

Quantifying disturbance and recovery in estuaries: tropical cyclones and high frequency measures of oxygen and salinity

CD Buelo^{1*}, AF Besterman², JA Walter^{1,3}, ML Pace¹, DT Ha¹, SJ Tassone¹

¹Department of Environmental Sciences, University of Virginia, Charlottesville, VA, USA

²Department of Biological Sciences, Towson University, Towson, MD, USA

³Center for Watershed Sciences, University of California, Davis, CA, USA

*Corresponding author, cbuelo10@gmail.com

Submitted to Estuaries & Coasts on 7/12/2023, accepted on 8/1/2023

This version of the article has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <https://doi.org/10.1007/s12237-023-01255-1>. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use <https://www.springernature.com/gp/open-research/policies/acceptedmanuscript-terms>

Abstract

Tropical cyclones impact estuaries via a variety of mechanisms including storm surge, flooding from precipitation, high winds, and strong wave action. Prior studies have documented disturbances caused by tropical cyclones, including prolonged periods of depressed salinity from high freshwater discharge and increased or decreased dissolved oxygen concentrations from increased loading of organic matter and/or nutrients. However, most studies of disturbance and recovery in estuaries have been limited to one or a few locations or storm events, limiting generalizations about tropical cyclone impacts and characteristic patterns of ecosystem response and recovery. We analyzed responses to 59 tropical cyclones across 19 estuaries in the eastern United States by applying a new method for detecting disturbance and recovery to long-term and high-frequency measurements of salinity and dissolved oxygen from NOAA's National Estuarine Research Reserve System. We quantified disturbance occurrence, timing, recovery time, and severity. Salinity disturbances generally started earlier and lasted longer than dissolved oxygen disturbances. Estuaries usually recovered within days, but some disturbances lasted weeks or months. Recovery time was positively correlated with disturbance severity for both variables. Tropical cyclone properties (especially precipitation) and location characteristics were both related to disturbance characteristics. Our findings demonstrate the power of high-frequency, long-term, and cross-system data, when combined with appropriate statistical methods, for analyzing hurricanes across many estuaries to quantify disturbances. Estuaries are resilient to hurricanes for the variables and time periods considered. However, persistent impacts can potentially damage resources provided by estuaries, eroding future resilience if hurricanes become more frequent and severe.

Keywords

Estuary, tropical cyclone, disturbance, resilience, salinity, dissolved oxygen

Acknowledgements

This work was supported by a Rapid Response Grant from the University of Virginia Environmental Resilience Institute as well as high performance computing resources provided by University of Virginia Research Computing's Rivanna system.

Introduction

Tropical cyclones (TCs) are large and severe meteorological events that have substantial impacts on coastal populations, communities, and ecosystems, as well as social-ecological systems (Parker et al. 2013, Crosswell et al. 2014, Danielson et al. 2017, Congdon et al. 2019, Armitage et al. 2020). Given these impacts as well as the observed and projected increase in severity and geographic extent of TCs with climate changes (Knutson et al. 2010, Sobel et al. 2016, Balaguru et al. 2022), there have been calls for systematic and coordinated study (Bortone 2006, Pruitt et al. 2019, Hogan et al. 2020). For estuarine ecosystems, understanding the effects of tropical cyclones is complicated by the variety and complexity of alterations in physical, chemical, and biological processes. Detailed studies of tropical cyclones have documented impacts for specific estuaries and storms (e.g., Paerl et al. 2001, Wetz and Paerl 2008, Patrick et al. 2020), but few studies have evaluated numerous storms and sites. One notable exception is Sanger et al. (2002), who examined high frequency water quality measurements from 18 estuaries in the eastern United States impacted by 24 TCs from 1995 – 2000. They found that

more intense storms were associated with stronger water temperature cooling prior to storm passage, and that most other water quality changes were short-lasting, though salinity declines occasionally persisted for months. More recently, Patrick et al. (2022) synthesized multiple response variables across multiple storms and sites. They determined that the relative resistance to disturbance was inversely related to recovery and demonstrated the potential for testing tropical cyclone impacts across many variables and storms to arrive at generalizations.

Quantifying the degree of disturbance and period of recovery is crucial to developing a generalized, and ideally predictive, understanding of tropical cyclone impacts on estuaries (Verdonschot et al. 2013). Such an understanding could inform management actions; for example, knowing the likely duration and severity of low salinity after TCs could be used with metapopulation modeling (Munroe et al. 2013) to choose locations for oyster restoration. Identifying disturbance and recovery requires separating event-driven changes from natural variability (Walter et al. 2022). However, doing so for tropical cyclone impacts on estuaries is challenging, because estuaries are inherently variable through time and space. At a given location within an estuary, what constitutes “normal” values of a variable is determined by interacting processes such as tidal, diel, and seasonal cycles, and weather, as well as local characteristics such hydrologic position, depth, and watershed land use (Tomasko et al. 2006, Wetz and Yoskowitz 2013, Perales-Valdivia et al. 2018, Scanes et al. 2020). The influence of these and other factors changes within and among estuaries (Sanger et al. 2002). As such, establishing baseline conditions requires either a strong understanding of what processes dominate at a given location or extensive prior data.

Despite the many difficulties that limit cross-system and cross-storm examination of disturbances, the data to do so are increasingly available for many estuaries and ecosystem

variables (Gaiser et al 2020, Mills et al. 2008). Long-term monitoring data have provided the opportunity to synthesize ecosystem responses across multiple storms. For example, in Apalachicola Bay, Florida, USA, Edmiston et al. (2008) also used a two-decade record of TC impacts to document contrasting or absent impacts to water depth, water quality, coastal erosion, sea turtle nest loss, and SAV and oyster populations, depending on storm size, speed, severity, landfall, surge height, and precipitation. Paerl et al. (2018) used two decades of monitoring data to distinguish how different storm types (wet vs. dry, windy vs. calm) led to different biogeochemical and phytoplankton responses in the Neuse River Estuary and Pamlico Sound in North Carolina, USA.

In addition to long-term data, developments in sensor technology and remote sensing platforms have made it possible to measure some variables at high frequency. In aquatic ecosystems, *in-situ* sensors can measure temperature, salinity, turbidity, pH, dissolved oxygen, phytoplankton pigment fluorescence, and nutrient concentrations on the scale of seconds to minutes (Glasgow et al. 2004, Fries et al. 2008). Satellite remote sensing can be used to infer shellfish and submerged aquatic vegetation (SAV) coverage (Nieuwhof et al. 2015, Wang et al. 2007), as well as the distribution of turbid waters (Doxaran et al 2006) and phytoplankton (Jiang et al 2020). For some locations, high-frequency, long-term measurements have been collected by monitoring programs like NOAA's National Estuarine Research Reserve System (NERRS; <https://coast.noaa.gov/nerrs/>). Data from this program as well as extensive weather data related to tropical cyclones are openly available.

While having high frequency and long-term data is a helpful first step to understanding patterns and controls of disturbances, appropriate statistical approaches are also required for generating insights from those data. Methods to objectively quantify and compare disturbance

timing, disturbance magnitude and time to recovery are required. For example, Patrick et al. (2022) developed response ratios for variables and scaled them by storm features (maximum winds and precipitation) to enable comparison of disturbance magnitudes. Systematic approaches coupled with sufficient data offer the potential to overcome the inherent spatial and temporal variability and heterogeneous responses to tropical cyclones that have limited comprehensive study of the patterns and controls on disturbance and recovery (Pruitt et al. 2019, Hogan et al. 2020, Patrick et al. 2022).

Here, we present a synthesis of disturbance and recovery measurements for tropical cyclones in estuaries monitored by the National Estuarine Reserve System (NERRS) of the U.S. National Oceanographic and Atmospheric Administration (NOAA) by taking advantage of their long-term data and the known tracks of many storms. We apply a new disturbance detection method designed to quantify the timing and magnitude of the disturbance and the length of recovery in high frequency data (Walter et al. 2022). We apply this method to continuous measures of salinity and dissolved oxygen where long-term observations provide a rich baseline for comparing “normal” variability to storm conditions. Based on hundreds of station-tropical cyclone-variable time series, we ask: 1) What are the characteristics (occurrence, timing, duration, severity) of tropical cyclone disturbances in estuaries? 2) What storm and site properties are associated with changes in disturbance characteristics? and 3) How are estuary resilience and resistance to tropical cyclone disturbances related?

Methods

Study Sites and Data – High frequency time series of water quality parameters from the NOAA’s NERRS program were analyzed to identify disturbance events associated with tropical

storms. NERRS is composed of 29 U.S. estuaries, with each site containing several monitoring stations that collect water temperature, chemistry, nutrient, and pigment data along with meteorology. For this study, we focus on measures of salinity and dissolved oxygen percent saturation (DO % sat) collected by automated sensors at high frequency (15 minutes since 2007, 30 minutes prior) at 19 Atlantic NERRS sites from 2000 to 2018 (Figure 1, Appendix S1 Table S1). Salinity is a critical determinant of habitat suitability for aquatic organisms, varies across most NERRS sites from coastal to inland stations, and is controlled entirely by physical processes. DO % sat is also a critical determinant of habitat suitability. DO % sat (as opposed to concentration) accounts for the effect of water temperature on oxygen solubility and is driven by biological (i.e., primary production and respiration) and physical-chemical processes (e.g., atmospheric exchange, chemical oxidation reactions). Data were obtained from the NERRS Central Data Management Office's Advanced Query System (<http://www.nerrsdata.org>) and all measurements with data quality flags were removed prior to analysis.

Storm Identification – We considered tropical cyclones (TCs) that potentially impacted salinity and oxygen at each NERRS. A two-step process was used to identify TCs and sites for analysis to limit the computationally intensive analyses to cases where TC disturbances were plausible. First, TCs that passed within 250 km of a specific NERRS site were identified using storm tracks from the *hurricaneexposure* and *hurricaneexposuredata* R packages (Anderson et al. 2020a, Anderson et al. 2020b). Second, for each identified TC, potential impacts were determined by visually inspecting plots of salinity and DO % saturation from 30 days prior to 60 days after the date the TC passed closest to the NERRS site. If any variable at any station within a NERRS site appeared to be affected by the TC (defined as an increase or decrease in the mean or variability relative to the 30 day pre-TC passage period), all stations and parameters for that

site and TC were classified as potentially impacted and included in further analyses. While the second step is subjective, including some TCs with minor impacts and possibly missing others with subtle impacts, we aimed to be inclusive in classifying potentially impacted sites/storms to allow the disturbance detection algorithm to quantitatively distinguish events that fell outside the range of historic variability (see below). Alternative methods for identifying TCs that potentially impacted water quality at sites were explored, such as thresholds in meteorological variables (high wind speed, heavy precipitation, drops in barometric pressure). However, the many potential mechanisms by which estuarine salinity and DO can be impacted by TCs (e.g., storm surge, wind driven waves, local precipitation, increased discharge from the watershed) as well as frequent missing meteorological data during TCs precluded the use of such a method. Ultimately, tropical cyclone-associated disturbances were not observed for over 60% of the 955 station-TC-variable combinations analyzed (see Results). Non-detection was expected given the minor impacts of weaker hurricanes and the expected decline in realized TC impacts with distance from the storm track. Our method identified TC impacts on salinity and dissolved oxygen in 38% of the cases and these detections were not limited to only events that caused widespread and severe disturbances.

Disturbance Detection – After potentially impacted NERRS sites for each TC were determined, individual stations within each site with sufficient data were identified for disturbance detection analysis. Station-TC-variable combinations with more than 25% total missing data or a 5 day or longer gap in measurements during the period from 14 days prior to 60 days after the TC was closest were excluded from further analysis, as were combinations with fewer than 8 other years of data to use as reference data (see below) during the same date range meeting the same gap length and total missing data requirements.

We applied a recently developed disturbance detection method designed for use with high frequency data (Walter et al. 2022). The method is implemented in an R package available on GitHub (<https://github.com/jonathan-walter/disturbhf>). It compares the distribution of a variable for rolling windows of the time series within a test period to a reference period using the empirical cumulative distribution function (ECDF). The analysis consists of three steps. First, the difference statistic time series $d_w(t)$ is calculated for each window within the test period:

$$d_w(t) = \sum_{i=1}^N \left| ECDF(x_{test,W}(t)) - ECDF(x_{ref,W}(t)) \right|_i * dx$$

where $x_{test,W}(t)$ are the variable values within a rolling window of width W centered at time t within the test period (Figure 2A), $x_{ref,W}(t)$ are the variable values in a reference period, N is the number of intervals at which to evaluate the EDCFs (here we use 1000) over the range of observed values in the test and reference windows, and dx is the width of those intervals (equal to $(x_{max} - x_{min}) / N$; Figure 2B,C). x_{ref} can be defined to be either fixed (all values within the reference period are used) or adaptive, where rolling windows of a specified width within the reference period are used to account for seasonal trends. As tropical cyclones occur from summer into late fall when seasonal changes in water quality might be expected, we use an adaptive reference period so rolling windows within the test period are compared to windows centered at the same day of year in the reference years.

In the second step, the d_w time series is rescaled based on the variability observed in the reference period. This is done by calculating $d_{w,ref}$ as above, but $x_{test,W}(t)$ is instead defined by rolling windows of observations within the reference period. The mean ($\mu_{dw,ref}$) and standard deviation ($\sigma_{dw,ref}$) of $d_{w,ref}$ are used to rescale $d_w(t)$ as a z-score (Figure 2D):

$$z(t) = (d_w(t) - \mu_{dw,ref}) / (\sigma_{dw,ref})$$

Finally, user-specified thresholds in $z(t)$ that define disturbance ($thresh_{dist}$) and recovery ($thresh_{recov}$) are applied to identify the timing of disturbance events (initiation and conclusion). Short disturbances and recoveries can optionally be combined or removed using minimum disturbance and recovery lengths. We used a test window width of 3 days and a reference window width of 6 days as a balance between the ability to detect shorter disturbances vs. power to accurately characterize variable distributions based on initial exploration of cases with obvious TC impacts. We also required that disturbances and recoveries last for at least 24 hours and set $thresh_{recov} = 0.5 * thresh_{dist}$. Based on a disturbance threshold sensitivity analysis (Appendix S1, Figure S1) we used $thresh_{dist} = 2$. Choice of $thresh_{dist}$ is subjective; our goal in using a single value was to allow for comparison across many TC-station-variable combinations that limited disturbances to those that were likely caused by storms. In studies focused on a smaller number of disturbances, site- or storm-specific values of $thresh_{dist}$ and test and reference window widths could be used. Lower $thresh_{dist}$ values will lead to more disturbances detected and generally longer durations (and vice versa). Narrower window widths are more likely to detect short disturbances but can lead to false detections due to low sample size.

The detection algorithm was iterated through the rolling window time steps from 14 days before TC passage to 60 days after. The shorter analysis window compared to the *Storm Identification* window was chosen to limit these computationally intensive calculations to only time periods when a disturbance was potentially caused by TC impacts; additionally, the longer pre-TC window was not needed because a minimum 8 years of data were used for the reference window. If a disturbance was detected in the initial analysis but no recovery occurred within 60 days, the post-TC window was extended for 60 additional days until a recovery was detected.

Disturbance Event Characteristics and Drivers – After setting the algorithm parameters

and identifying disturbance events, we quantified disturbance characteristics and explored potential relationships with station and TC variables. Disturbance event characteristics included disturbance occurrence, timing relative to when the center of the TC was closest, length of time between disturbance initiation and recovery, and severity (peak $z(t)$ during the disturbance). Disturbance events were limited to first occurring disturbances starting from 3 days before to 30 days after each TC was closest to each NERRS site. Thirty days after TC passage was chosen to attempt to capture all disturbances initiated by a storm, accounting for potential disturbance delays such as hydrologic lags and biological feedbacks. Following disturbance and recovery detection, mean salinity and dissolved oxygen during the disturbance were calculated and compared to means during the same date range in all other years of available data.

Potentially explanatory station variables included mean salinity and depth as proxies for relative location within the estuary (oceanic vs. inland), mean tidal range as a proxy for tidal influence, and standard deviation of salinity as a proxy for variability in the contribution of upstream vs. ocean water sources. Tropical cyclone traits included closest TC distance to the NERRS site, maximum wind gust speed, duration of wind gusts over 20 m s^{-1} , total TC precipitation, and storm surge height. TC distance and winds were determined from the *hurricaneexposure* R package (Anderson et al. 2020a); winds were from the population-weighted center of the closest county. This source of wind data was chosen because it provides consistently-modeled wind speeds across all TCs and locations in this study; while actual measurements at each location would be preferable, not all stations had meteorological data and often data were lost during TC impacts. Total TC precipitation was obtained from the PRISM reanalysis product (PRISM Climate Group 2021); daily precipitation totals from 3 days before to

7 days after the TC passed closest to the NERRS site were summed. Storm surge height was calculated from depth observations at each station as the difference between the maximum depth observed from 3 days before to 7 days after TC passage and the maximum depth from the preceding two weeks. While imperfect due to not accounting for whether storm surge occurred at high or low tide, or for longer period tidal cycles, the storm surge metric provides an indicator of how high water got at each location within an estuary. Total precipitation was square-root transformed and peak severity was log transformed prior to regressions to increase normality. Relationships between disturbance characteristics (occurrence, timing, length, and severity) and potential driver variables (TC precipitation, station mean tidal range, etc.) were assessed using multiple logistic and multiple linear regression; separate models were computed for each combination of disturbance characteristic (occurrence, timing, length, or severity) and ecosystem variable (salinity or dissolved oxygen) and the best model for each disturbance characteristic – ecosystem variable combination was determined using AIC and stepwise model selection (Venables and Ripley 2002).

Results

Disturbance Examples - The disturbance detection method identified different types and durations of anomalous salinity and oxygen conditions in the periods near TC passage. For example, Hurricane Florence in 2018 caused significant declines in salinity at North Inlet-Winyah Bay NERR's Debidue Creek station from over 35 to 1-2 psu, and it took 39 days to return to normal values (Figure 3A, Table 1). At Jacques Cousteau NERR, there was a sharp but short (~ 2 days) increase in salinity from 15-20 to 29 psu in response to storm surge from Hurricane Sandy (2012), after which salinity was lower than before the storm but still within the range of normal variability (Figure 3B). The distribution difference statistic also increased in

response to storms that impacted variability more strongly than mean values; at Great Bay NERR's Squamscott River station, Hurricane Hanna in 2008 significantly decreased the amplitude of dissolved oxygen oscillations but the mean remained near 85% saturation (Figure 3C). The normalized distribution difference statistic (z-score) increased to ~ 1.4 over this period, but not enough to trigger a disturbance event for a threshold of 2. A more severe disturbance in dissolved oxygen occurred at Rookery Bay NERR's Lower Henderson Creek station, where Hurricane Irma caused a crash in dissolved oxygen to near 0% saturation (Figure 3D). The z-score for this disturbance peaked at nearly 8 standard deviations above the reference period mean, indicating a highly anomalous event for the location.

Disturbance Event Characteristics – A wide range of disturbance timing, length, and severities were observed. Disturbances were detected for 40% of salinity cases and 37% of dissolved oxygen cases. Over 50% of detected disturbances began within 2.5 days (salinity) and 5.1 days (dissolved oxygen) of when the eye of a tropical cyclone passed closest to each NERRS site (Figure 4). Initiation of disturbances peaked for salinity from 0.5 to 1.5 days after TC passage, with the cluster of most frequently observed times from 3 days before to 4.5 days after (Figure 4A). For dissolved oxygen, disturbances most frequently began from 2.5 to 3.5 days after TC passage; the cluster of most frequent times peaked much lower and was wider relative to salinity (Figure 4B).

The majority of disturbances lasted less than a week (median of 6.3 days for salinity, 4.6 days for DO % sat; Figure 5), but the distribution of disturbance lengths had a long tail, with disturbances > 50 day seen for both DO% sat and salinity (Figure 5A, B). Salinity disturbances between 1 and 8 days accounted for 61% of salinity disturbances, with a large drop in the number of observed disturbances longer than 14 days (Figure 5A). The cluster of most common

disturbance lengths for dissolved oxygen was shifted to slightly lower values; the most common disturbance lengths was 1 – 6 days and a drop in observed disturbance lengths occurred after 12 days (Figure 5B). Of the 24 disturbances that lasted longer than 30 days, 20 were salinity disturbances (max = 98 days) and 4 were dissolved oxygen disturbances (max = 59 days; Table 1).

TC disturbances tended to decreased salinity and DO % values, though the relative proportion differed by variable. Only 8% of salinity disturbances had a higher mean value during the disturbance than in reference years, while 92% had a lower mean (Figure 6). For dissolved oxygen, 37% had a higher mean value during the disturbance and 63% were lower. Only 16.5% of all the dissolved oxygen disturbances had a mean DO % value greater than 100% (Figure 6).

There was a wide range in the peak severity value of each disturbance, with peak severity z-scores between 2 and 5 being most common (80% of cases for salinity, 84% for dissolved oxygen) and several values between 5 and 15 occurring for each variable (Figure 7). There was a significant, positive relationship between log-transformed peak severity and log-transformed disturbance length for both salinity and DO % saturation (Figure 7), with very similar correlation coefficients ($r = 0.51$ and 0.62 for salinity and oxygen, respectively, $p < 0.001$ for both variables).

Potential disturbance drivers – At least one potential driver was identified for all disturbance event characteristics for both dissolved oxygen and salinity based on regression analysis (Table 2; model coefficients are in Tables S2 – S9). Higher precipitation and mean salinity were associated with increased occurrence of salinity disturbances, while increases in tidal range were associated with lower salinity disturbance occurrence. For dissolved oxygen, TCs with higher precipitation and stations with greater depths had more disturbances, while

stations with larger tidal ranges had fewer disturbances. TCs with longer duration of wind gusts over 20 m/s were associated with later occurring salinity disturbances; increased precipitation led to earlier occurring salinity disturbances. For dissolved oxygen disturbance timing, stations with higher salinity had later starting disturbances while TCs with more precipitation that passed closer to the NERSS site had earlier disturbances. Longer salinity disturbances were positively correlated with tidal range, depth, wind gust duration, and maximum wind speed, and negatively correlated with mean station salinity and storm surge height. Longer dissolved oxygen disturbances were positively correlated with TC precipitation and station salinity variability, and negatively correlated with station mean salinity and TC wind gust duration. Several TC and station variables were positively and negatively related to peak severity for salinity disturbance (Table 2), while the only driver variable that had a significant relationship with peak disturbance severity for DO % sat was maximum storm surge height.

The mean magnitudes of change associated with regression model variables can be readily estimated for ranges typically associated with TCs. For example, the simplest model is for the severity of DO % sat (Table S9) which changes as a z-score from 4.2 to 6.8 standard deviations above the mean for storm surges of 1 and 4 meters, respectively. As a second example, TC precipitation often falls in the range of 100 to 300 mm (Table 1) but was as high as 680 mm in our data set. Assuming an estuary with a mean salinity of 20 psu and a salinity standard deviation of 2 psu, the predicted mean length of a DO disturbance would be 1.6 days for a storm precipitation of 100 mm, 2 days for a precipitation of 300 mm, and 2.4 days for a precipitation of 700 mm based on the model (Table S7).

Discussion

We identified and characterized tropical cyclone-associated disturbances across 19 estuaries and 59 different TCs in the eastern United States utilizing high-frequency, long-term data. The disturbance detection method distinguished diverse disturbance types for two important ecosystem state variables (Figure 3). In relation to our first question about disturbance characteristics, there was a wide range of disturbance timing, length, and severity (Figures 4, 5, 7). For our second question, both TC (especially precipitation) and site properties were associated with changes in disturbance characteristics (Table 2).

Across all TC-estuary combinations analyzed, dissolved oxygen saturation and salinity had a similar number of disturbances detected, though there were slightly more disturbances detected for salinity (40% vs. 37%). The lower number of disturbances detected prior to TCs for salinity relative to DO % saturation (Figure S1) suggests that disturbances in this variable are more closely tied to TC events, while additional mechanisms beyond TCs are also important generators of anomalous DO % saturation values. The roughly 60% of cases without detected disturbances highlight the importance of considering instances when TCs both do and do not cause impacts to provide a complete understanding. The common case study approach may present a bias towards the perception that TCs usually have large impacts, when in fact often they do not, at least for the variables we consider in this study.

Both increases and decreases in salinity due to TCs have been widely documented, but our findings indicate that freshwater inputs from precipitation and increased river/stream discharge is the dominant mechanism of salinity disturbance as opposed to storm surge (91% decreases vs. 9% increases; Figure 6). Biological processes that alter oxygen concentrations (i.e., respiration and primary production) are highly dynamic in time and related to several environmental drivers (e.g., nutrient loading, hydrology, temperature, solar radiation, and

salinity; Caffrey et al. 2014, Murrell et al. 2018, Tassone and Bukaveckas 2019), and are a likely explanation for the observation of oxygen anomalies not associated with TCs (Figure S1;). Biological mechanisms could also explain the observed lag in dissolved oxygen disturbance initiation relative to salinity. Dissolved oxygen disturbances were generally shorter in duration than salinity disturbances. While salinity and oxygen share many of the physical processes that promote recovery (e.g., stream and river discharge, tidal exchange), equilibration with atmospheric oxygen concentrations, especially under vigorous mixing in estuaries, could explain the faster recovery of DO % sat. relative to salinity (Kremer et al. 2003). While we were unable to assess water column mixing in this study because almost all NERRS stations have only a single, near-bottom sensor, future work looking at surface vs. bottom disturbances could provide important insights, especially for locations where increased freshwater inputs following TCs can lead to vertical salinity gradients and stratification (Mallin et al. 2002, Wetz and Yoskowitz 2013).

While short disturbances (< 7 days) were most frequent for both variables, longer disturbances were also common. Forty-seven percent of salinity disturbances lasted longer than 7 days, as did 28% of dissolved oxygen disturbances. For many organisms that inhabit estuaries, these disturbances likely represent prolonged periods of stress and require substantial movement for mobile species. Low oxygen can cause fish and shellfish die-offs as documented for some hurricanes (Paerl et al. 1998, Paerl et al. 2001, Parker et al. 2013). While many estuarine species tolerate relatively large salinity ranges, extreme salinity fluctuations associated with TC disturbances may exceed tolerances (Du et al. 2021). Though rare, we also found several cases where disturbances persisted for more than 30 days (Table 1). The extremely slow recovery rates for these events could arise from different mechanisms. Long duration, low salinity disturbances

can result from high freshwater discharge (Paerl et al. 2001, Du et al. 2019). For dissolved oxygen disturbances, extreme loading of organic matter from the watershed into estuaries can stimulate high respiration that draws down oxygen concentrations (Paerl et al. 2018). Alternatively, TCs can increase nutrient loading from internal or external sources, leading to algal blooms that increase oxygen concentrations (Shen et al. 2008). Though nutrient concentrations often decline quickly following hurricanes, recycling can maintain primary producer biomass (Peierls et al. 2003). The number of higher and lower dissolved oxygen concentrations during disturbances (39% and 61%, respectively; Figure 5) suggests that both organic matter driven respiration inputs and nutrient driven production can be important drivers of oxygen disturbances, but that oxygen consumption is usually greater than production during TC disturbances. Increased phytoplankton biomass, which has often been observed following TCs (Wetz and Paerl 2008, Herbeck et al. 2011, Philips et al. 2020) would be expected to cause DO saturation values over 100% during severe blooms , but such cases were relatively rare in our data set. TCs can also cause declines in phytoplankton due to light-limitation from high concentrations of suspended matter and organic carbon as well as losses from high flushing rates (Paerl et al. 1998, Malin et al. 2002, Paerl et al. 2018).

Our findings also offer insights into the overall time scales of TC disturbances at estuaries. NERR sites experienced from 1 to 18 TCs over the period of record considered in this study (Table S1). As an intermediate example, Apalachicola Bay had 9 TCs from 2000 to 2018. Assuming median disturbance lengths for each storm, this amounts to 57 and 41 days respectively of extreme conditions (greater than two standard deviations departure from means) for salinity and oxygen, respectively. These disturbance days are less than 1% of the 19-year record. Even if some rare, long-duration disturbances also occurred, this simple calculation

indicates estuaries are only impacted by TCs a small amount of the time in a cumulative sense. However, disturbances in salinity and dissolved oxygen may influence other processes (e.g., reproduction of longer-lived organisms, slow biogeochemical processes) that have lasting impacts. Thus, continued study of storm impacts is warranted, especially interactions that might ramify from short-term changes in easily measured physical-chemical variables considered in this study, such as salinity, dissolved oxygen, and turbidity impacts on benthic organisms, SAV, and fish populations (Paerl et al. 1998, Mallin et al. 1999, Mallin et al. 2002, Carlson et al. 2010). The topic gains significance if TCs increase in severity and/or geographic range with climate change as projected (Knutson et al. 2010, Sobel et al. 2016, Balaguru et al. 2022).

For our third question on the relationship between resilience and resistance, we found a positive relationship between disturbance length and disturbance severity. These two disturbance characteristics quantify metrics of ecosystem stability and recovery; resistance and resilience have taken several definitions and received considerable attention in ecology and other fields. Using the definitions of Pimm (1984), disturbance length is inversely related to resilience (longer disturbance lengths correspond to lower resilience) and quantifies the ability and speed of a system to recover after a perturbation, while disturbance severity is inversely related to ecosystem resistance (higher severity corresponding to low resistance) and measures a system's ability to oppose change. In an analysis of coastal ecosystem response to Hurricane Harvey in 2017 using similar definitions, Patrick et al. (2020) found a negative relationship between resilience and resistance for several types of estuary variables (hydrology, hydrography, biogeochemistry, biota) in the response of estuaries in Texas to Hurricane Harvey. Our disturbance length metric is similar to the return times Patrick et al. (2020) observed, though the resistance measures are not directly comparable (log response ratio vs. peak z-score). Another

recent synthesis of response to TCs for thousands of time series from 26 TCs in the Northern Hemisphere (Patrick et al. 2022) across different ecosystem and variable types also found a negative relationship between resistance and resilience using definitions that consider disturbance driver and system response magnitudes. Applying similar definitions to our results (resistance = $-1 * \ln(\text{peak z-score} / \text{TC precipitation})$, resilience = $\ln(\text{peak z-score} / \text{disturbance duration})$), we find weak negative correlations between resilience and resistance for salinity and dissolved oxygen salinity ($r = -0.14$ and -0.23 for salinity and oxygen, respectively). Taken together, our findings and other recent studies suggest that the relationship between resistance and resilience is likely dependent on the variable considered, the geographic scale (e.g., within/among estuaries), and ecosystem type; as well as the specific quantitative definition used for resistance and resilience.

By analyzing the impact of many tropical cyclones on several estuaries, we were able to examine the role that TC and site properties play in determining disturbance characteristics. Both TC and site properties were important. Total precipitation was the most common TC variable related to disturbance responses, with higher precipitation associated with more likely, earlier starting, longer, and more severe disturbances. This finding is important given the projected increase in tropical cyclone severity and precipitation amounts due to climate change (Patricola and Wehner 2018). Mean salinity was the most common station property associated with disturbance responses. Locations with higher salinity had more severe and earlier salinity disturbances, as well as later dissolved oxygen disturbances. TCs with higher maximum and longer duration wind speeds had earlier and longer lasting disturbances. These general patterns are a first step to developing a predictive understanding of disturbance characteristics and demonstrate the disturbance detection algorithm's ability to quantitatively characterize

disturbances across different locations, storms, and variables. However, additional storm and site (see below) and interactions between them are undoubtedly needed to explain the highly variable disturbance characteristics observed in this study. Identifying drivers of disturbances is important to identify current estuarine locations that are highly susceptible to disturbance and how future climate change may impact responses. Insights may offer ways to increase ecosystem resilience through management of infrastructure (e.g., water retention/release), habitat restoration to promote refuges and portfolio effects (Schindler et al. 2015), and watershed land-management to limit run-off during extreme storms.

While the disturbance detection method identified and characterized disturbances, it also has limitations. The method quantifies any difference between the distribution of values in the test and reference periods but does not distinguish between different types/directions of disturbances (e.g., if a change is to higher or lower values, or increased/decreased variability). There also is not a direct correspondence between the disturbance statistics (d_w or z-score) and physically meaningful ecosystem state values or thresholds (e.g., oxygen or salinity concentrations at which organisms are harmed). These issues can be addressed by first using the method to demarcate disturbances, then to compare differences in the mean, variance, etc. of the test and reference years within those periods (Figure 6). Finally, the method requires both high frequency and long-term measurements to identify disturbances, which limits the variables and locations to which it can be applied. However, these types of data are increasingly available from sensors that can measure important ecosystem state variables (Porter et al. 2012). The data from NOAA's National Estuarine Research Reserve System illustrates the immense value of long-term programs measuring the same variables at different sites, especially for events that are unpredictable but have large consequences like tropical cyclones.

Despite these limitations, our findings represent an advance in the study of patterns and drivers of disturbance at broad spatial and temporal scales in estuaries. By detecting disturbances that cause diverse deviations from baseline variability and quantifying several important disturbance characteristics, our method overcomes many of the constraints that have previously limited studies to one or a few TCs and/or locations (Pruitt et al. 2019). The approach fits naturally within proposed frameworks for understanding disturbance in ecological and socio-economic systems (Gaiser et al. 2020). For TCs specifically, Hogan et al. (2020) recently provided a framework for evaluating ecosystem component response to disturbances including a detailed conceptual diagram (see their Figure 2). This study includes many of the framework components and applies them to provide quantitative generalization across many locations and storms: disturbances to salinity and dissolved oxygen in estuaries generally start soon after tropical cyclones pass and typically recover within days, though weeks and months long disturbances do occur. Most (~90%) salinity disturbances cause declines in mean salinity driven by precipitation and discharge as opposed to increases from storm surge. In contrast, dissolved oxygen disturbances were more evenly split between increases and decreases. Properties of both tropical cyclones and the locations they impact are related to disturbance response. Future work could extend our findings to additional drivers and mechanisms including the role of upstream land use, antecedent conditions, estuary or habitat type, and hydrodynamics; the effect of disturbance impacts on specific biota; and other variables besides salinity and dissolved oxygen. For these and other ecosystem variables where measurements can be collected by sensors and analyzed in near-real time, knowledge that a disturbance is starting also offers the exciting possibility of quickly directing additional data collection and management actions to minimize impacts.

References

- Anderson, B, M Yan, J Ferreri, W Crosson, M Al-Hamdan, A Schumacher, D Eddelbuettel. 2020a. hurricaneexposure: Explore and Map County-Level Hurricane Exposure in the United States. R package version 0.1.1. <http://CRAN.R-project.org/package=hurricaneexposure>.
- Anderson B, A Schumacher, W Crosson, M Al-Hamdan, M Yan, J Ferreri, Z Chen, S Quiring, S Guikema. 2020b. hurricaneexposedata: Data Characterizing Exposure to Hurricanes in United States Counties. R package version 0.1.0. <https://github.com/geanders/hurricaneexposedata>.
- Armitage, AR, CA Weaver, JS Kominoski, SC Pennings. 2020. Resistance to Hurricane Effects Varies Among Wetland Vegetation Types in the Marsh–Mangrove Ecotone. *Estuaries and Coasts*. 43: 960–970.
- Balaguru, K, GR Foltz, LR Leung, W Xu, D Kim, H Lopez, R West. 2022. Increasing hurricane intensification rate near the US Atlantic coast. *Geophysical Research Letters*, 49(20), e2022GL099793.
- Bortone, SA. 2006. Recommendations on establishing a research strategy in the Gulf of Mexico to assess the effects of hurricanes on coastal Ecosystems. *Estuaries and Coasts*. 29(6A): 1062-1066.
- Caffrey, JM, MC Murrell, KS Amacker, JW Harper, S Phipps, MS Woodrey. 2014. Seasonal and Inter-annual Patterns in Primary Production, Respiration, and Net Ecosystem Metabolism in Three Estuaries in the Northeast Gulf of Mexico. *Estuaries and Coasts*. 37: S222-S241.
- Carlson, PR, LA Yarbrow, KA Kaufman, RA Mattson. 2010. Vulnerability and resilience of seagrass to hurricane and runoff impacts along Florida's west coast. *Hydrobiologia*. 649: 39-53.
- Congdon, VM, C Bonsell, MR Cuddy, KH Dunton. 2019. In the wake of a major hurricane: differential effects on early vs. late successional seagrass species. *Limnology and Oceanography Letters* 4: 155–163.
- Crosswell, JR, MS Wetz, B Hales, and HW Paerl. 2014. Extensive CO2 emissions from shallow coastal waters during passage of Hurricane Irene (August 2011) over the mid-Atlantic Coast of the U.S.A. *Limnology and Oceanography*. 59: 1651–1665.
- Danielson, TM, VH Rivera-Monroy, E Castañeda-Moya, H Briceño, R Travieso, BD Marx, E Gaiser, LM Farfán. 2017. Assessment of Everglades mangrove forest resilience: Implications for above-ground net primary productivity and carbon dynamics. *Forest Ecology and Management*. 404:115-125.
- Doxaran, D, P Castaing, SJ Lavender. 2006. Monitoring the maximum turbidity zone and detecting fine-scale turbidity features in the Gironde estuary using high spatial resolution

536 satellite sensor (SPOT HRV, Landsat ETM+) data. *International Journal of Remote*
537 *Sensing*. 27(11): 2303-2321.

538 Du, J, K Park, TM Dellapenna, JM Clay. 2019. Dramatic hydrodynamic and sedimentary
539 responses in Galveston Bay and adjacent inner shelf to Hurricane Harvey. *Science of the*
540 *Total Environment*. 653: 554-564.

541 Du, J, K Park, C Jensen, TM Dellapenna, WG Zhang, Y Shi. 2021. Massive oyster kill in
542 Galveston Bay caused by prolonged low-salinity exposure after Hurricane Harvey.
543 *Science of the Total Environment*. 774: 145132.

544 Edmiston, HL, SA Fahrny, MS Lamb, LK Levi, JM Wanat, JS Avant, K Wren, NC Selly. 2008.
545 Tropical storm and hurricane impacts on a Gulf Coast estuary: Apalachicola Bay, Florida.
546 *Journal of Coastal Research*. 55: 38-49.

547 Fries, DP, SZ Ianov, PH Bhanushali, JA Wilson, HA Broadbent, AC Sanderson. 2007.
548 Broadband, low-cost, coastal sensor nets. *Oceanography*. 20(4): 150-155.

549 Gaiser, EE, DM Bell, MCN Castorani, DL Childers, PM Groffman, CR Jackson, JS Kominoski,
550 DPC Peters, STA Pickett, J Ripplinger, JC Zinnert. 2020. Long-Term Ecological
551 Research and Evolving Frameworks of Disturbance Ecology. *BioScience*. 70(2): 141-
552 156.

553 Glasgow, HB, JM Burkholder, RE Reed, AJ Lewitus, JE Kleinman. 2004. Real-time remote
554 monitoring of water quality: a review of current applications, and advancements in
555 sensor, telemetry, and computing technologies. *Journal of Experimental Marine Biology*
556 *and Ecology*. 300: 409 – 448.

557 Herbeck, LS, D Unger, U Krumme SM Liu, TC Jennerjahn. 2011. Typhoon-induced
558 precipitation impact on nutrient and suspended matter dynamics of a tropical estuary
559 affected by human activities in Hainan, China. *Estuarine, Coastal and Shelf Science*. 93:
560 375 – 388.

561 Hogan JA, RA Feagin, G Starr, M Ross, T Lin, C O’Connell, TP Huff, BA Stauffer, KL
562 Robinson, and 34 others. 2020. A Research Framework to Integrate Cross-Ecosystem
563 Responses to Tropical Cyclones. *BioScience*. 70(6): 477–489.

564 Jiang, L, T Gerkema, JC Kromkamp, D van der Wal, PMC De La Cruz, K Soetaert. 2020.
565 Drivers of the spatial phytoplankton gradient in estuarine–coastal systems: generic
566 implications of a case study in a Dutch tidal bay. *Biogeosciences*. 17: 4135-4152.

567 Knutson, TR, JL McBride, J Chan, K Emanuel, G Holland, C Landsea, I Held, JP Kossin, AK
568 Srivastava, M Sugi. 2010. Tropical cyclones and climate change. *Nature Geoscience*. 3:
569 157-163.

570 Kremer, JN, A Reischauer, C D’Avanzo. 2003. Estuary-specific variation in the air-water gas
571 exchange coefficient for oxygen. *Estuaries*. 26(4A): 829-836.

572 Mallin, MA, MH Posey, GC Shank, MR McIver, SH Ensign, TD Alphin. 1991. Hurricane effects
573 on water quality and benthos in the Cape Fear watershed: natural and anthropogenic
574 impacts. *Ecological Applications*. 9(1): 350-362.

575 Mallin, MA, MH Posey, MR McIver, DC Parsons, SH Ensign, TD Alphin. 2002. Impacts and
576 recovery from multiple hurricanes in a piedmont-coastal plain river system. *BioScience*
577 52(11): 999-1010.

578 Mills, K, MJ Kennish, KA Moore. 2008. Research and Monitoring Components of the National
579 Estuarine Research Reserve System. *Journal of Coastal Research*. 10055: 1-8.

580 Munroe, D, A Tabatabai, I Burt, D Bushek, EN Powell, J Wilkin. 2013. Oyster mortality in
581 Delaware Bay: Impacts and recovery from Hurricane Irene and Tropical Storm Lee.
582 *Estuarine, Coast and Shelf Science*. 135: 209-219.

583 Murrell, MC, JM Caffrey, DT Marcovich, MW Beck, BM Jarvis, JD Hagy III. 2018. Seasonal
584 Oxygen Dynamics in a Warm Temperate Estuary: Effects of Hydrologic Variability on
585 Measurements of Primary Production, Respiration, and Net Metabolism. *Estuaries and*
586 *Coasts*. 41(3): 690-707.

587 Nieuwhof, S, PMJ Herman, N Dankers, K Troost, D van der Wal. 2015. Remote Sensing of
588 Epibenthic Shellfish Using Synthetic Aperture Radar Satellite Imagery. *Remote Sensing*.
589 7: 3710 – 3734.

590 Paerl, HW, JL Pickney, JM Fear, BL Peierls. 1998. Ecosystem responses to internal and
591 watershed organic matter loading: consequences for hypoxia in the eutrophying Neuse
592 River Estuary, North Carolina, USA. *Marine Ecology Progress Series*. 166: 17-25.

593 Paerl, HW, JD Bales, LW Ausley, CP Buzzelli, LB Crowder, LA Eby, JM Fear, M Go, BJ
594 Peierls, TL Richardson, JS Ramus. 2001. Ecosystem impacts of three sequential
595 hurricanes (Dennis, Floyd, and Irene) on the United States' largest lagoonal estuary,
596 Pamlico Sound, NC. *Proceedings of the National Academy of Sciences, USA*. 98(10):
597 5655-5660.

598 Paerl, HW, JR Crosswell, B Van Dam, NS Hall, KL Rossignol, CL Osburn, AG Hounshell, RS
599 Sloup, LW Harding Jr. 2018. Two decades of tropical cyclone impacts on North
600 Carolina's estuarine carbon, nutrient and phytoplankton dynamics: implications for
601 biogeochemical cycling and water quality in a stormier world. *Biogeochemistry*. 141:
602 307-332.

603 Parker, ML, WS Arnold, SP Geiger, P Gorman, EH Leone. 2013. Impacts of Freshwater
604 Management Activities on Eastern Oyster (*Crassostrea virginica*) Density and
605 Recruitment: Recovery and Long-Term Stability in Seven Florida Estuaries. *Journal of*
606 *Shellfish Research*. 32(3): 695–708.

607 Patrick CJ, L Yeager, AR Armitage, F Carvallo, VM Congdon, KH Dunton, M Fisher, AK
608 Hardison, JD Hogan, J Hosen, X Hu, B Kiel Reese, S Kinard, JS Kominoski, X Lin, Z
609 Liu, PA Montagna, SC Pennings, L Walker, CA Weaver, M Wetz. 2020. A System Level

610 Analysis of Coastal Ecosystem Responses to Hurricane Impacts. *Estuaries and Coasts*.
611 43: 943-959.

612 Patrick, CJ, JS Kominoski, WH McDowell, B Branoff, D Lagomasino, M Leon, and others.
613 2022. A general pattern of trade-offs between ecosystem resistance and resilience to
614 tropical cyclones. *Science Advances*. 8: doi: 10.1126/sciadv.abl9155.

615 Patricola, CM, MF Wehner. 2018. Anthropogenic influences on major tropical cyclone events.
616 *Nature*. 563: 339-346.

617 Peierls, BJ, RJ Christian, HW Paerl. 2003. Water quality and phytoplankton as indicators of
618 hurricane impacts on a large estuarine ecosystem. *Estuaries*. 26(5): 1329-1343.

619 Perales-Valdivia, H, R Sanay-Gonzales, A Valle-Levinson. 2018. Effects of tides, wind and river
620 discharge on the salt intrusion in a microtidal tropical estuary. *Regional Studies in Marine*
621 *Science*. 24: 400-410.

622 Philips, EJ, S Badylak, NG Nelson, KE Havens. 2020. Hurricanes, El Nino and harmful algal
623 blooms in two sub-tropical Florida estuaries: direct and indirect impacts. *Scientific*
624 *Reports*. 10:1910.

625 Pimm, SL. 1984. The complexity and stability of ecosystems. *Nature*. 307: 321-326.

626 Porter, JH, PC Hanson, C Lin. 2012. Staying afloat in the sensor data deluge. *Trends in Ecology*
627 *and Evolution*. 27(2): 121-129.

628 PRISM Climate Group, Oregon State University, <http://prism.oregonstate.edu>, 2021.

629 Pruitt, JN, AG Little, SJ Majumdar, TW Schoener, DN Fisher. 2019. Call-to-action: a global
630 consortium for tropical cyclone ecology. *Trends in Ecology and Evolution*. 34(7): 588-
631 590.

632

633 Sanger, DM, MD Arendt, Y Chen, EL Wenner, AF Holland, D Edwards and J Caffrey. 2002. A
634 Synthesis of Water Quality Data: National Estuarine Research Reserve System-wide
635 Monitoring Program (1995-2000). National Estuarine Research Reserve Technical
636 Report Series 2002: 3. South Carolina Department of Natural Resources, Marine
637 Resources Division Contribution No. 500. 135 p.

638 Scanes, E, PR Scanes. PM Ross. 2020. Climate change rapidly warms and acidifies Australian
639 estuaries. *Nature Communications*. 11:1803.

640 Schindler, DE, JB Armstrong, TW Reed. 2015. The portfolio concept in ecology and evolution.
641 *Frontiers in Ecology and the Environment*. 13(5): 257-263.

642 Shen, J, T Wang, J Herman, P Mason, GL Arnold. 2008. Hypoxia in a coastal embayment of the
643 Chesapeake Bay: A model diagnostic study of oxygen dynamics. *Estuaries and Coasts*,
644 31(4): 652-663.

645 Sobel, AH, SJ Camargo, TM Hall, C Lee, MK Tippet, AA Wing. 2016. Human influence on
646 tropical cyclone intensity. *Science*. 353(6296): 242-246.

647 Tassone, SJ, PA Bukaveckas. 2019. Seasonal, interannual, and longitudinal patterns in estuarine
648 metabolism derived from diel oxygen data using multiple computational approaches.
649 *Estuaries and Coasts*, 42(4): 1032-1051.

650 Tomasko, DA, C Anastasiou, C Kovach. 2006. Dissolved oxygen dynamics in Charlotte Harbor
651 and its contributing watershed, in response to Hurricanes Charley, Frances, and Jeanne –
652 impacts and recovery. *Estuaries and Coasts* 29(6A): 932-938.

653 Venables, WN, Ripley, BD. 2002. *Modern Applied Statistics with S*. New York: Springer (4th
654 ed).

655 Verdonshot, PFM, BM Spears, CK Feld, S Brucet, H Keizer-Vlek, A Borja, M Elliott, M
656 Kernan, RK Johnson. 2013. A comparative review of recovery processes in rivers, lakes,
657 estuarine and coastal waters. *Hydrobiologia*, 704(1):453-474.

658 Walter, JA, AF Besterman, CD Buelo, SJ Tassone, JW Atkins, ML Pace. 2022. An algorithm for
659 detecting and quantifying disturbance and recovery in high-frequency time series.
660 *Limnology and Oceanography: Methods*. Doi: 10.1002/lom3.10490

661 Wang, Y, M Traber, M Milstead, S Stevens. 2007. Terrestrial and Submerged Aquatic
662 Vegetation Mapping in Fire Island National Seashore Using High Spatial Resolution
663 Remote Sensing Data, *Marine Geodesy*, 30: 77-95.

664 Wetz, MS, HW Paerl. 2008. Estuarine Phytoplankton Responses to Hurricanes and Tropical
665 Storms with Different Characteristics (Trajectory, Rainfall, Winds). *Estuaries and Coasts*.
666 31: 419-429. Wetz, MS, DW Yoskowitz. 2013. An ‘extreme’ future for estuaries? Effects
667 of extreme climate events on estuary water quality and ecology. *Marine Pollution*
668 *Bulletin*. 69: 7-18.

Table 1. Tropical cyclone and station characteristics (calculated from all available years) of disturbance events lasting longer than 30 days.

NERRS Reserve	Station	Storm	Var.	Duration (days)	Precip (mm)	Depth (m)	Mean Salinity (psu)
Ashepoo Combahee Edisto Basin	Mosquito Creek	Andrea-2013	Sal	98.4	72.4	4	18.2
Chesapeake Bay Virginia	Taskinas Creek	Michael-2018	Sal	71.1	91.5	1.6	10.6
Guana Tolomato Mantanzas	San Sebastian	Irma-2017	Sal	58.9	266.8	5.1	33.8
Weeks Bay	Middle Bay	Ida-2009	DO	58.6	89.2	1.4	9.5
NorthInlet-Winyah Bay	Thousand Acre	Florence-2018	Sal	58.4	316.6	2.5	8
Weeks Bay	Weeks Bay	Lee-2011	DO	51.5	230.6	1.2	10
NorthInlet-Winyah Bay	Clambank	Florence-2018	Sal	50	316.6	2	32.7
Ashepoo Combahee Edisto Basin	Fishing Creek	Andrea-2013	Sal	46.9	72.4	2.7	9.4
Delaware	Blackbird Landing	Irene-2011	Sal	43.3	212.4	1.7	1.9
Weeks Bay	Magnolia River	Bill-2003	Sal	42.6	262.2	1.8	9.2
Weeks Bay	Middle Bay	Bill-2003	Sal	40.6	262.2	1.4	9.5
Great Bay	Lamprey River	Hanna-2008	Sal	39.4	146.8	2.2	12.1
Rookery Bay	Middle Blackwater River	Irma-2017	DO	38.8	255.5	1	30.6
NorthInlet-Winyah Bay	Debidue Creek	Florence-2018	Sal	38.7	316.6	2.5	32.1
North Carolina	Zeke's Basin	Florence-2018	Sal	37.7	680.2	0.7	22
Great Bay	Lamprey River	Charley-2004	DO	35.3	131	2.2	12.1
Guana Tolomato Mantanzas	Pine Island	Irma-2017	Sal	33.5	266.8	4.2	28.5

Sapelo Island	Cabretta Creek	Tammy-2005	Sal	32.5	307.2	3.1	31.8
Chesapeake Bay Virginia	Claybank	Isabel-2003	Sal	31.8	117.3	1.2	16.1
Chesapeake Bay Virginia	Goodwin Islands	Irene-2011	Sal	31.2	258.6	1	19.6
Delaware	Scotton Landing	Irene-2011	Sal	30.7	212.4	1.7	10.8
Jacques Cousteau	Chestnut Neck	Irene-2011	Sal	30.4	157.7	2.4	14.9
Chesapeake Bay Virginia	Taskinas Creek	Isabel-2003	Sal	30.3	117.3	1.6	10.6
Apalachicola Bay	Dry Bar	Dennis-2005	Sal	30.1	5.7	1.8	21.8

Table 2. Tropical cyclone (TC) and station variables related disturbance characteristics: occurrence/non-occurrence, timing (days after TC passage), length (days), and peak severity (z-score). TC characteristics are highlighted in blue; site/station characteristics are highlighted in green. Drivers are separated into positive and negative significant relationships with each disturbance event characteristic. * denotes $0.05 \leq p \leq 0.1$; other variable p-values are < 0.05 . See tables S2 – S9 in the Supplemental Information for coefficient values and standard errors.

Disturbance Characteristic	Salinity	DO % saturation
Occurrence	Positive: precipitation, mean(salinity) Negative: tidal range	Positive: precipitation, depth, TC distance* Negative: tidal range
Timing relative to TC	Positive: wind gust duration* Negative: precipitation	Positive: mean(salinity) Negative: precipitation, TC distance
Length	Positive: tidal range, depth, wind gust duration, wind gust max Negative: mean(salinity), storm surge height	Positive: precipitation, sd(salinity)* Negative:
Peak Severity	Positive: mean(salinity), precipitation Negative: sd(salinity)	Positive: storm surge height Negative:

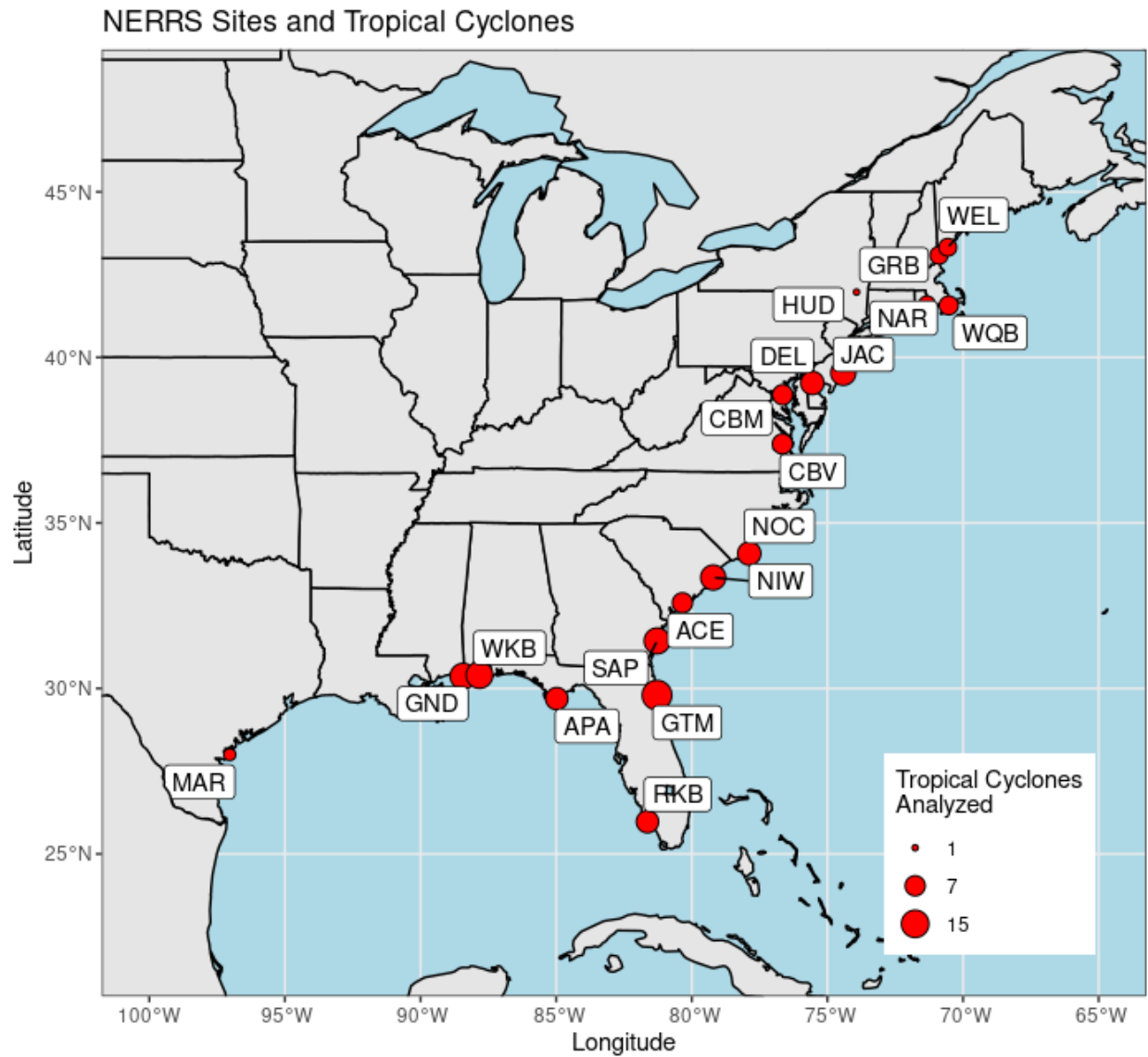


Figure 1. Location and site codes of NERRS sites analyzed for tropical cyclone impacts. Point size indicates the number of tropical cyclones analyzed.

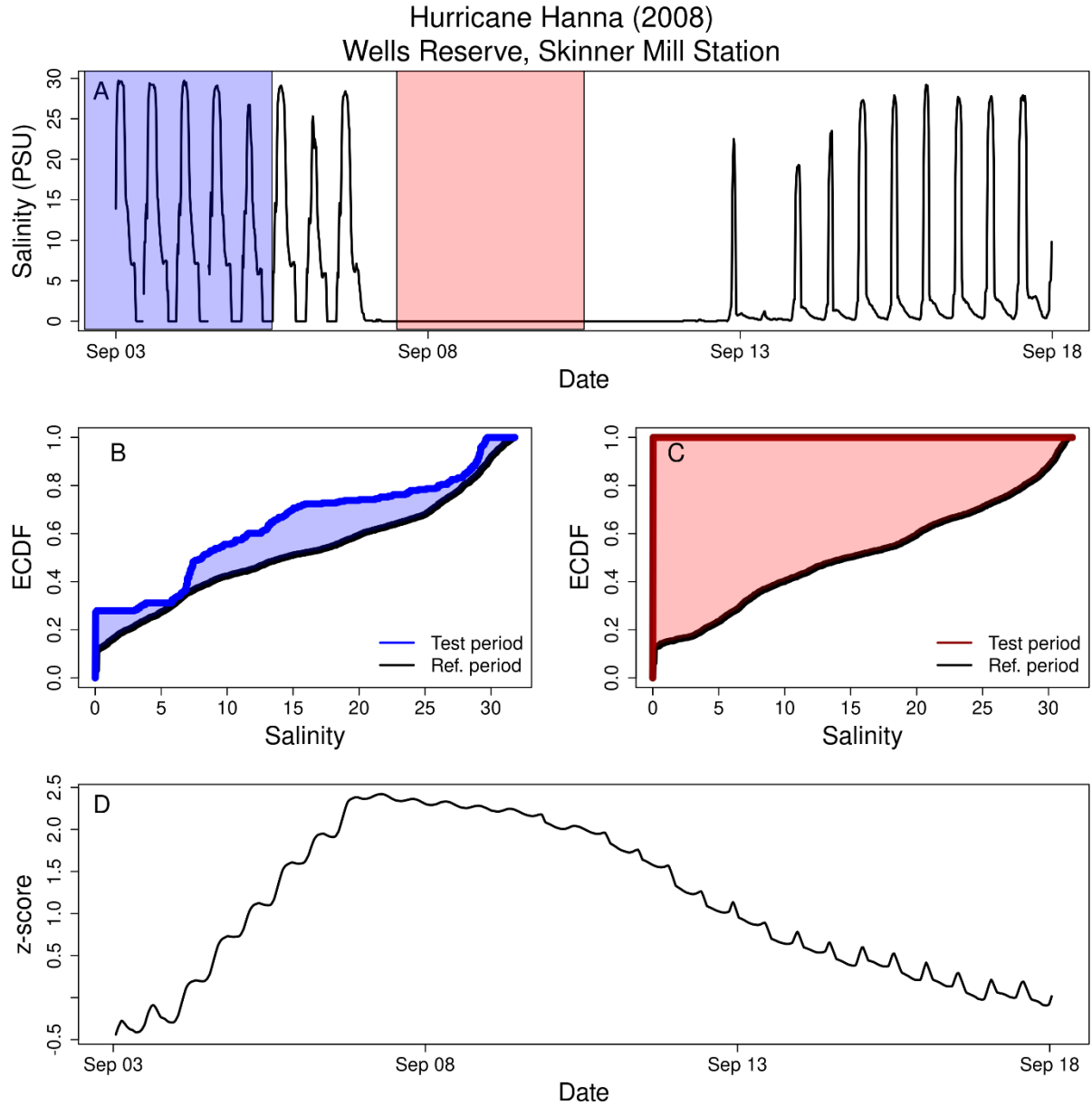


Figure 2. Illustration of the disturbance detection algorithm. A) Observations of salinity before and after Hurricane Hanna impacted the Wells Reserve (Maine, USA) in 2008. Shaded periods are example three-day test periods before (blue) and after (red) tropical cyclone passage. B) Empirical cumulative distributions (ECDF) for the before-impact example test period (blue) and corresponding reference periods in other years (black). The area of the shaded region corresponds to d_w , the distribution difference statistic. C) ECDFs for the after-impact example test period (red) and corresponding reference periods in other years (black). D) Time series of the normalized (z-score) distribution difference test statistic.

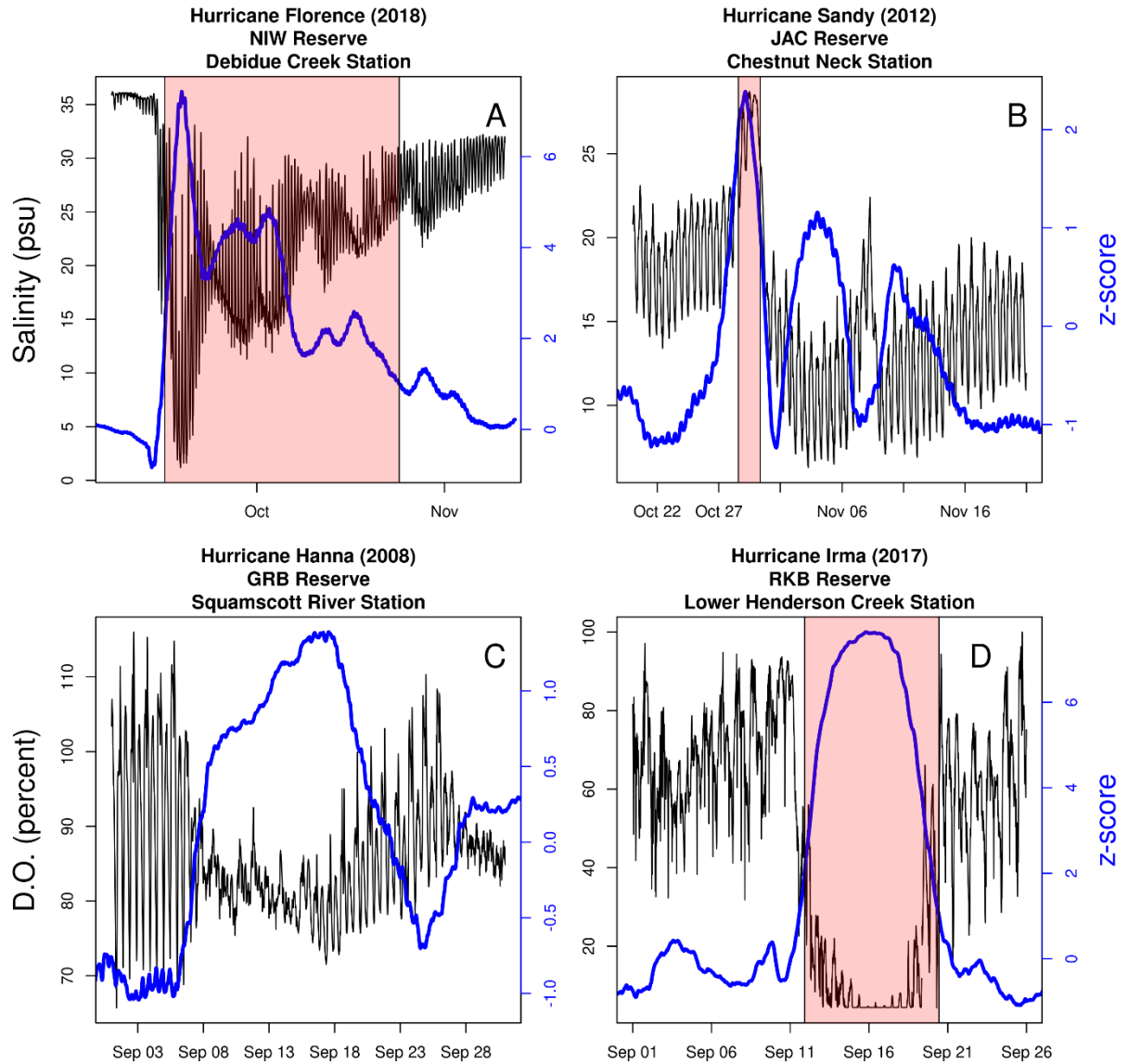


Figure 3. Example time series showing hurricane impacts and performance of the normalized distribution difference statistic. Black lines are observations of salinity (A, B) and dissolved oxygen percent saturation (C, D). Blue lines are the normalized distribution difference statistic for three-day wide rolling windows. Red shaded areas represent disturbances identified with z-score a disturbance threshold of 2 and recovery threshold of 1. Note different x and y axis scales.

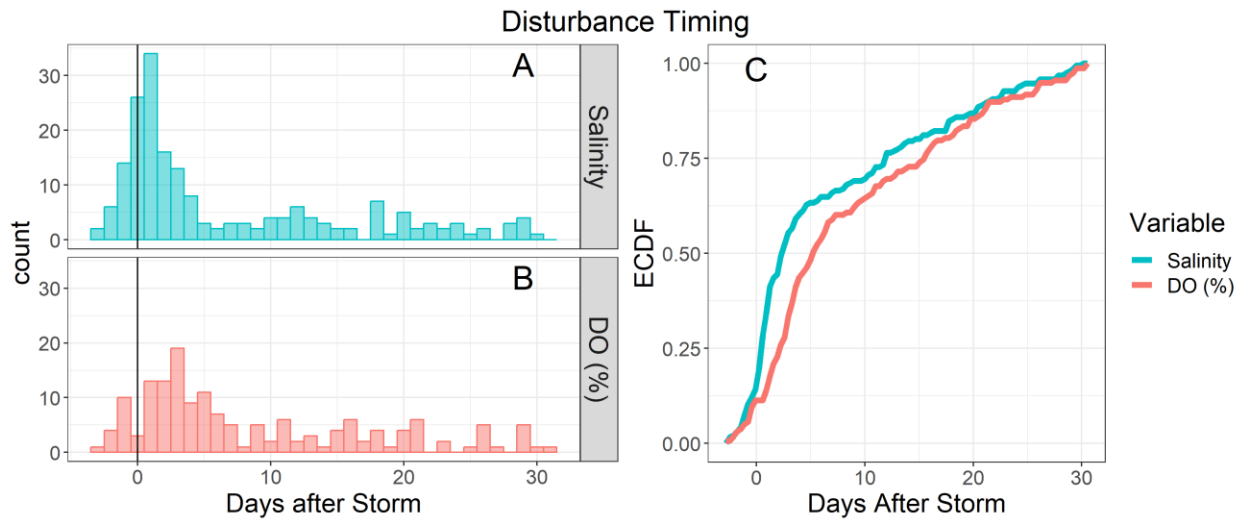


Figure 4. Histograms (A: dissolved oxygen, B: salinity) and empirical cumulative distribution curves (C) showing the timing of disturbance start relative to when the tropical cyclone was closest to each NERRS site.

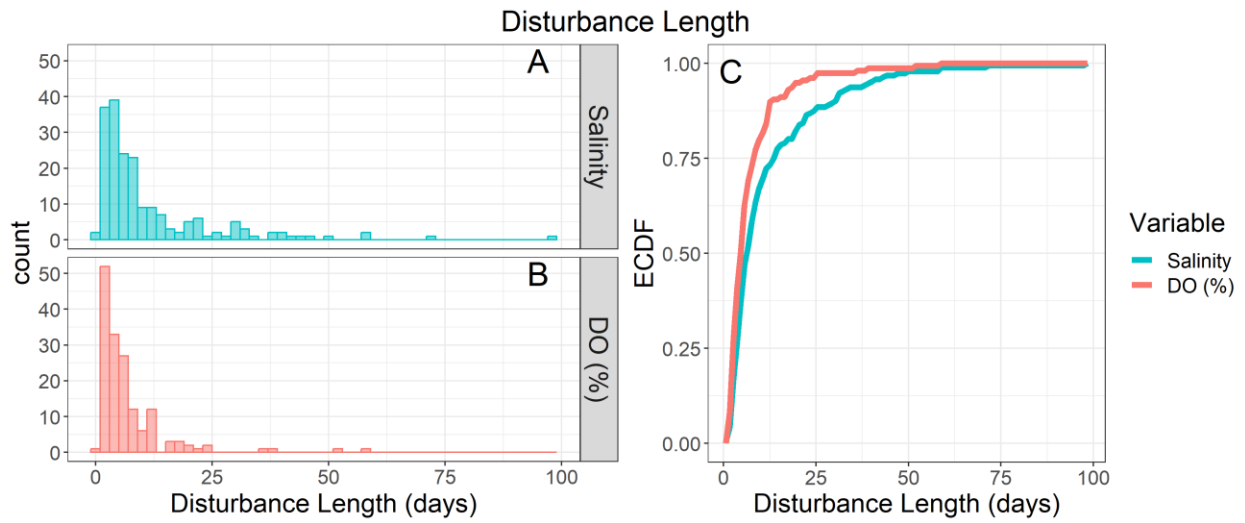


Figure 5. Histograms (A: dissolved oxygen, B: salinity) and empirical cumulative distribution curves (C) showing the length of time between disturbance initiation and recovery.

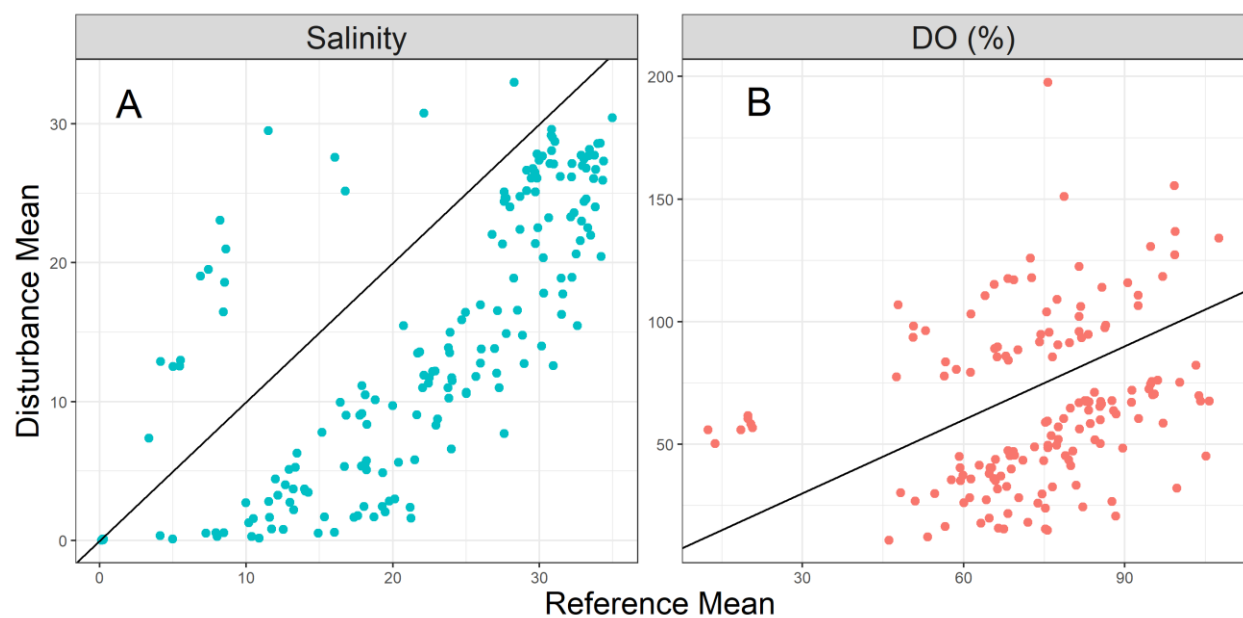


Figure 6. Mean salinity (A) and dissolved oxygen (B) during TC-associated disturbances vs. during the same periods in all other years. The black line is the 1:1 line, indicating no difference between the means during the disturbance and reference periods.

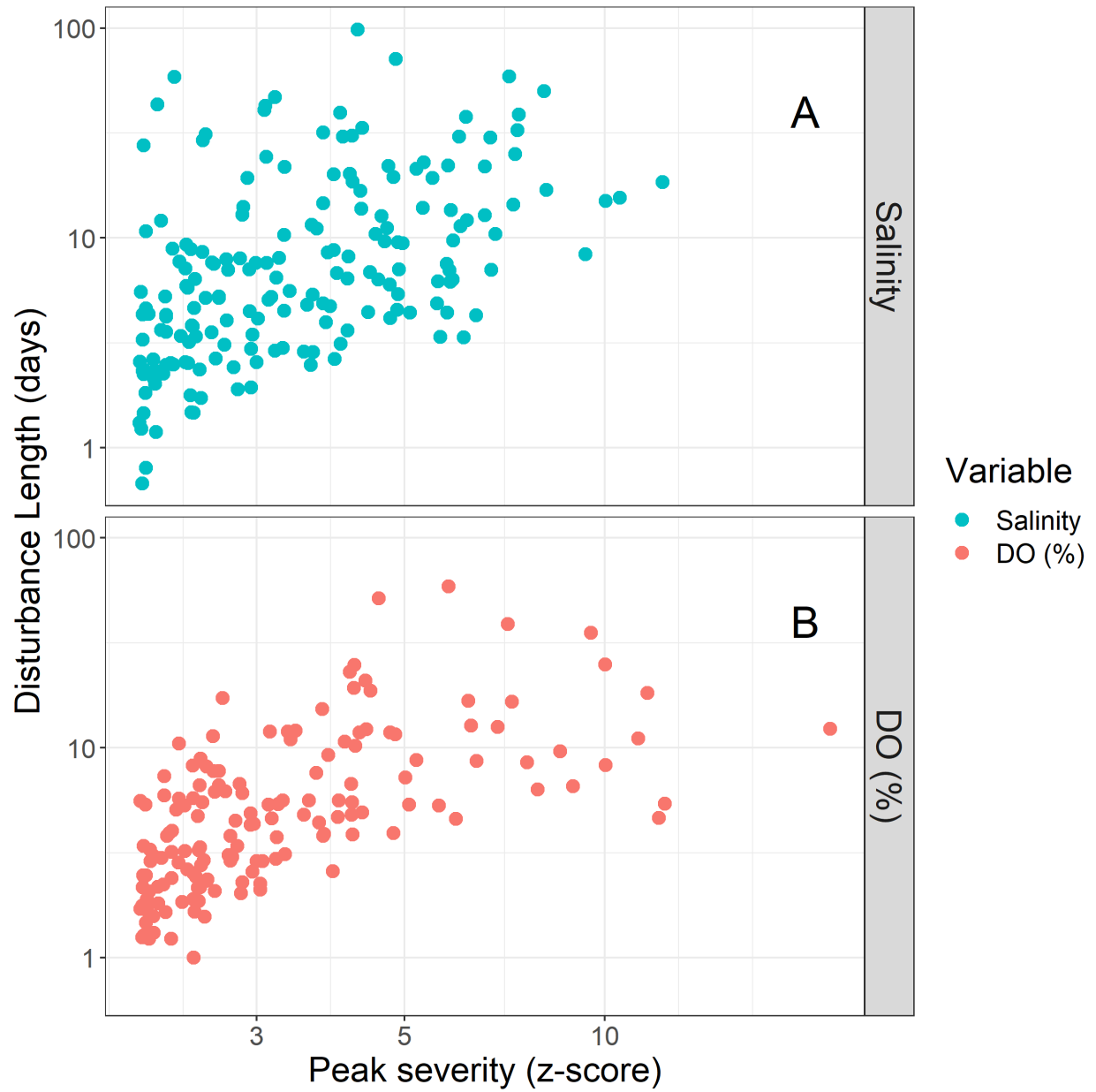


Figure 7. Relationship between disturbance length and peak disturbance severity for dissolved oxygen (A) and salinity (B).