

## Summary

We created three synthetic ‘pseudoproxies’ ( $\text{Sr}/\text{Ca}$ ,  $\delta^{18}\text{O}$ ,  $\delta^{11}\text{B}$ ) with various amounts of environmental information encoded into each, based on their theoretical dependence on SST, SSS and  $\text{pH}_{\text{sw}}$  (Fig. 2). Because we created this dataset with three coral variables and three potential climate targets, this is considered a square system (i.e., the number of predictor variables equals the number of unknown variables), which is important when considering regularization (see Sections 3.3.2 and 4.2). The means by which we calculated these three synthetic pseudoproxies are highly idealized, meaning that each pseudoproxy has a near-perfect relationship with its corresponding climate target(s). However, the uncertainty in each proxy’s relationship to the climate target is considered in our experimental design, which examines how uncertainty in SMITE-reconstructed SST and  $\text{pH}_{\text{sw}}$  estimates increases as the degree of Gaussian and autocorrelated noise increases. The magnitude of Gaussian noise and autocorrelation considered in our experiment (Table 1) is beyond the typical range observed in most coral-based paleoclimate studies (Hathorne et al., 2013; Jones et al., 2015; Standish et al., 2019). We expect therefore that all sources of uncertainty, and their subsequent impacts on SST and  $\text{pH}_{\text{sw}}$  estimates, are accounted for in this conservative analysis. The minimum uncertainty for each synthetic pseudoproxy was taken from the literature as analytical uncertainty. For synthetic  $\text{Sr}/\text{Ca}$  values, this was taken to be 0.009 mmol/mol, or approximately 0.1% RSD (Schrag, 1999). For synthetic  $\delta^{18}\text{O}$  values, we used an analytical uncertainty of 0.1‰ (Epstein & Mayeda, 1953). For synthetic  $\delta^{11}\text{B}$  values, analytical uncertainty was taken to be 0.09‰ (Stewart et al., 2021).

***Fig 1. Monthly time series of three synthetic pseudoproxies (B) calculated from environmental information spanning the 20<sup>th</sup> century off the east coast of Australia (A) (de Boissésón et al., 2018; Lenton et al., 2016). The gray shaded region in each panel, which is barely visible, indicates the minimum (analytical) uncertainty associated***

with each pseudoproxy. All three pseudoproxies exhibit an idealized relationship with their corresponding environmental variable(s) of interest ( $\text{Sr/Ca} \sim \text{SST}$ ;  $\delta^{18}\text{O} \sim \text{SST} + \text{SSS}$ ;  $\delta^{11}\text{B} \sim \text{SST} + \text{SSS} + \text{pH}$ ).

Table 1. Mean ( $\mu$ ), standard deviation ( $\sigma$ ), range of analytical errors ( $\epsilon$ ), and correlation coefficients ( $r$ ) for all synthetic proxies to each environmental variable.

Synthetic Proxy	$\mu$	$\sigma$	$\epsilon_{\min} - \epsilon_{\max}$ ( $\text{RSD}_{\min - \max}$ )	$r - \text{SST}$	$r - \text{SSS}$	$r - \text{pH}_{\text{sw}}$
Sr/Ca (mmol/mol)	8.98	0.12	0.009–0.180 (0.1–1.9%)	–1.00	0.68	0.54
$\delta^{18}\text{O}$ (‰)	–5.49	0.51	0.10–0.21 (1.8–3.7%)	–0.98	0.80	0.50
$\delta^{11}\text{B}$ (‰)	22.90	0.66	0.18–0.62 (0.78–2.68%)	–0.92	0.63	0.82

Monthly SST and  $\text{pH}_{\text{sw}}$  data from the Great Barrier Reef (18.5°S, 149.5°E) between 1900 and 2000 were acquired from Lenton et al. (2016), a 20<sup>th</sup> century reconstruction of SST, SSS and pH across the Great Barrier Reef ( $n = 1212$ ). SSTs ranged from 21.83°C to 29.69°C ( $\mu = 25.92^\circ\text{C} \pm 1.91, 1\sigma$ ), with a minor but significant warming trend of 0.08°C per decade ( $p < 0.0001$ ). Seawater pH ranged from 8.09 to 8.21 ( $\mu = 8.16 \pm 0.03$ ) and exhibits two significant negative trends pre-1950 (0.004 units per decade) and post-1950 (0.014 units per decade). Although SSS data are also available from this dataset for the same time interval, these data simply repeat the same annual cycle of SSS throughout the 20<sup>th</sup> century with no interannual or decadal variability. To better reproduce long-term changes in SSS, we used SSS data from the ORA20C dataset (de Boissésou et al., 2018) from the same location and time interval. This dataset extends back through the 20<sup>th</sup> century and is an advanced data assimilation product that tunes the output of the

European Center for Medium-range Weather Forecasts twentieth century reanalysis, ERA-20C, to *in situ* observations. SSS variations from this location in the ORA20C dataset exhibit a highly skewed left distribution ( $\mu = 33.69 \pm 0.43$  psu), with values ranging from 31.61 to 34.17 due to episodic freshwater runoff events.

According to hindcast archived data from the CSIRO Environmental Modelling Suite implemented by the Australian Institute of Marine Science (<https://research.csiro.au/cem/software/ems/>), SSS in this region between 2010 and 2022 ranged from 34.7 to 35.6 psu and exhibit a slightly skewed left distribution ( $\mu = 35.23 \pm 0.16$  psu). We acknowledge that the distributions of ORA20C and CSIRO SSS are statistically distinct from one another, both in terms of mean and variance ( $p < 0.001$ ). However, the purpose of including SSS in the synthetic experiment is to create interference in both synthetic  $\delta^{18}\text{O}$  and  $\delta^{11}\text{B}$  values for reconstructing SST and  $\text{pH}_{\text{sw}}$ , respectively. Thus, there are two important aspects of SSS that we wish to reproduce for the purposes of this experiment: long-term variability (interannual to decadal), and the covariance between SSS and SST. The ORA20C SSS dataset for this region exhibits substantial interannual and decadal-scale variability, while the Lenton et al. (2016) SSS dataset exhibits none. With respect to covariance, the CSIRO dataset shows that SST and SSS are moderately anti-correlated in this region ( $r = -0.58$ ). SST and SSS data from Lenton et al. (2016) exhibit a slightly weaker anticorrelation ( $r = -0.53$ ), while SST data from Lenton et al. (2016) and ORA20C SSS exhibit a slightly stronger anticorrelation ( $r = -0.68$ ). Since ORA20C SSS exhibits both long-term variability and similar covariance to SST as the observed CSIRO data, we chose to use the ORA20C SSS dataset for our synthetic experiment.

Synthetic Sr/Ca ratios were calculated as a function of SST using the mean slopes and intercepts for the Sr/Ca  $\sim$  SST relationship from Corrège (2006).

$$\frac{Sr}{Ca_c} = -0.0607(0.0090)T + 10.553(0.292) \quad (8)$$

Where  $T$  is temperature in degrees Celsius. Synthetic Sr/Ca ( $Sr/Ca_c$ ) ratios ranged from 8.75 to 9.23 mmol/mol ( $\mu = 8.98 \text{ mmol/mol} \pm 0.12$ ).

Synthetic  $\delta^{18}\text{O}$  values were calculated as a function of both SST and SSS using equation 1 from Thompson et al. (2011).

$$\delta^{18}\text{O}_c = -0.22\text{SST} + 0.27\text{SSS} \quad (9)$$

The regression slopes for SST and SSS were chosen using the same criteria from Thompson et al. (2011). The SST slope is the organic slope of the  $\delta^{18}\text{O}$  and SST relationship, while the SSS slope is based on basin-scale seawater  $\delta^{18}\text{O}$  and SSS regression estimates (LeGrande & Schmidt, 2006). Synthetic  $\delta^{18}\text{O}$  values ranged from  $-6.74$  to  $-4.52\text{‰}$  ( $\mu = -5.49 \pm 0.51\text{‰}$ ).

Synthetic  $\delta^{11}\text{B}$  values were calculated as a function of SST, SSS and  $\text{pH}_{\text{cf}}$ . They were determined by rearranging the pH-dependent equation from Zeebe and Wolf-Gladrow (2001) to solve for the boron isotope ratio of carbonate ( $\delta^{11}\text{B}_c$ ).

$$\text{pH} = \text{p}K_b - \log\left(\frac{\delta^{11}\text{B}_{\text{sw}} - \delta^{11}\text{B}_c}{\delta^{11}\text{B}_c - \delta^{11}\text{B}_{\text{sw}} + 1000(\alpha - 1)}\right) \quad (10)$$

Where the boron isotope ratio of seawater ( $\delta^{11}\text{B}_{\text{sw}}$ ) is  $39.61\text{‰}$  (Foster et al., 2010), and the mass fraction factor between boric acid and borate ion ( $\alpha$ ) is 1.0272 (Klochko et al., 2006). The negative log of the dissociation constant between boric acid and borate ion ( $\text{p}K_b$ ) is a function of both temperature and salinity (Dickson, 1990). We therefore calculated  $\text{p}K_b$  at each time interval by taking the negative log of the  $K_b$  equation from Dickson (1990). The values of  $\text{p}K_b$  ranged

between 8.56 and 8.64 given a temperature range between 21.83 and 29.69°C and a salinity range between 31.61 and 34.17 psu.

These calculations yield synthetic  $\delta^{11}\text{B}$  values between 18.30 and 19.64‰, which is expected given the pH of seawater. However, corals upregulate their internal pH relative to seawater (McCulloch et al., 2017) while also often exhibiting increased seasonal variance (Ross et al., 2017). Thus, to yield synthetic  $\delta^{11}\text{B}$  values consistent with those observed in coral aragonite, we calculated  $\text{pH}_{\text{cf}}$  from  $\text{pH}_{\text{sw}}$  using equation 13 from D’Olivo et al. (2019).

$$\text{pH}_{\text{cf}} = 0.49\text{pH}_{\text{sw}} + 4.93 - 0.02T \quad (11)$$

Note that the temperature sensitivity of synthetic  $\delta^{11}\text{B}$  values is realized in its dependence on both  $\text{pKb}$  as well as  $\text{pH}_{\text{cf}}$  as specified in equation 11. Meanwhile, the salinity sensitivity of synthetic  $\delta^{11}\text{B}$  values is only realized in its dependence on  $\text{pKb}$ . Synthetic  $\delta^{11}\text{B}$  values ranged from 22.21 - 24.10‰ ( $\mu = 23.17\text{‰} \pm 0.41$ ).

## Error Assessments

We use three metrics to quantitatively compare SMITE SST and  $\text{pH}_{\text{sw}}$  estimates with those derived from Sr/Ca ratios and  $\delta^{11}\text{B}$  values, respectively: the correlation coefficient ( $r$ ), the root-mean squared error (RMSE), and the standard error of prediction (SEP). Each metric provides a measure of the correlation, accuracy, and precision of the reconstruction, respectively. The SEP is defined as the uncertainty in derived SST estimates based on the uncertainty in both the climate target (SST,  $\text{pH}_{\text{sw}}$ ) as well as the uncertainty in the corresponding coral variable(s). Given that our SST measurements are derived from, or modeled after, temperatures derived from *in situ* loggers, uncertainty for temperature was fixed at 0.02°C (<https://www.onsetcomp.com/products/data-loggers/u22-001>). Uncertainty for our  $\text{pH}_{\text{sw}}$

measurements were fixed at 0.02 units. Uncertainty for SSS measurements (used in calculating  $pK_b$  for  $\delta^{11}\text{B}$ -derived  $\text{pH}_{\text{sw}}$  estimates) were based on BATS CTD measurements and also fixed at 0.02 psu.

The SEP for each climate target reconstruction is calculated using a bootstrap Monte Carlo approach. At each iteration ( $i = 1, \dots, 10000$ ), each individual measurement in both the coral variable and climate target fields are randomly resampled from a normal distribution with a mean equal to the given variable/climate target value ( $\mu_i$ ) and a standard deviation equal to the specified error ( $s_i$ ). Model parameters are then estimated from the perturbed coral variable and climate target fields, and SST/ $\text{pH}_{\text{sw}}$  estimates for each data point are stored. The 95% confidence interval for each predicted value is determined from the distribution of predicted values derived from each Monte Carlo iteration. The SEP is then determined as the average distance from the mean to the upper and lower bounds of the 95% confidence interval, divided by 1.96 (73). The 95% confidence interval for the SEP itself is then defined as the standard deviation of the SEP throughout each calibration dataset, multiplied by 1.96.

### Autocorrelated (Red) Noise

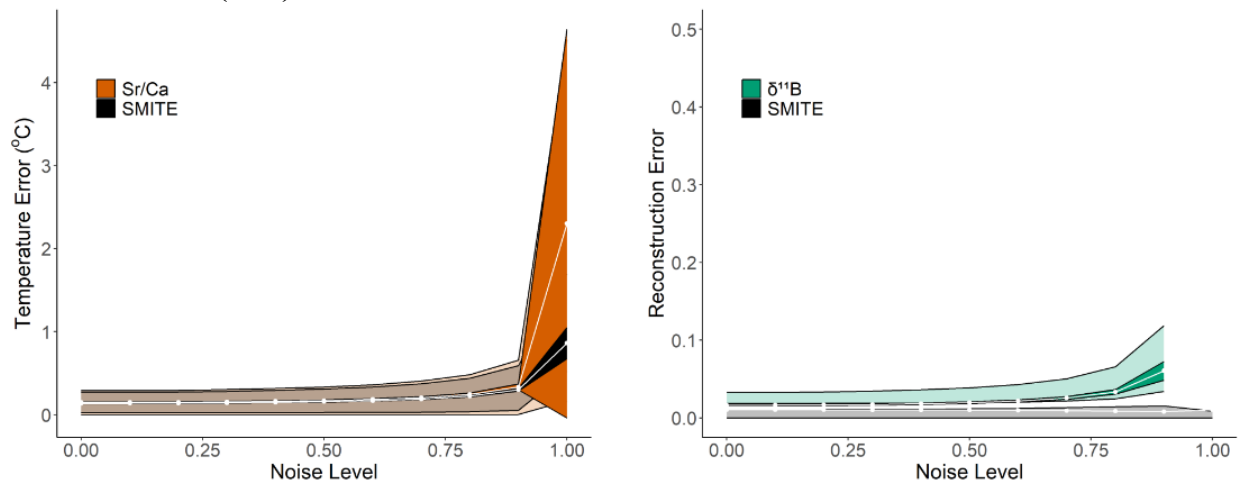
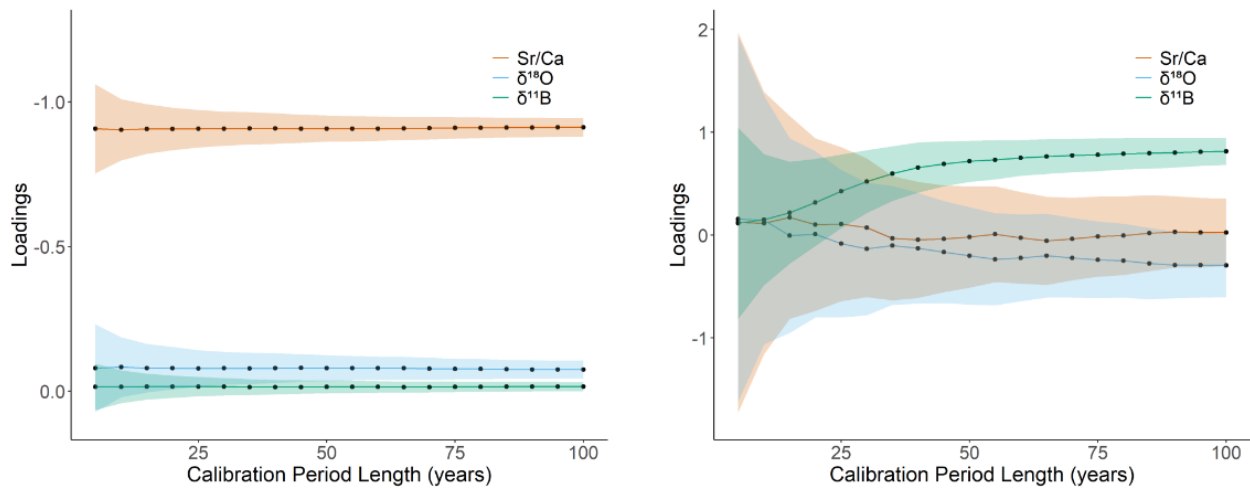


Figure 2. The precision (SEP; opaque envelope) and accuracy (RMSE; translucent envelope) of

*SMITE SST estimates (black) versus Sr/Ca-derived SST (left; orange) and  $\delta^{11}\text{B}$ -derived pH (right; green) over increasingly autocorrelated errors (fixed at analytical uncertainty).*

*Figure 3. SMITE model parameters, or loadings, for SST (left) and pH (right) over increasingly autocorrelated errors (fixed at analytical uncertainty). Shaded regions represent the Monte Carlo estimated 95% confidence interval for each parameter ( $i = 10,000$ ).*

### Calibration Period Length



*Figure 4. SMITE model parameters, or loadings, for SST (left) and pH (right) over different calibration period lengths. Shaded regions represent the Monte Carlo estimated 95% confidence interval for each parameter ( $i = 10,000$ ).*

## References

- de Boissésou, E., Balmaseda, M. A., & Mayer, M. (2018). Ocean heat content variability in an ensemble of twentieth century ocean reanalyses. *Clim. Dynam.*, 50(9), 3783-3798. <https://doi.org/10.1007/s00382-017-3845-0>
- Dickson, A. G. (1990). Thermodynamics of the dissociation of boric acid in synthetic seawater from 273.15 to 318.15 K. *Deep Sea Res. Part A*, 37(5), 755-766. [https://doi.org/10.1016/0198-0149\(90\)90004-F](https://doi.org/10.1016/0198-0149(90)90004-F)
- Epstein, S., & Mayeda, T. (1953). Variation of O18 content of waters from natural sources. *Geochim. Cosmochim. Acta*, 4(5), 213-224. [https://doi.org/10.1016/0016-7037\(53\)90051-9](https://doi.org/10.1016/0016-7037(53)90051-9)
- Foster, G. L., Pogge von Strandmann, P. A. E., & Rae, J. W. B. (2010). Boron and magnesium isotopic composition of seawater. *Geochem. Geophys. Geosyst.*, 11. <https://doi.org/10.1029/2010GC003201>
- Hathorne, E. C., Gagnon, A., Felis, T., Adkins, J., Asami, R., Boer, W., Caillon, N., Case, D., Cobb, K. M., Douville, E., DeMenocal, P., Eisenhauer, A., Garbe-Schönberg, D., Geibert, W., Goldstein, S., Hughen, K., Inoue, M., Kawahata, H., Kölling, M., . . . You, C.-F. (2013). Interlaboratory study for coral Sr/Ca and other element/Ca ratio measurements. *Geochem. Geophys. Geosyst.*, 14, 3730-3750. <https://doi.org/10.1002/ggge.20230>
- Jones, J. P., Carricart-Ganivet, J. P., Iglesias Prieto, R., Enríquez, S., Ackerson, M., & Gabitov, R. I. (2015). Microstructural variation in oxygen isotopes and elemental calcium ratios in the coral skeleton of *Orbicella annularis*. *Chem. Geol.*, 419, 192-199. <https://doi.org/10.1016/j.chemgeo.2015.10.044>
- Klochko, K., Kaufman, A. J., Yao, W., Byrne, R. H., & Tossell, J. A. (2006). Experimental measurement of boron isotope fractionation in seawater. *Earth Planet. Sc. Lett.*, 248, 276-285. <https://doi.org/10.1016/j.epsl.2006.05.034>
- LeGrande, A. N., & Schmidt, G. A. (2006). Global gridded data set of the oxygen isotopic composition in seawater. *Geophys. Res. Lett.*, 33(12). <https://doi.org/10.1029/2006GL026011>
- Lenton, A., Tilbrook, B., Matear, R. J., Sasse, T. P., & Nojiri, Y. (2016). Historical reconstruction of ocean acidification in the Australian region. *Biogeosciences*, 13, 1753-1765. <https://doi.org/10.5194/bg-13-1753-2016>
- McCulloch, M. T., D'Olivo, J. P., Falter, J., Holcomb, M., & Trotter, J. A. (2017). Coral calcification in a changing World and the interactive dynamics of pH and DIC upregulation. *Nat. Commun.*, 8, 15686. <https://doi.org/10.1038/ncomms15686>
- Ross, C. L., Falter, J. L., & McCulloch, M. T. (2017). Active modulation of the calcifying fluid carbonate chemistry ( $\delta^{11}\text{B}$ , B/Ca) and seasonally invariant coral calcification at sub-tropical limits. *Sci. Rep.*, 7, 1-11. <https://doi.org/10.1038/s41598-017-14066-9>
- Schrag, D. P. (1999). Rapid analysis of high - precision Sr/Ca ratios in corals and other marine carbonates. *Paleoceanography*, 14, 97-102. <https://doi.org/10.1029/1998PA900025>



- Standish, C. D., Chalk, T. B., Babila, T. L., Milton, J. A., Palmer, M. R., & Foster, G. L. (2019). The effect of matrix interferences on in situ boron isotope analysis by laser ablation multi-collector inductively coupled plasma mass spectrometry. *Rapid Commun. Mass. Sp.*, 33(10), 959-968. <https://doi.org/10.1002/rcm.8432>
- Stewart, J. A., Christopher, S. J., Kucklick, J. R., Bordier, L., Chalk, T. B., Dapoigny, A., Douville, E., Foster, G. L., Gray, W. R., Greenop, R., Gutjahr, M., Hemsing, F., Henehan, M. J., Holdship, P., Hsieh, Y.-T., Klevica, A., Lin, Y.-P., Mawbey, E. M., Rae, J. W. B., . . . Day, R. D. (2021). NIST RM 8301 Boron Isotopes in Marine Carbonate (Simulated Coral and Foraminifera Solutions): Inter-laboratory  $\delta^{11}\text{B}$  and Trace Element Ratio Value Assignment. *Geostand. Geoanal. Res.*, 45(1), 77-96. <https://doi.org/10.1111/ggr.12363>