Using SAR to estimate significant wave heights in the New Jersey coastal area

David Bazzett[1], Behzad Golparvar[1], and Roger Wang[1]

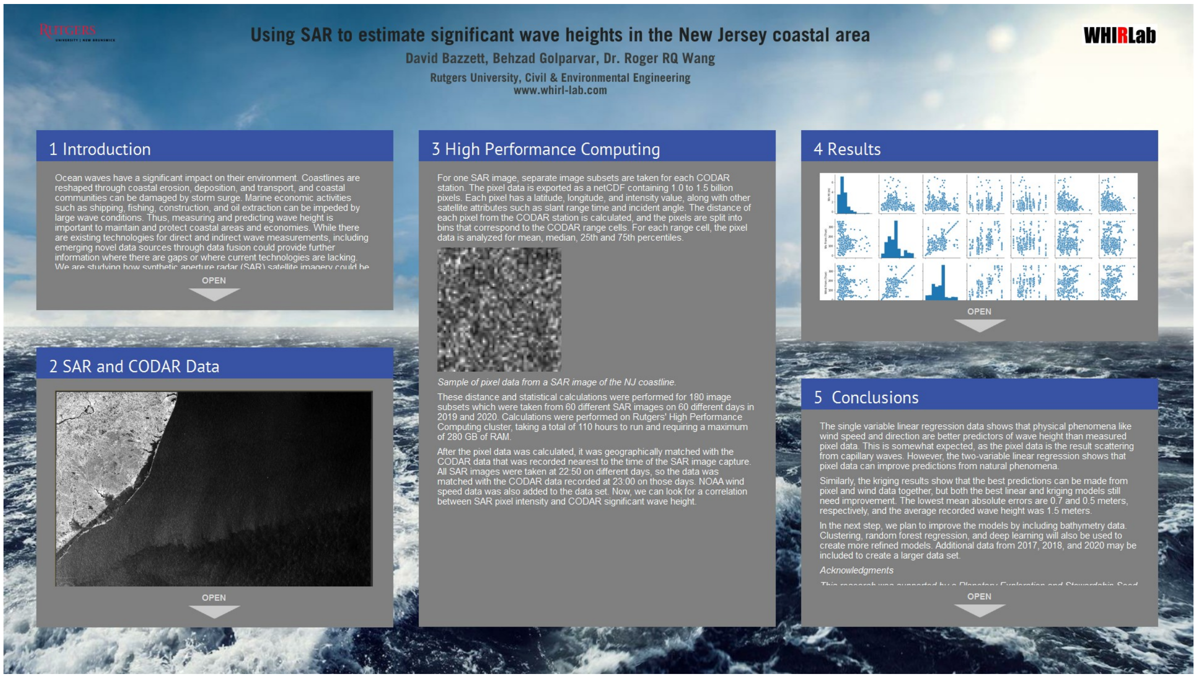
[1]Rutgers University

November 24, 2022

Abstract

Ocean waves interact with the environment in many ways. They transport energy and mass, and the resultant sea-surface roughness defines the drag coefficients that transmits wind energy to the ocean (Drennan et al., 2003). Through erosion and deposition, waves change the shape and landscape of coastal areas. Storm surge waves can cause flood damage in coastal areas. Recent studies revealed that wetlands are sensitive to the wave condition, which determines the retreat or growth of coastal ecosystems (Green and Coco, 2007; Mariotti and Fagherazzi 2010). Human activities rely on the condition of waves to conduct marine activities such as fishing, shipping, oil extraction, and offshore constructions. Thus, it is important to understand ocean waves to improve earth system modeling, protect the coastline, predict storm surge, preserve coastal ecosystems, and enhance the offshore business. This project will explore the application of synthetic aperture radar (SAR) imagery to predict significant wave height near the coast. High-frequency (HF) radar data of the ocean (aka CODAR) was used as ground truth data set to calibrate and validate the wave height estimator. Off-shore wind data was also included. The developed code will enhance the current capability to process the satellite data and create a new platform to monitor the coastal environment. The collected data will help further our understanding of the wave spectrum in a coastal environment and the data can support other research in the related topics, e.g. the interaction of waves and ice sheets, wetlands, shorelines, wind farm and aquaculture.

Using SAR to estimate significant wave heights in the New Jersey coastal area



David Bazzett, Behzad Golparvar, Dr. Roger RQ Wang

Rutgers University, Civil & Environmental Engineering
www.whirl-lab.com

PRESENTED AT:

EarthCube
2020 EarthCube Annual Meeting
Virtual – June 18, 2020

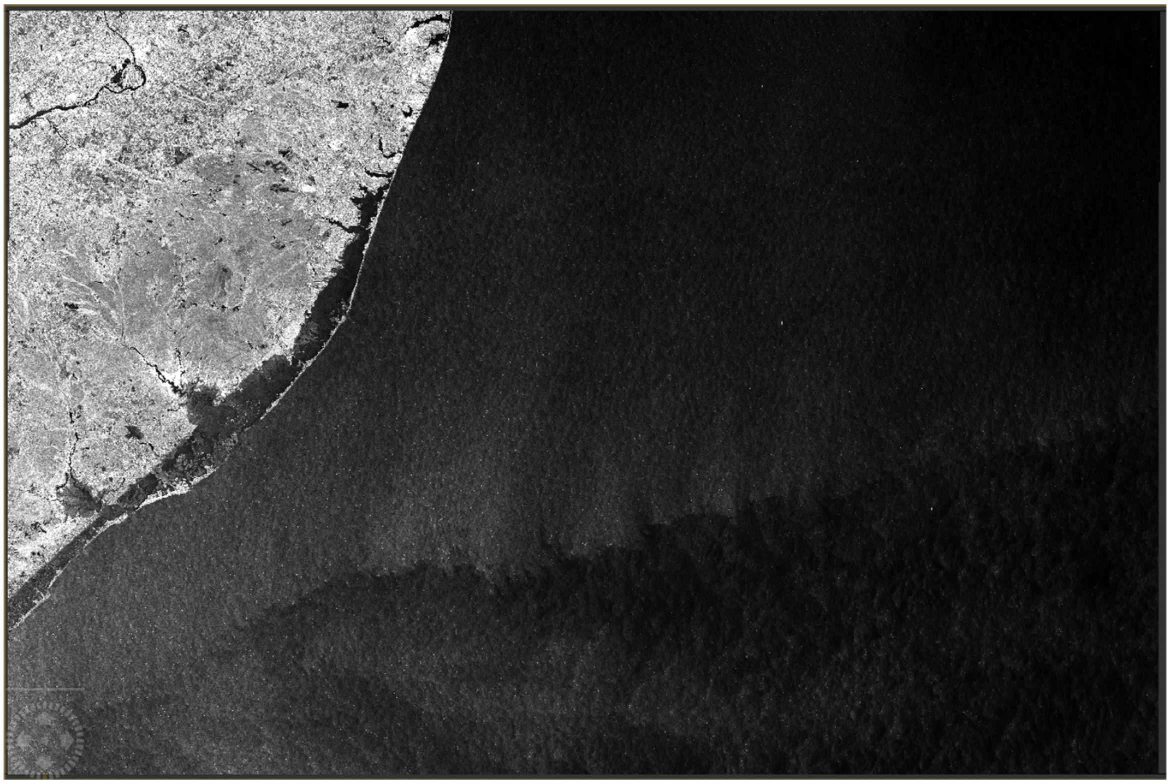
1 INTRODUCTION

Ocean waves have a significant impact on their environment. Coastlines are reshaped through coastal erosion, deposition, and transport, and coastal communities can be damaged by storm surge. Marine economic activities such as shipping, fishing, construction, and oil extraction can be impeded by large wave conditions. Thus, measuring and predicting wave height is important to maintain and protect coastal areas and economies. While there are existing technologies for direct and indirect wave measurements, including emerging novel data sources through data fusion could provide further information where there are gaps or where current technologies are lacking. We are studying how synthetic aperture radar (SAR) satellite imagery could be used for measuring waves.

In the last several years, researchers have shown that satellite data from the European Space Agency's (ESA) Sentinel 1A and 1B (S1A and S1B) satellites can be used to estimate significant wave height (Hs) and mean wave period through empirical relationships (Pleskachevsky, et al. 2019, Stopa and Mouche 2017). Pleskachevsky (2019) compared S1A Interferometric Wide Swath Mode (IW) images of the ocean to wave data measured from buoys developed using an empirical relationship where significant wave height (Hs) could be estimated from an image of the ocean in near-real-time (NRT). Similarly, Stopa and Mouche (2017) developed a machine learning algorithm that compared S1A Wave Mode (WV) data (not visual image data) to data from buoys and NOAA Wavewatch III (WW3), and thus developed an empirical relationship for estimating significant wave height (Hs) and average wave period. However, the spatial resolution of these methods is very low – the Pleskachevsky (2019) method returns Hs for an area of 2.5 km x 2.5 km, and the Stopa and Mouche (2017) method returns Hs and period for an area of 20 km x 20 km. Both of these methods are for areas in the open ocean, far from shore.

The objective of our research is to develop a similar empirical method for estimating Hs from S1 IW images at a higher resolution (less than 1 km by 1 km) for regions near the coast.

2 SAR AND CODAR DATA



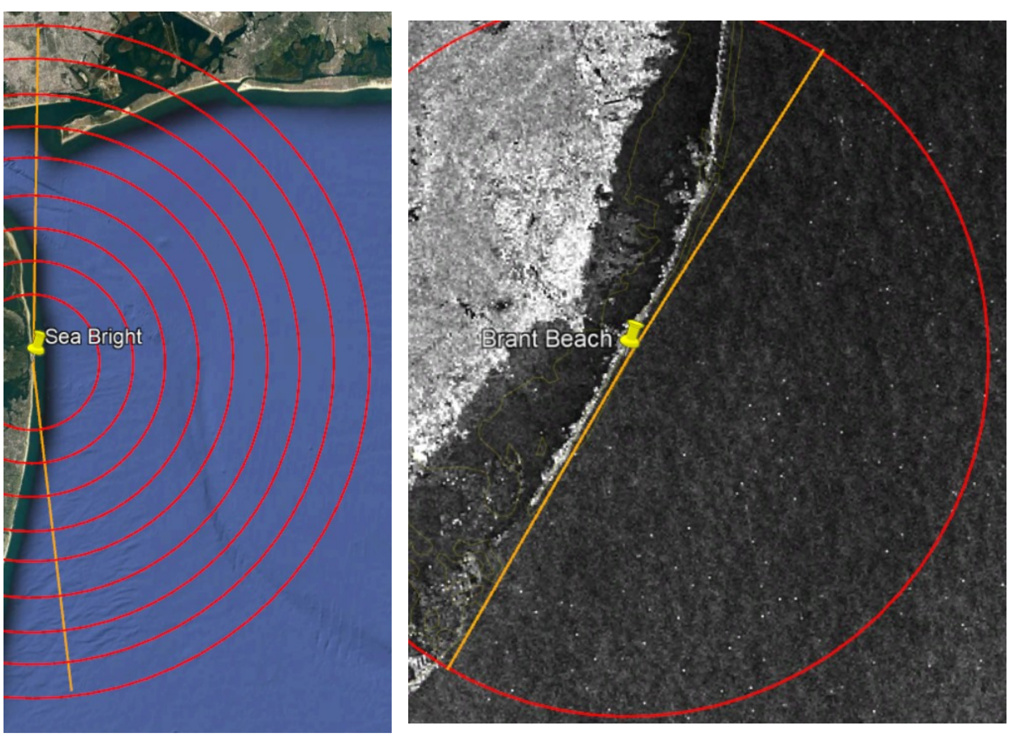
S1A image of the New Jersey coastline and ocean.

Sentinel 1A was launched on April 3, 2014, and Sentinel 1B was launched on April 25, 2016. These are C-band satellites that circle the earth in separate 12 day cycles (e.g. S1A will take an image of a given area once every 12 days). The images we worked with in this study were 260 km by 180 km with 10m pixel resolution. Data from S1A and S1B is stored in separate computer servers maintained by ESA Copernicus Data Center and NASA's Alaska Satellite Facility (ASF). The data is available to the public for downloading through ESA or NASA websites.

Satellite images can be processed by using a program created by ESA called Sentinel Application Platform (SNAP), which has tools for filtering, geometric processing, radiometric processing, and more. Additionally, we used an open-source GitHub project called pyroSAR (Truckenbrodt 2019) to automate the processing of images using python. Image processing generally consists of five steps: 1) thermal noise removal, 2) border noise removal, 3) subset to a specific location, 4) calibration of image data based on standard values, and 5) masking of land data. Processed images can then be analyzed.

Five CODAR stations along NJ Coast and NOAA buoy.

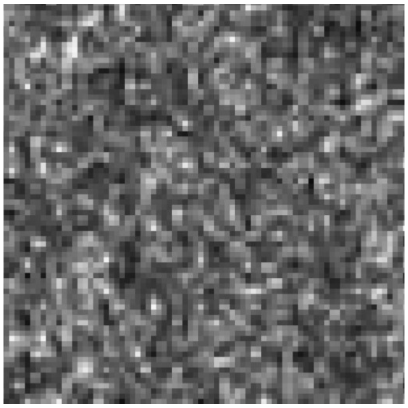
The Rutgers University Department of Marine and Coastal Sciences maintains five high frequency (HF) radar stations (aka CODAR) along the coast that measure ocean wave and wind data. These stations are able to measure data in "range cells" at different distances from the coast. The range cells are sectors that are concentric and spaced in increments of 3 km (e.g. range cell 2 from 4.5 to 7.5 km, range cell 3 from 7.5 to 10.5 km, etc.). This is shown below with the sectors in red circles bounded by the orange radial lines. Wave data is recorded every 10 to 60 minutes, averaged over time and averaged over the area of the sector. The recorded data includes significant wave height, period, direction, and wind direction.



CODAR with range cells (left) and CODAR with SAR overlay (right). The extents of these images are also what *might be captured in an image subset.*

3 HIGH PERFORMANCE COMPUTING

For one SAR image, separate image subsets are taken for each CODAR station. The pixel data is exported as a netCDF containing 1.0 to 1.5 billion pixels. Each pixel has a latitude, longitude, and intensity value, along with other satellite attributes such as slant range time and incident angle. The distance of each pixel from the CODAR station is calculated, and the pixels are split into bins that correspond to the CODAR range cells. For each range cell, the pixel data is analyzed for mean, median, 25th and 75th percentiles.

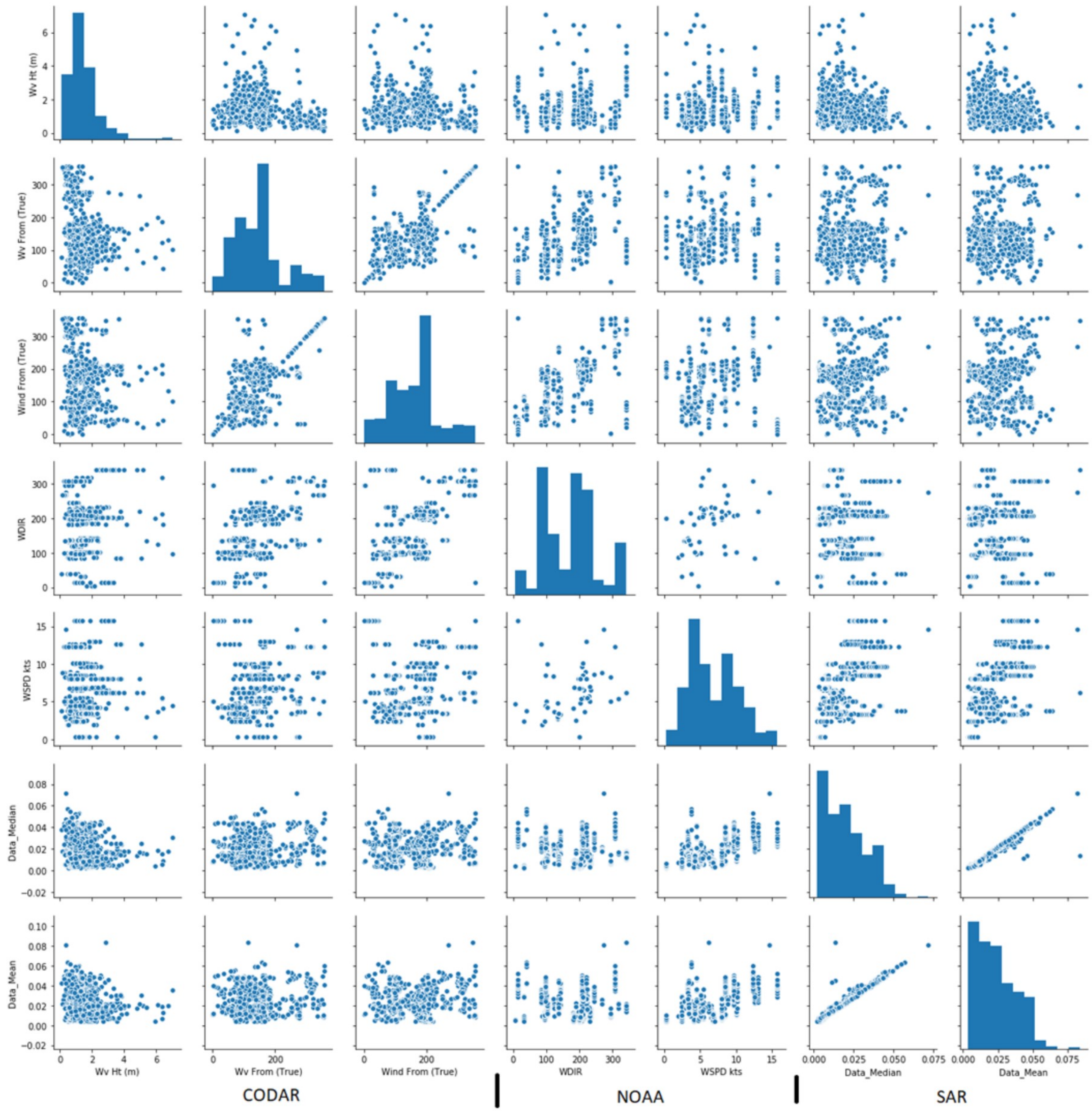


Sample of pixel data from a SAR image of the NJ coastline.

These distance and statistical calculations were performed for 180 image subsets which were taken from 60 different SAR images on 60 different days in 2019 and 2020. Calculations were performed on Rutgers' High Performance Computing cluster, taking a total of 110 hours to run and requiring a maximum of 280 GB of RAM.

After the pixel data was calculated, it was geographically matched with the CODAR data that was recorded nearest to the time of the SAR image capture. All SAR images were taken at 22:50 on different days, so the data was matched with the CODAR data recorded at 23:00 on those days. NOAA wind speed data was also added to the data set. Now, we can look for a correlation between SAR pixel intensity and CODAR significant wave height.

4 RESULTS



Using Python libraries, correlation matrix plots were examined and outliers in the data set were identified and removed, leaving a sample size of 619 points. A series of stepwise linear regressions were performed to identify the most accurate linear model for predicting significant wave height based on SAR pixel data, CODAR wind and wave direction, and NOAA wind speed and direction.

The most accurate linear model to predict the significant wave height had two input variables of NOAA wind speed and NOAA wind direction, with an R-squared of 0.66 and a mean relative error 67%. Models with three inputs were tested, but all had invalid P-values. Models with NOAA wind direction and pixel data performed almost as well, but the error indicates that all of these multivariate linear regression models are not sufficiently accurate for this dataset.

Linear Regression Results Independent variable: Wv Ht (m)								
Dependent variables	1 variable				2 variables, with NOAA Wind DIR			
	R^2 %	MAE [m]	MRE %	RMSE [m]	R^2 %	MAE [m]	MRE %	RMSE [m]
Pixel_StDev	0.01	1.51	98.1	1.78	Invalid P-value			
Pixel_25percentile	0.42	1.00	76.4	1.36	0.64	0.74	67.2	1.07
Pixel_Median	0.45	0.97	75.2	1.32	0.64	0.73	67.2	1.07
Pixel_Mean	0.48	0.94	73.3	1.29	0.65	0.73	66.8	1.06
Pixel_75percentile	0.48	0.95	74.4	1.29	0.65	0.73	67.1	1.06
CODAR Wv DIR	0.51	0.91	78.6	1.25	Invalid P-value			
CODAR Wind DIR	0.55	0.85	73.8	1.20				
NOAA Wind SPD	0.59	0.83	67.6	1.15	**0.66**	**0.72**	**66.7**	**1.03**
NOAA Wind DIR	0.63	0.76	68.2	1.09	Self			

MAE = mean abs. error, MRE = mean relative error, RMSE = root mean square error

As a second test, the data was checked with kriging (Gaussian process regression model), which is a nonparametric kernel-based probabilistic model. The most effective model was created by using three inputs: mean pixel intensity, wind speed, and wind direction. Using this model improved the predictions, with a root mean square error (in meter) of 0.64 and mean relative error of 40%. While the error is reduced, there is still room for improvement.

Kriging Results Independent variable: Wv Ht (m)					
Dependent variables		MRE, T %	MAE, P [m]	MRE, P %	RMSE, P [m]
single input	Pixel_StDev	1.50E-08	0.97	94.2	1.43
	Pixel_25percentile	9.10E-12	0.85	64.1	1.25
	Pixel_Median	2.00E-11	0.82	71.2	1.21
	Pixel_Mean	5.80E-12	0.87	87.5	1.26
	Pixel_75percentile	8.30E-12	0.78	65.1	1.05
	cellAvgDiffFromImg	9.70E-11	0.93	78.9	1.41
	CODAR Wv DIR	1.14E+03	13.72	1242.1	16.04
	CODAR Wind DIR	4.40E+02	5.02	449.8	6.42
	CODAR Wv Ht StdDv	1.68E+03	16.5	1515.8	22.8
	NOAA Wind SPD	4.03E+02	3.93	407.4	5.15
	NOAA Wind DIR	5.31E+02	5.88	540.0	7.84
3 input	Pixel_Mean NOAA Wind SPD NOAA Wind DIR	**5.50E-13**	**0.45**	**40.2**	**0.64**
6 input	Pixel_StDev Pixel_25percentile Pixel_Median Pixel_Mean Pixel_75percentile cellAvgDiffFromImg	3.10E-14	0.53	47.6	0.82

MAE = mean abs. error, MRE = mean relative error, RMSE = root mean square error
T = training, P = prediction

5 CONCLUSIONS

The single variable linear regression data shows that physical phenomena like wind speed and direction are better predictors of wave height than measured pixel data. This is somewhat expected, as the pixel data is the result scattering from capillary waves. However, the two-variable linear regression shows that pixel data can improve predictions from natural phenomena.

Similarly, the kriging results show that the best predictions can be made from pixel and wind data together, but both the best linear and kriging models still need improvement. The lowest mean absolute errors are 0.7 and 0.5 meters, respectively, and the average recorded wave height was 1.5 meters.

In the next step, we plan to improve the models by including bathymetry data. Clustering, random forest regression, and deep learning will also be used to create more refined models. Additional data from 2017, 2018, and 2020 may be included to create a larger data set.

Acknowledgments

This research was supported by a Planetary Exploration and Stewardship Seed Grant from the Rutgers Institute of Earth, Ocean, and Atmospheric Sciences.

The authors acknowledge the Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey for providing access to the Amarel cluster and associated research computing resources that have contributed to the results reported here. URL: https://oarc.rutgers.edu

ABSTRACT

Ocean waves interact with the environment in many ways. They transport energy and mass, and the resultant sea-surface roughness defines the drag coefficients that transmits wind energy to the ocean (Drennan et al., 2003). Through erosion and deposition, waves change the shape and landscape of coastal areas. Storm surge waves can cause flood damage in coastal areas. Recent studies revealed that wetlands are sensitive to the wave condition, which determines the retreat or growth of coastal ecosystems (Green and Coco, 2007; Mariotti and Fagherazzi 2010). Human activities rely on the condition of waves to conduct marine activities such as fishing, shipping, oil extraction, and offshore constructions. Thus, it is important to understand ocean waves to improve earth system modeling, protect the coastline, predict storm surge, preserve coastal ecosystems, and enhance the offshore business.

This project will explore the application of synthetic aperture radar (SAR) imagery to predict significant wave height near the coast. High-frequency (HF) radar data of the ocean (aka CODAR) was used as ground truth data set to calibrate and validate the wave height estimator. Off-shore wind data was also included. The developed code will enhance the current capability to process the satellite data and create a new platform to monitor the coastal environment. The collected data will help further our understanding of the wave spectrum in a coastal environment and the data can support other research in the related topics, e.g. the interaction of waves and ice sheets, wetlands, shorelines, wind farm and aquaculture.

REFERENCES

Andrey Pleskachevsky, Sven Jacobsen, Björn Tings & Egbert Schwarz (2019) Estimation of sea state from Sentinel-1 Synthetic aperture radar imagery for maritime situation awareness, International Journal of Remote Sensing, 40:11, 4104-4142, DOI:10.1080/01431161.2018.1558377

Stopa, J. E., and A. Mouche (2017), Significant wave heights from Sentinel-1 SAR: Validation and applications, J. Geophys. Res. Oceans, 122, 1827–1848, doi:10.1002/2016JC012364.

Truckenbrodt, John and Cremer, F. and Baris, I. and Eberle, J. (2019), pyroSAR: A Framework for Large-Scale SAR Satellite Data Processing, Big Data from Space, 197-200, doi:10.2760/848593