Depth in the Great Barrier Reef: The Comparison of LANDSAT 8 Images and Bathymetric Data

MIDN 1/C Bridget Kennedy

Background:

LANDSAT 8 imaging has been a breakthrough in remote sensing technology (USGS, 2017). Launched in 2013, as a product of collaboration between NASA and the United States Geological Survey, this satellite system consists of the Operational Land Imager(OLI), as well as the thermal Infrared Sensor (TIRS). These two sensors allow for 11 spectral bands to be examined within one image. The LANDSAT 8 system is returning over 400 scenes per day with a 30 meter resolution. The data obtained from LANDSAT 8 represents a revolutionary advance in technology and provided a number of research opportunities. Reefs represent one such opportunity and LANDSAT 8 data has proved useful in the investigation of geomorphologic zonation, reef community classification, and the bathymetry of reefs. Being able to study the depth of reefs, or more generally the depth of the ocean remotely provides significant scientific and military benefits. It can help with the optimal planning of ship routes, improving the safety of ships transit, and provide continually updated information, something paper charts cannot provide. It also allows for further investigation into how climate change is affecting the depth of the ocean, and reefs over time, a particularly timely topic.

Materials and Methods:

Three test areas in the Great Barrier Reef were selected from LANDSAT 8 Path 92 Row 74, and images were downloaded for April, August, and December, 2016. Those images were chosen from three distinct seasons - spring, summer, and winter – and three different areas, in order to examine whether seasonal variation or location effects the accuracy of predicting the depth of the reef. The images were downloaded from the USGS EarthExplorer website. In addition, bathymetric data from the Deep Reef Explorer Project (Beaman, 2010) was downloaded as a basis for comparison to our LANDSAT 8 predictions. MICRODEM, a freeware microcomputer mapping program, was used to compare the LANDSAT 8 images and the Deep Reef Explorer data. The bathymetric data was overlaid atop the LANDSAT 8 scenes; then, using a field correlation matrix, the band with the best reflectance-depth correlation was selected. **Table 1** shows the different bands utilized by LANDSAT 8. Only depths from 0 to 30 meters were included, due to the fact that there is little light penetration, and thus reflectance, below this 30 meter depth. Next, using MATLAB, a multiple linear regression was applied to each data set, generating an equation that could be used to predict water depth in each region. This model was then compared to the Deep Reef Explorer data (Beaman, 2010) in order to determine the accuracy of the model obtained from the LANDSAT data. The deviation between the predicted and actual depths was calculated using MATLAB and then plotted on the LANDSAT imagery using MICDODEM. The coefficients from the MATLAB equations were then compared to see whether the equations derived could accurately predict the water depth in the future. The MATLAB code used to calculate the multiple linear regression equation and deviation is shown below:

zestimate=b1*band1+b2*band2+b3*band3+b4*band4+b5*band5+b6*band6+b7*band7+b8*band8+b9*band9+b10*band10+b11*band11+b12

The equations for a linear model and an exponential model were also derived in order to determine which model best fit the data. The linear equation and subsequent exponential equations are:

zestimate= intercept+slope*band3
zestimate=a*exp(b*band3);

Table 1: LANDSAT 8 bands and wavelengths.

Bands	Wavelength (micrometers)
Band 1 - Ultra Blue (coastal/aerosol)	0.43 - 0.45
Band 2 - Blue	0.45 - 0.51
Band 3 - Green	0.53 - 0.59
Band 4 - Red	0.64 - 0.67
Band 5 - Near Infrared (NIR)	0.85 - 0.88
Band 6 - Shortwave Infrared (SWIR) 1	1.57 - 1.65
Band 7 - Shortwave Infrared (SWIR) 2	2.11 - 2.29
Band 8 - Panchromatic	0.50 - 0.68
Band 9 - Cirrus	1.36 - 1.38
Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19
Band 11 - Thermal Infrared (TIRS) 2	11.50 - 12.51

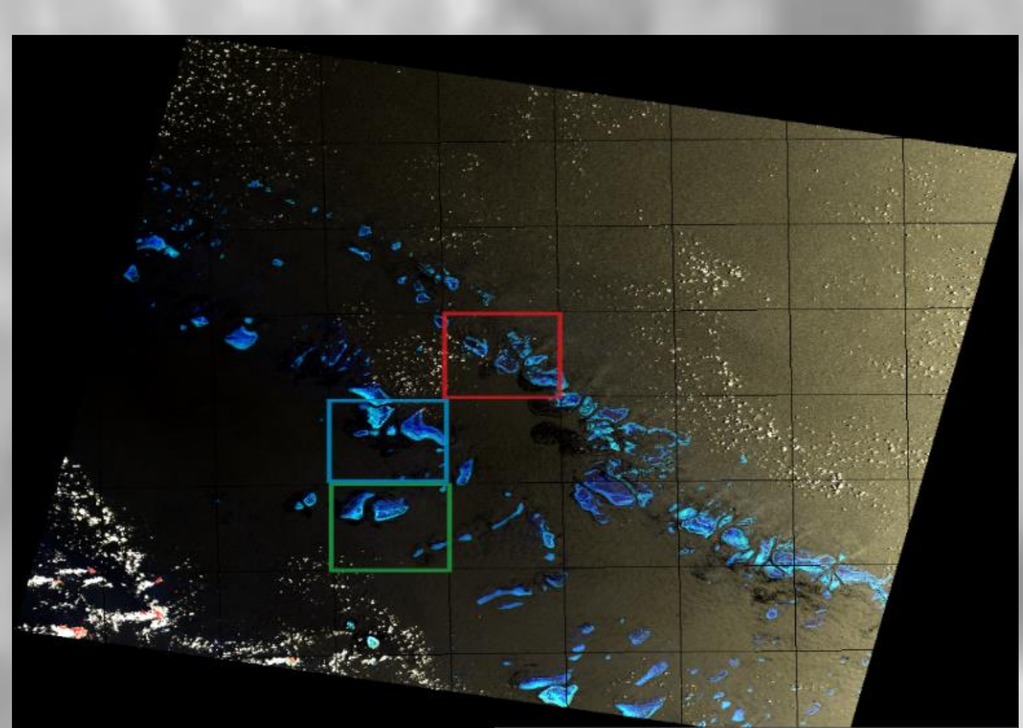

Figure 1. The locations of the data analyzed. Area 1 is within the blue box, area 2 is within the green box, and Area 3 is within the red box.

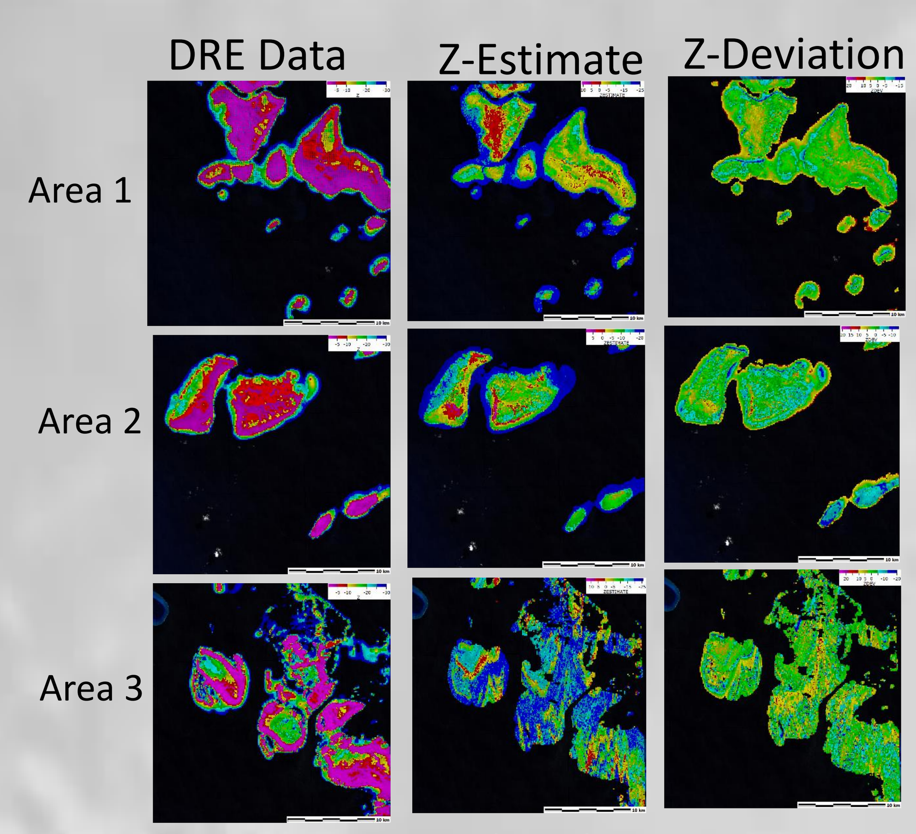

Figure 2. Areas 1, 2, and 3 in the Great Barrier Reef during April with the Deep Reef explorer data, the estimated depths from the multiple linear, and differences in depth between the Deep Reef Explorer Project data and the calculated depths.

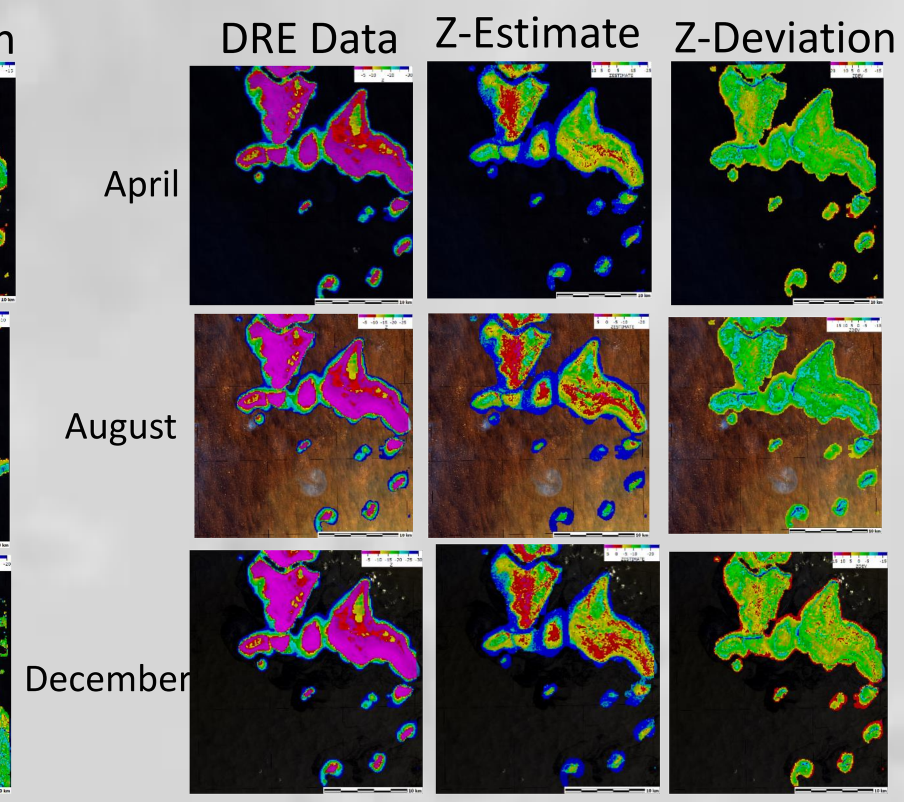

Figure 3. Area 1 in the Great Barrier Reef during April, August, and December with the Deep Reef explorer data, the estimated depths from the multiple linear, and differences in depth between the Deep Reef Explorer Project data and the calculated depths.

Table 2: The correlation coefficients and constant values from the multiple linear.

	April:				August:				December:		
Band:	Area 1	Area 2	Area 3		Area 1	Area 2	Area 3		Area 1	Area 2	Area 3
1	-0.03225	-0.0132	-0.07085		-0.02906	-0.01328	-0.06177		-0.03142	-0.0293	-0.04123
2	0.015658	0.008834	0.039012		0.017624	0.008894	0.036971		0.018095	0.01957	0.025116
3	0.009581	0.008866	0.014197		0.009101	0.008857	0.013779		0.007318	0.006248	0.00789
4	-0.00211	-0.00137	0.000349		-0.00223	-0.00131	-0.00312		-0.00066	-0.00035	0.001066
5	-0.00115	-0.00187	0.021232		0.006491	-0.00187	0.014927		0.002017	0.006292	0.007966
6	0.013057	-0.01721	0.006873		0.003638	-0.01757	-0.0015		-0.01031	0.001448	-0.00546
7	-0.02627	0.017912	-0.0398		-0.00555	0.018696	-0.01276		0.013586	-0.0083	0.000625
8	1.42E-05	-0.00083	-0.00198		-0.00191	-0.00085	-0.0037		-0.00154	-0.00238	-0.0023
9	-0.01044	-0.00047	-0.00346		-0.00569	-0.00026	0.004695		-0.0116	-0.01137	-0.02263
10	-0.01356	0.001569	0.003653		-0.00288	0.001308	-0.01458		0.001576	0.010173	0.013351
11	0.011837	0.000462	-0.00978		-0.00393	0.000728	-0.01565		0.004001	-0.00302	-0.02916
Constant	291.5947	-60.3938	436.8714		235.8911	-63.2036	916.1232		-12.3828	-67.8685	573.6368

Table 3. Coefficients from the exponential regression.

Exponential			
Area 1:	**April**	**August**	**December**
a	-1012	-4184	-705.4
b	-0.0006	-0.0007	-0.00046
Area 2:			
a	-5139	-5164	-643.2
b	-0.0007	-0.0007	-0.00044
Area 3:			
a	-42650	-112600	-2055
b	-0.0011	-0.0012	-0.00059

Table 4. Coefficients from the linear regression.

Linear:			
Area 1:	**April**	**August**	**December**
intercept	-48.51	-63.851	-49.4865
slope	0.0042	0.00616	0.003876
Area 2:			
intercept	-70.174	-70.12	-51.4367
slope	0.00687	0.00686	0.004057
Area 3:			
intercept	-58.44	-71.54	-55.2652
slope	0.00582	0.0073	0.004597

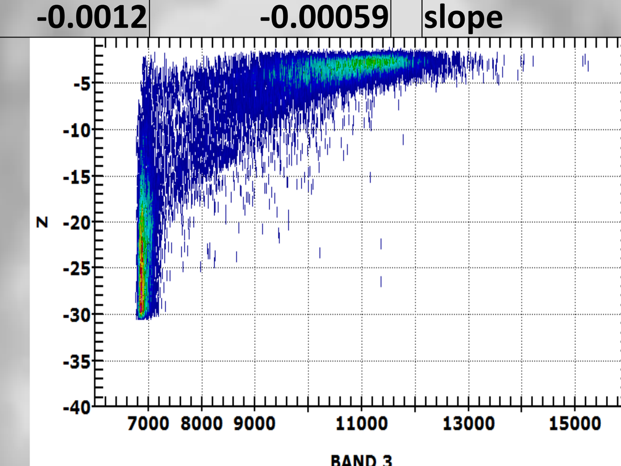

Figure 4. Band 3 correlation with depth for area 1 in April.

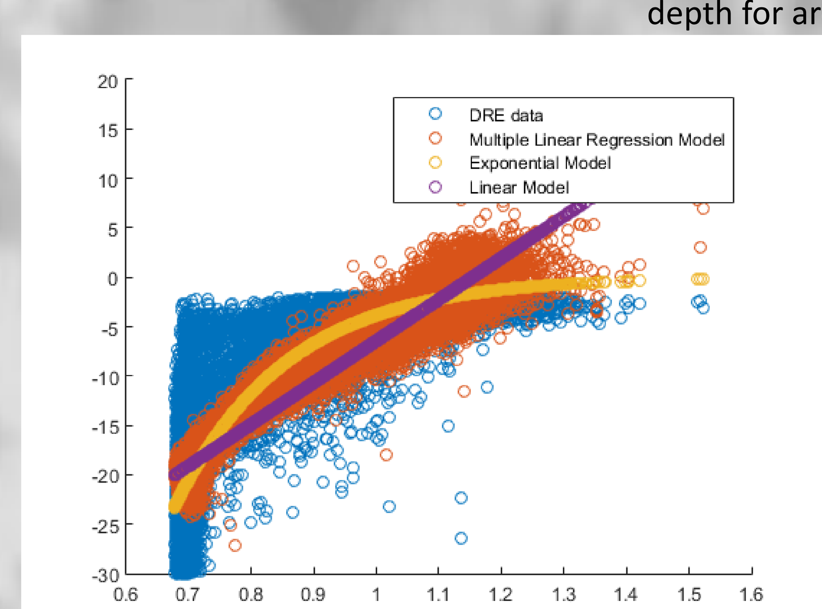

Figure 5. Comparison of multiple linear regression, exponential, and linear models for depths 0 to 30 meters.

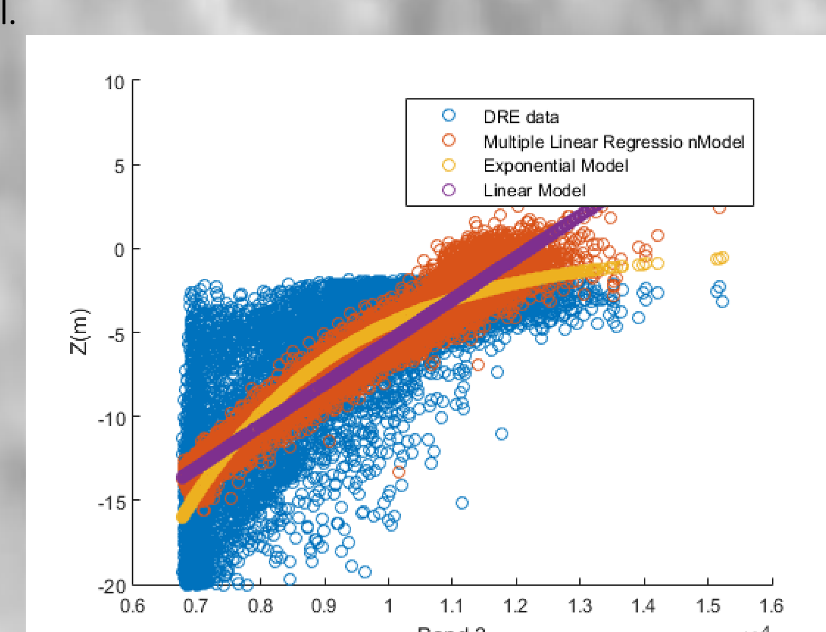

Figure 6. Comparison of multiple linear regression, exponential, and linear models for depths 0 to 20 meters.

Results:

For all three areas, during all three months of the year, Band 3(green) had the highest correlation value with depth. The correlation value was above 0.7145 for every area, which demonstrates a fairly strong linear correlation between Band 3 and depth **(Figure 4).**The green and red areas in this figure represent high point density concentrations. The two area of high point density follow two distinct linear patterns.

Figure 2 shows the April images. **Figure 2** shows relatively small variation between the actual and predicted depth in the center, or main parts of the reef. The green and yellow colors demonstrate a range of +/- 5 meters difference between the two data sets. The largest variation between the actual and predicted depth occurs on the edges of the reef, where the waves are breaking. This trend can also be seen in two other areas in the Great Barrier Reef, also during the month of April. The positive Z-estimate values in the middle image (red shading), in a region with negative depths in the ground truth, indicate a failure of the multiple linear regression.

In **Figure 3**, the differences in in predictably of an area based on the month of the year. All three images in all three of the rows look almost identical, with similar coloring in each part of the graphs. The same trends that **Figure 2** demonstrate are present. The data is most accurate in the center areas of the reefs away from where the waves are breaking, with a small range of deviation. Again, small parts of the reef have positive estimated z values, indicating a failure of the model.

Table 2 shows the coefficients for the multiple linear regression equation. After examining all the coefficients it is clear that the equations are not all alike, and thus the equations derived can not be used to predict the depths in every area for all times of the year. The coefficient values as well as the constants were variable.

Due to the variable coefficient and constant values in the multiple linear regression, as well as the positive z-estimate values, two more models were done to see if they would better represent the data. The equations for the linear and exponential models can be seen in Materials and Methods section. From these equation, coefficients for the equation, as well as a scatter plot were created to see which model best fit the Deep Reef Explorer data (Beaman, 2010). In **Table 3** and **Table 4**, after applying linear and exponential fits to the data, it becomes slightly more apparent that the multiple linear regression provides the best equation to model the data with the smallest variation. The coefficients in the multiple linear regression are more constant across different regions and times of year than the exponential and linear model coefficients. Which model best depicts the data is can be more easily seen in a scatter plot. After analyzing **Figure 5**, the multiple linear regression model and the exponential model looked like the best fit for the original z data. The main issue with the multiple linear regression model was that there were positive Z-estimate values, as described above. The exponential model did not have this issue. In order to minimize the amount of positive z-estimate values, the data was filtered to only go to depths of 20 meters instead of 30 meters. This was an attempt to get rid of some of the depths that may have been skewing the data, for light was not penetrating deep enough past 20 meters. **Figure 6** shows that this filtering of data causes the multiple linear regression to appear to be the best fit for the model. The amount of positive z-estimate values was decreased, and the model followed more of the overall trend of the original z-data.

Conclusion:

The predictability of ocean depth using LANDSAT imagery was not as accurate as hypothesized. Although there was a high correlation in each of the areas over the course of these three months with Band 3, this high correlation did not translate directly to the multiple linear regression. The coefficients and constants obtained from the multiple linear regression varied widely across the regions and months the LANDSAT image was taken. This may be due to various factors. One factor might be that the light penetrated even shallower than the 30 meters that was originally predicted. Limited light penetration could have skewed the values and caused the correlation not to be as linear as predicted.

However, as demonstrated by **Figure 2**, the accuracy of a given region during a specific period of time, between the Deep Reef Explorer data (Beaman, 2010) and the estimated depths, was very good. There was a small range of deviation for most of the reef areas. The largest deviations occurred where the waves were breaking over the reef, for these breaking waves would effect the reflectance value, and then skew the values obtained from the multiple linear regression prediction. However, the positive Z-estimate values provide for another source of error in the data. The positive values indicate that these parts of the coral reef are above water. The positive z values occur where the actual depths of the ocean are 5 meters deep or less. **Figure 3** also demonstrates that the time of year does not have a large effect on the predictability of the depths. Area 1, during all three times of the year displayed, similar estimated depths, and deviations from the Deep Reef Explorer data (Beaman, 2010).

The range of the z-deviation values, or the difference between the Deep Reef Explorer data (Beaman, 2010) and the z-estimate data, also provides another area of concern. The range of deviation in all three areas during the month of April is +15 meters above the reference depth, to -20 meters below sea level. Such a wide range of deviation and error may be attributable to errors in calculating the reflectance values.

Future research could be done to figure out at exactly what depth light penetrates accurately to in the Great Barrier Reef., in order to eliminate the presence of positive z-estimate values. In addition, the presence of white coral at low tides, as well as the waves breaking over the reef and producing white caps needs to somehow be taken into consideration. This will help to improve the predictability of the multiple linear regression. There is still much room for improvement and investigation in this area of study. If the ability accurately to predict water depths from LANDSAT 8 images is perfected, then the functionality of the United States Navy, and other militaries across the world might be significantly improved, as it will provide for real time imagery, accuracy, and route optimization.

References:

Beaman, R., 2010, Exploration & Discovery of Australia's Oceans: https://www.deepreef.org/, accessed 10 April 2017.

United States Geological Survey (USGS), 2017, Landsat 8 | Landsat Missions: https://landsat.usgs.gov/landsat-8, accessed 10 April 2017.