Determining predictors for a statistical tropical cyclone genesis tool based on GFS output

Daniel J. Halperin, Robert E. Hart, Henry E. Fuelberg, Joshua H. Cossuth

Florida State University

31 March 2014

This research is funded by the Joint Hurricane Testbed

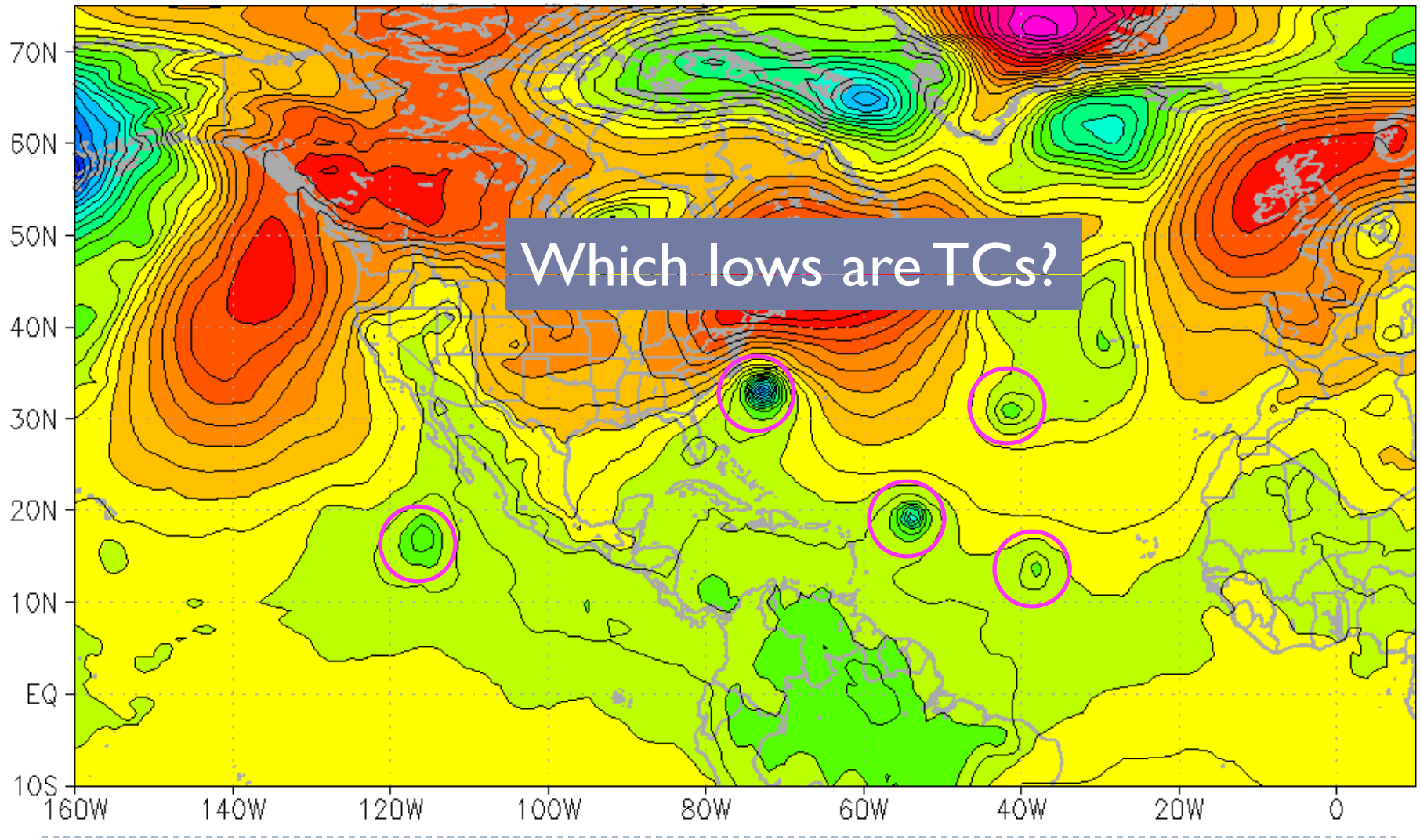
Background

- Prior research verified TC genesis forecasts out to 5 days in 5 global models over the NATL and EPAC during 2004-2013.
  - CMC, ECM, GFS, NGP, UKM
- Halperin, D.J., H.E. Fuelberg, R.E. Hart, J.H. Cossuth, P. Sura, and R.J. Pasch, 2013: An evaluation of tropical cyclone genesis forecasts from global numerical models. *Wea. Forecasting,* **28**, 1423-1445.
- Current research objective: Develop probabilistic forecasts of TC genesis based on global model output.

Background

- Not a new concept. Examples of research with a similar goal have already been presented in this session and more to come this week.
- A mix of model fields and observations
  - TCFP (DeMaria et al. 2001; Schumacher et al. 2009, 2014)
  - TCGI (Dunion et al. 2013)
- HFIP
  - Marchok (2012); Fiorino (2014); Peng et al. (2014)
- Climate-scale applications
  - Emanuel and Nolan (2004); Camargo et al. (2007); Emanuel (2010); Tippett et al. (2011); Bruyère et al. (2012)

Motivation


70N
60N
50N
40N
30N
20N
10N
EQ
10S
160W
140W
120W
100W
80W
60W
40W
20W
0
Which lows are TCs?

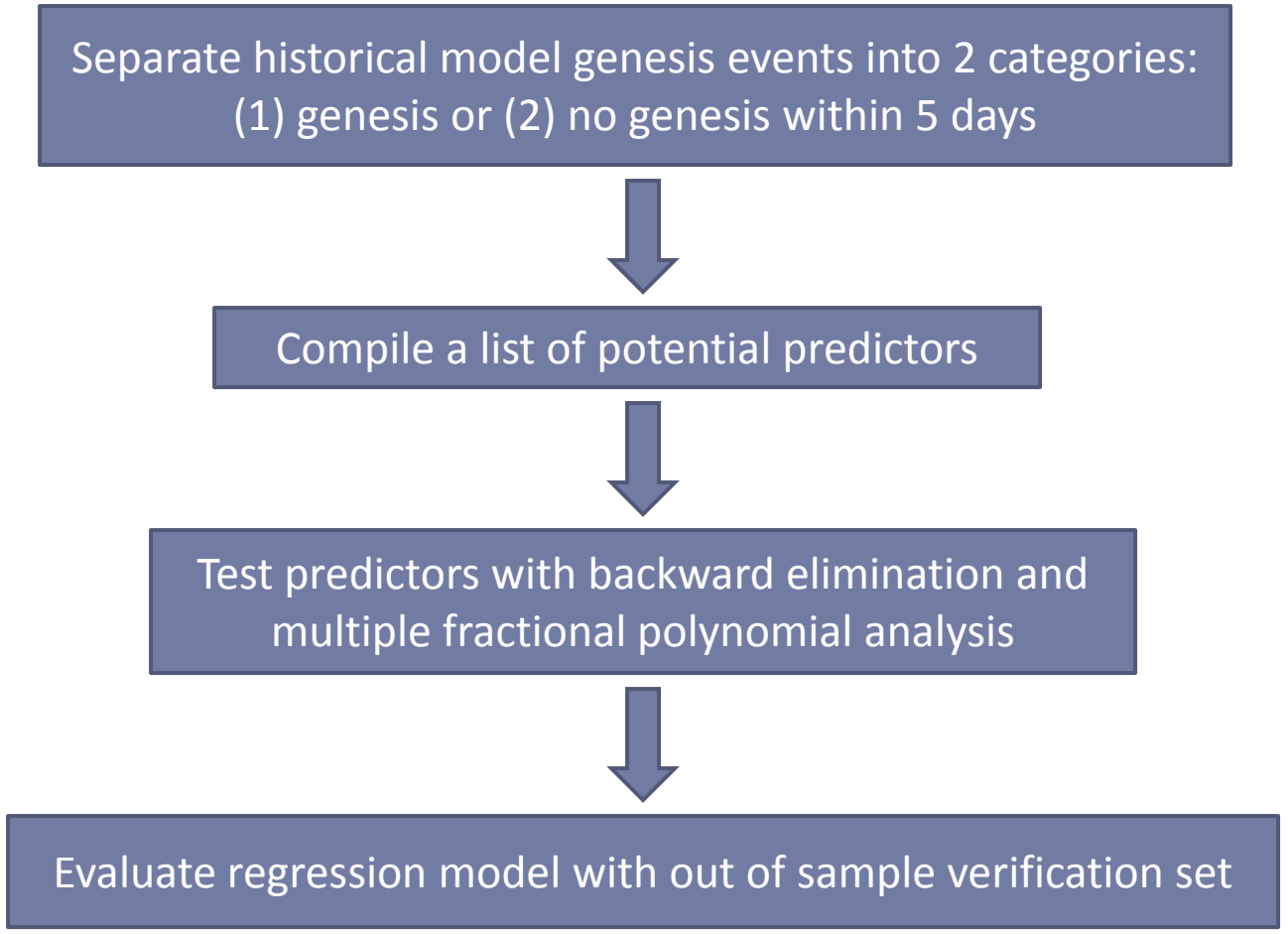
Summary of prior results

- Global operational models' performance:
  - Has generally improved since 2004.
  - Is better in the EPAC on average due to larger probability of detection values.
  - Varies among different subregions in the NATL.
  - Expectedly decreases with increasing forecast hour.
- The best performing model varies from year to year and basin to basin.
- For more information: **P. 146 on Thursday**. *Verification of TC genesis forecasts from global models: Updates and real-time applications.*

New Research Questions

- Prior results show spatial and temporal variations in model performance.
  - Other useful variables?
  - Is logistic regression suitable?
  - Do the predictors make sense physically?
  - Do the predictors provide insight regarding important TC genesis processes?
  - What large-scale factors might impact the performance of model genesis forecasts (e.g., SAL (Pratt and Evans 2009))?

Logistic regression model development

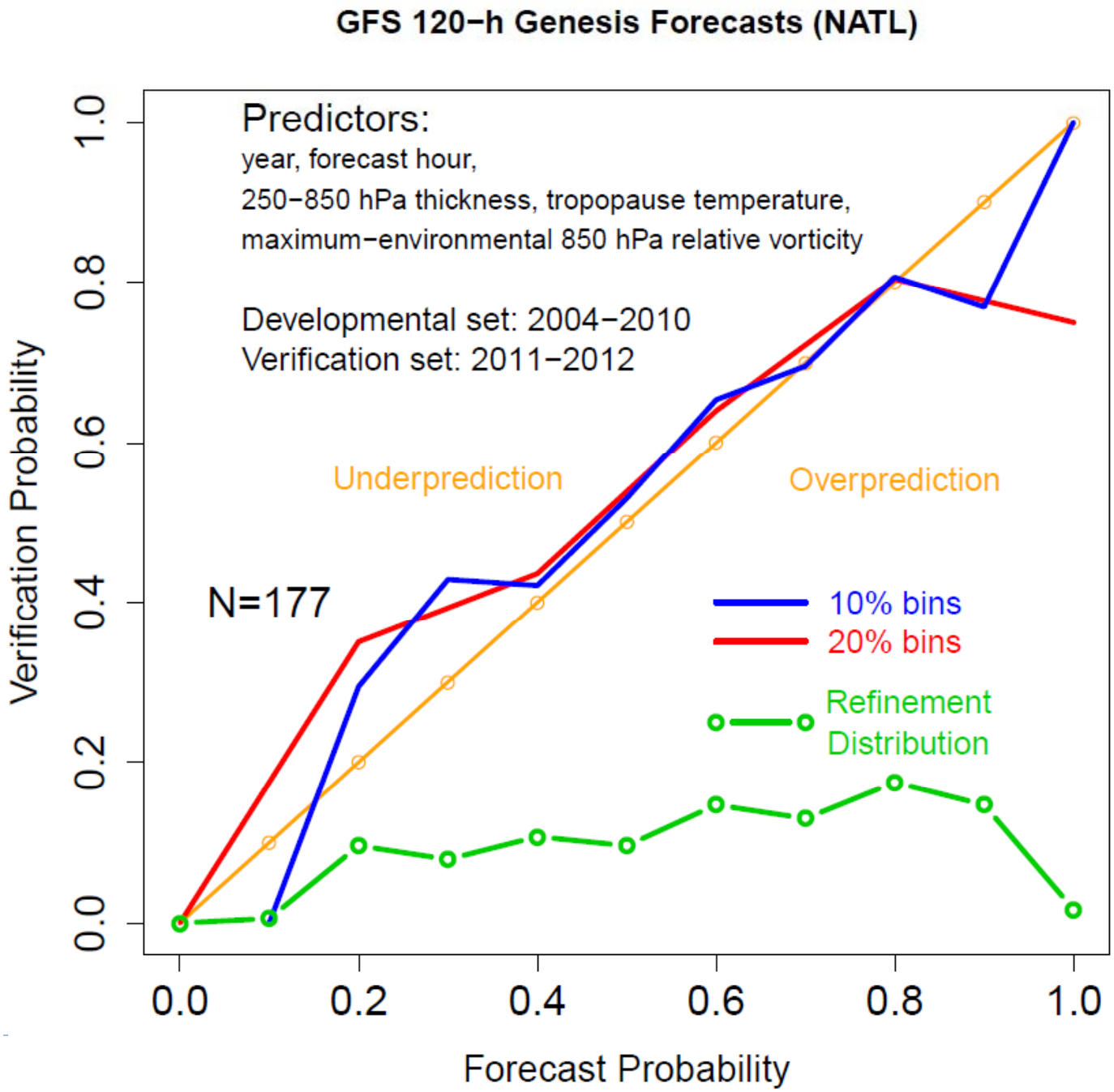


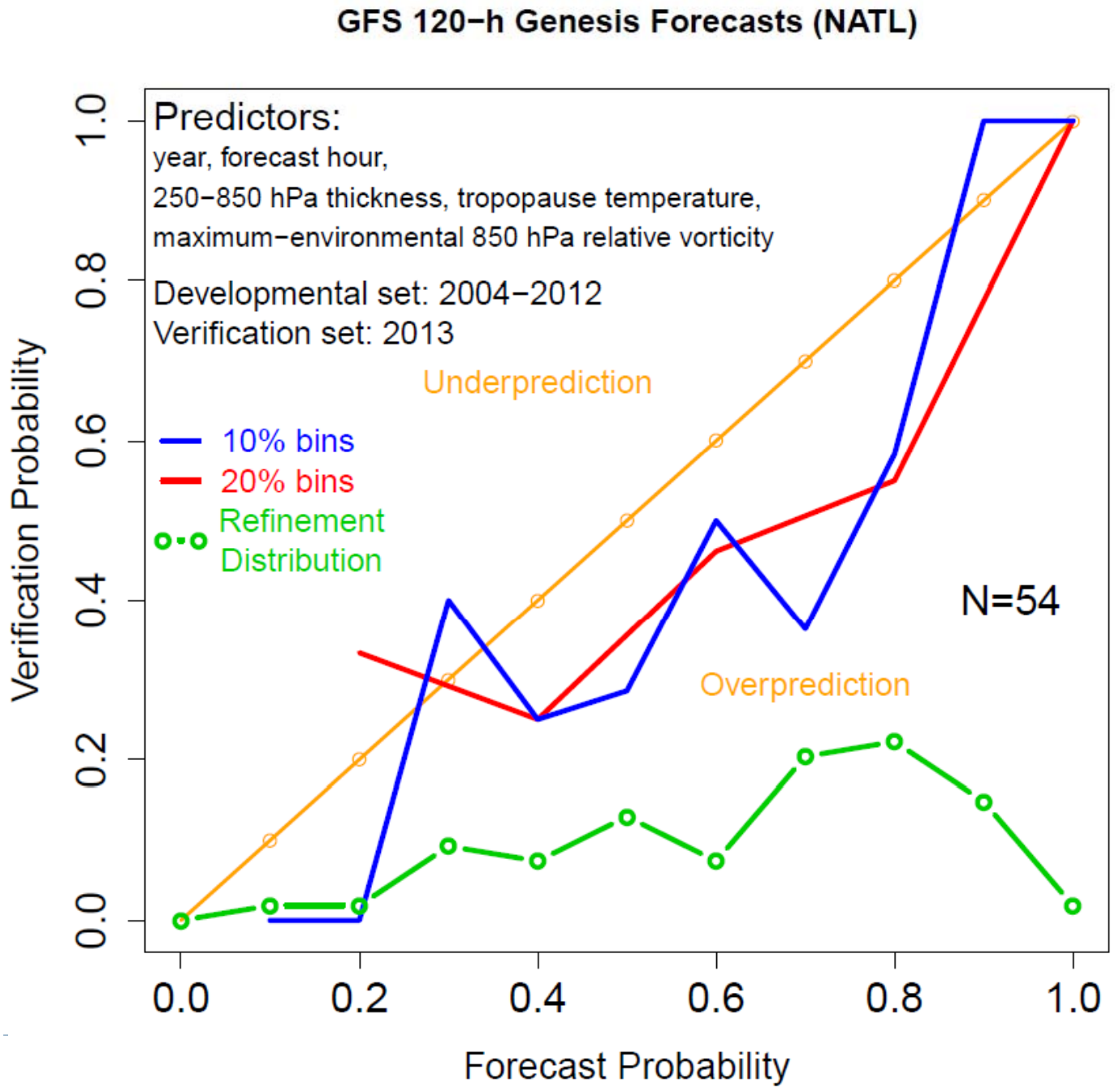
Selecting predictors

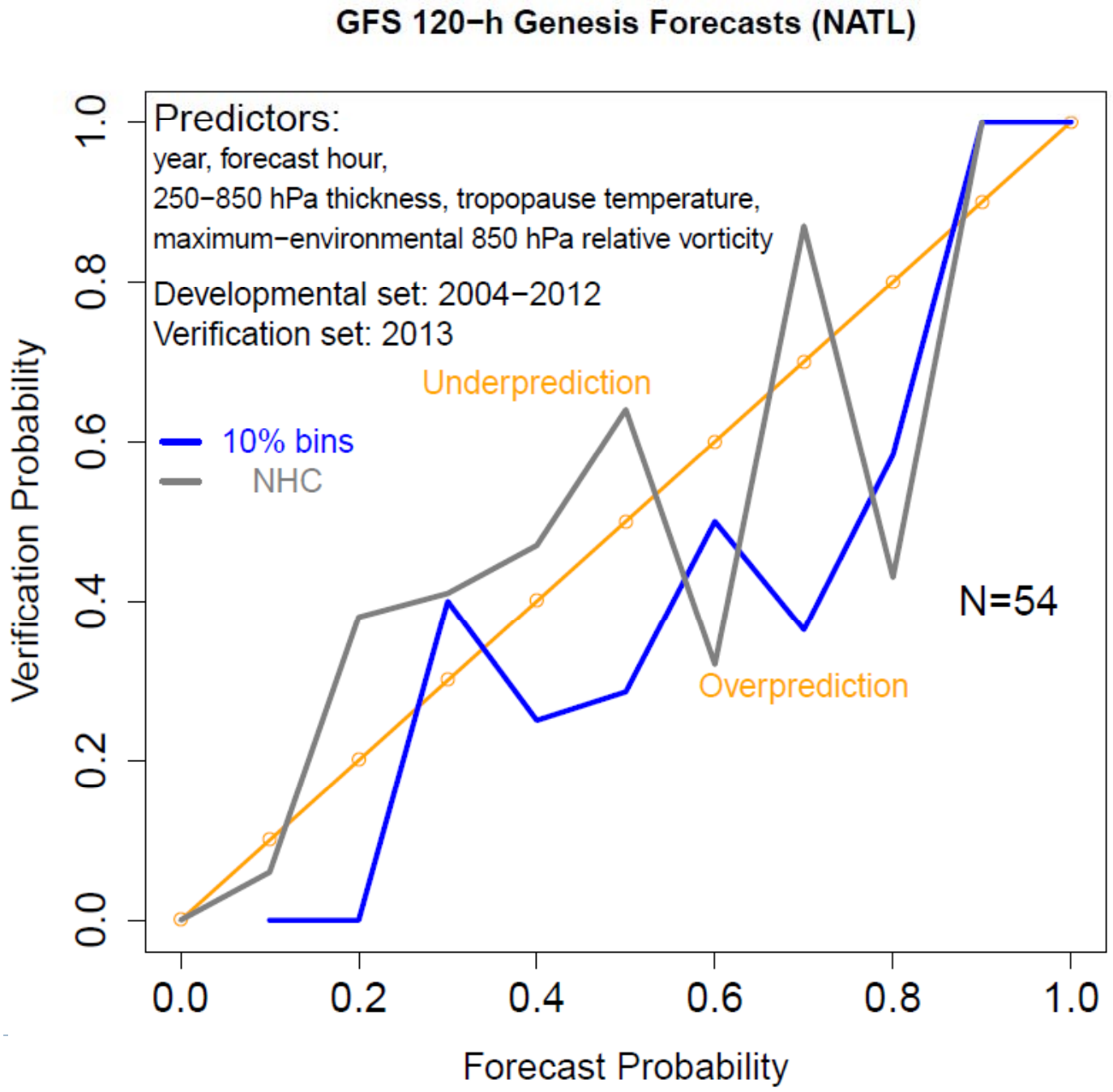
- 20 iterations of out-of-sample testing to determine significant predictors.
  - Each genesis event will be used in the verification set once.

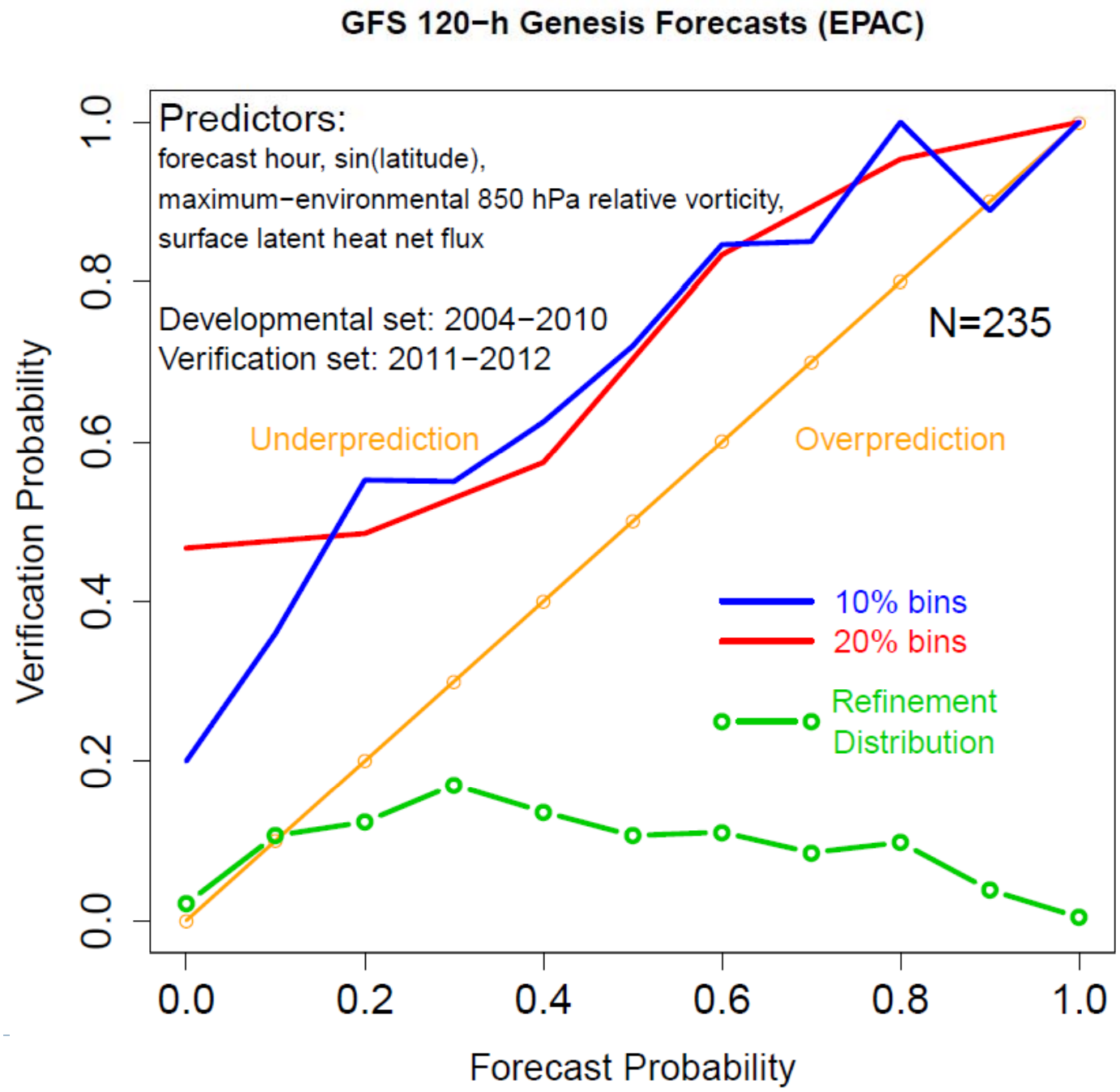
NATL	
Forecast hour	20
Year	20
250-850 mb ΔZ	20
850 mb ζ perturbation	20
Tropopause Temp	20
CAPE	20
Sfc latent heat net flux	20
Longitude	20

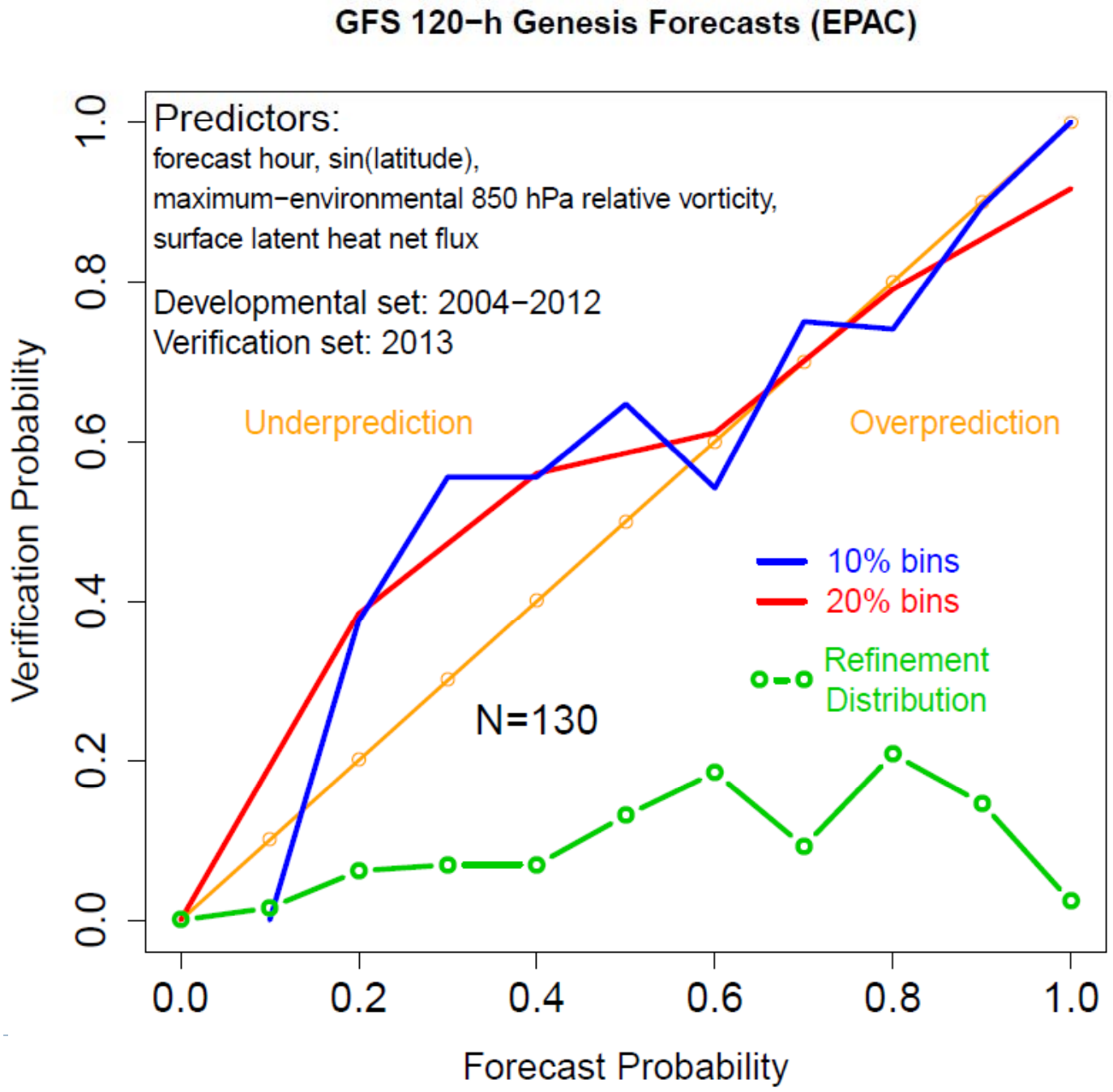
EPAC	
Forecast hour	20
Sfc latent heat net flux	20
sin(latitude)	20
850 mb ζ perturbation	18
PWAT	16
CIN	16

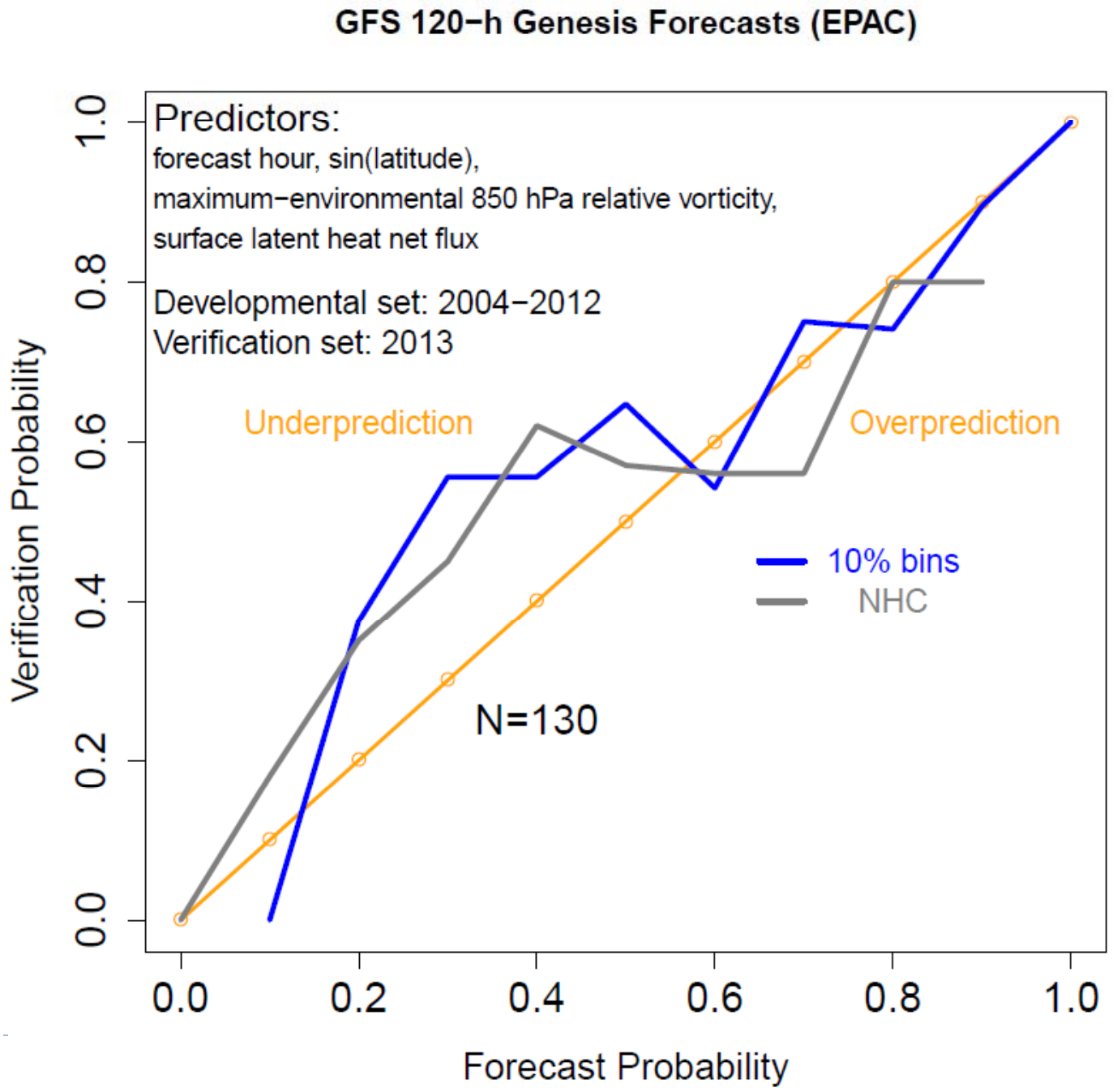

GFS 120−h Genesis Forecasts (NATL)
Predictors:
year, forecast hour,
250−850 hPa thickness, tropopause temperature,
maximum−environmental 850 hPa relative vorticity
Developmental set: 2004−2010
Verification set: 2011−2012
Underprediction
Overprediction
N=177
10% bins
20% bins
Refinement
Distribution
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0

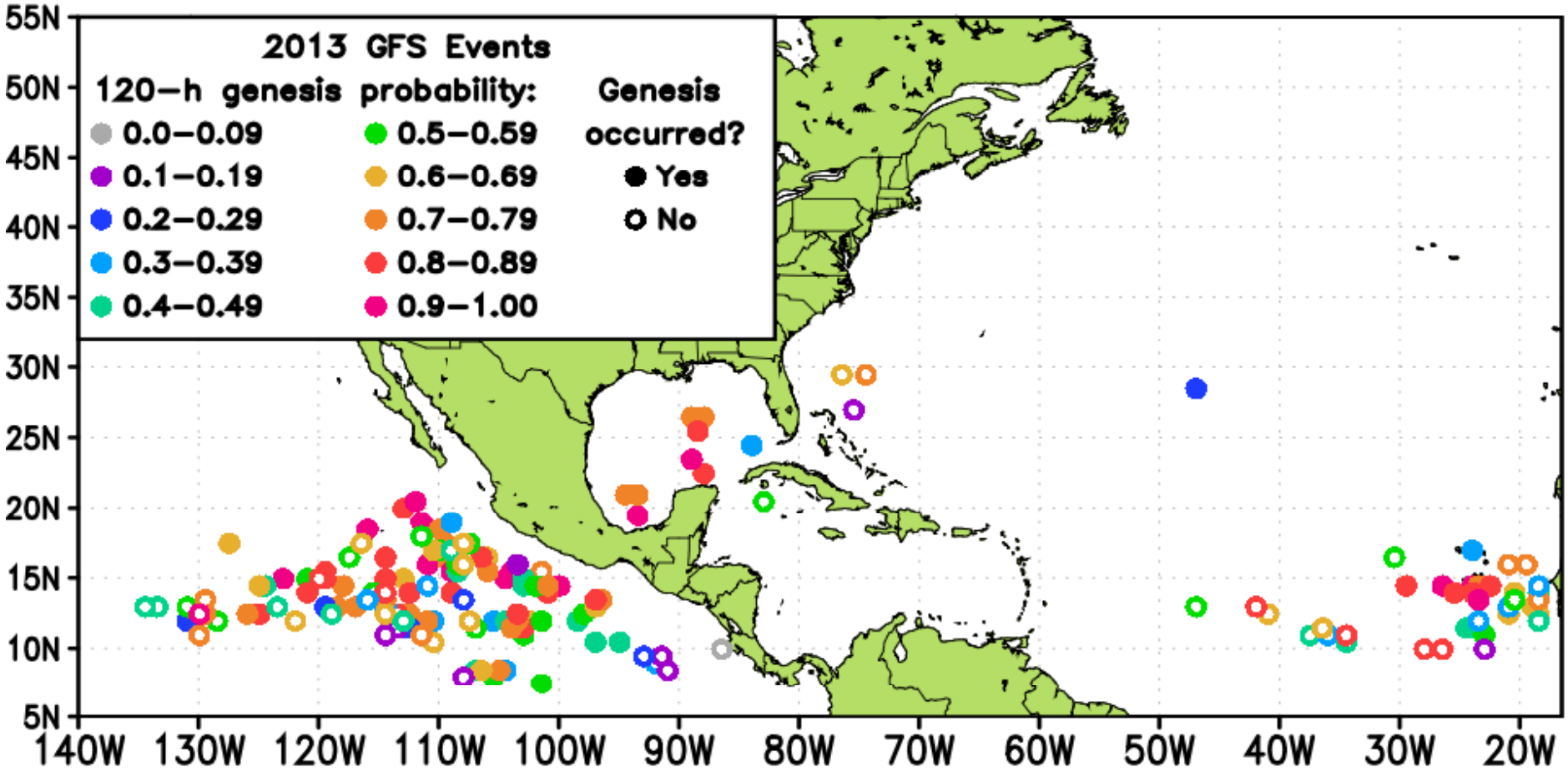

GFS 120−h Genesis Forecasts (NATL)
Predictors:
year, forecast hour,
250−850 hPa thickness, tropopause temperature,
maximum−environmental 850 hPa relative vorticity
Developmental set: 2004−2012
Verification set: 2013
Underprediction
10% bins
20% bins
Refinement
Distribution
N=54
Overprediction
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0

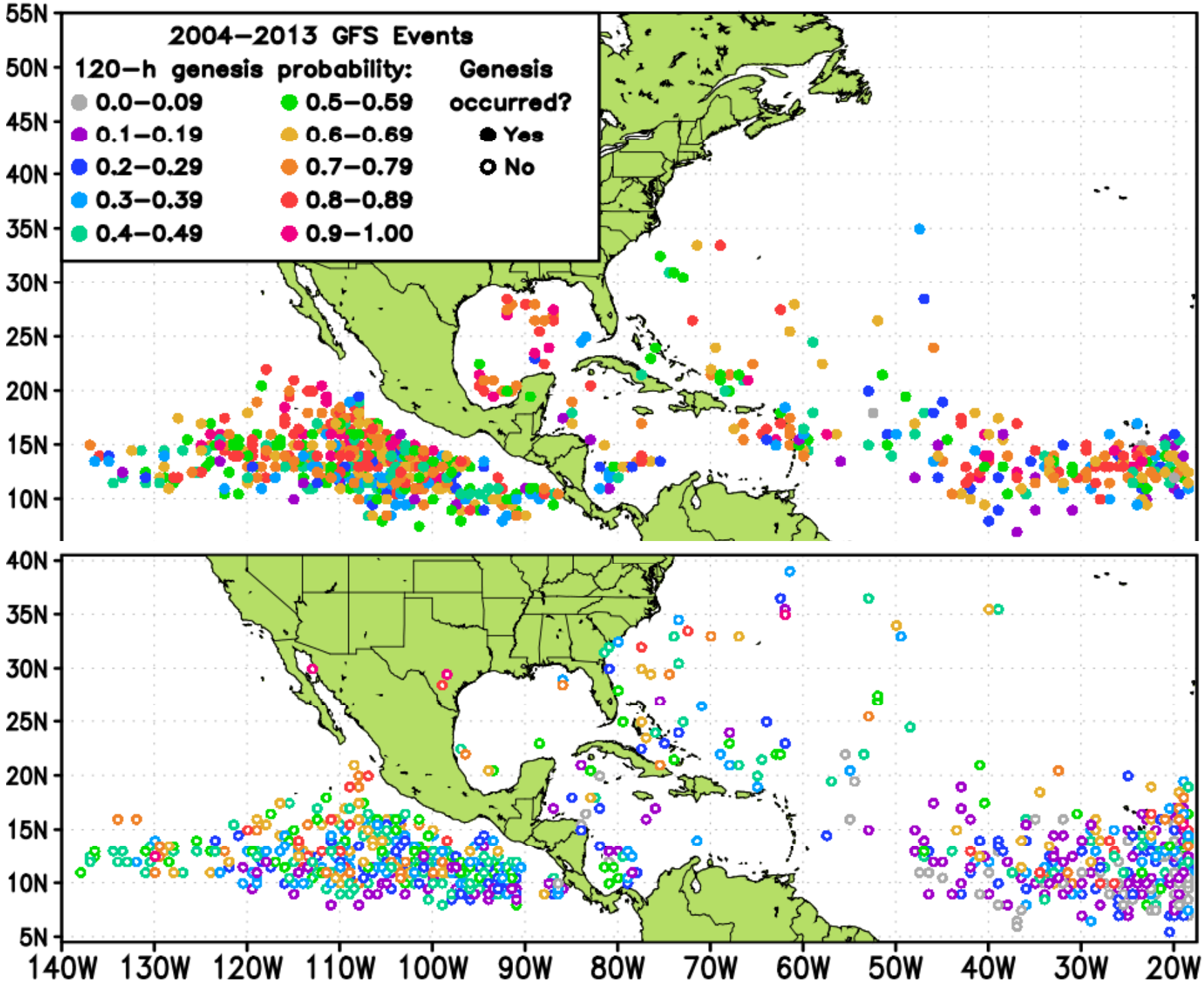

GFS 120−h Genesis Forecasts (NATL)
Predictors:
year, forecast hour,
250−850 hPa thickness, tropopause temperature,
maximum−environmental 850 hPa relative vorticity
Developmental set: 2004−2012
Verification set: 2013
Underprediction
10% bins
NHC
N=54
Overprediction
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0


GFS 120−h Genesis Forecasts (EPAC)
Predictors:
forecast hour, sin(latitude),
maximum−environmental 850 hPa relative vorticity,
surface latent heat net flux
Developmental set: 2004−2010
Verification set: 2011−2012
N=235
Underprediction
Overprediction
10% bins
20% bins
Refinement
Distribution
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0


GFS 120−h Genesis Forecasts (EPAC)
Predictors:
forecast hour, sin(latitude),
maximum−environmental 850 hPa relative vorticity,
surface latent heat net flux
Developmental set: 2004−2012
Verification set: 2013
Underprediction
Overprediction
10% bins
20% bins
Refinement Distribution
N=130
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0


GFS 120−h Genesis Forecasts (EPAC)
Predictors:
forecast hour, sin(latitude),
maximum−environmental 850 hPa relative vorticity,
surface latent heat net flux
Developmental set: 2004−2012
Verification set: 2013
Underprediction
Overprediction
10% bins
NHC
N=130
Verification Probability
Forecast Probability
0.0
0.2
0.4
0.6
0.8
1.0


2013 GFS Events
120–h genesis probability:
0.0–0.09
0.1–0.19
0.2–0.29
0.3–0.39
0.4–0.49
0.5–0.59
0.6–0.69
0.7–0.79
0.8–0.89
0.9–1.00
Genesis occurred?
Yes
No
55N
50N
45N
40N
35N
30N
25N
20N
15N
10N
5N
140W
130W
120W
110W
100W
90W
80W
70W
60W
50W
40W
30W
20W


2004–2013 GFS Events
120–h genesis probability:
0.0–0.09
0.1–0.19
0.2–0.29
0.3–0.39
0.4–0.49
0.5–0.59
0.6–0.69
0.7–0.79
0.8–0.89
0.9–1.00
Genesis occurred?
Yes
No
55N
50N
45N
40N
35N
30N
25N
20N
15N
10N
40N
35N
30N
25N
20N
15N
10N
5N
140W
130W
120W
110W
100W
90W
80W
70W
60W
50W
40W
30W
20W

Plans for 2014

- Develop regression models for the UKMET and CMC.
  - Develop a consensus approach.
- Run the regression models on forecast TCs in real-time.
- Post-season
  - Evaluate regression models.
  - Add 2014 forecasts to the database.
  - Potentially revisit predictor selection.