Improving NCEP's Global-Scale Wave Ensemble Averages using Neural Networks: Results and Next Steps

Ricardo M. Campos
Instituto Superior Técnico, University of Lisbon

ricardo.campos@centec.tecnico.ulisboa.pt
riwave@gmail.com

"Improving Global Wind-Wave Probabilistic Forecasts and Products Beyond Week 2"
2017-2019, Award Number: NA16NWS4680011
Vladimir Krasnopolsky, Jose-Henrique Alves, Steve Penny, Ricardo Martins Campos.

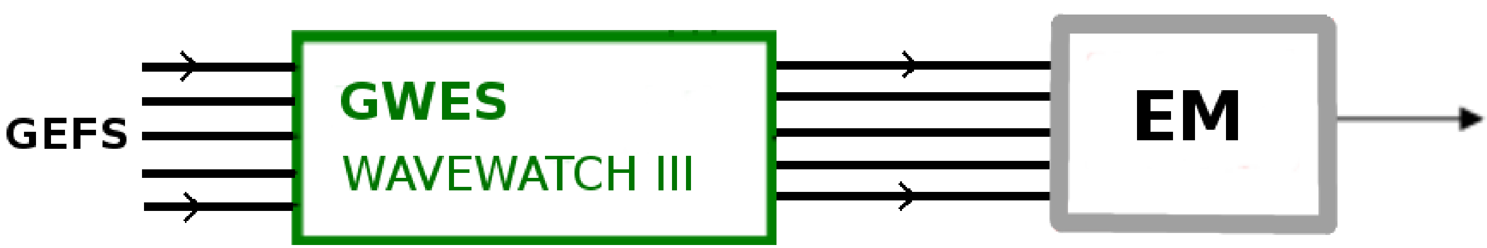
Outline

- Quick introduction to GWES
- MLP Neural Networks applied to non-linear ensemble averaging
- **(1)** First tests at single locations
  - NN Architectures
  - Tests with number of neurons, normalization etc

- NN spatial approach
  - NN Training Strategy
  - Spatial Distribution of Wind and Wave Climates
  - Assessment of GWES using NDBC buoys and Altimeters
  - Large sensitivity test: number of neurons, initialization, filtering
  - ➢ **(2)** GoM and **(3) Global** simulations

Global Wave Ensemble System (GWES)

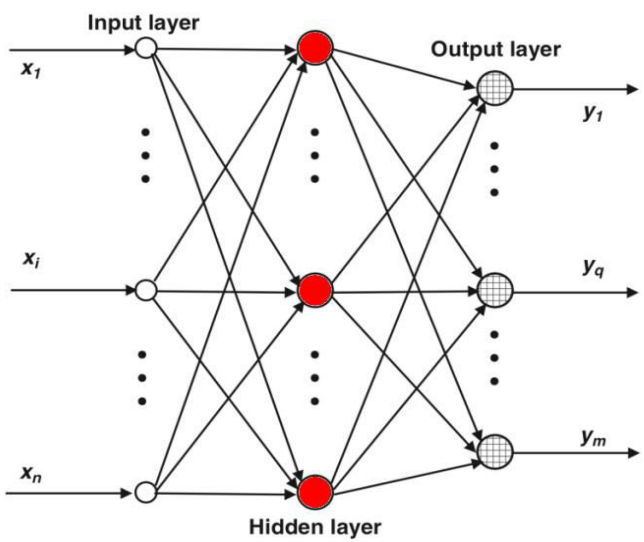
- The GWES was implemented in 2005 (Chen, 2006);
- 4 cycles per day;
- Resolution of 0.5 degree and 3 hours;
- Forecast range of 10 days;
- Total of 20 ensemble members plus a control member
- Forced by Global Ensemble Forecast System (GEFS) winds on WAVEWATCH III model (Tolman, 2016)
- Last major upgrade: 12/2015
- Arithmetic Ensemble Mean: $EM = \frac{1}{n}\sum_{i=1}^{n} x_i$



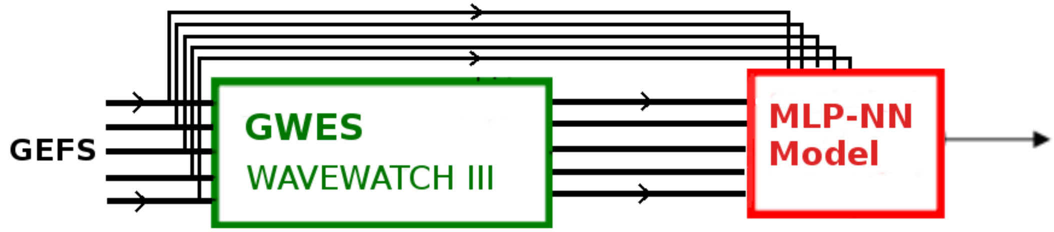
MLP Neural Networks

Multilayer perceptron model (MLP-NN) with hyperbolic tangent at the activation function. x_i is the input and y_q the output, a and b are the NN weights, n and m are the numbers of inputs and outputs respectively, and k is the number of nonlinear basis functions (hyperbolic tangents, or ¨neurons¨)

$$NN(x_1, x_2, \cdots, x_n; a, b) = y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \, . tanh\left(b_{j0} + \sum_{i=1}^{n} b_{ji} \, . x_i \right); \quad q = 1,2, \ldots, m$$



AI techniques provide a number of advantages, including easily generalizing spatially and temporally, handling large numbers of predictor variables, integrating physical understanding into the models, and discovering additional knowledge from the data (McGovern et al., 2017).

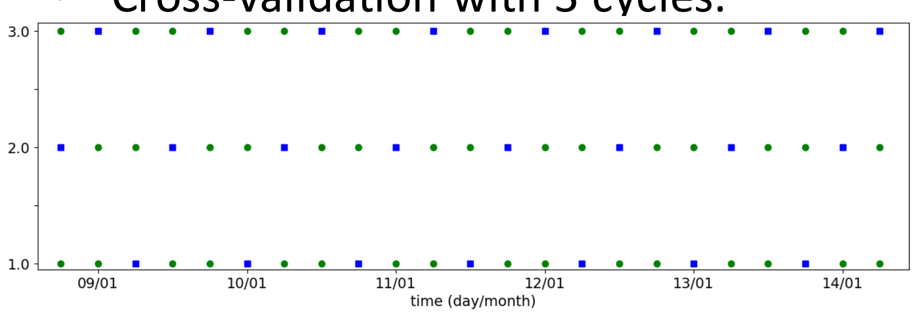


- Constructed based on Haykin (1999), Krasnopolsky (2013), and Krasnopolsky and Lin (2012)
- NNs have been used in a wide variety of meteorology applications since the late 1980s (Key et al. 1989), from cloud classification (Bankert 1994), tornado prediction and detection (Marzban and Stumpf 1996; Lakshmanan et al. 2005), damaging winds (Marzban and Stumpf 1998), hail size, precipitation classification, tracking storms (Lakshmanan et al. 2000), and radar quality control (Lakshmanan et al. 2007; Newman et al. 2013).

MLP Neural Networks

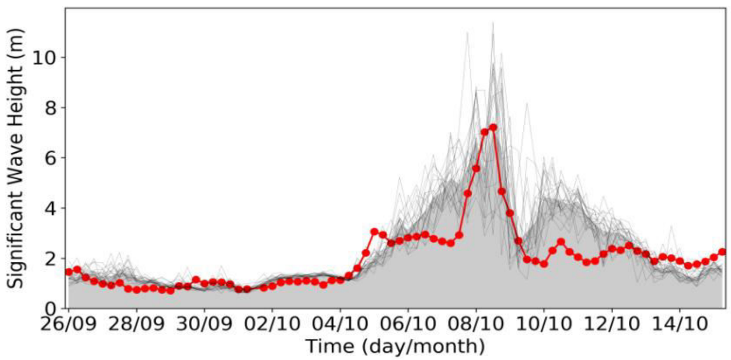
$$NN(x_1, x_2, \cdots, x_n; a, b) = y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} . tanh\left(b_{j0} + \sum_{i=1}^{n} b_{ji} . x_i\right); \quad q = 1,2, \dots, m$$

- Input variables: 10-meter wind speed (U10), significant wave height (Hs), peak wave period (Tp), mean period, wave height of wind-sea, wave period of wind-sea;
- Target variables: **U10, Hs**, Tp from measurements;
- Evaluated against buoy/altimeter observations during the training process;
- 21 ensemble members (20 plus the control member) per variable, plus the sin and cosine of time;
- Latitude and Longitude (sin,cos) are included as inputs during the regional analyses;
- One NN per forecast time / forecast time as new degree of freedom;
- Training (2/3) and test set (1/3);
- Cross-validation with 3 cycles.



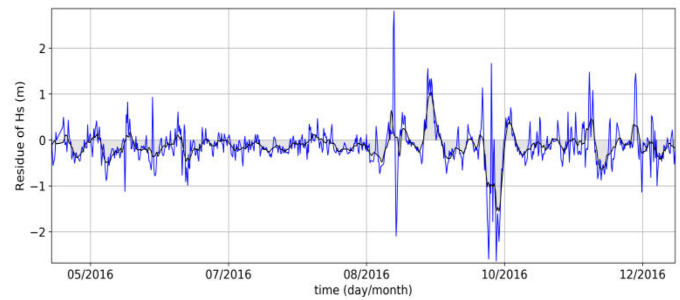
133 NN Inputs			2 NN Outputs	780 NNs
21 members	U10	Latitude	Residue U10	10 seeds
	Hs	Sine Longitude		
	Tp	Cosine Longitude		3 independent datasets
	Tm	Sine Time	Residue Hs	
	WsH	Cosine Time		26 different numbers of neurons
	Tws	Forecast lead time		
		NCEP/GWES cycle		

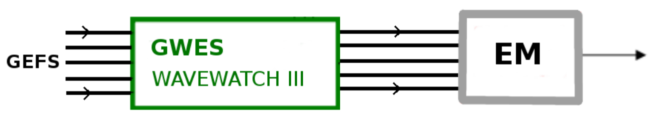
First tests at single locations

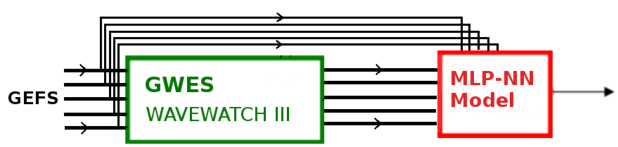


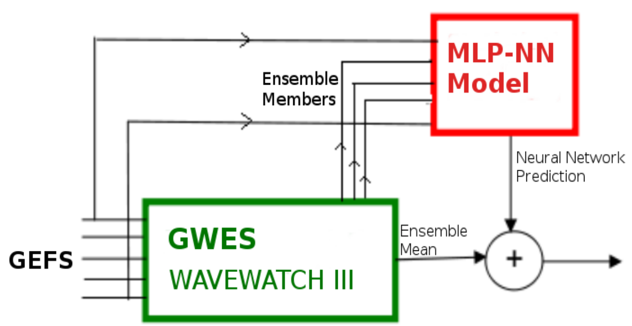
1. NN models are indicated primarily to nonlinear problems;
2. NN cannot deteriorate the EM!

Residue (measurements - model**) as the target variable**









$EM = \frac{1}{n}\sum_{i=1}^{n} p_i$	(1)
$NEM = NN(p_1, p_2, \cdots, p_n)$	(2)
$NEM = EM + NN_r(p_1, p_2, \cdots, p_n)$	(3)

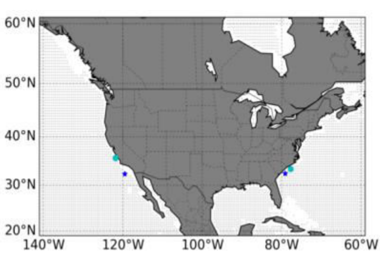
First tests at single locations

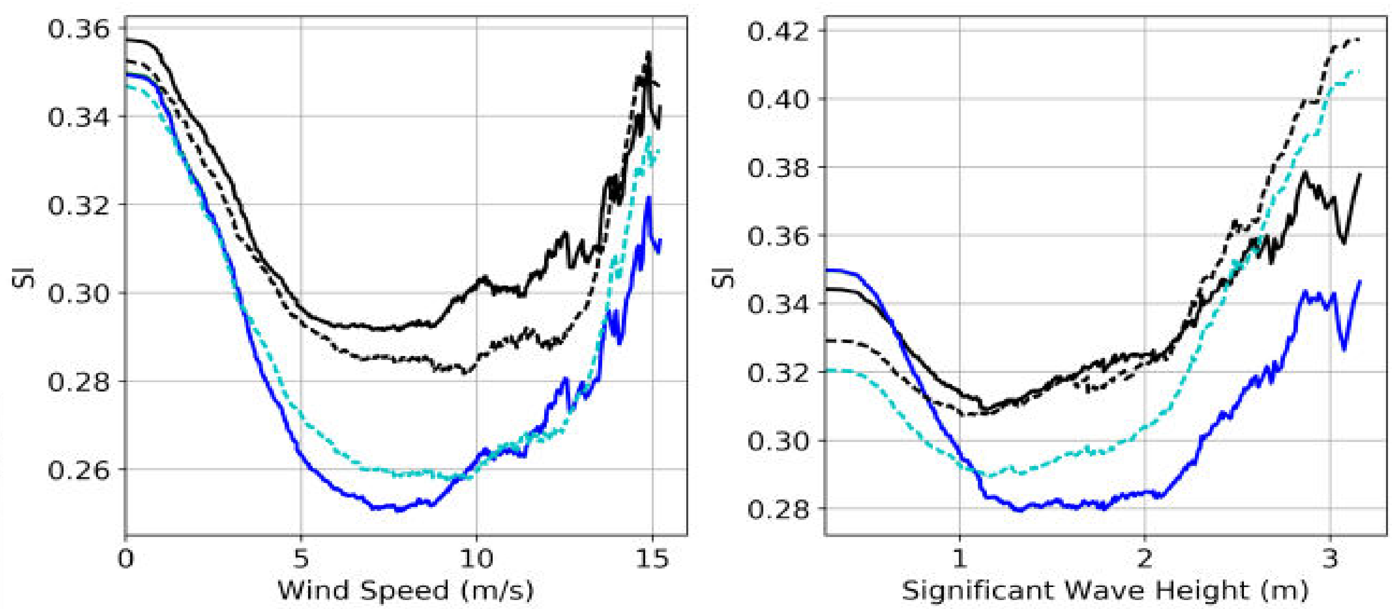
The best NN model: 11 neurons at the intermediate layer

41004	bias	RMSE	SI	CC
Best Member GWES	-0.101	0.526	0.427	0.724
EM GWES	-0.115	0.457	0.371	0.755
Linear Regression model	0.094	0.433	0.352	0.739
NN ensemble (5 members)	0.041	0.373	0.303	0.807

Reduction of the error with increasing quantiles.

Results of the NN simulation at the two Atlantic Ocean buoys. Curves of scatter indexes as a function of quantiles; black: arithmetic mean of ensembles (EM); blue: NN-training set (buoy 41004), cyan: NN-validation set (buoy 41013). Solid lines indicate buoy 41004, and dashed lines buoy 41013.



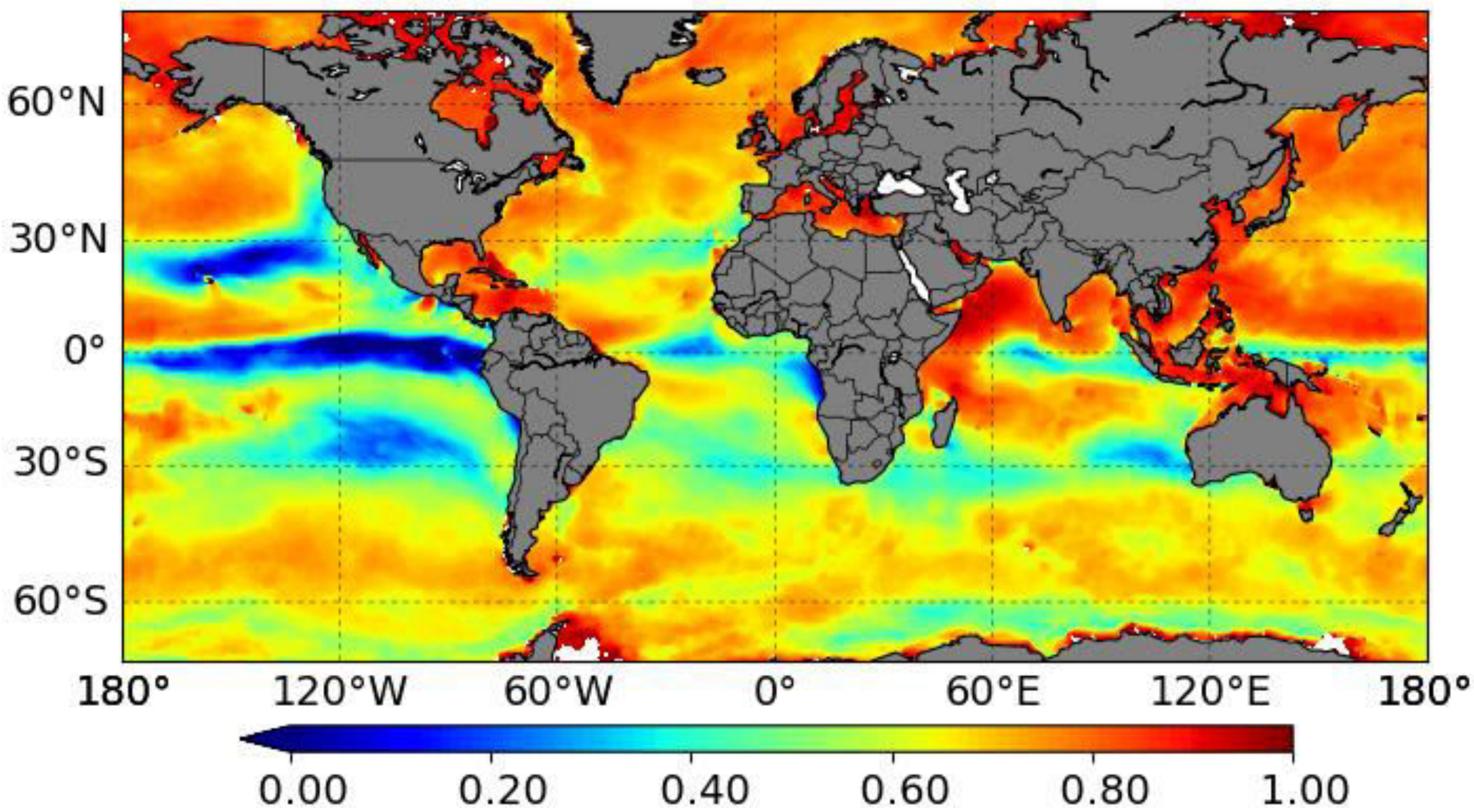


NN spatial approach

- Introduction of Lat/Lon as input variables instead of building one NN per grid point;
- Increase of NN complexity, Krasnopolsky (2013):

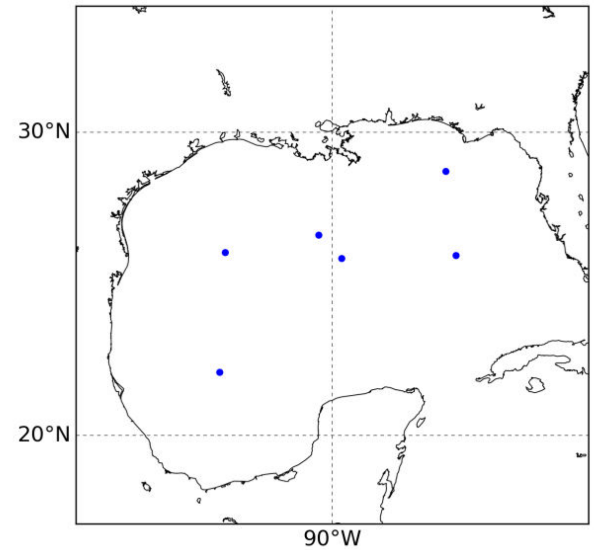
$$\boldsymbol{N_c = k.(n + m + 1) + m}$$

Different wind and wave climates. Correlation Coefficient Map of U10 and Hs



NN spatial approach - GOM

Simulation at the Gulf of Mexico. Sensitivity test:



- Total of 12 different numbers of neurons

N [2, 5, 10, 15, 20, 25, 30, 35, 40, 50, 80, 200]

- 8 different filtering windows

FiltW [0, 24, 48, 96, 144, 192, 288, 480] hours

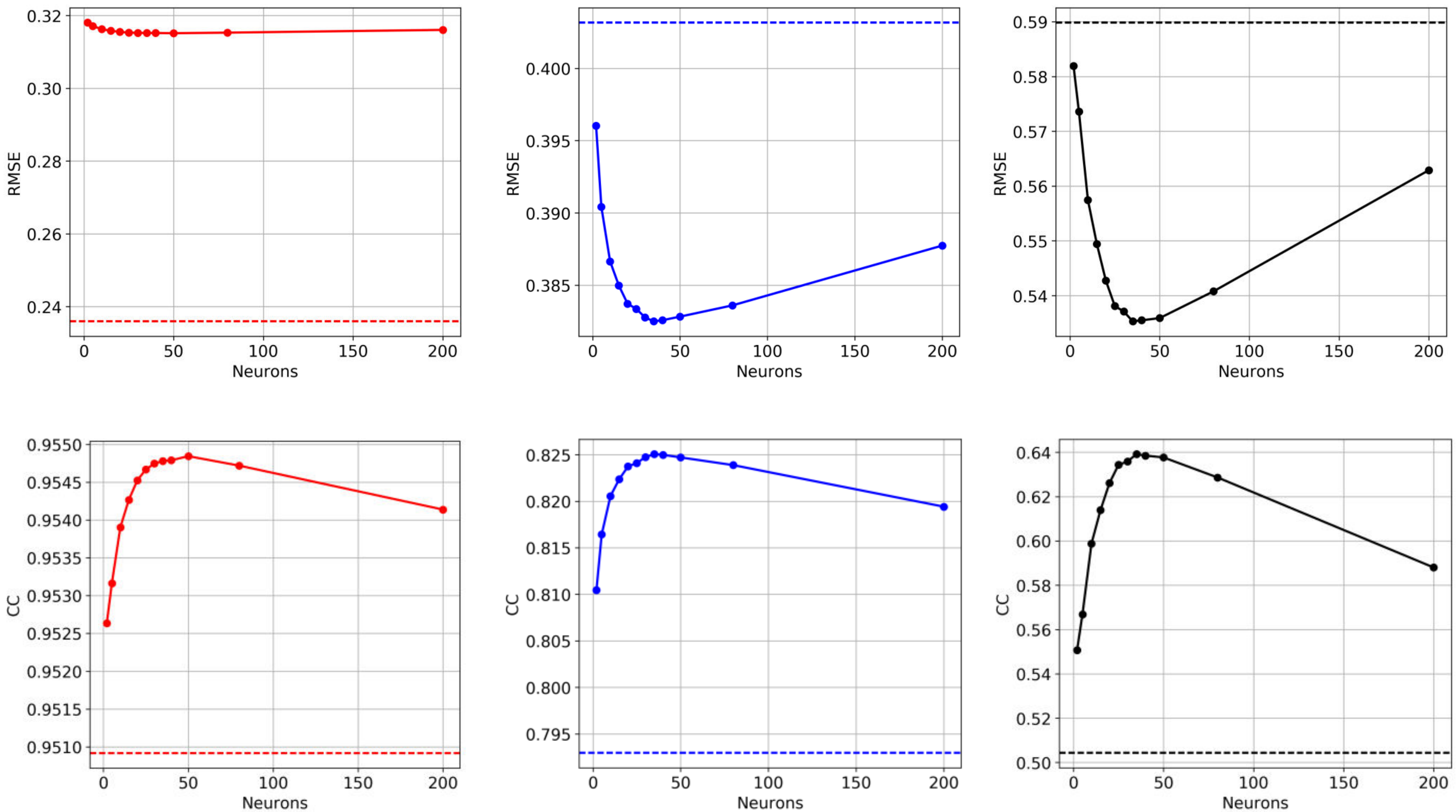
- 100 seeds for the random initialization

- Separated NNs for specific forecast days, from Day 0 to Day 10
- Total of **105,600** NNs
- NN training, 2/3 of inputs were selected for training and 1/3 for the test set, using a cross-validation scheme with 3 cycles
- *scikit-learn* (python) to reduce computational cost
- Six buoys appended to build the array with size 7913.

NN spatial approach - GOM

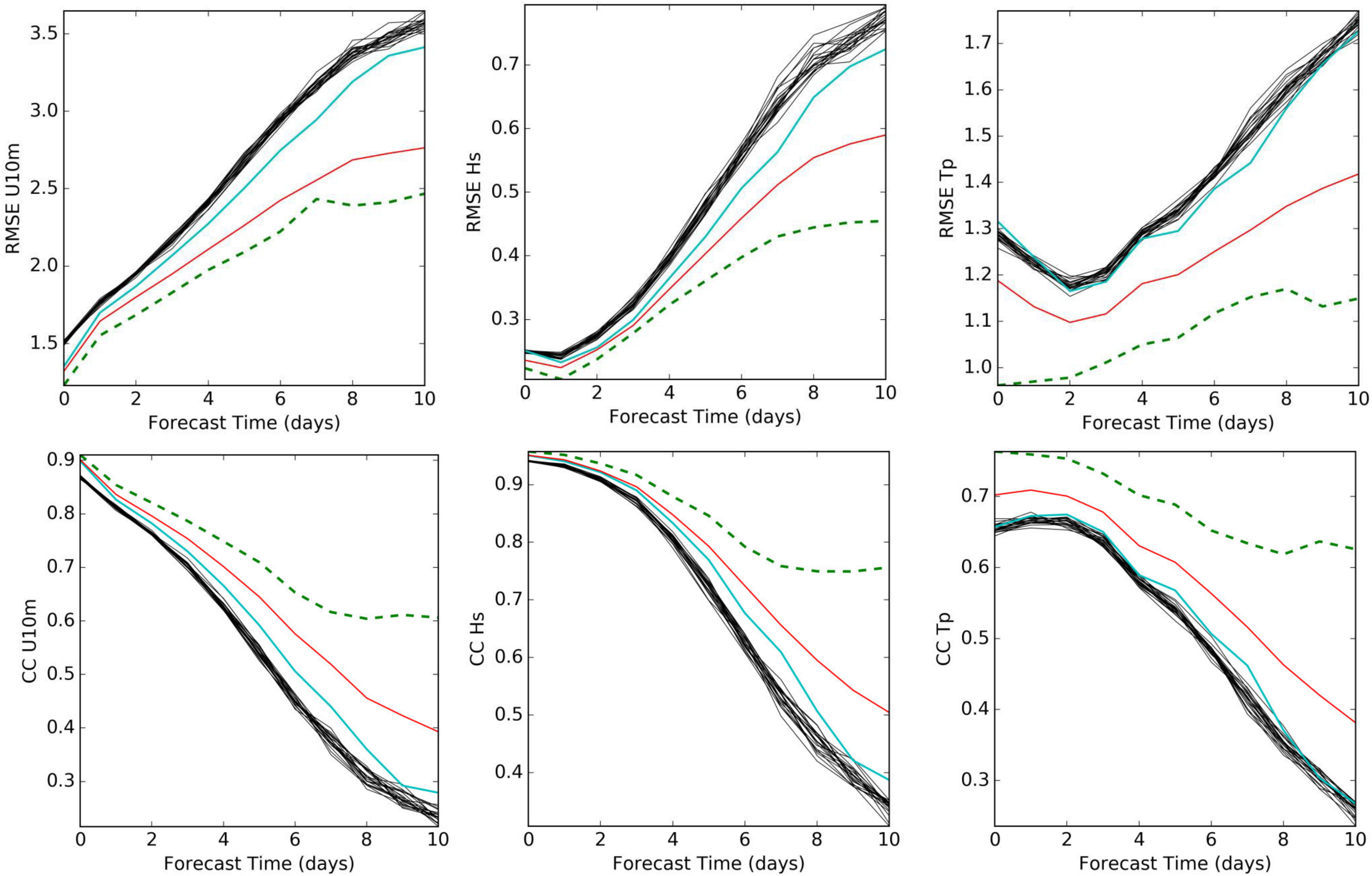
Hs

Day 0 Day 5 Day 10



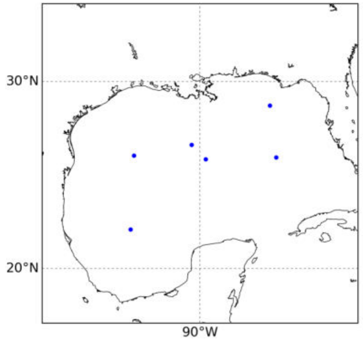
Results: NN spatial approach - GOM

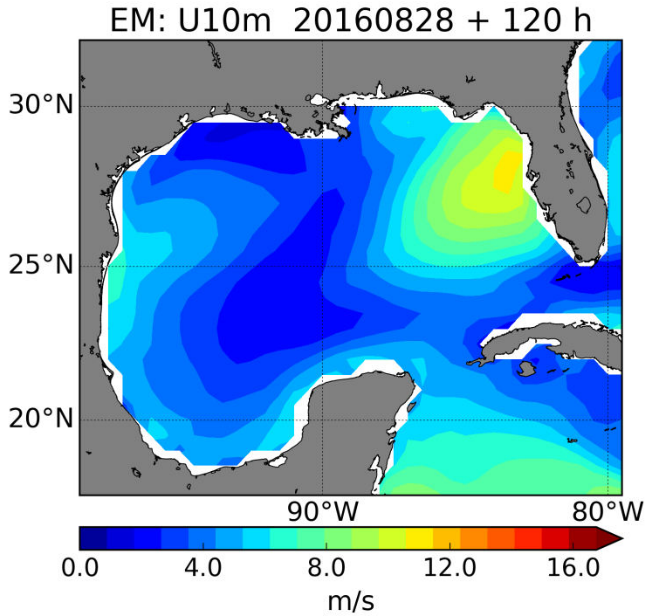
-Black: ensemble members
-Red: ensemble mean
-Cyan: control run
--Green: NN

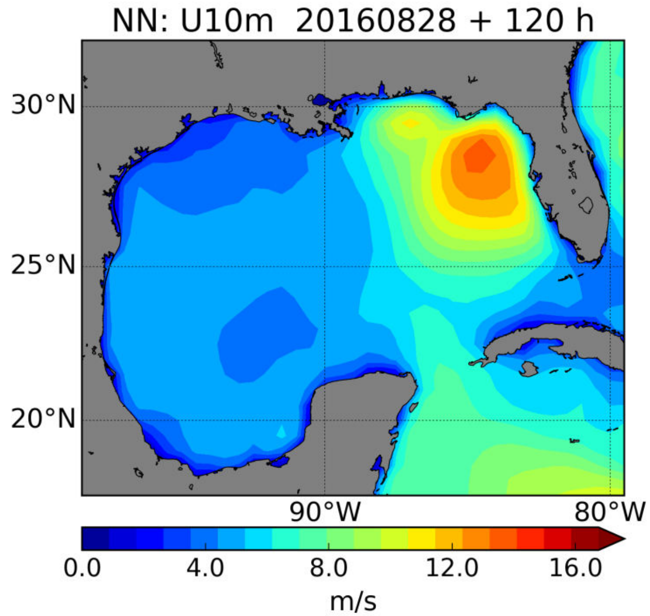


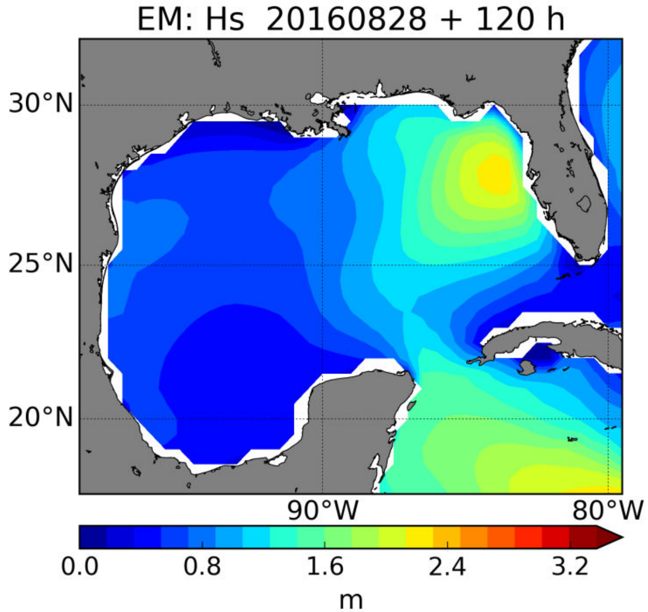
NN spatial approach - GOM

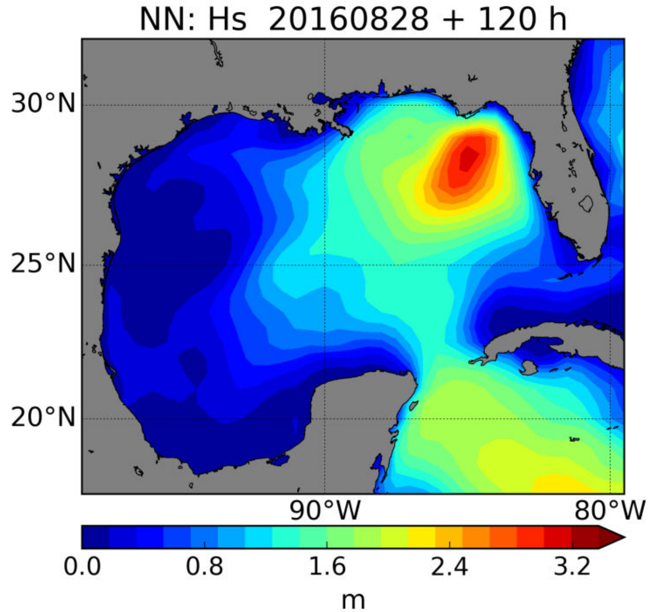
Hurricane Hermine
(September 02, 2016 – 00Z)
Highest winds (1-minute sustained): 80 mph (36 m/s)
Lowest pressure: 981 hPa





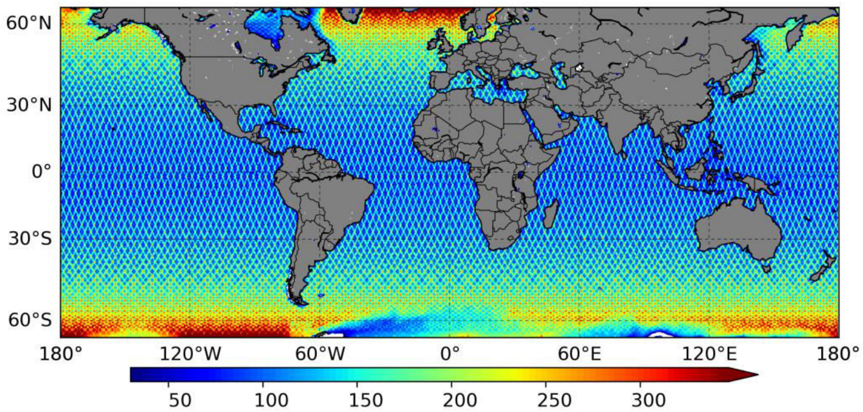


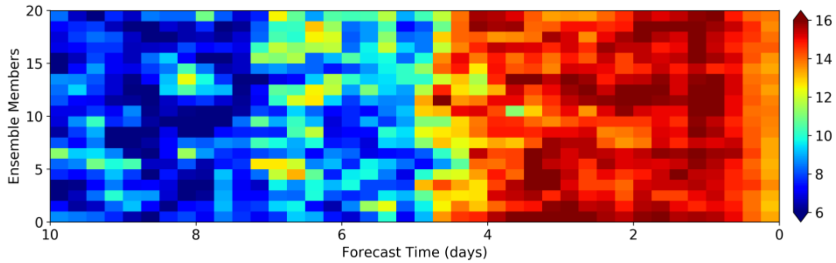




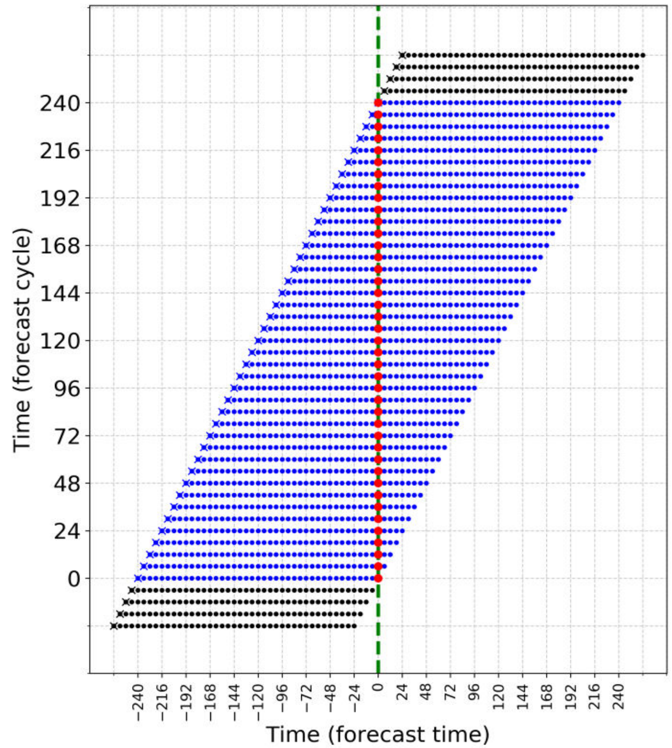
NN spatial approach - Global

- NN modeling (whole globe) using **<u>altimeter data</u>** and joining all forecast times into the training





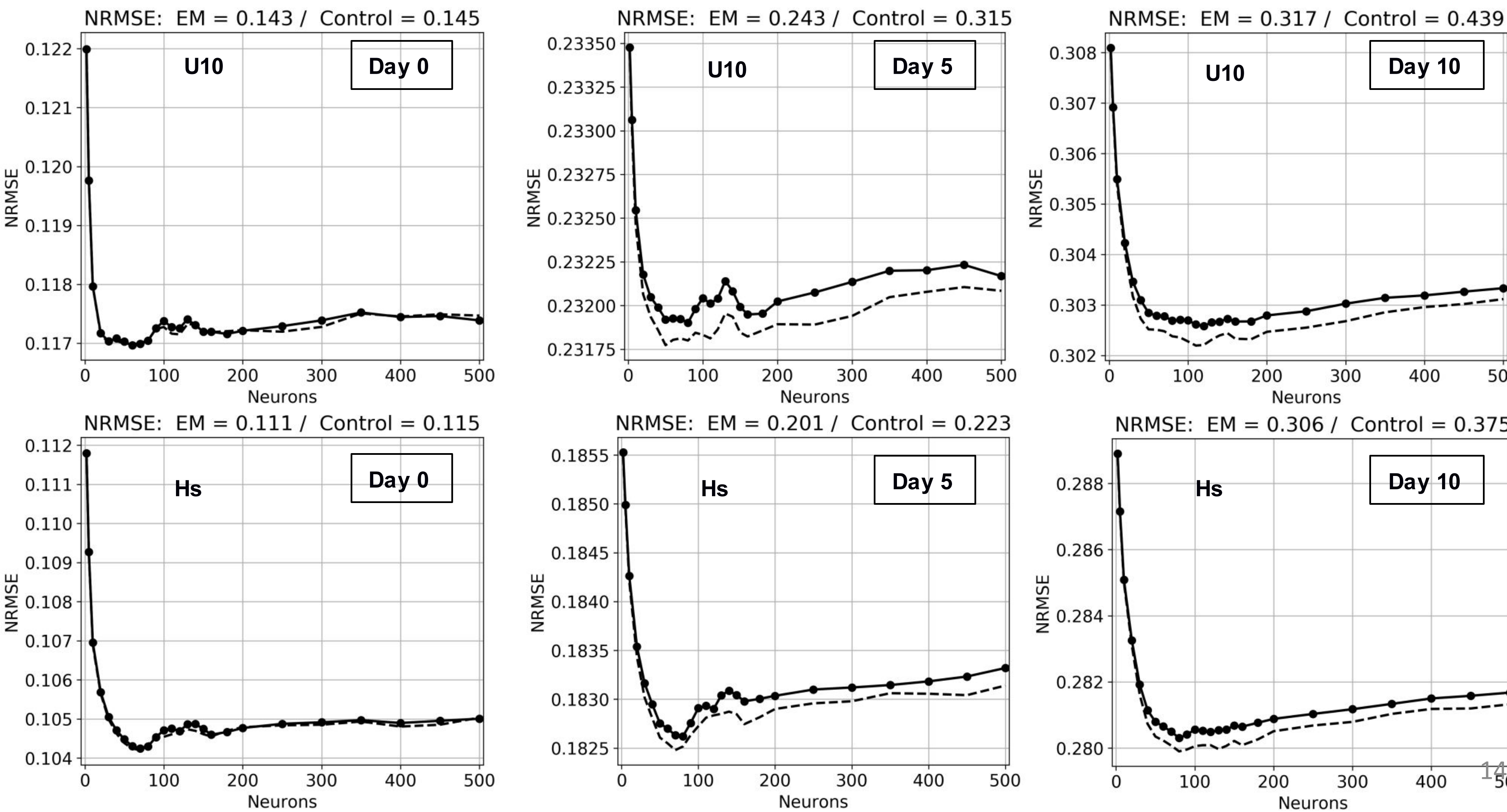
21 members x 41 fleads - GWES 54.4°S/74.5°W on 2017/06/10 12Z, HsAlt=13.8 m



- 07/2016 – 07/2017
- 4 satellite missions: **15,993,200** measurements
- **26 neurons [2-500], 10 seeds, and 3 datasets total of 780 NNs**

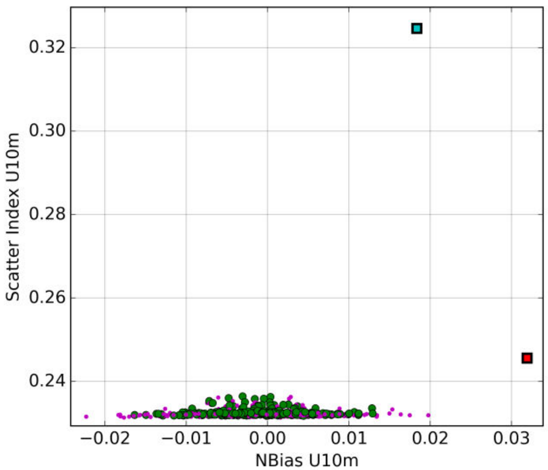
NN spatial approach - Global

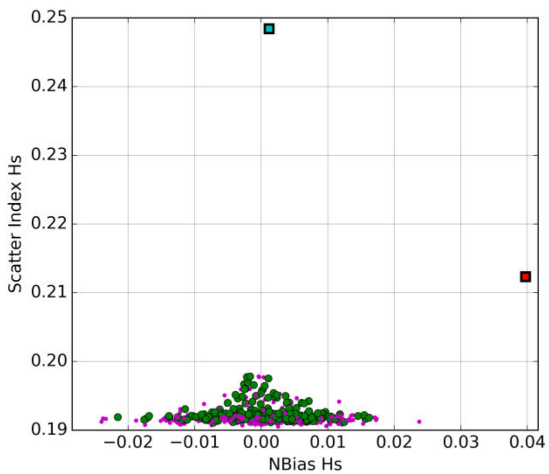
- Fday 0: sharp decay of the curve; 50 to 80 neurons
- Fday 5: second minimum is found around 160 to 180 neurons.
- Fday 10: best results using between 120 to 180 neurons. Higher RMSE for NNs with neurons equal to or less than 110.
- *The increasing scatter error of the surface winds at longer forecast ranges requires more complex NNs*. Distance between the NN curves of training and test set.

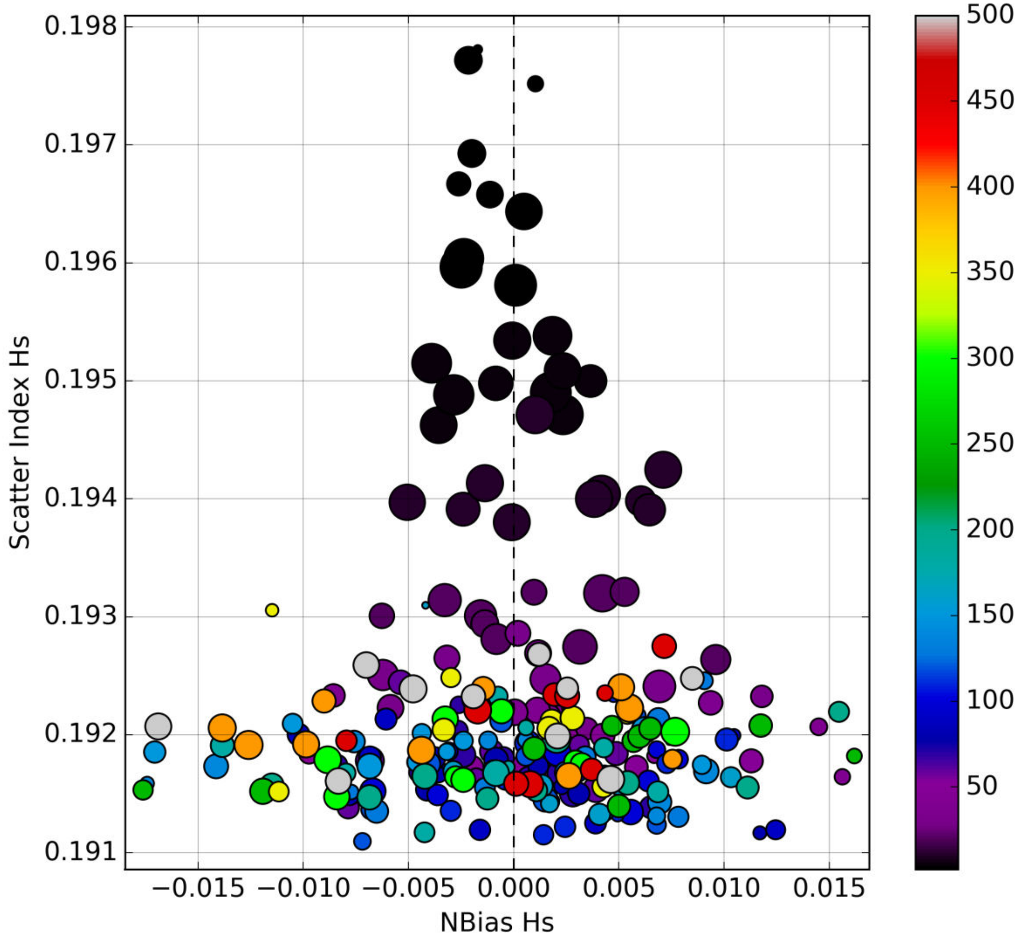


NN spatial approach - Global

- Results of 260 NNs, in terms of scatter error (y-axis) and systematic error (x-axis).
- Left plots: training set in magenta and test set in green, compared to the arithmetic EM (red square) and the control run (cyan square)
- Right plot: zoom-in the test set. Colorbar: number of neurons. Size of dots: normalized standard deviation of scatter error throughout different forecast ranges.

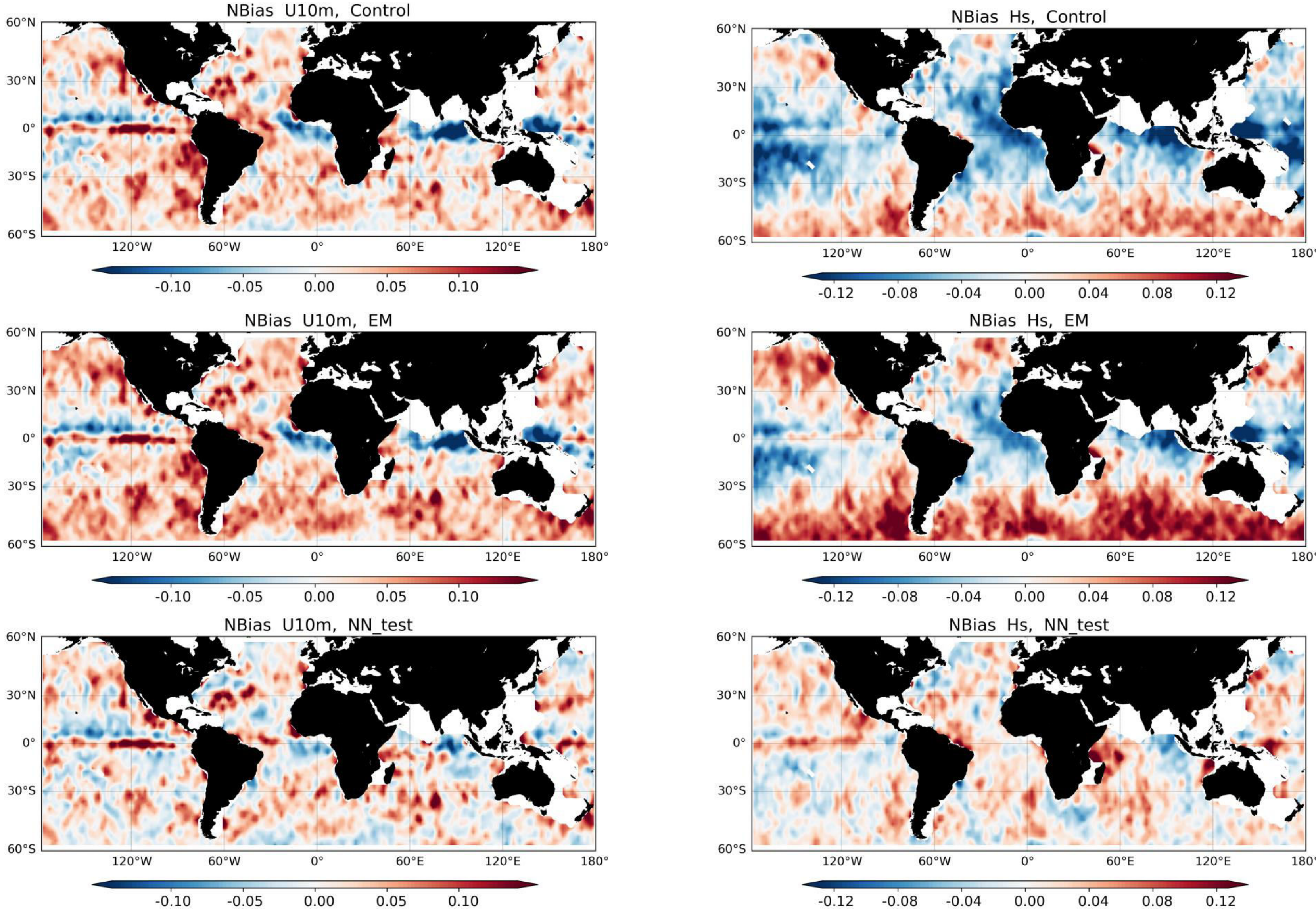






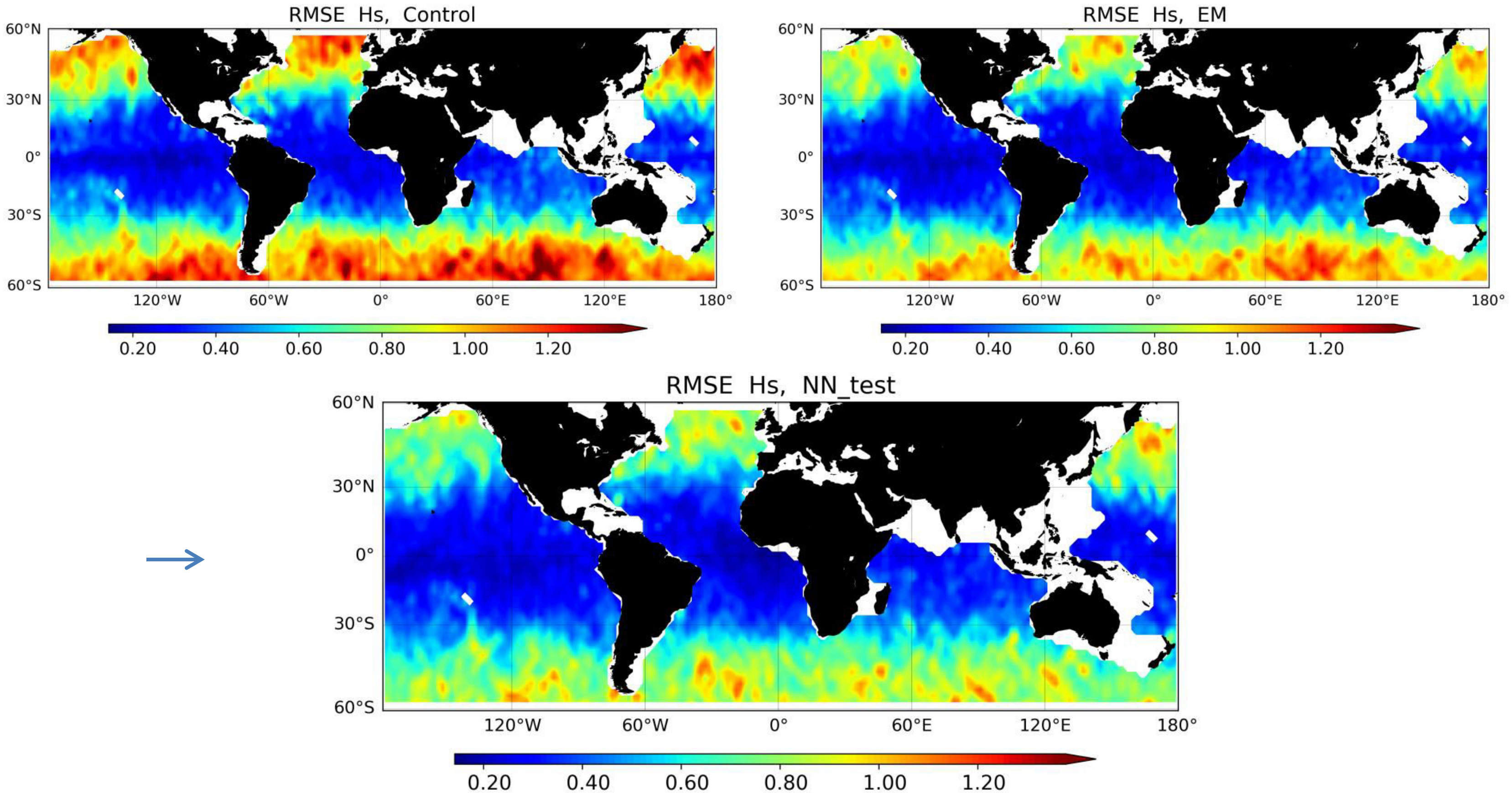
NN spatial approach - Global

- “Best” NN selected based on three steps (ranking and excluding large errors). Final NN containing 140 neurons at the hidden layer. Weights/Biases and normalization parameters saved.



NN spatial approach - Global

Significant Wave Height Hs

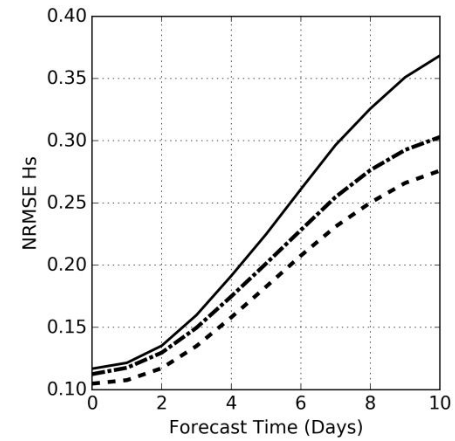


		U10m					Hs				
		NAtlantic	SAtlantic	Indian	NPacific	SPacific	NAtlantic	SAtlantic	Indian	NPacific	SPacific
Nbias	Control	0.016	0.018	0.013	0.010	0.029	-0.031	0.002	0.022	-0.017	0.008
	EM	0.029	0.034	0.030	0.019	0.041	0.001	0.041	0.065	0.017	0.048
	NN-Test	0.007	0.008	0.005	0.006	0.009	0.006	0.007	0.006	0.012	0.005
SI	Control	0.338	0.329	0.314	0.335	0.320	0.265	0.269	0.243	0.248	0.237
	EM	0.258	0.244	0.235	0.259	0.241	0.223	0.229	0.206	0.214	0.202
	NN-Test	0.245	0.231	0.223	0.242	0.229	0.208	0.209	0.183	0.197	0.182

Summary and Conclusions

- Wave Ensemble: Critical systematic and scatter errors are identified beyond the 6th- and 3rd- day forecasts, respectively.

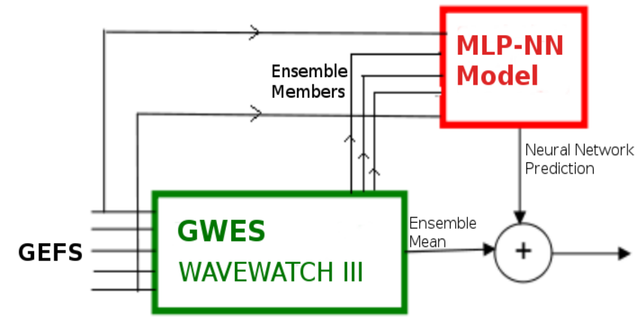
133 NN Inputs			2 NN Outputs	780 NNs
21 members	U10	Latitude	Residue U10	10 seeds
	Hs	Sine Longitude		
	Tp	Cosine Longitude		3 independent datasets
	Tm	Sine Time	Residue Hs	
	WsH	Cosine Time		26 different numbers of neurons
	Tws	Forecast lead time		
		NCEP/GWES cycle		



- Experiments demonstrated that one single NN model is able to improve the error metrics for the whole globe while covering all forecast ranges;
- The main advantage of the methodology: using NNs at longer forecast ranges beyond four days;
- Small number of neurons are sufficient to reduce the bias, while 140 neurons produce the greatest reduction in both the scatter and systematic errors (35 to 50 in the GoM);
- 60 to 80 neurons can minimize the errors of nowcast while 120 to 140 neurons are required to properly reduce the errors of day 10 forecast;
- The operational implementation of the nonlinear ensemble averaging using NN is simple! Files with Weights/Biases and normalization parameters.

Next steps

- The architecture/structure of the system is built and organized: can be re-trained with other numerical models and/or NN models;
- Limitation: still smoothing the fields at longer forecast leads, when systems (cyclones) become distant to each other in the GWES members;
- The performance of the hybrid modeling will be checked in terms of time spent after training.



- ➢ More tests with NN algorithms;
- ➢ Regionalization;
- ➢ Ensemble of Neural Networks (Krasnopolsky & Lin, 2012);
- ➢ Expand the forecast horizon to 16 days (or more);
- ➢ Multi-model ensemble including: GEFS/GWES, CMCE, FNMOC, ICON/DWD etc;
- ➢ Include more data for training;
- ➢ NNs, probabilistic domain;
- ➢ Additional efforts in the pre- and post- processing (Filters, Wavelets, EOF);
- ➢ NNs for best track of cyclones.

Thank you

Campos, R.M., Krasnopolsky, V., Alves, J.H.G.M., Penny, S.G., 2017. Improving NCEP's Probabilistic Wave Height Forecasts Using Neural Networks: A Pilot Study Using Buoy Data. National Oceanic and Atmospheric Administration - Office Note 490, http://doi.org/10.7289/V5/ON-NCEP-490

Campos, R.M., Alves, J.H.G.M., Penny, S.G., Krasnopolsky, V., 2018. Assessments of surface winds and waves from NCEP Ensemble Forecast System. AMS - Weather and Forecasting, 33, 1533-1546.

Campos, R.M., Krasnopolsky, V., Alves, J.H.G.M, Penny, S.G., 2019. Nonlinear Wave Ensemble Averaging in the Gulf of Mexico using Neural Networks. Journal of Atmospheric and Oceanic Technology, 36, 113-127.

Campos, R.M., Alves, J.H.G.M., Penny, S.G., Krasnopolsky, V., 2020. Global Assessments of the NCEP Ensemble Forecast System using Altimeter Data. Ocean Dynamics, 70, 405–419.

→ **Campos, R.M., Krasnopolsky, V., Alves, J.H.G.M, Penny, S.G., 2020. Improving NCEP's Global-Scale Wave Ensemble Averages Using Neural Networks. Ocean Modeling, 149.**

Nonlinear Wave Ensemble Averaging using Neural Networks

- Ricardo Campos, riwave@gmail.com ricardo.campos@centec.tecnico.ulisboa.pt

"Improving Global Wind-Wave Probabilistic Forecasts and Products Beyond Week 2"
Award Number: NA16NWS4680011

- Vladimir Krasnopolsky, Jose-Henrique Alves, Steve Penny, Ricardo Martins Campos.