed no consistent location bias.* QPF displacements at the hours of precipitation initiation were also examined to determine if they could be an indicator of the displacement that would occur for total accumulation. If a relationship existed, the displacement observed at initiation could possibly be used for ensemble updating in real time during an event.

Results showed that there is not a straightforward relationship between QPF displacement at the initiation hour and for the total period, but the correlation between initiation hour displacement and total hour displacement is stronger for some displacement directions (Kiel et al., in review). Therefore, the initiation hour displacement could give useful information on the final displacement if one takes into account the initial direction of the displacement. This idea is particularly promising for the HRRRE in which the displacement at total accumulation tends to occur in the same quadrant, or same general N-S direction, as at initiation (Figure 1).

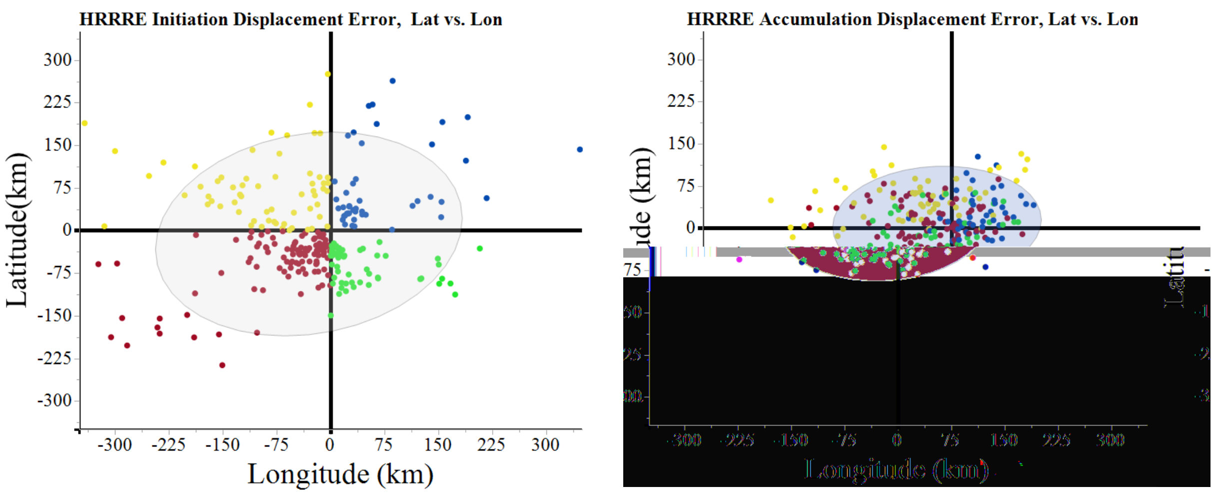

Figure 1: HRRRE initiation hour (left) and accumulation (right) displacement errors for all members of all cases studied, graphed by latitude and longitude in kilometers. Points are color-coded by the quadrant in which they were displaced at during the initiation hour. A 90% bivariate ellipse is drawn to indicate the general displacement bias.

Hugeback et al. (in prep) combines lessons learned from Carlberg et al. (2020) and the QPF analysis from Kiel et al. (in review) to test a refinement of the QPF shifting technique. In this study, the HRRRE was shifted using a randomly sampled direction and distance based on the actual HRRRE displacement errors (a proxy for a climatology of displacement errors) identified by Kiel et al. (in review). Further, a statistical analysis was conducted to identify the number of ensemble members that are needed to reasonably reproduce most of the displacement error variability of the HRRRE. It was found that 54 shifts were needed; while this still produces a large ensemble, it reduces the number of model runs by approximately 1/4 compared to the Carlberg et al (2020) method. The shifted QPF along with the original QPF were input into WRF-Hydro version 5.1.1 in a National Water Model 2.0 configuration to generate ensembles streamflow predictions for 50 stream gauges, over 29 events from the 2018 warm season. Weighting schemes were tested to adjust the streamflow ensembles based on the position of the ensemble members at convective initiation, a method suggested above.

Similar to Carlberg et al. (2020), the shifted ensemble produced higher peak flows than the original 9-member ensemble, which resulted in the shifted ensemble having a higher frequency of non-exceedance (FNE; how often the observation is captured within the bounds of the ensemble). *The streamflow* *ensembles produced from the randomly shifted QPF have better reliability and ranked probability scores for predictions of peak flow compared to the streamflow ensembles generated using original 9-member HRRRE* (Table 2). The accuracy of the predicted timing of peak streamflow was, on average, neither improved nor degraded by the shifting. The weighting scheme using convective initiation displacement information did not improved the forecast skill when compared to the equal weighting scheme (Table 2). We are still analyzing these results to determine why the expected improvement did not occur.

Table 2: Mean values for FNE and RPS across all events, for streamflow forecasts generated using the shifted and original ensembles. Results are shown for equal member weighting, weighting based on CI displacements, and weighting based on corrected CI displacements. The mean absolute error (MAE) is the mean difference between the forecast reliability and perfect reliability and is used to summarize the reliability results. Lower values of MAE are better values.

		Full Ensemble	9-mem HRRRE	Perfect score
Equal Weights	FNE	0.514	0.408	1
	RPS	0.423	0.454	0
	MAE	0.15	0.223	0
CI Weighted	FNE	0.514	0.408	1
	RPS	0.429	0.465	0
	MAE	0.174	0.234	0
CI Corrected	FNE	0.514	0.408	1
	RPS	0.428	0.469	0
	MAE	0.174	0.233	0

Using POP for probabilistic streamflow forecasting (Tasks 1 and 2)
Goenner et al. (2020) presents an analysis of the HRRRE POP for hydrologic forecasting using our methodology to convert the HRRRE POP to precipitation time series for hydrologic modeling. Precipitation time series were derived from POP and then input into a hydrologic forecasts model to generate streamflow predictions with probability thresholds of 5%, 10%, 25%, 75%, 90% and 95%. The NWS forecast model from the NCRFC was used to allow comparison to operational WPC-based forecasts generated with the same model and parameters. The High-Resolution Ensemble Forecast (HREF) system was included this study because of its operational status and usefulness as a comparison product to HRRRE. The method was evaluated for events that occurred in 2017 and 2018 in our study domain.

Probabilistic forecasts created using this method were poorly calibrated and largely overestimate the probability of major flooding. Similarly, predicted discharges associated with low probability of exceedance were much too large with no observations ever reaching the flow associated with the probability of exceedence threshold of 10% or less for HRRRE, and 25% or less for the HREF-based forecast. The high bias is likely due to the technique applying low probability of exceedance precipitation values (high rain rates) at every grid point in a basin.

When the ranked probability scores of the forecasts were evaluated, we found that the skill of the forecasts for prediction of the peak discharge was best for HRRRE, followed by the HREF (Figure 2). The NCRFC forecasts had the highest (poorest) RPS values. *Although our method provided more accurate probabilistic forecasts than the operational WPC-based forecasts issued from the NCRFC, we still do not recommend this approach without significant pre- and post-processing of the forecast information.* We tested a simple calibration on the probabilities that suggested the approach could work well with post-processing. Conducting proper statistical adjustments, however, requires a long record of forecasts, which is difficult to do with newer forecast products (Goenner et al., 2020).

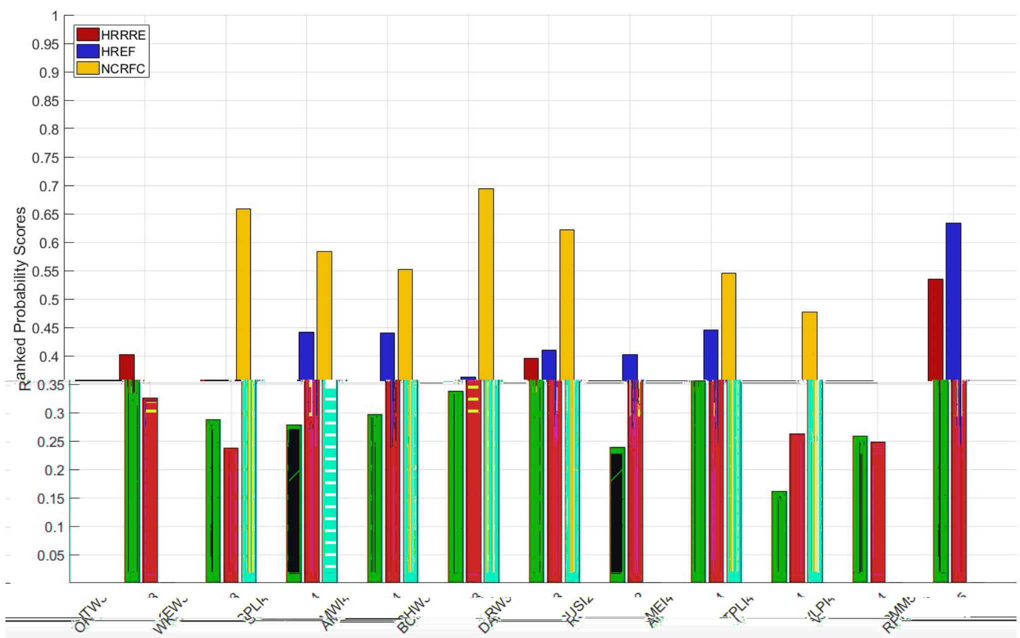

Figure 4 (from Goenner et al: RPS for the eleven basins (labeled using NCRFC abbreviations) arranged in order from the smallest (left) to the largest (right) in area. Probability thresholds used were the 5%, 50% and 95%.

Exploring model sensitivities (Task 5)

In consideration of NOAA's goal to deliver high-resolution reach-based river forecasts for the US, we undertook an analysis that evaluated the skill of a distributed rainfall-runoff model at multiple watershed scales (Madsen et al, 2020). Using the HL-RDHM, we tested the sensitivity of the model, which uses a 4-km grid spacing, to the resolution of the precipitation input. Two precipitation products were tested: Stage IV at ~4km resolution and NLDAS-2 at ~12.5 km resolution. We evaluated the accuracy of the model for subbasin scales ranging from 25 km^2 – 3493 km^2.

Discharge simulations were more accurate for the precipitation input with coarser spatial resolution, likely because the model structure is unable to represent the hydrologic effects produced by small-scale intense precipitation found in the 4-km precipitation product. Intense precipitation is more likely to produce infiltration excess runoff, whereas the model is better designed to handle saturation excess runoff. Model performance decreased with decreasing subbasin size, and performance dropped off significantly for subbasins smaller than 250km^2 or 30% of the total basin area. *Thus, understanding the* *role of input data resolution and the scale at which the hydrologic model can accurately resolve processes is an important consideration in forecasting.*

Publications

Carlberg, B., W. Gallus, K.J. Franz, 2018: A preliminary examination of WRF ensemble prediction of convective mode evolution, *Weather and Forecasting*, 33(3), 783-798, doi:10.1175/WAF-D-17-0149.1

Madsen, T., K. Franz, and T. Hogue, 2020: Evaluation of a distributed forecast model at multiple watershed scales, *Water*, 12, 1279, doi:10.3390/w12051279.

Carlberg, B., Franz K.J., W. Gallus, 2020: A Method to Account for QPF Spatial Displacement Errors in Short-Term Ensemble Streamflow Forecasting. *Water*, *12*, 3505.

Goenner, A.R.; Franz, K.J.; Jr, W.A.G.; Roberts, B., 2020: Evaluation of an Application of Probabilistic Quantitative Precipitation Forecasts for Flood Forecasting. *Water*, *12*, 2860.

Publications in prep, revision, or review

Kiel, B., W. Gallus, K.J. Franz, A climatology of precipitation placement errors in high-resolution ensembles, *in re-review for the Journal of Hydrometeorology.*

Hugeback, K., W. Gallus, K.J. Franz, Incorporating QPF displacement uncertainty in ensemble streamflow predictions using the WRF-Hydro, *in preparation for submission to Journal of Hydrometeorology*.

Presentations

Carlberg, B. R., K. J. Franz, W. A. Gallus Jr., Creating an ensemble streamflow forecast through the systematic shifting of QPF. Presentation. 29th Conference on Weather Analysis and Forecasting, Denver, CO, 2018.

Franz K.J., B. Carlberg, W. Gallus, Ensemble streamflow forecasts using spatially shifted QPF, Annual Meeting of the American Meteorological Society, Phoenix, AZ, January, 2019.

Goenner, A., K.J. Franz, W. Gallus, An approach to create probabilistic streamflow forecast from HRRRE quantitative precipitation forecasts, Annual Meeting of the American Meteorological Society, Phoenix, AZ, January, 2019.

Kiel, B., W. Gallus, K.J. Franz, and B. Carlberg, Convective system displacement errors in HRRRE and potential use for shifting QPF fields to improve flood forecasting, Annual Meeting of the American Meteorological Society, Phoenix, AZ, January, 2019.

Hugeback, K., B. Kiel, W. Gallus, K. J. Franz, Generation of WRF-Hydro probabilistic streamflow forecasts by shifting ensemble QPF based on a climatology of forecast rainfall displacement errors, Annual Meeting of the American Meteorological Society, Boston, MA, January, 2020.

Hugeback, K.K., W.A.Gallus, Jr., K.J. Franz, Accounting for displacement errors in HRRRE QPF to create short term ensemble streamflow forecasts, Annual Meeting of the American Meteorological Society, virtual, January 2021.

Gallus, Jr., W., K. K. Hugeback, K. Franz, Accounting for Spatial Displacement Errors in Ensemble Member QPF to Create Short-Term Ensemble Streamflow Forecasts, **Invited** presentation for the Flash Flood and Intense Rainfall Experiment seminar series, NOAA, July 13, 2021.

Other products

Hugeback, K., A Comparison of HREF and HRRRE Predictions for Ensemble Flash Flood Forecasting, Meteorology Senior Thesis, Iowa State University, 2018.

Kiel, B. A Climatology of Precipitation Displacement Errors in High-Resolution Ensembles, Meteorology Senior Thesis, Iowa State University, 2018.

Carlberg, B., Improving convective mode and streamflow forecasts through the use of convection-allowing ensembles, PhD Dissertation, Iowa State University, 2018.

Goenner, A., An approach to create probabilistic streamflow forecasts from HRRRE & HREF probabilistic quantitative precipitation forecasts, M.S. thesis, Iowa State University, 2019.

Hugeback, K., Accounting for Spatial Displacement Errors in HRRRE QPF to Create Short-Term Ensemble Streamflow Forecasts, M.S. thesis, Iowa State University, 2021