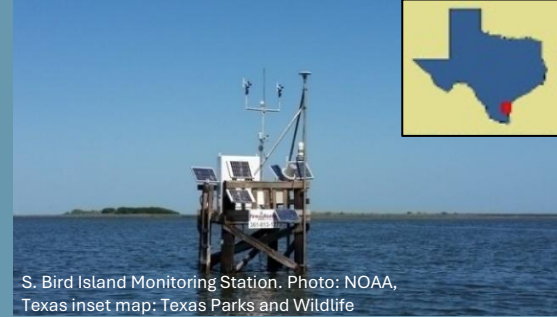


# Water Temperature Trends at South Bird Island, TX (2009 – 2023)

Spencer G. Womble – University of Georgia, Warnell School of Forestry and Natural Resources



S. Bird Island Monitoring Station. Photo: NOAA, Texas inset map: Texas Parks and Wildlife

## Background

The Laguna Madre provides critical habitat for fish, shellfish, and sea turtle species. Changes in water temperature regimes can stress marine organisms, disrupt breeding cycles, and alter ecosystem dynamics. Therefore, understanding temperature extremes and trends over time can improve identification of potential risks to the ecosystem. We used a percentile-based approach to identify days with extreme average temperatures and assess how trends in extreme day frequency, duration, and intensity changed over time. We also tested for an overall trend in water temperature by analyzing weekly means.

## Identifying Extreme Days

Daily mean water temperature was calculated from 15-min interval data collected at South Bird Island monitoring station (coordinates 27.484346, -97.318143) from 2009-2023. Dates were then grouped into either a reference period (first 10 years) or recent period (last five years) following PAIS vital signs protocols. Days with daily mean temperatures outside the 5<sup>th</sup> (11.9 °C; 53.4 °F) and 95<sup>th</sup> (30.9 °C; 87.6 °F) percentiles for the reference period were classified as cold or warm extremes, respectively. Days missing more than 20% of their possible observations were removed prior to

analysis. For the remaining days, an average of 31.6 days per year (15.8 cold and 15.8 warm) were identified as extreme temperature days for the reference period. Note that this percentile-based approach does not necessarily identify days where water temperatures were stressful to marine organisms, and some “extreme” days may still fall within an organism’s thermal tolerance.

## Extreme Days Analysis

Differences in frequency, duration, and intensity of extreme temperature days were compared between the reference and recent periods. Frequency was defined as the average number of days classified as extreme per year. Duration was assessed by counting runs of extreme temperatures, where each run represented two or more consecutive extreme temperature days. Intensity was calculated for extreme days only, defined as the absolute difference between a day’s mean temperature and its corresponding cold or warm threshold. Higher values indicate a greater deviation from the threshold.

## Temporal Trend Analysis

In addition to the extreme day analysis, a statistical test for an overall trend in water temperature was performed using weekly mean water temperature as the response

variable. Weekly mean water temperature was calculated for all weeks across the record, with weeks missing more than 20% of their possible observations removed prior to calculation. A generalized additive mixed model (GAMM) was then used to test for a temporal trend in weekly mean water temperature. Weekly means were used instead of daily means to reduce the risk of model bias due to confounding temporal autocorrelation.

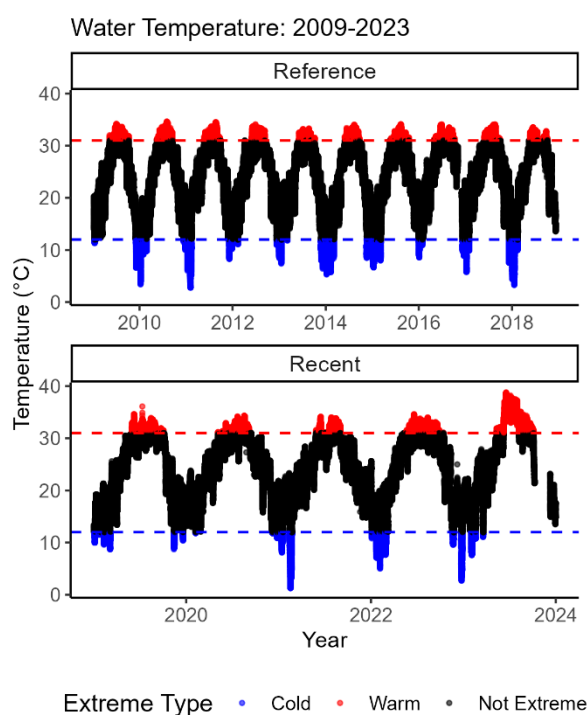


**Fig (i)** Water temperature observations from 2009-2023, colored by extreme day type. Dashed red and blue lines indicate thresholds for extreme warm and cold temperatures, respectively.

## Key Findings

The frequency of extreme warm days increased substantially in the recent period, while extreme cold days became less common. Intensity for both warm and

cold extremes was also higher in recent years. Warm run frequency rose slightly, and cold run frequency declined slightly, with no meaningful change in run duration for either type. However, no long-term trend was detected in mean weekly temperatures. Overall, these results suggest that even without a clear trend in average water temperature, extreme warm events are becoming more frequent, more intense, and somewhat more persistent. Further, although extreme cold days occurred less often, those that did occur were more intense in recent years.

### Extreme Day Frequency

- Between the reference and recent time periods, the average frequency of extreme warm days increased from 15.8 to 38.4 days per year, while average frequency of extreme cold days decreased from 15.8 to 13.5 days (Fig. ii(A)).

### Extreme Day Intensity

- For extreme warm days, average intensity increased from 0.4°C (0.5 °F) to 1.5°C (2.7 °F) between the reference and recent periods.
- For extreme cold days, average intensity increased from 2.2 °C (4.1 °F) to 2.9 °C (5.2 °F) between the reference and recent periods.

### Run Frequency

- The average number of warm runs per year increased from 3.1 to 4.8 between the reference and recent periods

- The average number of cold runs per year decreased from 3.7 to 2.4 between the reference and recent periods

### Run Duration

- The average duration of an extreme warm run increased from 4.5 days to 7.6 days between the reference and recent periods. However, this increase was primarily due to a heatwave during the summer of 2023. If 2023 data were removed, the average duration of warm runs in the recent period was 5.3 days.
- The average duration of an extreme cold run increased from 4.1 days to 4.4 days between the reference and recent periods

### Temporal Trend Analysis

- The estimated temporal trend was not statistically significant at  $\alpha = 0.05$ , with confidence intervals for the trend estimate ranging from a -0.01 to 0.15 °C change in mean water temperature per year

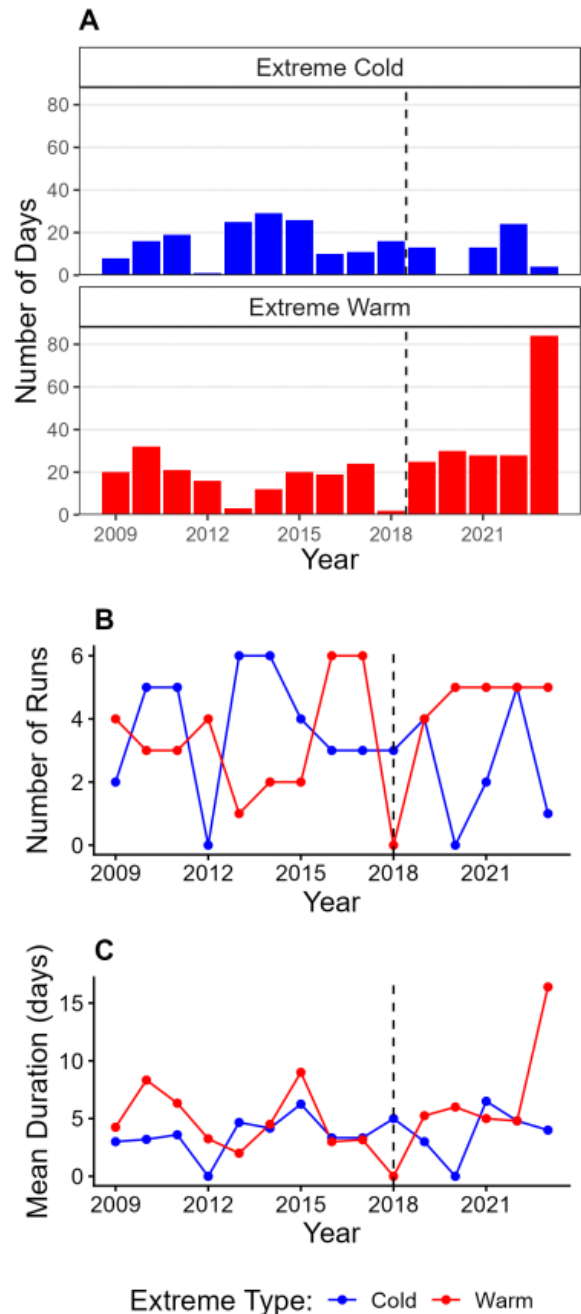


**Fig (ii)** Frequency of extreme days (**A**), run frequency (**B**), and mean run duration (**C**) by year. Vertical dashed black lines divide the reference and recent periods.

# Analysis of Water Temperature data from South Bird Island, TX (2009-2023)

**Submitted to:**

Inventory & Monitoring - Gulf Coastal Network

**Submitted by:**

Spencer G. Womble

University of Georgia, Warnell School of Forestry and Natural Resources,  
108 E Green Street, Athens, GA 30602-2152

**Date:**

December 16<sup>th</sup>, 2025

*This report is based on water temperature data recorded at South Bird Island from 2009-01-07 through 2023-12-31. These findings may be subject to change as future data are added through ongoing monitoring.*

# Contents

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Background .....                                                           | 1  |
| Methods.....                                                               | 2  |
| Data Collection Related Uncertainties in Long-term Time Series.....        | 2  |
| Summary of Missing Data.....                                               | 3  |
| Characterizing seasonality .....                                           | 6  |
| Extreme Days Analysis .....                                                | 7  |
| Extreme Days Classification and Frequency .....                            | 7  |
| Percentile Selection Considerations.....                                   | 7  |
| Extreme Temperature Runs .....                                             | 8  |
| Extreme Day Intensity.....                                                 | 8  |
| Temporal Trend Analysis .....                                              | 8  |
| Data Pre-processing for a Trend Analysis .....                             | 9  |
| Model Structure .....                                                      | 9  |
| Choosing the Basis Dimension.....                                          | 10 |
| Model Validation .....                                                     | 11 |
| Checking for Bias Due to 2023 Heatwave .....                               | 11 |
| An Exploratory Analysis Using a Nonlinear Term for the Trend Variable..... | 12 |
| Results.....                                                               | 12 |
| Characterizing seasonality .....                                           | 12 |
| Extreme Days Analysis .....                                                | 15 |
| Extreme Day Frequency.....                                                 | 16 |
| Extreme Day Runs .....                                                     | 18 |
| Extreme Day Intensity.....                                                 | 20 |
| Temporal Trend Analysis .....                                              | 21 |
| Key Take-aways .....                                                       | 23 |
| Extreme Days Analysis .....                                                | 23 |
| Temporal Trend Analysis .....                                              | 23 |
| References .....                                                           | 24 |
| Appendix I .....                                                           | 27 |
| Appendix II .....                                                          | 31 |

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| An Exploratory Analysis Using a GAMM with a Non-linear Term for the Trend Variable... | 31 |
| Some warnings about this method:.....                                                 | 34 |

## Figures

- Figure 1** Data completeness by month for all included years. Gray bars represent recorded observations, while red outlines show total possible observations. Missing data corresponds to the white space between gray bars and red outlines.....5
- Figure 2** Heatmap of missing data by month and year. Darker colors indicate a greater proportion of missing data. ....6
- Figure 3** Plot of raw temperature data by date. .... 13
- Figure 4** Daily mean temperature by day of year using the water year calendar. Colored shading used to indicate binary cold and warm seasons. The anomalously high points for June through August are the 2023 heatwave..... 14
- Figure 5** Density plot of daily mean temperature by time period (reference is 2009-2018 and recent is 2019-2023). Dashed black lines indicate the 5th and 95th percentile thresholds (11.9 and 30.9 °C, respectively) for the reference time period. Rug plots along the x-axis indicate individual days. .... 16
- Figure 6** Daily mean temperature for the reference and recent period. Red points are extreme warm days and blue points are extreme cold days. The red and blue dashed lines indicate the 11.9 °C and 30.9 °C thresholds for the 5<sup>th</sup> and 95<sup>th</sup> percentiles for the reference period, respectively..... 17
- Figure 7** Frequency of extreme days by year. Dashed black lines indicate the divide between reference and recent periods. .... 18
- Figure 8** Run frequency (**A**) and mean run duration (**B**) by year. Dashed black lines indicate the divide between recent and reference periods. .... 19
- Figure 9** Annual mean intensity of extreme days by extreme type. The dashed black line separates reference and recent periods. Note that 2020 did not have any extreme cold days..... 21
- Figure 10** Partial effect of fracyear on weekly mean temperature. This figure depicts the non-linear relationship between time of year and water temperature. The y-axis indicates the change from the model's global mean temperature (i.e., intercept) in °C. The lower x-axis shows the fractional time of year, and the upper x-axis shows the corresponding month. Note that month positions are approximate..... 22

## Tables

**Table 1** Summary statistics for water temperature (°C) by month and binary season (data combined across years). ..... 15

**Table 2** Summary statistics for the run frequency and duration analysis by period (reference = 2009-2018 and recent = 2019-2023) for years with runs. Units are in number of days long. .... 20

## Appendix Figures

**Appendix Figure 1** Observed versus fitted weekly mean temperatures for GAMM models with cubic cyclic spline basis dimensions of 3 and 5. Points closer to the diagonal indicate better model fit. The red line should be straight if no heteroskedasticity remains in the model. Note the slight curve in the red line when the basis dimension is set to 3. .... 27

**Appendix Figure 2** Intensity scores for cold and warm extreme days colored by time period (reference = 2009-2018 and recent = 2019-2023). Points are jittered along the x-axis by 0.1 years for easier visualization. The y-axis coordinates are not jittered..... 30

**Appendix Figure 3** Estimates of the slope of centered decimal date for the final model (labeled here as “Random Intercept”) and a model with a tensor smooth on centered decimal year. Note that the confidence intervals around the estimate for the “Smooth on Trend Variable” model are likely wider than were calculated by the slope estimation procedure. .... 33

## Appendix Tables

**Appendix Table 1** Summary of missing observations across months..... 28

**Appendix Table 2** Comparison of the GAMM with and without data recorded during the 2023 heatwave included. Estimates are coefficient estimates. The estimate for cent\_dec\_date is the trend estimate. Values for the CI column are the 2.5<sup>th</sup> and 97.5<sup>th</sup> percent confidence intervals, respectively. Note that the confidence intervals for cent\_dec\_year cross zero for both models, meaning the true slope may be zero. P is the p-value for the parameter. .... 29

# Background

Water temperature is a vital component of aquatic ecosystems due to its influence on ecosystem structure and function. However, changing water temperature conditions have caused concern for the future of aquatic ecosystems. Water temperatures are rising globally due to climate change, and warming rates are unequal across ecosystems (Barnett et al., 2005; Abraham et al., 2013; Gleckler et al., 2016). Of the aquatic ecosystems threatened by climate change, estuaries are especially vulnerable to increases in water temperature because of their shallow depths and limited exchange with the ocean (Scanes et al., 2020). Increasing estuary temperatures may disrupt these ecosystems by altering community structure and biogeochemical cycling (Hamilton, 2010; Grilo et al., 2011; Boucek and Rehage, 2014; Testa et al., 2018)..

The Laguna Madre is one of the largest hypersaline estuaries in the world and provides critical habitat for a variety of sea turtles, fish, and birds (Watson, K. 2005). This ecosystem may be particularly vulnerable to increasing temperatures due to its shallow depth and clear water. For instance, increasing water temperatures can be associated with reduced dissolved oxygen content (Matear and Hirst, 2003; Ni et al., 2019), excluding aquatic organisms with narrow thermal ranges (Madeira et al., 2012; Waltham and Sheaves, 2017), and damaging vital sea grass habitat (Echavarria-Heras et al., 2006; Collier and Waycott, 2014; Shields et al., 2019). Given Laguna Madre's ecological significance, water temperature has been measured continuously for over 15 years at South Bird Island, Texas. However, few studies have quantified how temperatures may have changed in the Laguna Madre over this timespan (Adams Tissot 2007).

To determine how average water temperature conditions have changed over time, we analyzed 15 years of water temperature data collected near South Bird Island in the Laguna Madre (coordinates 27.484346, -97.318143). Through discussions with the Network PI and Program Manager, and in conjunction with Network protocols, the following objectives were defined for this study as 1) characterization of the seasonality of the system, 2) comparisons in frequency, duration, and intensity of extreme temperature days between an initial reference time period (2009-2018) and a recent time period (2019-2023), and 3) performance of a statistical test for a temporal trend in average water temperatures. In combination, these analyses provide a multipronged approach to examining if and how water temperature dynamics have changed over time.

# Methods

To accomplish the above objectives, methodological steps included consideration of missing data (i.e., individual times without a temperature data record), seasonality, “extreme” days (i.e., the days with mean temperatures in the upper or lower percentiles), and estimating a temporal trend. Some of these analyses required additional data processing steps, which are described in more detail below. We also note some other reported or observed challenges for maintaining consistency across a multi-year continuous monitoring effort.

## Data Collection Related Uncertainties in Long-term Time Series

Long-term monitoring presents many challenges for maintaining data quality and consistency. Below are methodological challenges encountered in the dataset and the associated analytical decisions chosen for this report.

- Sampling frequency – In long-term monitoring, changes in sampling frequency may be necessary to account for environmental conditions or to achieve the desired resolution. In this dataset, temperatures from 2008-06-02 to 2009-01-06 were recorded every 30 minutes. However, the recording frequency was increased to 15-minute intervals after 2009-01-06. Observations collected prior to the change to 15-minute sampling intervals were removed to maintain resolution consistency across the time series.
- Missing data – Partial years of data (i.e., missing data) are common in long time series. While the full dataset included data from the years 2008 and 2024, 2008 only contained data from June through December, and 2024 only contained data from January through April. Given the potential sensitivity of the extreme days analysis to missing entire seasons, particularly the hottest or coldest times of the year, we removed all 2008 and 2024 data prior to analysis.
- Gear consistency – Updating sampling equipment can be necessary for reasons such as device obsolescence or damage. However, equipment changes can introduce uncertainty due to differences in precision and accuracy.
  - Data from 2008-06-02 to 2022-03-04 were collected using a YSI 6600 multiparameter sonde. This sonde model was later replaced by a YSI EXO2 sonde which was deployed on 2022-03-14 at 9:15pm. We were unable to directly assess the influence of sonde model on the temperature records because overlapping data for both sonde

models were not available. If both sonde models were deployed side by side (e.g., for at least a month), then the difference in variability due to the equipment change could be compared more directly. Data for both sondes were compared visually and with summary statistics, and no obvious changes in temperature variability, precision, or patterns were observed.

- In addition to changing sonde models, an anti-fouling device was also installed in April 2013 to control biofouling. The principal investigator (PI) at the time of writing this report indicated that the device greatly reduced the frequency at which sensors failed due to biofouling. No clear change in temperature variability around the time when the device was installed was observed during data exploration, and we did not incorporate this change into the trend analysis model.
- Quality control and assurance (QAQC) – QAQC protocols are also sometimes changed as long-term monitoring progresses. Notes accompanying the data in Aquarius, the National Park Service's (NPS) continuous water quality parameters database, indicated there were intermittent periods where QAQC was not applied to the data. However, it was not always clear which data did not undergo QAQC. Further, the details of when this period of uncertain QAQC occurred were not explicitly provided to the author of this report, possibly because this information was unknown. All information regarding QAQC uncertainty was provided via email by the PI for the Inventory and Monitoring Division – Gulf Coastal Network at the time of writing this report.

## Summary of Missing Data

Nonrandom missingness of temporal data can be influential to its interpretation and analysis. Therefore, we summarize temperature missingness at multiple levels including observational (15-min), monthly, and yearly intervals. After removing 2008 and 2024 data, temperature data were missing for 50,260 of a possible 525,216 observations (~10% missing). Missingness was unequal across months, ranging from 5% for January (aggregated across all years) to 21% for August (Fig. 1). August and November had the most missing data with 19% missing for November. The high November missingness rate was driven primarily by the years 2016, 2018, and 2023 where 92%, 70%, and 100% of their possible observations were missing, respectively (Fig. 2). August experienced 8 years with at least 10% of data missing. However, the remaining 7 years contained little to no missing

data for August, and these years were distributed throughout the full record rather than temporally clustered.

It is currently unclear why August and November had such higher rates of missingness or what caused these data gaps given that missingness rates for other summer and autumn months were below 13%, with most months having less than 10% missingness (Appendix Table 1). The high frequency of missingness for August may indicate systemic (i.e., nonrandom) missingness in the data. However, because August missingness was not consistent across consecutive years, the temporal trend model can still be estimated through years with incomplete August data, reducing the potential for systematic bias. Unlike the temporal trend analysis, the extreme days analysis directly compares observed counts between time periods without modeling underlying trends. If August data are disproportionately missing in one period, this comparison may be biased because there is no way to estimate extreme day frequency during periods of missing data.

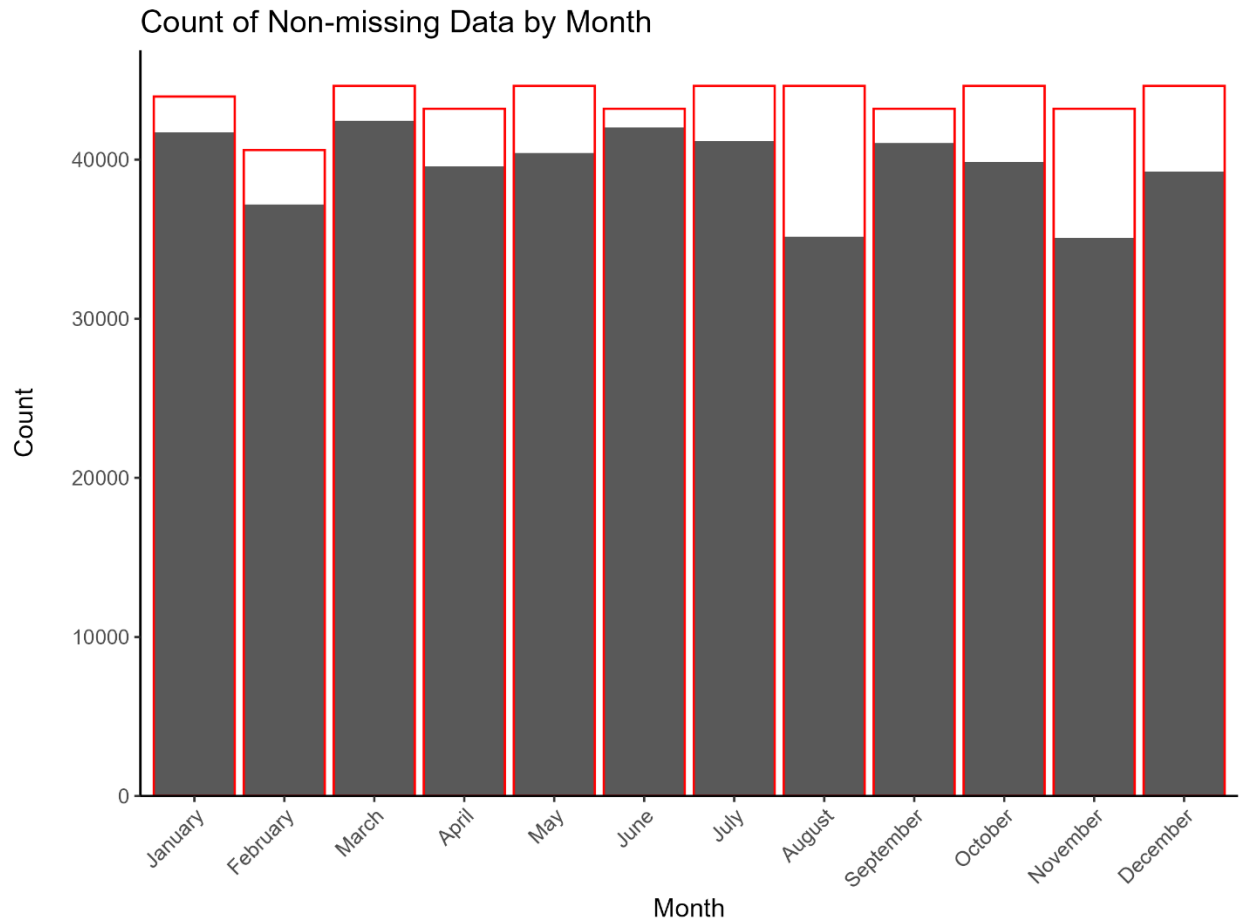


**Figure 1** Data completeness by month for all included years. Gray bars represent recorded observations, while red outlines show total possible observations. Missing data corresponds to the white space between gray bars and red outlines.

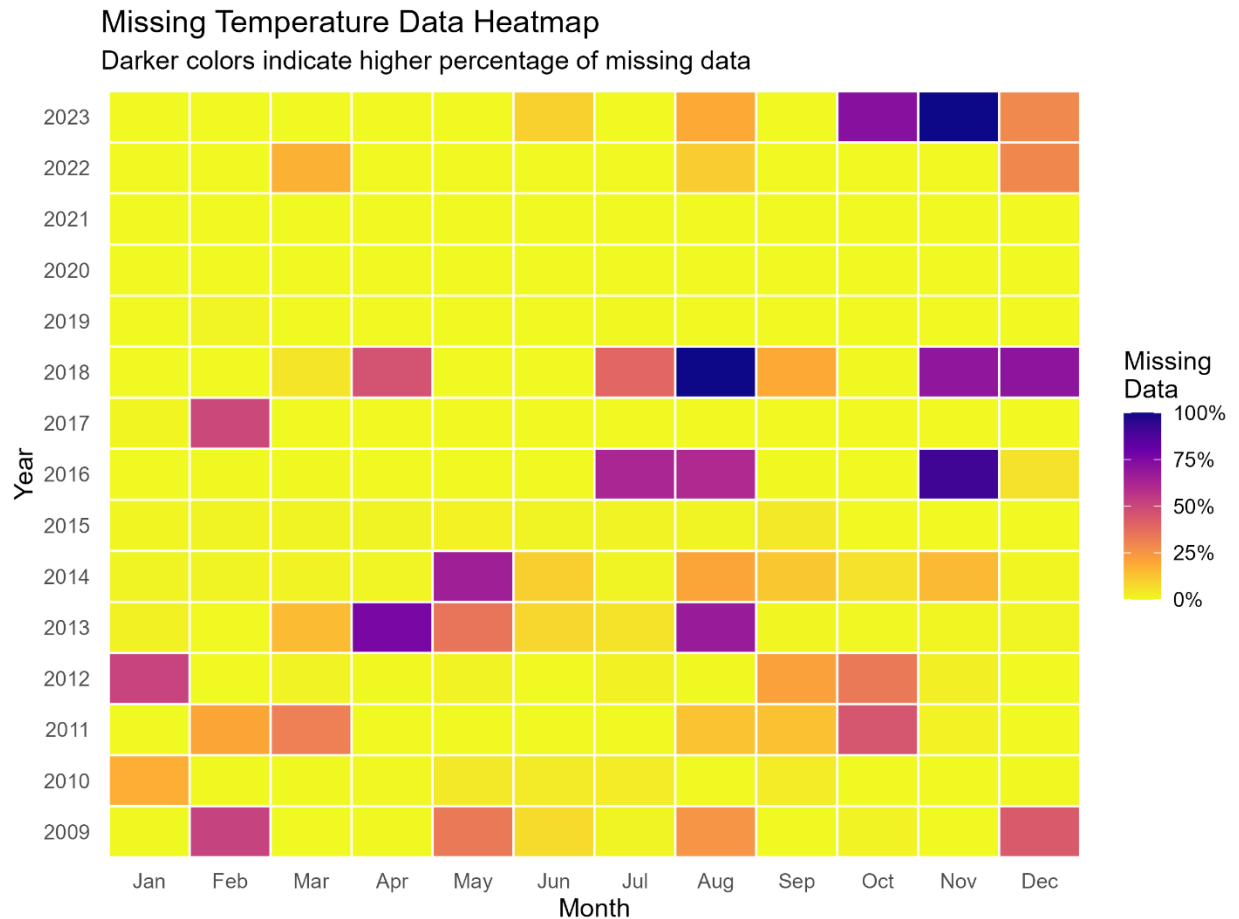


**Figure 2** Heatmap of missing data by month and year. Darker colors indicate a greater proportion of missing data.

## Characterizing seasonality

Seasonality was characterized using a combination of plotting and summary statistics. Monthly means, medians, ranges, and interquartile ranges were calculated. Raw data were plotted by day of year to evaluate how seasonal patterns differed across years. These comparisons allowed us to describe both the timing and magnitude of seasonal variation in temperature.

## Extreme Days Analysis

### Extreme Days Classification and Frequency

A percentile approach was used to classify days as having non-extreme or extreme temperatures. Note that the designation of a day as non-extreme or extreme using this method is a purely statistical approach and not related to ecological thermal thresholds for a specific species or aquatic community. First, dates missing more than 20% of their observations were excluded from the analysis to ensure robust estimates of daily mean temperatures. Daily mean water temperatures were then calculated. The time series was then divided into a reference or a recent time period. Data collected from 2009-2018 were classified as reference while data from 2019-2023 were classified as recent, yielding 3,160 days for the reference period and 1,734 days for the recent period. Percentiles for mean daily water temperatures for the reference period were then calculated. Days with mean daily temperatures below the 5<sup>th</sup> percentile were considered extreme cold days. Days above the 95<sup>th</sup> percentile were considered extreme warm days. This means that approximately 316 (10%) of the days from 2009-2018 were considered extreme. Assuming that the distribution of water temperatures was consistent across years for the reference period, this reduces to around 15.8 days per year for each extreme type (31.6 total extreme days per year).

The 5<sup>th</sup> and 95<sup>th</sup> percentiles temperature thresholds calculated using the reference period data were then compared to daily mean temperatures for the recent period. If there were no differences between reference and recent periods, then around 10% of the mean daily temperatures (around 174 days) would be expected to fall outside the 5<sup>th</sup> and 95<sup>th</sup> percentile thresholds of the reference data. When extreme temperature days (either warm or cold) exceeded 5% of the days for the reference period, that extreme type showed an increase. When they fell below 5%, that extreme type showed a decrease compared to the reference period.

### Percentile Selection Considerations

The choice of which percentiles to use to define extreme events inherently has some degree of subjectivity. The Intergovernmental Panel on Climate Change suggest using the 10<sup>th</sup> and 90<sup>th</sup> percentiles as thresholds for extreme events, although they acknowledge that these thresholds may not apply for all datasets, and domain knowledge should ultimately guide the decision (IPCC, 2014; Camuffo et al., 2020). We selected the 5<sup>th</sup> and 95<sup>th</sup> percentiles to have a relatively conservative threshold while still retaining enough data to make a robust comparison between reference and recent conditions. We believe this decision was appropriate in the context of this study, but we recognize that these thresholds may not be appropriate for other datasets or the current dataset once additional data are added. An iterative approach where different percentile thresholds are compared would provide an estimate of how results might change depending on the

definition of an extreme value. However, this type of sensitivity analysis was beyond the scope of this study. If the goal of a similar analysis is to determine if trends in water temperature are affecting specific species of concern, thresholds should be determined by the species' thermal tolerance range rather than by water temperature percentiles.

### Extreme Temperature Runs

The quantity and duration of extreme temperature days was evaluated by calculating the number of runs, defined as two or more consecutive days of extreme temperatures, and the number of days each run persisted. While the number of extreme cold or warm days in the reference period was defined at 5% for each extreme type, the number of runs depends on the timing of extreme days and can differ across years. Mean run frequency and duration were calculated for reference and recent years and compared. Yearly counts of runs were also calculated to compare within-period variability.

### Extreme Day Intensity

The magnitude of extreme days was also evaluated for the reference and recent periods using an intensity score. For each extreme day, intensity was calculated by subtracting the daily mean from the 5<sup>th</sup> and 95<sup>th</sup> percentile threshold values for the reference period and then taking the absolute value of the result (Equation 1). This means that intensity values are always positive, with larger values indicating a greater difference between the threshold vs observed temperature. Average intensity scores for both extreme types were then compared between periods to quantify changes in intensity over time, as well as across years more generally. The ecological significance of increases in extreme event intensity is highly dependent on the species or ecosystem process of interest, and the potential ecological effects of changing intensity are not addressed in this report.

Equation 1:

$$Intensity_i = |reference\ temp.\ threshold - daily\ mean\ temp_i|$$

where  $Intensity_i$  is the intensity score for an extreme day ( $i$ ),  $daily\ mean\ temp_i$  is the daily mean temperature of an extreme day, and  $reference\ temp.\ threshold$  is the value of either the 5<sup>th</sup> or 95<sup>th</sup> percentile of the daily mean temperatures for the reference period. The absolute value of the difference between  $daily\ mean\ temp_i$  and  $reference\ temp.\ threshold$  is taken so that all intensity values are positive.

### Temporal Trend Analysis

While the extreme days analysis helps inform how temperature extremes are changing, a single value representing the relative change in mean temperature is also often desired due to its intuitive interpretation. Therefore, we tested for a temporal trend in the average change in water temperature per year. Weekly means were selected over

individual points or daily averages to reduce temporal autocorrelation which could not be robustly controlled otherwise. Averaging across weeks does reduce data resolution and statistical power, but this approach provides greater confidence in model results. Changes in mean water temperature related to seasonality were accounted for by the inclusion of a covariate representing the time of year during which an observation was recorded

## Data Pre-processing for a Trend Analysis

Weekly mean water temperature was calculated using data collected from 2009-01-07 through 2023-12-31. Sundays were selected to represent the start date for each week. Weeks missing more than 20% of the possible 672 observations were excluded from the analysis. We removed observations from the analysis if they were recorded before the first Sunday or after the last Sunday in the dataset. This resulted in weekly mean data for 2009-01-07 and 2023-12-31 being removed.

Several additional variables were calculated for inclusion in the temporal model to create a “trend” variable, account for seasonality, and account for the temporal autocorrelation between weeks. First, dates for each Sunday were converted to decimal dates. A decimal date is a way of representing a date as a single decimal number, where the integer part represents the year and the fractional part represents the position within that year. For example, 2009-01-01 would be represented as 2009.000 (the start of the year), while 2009-12-31 would be approximately 2009.997. To account for seasonality, the fractional part of the decimal dates (i.e., values after the decimal) was extracted by subtracting the year component from each decimal date. The results were assigned to a variable called “fraction of the year” (abbreviated as *fracyear*) which represented the time of year that a sample was recorded. To create the “trend” variable, decimal dates were centered by subtracting the overall mean decimal date, yielding a variable called “centered decimal date”. A statistically significant relationship between centered decimal date and weekly mean temperature would indicate the presence of a temporal trend in weekly mean temperature. A third variable called “week index” was created to identify the chronological position of a week in the record. This variable was needed to inform the model of where gaps between consecutive weeks occurred when decomposing the temporal autocorrelation in the model. Lastly, a variable called “binary seasonality” was created to account for unequal variance in temperatures between a warm season (May-September) and cold season (October-April) identified during data exploration.

## Model Structure

We modeled weekly mean temperature using a generalized additive mixed model (GAMM) with a fixed linear effect of centered decimal date, a cyclic smooth of *fracyear* with a basis dimension of  $k = 5$ , and a random intercept for year. Sonde model was originally included as a covariate to account for the replacement of the YSI 6600 with the new YSI

EXO2 but ultimately was removed due to collinearity with centered decimal date. Residual errors were modeled with a Gaussian distribution, allowing for unequal variance by binary season. A first-order autoregressive (AR(1)) temporal correlation structure using week index as a time covariate was also included to account for temporal dependence between samples.

The final model was (equation 2):

$$Y_{i,t} \sim \beta_0 + \beta_1 \text{Centered Decimal Date}_{i,t} + f(\text{fracyear}_i) + b_t + \epsilon_{i,t}$$

where  $Y_{i,t}$  is the mean temperature for week ( $i$ ) of year ( $t$ ),  $\beta_0$  is the global intercept,  $\beta_1$  is the fixed effect slope for centered decimal date,  $f(\text{fracyear}_i)$  is a cyclic smooth function for fracyear, modeled with a cubic spline basis ( $k = 5$ ),  $b_t$  is a random intercept adjustment for each year.

The residual error ( $\epsilon_{i,t}$ ) structure of the model is assumed to follow a Gaussian distribution with (equation 3):

$$\epsilon_{i,t} \sim N(0, \sigma_{\epsilon_s}^2)$$

where  $\sigma_{\epsilon_s}^2$  varies by binary season ( $s$ ). The residuals follow an AR(1) correlation structure (equation 4):

$$\epsilon_s = \rho \epsilon_{s-1} + \eta_s$$

Where  $\epsilon_s$  is the residual at time  $s$ ,  $\rho$  is the autocorrelation coefficient,  $\epsilon_{s-1}$  is the residual of the preceding time point, and  $\eta_s$  is random noise

All models were built using the *mgcv* package in the statistical software program R.

### Choosing the Basis Dimension

The basis dimension  $k$  (i.e., how “wiggly” the regression line is allowed to be) is critical for covariates that include smoothers as the choice can influence the trend estimate and model validation. The basis dimension for fracyear was selected using leave-one-out blocked cross-validation where years were considered blocks (Bergmeir and Benítez, 2012; Roberts et al., 2017; Liu and Zhou, 2024). We first defined a grid of basis dimensions  $k \in \{k_i, \dots, k_{10}\}$  where  $k_i$  is the minimum number of basis dimensions needed for a cyclic spline. For our model,  $k$  ranged from 3 to 10. For each unique year  $y$ , we fit the model to all data excluding year  $y$  (training set) and evaluated predictions on the held-out year  $y$  (test set). This preserves temporal dependence within training/test splits. Each iteration of this process is considered a “fold”. For the held-out year, we generated marginal predictions (i.e., predictions from the GAM component, without year-specific random effects) and calculated the fold-level prediction root mean square error (RMSE). For each  $k$ , we averaged fold-level errors to obtain the cross-validated error  $\bar{E}(k)$ . The

standard error of the cross-validated error for each basis dimension,  $SE(k)$ , was calculated as (equation 5):

$$SE(k) = \frac{SD\{E_y(k)\}}{\sqrt{N_{\text{folds}}}}$$

where  $SD\{E_y(k)\}$  is the standard deviation of the cross-validation errors obtained across folds for a given basis dimension  $k$ , and  $\sqrt{N_{\text{folds}}}$  is the square root of the number of cross-validation folds. The lowest RMSE observed was  $k = 5$  while  $k = 3$  had the lowest SE. In cross-validation, it is common to choose the simplest model that is within 1 SE of the minimum cross-validation error, ergo indicating that  $k = 3$  would be preferred. However, heteroskedasticity was present when  $k = 3$  (Appendix Figure 1). Therefore, we defaulted to  $k = 5$  as this minimized the cross-validation error and lessened the apparent heteroskedasticity.

## Model Validation

The normality assumption was evaluated using qqplots of deviance residuals. Normalized residuals (i.e., response residuals scaled to account for random effects) for the parametric portion of the model were extracted and used to verify homogeneity in the overall model residuals and when plotted against each candidate variable. Normality of the random intercepts for year was verified via density plot. Temporal autocorrelation was assessed via autocorrelation and partial autocorrelation function plots of the normalized residuals. No major violations of model assumptions were observed.

## Checking for Bias Due to 2023 Heatwave

A prolonged heatwave occurred during the summer of 2023 during which water temperatures were extremely high (daily means ranging from 34.2 to 36.6 °C). We defined the dates of the heatwave based on the criteria for an extreme temperature run as outlined in the Extreme Days Analysis section of the methods. Based on this criterion, the dates for the extreme temperature run were 2023-06-16 through 2023-08-1 (44 days). This extreme warm run was hotter and persisted longer than any other run in the dataset.

Given the presence of this run near the end of the record and its severity, we further considered its influence when applying a linear regression model to the data. Linear regression can be sensitive to extreme values or unusual patterns which occur near the beginning or end of the record, resulting in regression slopes being pulled higher or lower than would be if the anomaly was not present. Such influence can lead to incorrect conclusions about the presence or absence of a monotonic temporal trend. Therefore, we reran the model with the 2023 extreme warm run removed. The estimates from the models with and without the 2023 heatwave were then compared to evaluate the influence of those weeks. Excluding the 2023 heatwave did not change the direction or statistical significance of the trend estimate (Appendix Table 2), and neither model violated model

assumptions. Therefore, dates recorded during the heatwave were retained in the subsequent results presented here.

### An Exploratory Analysis Using a Nonlinear Term for the Trend Variable

During validation of the final model, some minor non-linearity remained when plotting model residuals against the trend variable. To ensure that this minor violation of model assumptions was not biasing results, an alternative method in which centered decimal date was modeled with a nonlinear term was also explored. Model comparison indicated that both models performed similarly, and the simpler model with a linear term for centered decimal year was preferred to maintain clearer interpretability. The details of this alternate model, its advantages to the current model, and limitations can be found in Appendix II.

## Results

### Characterizing seasonality

Water temperature at S. Bird Island was characterized by hot, stable summers and variable winters with occasional cold snaps (Fig. 3). The site experiences two distinct seasonal periods, a warm season from May through late September and a cool season from early October through April (Fig. 4). A key characteristic distinguishing the warm vs cool season temperature dynamics is the difference in temperature variability. Beginning around late April, temperatures steadily rise while variability decreases (Fig. 5, Table 1). Temperatures then generally remain high and stable through September. By October, temperatures steadily decrease and become more variable, with peak variability occurring in January and February. This increase in variability is likely due to intermittent cold fronts which rapidly reduce air temperatures, cooling the waters of the upper Laguna Madre (Tissot et al. 2007).

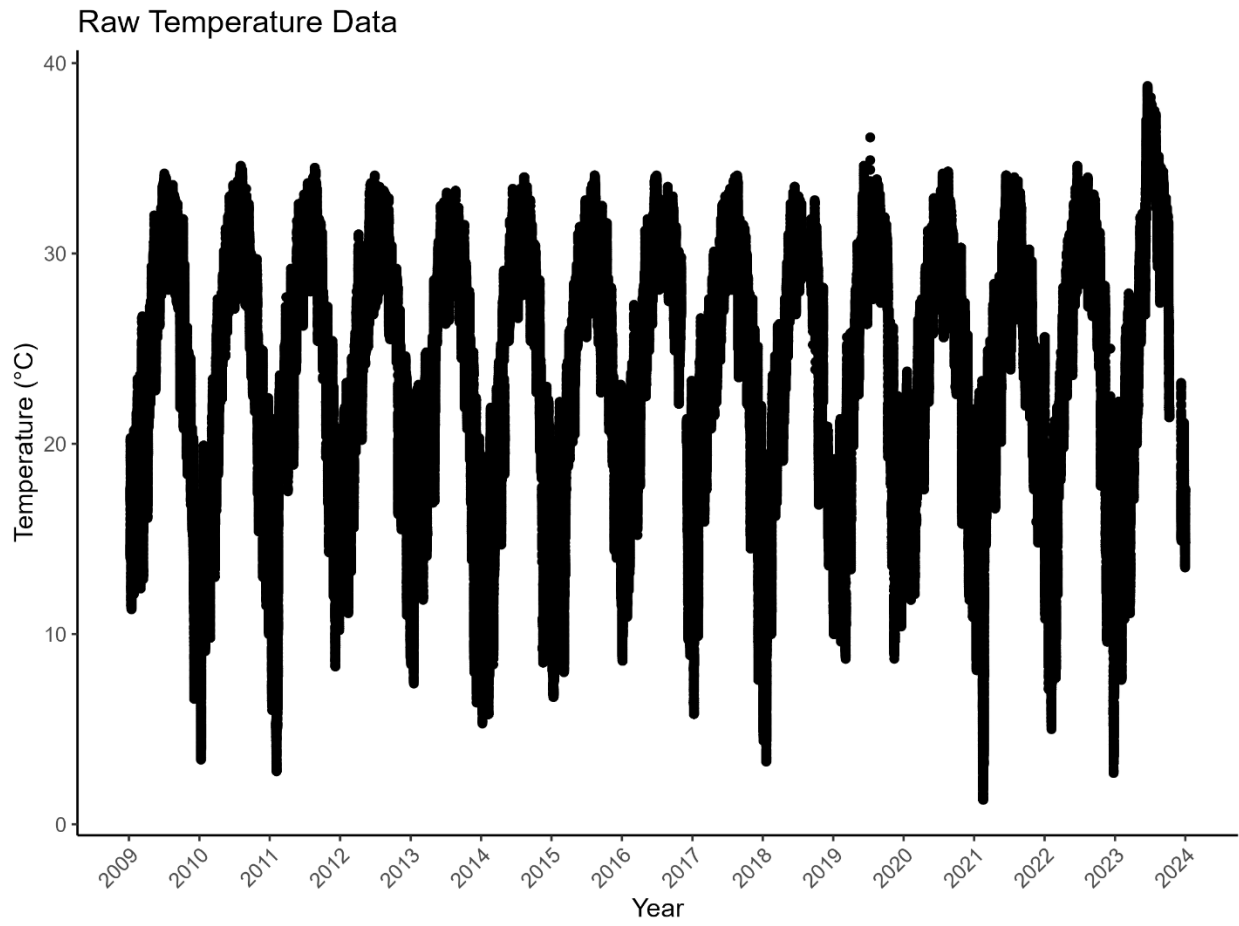


**Figure 3** Plot of raw temperature data by date.

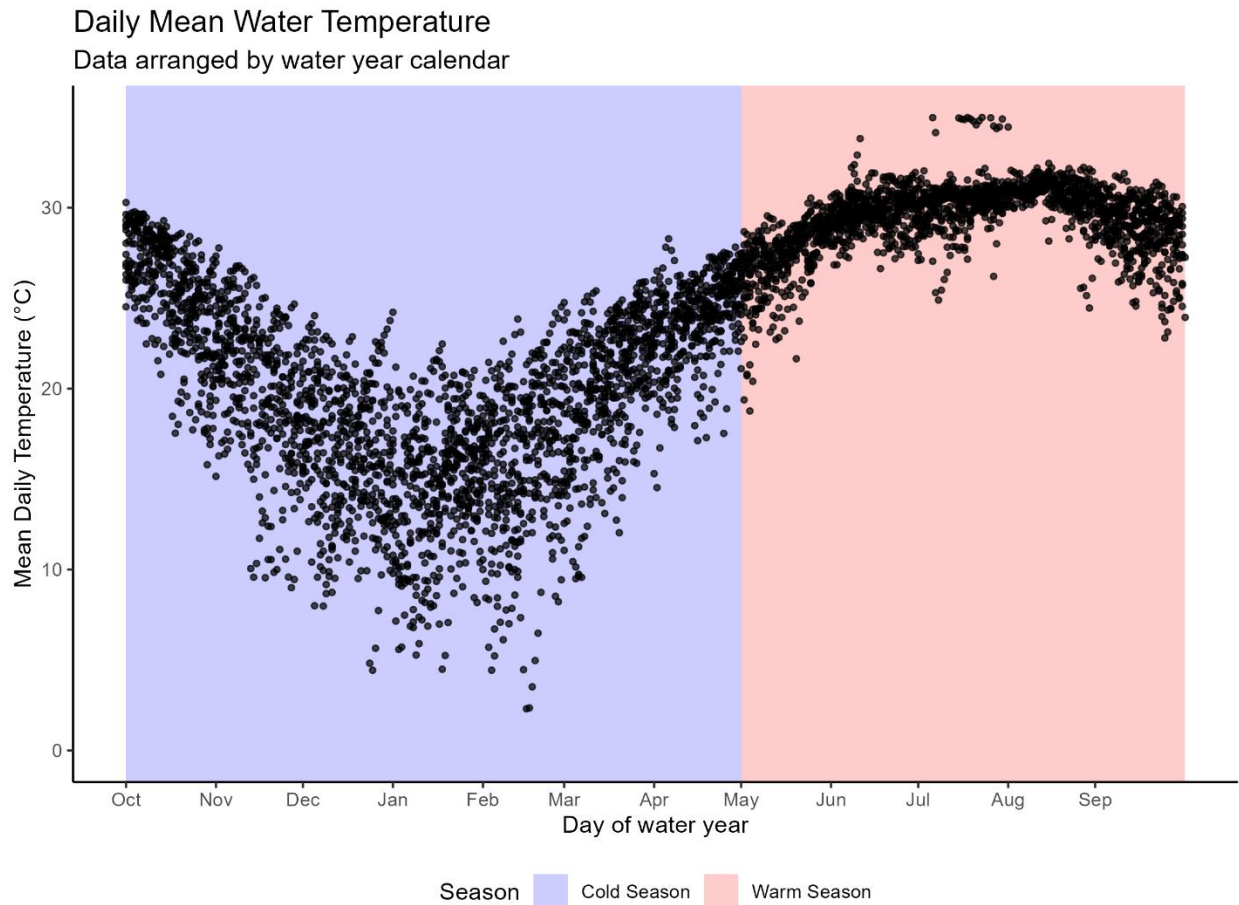


**Figure 4** Daily mean temperature by day of year using the water year calendar. Colored shading used to indicate binary cold and warm seasons. The anomalously high points for June through August are the 2023 heatwave.

**Table 1** Summary statistics for water temperature (°C) by month and binary season (data combined across years).

| Period                      | Possible Obs | Non-missing Obs | Mean | Median | Min  | Max  | Range | Std Dev | Q1   | Q3   | IQR <sup>2</sup> | Missing (%) |
|-----------------------------|--------------|-----------------|------|--------|------|------|-------|---------|------|------|------------------|-------------|
| <b>Monthly</b>              |              |                 |      |        |      |      |       |         |      |      |                  |             |
| January                     | 43,968       | 41,738          | 14.7 | 15.0   | 3.3  | 25.6 | 22.3  | 3.7     | 12.4 | 17.3 | 4.9              | 5.1         |
| February                    | 40,608       | 37,196          | 16.2 | 16.7   | 1.3  | 26.6 | 25.3  | 4.2     | 13.7 | 19.3 | 5.6              | 8.4         |
| March                       | 44,640       | 42,444          | 20.3 | 20.9   | 8.0  | 28.0 | 20.0  | 3.5     | 18.2 | 22.9 | 4.7              | 4.9         |
| April                       | 43,200       | 39,564          | 23.9 | 24.2   | 13.4 | 31.0 | 17.6  | 2.5     | 22.6 | 25.7 | 3.1              | 8.4         |
| May                         | 44,640       | 40,413          | 27.2 | 27.5   | 17.0 | 32.6 | 15.6  | 2.1     | 26.1 | 28.6 | 2.5              | 9.5         |
| June                        | 43,200       | 42,017          | 30.0 | 29.8   | 25.8 | 38.8 | 13.0  | 1.9     | 28.8 | 31.0 | 2.2              | 2.7         |
| July                        | 44,640       | 41,174          | 30.5 | 30.2   | 23.9 | 38.2 | 14.3  | 2.0     | 29.2 | 31.6 | 2.4              | 7.8         |
| August                      | 44,640       | 35,143          | 30.7 | 30.6   | 23.5 | 35.6 | 12.1  | 1.5     | 29.8 | 31.8 | 2.0              | 21.3        |
| September                   | 43,200       | 41,051          | 28.8 | 29.0   | 21.9 | 34.3 | 12.4  | 1.9     | 27.7 | 30.0 | 2.3              | 5.0         |
| October                     | 44,640       | 39,876          | 25.4 | 25.9   | 14.5 | 32.0 | 17.5  | 3.1     | 23.7 | 27.8 | 4.1              | 10.7        |
| November                    | 43,200       | 35,087          | 20.3 | 20.8   | 8.0  | 29.8 | 21.8  | 3.9     | 17.9 | 23.1 | 5.2              | 18.8        |
| December                    | 44,640       | 39,253          | 16.5 | 16.6   | 2.7  | 25.1 | 22.4  | 3.7     | 13.9 | 19.4 | 5.5              | 12.1        |
| <b>Seasonally</b>           |              |                 |      |        |      |      |       |         |      |      |                  |             |
| Warm Season<br>(May-Sept.)  | 220,320      | 199,798         | 29.4 | 29.5   | 17.0 | 38.8 | 21.8  | 2.3     | 28.2 | 30.8 | 2.6              | 9.3         |
| Cool Season<br>(Oct.-April) | 304,896      | 275,158         | 19.6 | 19.9   | 1.3  | 32.0 | 30.7  | 5.2     | 16.0 | 23.6 | 7.6              | 9.8         |

<sup>1</sup>CV = Coefficient of Variation (Standard Deviation / Mean × 100)

<sup>2</sup>IQR = Interquartile Range (Q3 - Q1)

Data period: 2009-01-08 to 2023-12-31

## Extreme Days Analysis

Between reference and recent time periods, extreme temperature patterns were characterized by general shift towards warmer summers and winters, with fewer extreme cold days and more frequent extreme warm days. Overall daily mean temperature increased by 0.5 °C between periods. However, this general warming trend was also accompanied by an increase in the magnitude of extreme cold and warm days, along with an overall increase in variability. Extreme warm days occurred more frequently, tended to be warmer (overall daily mean increased from 31.3°C to 32.4°C), and occurred in runs more often. Extreme cold days tended to occur less frequently and were generally warmer (overall daily mean decreased from 9.7 °C to 9.0 °C) but were more variable and had a higher overall intensity (Fig. 5)

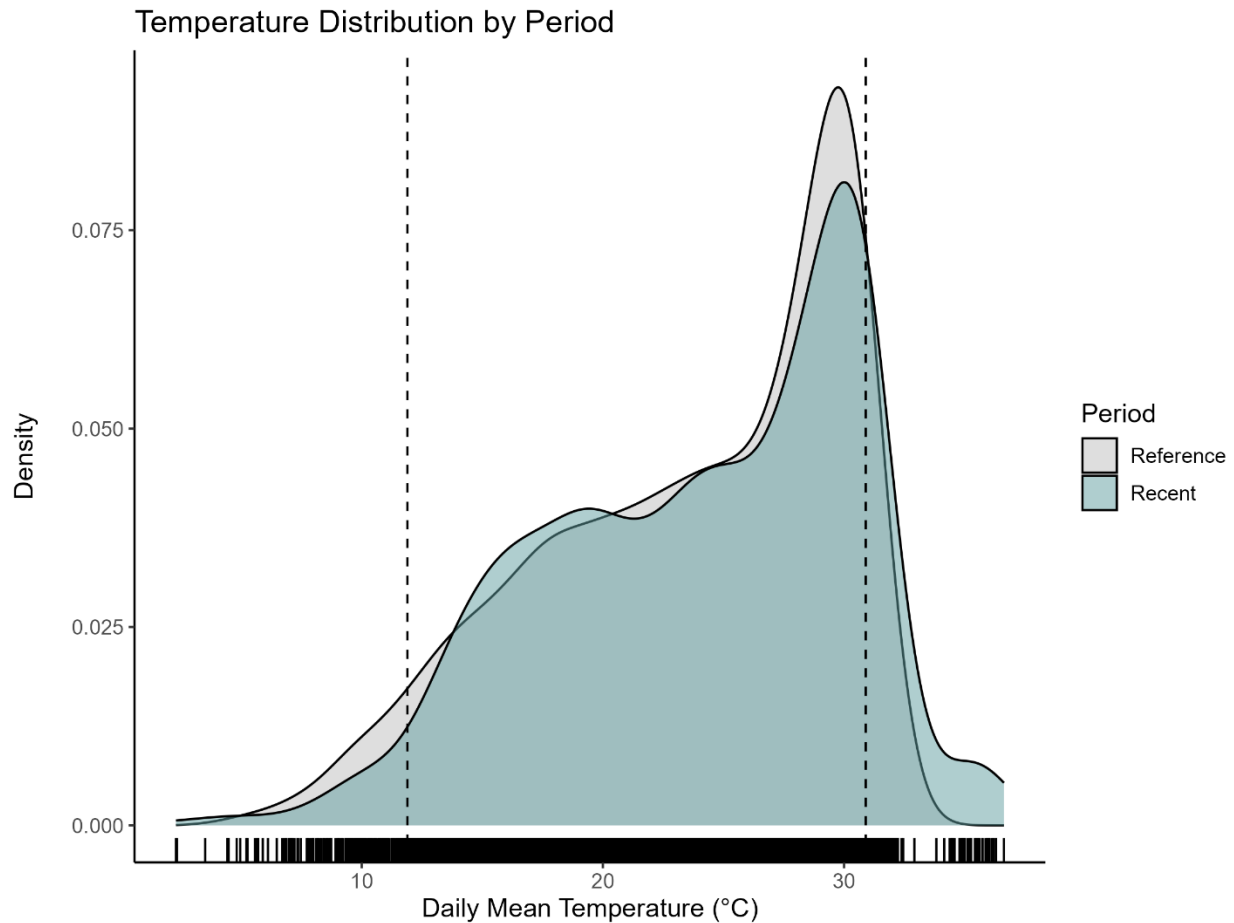


**Figure 5** Density plot of daily mean temperature by time period (reference is 2009-2018 and recent is 2019-2023). Dashed black lines indicate the 5<sup>th</sup> and 95<sup>th</sup> percentile thresholds (11.9 and 30.9 °C, respectively) for the reference time period. Rug plots along the x-axis indicate individual days.

### Extreme Day Frequency

Daily mean temperatures for the 5<sup>th</sup> and 95<sup>th</sup> percentiles of the reference period were 11.9 °C and 30.9 °C, respectively (Fig. 6). The annual average number of extreme warm days more than doubled (15.8 to 38.4), while extreme cold days declined (15.8 to 13.8; Fig. 7). The 2023 heatwave inflated the average number of extreme warm days per year. When 2023 is excluded, the average annual frequency of extreme warm days decreases to 27.0. Overall, mean annual extreme cold day occurrence showed a moderate decline, whereas extreme warm day occurrences per year increased.

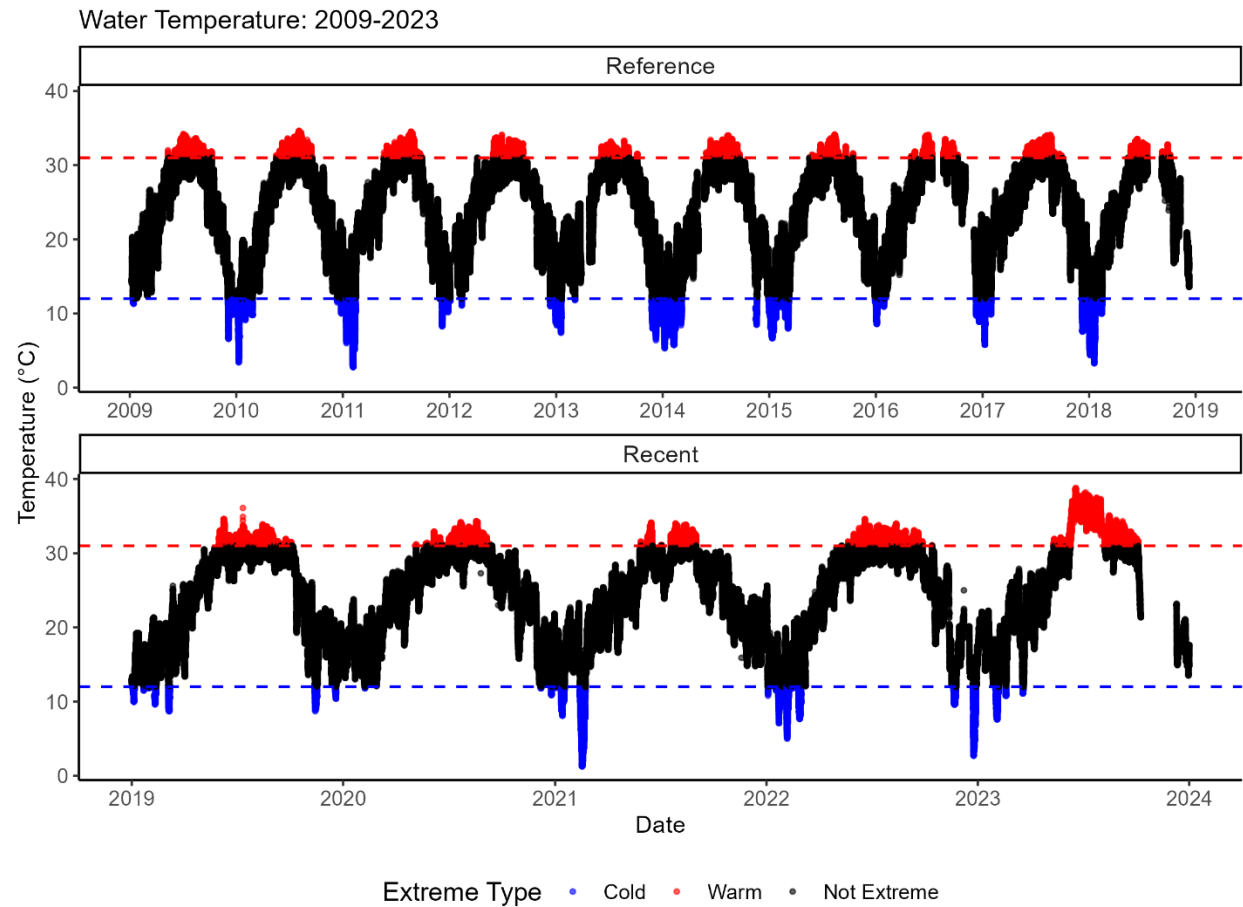


**Figure 6** Daily mean temperature for the reference and recent period. Red points are extreme warm days and blue points are extreme cold days. The red and blue dashed lines indicate the 11.9 °C and 30.9 °C thresholds for the 5<sup>th</sup> and 95<sup>th</sup> percentiles for the reference period, respectively.

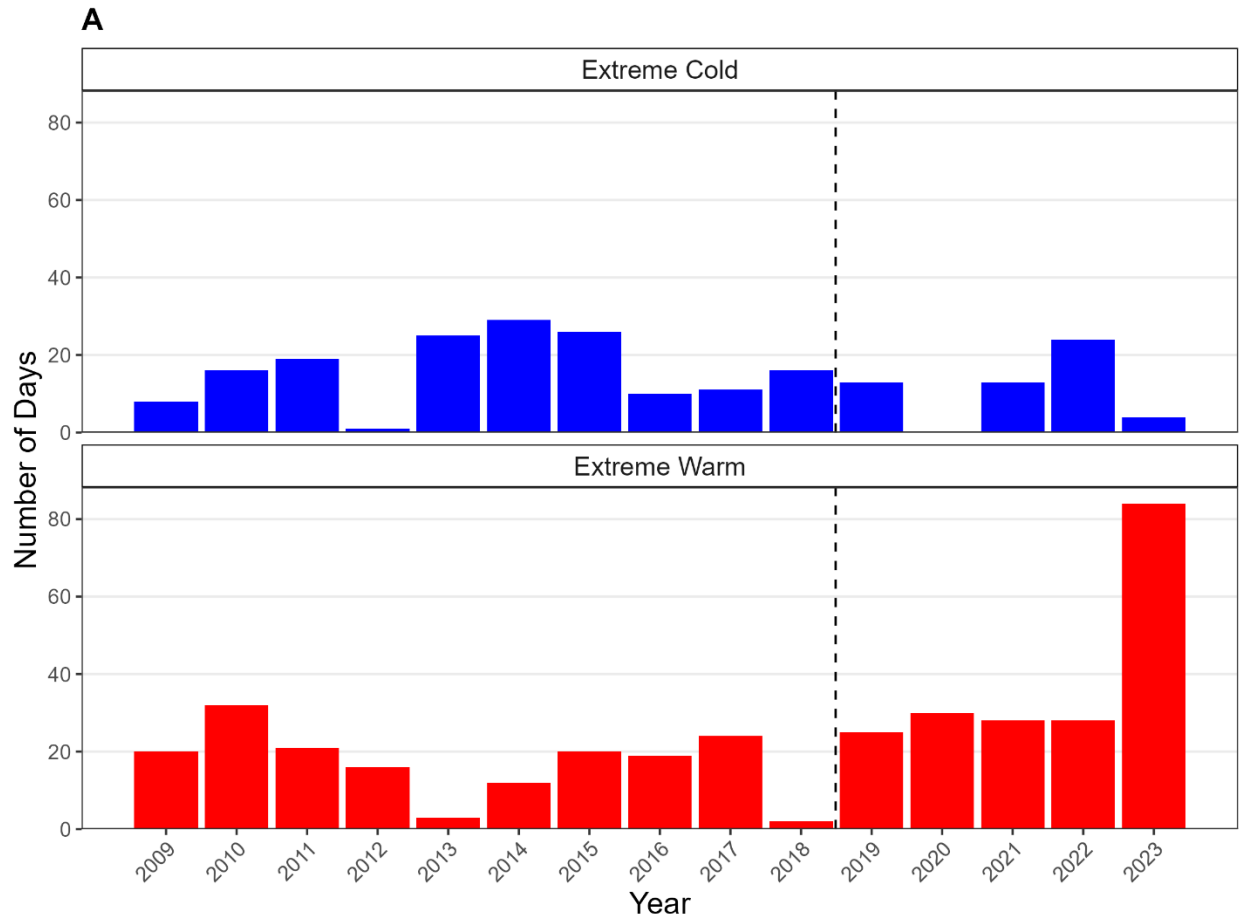


**Figure 7** Frequency of extreme days by year. Dashed black lines indicate the divide between reference and recent periods.

### Extreme Day Runs

Runs, defined as two or more consecutive days of average extreme temperatures, were quantified as number of runs per year (i.e., frequency) and number of days per run (i.e., duration) for years where runs were present. Run frequency and duration differed between reference and recent time periods. Warm-run frequency increased, with mean annual run count rising from 3.4 to 4.8 runs per year while cold run frequency decreased from 4.1 to 3.0 runs per year. The pattern in run duration for warm extremes was strongly influenced by the 2023 heatwave (Fig. 8, Table 4). Including 2023, mean warm-run duration increased from 4.5 to 7.6 days between the reference and recent time periods, driven by a 44-day event in 2023 that substantially inflated the average. Excluding 2023, mean warm-run duration in the recent period was 5.3 days and cold run duration was 4.5 days.

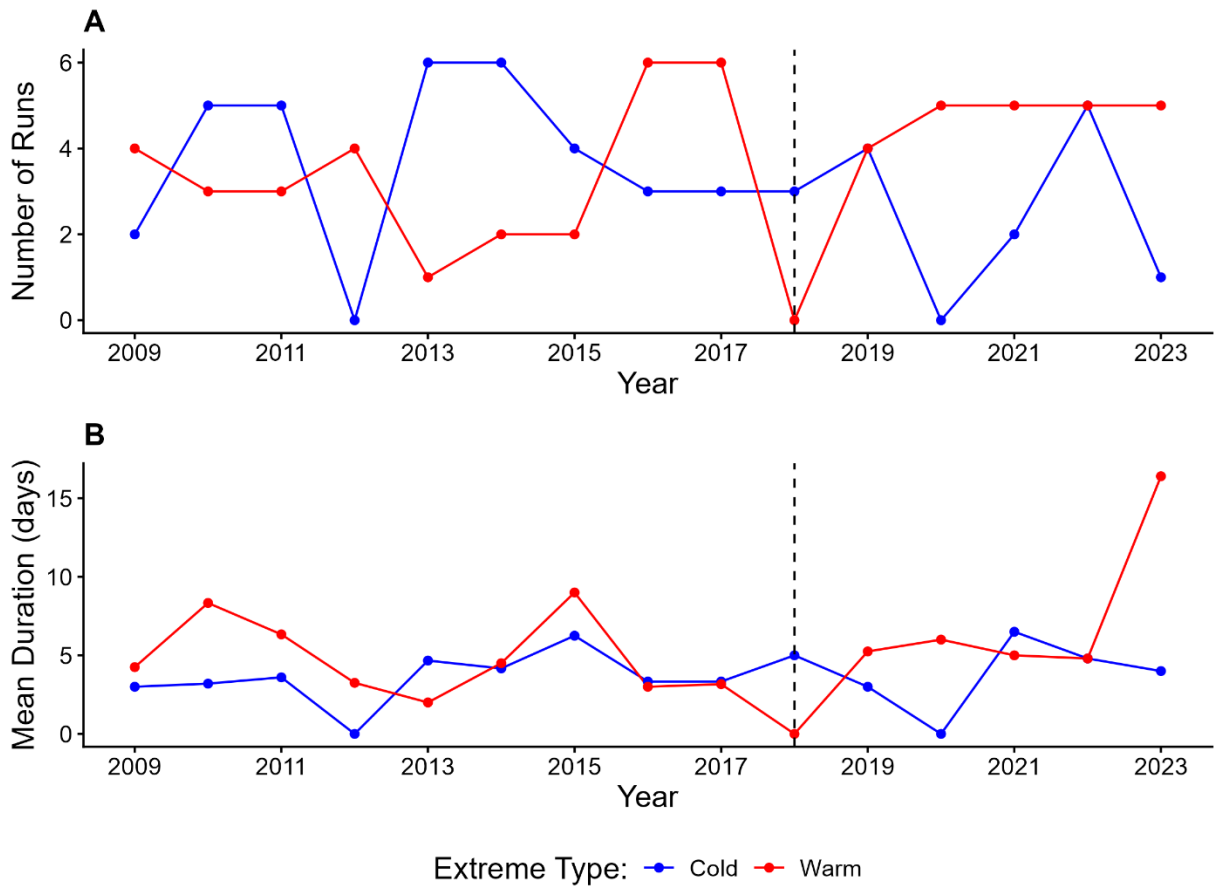


**Figure 8** Run frequency (**A**) and mean run duration (**B**) by year. Dashed black lines indicate the divide between recent and reference periods.

**Table 2** Summary statistics for the run frequency and duration analysis by period (reference = 2009-2018 and recent = 2019-2023) for years with runs. Units are in number of days long.

| <b>Run Frequency</b>    |                       |           |            |            |                       |           |            |            |
|-------------------------|-----------------------|-----------|------------|------------|-----------------------|-----------|------------|------------|
| <b>Period</b>           | <b>Extreme Warm</b>   |           |            |            | <b>Extreme Cold</b>   |           |            |            |
|                         | <b>Annual average</b> | <b>SD</b> | <b>Min</b> | <b>Max</b> | <b>Annual average</b> | <b>SD</b> | <b>Min</b> | <b>Max</b> |
| Reference               | 3.4                   | 1.7       | 1          | 6          | 4.1                   | 1.5       | 2          | 6          |
| Recent                  | 4.8                   | 0.4       | 4          | 5          | 3.0                   | 1.8       | 1          | 5          |
| Recent – excluding 2023 | 4.8                   | 0.5       | 4          | 5          | 2.8                   | 3.7       | 2          | 5          |

| <b>Run Duration</b>   |                       |           |            |            |                       |           |            |            |
|-----------------------|-----------------------|-----------|------------|------------|-----------------------|-----------|------------|------------|
|                       | <b>Extreme Warm</b>   |           |            |            | <b>Extreme Cold</b>   |           |            |            |
|                       | <b>Annual average</b> | <b>SD</b> | <b>Min</b> | <b>Max</b> | <b>Annual average</b> | <b>SD</b> | <b>Min</b> | <b>Max</b> |
| Reference             | 4.5                   | 3.7       | 2          | 18         | 4.4                   | 3.0       | 2          | 18         |
| Recent                | 7.6                   | 8.9       | 2          | 44         | 4.1                   | 1.8       | 2          | 9          |
| Recent excluding 2023 | 5.3                   | 4.4       | 2          | 19         | 4.5                   | 1.9       | 2          | 9          |

### Extreme Day Intensity

Extreme day intensity increased between reference and recent periods, with cold extremes having greater magnitude and variability than warm extremes (Fig. 10, Table 5, Appendix Fig. 2). Mean cold-day intensity rose from 2.2°C to 2.9°C, while mean warm-day intensity increased from 0.4°C to 1.5°C.

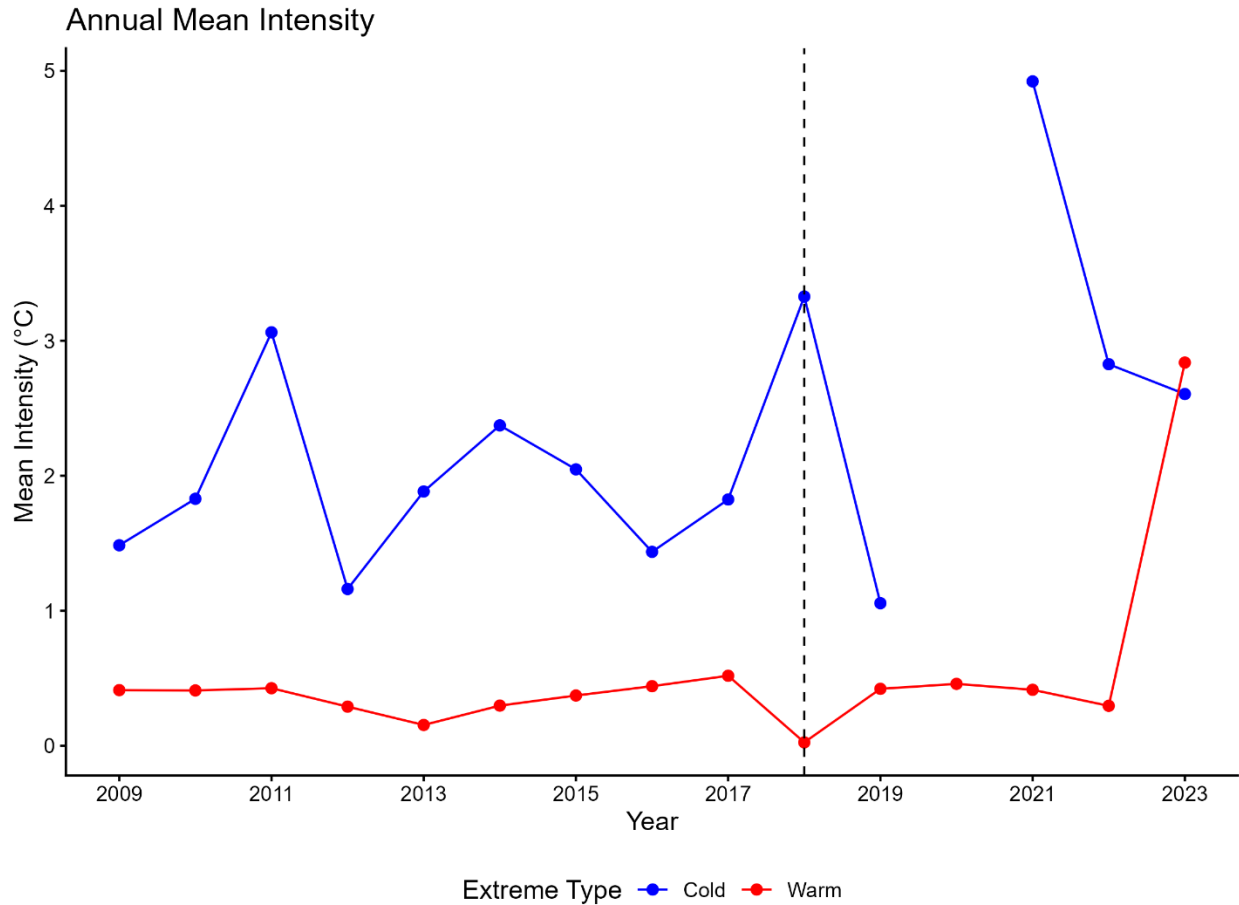


**Figure 9** Annual mean intensity of extreme days by extreme type. The dashed black line separates reference and recent periods. Note that 2020 did not have any extreme cold days.

## Temporal Trend Analysis

The GAMM model of weekly mean water temperatures explained 87% of the variability in the data ( $\text{adj-}R^2 = 0.87$ ). The estimated temporal trend was not statistically significant at  $\alpha = 0.05$ , with confidence intervals for the trend estimating ranging from a - 0.01 to 0.15 °C change in mean water temperature per year ( $t = 1.65s$ ,  $p = 0.099$ ). There was also a strong non-linear relationship between fracyear and weekly mean water temperature (effective degrees of freedom = 2.99,  $F = 876.7$ ,  $p < 0.001$ ), meaning that water temperature was strongly influenced by time of year (Fig. 11).

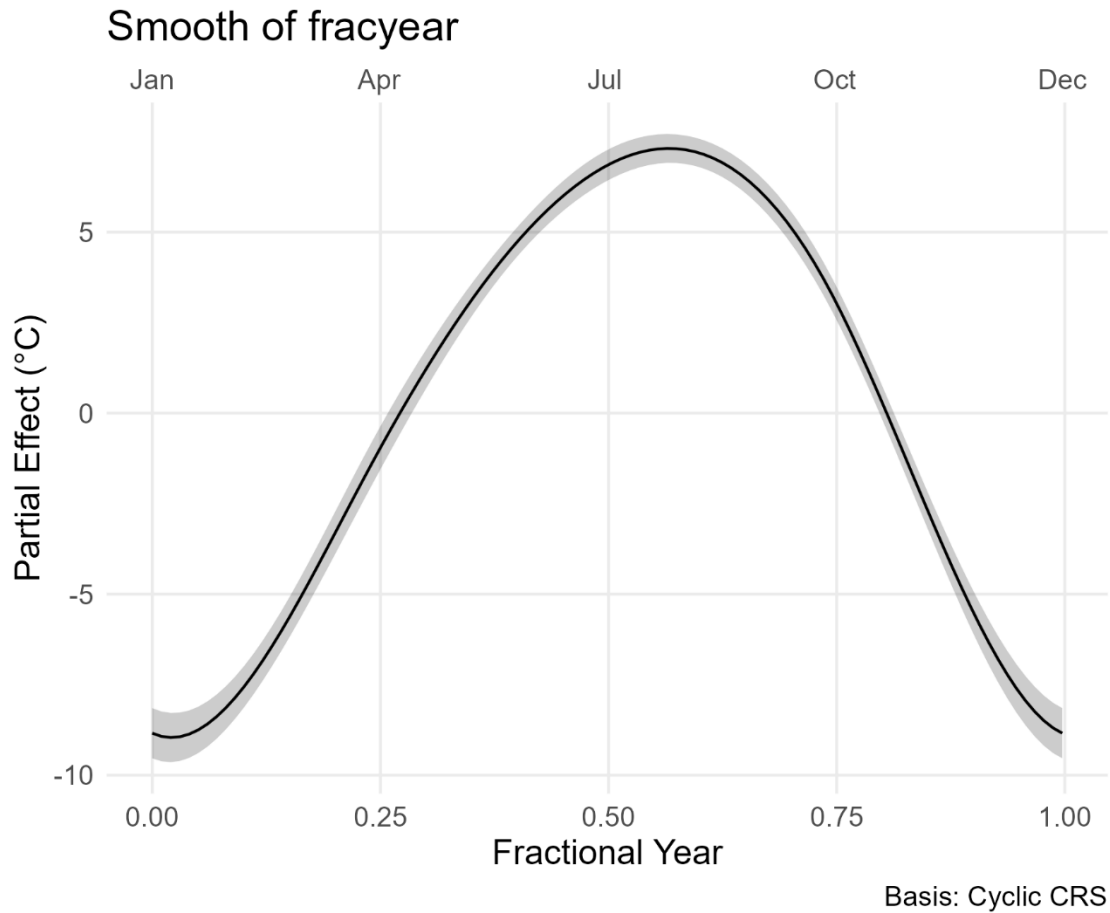


**Figure 10** Partial effect of fracyear on weekly mean temperature. This figure depicts the non-linear relationship between time of year and water temperature. The y-axis indicates the change from the model’s global mean temperature (i.e., intercept) in °C. The lower x-axis shows the fractional time of year, and the upper x-axis shows the corresponding month. Note that month positions are approximate.

The random effects portions of the model also provided important insights into the temperature data. Year-to-year random variation was high ( $SD = 0.63\text{ }^{\circ}\text{C}$ ), indicating some interannual variability beyond the linear trend. Among individual years, 2021 and 2023 were the most distinct, with annual means  $1.06\text{ }^{\circ}\text{C}$  below and  $1.44\text{ }^{\circ}\text{C}$  above the global mean, respectively. Seasonal variability also differed substantially. Residual standard deviation for the warm season was 61% lower than in the cold season (scaling factor = 0.39), reflecting more stable warm-season temperatures. Further, variability due to temporal autocorrelation in the residuals was adequately modeled by an AR(1) correlation structure, which estimated a correlation coefficient of  $\phi = 0.43$ .

# Key Take-aways

## Extreme Days Analysis

- There were more extreme warm days and fewer extreme cold days per year in the recent period (2019-2023) compared to the reference period (2009-2018).
- Mean intensity increased by 1.1 °C for warm extremes and 0.7 °C for cold extremes.
- Run frequency for extreme warm days increased between periods from an average of 3.4 to 4.8 runs per year.
- Run frequency for extreme cold days decreased between periods from an average of 4.1 to 3.0 runs per year.
- Run duration for extreme warm days increased between periods from an average of 4.5 to 7.6 days per year. This increase was partially driven by an unusually warm run that occurred in 2023. If 2023 were removed from the data, the average run duration for the reference period would have increased to 5.3 runs per year.
- Run duration for extreme cold days decreased between periods from an average of 4.4 to 4.1 days per year.

## Temporal Trend Analysis

- There was no statistically significant overall trend in weekly mean water temperature.
- The fixed components of the model (centered decimal year and fracyear) explained 87% of the variability in the data.
- The residual standard deviation for the within-year warm season was 61% lower than the within-year cold season, meaning the cold season generally experiences substantially more variable temperatures than the warm season.

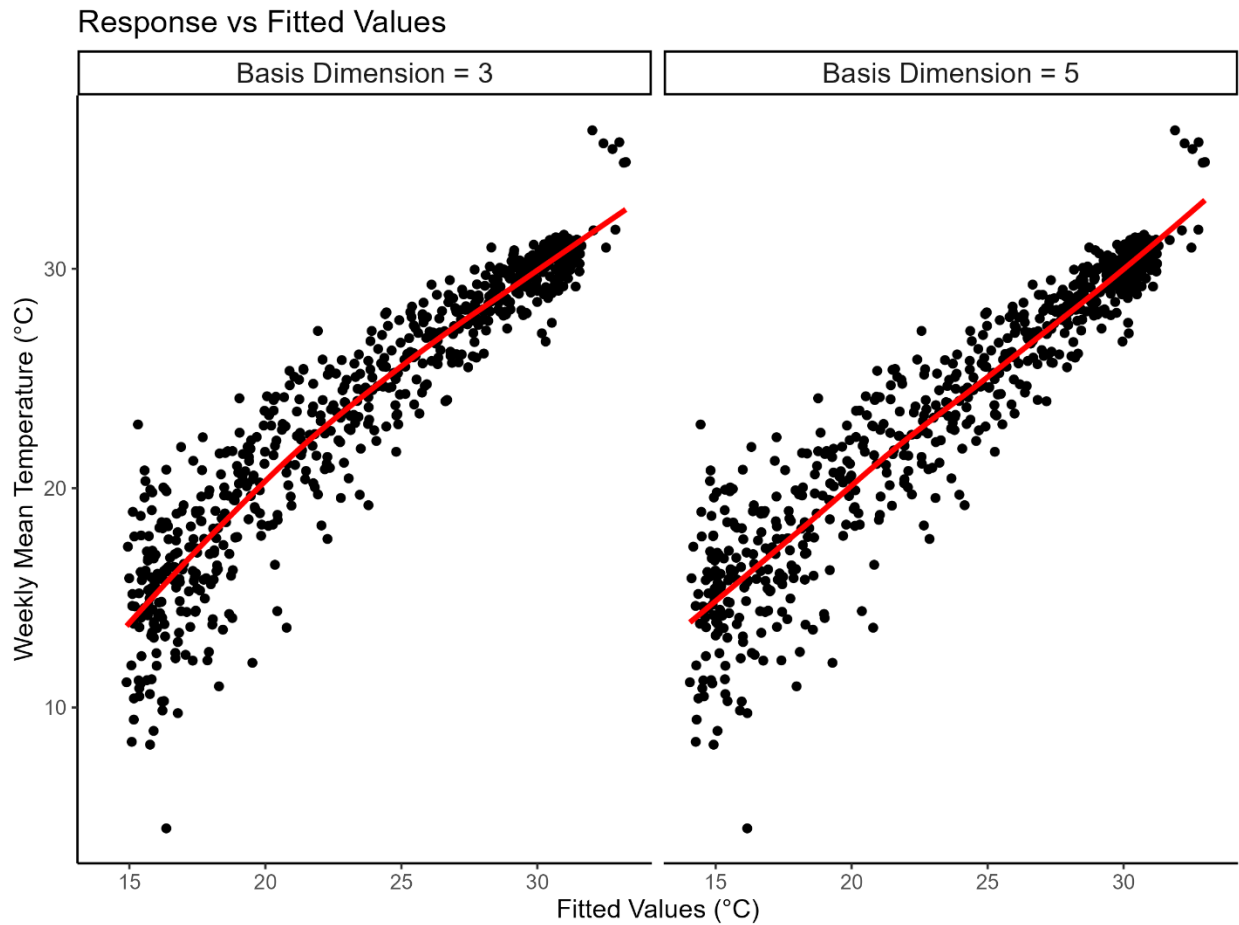
## References

- Abraham, J.P., Baringer, M., Bindoff, N.L., Boyer, T., Cheng, L.J., Church, J.A., Conroy, J.L., Domingues, C.M., Fasullo, J.T., Gilson, J., Goni, G., Good, S.A., Gorman, J.M., Gouretski, V., Ishii, M., Johnson, G.C., Kizu, S., Lyman, J.M., Macdonald, A.M., Minkowycz, W.J., Moffitt, S.E., Palmer, M.D., Piola, A.R., Reseghetti, F., Schuckmann, K., Trenberth, K.E., Velicogna, I., Willis, J.K., 2013. A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Reviews of Geophysics* 51, 450–483.  
<https://doi.org/10.1002/rog.20022>
- Adams, J., & Tissot, P. (2007). Monitoring and predictive modeling of water temperatures in the Laguna Madre (Federal Aid Grant No. T-9). Texas Parks and Wildlife Department.
- Barnett, T.P., Pierce, D.W., AchutaRao, K.M., Gleckler, P.J., Santer, B.D., Gregory, J.M., Washington, W.M., 2005. Penetration of Human-Induced Warming into the World's Oceans. *Science* 309, 284–287. <https://doi.org/10.1126/science.1112418>
- Bergmeir, C., Benítez, J.M., 2012. On the use of cross-validation for time series predictor evaluation. *Information Sciences* 191, 192–213.  
<https://doi.org/10.1016/j.ins.2011.12.028>
- Boucek, R.E., Rehage, J.S., 2014. Climate extremes drive changes in functional community structure. *Glob Change Biol* 20, 1821–1831. <https://doi.org/10.1111/gcb.12574>
- Camuffo, D., Becherini, F., Della Valle, A., 2020. Relationship between selected percentiles and return periods of extreme events. *Acta Geophys.* 68, 1201–1211.  
<https://doi.org/10.1007/s11600-020-00452-x>
- Collier, C.J., Waycott, M., 2014. Temperature extremes reduce seagrass growth and induce mortality. *Marine Pollution Bulletin* 83, 483–490.  
<https://doi.org/10.1016/j.marpolbul.2014.03.050>
- Echavarria-Heras, H.A., Solana-Arellano, E., Franco-Vizcaíno, E., 2006. The Role of Increased Sea Surface Temperature on Eelgrass Leaf Dynamics: Onset of El Niño as a Proxy for Global Climate Change in San Quintín Bay, Baja California. *Bulletin, Southern California Academy of Sciences* 105, 113–127.  
[https://doi.org/10.3160/0038-3872\(2006\)105\[113:TROISS\]2.0.CO;2](https://doi.org/10.3160/0038-3872(2006)105[113:TROISS]2.0.CO;2)
- Feher, L.C., Osland, M.J., Johnson, D.J., Grace, J.B., Guntenspergen, G.R., Stewart, D.R., Sklar, F.H., 2024. Nonlinear patterns of surface elevation change in coastal wetlands: The value of generalized additive models for quantifying rates of change. *Estuaries and Coasts* 47, 1893–1902.
- Gleckler, P.J., Durack, P.J., Stouffer, R.J., Johnson, G.C., Forest, C.E., 2016. Industrial-era global ocean heat uptake doubles in recent decades. *Nature Clim Change* 6, 394–398. <https://doi.org/10.1038/nclimate2915>
- Grilo, T.F., Cardoso, P.G., Dolbeth, M., Bordalo, M.D., Pardal, M.A., 2011. Effects of extreme climate events on the macrobenthic communities' structure and functioning of a temperate estuary. *Marine Pollution Bulletin* 62, 303–311.  
<https://doi.org/10.1016/j.marpolbul.2010.10.010>

- Hamilton, S.K., 2010. Biogeochemical implications of climate change for tropical rivers and floodplains. *Hydrobiologia* 657, 19–35. <https://doi.org/10.1007/s10750-009-0086-1>
- Hutchings, A.M., de Vries, C.S., Hayes, N.R., Orr, H.G., 2024. Temperature and dissolved oxygen trends in English estuaries over the past 30 years. *Estuarine, Coastal and Shelf Science* 306, 108892. <https://doi.org/10.1016/j.ecss.2024.108892>
- IPCC. 2014. Climate change 2014: synthesis report. In: Mach KJ, Meyer LA, Pachauri RK, Planton S, von Stechow C (eds) Annex II: Glossary. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva
- Liu, S., Zhou, D.J., 2024. Using cross-validation methods to select time series models: Promises and pitfalls. *Brit J Math & Statis* 77, 337–355. <https://doi.org/10.1111/bmsp.12330>
- Madeira, D., Narciso, L., Cabral, H.N., Vinagre, C., 2012. Thermal tolerance and potential impacts of climate change on coastal and estuarine organisms. *Journal of Sea Research* 70, 32–41. <https://doi.org/10.1016/j.seares.2012.03.002>
- Matear, R.J., Hirst, A.C., 2003. Long-term changes in dissolved oxygen concentrations in the ocean caused by protracted global warming. *Global Biogeochemical Cycles* 17, 2002GB001997. <https://doi.org/10.1029/2002GB001997>
- Murphy, R.R., Perry, E., Harcum, J., Keisman, J., 2019. A Generalized Additive Model approach to evaluating water quality: Chesapeake Bay case study. *Environmental Modelling & Software* 118, 1–13. <https://doi.org/10.1016/j.envsoft.2019.03.027>
- Ni, W., Li, M., Ross, A.C., Najjar, R.G., 2019. Large Projected Decline in Dissolved Oxygen in a Eutrophic Estuary Due to Climate Change. *JGR Oceans* 124, 8271–8289. <https://doi.org/10.1029/2019JC015274>
- Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40, 913–929. <https://doi.org/10.1111/ecog.02881>
- Scanes, E., Scanes, P.R., Ross, P.M., 2020. Climate change rapidly warms and acidifies Australian estuaries. *Nat Commun* 11, 1803. <https://doi.org/10.1038/s41467-020-15550-z>
- Shields, E.C., Parrish, D., Moore, K., 2019. Short-Term Temperature Stress Results in Seagrass Community Shift in a Temperate Estuary. *Estuaries and Coasts* 42, 755–764. <https://doi.org/10.1007/s12237-019-00517-1>
- Testa, J.M., Murphy, R.R., Brady, D.C., Kemp, W.M., 2018. Nutrient- and Climate-Induced Shifts in the Phenology of Linked Biogeochemical Cycles in a Temperate Estuary. *Front. Mar. Sci.* 5, 114. <https://doi.org/10.3389/fmars.2018.00114>
- Texas A&M University Natural Resources Institute, 2014. <http://txlandtrends.org/data>.
- Tissot, P. E., Ball, R., & Adams, J. S. 2007. Monitoring and Predictive Modeling of Water Temperatures in the Laguna Madre. Technical Report for Texas Parks & Wildlife and the Coastal Conservation Association.

- Waltham, N.J., Sheaves, M., 2017. Acute thermal tolerance of tropical estuarine fish occupying a man-made tidal lake, and increased exposure risk with climate change. *Estuarine, Coastal and Shelf Science* 196, 173–181.  
<https://doi.org/10.1016/j.ecss.2017.06.032>
- Watson, J. Keith. 2005. Integration of the North American Bird Conservation Initiative (NABCI) into southeast National Park Service planning and operations. In: Ralph, C. John; Rich, Terrell D., editors 2005. *Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference*. 2002 March 20-24; Asilomar, California, Volume 2 Gen. Tech. Rep. PSW-GTR-191. Albany, CA: U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Research Station: p. 1169-1172

## Appendix I



**Appendix Figure 1** Observed versus fitted weekly mean temperatures for GAMM models with cubic cyclic spline basis dimensions of 3 and 5. Points closer to the diagonal indicate better model fit. The red line should be straight if no heteroskedasticity remains in the model. Note the slight curve in the red line when the basis dimension is set to 3.

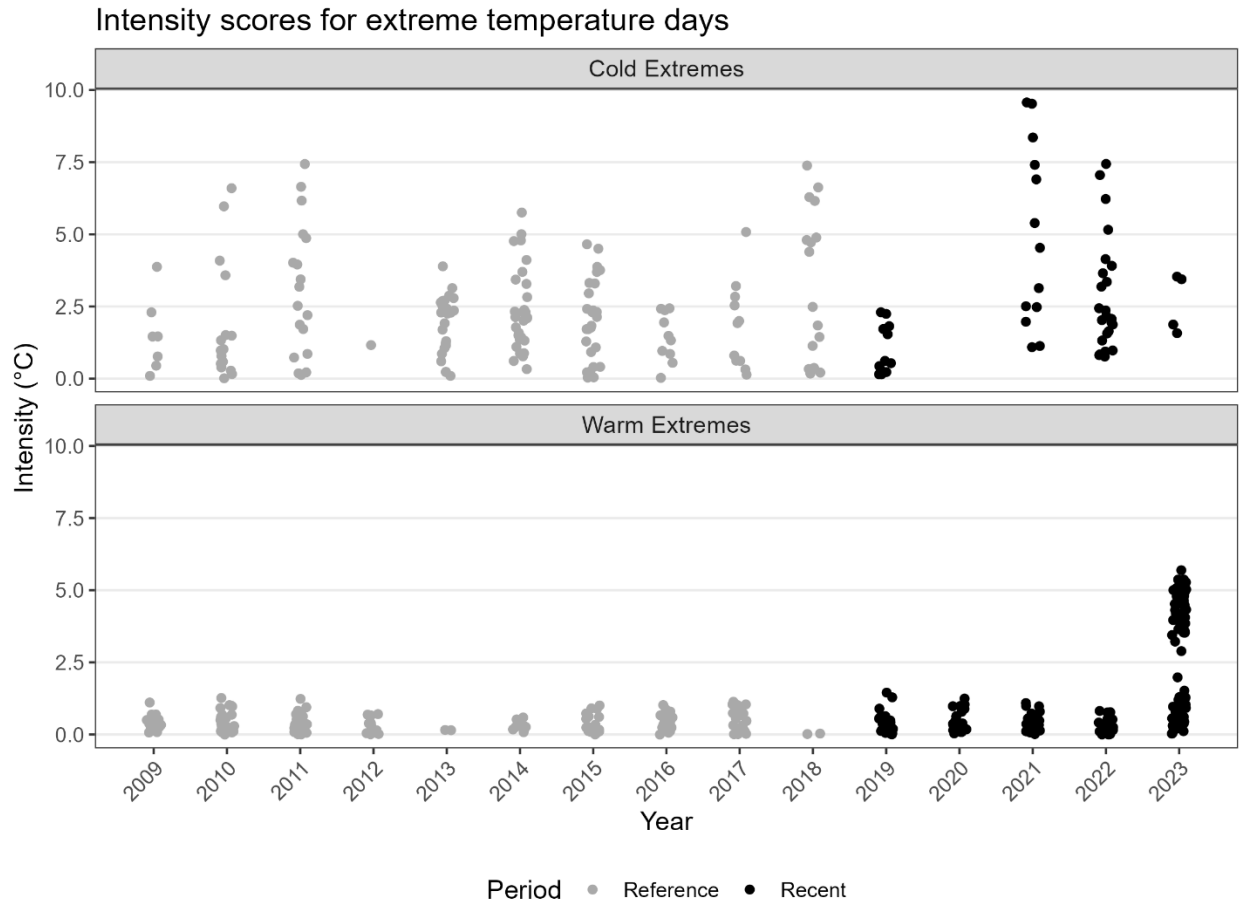
**Appendix Table 1** Summary of missing observations across months.

| Month     | Total Possible Obs. <sup>1</sup> | Missing Values <sup>2</sup> | Non-Missing Values | % Missing <sup>3</sup> |
|-----------|----------------------------------|-----------------------------|--------------------|------------------------|
| January   | 43,968                           | 2,230                       | 41,738             | 5%                     |
| February  | 40,608                           | 3,412                       | 37,196             | 8%                     |
| March     | 44,640                           | 2,196                       | 42,444             | 5%                     |
| April     | 43,200                           | 3,636                       | 39,564             | 8%                     |
| May       | 44,640                           | 4,227                       | 40,413             | 9%                     |
| June      | 43,200                           | 1,183                       | 42,017             | 3%                     |
| July      | 44,640                           | 3,466                       | 41,174             | 8%                     |
| August    | 44,640                           | 9,497                       | 35,143             | 21%                    |
| September | 43,200                           | 2,149                       | 41,051             | 5%                     |
| October   | 44,640                           | 4,764                       | 39,876             | 11%                    |
| November  | 43,200                           | 8,113                       | 35,087             | 19%                    |
| December  | 44,640                           | 5,387                       | 39,253             | 12%                    |
|           |                                  |                             |                    |                        |

<sup>3</sup>Color coding: Blue (low missing %), Yellow-Orange (moderate), Red- Orange (high missing %)

**Appendix Table 2** Comparison of the GAMM with and without data recorded during the 2023 heatwave included. Estimates are coefficient estimates. The estimate for cent\_dec\_date is the trend estimate. Values for the CI column are the 2.5<sup>th</sup> and 97.5<sup>th</sup> percent confidence intervals, respectively. Note that the confidence intervals for cent\_dec\_year cross zero for both models, meaning the true slope may be zero. *p* is the *p*-value for the parameter.

| <i>Model with heatwave included</i>    |                  |               |                  |
|----------------------------------------|------------------|---------------|------------------|
| <b>Parametric coefficients</b>         | <b>Estimates</b> | <b>CI</b>     | <b><i>p</i></b>  |
| (Intercept)                            | 23.76            | 23.37 – 24.14 | <b>&lt;0.001</b> |
| cent_dec_yr                            | 0.07             | -0.01 – 0.15  | 0.990            |
| <u><i>Smooth terms:</i></u>            | <b>EDF</b>       | <b>Ref.df</b> | <b><i>p</i></b>  |
| s(fracyear)                            | 2.99             | 3.00          | <b>&lt;0.001</b> |
| Observations                           | 682              |               |                  |
| R <sup>2</sup>                         | 0.87             |               |                  |
| <i>Model without heatwave included</i> |                  |               |                  |
| <b>Predictors</b>                      | <b>Estimates</b> | <b>CI</b>     | <b><i>p</i></b>  |
| (Intercept)                            | 23.69            | 23.36 – 24.02 | <b>&lt;0.001</b> |
| cent_dec_yr                            | 0.05             | -0.02 – 0.12  | 0.128            |
| <u><i>Smooth terms:</i></u>            | <b>EDF</b>       | <b>Ref.df</b> | <b><i>p</i></b>  |
| s(fracyear)                            | 2.99             | 3.00          | <b>&lt;0.001</b> |
| Observations                           | 677              |               |                  |
| R <sup>2</sup>                         | 0.88             |               |                  |



**Appendix Figure 2** Intensity scores for cold and warm extreme days colored by time period (reference = 2009-2018 and recent = 2019-2023). Points are jittered along the x-axis by 0.1 years for easier visualization. The y-axis coordinates are not jittered.

## Appendix II

### An Exploratory Analysis Using a GAMM with a Non-linear Term for the Trend Variable

For the above trend analysis, including year as a random effect gave the model enough flexibility to meet the assumption of linearity for the centered decimal date variable. However, this approach may become less viable as additional data are added to the record. If future data are such that the model cannot adequately decompose the non-linearity using random effects and a smoother on a variable representing within-year seasonality (e.g., *fracyear*), then a more complicated non-linear GAMM with a smoother on the “trend” variable (e.g., centered decimal date) may be a viable alternative.

Adding a non-linear smoother to the long-term “time” variable would add complexity to model fitting and inference. First, the nonlinear trend cannot be directly interpreted in the context of an overall increase or decrease. Second, the methods available to estimate an overall trend are much more intricate than for a traditional linear model, and it may be difficult for analysts with limited statistical or coding experience to follow or replicate. Third, there are limitations to our ability to estimate a global trend from a smoothed variable, especially if random effects like random intercepts, variance structures, or correlation structure are included (which will likely occur). This results in additional uncertainty in deriving a global trend, particularly uncertainty in the confidence intervals around the basis functions (i.e., the series of slopes which make up the regression line for smoothed parameters).

Despite these added complexities, we have provided an R function to derive a global trend estimate from a model with a smoothed centered decimal date parameter using bootstrapped sampling of the posterior distribution for centered decimal date. This method is an adaptation of the methods used in Murphy et al. 2019, Feher et al. 2023, and Hutchings et al. 2024. The steps are as follows:

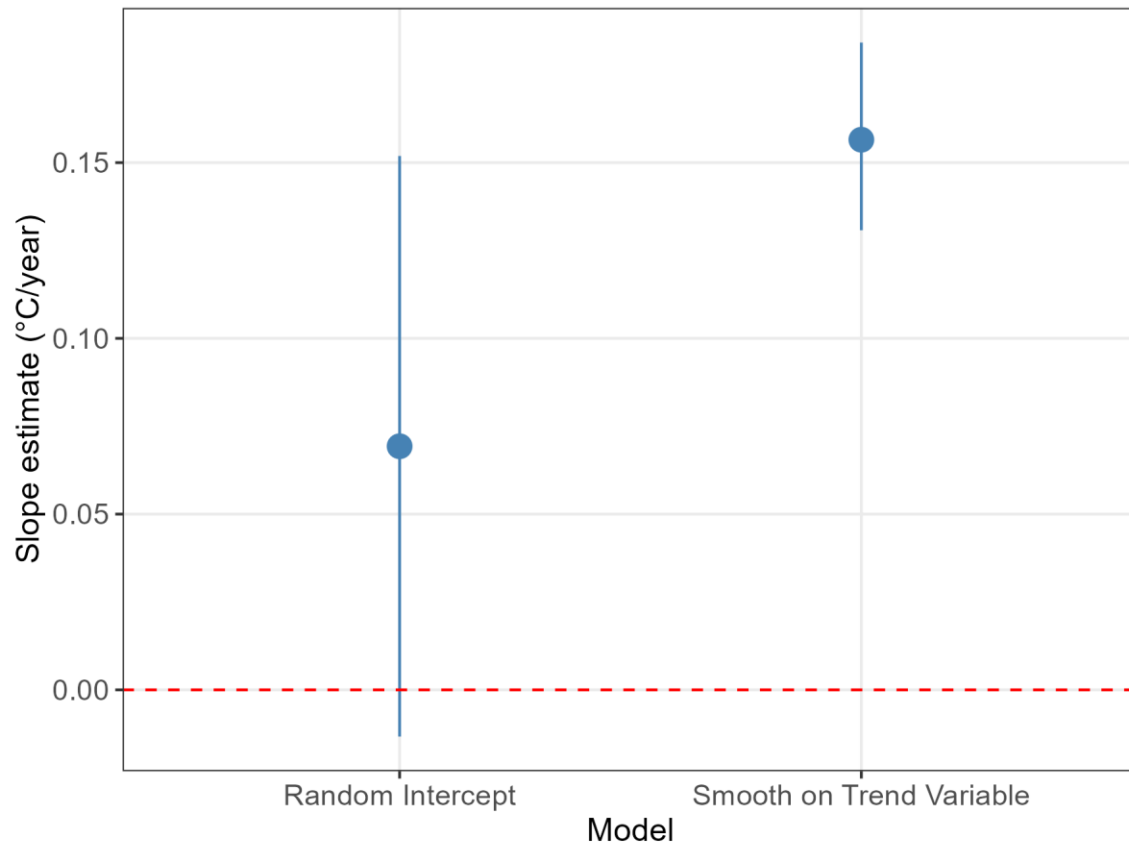
1. Extract the observed values of the predictor variable of interest (e.g., centered decimal date) and create a data frame containing a grid of evenly spaced values ranging from the minimum and maximum of the observed values. Include all other predictors but with their values held at their means.
2. Using the *gratia* package, calculate the first derivatives (i.e. slope) of the basis functions and their standard errors for each grid point (function defaults to 400 grid points).
3. For each grid point, simulate uncertainty in the new data frame by drawing *x* many samples from the normal distribution centered on that grid point (function

defaults to 2,000 draws). This creates a posterior distribution of possible slopes at each grid point.

4. Apply density weighting if desired. Weighting by density gives regions in the record with many observations more weight than data sparse regions when calculating the global slope. To accomplish this, the density function is used to compute a smooth estimate of the distribution of the observed data over the grid values range. The density values for each grid point are assigned weights (their density value) which are then normalized by dividing the density value for a grid point by the sum of the weights. If weighting by density is not desired, set the “density” argument in the function to density = “uniform”. Use density weighting when you want the trend estimate to reflect periods with more observations (and thus more certainty), and use uniform weighting when you want all time periods to contribute equally regardless of how much data were collected.
5. For each grid point, compute an average of the slopes from the x many draws from the normal distribution centered on that grid point. This gives you x many possible slopes for each grid point.
6. Then take the mean of all possible slopes across all the grid points. If using the default settings, this would be the mean of 2,000 draws \* 400 grid points = 800,000 possible slopes.
7. Calculate the 2.5% and 97.5% quantiles of the 800,000 possible slopes to get 95% confidence intervals on the estimate of the global slope.

When this method of estimating a global slope was applied to the GAMM model with a smoother on centered decimal year, the returned estimate indicated a significant warming trend in water temperature of 0.16 °C per year (CI = 0.13 and 0.18 °C) (Appendix Fig. 3). In this report, we presented the model without a smooth on centered decimal date rather than this model because the model without a smoother on centered decimal date was valid and had a more intuitive interpretation. However, this method of adding a smoother to the “trend” variable may be preferred in the future as climate conditions change if the relationship between time and water temperature becomes more non-linear.

### Global Slope Estimates for Random Intercept and Smoothed Models



**Appendix Figure 3** Estimates of the slope of centered decimal date for the final model (labeled here as “Random Intercept”) and a model with a tensor smooth on centered decimal year. Note that the confidence intervals around the estimate for the “Smooth on Trend Variable” model are likely wider than were calculated by the slope estimation procedure.

## Some warnings about this method:

1. The *gratia* package does not currently have a method to account for the random portion of the error structures of GAMMs when calculating the first derivatives and their standard errors. This means that the standard errors will likely be underestimated, meaning confidence intervals will also be narrower than expected.
2. Because the temporal autocorrelation error structure of the model cannot be accounted for when taking the first derivatives (presently), there will be residual temporal dependence between grid points (i.e., a violation of the assumption of independent samples). This will result in lower standard errors for the estimated slopes for each grid point, which will result in narrower confidence intervals. Whether this is a severe enough violation to nullify model results may be case dependent.

### **Potential Method to Correct for Narrow Confidence Intervals**

To address these limitations related to artificially narrow confidence intervals, a sensitivity analysis could be conducted where confidence intervals for the estimated overall slope are iteratively widened across a user-defined range. If slope estimates and confidence intervals are relatively stable across iterations, this provides evidence that inference about the overall slope is probably valid. Large changes in the overall slope estimate and confidence intervals across iterations would imply that inferences about an overall slope would be unreliable. Note that a criticism of this method is that the inflation values are inherently subjective.