A Local Verification Study of Convective Allowing Model Performance During Convective Events in Eastern New York and Western New England

Part II: Environmental Breakdown

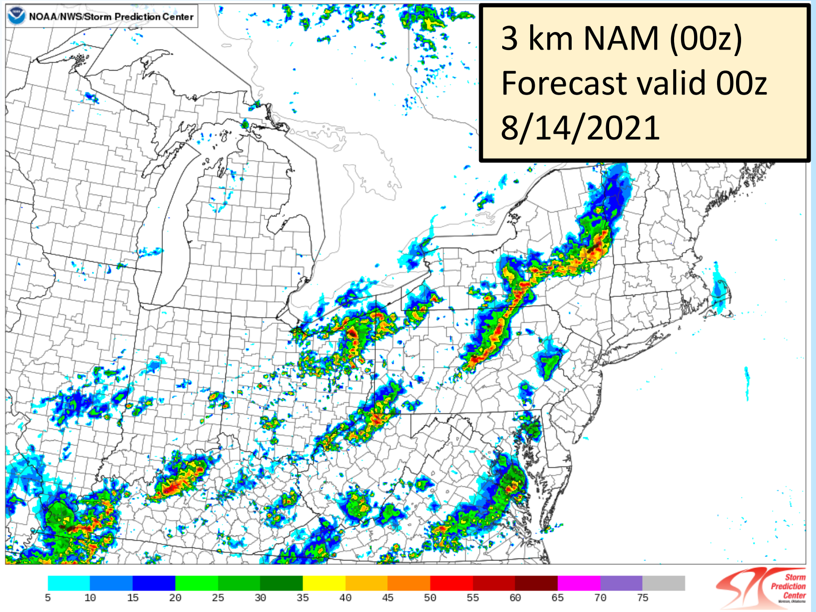
Michael Main and Michael Evans

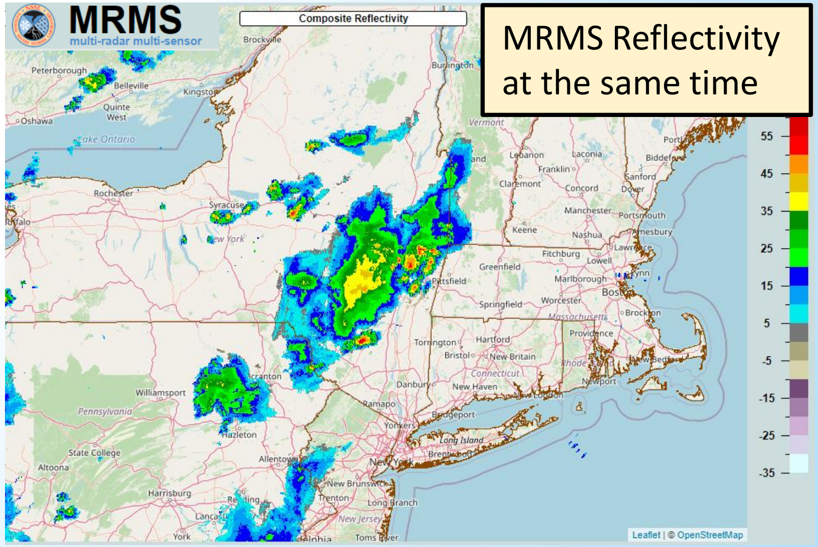
NOAA/NWS Weather Forecast Office, Albany, NY

Motivation (1/2)

- Part I of this study examined the performance of the HRRR and 3 km NAM in forecasting the coverage, timing, and evolution of convection.

- Results showed that:
    - Coverage was overdone for non-severe cases and underdone for severe cases.
    - Both models had a slow bias, especially for severe events.
    - Neither model was superior to the other.
    - There was little improvement from the 00z to 12z model runs for severe cases.

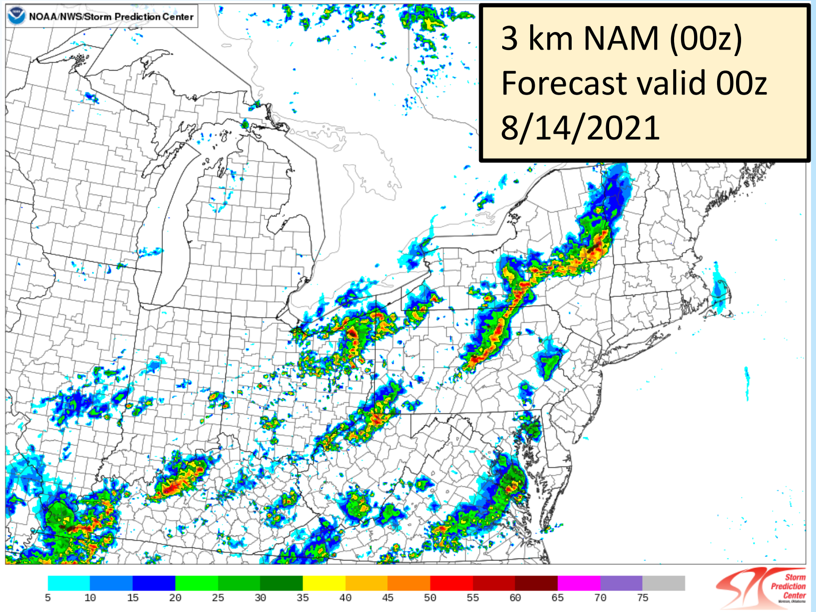


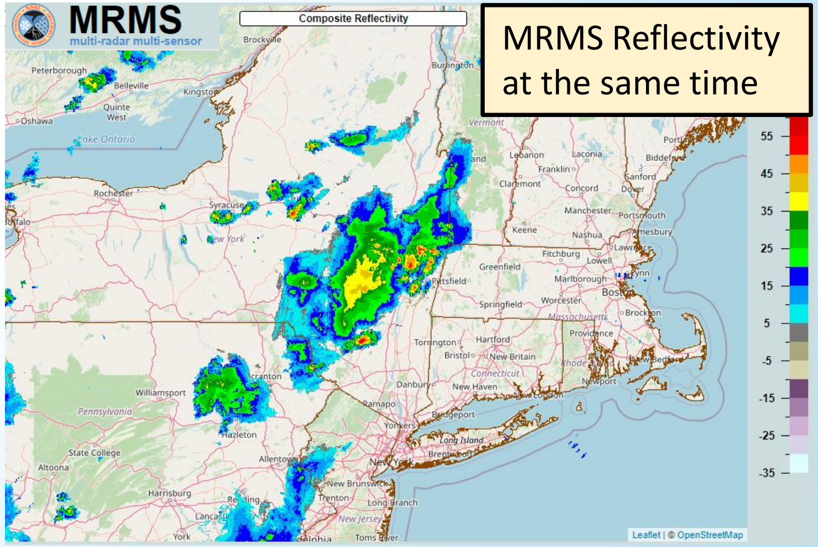


Motivation (2/2)

- These results are for all convective environments.

- Do the results change for certain types of environments?

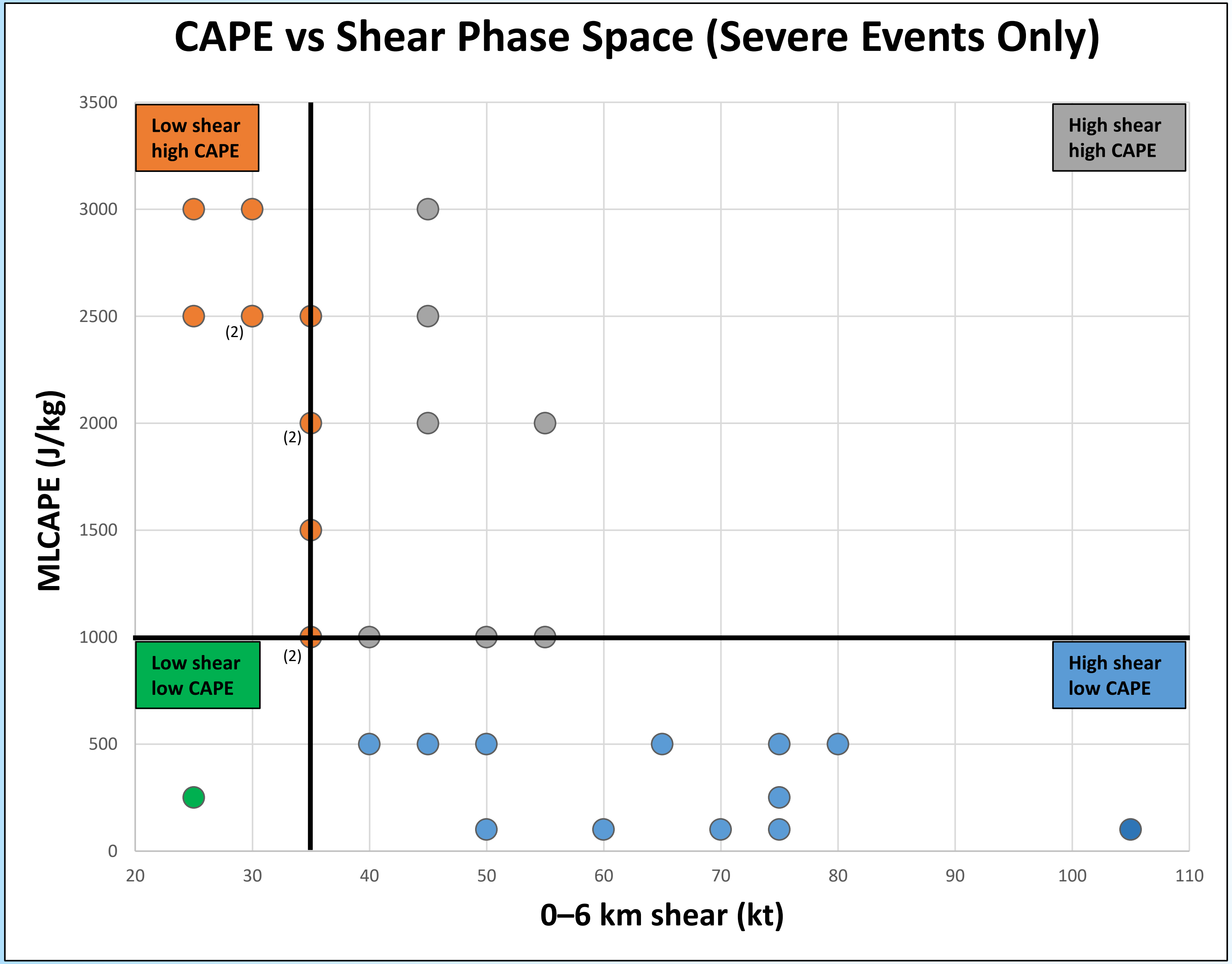




Methodology (1/3)

- The same 32 severe events and corresponding forecaster evaluations from Part I were used here. Null events were not included.
- For each event, the maximum MLCAPE and 0–6-km shear values were recorded from the SPC mesoanalysis archive.
- Convective environments were broken down into 4 categories:
  - High shear high CAPE (HSHC)
  - High shear low CAPE (HSLC)
  - Low shear high CAPE (LSHC)
  - Low shear low CAPE (LSHC)
- Environments were classified using both the Sherburn et al. (2016) criteria and the Vaughan et al. (2017) criteria.

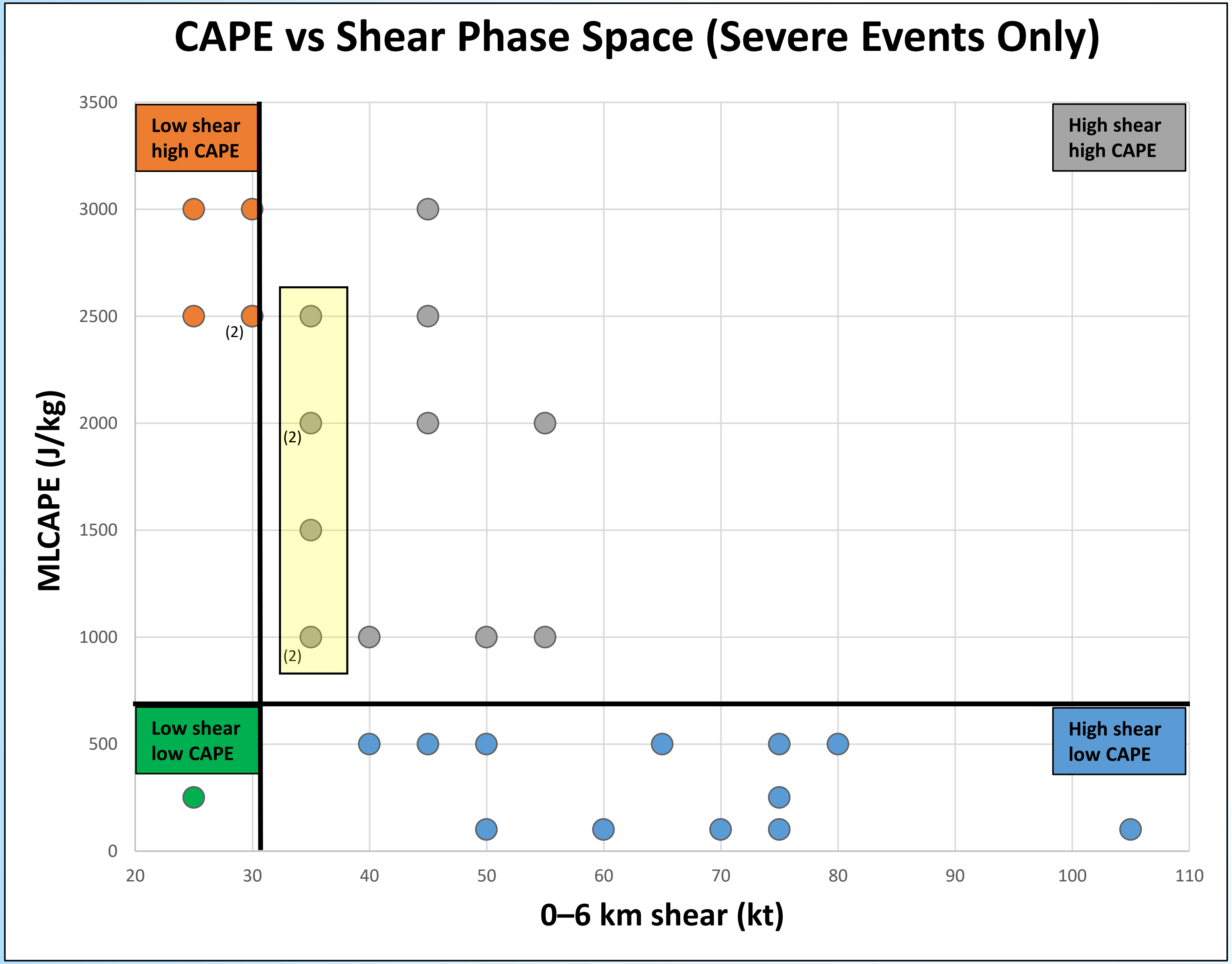
Methodology (2/3)



Sherburn et al. Criteria

Category	High shear low CAPE	High shear high CAPE	Low shear high CAPE	Low shear low CAPE
# of Events	12	7	11	1
Cape Limit (J/kg)	< 1000	≥ 1000	≥ 1000	< 1000
Shear Limit (kt)	> 35	> 35	≤ 35	≤ 35
Severe Reports	68	138	235	0
Flash Flood Reports	71	22	8	2

Methodology (3/3)

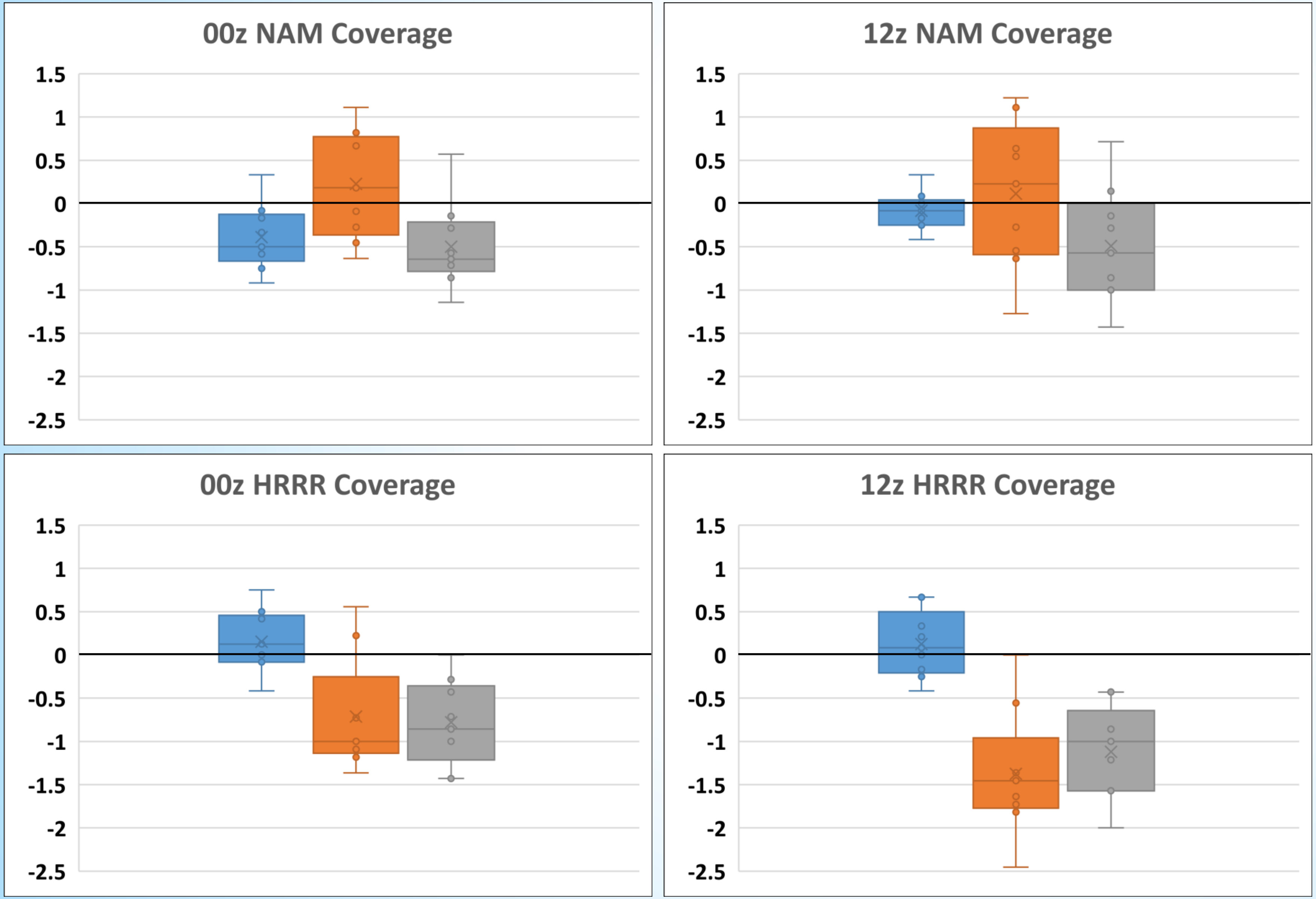


Vaughan et al. Criteria

Category	High shear low CAPE	High shear high CAPE	Low shear high CAPE	Low shear low CAPE
# of Events	12	13	5	1
Cape Limit (J/kg)	< 662	≥ 662	≥ 662	< 662
Shear Limit (kt)	> 31	> 31	≤ 31	≤ 31
Severe Reports	68	324	79	0
Flash Flood Reports	71	26	4	2

- **Changing the criteria shifts some cases from the LSHC category to HSHC**

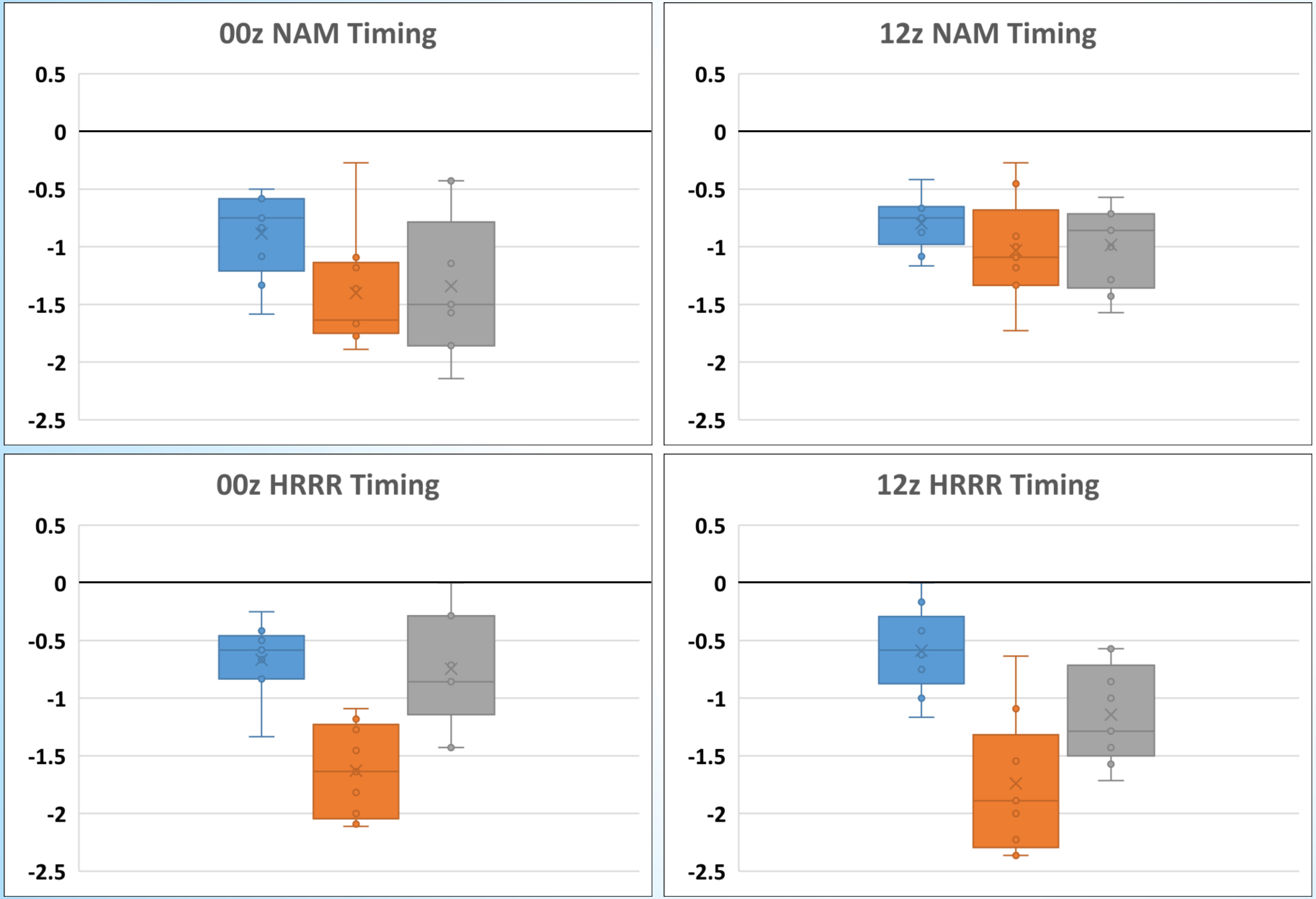
Results: Coverage (Sherburn et al. Criteria)



- NAM Coverage shows a small high bias for LSHC events
- HRRR Coverage is too low for LSHC and HSHC events
- NAM coverage overall is better than the HRRR
- There is little improvement in coverage from the 00z to 12z runs, especially with the HRRR

Legend
- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

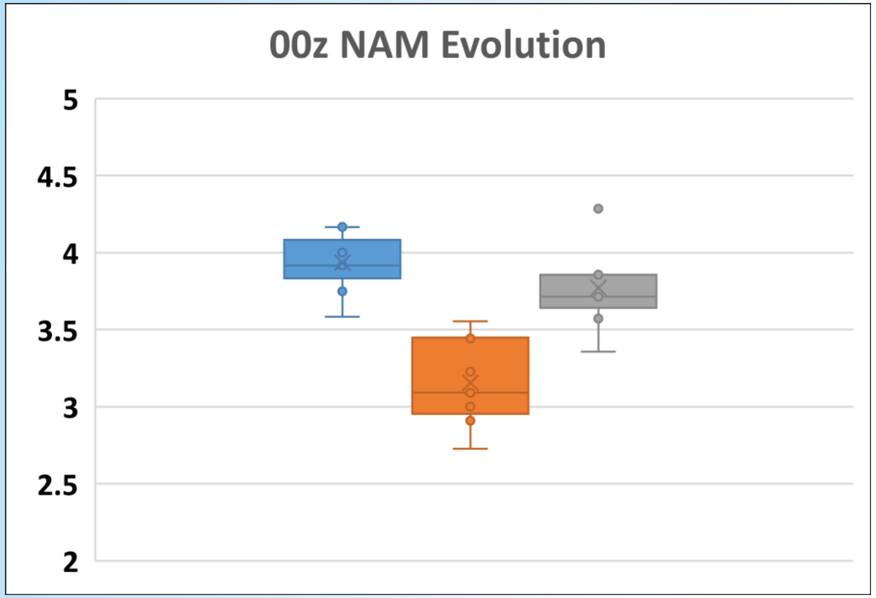
Results: Timing (Sherburn et al. Criteria)

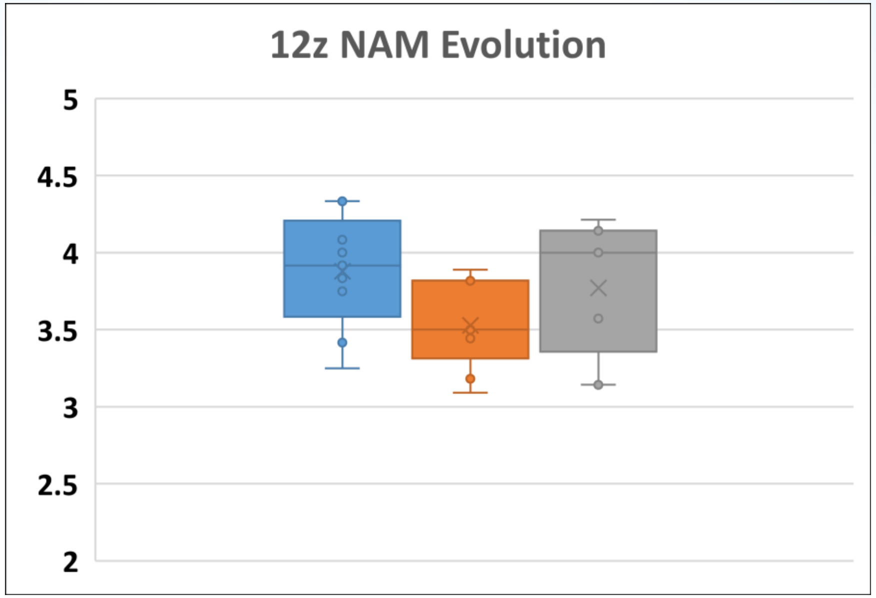


- CAMs are too slow with timing across all environments
- The HRRR slow bias is worse for LSHC events
- The NAM shows improvement from 00z to 12z while the HRRR does not
- High shear low CAPE events have the smallest slow bias for both models

Legend
- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

Results: Evolution (Sherburn et al. Criteria)



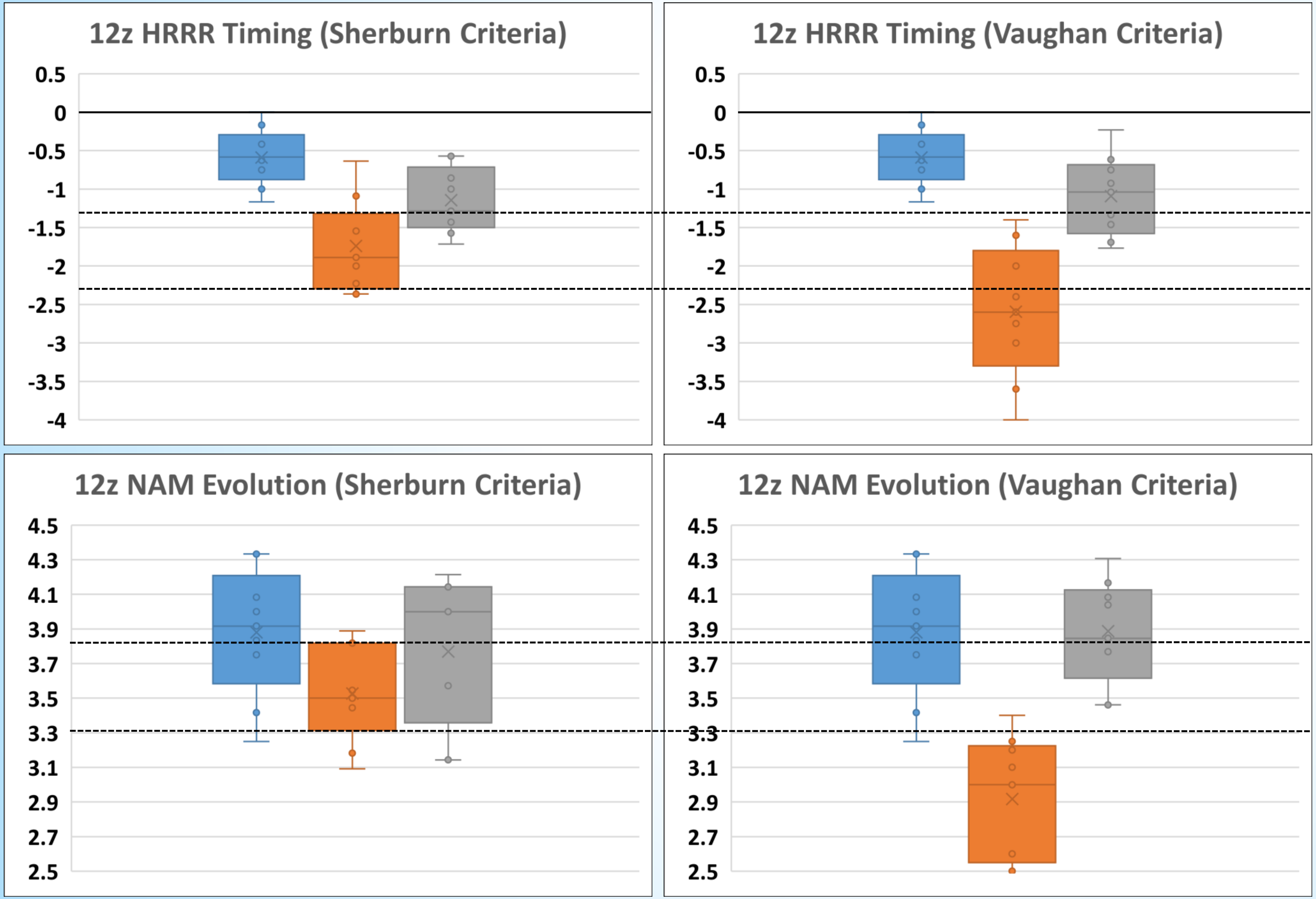


- **HSLC events received the highest evolution scores for both models**
- **LSHC events have the lowest evolution scores for both models**
- **The NAM shows improvement from 00z to 12z for LSHC cases while the HRRR actually performs worse at 12z vs 00z**

Legend

- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

Results: Sherburn vs Vaughan Criteria



- Changing criteria separated out the lowest shear events
- Changing the thresholds used to differentiate convective environments did not change the model biases noted above
- However, it did highlight these model biases
- This was true of both models

Legend
- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

Methodology: Forcing Strength

Category	Strongly Forced	Moderate Forcing	Weakly Forced
# of Events	7	11	14
500 mb Height Falls (m in 12 hr)	≥ 50	< 50 and > 0	≤ 0
Average CAPE (J/kg)	365	1560	1560
Average 0–6 km bulk shear	74	42	40
Severe Reports	56	268	137
Flash Flood Reports	0	69	34

- Same 32 severe events included
- Does not account for low-level forcing (i.e. cold fronts, convergence boundaries…)
- Strongly forced events are all of the HSLC variety
- Very similar environments for moderate and weak forcing
- Moderately forced events have the greatest number of reports per event

Methodology: Forcing Strength

Category	Strongly Forced	Moderate Forcing	Weakly Forced
# of Events	7	11	14
500 mb Height Falls (m in 12 hr)	≥ 50	< 50 and > 0	≤ 0
Average CAPE (J/kg)	365	1560	1560
Average 0–6 km bulk shear	74	42	40
Severe Reports	56	268	137
Flash Flood Reports	0	69	34

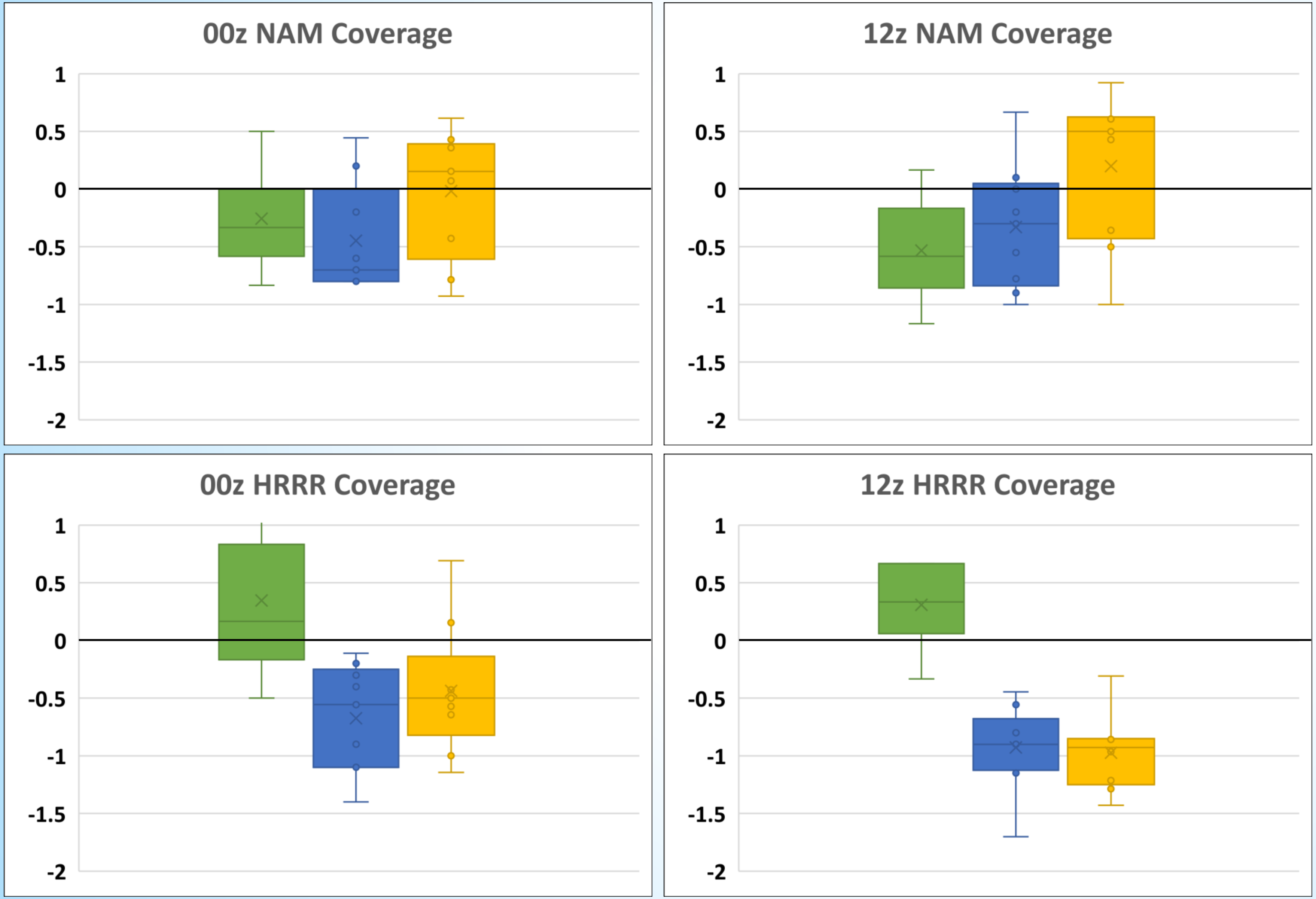
- Same 32 severe events included
- Does not account for low-level forcing (i.e. cold fronts, convergence boundaries…)
- Strongly forced events are all of the HSLC variety
- Very similar environments for moderate and weak forcing
- Moderately forced events have the greatest number of reports per event

Methodology: Forcing Strength

Category	Strongly Forced	Moderate Forcing	Weakly Forced
# of Events	7	11	14
500 mb Height Falls (m in 12 hr)	≥ 50	< 50 and > 0	≤ 0
Average CAPE (J/kg)	365	1560	1560
Average 0–6 km bulk shear	74	42	40
Severe Reports	56	268	137
Flash Flood Reports	0	69	34

- Same 32 severe events included
- Does not account for low-level forcing (i.e. cold fronts, convergence boundaries…)
- Strongly forced events are all of the HSLC variety
- Very similar environments for moderate and weak forcing
- Moderately forced events have the greatest number of reports per event

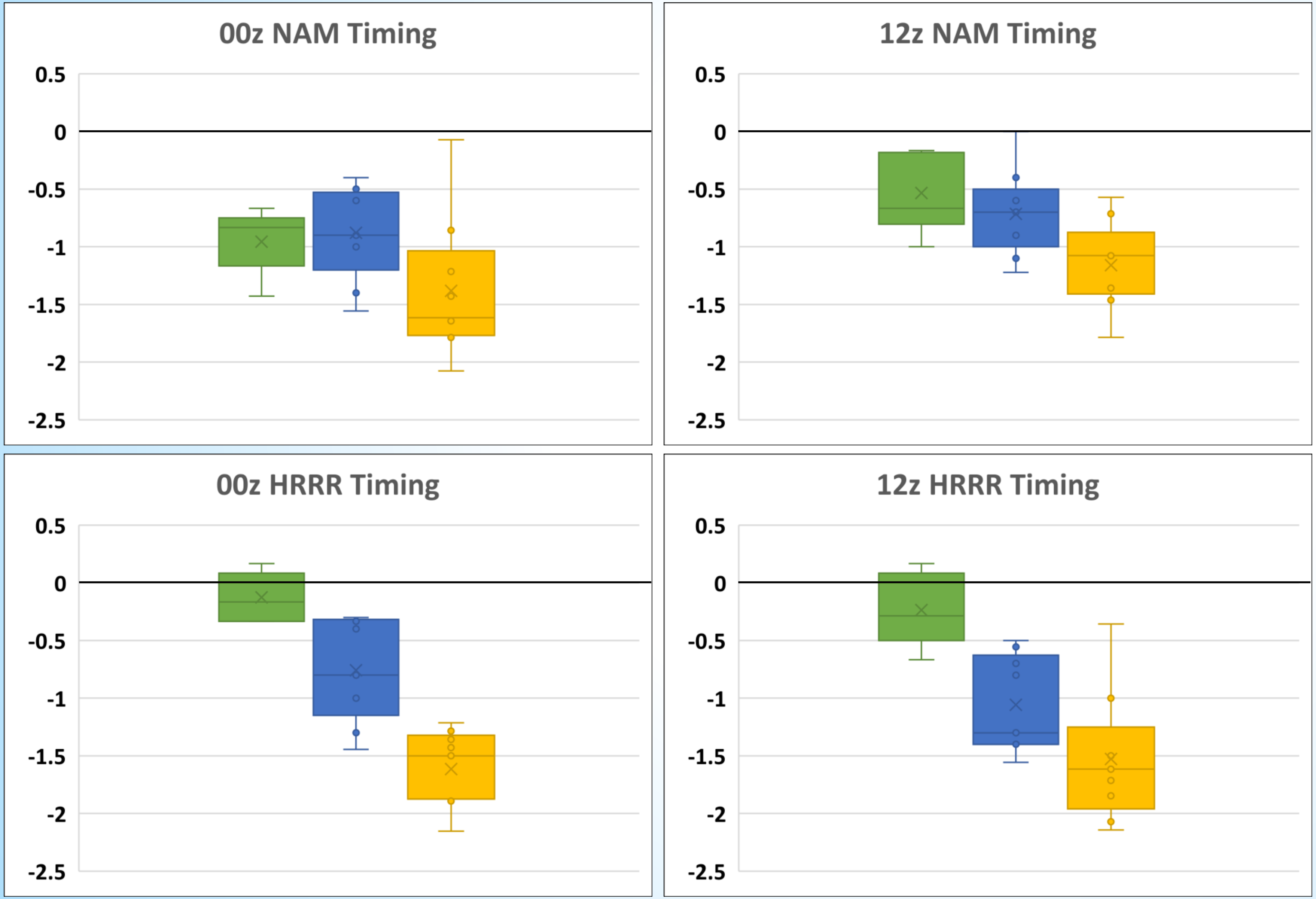
Results: Coverage (Forcing Magnitude Criteria)



- The NAM has a low bias for strong and moderate forcing but a high bias for weak forcing
- The HRRR over forecasts coverage for strongly forced events
- The HRRR under forecasts coverage for moderate and weak forcing, and is worse at 12z vs 00z

Legend
- Strongly Forced
- Moderately Forced
- Weakly Forced

Results: Timing (Forcing Magnitude Criteria)

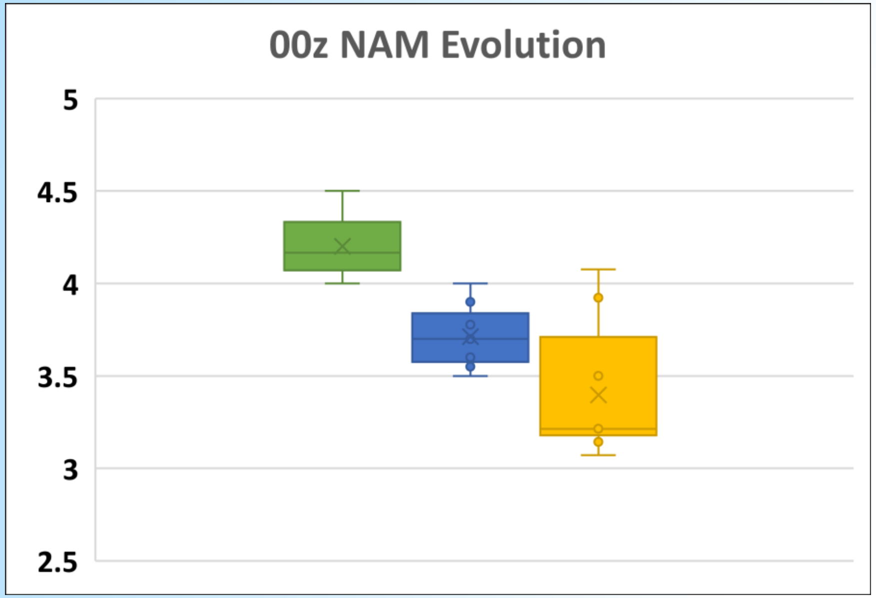


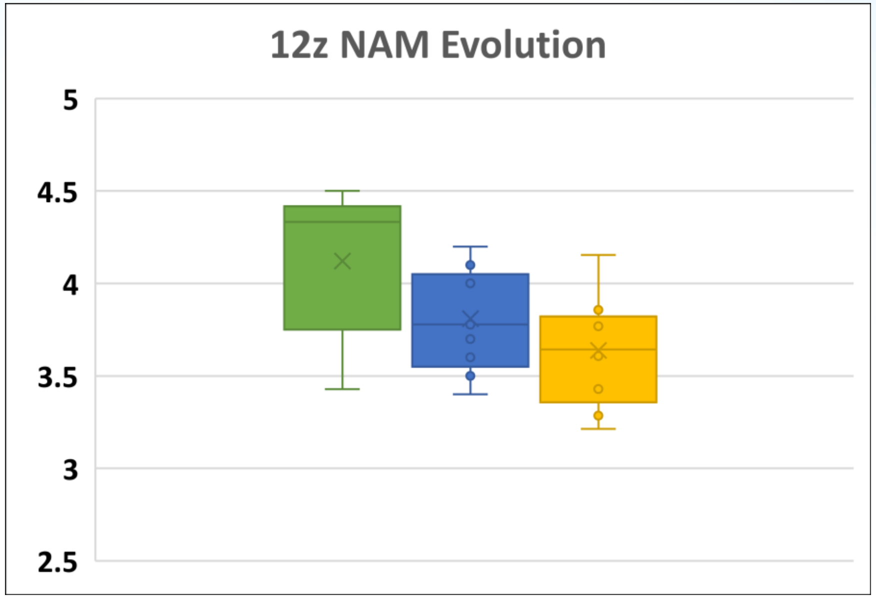
- CAMs are too slow for all events with timing
- The slow bias increases as forcing strength decreases
- HRRR timing is better for strongly forced events and worse for weakly forced events compared to the NAM
- The NAM improved from 00z to 12z but the HRRR did not

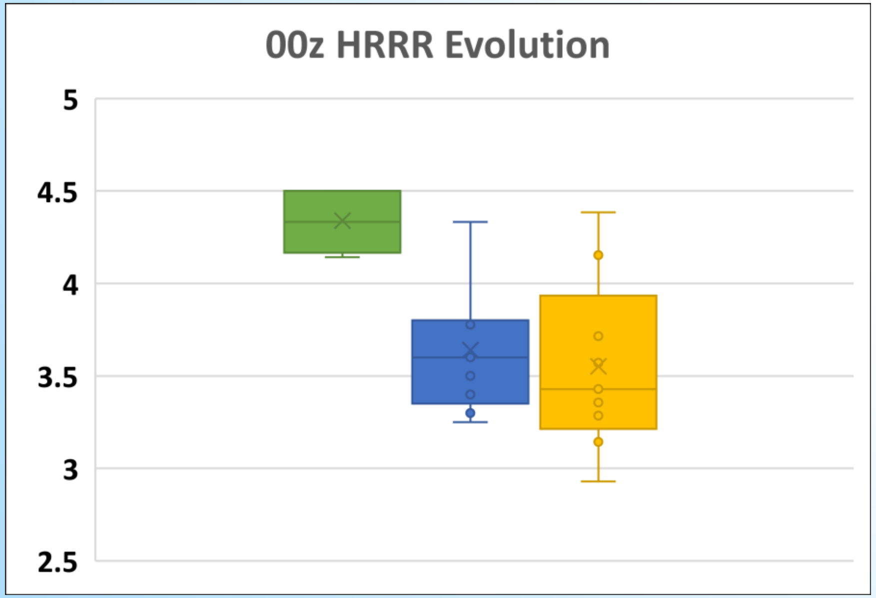
Legend

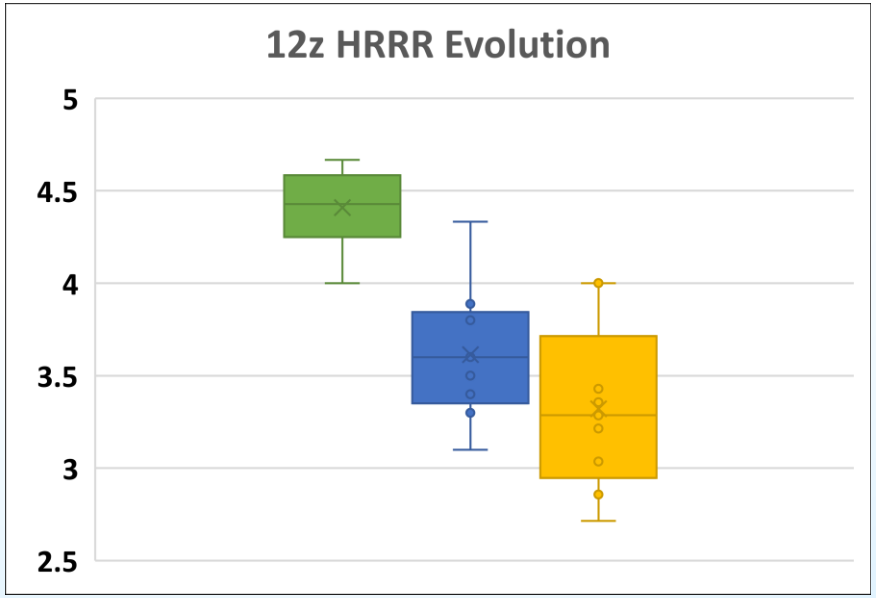
- Strongly Forced
- Moderately Forced
- Weakly Forced

Results: Evolution (Forcing Magnitude Criteria)









- **Evolution scores decrease for moderate and weakly forced events for both models**
- **The NAM scored better than the HRRR for moderately forced events while the HRRR scored better for strong forcing**
- **The HRRR showed slight improvement from 00z to 12z for strong forcing only**

Legend
- **Strongly Forced**
- **Moderately Forced**
- **Weakly Forced**

Conclusions

- The slow bias in both the HRRR and 3 km NAM is evident for all convective environments, but is worse when shear and upper-level forcing area weaker.
- Model simulated convective coverage is overdone by the 3 km NAM and underdone by the HRRR in convective environments with weak shear and weak upper-level forcing.
- Model-simulated convective evolution received the highest scores in convective environments with high shear and strong upper-level forcing.
- Changing the criteria for the convective environment magnified the model biases but did not change the overall results.

Future Work

- Examine model biases in temperature, dew point, CAPE… to see if there is any correlation between biases in these fields and the biases noted here.
- Perform the breakdown by forecaster scores instead of by the type of convective environment
  - May give us thresholds where models can be trusted more
  - Identify physical features/processes that may cause models to struggle
- Expand the analysis to the non-severe cases
- Add additional cases to increase sample size

References and Acknowledgements

NCEI, 2022: Storm Events database. NOAA/National Centers for Environmental Information, https://www.ncdc.noaa.gov/stormevents/.

Sherburn, K. D., M. D. Parker, J. R. King, and G. M. Lackmann, 2016: Composite environments of severe and nonsevere high-shear, low-CAPE convective events. *Wea. Forecasting*, **31**, 1899–1927, https://doi.org/10.1175/WAF-D-16-0086.1.

Vaughan, M. T., Tang, B. H. & Bosart, L. F. Climatology and analysis of high-impact, low predictive skill severe weather events in the northeast United States. *Wea. Forecast*. **32**, 1903–1919 (2017).

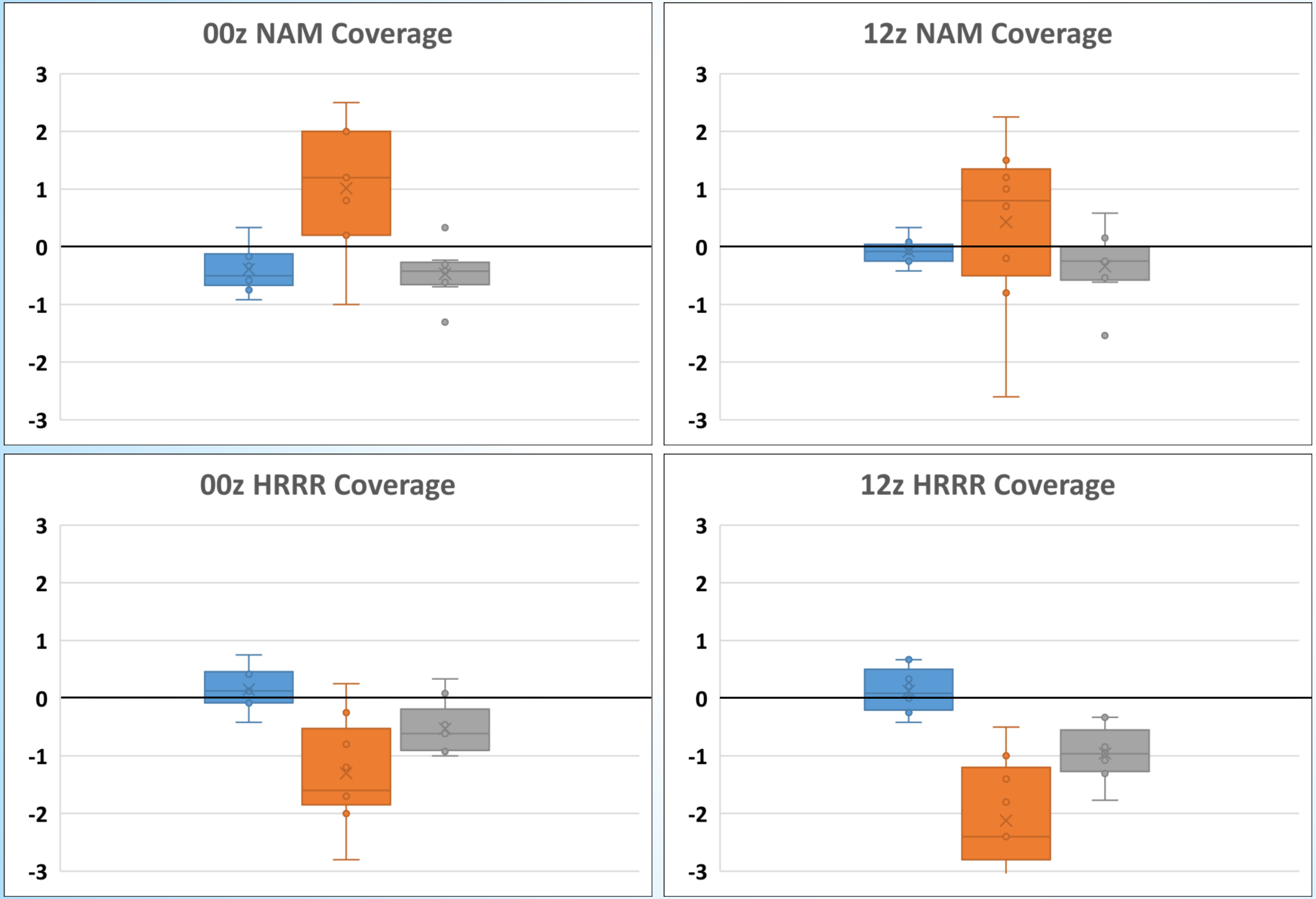
This study was lead by Mike Evans from NWS Albany. This study was made possible through use of data hosted in the SPC HREF archive, the SPC Mesoanalysis archive, and the MRMS radar archive. Also, thank you to all of the forecasters and UAlbany student interns who took part in the model evaluation portion of this study. Finally, thank you to Lee Picard for completing the environmental analysis for the 32 severe cases.

Conclusions

- The slow bias in both the HRRR and 3 km NAM is evident for all convective environments, but is worse when shear and upper-level forcing area weaker.
- Model simulated convective coverage is overdone by the 3 km NAM and underdone by the HRRR in convective environments with weak shear and weak upper-level forcing.
- Model-simulated convective evolution received the highest scores in convective environments with high shear and strong upper-level forcing.
- Changing the criteria for the convective environment magnified the model biases but did not change the overall results.

Supplemental Slides

Results: Coverage (Vaughan et al. Criteria)

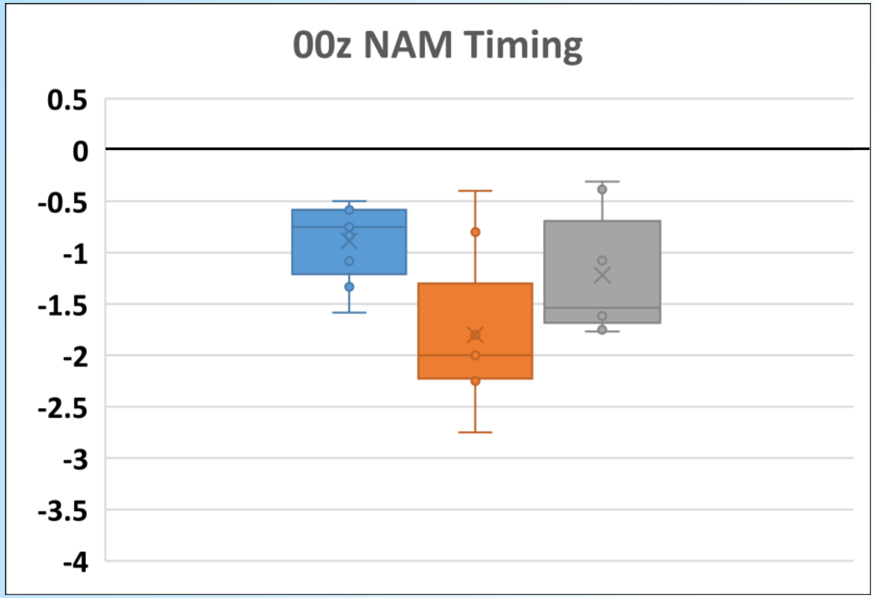


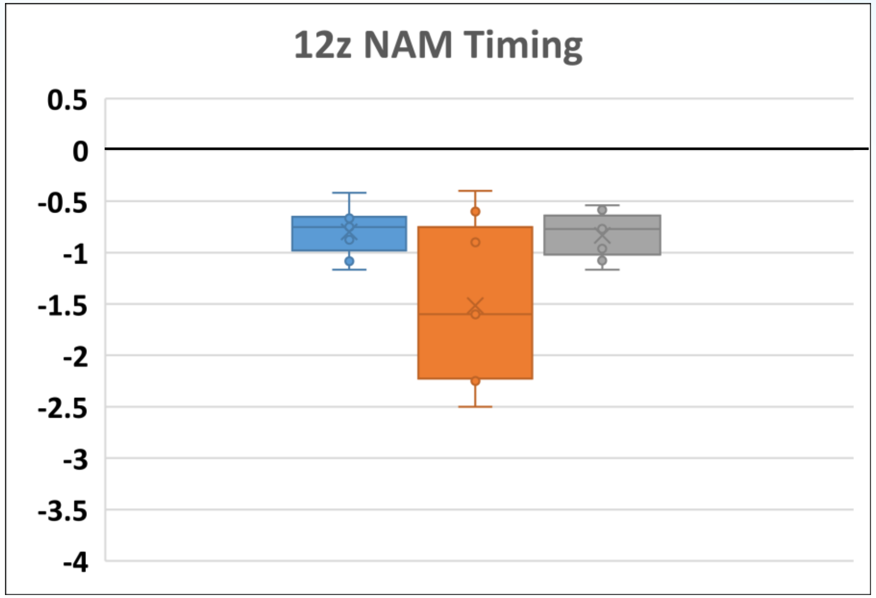
- HRRR Coverage is too low for LSHC events while the NAM is too high
- The HRRR does better with HSLC events vs HSHC while the NAM is similar for both types
- The NAM shows more improvement from 00z to 12z than the HRRR

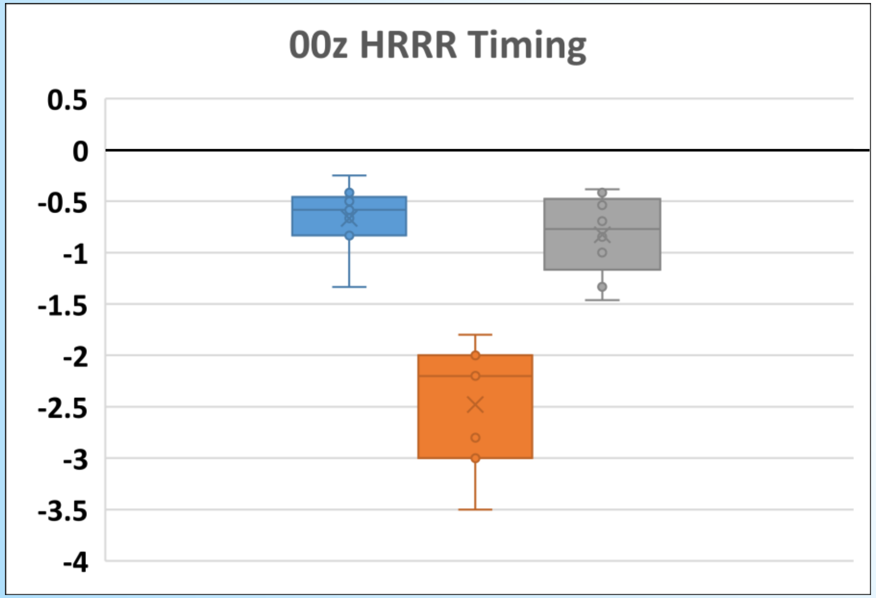
Legend

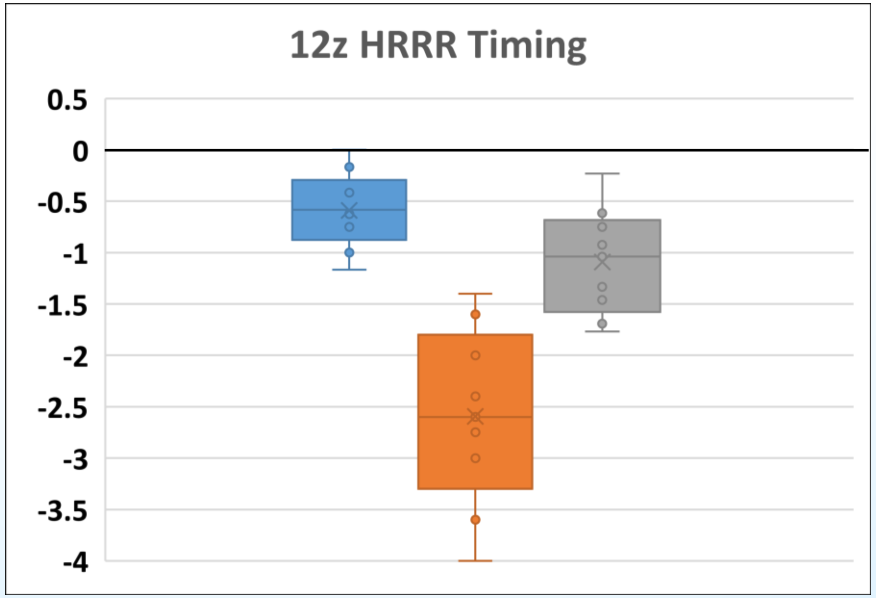
- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

Results: Timing (Vaughan et al. Criteria)







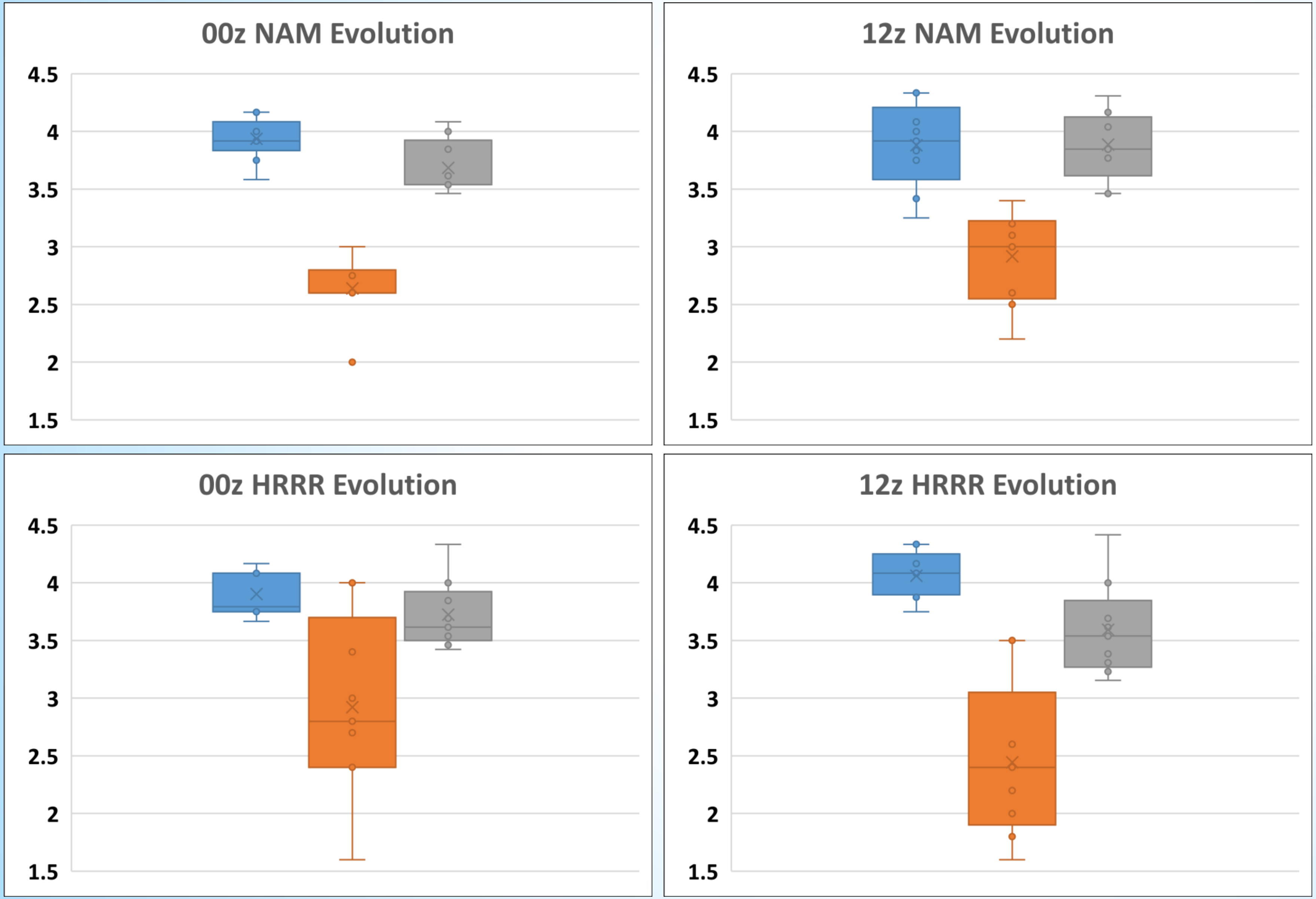


- CAMs are too slow across the board with timing
- Slow bias is worse for LSHC events.
- Slow bias is worse in the HRRR compared to the NAM
- The NAM improved from 00z to 12z for HSHC events but the HRRR was worse at 12z

Legend

- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE

Results: Evolution (Vaughan et al. Criteria)



- HSLC events scored the highest evolution for both CAMS
- LSHC events have the worst evolution for both models
- From 00z to 12z, the HRRR evolution of HSLC events improved, but it decreased for LSHC and HSHC events.
- There was little change in the NAM from 00z to 12z.

Legend

- High Shear Low CAPE
- Low Shear High CAPE
- High Shear High CAPE