Probabilistic Forecasting of Extreme Waves:
Methodology, Validation, and Post-Processing with Machine Learning

Ricardo M Campos[1], Darin Figurskey[2], Avichal Mehra[2]

1. Lynker at Ocean Prediction Center, NCEP/NWS/NOAA, ricardo.campos@noaa.gov
2. Ocean Prediction Center, NCEP/NWS/NOAA

Introduction

- A new oceanic hazards outlook (week 2) has been developed based on the NOAA's Global Ensemble Forecast System (GEFSv12).

https://ocean.weather.gov/week2/

It contains delineations of where winds and waves are expected to have the potential of posing a hazard to either life or property for vessels at sea.

- Over the last three years, we have evaluated the results through daily qualitative assessments as well as quantitative statistical validation.

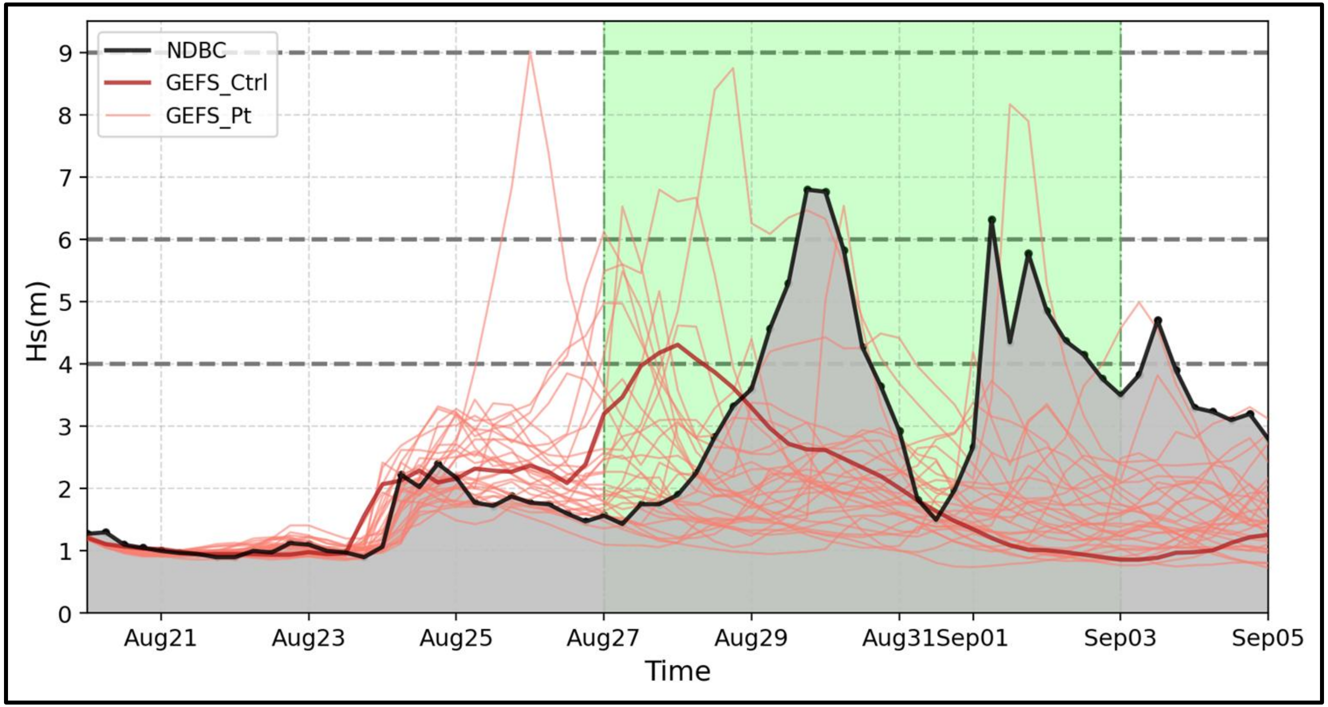
"Probabilistic wave forecast for week two and beyond based on NOAA's Global Ensemble Forecast System" https://doi.org/10.1175/WAF-D-24-0154.1

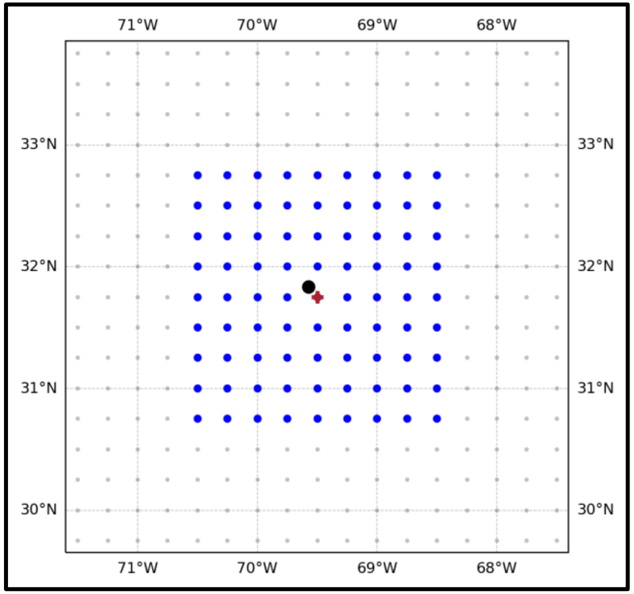
Main goal: develop and implement soft-computing post-processing methods to improve the skill in extreme events.

Requirements: low computational cost, few dependencies, good portability, ease of re-training and implementation, and simplicity of editing.

Oceanic Hazards Outlook - Probability Maps

- Spatio-temporal analysis: Time (**week 2**) and Window (**2°X2°**);
- N-max selection (=2): utilize not only the single max value of the time-series for each member but the n-max ones;
- Distribution threshold (= 87th percentile): focus on members and locations with the most severe conditions.

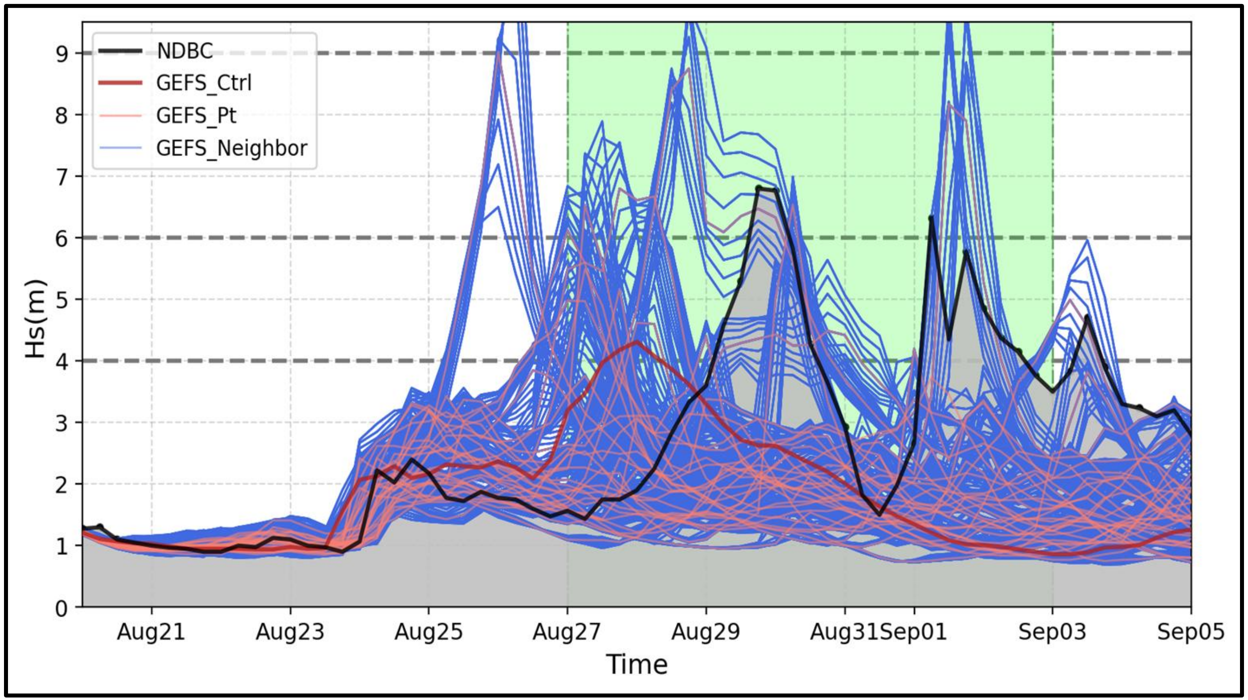


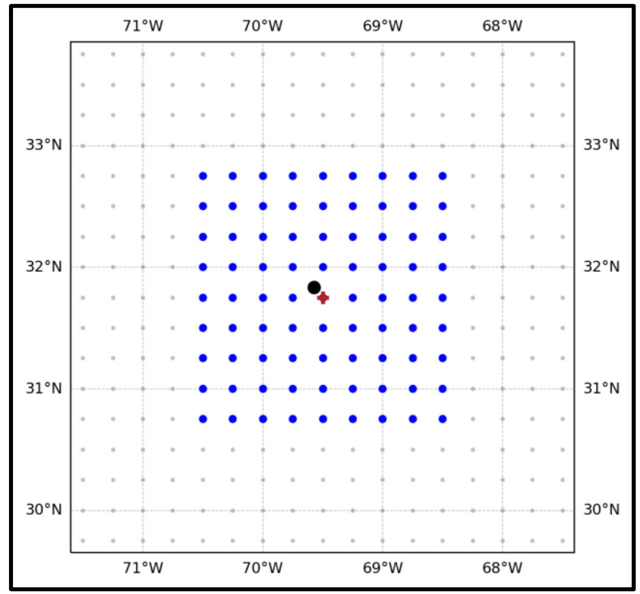


Example 20-Aug-2023 (Atlantic)

Probability Map Algorithm

- Maximum values (6h-resolution) are selected for each 7-day segment;
- 30+1 members, 2°X2° neighbors (81 points for a 0.25° resolution grid);
- The resulting array has dimensions of 2 x 31 x 81;
- A threshold is applied to select the tail of the ECDF (> 87th percentile);
- Probabilities are computed for each level of Hs and U10 at the grid points.



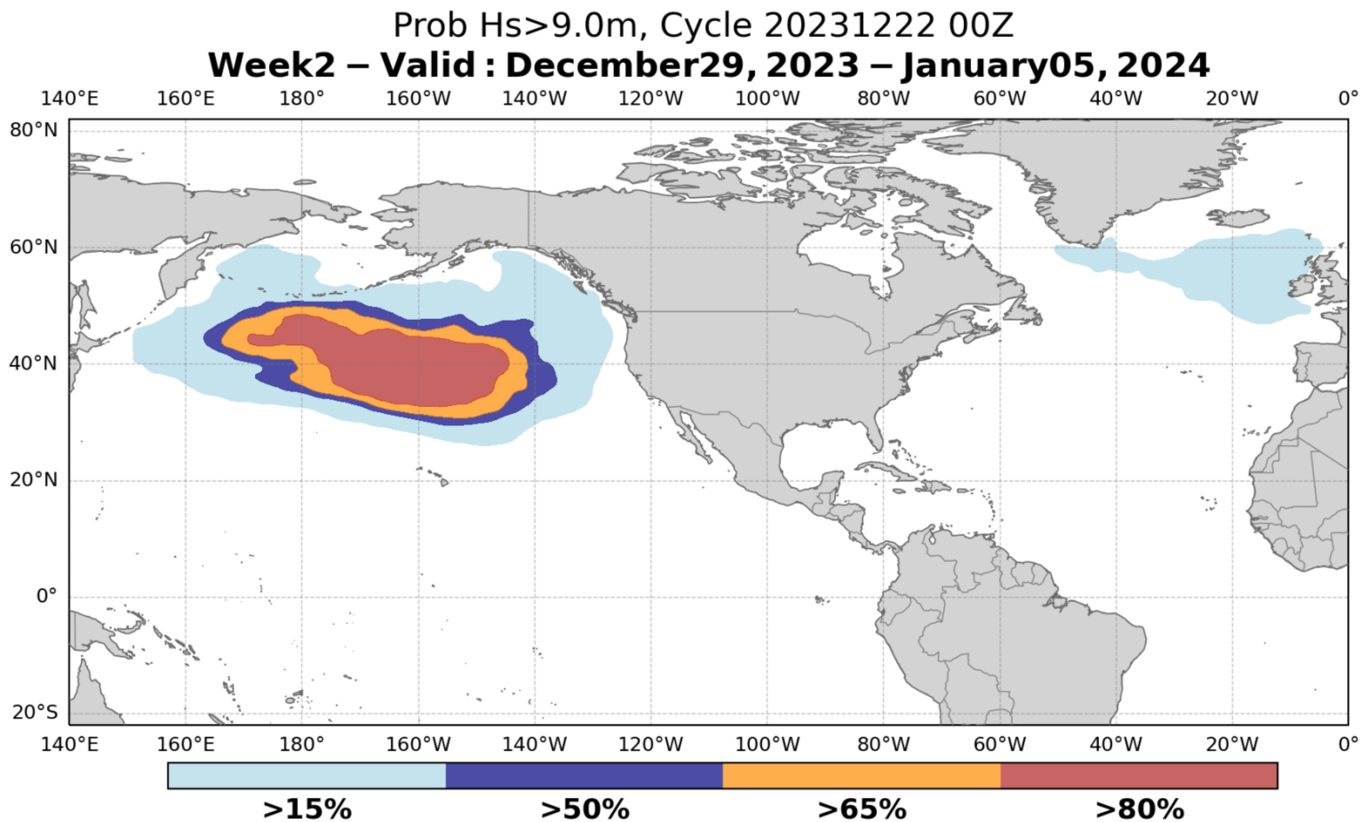


Probability Map Algorithm

Levels for the probability plots:

- 10-m wind speed (knots): **34, 48, 64**;
- Significant wave height (meters): **4, 6, 9, 14**;

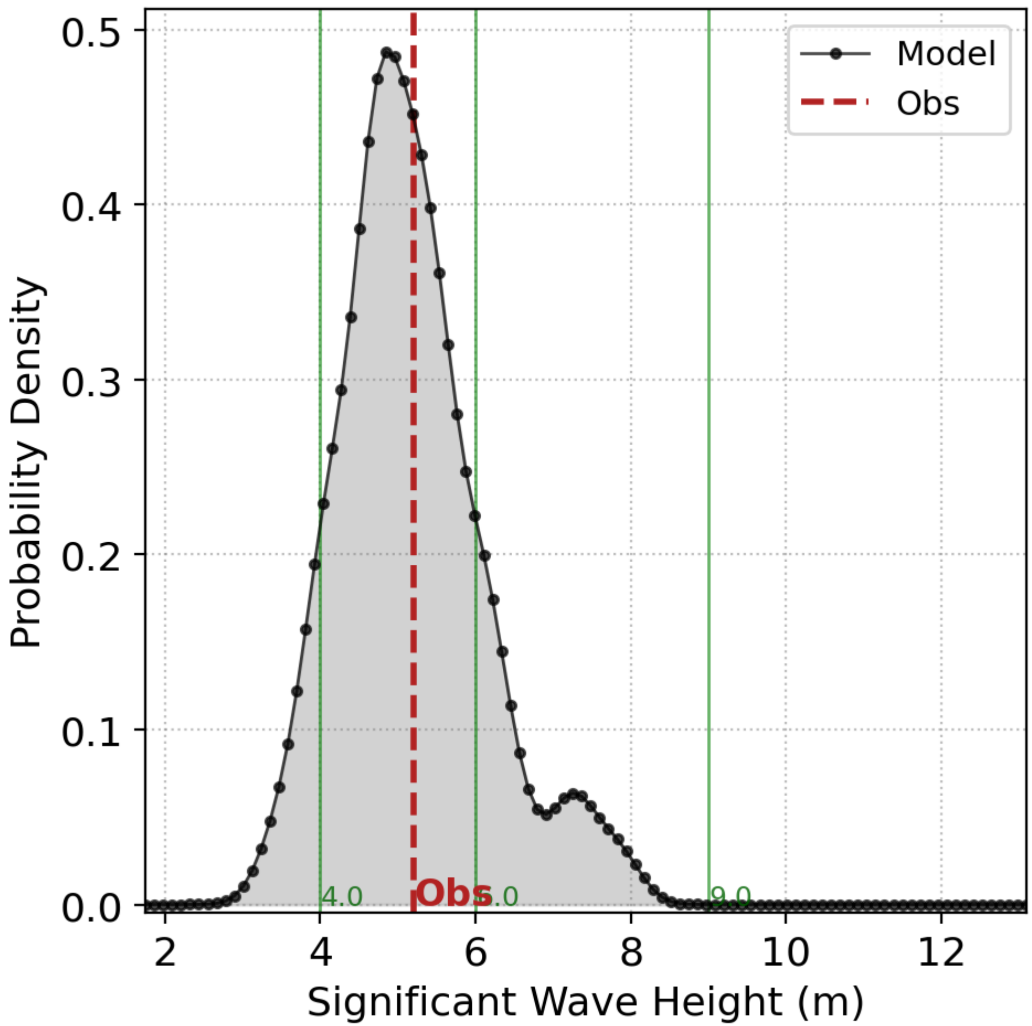
Probabilities: **15%, 50%, 65%, and 80%**.



Large waves along Ventura, CA shoreline.

Examples of PDF:

Point NDBC 46006, 40.764°N / 137.377°W



Recent and ongoing developments

Summary of research lines composing the hybrid model:

- Uncertainty analysis of observations using triple collocation;
- Spatial assessments of the week-2 probabilistic forecast;
- Optimization of statistical parameters as a function of space and time;
- Ensemble lag;
- Multi model ensemble;
- Bias correction and machine learning post-processing models.

Triple Collocation analysis of wind and wave observations

Moored, heave-pitch-roll wave buoys:

- National Data Buoy Center (**NDBC**, NOAA), 3-meter discus buoys;
- Coastal Data Information Program (**CDIP,** SIO), Datawell Mark III Directional Waveriders

Free-drifting buoys (small dimensions) with GPS-derived data:

- **microSWIFT** (APL-UW): 9.2 cm diameter, 21.4 cm length, and 0.9 kg weight
- **DWSD** (LDL-Scripps): spherical hull with a 39 cm diameter and 12 kg weight
- **Spotter** (Sofar Ocean): spherical hull with a 42 cm diameter and 7.5 kg weight

Satellite Altimeters (Australian Ocean Data Network):

- Post-processed and calibrated by **AODN**: CFOSAT, SARAL, CRYOSAT2, HY2B, JASON3, SENTINEL3A, SENTINEL3B, and SENTINEL6A

Saildrones (PMEL and AOML, NOAA).

Model Global Data Assimilation System (**GDAS**), GFS NOAA, 0.16° and 1h res.

Triple Collocation analysis of wind and wave observations

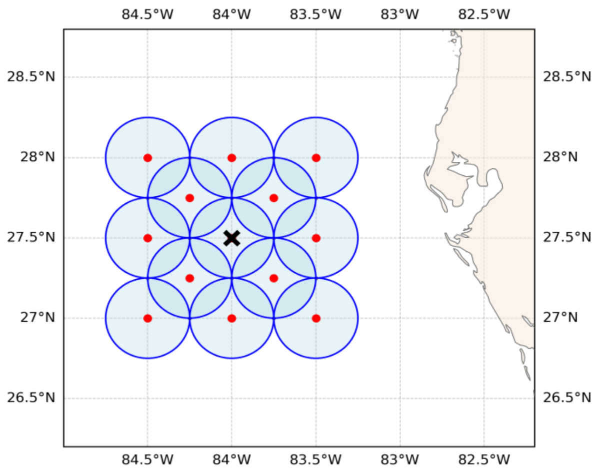
Janssen, P. A., S. Abdalla, H. Hersbach, and J. R. Bidlot, 2007. https://doi.org/10.1175/JTECH2069.1
Houghton, I.A et al. 2021. https://doi.org/10.1175/JTECH-D-20-0187.1

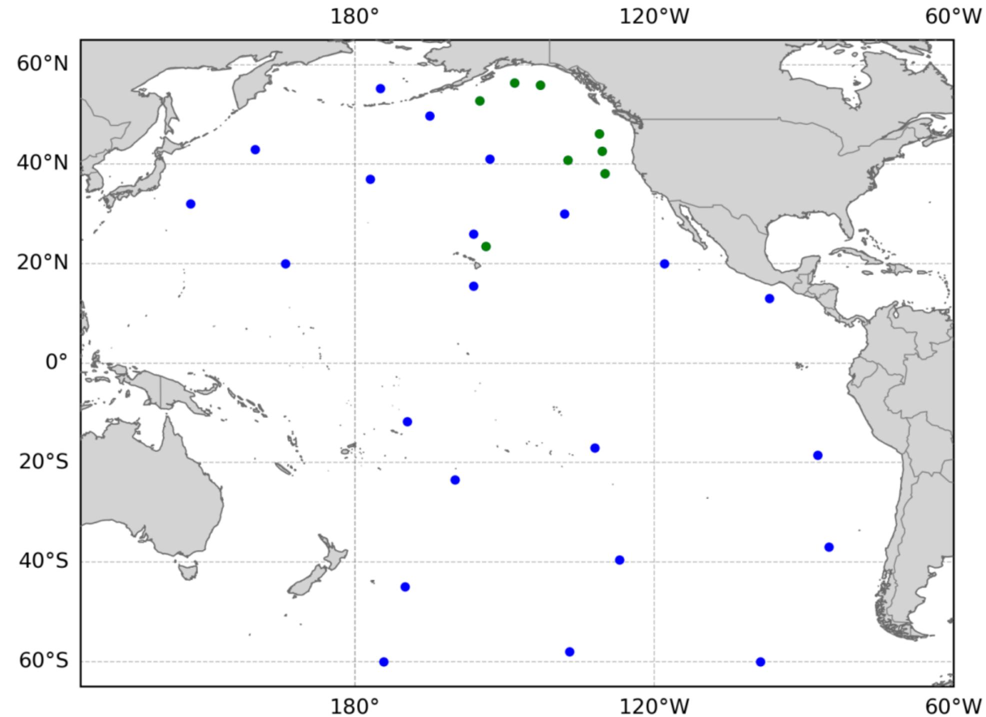
- Analysis (GDAS NCEP/NOAA); Altimeter (AODN); buoys and saildrones
- 3 years of data: 2022, 2023, 2024

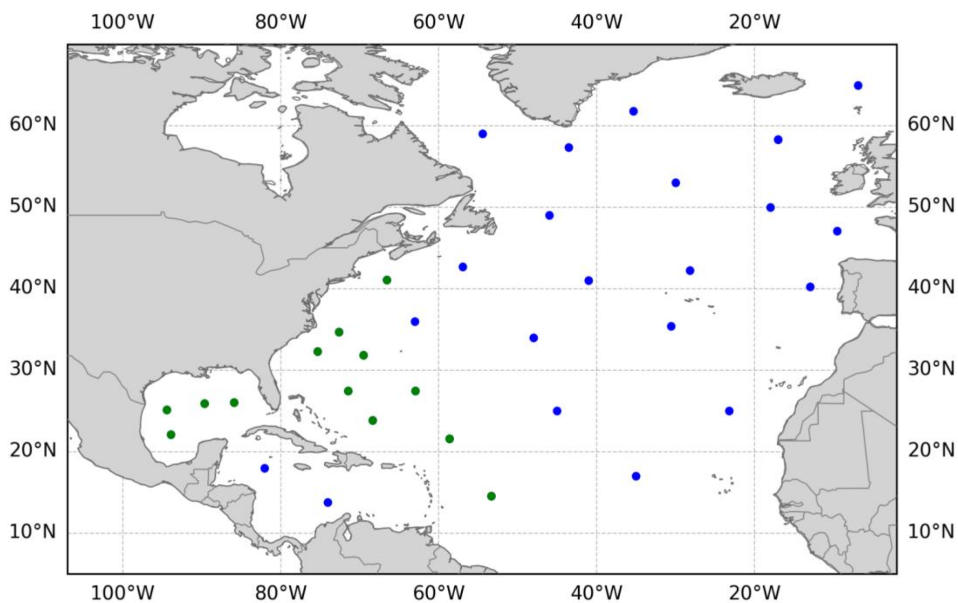
Observation	Residual Error (m)
CDIP	**0.131**
NDBC	**0.174**
SPOTTER	**0.140**
DWSD	**0.298**
MICROSWIFT	**0.303**
SAILDRONE	**0.135**
ALTIMETER	**0.12 to 0.21**

Observation (Hs>4)	Residual Error (m)
CDIP	**0.331**
NDBC	**0.378**
ALTIMETER	**0.17 to 0.23**

Validation and spatial distribution of the forecast performance



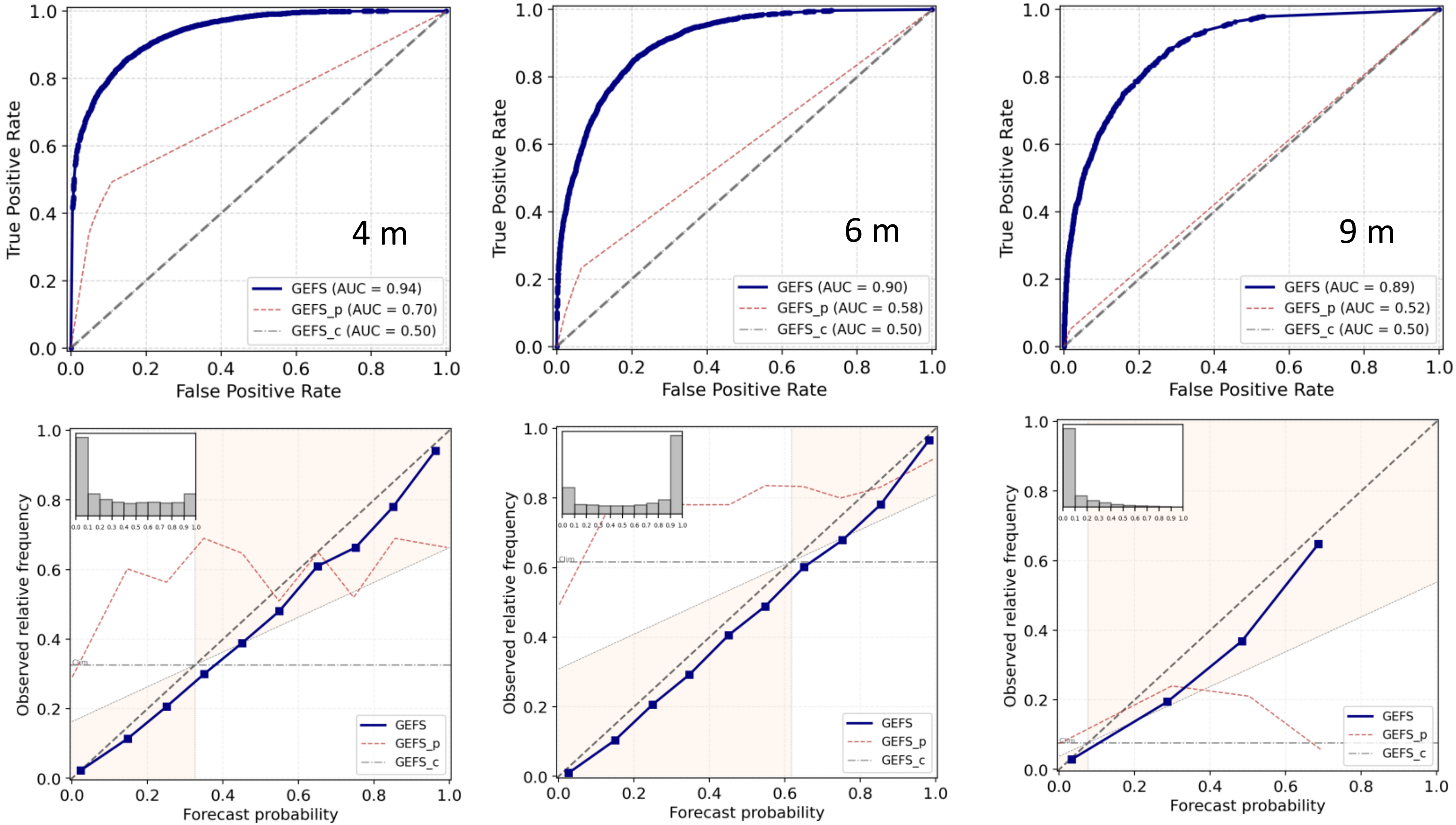




Statistical analysis and validation were conducted for each point and cluster. The quality of the forecast is both site-dependent and season-dependent.

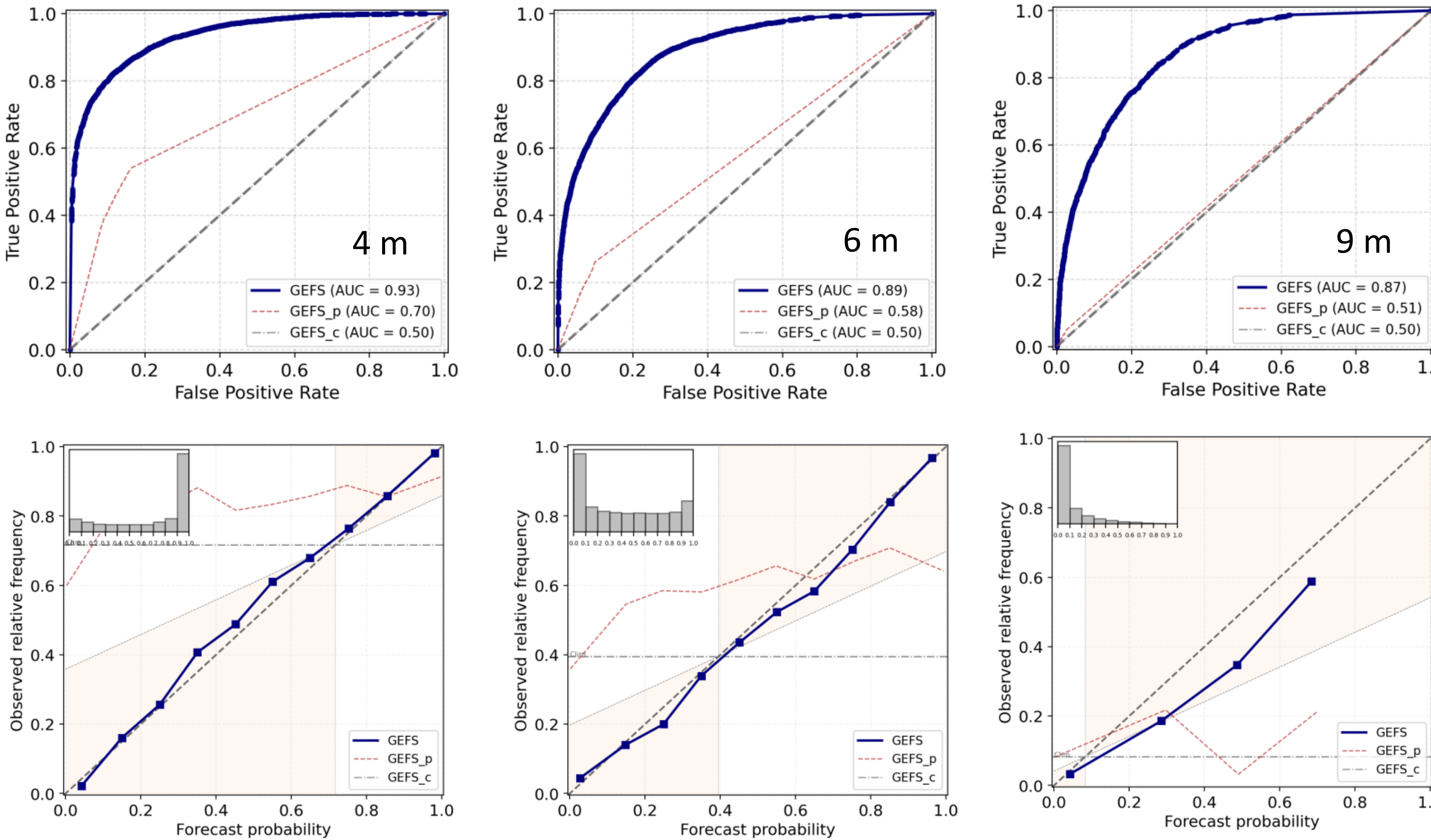
Validation and spatial distribution of the forecast performance

North Pacific – Center North (Hs = 4, 6, 9 m)



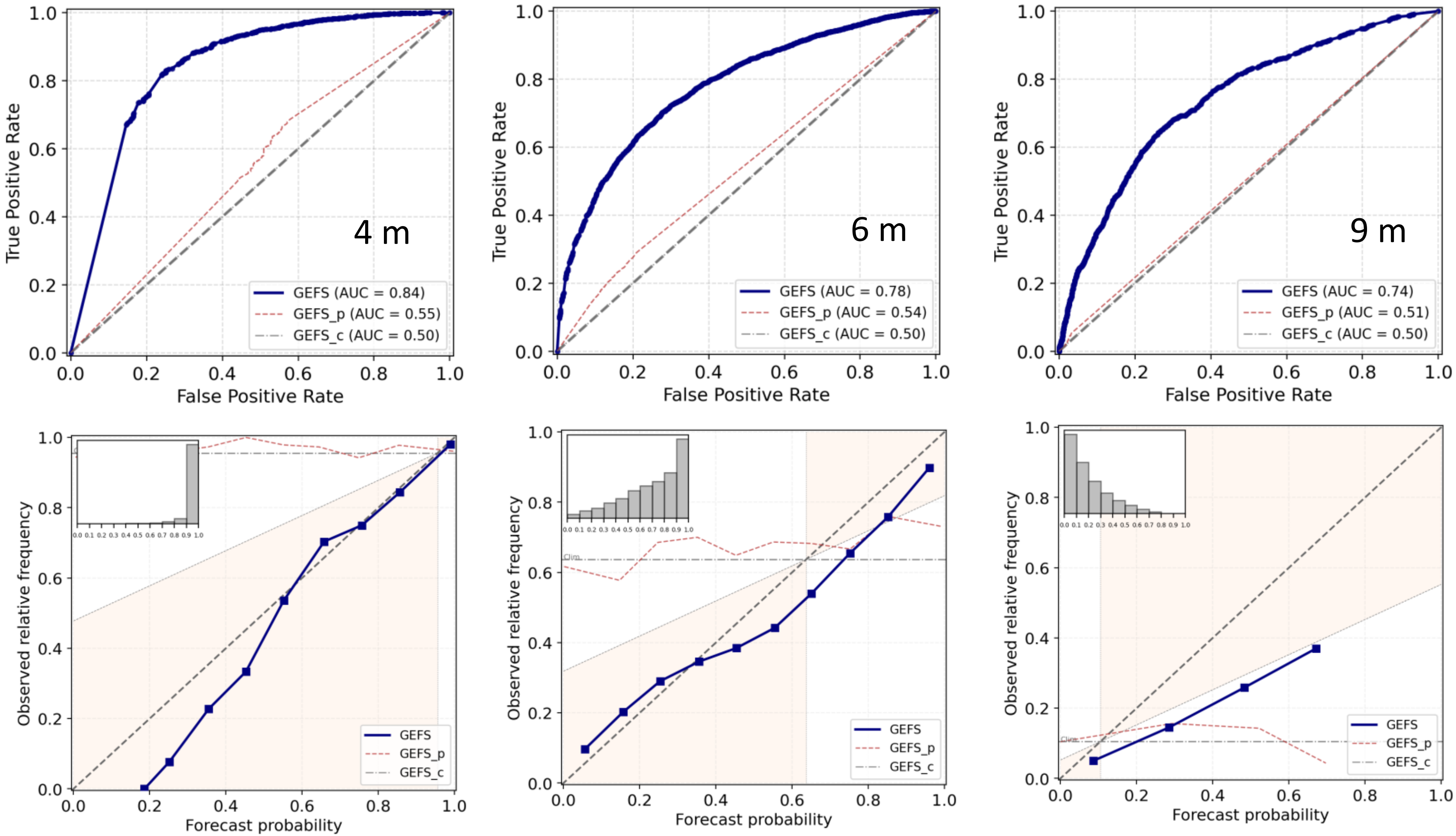
Validation and spatial distribution of the forecast performance

North Atlantic – Center North (Hs = 4, 6, 9 m)



Validation and spatial distribution of the forecast performance

South Pacific – South, around 60S



Ensemble lag

Combination of ensemble cycles – GEFS 00Z and 12Z

A simple example of ensemble expansion:

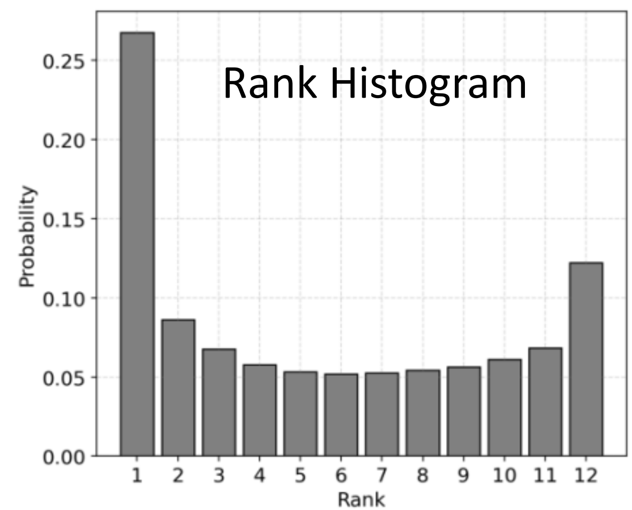
1. 00Z from cycle 01/26: **31 members**
2. 00Z from cycle 01/26 + 12Z from cycle 01/25 : **62 members**
3. 00Z from cycle 01/26 + 12Z from cycle 01/25 + 00Z from cycle 01/25 : **93 members**
4. ...

Increase spread

Potentially improve the ensemble forecast for longer forecast lead times

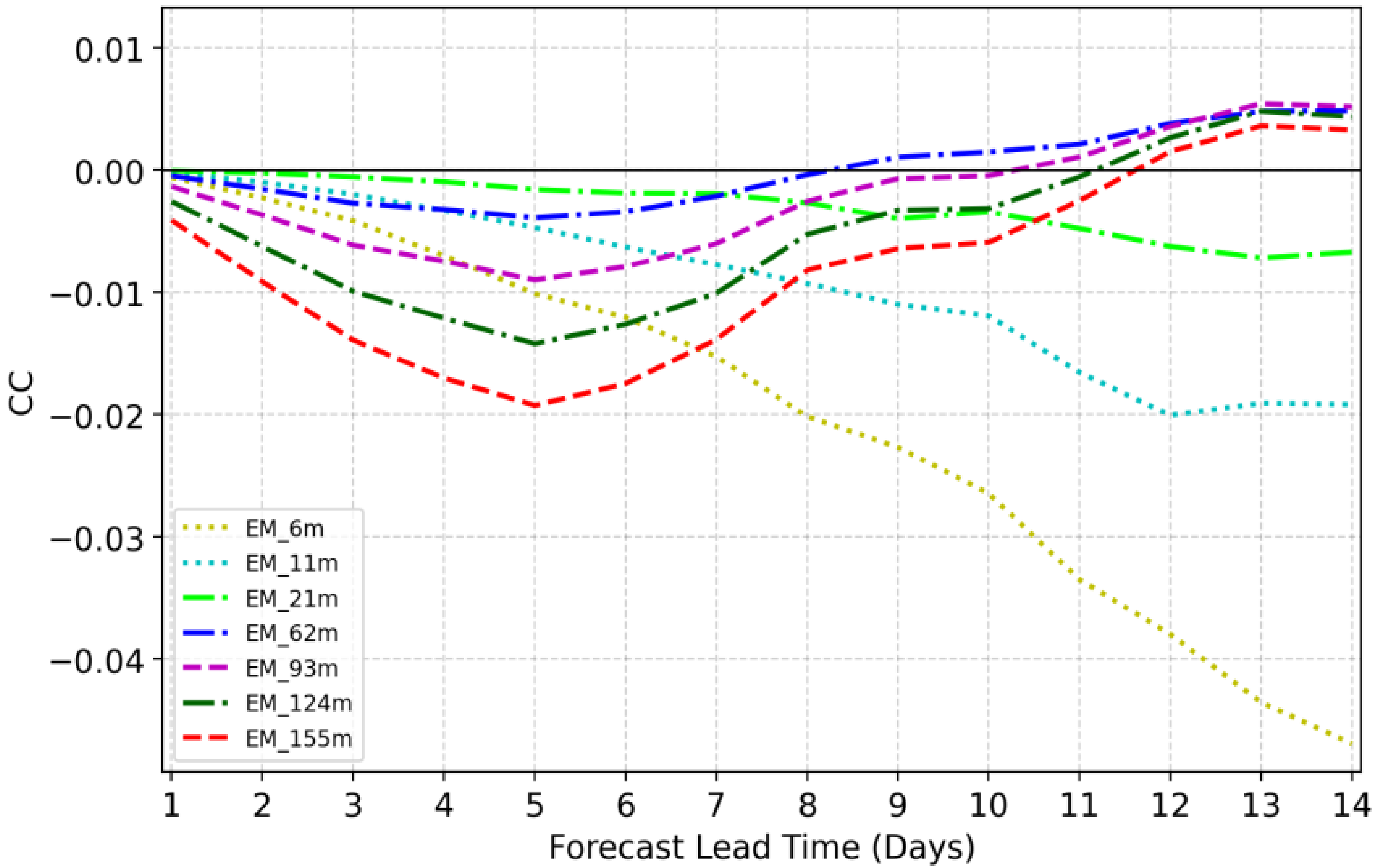
Usual problem of wave forecast from GEFS, under-spread

U-shape indicates over-confidence of the model



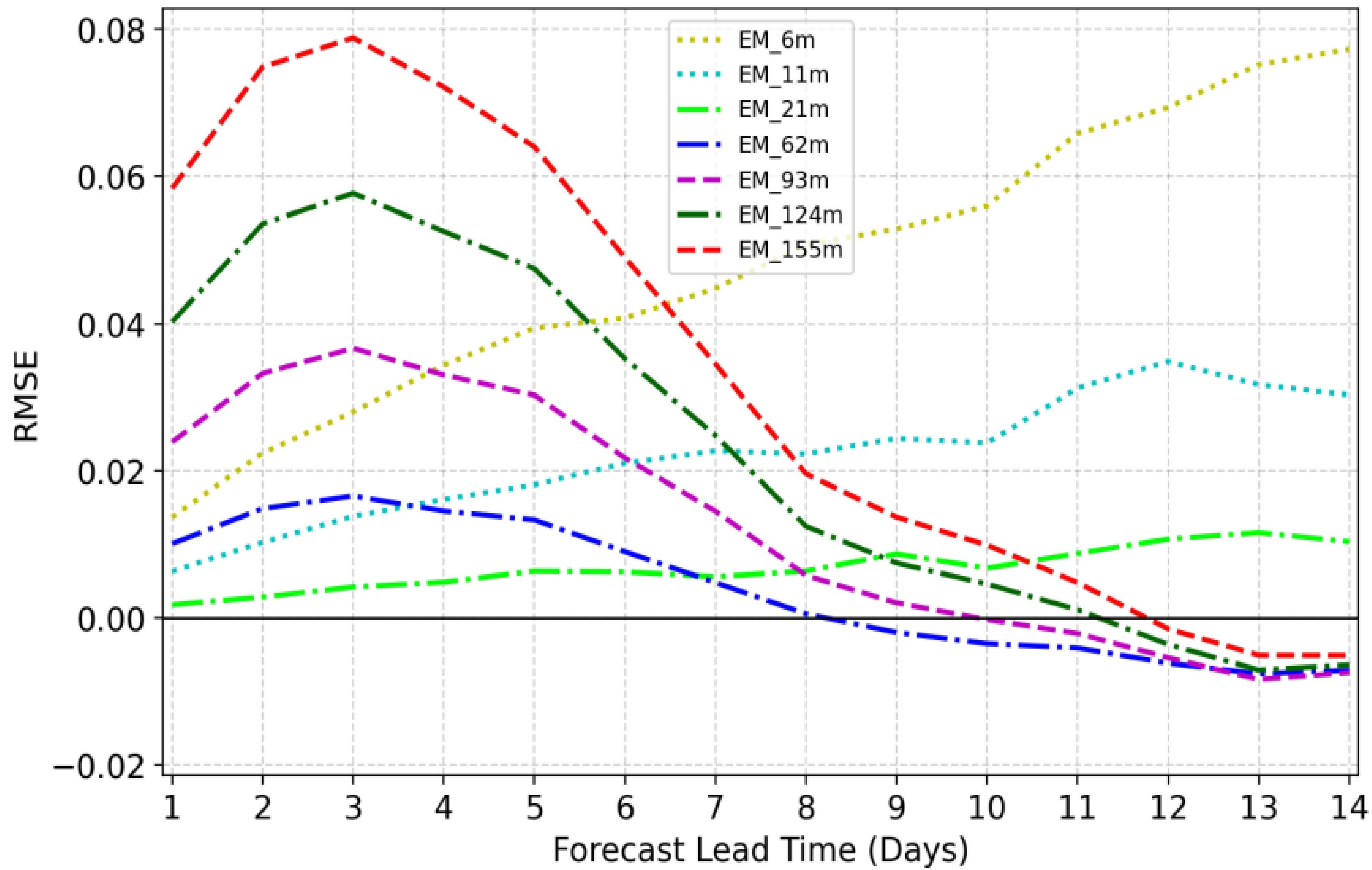
Ensemble lag

Correlation Coefficient (CC) and RMSE: compared against the default (31 members)



Ensemble lag

Correlation Coefficient (CC) and RMSE: compared against the default (31 members)



Multi-model ensemble

Combination of 3 wave ensembles: total of 103 members:

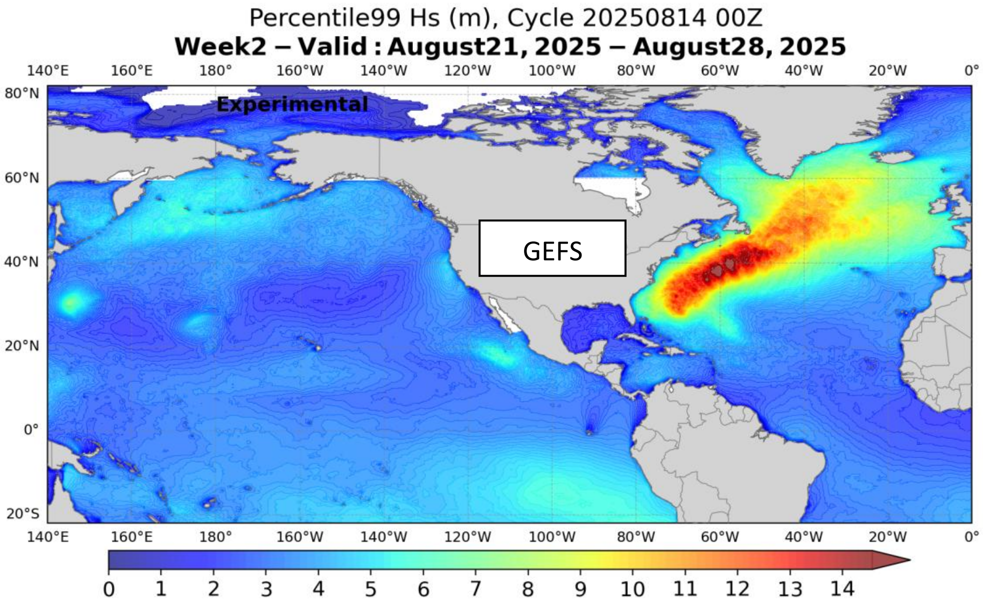
- GEFSv12 NCEP/NOAA: 31 members
- CMCE (Env Canada): 21 members
- ENS15 ECMWF: 51 members

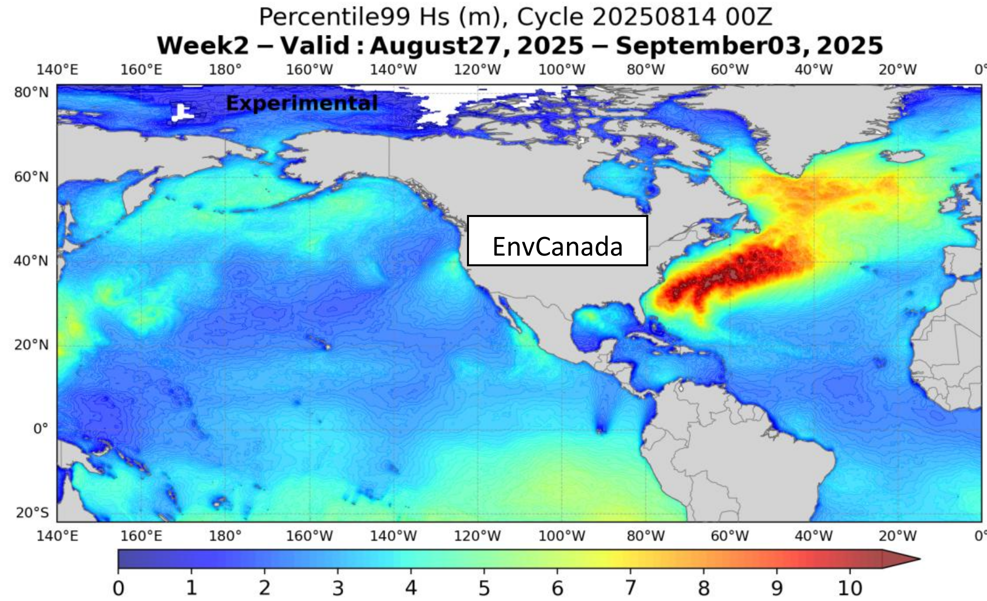
Benefits: Different physics and numerical models – increase the probability of detecting events.

Challenges: To combine different PDFs and multimodal distribution. Different resolutions.

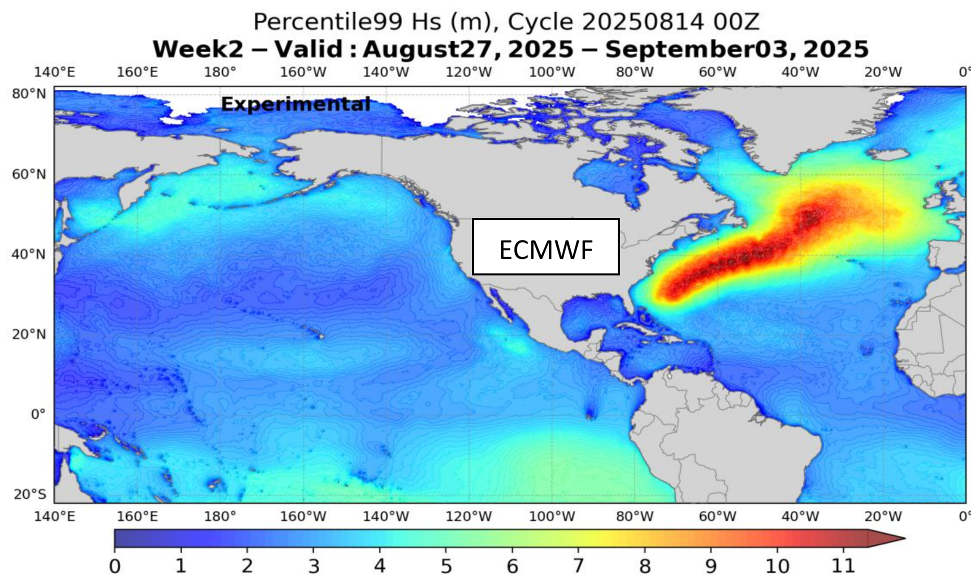
Multi-model ensemble

99th Percentile, Hs (m)



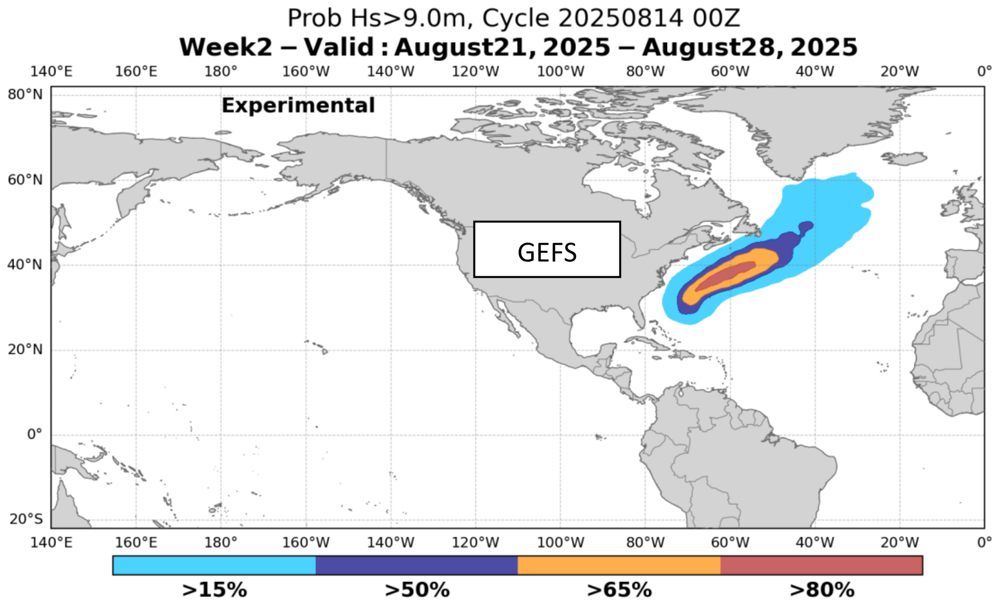


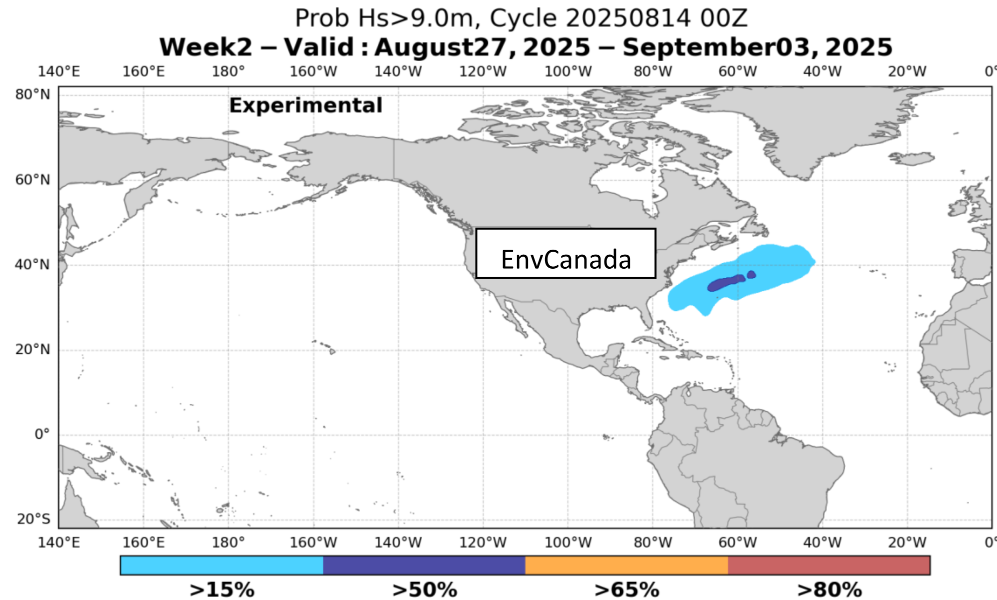
Hurricane Erin (Cat 5)



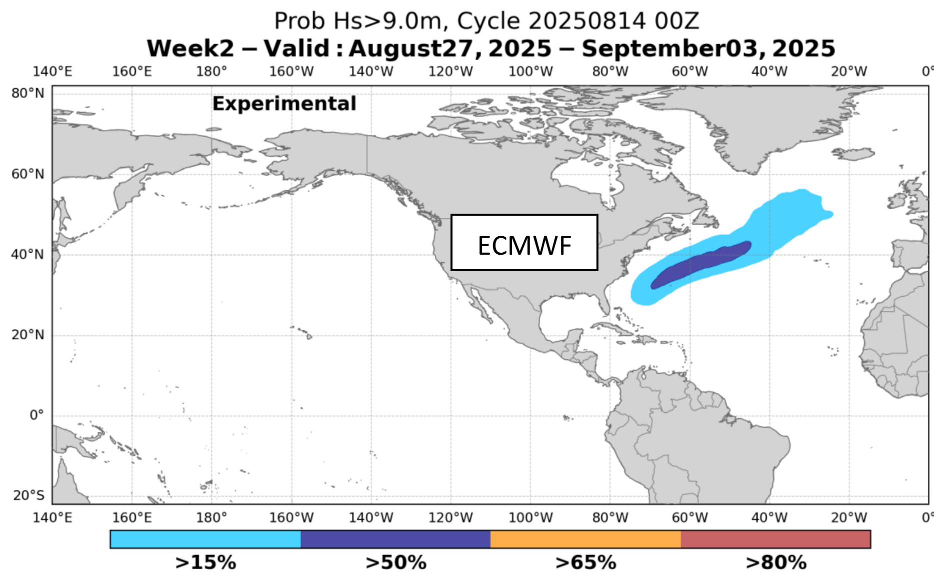
Multi-model ensemble

99th Percentile, Hs (m)





Hurricane Erin (Cat 5)

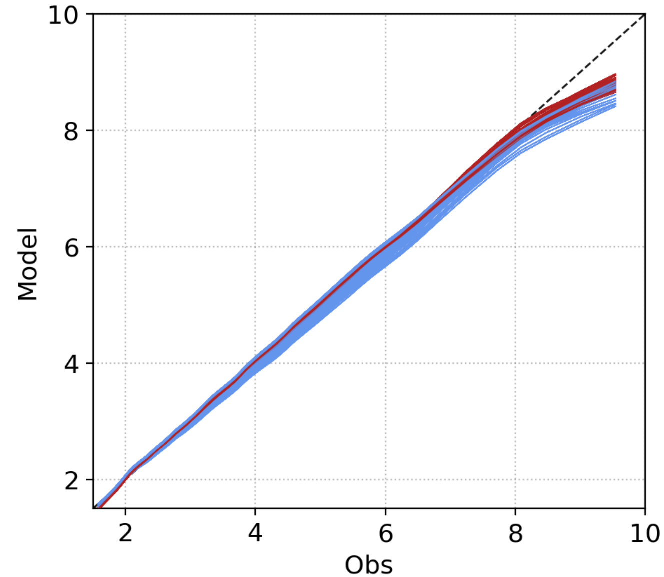


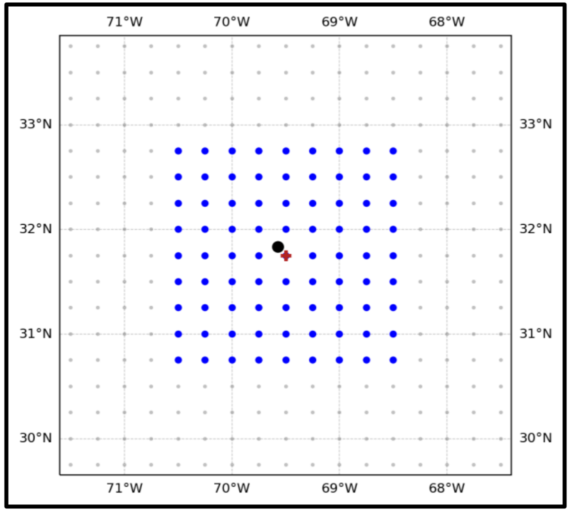
Univariate Linear Bias Correction

Quantile Mapping Method (QMM) applied to GEFSv12

POD		Hs>4.0	Hs>6.0	Hs>9.0
P>15%	GEFSv12	1.00	0.96	0.64
	GEFSv12 BC(QMM)	1.00	0.98	0.71
P>50%	GEFSv12	0.97	0.69	0.08
	GEFSv12 BC(QMM)	0.98	0.77	0.13
P>65%	GEFSv12	0.92	0.48	0.00
	GEFSv12 BC(QMM)	0.95	0.57	0.02
P>80%	GEFSv12	0.83	0.28	0.00
	GEFSv12 BC(QMM)	0.87	0.35	0.00

CSI		Hs>4.0	Hs>6.0	Hs>9.0
P>15%	GEFSv12	0.84	0.51	0.19
	GEFSv12 BC(QMM)	0.84	0.50	0.16
P>50%	GEFSv12	0.86	0.52	0.07
	GEFSv12 BC(QMM)	0.85	0.54	0.11
P>65%	GEFSv12	0.85	0.41	0.00
	GEFSv12 BC(QMM)	0.86	0.47	0.02
P>80%	GEFSv12	0.78	0.27	0.00
	GEFSv12 BC(QMM)	0.82	0.33	0.00





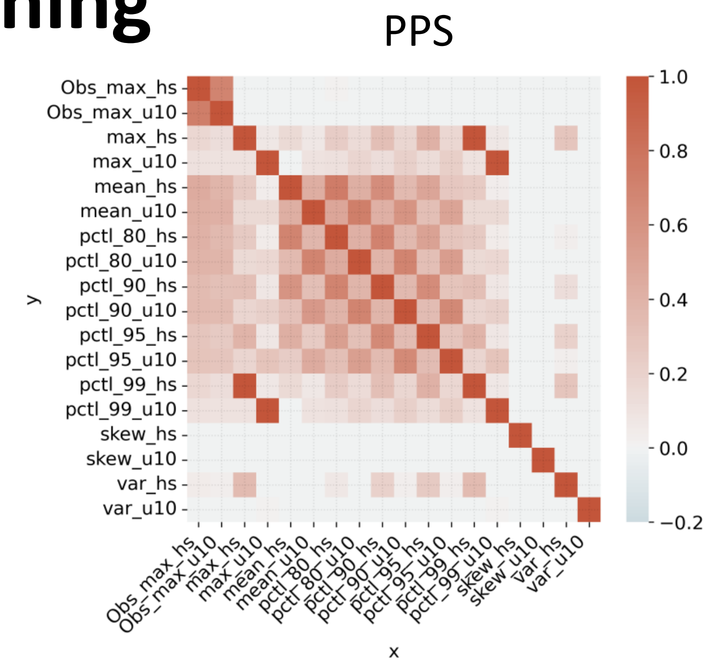
Post-processing using machine learning

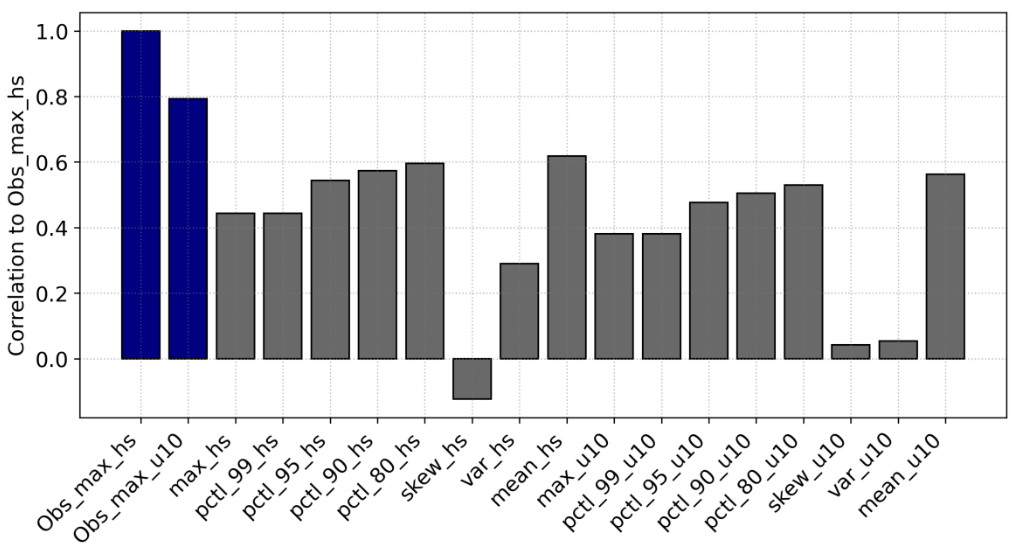
Pre-processing

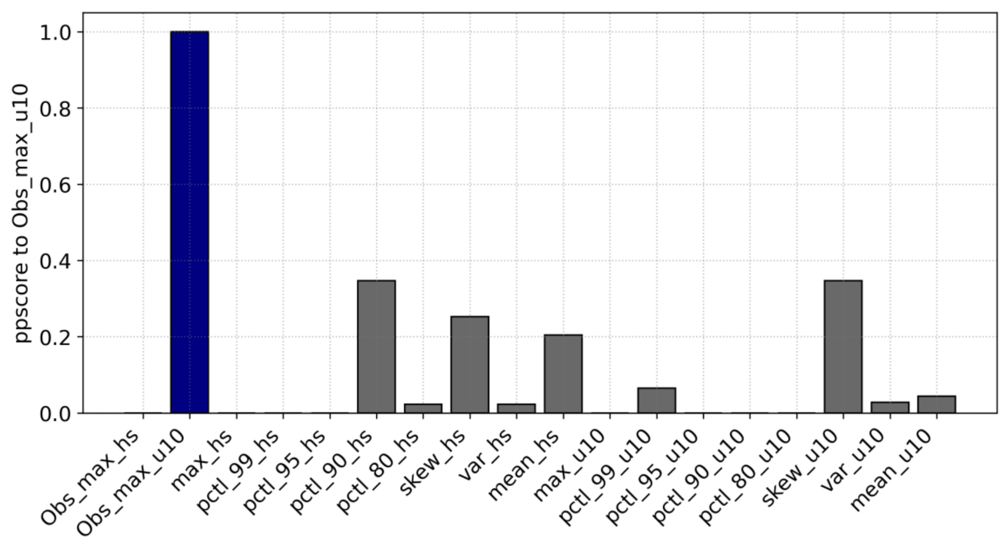
Initial data screening and **feature selection**:

- Pearson correlation coefficient;
- Predictive power score (PPS);
- Recursive feature elimination with cross validation (RFECV);

Variables: *mean, variance, skewness, and percentiles (80, 90, 95, 99) of <u>Hs and U10</u>* (adding Tp)







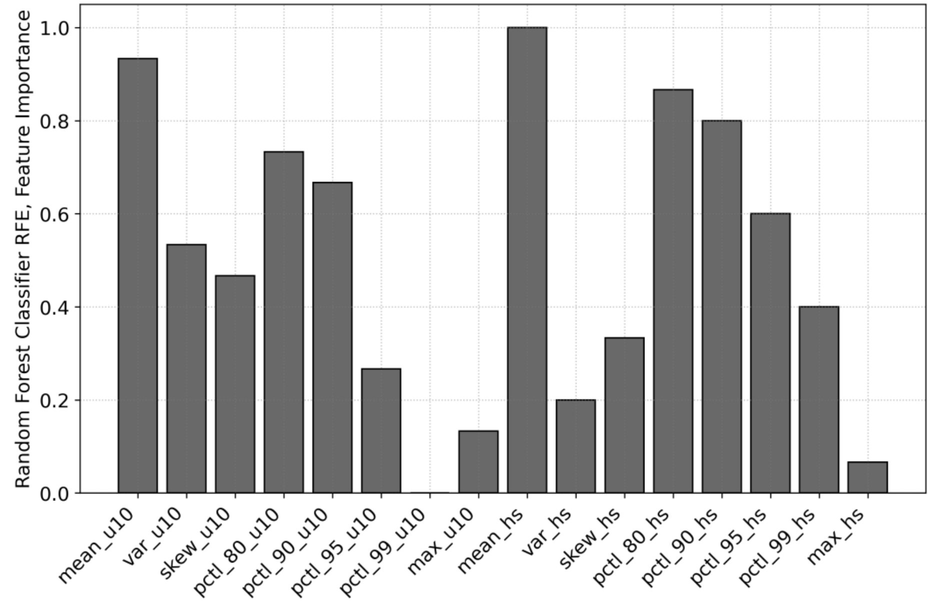
Post-processing using ML

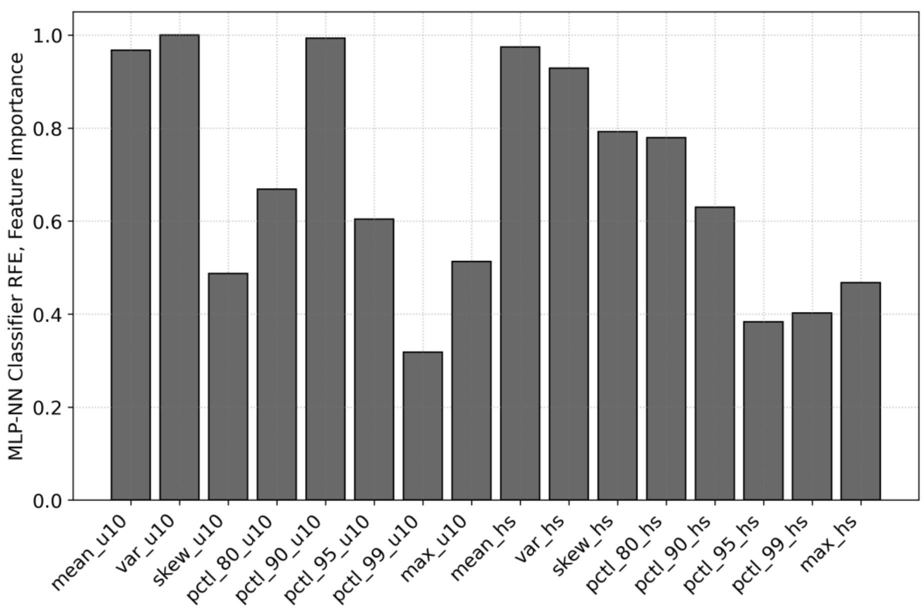
Pre-processing

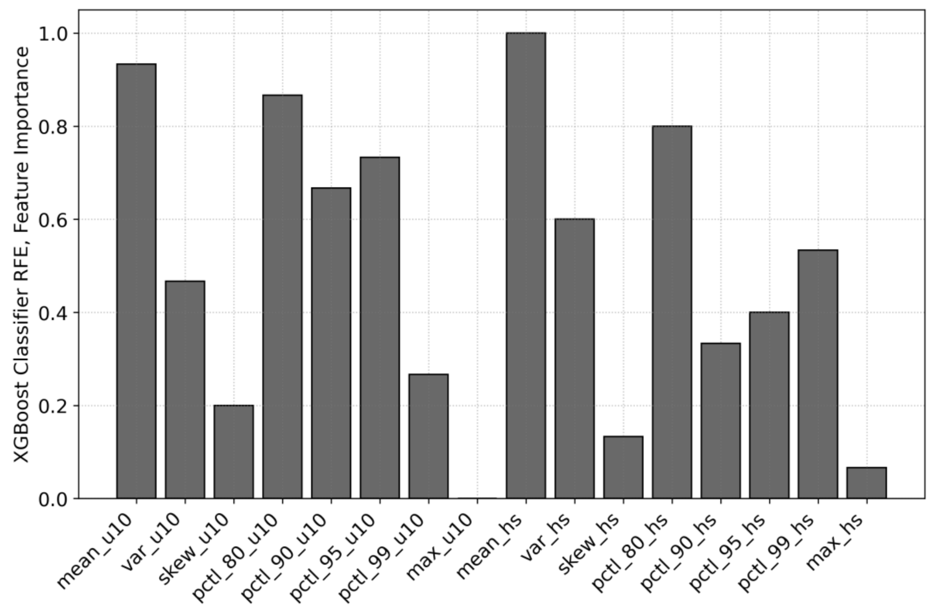
Recursive feature elimination with cross validation (RFECV):

Optimal feature selection for each ML model

Variables: *mean, variance, skewness, and percentiles (80, 90, 95, 99) of Hs and U10*







Post-processing using ML

Models, Architecture, and Optimization

Optimization of hyperparameters using <u>GridSearchCV and KFold</u>:
Scoring metrics: *roc_auc, accuracy, neg_log_loss*

- **Random Forest**: *n_estimators, max_depth, min_samples_split, min_samples_leaf*
- **XGBoost**: *learning_rate, n_estimators, max_depth, min_child_weight, subsample, colsample_bytree, gamma, reg_alpha, reg_lambda*
- **MLP-NN**: hidden_layer_sizes, activation, max_iter, alpha

Optimization Algorithm: Adaptive Moment Estimation (Adam).

This step is the most time-consuming!

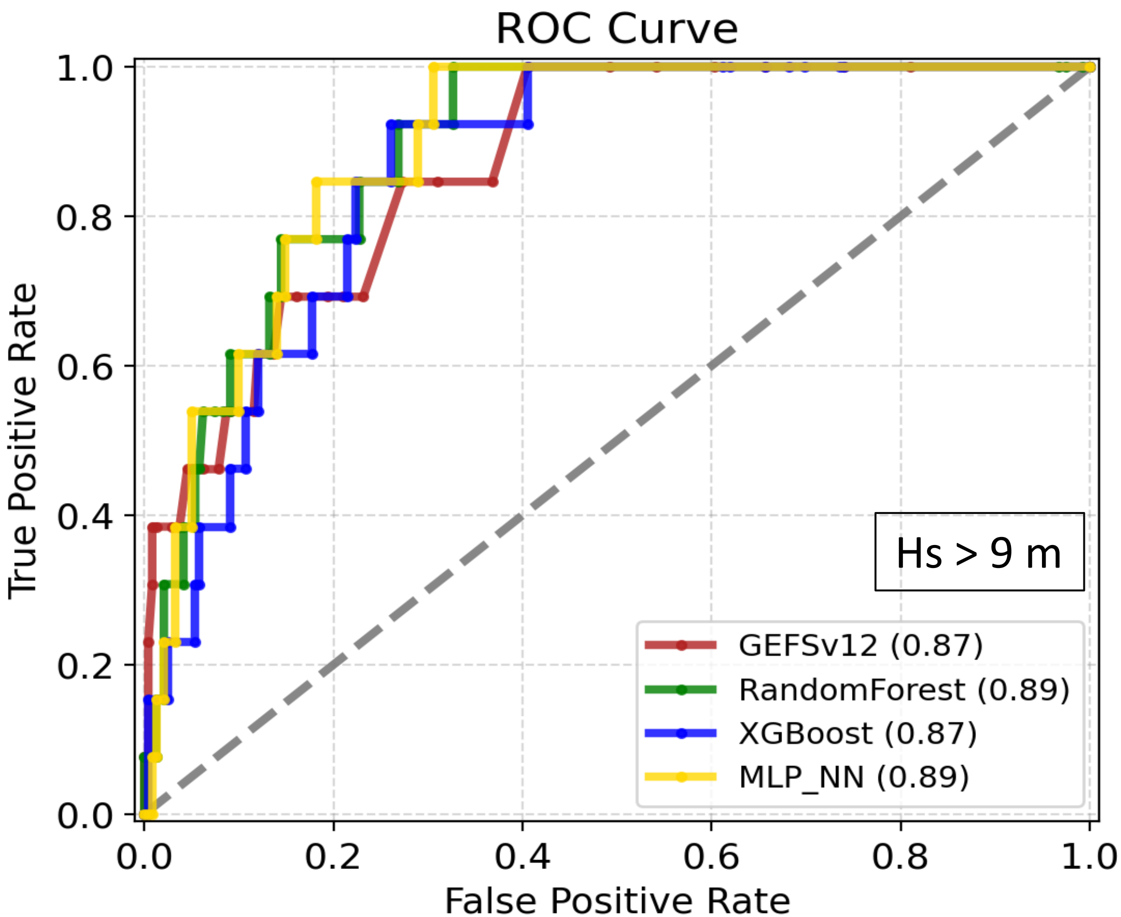
The optimized hyperparameters were tested with various different random initial conditions.

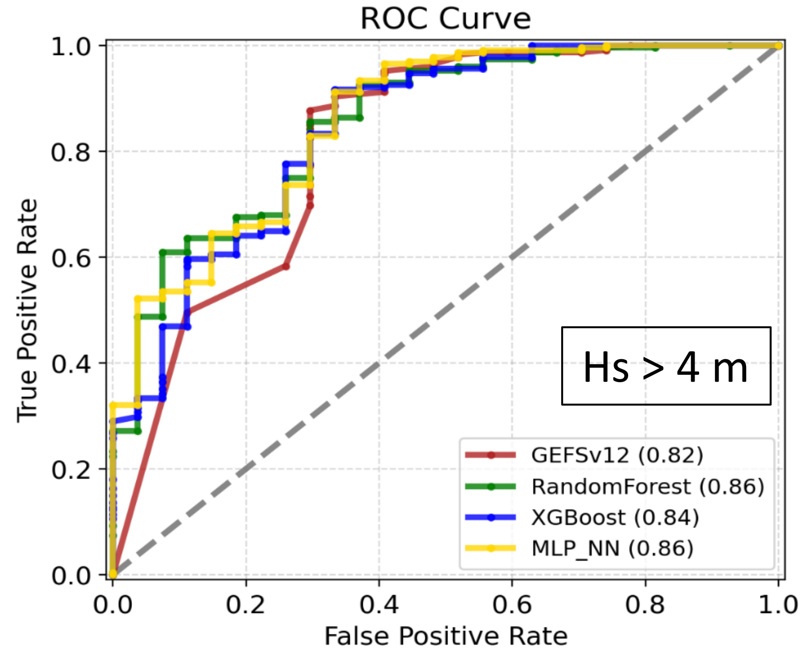
We are recently expanding the dataset, including mode variables, and moving to the AWS cloud (Amazon Q support) with GPU processing.

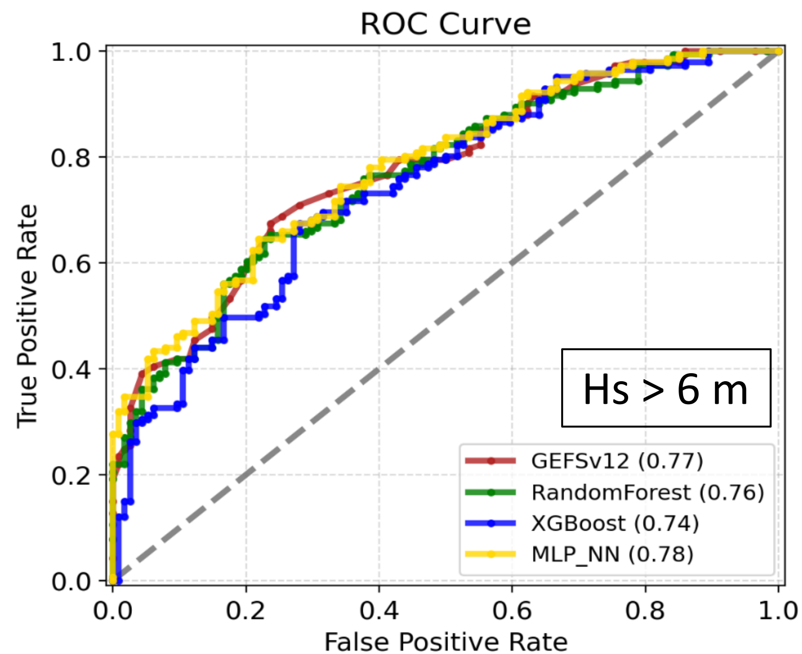
Post-processing using ML

Preliminary results

Validation set: 1-Oct-2023 to 31-Dez-2023



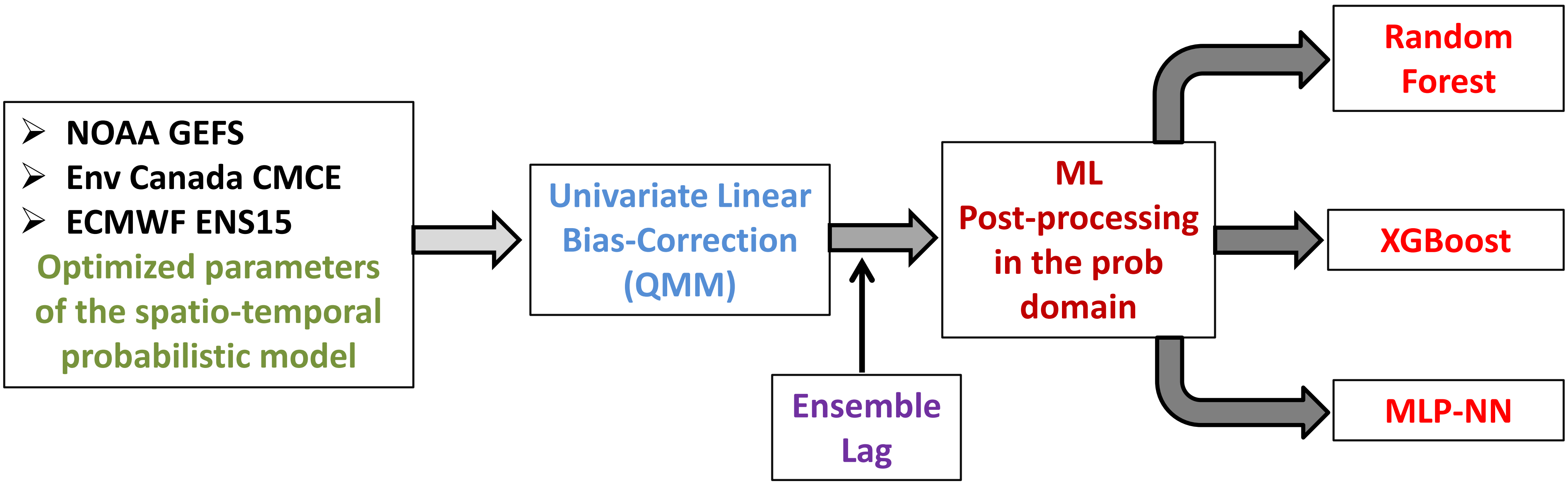




Validation set is too small (255): 3 months and 3 points

Final remarks - Ongoing and Next steps

- Add more data, run the optimization, and re-train the models;
- Process tropical cyclones separately;
- Lagged ensembles and multi-model;
- Combine sequential layers of QMM and ML models;
- Include it in the experimental operational forecast.



Thank you

ricardo.campos@noaa.gov

Acknowledgments:
I acknowledge the support of OPC, NHC, EMC, CIMAS, and AOML for their significant contributions. This work utilized computational resources from the Orion supercomputer at Mississippi State University (in partnership with NOAA) and AWS.