Co-occurrence of aquatic heatwaves with atmospheric heatwaves, low dissolved oxygen, and low pH events in estuarine ecosystems

Spencer J. Tassone[1*], Alice F. Besterman[1,2,3], Cal D. Buelo[1], Jonathan A. Walter[1], Michael L. Pace[1]

[1]Department of Environmental Sciences, University of Virginia, Charlottesville, VA

[2]Woodwell Climate Research Center, Woods Hole, MA

[3]Buzzards Bay Coalition, New Bedford, MA

*Corresponding Author:

Spencer J. Tassone (ORCID iD: https://orcid.org/0000-0002-9340-7170)

Department of Environmental Sciences

University of Virginia

291 McCormick Rd

Box 400123

Charlottesville, VA 22904

sjt7jc@virginia.edu

Acknowledgments

This work was supported by funding from the University of Virginia's Environmental Resilience Institute and an NSF grant supporting the LTER-Virginia Coast Reserve project (DEB 1832221). We thank Dr. Patricia L. Wiberg for discussions on heatwave and Fourier transform analysis, Dr. Christopher Sherwood for discussions on Fourier transform analysis and code, Dat Ha for early revisions, and helpful, positive feedback received by two anonymous reviewers. Associate Editor Dr. Arnoldo Valle-Levinson provided guidance and raised questions that improved our analysis and revision of the initial manuscript. We thank the NERRS staff for assistance and advice.

Summary of Comments on Tassone_EstuarineHW_ESCO_R3_final.pdf

Page: 1

Number: 1 Author: Reviewer Subject: Sticky Note Date: 6/16/2023 12:03:09 PM

Tassone, S. J., Besterman, A. F., Buelo, C. D., Walter, J. A., & Pace, M. L. (2022). Co-occurrence of aquatic heatwaves with atmospheric heatwaves, low dissolved oxygen, and low PH events in estuarine ecosystems. *Estuaries and Coasts*, *45*(3), 707-720. https://doi.org/10.1007/s12237-021-01009-x

Conflict of Interest

The authors declare that they have no conflict of interest.

Abstract

Heatwaves are increasing in frequency, duration, and intensity in the atmosphere and marine environment with rapid changes to ecosystems occurring as a result. However, heatwaves in estuarine ecosystems have received little attention despite the effects of high temperatures on biogeochemical cycling and fisheries and the susceptibility of estuaries to heatwaves given their low volume. Likewise, estuarine heatwave co-occurrence with extremes in water quality variables such as dissolved oxygen (DO) and pH have not been considered and would represent periods of enhanced stress. This study analyzed 1,440 station years of high-frequency data from the National Estuarine Research Reserve System (NERRS) to assess trends in the frequency, duration, and severity of estuarine heatwaves and their co-occurrences with atmospheric heatwaves, low DO, and low pH events between 1996-2019. Estuaries are warming faster than the open and coastal ocean, with an estuarine heatwave mean annual occurrence of 2 ± 2 events, ranging up to 10 events per year, and lasting up to 44 days (mean duration = 8 days). Estuarine heatwaves co-occur with an atmospheric heatwave 6-71% of the time, depending on location, with an average estuarine heatwave lag range of 0-2 days. Similarly, low DO or low pH events co-occur with an estuarine heatwave 2-45% and 0-18% of the time, respectively, with an average low DO lag of 3 ± 2 days and low pH lag of 4 ± 2 days. Triple co-occurrence of an estuarine heatwave with a low DO and low pH event was rare, ranging between 0-7% of all estuarine heatwaves. Amongst all the stations, there have been significant reductions in the frequency, intensity, duration, and rate of low DO event onset and decline over time. Likewise, low pH events have decreased in frequency, duration, and intensity over the study period, driven in part by reductions in all severity classifications of low pH events. This study provides the first

baseline assessment of estuarine heatwave events and their co-occurrence with deleterious water quality conditions for a large set of estuaries distributed throughout the U.S.

Keywords

Heatwave, Disturbance, Estuary, Co-occurrence, Water Quality, National Estuarine Research Reserve System (NERRS)

Introduction

Greenhouse gas concentrations are expected to continue to increase in the atmosphere throughout the first half of the 21st century, which will lead to climate warming and an increase in water temperatures (Hoegh-Guldberg et al. 2014). As aquatic systems absorb excess heat, a higher frequency, duration, intensity, and spatial extent of discrete extreme warm-water periods known as heatwaves is expected (Frölicher et al. 2018; Oliver et al. 2018; Collins et al. 2019; Oliver et al. 2021). Many physiological, ecological, and biogeochemical processes are temperature-dependent, so extremes in temperature, especially at the higher end, can compound other stressors within aquatic ecosystems. The solubility of dissolved gasses is inversely related to temperature, such that as temperature increases, dissolved oxygen (DO) concentration decreases. Furthermore, ecosystem respiration is a temperature-dependent process that consumes DO and produces carbon dioxide (CO_2) which when mixed with water forms carbonic acid, potentially resulting in lower pH depending on buffering conditions. Aquatic heatwaves push organisms past their thermal tolerance (Madeira et al. 2012) and may trigger multiple stressors in aquatic ecosystems such as low DO and low pH events, which are known to cause local population extinction (Llansó 1992, Baird et al. 2004), regional mass mortality (Garrabou et al. 2009), and trophic cascades (Piatt et al. 2020).

While many studies have reported positive temperature trends for aquatic environments, only recently have studies examined extremes in water temperature referred to in coastal and oceanic environments as marine heatwaves. Marine heatwaves (MHWs) have a standardized definition as discrete periods when water temperature is above a seasonally referenced 90th percentile for five or more consecutive days (Hobday et al. 2016). These relatively short-lived, extreme events have attracted research interest, in part because of their potential for disproportionate impacts on ecosystems (Easterling et al. 2000, Jentsch et al. 2007) and because future warming is likely to increase the frequency, intensity, duration, and spatial extent of heatwaves. MHWs cause shifts in habitat-forming keystone species (e.g., seagrass, salt marshes, and mangroves), non-native species invasions, and regime shifts (Lima and Wethey 2012; Marbà and Duarte 2010; Osland et al. 2013; Smale and Wernberg 2013; Sorte et al. 2013; Wernberg et al. 2016; Aoki et al. 2021). Relative to coastal and oceanic ecosystems, little attention has been given to heatwaves within estuaries despite their vulnerability due to lower water volume and greater heat exchange by tidal, atmospheric, and fluvial forces.

Many organisms use habitat in the river-estuary-coastal ocean continuum and are sensitive to seasonal extremes in these different locations. This variation among habitats requires better understanding of when, where, and how aquatic heatwaves are changing with implications for species distributions, fisheries management, and conservation. Prior studies of MHWs have relied on satellite-derived sea surface temperature (SST) for broad-scale analysis however, satellite-derived SST measures have limitations. Remotely sensed SST measures are collected at daily to monthly time intervals, are limited by cloud cover, and have low spatial resolution (1-50 km) limiting their use to large bodies of water with low land contamination (i.e., land cover within pixel < 50%; Lima and Wethey 2012). Direct measurements of temperature using in-situ thermistors provide another approach for measuring heatwaves and have the benefit of high temporal resolution (minutes to hours) that allow study of site-specific changes in smaller waterways that would otherwise be contaminated by land using satellite observations, are measured in conjunction with other water quality variables, and can be extrapolated to broad scales through monitoring networks. This study 1) provides a baseline assessment of estuarine heatwaves (EHWs) using in-situ water temperature measurements from the last 24 years (1996-

2019) made at U.S. National Estuarine Research Reserve Sites, 2) examines temporal relationships between atmospheric heatwaves (AHWs) and EHWs, and 3) examines the degree to which EHWs and extremes of low DO and pH co-occur.

Methods

Study Sites

The U.S. National Oceanic and Atmospheric Administration's (NOAA) National Estuarine Research Reserve System (NERRS) maintains a system-wide monitoring program that measures meteorological and water quality conditions using high-frequency automated sensors throughout North America. Of the 29 reserves managed by NERRS, we identified 12 reserves (Figure 1) with the longest and most complete meteorological station and water quality station records, beginning in 2002 and 1996, respectively (NOAA 2020). The selected reserves occur on both coasts of the contiguous U.S., including the Gulf of Mexico and Puerto Rico and represent five distinct estuarine habitat types (open water, submerged aquatic vegetation (SAV), upland, marsh, and mangrove) as described in earlier NERRS publications (Wenner et al. 2001; Sanger et al. 2002). Each reserve maintains a minimum of four water quality monitoring stations; within the 12 reserves, we identified 17 water quality monitoring stations with water temperature, dissolved oxygen, and pH records from 1996-2019 that had relatively few missing data (max data gap < 4.2% of timeseries; Table 1). Each reserve follows standard operating procedures and QAQC protocols that allow for comparison across the reserve system. Automated atmospheric and water quality stations collect samples at 30-minute (prior to 2007) and 15-minute intervals (post 2007). YSI water quality sondes (model 6600 until 2013, EXO2 thereafter) are calibrated every 1-4 weeks depending on site and season to reduce sensor drift and cleaned to reduce data loss due to biofouling. All data were accessed using the NERRS Centralized Data Management Office's Advanced Query System (http://cdmo.baruch.sc.edu/).

Extreme Event Detection

Atmospheric and water quality data flagged during the QC protocol as outside sensor range, missing, rejected, or suspect were removed from the analysis. Tidal and diurnal signals in the high-frequency water quality data were removed using Fourier transform (aka transform filter) as a low-pass filter (Walters and Heston 1982; Thomson and Emery 2014). This was done to account for advective fluxes from tidal processes that may impact the water quality time series. Gibbs' phenomenon was suppressed using a three-point taper between the pass-band and stop-band (Forbes 1988; Thomson and Emery 2014). Inspection for Gibbs' phenomena contamination in the filtered data did not reveal any obvious introduced spurious oscillations. Daily averages for each parameter were then calculated from the low-pass filtered, high-frequency data if ≥ 75% of a day's high-frequency observations were present using the R package 'openair' version 2.8-3 (Carslaw and Ropkins 2012). The resulting 216 station years of daily average air temperature and 408 station years of daily average water temperature, DO concentration, and pH were then used to identify extreme events using the R package 'heatwaveR' version 0.4.4 (Schlegel and Smit 2018; R Core Team 2020). While the datasets used in our analysis do not provide the recommended 30 consecutive years of data suggested for this sort of analysis (Hobday et al. 2016), the data are the best available source of high-frequency measures of several estuarine water quality variables for over two-decades across broad geographic regions. In addition, Schlegel et al. (2019) presented a sensitivity analysis of the heatwave algorithm which demonstrated that time series as short as 10-years provided event duration and intensity measures similar to 30-year time series.

Estuarine heatwaves followed the Hobday et al. (2016) definition of marine heatwaves as those times when the daily mean water temperatures exceeded the seasonally adjusted 90th percentile threshold (e.g., the upper-tail of the water temperature distribution) for a period ≥ 5 days without a drop below the threshold for ≥ 2 days. Extreme low DO and pH events followed the same definition as estuarine heatwaves, with the exception that these extremes were those times when DO or pH were below the 10th percentile. We identified low pH (i.e., more acidic) rather than high pH conditions of interest given future projections of ocean acidification (Doney et al. 2020). Atmospheric heatwaves were defined similar to estuarine heatwaves, with the exception that atmospheric heat waves last for a period ≥ 3 consecutive days (i.e., no daily gaps below the seasonal threshold during an event; Perkins and Alexander 2013) relative to EHW which last ≥ 5 days and that allow for ≤ 2 gap days below the seasonal threshold during an event (Hobday et al. 2016). These definitions follow convention in the two fields and are related to the greater atmospheric temperature variability relative to water. An EHW was classified as co-occurring with an AHW when the EHW event started while there was an active AHW event occurring. Similarly, low DO and low pH events co-occurred with an EHW if they started while there was an active EHW. Our approach to identifying co-occurrence is conservative given the potential for lags among the variables, but in using this approach we avoided the problem of specifying lags, which likely vary. Since the AHW timeseries was shorter relative to the EHW's, the co-occurrence and lag analyses were assessed between 2002-2019. As EHW and extreme water quality conditions had the same timeseries length, their co-occurrence and lag analyses spanned the entire 1996-2019 timeseries. Severity classification (moderate, strong, severe, extreme) of each extreme event followed the Hobday et al. (2018) method which is based on event peak intensity and multiples of the 90th percentile difference from the seasonally adjusted mean.

Statistical Analysis

When extreme events were identified, we also quantified event duration, cumulative intensity above the 90th percentile (or below the 10th percentile for DO and pH), rate of event onset, and event decline rate. Event frequency, defined here as the annual sum of events per station, was derived to examine long-term changes in extreme events. Events were further categorized based on reserve region, season (winter = Dec-Feb, spring = March-May, summer = June-Aug, fall = Sept-Nov), and mean salinity during the event (freshwater = 0-0.5 g kg^{-1}, oligohaline = 0.5-5 g kg^{-1}, mesohaline = 5-18 g kg^{-1}, polyhaline = 18-30 g kg^{-1}, seawater = 30+ g kg^{-1}) to identify spatial, seasonal, and salinity-based patterns. While these seasonal definitions may have limitations given differences in season length between regions, they represent Northern hemisphere meteorological seasons (NOAA 2016) and have been used in estuarine and atmospheric studies of a similar spatial scale (Caffrey 2004; Lau and Nath 2012). Annual trends among events with regard to category and variable of interest were calculated using the non-parametric Mann-Kendall trend test with Sen's slope estimator, which are robust against outliers and non-normally distributed data, and have been used in other long-term trend analyses (Hirsch et al. 1982; Webb and Nobilis 1995; Kaushal et al. 2010; Perkins and Alexander 2013). For reserves with two stations, we present averages and associated range. Due to the multiple comparisons, a 10% False Discovery Rate (FDR, also referred to as Benjamini-Hochberg correction) was applied to reduce type 1 errors upon statistical inference (Benjamini and Hochberg 1995). All analyses were performed in the R environment for statistical computing (R

Core Team 2020) with code available on GitHub (https://github.com/spencer-tassone/EstuarineHeatwaves).

Results

Co-Occurrence and Lag-Time of Events

Relatively low intensity, low duration AHW events are able to produce EHWs when water temperature is near its 90th percentile threshold (Figure 2a,b). Estuaries can retain this excess heat and, in some cases, trigger the co-occurrence of a low DO event (Figure 2c) and low pH event (Figure 2d). The proportion of triple co-occurrence of stressful water quality conditions where there was a low DO event and low pH event that began during an EHW was rare, occurring on average (± SD) in only 2 ± 2% of EHWs but ranging up to 7% at some stations. The proportion of EHW that co-occurred during an active AHW event averaged 36% across all stations and ranged between 6-71% (Figure 3). The Mid-Atlantic and Northeast regions had the greatest proportional co-occurrences of atmospheric and estuarine heatwaves, averaging 51 ± 3% and 39 ± 21%, respectively. On average, 17 ± 11% of EHWs had co-occurring low DO events, and 6 ± 5% of EHWs had co-occurring low pH events (Figure 3). Amongst co-occurring events, the average lag-time between an AHW and the onset of an EHW ranged between 0-2 days (Figure 4). Similarly, the average lag-time between co-occurring EHW and the onset of both low DO and low pH events was 3 ± 2 and 4 ± 2 days respectively.

Atmospheric Heatwaves (AHWs)

Between 2002-2019, across NERRS reserves the number of AHW events did not increase over time. During the 18 years observed, the 12 NERRS reserves had a total of 899 AHW events, with a mean reserve total of 75 ± 11 AHWs (Figure 5). Similarly, during the 2002-2019 period, there was a larger number of AHW than EHW or low DO and low pH events. The 18-year mean air temperature trend among the reserves ranged from -0.10 to 0.12 $^{\circ}$C year^{-1} with 8 of the 12 reserves having positive trends. The trend in annual standard deviation was lower than that of the mean air temperature, ranging from -0.03 to 0.05 $^{\circ}$C year^{-1} with 6 of the 12 reserves having positive increases (SI Figure 1a).

Estuarine Heatwaves (EHWs)

Among the 17 water quality stations, there was a total of 900 EHW events between 1996-2019 (Table 1) with a mean station total number of EHWs of 53 ± 8 (Table 1). The mean annual frequency and duration of EHWs ranged between 1-4 events year^{-1} and 6-12 days respectively. The average onset rate of EHWs declined by -0.01 $^{\circ}$C day^{-1} year^{-1} (p-value = 0.001; Table 2) among the NERRS stations. There has been no significant trend in severity classification of EHW events (Figure 6a), although there has been a non-significant increasing trend in the total number of EHW days (Figure 6d). The 24-year mean water temperature trend among the reserves ranged between -0.07-0.09 °C year^{-1} which was greater than the trend in the annual standard deviation (-0.02-0.04 $^{\circ}$C year^{-1}; SI Figure 1b). Patterns in mean EHW duration, total number of events, and event frequency as a function of mean depth, mean tidal range, and habitat type were non-significant (SI Figure 2).

Low Dissolved Oxygen Events

The frequency (-0.05 events year^{-1}, p-value = 0.008), average cumulative intensity (-0.20 mg L^{-1} days year^{-1}, p-value = 0.031), average duration (-0.12 days year^{-1}, p-value = 0.003), onset and decline rates (-0.01 mg L^{-1} day^{-1} year^{-1}, p-value = 0.003, -0.02 mg L^{-1} day^{-1} year^{-1}, p-value = 0.003 respectively) of low DO events decreased over time throughout the NERRS stations (Table 2). There has been a decrease in the total number of low DO events driven by reductions in the strong, severe, and extreme severity classifications (Figure 6b). Similarly, there has been a significant reduction in total low DO days, decreasing by 11.9 days $year^{-1}$ (Figure 6e). Regionally, the Northeast had significant declines in average cumulative intensity (-0.37 mg L^{-1} days $year^{-1}$, p-value = 0.001) and event duration (-0.20 days $year^{-1}$, p-value = 0.004). Likewise, the Northeast and Mid-Atlantic had decreased average onset (both -0.01 mg L^{-1} day^{-1} $year^{-1}$, p-value ≤ 0.012) and decline rates of low DO events (-0.02 mg L^{-1} day^{-1} $year^{-1}$, p-value = 0.011, -0.04 mg L^{-1} day^{-1} $year^{-1}$, p-value < 0.001 respectively). The frequency (-0.03 events $year^{-1}$, p-value < 0.001), average duration (-0.14 days $year^{-1}$, p-value = 0.005), and average decline rate (-0.02 pH $days^{-1}$ $year^{-1}$, p-value = 0.012) declined significantly throughout the summer. Patterns in mean low DO duration, total number of events, and event frequency as a function of mean depth, mean tidal range, and habitat type were mostly non-significant (SI Figure 3).

Low pH Events

The frequency (-0.09 events $year^{-1}$, p-value ≤ 0.001), average duration (-0.13 days $year^{-1}$, p-value = 0.04), and average cumulative intensity (-0.07 pH days $year^{-1}$, p-value = 0.006) of low pH events significantly decreased over time (Table 2). Likewise, there has been a decrease in the total number of low pH events driven by significant reductions in all categories of extreme event classification (Figure 6c). Similarly, there has been a significant reduction in total low pH days, decreasing by 24.9 days $year^{-1}$ (Figure 6f). All seasons had declines in the frequency of low pH events, with spring and summer reduced by -0.03 events $year^{-1}$ (p-values ≤ 0.008) and fall and winter reduced by -0.02 events $year^{-1}$ (p-values ≤ 0.012). Furthermore, the average cumulative intensity (-0.12 pH days $year^{-1}$, p-value = 0.001) and duration (-0.29 days $year^{-1}$, p-value = 0.004) of low pH events decreased in the Spring. The onset rate of low pH events was reduced in the fall (-0.004 pH day^{-1} $year^{-1}$, p-value = 0.016). Among salinity gradients, the frequency of low pH events declined at the salinity end-members (tidal fresh = -0.08 events $year^{-1}$, p-value =

0.009; Seawater = -0.10 events $year^{-1}$, p-value = 0.003). Mesohaline, polyhaline, and seawater segments had significant declines in the average cumulative intensity of low pH events (slopes between -0.11 and -0.08 pH days $year^{-1}$, p-values ≤ 0.011). Furthermore, the average duration (-0.27 days $year^{-1}$, p-value = 0.027) and decline rate (-0.004 pH day^{-1} $year^{-1}$, p-value = 0.007) of low pH events were reduced in the seawater end-member. Regionally, the West coast (-0.20 events $year^{-1}$), Northeast, Mid-Atlantic (both slopes = -0.10 events $year^{-1}$), and Southeast (-0.09 events $year^{-1}$) had significant reductions in low pH event frequency (all p-values ≤ 0.03). Similarly, the West coast, Northeast, and Southeast regions had reduced average cumulative intensity (slopes between -0.10 and -0.06 pH days $year^{-1}$, p-values ≤ 0.027) and average decline rates (slopes between -0.01 and -0.002 pH day^{-1} $year^{-1}$, p-values ≤ 0.014). The Northeast region was the only region to have reduced average low pH event onset rate (-0.002 pH day^{-1} $year^{-1}$, p-value = 0.027). Patterns in mean low pH duration, total number of events, and event frequency as a function of mean depth, mean tidal range, and habitat type were non-significant (SI Figure 4).

Discussion

The co-occurrence of and lag-time between extreme events (i.e., EHW, low DO, and low pH) potentially compounds the stressful conditions that a single event may cause. Atmospheric and estuarine heatwaves had the greatest proportion of co-occurrence (6-71%; Figure 3) and shortest lag-time (average = 1 ± 1 day; Figure 4), suggesting that AHW can push estuaries into heatwave conditions rapidly. Given the relatively high proportion of co-occurrence of AHW and EHW and their short lag times, estuarine heatwaves may interact nonlinearly with ecosystem processes as has been observed in other aquatic ecosystems (Stenseth and Mysterud 2002; Caissie 2006; Wilhelm and Adrian 2008). Estuarine heatwave and low DO co-occurrence were twice as likely throughout the Atlantic coast (average = 20 ± 9%) relative to the Pacific, Gulf of

Mexico, and Caribbean (average = 8 ± 2%; Figure 3). Estuarine heatwave and low pH co-occurrence were relatively low and similar across all regions (6 ± 4%). Warm water temperatures stimulate metabolic activity that can cause diel fluctuations in pH (Duarte et al. 2013) however watershed inputs of bicarbonate alkalinity have been increasing in many rivers (Kaushal et al. 2013), increasing the buffering capacity of receiving water bodies against low pH events including those brought on by relatively short-lived extreme EHW events. Likewise, increased production during periods of elevated water temperature could reduce low pH stress in estuaries that are experiencing increased loading of acidic waters, or reduced watershed loading via drought or water withdrawals. The average lag-time between an EHW and low DO or low pH was 3 ± 2 days and 4 ± 2 days respectively, suggesting a possible connection between extreme warm water conditions and ecosystem respiration. The triple co-occurrence of an estuarine heatwave, low DO, and low pH event was low amongst most stations, averaging 2 ± 2%, however Tivoli South Bay, Hudson River, NY had a 7% rate of triple co-occurrence. The relatively high co-occurrence of EHW and low pH events, as well as triple co-occurrences at Tivoli South Bay are likely due this station being connected to a tidal freshwater swamp. The co-occurrence of low DO and low pH appears to have both additive and synergistic negative effects on fish and bivalves (Gobler and Baumann 2016; Chan et al. 2019). These low DO and low pH conditions have become more common in stratified coastal zones due in part to rising water temperature (Gruber 2011; Doney et al. 2012). The NERRS datasets do not allow us to determine if stratification is present in these shallow estuaries (mean water column depth range between 0.68-5.51 meters) as measurements are collected at a single fixed position in the water column. The water quality data produced by NERRS come from ‘bottom waters’ as the water quality sondes are fixed in elevation above the sediment surface between 0.15-1.0 meters depending on station (1.7 m for APAES which is the lone ‘surface water’ station). If stratification were present, the AHW-EHW co-occurrence results presented herein are likely conservative estimates given the position of the thermistor sensor in the water column. Stratification may also cause the EHW and low DO/pH events to be conservative estimates again due to sensor position as surface and bottom waters are not exchanged (i.e., surface water experiencing an EHW while bottom water is not). Nonetheless, our results represent observed estuarine water quality conditions where the triple co-occurrence of extreme events was not frequent at most stations. However, these relatively rare periods represent times of enhanced stress where potentially large and enduring changes may occur.

Heating and cooling trends of annual estuarine water temperature were observed throughout the 408 station years available. Of the 17 stations examined in this analysis, 10 showed positive trends in annual water temperature with four being significant (p-value < 0.05) ranging from 0.03-0.09 $^{\circ}$C year^{-1}. These warming rates are similar to what others have observed in streams, rivers, and lakes throughout the U.S. (Kaushal et al. 2010; Ding and Elmore 2015; O’Reilly et al. 2015) and greater than warming rates observed in coastal and open ocean studies (Burrows et al. 2011; Lima and Wethey 2012). This suggests that inland aquatic ecosystems, including estuaries, are among the most vulnerable to warming and extreme temperature events which can cause shifts in habitat-forming keystone species (e.g., seagrass, salt marshes, and mangroves), possibly alter greenhouse gas production, local extinction events, non-native species invasions, and regime shifts (Borges et al. 2016; Lima and Wethey 2012; Marbà and Duarte 2010; Marotta et al. 2014; Osland et al. 2013; Smale and Werberg 2013; Bartosiewicz et al. 2016; Wernberg et al. 2016; Kendrick et al. 2019). The standard deviation in annual estuarine water temperature increased at 13 of the 17 stations (SI Figure 1b), suggesting that water temperature variability is increasing at the upper-tail of the water temperature distribution (i.e., distribution width increasing). Reserves in the lower Northeast, Mid-Atlantic, and Northern West coast had trends in annual water temperature standard deviation increasing quicker than the mean. Similar results have been observed for atmospheric heatwaves in Europe (Schär et al. 2004). Throughout the U.S., estuaries will continue to experience extremely high water temperatures, driven in part by projected increases in atmospheric and marine heatwave frequency, intensity, duration, and spatial extent (Lau and Nath 2012, Collins et al. 2019). Additional long-term data from additional sites are needed to refine quantitative relationships between habitats, heatwaves, and water quality variables. The NERRS dataset provides a valuable long-term record and a baseline for future analyses of extreme temperature impacts on these ecosystems.

All significant trends in extreme low DO events were negative, suggesting an improvement in estuarine water quality. Low DO events in estuaries are, in general, decreasing in frequency, mean cumulative intensity, mean duration, and rates of onset and decline (Table 2). There has been a significant reduction in the total number of annual low DO events (Figure 6b) and the number of low DO days (Figure 6e) with significant reductions in strong, severe, and extreme classifications of low DO events. Significant management improvements in total nitrogen runoff and wastewater treatment have been attributed to improvements in water quality including rising DO concentrations (Tomaso and Najjar 2015; Whitney and Vlahos 2021); however, climate forecasts indicate these improvements in DO will be reduced given lower oxygen solubility and increased respiration under warming estuarine conditions (Whitney and Vlahos 2021). Of the 846 low DO events that were observed in our analysis, 54.7% reached hypoxia (≤ 2.0 mg L^{-1}) with oxygen dropping to essentially zero (< 0.01 mg L^{-1}) in some cases.

Similarly, the mean low DO event duration was 9 days but ranged up to 52 days. While extreme low DO events in estuaries are becoming less frequent, intense and shorter in duration, they co-occur with EHWs in some cases. Improvements in water quality conditions might be reversed in the future if water temperature continues to increase leading to more co-occurring extremes. Likewise, if projections of increase precipitation led to increased freshwater run-off and salinity stratification, then low DO events could become more frequent in bottom water sites increasing the probability of co-occurring water quality extremes.

Similar to the trends in low DO events, all significant trends related to low pH events were negative thereby providing further evidence of improvements in estuarine water quality. In general, low pH events are decreasing in frequency, duration, and intensity over time within U.S. estuaries with frequencies being reduced during all seasons (Table 2). Furthermore, there have been significant reductions in all severity classifications of low pH events (Figure 6c) and a significant reduction in the total number of low pH days (Figure 6f). However, there is regional heterogeneity in low pH frequency with all regions except the Gulf of Mexico showing significant reductions. Similarly, the West coast, Northeast, and Southeast were the only regions to have significant annual declines in mean cumulative intensity and mean decline rate while the Northeast was the only region to have reduced mean low pH event onset rates. Of the 815 low pH events observed in our analysis 46.4% had a pH < 7.0, with some reaching a pH of 6.0, and lasting up to 97 days (mean duration = 12 days). While atmospheric CO_2 emissions continue to push pH in the marine environment lower (Doney et al. 2009), estuarine pH is controlled by many complex interactions, including watershed export of alkalinity (via carbonate lithology, acid deposition, topography, land use/cover, and wastewater effluent), organic matter, CO_2, nutrient inputs, stratification, and metabolic rates of production and respiration (Cai et al. 2011;

Kaushal et al. 2013; Duarte et al. 2013). We show that EHWs co-occur with extreme low pH conditions relatively rarely (≤ 13% of EHWs) for most NERRS stations, but not for all (max co-occurrence = 18% for Tivoli South Bay, Hudson River, NY). Our results provide evidence of improvements in estuarine water quality via reductions in the frequency, duration, and severity of low pH (i.e., acidic) conditions on a broad scale.

Atmospheric heatwaves at the NERRS reserves had no significant trends in heatwave characteristics (Table 2). However, there were 899 AHWs observed between 2002-2019 at the 12 NERRS reserves, with West coast and Gulf of Mexico reserves having the greatest number of AHWs relative to the number of reserves in those regions (88 and 82 events reserve^{-1}). AHW events were relatively frequent, with a mean annual occurrence of 4 ± 3 events, ranging up to 18 events year^{-1}, and lasting up to 15 days (mean duration = 4 days). Seasonally, AHW were on average most frequent in the summer and least frequent in the winter however, winter cumulative intensity was greater on average than summer cumulative intensity (2.14 and 1.28 °C days respectively). Likewise, simple linear regression of mean annual air temperature did show significant increases in air temperature, with positive slopes at 8 of the 12 reserves (SI Figure 1a). These findings follow prior studies of North American atmospheric temperature trends that have shown a lengthening of the growing season driven in part by warming fall temperatures since the mid-twentieth century (Piao et al. 2008; Burrows et al. 2011; Barichivich et al. 2012). A warming atmosphere will increase diffusive thermal fluxes across the atmosphere-surface water boundary leading to increased temperature stress within estuaries.

Estuarine heatwaves had no significant trends in EHW characteristics across a broad scale with the exception of reduced average onset rate (Table 2). This lack of significant trends is in contrast to studies of the coastal and open oceans that have reported increasing frequencies and durations of marine heatwaves (Lima and Wethey 2012; Frölicher et al. 2018; Oliver et al. 2018). Nonetheless, EHWs occurred relatively frequently, with a mean annual occurrence of 2 ± 2 events, ranging up to 10 events $year^{-1}$, and lasting up to 44 days (average duration = 8 days). Trends in estuarine heatwaves along salinity gradients were non-significant however, there were between 81-268 total events among the salinity segments over the study period with an average duration spanning 7-9 days. Several complex interactions between estuarine attributes (i.e., freshwater discharge, oceanic exchange, depth, residence time, geomorphology, latitude, salinity), catchment characteristics (i.e., riparian cover, impervious surface area, wastewater effluent), and regional climate (i.e., cloud cover, precipitation regime) influence the thermal properties of estuaries, which lead to estuarine segments experiencing differential thermal conditions (Scanes et al. 2020). The onset rate of estuarine heatwaves declined over the study period, suggesting that estuarine heatwaves are reaching peak heatwave intensity less rapidly. This reduction in event onset rate may benefit organisms as it suggests the change from non-heatwave conditions to peak heatwave conditions is becoming more gradual (i.e., less abrupt change). Estuarine heatwaves will likely accelerate due to increased diffusive air-water boundary heat flux and horizontal advective thermal fluxes from a warming coastal ocean and fluvial discharge. This will have cascading impacts on estuarine ecosystem function and services, including ecologically and commercially important organisms.

This study provides the first baseline assessment of estuarine heatwaves and their co-occurrence with low DO and pH for the NERRS sites with the longest records. Future studies should consider the three-dimensional structure of aquatic heatwave events as studies have been limited to surface waters (i.e., satellite data) or fixed water depths (i.e., current study). Understanding the depth profile and the spatial areal extent of these disturbance events would help resolve their impacts on species' movement and range dynamics, as extreme events can cause barriers to migration if they extend through the whole water column (Major and Mighell 1967, Cairns 1971) and eliminate or severely limit thermal refuges. Similarly, water circulation (i.e., residence time) and stratification are known to influence water quality conditions. Further analysis of EHWs should consider estuary specific residence time and salinity stratification as potential drivers of EHW characteristics. Likewise, this study considered a single tidal filtering approach however, further analyses should consider other tidal filters (i.e., weighted regression) that could be applied to high frequency estuarine measures (Beck et al. 2015). Additionally, benthic sediments are considered thermally stable relative to the overlying water column; however, benthic respiration increases with temperature, and it remains unclear if EHWs increase sediment temperature possibly affecting carbon storage. Developing a thorough understanding of depth profiles during heatwaves would benefit resource managers working to restore aquatic ecosystems, conserve fisheries, limit non-native species introductions, and model carbon sources and sinks from shallow aquatic ecosystems (Caissie 2006; Aoki et al. 2021). Furthermore, water temperatures are increasing in streams and rivers driven in part by a warming climate, discharge regulation, and increased impervious surface area (Webb and Nobilis 1995; Kaushal et al. 2010; Ding and Elmore 2015), yet there is a lack of understanding regarding the frequency, intensity, and duration of aquatic heatwaves in these inland lotic ecosystems that affect thermal inputs to estuaries. Lastly, comparisons among in-situ studies and those utilizing satellite-derived SST may provide an opportunity for a thorough characterization of aquatic heatwaves events in space and time as heatwaves are often short-lived, intense, and occurring at varying spatial scales. Combining in-situ and satellite temperature offers the opportunity to document more events with coordinated study of impacts as the heating of estuaries increases in the future.

References

Aoki, L. R., K. J. McGlathery, P. L. Wiberg, M. P. J. Oreska, A. C. Berger, P. Berg, R. J. Orth. 2021. Seagrass recovery following marine heat wave influences sediment carbon stocks. *Frontiers in Marine Science* 7: 1170.

Baird, D., R. R. Christian, C. H. Peterson, and G. A. Johnson. 2004. Consequences of hypoxia on estuarine ecosystem function: energy diversion from consumers to microbes. *Ecological Applications* 14(3): 805–822.

Barichivich, J., K. R. Briffa, T. J. Osborn, T. M. Melvin, and J. Caesar. 2012. Thermal growing season and timing of biospheric carbon uptake across the Northern Hemisphere. *Global Biogeochemical Cycles* 26(4): GB4015

Bartosiewicz, M., I. Laurion, F. Clayer, and R. Maranger. 2016. Heat-Wave Effects on Oxygen, Nutrients, and Phytoplankton Can Alter Global Warming Potential of Gases Emitted from a Small Shallow Lake. *Environmental Science & Technology* 50(12): 6267–6275.

Beck, M. W., J. D. Hagy III, M. C. Murrell. 2015. Improving estimates of ecosystem metabolism by reducing effects of tidal advection on dissolved oxygen time series. *Limnology and Oceanography: Methods* 13(12): 731-745.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1): 289–300.

Borges, A. V., W. Champenois, N. Gypens, B. Delille, and J. Harlay. 2016. Massive marine methane emissions from near-shore shallow coastal areas. *Scientific Reports* 6(1): 27908.

Burrows, M. T., D. S. Schoeman, L. B. Buckley, P. Moore, E. S. Poloczanska, K. M. Brander, C. Brown, J. F. Bruno, C. M. Duarte, B. S. Halpern, J. Holding, C. V. Kappel, W. Kiessling, M. I. O'Connor, J. M. Pandolfi, C. Parmesan, F. B. Schwing, W. J. Sydeman, and A. J. Richardson. 2011. The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science* 334(6056): 652–655.

Caffrey, J. M. 2004. Factors Controlling Net Ecosystem Metabolism in U. S. Estuaries. *Estuaries* 27(1): 90–101.

Cai, W. J., X. Hu, W. J. Huang, M. C. Murrell, J. C. Lehrter, S. E. Lohrenz, W. C. Chou, W. Zhai, J. T. Hollibaugh, Y. Wang, P. Zhao, X. Guo, K. Gundersen, M. Dai, G. C. Gong. 2011. Acidification of subsurface coastal waters enhanced by eutrophication. *Nature Geoscience* 4(11): 766-770.

Cairns, J. 1971. Thermal Pollution: A Cause for Concern. *Journal (Water Pollution Control Federation)* 43(1): 55-66.

Caissie, D. 2006. The thermal regime of rivers: a review. *Freshwater Biology* 51(8): 1389–1406.

Carslaw, D. C., and K. Ropkins. 2012. openair --- an R package for air quality data analysis. *Environmental Modelling & Software* 27-28: 52-61.

Chan, F., J. A. Barth, K. J. Kroeker, J. Lubchenco, and B. A. Menge. 2019. The dynamics and impact of ocean acidification and hypoxia. *Oceanography* 32(3): 62-71.

Collins, M., M. Sutherland, L. Bouwer, S.-M. Cheong, H. J. D. Combes, M. K. Roxy, I. Losada, K. McInnes, B. Ratter, E. Rivera-Arriaga, R. D. Susanto, D. Swingedouw, L. Tibig, P. Bakker, C. M. Eakin, K. Emanuel, M. Grose, M. Hemer, L. Jackson, A. Kääb, J. Kajtar, T. Knutson, C. Laufkötter, I. Noy, M. Payne, R. Ranasinghe, G. Sgubin, M.-L. Timmermans, A. Abdulla, M. H. González, and C. Turley. 2019. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*: Extremes, Abrupt Changes and Managing Risks 6: 589-655.

Ding, H., A. J. Elmore. 2015. Spatio-temporal patterns in water surface temperature from Landsat time series data in the Chesapeake Bay, U.S.A. *Remote Sensing of Environment* 168: 335-345.

Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas. 2009. Ocean Acidification: The Other CO_2 Problem. *Annual Review of Marine Science* 1(1): 169–192.

Doney, S. C., M. Ruckelshaus, J. Emmett Duffy, J. P. Barry, F. Chan, C. A. English, H. M. Galindo, J. M. Grebmeier, A. B. Hollowed, N. Knowlton, J. Polovina, N. N. Rabalais, W. J. Sydeman, and L. D. Talley. 2012. Climate Change Impacts on Marine Ecosystems. *Annual Review of Marine Science* 4(1): 11–37.

Doney, S. C., D. S. Busch, S. R. Cooley, and K. J. Kroeker. 2020. The Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities. *Annual Review of Environment and Resources* 45(1): 83–112.

Duarte, C. M., I. E. Hendriks, T. S. Moore, Y. S. Olsen, A. Steckbauer, L. Ramajo, J. Carstensen, J. A. Trotter, and M. McCulloch. 2013. Is Ocean Acidification an Open-Ocean Syndrome? Understanding Anthropogenic Impacts on Seawater pH. *Estuaries and Coasts* 36(2): 221–236.

Easterling, D. R. 2000. Climate Extremes: Observations, Modeling, and Impacts. *Science* 289(5487): 2068–2074.

Forbes, A. M. G. 1988. Fourier transform filtering: A cautionary note. *Journal of Geophysical Research: Oceans* 93(C6): 6958-6962.

Frölicher, T. L., E. M. Fischer, and N. Gruber. 2018. Marine heatwaves under global warming. *Nature* 560(7718): 360–364.

Garrabou, J., R. Coma, N. Bensoussan, M. Bally, P. Chevaldonné, M. Cigliano, D. Diaz, J. G. Harmelin, M. C. Gambi, D. K. Kersting, J. B. Ledoux, C. Lejeusne, C. Linares, C. Marschal, T. Pérez, M. Ribes, J. C. Romano, E. Serrano, N. Teixido, O. Torrents, M. Zabala, F. Zuberer, and C. Cerrano. 2009. Mass mortality in Northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. *Global Change Biology* 15(5): 1090–1103.

Gobler, C. J., and H. Baumann. 2016. Hypoxia and acidification in ocean ecosystems: coupled dynamics and effects on marine life. *Biology Letters* 12(5): 20150976.

Gruber, N. 2011. Warming up, turning sour, losing breath: ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369(1943): 1980–1996.

Hirsch, R. M., J. R. Slack, and R. A. Smith. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research* 18(1): 107–121.

Hobday, A. J., L. V. Alexander, S. E. Perkins, D. A. Smale, S. C. Straub, E. C. J. Oliver, J. A. Benthuysen, M. T. Burrows, M. G. Donat, M. Feng, N. J. Holbrook, P. J. Moore, H. A. Scannell, A. Sen Gupta, and T. Wernberg. 2016. A hierarchical approach to defining marine heatwaves. *Progress in Oceanography* 141: 227–238.

Hobday, A., E. Oliver, A. Sen Gupta, J. Benthuysen, M. Burrows, M. Donat, N. Holbrook, P. Moore, M. Thomsen, T. Wernberg, and D. Smale. 2018. Categorizing and Naming Marine Heatwaves. *Oceanography* 31(2): 162-173.

Hoegh-Guldberg, O., R. Cai, E.S. Poloczanska, P.G. Brewer, S. Sundby, K. Hilmi, V.J. Fabry, and S. Jung., 2014. The Ocean: Climate Change 2014: Impacts, Adaptation, and Vulnerability. *Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 1655-1731.

Jenuch, A., J. Kreyling, and C. Beierkuhnlein. 2007. A New Generation of Climate-Change Experiments: Events, Not Trends. *Frontiers in Ecology and the Environment* 5(7): 365-374.

Kaushal, S. S., G. E. Likens, N. A. Jaworski, M. L. Pace, A. M. Sides, D. Seekell, K. T. Belt, D. H. Secor, and R. L. Wingate. 2010. Rising stream and river temperatures in the United States. *Frontiers in Ecology and the Environment* 8(9): 461–466.

Kaushal, S. S., G. E. Likens, R. M. Utz, M. L. Pace, M. Grese, and M. Yepsen. 2013. Increased river alkalinization in the Eastern U.S. *Environmental Science & Technology* 47(18): 10302-10311

Kendrick, G. A., R. J. Nowicki, Y. S. Olsen, S. Strydom, M. W. Fraser, E. A. Sinclair, J. Statton, R. K. Hovey, J. A. Thomson, D. A. Burkholder, K. M. McMahon, K. Kilminster, Y. Hetzel, J. W. Fourqurean, M. R. Heithaus, and R. J. Orth. 2019. A Systematic Review of How Multiple Stressors from an Extreme Event Drove Ecosystem-Wide Loss of Resilience in an Iconic Seagrass Community. *Frontiers in Marine Science* 6: 455.

Lau, N. C., and M. J. Nath. 2012. A Model Study of Heat Waves over North America: Meteorological Aspects and Projections for the Twenty-First Century. *Journal of Climate* 25(14): 4761–4784.

Lima, F. P., and D. S. Wethey. 2012. Three decades of high-resolution coastal sea surface temperatures reveal more than warming. *Nature Communications* 3(1): 704.

Llansó, R. J. 1992. Effects of hypoxia on estuarine benthos: the lower Rappahannock River (Chesapeake Bay), a case study. *Estuarine, Coastal and Shelf Science* 35(5): 491-515.

Madeira, D., L. Narciso, H. N. Cabral, and C. Vinagre. 2012. Thermal tolerance and potential impacts of climate change on coastal and estuarine organisms. *Journal of Sea Research* 70: 32–41.

Major, R. L., Mighell, J.L., 1967. Influence of rocky reach dam and the temperature of the okanogan river on the upstream migration of sockeye salmon. *Fisheries Bulletin*, 66: 131-147.

Marbà, N., and C. M. Duarte. 2009. Mediterranean warming triggers seagrass (Posidonia oceanica) shoot mortality. *Global Change Biology* 16(8): 2366–2375.

Marotta, H., L. Pinho, C. Gudasz, D. Bastviken, L. J. Tranvik, and A. Enrich-Prast. 2014. Greenhouse gas production in low-latitude lake sediments responds strongly to warming. *Nature Climate Change* 4(6): 467–470.

NOAA National Centers for Environmental Information. 2016. Meteorological versus astronomical seasons. https://www.ncei.noaa.gov/news/meteorological-versus-astronomical-seasons

NOAA National Estuarine Research Reserve System (NERRS). 2020. System-wide Monitoring Program. Data accessed from the NOAA NERRS Centralized Data Management Office website: http://www.nerrsdata.org

Oliver, E. C. J., M. G. Donat, M. T. Burrows, P. J. Moore, D. A. Smale, L. V. Alexander, J. A. Benthuysen, M. Feng, A. Sen Gupta, A. J. Hobday, N. J. Holbrook, S. E. Perkins-Kirkpatrick, H. A. Scannell, S. C. Straub, and T. Wernberg. 2018. Longer and more frequent marine heatwaves over the past century. *Nature Communications* 9(1): 1-12.

Oliver, E. C. J., J. A. Benthuysen, S. Darmaraki, M. G. Donat, A. J. Hobday, N. J. Holbrook, R. W. Schlegel, A. Sen Gupta. 2021. Marine heatwaves. *Annual Review of Marine Science* 13: 313-342.

O'Reilly, C. M., S. Sharma, D. K. Gray, S. E. Hampton, J. S. Read, R. J. Rowley, P. Schneider, J. D. Lenters, P. B. McIntyre, B. M. Kraemer, G. A. Weyhenmeyer, D. Straile, B. Dong, R. Adrian, M. G. Allan, O. Anneville, L. Arvola, J. Austin, J. L. Bailey, J. S. Baron, J. D. Brookes, E. Eyto, M. T. Dokulil, D. P. Hamilton, K. Havens, A. L. Hetherington, S. N. Higgins, S. Hook, L. R. Izmest'eva, K. D. Joehnk, K. Kangur, P. Kasprzak, M. Kumagai, E. Kuusisto, G. Leshkevich, D. M. Livingstone, S. MacIntyre, L. May, J. M. Melack, D. C. Mueller-Navarra, M. Naumenko, P. Noges, T. Noges, R. P. North, P. Plisnier, A. Rigosi, A. Rimmer, M. Rogora, L. G. Rudstam, J. A. Rusak, N. Salmaso, N. R. Samal, D. E. Schindler, S. G. Schladow, M. Schmid, S. R. Schmidt, E. Silow, M. E. Soylu, K. Teubner, P. Verburg, A. Voutilainen, A. Watkinson, C. E. Williamson, and G. Zhang.

2015. Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters* 42(24): 10773-10781

Osland, M. J., N. Enwright, R. H. Day, and T. W. Doyle. 2013. Winter climate change and coastal wetland foundation species: salt marshes vs. mangrove forests in the southeastern United States. *Global Change Biology* 19(5): 1482–1494.

Perkins, S. E., and L. V. Alexander. 2013. On the Measurement of Heat Waves. *Journal of Climate* 26(13): 4500–4517.

Piao, S., P. Ciais, P. Friedlingstein, P. Peylin, M. Reichstein, S. Luyssaert, H. Margolis, J. Fang, A. Barr, A. Chen, A. Grelle, D. Y. Hollinger, T. Laurila, A. Lindroth, A. D. Richardson, and T. Vesala. 2008. Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* 451(7174): 49–52.

Piatt, J. F., J. K. Parrish, H. M. Renner, S. K. Schoen, T. T. Jones, M. L. Arimitsu, K. J. Kuletz, B. Bodenstein, M. García-Reyes, R. S. Duerr, R. M. Corcoran, R. S. A. Kaler, G. J. McChesney, R. T. Golightly, H. A. Coletti, R. M. Suryan, H. K. Burgess, J. Lindsey, K. Lindquist, P. M. Warzybok, J. Jahncke, J. Roletto, and W. J. Sydeman. 2020. Extreme mortality and reproductive failure of common murres resulting from the northeast Pacific marine heatwave of 2014-2016. *PloS one* 15(1): e0226087.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Sanger, D. M., M. D. Arendt, Y. Chen, E. L. Wenner, A. F. Holland, D. Edwards, J. Caffrey. 2002. A Synthesis of Water Quality Data: National Estuarine Research Reserve System-wide Monitoring Program (1995-2000). *National Estuarine Research Reserve Technical Report Series 3*. South Carolina Department of Natural Resources, Marine Resources Division Contribution No. 500. 135 p.

Scanes, E., P. R. Scanes, and P. M. Ross. 2020. Climate change rapidly warms and acidifies Australian estuaries. *Nature Communications* 11(1): 1-11.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller. 2004. The role of increasing temperature variability in European summer heatwaves. *Nature* 427(6972): 332–336.

Schlegel, R. W., and A. J. Smit. 2018. heatwaveR: A central algorithm for the detection of heatwaves and cold-spells. *Journal of Open Source Software* 3(27): 821

Schlegel, R. W., E. C. J. Oliver, A. J. Hobday, and A. J. Smit. 2019. Detecting Marine Heatwaves with Sub-Optimal Data. *Frontiers in Marine Science* 6: 737.

Stenseth, N. C., and A. Mysterud. 2002. Climate, changing phenology, and other life history traits: Nonlinearity and match-mismatch to the environment. *Proceedings of the National Academy of Sciences* 99(21): 13379–13381.

Thomson, R. E., W. J. Emery. 2014. Digital Filters. In *Data analysis methods in physical oceanography*, Third ed. 593-637. Waltham, MA: Elsevier B.V.. doi: 10.1016/B978-0-12-387782-6.00006-5

Tomaso, D. J., and R. G. Najjar. 2015. Long-term variations in the dissolved oxygen budget of an urbanized tidal river: The upper Delaware Estuary. *Journal of Geophysical Research: Biogeosciences* 120(6): 1027–1045.

Walters, R. A., and C. Heston. 1982. Removing tidal-period variations from time-series data using low-pass digital filters. *Journal of Physical Oceanography* 12(1): 112-115.

Webb, B. W., and F. Nobilis. 1995. Long term water temperature trends in Austrian rivers. *Hydrological Sciences Journal* 40(1): 83–96.

Wenner, E. L., A. F. Holland, M. D. Arendt, Y. Chen, D. Edwards, C. Miller, M. Meece, J. Caffrey. 2001. A synthesis of water quality data from the National Estuarine Research Reserve System-wide monitoring program. *Final Report to The Cooperative Institute for Coastal and Estuarine Environmental Technology* NOAA Grant No.: NA97OR0209SC. South Carolina Department of Natural Resources, Marine Resources Division, Contribution No. 459.

Wernberg, T., S. Bennett, R. C. Babcock, T. de Bettignies, K. Cure, M. Depczynski, F. Dufois, J. Fromont, C. J. Fulton, R. K. Hovey, E. S. Harvey, T. H. Holmes, G. A. Kendrick, B. Radford, J. Santana-Garcon, B. J. Saunders, D. A. Smale, M. S. Thomsen, C. A. Tuckett, F. Tuya, M. A. Vanderklift, and S. Wilson. 2016. Climate-driven regime shift of a temperate marine ecosystem. *Science* 353(6295): 169–172.

Whitney, M. M., and P. Vlahos. 2021. Reducing Hypoxia in an Urban Estuary Despite Climate Warming. *Environmental Science & Technology* 55(2): 941–951.

Wilhelm, S., and R. Adrian. 2008. Impact of summer warming on the thermal characteristics of a polymictic lake and consequences for oxygen, nutrients and phytoplankton. *Freshwater Biology* 53(2): 226-237.

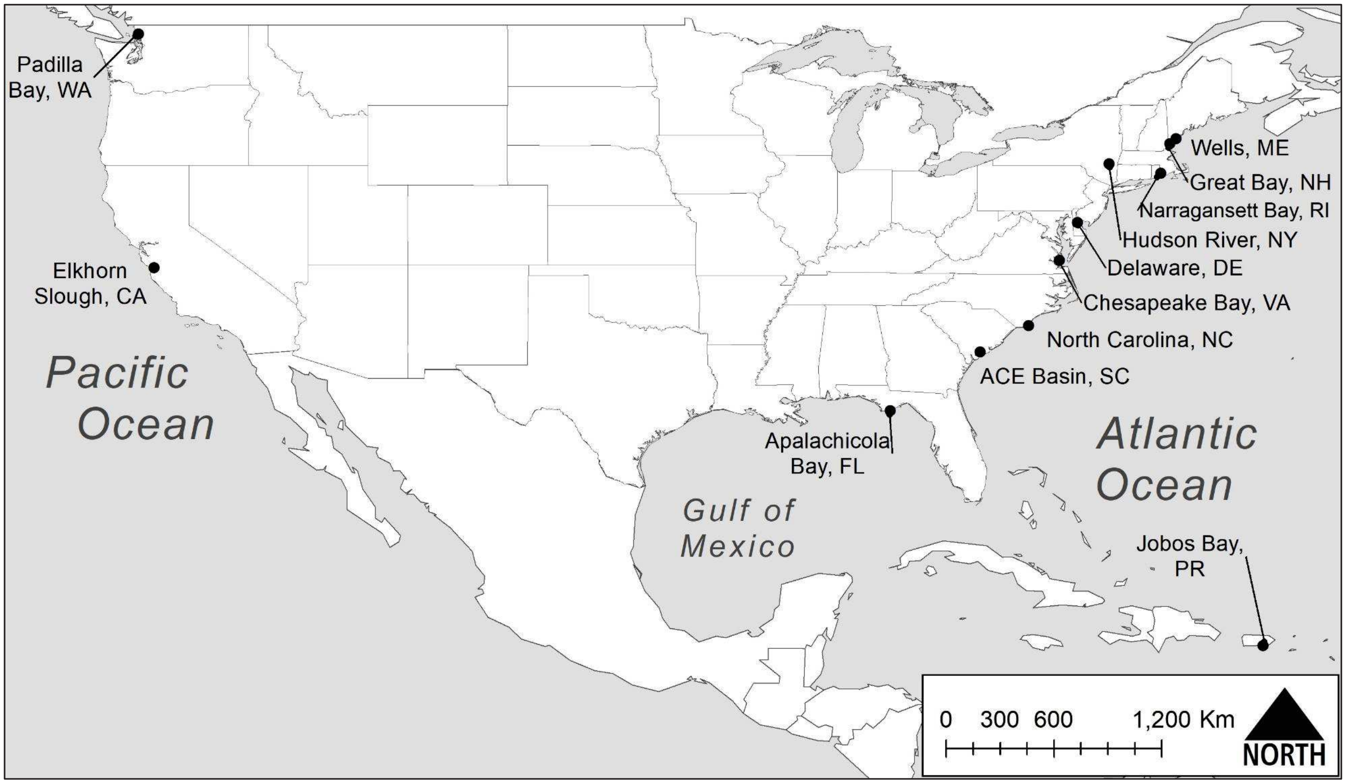
Figures

Table 1. NERRS reserve and station descriptions. Total number of estuarine heatwave (EHW) events, low dissolved oxygen (DO) and low pH events, mean salinity, mean depth and mean tidal range reflect the time period 1996-2019. Habitat type was characterized by Wenner et al. 2001 and Sanger et al. 2002.

Estuary	Region	Reserve Code	Station ID	Total EHW	Total Low DO	Total Low pH	Habitat Type	Mean Salinity (± SD, g kg^{-1})	Mean Depth (m)	Mean Tidal Range (m)
Wells, ME	Northeast	WEL	IN	48	55	46	SAV	30.8 ± 1.9	4.08	1.82
Great Bay, NH		GRB	GB	45	43	37	SAV	23.6 ± 5.6	5.51	1.40
Hudson River, NY		HUD	SK	38	40	31	Upland	0.2 ± 0.0	1.04	0.01
			TS	43	40	44	Marsh	0.1 ± 0.0	1.55	0.80
Narragansett Bay, RI		NAR	PC	44	30	41	Open Water	28.5 ± 1.9	2.80	0.38
Delaware, DE	Mid-Atlantic	DEL	BL	57	60	59	Marsh	2.0 ± 2.1	1.67	0.87
			SL	60	60	47	Marsh	10.8 ± 6.9	1.71	0.82
Chesapeake Bay, VA		CBV	TC	55	51	47	Marsh	10.6 ± 4.5	1.61	0.29
North Carolina, NC	Southeast	NOC	EC	46	49	52	Marsh	22.1 ± 6.7	1.13	0.76
			RC	55	43	44	Marsh	30.1 ± 4.3	1.55	0.88
ACE Basin, SC		ACE	SP	62	61	55	Marsh	29.0 ± 4.5	2.47	1.20
Jobos Bay, PR		JOB	09	56	61	52	Mangrove	36.5 ± 2.7	1.03	0.02
Apalachicola, FL	Gulf of Mexico	APA	EB	60	51	51	Open Water	11.0 ± 8.5	2.16	0.12
			ES	61	47	55	Open Water	9.7 ± 7.9	2.16	0.12
Padilla Bay, WA	West	PDB	BY	63	43	45	SAV	29.0 ± 1.1	3.36	0.84
Elkhorn Slough, CA		ELK	AP	46	54	52	Marsh	31.5 ± 4.6	0.68	0.06
			SM	61	58	57	Marsh	31.8 ± 2.7	1.79	0.71

Table 2. Results of Mann-Kendall and Sen's slope (MK-SS) analysis related to atmospheric and estuarine heatwaves (AHW and EHW respectively), low DO and low pH events. Total tests run indicates the number of MK-SS analyses for each test (AHW = 5 variables*5 regions + 5 variables*4 seasons + 5 variables*(1)time = 50 total tests run; additionally EHW, low DO and low pH included 5 variables*5 salinity classes = 75 total tests run). Only statistically significant tests are presented.

Test	Total Tests Run	Test Type	Variable	Category	Slope	MK p-value	10% FDR
AHW	50	-	-	-	-	-	-
EHW	75	Time	Avg. Onset Rate		-0.01	0.001	0.020
Low DO	75	Time	Avg. Onset Rate		-0.01	0.003	0.020
			Avg. Decline Rate		-0.02	0.003	0.060
			Avg. Duration		-0.12	0.003	0.040
			Avg. Cumulative Intensity		-0.20	0.031	0.100
			Frequency		-0.05	0.008	0.080
		Season	Frequency	Summer	-0.03	< 0.001	0.005
			Avg. Duration		-0.14	0.005	0.010
			Avg. Decline Rate		-0.02	0.012	0.015
			Avg. Cumulative Intensity	Fall	-0.20	0.016	0.020
		Region	Avg. Cumulative Intensity	NE	-0.37	0.001	0.008
			Avg. Duration		-0.20	0.004	0.012
			Avg. Onset Rate	NE	-0.01	0.012	0.024
				Mid-Atlantic	-0.01	0.005	0.016
			Avg. Decline Rate	NE	-0.02	0.011	0.020
				Mid-Atlantic	-0.04	< 0.001	0.004
Low pH	75	Time	Frequency		-0.10	< 0.001	0.020
			Avg. Duration		-0.13	0.040	0.060
			Avg. Cumulative Intensity		-0.07	0.006	0.040
		Season	Frequency	Winter	-0.02	0.002	0.015
				Spring	-0.03	0.008	0.025
				Summer	-0.03	< 0.001	0.005
				Fall	-0.02	0.012	0.030
			Avg. Onset Rate		-0.004	0.016	0.035
			Avg. Duration	Spring	-0.29	0.004	0.020
			Avg. Cumulative Intensity		-0.12	0.001	0.010
		Salinity	Frequency	Tidal Fresh	-0.08	0.009	0.020
				Seawater	-0.10	0.003	0.004
			Avg. Decline Rate		-0.004	0.007	0.016
			Avg. Duration		-0.27	0.027	0.028
			Avg. Cumulative Intensity		-0.10	0.003	0.008
				Polyhaline	-0.11	0.011	0.024
				Mesohaline	-0.08	0.006	0.012
		Region	Frequency	West	-0.20	0.002	0.004
				NE	-0.10	0.004	0.008
				Mid-Atlantic	-0.10	0.026	0.032
				SE	-0.09	0.030	0.044
			Avg. Cumulative Intensity	West	-0.06	0.024	0.028
				NE	-0.10	0.014	0.024
				SE	-0.07	0.027	0.036
			Avg. Decline Rate	West	-0.004	0.007	0.016
				NE	-0.002	0.014	0.020
				SE	-0.01	0.007	0.012
			Avg. Onset Rate	NE	-0.002	0.027	0.040



663 Figure 1. NERRS Reserve locations used for this analysis included all continuous water quality stations with records from 1996-2019.

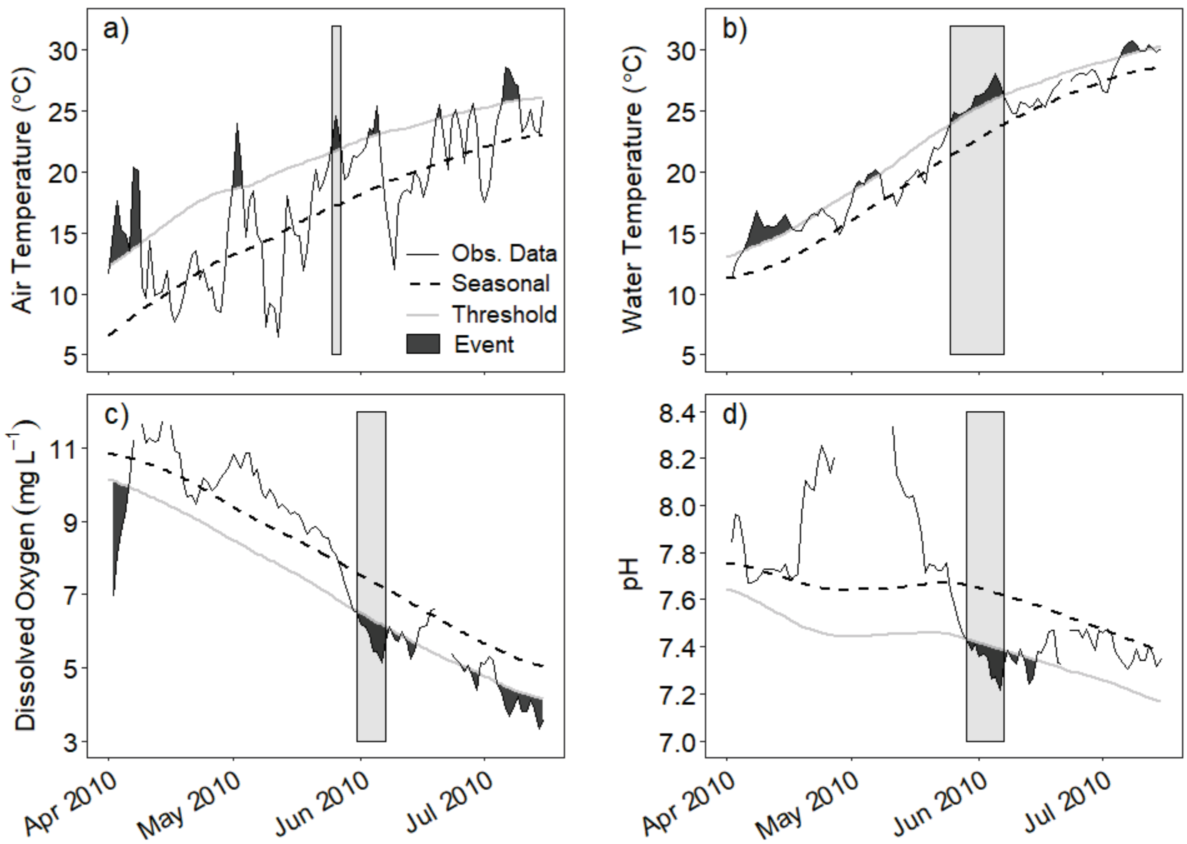


Figure 2. Example of atmospheric heatwaves (a), estuarine heatwaves (b), low dissolved oxygen (c), and low pH (d) events. The shaded bars in each figure are an example of a time when an estuarine heatwave co-occurred with an atmospheric heatwave, which co-occurred with a low DO and low pH event. Extreme events in the four variables are highlighted in dark grey. Seasonal refers to the predicted seasonal average value, while threshold refers to the position of the 90%-tile; both are based on a 24-year reference period for the water quality variables and a 18-year reference period for air temperature. Data come from Hudson River, NY – Tivoli South Bay (HUDTS) NERR.

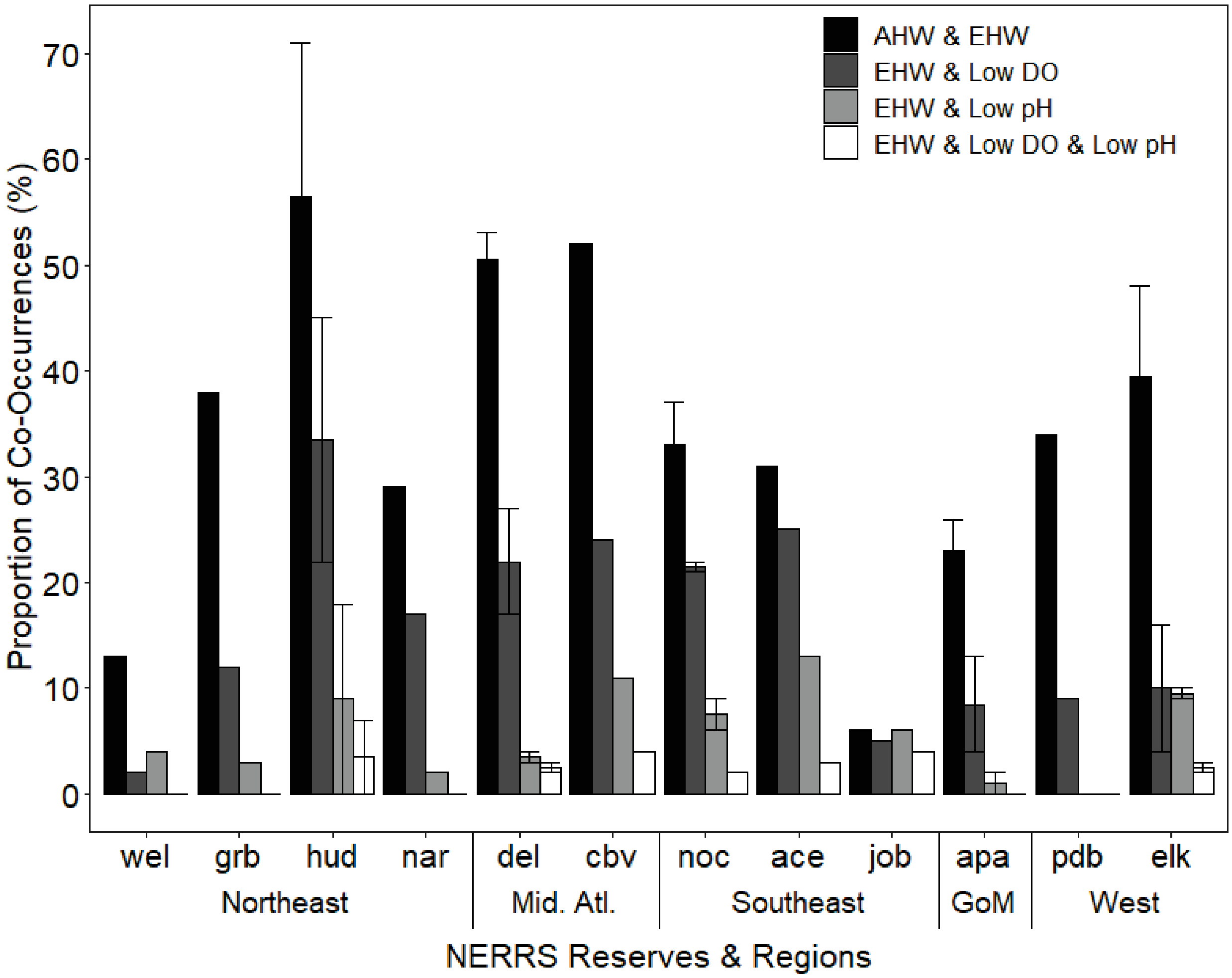


Figure 3. Proportion of co-occurrences between atmospheric & estuarine heatwaves (AHW and EHW, respectively), EHW & low DO events, EHW & low pH events, and EHW, low DO & low pH events. Error bars represent the range between multiple stations within a reserve.

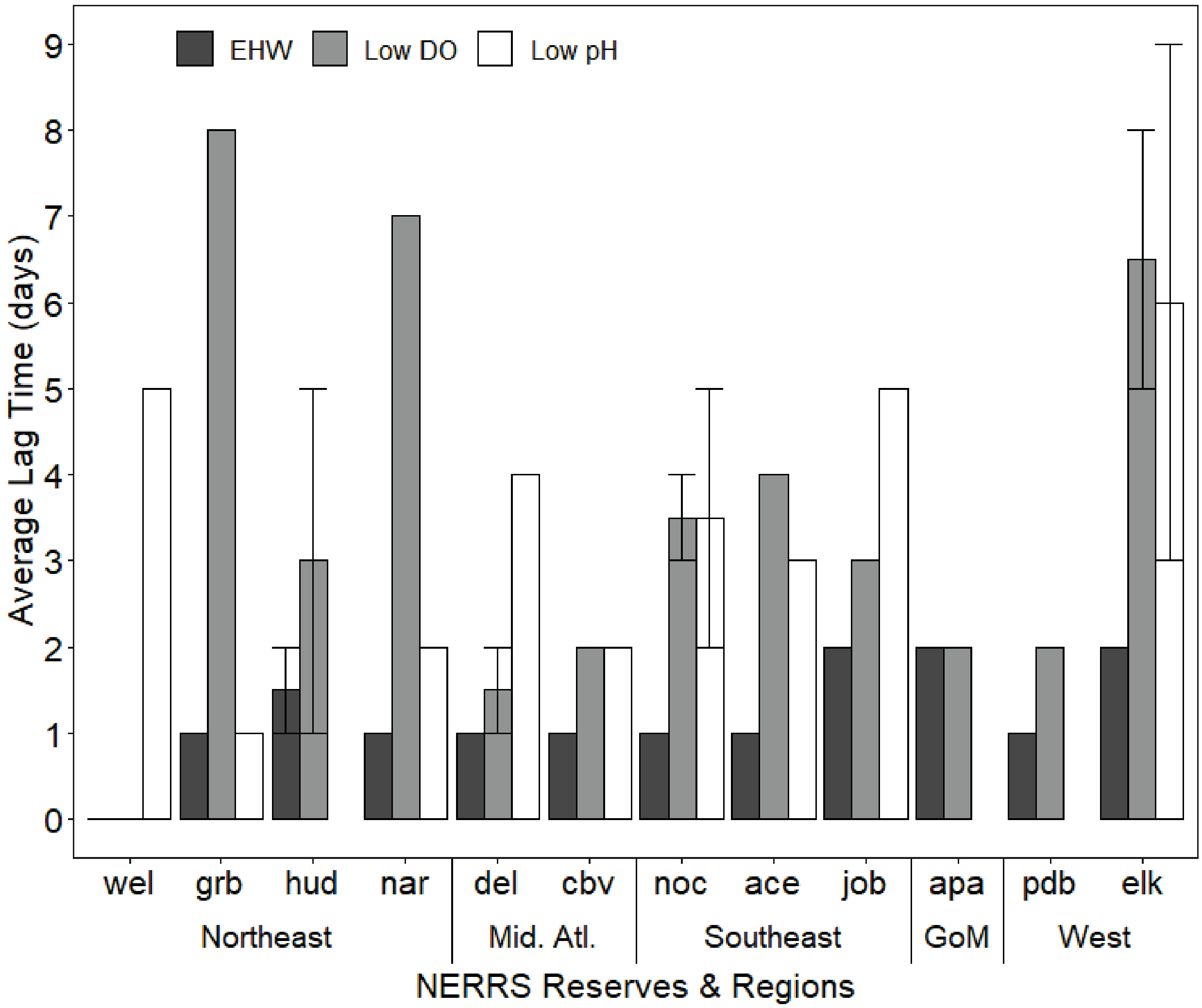
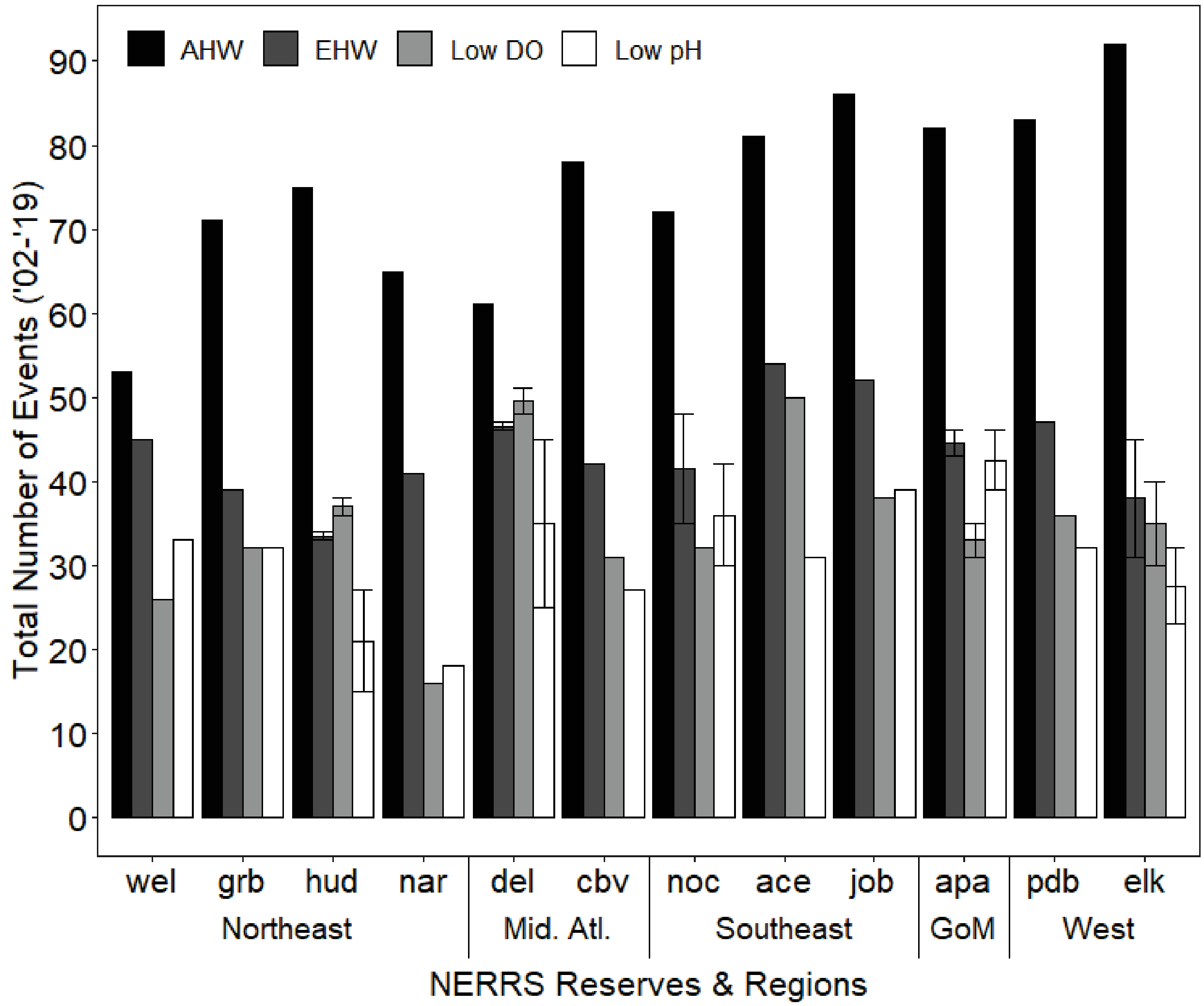


Figure 4. Average lag time between atmospheric-estuarine heatwave (EHW) events (dark grey), EHW-low DO events (light grey), and EHW-low pH events (white). Error bars represent the range between multiple stations within a reserve.



680 Figure 5. Total number of extreme events between 2002-2019 across NERRS reserves and U.S.
681 regions. Error bars represent the range between multiple stations within a reserve.

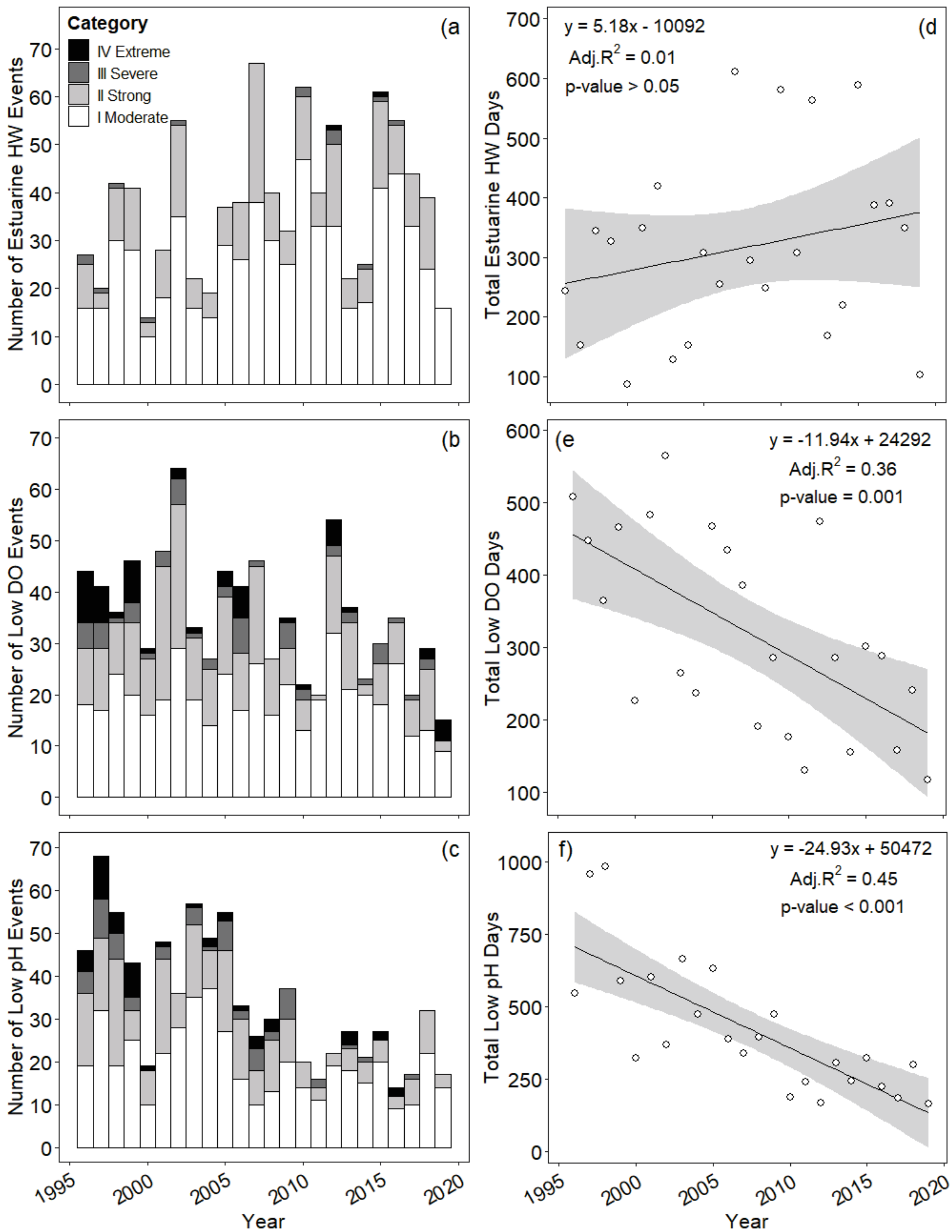
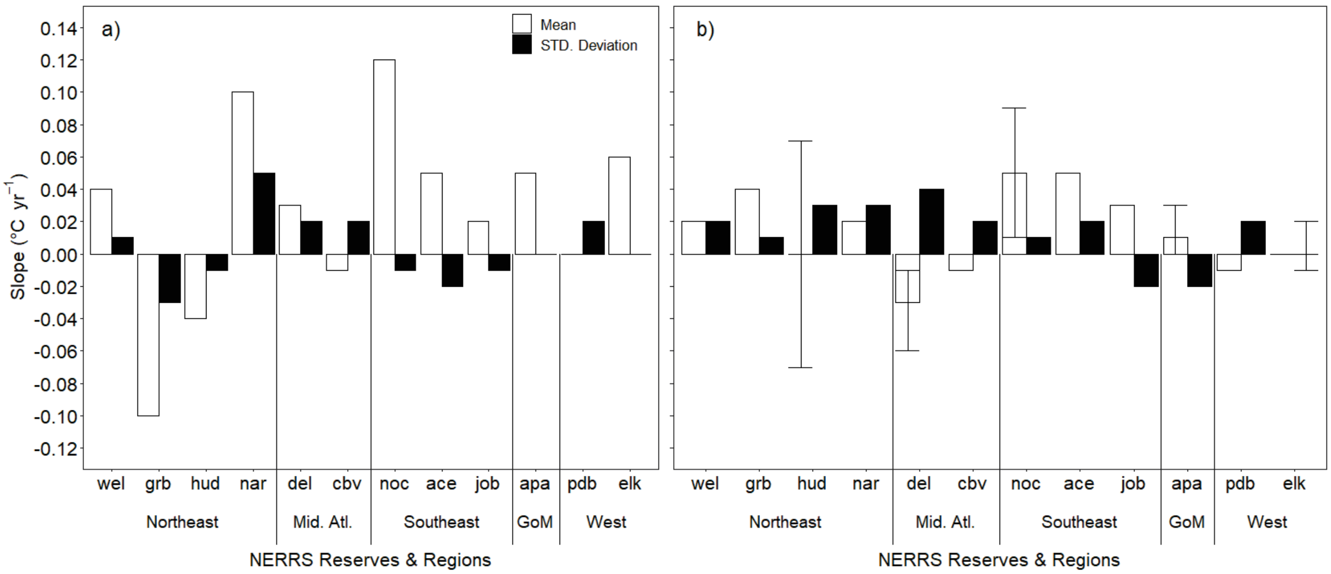


Figure 6. Severity classification of the annual total number of estuarine heatwaves (a), low DO events (b), and low pH events (c) throughout the 17 NERRS stations used in this analysis. Linear regression of total number of estuarine heatwave days (d), low DO days (e), and low pH days (f) amongst the 17 NERRS stations considered.

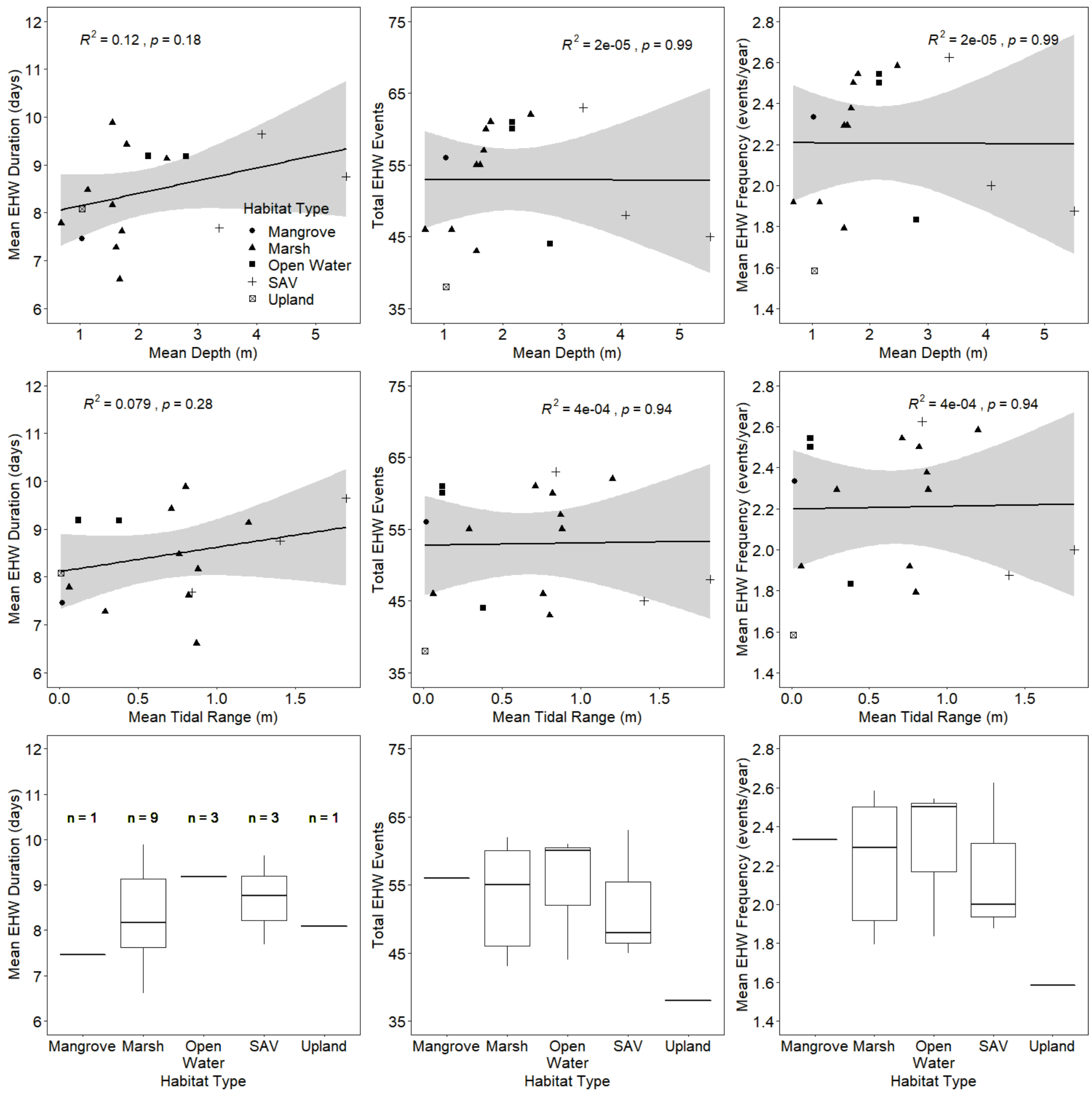
Supplemental Information

SI Table 1. Linear regression trends for 24-years of annual total duration (days $year^{-1}$ ± SE) of estuarine heatwaves (EHW), low DO, and low pH events and associated p-values (p-val.). Statistically significant (p-value < 0.05) trends are bolded and italicized.

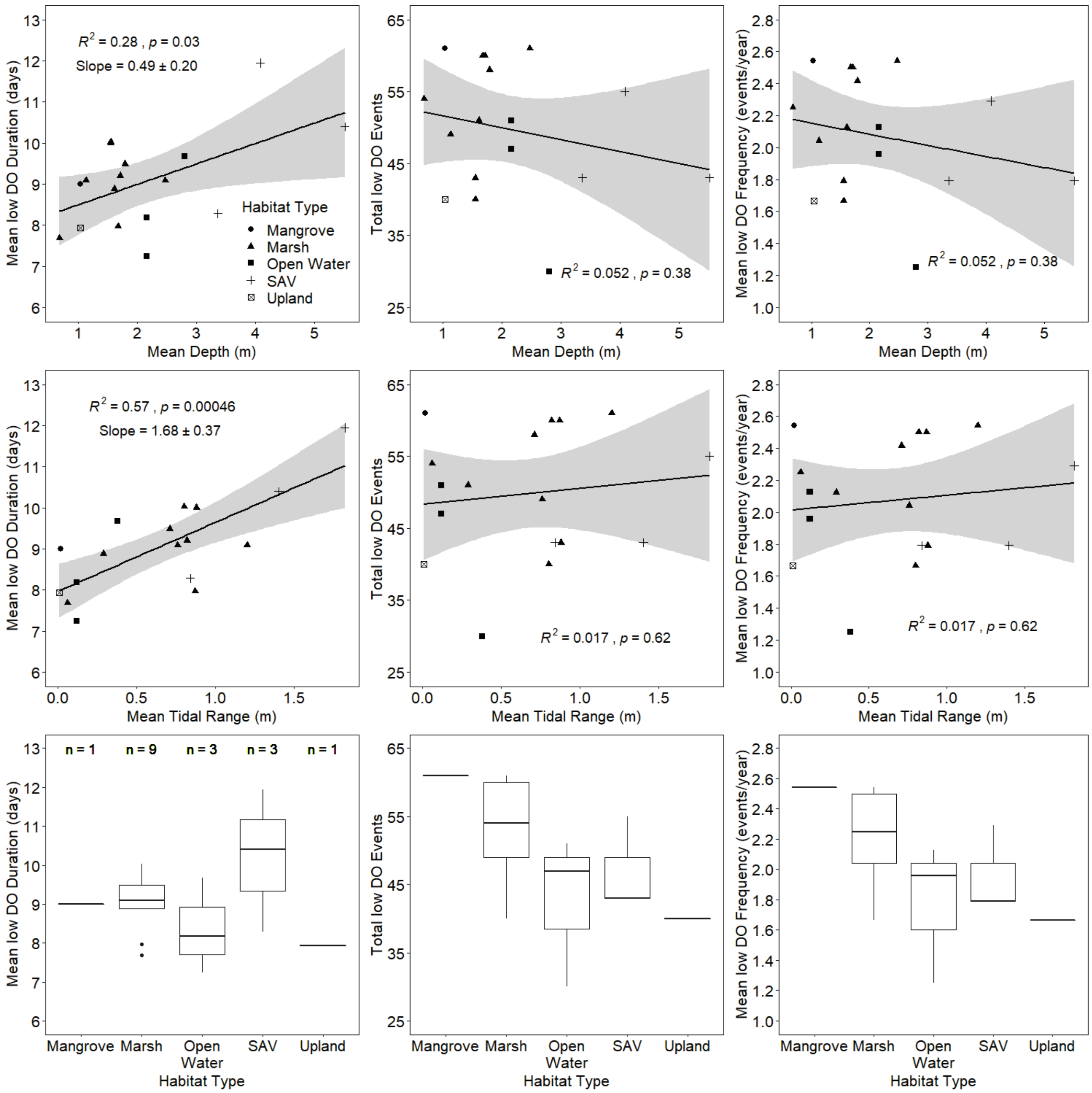
Estuary	Region	Reserve Code	Station ID	EHW	p-val.	Low DO	p-val.	Low pH	p-val.
Wells, ME	Northeast	WEL	IN	1.77 ± 1.12	0.138	-3.61 ± 2.93	0.249	-2.52 ± 2.45	0.325
Great Bay, NH		GRB	GB	0.56 ± 0.57	0.346	-0.83 ± 0.81	0.323	0.01 ± 1.13	0.990
Hudson River, NY		HUD	SK	0.82 ± 0.59	0.185	0.67 ± 0.56	0.251	-1.56 ± 0.91	0.114
			TS	-0.63 ± 0.95	0.518	-0.49 ± 0.92	0.606	***-1.53 ± 0.47***	***0.004***
Narragansett Bay, RI		NAR	PC	1.37 ± 0.85	0.127	***-2.08 ± 0.81***	***0.028***	-3.65 ± 2.16	0.120
Delaware, DE	Mid-Atlantic	DEL	BL	0.20 ± 0.30	0.510	***-0.94 ± 0.43***	***0.042***	-1.49 ± 1.01	0.159
			SL	-0.20 ± 0.37	0.584	0.13 ± 0.66	0.843	***-2.41 ± 1.08***	***0.038***
Chesapeake Bay, VA		CBV	TC	-0.13 ± 0.31	0.687	-0.82 ± 0.58	0.176	-0.25 ± 0.93	0.792
North Carolina, NC	Southeast	NOC	EC	0.05 ± 0.36	0.884	0.14 ± 0.61	0.817	0.24 ± 0.85	0.778
			RC	0.34 ± 0.61	0.585	-0.42 ± 0.91	0.648	0.01 ± 1.01	0.994
ACE Basin, SC		ACE	SP	0.59 ± 0.61	0.349	-0.50 ± 0.77	0.527	-2.37 ± 1.16	0.062
Jobos Bay, PR		JOB	09	***1.23 ± 0.47***	***0.018***	-3.04 ± 1.53	0.070	-0.75 ± 0.74	0.329
Apalachicola, FL	Gulf of Mexico	APA	EB	-0.66 ± 0.51	0.212	0.25 ± 0.63	0.697	-0.04 ± 0.63	0.952
			ES	-0.21 ± 0.57	0.715	-0.49 ± 0.41	0.258	0.13 ± 0.60	0.833
Padilla Bay, WA	West	PDB	BY	0.86 ± 0.63	0.190	0.06 ± 0.44	0.885	-2.59 ± 1.61	0.132
Elkhorn Slough, CA		ELK	AP	0.07 ± 0.73	0.927	***-1.14 ± 0.44***	***0.019***	***-4.07 ± 1.45***	***0.015***
			SM	-0.43 ± 0.82	0.607	-0.58 ± 0.59	0.338	***-3.03 ± 1.20***	***0.022***



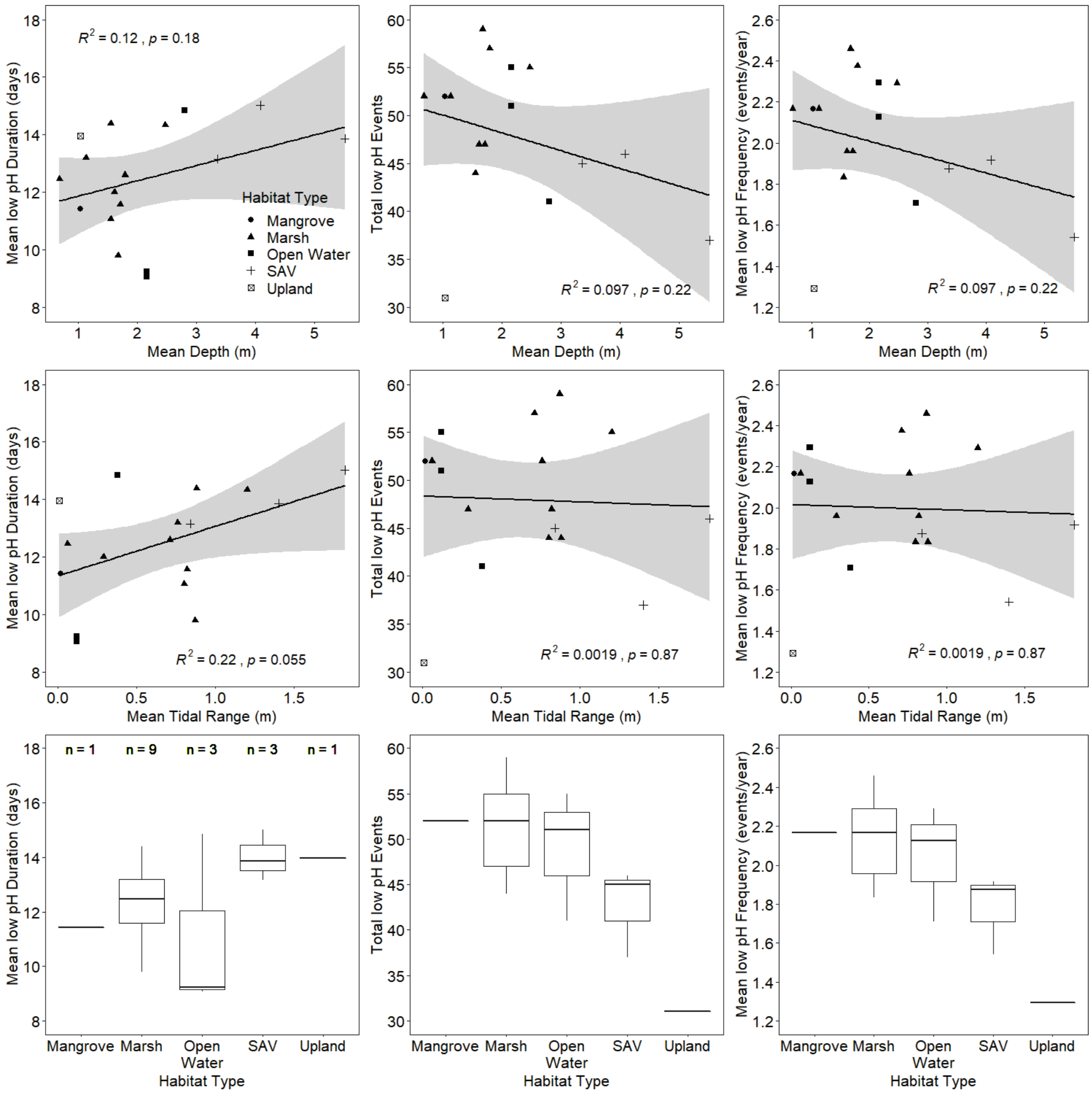
691 SI Figure 1. Slope of the linear regression between mean annual air (a) and water (b) temperature (white columns) and annual standard
692 deviation (black columns) for each NERRS reserve. Error bars represent the range between multiple stations within a reserve.



SI Figure 2. Linear regression and boxplot results for EHW characteristics as a function of NERRS station mean depth, mean tidal range (top and middle row respectively), and estuarine habitat type (bottom row).



SI Figure 3. Linear regression and boxplot results for low DO event characteristics as a function of NERRS station mean depth, mean tidal range (top and middle row respectively), and estuarine habitat type (bottom row).



SI Figure 4. Linear regression and boxplot results for low pH event characteristics as a function of NERRS station mean depth, mean tidal range (top and middle row respectively), and estuarine habitat type (bottom row).