

Factors Controlling Coral Skeletal U/Ca Ratios with Implications for their Use as a Proxy for Past Ocean Conditions

Emily Patterson, Spencer Eanes, Penelope Lancrete, Anne Gothmann^a, & Paul Roback^b

^aDepartment of Environmental Studies and Physics, St. Olaf College, Northfield, MN

^bDepartment of Mathematics, Statistics, and Computer Science, St. Olaf College, Northfield, MN

<https://doi.org/10.33697/ajur.2020.031>

Students: patter5@stolaf.edu, eanes1@stolaf.edu, lancre1@stolaf.edu

Mentors: gothma1@stolaf.edu*, roback@stolaf.edu

ABSTRACT

Seawater temperature, salinity and carbonate chemistry have been shown to influence the uranium/calcium (U/Ca) ratios of scleractinian coral skeletons. This apparent sensitivity of U/Ca to multiple environmental parameters calls into question whether there is one environmental variable that most strongly controls coral U/Ca, and whether U/Ca can be straightforwardly applied as a paleoenvironmental proxy due to the tendency of environmental variables to covary in space and time. In this study, uranium concentration data from an existing compilation of tropical scleractinian coral U-series measurements is paired with environmental data from the World Ocean Atlas (WOA) and the Global Ocean Data Analysis Project (GLODAP) to examine the sensitivity of coral skeletal U/Ca to multiple seawater properties including temperature, salinity, pH, and saturation state. First, univariate linear regressions and multiple linear regressions were used to compare relationships between uranium and environmental parameters in the dataset with relationships observed in previous studies. Next, principal component analysis and regularized regression were used to identify the most likely predictors of coral U/Ca in order to create a multiple linear regression model. Results indicate that pH, Ω , alkalinity, and temperature are all significant predictors of uranium concentrations in coral. The magnitude and strength of relationships between U/Ca and environmental variables also differ across different genera. Seawater properties with strong correlations and small ranges make interpretation of these results difficult. However, results of these analyses indicate that U/Ca is dependent on multiple environmental parameters and that previously developed univariate regressions may be insufficient to characterize the full range of variables that influence coral [²³⁸U].

KEYWORDS

Coral; Paleooceanography; Proxy Calibration; Uranium; Multiple Linear Regression; Regularized Regression; Environmental Change; Oceanographic Databases

INTRODUCTION

Understanding how and why the environment has changed naturally over Earth's history can provide valuable insight into how Earth's climate system will respond to ongoing anthropogenic perturbations. Instrumental records of oceanic changes do not exist beyond the last ~100 years, however, so to access information about ocean changes further in the past, indirect indicators of environmental change, referred to as 'proxies', are used. Scleractinian corals skeletons are one of the most fruitful proxies for past climate and environmental change because their age can be precisely determined, they tend to record environmental properties of the seawater in which they grow, it is possible to make multiple measurements of environmental properties on the same sample due to their relatively large size, and they can capture sub-annual environmental variability.^{1,2} In recent years, the U/Ca ratio of corals has been explored as a proxy that responds linearly to changes in seawater properties including seawater temperature,³ salinity,⁴ pH,⁵ and carbonate ion concentration.⁶ The ability of coral U/Ca to capture past seawater salinity, pH, and/or carbonate ion concentration, in particular, would be a valuable addition to the arsenal of coral-based proxies, due to the relative lack of proxies for these important environmental variables. In particular, indicators of pH and carbonate ion concentration could help us better understand the relationship between changes in atmospheric CO₂ and ocean acidification.⁷ An understanding of past changes in seawater salinity can provide insight into the relationship between salinity and ocean circulation and also can be used to learn about the relationship between changes in climate and the hydrological cycle.⁸

There is reason to expect U/Ca in corals to be sensitive to seawater pH and/or carbonate ion concentration. Inorganic aragonite precipitation experiments conducted by DeCarlo *et al.* (2015) suggest that U is incorporated into the aragonite mineral in proportion to the ratio of [U]/[CO₃²⁻] in seawater.⁹ These results support the idea that corals, which make aragonite skeletons, may also incorporate uranium in proportion to the concentration of carbonate ion in seawater. A potential problem is the fact

that corals do not appear to precipitate their aragonite skeletons directly from seawater, but rather from a semi-enclosed space referred to as the calcifying fluid.¹⁰⁻¹² Because corals are able to manipulate the chemical composition of this calcifying fluid (through active pumping of protons, for example), the calcifying fluid has a carbonate chemistry that is distinct from external seawater.¹³⁻¹⁵ As a result, coral U/Ca could be more dependent on other environmental parameters that have strong effects on the coral skeletal growth process.

There is also evidence that coral U/Ca may be dependent on sea surface salinity. First, Swart and Hubbard (1982) showed that coral U/Ca ratios appear to be dependent mainly on the absolute concentration of U in seawater.¹⁶ As a result, they postulated that U/Ca in coral could record salinity variations as a result of changes in the absolute concentration of U. Results of Shen and Dunbar (1995), showing that U/Ca in tropical coral skeletons decreased during times of greater local rainfall, generally support this idea.¹⁷

Many published studies have also examined relationships between U/Ca and seawater temperature.^{3, 18-23} While it does not appear that U/Ca is sensitive to temperature in inorganic aragonite,⁷ U/Ca is correlated with temperature in natural samples over seasonal cycles.²⁴ As a result, U/Ca may be related to seawater temperature through an effect of temperature on the coral skeletal growth process.

To date, the studies calibrating environmental proxies based on coral skeletal U/Ca ratios come from natural samples collected from a single geographic location,^{17, 24} or from laboratory-based culture experiments in which one environmental variable is varied while others are held constant.⁵ However, if paleoclimatologists hope to apply proxies over a broad geographic range and far into the past (when seawater boundary conditions may have looked quite different), it is important to quantify whether U/Ca depends on a single, or multiple, environmental parameters. It is also important to quantify how precisely environmental variations could be reconstructed from coral U/Ca.

One common approach in proxy development is the ‘space-for-time’ approach, which trains proxies based on modern spatial variability in environmental parameters. For example, in the case of ocean sediment cores, measurements of chemical, geological, physical or biological variables are made in the top few centimeters of ocean sediment and are examined for correlations with environmental variables in the water column overlying the sediment core location. Recent studies have combined core-top measurements of Mg/Ca ratios in foraminifera with oceanographic databases to refine the foraminifera-based Mg/Ca paleothermometer.^{25, 26} An assumption embedded in the ‘space-for-time’ calibration strategy is that the relationships between environmental changes and the proxy of interest are common across space and at a single location over hundreds, thousands, or even millions of years.

This study attempts to conduct a calibration of U/Ca’s sensitivity to environmental changes using a ‘space-for-time’ approach whereby modern and Recent (Holocene and younger) tropical coral U data are compared with environmental data from the World Ocean Atlas (WOA) and the Global Ocean Data Analysis Project (GLODAP). A wealth of coral U data exists in the literature due to the fact that tropical corals have been measured by U-series dating as a way to generate sample ages for decades (e.g. Robinson *et al.* 2014).²⁷ In order to obtain U-series ages from coral, the abundance of ²³⁸U, which makes up >99% of the U in the coral skeleton, is determined. As a result, U-series measurements made over the past few decades also produce measurements of coral [U]. Recently, Chutcharavan *et al.* (2018) compiled coral U-series measurements from the literature and analyzed the data to determine whether seawater ²³⁴U/²³⁸U ratios appear to have changed over time.²⁸ In the present study, the compilation of Chutcharavan *et al.* (2018) is used to investigate whether the [U] of recent corals appears to vary significantly with environmental variables such as temperature, salinity, pH, and/or carbonate ion concentration.

METHODS AND PROCEDURES

Data Compilation

Coral U-series data compiled in Chutcharavan *et al.* (2018) was used to evaluate the relationship between coral skeleton [U] and the environmental properties of seawater in which the corals grew.²⁸ The main variables of interest alongside the genus, age, calcite abundance (as % calcite), and ²³⁸U of the corals include sea surface temperature (SST), salinity (SSS), pH, total alkalinity, dissolved inorganic carbon, and saturation state with respect to aragonite (Ω). The two most abundant genera from the coral U-series data were included in our analyses. A summary of these variables for these two genera can be found in **Table 1**. Coral ²³⁸U data is converted to coral U/Ca molar ratios using the atomic weight of U and by assuming that Ca concentrations in the tropical coral skeletons are 9.52 mmol Ca/g coral aragonite. Analyses of Ca concentrations in coral skeletons show that [Ca] in tropical coral skeletons varies by ~0.4%.^{29, 30} As such, U/Ca variations in coral skeletons are dominated by variability in ²³⁸U. All samples for which ²³⁸U were labeled as “not reported” and/or genus was reported as “nd”, “unknown”, or “unidentified” were excluded from the analysis.

While the coral samples in Chutcharavan *et al.* (2018) range in age from modern to ~55 kyr, we restrict our investigation to $n=737$ samples of Holocene age, specifically <10 kyr. These samples indicate stable seawater $[^{234}\text{U}]/[^{238}\text{U}]$ ratios over the last 10 kyr. Restricting our samples to those of Holocene age and younger ensures that we retain a relatively large number of samples for our analyses. Additionally, the Holocene is a time of relatively stable climate, such that we expect the modern spatial variability in environmental variables covered by our samples from different sites to be much greater than the Holocene environmental changes at any given site. Although recent studies have demonstrated variability in SSTs of up to 1°C during the Holocene at high latitudes, SSTs in the tropics appear to vary by ~0.5°C during this time.^{31,32} Discrepancies in environmental records from the Holocene generated by different proxies also remain, making it difficult to explicitly correct for climatic variability in Holocene-age samples.³²

Variable	5 th Percentile	Median	Mean	95 th Percentile
$[^{238}\text{U}]$ (ppm)	0.93	1.14	1.20	1.66
Age (ka)	0.01	0.46	2.4	9.00
Calcite Abundance (%)	0.00	0.50	0.44	1.00
Salinity (g/kg)	33.05	35.11	34.72	35.94
Temperature (°C)	24.29	26.65	26.63	29.35
pH (total scale)	8.05	8.11	8.10	8.13
Total Alkalinity (μmol/kg)	2180.89	2264.88	2272.45	2364.90
TCO ₂ (μmol/kg)	1870.91	1953.53	1943.54	2001.25
Ω	3.16	3.76	3.69	4.18
δ ²³⁴ U initial (‰)	141.23	144.83	144.66	148.51

Table 1. Data summary of key coral and environmental variables.^{29–32} These values include all non-missing values with age under 10,000 years from the two most abundant genera in the Chutcharavan *et al.* (2018) compilation.

Complementary environmental data from the World Ocean Atlas (WOA) and the Global Ocean Data Analysis Project (GLODAP) was paired with the coral U-series dataset. Both projects have been updated with recent oceanographic data, mapped at 1° by 1° gridded points with latitudes, longitudes, and depths associated with each point. This feature makes the WOA and GLODAP datasets appropriate to compare directly with one another and with our coral dataset. The mean oceanographic temperature and salinity data from the WOA is based on the analysis of historical oceanographic profiles and select surface-only data as described in greater detail in Zweng *et al.* (2018) and Locarnini *et al.* (2018).^{33,36} All other environmental variables were accessed as mapped climatologies from GLODAPv2.^{34,35} Additional detail on the methodologies used to create these mapped climatologies can be found in Lauvset *et al.* (2016).³⁴

To pair coral data from a particular geographic location with environmental data, a distance function was created which measures the distance from each coral sample included in Chutcharavan *et al.* (2018) to the grid locations where each of the environmental variables were measured. Since the GLODAP and WOA data are built on a grid and use interpolation from measured data, coral environmental parameters were estimated by taking the weighted average of the three closest values to each fossil coral sample. The function was coded in C++ for efficiency and run in R using the Rcpp library.^{37–39} The function measured the distance between any two points on earth, that is the coral and each environmental grid location, measured in latitude and longitude, taking the earth's curvature into account. In the case that latitudes or longitudes were not reported in Chutcharavan *et al.* (2018) or in the original scientific papers cited by Chutcharavan *et al.*, the original scientific papers reporting coral U-series data were searched for coral collection sites, and the longitudes and latitudes corresponding to these sites were identified.

Data Cleaning

The coral samples compiled in Chutcharavan *et al.* (2018) were constrained to those of Holocene age (*i.e.* ~10,000 years of age and younger). At ages beyond about 15,000 years, there are fewer coral samples in the compilation for a given time point such that the data are sparser. After filtering for age, the resulting dataset contains 737 of the original 1860 samples. Filtering the coral samples to those younger than ~10,000 years results in $[^{238}\text{U}]$ vs. Age slopes close to zero as seen in **Figure 1** below.

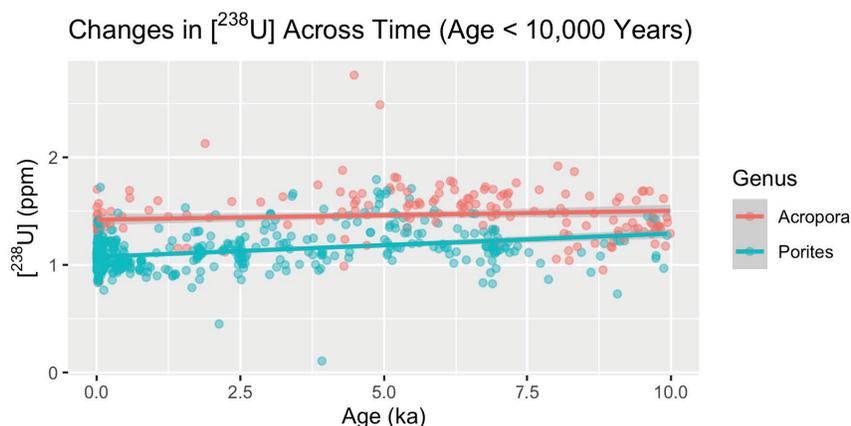


Figure 1. This plot uses coral ^{238}U data as a function of age for samples of the genera *Acropora* and *Porites* across the entire Chutcharavan *et al.* (2018) dataset, filtered to remove NA's and filtered for samples younger than 10,000 years. Age slopes were close to zero for each genus: *Acropora* (0.008 (+/- 0.004)) and *Porites* (0.022* (+/- 0.003)), and when including both genera (0.03* (+/- 0.003)).

*indicates significance at the .05 level. Results include coefficient estimates followed by standard errors.

Next, the relationship between ^{238}U and % calcite was explored. The presence of calcite in the coral skeleton can overprint the original geochemical composition of coral skeletons at the time of growth because U incorporation in calcite and coral aragonite differ.⁴⁰ Some labs reported % calcite as a range of values, in which case the maximum reported value was used in order to avoid underestimating the % calcite of the sample. The data was assessed on variable and genus levels for differences between corals with different amounts of calcite. Calcite values were divided into three groups: less than or equal to one percent, greater than one percent, and those with no values listed ($\leq 1 = 120$ samples, $> 1 = 37$ samples, N/A = 580 samples). The threshold chosen was 1% because this value corresponds to the detection limit of many X-ray diffraction measurements that quantify calcite abundance.²⁸ The directions of the slopes for the group where calcite $> 1\%$ differs from that of the calcite $\leq 1\%$ group for pH and Ω . There were no significant differences in slopes between calcite groups for temperature or salinity (**Figure A1** in Appendix). To optimize U data quality, the 37 samples with values of calcite greater than 1% were removed from the working dataset. We also note that the majority of samples younger than 1 kyr did not record percent calcite.²⁸ The original studies corresponding to these published samples assess preservation by qualitative measures including scanning electron microscopy, or otherwise note that there are no detectable traces of calcite in the samples.^{41,42} Our univariate regression results for temperature, salinity, and pH all did not change significantly between the case where samples were restricted to only those with calcite values $\leq 1\%$ (*i.e.* samples with calcite values of N/A were removed) and the case where samples with calcite values of N/A were included. As such, the final dataset presented here contains samples with calcite values of $\leq 1\%$ and N/A, in order to obtain a more robust analysis with greater sample sizes.

The next data integrity check was to examine ^{238}U values at the lab level for outliers. Labs 1, 3, 9, 11 and 16 were investigated as their means fell greater than one standard deviation from the mean of all the samples. A multiple linear regression model was trained on all the other lab data and used to create prediction intervals for the data from those five more extreme labs. Only six out of 104 observations fell outside of a 95% prediction interval across all sites for those five outlying labs, so no lab exclusions were made (**Figure 2**). Additional details can be found in the project GitHub repository, a link to which is provided in the Appendix.

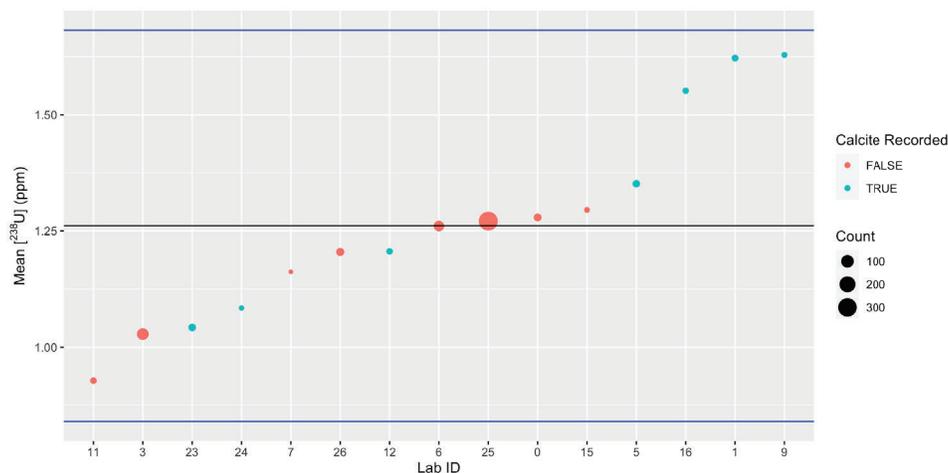


Figure 2. Each lab's mean $[^{238}\text{U}]$ was calculated for all samples, and points sized by number of samples. Points are colored by whether or not the lab reported calcite values for their sample. The black horizontal line is the global mean, and the blue lines are +/- two standard errors from the lab aggregated mean.

An additional analysis conducted to validate the data was to check if any sites had unusual or strongly correlated environmental data values. Since environmental data is resolved at the site level, there are 51 unique sites, and thus 51 values for each of the environmental variables (temperature, salinity, pH, Ω , alkalinity, TCO_2). As seen in Figure 3, salinity values above and below 34 g/kg seemed to cluster in regard to their pH and temperature values, leading to a further investigation of the sites with low salinity values.

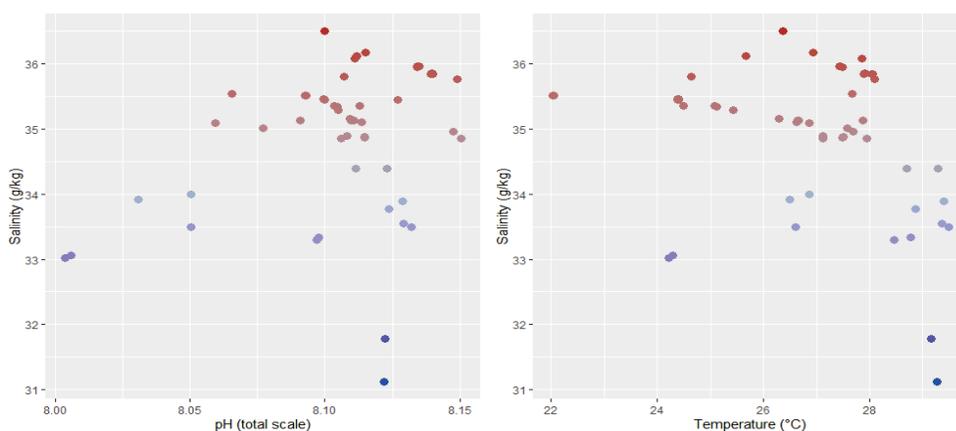


Figure 3. Salinity values associated with each geographic site in our data set, plotted against pH and temperature, colored by salinity.

These points of interest came from a broad region of low mean salinity waters in the Indo-Pacific, and their salinity, temperature, and pH values are reasonable for this geographic location (Figure 4).

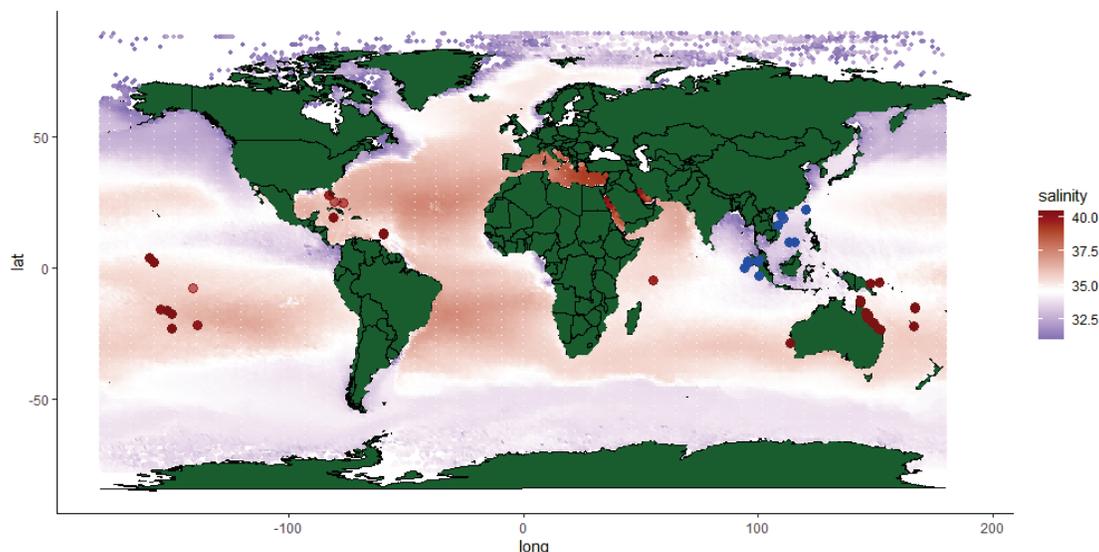


Figure 4. Global map with gridded ocean salinity data colored from blue (low) to red (high).³⁶

Variability in relationships between [U] and environmental variables were also examined by genus; only corals with genera that had more than 50 samples were selected. It is notable that 297 samples had no genus listed and were excluded. After selecting the top two genera, filtering for age and calcite abundance, and finding no issues with locations or labs, 700 samples remained out of the original 1860 (Table 2). After data cleaning, 51 sites remained out of the original 78. The original data distribution can be found in Table A1 in Appendix.

Genus	Number of Samples	Percent of Total
<i>Pozites</i>	533	76%
<i>Acropora</i>	167	24%
Both Genera	700	100%

Table 2. After filtering for age and calcite and keeping genera with more than 50 samples.

Statistical Methods

Univariate linear regression was used to test if [²³⁸U] can be used as a proxy for temperature, salinity, and pH separately, as has been done in previous studies.^{3, 6, 18–24, 43, 44} Additionally, multiple linear regression was used in order to account for additional variables which cannot be controlled for in an observational setting. In order to perform multiple linear regression, predictors need to be chosen, a process known as feature selection. Regularized regression (e.g. lasso and ridge regression) has been widely used in recent years due to its demonstrated prediction prowess on data with many predictors of unknown relationships. Regularized regression is suited to this task because the underlying relationships are expected to be linear, not categorical or tree-like. The technique of regularized regression solves a linear regression minimization problem for a best fit linear model to data but includes a penalty parameter proportional to a measure of coefficient size. Thus, this technique prefers more parsimonious models with coefficients of smaller magnitude. There are multiple flavors of regularized regression with different methods to measure the magnitude of models’ coefficients, the two most common being ridge regression, which uses the sum of the squared coefficients, and lasso regression, which uses the sum of the absolute value of the coefficients. Elastic net is a more general form of regularized regression that uses a linear combination of squared and absolute value coefficients.⁴⁵

To compare the regularized regression methods, two stages of 10-fold cross validation were employed. The first stage is used to reserve some out-of-bag test data in order to compare regularized regression techniques. For each of the folds of the first-stage, 10-fold cross-validation is again performed to select the minimum and one standard error penalty parameter for each of the three techniques: ridge, lasso, and elastic net. Then, minimum and one standard error models for each of the techniques were used to compute a root mean squared error (RMSE) on the initial reserved out-of-bag data from the first stage of cross-validation. RMSE is a technique used to compare model accuracy in regression by taking the mean of the squared differences of actual and predicted values, and then taking the square root of that mean. A lower RMSE indicates a more accurate model. The RMSE is then averaged for the 10 folds of the first stage of cross-validation. The minimum penalty parameter always performed better than the

one standard error penalty parameter across the three models, and the RMSE was within .01% for the minimum penalty parameter across the three methods. Thus, since there was on average essentially no difference in RMSE between the techniques, lasso was chosen for simplicity.

Bootstrapping, the technique of resampling the data with replacement, was employed to ensure stable estimates from variable selection. Due to the randomness of sampling in cross-validation of lasso modeling, bootstrapping the data was necessary to quantify random differences and stabilize the results. The data was bootstrapped 5,000 times, a lasso model was fit to the bootstrap sample, and the selected variables, their coefficients, and the penalty parameter were all recorded.

Best subset selection was implemented for comparison with the bootstrapped lasso approach. This involved fitting every possible combination of linear models with a given set of predictors, of which there are 2^p possible models, where p is the number of predictors in the data. These models can then be compared with a selection of characteristics, such as adjusted R-squared, or information criterion like AIC (Akaike information criterion) and BIC (Bayesian information criterion).⁴⁶ Information criterion are frequently used and often preferred when comparing models with different sets of predictors, as they can help account for the differing complexities between models.⁴⁷ BIC was selected since it more heavily penalizes additional predictors than AIC, and it was valuable to achieve a parsimonious model to compare against the results of bootstrap lasso.

Finally, principal component analysis (PCA) was performed on the data to help understand the relationships between the predictors. PCA has previously been applied to questions in the broad field of paleoceanography to identify structure in data when multiple variables are involved.⁴⁸⁻⁵⁰ For example, Lough (2004) used PCA to determine whether scleractinian coral proxy-based sea-surface temperature records captured variability in El-Niño Southern Oscillation.⁵⁰ PCA allows the collapse of high dimensional data into a few independent components. Singular value decomposition is applied to the input data and results in a linear transformation of the data such that the first new predictor describes most of the variability in the data, and each subsequent predictor describes less. The original predictors can also be visualized in “PCA-space” to give some understanding of how they relate to one another.

RESULTS

*Comparison to Previous Studies
Temperature*

Linear regression was used to test if there was a significant association between explanatory variables and $[^{238}\text{U}]$, when applying an age restriction of 10,000 years and excluding percent calcite values greater than 1%. The first test was between temperature and $[^{238}\text{U}]$ (Figure A2). When including both genera of interest (*Porites* and *Acropora*), there was significant evidence of a negative association between temperature and $[^{238}\text{U}]$ ($t = -16.66$, $p\text{-value} < 0.001$). When looking at each individual genus, a highly significant negative association between temperature and $[^{238}\text{U}]$ was found for both *Porites* ($t = -10.79$, $p\text{-value} < 0.001$) and for *Acropora* ($t = -8.89$, $p\text{-value} < 0.001$). Negative relationships between $[^{238}\text{U}]$ and temperature are consistent with results from previous studies on *Porites* coral (Table 3, Figure 5). Comparisons between the results in this study with and without filtering for age can be found in Table A2.

Primary Results			Previous Lab Studies			
Intercept	Temperature	Genus	Intercept	Temperature	Genus	Source
2.295* (+/- 0.108)	-0.044* (+/- 0.004)	<i>Porites</i>	1.928	-0.033*	<i>Porites</i>	Ourbak <i>et al.</i> (2006)
3.600* (+/- 0.239)	-0.083* (+/- 0.009)	<i>Acropora</i>	2.232	-0.0456*	<i>Porites</i>	Min <i>et al.</i> (1995)
3.169* (+/- 0.116)	-0.074* (+/- 0.004)	Both Genera	2.106	-0.0367*	<i>Porites</i>	Correge <i>et al.</i> (2000)
			1.488	-0.0212*	<i>Porites</i>	Armid <i>et al.</i> (2011)
			1.957	-0.029*	<i>Porites</i>	Wei <i>et al.</i> (2000)
			2.057	-0.034*	<i>Porites</i>	Felis <i>et al.</i> (2009)
			2.26	-0.044*	<i>Porites</i>	Fallon <i>et al.</i> (1999)
			2.24	-0.046*	<i>Porites</i>	Sinclair <i>et al.</i> (1998)

Table 3. Primary results, which showed a negative association between temperature and $[^{238}\text{U}]$, compared to previous lab studies. *indicates significance at the .05 level. Primary results include coefficient estimates followed by standard errors.

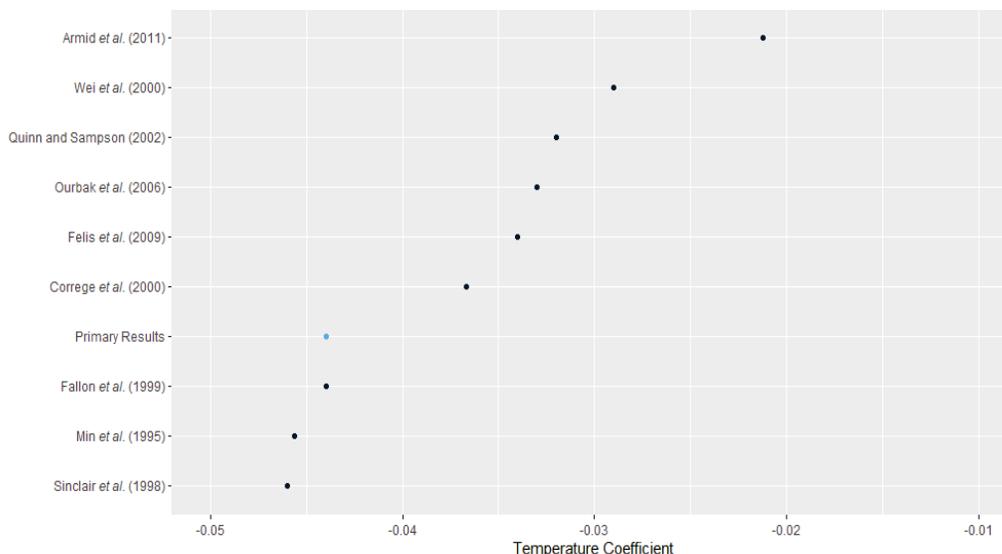


Figure 5. Coefficients from primary result using only *Porites* compared to other reported *Porites* slopes using univariate OLS regression.

Salinity

Salinity was added to the previous set of models to once again examine the relationship between temperature and $[^{238}\text{U}]$, this time holding salinity constant. Adjusting for salinity in this set of models gives insight into the true relationship between temperature and $[^{238}\text{U}]$ since salinity is more variable in this large-scale observational study than in lab studies. The model with both genera included revealed that a significant negative relationship between temperature and $[^{238}\text{U}]$ remained after accounting for salinity ($t = -12.13$, $p\text{-value} < 0.001$). When breaking the analysis down by genus, this negative relationship was still significant for both *Porites* ($t = -7.89$, $p\text{-value} < 0.001$) and for *Acropora* ($t = -8.04$, $p\text{-value} < 0.001$). Compared to Ourbak et al. (2006), both the *Porites* model and the larger model including both genera in this study followed a similar pattern, revealing a significant negative relationship between temperature and $[^{238}\text{U}]$ after accounting for salinity.

After accounting for temperature, the model with both genera showed a positive effect of salinity on $[^{238}\text{U}]$ ($t = 6.58$, $p\text{-value} < 0.001$). However, when breaking the analysis down by genus, this positive effect of salinity after adjusting for temperature was only significant for *Porites* ($t = 4.59$, $p\text{-value} < 0.001$) and not for *Acropora* ($t = 1.10$, $p\text{-value} = 0.271$). Salinity results for *Porites* revealed a significant positive relationship between salinity and $[^{238}\text{U}]$ after adjusting for temperature, which is directionally consistent with Ourbak et al. (2006) (Table 4).

Ourbak et al. (2006)			
Intercept	Temperature	Salinity	Genus
0.162 (+/- 0.019)	-0.022* (+/- 0.003)	0.162* (+/- 0.019)	<i>Porites</i>
Primary Results			
Intercept	Temperature	Salinity	Genus
0.986* (+/- 0.299)	-0.035* (+/- 0.031)	0.031* (+/- 0.007)	<i>Porites</i>
1.599 (+/- 1.804)	-0.080* (+/- 0.010)	0.054 (+/- 0.048)	<i>Acropora</i>
0.887* (+/- 0.359)	-0.059* (+/- 0.005)	0.054* (+/- 0.008)	Both Genera

Table 4. Primary results for temperature and salinity compared to previous lab studies.

*indicates significance at the 0.05 level. Primary results include coefficient estimates followed by standard errors.

pH

Linear regression was also used to test whether pH was significantly associated with $[^{238}\text{U}]$ (Figure A2). For the model including both genera, there was evidence that pH was significantly positively associated with $[^{238}\text{U}]$ ($t = 3.24$, $p\text{-value} = 0.001$). When looking at each individual genus, pH was significantly related to $[^{238}\text{U}]$ for both *Porites* ($t = 2.17$, $p\text{-value} = 0.031$) and *Acropora* ($t = -5.34$, $p\text{-value} < 0.001$), although this relationship was positive for *Porites* and negative for *Acropora*. The negative relationship

between pH and $[^{238}\text{U}]$ in the *Acropora* model was directionally similar to the results of Inoue *et al.* (2011) which also modeled only *Acropora*. However, the magnitude of our primary results was about 28 times larger than was seen in Inoue *et al.* for the intercept, and about 47 times larger for the slope (Table 5). After accounting for temperature or temperature and salinity in the *Acropora* models, the intercept was large but insignificant, and the models revealed no significant relationship between $[^{238}\text{U}]$ and pH in *Acropora* (Table 6).

Inoue <i>et al.</i> (2011)		
Intercept	pH	Genus
2.96	-0.21*	<i>Acropora</i>
Primary Results		
Intercept	pH	Genus
-2.451 (+/- 0.836)	0.441* (+/- 0.199)	<i>Porites</i>
82.319* (+/-14.961)	-9.974* (+/- 1.845)	<i>Acropora</i>
-6.116* (+/- 2.215)	0.903* (+/- 0.274)	Both Genera

Table 5. Primary results for pH compared to previous lab studies. *indicates significance at the 0.05 level. Primary results include coefficient estimates followed by standard errors.

pH <i>Acropora</i> models			
Intercept	pH	Temperature	Salinity
82.319* (+/-14.961)	-9.974* (+/- 1.845)		
11.854 (+/- 19.386)	-1.932 (+/- 2.423)	-0.091* (+/- 0.014)	
-9.362 (+/- 19.564)	1.399 (+/- 2.486)	-0.086* (+/- 0.015)	0.048 (+/- 0.050)

Table 6. Comparing univariate pH model to multivariate models for *Acropora*. *indicates significance at 0.05 level. Primary results include coefficient estimates followed by standard errors.

Further Analysis - Variable Selection

In this observational study, $[^{238}\text{U}]$ is not adequately explained by one or two variables as described in previous studies. Multiple regression modeling can be utilized in this context as an attempt to replicate how certain conditions can be carefully controlled in lab studies (Table 7). Three multiple regression models were built, one for each genus of corals, and one overall (for both genera combined), controlling for all six variables of interest simultaneously. These multiple regression models confirmed some relationships the univariate models in this study have previously shown; for example, pH has a positive relationship with $[^{238}\text{U}]$ for both genera, modeled individually and overall, when controlling for temperature, Ω , total alkalinity, salinity and total CO_2 . Interestingly, this positive relationship is significant in *Acropora* when controlling for all other parameters, contrary to the negative linear relationship found between pH and $[^{238}\text{U}]$ in the univariate regression results (Tables 6 and 7). The relationship between temperature and $[^{238}\text{U}]$ is negative in both genera modeled individually and overall, after accounting for the other five parameters, which is consistent with our primary results in Tables 3 and 4.

	<i>Porites</i>	<i>Acropora</i>	Both Genera
Intercept	-43.853* (+/-13.323)	-303.493* (+/-107.097)	-69.259* (+/- 16.004)
pH	5.469* (+/-1.581)	36.652* (+/-12.814)	8.622* (+/- 1.896)
Temperature	-0.044* (+/-0.007)	-0.018 (+/-0.034)	-0.065* (+/- 0.008)
Ω	-0.620* (+/-0.189)	-3.613* (+/-0.839)	-1.098* (+/- 0.222)
Total Alkalinity	0.003* (+/-0.001)	0.002 (+/-0.009)	0.005* (+/- 0.002)
Salinity	-0.030* (+/-0.013)	0.142 (+/-0.080)	-0.014 (+/- 0.015)
TCO ₂	-0.001 (+/-0.002)	-0.006 (+/-0.012)	-0.003 (+/- 0.002)

Table 7. Multiple Linear Regression with all variables for the two major genera. Amount of explained variance for the three models: *Porites* adj. R² = 0.317, *Acropora* adj. R² = 0.389, Both Genera adj. R² = 0.435.

*indicates significance at the 0.05 level. Primary results include coefficient estimates followed by standard errors.

While it is beneficial to see which environmental parameters are most significantly related to [238U] when including all variables of interest in multiple linear regression, feature selection can be used to further identify the subset of predictors that are most significantly related to the response, [238U]. First, principal component analysis was performed to identify how predictors are related to one another (Figure 6, Figure A3). The first PCA component captures 51.7% of the variability in the explanatory variables, and the second captures 38.7% (Figure A4). This means that the set of explanatory variables can be projected into two dimensions while capturing 90.4% of the variability in the data. One notable piece of information derived from PCA is variable loading, or how strongly predictors align with one PCA dimension. It is clear that pH is very highly loaded on the second dimension because it points nearly parallel to that axis. Salinity and total alkalinity are highly loaded on dimension one, though to a lesser extent than pH is loaded on dimension two. Finally, temperature and TCO₂ measure a very similar parameter, but are negatively correlated.

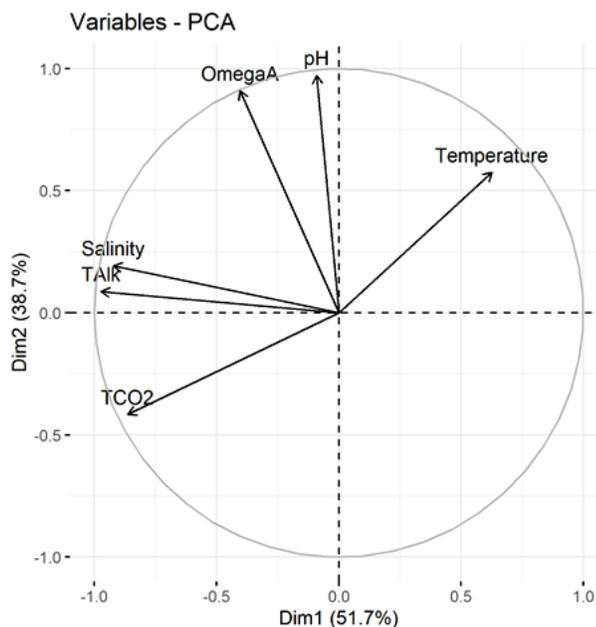


Figure 6. Visualization of predictors projected onto the first two PCA dimensions, which shows how predictors relate to each other in a reduced dimension space.

To perform feature selection, lasso models were fit from many bootstrapped samples of the data to account for instabilities in lasso estimates. The variables of interest in these analyses were temperature, salinity, pH, Ω, total alkalinity, and TCO₂. The data

was bootstrapped, and each bootstrap sample was fit with a lasso model to create a bootstrapped confidence interval for predictor coefficients (Figure 7).

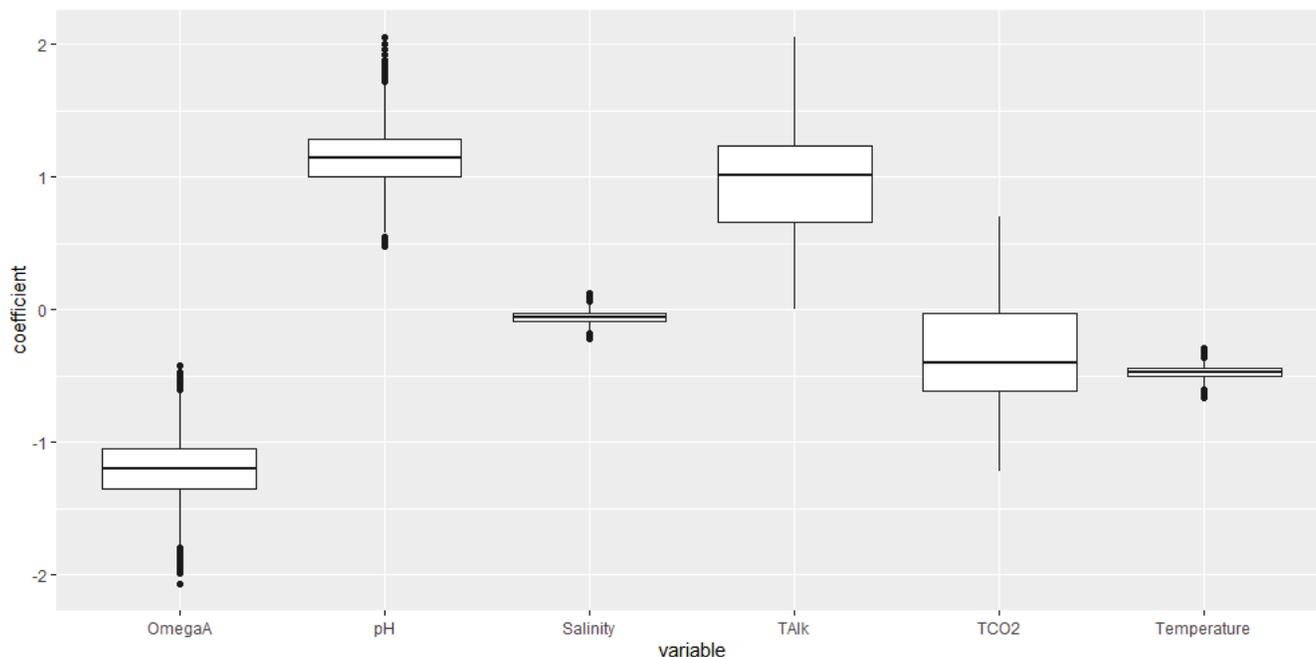


Figure 7. Box plots of slopes of standardized variables resulting from minimum penalty parameter lasso run on 5,000 bootstraps of the data. Lasso sets unimportant variables to zero, thus plots that are centered or overlapping with zero are considered unimportant, while those that do not overlap zero are selected.

It can be seen that when using the minimum error penalty parameter, Ω , pH, temperature, and alkalinity were always non-zero (that is, selected by the model). Best subset selection identified the same subset of predictors when using all possible combinations of predictors and using both AIC or BIC as the criterion. Thus, results indicate that Ω , pH, temperature, and alkalinity all control ^{238}U to some extent, and a final multiple linear regression model was built using all four as predictors for both genera of interest to quantify these relationships. This final model reveals that ^{238}U has negative linear relationships with temperature and Ω , and positive linear relationships with pH and alkalinity (Table 8). The negative linear relationship between temperature and ^{238}U is consistent with previous calibrations. While the relationship between pH and ^{238}U is not consistent with past studies, this is likely due to measuring the effect of pH while controlling Ω , which is very difficult to measure in practice as these variables are closely coupled. These relationships also likely differ by genus.

	Both Genera
Intercept	-84.511* (+/- 12.917)
pH	10.428* (+/- 1.529)
Temperature	-0.062* (+/- 0.008)
Ω	-1.125* (+/- .220)
Total Alkalinity	0.003* (+/- 0.001)

Table 8. Multiple linear regression with variables selected from lasso and best subset selection. *indicates significance at the 0.05 level. Primary results include coefficient estimates followed by standard errors.

DISCUSSION

This study reveals that using a large observational dataset of coral fossils yields results consistent with several existing, lab-based studies measuring temperature and salinity with smaller sample sizes. Specifically, this study confirms previous lab calibrations that describe the univariate relationship between ocean temperature and ^{238}U in fossil coral.^{3, 4, 18–24} After accounting for salinity, a significant negative relationship remains between temperature and ^{238}U . Additionally, after adjusting for temperature, a significant positive relationship was found between salinity and ^{238}U . Both relationships are consistent with the results of Ourbak *et al.* (2006).⁴ Interestingly, the relationship between pH and ^{238}U differed from past lab studies. Specifically, an overall positive relationship between pH and ^{238}U was found, inconsistent with results from lab studies such as Inoue *et al.* (2011).⁵ This discrepancy was found in the model for *Porites* alone and when modeling both genera together, but not in the model of only *Acropora*. However, after adjusting for temperature, there was no statistically significant relationship between pH and ^{238}U in the *Acropora* model. This may be reconciled in two ways. First, univariate relationships are generally insufficient to describe this observational data. Secondly, existing lab studies⁵ did not decouple pH, Ω , and other carbonate system variables, making it difficult to identify which variable might in reality be controlling U/Ca. This analysis reveals a negative relationship between Ω and ^{238}U , generally consistent with Inoue *et al.* (2011).⁵ Indeed, previous univariate model results showing pH as a good predictor of ^{238}U may be related to pH's high loading on PCA dimension 2. The observation that pH corresponds nearly exactly with a PCA dimension 2 like Ω and temperature. Thus, slopes reported for pH from previous experiments may indicate control by either Ω and/or temperature and not by pH itself.

Based on this work, univariate relationships are insufficient to fully describe the complex interactions of corals and seawater conditions. This result suggests that it would likely be challenging to apply a simple proxy for inferring past oceanic conditions as represented by coral ^{238}U , since many parameters control ^{238}U . However, if there is evidence that other variables remain constant in some regions or time periods, it could be possible to predict other variables from ^{238}U . Specifically, since the relationship between temperature and ^{238}U remained even after accounting for salinity, it is possible that ^{238}U could be used as a proxy for measuring temperature. The relationship between pH and ^{238}U does not appear as simple.

A major difference with this observational data compilation and lab studies is that lab studies often fix all experimental parameters besides the one or two they are studying. Furthermore, a study to determine which combination of many possible parameters best describe ^{238}U in corals has not been done. However, the difficulty is that in observational ocean data, and in many lab experiments, environmental parameters may covary. While it would be valuable to further test these results in additional laboratory experiments, this difficulty can be somewhat addressed with careful variable selection by statistically approximating the control of parameters using multiple linear regression, although when two parameters are highly correlated it is difficult to evaluate the effect of changes in one while holding the other constant. When including all parameters of interest in variable selection through bootstrap lasso and best subset selection, pH, Ω , alkalinity, and temperature are all significant predictors of ^{238}U . Genus also seems to play a key role in these relationships.

The authors acknowledge that the current study still suffers from limitations including the need to filter by genus (rather than by species) to include large enough sample sizes. Hence, there could be species-specific effects that are not captured. Additionally, a stable climate was assumed for the last 10,000 years, but this is an oversimplification. Previous work has highlighted regional changes in climate over this time period that are unaccounted for in this examination, but may be possible to include in future studies once regional Holocene climate changes are better constrained.^{31, 32} Furthermore, many of the environmental variables of interest are highly correlated, which can make regression results difficult to interpret. Finally, the range in the environmental variables is relatively small (*e.g.* pH values from 8.00 to 8.15), especially compared to lab calibrations. Though significant conclusions can still be drawn, values over larger data ranges, as can be obtained through further lab experiments, would help to more fully understand the relationships identified in this study.

CONCLUSIONS

In this study, a compilation of U-series measurements in tropical corals from a range of geographic locations was used in combination with environmental variables from two oceanographic databases to quantify relationships between coral ^{238}U and seawater temperature, salinity, aragonite saturation state, and pH. Univariate linear regressions and multiple linear regressions were used to compare relationships between U and environmental parameters. Though there is some fragility to multiple linear regression models on these data, results of variable selection analyses indicate that U is dependent on multiple environmental parameters and that previously developed univariate regressions may be insufficient to characterize the full range of variables that influence coral ^{238}U . In addition, relationships between ^{238}U and environmental variables vary by genus. Laboratory culture experiments, in which larger ranges of environmental variability can be explored, may prove useful in further testing the multivariate relationships found here and for identifying the physical, chemical, and biological mechanisms driving the dependences of coral uranium abundance on environmental change.

ACKNOWLEDGEMENTS

The authors thank Ella Hagopian (Class of 2020), Henry Henson (Class of 2020), Brendan Ireland (Class of 2019), Katie Murney (Class of 2019), and Matthew Richey (Professor of Mathematics, Statistics and Computer Science) from St. Olaf College for their contributions to the initial stages of this research. They also acknowledge the Center for Interdisciplinary Research at St. Olaf College and its Director, Katie Ziegler-Graham, for support.

REFERENCES

1. Druffel, E. (1997) Geochemistry of corals: Proxies of past ocean chemistry, ocean circulation, and climate. *Proc. Natl. Acad. Sci.* 94, 8354–8361. <https://doi.org/10.1073/pnas.94.16.8354>
2. Edwards, R.L., Gallup, C.D., Cheng, H. (2003) Uranium-series Dating of Marine and Lacustrine Carbonates. *Rev. Mineral. Geochem.* 52, 363–405. <https://doi.org/10.2113/0520363>
3. Correge, T., Delcroix, T., Recy, J., Beck, W., Cabioch, G., Le Cornec, F. (2000), Evidence for stronger El Nino-Southern Oscillation (ENSO) events in a mid-Holocene massive coral, *Paleoceanography* 15, 465–470. <https://doi.org/10.1029/1999PA000409>
4. Ourbak, T., Correge, T., Malaize, B., Le Cornec, F., Charlier, K., and Peypouquet, J.-P. (2006) A high-resolution investigation of temperature, salinity, and upwelling activity proxies in corals. *Geochem. Geophys. Geosyst.* 7, Q03013. <https://doi.org/10.1029/2005GC001064>
5. Inoue, M., Suwa, R., Suzuki, A., Sakai, K., and Kawahata, H., (2011). Effects of seawater pH on growth and skeletal U/Ca ratios of *Acropora digitifera* coral polyps. *Geophys. Res. Lett.* 38, 12801–12804. <https://doi.org/10.1029/2011GL047786>
6. Anagnostou, E., Sherrill, R. M., Gagnon, A. C., LaVigne, M., Field, M. P., and McDonough, W. F., (2011). Seawater nutrient and carbonate ion concentrations recorded as P/Ca, Ba/Ca, and U/Ca in the deep-sea coral *Desmophyllum dianthus*. *Geochim. Cosmochim. Acta* 75, 2529–2543. <https://doi.org/10.1016/j.gca.2011.02.019>
7. Rae, J.W.B., Foster, G.L., Schmidt, D.N., Elliot, T. (2011) Boron isotopes and B/Ca in benthic foraminifera: Proxies for the deep ocean carbonate system. *Earth. Planet. Sci. Lett.* 302, 403–413. <https://doi.org/10.1016/j.epsl.2010.12.034>
8. Nyadjro, E. S., & Subrahmanyam, B. (2016). Spatial and temporal variability of central Indian Ocean salinity fronts observed by SMOS. *Remote Sensing of Environment*, 180, 146–153. <https://doi.org/10.1016/j.rse.2016.02.049>
9. DeCarlo, T.M., Gaetani, G.A., Holcomb, M., Cohen, A.L., (2015). Experimental determination of factors controlling U/Ca of aragonite precipitated from seawater: Implications for interpreting coral skeleton. *Geochimica et Cosmochimica Acta* 162, 151–165. <https://doi.org/10.1016/j.gca.2015.04.016>
10. McConnaughey, T. (1989) ¹³C and ¹⁸O isotopic disequilibrium in biological carbonates: II. *In vitro* simulation of kinetic isotope effects. *Geochim. Cosmochim. Acta* 53, 163–171. [https://doi.org/10.1016/0016-7037\(89\)90282-2](https://doi.org/10.1016/0016-7037(89)90282-2)
11. Adkins J. F., Boyle E. A., Curry W. B. and Lutringer A. (2003) Stable isotopes in deep-sea corals and a new mechanism for “vital effects.” *Geochim. Cosmochim. Acta* 67, 1129–1143. [https://doi.org/10.1016/S0016-7037\(02\)01203-6](https://doi.org/10.1016/S0016-7037(02)01203-6)
12. Tambutté, E., Tambutté, S., Segonds, N., Zoccola, D., Venn, A., Erez, J. and Allemand, D. (2012) Calcein labelling and electrophysiology: insights on coral tissue permeability and calcification. *Proc. R. Soc. B Biol. Sci.* 279, 19–27. <https://doi.org/10.1098/rspb.2011.0733>
13. Comeau, S., Tambutte, E., Carpenter, R., Edmunds, P., Evensen, N., Allemand, C., Ferrier-Pages, C., Tambutte, S., Venn A. (2017) Coral calcifying fluid pH is modulated by seawater carbonate chemistry not solely seawater pH. *Proc. R. Soc. B Biol. Sci.* 284, 20161669. <https://doi.org/10.1098/rspb.2016.1669>
14. Al-Horani, F.A., Al-Moghrabi, S.M., de Beer, D. (2003) The mechanism of calcification and its relation to photosynthesis and respiration in the scleractinian coral *Galaxea fascicularis*. *Mar. Biol.* 142, 419–426. DOI 10.1007/s00227-002-0981-8
15. Zoccola D., Tambutté E., Kulhanek E., Puverel S., Scimeca J.-C., Allemand D. and Tambutté S. (2004) Molecular cloning and localization of a PMCA P-type calcium ATPase from the coral *Stylophora pistillata*. *Biochim. Biophys. Acta - Biomembr.* 1663, 117–126. <https://doi.org/10.1016/j.bbmem.2004.02.010>
16. Swart, P. K., and Hubbard, J. A. E. B., (1982). Uranium in Scleractinian Coral Skeletons. *Coral Reefs* 1, 13–19. <https://doi.org/10.1007/BF00286535>
17. Shen, G. T., and Dunbar, R. B., (1995). Environmental controls on uranium in reef corals. *Geochim. Cosmochim. Acta* 59, 2009–2024. [https://doi.org/10.1016/0016-7037\(95\)00123-9](https://doi.org/10.1016/0016-7037(95)00123-9)
18. Quinn, T.M. and Sampson, D.E. (2002) A multiproxy approach to reconstructing sea surface conditions using coral skeleton geochemistry, *Paleoceanogr. Paleoclimatol.* 17. doi:10.1029/2000PA000528
19. Armid, A., Asami, R., Fahmiati, T., Sheikh, M.A., Fujimura, H., Higuchi, T., Tair, E., Shinjo, R., Oomori, T. (2011) Seawater temperature proxies based on DSr, DMg, and DU from culture experiments using the branching coral *Porites cylindrica*. *Geochem. Cosmochim. Acta* 75, 4273–4285. <https://doi.org/10.1016/j.gca.2011.05.010>
20. Wei, G., Sun, M., Li, X., Nie, B. (2000) Mg/Ca, Sr/Ca, and U/Ca ratios of a porites coral from Sanya Bay, Hainan Island, South China Sea and their relationships to sea surface temperature. *Paleogeog., Paleoclim., Paleoecol.* 162, 59–74. [https://doi.org/10.1016/S0031-0182\(00\)00105-X](https://doi.org/10.1016/S0031-0182(00)00105-X)

21. Felis, T., Suzuki, A., Kuhnert, H., Dima, M., Lohmann, G., Kawahata, H. (2009) Subtropical coral reveals abrupt early-twentieth-century freshening in the western North Pacific Ocean. *Geology* 37, 527–530. <https://doi.org/10.1130/G25581A.1>
22. Fallon, S.J., McCulloch, M.T., van Woessik, R., Sinclair, D.J. (1999) Corals at their latitudinal limits: laser ablation trace element systematics in *Porites* from Shirigai Bay, Japan. *Earth Planet. Sci. Lett.* 172, 221–238. [https://doi.org/10.1016/S0012-821X\(99\)00200-9](https://doi.org/10.1016/S0012-821X(99)00200-9)
23. Sinclair, D.J., Kinsley, L.P.J., McCulloch, M.T. (1998) High resolution analysis of trace elements in corals by laser ablation ICP-MS. *Geochim. Cosmochim. Acta* 62, 1889–1901. [https://doi.org/10.1016/S0016-7037\(98\)00112-4](https://doi.org/10.1016/S0016-7037(98)00112-4)
24. Min, G. R., Edwards, R. L., Taylor, F. W., Recy, J., Gallup, C. D., and Beck, J. W., (1995). Annual cycles of U/Ca in coral skeletons and U/Ca thermometry. *Geochim. Cosmochim. Acta* 59, 2025–2042. [https://doi.org/10.1016/0016-7037\(95\)00124-7](https://doi.org/10.1016/0016-7037(95)00124-7)
25. Saenger, C.P., Evans, M.N. (2019) Calibration and Validation of Environmental Controls on Planktic Foraminifera Mg/Ca Using Global Core-Top Data. *Paleoceanogr. Paleoclimatol.* 34, 1249–1270. <https://doi.org/10.1029/2018PA003507>
26. Tierney, J.E., Malevich, S.B., Gray, W., Vetter, L., Thirumalai K. (2019) Bayesian Calibration of the Mg/Ca Paleothermometer in Planktic Foraminifera, *Paleoceanogr. Paleoclimatol.* 34, 2005–2030. <https://doi.org/10.1029/2019PA003744>
27. Robinson, L.F., Adkins, J.F., Frank, N., Gagnon, A.C., Prouty, N.G., Roark, E.B., van de Flierdt, T. (2014) The geochemistry of deep-sea coral skeletons: A review of vital effects and applications for paleoceanography. *Deep Sea Research Part I: Topical Studies in Oceanography* 99, 184–198. <https://doi.org/10.1016/j.dsr2.2013.06.005>
28. Chutcharavan, P. M., Dutton, A., Ellwood, M. J., (2018). Seawater ²³⁴U/²³⁸U recorded by modern and fossil corals. *Geochim. Cosmochim. Acta* 224, 1–17. <https://doi.org/10.1016/j.gca.2017.12.017>
29. Sun, Y., Sun M., Lee, T., Nie, B. (2005) Influence of seawater Sr content on coral Sr/Ca and Sr thermometry. *Coral Reefs* 24, 23–29. [doi 10.1007/s00338-004-0467-x](https://doi.org/10.1007/s00338-004-0467-x)
30. Allison, N., Finch, A.A., Sutton, S.R., Newville, M. (2001) Strontium heterogeneity and speciation in coral aragonite: Implications for the strontium paleothermometer. *Geochim. Cosmochim. Acta* 65, 2669–2676. [https://doi.org/10.1016/S0016-7037\(01\)00628-7](https://doi.org/10.1016/S0016-7037(01)00628-7)
31. Marcott, S.A., Shakun, J.D., Clark, P.U., Mix, A.C. (2013) A Reconstruction of Regional and Global Temperature for the Past 11,300 Years. *Science* 339, 1198–1201. [doi: 10.1126/science.1228026](https://doi.org/10.1126/science.1228026)
32. Rodriguez, L. G., Cohen, A. L., Ramirez, W., Oppo, D. W., Pourmand, A., Edwards, R. L., et al. (2019). Mid- Holocene, coral-based sea surface temperatures in the western tropical Atlantic. *Paleoceanogr. Paleoclimatol.*, 34, 1234–1245. <https://doi.org/10.1029/2019PA003571>
33. Locarnini, R. A., A. V. Mishonov, O. K. Baranova, T. P. Boyer, M. M. Zweng, H. E. Garcia, J. R. Reagan, D. Seidov, K. Weathers, C. R. Paver, and I. Smolyar, (2018). *World Ocean Atlas 2018, Volume 1: Temperature*. A. Mishonov Technical Ed.; NOAA Atlas NESDIS 81, 52 pp. <https://archimer.ifremer.fr/doc/00651/76338/>
34. Lauvset, S.K., Key, R.M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterstrom, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F., Watelet, S. (2016) A new global interior ocean mapped climatology: the 1 x 1 GLODAP version 2, *Earth Syst. Sci. Data* 8, 325–340. DOI 10.5194/essd-8-325-2016
35. Olsen, A., Key, R.M., van Heuven, S., Lauvset, S.K., Velo, A., Lin, Z., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterstrom, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F., Suzuki, T. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2) - an internally consistent data product for the world ocean *Earth Syst. Sci. Data* 8, 297–323. <https://doi.org/10.5194/essd-8-297-2016>
36. Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, and I. Smolyar, (2018) *World Ocean Atlas 2018, Volume 2: Salinity*. A. Mishonov Technical Ed.; NOAA Atlas NESDIS 82, 50 pp. <https://archimer.ifremer.fr/doc/00651/76339/>
37. Eddelbuettel D, François R (2011). Rcpp: Seamless R and C++ Integration. *J. Stat. Softw.* 40, 1–18. [doi: 10.18637/jss.v040.i08](https://doi.org/10.18637/jss.v040.i08).
38. Eddelbuettel D (2013). *Seamless R and C++ Integration with Rcpp*. Springer, New York. doi: 10.1007/978-1-4614-6868-4, ISBN 978-1-4614-6867-7.
39. Eddelbuettel D, Balamuta JJ (2017). “Extending extitR with extitC++: A Brief Introduction to extitRcpp.” *PeerJ Preprints*, 5, e3188v1. ISSN 2167-9843, [doi: 10.7287/peerj.preprints.3188v1](https://doi.org/10.7287/peerj.preprints.3188v1), <https://doi.org/10.7287/peerj.preprints.3188v1>.
40. Gothmann, A.M., Higgins, J.A., Adkins, J.F., Broecker, W., Farley, K.A., McKeon, R., Stolarski, J., Planavsky, N., Wang, X., Bender, M.L. (2019) A Cenozoic record of seawater uranium in fossil corals. *Geochim. Cosmochim. Acta* 250, 173–190. <https://doi.org/10.1016/j.gca.2019.01.039>
41. Rashid, R., A. Eisenhauer, P. Stocchi, V. Liebetrau, J. Fietzke, A. Ruggeberg, and € W.-C. Dullo (2014), Constraining mid to late Holocene relative sea level change in the southern equatorial Pacific Ocean relative to the Society Islands, French Polynesia, *Geochim. Geophys. Geophysyst.*, 15, 2601–2615, [doi:10.1002/2014GC005272](https://doi.org/10.1002/2014GC005272)
42. Cobb, K.M., Charles, C.D., Cheng, H., Kastner, M., Edwards, R.L. (2003) U/Th-dating living and young fossil corals from the central tropical Pacific. *Earth Planet. Sci. Lett.* 210, 91–103. [https://doi.org/10.1016/S0012-821X\(03\)00138-9](https://doi.org/10.1016/S0012-821X(03)00138-9)
43. Armid, A., Takaesu, Y., Fahmiati, T., Yoshida, S., Hanashiro, R., Fujimura, H., Higuchi, T., Taira, E., and Oomori, T., (2008). U/Ca as a possible proxy of carbonate system in coral reef in *Proceedings of the 11th International Coral Reef Symposium*, 92–96.

44. Inoue, M., Nakamura, T., Tanaka, Y., Suzuki, A., Yokoyama, Y., Kawahata, H., Sakai, K., Gussone, N. (2018) A simple role of coral-algal symbiosis in coral calcification based on multiple geochemical tracers. *Geochim. Cosmochim. Acta* 235, 76–88. <https://doi.org/10.1016/j.gca.2018.05.016>
45. Zou, H., Hastie, T. (2005) Regularization and variable selection via the elastic net. *J. R. Statist. Soc. B.* 67, 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
46. Efron B., Hastie, T. (2016) *Statistical Inference: Algorithms, Evidence, and Data Science*. 1st ed., 91-104, Cambridge University Press.
47. Hastie, T., Tibshirani, R., Friedman, J. (2008) *The Elements of Statistical Learning*, 2nd ed., 605-624.
48. Yuan, Z., Liu, D. Masque, P., Zhao, M., Song, X., Keesing, J.K. (2020) Phytoplankton responses to climate-induced warming and interdecadal oscillation in North-Western Australia. *Paleoceanogr. Paleoclimatol.*, 35, e2019PA003712. <https://doi.org/10.1029/2019PA003712>
49. Iwamori, H., K. Yoshida, H. Nakamura, T. Kuwatani, M. Hamada, S. Haraguchi, and K. Ueki (2017), Classification of geochemical data based on multivariate statistical analyses: Complementary roles of cluster, principal component, and independent component analyses, *Geochem. Geophys. Geosyst.*, 18, 994– 1012, [doi:10.1002/2016GC006663](https://doi.org/10.1002/2016GC006663)
50. Lough, J.M. (2004) A strategy to improve the contribution of coral data to high-resolution paleoclimatology. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 204, 115-143, [https://doi.org/10.1016/S0031-0182\(03\)00727-2](https://doi.org/10.1016/S0031-0182(03)00727-2)

ABOUT STUDENT AUTHORS

Emily Patterson recently graduated from St. Olaf College in May, 2020, with a B.A. degree in Psychology and a concentration in Statistics and Data Science. She is now working as an Associate Data Scientist at the Allant Group. In the future, she plans to pursue a career in data science with a focus on the fields of healthcare or environmental science.

Spencer Eanes graduated from St. Olaf College in May, 2020, with a B.A. degree in Mathematics and Computer Science, and a concentration in Statistics and Data Science. He is now working as a Software Developer at Epic Systems, where he develops a telemedicine video conferencing solution.

Penelope Lancrete graduated from St. Olaf College in May, 2020, with a B.A. degree in Mathematics and Ancient Studies, and a concentration in Statistics and Data Science. She is now working as a Catastrophe Risk Analyst at TigerRisk Partners.

PRESS SUMMARY

An understanding of past environmental conditions provides context for evaluating modern climate changes, but there is a lack of data for periods before the existence of instrument-based environmental records. To reconstruct conditions hundreds to millions of years into the past, proxies are needed - indirect indicators of past conditions measured from data or samples that are currently accessible. Coral skeletons are one good example of an environmental proxy. They are globally distributed, can be reliably dated, and have potential to record high-resolution climate information through their chemical composition and physical structure.

Recently, the uranium/calcium ratio in fossil coral skeletons has been investigated as a proxy for past ocean pH, temperature, salinity, and carbonate ion concentration. However, existing calibrations are either lab-based or were developed using small numbers ($n \sim 15$) of natural samples. This study explores whether existing calibrations can describe the relationship between uranium concentrations and seawater environmental conditions, such as seawater pH, temperature, salinity, and carbonate ion concentration for a large observational dataset of 700 fossil coral samples. Results suggest that no single variable most strongly controls U/Ca, but rather that multiple environmental variables are needed to capture it. Specifically, pH, Ω , alkalinity, and temperature are all significant predictors of U/Ca. Genus also plays a key role in these relationships. These results provide further insight into whether the fossil coral uranium proxy can be used to accurately and precisely reconstruct past ocean conditions.