A Multicentennial Proxy Record of Northeast Pacific Sea Surface Temperatures From the Annual Growth Increments of *Panopea generosa*

David C. Edge, David J. Reynolds, Alan D. Wanamaker, Daniel Griffin, Dominique Bureau, Christine Outridge, Bethany C. Stevick, Richard Weng, and Bryan A. Black

Supporting Information: Supporting Information may be found in the online version of this article.

Abstract  Growth-increment widths of Pacific geoduck (*Panopea generosa*), a long-lived bivalve, are used to develop the first marine-based, multicentennial, annually resolved, and exactly dated archive of Northeast Pacific sea surface temperatures (SST). The chronology is sampled from the Tree Nob Islands, British Columbia, Canada, continuously covers 1725–2008, and also contains nine older radiocarbon-dated segments, which together span 58% of the last 1,500 years. Age-related growth declines were removed by aligning all increments relative to age of increment formation and fitting with a single detrending curve to preserve low-frequency signals. The geoduck chronology was used to reconstruct local SST variability across the seasonal window of April through November. The chronology at both the concurrent (lag-0) and following (lag+1) year are both highly significant predictors of SST in a stepwise multiple linear regression, explaining 54% of the variance in the period of instrumental overlap (1940–2001), passing strict tests of calibration-verification. Reconstructed SSTs contained significant spectral power at periods from 3 to 64 years, suggesting that 20th century variability in these periodicities is not unusual in the longer-term context. The period of lowest growth coincided with the Dalton minimum, an episode of reduced solar irradiance from 1790–1830, as well as the 1809 Unknown eruption, suggesting that solar and volcanic signals are present in the SST history. The most conspicuous aspect of the reconstruction is the steady and unprecedented warming trend that began in the mid-1800s and continues through present. The post-1976 interval includes the two warmest decades of the reconstruction.

1. Introduction

In terrestrial environments, tree-ring data are a critically important indicator of long-term climate and environmental variability, especially at mid latitudes. A distinguishing characteristic of tree-ring data amongst climate proxies is that they are well replicated, annually resolved, and absolutely dated through the process of crossdating, in which synchronous patterns induced by climate are matched among individuals of a given species and location (Douglass, 1941; Fritts, 1971; Glock & Pearson, 1937). By this method, growth irregularities such as false rings, micro-rings, or locally absent rings can be identified such that each increment in the data set is assigned its correct year of formation (Fritts, 1976). For live-collected samples, the year of collection anchors the absolute dating of the chronology. Where available, dead-collected material of unknown antiquity can be crossdated with one another and the live collected record. In doing so, chronologies can be generated that are much longer than the average individual lifespan for the species, in some cases spanning multiple millennia (e.g., Ferguson & Graybill, 1983; Pilcher et al., 1984). Such exact dating facilitates seamless integration of chronologies with one another and instrumental records (Briffa et al., 1996; Mann & Jones, 2003).

Over the past two decades, this same crossdating technique has been applied to growth increments formed in the hard parts of marine organisms to reconstruct environmental variability prior to the start of instrumental records (Black et al., 2019). This approach has been especially successful with long-lived bivalves *Arctica islandica* and *Glycymeris glycymeris* in the North Atlantic, resulting in a network of continuous,
multicentennial chronologies from the Gulf of Maine through northern Norway (Butler et al., 2009, 2010; Reynolds et al., 2017; Wanamaker et al., 2012, 2019). A chronology near Iceland spans more than a thousand years and provides exceptionally long-term perspectives of North Atlantic marine dynamics (Butler et al., 2013; Reynolds et al., 2016). In the Northeast Pacific (NEP), a network of crossdated growth-increment width chronologies with strong sensitivity to sea surface temperature (SST) has been developed from the long-lived bivalve geoduck (*Panopea generosa*; Black et al., 2009; Strom et al., 2004). However, these chronologies have involved only live-collected material and the temporal extent is therefore limited by the maximum longevity of the species, which is ~150 years.

Although live-collected geoduck often doubles the length of the observational record in the coastal NEP (Black et al., 2009), longer-term perspectives on SST are needed. NEP SSTs vary at low frequencies and are characterized by rapid regime shifts associated with the restructuring of marine ecosystem species composition, food web structure, and energy flows (Chavez et al., 2003; Clark et al., 1999; Hare et al., 1999) while also synchronous with drought, snowpack, and fire frequency in western North America (Berkelhammer, 2019; Kitzberger et al., 2007; Mote, 2006). On even broader spatial scales, NEP SSTs have linkages to variability in global-scale temperature patterns (England et al., 2014; Meehl et al., 2011, 2013; Thompson et al., 2015; Yin et al., 2018). Given the importance of NEP SSTs, there have been numerous attempts to reconstruct its longer-term history, largely from tree-ring data, but these histories poorly agree with one another prior to the twentieth century, leaving considerable uncertainty regarding pre-industrial ranges of climate variability (Henley, 2017; Kipfmüller et al., 2012; Newman et al., 2016).

Here, we use growth-increment widths of live- and dead-collected Pacific geoduck (*P. generosa*) to construct the first marine-based, multicentennial, annually resolved, and exactly dated archive of SST in the NEP. The crossdated chronology continuously spans 1725–2008 and also contains well-replicated, radiocarbon-dated segments that in combination span 58% of the last 1,500 years. We explore ways to preserve not only the high-frequency variability inherent in the crossdated chronologies, but also use the robust replication and temporal depth of the sample population to preserve low-frequency variability. In so doing, we develop a reconstruction of regional SSTs along the British Columbia, Canada coast to address pre-1900 SST variability and evaluate the magnitude of 19th and 20th century warming.

### 2. Materials and Methods

#### 2.1. Sample Collection and Preparation

A Pacific geoduck growth-increment width chronology had been previously developed using live-collected samples from the Tree Nob Islands in northern British Columbia (Figure 1; Black et al., 2009). Given the consistently great longevity of live-collected samples at this site, strong relationship with SST, and presence of dead shell on the ocean floor, Tree Nob was chosen as an ideal site for developing an extended chronology with dead-collected material. On June 26, 2018, professional divers from Trident Biologicals, Nanaimo, BC, Canada, excavated a pit ~2 m wide and 2 m deep at ~10 m depth at 54°13'2.53"N, 130°47'21.59"W (Figure 1).

A Venturi device was employed to move large quantities of sediment while geoduck shells were manually collected. Due to the instability of the bottom sediments, shells often slid from their original locations before they could be collected, complicating efforts to maintain stratigraphic order of the samples. Approximately 500 shells were gathered, occurring in various taphonomic states from highly preserved with a complete periostracum to heavily eroded with a relatively friable, “chalky” shell texture. In a small percentage (<5%) valves remained attached by an intact ligament, but the vast majority were single valves assumed to be unique individuals.

The length (posterior to anterior), height (dorsal to ventral), width (shell margin to outer shell apex), and weight (mass in grams) of each valve was measured. We also used a subjective 5-level classification system for recording the general states of bioerosion, shell margin deterioration, and periostracum of each shell as potential proxy measures of shell antiquity (Butler et al., 2020).

The methods for shell preparation and acetate peel preparation were adapted from Richardson (2001). A 2 cm transverse section containing the midline was cut from the dorsal to ventral margins of the valve. The valve section was then embedded in two-part EpoxyCure from Buhler within a silicone mold. The embedded sample was cut along the midline of the umbo (Scourse et al., 2006, Figure 2) with an IsoMet 5000 precision
saw. One of the resulting halves was polished on an EcoMet 30 grinding wheel with 600-, 800-, 1000-, 1400-grit paper followed by 3-micron diamond paste. The polished sample was etched by bathing in 0.1 molar HCl for 135 s and immediately rinsed in tap water. Once completely dry, the minimum amount of acetone required to coat the surface was applied. A 125 µm acetate sheet was immediately placed over the acetone. After 25 min, the acetate sheet, partially dissolved by the acetone into the etched shell surface, was peeled away and placed between two glass microscope slides (Figure 1b, additional details in Figure S1).

Samples were imaged at 80-200X using a Leica M205 C stereo microscope and a Leica DMC 5400 20-megapixel camera. Increment widths were measured using the caliper tool in Image Pro Premier 9.3 (Media Cybernetics). A polyline was drawn along the curved axis of maximum growth, perpendicular to each increment. Increment boundaries were manually marked at the outer edge of the winter line (dark bands in Figure 1b). Each individual was measured twice along separate axes of measurement. Although these axes were by necessity in close proximity to one another, the repeated measurements helped reduce error due to placement of increment boundaries on the winter line, which could occasionally be diffuse, especially in very narrow increments. All measurements were exported with micron precision.

2.2. Crossdating

The existing absolutely dated Tree Nob chronology (TNC) was used as a starting point for our dating (Black et al., 2009). Dead collected samples were crossdated against the live-collected chronology and then other dead-collected samples using a general “skeleton plotting” approach. In skeleton plotting, the high-frequency (year-to-year) growth pattern in the sample is transferred to graph paper. Then skeleton plots are manually slid past one another to search for matches (Stokes & Smiley, 1968). Given the many pairwise comparisons possible, this can be a very time intensive procedure. To greatly expedite this process, a cross-correlation procedure analogous to skeleton plotting was performed in RingdateR (Figure S2; Reynolds et al., 2021), which extracts high-frequency (year-to-year) variability from each set of measurements and then explores all pairwise combinations and lags for possible matches. Because RingdateR automatically recommends matches from datasets involving many measurement time series, the time spent crossdating is primarily in confirming matches and revising missed/false rings. Time series plots and heat maps of running correlations can be used to help identify any missing or falsely added rings that can be visually confirmed by re-inspecting the sample (see Reynolds et al., 2021, Figure 2b). Ultimately, we employed no statistical threshold for crossdating, rather, the suggested crossdating was confirmed by meeting three criteria: (a) one and only
one lag produced a correlation far greater than all other possible lags, (b) running correlations were highly positive and stable throughout the period of overlap, and (c) visual inspection of the sample identified no errors in growth-increment interpretation or major distortions in growth.

Samples were initially added to the live-collected TNC or compared against one another to form chronologies “floating” in time. Samples that remained undated were regularly reexamined with the addition of new series. Crossdated series were checked for errors in RingdateR, and crossdating was additionally verified in Cofecha (Holmes, 1983). The default spline length of 32 years was reduced to 22 years to improve the common signal amongst the standardized series (Black et al., 2008). Mean sensitivity, a coarse indicator of year-to-year variability (Bunn et al., 2013; Strackee & Jansma, 1992), as well as series intercorrelation was calculated for the absolutely dated and floating chronologies using the Dendrochronology Program Library in R (dplR; Bunn, 2008).

2.3. Radiocarbon Dating

Radiocarbon sampling and analysis was performed to date floating chronologies with replication greater than two individuals. Fourteen carbonate samples were collected for radiocarbon dating. Samples containing ~10 growth years were obtained from the ventral margin (area of newest growth) using a diamond-grit cutting wheel, variable speed rotary tool. Two samples were taken from the absolutely dated (Modern) chronology, two each from Floating Chronologies 1–3 (the floating chronologies with replication >3 individuals), and one each from the remaining floating chronologies. Radiocarbon (14C) was analyzed at the National Ocean Sciences Accelerator Mass Spectrometry (NOSAMS) facility located in Woods Hole, MA, USA. All radiocarbon determinations were calibrated using the Marine20 radiocarbon curve (Heaton et al., 2020) using OxCal Version 4.3 (Ramsey, 1995).

In the paired samples taken from Floating Chronologies 1–3, one was taken near the beginning and one near the end of each chronology. The offset in the 14C dates between the paired samples was checked for consistency with the number of intervening years along the chronology, which provided a coarse test of crossdating accuracy. Subsequently, a Bayesian “wiggle matching” approach was implemented in OxCal to further refine the 14C age estimates and uncertainties using the paired radiocarbon probability distributions and their respective known sclerochronological offsets (Figure 4 inset; de Vries, 1958; Ramsey et al., 2001). The two samples from the Modern chronology (TND031 and TND045) were used to calculate the local reservoir age correction (ΔR) following methods outlined by Lower-Spies et al. (2020), which was applied to all floating segments to improve dating accuracy.

2.4. Chronology Construction

Once verification of crossdating was complete, measurement time series were standardized and averaged into population-level growth chronologies. Similar to trees, the radial growth increments of geoduck exhibit exponential decay, and these ontogenetic trends must be removed to isolate environmental signals. We explored two detrending techniques, the first of which was negative exponential (EXP) wherein a separate modified negative exponential function was fit to each measurement time series in the data set after which observed values were divided by those predicted (Black et al., 2009; Butler et al., 2010, 2013). The other approach was regional curve standardization (RCS) in which all measurement data were aligned with respect to biological age and fit with a single age-varying spline (called the regional curve, RC; additional details in Figure S4; Melvin et al., 2007) prior to dividing observed values by those predicted (Butler et al., 2010; Mitchell, 1967). Biases, however, could be introduced into an RCS chronology if long-lived individuals tend to grow more slowly than their shorter-lived counterparts, especially if growth from long-lived individuals dominates the early portion of the data set (Black et al., 2008; Schulman, 1954). To identify any relationships between growth and lifespan, measurement series were subdivided into groups with respect to longevity among which age-specific increment width was compared (Esper et al., 2003). Bias may also be introduced if a climatic trend is imprinted on the RC, as could happen in the case of contemporaneous birth of a significant subset of samples (Melvin & Briffa, 2008). To address this issue, Melvin and Briffa developed “signal-free detrending,” an iterative process in which each measurement time series is first divided by the mean chronology to remove the common, population-wide “signal.” This “signal-free” measurement series
is then detrended, and the process is repeated until there is minimal difference between the current and prior detrending curve.

Geoduck growth is characterized by rapid decline during its first $\sim 15$ years followed by many decades of extremely slow growth. In trees, long periods of very slow growth can affect the fit of negative exponential detrending curves, inducing artifacts in variance and mean trends, especially near the beginning and end of the chronology (Cook & Peters, 1997). A powerful and common solution is to apply an adaptive power transformation (EXPapt, RCSapt) to each measurement time series and then calculate residuals (subtraction) from the best-fit curve rather than indices (division; Cook & Peters, 1997). A single chronology was developed for each of these approaches by averaging all standardized time series with respect to calendar year. The robust bi-weight mean was used to reduce the impact of outliers (Cook & Kairiukstis, 1990). All detrending and chronology construction was performed in R (R Core Team, 2020) using many tools of dplR (Bunn, 2008).

We reviewed the results of EXP and RCS detrending (with/without APT) on the detrended series and resultant chronologies to determine which approach best suited the data. This was accomplished by examining the detrended time series for evidence of artifacts such as ontogenetic trends in variance or clearly spurious trends in mean. The strength of the patterns shared among samples was quantified by $30$-year-windowed (15-year overlap) mean interseries correlation ($\bar{r}$) and expressed population signal (EPS; Wigley et al., 1984, Equation 3.44 in Cook & Kairiukstis, 1990). EPS is calculated as

$$\frac{N \times \bar{r}}{1 + (N - 1) \times \bar{r}}$$

where $N$ is the total sample size. An EPS exceeding 0.85 has been used canonically as a minimum threshold for dendroclimatic reconstruction, though this value is arbitrary and must be interpreted with caution given that non-climatic processes could also give rise to common variance (Buras, 2017). In this study, EPS is also used as a metric of shared variance to interpret the relative efficacy of various detrending techniques. In an important limitation, the EPS calculation does not account for the difference in mean between individual series resulting from RCS detrending (Jones et al., 2009). Thus, a more rigorous standard was applied in which the EPS values of RCS chronologies were adjusted based on the ratio of variance in the indices to the variance in the corresponding 21-year spline-detrended indices ($\text{EPS}_{\text{adj}}$; Edge, 2021; Melvin & Briffa, 2014). We employed the more flexible 21-year spline (which tends to further reduce the $\text{EPS}_{\text{adj}}$ values) rather than the recommended 50-year spline to account for the shorter longevity and steep ontogenetic trend in geoduck growth compared with most tree species. Buras (2017) also suggests the use of subsample signal strength rather than EPS, however, we opt for a minimum sample size cutoff based on bootstrapped reconstruction residuals (detailed in Section 2.6). The final chronologies were judged based on measures of chronology variance and by analysis of a smoothed periodogram with variable taper.

### 2.5. Climate-Growth Relationships

Both temperature and sea surface salinity (SSS) have been connected to bivalve shell growth (Hiebenthal et al., 2012; Noakes & Campbell, 1992; Pourmozaffar et al., 2020; Strom et al., 2004), and SST is known to covary with Pacific geoduck shell growth at Tree Nob (Black et al., 2009). Lewis and Cerrato (1997) hypothesized that shell growth in many bivalves is related to temperature through metabolic rate based on laboratory manipulations of temperature and food, wherein shell growth was found to continue despite loss of tissue mass in *Mya arenaria*. And Storr et al. (1982) found shell growth to be directly related to temperature up to $13^\circ C$ in another subtropical bivalve, *Mercenaria mercenaria*. We examined instrumental records of both SST and SSS for relationships to the TNC. SST and SSS have been recorded daily at Langara Lighthouse, $\sim 145$ km northwest of Tree Nob, since 1936 (Figure 1, http://dfo-mpo.gc.ca). The period of 1940–2001 was selected for the climate-growth analysis based on replication in the chronology ($n \geq 7$) and data continuity at Langara. Daily data were aggregated to monthly means, and years with greater than three consecutive missing monthly values were discarded. Missing monthly SST and SSS values were filled using an interpolation algorithm in the forecast package in R that models seasonal patterns (Hyndman & Koohikar, 2007). We calculated correlations between the TNC and monthly averaged SSS and SST over a 24-month period spanning January of the year prior to ring formation through December of the concurrent year.
Significance of individual correlations was tested ($\alpha = 0.05$) using the “exact” simulation method to account for autocorrelation, wherein nonparametric estimates of series spectra were used to generate surrogates for a bootstrapping technique (Meko et al., 2011; Percival & Constantine, 2006). Those months that correlate most strongly with a chronology are generally sequential. Thus, we averaged across those months with peak correlations within SST and then within SSS. The SST mean series and the SSS mean series were entered as predictors of the TNC in a multiple linear regression using the MASS package in R with a $p < 0.01$ level to enter (Ripley et al., 2013). The parsimony of the model was further tested using Akaike Information Criterion (Sakamoto et al., 1986). This regression procedure allowed us to identify whether SST, SSS, or their combination were significant predictors of the TNC.

### 2.6. SST Reconstruction

We used linear regression over the 1940–2001 interval to generate the reconstruction (note that the reconstruction ends in 2000 due to the use of forward-lagged chronology values [lag+1]). First, however, we log transformed the TNC to normalize a right-skewed distribution characterized by outlier years of positive growth (Fritts, 1976; Menesguen & Dreves, 1987). Regression coefficients were used to reconstruct SST at Langara Lighthouse over all chronology intervals with sufficiently high EPS values, including floating segments. We calculated three skill metrics wherein both the early (1940–1970) and late (1971–2000) periods were used alternatively as independent calibration and verification intervals (Mann & Rutherford, 2002). The first skill metric is mean squared error (MSE), which is a measure of the difference between the predicted and observed values (Gauss, 1821; Lehmann & Casella, 2006). The second is the reduction of error (RE) statistic, which is a measure of the fit of the reconstruction in the calibration interval relative to the fit of the mean of the target data (Fritts, 1976). Finally, the coefficient of efficiency (CE) is calculated similarly to RE but in the verification interval (Cook et al., 1994). A value $>0$ for the RE and CE statistics indicates the reconstruction is a better predictor than the mean of the target, though the target mean is not an appropriate threshold to test skill when the target time series contains significant autocorrelation (Macias-Fauria et al., 2012). Therefore, we used a Monte Carlo phase randomization simulation in which a random time series of the same length and spectral properties as SST (Ebisuzaki, 1997) was generated using the Astrochron package in R (Meyers, 2014). The RE, CE, and MSE were calculated from this simulated SST series and the true SST series. This process was repeated 10,000 times, and the 99th percentile values were retained as significance thresholds ($\alpha = 0.01$) (Edge, 2021; Macias-Fauria et al., 2012). Also, the empirical prediction intervals of the reconstruction were calculated as the median and 90-percentile residuals obtained between the reconstruction and the instrumental SST in the verification interval (Lee & Scholtes, 2014).

The reconstruction period was limited to years with a sample size (number of individuals representing a single calendar year; Figure 2) sufficient to represent the population. To find the minimum necessary sample size, we used 1000 bootstrapped simulations at each possible sample size to determine at what sample size the MSE nears an asymptotic minimum. The simulated chronologies were created by sampling with replacement at each year to build chronologies of uniform sample size. The MSE was calculated for each simulated chronology over the full period of instrumental overlap. The median MSE of the 1000 simulations at each sample size was used to represent the error present in a chronology with the given sample size. A time series of these representative errors was used to find the minimum sample size required to adequately represent the population.

Wavelet analysis was performed on each segment of the reconstruction to assess frequency characteristics through time (Torrence & Compo, 1998) using the WaveletComp package in R (Roesch & Schmidbauer, 2018). Significance ($p$-value) was calculated by comparing the wavelet results against the wavelets of 10,000 simulated time series with the same autocorrelation and moving average structure as the reconstruction (Hyndman & Khandakar, 2007).

### 3. Results

#### 3.1. Sample Collection and Preparation

In total, 262 shells were processed. After processing the first 50 randomly selected shells we determined whether biometric measurements could be used to broadly estimate longevity, defined as total number of...
growth increments in the sample. To provide precise ages, only samples that could be crossdated and measured through the terminal growth year were used. Damaged shells were also discarded from this calculation so that only samples with accurate masses were used. We found a significant relationship between sample mass and longevity ($R^2 = 0.383$, $n = 15$, $p = 0.014$), so going forward we prioritized the processing of those shells over 70g to eliminate most individuals less than $\approx 50$ years in age. We found that, in general, measurement time series of at least 50 yr were necessary to extend the chronologies.

3.2. Crossdating

Of the 262 samples processed, 115 were successfully crossdated into 10 chronologies spanning more than 1,200 years (Figure 2). The median segment length of the crossdated individuals is 55.5 years, while the maximum longevity is 179 years. This is the longest-lived geoduck known, surpassing a 168-year-old individual (Bureau, 2002). Indices of bioerosion, shell margin deterioration, and periostracum preservation were compared to date of death, as established via crossdating and radiometric techniques. However, the relationships were not statistically significant. The series intercorrelations (the mean of correlation between each individual and the average of all others) of the four well-replicated chronologies averages $\sim 0.8$ (Table 1). The Modern chronology spans the interval from 1725 to 2008.

3.3. Radiocarbon Dating

In Floating Chronologies 1–3, the number of years between the early and late increments estimated from the radiometric analysis generally matched the number of intervening years established by crossdating.
(Table 2). All discrepancies between crossdated/radio carbon offset for paired samples were found to be within the 1σ error of the radiocarbon ages. When accounting for the local ΔR, Floating Chronologies 1, 2, and 3 have start dates of 699, 1010, and 1248 CE, respectively. The remaining floating chronologies were dated to various intervals during the last three millennia, but with particularly strong coverage from approximately 500 CE to present (Table 2; complete radiocarbon probability distributions in Figure S3). The measured/Marine20 radiocarbon ages (BP) of the samples used for the ΔR calculation are 860/606 (TND031) and 995/709 (TND045) for a regional offset of 270 years (Figure 2 Inset; σ = 22).

3.4. Chronology Construction

The RCS and EXP detrending methods (without adaptive power transform) produced chronologies with similar patterns in the year-to-year, high frequency domain (Figure 3a). The indices resulting from RCS detrending are more variable with greater spread among individuals relative to indices resulting from EXP detrending (Figure 3b). The EXP/RCS methods produced high average rbar (0.57/0.58, the average of correlation coefficients from all possible sample pairings), EPS (0.97/0.94), and EPSadj (0.91) statistics (Figure 3c).

<table>
<thead>
<tr>
<th>Chronology segment</th>
<th>Chronology interval*</th>
<th>Chronology length (years)</th>
<th>Total rings (years)</th>
<th>Number of series/samples</th>
<th>Series inter-correlation</th>
<th>Mean sensitivity</th>
<th>Mean segment length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modern</td>
<td>1725–2008</td>
<td>284</td>
<td>8,808</td>
<td>154/77</td>
<td>0.801</td>
<td>0.302</td>
<td>57</td>
</tr>
<tr>
<td>Floating 1</td>
<td>699–841</td>
<td>143</td>
<td>2,161</td>
<td>32/16</td>
<td>0.79</td>
<td>0.27</td>
<td>67</td>
</tr>
<tr>
<td>Floating 2</td>
<td>1010–1145</td>
<td>136</td>
<td>3,208</td>
<td>46/23</td>
<td>0.81</td>
<td>0.294</td>
<td>69</td>
</tr>
<tr>
<td>Floating 3</td>
<td>1248–1376</td>
<td>129</td>
<td>1115</td>
<td>18/9</td>
<td>0.863</td>
<td>0.292</td>
<td>61</td>
</tr>
</tbody>
</table>

Note. Only chronology segments with sample size greater than 3 included in table. *Chronology Intervals of floating chronologies are estimated based on radiocarbon dating. See Table 2 for radiocarbon data.

Figure 3. Empirical comparison of detrending methods in the modern interval. (a) Black line: regional curve standardization (RCS)-detrended average ring width chronology, Blue line: Negative exponential (EXP) chronology, Gray shading: Sample count, (b) Individual detrended ring width indices, EXP on left, RCS on right, (c) rbar, expressed population signal (EPS), and EPSadj statistics calculated over 30-year windows, 10-year overlap. (d) Smoothed periodogram with 0%, 5%, and 10% taper.
though the RCS chronology contains much more spectral power at lower frequencies relative to the EXP chronology, especially in the 100-year domain (Figure 3d; see also Figure S5). The retention of low-frequency variability is also apparent in a direct comparison of the two chronologies, especially the long-term 20th century increase that is captured by the RCS chronology but not the EXP chronology (Figure 3a). There were no apparent age-specific differences in growth rate among cohorts of different longevities (Figure S6), suggesting that fast-growing individuals are not more likely to be short-lived relative to slow-growing individuals. This allowed detrending of all series by a single ontogenetic RCS growth curve. Note that this single curve was developed using all crossdated time series (Figure 2), including live- and dead-collected individuals. This one RCS curve was used to produce a single chronology discontinuous in time, but such that segments were scaled relative to one another. We also produced a “signal-free” RC that proved to be very similar to the standard RC. We did not use signal-free detrending in our final chronology due to the complexity of implementation with a discontinuous chronology and its similarity with the standard RC (Figure S6). Combining RCS detrending with APT altered low-frequency signals in the chronology by offsetting the positions of measurement time series relative to one another (Figure S7). We developed a method to return all series to their correct positions to confirm the cause of this artifact (Edge, 2021). Thus, we utilized the RCS method without APT for climate-growth relationships and reconstruction given its exceptional ability to retain low-frequency variability while retaining robust chronology statistics.

### Table 2

Radiocarbon Dates and Associated Data

<table>
<thead>
<tr>
<th>Segment</th>
<th>Lab #</th>
<th>Sclero Age</th>
<th>F Modern</th>
<th>Fm Err</th>
<th>δ13C</th>
<th>Δ14C-corrected Median Radiocarbon</th>
<th>Wiggle-Matched Median</th>
<th>ΔR</th>
<th>Sclero Offset</th>
<th>Radiocarbon Offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS-157458</td>
<td>0.66</td>
<td>0.0015</td>
<td>2.35</td>
<td>-749</td>
<td></td>
<td></td>
<td></td>
<td>80</td>
<td></td>
<td>107</td>
</tr>
<tr>
<td>OS-157459</td>
<td>0.6635</td>
<td>0.0015</td>
<td>2.29</td>
<td>-698</td>
<td></td>
<td></td>
<td></td>
<td>70</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>OS-157461</td>
<td>0.6688</td>
<td>0.0015</td>
<td>3.06</td>
<td>-623</td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>OS-157487</td>
<td>0.7675</td>
<td>0.0025</td>
<td>2.38</td>
<td>697</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floating 1</td>
<td>OS-151095</td>
<td>0.775</td>
<td>0.0017</td>
<td>2.93</td>
<td>770</td>
<td>780</td>
<td></td>
<td>80</td>
<td></td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>OS-151062</td>
<td>0.7847</td>
<td>0.0018</td>
<td>3.16</td>
<td>877</td>
<td>860</td>
<td></td>
<td>70</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Floating 2</td>
<td>OS-157493</td>
<td>0.7982</td>
<td>0.0016</td>
<td>1.48</td>
<td>1022</td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>OS-151066</td>
<td>0.8045</td>
<td>0.0023</td>
<td>3.11</td>
<td>1092</td>
<td>1072</td>
<td></td>
<td>1124</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS-151065</td>
<td>0.8078</td>
<td>0.0019</td>
<td>1.63</td>
<td>1123</td>
<td>1142</td>
<td></td>
<td>1406</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floating 3</td>
<td>OS-151063</td>
<td>0.8384</td>
<td>0.0025</td>
<td>3.75</td>
<td>1382</td>
<td>1356</td>
<td></td>
<td>1702</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS-151064</td>
<td>0.8388</td>
<td>0.002</td>
<td>3.24</td>
<td>1385</td>
<td>1406</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS-157460</td>
<td>0.8767</td>
<td>0.0019</td>
<td>2.23</td>
<td>1702</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern</td>
<td>OS-157457</td>
<td>1789</td>
<td>0.8837</td>
<td>2.54</td>
<td>1770</td>
<td>-98.921</td>
<td>254</td>
<td>126</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS-148017</td>
<td>1915</td>
<td>0.8985</td>
<td>2.39</td>
<td>1885</td>
<td>-97.688</td>
<td>286</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Segment: chronology interval location of sample, yellow, green, red, and blue coloring shows membership to a named, well-replicated chronology segment while gray indicates membership to one of the unnamed, poorly-replicated segments. Lab#: Sample ID given by NOSAMS. Sclero Age: Sample calendar age based on crossdating. F Modern: Sample's radiocarbon content relative to modern standard as reported by NOSAMS. Fm Err: NOSAMS reported 1σ error of Fraction modern. δ13C: ratio of 13C:12C, reported in parts per thousand. Δ14C-corrected Median Radiocarbon: median date based on F Modern after ΔR correction in OxCal. Δ4C: radiocarbon age based on the equation $\Delta^{14}C = (Fm \times e^{(y-x) \cdot 1000} - 1) \times 1000$. ΔR: local reservoir correction, relative to Marine20 (Heaton et al., 2020). Sclero Offset: Span in years between the innermost and outermost radiocarbon sampled in a series. Values are rounded to nearest 5 when the small, outer rings of sample could not be precisely counted. Radiocarbon Offset: Span in years between ΔR-corrected Median Radiocarbon dates of the sample pair. Note that wiggle matched median dates are used to place the floating chronologies in time (e.g., OS-151065 dates to 1142, so the estimated median ring sampled from this series is assigned the date 1142, thereby affixing the Floating 2 chronology in time.)
3.5. Climate-Growth Relationships

Significant climate-growth relationships were found across nearly all months of the concurrent and prior year for both SSS and SST (Figure 4). We found that the average of prior July – concurrent November provided the highest correlations with geoduck growth for both SSS and SST. A stepwise multiple linear regression showed that the average of prior July – concurrent November SST was the only significant predictor ($p = 1.1 \times 10^{-10}$, $R^2 = 0.61$) and that the average of prior July – concurrent November SSS did not contribute any additional explained variance ($p = 0.427$). The correlation with SST was significant after accounting for autocorrelation using Ebisuzaki surrogates (1997) in a 10000-iteration simulation ($r = 0.78$, $\alpha = 0.01$; Meyers, 2014).

3.6. SST Reconstruction

Although the average of prior July – concurrent November SST provided the highest correlations with the TNC, we chose a narrower window not exceeding a year in width for greatest utility and ease of interpretation. However, when the seasonal window was narrowed, the TNC lagged by one year also became significant in reconstruction models (SST vs. lag-1 TNC $r = 0.629$; SST vs. lag-0 TNC $r = 0.635$; first-order autocorrelation of SST $= 0.456$). The need for this lag was consistent with the significant correlations observed for prior year SST (Figure 4) and that a prior July – concurrent November (seventeen-month window of SST) generated the highest single-variable correlation with the TNC. SST averaged from April through November provided the greatest variance explained when including concurrent and lagged (lag+1) TNC in a model and was thus chosen as the target for the reconstruction. In a stepwise multiple linear regression, both concurrent and lagged TNC proved significant ($p = 6.51 \times 10^{-6}$ and 0.00262 respectively), explaining 47% and 8% of variance, respectively. This model also produced the optimal Akaike Information Criterion given all SST and SSS predictors (SSS$_0$, SSS$_{+1}$, SSS$_{+2}$, SST$_0$, SST$_{+1}$, SST$_{+2}$). Regression residuals showed no significant linear trend ($p = 0.496$), lag-1 autocorrelation (annual; $p = 0.11$), or deviation from normality ($p = 0.06$). The split period calibration-verification resulted in significant RE and CE statistics.
for both the early (RE = 0.54 > RE₀₁ = 0.13; CE = 0.49 > CE₀₁ = 0.04) and late (RE = 0.57 > RE₀₁ = 0.14; CE = 0.53 > CE₀₁ = 0.05) calibration intervals. The adjusted $R^2$ (Wherry Formula 1 from Yin & Fan, 2001) for the full period of overlap is 0.54 ($n = 62$). The 50- and 90-percentile prediction intervals are ±0.25 and ±0.66, respectively. The reconstruction interval was limited to years with a sample size of seven or more based on the bootstrapped error estimation.

Figure 5. Sea surface temperature (SST) reconstruction. (a) Seasonal (April–November) SST at Langara Lighthouse reconstructed from Tree Nob geoduck. Red: Instrumental SST (April–November average) as measured at Langara Lighthouse. Black: Reconstructed SST from TNC, time interval truncated to sample size ≥7. Gray: 40-year, 50% frequency cutoff cubic smoothing spline of reconstruction (and instrumental data after 2000). Dark and light blue bands: 50% and 90% prediction intervals based on validation interval error. The first three chronology intervals are dated from radiocarbon samples, calendar dates shown for these intervals are approximate. See Supporting Information S1 for age probability distributions. (b) Black line segments show median reconstructed SST over 12-year window with no overlap. Gray shading shows 1σ range. (c) Expressed population signal (EPS) (orange) and adjusted EPS (blue) calculated over a running 20-year window. (d) Chronology sample count (discrete shells). (e) Wavelet of reconstructed SST. Coloring based on power. White circles enclose regions of significance based on Monte Carlo simulations ($p < 0.01$).
Reconstructed temperatures fluctuate at interannual to centennial scales and generally fall within 9–10°C, at least until late in the 20th century when warming exceeds this pre-industrial envelope (Figure 5a). Warming trends are also apparent in the Floating 2 segment, though the trend in the Modern interval is longer, beginning in the mid-1800s and spanning to present, while also of greater magnitude. The Modern segment contains both the coldest and warmest reconstructed temperatures in 1810 and 1998, respectively. The 1976-1977 Pacific regime shift (Miller et al., 1994) is captured by the reconstruction, and its magnitude is the most extreme in the record (Figure 5a). The maximum and minimum temperatures of the Modern interval fall within the 90% prediction interval of some years in prior centuries, so there is some possibility that comparable extremes occurred prior to modern times. The final and penultimate 12-year median SSTs fall outside the 1σ range of all prior intervals (Figure 5b). The EPS and EPS\text sub{adj} are high throughout, and those intervals with a minimum sample size of seven individuals are shown (Figures 5b and 5c). Wavelet analysis shows significant ($p < 0.1$) and consistent power at periods of 4–8 years across all reconstruction intervals (Figure 5d). Power at periods 16- to 64-years is also significant in some intervals, though the power is intermittent and the band variable, throughout the reconstruction. Note, however, that these are non-stationary time series and the null models may therefore not accurately represent significance. Significance is displayed to show exceptional power but should be interpreted with caution.

4. Discussion

In the marine realm, well-replicated, crossdated chronologies are expanding in spatial coverage and species diversity to address impacts of environmental variability on growth, environmental history, and interactions among species and trophic levels (Black et al., 2019). However, absolutely dated chronologies extending prior to 1800 CE remain less common and currently number less than a dozen (Black et al., 2019). Although six geoduck chronologies have been published, all were developed from live-collected material and thus extend, at most, into the mid-1800s CE (Black et al., 2009; Strom et al., 2004). The TNC is thus the first multicentennial, crossdated chronology of any species in the NEP and was made possible by combining live-collected individuals with overlapping dead-collected shells. Growth synchrony, or covariance, among samples was quite strong, facilitating crossdating and greatly reducing the chances of spurious matches among dead-collected individuals. Indeed, dead-collected individuals had only one very conspicuous placement in time according to cross-correlation analyses, and multiple radiometric dates verify the dates obtained by crossdating. The mean correlation between each individual and the average of all others was consistently above 0.8, which is as high as is observed in other bivalve datasets including A. islandica or G. glycymeris (e.g., Butler et al., 2009) and is among the highest values recorded for tree-ring data (e.g., Stahle et al., 2013). RingdateR greatly improved the pace of extending the chronology, which ultimately yielded a temporal coverage of $\sim$1,200 of the last 3,000 years.

Exact crossdating ensures that high-frequency, year-to-year variability is fully expressed in the final chronology, but there may still be issues retaining low-frequency variability, especially at multidecadal and longer timescales. When fitting a separate function to each set of measurements to remove age-related growth declines (Cook & Kairiukstis, 1990) any trends longer than the measurement series are also removed. Therefore, low-frequency variability in the resulting chronology is limited to timescales that are less than the average series length (Cook et al., 1995). Applying RCS allowed us to avoid this “segment length curse” by comparing each set of measurements to a single “regional curve” of age-related growth (Mitchell, 1967). Given that the analysis involved living and dead samples, this single curve provided a universal benchmark of average growth that spanned many cohorts and environmental regimes (Briffa et al., 1992). Individuals that lived through poor (favorable) environmental regimes would have relatively slower (faster) growth compared to the curve, and this information is retained in the detrending process and incorporated into the chronology.

Although RCS, under the correct circumstances, provides a solution to the segment-length curse, it introduces several new assumptions and complications (Briffa & Melvin, 2011). Biases arise if longer-lived individuals inherently grow more slowly than short-lived counterparts, especially if those slower-growing individuals are over-represented early in the chronology (Black et al., 2008; Esper et al., 2003; Schulman, 1954). RCS also requires individuals that represent a range of environmental regimes and thus spread across a
long timespan. Over-representation of samples in a narrow or fixed time window could result in a single climatic pattern being heavily imprinted on the regional curve, thereby distorting long-term trends (Melvin & Briffa, 2008). Finally, RCS tends to introduce relatively high levels of variance among indices and thus requires large sample counts to maintain a given level of signal strength in the chronology (Esper et al., 2003; Fritts, 1976; Melvin & Briffa, 2014).

The Tree Nob data set appears to meet the underlying assumptions necessary for an RCS analysis. Ontogenetic growth declines are highly geometric, and geoduck lack the sustained growth pulses or suppressions common in trees following disturbance, allowing the development of a robust regional curve. Also, our analyses indicate that there are no strong relationships between growth rate and longevity. Sample counts are large, and shells were collected from a geographically focused area that should be climatologically homogenous. Moreover, samples represent intervals of time that span the majority of the past 1,500 years and thus capture the necessary diversity of environmental regimes. The RCS method has been previously used on geoduck (Strom et al., 2005), but with all live-collected individuals, which resulted in a chronology that did not substantially differ from negative exponential detrending. However, the approach has been effective at resolving low-frequency variability in *A. islandica* chronologies developed from live- and dead-collected samples (Butler et al., 2010, 2013). In the TNC, evidence for preservation of low frequency is apparent in the chronology periodogram (Figure 3d) as well as the steep increase over the most recent century (Figure 3a). The spectral character of the RCS chronology is also a better approximation of the Langara SST record than the EXP chronology, particularly in the low-frequency domain (Figure S5). Poor fits of the regional curve functions tend to be most pronounced at the ends of the time series and can be amplified if low-frequency variability causes these end data points to exert leverage in the curve-fitting regression. However, signal-free detrending techniques did not make a difference in the Tree Nob RCS analysis, suggesting that there was no climate signal in the ontogenetically aligned data that could have biased the regional curve fit.

The adaptive power transformation has been commonly employed in tree rings (Esper et al., 2003; Fan et al., 2008; Panthi et al., 2017) and bivalves (Butler et al., 2010, 2013; Marali & Schöne, 2015) for the stabilization of variance. This tool is particularly useful when paired with exponential detrending, especially in situations where absolute growth rate drops to very low levels and remains in that state for prolonged periods (Cook & Peters, 1997), as is common in bivalve datasets. APT has been paired with RCS in tree-ring research (Büntgen et al., 2005; Esper et al., 2003), but this detrending combination introduced low-frequency artifacts in the Tree Nob data because transformation of individual series altered their positions relative to the regional curve. This vertical offset of individual series radically altered low-frequency signals. Thus, care must be taken in implementing APT in combination with RCS when analyzing bivalve growth-increment data.

The TNC correlates very strongly with SST, especially over the seasonal window of April through November, which is consistent with earlier findings at Tree Nob (Black et al., 2008) and generally matches with the warm-season correlations of March-October in the Protection Island chronology in the Strait of Juan de Fuca (Strom et al., 2004). A key difference, however, is that in our reconstruction, the TNC lagged by one year was also a significant predictor of SST in a stepwise regression. This suggests there is “biological memory” of the prior growing season in geoduck growth increments at this site, although memory at a point lower in the food chain cannot be ruled out. To further confirm this memory, we correlated the geoduck chronology against a much wider seasonal window of prior July through the concurrent November. When this window was used as a reconstruction target, only the current year of geoduck growth was significant. Thus, to reconstruct SST in the seasonal window of current April through November, subsequent-year geoduck growth significantly increased predictive skill. Biological memory has been well documented in tree-ring records and can arise from the multi-year persistence of leaves in non-deciduous species or use of stored energy (Esper et al., 2015; Matalas, 1962; Schulman, 1956). There may be an analogous situation in bivalves in which stored glycogen, for which peak reserves occur in July and August, and is carried over from prior growing seasons (Feldman et al., 2004; Soudant et al., 1996).

The SST reconstruction from geoduck is characterized by considerable interannual to interdecadal variability and is linked to broad-scale climate patterns in the North Pacific basin (Figure S8). The most energetic of these is the El Niño Southern Oscillation (ENSO), which is teleconnected to the Aleutian Low (Newman et al., 2016; Schneider & Cornuelle, 2005) and thus is expected to be related to the geoduck SST.
reconstruction. The Niño3 index from the Hadley Centre Sea Ice and Sea Surface Temperature data set (Rayner et al., 2003) significantly correlates with Langara SST (1940–2017, $r = 0.45$, $p = 2.8e^{-5}$) and also significantly correlates with the reconstruction over the interval 1870–2000 ($r = 0.31$, $p = 3.0e^{-4}$). Warm years in the reconstruction correspond to major El Niño events including 1941 and 1998 (Trenberth & Stepaniak, 2001). This may also explain some of the power in the 3- to 7-year window of the wavelet analysis of the SST reconstruction (Figure 5e). Lower-frequency variability in the reconstruction is also consistent with the Pacific Decadal Oscillation (PDO), defined as the leading empirical orthogonal function of gridded SST north of 20° latitude in the North Pacific after the global warming trend has been removed (Mantua et al., 1997). Over the 20th century, the PDO varies primarily at 10–80-year timescales (Newman et al., 2016; Wills et al., 2018). The Langara reconstruction does show some power in these frequencies over the past 1,500 years, suggesting energy in these domains is characteristic of North Pacific SST variability over the past several centuries; however, the low-frequency power is variable in peak periodicity, intermittent in time, and less significant on average than peaks in the 4–8-year band. The Langara instrumental record significantly correlates with PDO ($r = 0.71$, $n = 78$, $p < 1.0e^{-5}$) as does the Langara reconstruction from geoduck ($r = 0.49$, $n = 108$, $p < 1.0e^{-5}$). The PDO index, however, is detrended to remove the global warming trend and accentuate interdecadal variability (Mantua et al., 1997) while the geoduck chronology was generated to preserve these longer-term trends. Thus, a more appropriate comparison is with the “Arc” pattern, calculated as the dominant mode of SST variability in the northeastern Pacific (east of 180° longitude) without detrending. The Arc pattern is more closely focused on the northeastern Pacific and is dynamically linked to the atmosphere (Johnstone & Mantua, 2014). Correlation between the Arc pattern and the Langara reconstruction is $r = 0.62$ ($p < 1.0e^{-5}$), underscoring the strength of this broad-scale climate pattern in the geoduck increment widths that includes the 20th century warming trend.

Multiple studies have targeted the PDO or Gulf of Alaska sea surface temperatures for reconstruction. Analysis involving annually resolved archives are dominated by tree-ring datasets (Biondi et al., 2001; D’Arrigo & Wilson, 2006; D’Arrigo et al., 2001; Gedalof & Smith, 2001; MacDonald & Case, 2005), though long-term instrumental precipitation records from China have also been used (Shen et al., 2006). These reconstructions, however, poorly agree with one another before 1900 (Henley, 2017; Kipfmueller et al., 2012), and thus there is considerable uncertainty about the variability of the North Pacific prior to the 20th century (Cook, 2009; Newman et al., 2016). Although the geoduck-based reconstruction contains 20th century warming trends and other reconstructions may not, especially if the target is explicitly the PDO, we compared the Langara history to six annually resolved proxy records that are relevant to northeastern Pacific SST (Biondi et al., 2001; D’Arrigo & Wilson, 2006; D’Arrigo et al., 2001; Gedalof & Smith, 2001; MacDonald & Case, 2005; Shen et al., 2006). Among these only the Gedalof and Smith (2001) reconstruction significantly and stably correlates ($r = 0.39$, $n = 250$, $p < 1.0e^{-5}$), with comparable relationships pre- and post- 1900 with the Langara reconstruction. The combination of the Langara reconstruction with that from Gedalof and Smith by simple average produces a series that correlates much more strongly with the PDO index ($r = 0.66$, $n = 82$, $p < 1.0e^{-5}$) than either series individually. This result is consistent with earlier findings that a composite of multiple geoduck and SST-sensitive tree-ring chronologies could explain greater quantities of SST variance when combined (Black et al., 2009). The agreement between the Langara reconstruction and that of Gedalof and Smith (2001) suggests that trees and bivalves likely share common patterns, and that with careful selection of chronologies, could be combined to generate much longer multi-proxy histories. Previous SST-related reconstruction attempts may suffer from some combination of instability in the dominant modes of climate variability (Bond et al., 2003; Di Lorenzo et al., 2008; Gedalof et al., 2002), spatial limitations of proxy networks, inconsistencies of teleconnections (Batehup et al., 2015; Gallant et al., 2013), idiosyncrasies of reconstruction methods, and incompatibility of frequency characteristics (Franke et al., 2013). The geoduck chronology could provide a much longer, annually resolved, marine-based estimate of SST variability with which to screen tree-ring chronologies for stability over multicentennial timescales, which can be difficult to determine given the brevity of the instrumental record, the relatively low number of multidecadal cycles, and the possibility of changing climate-growth relationships with anthropogenic warming (Esper & Frank, 2009; Frank et al., 2007; Wilson et al., 2007). However, the TNC originates from one point in the Pacific and, although SST in this region is strongly correlated with the Arc and PDO patterns in the instrumental period, there is the possibility that SST at this location may not be coupled with basin-scale patterns in earlier time periods.
In addition to broad-scale teleconnected climate processes, the geoduck-based reconstruction also appears to capture cooling associated with minima in solar activity and major volcanic eruptions. The coldest decade of the reconstruction centered around 1810 coincides with the Dalton Minimum as well as the coldest decade of the last 500 years reported in a review of Northern Hemisphere (NH) temperature reconstructions (Cole-Dai et al., 2009). Volcanism is also a likely contributor to the cold, and thus the slow growth that geoduck exhibit during this period (Wagner & Zorita, 2005). The two largest NH volcanic eruptions since 1735 occurred in 1809 (Unknown) and 1815 (Tambora; Gao et al., 2008). At least one of these appears to have a signature in the geoduck-based SST reconstruction, indicated by the three coldest years on record in 1809–1811. The year 1816, however, is relatively warm in the Langara reconstruction and was likely associated with an El Niño event (Li et al., 2013), though the reconstruction returns to conspicuously cool conditions in 1817 and 1818. The 1982 and 1991 eruptions of El Chichón and Pinatubo also foretold brief, sharp cooling events in the warmest reconstruction interval. Thus, volcanic activity may have an influence on the regional SST but may also be overridden by other climatic processes (Adams et al., 2003).

One of the most conspicuous features of the reconstruction is the warming trend that began around 1850 and continues through the end of the record. Indeed, the decades since the 1976 shift to a warm regime (Latif & Barnett, 1996; Nitta & Yamada, 1989; Trenberth, 1990) are the warmest of the 50 decades in the reconstruction (Figure 5b), consistent with NH surface temperature and global SST reconstructions (Mann et al., 2009). Although the reconstruction ends in 2000 and is specific to a nearshore location, it does underscore the long-term warming that has occurred in the northeastern Pacific over the past century or longer and thereby provides context for the recent and apparently unprecedented marine heatwaves known colloquially as “The Blob” (Bond et al., 2015; Di Lorenzo & Mantua, 2016). The first of these warm-water events occurred between 2014 and 2015 followed by another in 2019. These heatwaves were characterized by significant reductions in krill and forage fish, species redistributions, harmful algal blooms, seabird mortality, broad-scale marine mammal strandings, and the closure of multiple fisheries (Amaya et al., 2020; Cornwall, 2019; Gewin, 2015; Kintisch, 2015). Peak warming occurred primarily offshore and is thus not reflected well in the Langara instrumental record, but the warming trends that culminated in these apparently unprecedented heatwaves are clear in the geoduck-based reconstruction.

Ultimately, the TNC continuously extends between 1725 and 2008 while also providing multiple windows that span 58% of the past 1,500 years. As such, this is the longest annually resolved marine archive currently available in the North Pacific, and with a high probability of being substantially extended. Indeed, with additional collections, there is the distinct possibility that gaps in the chronology could be filled to yield a continuous millennial to 1,500-year record. From the limited number of radiocarbon dates of randomly selected, undated shells, the minimum preservation is 3,000 years, but may be much greater considering that A. islandica that died more than 10,000 years BP have been retrieved from the North Atlantic seafloor (Butler et al., 2020). Given the abundance of geoduck along the NEP coast and typical average lifespan, chronologies of similar length could be developed from the Strait of Juan de Fuca to Kodiak, Alaska. To better constrain past climate variability, the exactly dated framework of the TNC could also be sampled for isotopic or possibly microchemical analysis. Oxygen isotope ratios have proven to be a robust indicator of past hydrography (i.e., combined effects of salinity and temperature changes), especially for A. islandica and may also be useful for geoduck (Reynolds et al., 2013, 2016; Wanamaker et al., 2008).

On broader spatial scales, NEP SST modulates global temperatures via fluxes of heat storage and release, especially on decadal timescales (England et al., 2014; Meehl et al., 2011, 2013; Thompson et al., 2015; Yin et al., 2018). Thus, better constraining the past may afford some predictive skill for the future. For example, decadal climate prediction skill for North America, and associated climate adaptation planning, is also contingent on the predictability of NEP SST regime (Meehl et al., 2009). Moreover, an annually resolved reconstruction of basin-wide NEP SST, possibly generated through multi-proxy approaches, could also be used to test hypotheses regarding the origins of NEP SST decadal-scale variability. NEP SST regimes have been interpreted by some as oscillatory, with forcing mechanisms related to Rossby wave propagation time or gyre circulation (Chao et al., 2000; Mantua & Hare, 2002; Minobe, 1999). However, an increasing body of evidence suggests that NEP SST is better described as red noise resulting from higher-frequency signals of tropical (e.g., the El Niño Southern Oscillation) and extratropical (e.g., variability in the Aleutian Low) origin being buffered in the water column and circulation patterns of the NEP (Mestas-Nunez & Miller, 2006;
Newman et al., 2003; Pierce, 2001; Schneider & Cornuelle, 2005; Wills et al., 2019). A basin-scale NEP reconstruction could be compared to independent, annually resolved ENSO reconstructions to test these relationships over multicentennial timescales. Also, residual variance could be compared to eastern Asian climate or that of the western Pacific to better identify extratropical influences on NEP SST. As such, geoduck represent a key archive that could be greatly expanded in temporal and spatial scope to address basic questions of past NEP variability and its drivers.

Data Availability Statement

The growth increment measurements and related “pith offsets”, master chronology, and reconstruction data are archived at NOAA Paleoclimatology (https://www.ncdc.noaa.gov/paleo/study/33312).

Acknowledgments

Shell collection was performed by Trident Biological Consulting Ltd., and we thank Mike Atkins for his thoughtful and professional services that provided this unique collection. We also thank Hank Carson at the Washington Department of Fish and Wildlife for his support in the sample collection process. This work is funded by the National Science Foundation (AGS Award Number: 1855628 to B.A. Black; Award Number: 1602751 to A.D. Wanamaker; Award Number: 1602633 to P.G. Butler). We thank Andy Bunn, Rob Wills, and Luis Calderon-Aguilera for their substantive comments that significantly improved this manuscript.

References


Gauss, C. (1821). Theory of the combination of observations which leads to the smallest errors. AKADEMIE VAN WETENSCHAPPEN B, 7(2), 193. https://doi.org/10.1017/s0033822200005592


