Multi-model Ensemble Forecasts of Ground Magnetic Perturbations

Kathryn Wilbanks, Daniel Welling

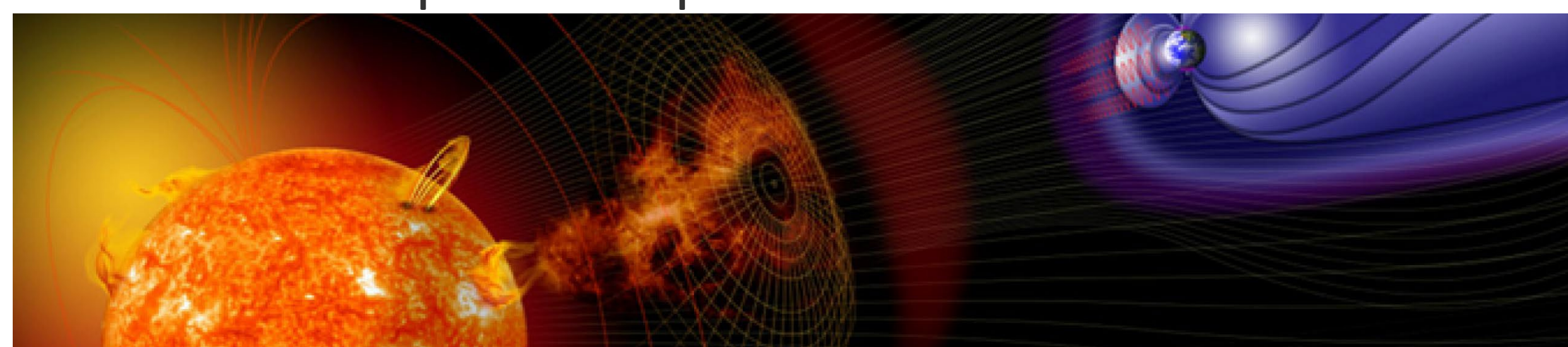University of Michigan Climate and Space Sciences and Engineering

Space Weather Forecasting

From aviation, high frequency radios, GPS, to power grids, space weather has impacts to every aspect of human's daily lives. The creation of a robust forecasts is needed to help lessen societal impacts of space weather.

Community-led model validations resulted in the implementation of a deterministic forecast at the Space Weather Prediction Center, SWPC[1]. This current operational model is the Space Weather Modeling Framework (SWMF). While many advancements have been made — Space Weather forecasting still lags behind terrestrial forecasting.

Ensemble Forecasting

Meteorologist utilize ensemble forecasting, a combination of multiple models, to build a probabilistic forecast. This can be se
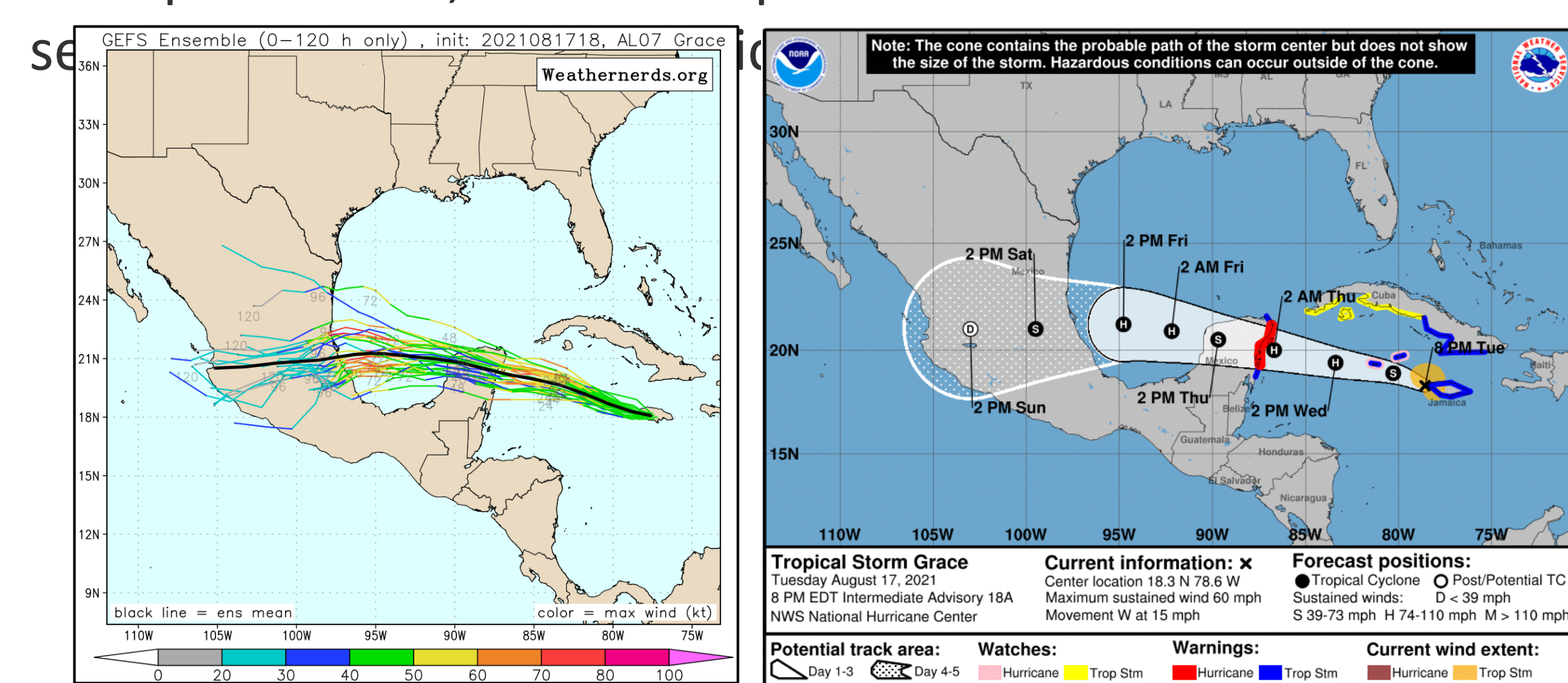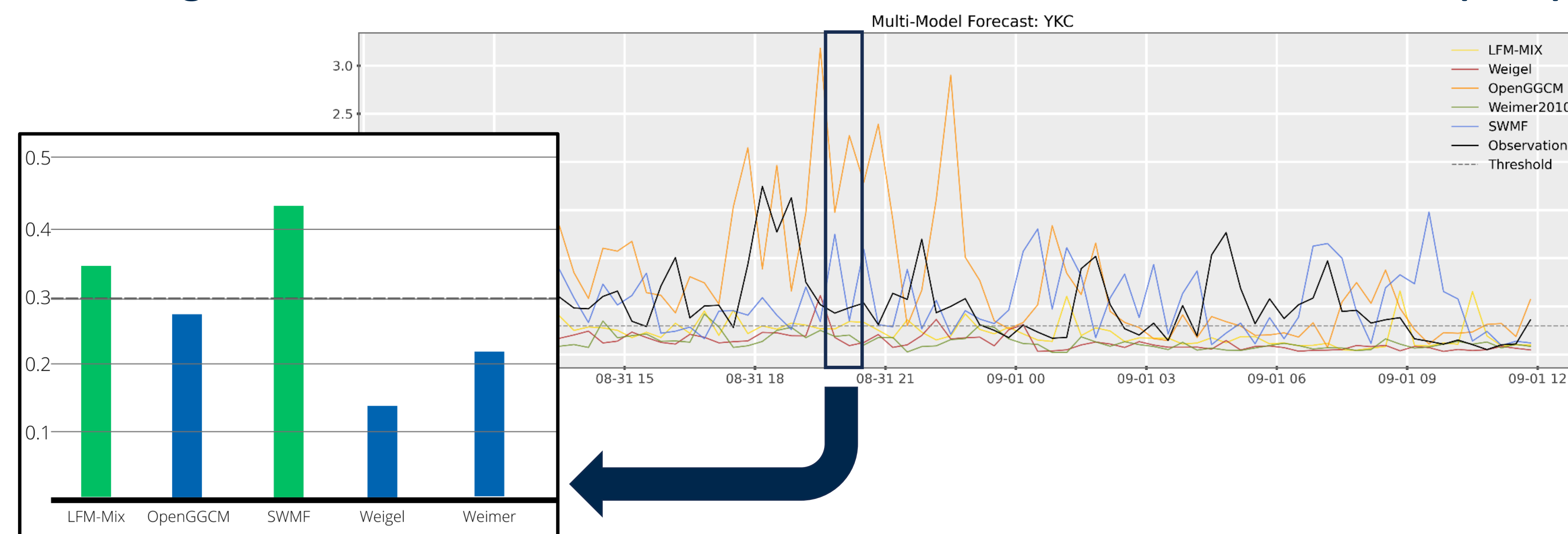

Perturbed input ensemble modeling, where a single model's initial conditions are changed to create multiple model runs, has been explored[2]. This work showed in an improvement in the skill scores of forecasts.
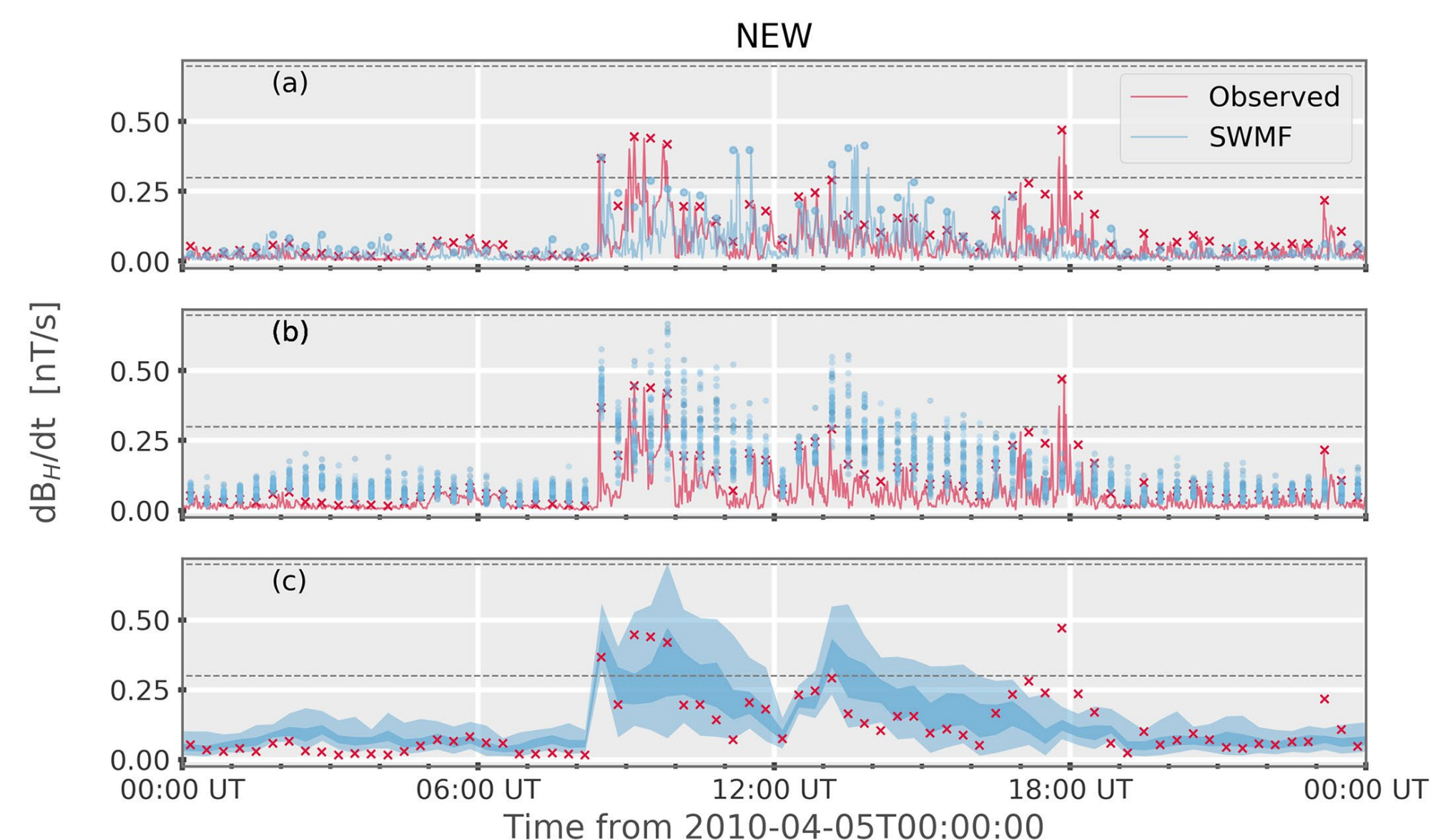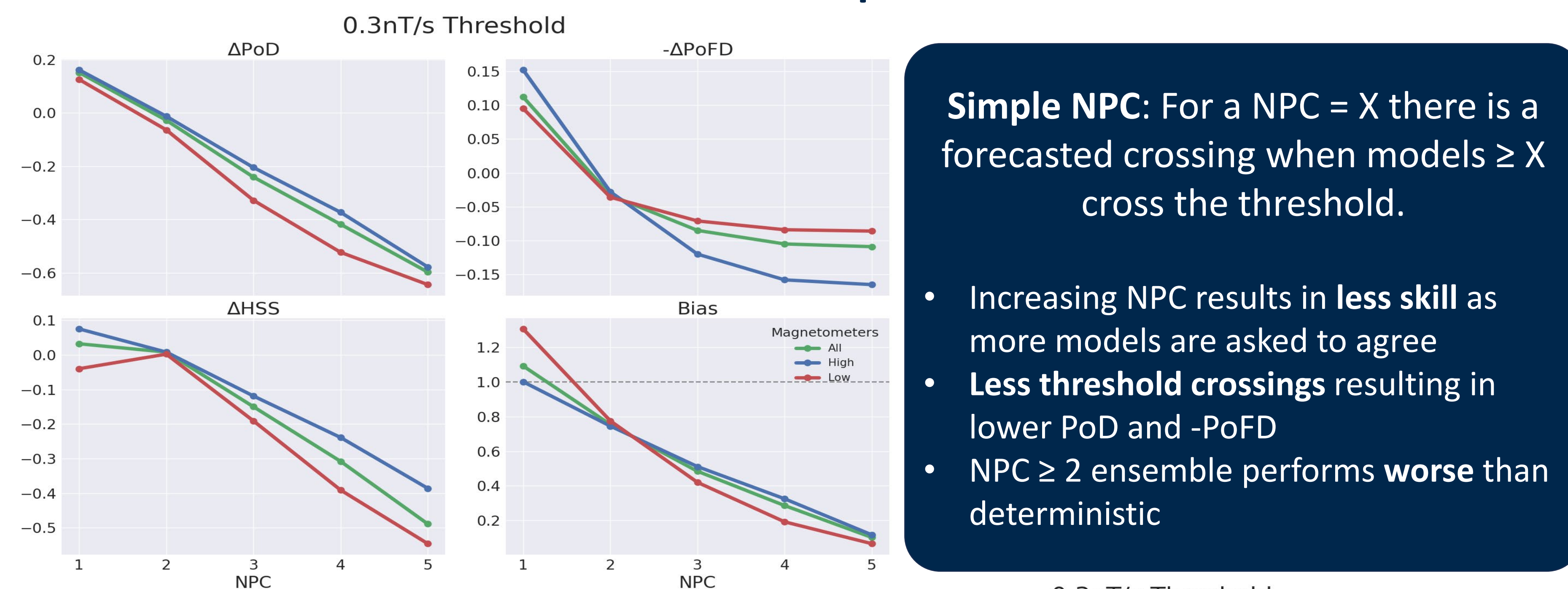

This work looks at utilizing similar methodologies of the perturbed input ensemble by creating a multi-model ensemble based on the same SWPC challenge validation from Pulkkinen 2013. By taking the six models used in the Pulkkinen challenge to create a multi-model ensemble using a naïve probabilistic classifier (NPC) and compare the ensembles metrics and skill score to that of the operational deterministic forecast.

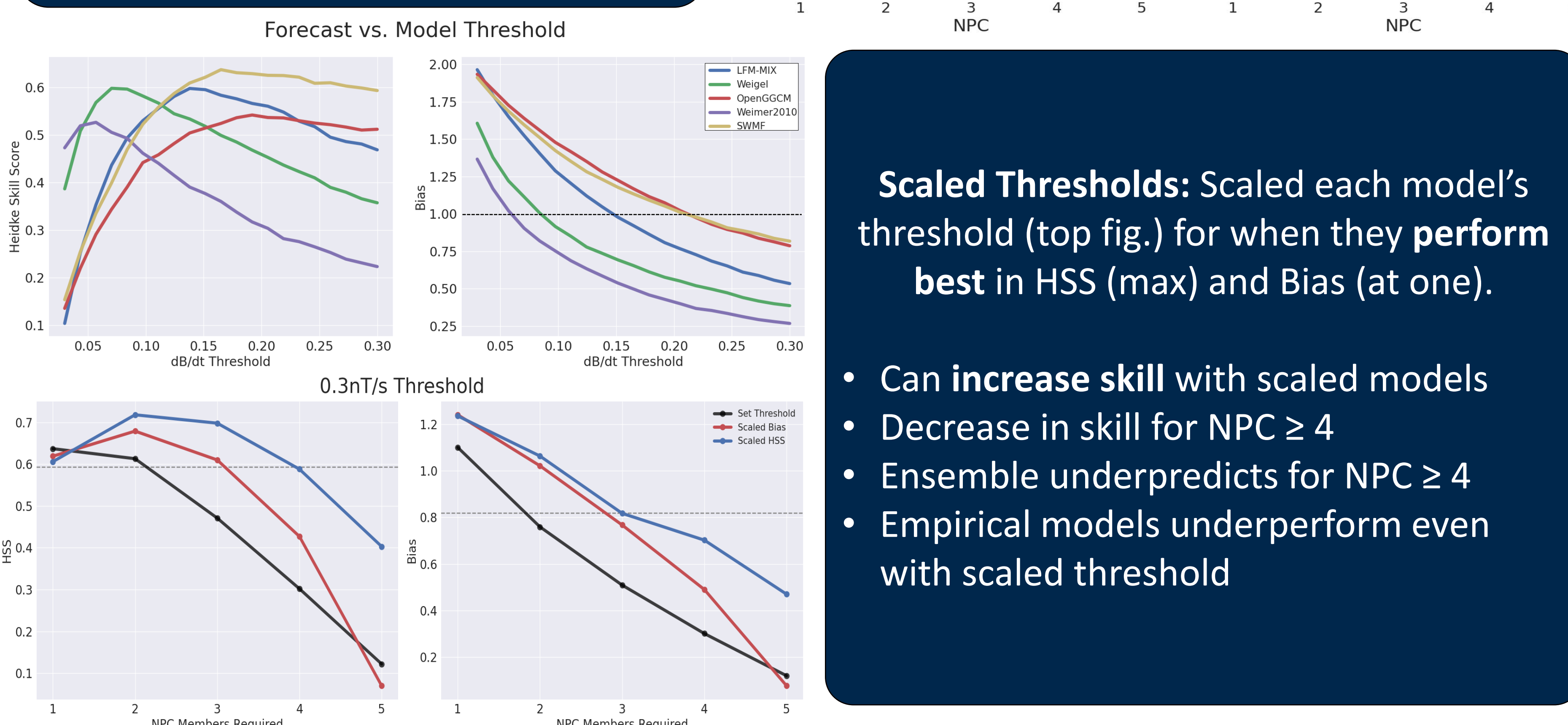## Creating a Multi-model Ensemble Forecast: Naïve Probabilistic Classifier (NPC)

Multi-model Ensemble Compared to Deterministic

Simple NPC: For a NPC = X there is a forecasted crossing when models ≥ X cross the threshold.

- Increasing NPC results in **less skill** as more models are asked to agree
- **Less threshold crossings** resulting in lower PoD and -PoFD
- NPC ≥ 2 ensemble performs **worse** than deterministic

Best of NPC: For a forecasted threshold crossing, either NPC ≥ X or the deterministic model is used.

- Convergence of PoD, PoFD, and HSS to zero show the deterministic taking over for NPC ≥ 3
- Best of **underpredicts** for NPC ≥ 2 also due to deterministic underpredicting

Scaled Thresholds: Scaled each model's threshold (top fig.) for when they **perform best** in HSS (max) and Bias (at one).

- Can **increase skill** with scaled models
- Decrease in skill for NPC ≥ 4
- Ensemble underpredicts for NPC ≥ 4
- Empirical models underperform even with scaled threshold

Methodology

Recreated the binary event analysis seen in Pulkkinen 2013 SWPC Challenge with the same six events, magnetometer sets, and five models (two empirical and three global magnetohydrodynamic models).

Event Num.	Date Range	Min. (*Dst*)	Max (*Kp*)
1	29 Oct. – 30 Oct. 2003	-353 nT	9
2	14 Dec. – 16 Dec. 2006	-139 nT	8
3	31 Aug. -1 Sep. 2001	-40 nT	4
4	31 Aug. – 1 Sep. 2005	-131 nT	7
7	5 Apr. – 6 Apr. 2010	-73 nT	8
8	5 Aug. – 6 Aug. 2011	-113 nT	8

- Binned time into 20 minute intervals with thresholds of 0.3, 0.7, 1.1, and 1.5 nT/s
- Compared observed and simulated dB_h/dt peaks of ground magnetometers using three metrics and one skill score built off the following 2x2 contingency table:

Threshold Crossing?	Observed Crossing	No Observed Crossing
Forecasted Crossing	Correct Forecast	False Negative
No Forecasted Crossing	False Forecast	Correct negative

- Table was then used to calculate 3 metrics and 1 skill score: Probability of Detection (PoD), Probability of False Detection (PoFD), Heidke Skill Score (HSS) and Bias

Conclusion

First multi-model ensemble forecast for geomagnetic disturbances (GMDs).

Mean and Median forecast have a decrease in skills, but can be useful for determining **forecast uncertainty**.

There is a **decrease** in skill and metrics as more NPC members are required. This is due to all five models underpredicting, especially the two empirical models.

See **skill improvements** depending on how we construct the ensemble. Specifically, Best of for NPC<3 and Scaled Thresholds.

References
[1] A. Pulkkinen, L. Rastätter, M. Kuznetsova, H. Singer, C. Balch, D. Weimer, G. Toth, A. Ridley, T. Gombosi, M. Wiltberger, J. Raeder, and R. Weigel, Space Weather **11**, 369 (2013).
[2] S.K. Morley, D.T. Welling, and J.R. Woodroffe, Space Weather **16**, 1330 (2018).
[3] D.T. Welling, B.J. Anderson, G. Crowley, A.A. Pulkkinen, and L. Rastätter, Space Weather **15**, 192 (2017).