

Persistence and intervening states of warm low oxygen episodes at eight southeastern United States estuarine stations

Alberto Franco Cerna Cueva^{1*}, José Kalión Guerra Lú¹,
Casiano Aguirre Escalante¹, Luis Eduardo Oré Cierito¹,
Oscar Alexander Rosales Pachamango¹, Leiwer Flores Flores²

¹ Facultad de Recursos Naturales Renovables, Universidad Nacional Agraria de la Selva, Carretera Central km 1.21, Tingo María 10131, Huánuco, Peru

² Facultad de Ciencias Agrarias, Universidad Nacional de Cajamarca – Filial Jaén, Jaén, Peru

* Corresponding author's e-mail: alberto.cerna@unas.edu.pe

ABSTRACT

This study describes the temporal organization of simultaneous warmth and low oxygen in eight fixed near-bottom estuarine station records. Original 15-minute temperature–oxygen observations from May–September 2018–2023 supplied 35 eligible station-seasons and 510,454 valid pairs (99.29% of scheduled observations within those seasons). Warmth exceeded a fixed station-specific P90 of valid annual temperatures during 2018–2023; low oxygen was below 2 mg/L. Four observed states distinguished neither condition, warmth only, low oxygen only, and their overlap. The records contained 417 compound episodes and 412 nominal occupancy hours. Episode count and the fraction of occupancy assigned to sustained runs described different aspects of persistence. Among 384 fully observed interepisode intervals, 207 lacked a neither-state reading. A post hoc partition identified 183 entirely warm-only intervals, 22 entirely low-oxygen-only intervals, and two containing both single-condition states. These intervals represented 53.9% of the count but only 7.1% of total interval hours. Their count fraction fell to 40.9% under stricter coverage and 43.7% under conservative deployment screening. An exploratory two-common-year station comparison was sensitive to P90 versus P95 and does not establish a persistent station difference. The contribution is a descriptive four-state characterization of monitored sequences, separating episode counts, occupancy-weighted persistence, and intervening conditions. It does not identify ecological mechanisms or organismal responses.

Keywords: estuarine monitoring, dissolved oxygen, water temperature, compound episodes, persistence, interepisode states.

INTRODUCTION

Environmental records with the same total time in low oxygen need not have the same temporal organization. That time may occur in sustained runs or many brief excursions. When high temperature coincides with low oxygen, the interval after their overlap ends may still contain either warm water or low oxygen. Describing these sequences is relevant to interpreting monitoring records because a compound-event count alone

gives neither the duration of its constituent runs nor the conditions separating them.

Subdaily variation in estuarine oxygen is well established. Fixed-site networks have characterized hypoxic-event duration, intermittency, and interannual variability (Codiga et al., 2009). Observations and process analyses in subtropical estuaries show that diel oxygen patterns can reflect both biological production–respiration cycles and physical transport or mixing (Duvall et al., 2022). In Weeks Bay, continuous

measurements and modeling resolved low oxygen lasting from minutes to hours and substantial vertical gradients close to the bed (Jarvis et al., 2023). A recent National Estuarine Research Reserve System (NERRS) assessment examined long-term oxygen patterns and derived monthly low-oxygen frequencies from 15-minute observations (Reinl et al., 2026). Temporal variation, observation frequency, and sensor position therefore have established roles in interpreting estuarine oxygen records.

Temperature–oxygen co-occurrence also has direct precedents. Tassone et al. (2022) analyzed estuarine heatwaves and coincident low oxygen and low pH in NERRS records, using filtered daily observations and seasonally varying percentile thresholds with multiday persistence criteria. Lucey et al. (2022) quantified the timing and duration of compound temperature–oxygen conditions on tropical reefs alongside physiological experiments. At a different spatial scale, Wong et al. (2024) characterized temperature, oxygen, and acidity extremes that extended through substantial parts of the ocean water column. Freeman et al. (2026) likewise quantified frequency, intensity, and duration of modeled physical–biogeochemical extremes in the California Current. In NERRS records, Nardi et al. (2026) characterized warm temperature anomalies before and after marine heatwaves, demonstrating that thermal exposure extends beyond conventionally bounded events. The present analysis provides a geographically bounded methodological and descriptive extension: four-state sequences at the original 15-minute scale, run counts alongside occupancy-weighted persistence, and the composition of intervals between compound runs. Temperature–hypoxia co-occurrence itself is already established in NERRS records. It does not redefine marine heatwaves or infer the biological consequences investigated experimentally in other systems.

A useful distinction is between three quantities: the fraction of observed low-oxygen time that is also locally warm; the persistence of the resulting compound runs; and the composition of the intervals separating those runs. Here, “compound” specifically denotes simultaneous warmth and low oxygen; it differs from the thermal-event terminology used for overlapping marine heatwaves and adjacent warm phases by Nardi et al. (2026). The first quantity measures overlap, the second indicates whether occupancy

is concentrated in brief or sustained sequences, and the third establishes which measured condition persists after overlap ceases. None can be inferred from another. In particular, a large conditional overlap can arise from very little low-oxygen time, and many isolated compound readings can contribute little to total occupancy.

The North Inlet–Winyah Bay (NIW), South Carolina, and Guana Tolomato Matanzas (GTM), Florida, reserves provide official monitoring observations of co-located temperature and dissolved oxygen with accompanying deployment documentation. Their stations occupy different tidal creeks and estuarine settings, and the records allow comparisons among monitored locations without treating a sensor as representative of an entire estuary. Long-term observations from adjacent North Inlet and Winyah Bay show contrasting water-quality trajectories between ocean- and river-dominated settings (Dunn et al., 2023). Water-quality observations elsewhere in the GTM region have also emphasized how location relative to freshwater and marine influences complicates simple site classifications (Chaya et al., 2023).

The objective was to describe how locally warm, low-oxygen conditions were organized across eligible May–September seasons during 2018–2023. The analysis quantified station-season occupancy, run-count and occupancy-weighted persistence, and the thermal–oxygen states between consecutive compound episodes. Recurrence within fully observed 24-h and 72-h horizons was a secondary descriptor. Coverage, threshold, and deployment sensitivities were used to delimit the findings rather than to select a preferred site ranking. The target is the water observed at the eight sensors, not a regional estimate of estuarine habitat exposure. Station contrasts are exploratory descriptions of the observed support; no persistent station difference was specified as a confirmatory hypothesis.

MATERIALS AND METHODS

Monitoring records and study stations

Original annual comma-separated files were obtained from the NERRS System-wide Monitoring Program (SWMP), through the Centralized Data Management Office (CDMO; NOAA NERRS, 2019), on 30 September 2026. The

dataset DOI is 10.25921/vw8a-8031. The primary source comprises 48 station-year files of original instrument observations for eight stations during 2018–2023. Annual metadata reports were used to interpret deployment and quality-control history; environmental outcomes were calculated from the instrument CSV records rather than extracted from annual-report summaries. Temperature in °C and dissolved oxygen concentration in mg/L were paired only at the same station and timestamp. The source interval is 15 min. Source times were interpreted as fixed Eastern Standard Time (EST; UTC–05:00), without daylight-saving adjustment or a second correction of clocks already corrected by the provider.

Station selection preceded environmental outcome estimation and used official primary status, operational history, availability of paired variables, metadata, and coverage. Feasibility screening encompassed 12 stations in NIW, GTM, and Sapelo Island during 2017–2024. Sapelo Island was excluded because of insufficient within-month coverage and heterogeneous continuity. The primary analysis excludes 2024 because of provisional or mixed authentication status and begins in 2018 after the December 2017 Clambank Landing instrument transition. No station or season was removed because it had few or no compound detections.

Clambank Landing, Debidue Creek, and Oyster Landing are in North Inlet; Thousand Acre is in Winyah Bay. Their sondes are fixed to pilings, nominally 0.30 m above the bed. GTM includes Fort Matanzas and Pellicer Creek in the Matanzas system, Pine Island on the Tolomato River,

and San Sebastian at the San Sebastian–Matanzas confluence. GTM instruments are fixed near the bed, with nominal positions within 1 m of it. Water level changes the distance below the surface. These are point measurements rather than vertical profiles or integrated habitat-volume measurements. Table 1 identifies the stations and eligible years; supplementary station information records coordinates and installation qualifications.

Annual metadata identify YSI EXO2 sondes and optical dissolved oxygen sensors model 599100-01. Temperature/conductivity hardware and antifouling arrangements changed within the record. GTM deployment tables show alternation of temperature/conductivity models before the later common wiped-conductivity/temperature configuration; NIW documents contain inconsistent earlier specifications. Consequently, uniform temperature precision across all years was not assumed. The later 599827 specifications report ± 0.2 °C, and optical oxygen specifications in the relevant concentration range give the larger of ± 0.1 mg/L or $\pm 1\%$ of the reading. Small threshold perturbations used these specifications as classification checks, not as a full measurement-error model. A 2021–2023 sensitivity addressed the later configuration.

Quality screening and observation support

Only authenticated records with historical = 1 and numerical parameter flag 0 were retained for the relevant variable. The numeric flag was parsed separately from appended comments. Environmental comments such as CDA, indicating

Table 1. Study stations and primary eligible May–September seasons

Code	Station	Waterbody	Eligible years	P90 (°C)
NIW-CB	Clambank Landing	North Inlet	2018, 2020, 2021, 2022	29.1
NIW-DC	Debidue Creek	North Inlet	2020, 2021, 2022, 2023	29.5
NIW-OL	Oyster Landing	North Inlet	2019, 2020, 2022, 2023	29.5
NIW-TA	Thousand Acre	Winyah Bay	2019, 2020, 2021, 2022, 2023	29.1
GTM-FM	Fort Matanzas	Matanzas River estuary	2018, 2020, 2021, 2022, 2023	29.5
GTM-PC	Pellicer Creek	Pellicer Creek	2018, 2019, 2020, 2021, 2023	30.7
GTM-PI	Pine Island	Tolomato River	2018, 2020, 2022, 2023	30.5
GTM-SS	San Sebastian	Confluence of San Sebastian and Matanzas Rivers	2018, 2020, 2021, 2022	29.7

Note: NIW: North Inlet–Winyah Bay reserve, South Carolina; GTM: Guana Tolomato Matanzas reserve, Florida. Source identifiers are the lowercase reserve code plus station code and wq, for example niwolwq. P90 uses annual temperature-valid data from 2018–2023, independently of paired oxygen availability. Sensor geometry and coordinate qualifications are given in Table S1.

low oxygen, and CAB, indicating algal blooms, were retained. Missing, nonfinite, rejected, or suspect values were excluded, as were physically impossible negative oxygen values; valid oxygen values of zero remained admissible. No temperature or dissolved oxygen flags 4 or 5 occurred in the eight-station 2018–2023 source cohort. The treatment follows the source quality-control semantics, supplemented by documented deployment information (NERRS CDMO, n.d.).

An 87-row source-derived exclusion ledger was fixed before the first environmental estimate. It distinguishes primary exclusions from optional conservative deployment exclusions and applies masks to the affected parameter or both parameters as appropriate. Documented displacement, exposure, and explicit sensor failures can invalidate otherwise accepted numerical flags. An isolated depth/level calibration problem does not invalidate temperature or oxygen. Uncertain failure times were handled as explicitly conservative exclusion intervals, not as reconstructed measurements. The complete ledger and source references are retained with the reproducibility files.

The analytical window was 1 May 00:00 to 30 September 23:45 in each year. A station-season qualified if at least 90% of expected grid positions had jointly valid observations and every constituent month reached at least 80%. Primary station inclusion required at least three eligible seasons. A stricter within-cohort analysis required 95% seasonal and 90% monthly coverage. These pragmatic coverage rules restrict support; they do not establish that missing observations are environmentally random. Data were never interpolated or carried across gaps. All 48 candidate seasons, their monthly and seasonal coverage, and reasons for exclusion are reported in the supplement and associated tables.

Thermal reference and four observed states

For each station, a fixed empirical 90th percentile (P90) was calculated from all valid temperature readings in the full annual 2018–2023 reference window using linear interpolation between order statistics. Valid temperature did not require contemporaneously valid oxygen. The temperature-specific ledger masks were applied before threshold calculation. Each calendar month, pooled across the six reference years, required at least 75% temperature coverage. The

threshold is retrospective and local: the same observations may contribute to both the reference distribution and the descriptive analysis. It is neither a long-term climatology nor a thermal-tolerance or marine-heatwave threshold.

At a jointly valid timestamp, the warm condition W was temperature strictly above the station P90, and low oxygen L was dissolved oxygen strictly below 2 mg/L. Their intersection C was compound. Four mutually exclusive states were assigned: neither, warm only, low oxygen only, and compound. Temperature equal to P90 does not satisfy W; oxygen equal to 2 mg/L does not satisfy L. The neither state means only that neither operational threshold is exceeded. No biological safety interpretation is assigned to it.

The principal occupancy estimand was the compound fraction of low-oxygen readings, $n(C)/n(L)$, by station-season. It remains undefined when $n(L)$ is zero. Companion quantities were $n(C)/n(V)$, $n(L)/n(V)$, and $n(W)/n(V)$, where V denotes jointly valid readings. Nominal occupancy hours equal the classified reading count multiplied by 0.25 h. This conversion represents occupancy on a regular observation grid and is not a continuous measurement of exposure between samples. Pooled station estimates sum numerators and denominators across eligible seasons; they are not means across estuaries. Table 2 provides absolute denominators and all four pooled ratios.

Compound runs and intervals between them

A compound episode was a maximal run of consecutive 15-min compound readings. Any valid non-compound reading or missing grid position terminates a run. The first-to-last observed span is $(N - 1) \times 0.25$ h for N compound readings, whereas nominal occupancy is $N \times 0.25$ h. A singleton is therefore a retained detection with zero span and 0.25 nominal occupancy hours. Actual threshold-crossing times are bounded by adjacent samples and remain unresolved within those brackets. Each run was classified as complete only if its immediately preceding and following grid positions contained valid non-compound readings; otherwise the affected boundary was recorded as gap- or window-truncated.

Persistence was summarized both by the distribution of complete-run spans and by the proportion of compound occupancy associated

with runs longer than specified spans. The latter weights runs by their number of compound readings. Empirical cumulative distributions describe observed complete runs and are not survival estimates. No independent-censoring assumption or biological exposure-response model was imposed.

For adjacent runs at a station, the exit anchor was the first valid non-compound reading after the earlier run and the next-entry anchor was the first compound reading of the next run. An interval was eligible only when all positions between those anchors were valid and within one analytical window. Composition includes the exit anchor and excludes the next-entry reading. Counts of warm-only, low-only, and neither readings, multiplied by 0.25 h, sum to the elapsed time between anchors. The longest uninterrupted neither-state span was recorded separately. Intervals crossing gaps or separate seasonal windows were unavailable, not assigned a benign state.

Two different summaries were retained: the fraction of eligible intervals containing no neither-state reading, and the fractions of total interval hours in each state. The first weights intervals equally; the second weights by time and was displayed in ≤ 6 , >6 –24, >24 –72, and >72 h bins. Reserve-pooled displays summarize the sampled intervals only, with station-specific counterparts in the supplement. Secondary recurrence was the fraction of eligible exits followed by a new compound entry within 24 or 72 h. Every timestamp from the exit anchor through the complete horizon had to be valid within the same window, even when an early re-entry was observed.

Prespecified sensitivity analyses and reproducibility

The analysis plan, configuration, source-file hashes, code, and ledger were frozen on 30 September 2026 at 22:42:30 UTC; the first environmental estimate was recorded at 22:42:51 UTC. This was an internal pre-estimation specification, not a public preregistration. The primary analysis is descriptive and has no primary hypothesis test. Readings, runs, and interepisode intervals are temporally dependent and were not treated as independent ecosystem replicates. This timing establishes the order of the internal specification and estimation; it does not show that hypotheses were formulated without prior exploratory knowledge of the source records.

Prespecified alternatives included P85 and P95 thermal thresholds; year-specific leave-one-year-out reference P90 values applied only to the omitted year; thermal offsets of ± 0.2 °C; oxygen thresholds of 1.9 and 2.1 mg/L; broader low-oxygen definitions of 3 and 4 mg/L; stricter coverage; conservative removal of deployments with possible wiper/fouling effects; and the 2021–2023 subset. Conservative exclusions triggered renewed checks of the reference monthly coverage and the minimum three eligible seasons. All authenticated seasons were also summarized as a secondary observed-support comparison. Minimum run spans of 0.25 and 0.5 h were used to report retention of detections and occupancy, without replacing the unfiltered primary runs.

Station-season estimates and their ranges provide the main representation of annual variation. Exploratory percentile intervals used 2,000 resamples of the six calendar years, synchronized across stations and preserving the eligibility matrix. They condition on the fixed empirical thermal thresholds, retain undefined denominators as missing, and do not represent regional sampling uncertainty or total measurement uncertainty. The supplement reports valid and undefined bootstrap draws. A secondary daily comparison required $\geq 90\%$ daily joint coverage and representation in all four 6-h blocks; it distinguished compound daily means, same-day extrema, and any simultaneously compound reading.

The original deterministic analysis used Python 3.12.14, NumPy 2.3.5, and pandas 2.2.3. Thirty-two pipeline tests cover artificial runs, gaps, ties, zero oxygen, comments, boundaries, and recurrence censoring. They test software behavior on constructed inputs and do not establish ecological validity. A separate implementation re-reads all 48 annual source files and the same exclusion ledger, recalculates percentiles, and detects episodes with a sequential algorithm. Agreement of its 49 comparison groups is a computational consistency check within the same AI-assisted workflow, rather than independent scientific validation. For this revision, the archive includes the original CSVs, source-valued state grids, annual metadata, four recovered historical source-audit examples, and seven additional post hoc R1 examples. Original values, flags, exclusions, classifications, neighboring readings, selection rules and distinct audit dates are documented. Complete reconstruction instructions and execution reports accompany the source snapshot.

Post hoc comparisons of temporal support

After the primary results were available, a separate protocol dated 1 October 2026 specified additional descriptive checks of common observation years and the influence of individual outcome years. A second post hoc elaboration classified eligible interepisode intervals as entirely warm only, entirely low oxygen only, containing both single-condition states without neither, or containing at least one neither reading. These post hoc analyses retain the original seasonal eligibility, accepted readings, fixed P90 thresholds, episode boundaries, and interval definitions. They are separate from the pre-estimation specification of 30 September 2026. The common-year station contrast and the four-way interval partition therefore serve as exploratory elaborations of the primary descriptors, not confirmatory tests.

For Oyster Landing, Clambank Landing, and Thousand Acre, eligible calendar years were intersected for each station pair and for the three-station comparison. Pooling then used only those common years, with annual counterparts retained. These restricted comparisons do not impose a new three-season station-selection rule. Compound occupancy and occupancy in runs spanning more than 3 h were also expressed per 1,000 valid observed hours. A separate diagnostic restricted occupancy to timestamps jointly valid at both stations within the common years. Original run and interval profiles were kept distinct from this occupancy diagnostic. A later post hoc check redetected runs on the exact shared-valid grid, treating any position missing at either station as a gap and reassessing complete or truncated bounds; it retained fragment-to-original-run correspondences. These matched-grid fragments were not assumed to recover unobserved durations.

One outcome year at a time was removed from each station, the network pool, and each common-year pair, while all thermal thresholds remained fixed. This is a year-influence analysis, distinct from the prespecified leave-one-reference-year-out calculation that re-estimates P90 and applies it to the omitted year. Every deletion was retained, including single-year residual comparisons and undefined denominators. No intervals were joined across deleted years. The resulting ranges describe influence of the observed years and are not confidence intervals. No hypothesis tests or bootstrap comparisons were applied to these small common-year sets. A

further post hoc check crossed the common 2020 and 2022 support with the already calculated P85, P95, and oxygen <1.9 and <2.1 mg/L alternatives, retaining both pooled and annual results without recalculating the reference thresholds.

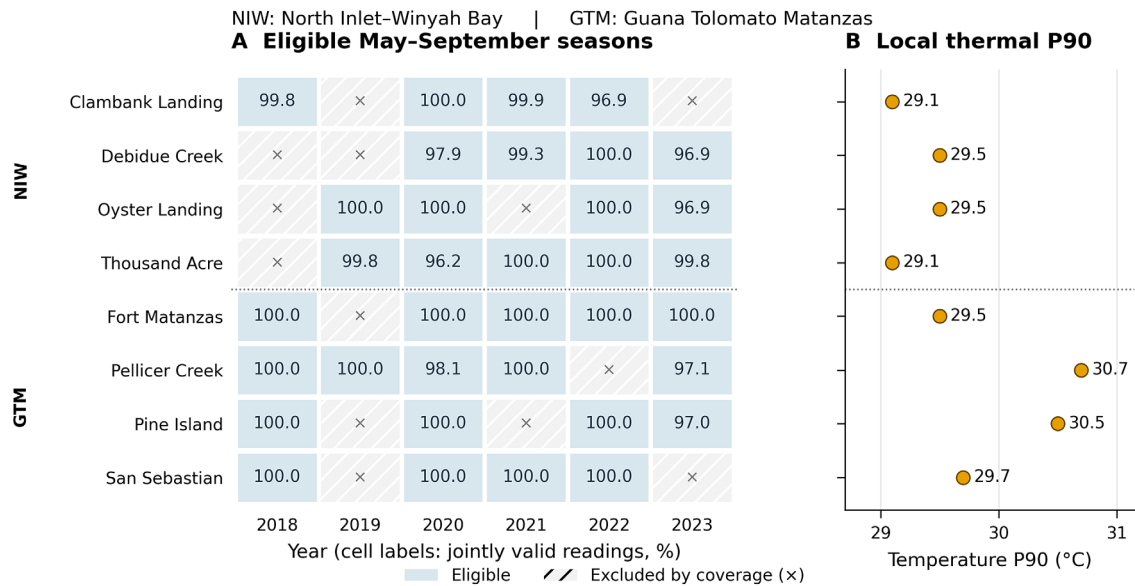
RESULTS

Eligible support and occupancy

Thirty-five station-seasons met the primary coverage rules, with four or five seasons at every station (Figure 1). They contained 510,454 jointly valid readings out of 514,080 scheduled positions, equivalent to 127,613.5 nominal observed hours and 99.2947% coverage. The minimum eligible seasonal coverage was 96.21%, and the lowest individual-month coverage within an eligible season was 83.13%. Eligible seasons still contained 114 missing runs, totaling 3,626 scheduled readings, with the longest missing run containing 125.25 scheduled hours of absent observations. Twenty-eight seasons met the stricter rules. No station qualified in every year; 2020 was the only year in which all eight stations qualified. The pooled-month temperature reference coverage exceeded 83.3% at every station, and fixed P90 values ranged from 29.1 to 30.7 °C.

Compound conditions occupied 1,648 readings or 412 nominal hours. Oyster Landing contributed 199 compound hours within 811.75 low-oxygen hours, so 24.5% of its observed low-oxygen occupancy coincided with locally warm conditions. Clambank Landing had 80.25 compound hours within 214.5 low-oxygen hours (37.4%), and Thousand Acre had 52.25 within 286.5 h (18.2%). Thus, Oyster Landing combined more absolute compound occupancy with a smaller conditional fraction than Clambank Landing (Figure S5). Compound occupancy represented 1.37%, 0.55%, and 0.29% of valid observations at those three stations, respectively.

GTM records had smaller absolute low-oxygen totals. Pellicer Creek had 23 compound hours among 39 low-oxygen hours (59.0%). Pine Island had 53.5 among 74.5 h (71.8%), but 49.5 of its compound hours occurred in 2023 and 4 h in 2022; its eligible 2018 and 2020 seasons had none. Fort Matanzas had 3.25 low-oxygen hours without compound detections. San Sebastian had only three low-oxygen readings, of which one was compound; its 33.3% conditional fraction



35 eligible station-seasons; 510,454 valid paired readings. P90 uses valid annual temperatures, 2018–2023. Fixed near-bottom sensors; coverage does not establish whole-water-column or whole-estuary conditions.

Figure 1. Observation support and local thermal thresholds

(A) Primary eligibility of each candidate station-season after authenticated-data, parameter-flag, manual-ledger, and coverage screening. An eligible season requires at least 90% joint valid readings over May–September and at least 80% in every constituent month. Values show joint coverage as a percentage of expected 15-min readings; crossed hatched cells denote excluded seasons. Rounding to 100.0% does not assert zero missing readings.

(B) Fixed empirical station-specific annual-reference P90 temperature, estimated independently of contemporaneous oxygen availability. The thermal threshold is not a thermal tolerance or a long-term climatology. The first four stations are NIW and the last four GTM. This order is retained rather than ranking stations by outcomes.

therefore represents 0.25 compound hours. Debidue Creek similarly had limited support for an event profile, with 3.75 compound hours. Full state compositions and annual occupancy are shown in Figures S1 and S4.

Annual conditional fractions varied within stations. Oyster Landing ranged from 9.0% to 36.7% and Clambank Landing from 21.4% to 52.3%. Thousand Acre ranged from 0% to 24.0%. These comparisons retain the differing eligible-year sets in Table 1. Seasons with no low oxygen have an undefined conditional fraction, including three of five Fort Matanzas seasons; they remain present in the occupancy tables.

Exploratory common-year comparison and year influence

The post hoc Oyster Landing–Clambank Landing comparison retained their two common eligible seasons, 2020 and 2022. Oyster Landing had 59.5 compound hours in 7,341.75 valid hours and Clambank Landing had 53.75 in 7,229.25 h, equivalent to 8.10 and 7.44 compound hours per

1,000 observed hours (Figure 2). Thus, the large difference in absolute occupancy in the unequal-year pools was much smaller on common years and an observation-time denominator. These values describe the observed records and do not establish statistical equivalence.

Across the full Oyster Landing cohort, deleting one outcome year at a time left 41.9–53.6% of compound occupancy in runs spanning more than 3 h. Clambank Landing retained 0%. Thousand Acre differed: all its >3-h occupancy occurred in 2021, and its contribution was zero in the four eligible years shared with Oyster Landing. Full deletion profiles and pairwise comparisons are retained in the supplementary results and machine-readable files (Table S9).

Persistence of compound episodes

All 417 detected primary episodes had valid non-compound observations on both boundaries. Their completeness does not describe unobserved events within data gaps. Singletons accounted for 158 episodes (37.9%) and 39.5 nominal

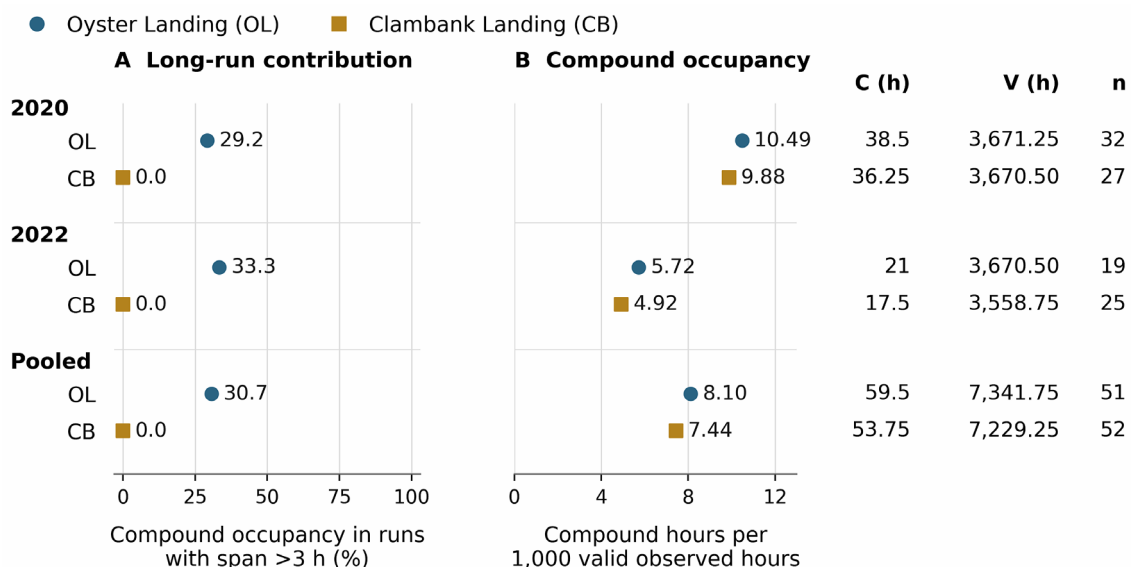
Table 2. Pooled occupancy and denominators within primary eligible seasons

Station	V (h)	L (h)	C (h)	C/L (%)	C/V (%)	L/V (%)	W/V (%)
NIW-CB	14,562.5	214.5	80.25	37.4	0.551	1.473	21.8
NIW-DC	14,470	9.25	3.75	40.5	0.026	0.064	21.8
NIW-OL	14,571	811.75	199	24.5	1.366	5.571	25.1
NIW-TA	18,206.5	286.5	52.25	18.2	0.287	1.574	23.8
GTM-FM	18,358.25	3.25	0	0.0	0.000	0.018	24.1
GTM-PC	18,181.75	39	23	59.0	0.127	0.215	22.0
GTM-PI	14,577.5	74.5	53.5	71.8	0.367	0.511	27.3
GTM-SS	14,686	0.75	0.25	33.3	0.002	0.005	24.4

Note: V: valid paired observations; L: oxygen <2 mg/L; W: temperature >P90; C: both conditions. Hours are nominal occupancy (reading count × 0.25 h). Pooled fractions sum numerators and denominators over the eligible years in Table 1. All pooled L denominators are positive; seasonal undefined fractions are retained as NA in Table S4. Rounding is for display only.

compound hours (9.6%). Retaining only spans of at least 0.25 h left 259 episodes and 372.5 h; a 0.5-h minimum left 187 episodes and 336.5 h. Frequent short detections therefore contributed much less occupancy than their share of the episode count.

Oyster Landing had 121 episodes with median span 0.75 h, 90th percentile 3.75 h, and maximum 6 h. Runs spanning more than 1 h carried 80.4% of its compound occupancy, and runs spanning more than 3 h carried 48.9%. Clambank Landing had 88 episodes, median span 0.25 h, and 90th



C = nominal compound hours; V = valid observed hours; n = complete runs. Annual and pooled points reuse observations and are not independent replicates.

Figure 2. Common-year occupancy and long-run contribution Oyster Landing and Clambank Landing are compared on their two common primary eligible seasons, 2020 and 2022, and on the pooled sum of those seasons

This post hoc comparison retains the original fixed station-specific P90, oxygen <2 mg/L, accepted readings, and episode boundaries. (A) Percentage of compound occupancy contributed by runs with first-to-last span >3 h. (B) Compound occupancy per 1,000 valid observed hours.

C and V give the corresponding nominal-hour numerator and observation denominator; n is the number of complete runs. Annual and pooled points reuse observations and are not independent replicates. Matching eligible years does not imply matching every valid timestamp; the separate exact-timestamp occupancy diagnostic is reported in the supplement.

percentile 1.825 h. Although runs longer than 1 h carried 55.8% of its occupancy, no primary run exceeded 3 h. Thousand Acre had 73 episodes, median span 0.25 h, and 13.9% of occupancy in runs longer than 3 h (Table 3). Figure 3 shows the full span distributions and the complementary occupancy fraction in runs longer than 1 h.

Pellicer Creek illustrates the distinction between run-count and time weighting: 21 of its 34 episodes were singletons, giving a median span of zero, yet runs longer than 1 h supplied 54.3% of its compound occupancy. Pine Island had 93 episodes with median span 0.25 h and 18.2% of occupancy in runs longer than 1 h. Debidue Creek

had seven episodes and San Sebastian one singleton; neither supported a substantial long-run profile. Fort Matanzas had no run distribution. All stations, including those sparse or absent distributions, are retained in Table 3.

All primary runs are fully bounded by valid non-compound readings. Span equals $(N - 1) \times 0.25$ h; occupancy equals $N \times 0.25$ h. Span quantiles weight episodes equally. Occupancy fractions weight each run by its compound-reading count. NA denotes no compound run or no compound occupancy denominator. Data: `independent_station_profiles_all_variants.csv` and `events.csv`; sensitivity profiles are in Table S5.

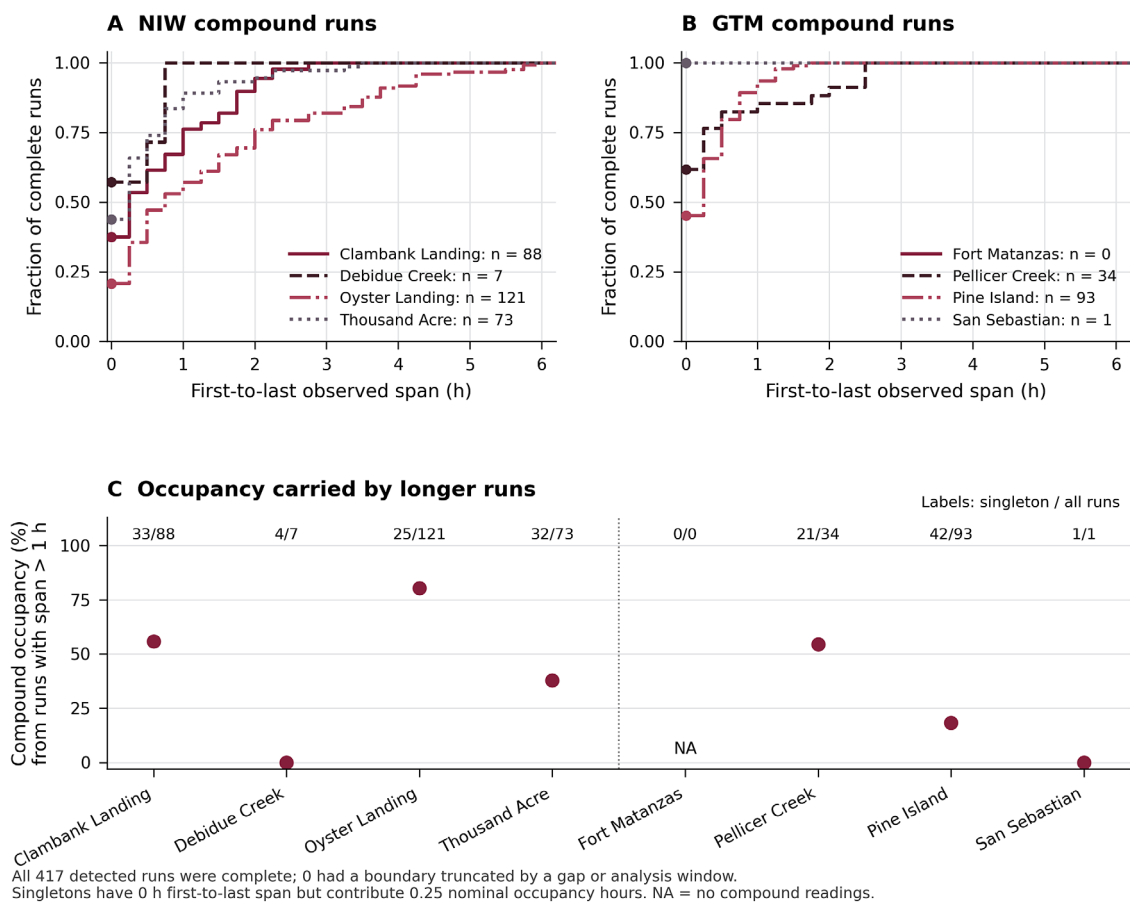


Figure 3. Compound-run spans and their occupancy contribution

(A–B) Empirical cumulative distributions of complete compound runs at NIW and GTM stations.

A run is a maximal uninterrupted sequence of consecutive compound readings. Its observed span is the elapsed time between first and last compound readings, $(N-1) \times 0.25$ h. Singletons therefore have zero observed span but retain their 0.25 nominal occupancy hours. Curves are descriptive ECDFs, not survival estimates, and begin at the singleton fraction. Legend n denotes complete runs. Fort Matanzas has no detected run and therefore no curve. All 417 detected primary runs have observed non-compound readings on both boundaries; none has a gap- or window-truncated boundary. (C) Percentage of each station's compound occupancy carried by runs with first-to-last span > 1 h. This time-weighted representation differs from the run-count weighting in A–B. Labels give singleton and total run counts, not an estimated biological rate.

NA marks the no-compound denominator at Fort Matanzas. A zero at Debidue Creek or San Sebastian is an observed zero contribution from runs exceeding 1 h.

Table 3. Primary compound-run spans and occupancy-weighted persistence

Station	Runs	Single-tons	Median span (h)	P90 span (h)	Maximum span (h)	Occupancy in span >1 h (%)	Occupancy in span >3 h (%)
NIW-CB	88	33	0.25	1.825	2.75	55.8	0.0
NIW-DC	7	4	0	0.75	0.75	0.0	0.0
NIW-OL	121	25	0.75	3.75	6	80.4	48.9
NIW-TA	73	32	0.25	1.4	3.5	37.8	13.9
GTM-FM	0	0	NA	NA	NA	NA	NA
GTM-PC	34	21	0	1.925	2.5	54.3	0.0
GTM-PI	93	42	0.25	0.95	1.75	18.2	0.0
GTM-SS	1	1	0	0	0	0.0	0.0

States separating episodes and recurrence

There were 410 adjacent-run pairs within stations. Fourteen crossed separate seasonal windows and 12 within-season pairs encountered missing observations, leaving 384 fully observed intervals. Of these, 207 (53.9%) contained no neither-state reading. The station counts were 44/112 at Oyster Landing, 32/79 at Clambank Landing, 35/68 at Thousand Acre, 16/30 at Pellicer Creek, 79/91 at Pine Island, and 1/4 at Debidue Creek. Fort Matanzas and San Sebastian provided no eligible pair. These are fractions of intervals, not fractions of elapsed time. A post hoc partition showed that 183 of the 207 intervals without neither were warm only throughout, 22 were low oxygen only throughout, and two contained both single-condition states without a neither reading (Table S10). Together these 207 intervals occupied 1,237 of 17,354.5 interepisode hours (7.1%); their frequency by count therefore does not describe their share of elapsed time.

Time-weighted composition differed with interval length (Figure 4). NIW intervals of ≤ 6 h contained 51.8% warm-only, 44.5% low-only, and 3.6% neither hours. NIW intervals longer than 72 h contained 58.1% neither hours. All 116.5 h in the 59 GTM intervals of ≤ 6 h were warm only, whereas neither occupied 58.7% of GTM interval hours in the >72 -h bin. Such reserve-pooled compositions reflect both the observed station contributions and their interval lengths; Figure S3 provides the station-specific distributions.

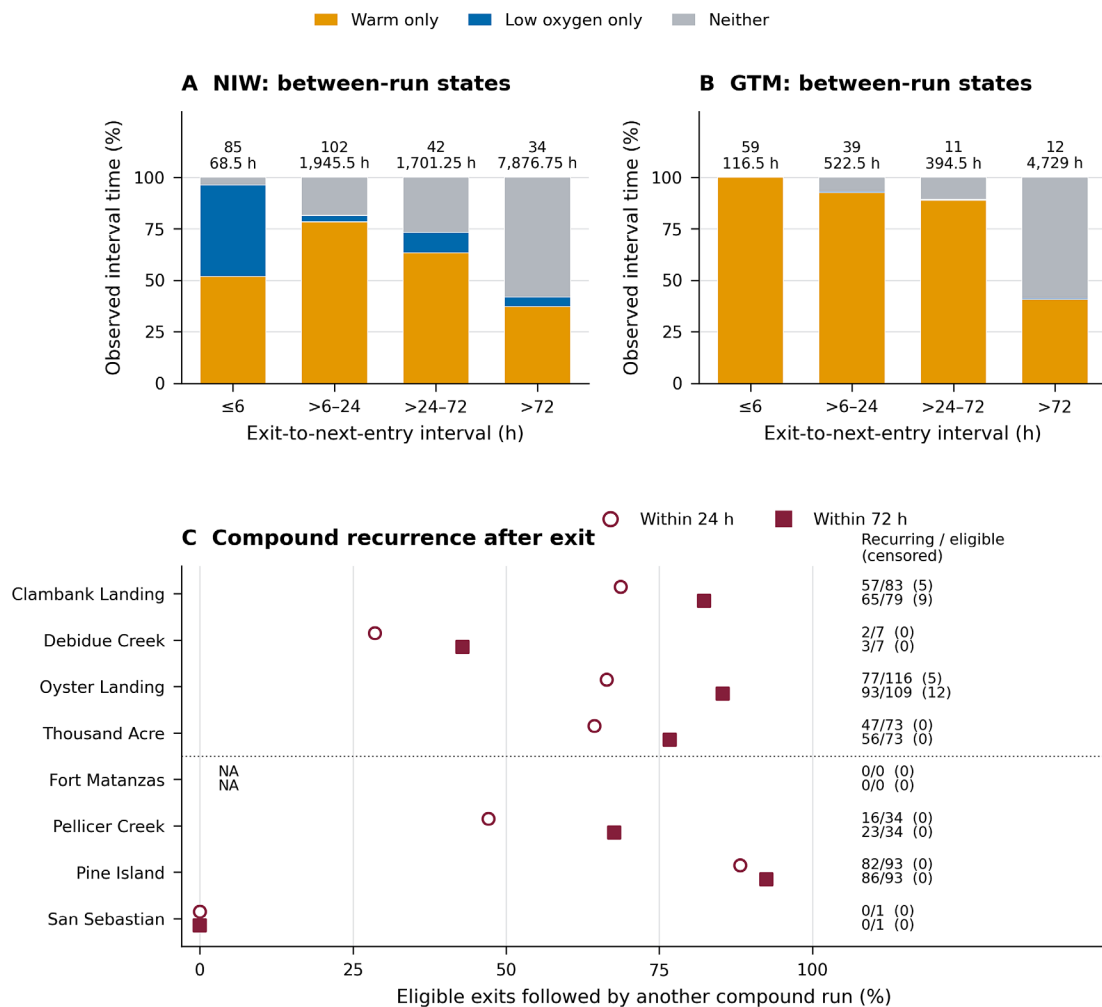
Of the 417 exits, 407 had a fully observable 24-h horizon and 396 had a fully observable 72-h horizon. At Oyster Landing, recurrence occurred after 77/116 eligible exits within 24 h and 93/109 within 72 h; corresponding Clambank Landing counts were 57/83 and 65/79. Thousand Acre

had 47/73 and 56/73. Pine Island had 82/93 and 86/93, but its coverage sensitivity substantially reduced the 24-h estimate, as described below. These fractions describe return to the measured compound state and have no assigned physiological timescale.

Figure 5 shows the longest complete run in each reserve, selected by a fixed rule. Oyster Landing on 21 August 2023 had 25 compound readings between 02:45 and 08:45 EST, giving 6 h of span and 6.25 nominal occupancy hours. Within that run, temperature was 29.9–30.7 °C and oxygen was 0.8–1.9 mg/L; the flanking oxygen readings were 2.0 mg/L. Pellicer Creek on 15 August 2023 had a 2.5-h span. These examples show the state sequence around deliberately selected long runs rather than typical station conditions.

Sensitivity of magnitudes and temporal profiles

The Oyster Landing long-run contribution persisted across several prespecified perturbations. With stricter coverage, 79 episodes remained, the 90th percentile span was 3.5 h, and 41.9% of compound occupancy occurred in runs longer than 3 h. That contribution was 57.5% under P85, 50.0% under P95, and 54.4% in 2021–2023. The conservative deployment mask left its primary results unchanged. Clambank Landing continued to have no >3 -h run under these thermal, coverage, and later-period alternatives. Small oxygen-margin perturbations to 1.9 and 2.1 mg/L also preserved this unequal-year pooled contrast. Using 3 or 4 mg/L defines a broader oxygen condition and produces longer runs at Clambank Landing as well; it does not test an identical phenomenon. These full-cohort results cannot be transferred automatically to the common-year comparison.



A–B: sums of readings in fully observed intervals, pooled within each reserve; labels show intervals and nominal hours.
 C: the full 24/72-h horizon must be observed. Missing horizons are censored, even when a re-entry is seen earlier.
 The four threshold states characterize measured environmental conditions; organism-level responses were not measured.

Figure 4. States between compound runs and fully observed recurrence

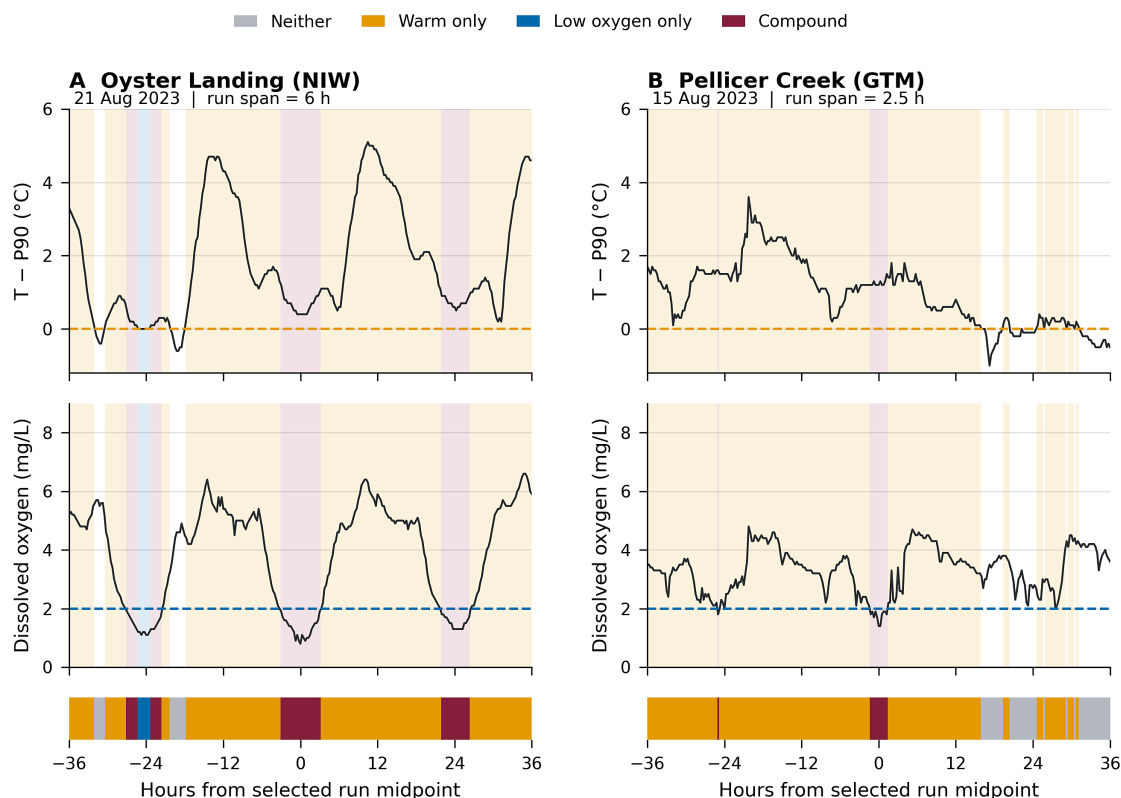
(A–B) Composition of observed intervals between successive compound runs, grouped into intervals ≤6, >6–24, >24–72, and >72 h and pooled within NIW or GTM. The exit anchor is the first non-compound reading following a run; the next-entry anchor is the first compound reading of the next run. State composition includes exit-anchor readings and excludes the next-entry reading. Colored shares are sums of state-specific reading-hours divided by the sum of fully observed interval hours within each bin. They are time-weighted compositions, not fractions of intervals or mean station values. Labels give interval count and nominal observed hours; all 384 eligible intervals contribute, whereas 26 other adjacent-run pairs are unavailable owing to separate windows or missing observations. Fort Matanzas and San Sebastian do not provide an eligible pair.

Station-specific compositions appear in Figure S3. (C) Fraction of eligible exits followed by a compound re-entry within 24 or 72 h. Eligibility requires every 15-min timestamp from exit through the specified horizon to be valid and inside the same observation window. Even an early observed re-entry does not make a later-incomplete horizon eligible. Labels give recurring/eligible exits, with censored counts in parentheses. Open circles denote 24 h and filled squares 72 h. NA indicates an absent denominator.

On the common 2020 and 2022 support, the post hoc crossed sensitivity gave Oyster Landing >3-h contributions of 47.3% under P85, 18.8% at oxygen <1.9 mg/L, and 24.8% at <2.1 mg/L, compared with 0% at Clambank Landing. Under P95, neither station had a >3-h run: Oyster Landing retained 10.75 compound hours in eight

episodes and Clambank Landing 5 h in five episodes. Oyster Landing had no P95 compound reading in 2022, so its annual persistence fraction was undefined. The <1.9 mg/L contrast also disappeared in 2022 (Table S11).

Occupancy magnitudes were strongly threshold-dependent (Figure S2). At Oyster Landing, the



Selection: longest complete run in each reserve; ties resolved by earliest start, then station ID.
Both windows span 72 h. Shading and bottom strip show the four observed states at 15-min resolution.
Source clock: fixed EST (UTC–5). Threshold crossings are bracketed by readings; continuous state is not asserted.

Figure 5. Reproducibly selected thermal–oxygen state sequences.

The longest complete compound run in each reserve was selected; ties were resolved by earliest start time and then station identifier. The selected runs are Oyster Landing, 21 August 2023 02:45–08:45 fixed EST (6-h first-to-last span), and Pellicer Creek, 15 August 2023 13:00–15:30 fixed EST (2.5 h). Each panel spans 72 h centered on that run's midpoint. Upper traces show temperature relative to the station-specific P90; lower traces show dissolved oxygen. Dashed lines mark the classification thresholds. Background shading and the categorical strip depict observed states; unobserved readings, if present, remain gaps and are not interpolated.

Both reserves use common axes. These deliberately selected maximum-span examples illustrate sequence structure rather than typical conditions; the complete distributions are given in Figure 3.

Fixed EST is UTC–5 without daylight-saving adjustment.

compound fraction of low-oxygen time changed from approximately 55.3% under P85 to 4.0% under P95; Clambank Landing changed from 60.4% to 3.4%. These alternatives do not support a universal ranking independent of the chosen thermal definition. The year-specific leave-one-year-out analysis and all other prespecified variants are retained in the supplementary results.

Under stricter coverage, 90/220 intervals (40.9%) had no neither-state reading. The conservative deployment alternative gave 128/293 (43.7%). At Oyster Landing, Clambank Landing, and Thousand Acre, stricter-coverage counts were 28/74, 19/55, and 35/68, respectively. In the post hoc outcome-year analysis, omitting 2023

reduced the pooled count to 75/202 (37.1%); the other single-year omissions gave 54.6–57.3%. Persistence of at least one exceedance between episodes was therefore a material feature of the observed sequences, but a majority of intervals was not a uniform result across specifications.

Pine Island was especially sensitive to observed support. Stricter coverage reduced its 24-h recurrence from 82/93 to 3/7 eligible exits. Under conservative deployment exclusions it had only two eligible seasons and failed the pre-defined three-season station requirement. None of its 93 primary episodes overlapped the optional excluded deployment windows; its loss of eligibility therefore does not demonstrate that those

episodes were artifacts. These findings limit interpretation of the pooled Pine Island profile as a stable multiyear station characteristic.

The secondary daily analysis included 5,304 eligible station-days. Daily mean temperature and oxygen were never jointly beyond the primary thresholds, 256 days contained a simultaneous compound reading, and 609 met the separate same-day extrema criteria. These quantities answer different aggregation questions; they are reported as an observation comparison rather than as the central environmental result.

DISCUSSION

Occupancy and the longer run contribution describe different features

The four-state description separates episode counts, occupancy-weighted persistence, and conditions between episodes. Each quantity refers to the recorded sequences, and their interpretation depends on the threshold and observation support. The post hoc Oyster Landing–Clambank Landing comparison illustrates that distinction on only two common eligible years, 2020 and 2022: compound occupancy rates were 8.10 and 7.44 h per 1,000 valid observed hours, while >3-h runs supplied 30.7% and 0% of compound occupancy under P90. Both stations had no >3-h run under P95 on the same support. Under stricter coverage, only one common year remained. This example is consequently insufficient evidence of a persistent station difference, a stable ranking, or a difference in ecological risk. The pooled P90 contrast cannot be generalized from its occurrence in these two observed years; threshold alternatives and annual counterparts delimit this exploratory example.

Single-reading detections require similar care. They were common among episode counts but supplied less than one tenth of total compound occupancy. Pellicer Creek had a median episode span of zero while more than half its occupancy occurred in runs exceeding 1 h. A statement that a station mainly experiences short episodes can therefore depend on whether “mainly” refers to the number of episodes or to the time assigned to them. Reporting both weights preserves this distinction without discarding accepted brief detections or inventing continuity between measurements.

Tassone et al. (2022) showed the occurrence of coincident estuarine extremes using a multiday event definition. Reinl et al. (2026) retained information from 15-minute oxygen observations when calculating monthly low-oxygen frequencies across NERRS. Nardi et al. (2026) further showed why conditions surrounding an event deserve attention, using temperature records and a daily, seasonally varying marine-heatwave framework. The contribution here is methodological and descriptive: a geographically bounded description of measured four-state sequences that links compound occupancy to run persistence and the joint thermal–oxygen composition between runs. We make no claim to discovering temperature–hypoxia co-occurrence or a new ecological mechanism. Earlier work already addresses event duration, warmth outside a heatwave, and spacing between thermal events; the present focus is which of the two measured conditions remains between compound runs. The results do not establish that the two variables coincide more often than expected after accounting for common seasonal or diel structure. That claim would require a separate dependence-preserving null model.

Compound cessation can leave one measured condition in place

The interval analysis gives environmental meaning to the end of a compound run. An exit can occur because temperature falls to or below P90 while oxygen remains below 2 mg/L, or because oxygen rises to or above 2 mg/L while temperature remains above P90. In 207 primary intervals, at least one exceedance was observed at every intermediate timestamp. Even when a neither-state reading occurred, its continuity and duration remained separate properties from the total interval length. This explains why time outside the compound state should not automatically be interpreted as time in the neither state.

The fraction of intervals lacking neither was 53.9% in the primary specification, 40.9% under stricter coverage, and 43.7% under conservative deployment screening. The consistent descriptive finding is a material subset of intervals with continuing exceedance, rather than a universal majority. These intervals were predominantly short: their combined lengths represented 7.1% of total eligible interval hours in the primary analysis, 4.6% under stricter coverage, and 5.0% under conservative screening. Those percentages

describe the entire lengths of intervals without neither; they do not include warm-only or low-only portions within intervals that also contained neither. Longer observed intervals contained larger neither-state shares in the reserve-pooled display, but those shares are time-weighted and reflect unequal station and year contributions. They should not be substituted for the station-specific interval counts or interpreted as a population recovery probability.

For monitoring assessments, the results support retaining separate temperature and oxygen state indicators when a compound indicator ends, with missing observations displayed as unknown. In the post hoc interval partition, warm-only intervals accounted for 183/207 (88.4%) of those without neither, compared with 22/207 (10.6%) that remained low oxygen only; two intervals contained both single-condition states. A station summary can report observed hours and seasonal support alongside compound occupancy, the share of that occupancy in sustained runs, and both the number and time composition of fully observed intervals. This identifies which measured condition remains after the overlap ends. The same distinctions can inform the environmental sequences selected for subsequent experiments. These descriptors characterize the monitored water; they do not establish an alarm threshold, an intervention benefit, or a biological response. Biological interpretation would additionally require organismal location, relevant thresholds, and responses over time. Lucey et al. (2022), for example, combined field records with species-specific experiments; those experimental responses cannot be transferred to the present thresholds and estuarine stations. No physiological recovery, mortality, or safe-habitat estimate follows from these records alone.

Environmental setting and observation scale

Different run profiles can be compatible with several combinations of water movement, mixing, and metabolism. Process-based work shows that the balance of mechanisms behind diel oxygen variability can change over short distances and with depth (Duvall et al., 2022). High-resolution measurements in Weeks Bay further demonstrate that even near-bottom oxygen gradients may be substantial over less than a meter (Jarvis et al., 2023). The present analysis did not estimate

process contributions, and station contrasts cannot be assigned to tidal transport, freshwater influence, nutrient loading, or local respiration from the state sequences alone.

Sensor geometry is particularly important. The NIW and GTM records describe nominally fixed near-bottom positions with documented deployment exceptions. They do not resolve the vertical extent of warm low-oxygen water or the alternative water masses available to organisms. This scope differs from the explicit water-column criterion of Wong et al. (2024). A station with no detected compound run cannot be designated a safe refuge, and an episode at one sensor cannot be expanded into an estuary-wide habitat-volume estimate.

Coverage and threshold choices limit generalization

High pooled completeness coexisted with unequal seasonal support. None of the eight stations had all six years eligible, and only one year supplied a complete eight-station comparison. The pooled profiles therefore combine spatial and temporal differences in the available record. Matching eligible years substantially reduced the Oyster Landing–Clambank Landing occupancy contrast, while retaining a P90-specific difference in the >3-h contribution that disappeared under P95. This threshold dependence and the two-year support prevent interpretation as a persistent site effect. Deleting outcome years also showed that Thousand Acre’s pooled long-run contribution depended entirely on 2021. Missing runs can also occur within eligible seasons, and the complete observed episodes do not rule out unobserved episodes during those gaps. Excluding gap-crossing intervals and incomplete recurrence horizons prevents a known misclassification, but does not recover the missing environmental history.

Pine Island shows why this qualification matters. Its compound occupancy was concentrated in 2023 and its recurrence estimate changed sharply with coverage selection. Removing possible problematic deployments reduced multiyear support without directly removing primary compound episodes. This supports caution about temporal stability and comparability; it supplies no basis for asserting that the 2023 signal was caused by a sensor fault. At Fort Matanzas, San Sebastian, and Debidue Creek, small absolute oxygen or episode counts

constrain the precision and meaning of conditional fractions even where coverage is high.

The six-year P90 reference includes the annual temperature distribution and was reused retrospectively in the analysis. It does not identify temperature anomalies relative to each day of year, and a warm state in this study is not a marine heatwave. Threshold shifts changed the amount of compound occupancy substantially. Broader oxygen cutoffs changed the question and could alter persistence contrasts. The sensitivity results therefore delimit which observed features survive reasonable perturbations; they are not interchangeable estimates of one threshold-free exposure quantity.

Finally, the 15-min interval bounds crossings but cannot establish a constant state between observations. Small threshold-margin checks cannot account for all calibration or deployment uncertainty, and six synchronized years provide limited information for resampling. The exploratory bootstrap intervals are conditional summaries rather than independent confirmation of between-station differences. The study supports local descriptive comparisons of observed sequences, without causal attribution, national representativeness, or climatic trend inference.

CONCLUSIONS

The eight station records contained 417 compound episodes and 412 nominal occupancy hours across 35 eligible seasons. Reporting episode counts alongside occupancy-weighted persistence and intervening states distinguishes different properties of the monitored sequences. Among 384 fully observed interepisode intervals, 207 lacked a neither-state reading: 183 were entirely warm only, 22 entirely low oxygen only, and two contained both single-condition states. Those intervals represented 53.9% of the count and 7.1% of total interval hours; a majority was not maintained across coverage, deployment, and outcome-year alternatives. The exploratory Oyster Landing–Clambank Landing comparison used only 2020 and 2022, and its >3-h contrast disappeared at P95. It does not establish a persistent station difference. The findings provide a local methodological and descriptive characterization of the water observed at these sensors, without causal, physiological, or estuary-wide habitat inference.

Acknowledgements

We acknowledge the North Inlet–Winyah Bay National Estuarine Research Reserve, including Erik Smith, Robert Dunn, Tracy Buck, and Baker Stevens, and the Guana Tolomato Matanzas National Estuarine Research Reserve, including principal investigator Nikki Dix and SWMP personnel Shannon Dunnigan, Silas Tanner, Kathryn M. Petrinc, and Megan Howkins, for collecting, quality controlling, and maintaining the 2018–2023 water-quality data used in this study. The data were provided through the National Estuarine Research Reserve System’s System-wide Monitoring Program and Centralized Data Management Office. The original monitoring research and data collection were conducted under an award from the Estuarine Reserves Division, Office for Coastal Management, National Oceanic and Atmospheric Administration.

REFERENCES

1. Chaya, T. M., Schafer, T. B., & Osborne, T. Z. (2023). Assessing water quality dynamics in tidally influenced blackwater rivers along a rural–urban gradient. *Water*, 15(23), 4154. <https://doi.org/10.3390/w15234154>
2. Codiga, D. L., Stoffel, H. E., Deacutis, C. F., Kiernan, S., & Oviatt, C. A. (2009). Narragansett Bay hypoxic event characteristics based on fixed-site monitoring network time series: Intermittency, geographic distribution, spatial synchronicity, and interannual variability. *Estuaries and Coasts*, 32(4), 621–641. <https://doi.org/10.1007/s12237-009-9165-9>
3. Dunn, R. P., Krask, J. L., Pinckney, J. L., & Smith, E. M. (2023). Contrasting trends in water quality between adjacent ocean- and river-dominated estuaries: Evidence for marsh porewaters as a source of nutrient enrichment? *Limnology and Oceanography*, 68(9), 2040–2058. <https://doi.org/10.1002/lno.12405>
4. Duvall, M. S., Jarvis, B. M., Hagy, J. D., III, & Wan, Y. (2022). Effects of biophysical processes on diel-cycling hypoxia in a subtropical estuary. *Estuaries and Coasts*, 45(6), 1615–1630. <https://doi.org/10.1007/s12237-021-01040-y>
5. Freeman, N. M., Hervieux, G., Alexander, M. A., Jacox, M. G., Amaya, D. J., Scott, J. D., & Capotondi, A. (2026). Characterizing compound physical and biogeochemical extremes in the California Current Large Marine Ecosystem. *PLOS Climate*, 5(2), e0000638. <https://doi.org/10.1371/journal.pclm.0000638>

6. Jarvis, B. M., Hagy, J. D., III, Melendez, W., Simmons, C. W., & Wan, Y. (2023). Measuring and modeling diel oxygen dynamics in a shallow hypereutrophic estuary: Implications of low oxygen exposure on aquatic life. *Science of The Total Environment*, 882, 163474. <https://doi.org/10.1016/j.scitotenv.2023.163474>
7. Lucey, N., Aube, C., Herwig, A., & Collin, R. (2022). Compound extreme events induce rapid mortality in a tropical sea urchin. *The Biological Bulletin*, 243(2), 239–254. <https://doi.org/10.1086/722283>
8. Nardi, R. U., Mazzini, P. L. F., Walter, R. K., & Shen, J. (2026). Persistent warm water anomalies before and after marine heatwaves amplify heat exposure and associated risks. *Communications Earth & Environment*, 7, Article 776. <https://doi.org/10.1038/s43247-026-03739-x>
9. National Estuarine Research Reserve System Centralized Data Management Office. (n.d.). SWMP data quality assurance and quality control. Retrieved September 30, 2026, from <https://nerrsdata.org/data/qaqc.cfm>
10. NOAA National Estuarine Research Reserve System. (2019). NOAA National Estuarine Research Reserve (NERR) System-wide Monitoring Program meteorological, water quality and nutrient/pigment data from 1994 to 2024 (NCEI Accession 0200366) [Data set]. NOAA National Centers for Environmental Information. Retrieved September 30, 2026, from <https://doi.org/10.25921/vw8a-8031>
11. Reinl, K. L., Dunn, R. P., Cressman, K. A., Collins, T., DeBose, J. L., Friedrichs, C. T., Goldstein, J. S., Helms, A. R., Kinkade, C., Krask, J. L., Parrish, D. B., Nicklay, H. N., Ridge, J. T., Sanger, D. M., Cianci-Gaskill, J. A., Dix, N. G., Grothues, T. M., McMurray, S. E., & Peter, C. (2026). In situ observations reveal continental-scale warming, oxygen decline, and eutrophication in U.S. estuaries. *Communications Earth & Environment*, 7, Article 716. <https://doi.org/10.1038/s43247-026-03934-w>
12. 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>
13. Wong, J., Münnich, M., & Gruber, N. (2024). Column-compound extremes in the global ocean. *AGU Advances*, 5(3), e2023AV001059. <https://doi.org/10.1029/2023AV001059>