Verification of Ensemble-based Probabilistic Intensity Change Forecasts

Ryan D. Torn, Univ. Albany

Acknowledgement:
Paul Kucera & DTC Team

Contributions from various Ensemble Forecasting Groups & NHC Forecasters

Overview

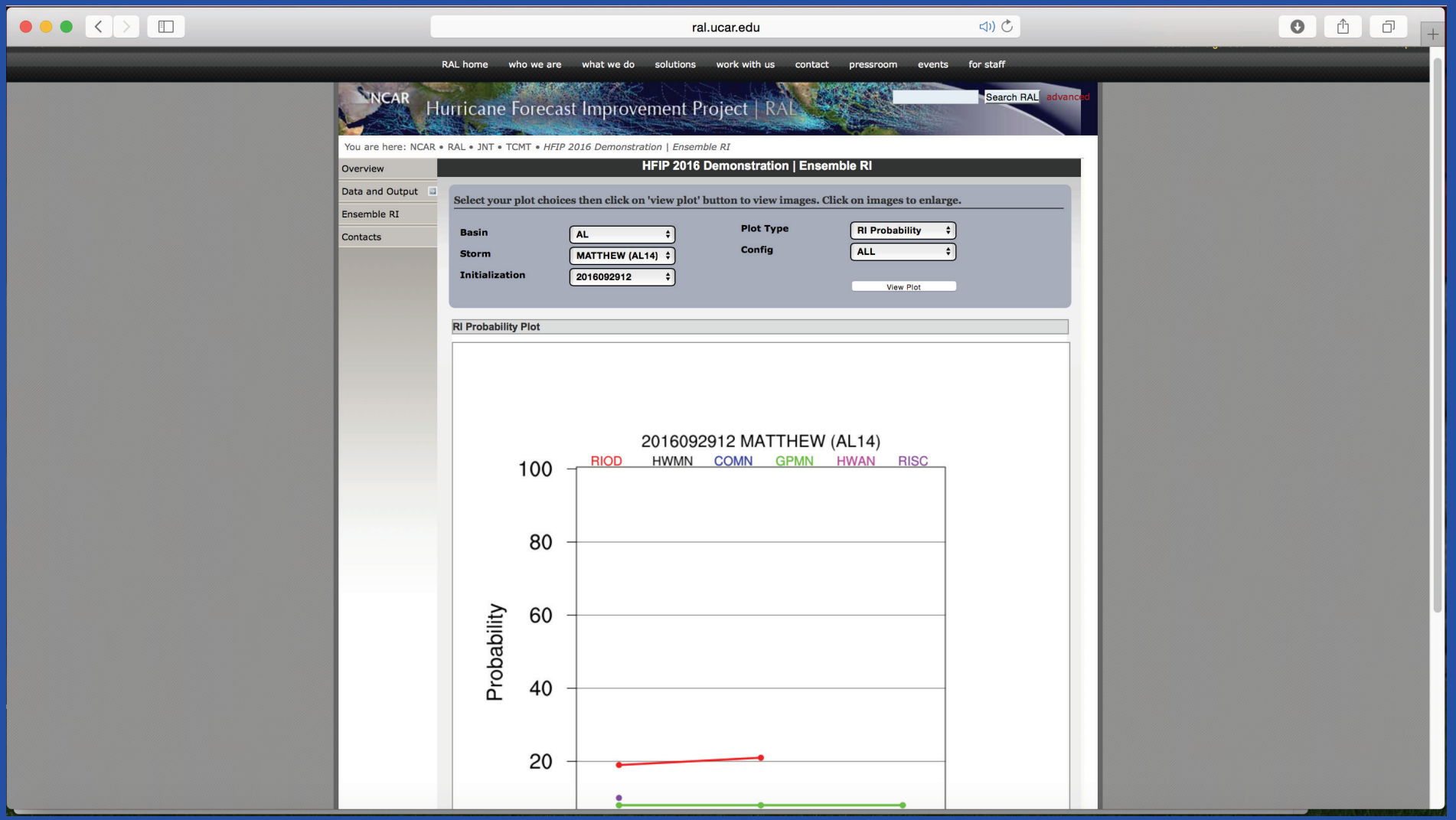
- **Charge**: Develop new ensemble-based products beyond the typical mean/standard deviation that could be used by NHC forecasters to improve forecasts
- For 2016: Develop an ensemble-based probabilistic RI product similar to the SHIPS RI product
  - HWRF Ensemble (Dyn.)
  - GFDL Ensemble (Dyn.)
  - COAMPS-TC Ensemble (Dyn.)
  - SPCE (Stat)
  - HWRF Analog (Stat)

Accomplishments

- Extension of ATCF e-deck format to intensity change probabilities
- Software to compute probabilities and write e-decks
- Retrospective forecast probabilities submitted for AL & EP 2013-2015 seasons
- Expansion of MET-TC software to validate RI probabilities
- Validation of retrospective forecasts
- Submission of real-time ensemble-based probabilities to DTC for demonstration project
- Validation of demonstration forecasts
- Development of real-time webpage to plot probabilities

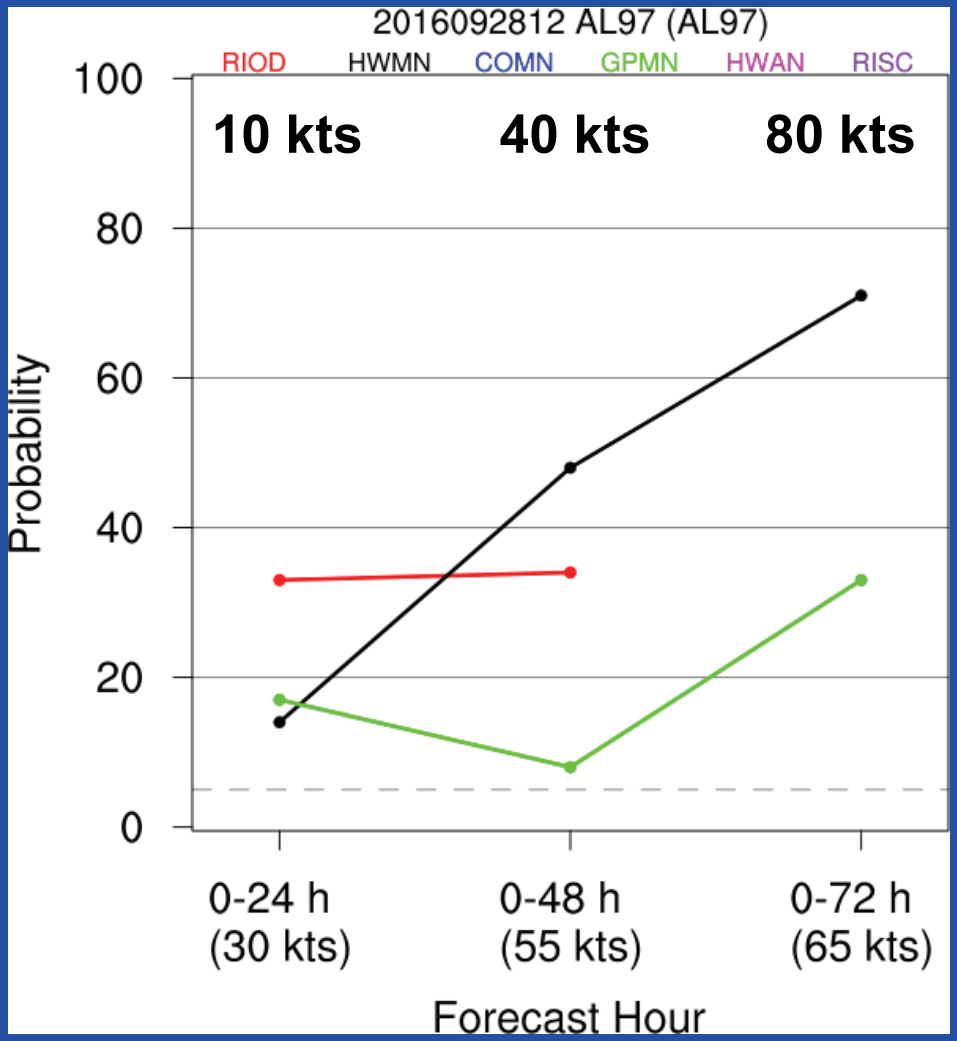
Real-Time Implementation

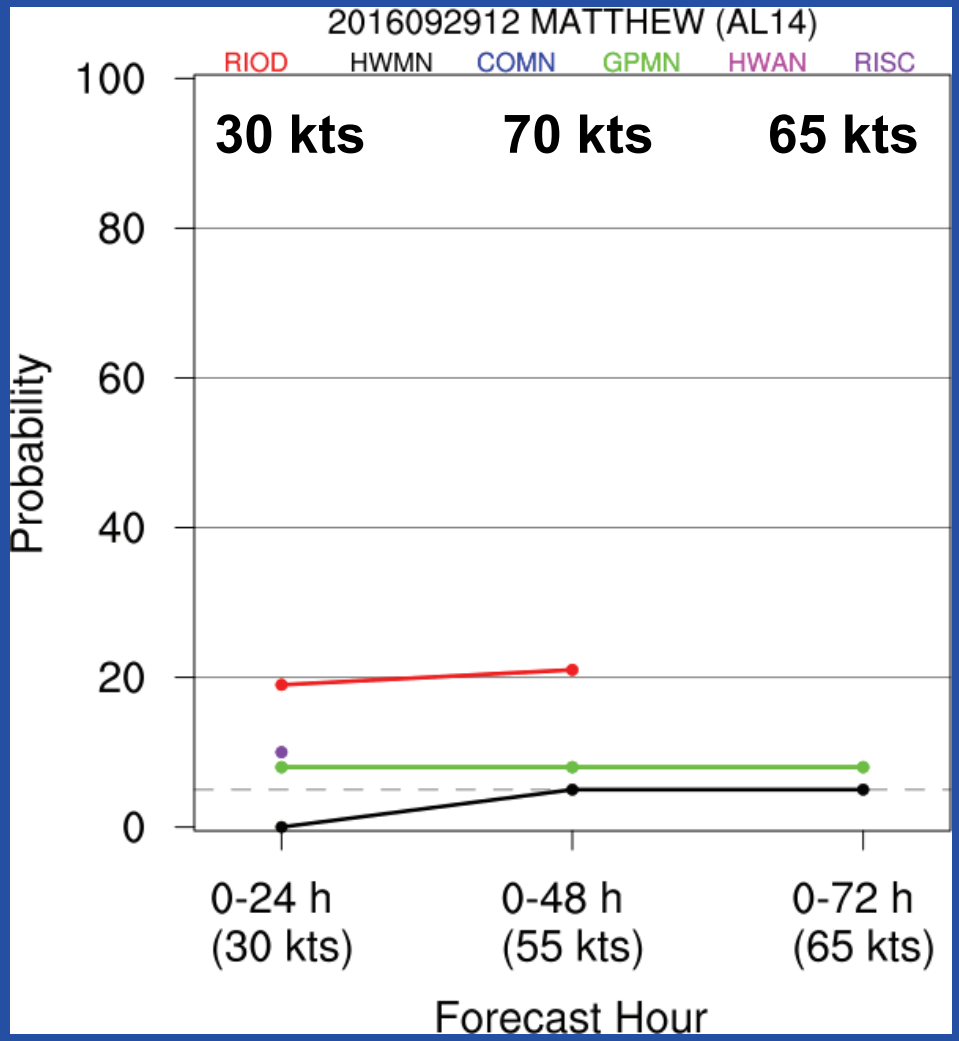
- Groups generated forecasts for storms during computing slot
- Intensity change rate interpolated to fit NHC forecast windows
  - XXMN – Raw intensity changes
  - XXXI – 6 h interpolated intensity changes
  - XXX2 – 12 h interpolated intensity changes
    - Example: COMN, COMI, COM2



https://ral.ucar.edu/projects/hfip/d2016/ensRI/

Matthew Example

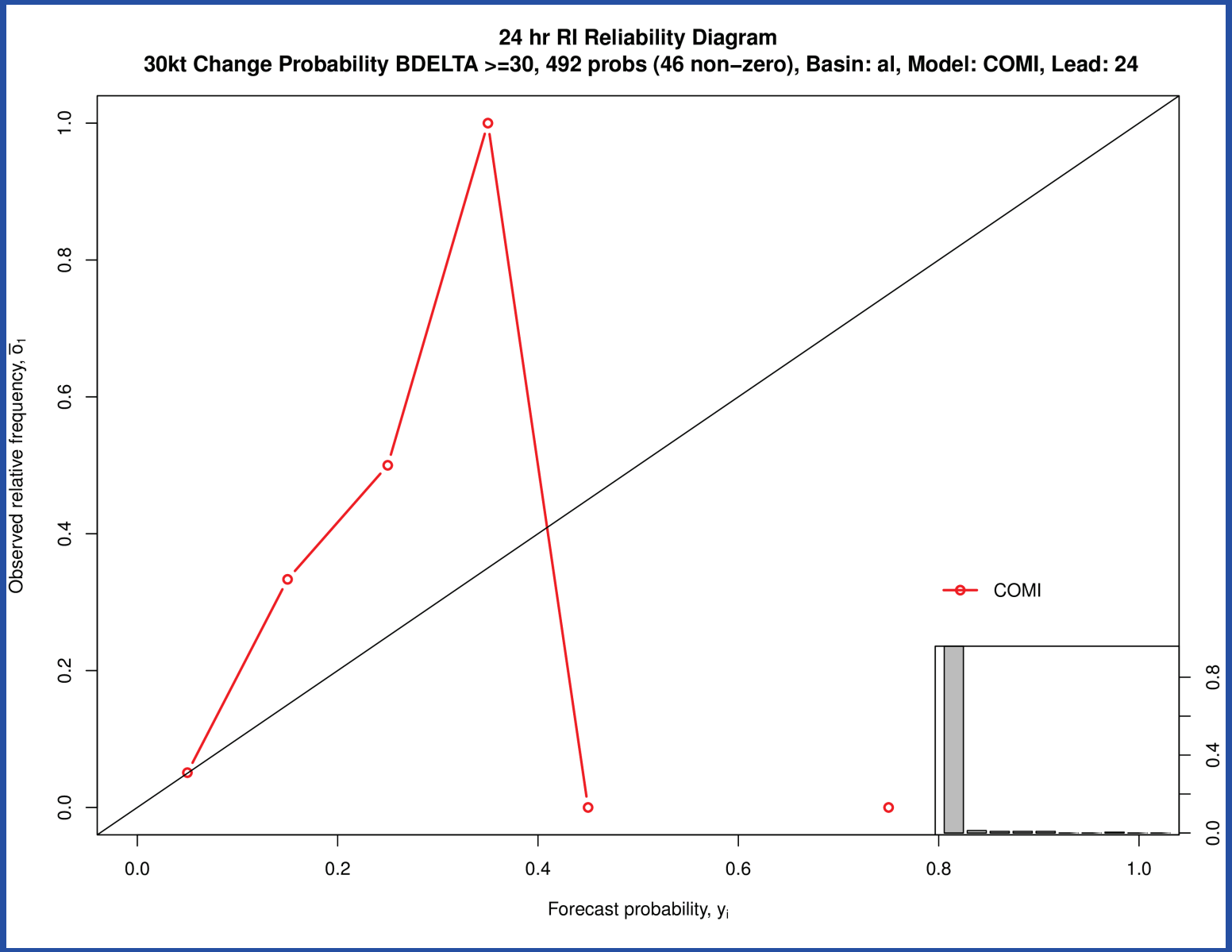




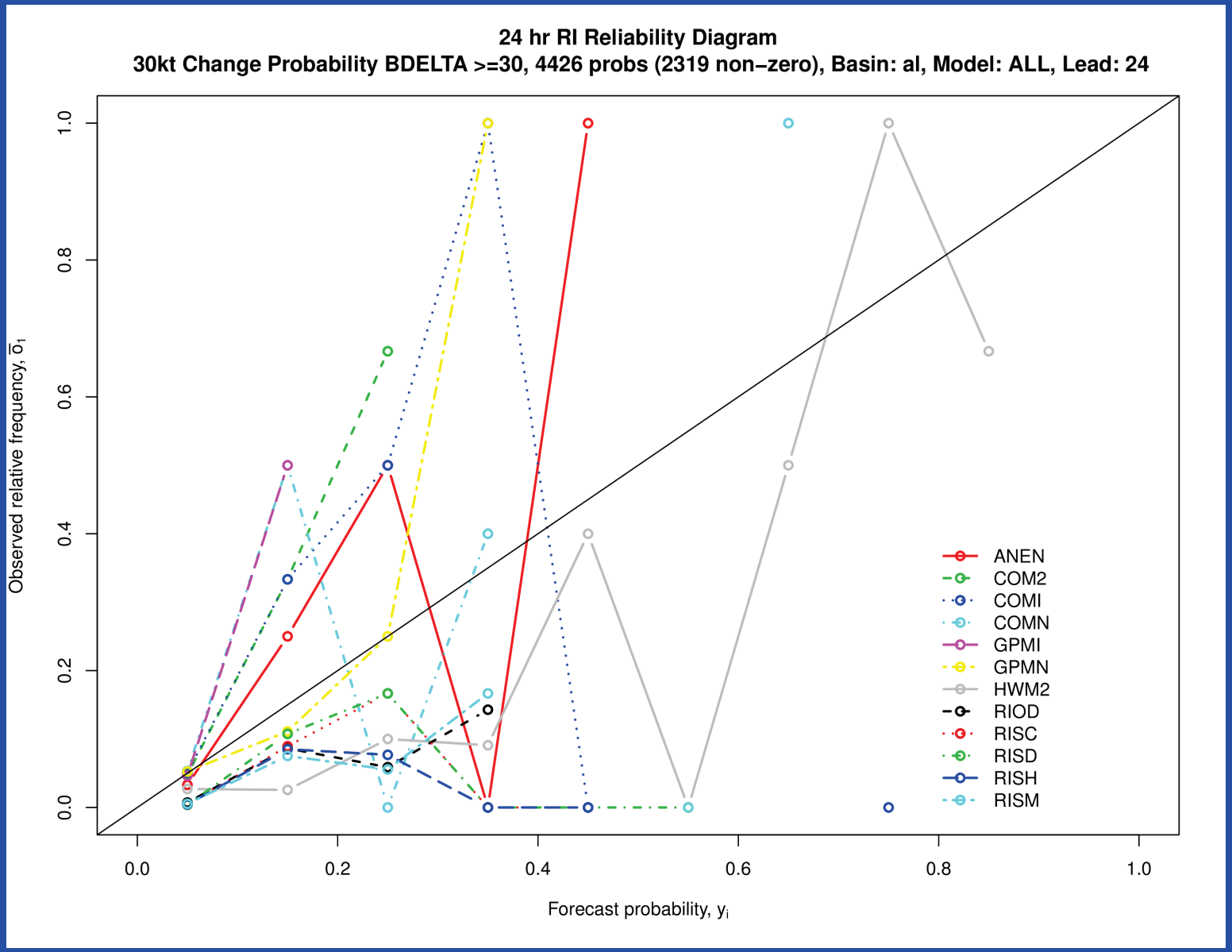
Verification

- Focus of this talk is on verification of the ensemble probabilities
- Primarily will discuss 2013-2015 retrospectives due to larger number of cases available
  - Exception is HWRF, which is based on past real-time ensemble forecasts, so case count is lower

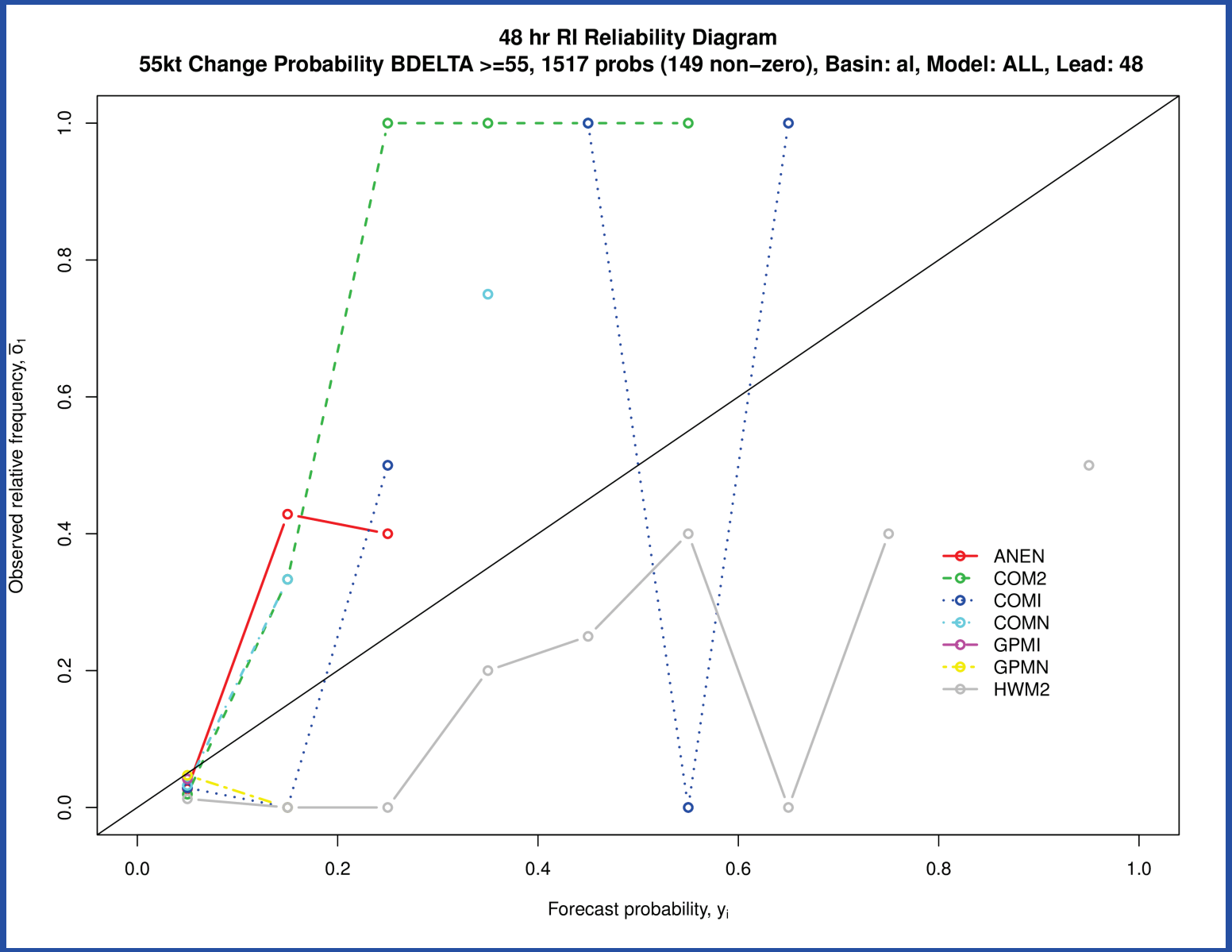
0-24 h RI Probability



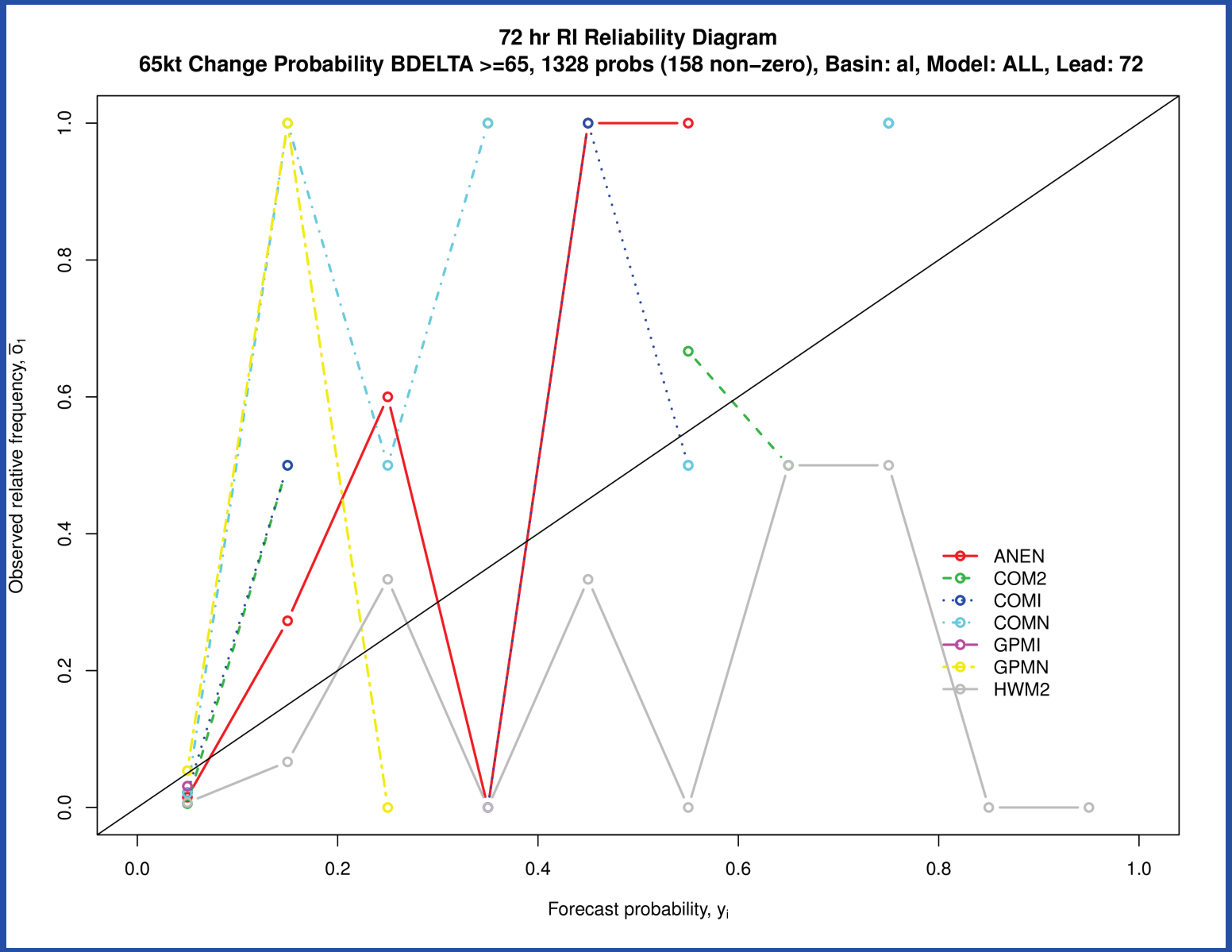
00-24 h AL RI Reliability



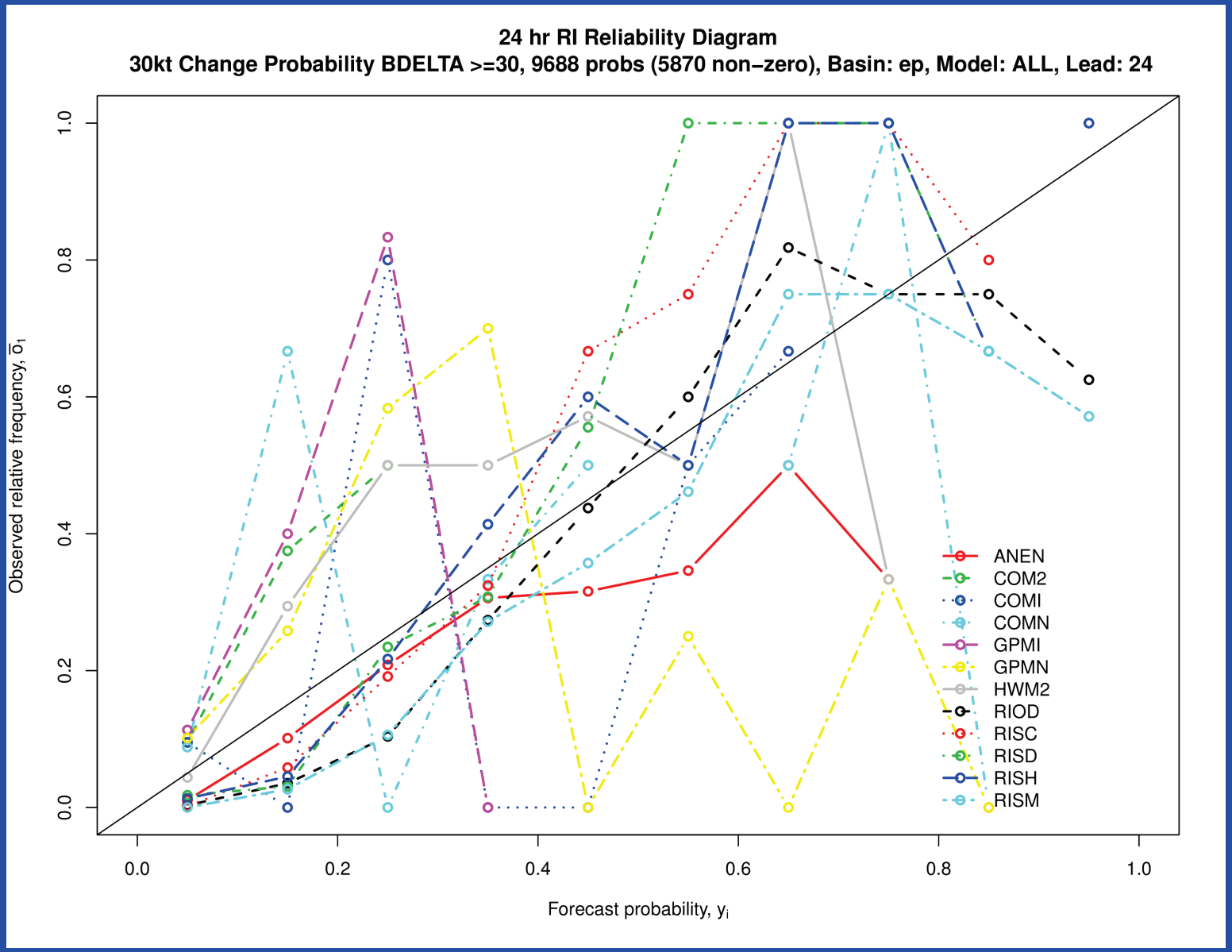
00-48 h AL RI Reliability



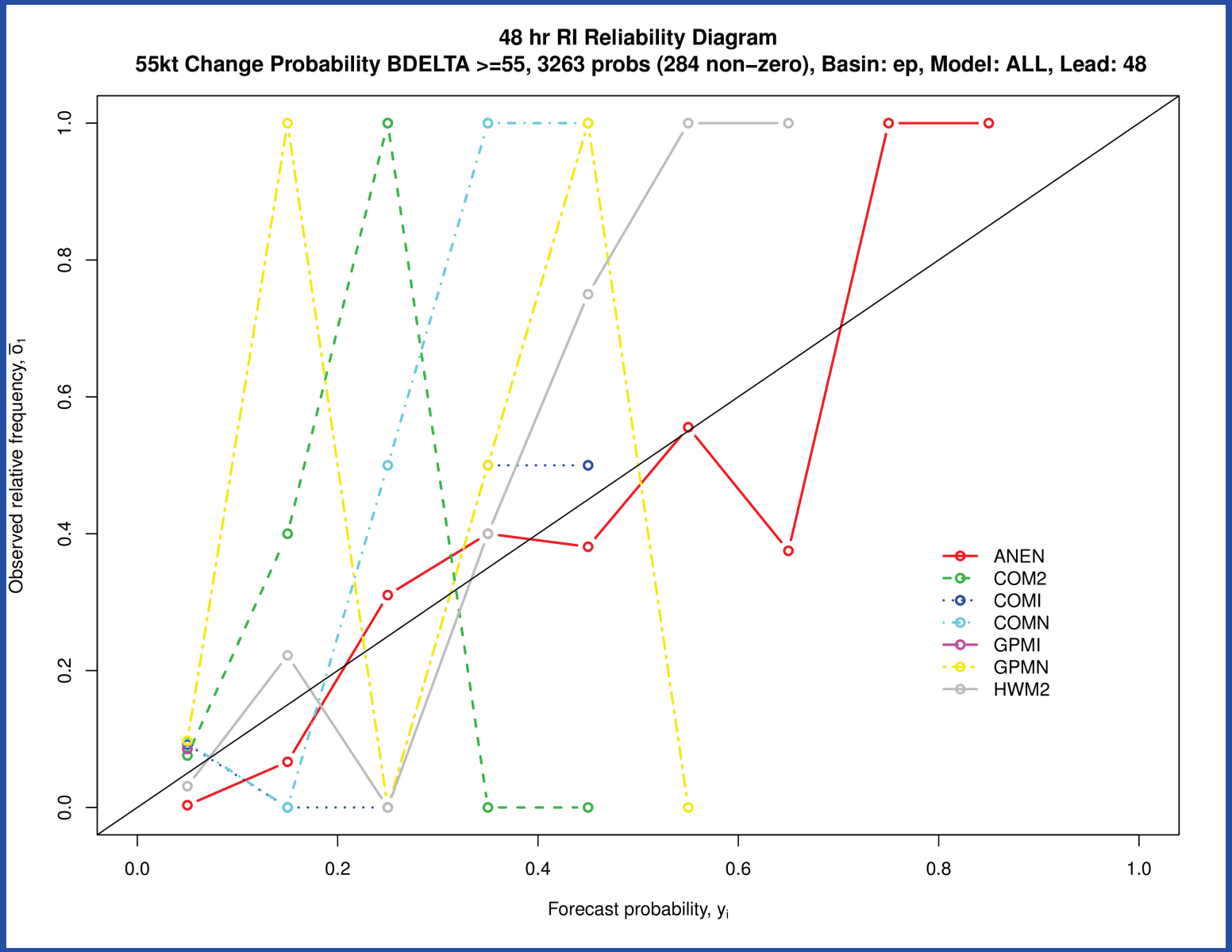
00-72 h AL RI Reliability



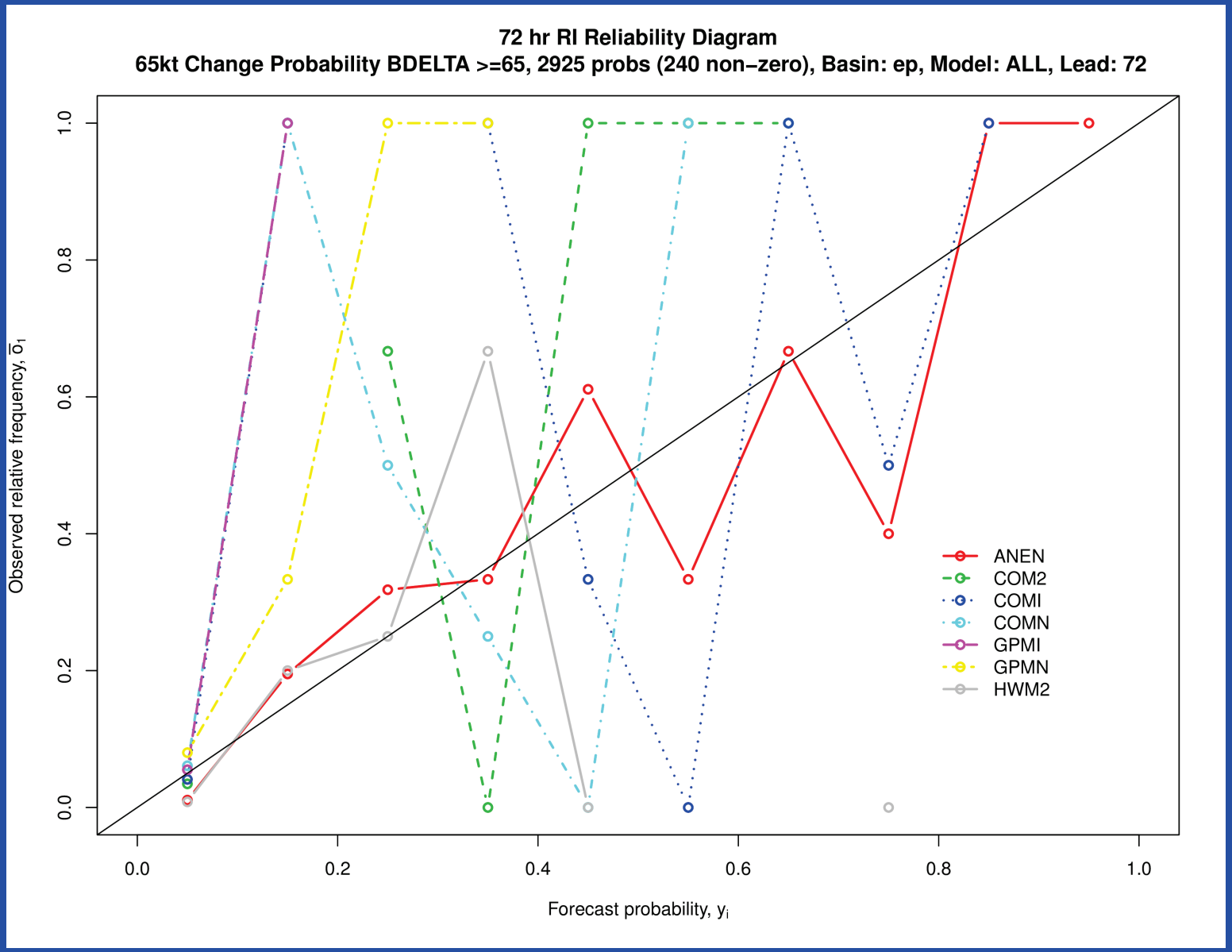
00-24 h EP RI Reliability



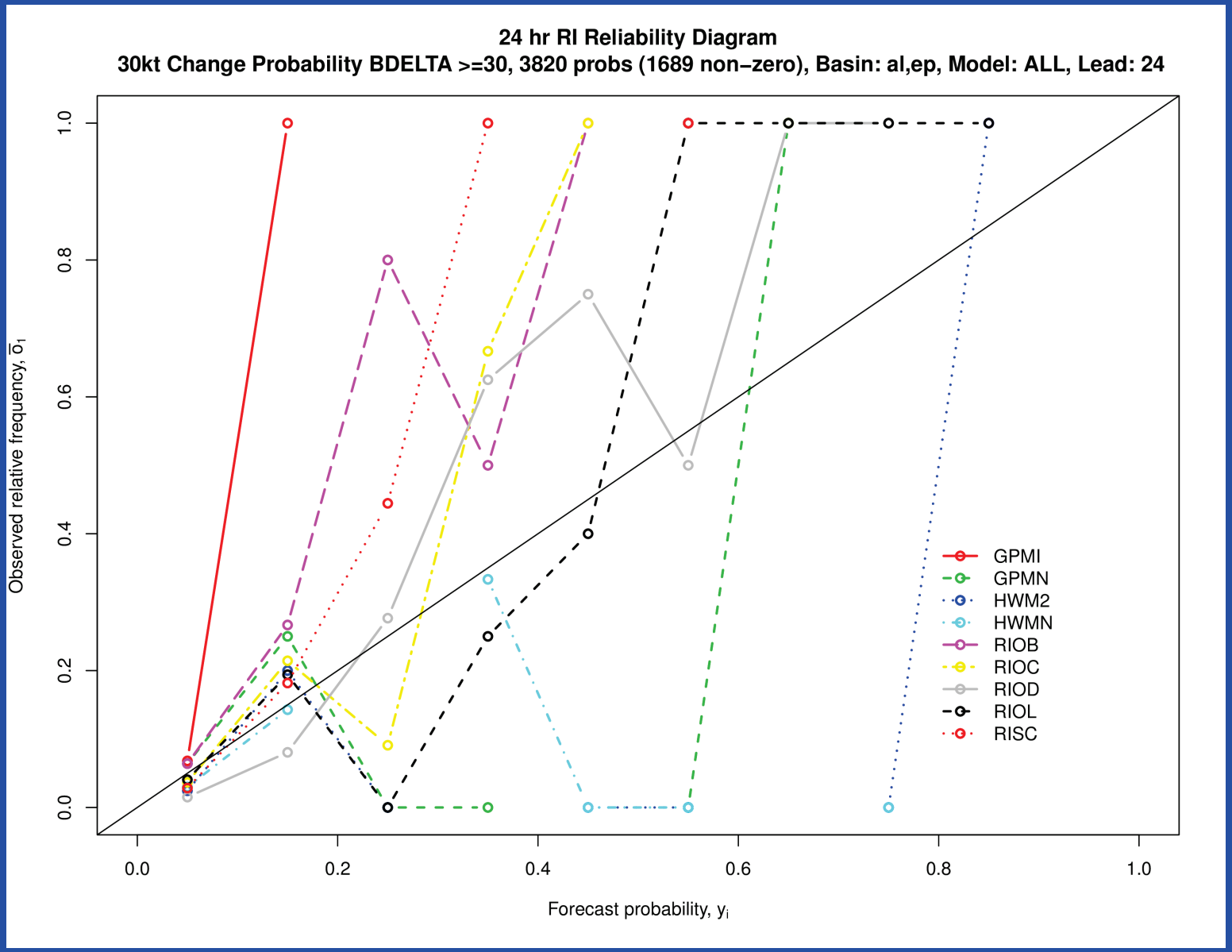
00-48 h EP RI Reliability



00-72 h EP RI Reliability



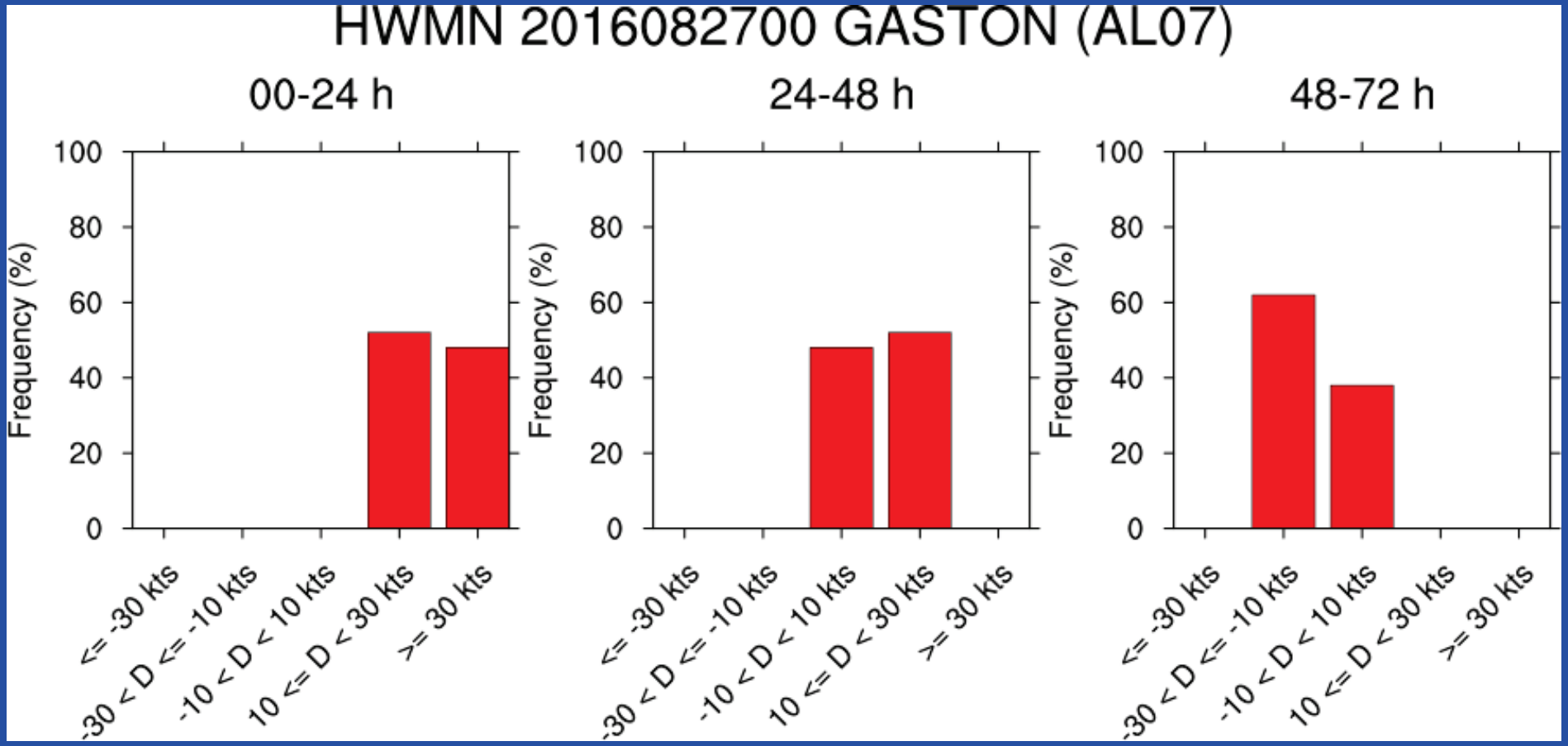
00-24 h AL, EP RI Demo



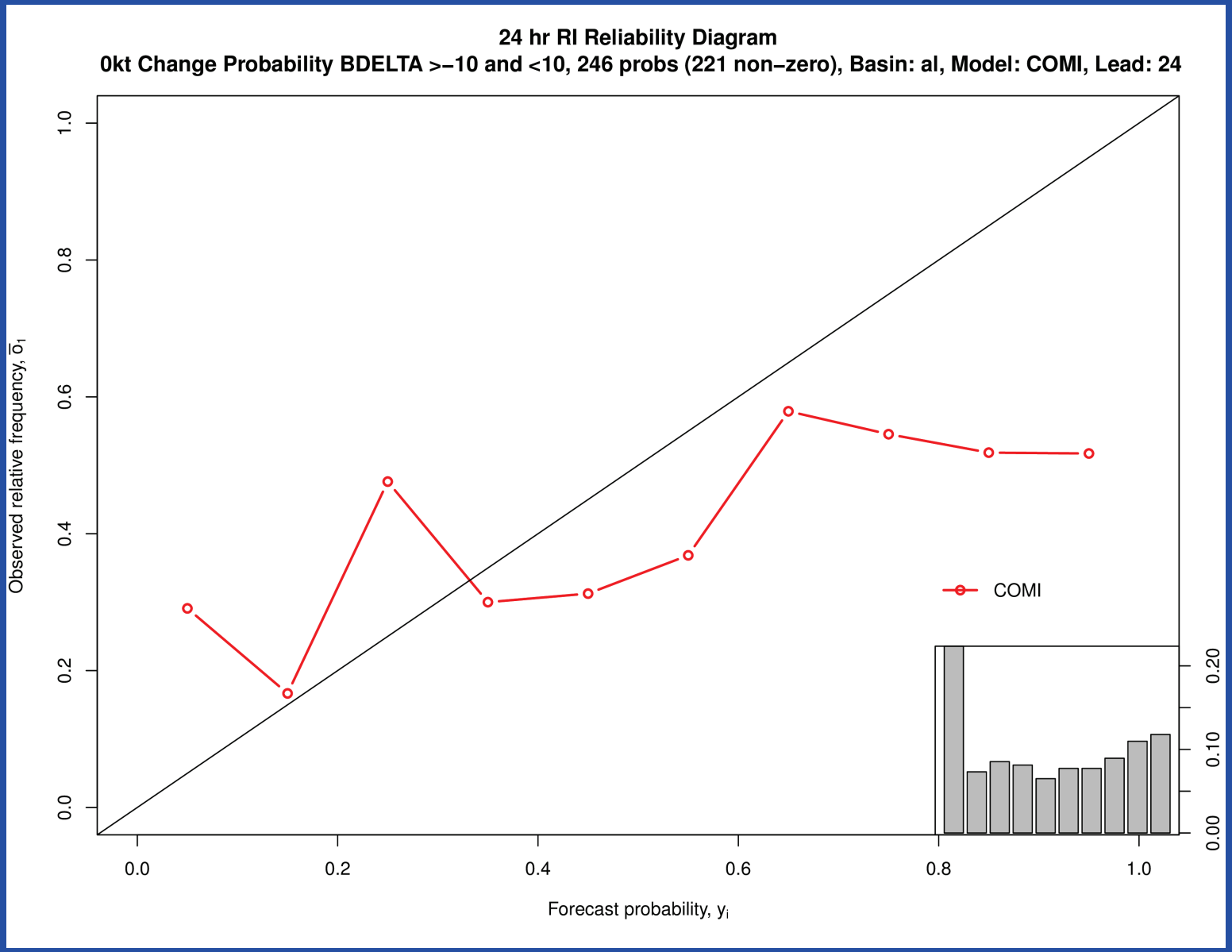
Other Intensity Categories

- Difficult to validate RI probabilities because it is by definition they are a relatively rare event
- Do any EPS have skill in other intensity change categories:
  - <= -30 kt 24 h^{-1} (rapid weakening)
  - -30 kt 24 h^{-1} < δ <= -10 kt 24 h^{-1} (weakening)
  - -10 kt 24 h^{-1} < δ < 10 kt 24 h^{-1} (steady)
  - 10 kt 24 h^{-1} <= δ < 30 kt 24 h^{-1} (intensifying)
  - >= 30 kt 24 h^{-1} (rapid intensification)

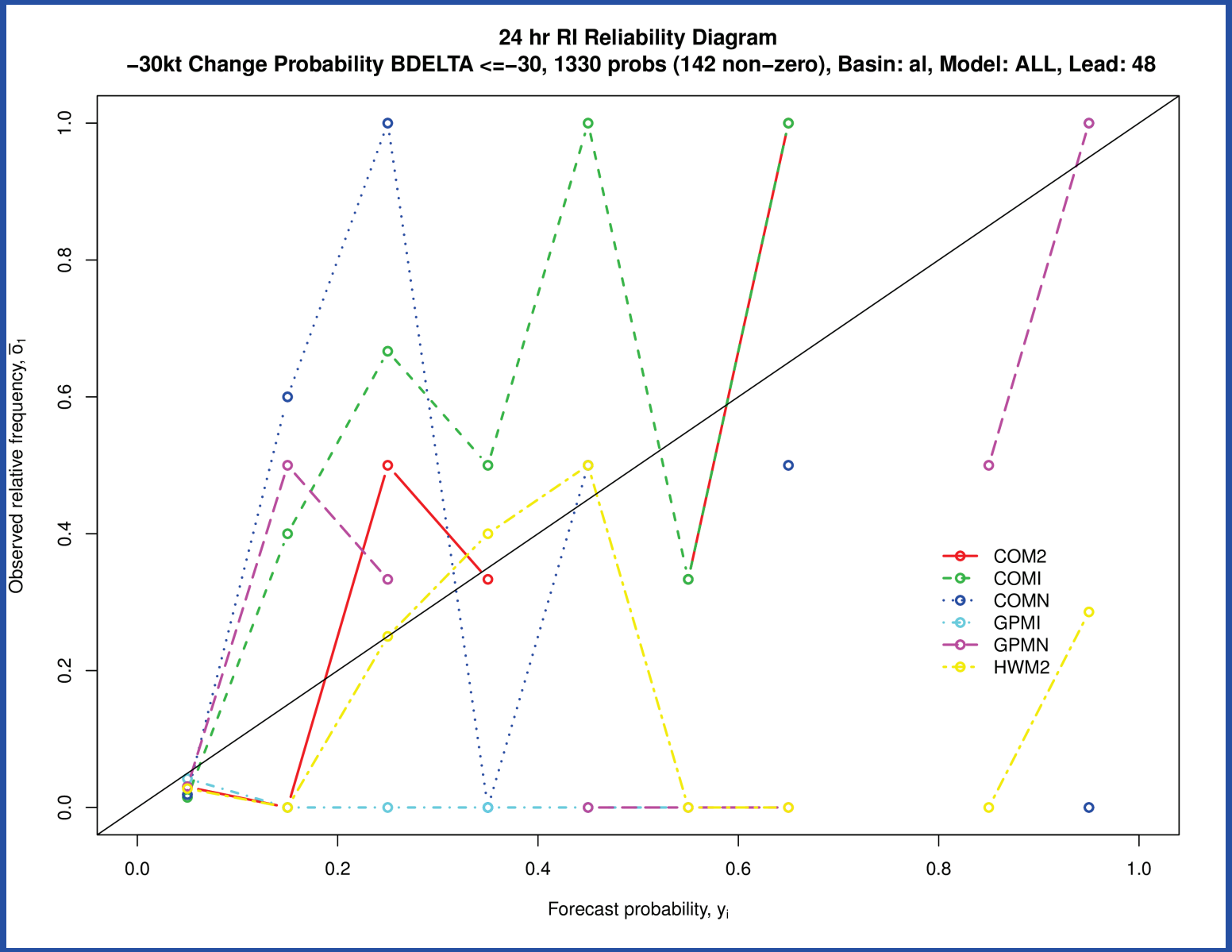
Multiple Category Example



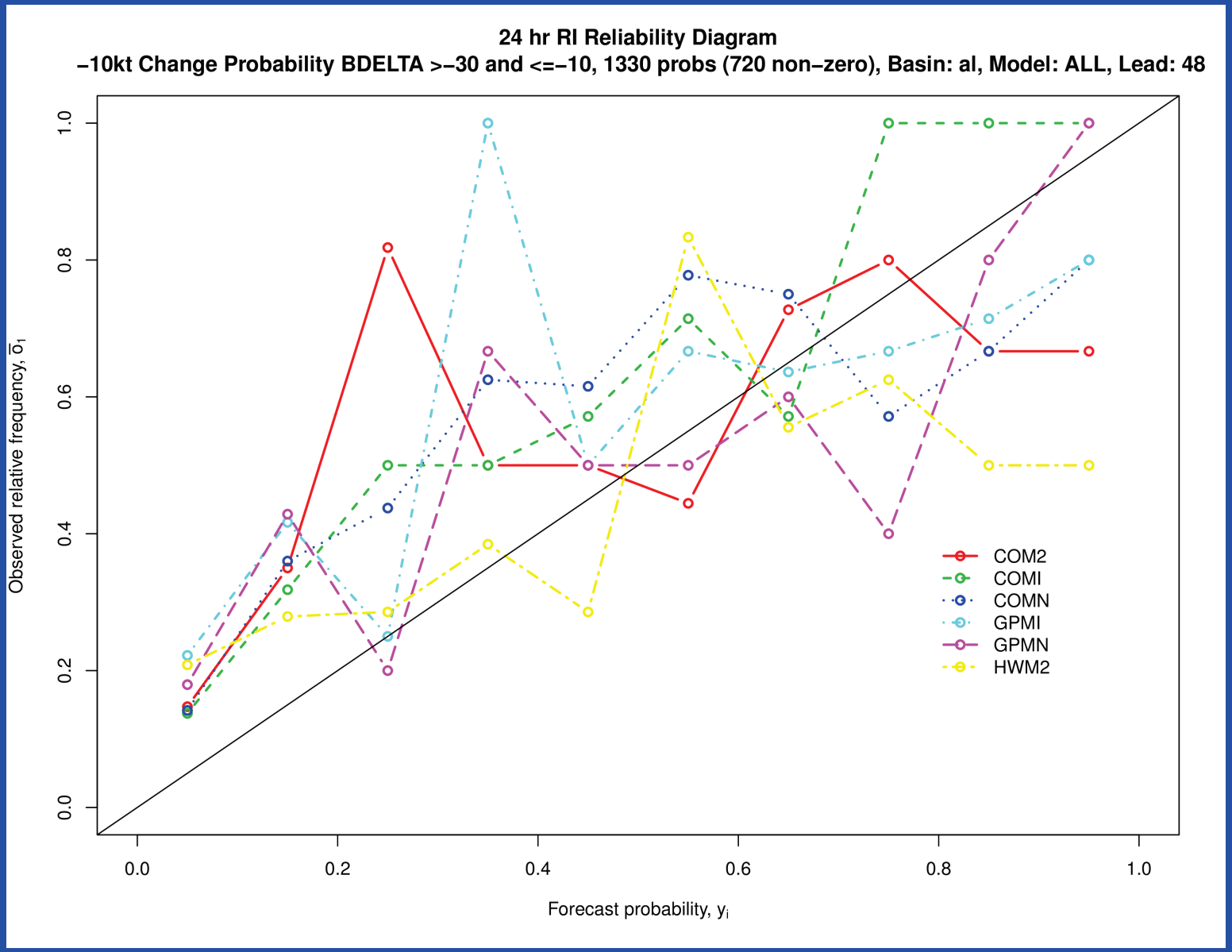
24-48 h AL, Steady



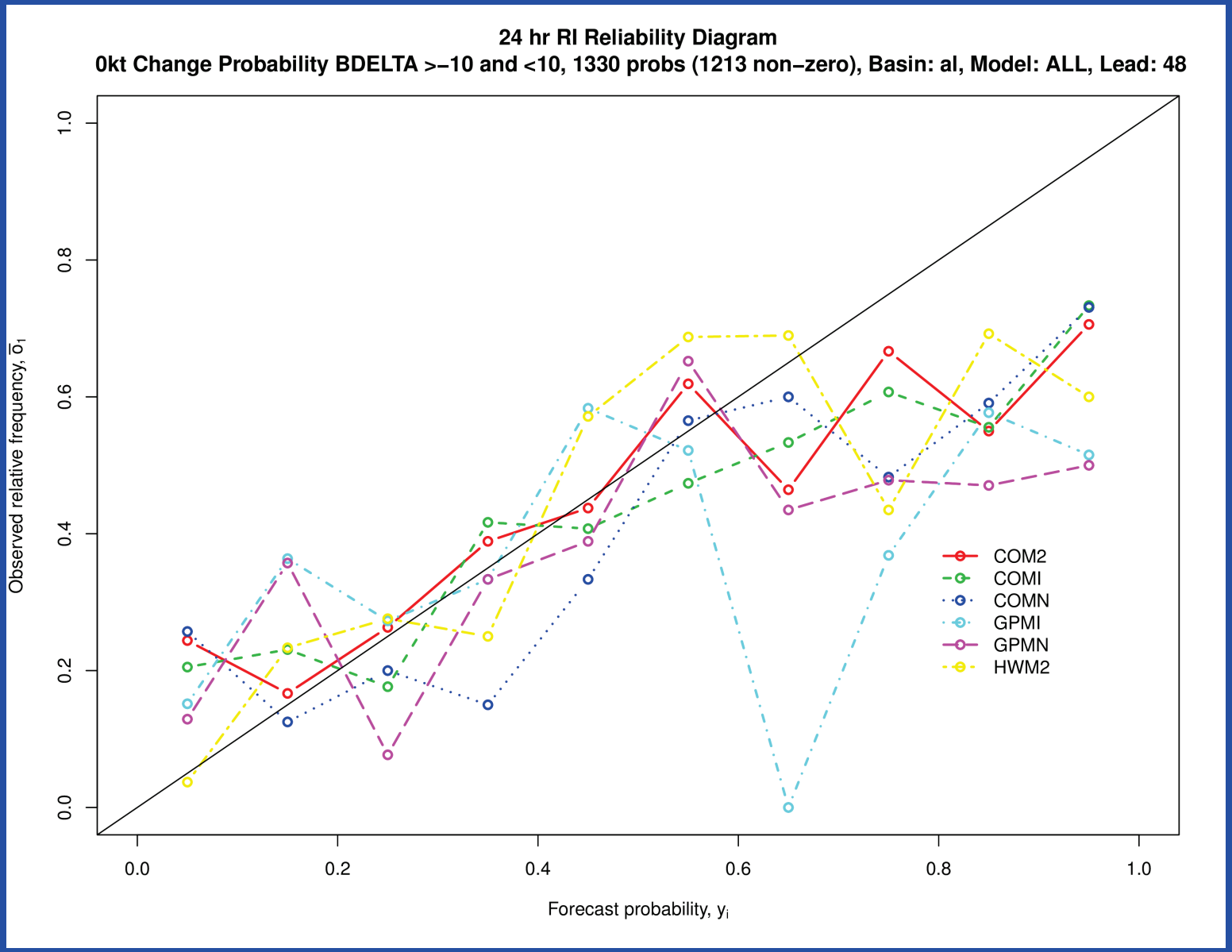
24-48 h AL, Rapid Weakening



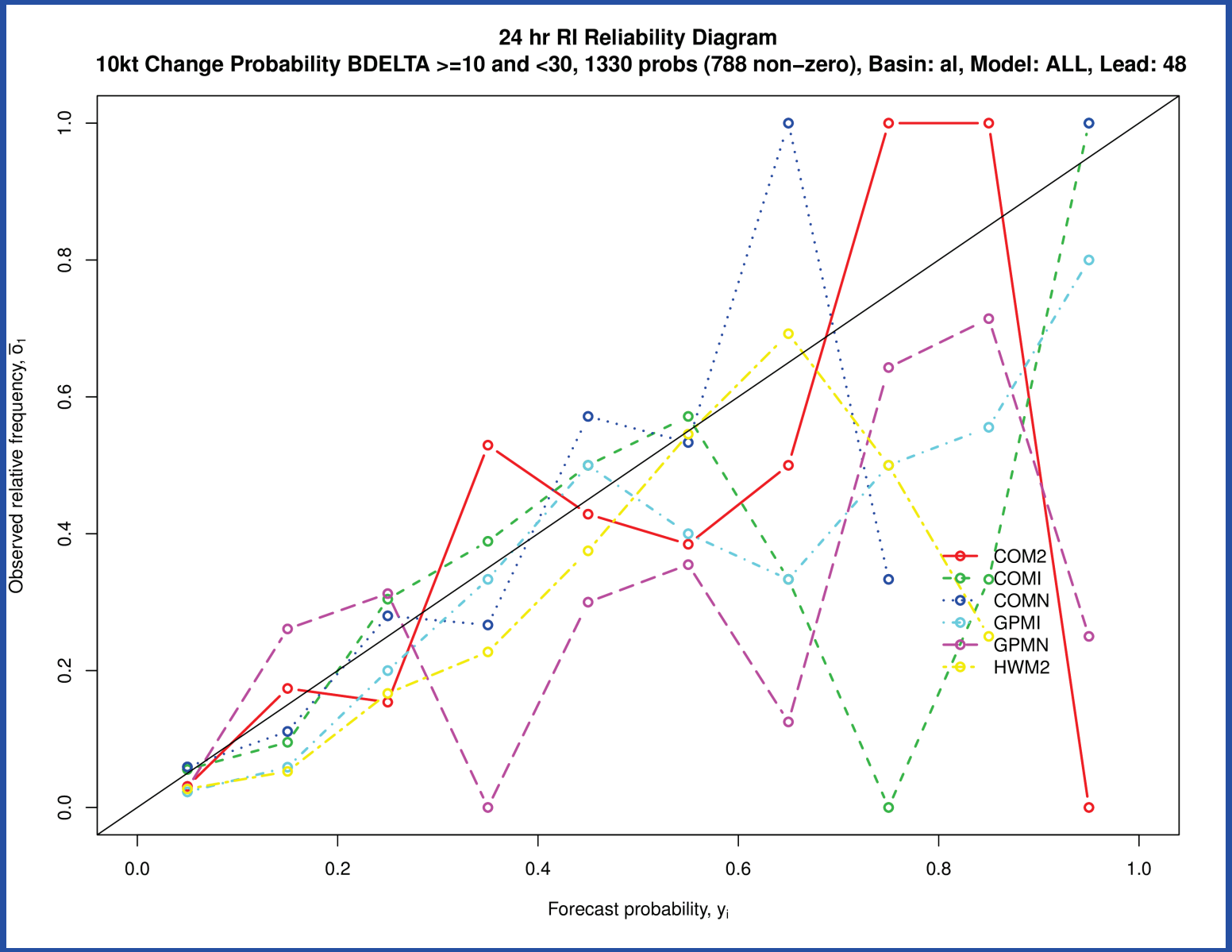
24-48 h AL, Weakening



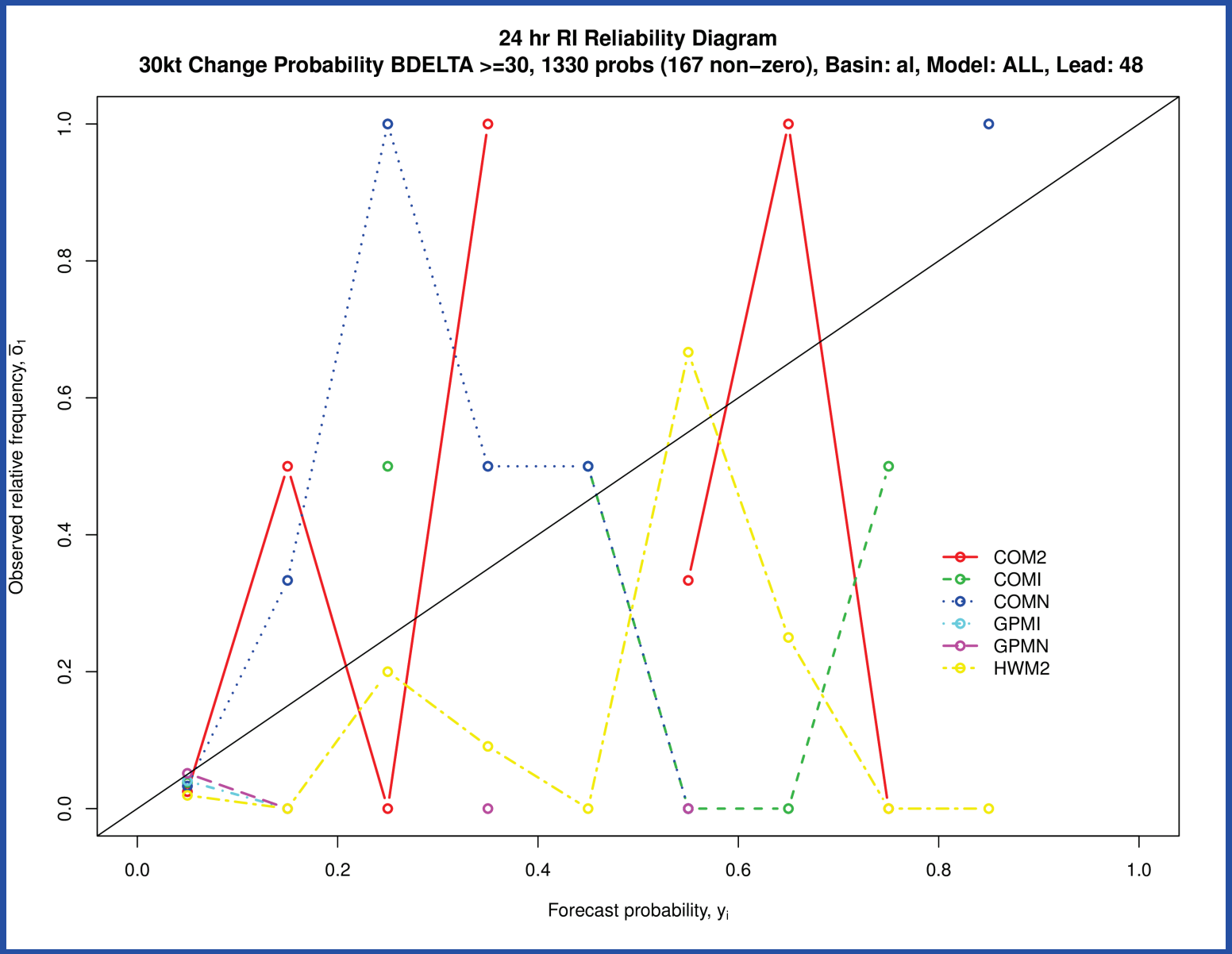
24-48 h AL, Steady



24-48 h AL, Intensifying



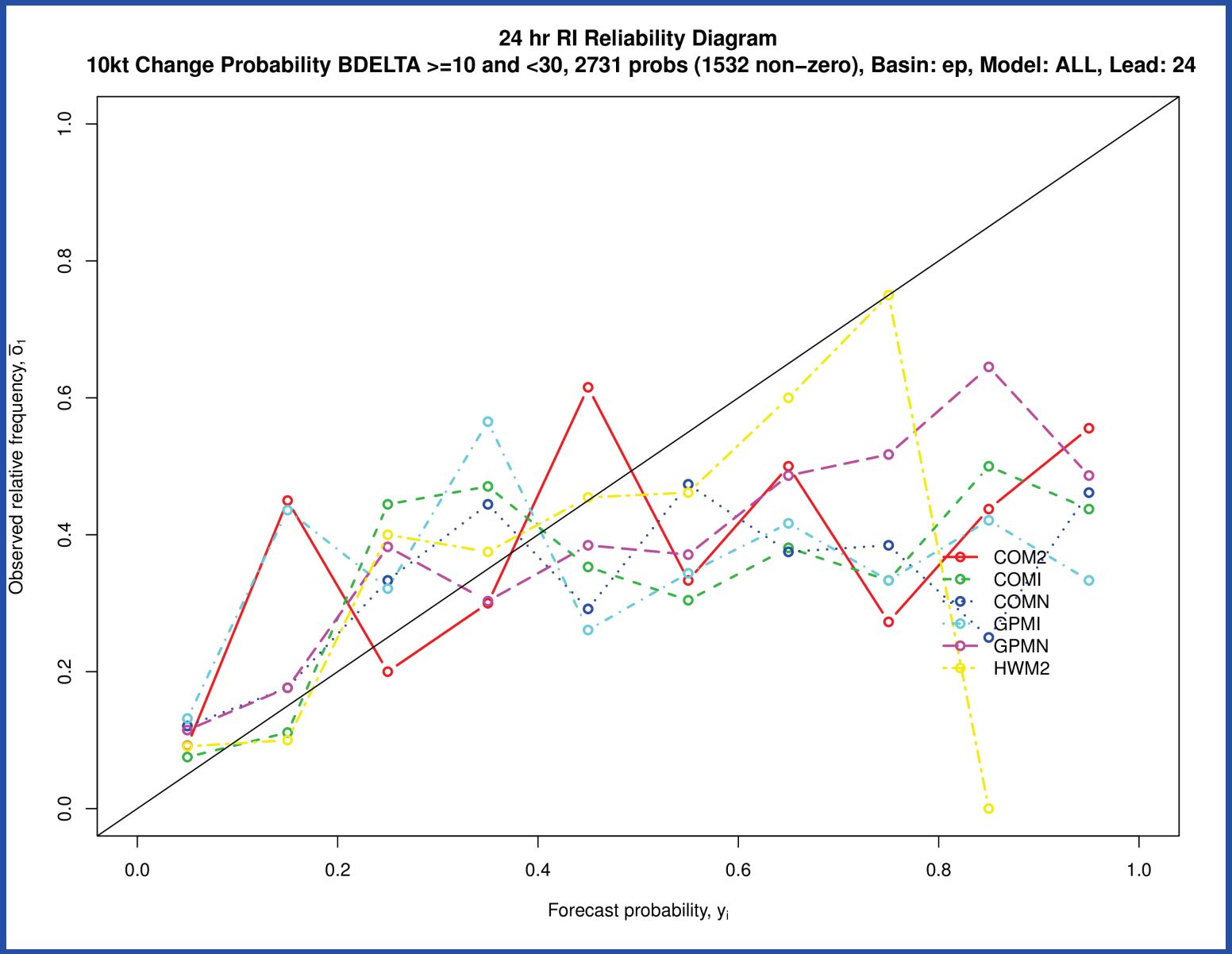
24-48 h AL, RI


24 hr RI Reliability Diagram
30kt Change Probability BDELTA >=30, 1330 probs (167 non−zero), Basin: al, Model: ALL, Lead: 48
Observed relative frequency, $\bar{o}_1$
Forecast probability, y_i
0.0
0.2
0.4
0.6
0.8
1.0
COM2
COMI
COMN
GPMI
GPMN
HWM2

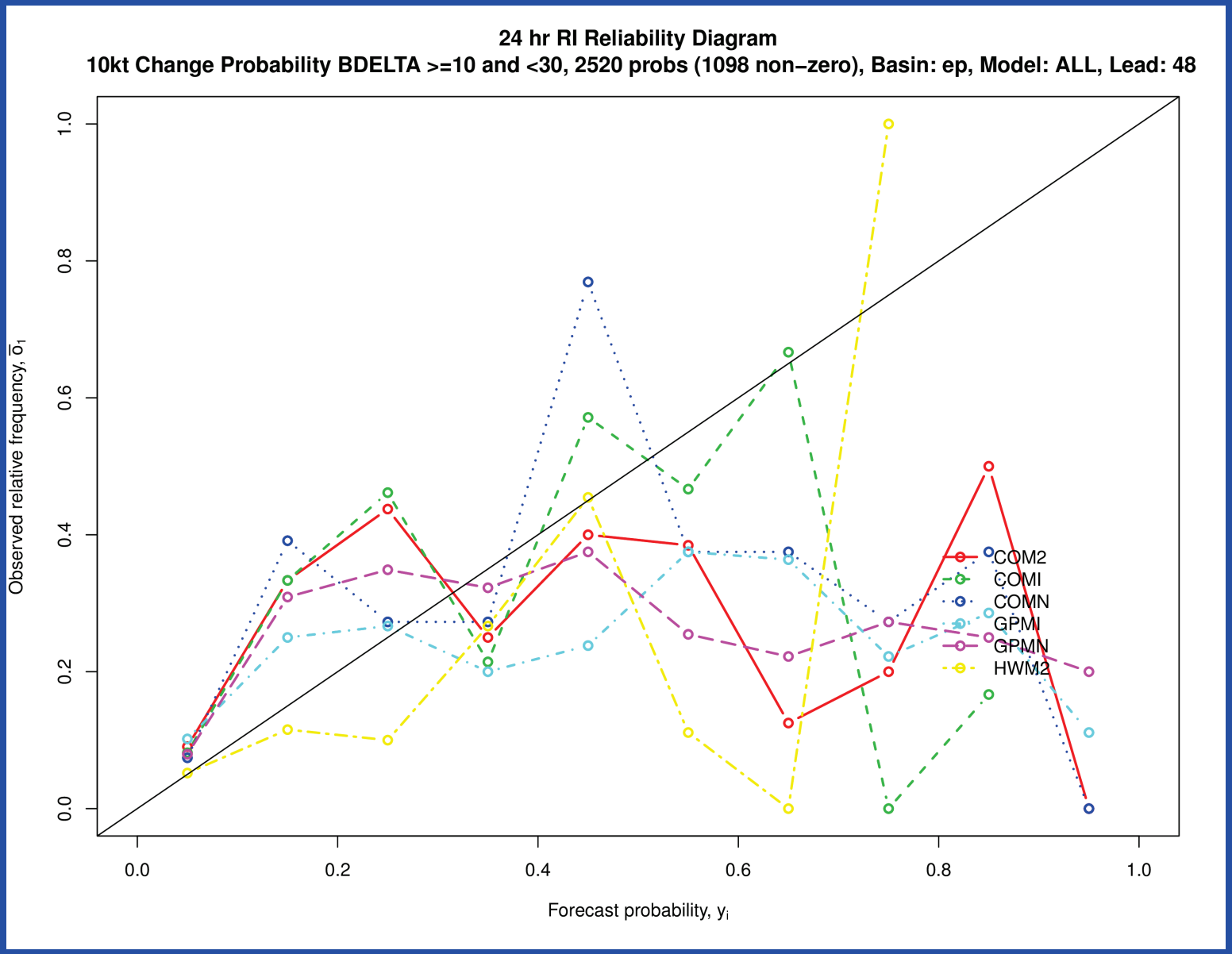
EP Intensifying

- For the most part, little difference in performance for non-rapid intensity changes during 0-24, 24-48, 48-72 h periods
- Exception is in EP Intensifying cases:

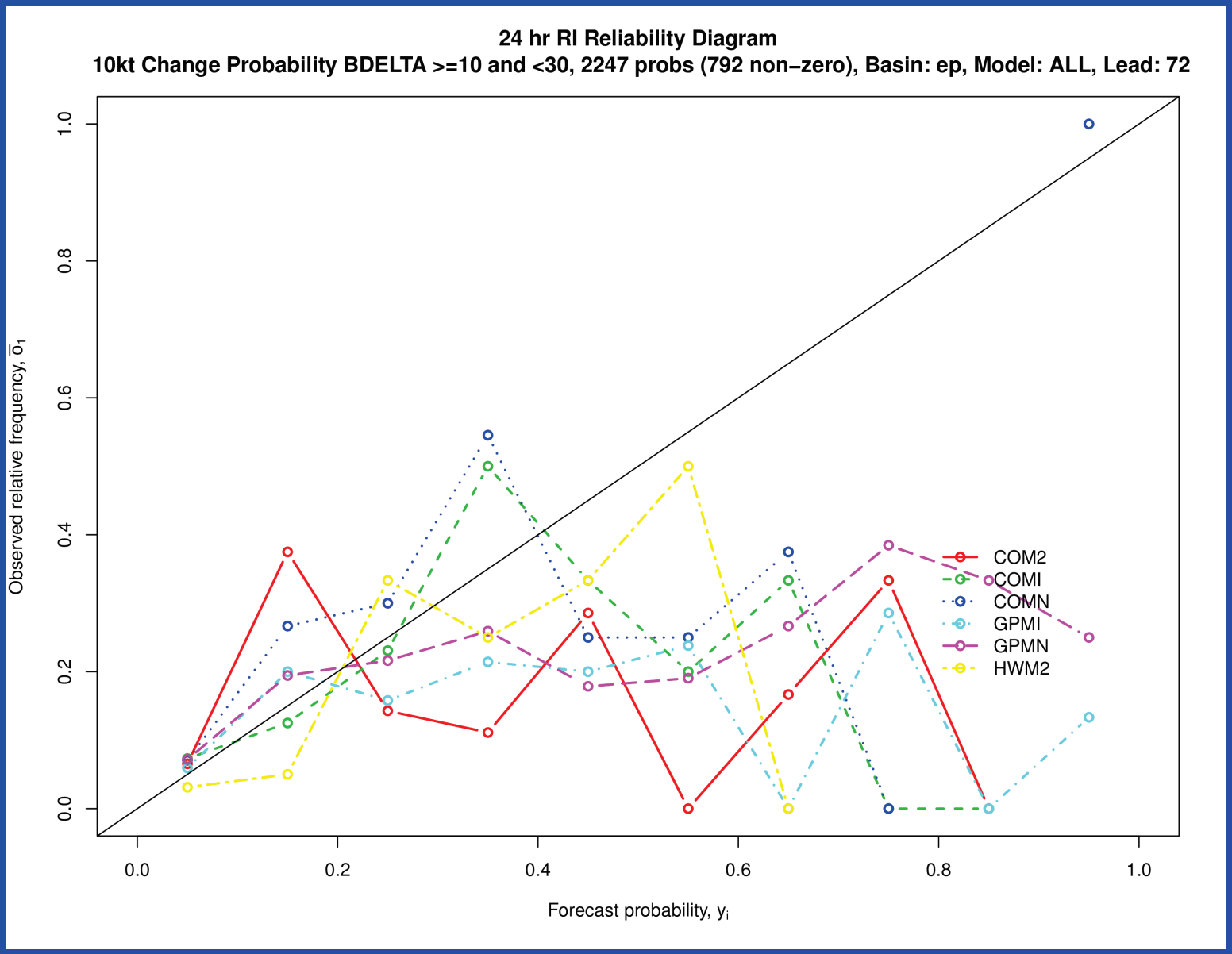
00-24 h EP, Intensifying



00-24 h EP, Intensifying



48-72 h EP, Intensifying



Summary

- Technical aspects in place to deliver ensemble-based probabilistic intensity change to NHC
- Retrospective forecasts show varied skill at predicting varied RI categories
- More skill at predicting more typical intensity change categories
  - Evidence of over-confidence
- Eastern Pacific shows some probability drift with time for intensifying TCs

Future Directions

- In-depth validation of probabilities
- Increase Forecaster participation
- Increase reliability of forecast delivery
- Understanding what situations lead to over-confidence in the ensemble (i.e., near 100% probabilities)
- Multi-model probabilistic intensity change (i.e., combination of HWRF, COAMPS-TC)