Exploring Regional FV3-based Configurations during the 2020 Virtual Spring Forecasting Experiment

BURKELY T. GALLO[1,2], ISRAEL JIRAK[2], ADAM J. CLARK[3], BRETT ROBERTS[1,2,3], KENT KNOPFMEIER[1,3], DAVID HARRISON[1,2], YUNHENG WANG[1,3], JEFF BECK[4], CURTIS ALEXANDER[4], JACOB CARLEY[5], LUCAS HARRIS[6], AND CAROLINE BAIN[7]

[1] COOPERATIVE INSTITUTE FOR MESOSCALE METEOROLOGICAL STUDIES, UNIVERSITY OF OKLAHOMA, NORMAN, OKLAHOMA

[2] NOAA/NWS/NCEP STORM PREDICTION CENTER, NORMAN, OKLAHOMA

[3] NOAA/OAR NATIONAL SEVERE STORMS LABORATORY, NORMAN, OKLAHOMA

[4] NOAA/OAR/ESRL GLOBAL SYSTEMS LABORATORY, BOULDER, COLORADO

[5] NOAA/NWS/NCEP ENVIRONMENTAL MODELLING CENTER, COLLEGE PARK, MARYLAND

[6] NOAA/OAR GEOPHYSICAL FLUID DYNAMICS LABORATORY, PRINCETON, NEW JERSEY

[7] UNITED KINGDOM MET OFFICE, EXETER, UNITED KINGDOM

Virtual

The 2020 Spring Forecasting Experiment

See Adam Clark's presentation from Monday

SFE's main goals are to accelerate R2O↔O2R efforts and foster collaboration between researchers and forecasters

Decided in mid-March to shift the experiment online rather than in NOAA's Hazardous Weather Testbed – kudos to the team for pulling it off!

Independent model evaluations by participants followed by group discussions

Participants rotated between two groups throughout the week

Goal: to maintain momentum in key research areas (SAR-FV3, CAM ensemble development, Warn-on-Forecast, etc.)

Model Evaluations during the 2020 SFE

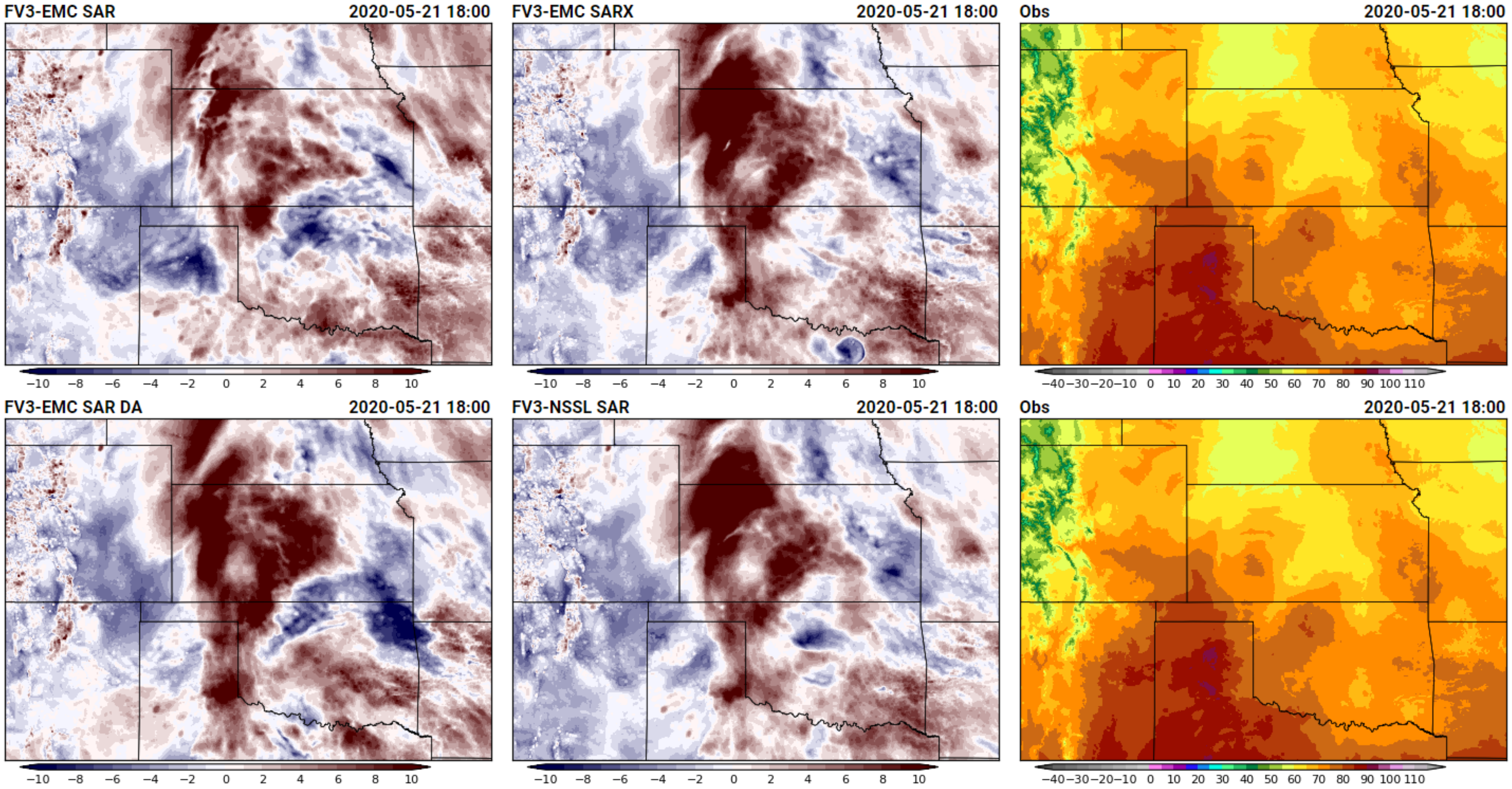
All models initialized at 0000 UTC

Rated composite reflectivity and UH at three separate times:
- 1800 UTC
- 2300 UTC
- 0400 UTC

Rated environmental fields at two separate times:
- 1800 UTC
- 0000 UTC

Used a scale of 1 (Very Poor) to 10 (Very Good)
- Asked about usefulness to forecasters in forecasting severe convection

Results today will summarize 3 comparisons



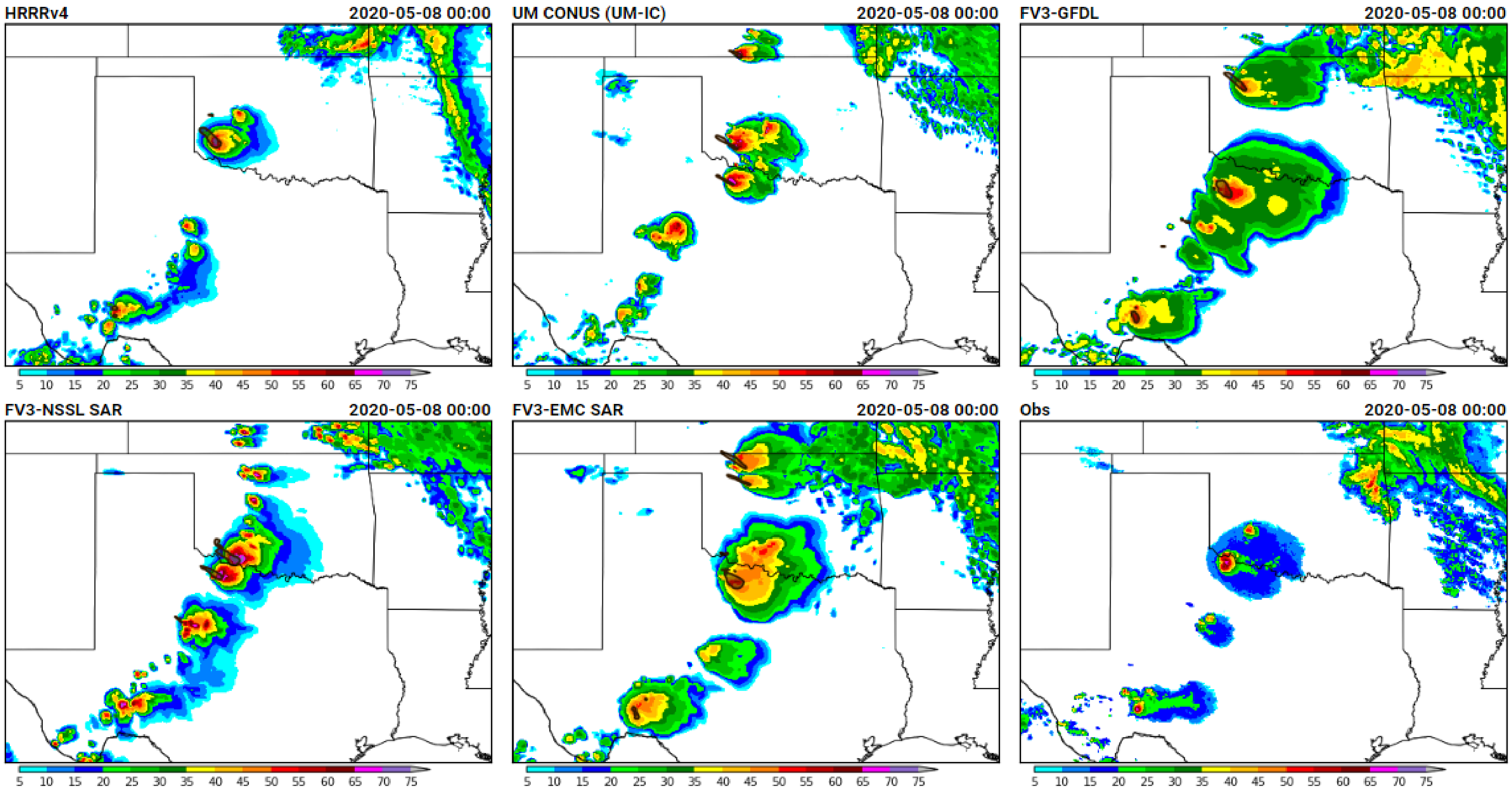
Deterministic Flagships

"Bake-off" between state-of-the-art guidance from different modelling centers

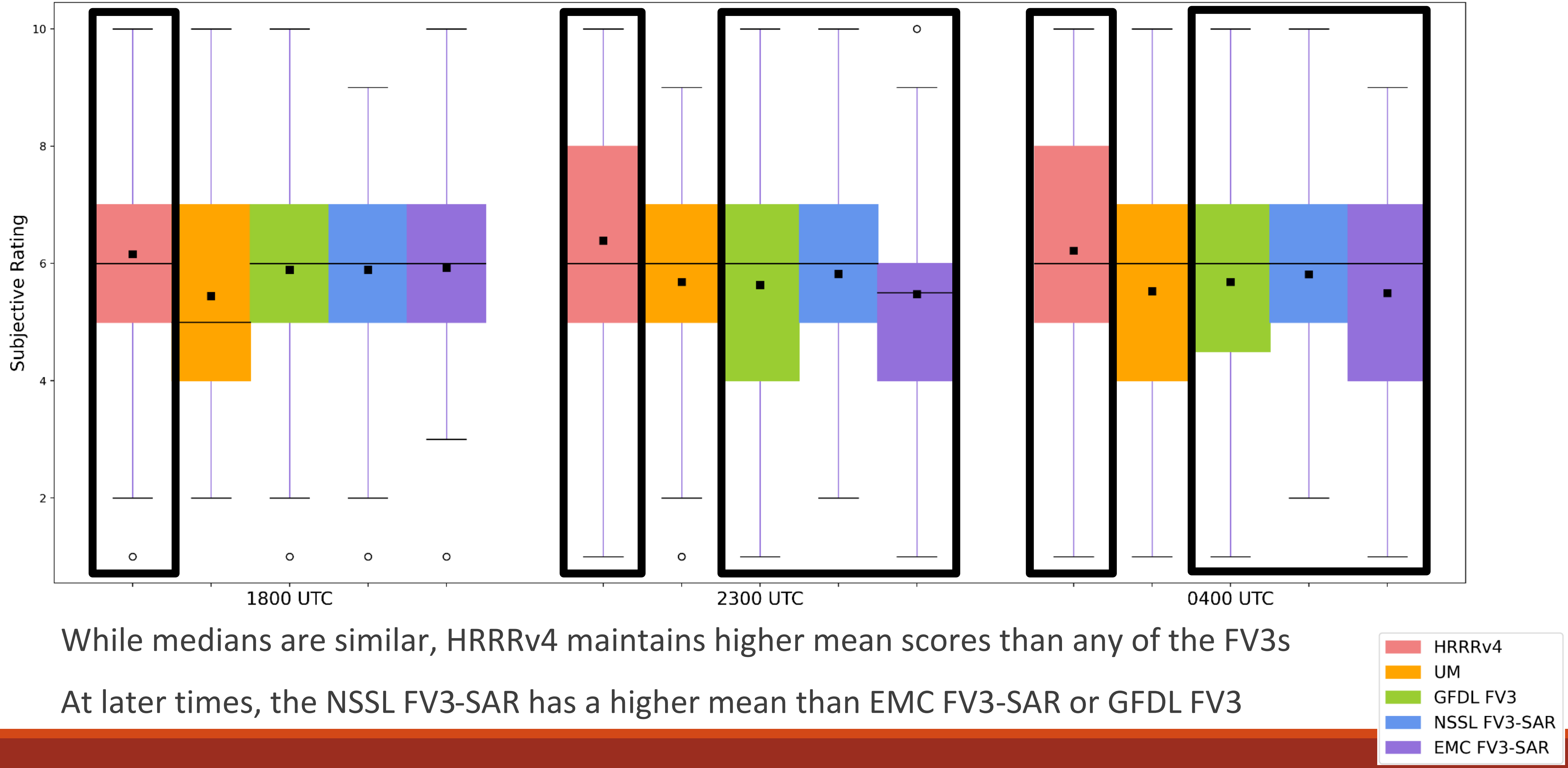
- GSL, Met Office, GFDL, NSSL, and EMC

Comparing guidance to soon-to-be operational guidance in the HRRRv4

See operations plan at: https://hwt.nssl.noaa.gov/sfe/2020/docs/HWT_SFE2020_operations_plan.pdf for more details on configuration specifications



Deterministic Flagships: Composite Reflectivity and UH



While medians are similar, HRRRv4 maintains higher mean scores than any of the FV3s

At later times, the NSSL FV3-SAR has a higher mean than EMC FV3-SAR or GFDL FV3

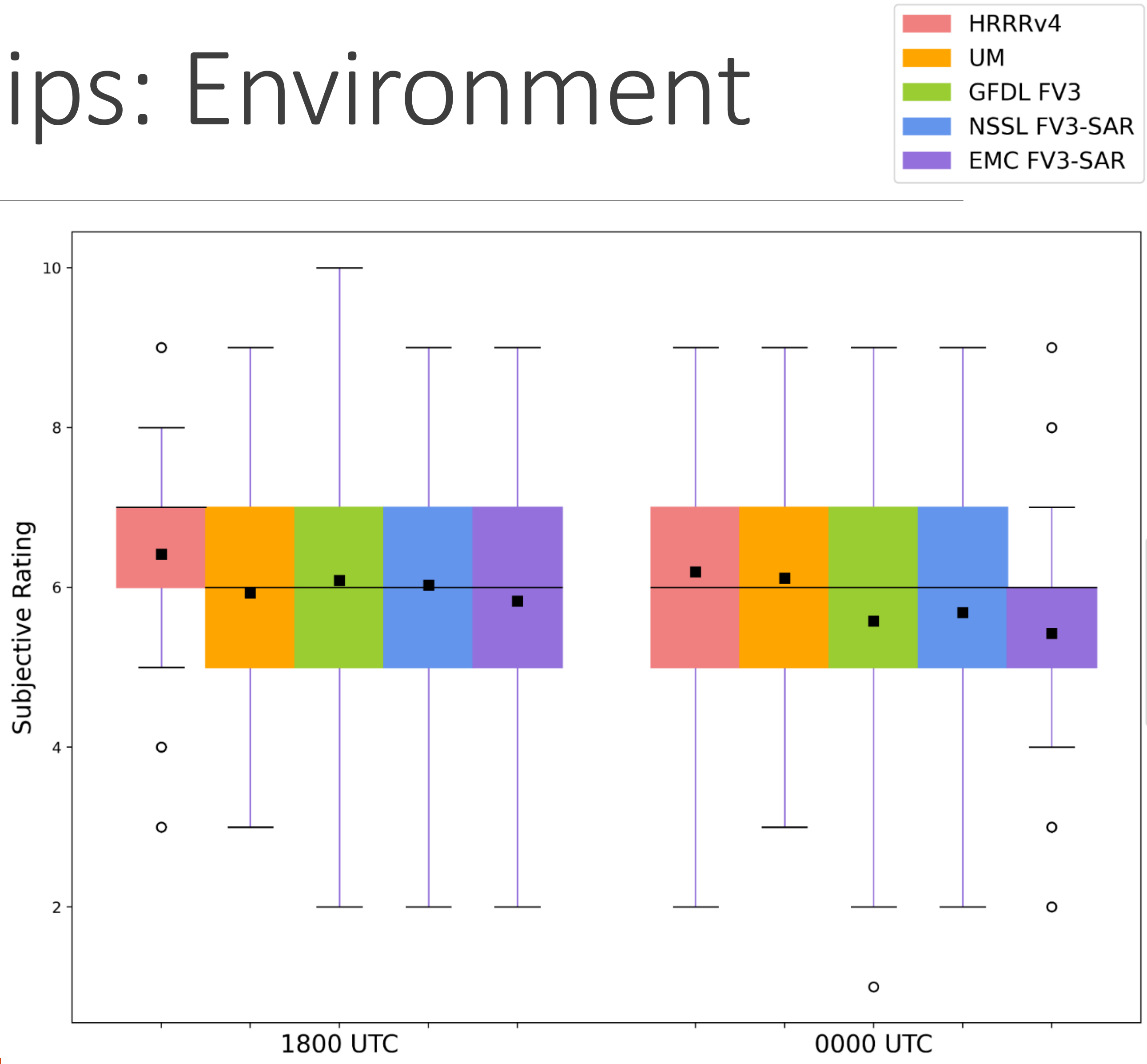
Deterministic Flagships: Environment

Differences in environment ratings were smaller than composite reflectivity and UH but showed similar trends

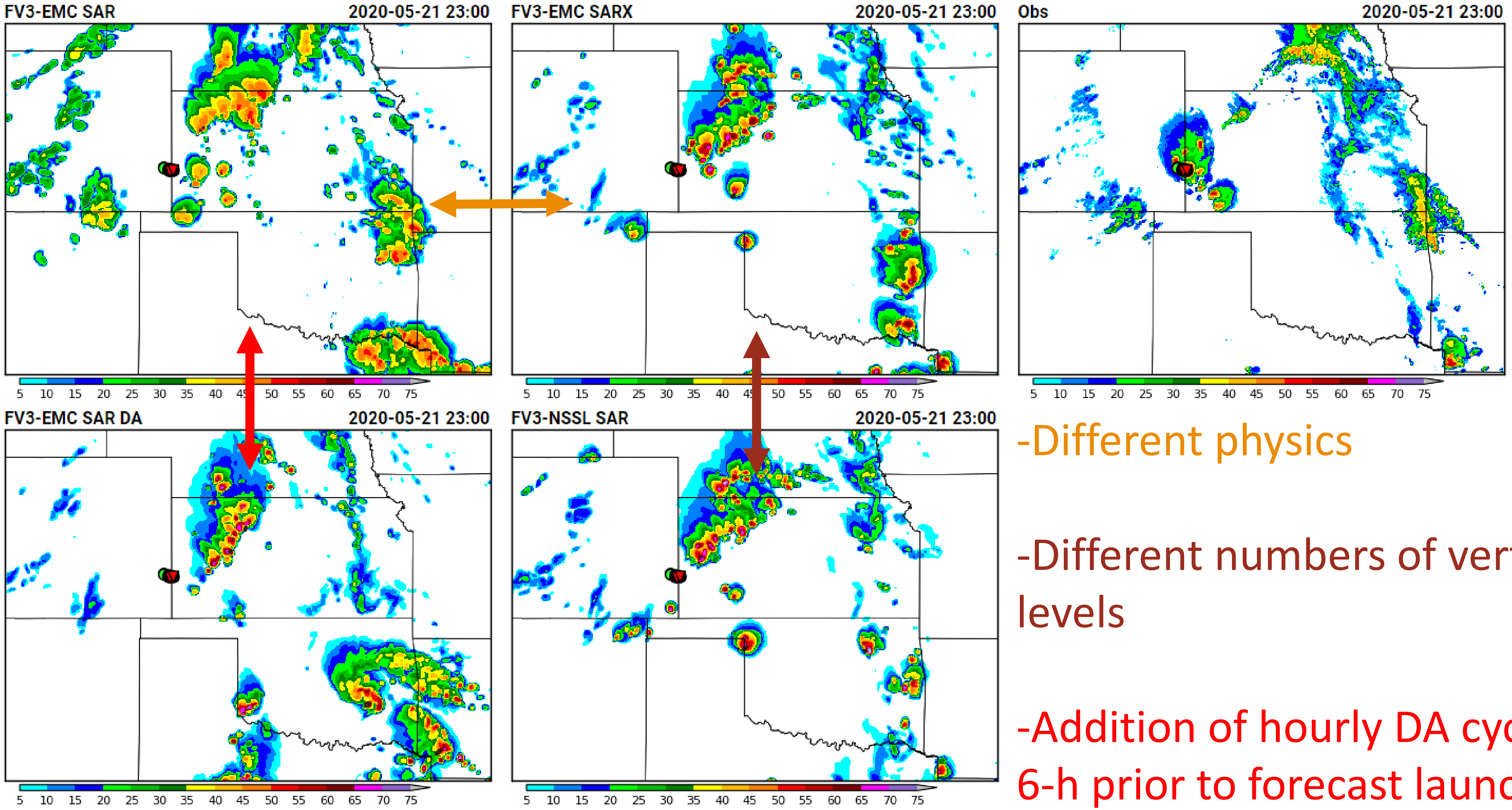
- 0000 UTC environments tended to be dominated by convective processes

Participant comments

- GFDL FV3 cold pools and CAPE often mentioned as doing well
- Overall, cool and moist biases in FV3-based models persist
- Low bias in instability compared to analyses
- Warm and dry bias mentioned in the HRRRv4



A6. FV3-SAR Physics/DA/Vertical Levels



-Different physics

-Different numbers of vertical levels

-Addition of hourly DA cycle over 6-h prior to forecast launch

FV3-SAR Physics/DA/Vertical Levels: Composite Reflectivity and UH

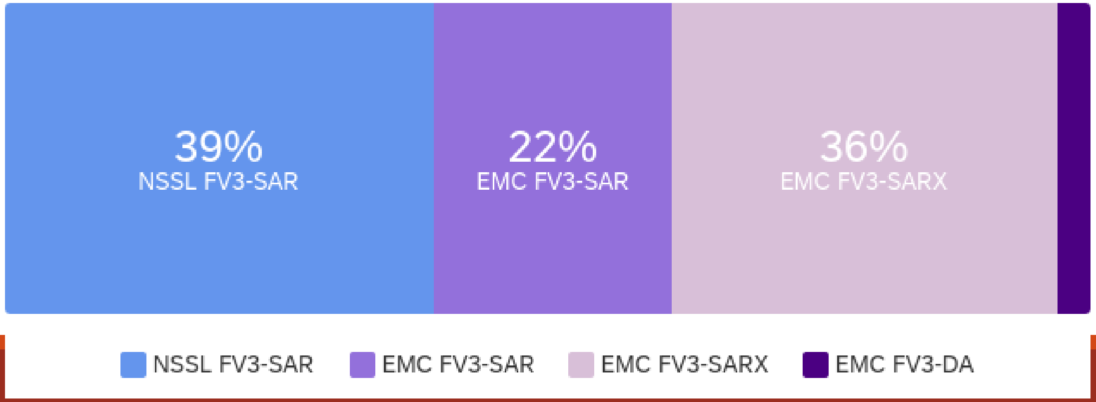
NSSL FV3-SAR
EMC FV3-SAR
EMC FV3-SARX
EMC FV3-SAR DA

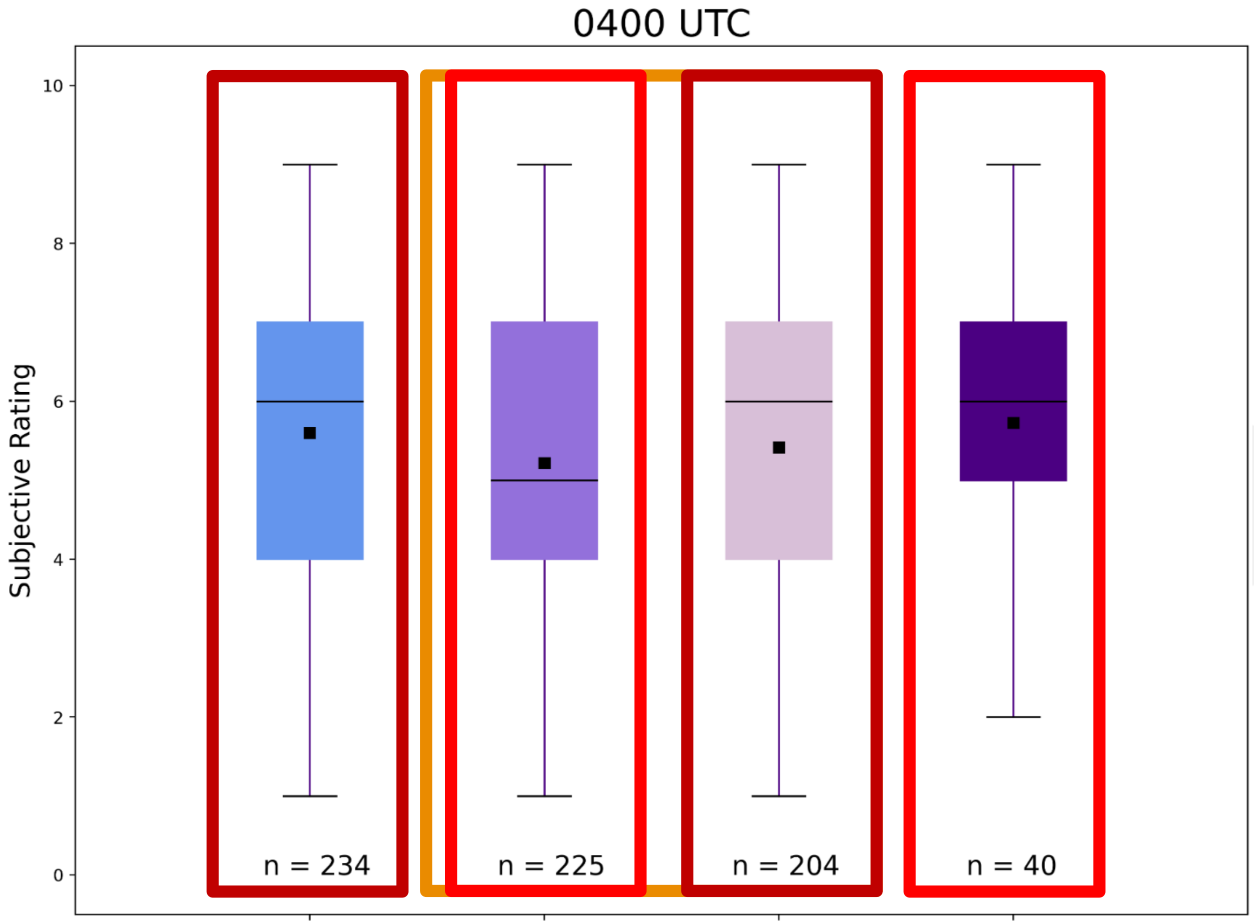
Different physics comparison:

- EMC FV3-SARX physics better than the EMC FV3-SAR throughout
- Physics same as in NSSL FV3-SAR

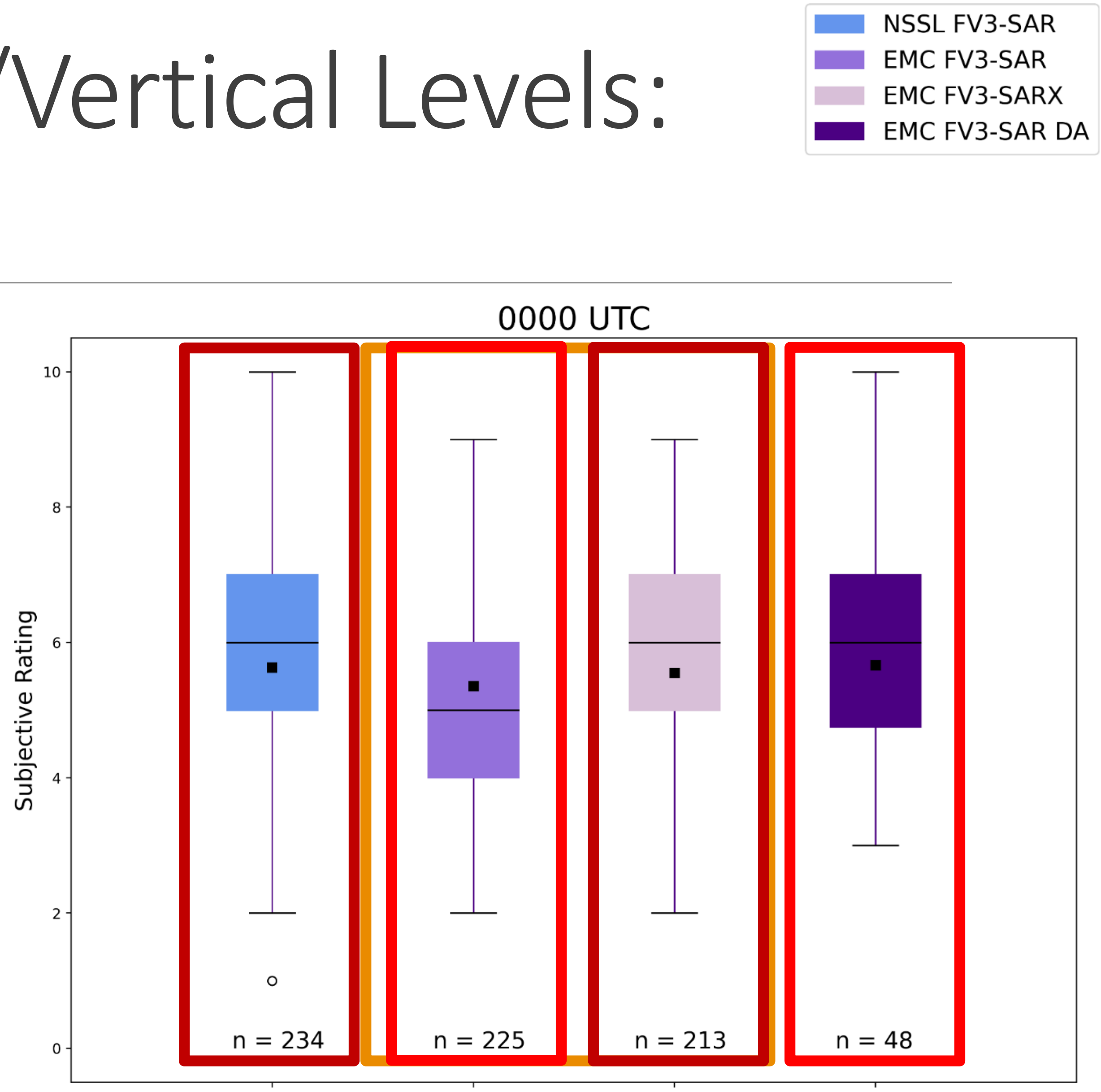
Different vertical levels comparison:

- Nearly identical performance
- EMC FV3-SARX (50 levels) has higher mean earlier; NSSL FV3-SAR (80 vertical levels) has higher mean later





FV3-SAR Physics/DA/Vertical Levels: Environment



Different physics comparison:

- EMC FV3-SARX performs better than EMC FV3-SAR, especially at later times

Different vertical levels comparison:

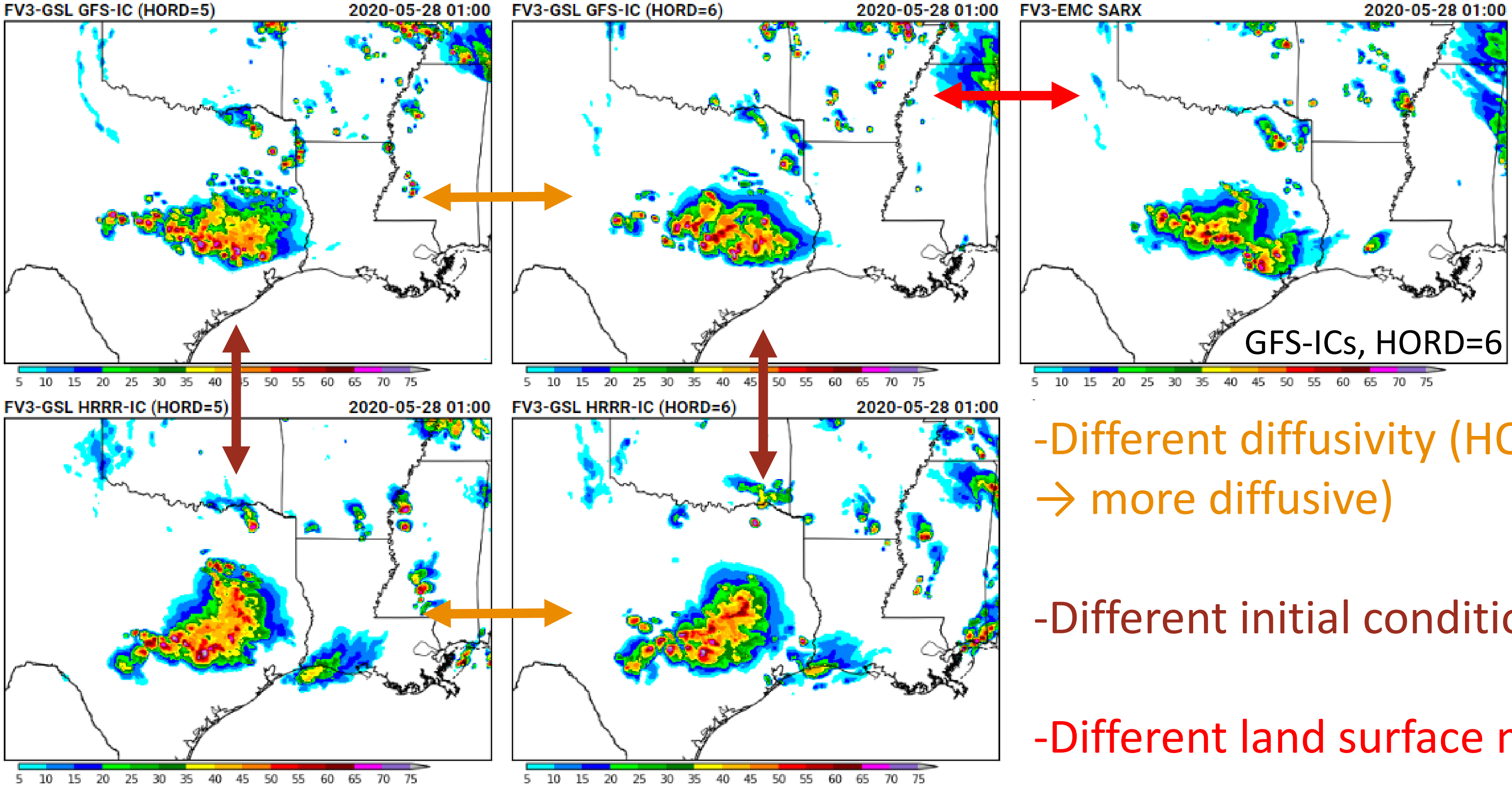
- Very similar performance between NSSL FV3-SAR (80 levels) and EMC FV3-SARX (50 levels)
- NSSL FV3-SAR mean slightly higher at both times

Different DA comparison:

- EMC FV3-SAR DA better at later times, but sample size limited

A7. FV3-SAR IC/Hord/LSM

Participants were reminded of their ratings for the FV3-EMC SARX while rating



-Different diffusivity (HORD=6 → more diffusive)

-Different initial conditions

-Different land surface model

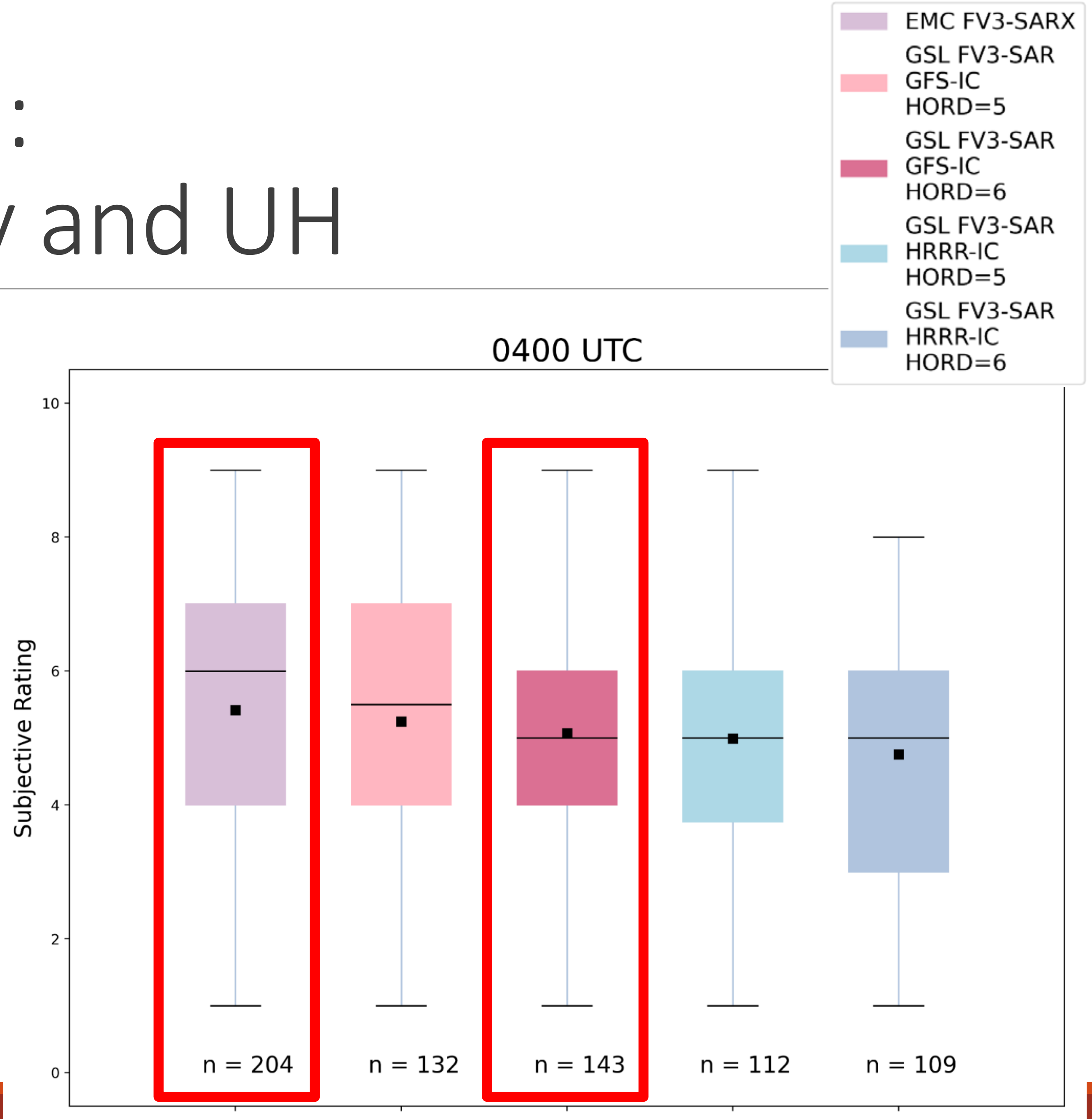
FV3-SAR IC/Hord/LSM: Composite Reflectivity and UH

Differences don't appear until 0400 UTC
- Exception: EMC FV3-SARX performs best

Different diffusivity comparison:
- HORD=5 generally higher than HORD=6
- Less diffusivity performing higher for both IC configurations

Different IC comparison:
- GFS ICs higher than HRRR ICs
- GFS ICs with HORD=6 almost identical to HRRR ICs with HORD=5
- Slightly more GFS-IC cases available

Different LSM comparison:
- NOAH in EMC FV3-SARX outperformed RUC in GSL FV3-SAR with GFS-ICs and HORD=6



FV3-SAR IC/Hord/LSM: Environment

Larger differences at 0000 UTC compared to 1800 UTC

- Exception: EMC FV3-SARX performs best

Different diffusivity comparison:

- Almost no difference in ratings

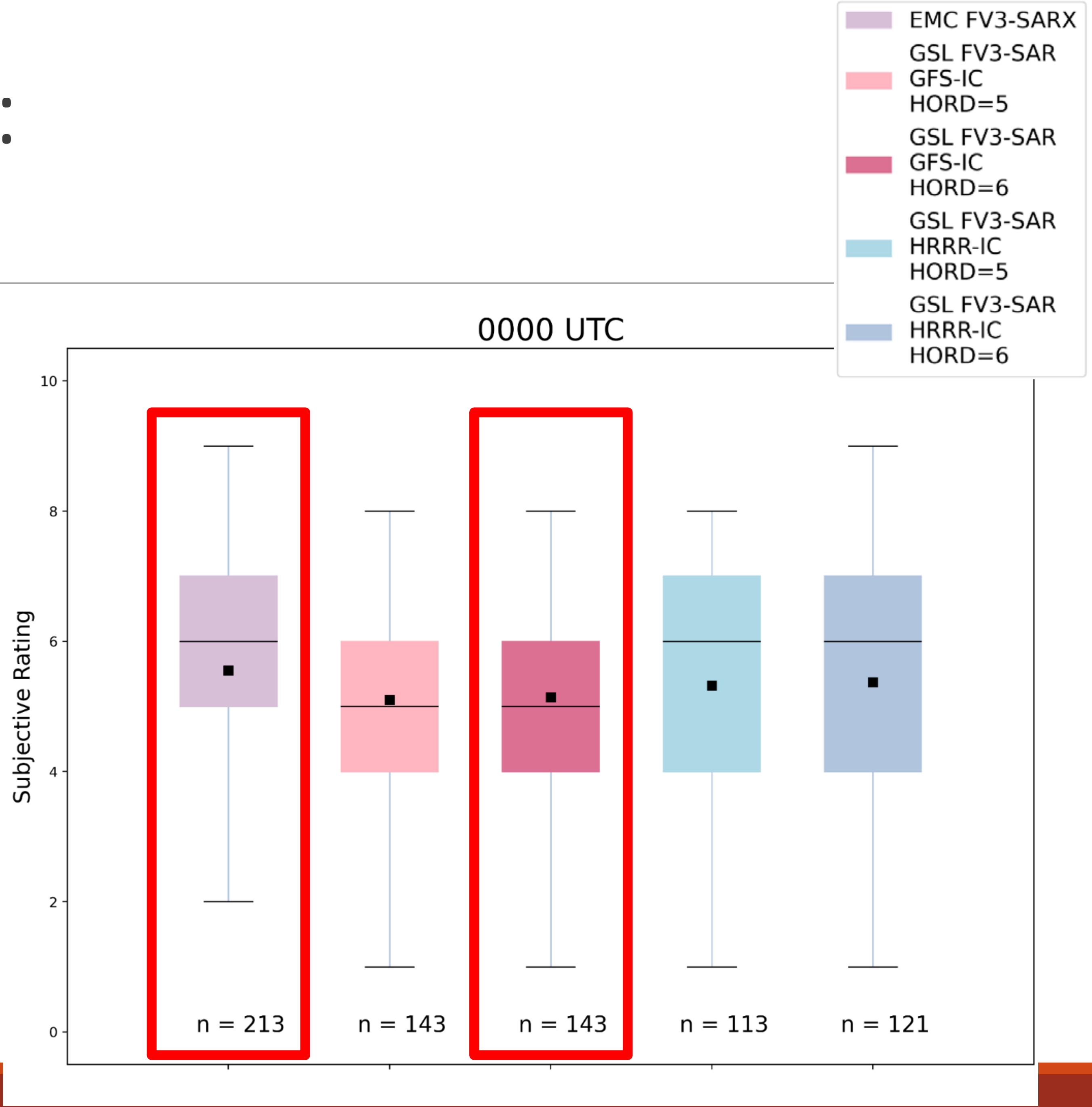
Different ICs comparison:

- HRRR ICs perform better than GFS-ICs (although means are quite close)

Different LSMs Comparison:

- NOAH in EMC FV3-SARX outperformed RUC in GSL FV3-SAR with GFS-ICs and HORD=6

Overall, similar results to UH and composite reflectivity except for ICs



Conclusions

Development of FV3-based CAMs has rapidly accelerated over the past few years
- Major progress has been made

Types of changes that had the largest impacts on subjective model performance
- Using more advanced physics (such as Thompson, MYNN) ↑
- NOAH LSM ↑
- Increased vertical levels ↑ (mostly at earlier times)
- Initial conditions ↔

Persistent cool, moist bias in FV3-based CAMs, but seems improved from prior years

These are subjective evaluations – objective verification will be taking place to complement these analyses

Summary report will be forthcoming and available on the SFE's homepage: https://hwt.nssl.noaa.gov/sfe/2020/
- Goal is to have it complete by end of August

More questions? Feel free to contact me at burkely.twiest@noaa.gov