GFS Wave vs Multi-1 Evaluation Summary

Deanna Spindler, Todd Spindler, Roberto Padilla

Summary Statistics from the websites

Buoys: Evaluate near coastal areas

https://polar.ncep.noaa.gov/waves/validation/gfsv16/buoys/

Sig. wave height (Hs), Peak period (Tp), wind speed (u10) and direction (udir)

Click on “Aggregated Buoys”, then “Statistics”

- Subjective based on the aggregated stats

Click on “Summary Stats”, then “By Month” or “By Forecast”

- By month: aggregated buoys for the entire month.
- By forecast hour: aggregated buoys by forecast hour for June 2019 - Aug 2020.

Statistics:

Bias, RMSE, Cross-Correlation, Scatter Index, 95th Percentile

Satellites: Evaluate over the open ocean

https://polar.ncep.noaa.gov/waves/validation/gfsv16/satellite/

Click on “Summary”, then “Significant Wave Height” or “Wind Speed”

Sig. wave height, wind-speed

- By month

Same website as above, then choose “Click here for monthly”, then “Field Maps”

For Sig. wave height:

- Per month, by forecast hour: daily averages of field maps of Bias and RMSE (produced with Global Ocean L4 Sig. Wave Height from NRT Satellite Measurements, generated using EU Copernicus Marine Service Information)

Statistics: **Bias, RMSE, 95th Percentile**

Rating Procedure

- Individually examine the **nowcast, 24hr, 96 hr, 168 hr** forecasts.
- For each variable and statistic, subjectively rate whether GFS Wave or Multi-1 was better, assigning values from 0 to 2. Most were 0 or 1, 2 were given if much better.
  - RMSE, Cross-correlation, and Scatter Index were given higher importance, with bias and 95% used as tie-breakers.
- Add up the scores across stats per variable, and get a winner for that forecast hour: if most of the stats are for model A then +A (3 points), if model A is better A (2 points), just slightly better then -A (1 point).
- Add up the +A, A, -A scores for the final overall score.
- For satellites: Global, North Atlantic (at_10m), and West Coast of US (wc_10m).

Buoys Subjective

Orange => MULTI-1(M) Blue=> GFS WAVE (G) T = Tie

NOWCAST

Nowcast	Hs	u10	Tp
Bias	M	M	G
RMSE	M	G	G
CC	T	G	G
SI	G	G	G
95p	M	M	G
OverAll	M	+G	+G

24 HR

24 hr	Hs	u10	Tp
Bias	M	M	G
RMSE	M	T	G
CC	M	G	G
SI	T	G	M
95p	M	M	G
OverAll	+M	-G	+G

96 HR

96 hr	Hs	u10	Tp
Bias	M	M	G
RMSE	G	T	G
CC	T	G	G
SI	G	T	M
95p	M	T	G
OverAll	-G	T	+G

168 HR

168 hr	Hs	u10	Tp
Bias	G	M	G
RMSE	T	M	G
CC	M	G	G
SI	T	G	M
95p	G	M	G
OverAll	-M	-G	+G

Buoys Objective, by month

Orange => MULTI-1(M) Blue=> GFS WAVE (G) T = Tie

NOWCAST

Nowcast	Hs	u10	Tp
Bias	M	M	G
RMSE	M	G	G
CC	M	G	G
SI	M	G	M
95p	M	M	G
OverAll	+M	+G	+G

24 HR

24 hr	Hs	u10	Tp
Bias	M	M	G
RMSE	M	T	G
CC	M	G	G
SI	M	G	M
95p	M	M	G
OverAll	+M	G	+G

96 HR

96 hr	Hs	u10	Tp
Bias	M	M	G
RMSE	T	T	G
CC	T	G	G
SI	G	G	M
95p	M	M	G
OverAll	T	-G	+G

168 HR

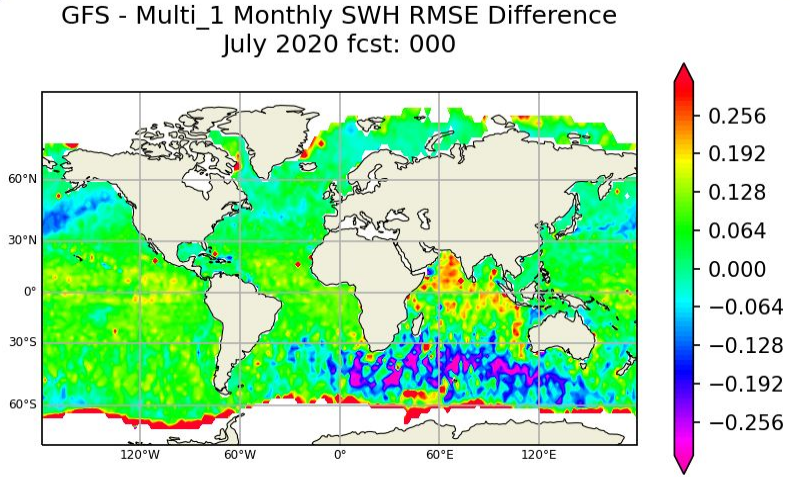
168 hr	Hs	u10	Tp
Bias	M	M	G
RMSE	G	T	G
CC	M	G	T
SI	T	T	M
95p	M	T	G
OverAll	-M	T	-G

Satellite Summary from website statistics

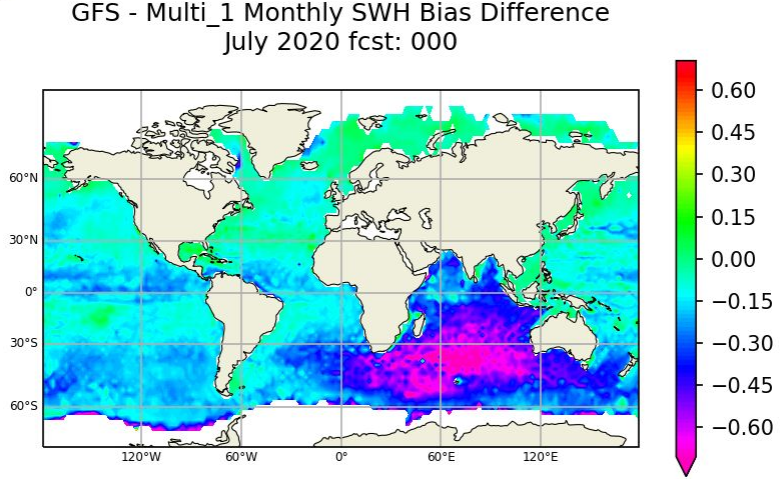
(this includes nowcast, 24, 96, and 168 hour forecasts)

	JASON-3	JASON-3	SARAL	SARAL	CRYOSAT-2	CRYOSAT-2	SENTINEL-3A	SENTINEL-3A
GLOBAL	Hs	u10	Hs	u10	Hs	u10	Hs	u10
MULTI-1		Marg Better		Tie	Tie		Better	
GFS WAVE	Much Better		Much Better	Tie	Tie	Much Better		Better
ATLANTIC								
MULTI-1	Better	Better		Much Better	Better	Tie	No Data	No Data
GFS WAVE			Marg Better			Tie	No Data	No Data
WEST COAST								
MULTI-1	Marg Better	Tie	No Data	No Data	Much Better	Better	Much Better	Tie
GFS WAVE		Tie	No Data	No Data				

Satellite, by month and fcst hour



Obs: Global Ocean L4 Significant Wave Height from NRT Satellite Measurements
(WAVE_GLO_WAV_L4_SWH_NRT_OBSERVATIONS_014_003)
Generated using EU Copernicus Marine Service Information



Obs: Global Ocean L4 Significant Wave Height from NRT Satellite Measurements
(WAVE_GLO_WAV_L4_SWH_NRT_OBSERVATIONS_014_003)
Generated using EU Copernicus Marine Service Information

Summary from website statistics

Buoys Overall	Sig Wave Height	Wind Speed	Peak Period
Subjective	Multi-1	GFS Wave	GFS Wave
Objective/month	Multi-1	GFS Wave	GFS Wave
Objective/fcst	Tie	Tie	GFS Wave

Satellite Overall	Sig Wave Height	Wind Speed
Global	GFS Wave	GFS Wave
North Atlantic (at_10m)	Multi-1	Multi-1
US West Coast (wc_10m)	Multi-1	Multi-1

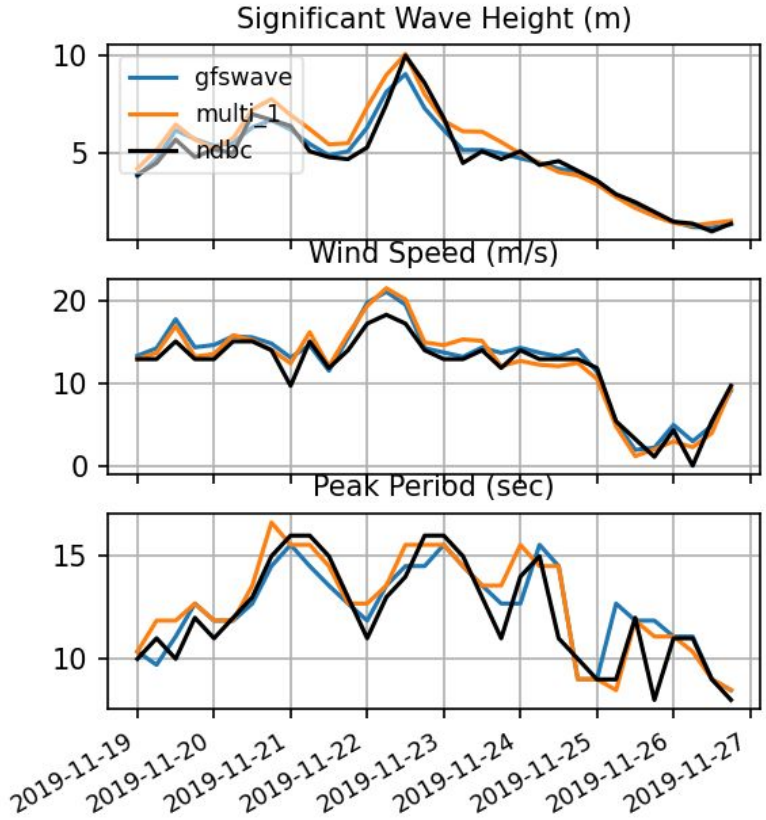
Conclusions from website statistics

- Globally, GFS Wave Sig. Wave Height has lower Bias and RMSE than Multi-1.
- Globally, GFS Wave is better at predicting significant wave height, peak period, and wind speed than Multi-1.
- Regionally, Multi-1 has better wind speeds and does better at predicting sig. wave height. GFS Wave does better predicting the peak period.
- The GFS Wave bias of sig. wave height tends to be too low. Seen but not shown here, is that GFS Wave sometimes gets the timing of the peak wave height better than Multi-1, but Multi-1 gets the actual height better.
- Near the coastal US, Multi-1 does better. The loss of the 4 arcmin grids also means the loss of the highest resolution winds and bathymetry. Multi-1 is two-way nested between the 4 arcmin and 10 arcmin grids, meaning that the 10 arcmin grids get the benefit of the higher resolution winds and bathymetry leading to better wave prediction, as seen in the satellite summary.

Request to evaluate Hs for dangerous seas

- Dangerous seas cases: Hs >= 4m (12 ft), and Hs >= 7 m (23 ft)
- The 95th quantile for Hs is ~ 3 meters
  - This significantly limits the amount of data available
- Website statistics for buoys are calculated for all buoys for that day, and for satellites the entire track for the day is used
  - This means Hs < 4m dominate the statistics presented before
- Caveats for dangerous seas stats:
  - Using buoys as “truth” when Hs >= 7 m
  - Satellites significant wave height cut-off value = 11 m

310NM SSW of Kodiak, AK, November 2019



Hs

	GFSWAVE	MULTI_1
Bias	0.0236	0.4609
RMSE	0.5268	0.8083
Corr	0.9663	0.9554
SI	10.6104	12.0153

u10

	GFSWAVE	MULTI_1
Bias	0.9368	0.6601
RMSE	1.4993	1.5922
Corr	0.9719	0.9684
SI	5.1335	9.5479

Tp

	GFSWAVE	MULTI_1
Bias	0.0675	0.3524
RMSE	1.2421	1.0856
Corr	0.8306	0.8952
SI	9.7416	7.5679

For the time period in the plot

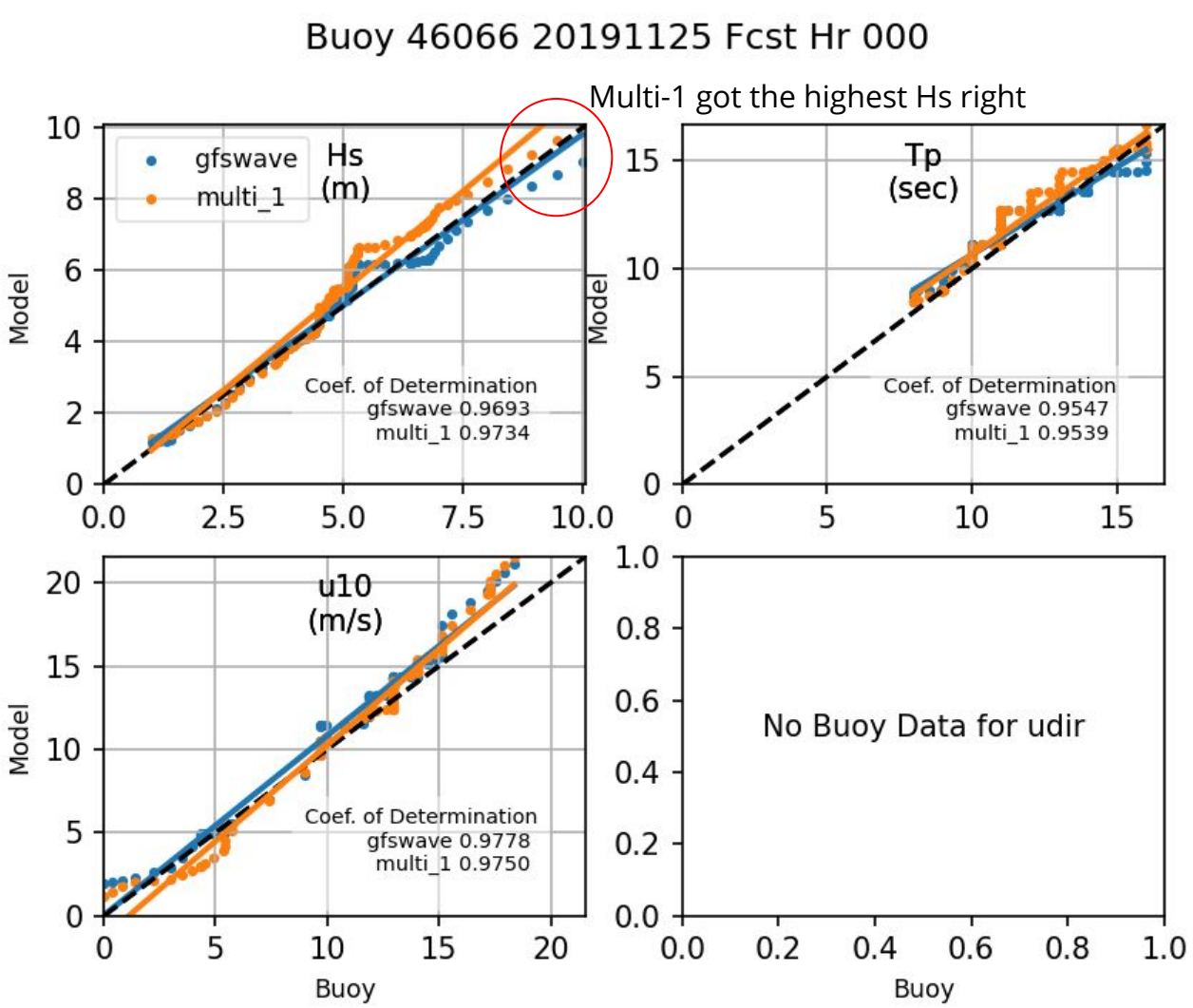
Bias of the Sig. Wave Height

Fcst	GFSWave	Multi-1
000	0.0236	0.4609
024	-0.0888	0.4528
048	0.0389	0.4798
072	-0.0383	0.3936

RMSE of the Sig. Wave Height

Fcst	GFSWave	Multi-1
000	0.5268	0.8083
024	0.5423	0.7699
048	0.5914	0.8464
072	0.9750	1.2413

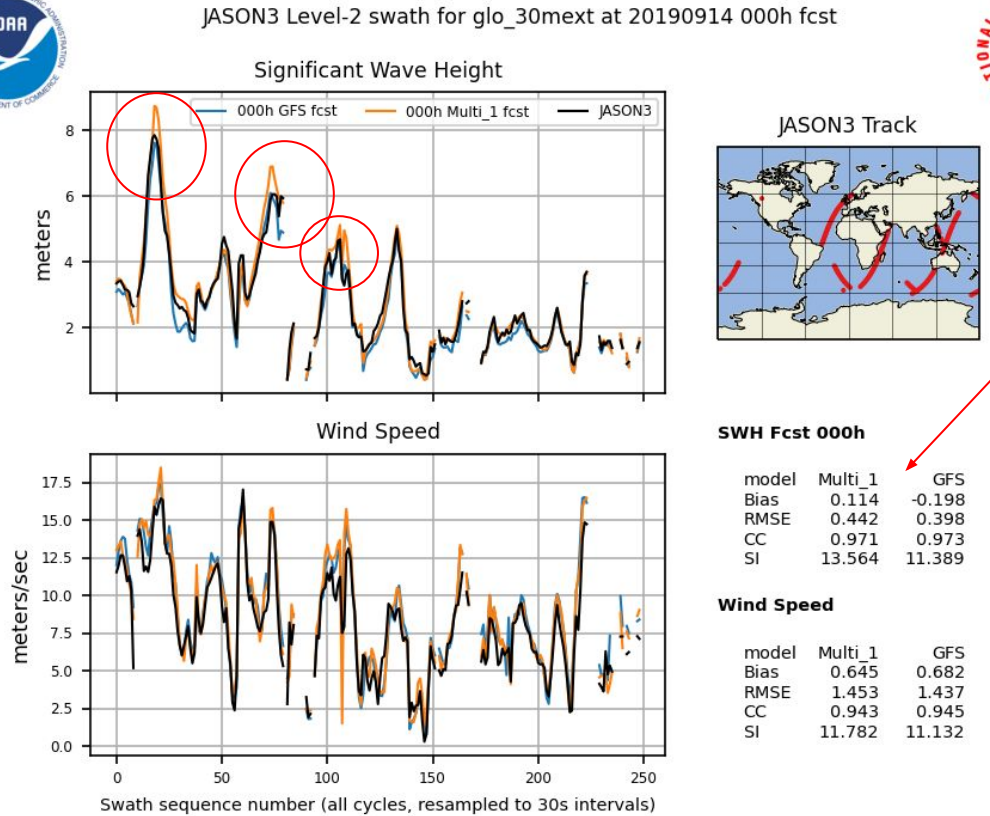
Note that most of this period has Hs >= 4 m



GFSWave
- Does really well Hs<5m, and 7m<=Hs<=8m
- Overestimates 5m<=Hs<=6m
- Under-estimates Hs > 8m

Multi-1
- Slightly underestimates Hs<4.5m
- Overestimates Hs>5m
- Slightly overestimates Hs>8m

Satellite tracks with dangerous seas: Hs > 4m



SWH Fcst 000h

model	Multi_1	GFS
Bias	0.114	-0.198
RMSE	0.442	0.398
CC	0.971	0.973
SI	13.564	11.389

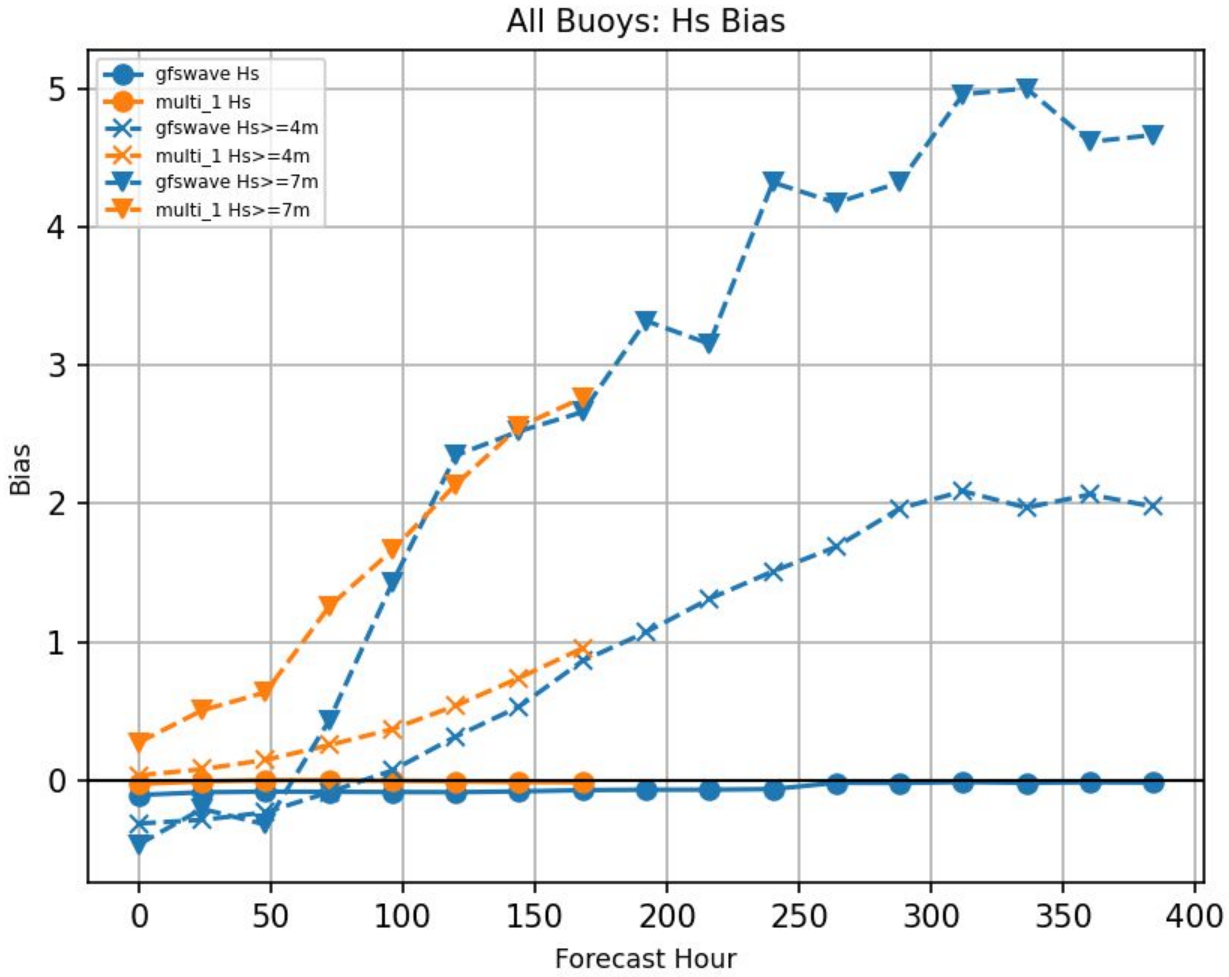
Wind Speed

model	Multi_1	GFS
Bias	0.645	0.682
RMSE	1.453	1.437
CC	0.943	0.945
SI	11.782	11.132

- Stats are based on the entire track for the day, and are dominate by Hs < 4m

- Multi-1 overestimates the highest waves, GFS Wave gets the correct height

Orange = Multi-1 Blue = GFS Wave



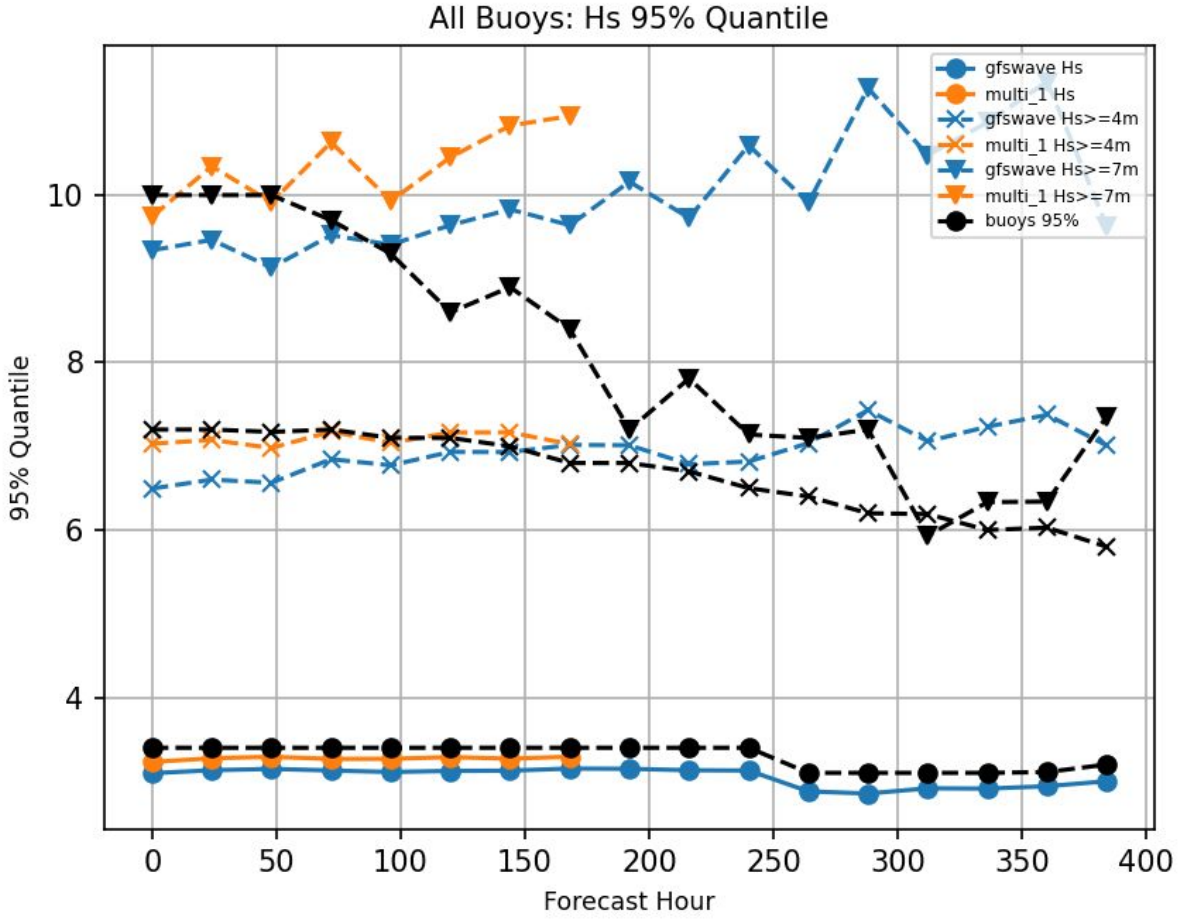
Sig. Wave Height Bias per forecast hour

- All buoy and models match-ups from June 2019 through Aug 2020

● ● => all Hs
X X => Hs >= 4m
▲ ▲ => Hs >= 7m

GFS Wave has lower bias for Hs >= 4m and for Hs >= 7m

Orange = Multi-1 Blue = GFS Wave



Sig. Wave Height
95% Quantile
per fcst hour

● ●=> all Hs
X X => Hs >= 4m
▲ ▲=> Hs >= 7m

Multi-1 has a slight edge over all Hs and Hs >= 4 m

For Hs >= 7, **sample size is quite small**, and both Multi-1 and GFS-Wave overpredict “dangerous seas” for fcsts > 100 hr

For fcsts < 50 hrs, GFS Wave underpredicts Hs >= 7m

Conclusions for dangerous seas

- The statistics shown in the evaluation websites are dominated by the Hs < 4m
- Similarly to the website stats, specific cases show that GFS Wave has lower bias and RMSE than Multi-1 for Hs >= 4m
- GFS Wave does really well predicting significant wave heights where 6 m <= Hs <= 8m. Multi-1 does better at Hs > 8m.
- For Hs >= 7m the sample size is quite small, but both Multi-1 and GFS-Wave overpredict these dangerous seas for fcsts > 100 hr