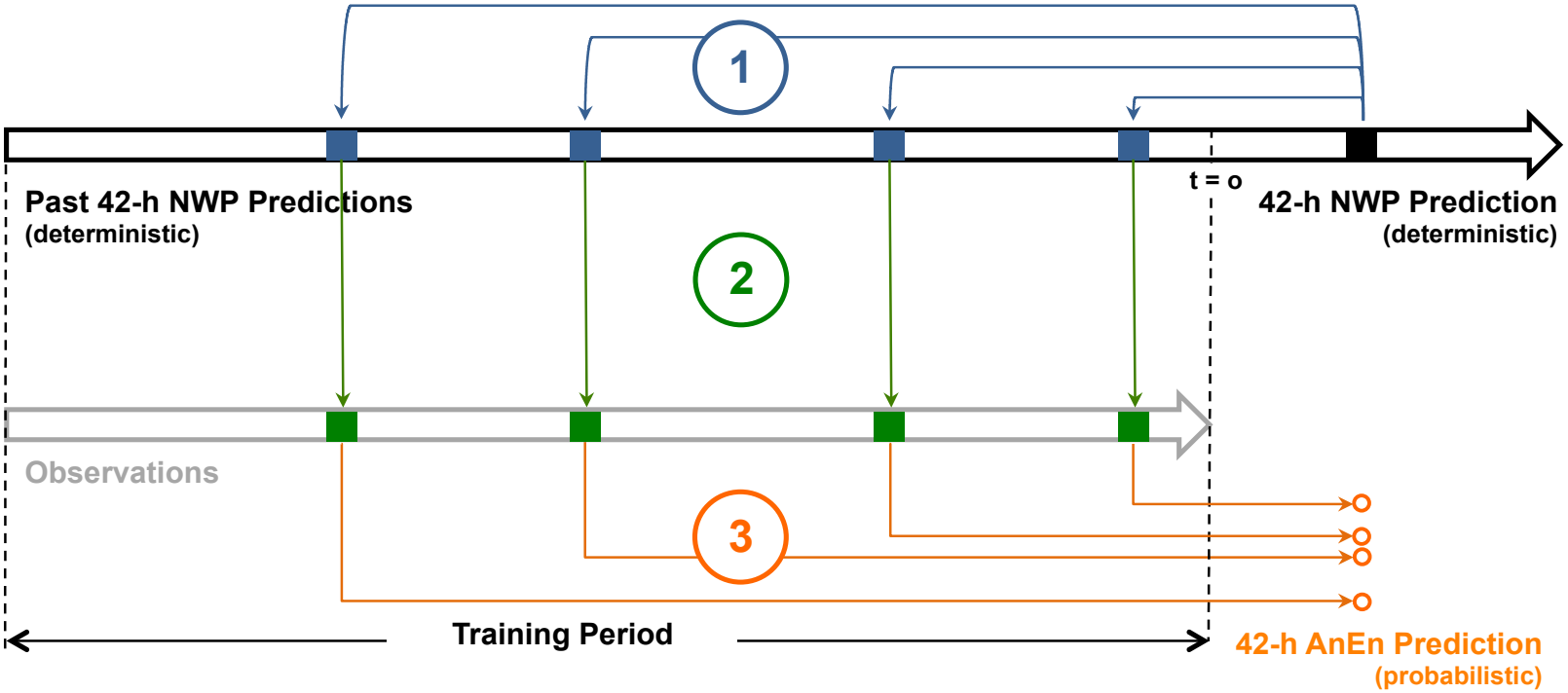
A Hybrid, NWP-Analog Ensemble

F. Anthony Eckel
Microsoft, Bellevue, WA

Luca Delle Monache
National Center for Atmospheric Research, Boulder, CO

Analog Ensemble (AnEn)*

1. From today's model forecast, find *n* similar past predictions by the same model
2. Obtain the verifying observation from each analog
3. Each observation is an ensemble member for today's forecast



*Luca Delle Monache, F. Anthony Eckel, Daran L. Rife, Badrinath Nagarajan, and Keith Searight, 2013: **Probabilistic Weather Prediction with an Analog Ensemble**, *Mon. Wea. Rev.*, **141**, 3498–3516.

Delle Monache et al. (2013) Results

AnEn: Analog Ensemble

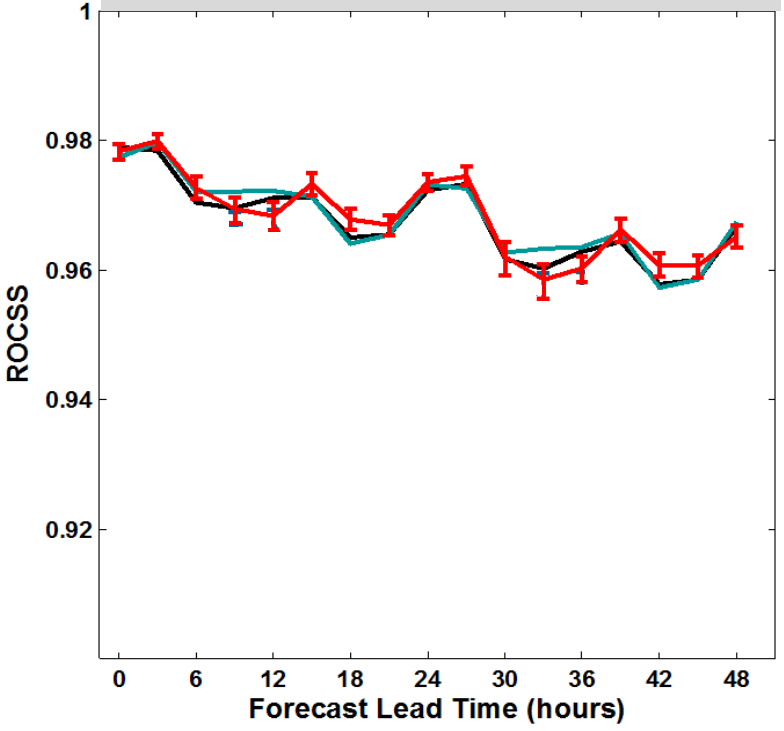
LR: Logistic Regression

Used deterministic, higher-resolution model forecast

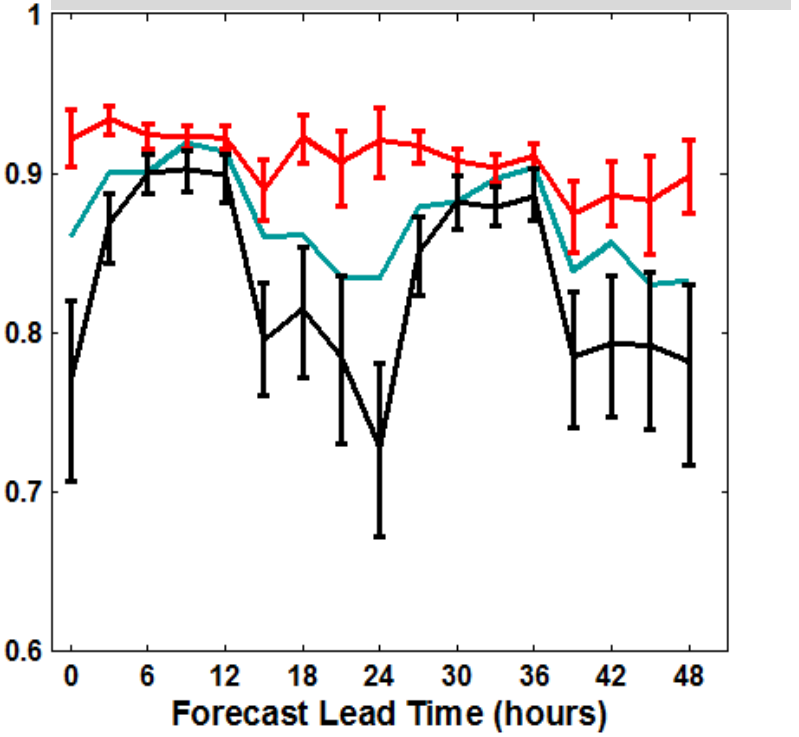
EMOS: Ensemble MOS

*Used 21-member NWP ensemble of lower-resolution model forecasts, **at ~2× the cost***

Pr(2-m temperature < 15 °C)

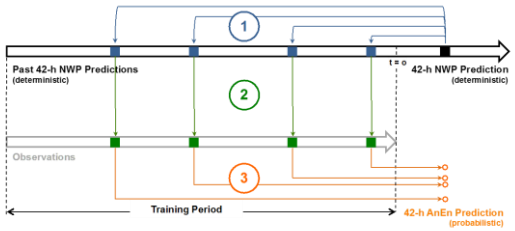


Pr(10-m wind speed > 10 m s^{-1})

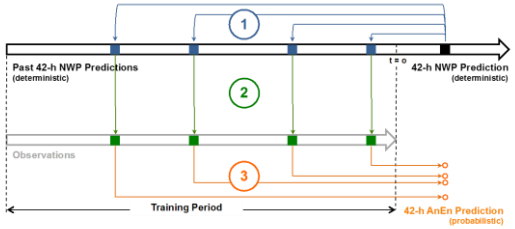


Hybrid NWP-Analog Ensemble (HyEn)

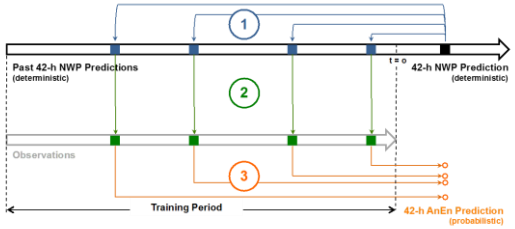
analog ensemble w/ REPS #1



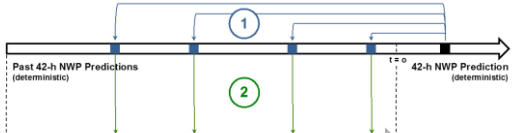
analog ensemble w/ REPS #2



analog ensemble w/ REPS #3



analog ensemble w/ REPS #4



NWP Ensemble – dynamically capture flow- dependent error growth

…and then…

Analog Members – provide additional sampling, and down-scaling calibration

HyEn

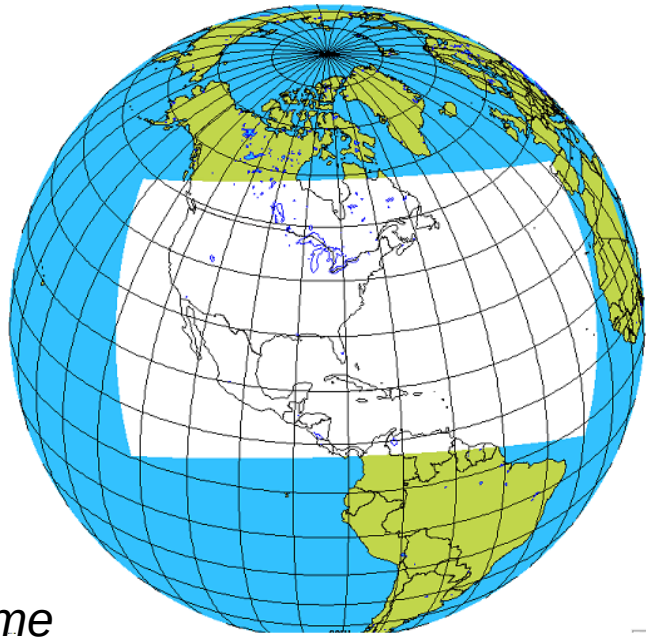
Sampling strategy?

Forecast Data

for AnEn

High-res Regional Global Environment Multiscale (GEM)

- Model: *GEM 4.2.0*
- Grid Spacing: ***~15 km***



≅ $$$

Regional Ensemble Prediction System (REPS*)

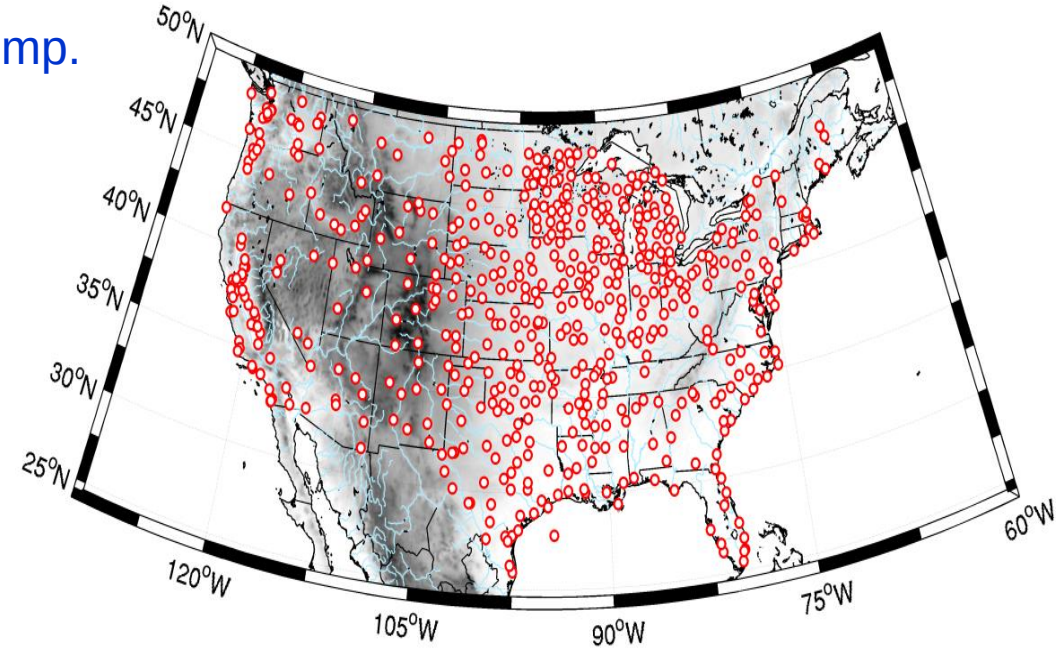
for EMOS & HyEn

- Model: *GEM 4.2.0*
- Grid Spacing: ***~33 km***
- Forecasts: *Used only 12Z cycle, 0 to 48-h lead time*
- # of Members: *Used only **10** (of 21)*
- Initial & Boundary Conditions: *21-member Global EPS*
- Stochastic Physics: *Markov Chains on physical tendencies*

*Li, X., M. Charron, L. Spacek, and G. Candille, 2008: A regional ensemble prediction system based on moist targeted singular vectors and stochastic parameter perturbations. *Mon. Wea. Rev.*, **136**, 443–462.

Ground Truth Data

- Locations: 550 hourly METAR Surface Observations within CONUS
- Period: ~15 months,1 May 2010 – 31 July 2011
- Variables: 10-m wind speed, 2-m temp.



Member Selection:
Repeat vs. No-Repeat

Cycle
12Z on June 4, 2011
(Date #**401**)

Location
KSEA (SeaTac Apt., WA)

Lead Time
36-h

Variable
2-m Temperature

Analog Date #

Analog Rank →

REPS Member #	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8* ...
1	91	109	100	74	52	56	110	49 ...
2	72	147	127	153	123	49	110	120 ...
3	58	56	72	51	99	73	101	97 ...
4	73	87	98	64	72	82	94	99 ...
5	100	97	98	47	78	72	112	94 ...
6	369	31	121	131	63	64	62	29 ...
7	110	153	147	54	124	93	399	152 ...
8	93	72	92	73	110	52	78	99 ...
9	120	110	127	36	98	58	148	146 ...
10	82	53	73	72	36	52	99	55 ...

HyEn10×7

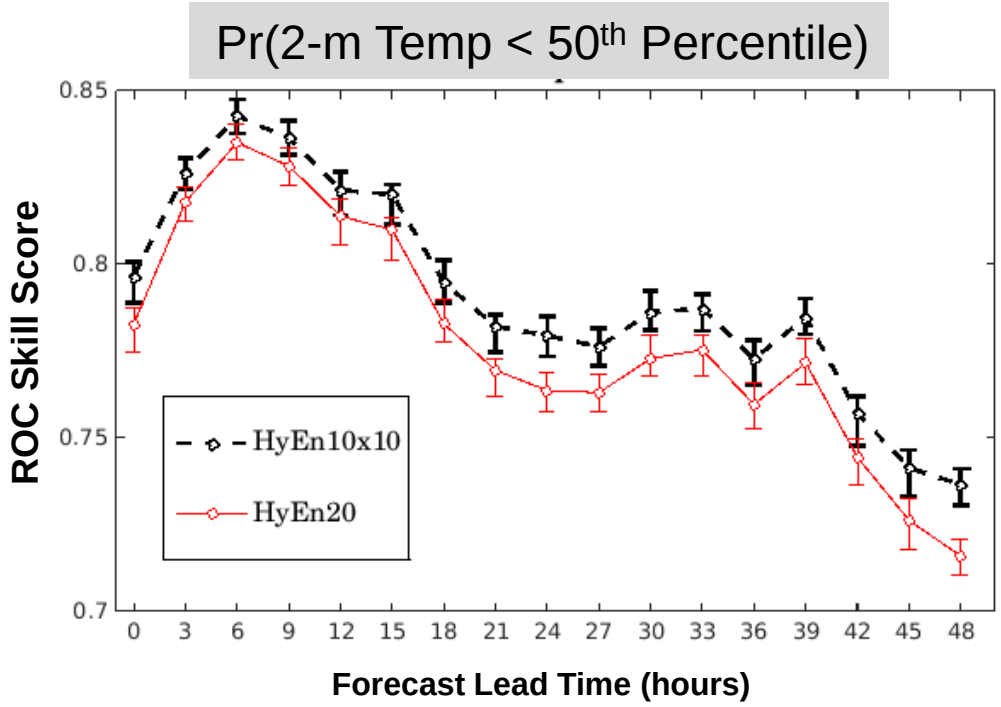
Verifying Obs. (°C)

REPS Member #	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8* ...
1	20	21	19	26	23	18	19	13 ...
2	18	23	17	17	21	13	19	17 ...
3	20	18	18	17	21	21	21	18 ...
4	21	25	17	16	18	18	23	21 ...
5	19	18	17	16	18	18	21	23 ...
6	14	16	21	16	19	16	17	14 ...
7	19	17	23	23	24	22	25	22 ...
8	22	18	19	21	19	23	18	21 ...
9	17	19	17	16	17	20	17	19 ...
10	18	24	21	18	16	23	21	18 ...

HyEn40
(No-Repeat members)

Member Selection:
Repeat vs. No-Repeat

✓ *Allowing repeat (auto-weighted) members works best*



10-m Wind Speed Results

AnEnφ: Analog Ensemble
φ (*optimal*) members

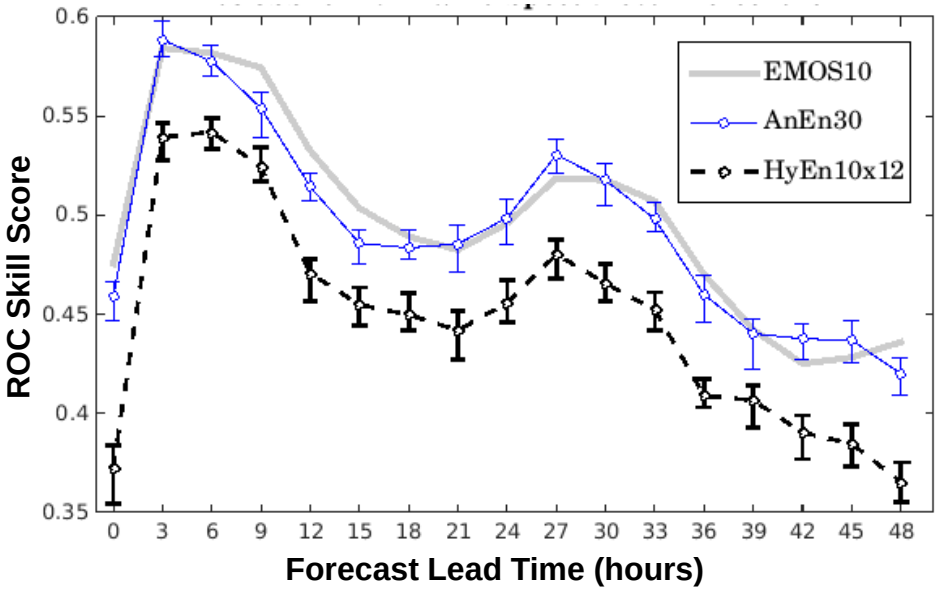
Used deterministic, higher-resolution model forecast

EMOS10: Ensemble MOS
using ensemble mean & spread

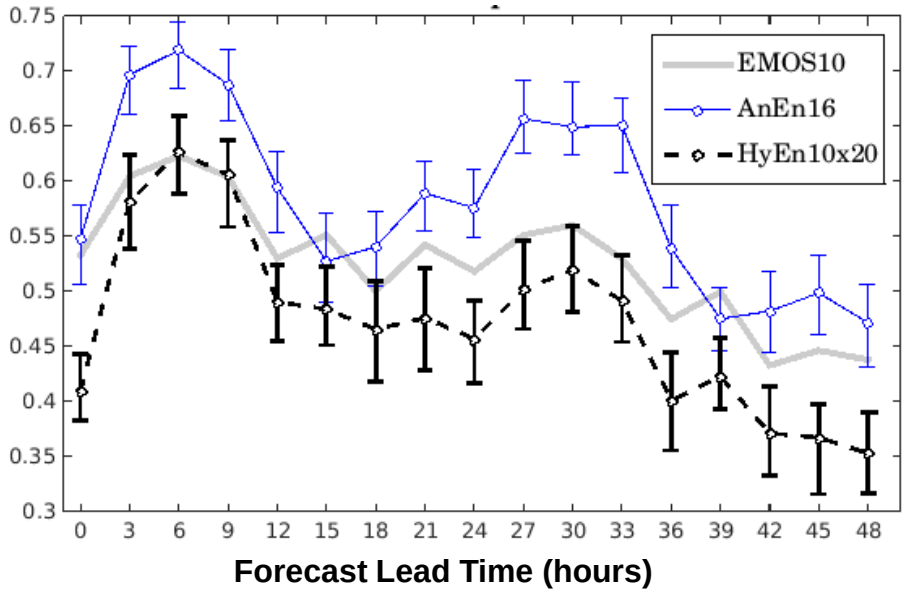
HyEn10×γ: Hybrid Ensemble
γ (*optimal*) analogs on each NWP mbr

*Used 10 members from NWP ensemble of lower-resolution model forecasts, **at ~same cost***

Pr(10-m Wind Speed > 50th Percentile)



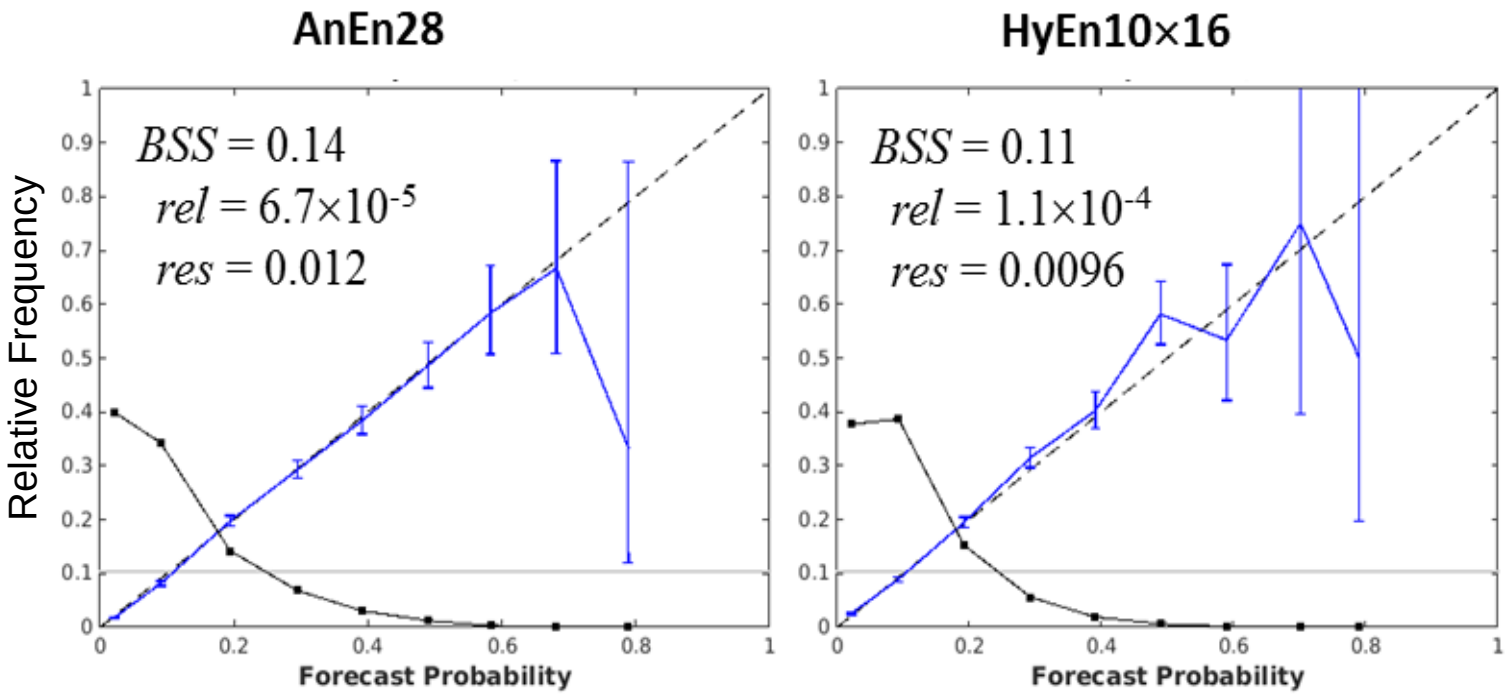
Pr(10-m Wind Speed > 98th Percentile)



10-m Wind Speed Results

10-m Wind Speed ≥ 90th percentile

30-h Lead Time



Somewhat overdispersive (too much spread)

rel & *res* worse than AnEn

2-m Temperature Results

AnEnφ: Analog Ensemble
φ (*optimal*) members

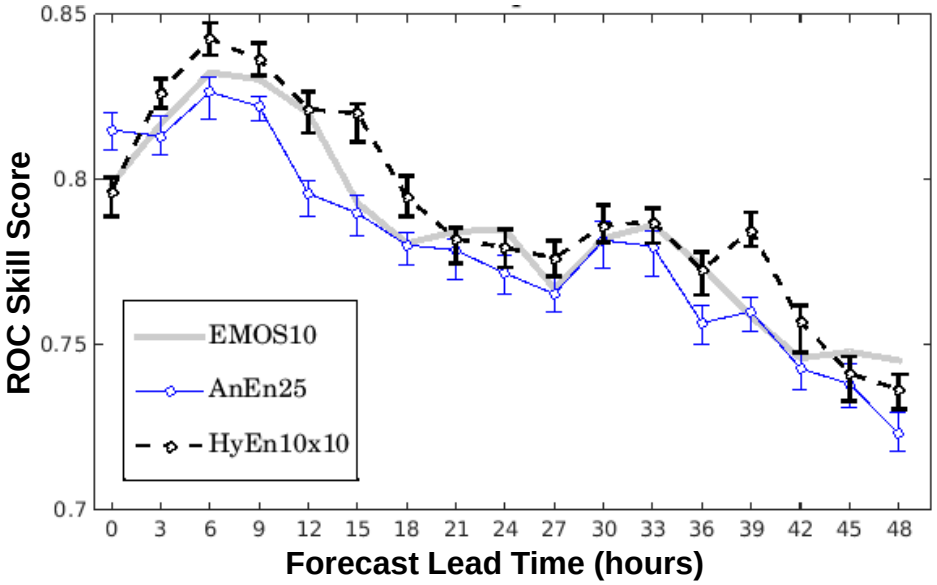
Used deterministic, higher-resolution model forecast

EMOS10: Ensemble MOS
using ensemble mean & spread

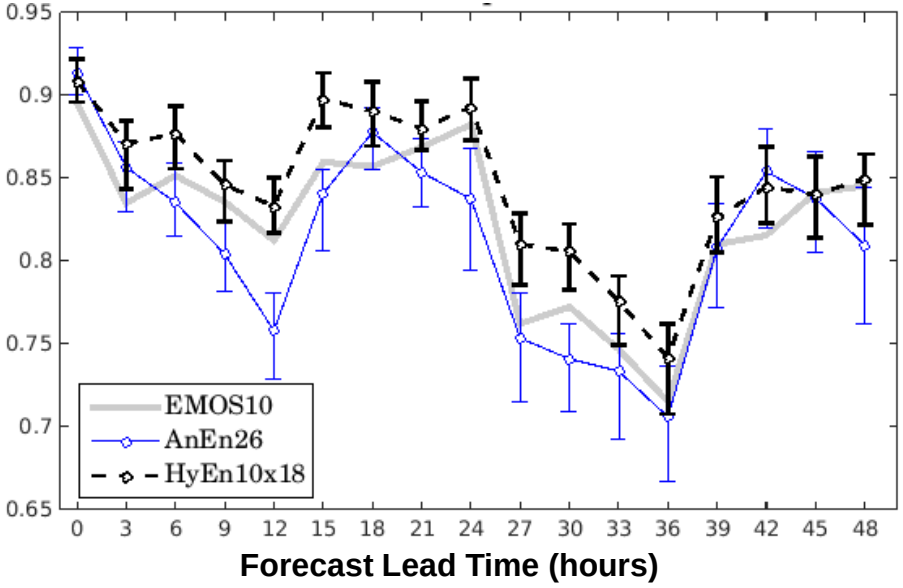
HyEn10×γ: Hybrid Ensemble
γ (*optimal*) analogs on each NWP mbr

*Used 10 members from NWP ensemble of lower-resolution model forecasts, **at ~same cost***

Pr(2-m Temp < 50th Percentile)



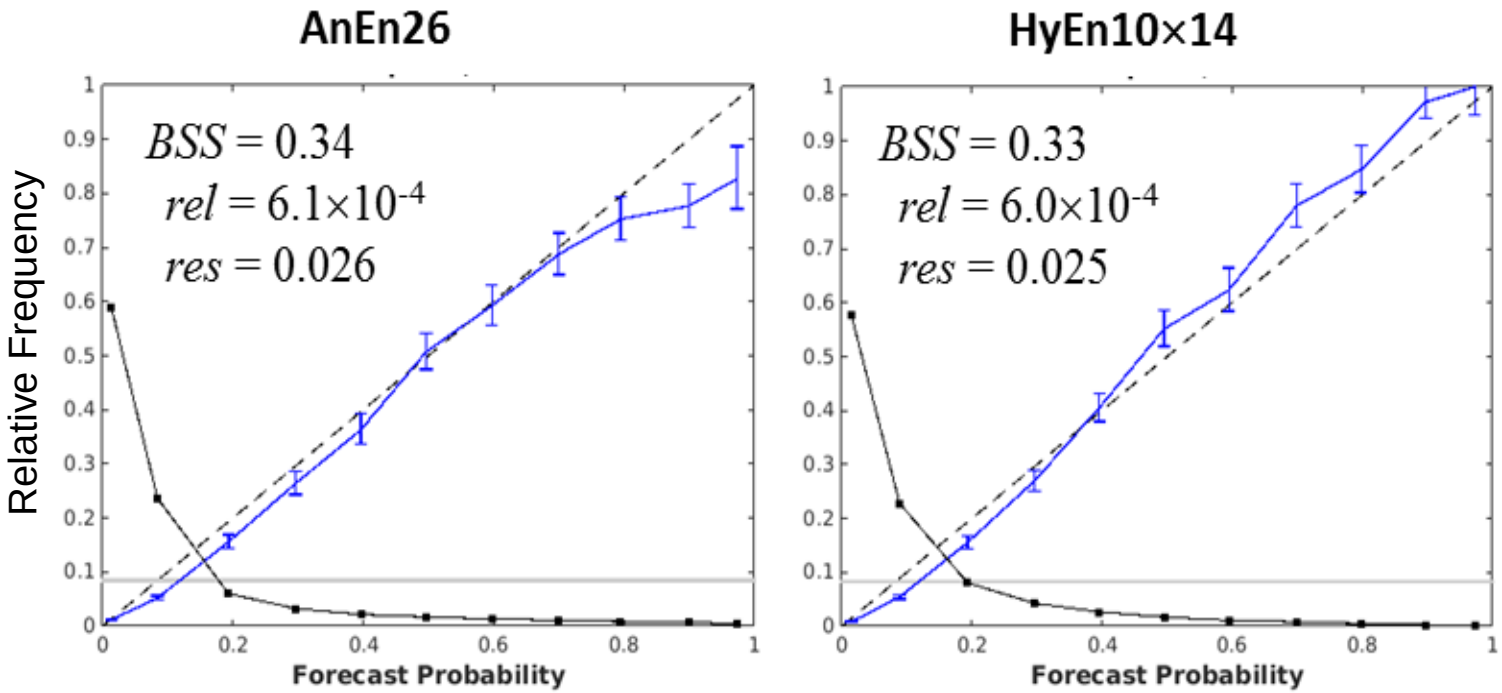
Pr(2-m Temp < 2nd Percentile)



2-m Temperature Results

2-m Temperature ≤ 10th percentile

30-h Lead Time



overdispersive
(too much spread)

rel & *res* on par with AnEn

Conclusions

HyEn – *may be too unreliable?*

- Can out perform **AnEn**, perhaps when the NWP members simulate flow-dependent uncertainty well
- Unpredictably prone to being less sharp and overdispersive

AnEn – *likely best method moving forward, due to its advantages*

- Straightforward approach more easy to tune and improve
- Enables use of higher-resolution NWP
- Better opportunity for longer training dataset

Ongoing Investigation (of **AnEn** & **HyEn**)

- Expand performance analysis with more data, more variables, more seasons, etc.
- Try **HyEn** on a multi-model NWP ensemble
- Improve analog approach
  - Tune analog metric to make best use of more predictors
  - Make # of analogs adaptive to current forecast situation, *rather than based on rarity of the probabilistic event threshold*

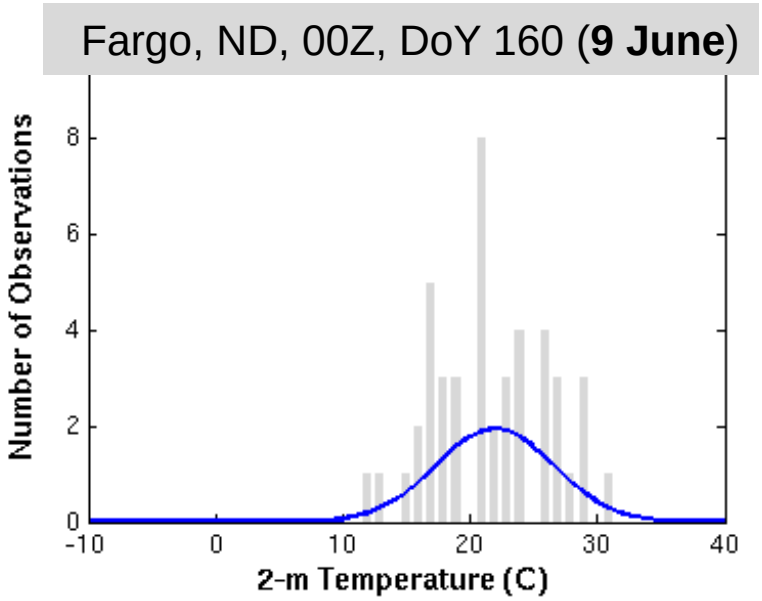

Appendix!

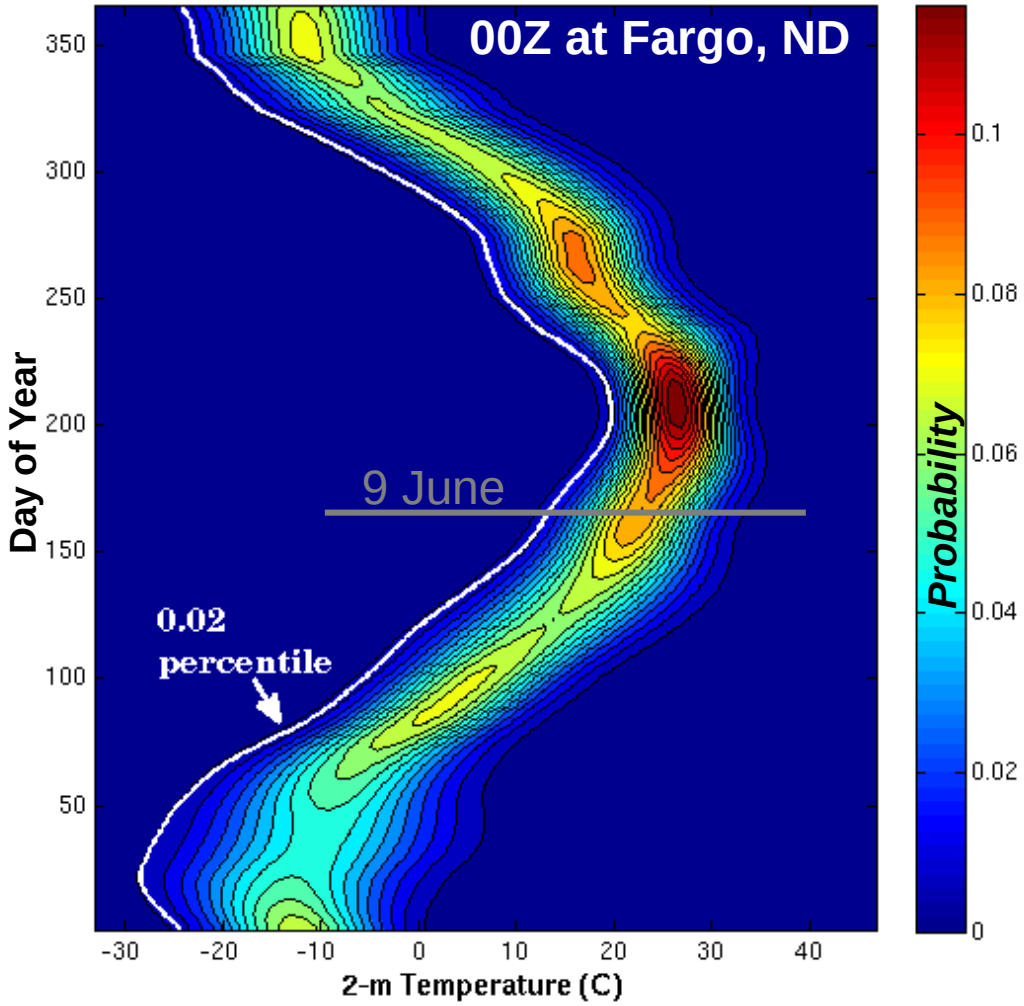
Finding Climate Percentiles

Goal: Climate PDFs, for the research dataset, stratified by:

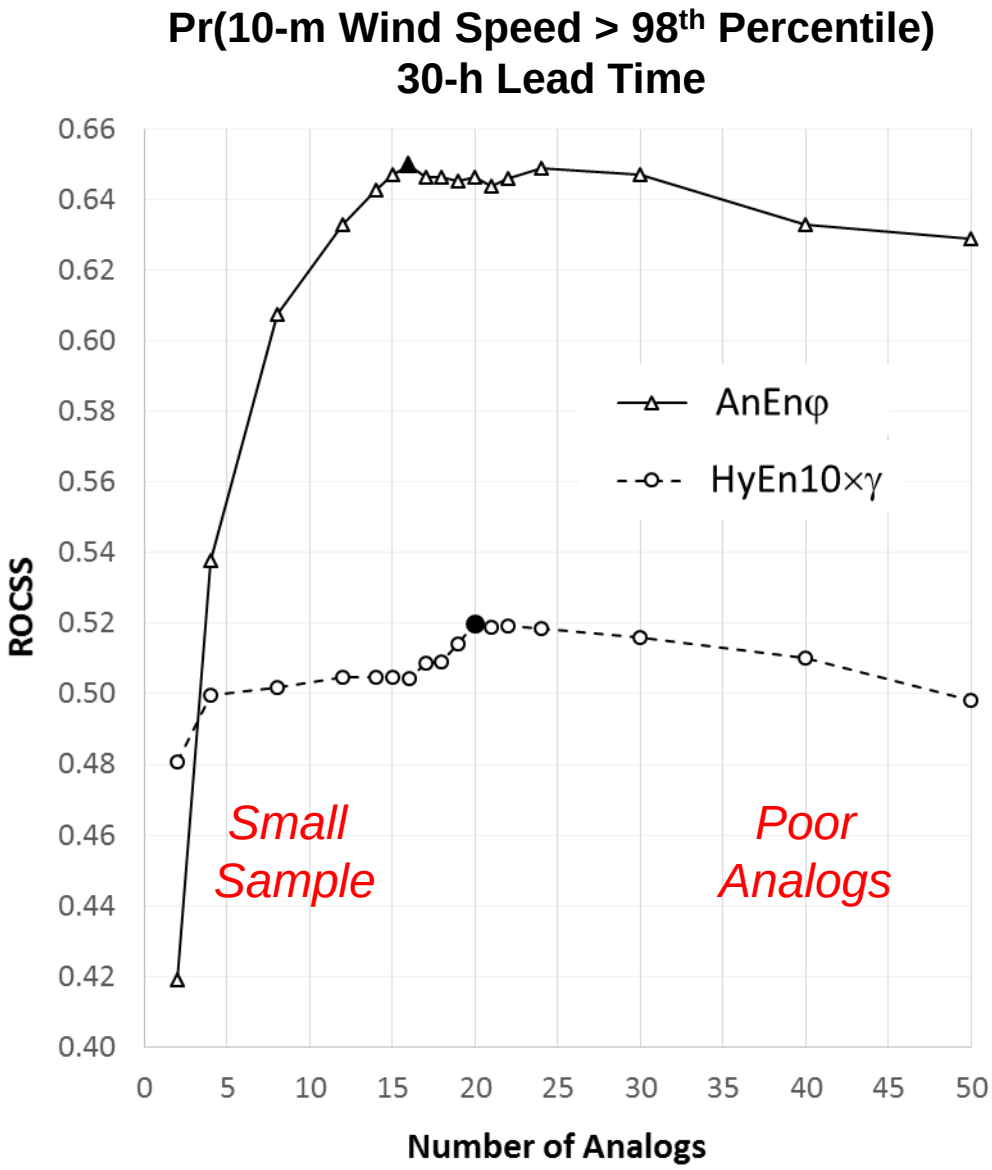
- Location
- Day of year (DoY)
- Time of day

Method: Using the 15 months of obs., fit all obs. within 15 days of the DoY to an assumed PDF





Empirical Optimization of # of Members



Optimal # of Members

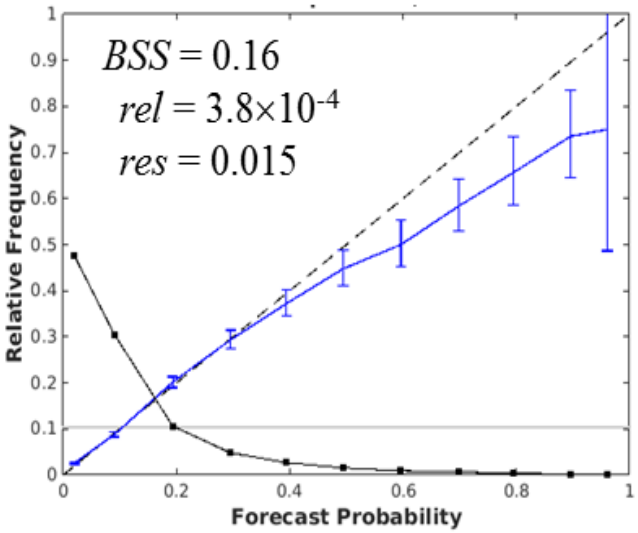
	50th	90th	98th
AnEnφ	**30**	**28**	**16**
HyEn10×γ	**12**	**16**	**20**

10-m Wind Speed Results

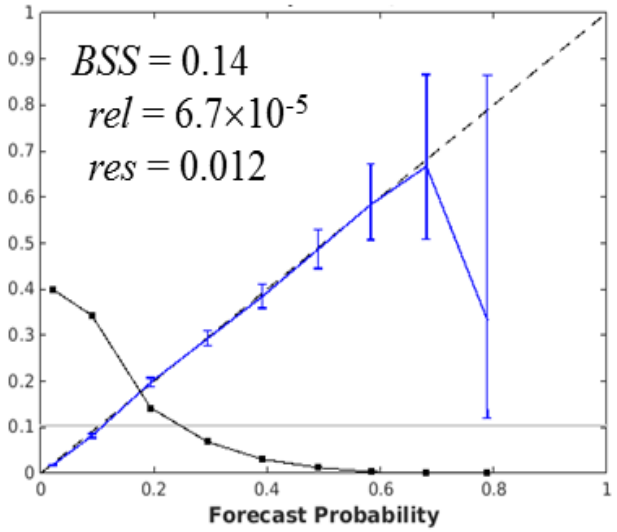
10-m Wind Speed ≥ 90th percentile

30-h Lead Time

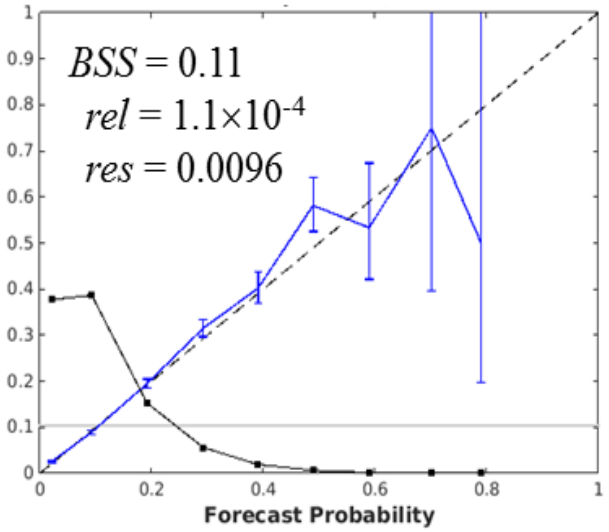
EMOS10



AnEn28



HyEn10×16



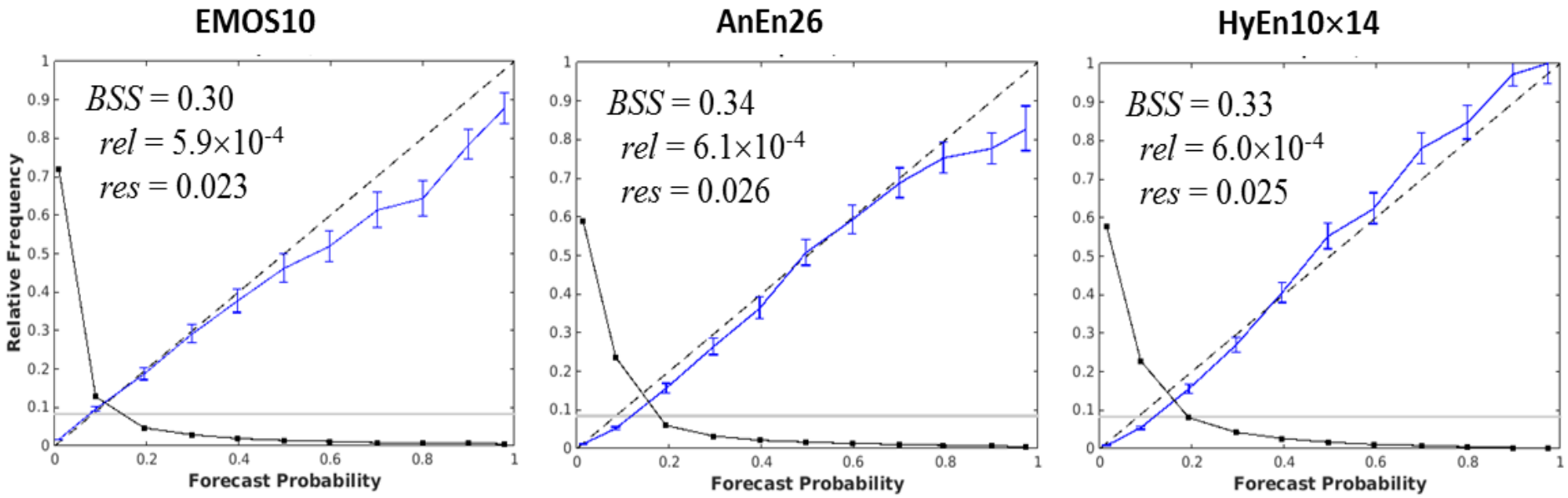
Somewhat overdispersive (too much spread)

rel & *res* worse than AnEn

2-m Temperature Results

2-m Temperature ≤ 10th percentile

30-h Lead Time



Somewhat underdispersive (not enough spread)

rel & *res* on par with AnEn