

SPATIAL ALIGNMENT OF CAM ENSEMBLE TO IMPROVE ENSEMBLE CONSENSUS PRECIPITATION

ChangJae Lee*¹ and Keith A. Brewster²

¹Korea Meteorological Administration
Seoul, Korea

²Center for Analysis and Prediction of Storms,
University of Oklahoma, Norman, OK, USA

1. INTRODUCTION

Several methods have been devised to find consensus among ensemble forecast members. For example, the ensemble mean, a simple point-wise arithmetic average of ensemble members is commonly used in operational ensembles. However, due to the difference in spatial distribution and intensity of precipitation features in each ensemble member, the ensemble mean of precipitation forecasts tends to reduce the magnitude of forecast maxima while expanding the areal coverage of light precipitation.

The probability-matched (PM, Ebert, 2001) ensemble mean and localized PM (LPM, Clark, 2017, Snook et al., 2020) mean methods have been introduced to overcome these problems. The PM and LPM use ensemble members' probability density function (PDF) to preserve the ensemble forecast's maxima. PM and LPM methods redistribute the values of each grid point of the ensemble mean but may not preserve the spatial structures of the features themselves, which can be blurred if there are offsets of feature locations among the members.

This study aims to find a way to improve ensemble consensus precipitation by directly considering the spatial offsets among ensemble members. This study uses the phase-correcting method to align the fields of each ensemble member to a common location. Offsets are found for each ensemble member with respect to other members in pairs, and the vector mean offset is calculated from all pairs. The fields are shifted by this offset, moving them toward a common central location, and the Spatially Aligned Mean (SAM) is obtained by averaging the re-aligned members.

2. METHOD

The algorithm used in Spatially Aligned Mean (SAM) is based on the Phase-Correcting Data Assimilation method (Brewster, 2003), as diagramed in Fig. 1. This algorithm is to find shift vectors that minimize a squared difference sum among a pair of ensemble members. Once shift vectors are obtained and averaged among all pairs, each member's field or fields can be adjusted using shift vectors.

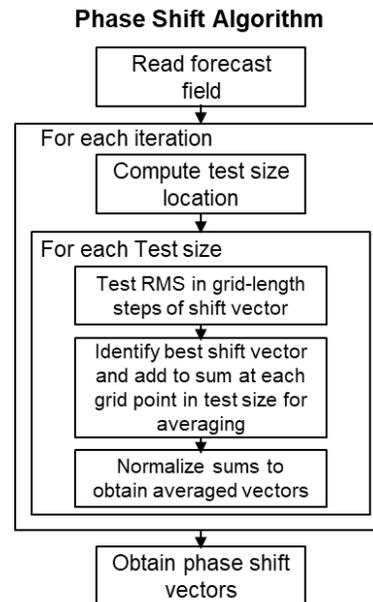


Figure 1. Flowchart for obtaining phase-shift vector fields

Since spatial offsets can vary across the model domain, the algorithm proceeds by dividing the domain into overlapping patches. The patch size is flexible and can be set while considering horizontal scales like synoptic-scale, meso-alpha, and meso-beta scale. For each test patch, shift vectors are determined by finding the offset of grid points in the

*Corresponding author contact: ChangJae Lee,
changjae.lee.3789@gmail.com

VALID: 2022.09.28.15UTC (+015H) / RUN: 2022.09.28.00UTC / 3hr Precipitation

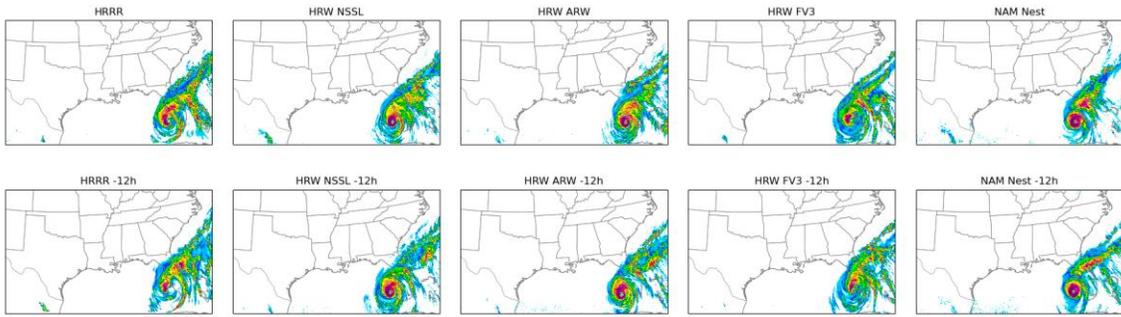


Figure 2. HREF result of 2022 Hurricane Ian case (original field)

VALID: 2022.09.28.15UTC (+015H) / RUN: 2022.09.28.00UTC / 3hr Precipitation / 2 pass

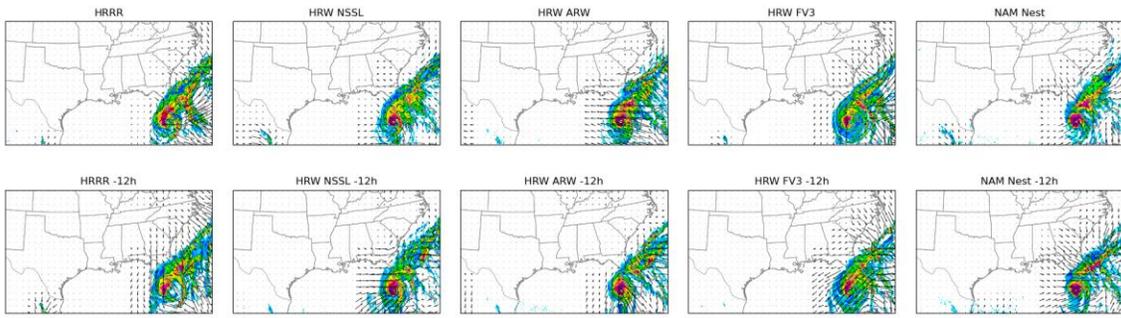


Figure 3. Shifted fields of HREF result of 2022 Hurricane Ian case with shift vectors

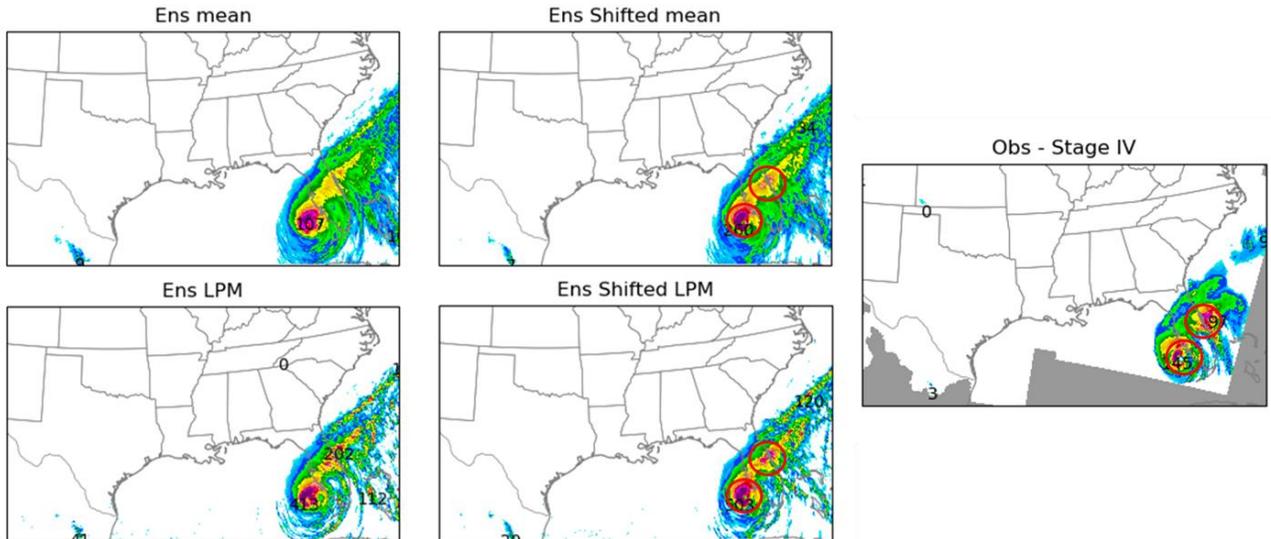


Figure 4. Ensemble mean, Spatially Aligned Mean, LPM, and SAM-LPM field of HREF result of 2022 Hurricane Ian case

x/y directions, which minimizes the RMS differences between each pair of members, including a penalty for large offset distance. The entire domain's shift vector field is obtained by averaging all the overlapping test patches' shift vectors. An iterative approach using a cascade of test patch sizes can be applied so that both large-scale and small-scale phase errors can be corrected.

After determining the shift vectors for each member-member pair, the vectors are then averaged to find the offset for each member to bring the fields to a central location. After moving the field using the calculated and averaged shift vectors, the field values are reassigned using the PDF from the original field to restore its intensity. This is necessary because moving the grid involves interpolation which has the effect of smoothing and losing the maxima.

To illustrate this, we present an example of the technique applied to US operational models. Fig. 2. is the 3-h precipitation accumulation from the individual members of the US High-Resolution Ensemble Forecast (HREF) for the case of Hurricane Ian in 2022, and Fig. 3. is the result after SAM has been applied. Applying the LPM technique to the average of all phase-shifted members (SAM-LPM) results in the fields shown in Fig. 4. The forecasted precipitation fields have a more clear eye structure and the spiral band of the hurricane to the northeast of the center, which has intense rainfall on the east coast of Florida, closely matching the observed rainfall.

3. EXPERIMENT DESIGN

In this experiment, the Spatially Aligned Mean (SAM) is applied to 3-hour accumulated precipitation output from an operational high-resolution (~3 km) Convection-Allowing Model (CAM) ensemble, the US High-Resolution Ensemble Forecast (HREF), which has ten members, including the 12-hour time-lagged members. Also, to preserve the ensemble forecast maxima, LPM is applied to the SAM results.

The proposed SAM-LPM technique is applied 3- for lead times of 15 to 36 hours over the contiguous United States (CONUS) and verified using Stage IV (4km resolution) precipitation data. Testing is done for four weeks (Jun 20 ~ Jul 3, Jul 10 ~ Jul 24) in the summer of 2022 corresponding to the Hydrometeorology Testbed (HMT) Flash Flood and

Intense Rainfall (FFaIR) experiment. Both 00Z and 12Z ensemble members are used.

In this experiment, two cascading test patch sizes are used. For the first pass, the patch size is 600 km (synoptic scale), and 225 km (meso-alpha scale) is used for the patch size in the second pass. Both results of SAM and SAM-LPM are evaluated for both the first and the second passes.

Point-wise and spatial feature verifications are performed with several precipitation thresholds using the Meteorology Evaluation Tools (MET) program. The neighborhood method is used with a 32 km width for point-wise verification, and Method for Object-based Diagnostic Evaluation (MODE) is used for spatial verification.

Detailed results of each case are presented on the Web via the links at https://caps.ou.edu/clee/ens/ens_view.php.

4. VERIFICATION RESULTS

The verification is performed with several 3h precipitation thresholds (1 mm, 5 mm, 10 mm, 15 mm, 20 mm, and 25 mm) in order to evaluate performance for weak through intense rainfall. Since these verification results have a diurnal cycle corresponding to the diurnal variation in 3-hour rainfall, the overall verification result will also be provided in the figures.

Figures 5 and 6 show Frequency Bias results for the 10 mm and 25 mm 3h rainfall threshold over the 2022 FFaIR period. At these thresholds the frequency bias for SAM is higher (and closer to 1.0, ideal) than that of the regular mean. Also, the frequency bias for SAM-LPM is higher than the standard LPM, which means SAM-LPM has more compact high-intensity values because the neighborhood method is used for the verification with the same PDF. In the case of intense rainfall with a 25mm/3h threshold, the overall Frequency Bias for SAM-LPM has a score close to 1.

Figures 7 and 8 show the Probability of Detection (POD) results for 10mm and 25mm 3-h rainfall thresholds over the 2022 FFaIR. POD for SAM and SAM-LPM increases significantly compared to the regular mean and LPM, especially in high-impact rainfall. These results also include a diurnal cycle, and in the case of heavy rain, night-time precipitation generally scores better because convection initiation in the afternoon is hard to predict with the right position and timing.

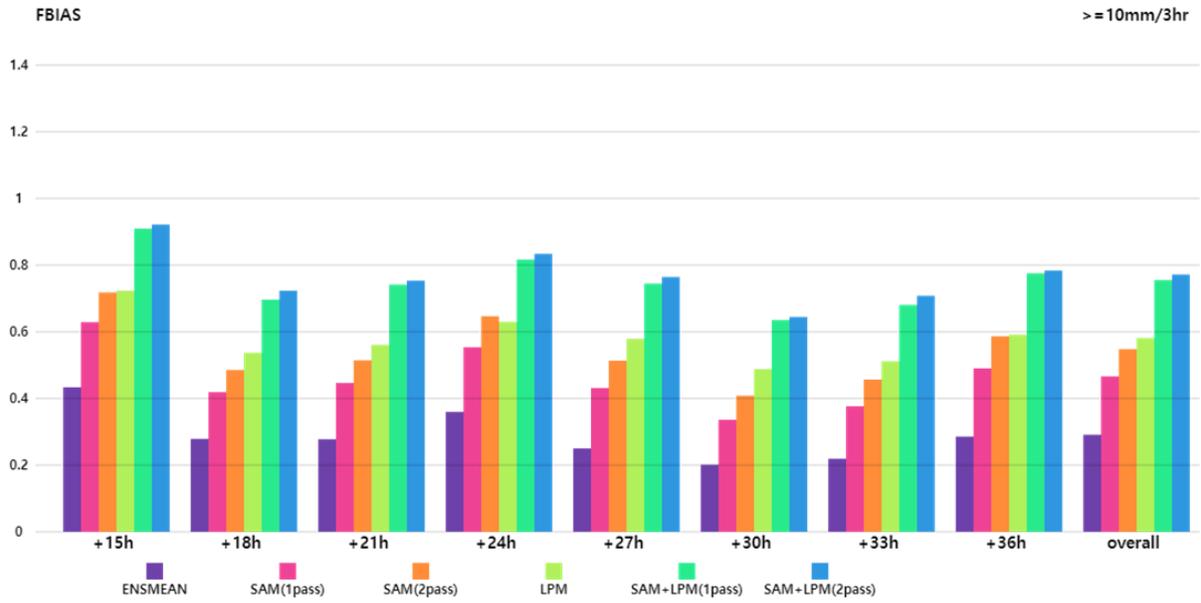


Figure 5. Frequency Bias results for 10mm/3h rainfall threshold over the 2022 FFaIR

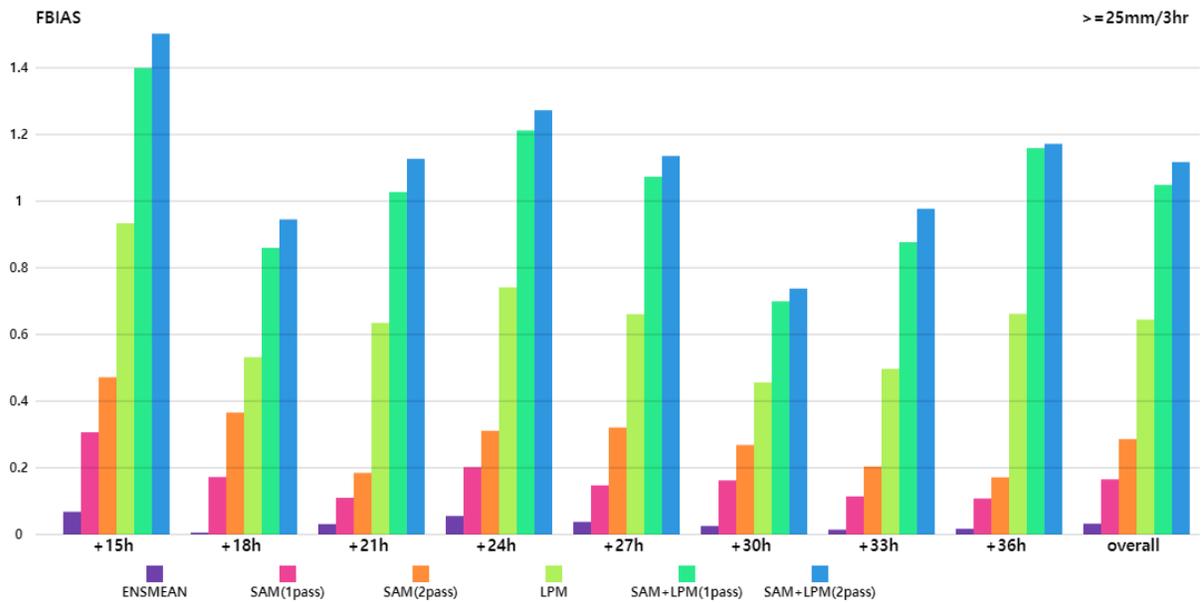


Figure 6. Frequency Bias results for 25mm/3h rainfall threshold over the 2022 FFaIR

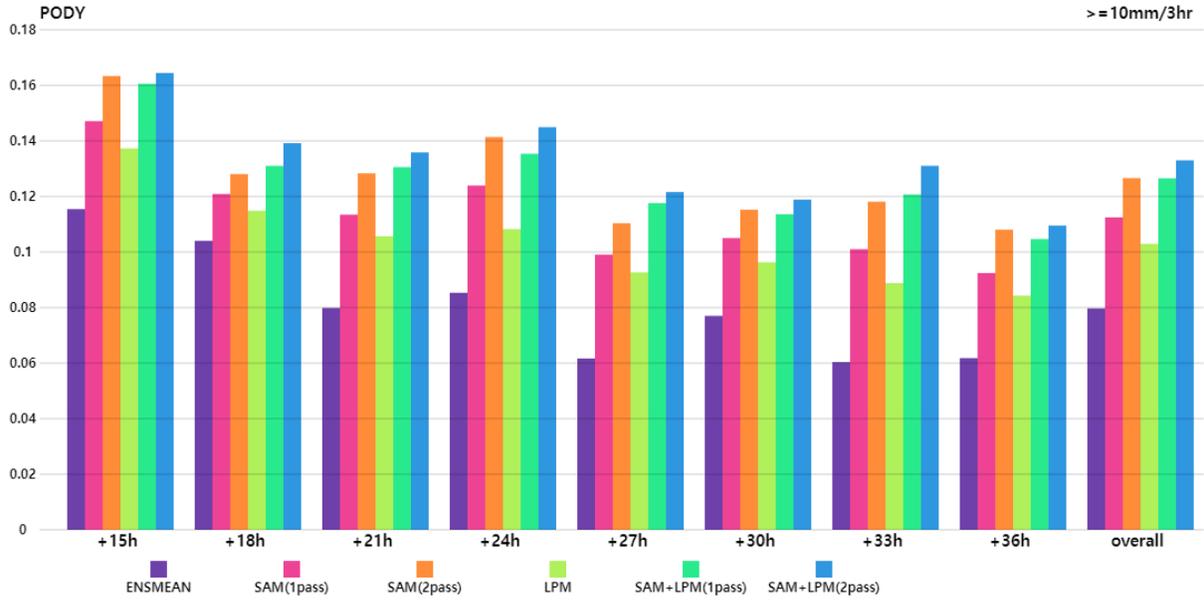


Figure 7. Probability of Detection results for 10mm/3h rainfall threshold over the 2022 FFaIR

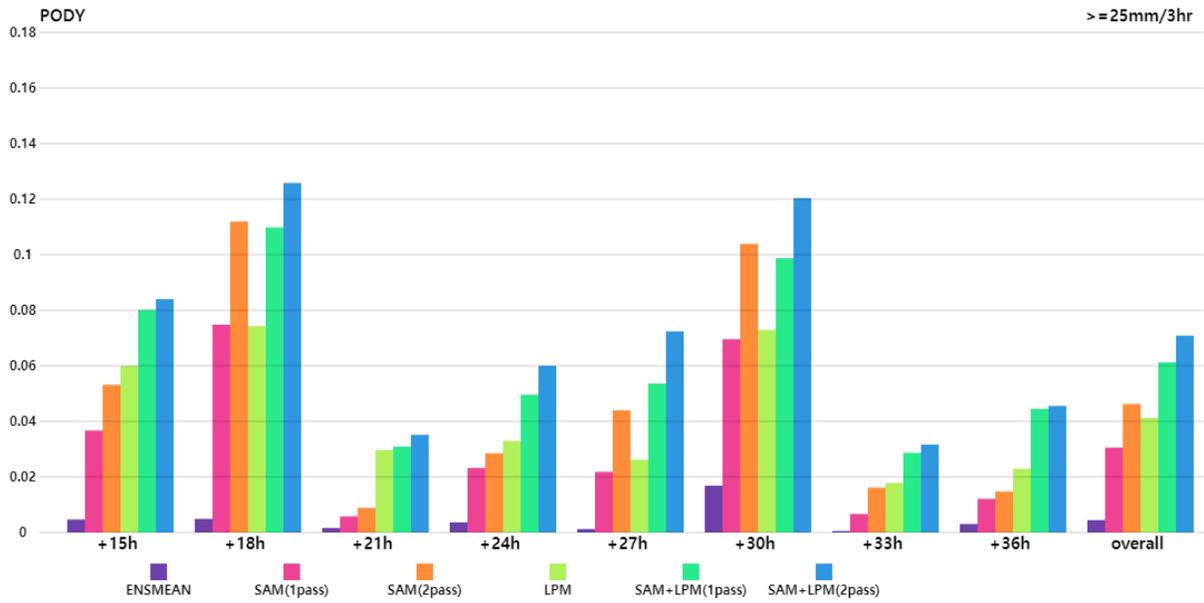


Figure 8. Probability of Detection results for 25mm/3h rainfall threshold over the 2022 FFaIR

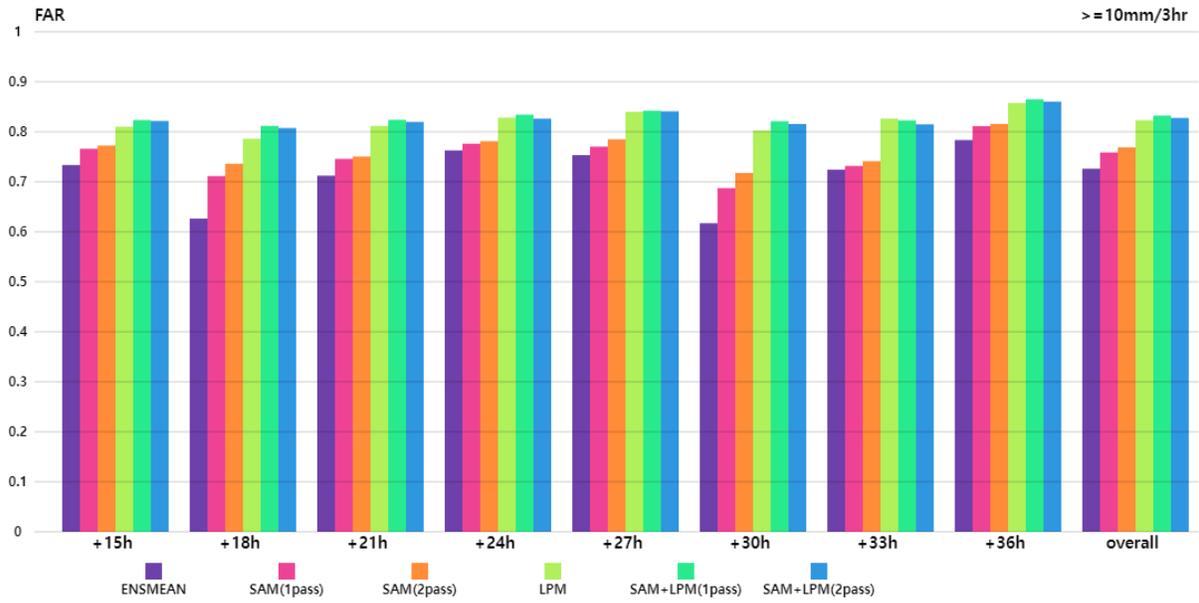


Figure 9. False Alarm Ratio results for 10mm/3hr rainfall threshold over the 2022 FFaIR

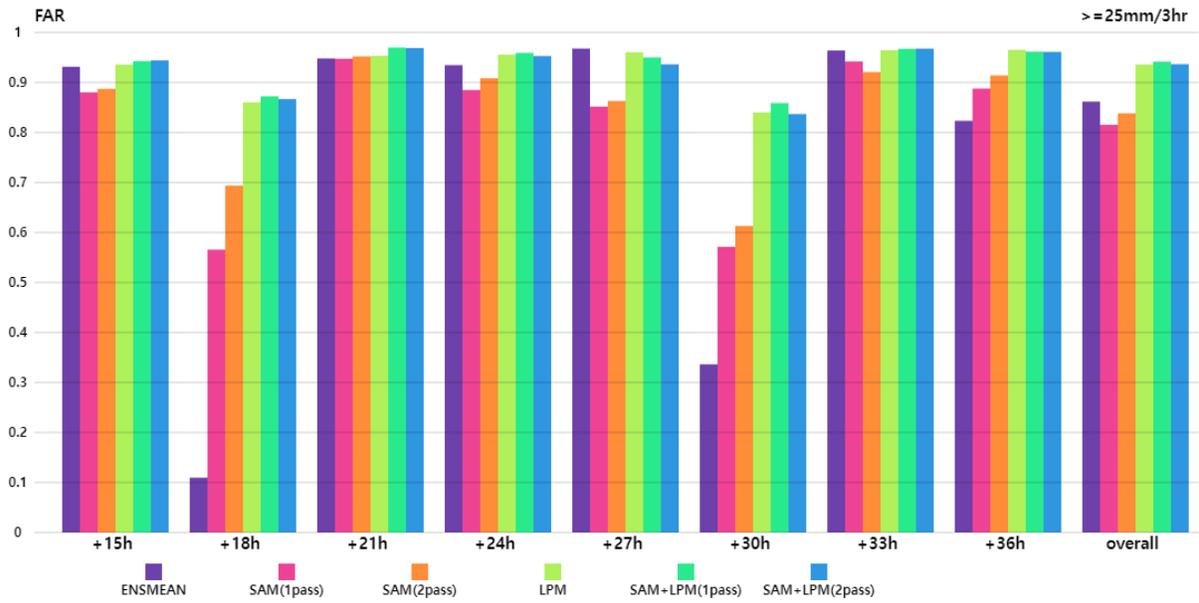


Figure 10. False Alarm Ratio results for 25mm/3hr rainfall threshold over the 2022 FFaIR

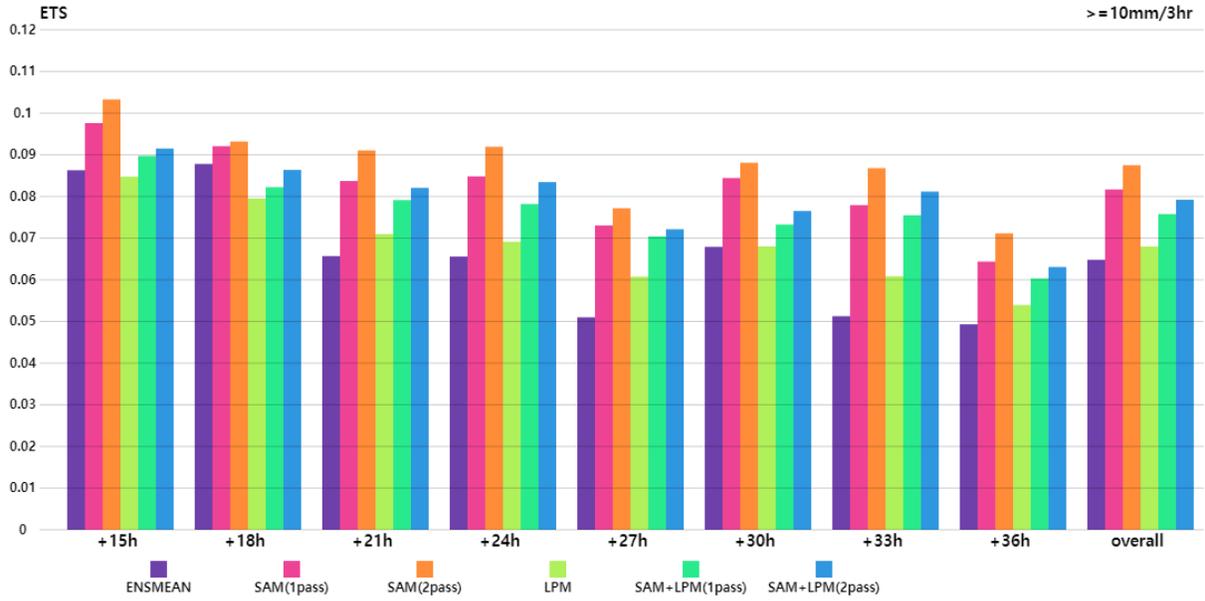


Figure 11. Equitable Threat Score results for 10mm/3hr rainfall threshold over the 2022 FFaIR

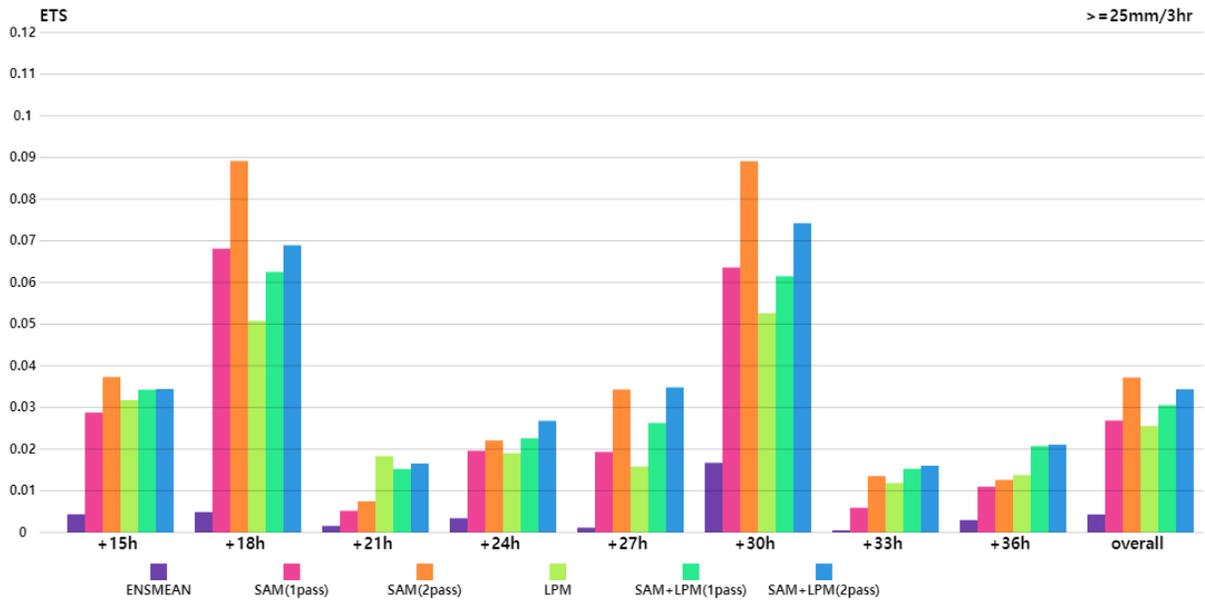


Figure 12. Equitable Threat Score results for 25mm/3hr rainfall threshold over the 2022 FFaIR

Figures 9 and 10 show the False Alarm Ratio (FAR) results for 10 mm and 25 mm 3-h rainfall threshold over the 2022 FFaIR. At the lower threshold (10 mm), FAR for SAM is higher than the regular mean, but in the higher threshold (25 mm), FAR for SAM is similar. FAR for SAM-LPM is similar to regular LPM for all thresholds, which means the spatial alignment technique does not worsen the results for this metric.

Figures 11 and 12 show the Equitable Threat Score (ETS) results for 10mm and 25mm 3-h rainfall thresholds over the 2022 FFaIR period. ETS for SAM and SAM-LPM is higher than the simple mean and LPM for all time periods. ETS for SAM is slightly better than SAM-LPM because although the SAM-LPM has a much higher POD, it also has a higher FAR than SAM at this threshold.

5. SPATIAL VERIFICATION

In this section, spatial verification using Method for Object-Based Diagnostic Evaluation MODE (Bullock et al. 2016) as part of the MET Toolkit (Jensen et al., 2023) will be shown for a few cases. These cases are examined to see if SAM improves the spatial features of ensemble consensus.

The first case is one of a squall line along a front from 12 UTC on 25 July 2022. Due to the spatial difference of the location of the boundary in the 24h forecasts, the standard LPM in Fig. 13 shows sections of the line in two parts, an artifact of position differences among individual members. SAM-LPM in Fig. 14. has a consistent single-line structure for the length of the squall line, similar in structure to the observed squall line.

The improvement in forecast structure for SAM-LPM vs LPM is measured quantitatively by the Interest metric of MODE (Table 1.), which is the sum of normalized feature scores such as intersection ratio, centroid distance, and angle distance. The Interest score for the squall line feature is higher in SAM-LPM, due especially to a closer centroid distance and larger intersection area.

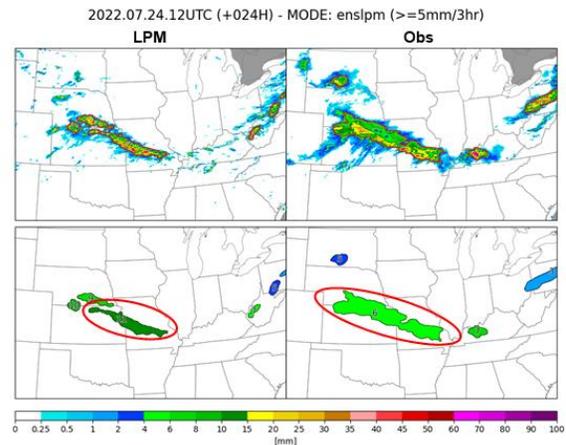


Figure 13. MODE results ($\geq 5\text{mm}/3$) of LPM and Observation for squall line case (24h forecast valid at 12 UTC 25 July 2022)

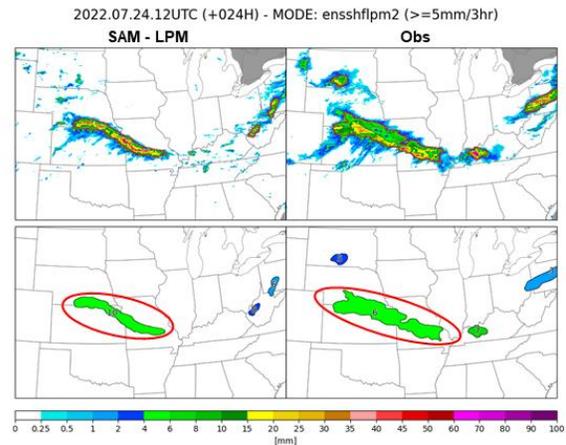


Figure 14. MODE results ($\geq 5\text{mm}/3$) of SAM-LPM and Observation for squall line case (24h forecast valid at 12 UTC 25 July 2022)

	LPM	SAM-LPM
The interest of MODE	0.92244	0.9605
Centroid distance (km)	35.35783	5.95921
Intersection area (km ²)	2221	3357

Table 1. The scores of MODE metrics ($\geq 5\text{mm}/3$) of LPM and SAM-LPM for the squall line case (24h forecast valid at 12 UTC 25 July 2022)

A case from 21 UTC 18 July 2022 shows phase difference in the propagation of a mesoscale convective system evident among 33h HREF member forecasts. Some members have intense rainfall along the border between Missouri and Iowa; the others have a convection positioned south and east of that border. Though the members have disagreements about the location of the convective system, SAM is able to bring all the forecasts to a common central position.

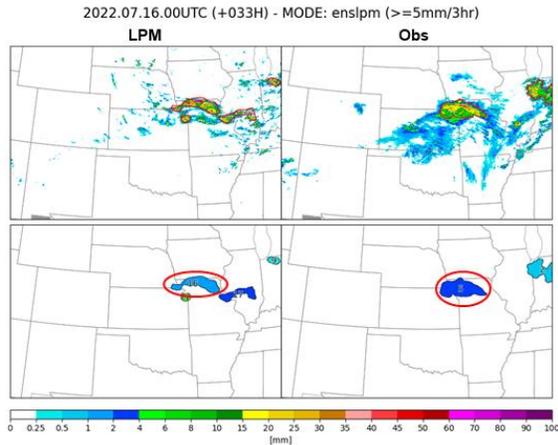


Figure 15. MODE results ($\geq 5\text{mm}/3$) of LPM and Observation for convective system case (33h forecast valid at 21 UTC 18 July 2022)

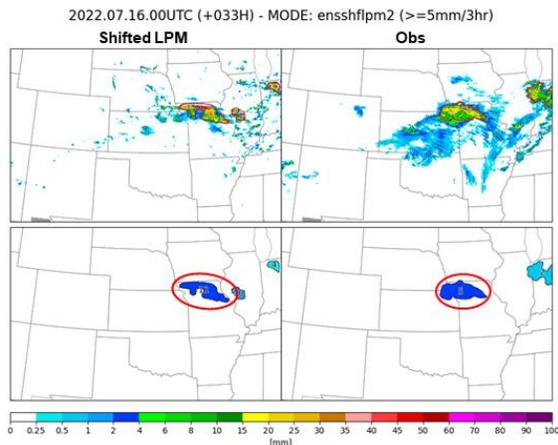


Figure 16. MODE results ($\geq 5\text{mm}/3$) of SAM-LPM and Observation for convective system case (33h forecast valid at 21 UTC 18 July 2022)

	LPM	SAM-LPM
The interest of MODE	0.97723	0.98144
Area ratio	0.67505	0.80354
Intersection area (km^2)	702	877

Table 2. The scores of MODE metrics ($\geq 5\text{mm}/3$) of LPM and SAM-LPM for the convective system case (33h forecast valid at 21 UTC 18 July 2022)

The standard LPM in Fig. 15. shows separated objects, but SAM-LPM in Fig. 16. has a more compact object in Missouri. As a result, SAM-LPM has a precipitation feature ($> 5\text{ mm}/3\text{h}$) better matching the observed event, and the MODE scores (Table 2.) of SAM-LPM are improved for area ratio and intersection area leading to better overall Interest score.

6. SUMMARY

From the verification results for testing on HREF 3h precipitation during four weeks of the summer of 2022, the Spatial Aligned Mean (SAM) ensemble consensus technique outperforms the simple ensemble mean, and Spatially aligned LPM (SAM-LPM) also outperforms the standard LPM method. The results show that the spatial alignment technique improves the ensemble consensus in common metrics such as ETS.

Spatially aligned LPM (SAM-LPM) improves the structure of the mean as demonstrated in MODE results while preserving the ensemble forecast maxima, thus seems to be the best candidate for calculating an ensemble consensus for these fields.

7. ACKNOWLEDGMENTS

Access to the HREF individual member forecasts is provided to CAPS by Kent Knofmeier of the National Severe Storms Laboratory.

Verification is performed using MET toolkit obtained from the Developmental Testbed Center, a collaborative effort among NOAA, JCSDA, National Science Foundation and contributors in the academic community.

Partial support for this work is from the NOAA Testbed program, under grants NA19OAR4590141

and NA22OAR4590522, and a Korean Government Long-Term Fellowship for Overseas Studies awarded to Mr. Lee.

8. REFERENCES

Bullock, R., T. Fowler, and B. Brown, 2016: Method for Object-Based Diagnostic Evaluation. NCAR Technical Note NCAR/TN-532+STR, 66 pp.

Clark, A. J., 2017. Generation of ensemble mean precipitation forecasts from convection-allowing ensembles. *Weather and Forecasting*, 32(4), 1569–1583. <https://doi.org/10.1175/WAF-D-16-0199.1>

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.

Jensen, T., J. Prestopnik, H. Soh, L. Goodrich, B. Brown, R. Bullock, J. Halley Gotway, K. Newman,

J. Opatz, 2023: The MET Version 11.1.0 User's Guide. Developmental Testbed Center. Available at: <https://github.com/dtcenter/MET/releases>

Nelson, B. R., O. P. Prat, D. J. Seo, and E. Habib, 2016: Assessment and implications of NCEP stage IV quantitative precipitation estimates for product intercomparisons. *Weather and Forecasting*, 31, 371–394, <https://doi.org/10.1175/WAF-D-14-00112.1>.

Snook, N., F. Kong, A. Clark, B. Roberts, K. A. Brewster, and M. Xue, 2020: Comparison and Verification of Pointwise and Patchwise Localized Probability Matched Mean Algorithms for Convective Ensemble Forecasting. *Geophysical Research Letters*, 47, <https://doi.org/2020GL087839>.