ERDC/CHL CHETN-IV-127
February 2021

US Army Corps of Engineers®

altWIZ: A System for Satellite Radar Altimeter Evaluation of Modeled Wave Heights

by Clarence O. Collins III and Tyler Hesser

PURPOSE: This Coastal and Hydraulics Engineering Technical Note (CHETN) describes the design and implementation of a wave model evaluation system, altWIZ, which uses wave height observations from operational satellite radar altimeters. The altWIZ system utilizes two recently released altimeter databases: Ribal and Young (2019) and European Space Agency Sea State Climate Change Initiative v.1.1 level 2 (Dodet et al. 2020). The system facilitates model evaluation against 1 Hz[1] altimeter data or a product created by averaging altimeter data in space and time around model grid points. The system allows, for the first time, quantitative analysis of spatial model errors within the U.S. Army Corps of Engineers (USACE) Wave Information Study (WIS) 30+ year hindcast for coastal United States. The system is demonstrated on the WIS 2017 Atlantic hindcast, using a 1/2° basin scale grid and a 1/4° regional grid of the East Coast. Consistent spatial patterns of increased bias and root-mean-square-error are exposed. Seasonal strengthening and weakening of these spatial patterns are found, related to the seasonal variation of wave energy. Some model errors correspond to areas known for high currents, and thus wave-current interaction. In conjunction with the model comparison, additional functions for pairing altimeter measurements with buoy data and storm tracks have been built. Appendices give information on the code access (Appendix I), organization and files (Appendix II), example usage (Appendix III), and demonstrating options (Appendix IV).

INTRODUCTION

Model Evaluation Methods. Third-generation spectral (i.e., phase-average) wave models solve the action balance equation for the evolution of wavenumber spectra in space and time (e.g., Cavaleri et al. 2007). The performance of a typical model simulation is evaluated by comparison to measurements. Some in situ measurements can provide two-dimensional (2D) frequency-direction spectra, $E(f,\theta)$. The 2D spectra contain statistical information about the stochastic wave field. Spectra are overly cumbersome to compare because of the dimensions $f \times \theta \times time$. No satisfactory technique has emerged for comparing large amounts of spectra, so it is customary to reduce the dimensions of the problem by calculating spectral metrics as a function of time. This collapses the frequency-direction axes, leading to timeseries of a small number of parameters that can be easily digested with statistics and diagrams (e.g., Krogstad et al. 1999; Collins 2012).

Disregarding the frequency-direction information, in situ observations provide hourly time-sampled data at a fixed point in space or along a drift track. If one is interested in resolving spatially dependent errors of a model, remotely sensed data are needed. Although there are hybrid approaches — such

[1] For a full list of the spelled-out forms of the units of measure used in this document, please refer to *US Government Publishing Office Style Manual*, 31st ed. (Washington, DC: US Government Publishing Office 2016), 248-52, https://www.govinfo.gov/content/pkg/GPO-STYLEMANUAL-2016/pdf/GPO-STYLEMANUAL-2016.pdf.

as X-band marine radar, stereo-video, scanning altimeter, scanning LIDAR, and satellite synthetic aperture radar, all of which provide limited space-time data-regional, basin, or global scale coverage is only available from spaceborne radar altimeters. The compromise with altimeter data is poor time-sampling at one location. Specifically, a single point in space is measured at best once a week. Also, there is no spectral information, only significant wave height (Hs).

Altimeter data provides Hs along a track directly beneath the satellite so that the along-track coverage is very good (~10 km). However, since most of these satellites are in polar orbits, the cross-track resolution is poor, ~100 – 400 km at the equator. Still, given approximately 7 to 14 days, depending on the satellite track, there is nearly complete global coverage. After 33+ years of collecting data, and depending on the variables served, this results in approximately 0.1 – 0.5 Tb of data, necessitating a system for handling the pairing of data.

An evaluation system for wave models is presented here. It was developed and implemented for model evaluation in general but specifically intended for use with the USACE WIS 30+ year hindcasts of the United States and territories coastal areas. The system complements evaluation against in situ buoys and other means, which have already been implemented for decades[1].

Altimeter Primer. Since the 1970s, there have been more than a dozen satellite missions carrying radar altimeters. Some satellites function individually, but others have been linked together in a constellation to optimize data coverage and collection. The repeat cycle, that is the time interval between measuring the same location in sequence, ranges from mission to mission, from ~month to ~week. There is an inherent tradeoff between the repeat cycle length and the cross-track distance between sequential tracks where a longer cycle allows closer spaced tracks while a shorter cycle necessitates longer distances between tracks. Each satellite carries altimeters that differ in detail but provide similar information to be processed.

Radar altimeters are mounted in a nadir-looking orientation on satellites following near polar orbits. They send out microwave pulses at ~1700 Hz, usually in the Ka band, but other bands are also used. The pulse spreads concentrically in space as it hits Earth's surface. The echo, or return signal, of a pulse has a characteristic waveform. On a flat surface, this waveform resembles a step function. As the surface roughness increases, the waveform becomes skewed with sloped leading and trailing edges. The shape of the waveform can be modeled with six parameters. Two of these parameters, the amplitude of the waveform and the slope of the leading edge, are related to wind speed and significant wave height, respectively.

Altimeter wave heights are calibrated against the significant wave height of operational buoy networks including the National Oceanic and Atmospheric Administration National Data Buoy Center and also cross-calibrated against concurrent satellite missions. Because there are so many satellites managed by different agencies, just getting all the data in one place, in a single format, is a massive undertaking, as well as ensuring the data are consistent across missions and time. Fortuitously, there are groups dedicated to just such a purpose, and they have produced consistent, calibrated, or otherwise adjusted altimeter data.

[1] Hesser T., A. Cialone, C. O. Collins, A. Cox, and R. Jensen. Under revision. "Wave Information Study: A 35+ Year Hindcast." *Journal of Coastal Research*.

One such project is the European Space Agency Sea State Climate Change Initiative project (henceforth, ESA), which is the successor to the GLOBWAVE project. The current release of the dataset is version 1.1, and it covers 10 missions — ERS-1, TOPEX, ERS-2, GFO, JASON-1, ENVISAT, JASON-2, CRYOSAT-2, SARAL, and JASON-3 — from August 1991 to December 2018 (Dodet et al. 2020). Along with their calibrated wave height (called "adjusted" because it is an adjustment to the original calibration), they also offer a product called "denoised significant wave height" (Quilfen and Chapron 2019 [QC19], Quilfen and Chapron 2020). QC19 used empirical mode decomposition as a noise filter. They were able to use this filter to produce a denoised time series.

Another such dataset was put together by Ribal and Young (2019 [RY19]). RY19 curated and inter-calibrated 13 altimeter missions, including all those within the ESA Sea State, as well as GEOSAT, HY-2A, and SENTINEL-3A, altogether spanning the 33 years between 1985 and 2018.

As mentioned, RY19 and ESA offer several wave height products, all served at 1 Hz. Both offer the original version of the data and recalibrated data. The calibration methods differ in detail, and thus there is some (usually linear) relationship between the calibrated wave heights from each database. This relationship is derived on a per-mission basis. Although the calibrated wave heights from each database may not match exactly, it is assumed that in either case, the recalibrated data are more consistent across time and missions than the original data.

altWIZ

Altimeter Data and Code Design. First, the altimeter dataset must be selected: either the RY19 or ESA databases can be implemented through specifying paths and options within the altWIZ system. Because these datasets are more than 100 gigabytes of data, part of the challenge is breaking down the data into sets of suitable size for processing. Both databases use netCDF files, and both have directories for each mission, but the organization differs from there. For RY19, the full time series is binned into 1°x1° files. Therefore, when a file is loaded, the entire time series is loaded for that 1°x1° bin. For ESA, each mission has a directory that is further organized into subdirectories by year, month, and day. There are multiple files for each day containing the complete sequential time series.

Anecdotally, the time it takes to retrieve altimeter data for model results (i.e., spatially distributed data from a month to years) is similar using either dataset, but retrieving data from a buoy (i.e., a multi-year time series at a single location) takes substantially longer with ESA data (presumably due to its organization).

The inputs into altWIZ are options for altimeter data and paths for the model directory and altimeter data directory. The end result is a MATLAB structure containing shared coordinates, time stamp, altimeter wave height, model (or buoy) wave height, averaging method, and the detailed metadata with options and paths used.

The logic of the altWIZ system is given in a simple flow chart in Figure 1 and goes as follows: (1) Given a monthly model file, create a list of altimeters missions active during that month. (2) Search and load all of the altimeter missions for data that overlap with the observations and filter the data according to the quality control flags (set by the user as an option). (3) Further

reduce the data by retaining measurements within a tolerance of the model time and space range. (4) Average the altimeter data (or not). (5) Three-dimensionally interpolate the model to match the altimeter sampling.

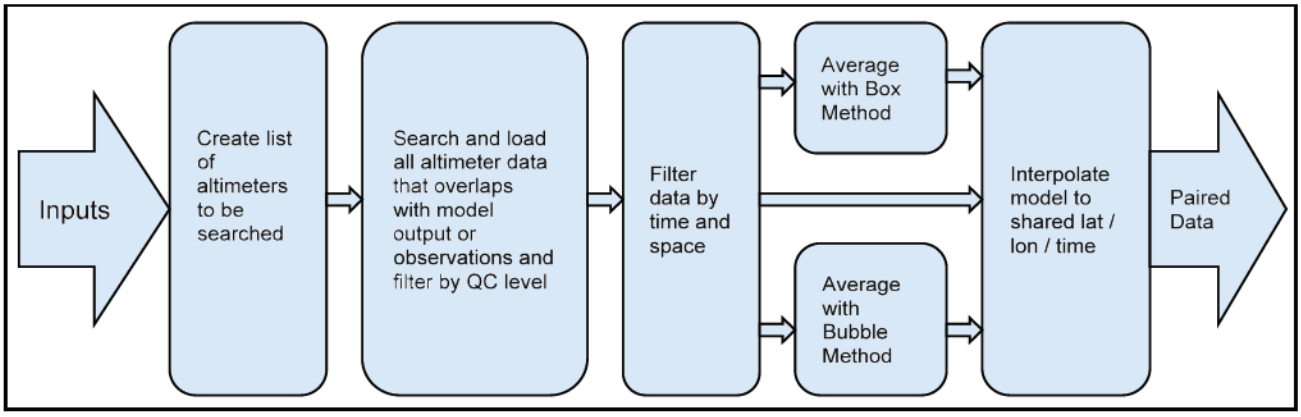


Figure 1. Flow chart for altWIZ processing of altimeter data.

Step 2 is handled somewhat differently depending on the altimeter database. Since the RY19 altimeter data are binned spatially, it is determined if the model grid is active within in the 1°×1° altimeter file. If so, the file is loaded; if not, a loop continues to check the rest of the spatially binned files (and the remaining altimeter missions). For ESA data, all the files within the day-directories that overlap with the model are loaded.

The last step in the flow chart is interpolation. The model could be interpolated to the native altimeter output sampling frequency, 1 Hz, or it could be interpolated to altimeter data that have been averaged in space and time. There are two methods for averaging altimeter data: the box and the bubble methods. For either method, a minimum number of observations can be set, and the mean and standard deviation wave heights are given. Usually, there is a single altimeter pass within the time sampling (~30-minute intervals for a buoy or ~1-hour interval for the model output time), but sometimes there are multiple passes. On the whole, the averaging can be thought of as an along-track average, which produces a wave height measurement that is smoothed over space. A spatial average is closer to the spatial resolution of a model. For buoy data, one may only want the nearest measurement of each pass instead of an average, so both are given in a version of the code that is adapted for pairing with point measurements.

One method for averaging is referred to as the box method. The box method uses the model grid and output time-step from the model. A model gives wave height at grid points, and these grid points become the center of boxes in this 2D + time space. Altimeter data that lie in a particular box are binned together, averaged, and paired with that grid point. A disadvantage of this method is that for regular grid bins, the size of the box changes as a function of latitude. Another is that the distances considered for binning are not equal in every direction. Furthermore, depending on the grid resolution, some of the data in the cube may be further out than what one might want to allow (e.g., 50 or 100 km as a maximum distance). The main advantage is that it is much faster than the second averaging method, the bubble method.

For the bubble method, one decides a priori the maximum time difference and spatial distance allowed of altimeter data from each grid point in space-time. Instead of a box, this can be visualized as a spheroid around each grid point in which all the altimeter data are averaged. Currently, this method takes considerably longer for model data, but it is mitigated somewhat by increasing the number of subdomains — the number of subdomains is set as an option in the code. In addition, if the grid resolution is finer than the size of the bubble, there is a risk of overlapping the bubbles and oversampling the altimeter data. Generally, one would want to keep the bubble smaller than the model resolution to take advantage of the precise ability to decide which data are to be included in the comparison.

In Appendix IV, results from the two averaging methods and no averaging are compared. It is found, generally, that it matters little whether or not one averages (or what method one chooses). Since this is the case, the main advantage of averaging is to reduce the paired dataset to something closer to the model resolution. Otherwise, it may be quicker to forgo averaging.

Exploring Options. It has been implied that there are a few available "options." Option choices affect the resulting altimeter data that are retrieved. The full list of options is specified in a MATLAB structure as follows:

```
options.altDatabase = 'RY19' | 'ESA'
```

This options.altDatabase specifies the desired database and triggers switches in the code. By default, the calibrated Hs is taken for RY19, and the denoised Hs is taken for ESA.

```
options.loopSize = x
```

This options.loopSize allows one to subset the model domain to reduce memory constraints. The number of subdomains calculates out as the square of this loopSize number, so if x = 2, the domain will be quartered; if x = 5, the domain will be split into 25 subdomains.

```
options.averagingMethod = 'box' | 'bubble' | 'none'
```

This specifies the averaging method as covered above.

```
options.minNumberObs = x
```

This number is the minimum number of observations that will be considered for an average; otherwise, the observations are ignored and the iteration skipped.

```
options.maxTimeDiff = x [days]
```

This specifies the maximum time difference between altimeter data and the model grid point that it is to be paired with. Logically, this should be half of the model output time. Therefore, if the model output is hourly, then this should be set to 30 minutes. This option is used to reduce altimeter data to within the ranges of the model but also as a limiter in the bubble method.

```
options.maxDistance = x [km]
```

If using bubble method, one must additionally specify the maximum allowable spatial difference between altimeter data and the model grid point that it is to be paired with. Again, logically, this should be half the model resolution or less.

```
options.QC = 1 | 2
```

This specifies the desired level of quality control.

```
Options.save = 0 | 1
```

This is a logical variable that indicates whether or not the results of the pairing should be saved to a .mat file when the process completes. This option structure is an input into the main pairing function. The only other inputs are paths designating the locations of the altimeter database, the processing code, the model data files, and, if results are to be saved, the save file.

The resulting model-altimeter comparison will be sensitive to options chosen, primarily on the database, but to some degree the averaging method and quality control (QC) level. An exhaustive test of the sensitivity of these results to different options is beyond the scope of this work, but some tests are shown in the appendices.

WIS 2017 ATLANTIC BASIN EVALUATION: WIS is a USACE-sponsored effort to hindcast wave conditions, with the central application of defining the wave climate near USACE projects. WIS began in the 1970s with a focus on the Great Lakes region and has since expanded to all coastal regions in the United States and territories. Currently, WIS is split into basin scale regions, where wave models, grids, and settings are tailored to a particular wave climate. For each region, there are multiple grids: usually a lower-resolution basin scale grid and multiple, high-resolution regional nested grids. In the following example, the WIS hindcast for the year 2017 in the North Atlantic Ocean is presented, in particular the basin grid and a higher-resolution grid that focuses on the East Coast.

Atlantic 2017 Basin Level 1. The basin level 1 grid is a 1/2° grid from -84 - 20° E and 0 - 72° N. The entire year of 2017 was run with Oceanweather Inc. winds at 1/2° resolution to force the WAVEWATCH III model using ST4 source terms with default parameters and tunings.

The altWIZ options were set to use RY19 database, QC level 1, and averaged in time and space with the box method. Figure 2 shows the overall scatter plot, including the density of pairs, and a q-q plot.

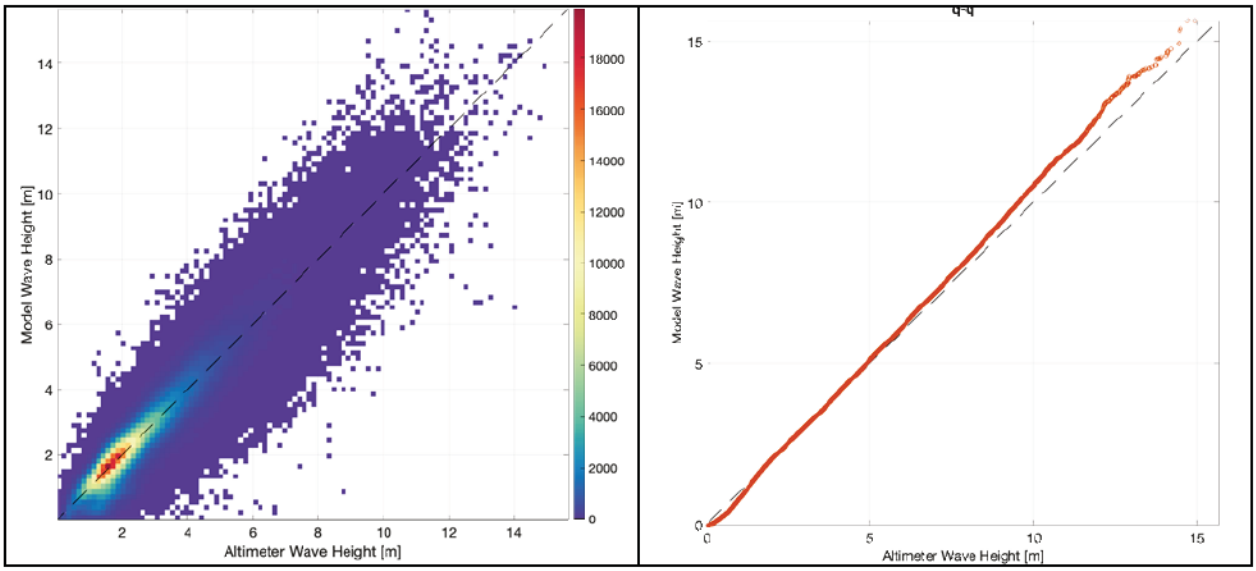


Figure 2. Altimeter observations on the x-axis and model results on the y-axis. Left: scatter plot with density of points shown in color; and right: q-q plot.

Table 1 gives the error statistics for each grid level. The comparison shows good overall agreement with a correlation coefficient, $R^2 = 0.93$, bias = 0.02 m, and root-mean-square-error, rmse = 0.49 m. In the q-q plot, the wave model starts to diverge from the altimeter observations, giving higher wave heights, approximately Hs = 6 m, and this divergence increases as a function of Hs until approximately 12 m.

Table 1. Error statistics of correlation coefficient, bias, and root-mean-square-error for various grid levels.

	Basin Level 1	Basin Level 1 (15 - 60 N)	East Coast Level 2 (RY19, QC 2, no averaging)	East Coast Level 2 (ESA, QC1, no averaging)
R^2	0.93	0.95	0.93	0.94
Bias [m]	0.02	-0.07	-0.07	0.02
rmse [m]	0.49	0.43	0.43	0.39

In Figure 3, one can see how errors are distributed spatially by binning the values and taking averages of these metrics. Compared to the altimeter data, the model is biased low around most model boundaries (except the eastern Gulf of Mexico). Correspondingly, there are higher values of rmse in these same areas. This is because there were no active boundaries in the model. Indeed, by only considering pairs between 15° and 60° N, the comparison is improved with $R^2 = 0.95$, bias = -0.07 m, and rmse = 0.43 m. The model is biased high around the Caribbean and in between 45° and 60° N, where the rmse is relatively low with isolated areas of high rmse to the north of the northern coast of South America and in the eastern Caribbean. Some of the increased rmse is presumably due, in part, to the presence of islands and other coastal features that lead to artifacts in the altimeter waveforms and errors in estimating wave height. There is low bias and rmse in the

southern portion of the North Atlantic, south of approximately 30° N. There is a swath of increased rmse that starts as a thin band off the coast of North Carolina and expands towards the east. The shape and location of this area of increased rmse corresponds neatly with the location and extent of Gulf Stream extension. The Gulf Stream and its extension is well known for strong currents and energetic eddies. The interaction of waves with the current introduces variation in wave height through refraction. Since there is no current input into WIS hindcasts, a lack of wave height variation in the model would result in higher rmse compared to altimeter.

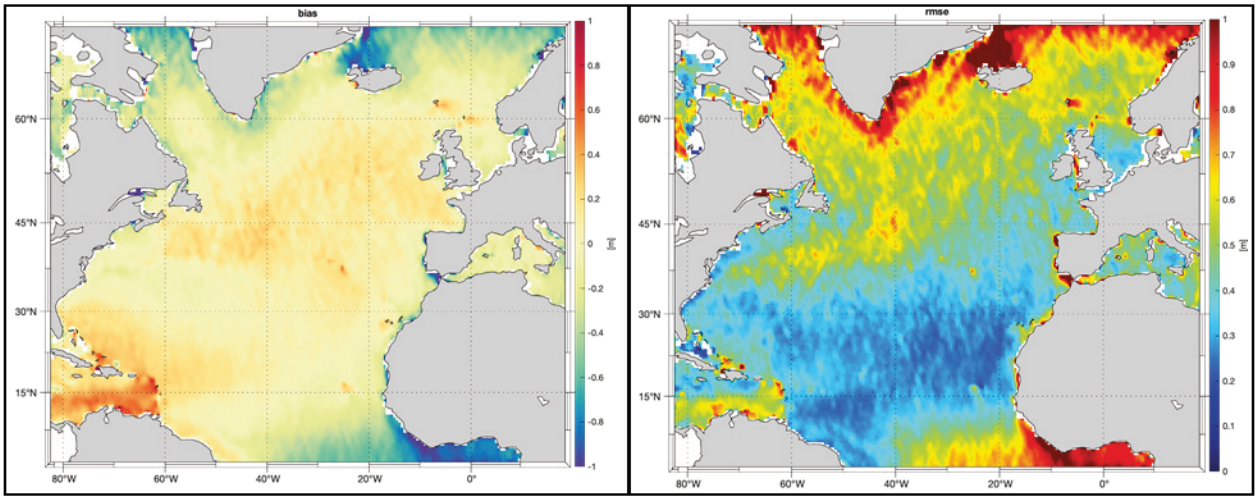


Figure 3. Bias (left) and rmse (right) with color indicating value.

In Figures 4 and 5, data were collated into the four northern seasons, where spring includes February, March, and April, etc. Generally, the patterns are similar to the year average, but they are strengthened in spring and winter and weakened in summer. This is presumably related to the overall increase of energy in the northern hemisphere winter months.

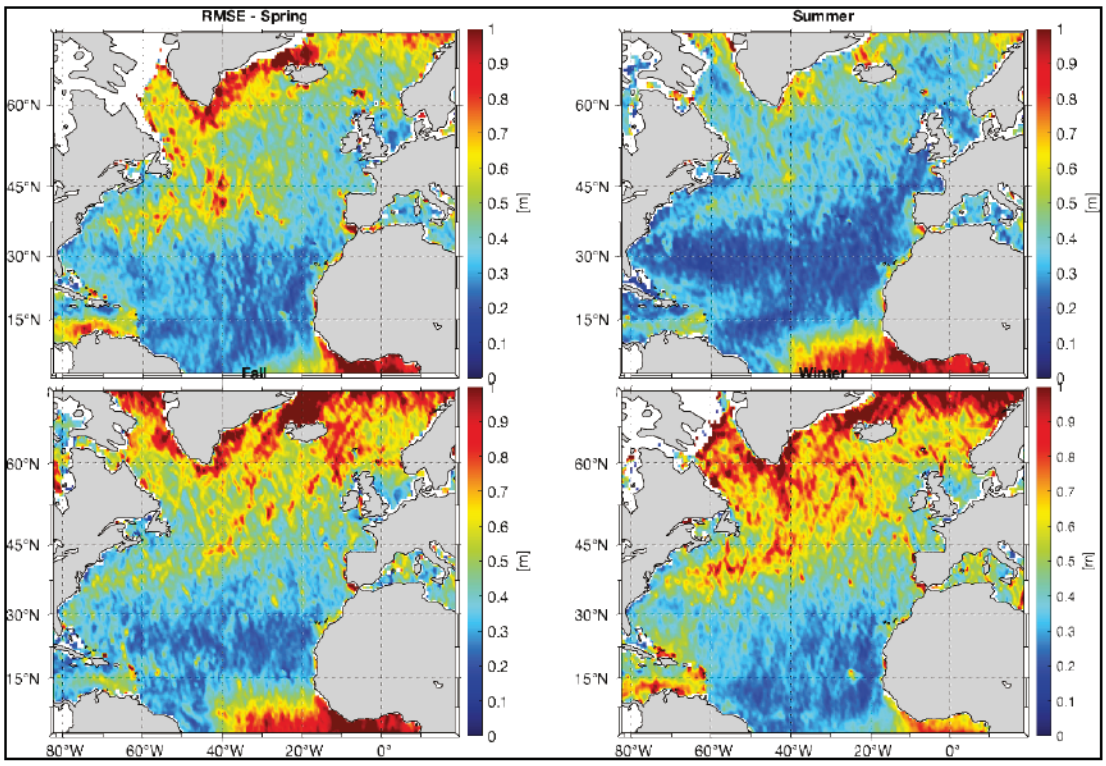


Figure 4. Rmse binned in time clockwise from top left: spring, summer, winter, and fall.

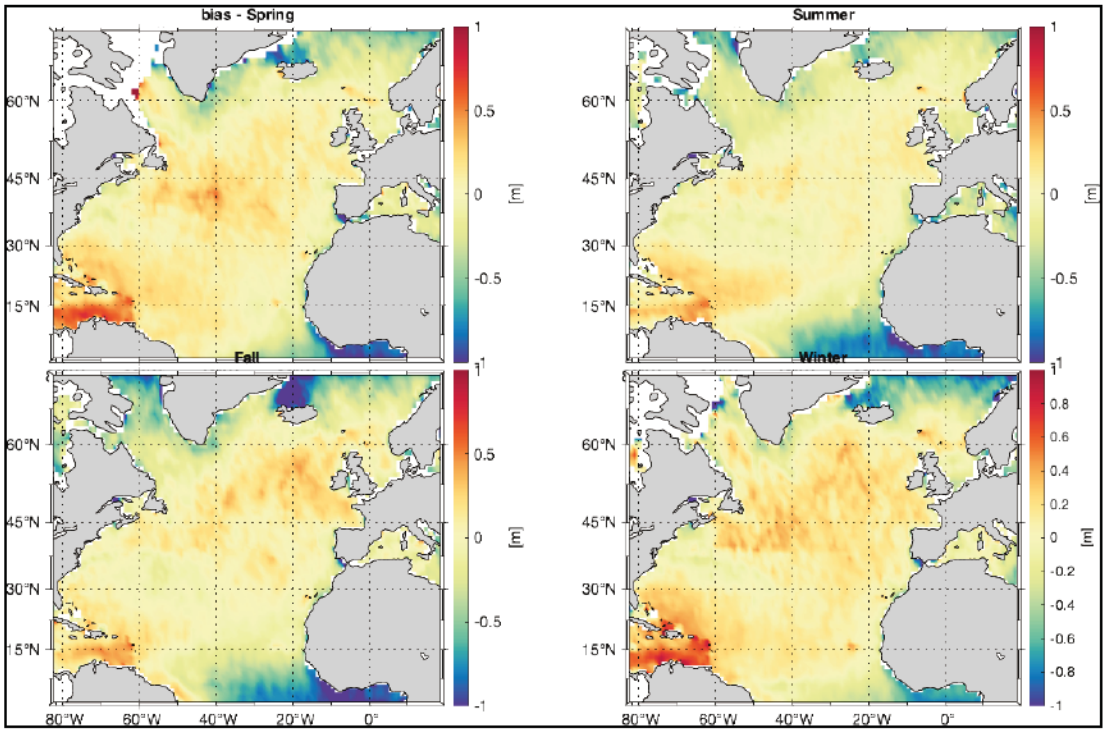


Figure 5. As in Figure 4 for bias.

Atlantic 2017 East Coast Level 2. For the East Coast Level 2 grid, which has a 1/4° resolution, the RY19 data were used, QC level 2 (to include coastal data), and no spatial averaging before interpolation. For this grid, the boundaries are active and come from the basin level 1 grid.

The overall statistics are similar to the basin level restricted domain. The q-q plot in Figure 6 shows that the wave height remains close to a 1:1 relationship. The map of bias in Figure 7 shows values less than 0.5 m over most of the domain. The southern portion of the domain, east of Florida and the Caribbean, and the northeastern corner of the domain trend high whereas the central portion of the domain trends low. Again, the rmse values are increased in the area of the Gulf Stream extension (Figure 7).

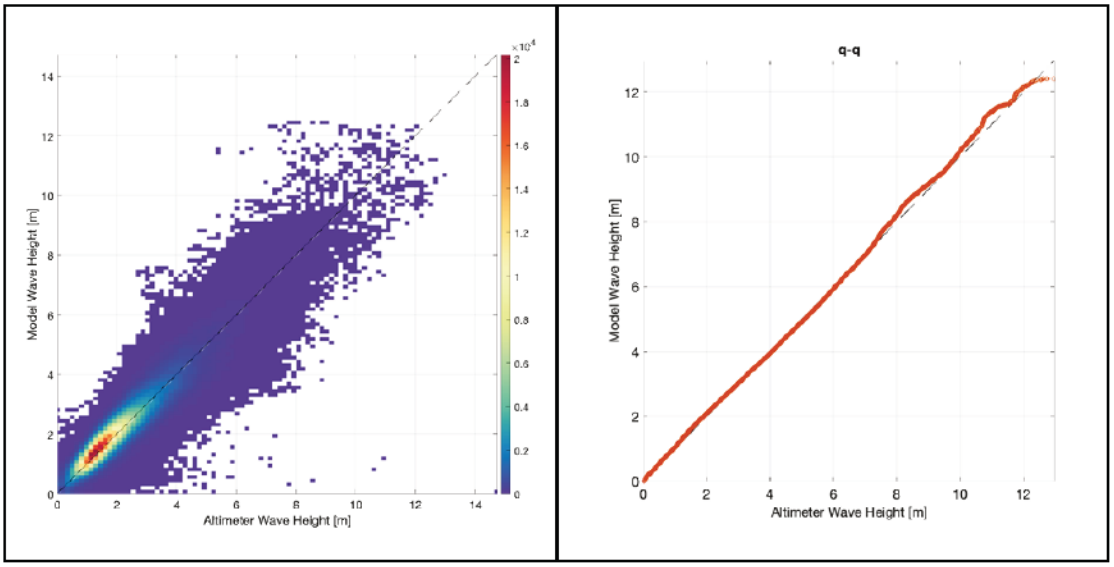


Figure 6. As in Figure 2 for East Coast level 2 grid.

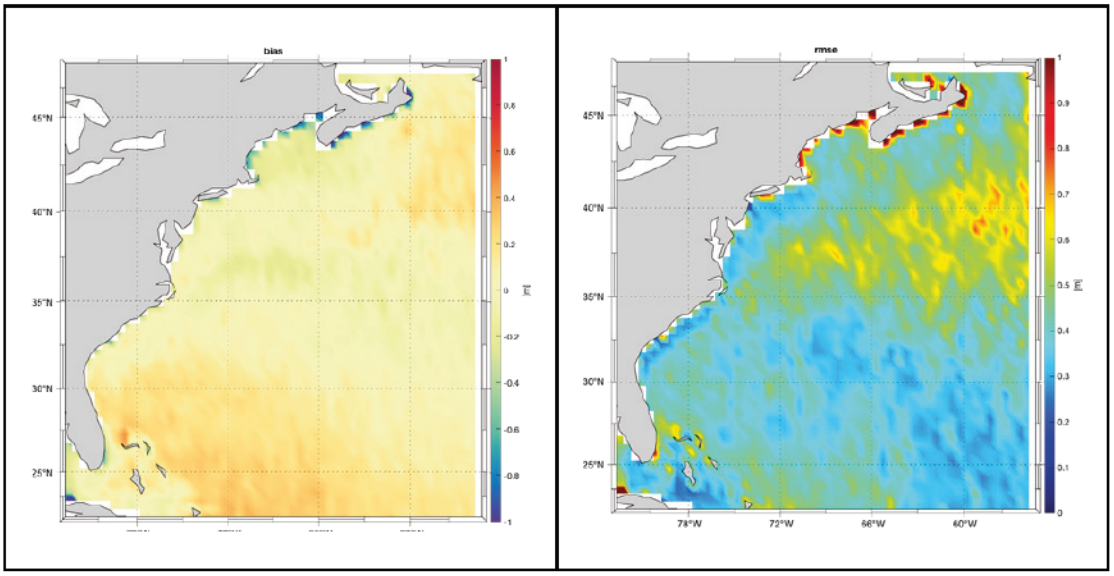


Figure 7. As in Figure 3 or East Coast level 2 grid.

Seasonal binning shows much the same spatial patterns but strengthened (weakened) in the more energetic spring / winter (summer). Data were binned monthly and into two 10° latitude-bands, as shown in the statistics in Figure 8. The correlations are generally above 0.85, except in summer between 25° and 35° N. This could possibly be related to an increase in proportion of swell coming from the southern hemisphere (missing in the model), although bias in this time period and location are nearly negligible. Further investigation is needed to confirm or reject this hypothesis.

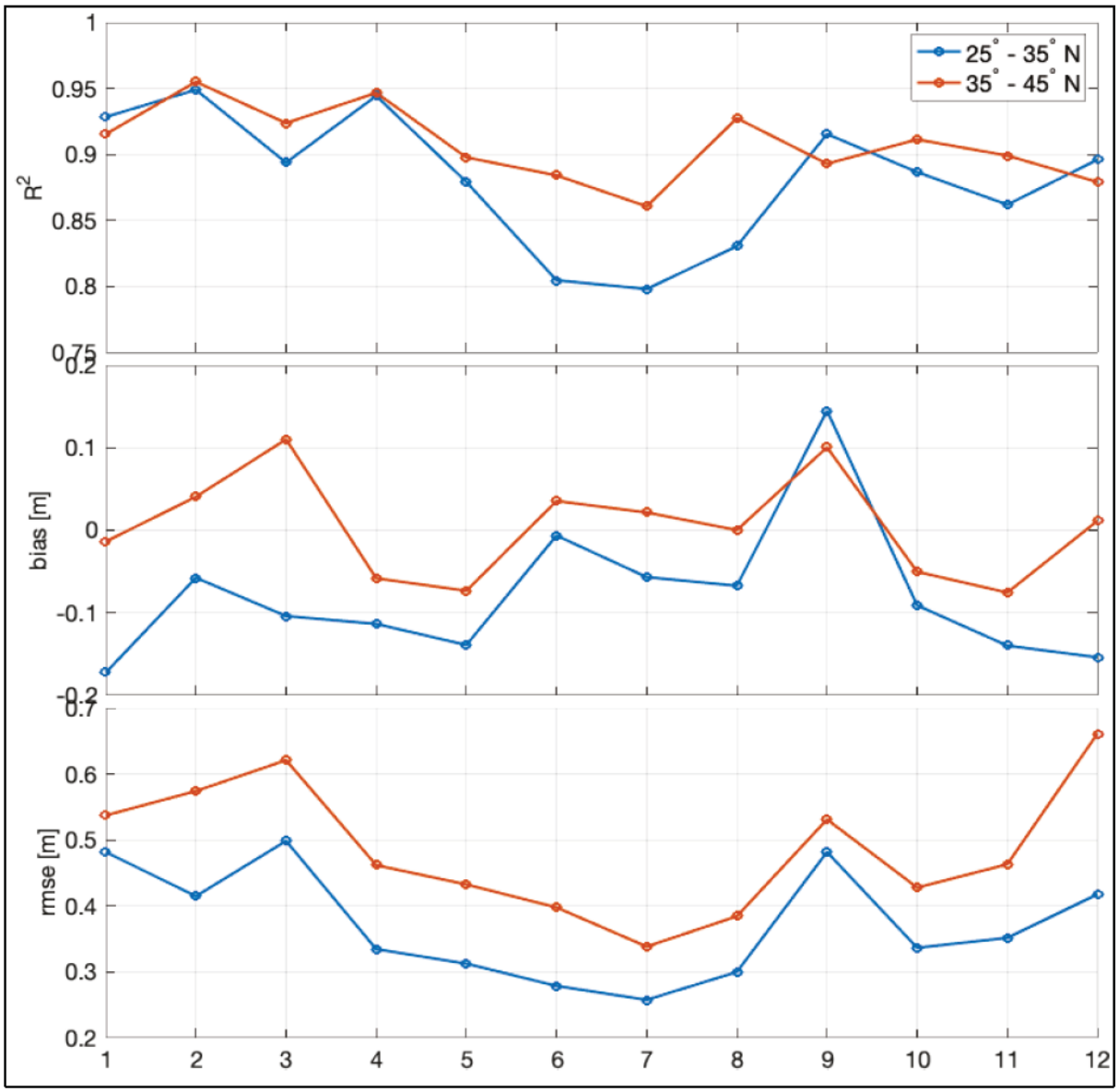


Figure 8. Statistics for comparison made between 25° and 35° N are shown in blue and between 35° and 45° N in red over 12 months. Top to bottom is R^2, bias, and rmse.

The bias is consistently negative between 25° and 35° N, except for September when it is more than 0.1 m positive. Between 35° and 45°, the bias alternates between negative and positive. Rmse patterns are similar for both latitude regions but consistently ~0.1 m higher in the higher latitude band, which contains the Gulf Stream extension.

CONCLUSION: Presented and demonstrated here is a system for comparing WIS model results to altimeter data, called altWIZ. altWIZ allows for a variety of options and incorporates two different, curated, and calibrated altimeter datasets: (1) Ribal and Young (2019) and (2) ESA Sea State v1.1 level 2 (Dodet et al. 2020). The altWIZ capabilities were demonstrated using the WIS 2017 Atlantic hindcast, using both a lower-resolution basin grid and a higher-resolution East Coast regional grid. The overall comparison was favorable but revealed consistent spatial patterns of errors. These patterns strengthened and weakened seasonally. This was related to the seasonal variability of the wave climate, with higher wave heights in the winter and lower wave heights in the summer. There were significant errors at the boundaries because WIS lacks global boundary conditions, but these tend to lessen towards the U.S. coastlines where WIS results are applied. Future work will implement the altWIZ into the WIS workflow to produce comparisons for each WIS basin and grid across the extent of the altimeter time series.

ADDITIONAL INFORMATION: This CHETN was prepared as part of the USACE Coastal and Ocean Data System (CODS) program by Dr. Clarence O. Collins III and Dr. Tyler Hesser, US Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Coastal Analysis and Observations Branch, Duck, NC. Questions pertaining to this CHETN may be directed to Dr. Clarence O. Collins III (*Clarence.O.Collins@usace.army.mil*) or to the USACE CODS Program Manager, Dr. A. Spicer Bak (*Spicer.Bak@usace.army.mil*).

Sincere appreciation is expressed to Matt Malej for his review and to Candice Hall for her review and detailed editing.

This ERDC/CHL CHETN should be cited as follows:

> Collins, Clarence O., III, and Tyler Hesser. 2021. *altWIZ: A System for Satellite Radar Altimeter Evaluation of Modeled Wave Heights*. ERDC/CHL CHETN-IV-127. Vicksburg, MS: U.S. Army Engineer Research and Development Center. *http://dx.doi.org/10.21079/11681/39699*

REFERENCES

Cavaleri, L., J.-H. G. M. Alves, F. Ardhuin, A. Babanin, M. Banner, K. Belibassakis, M. Benoit, M. Donelan, J. Groeneweg, T. H. C. Herbers, P. Hwang, P. A. E. M. Janssen, T. Janssen, I. V. Lavrenov, R. Magne, J. Monbaliu, M. Onorato, V. Polnikov, D. Resio, W. E. Rogers, A. Sheremet, J. M. Smith, H. L. Tolman, G. van Vledder, J. Wolf, and I. Young. 2007. "Wave Modelling – The State of the Art." *Progress in Oceanography* 75(4): 603–674.

Collins III, C. O. 2012. *In Situ Wave Measurements: Sensor Comparison and Data Analysis*. Open Access Theses. University of Miami.

Collins III, C. O., B. Lund, R. J. Ramos, W. M. Drennan, and H. C. Graber. 2014. "Wave Measurement Intercomparison and Platform Evaluation During the ITOP (2010) Experiment." *Journal of Atmospheric and Oceanic Technology* 31(10): 2309–2329.

Dodet, G., J. F. Piolle, Y. Quilfen, S. Abdalla, M. Accensi, F. Ardhuin, E. Ash, J. R. Bidlot, C. Gommenginger, G. Marechal, and M. Passaro. 2020. "The Sea State CCI Dataset v1: Towards a Sea State Climate Data Record Based on Satellite Observations." *Earth Syst. Sci. Data* 12(3): 1–28.

Krogstad, H. E., J. Wolf, S. P. Thompson, and L. R. Wyatt. 1999. "Methods for Intercomparison of Wave Measurements." *Coastal Engineering* 37(3–4): 235–257.

Quilfen, Y., and B. Chapron. 2019. "Ocean Surface Wave-Current Signatures from Satellite Altimeter Measurements." *Geophysical Research Letters* 46(1): 253–261.

Quilfen, Y., and B. Chapron. 2020. "On Denoising Satellite Altimeter Measurements for High-Resolution Geophysical Signal Analysis." *Advances in Space Research*. https://doi.org/10.1016/j.asr.2020.01.005

Ribal, A., and I. R. Young. 2019. "33 Years of Globally Calibrated Wave Height and Wind Speed Data Based on Altimeter Observations." *Scientific Data* 6(1): 1–15.

Thomson, J., E. A. D'Asaro, M. F. Cronin, W. E. Rogers, R. R. Harcourt, and A. Shcherbina. 2013. "Waves and the Equilibrium Range at Ocean Weather Station P." *Journal of Geophysical Research: Oceans* 118(11): 5951–5962.

APPENDIX I: CODE ACCESS

The code is managed through a github repository at *https://github.com/Tripphysicist/altWIZ.git*.

APPENDIX II: FUNCTION LIST

`altimeterModelPairing` - main function that pair altimeter data with model data
- `defineSatListESA` - define a list of ESA satellites active for the input time
- `defineSatListRY19` - define a list of RY19 satellites active for the input time
- `getRY19AltimeterObsModel` - load in altimeter data from RY19 database
- `getESAAltimeterObs` - load in altimeter data from ESA database
- `boxMethod` - bin data using the model grid and make averages
- `bubbleMethod` - bin data using the user defined maximum distance and time difference and make averages
  - `Latlon2dist` - calculate distance between two coordinates using Haversine formula

Two other versions of the main function were developed. One, `altimeterBuoyPairing`, for pairing with a point source measurement, that is mobile or stationary, and `altimeterStormPairing` for pairing data centered around tropical systems. The implementations are similar, but the details of the options are different depending on the application. See the source code for details.

`altimeterBuoyPairing` - beta version of the function that pairs altimeter data with buoy data
- `defineSatListESA` - define a list of ESA satellites active for the input time
- `defineSatListRY19` - define a list of RY19 satellites active for the input time
- `getRY19AltimeterObsBuoy` - load in altimeter data from RY19 database
- `getESAAltimeterObs` - load in altimeter data from ESA database
  - `bubbleMethodBuoy` - bin data using the user defined maximum distance and time difference and make averages
    - `Latlon2dist`- calculate distance between two coordinates using Haversine formula

`altimeterStormPairing` - beta version of the function that pairs altimeter data with tropical storm tracks from Atlantic HURDAT2
- `getStormTrack.m` - get track information for a storm based on year and name
- `defineSatListESA` - define a list of ESA satellites active for the input time
- `defineSatListRY19` - define a list of RY19 satellites active for the input time
- `getRY19AltimeterObsBuoy` - load in altimeter data from RY19 database
- `getESAAltimeterObs` - load in altimeter data from ESA database

There are also some plotting scripts in the repository for visualizing the paired data, but these require additional dependencies and are not covered here.

APPENDIX III: USAGE EXAMPLE

The code is written in MATLAB. An example is given here, but please refer to the internal code documentation of the latest code version from Github. This could be run as a script or entered into the command line. The main function is `altimeterModelPairing`.

```
%This is code to pair altimeter data from Ribal and Young (2019) to
%WIS model results

%specify path to code
codePath = '/Users/tripp/D/Analysis/altimeterComparison/altVmodel';

%specify path to model files
mdPath = '/Users/tripp/D/Analysis/altimeterComparison/modelData/AtlanticYearRun/';

%specify path to altimeter files
altPath = '/Users/tripp/D/Datasets/Satellite/Altimeter/Ribal_Young_2019/';

%specify path to save file
savePath = '/Users/tripp/D/Analysis/altimeterComparison/data/atlantic2016May-2017Nov';

%specify options

 options.altDatabase = 'RY19';
 options.averagingMethod = 'box';
 options.maxTimeDiff = 30/(24*60);
 options.maxDistance = 25;
 options.minNumberObs = 7;
 options.loopSize = 5;
 options.QC = 1;
 options.save = 1;

%run the pairing function
[pData] = altimeterModelPairing(codePath,mdPath,altPath,savePath,options);
```

APPENDIX IV: SOME EXAMPLES COMPARING OPTIONS

Comparing Averaging Methods. The box and bubble averaging methods consider slightly different geometries in 2D + time space. For the box method, the averaging space is tied to the underlying model grid resolution. For the bubble method, the distances are set by the user. In both methods, the model grid points are central to the averaging-space. Therefore, as the size of the two averaging-spaces become similar, the observation sets should also become similar (i.e., share most members [besides a few points due to geometric differences]). Thus, the statistical comparison between the two averaging methods and the model should come close to converging. To demonstrate their similarity, and to show how the resulting statistics compares to no-averaging, both methods were compared using a relatively small amount of data: nine days of model data, from October 1, 2010, 00:00:00 UTC to October 9, 2010, 07:00:00 UTC, from a grid area in the Pacific Ocean, 18-24°N by 212-236° W, with a 1/2° grid resolution. Figure A-1 shows the significant wave heights from altimeter and the model, as interpolated to the altimeter. The second and third rows show that the averaging methods discretize the data in space and time (to the native model grid).

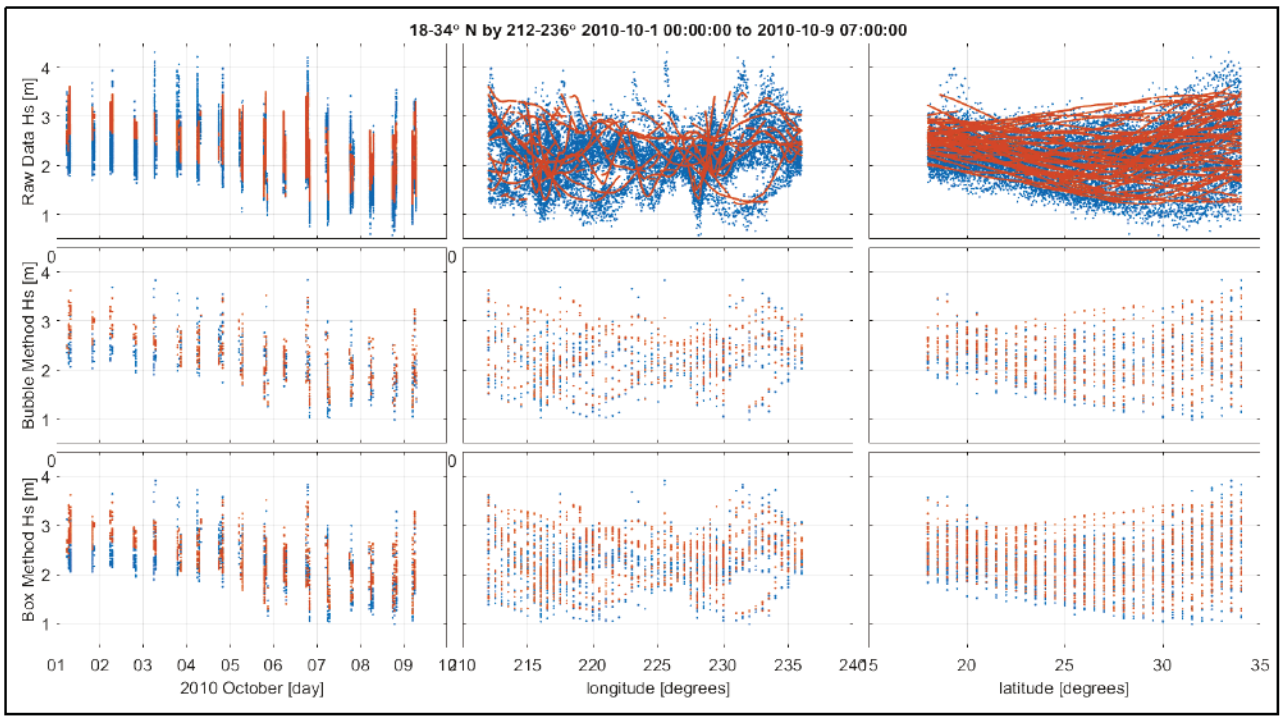


Figure A-1. Significant wave height from a wave model (red dots) and satellite altimeter (blue dots).

The first column is as a function of time, the second column is as a function of longitude, and the third column is as a function of latitude. The first row is raw altimeter data and interpolated model data. The second row is the bubble method of averaging altimeter data, and the third row is the box method of averaging altimeter data. Statistics for each row are in Figure A-2 and Table A-1.

Table A-1. Statistics from paired data.

Statistic/Method	No Averaging	Bubble Method	Box Method
Number of Pairs	12660	606	1168
bias (m)	0.15	0.15	0.14
rmse (m)	0.38	0.35	0.35
R^2	0.76	0.80	0.79
SI	0.15	0.15	0.14

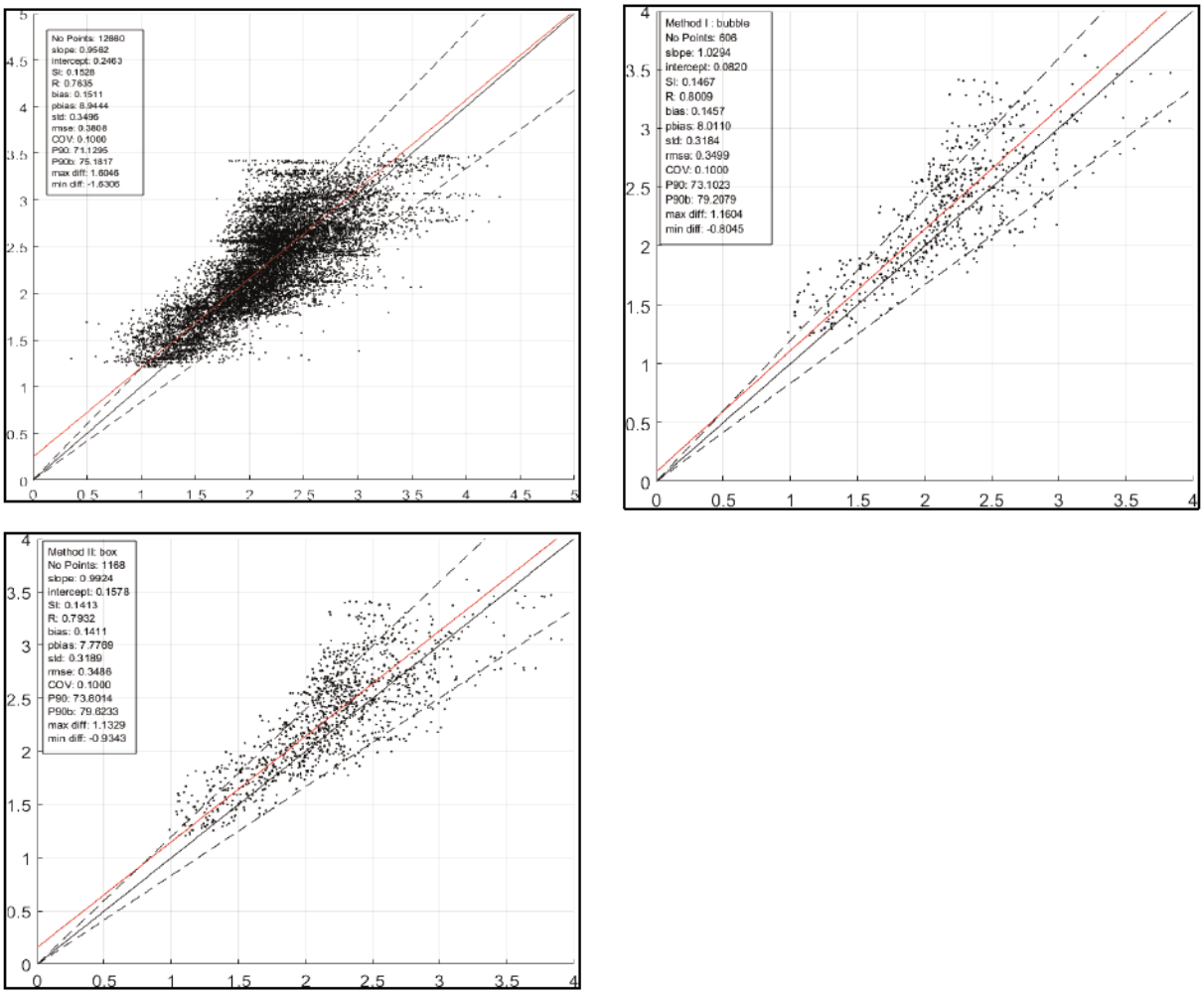


Figure A-2. Scatter plot and statistics altimeter model comparison, clockwise from bottom left, box method, no averaging, and bubble method.

To take this a step further, a similar analysis is performed on a much larger dataset: 1 month of data, February 2017, covering the Atlantic Ocean 83.5° W to 20° E and 0° N to 72° N, with a 1/2° grid resolution. Table A-2 and A-3 show bias and rmse as a function of space averaged into 1° bins. Figure A-3 shows the corresponding scatter plots.

Table A-2. Bias [m] and rmse [m] maps for box, bubble, and no averaging.

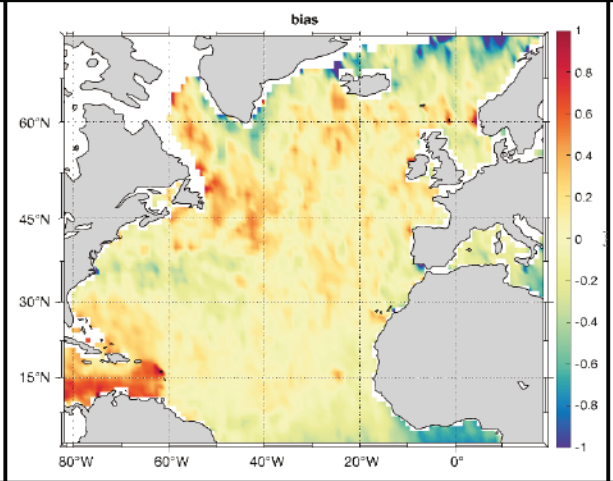
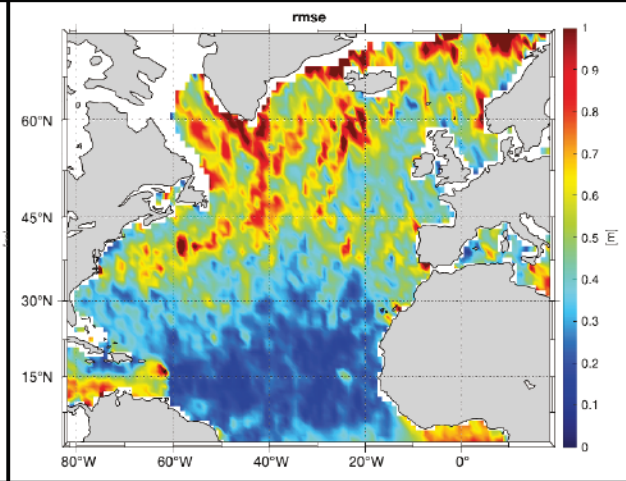
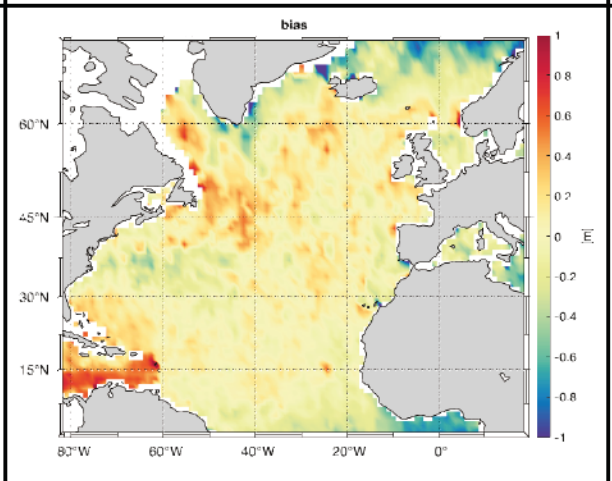
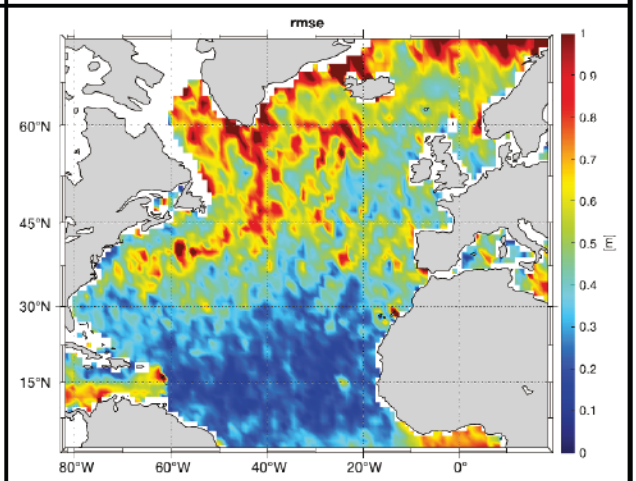
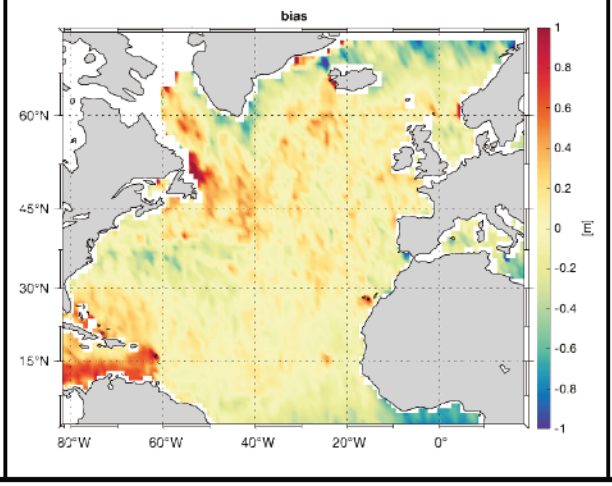
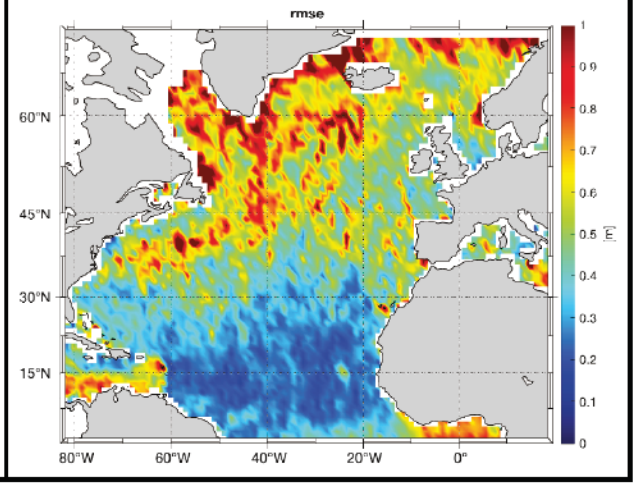
Method/ Metric	bias [m]	rmse [m]
Box		
Bubble		
None		

Table A-3. Statistics from paired data.			
	Box	Bubble	Raw QC1
Number of pairs	70,307	66,105	852,116
R^2	0.96	0.95	0.95
Bias [m]	-0.01	0.02	0.01
Rmse [m]	0.46	0.54	0.52

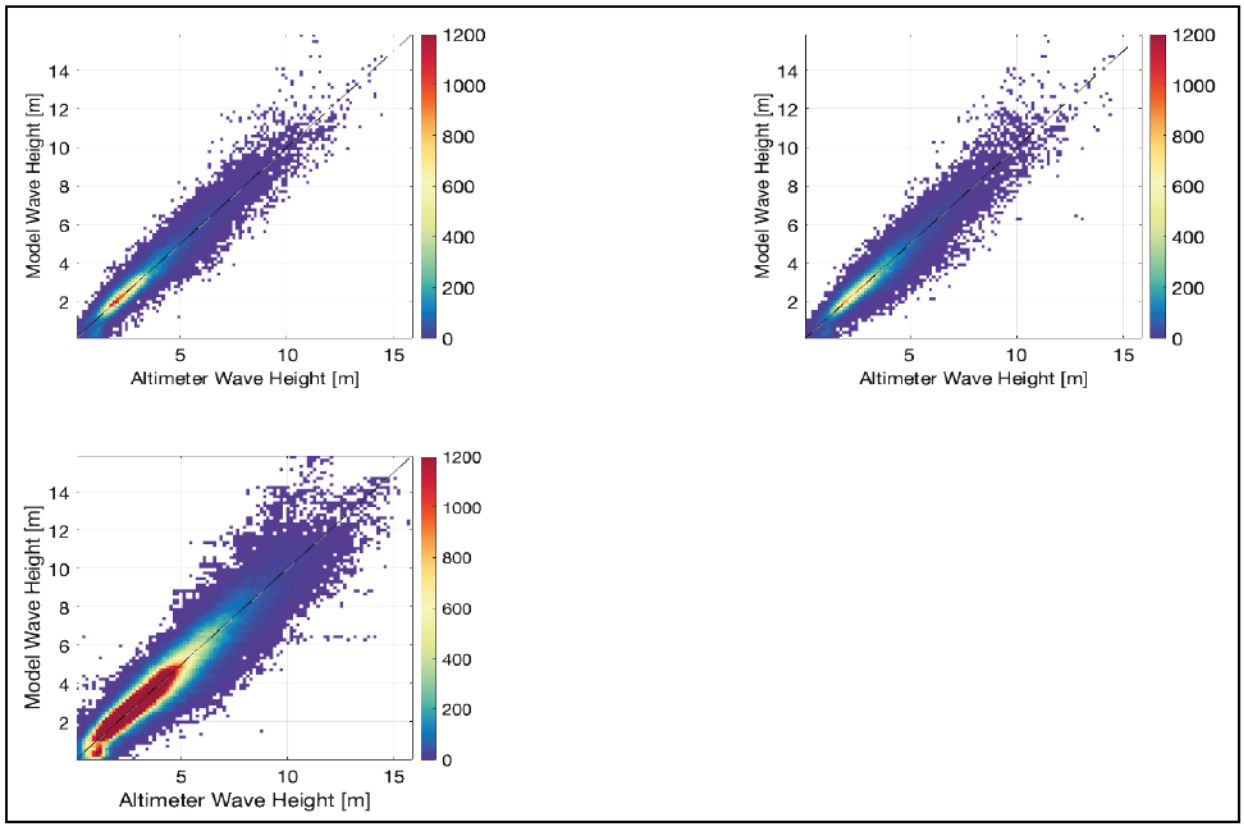


Figure A-3. Scatter plots for various averaging methods, clockwise from bottom left: no averaging, bubble method, and box method.

From calculating statistics and plotting the maps of bias and rmse, one would draw similar conclusions regardless of if and how the data were averaged. At least qualitatively, and quantitatively in a spatially averaged sense, these methods can be used interchangeably. The details do vary if one looks at a spatial map of the difference between metrics, bias or rmse, for example (not shown).

Ribal and Young (2019) vs. ESA Sea State. A similar analysis that swaps the RY19 data for ESA data is presented here. This is not a direct comparison of the two databases but inferences are made from the differences between making model comparisons with each. The ESA data use a denoised Hs value created using empirical mode decomposition as a noise filter (Quefflen and

Chapron 2019, 2020). Through this process, they were able to reveal signals of persistent wave current interaction (e.g., in the vicinity of the Gulf Stream extension and the Agulhas current). It is expected that the denoised data may reduce the apparent rmse of the model. The difference in the spatially binned bias and rmse are shown on maps in Figure A-4. There is an offset (i.e., a bias), where the bias for RY19 is generally higher than ESA, on average by 0.09 m, but does not otherwise vary greatly with a standard deviation of 0.07 m. There is no clear spatial relationship or relationship to the magnitude of the bias. This is likely just reflecting the different calibrations or adjustments used between the two datasets. The difference between rmse is also minor; on average, the RY19 rmse is higher by 0.04 m, and the standard deviation between the two is 0.08 m. In contrast to bias, there is a definite spatial pattern of the rmse differences. ESA shows decreased rmse south of 30°N and along the coast, but there are patches of increased rmse between 30° and 45° N. The reason for this is unclear but certainly related to the ESA denoised product.

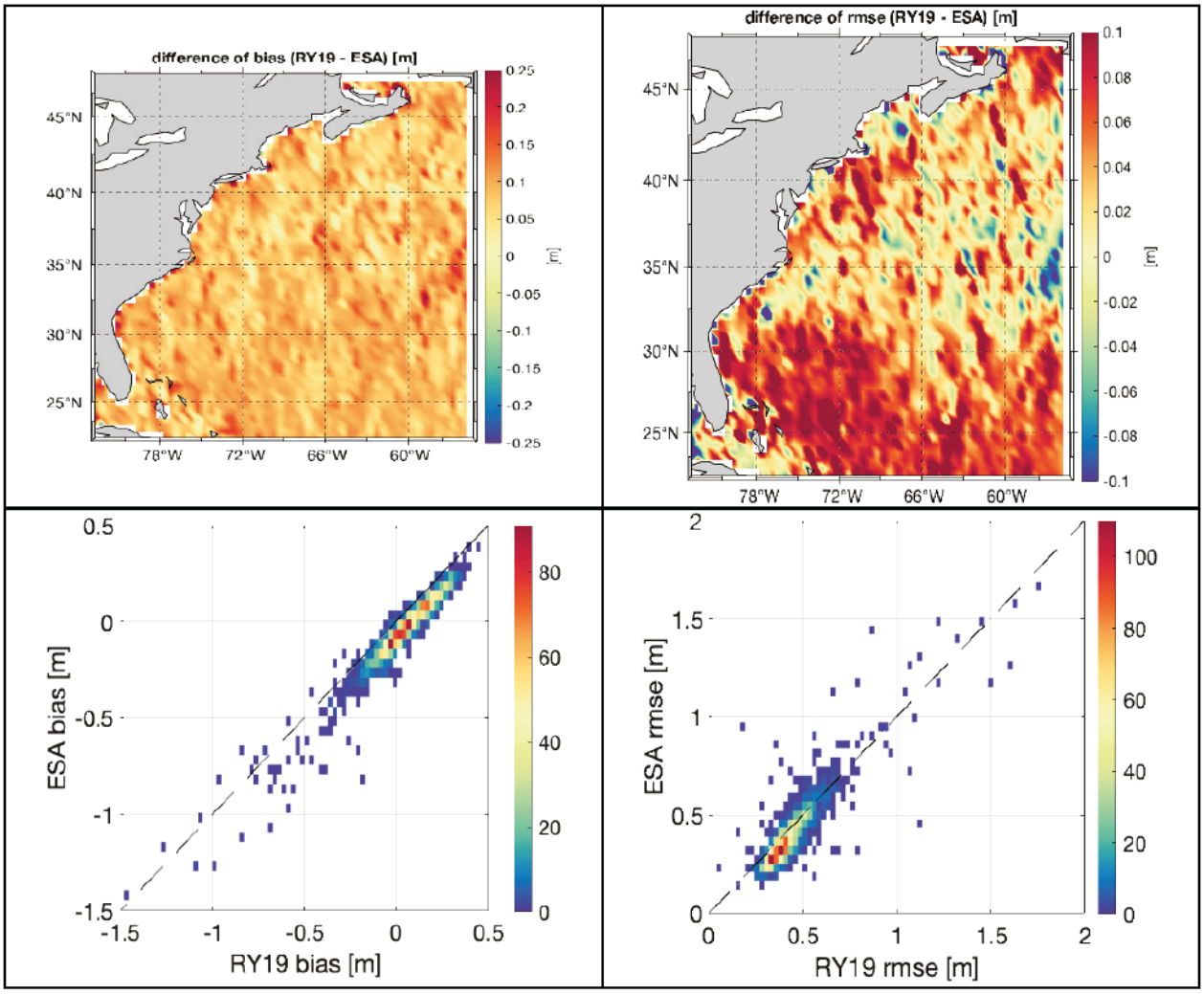


Figure A-4. Comparison of bias and rmse between the model and RY19 and the model and ESA. The top left is the difference in bias, values shown in color. The scatter plot of bias is shown directly below. Same for rmse on the right side.

Sensitivity to Distance when Pairing with Buoys. Although the focus of this CHETN is on comparing satellite data with model results, a module has been built to compare the satellite data against buoy data, either stationary or drifting. The bubble method is the logical choice for averaging data around a buoy location. Some sensitivity tests were run for the maximum distance parameter that can be seen in Table A-4. To test this, averaged Hs and a nearest single value Hs data from Ocean Station Papa (e.g., Thomson et al. 2013) are examined against both the RY19 and the ESA data. Of course, the numbers here do not tell the entire story of the differences between the two datasets.

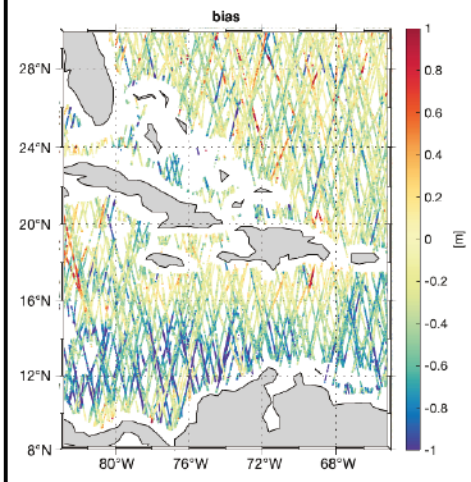
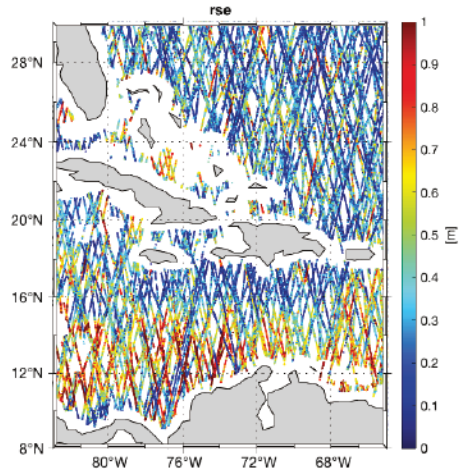
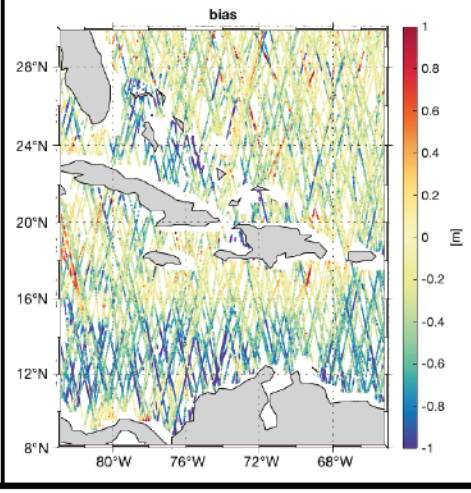
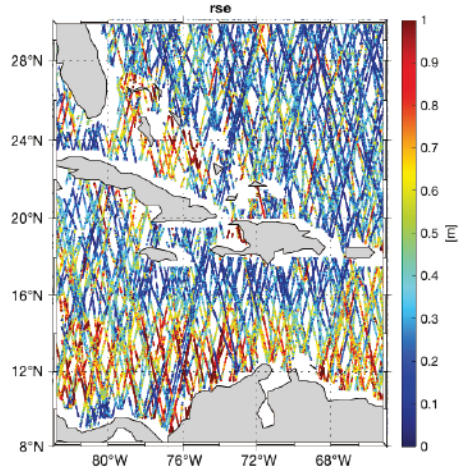
Table A-4. Statistics for paired data between altimeter datasets, RY19 and ESA, with Ocean Station Papa. The values in parentheses are calculated with the nearest observation instead of an average.

Dataset	RY19	ESA	RY19	ESA	RY19	ESA
distance [km]	% of buoy data paired	% of buoy data paired	bias [m]	bias [m]	rmse [m]	rmse [m]
25	0.38	0.31	0.11 (0.10)	0.04 (0.04)	0.23 (0.27)	0.18 (0.18)
50	0.70	0.61	0.11 (0.12)	0.05 (0.05)	0.24 (0.28)	0.21 (0.22)
100	1.37	1.28	0.12 (0.12)	0.06 (0.06)	0.30 (0.33)	0.28 (0.29)
200	1.83	2.74	0.13 (0.12)	0.06 (0.06)	0.37 (0.37)	0.37 (0.37)

There is excellent agreement if the altimeter data are confined to a radius of 25 km, but this severely reduces the number of points compared (0.38% of all buoy data for RY19 and 0.31% for ESA). The results are similar between RY19 and ESA and similar to what was observed with the model comparison with ESA, but here the bias is significantly reduced for ESA. Choosing the nearest value over the mean value increases the rmse slightly, but otherwise the numbers are similar. The bias is relatively stable for each distance tested, but the rmse increases steadily with distance. The optimal distance seems to be approximately 100 km, which matches up with expectations from Collins et al. (2014). This test uses just over 7 years of buoy data, but if one is interested in specific events (e.g., storms), the distance criteria may need to be relaxed to get data points for the time of interest. This demonstrates the implicit trade-off one makes when including further-afield data.

QC and Approaching the Coasts with RY19. Generally, it is advised to use altimeter data that pass the strictest QC level; in the code, this option would be level 1. However, there are some applications where one may want to alter this, near the coast, for example. As the altimeter footprint approaches the coast, the return echo of the radar pulse is affected, and eventually dominated, by signals unrelated to the sea surface, making the data unreliable near the coast. In the RY19 dataset, any data within 50 km of the coast are flagged with QC code 2, "probably good data," vs. QC code 1, "good data." The difference in QC-1 and QC-2 data is demonstrated in Table A-5, which show maps of bias and root-squared-error (rse) around the Caribbean.

Table A-5. Bias and rse for RY19 QC flag 1 and QC flag 2.

	bias [m]	rse [m]
QC flag 1		
QC flag 2		

These data should be examined on a case-by-case basis to see if it is of acceptable quality and to determine if including these data in the comparison is worth the potential increase of data uncertainty. In some cases, it will be important to have this extra data near the coast (e.g., Hurricane Matthew whose track ran through the Caribbean).

APPENDIX V: METRIC MAP RESOLUTION

This is one month of comparison for February 2017. The model grid had a 1/2° resolution. Table A-6 (pages 23 – 25) shows how different levels of spatial averaging result in different error maps.

Table A-6. Maps or bias and rmse for various levels of spatial averaging.

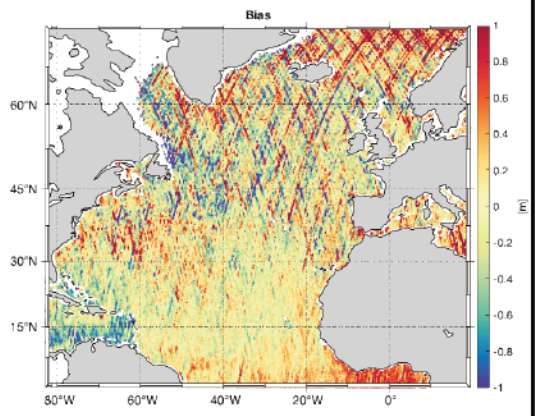
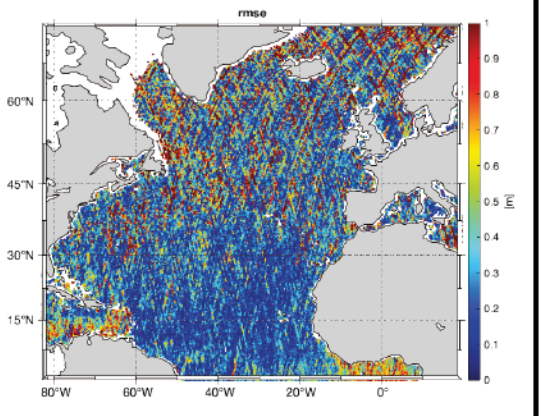
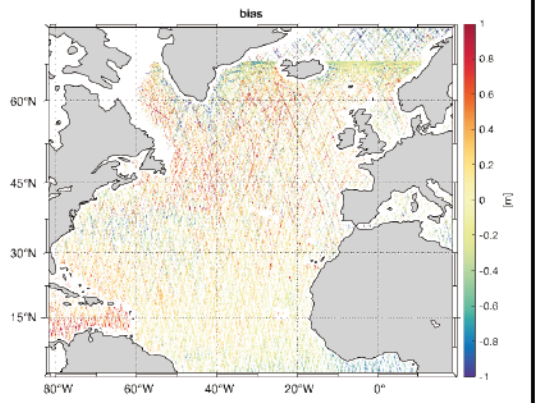
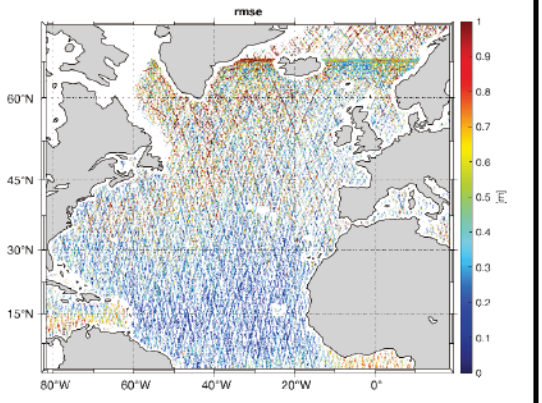
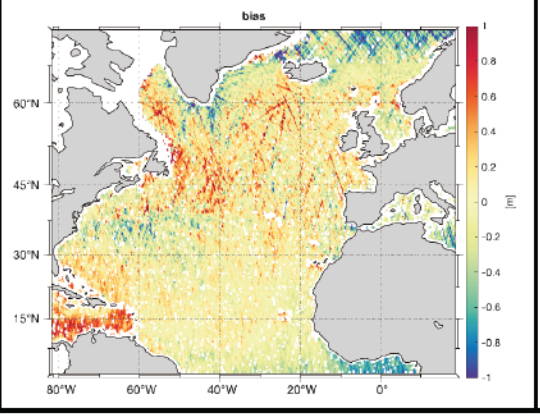
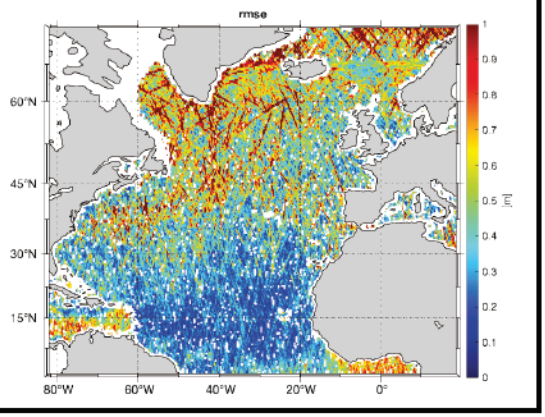
Spatial averaging	bias [m]	rms or rmse [m]
No spatial averaging		
1/10° bins		
1/4° bins		

Table A-6. Maps or bias and rmse for various levels of spatial averaging.

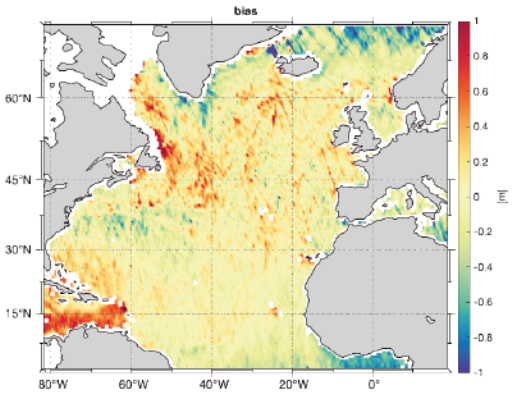
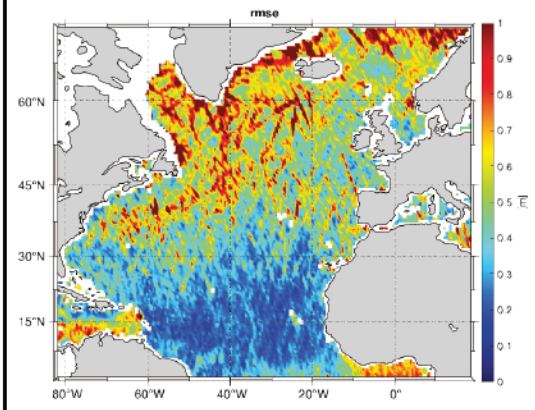
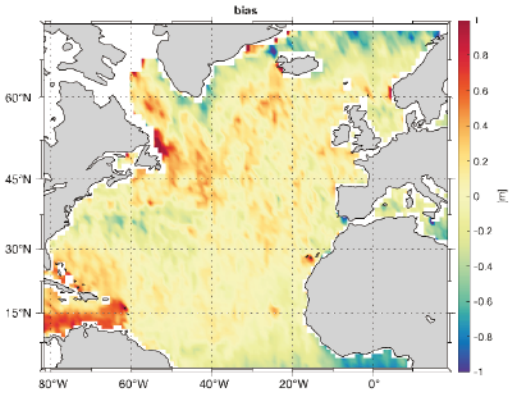
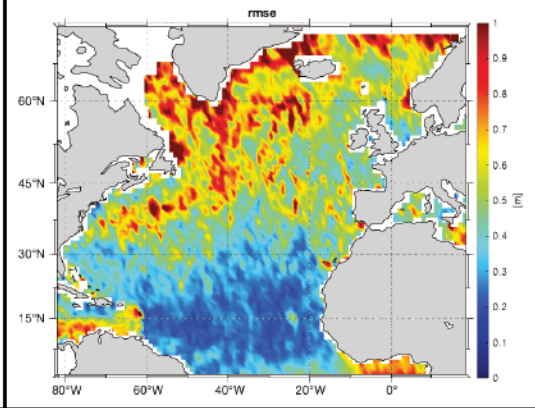
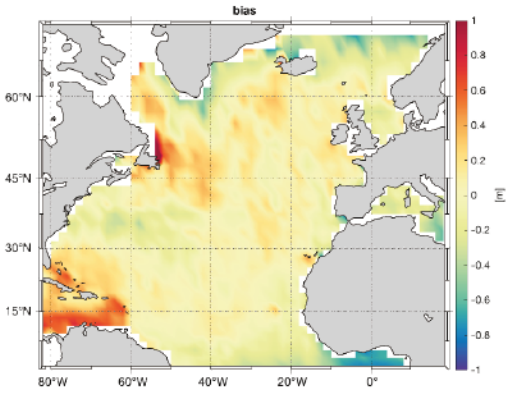
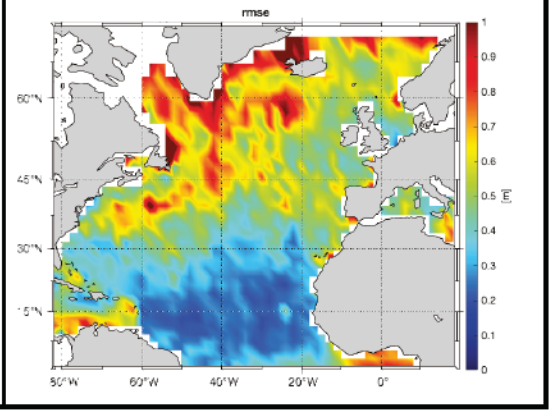
Spatial averaging	bias [m]	rms or rmse [m]
1/2° bins		
1° bins		
2° bins		

Table A-6. Maps or bias and rmse for various levels of spatial averaging.

Spatial averaging	bias [m]	rms or rmse [m]
5° bins