1. INTRODUCTION

Marine heatwaves (MHWs) are a recently identified phenomenon (Hobday et al. 2016, Scannell et al. 2016, Carvalho et al. 2021) that are intensifying as the global climate warms as a direct response to human actions (Frölicher et al. 2018, Smale et al. 2019, Laufkötter et al. 2020). Prolonged MHWs have caused severe disruptions to associated marine ecosystems (Smith et al. 2023). In the Northeast Pacific,
the 2014–2016 MHW that occurred as a result of anomalously low rates of heat loss from surface ocean waters is considered to be the largest and most prolonged MHW recorded to date (Holbrook et al. 2020). With peak extent covering an ocean area the size of Canada and intensity more than 4°C above climatology, the ecosystem impacts were profound, with documented effects on phytoplankton through to whales (Di Lorenzo & Mantua 2016, Suryan et al. 2021).

However, other periodic ocean warming phenomena are longstanding in this region. The El Niño–Southern Oscillation (ENSO) has been a stable feature on decadal to century scales (Braganza et al. 2009), with ecosystem impacts ranging up through marine birds and marine mammals (Hodder & Graybill 1985, Bertram et al. 2005, Häussermann et al. 2017). There is some evidence that these phenomena are linked, such that the impact of MHWs in this region may be extended by ENSO onset 6–12 mo later (Capotondi et al. 2022). To the north, within the Bering and Chukchi Seas, the transition to a multi-year warm stanza in 2014 (Stabeno et al. 2017), concurrent with the Northeast Pacific MHW of 2014–2016, was associated with dramatic shifts in sea ice extent, including the lowest sea ice minimum (2017–2018) and cold pool extent on record (1980–2019; Stabeno & Bell 2019). Arctic marine ecosystems generally (Stocker et al. 2013), and the Arctic region in particular (Grebmeier & Maslowski 2014, Huntington et al. 2020), appear to be in a period of rapid transition (Mueter et al. 2017, Stabeno & Bell 2019). Arctic sea surface temperatures (SSTs) are rising rapidly, and recent observations suggest that the Arctic climate is entering a ‘new normal’ (Jeffries et al. 2013, Wood et al. 2015).

Although MHWs are generally identified as discrete periods of prolonged (days-to-months) and significantly elevated ocean temperatures (Hobday et al. 2016), operational definitions now provide a standardized method for identifying and characterizing MHW events. Typically, MHWs are identified according to a minimum number of consecutive days during which temperatures exceed the 90th (or 99th; Holbrook et al. 2020) percentile of the 30 yr SST climatology (Hobday et al. 2016, Smale et al. 2019). In addition to capturing truly extensive, persistent events, such as the Northeast Pacific MHW of 2014–2016, the operational definition encompasses events across multiple ocean warming phenomena, including ENSO and Arctic warming, which collectively affect North Pacific marine ecosystems in profound ways.

Holbrook et al. (2020) suggested that ‘skillful MHW prediction’ could assist natural resource management, informing adaptation toward negative (e.g. exceeding thermal tolerances, Ledet et al. 2021; trophic mismatch, Piatt et al. 2020; increasing human–wildlife conflicts, Santora et al. 2020, Samhouri et al. 2021) and potentially positive (e.g. range extension; Sanford et al. 2019) impacts. However, at the spatio-temporal scale and intensity of the largest MHWs, what may be more important is understanding and predicting immediate and long-term ecosystem responses (e.g. Beaugrand et al. 2019) and whether and where tipping points of ecosystem change reside (Heinze et al. 2021). Understanding and forecasting those effects may be particularly important given that the incidence of MHWs has dramatically increased in the Northeast Pacific over the last 2 decades, with a 3-fold increase in intensity and a 9-fold increase in duration for those occurring in 2000–2022 compared to 1982–1999 (Barkhordarian et al. 2022).

In general, lower trophic levels should respond to bottom-up forcing more quickly than upper-level organisms. For instance, during the Northeast Pacific MHW of 2014–2016, abundance and size distribution of mesozooplankton in the Gulf of Alaska changed in 2014 soon after the MHW onset (Suryan et al. 2021, Batten et al. 2022). By contrast, upper trophic impacts were not apparent until 2015 (seabirds; Piatt et al. 2020) or beyond (Pacific cod Gadus microcephalus; Barbeaux et al. 2020). The delay in upper trophic response existed despite suggested differences in physiological response at the organismal level between homeotherms and poikilotherms, the latter of which may be expected to respond earlier, with potential to exacerbate effects on the former (Piatt et al. 2020). Thus, upper trophic response can be seen as an indicator of marine ecosystem impact following a MHW, as these long-lived and often wide-ranging species have life-history and natural-history flexibility ameliorating short-term environmental forcing (Burger & Piatt 1990, Johns et al. 2022).

Marine birds are excellent indicators of ecosystem change in response to both environmental and anthropogenic system forcing (Furness & Camphuysen 1997, Sydeman et al. 2021), as they are numerous, particularly in productive coastal waters (Nur et al. 2011), and are more easily observed and enumerated than their prey (Piatt et al. 2007a). As long-lived, strong trophic interactors (Cairns 1988, Österblom et al. 2006), marine birds integrate bottom-up forcing across space and time. Thus, marine bird vital rates are frequently used to infer changes in ocean productivity (Lee et al. 2007, Jessup et al. 2009, Avery-
Gomm et al. 2012). Sydeman et al. (2021) advocated for expanded monitoring of seabird breeding productivity, as this metric is correlated with ocean warming, particularly for piscivorous, surface-foraging species in the Northern Hemisphere. However, while breeding depression is most often provoked by bottom-up shifts in ocean productivity (Frederiksen et al. 2006, Suryan et al. 2006), other factors can also depress reproductive success. Among those, top-down forcing (e.g. Parrish et al. 2001, Gladics et al. 2015, Piatt et al. 2020) and direct environmental perturbations (i.e. storms and/or heat stress; Boersma & Rebstock 2014, Cook et al. 2020) can reduce breeding productivity, such that breeding depression is not always an indication of altered trophic dynamics. In addition, many species are buffered against ecosystem change via behavioral modification (Piatt et al. 2007b), dietary breadth (Sinclair et al. 2008), and/or prey switching (Watanuki et al. 2022). In comparison, a mortality signal in marine bird species, and specifically those that are an order, or orders, of magnitude above normal, is a clear indication of significant and substantial ecosystem impact.

In this paper, we gathered and analyzed 29 yr (1993–2021) of beached bird monitoring data (~90 000 surveys) to identify and characterize mortality events of marine birds throughout the Northeast Pacific and Alaska. Collated from 4 large-scale programs, the compiled data represent the most extensive record, in time and across space, of beached bird stranding events ever analyzed. We explored the association between marine bird mortality events and measures of ocean warming, as well as more specific characteristics associated with MHWs. In addition to definitively demonstrating that warmer than normal water is associated with marine bird mortality events, our goals were to understand whether there are thresholds to upper trophic response, including onset timing, duration, spatial extent, and overall mortality event magnitude, and whether there are fundamental response patterns across multiple mortality events.

2. MATERIALS AND METHODS

2.1. Beached bird data and processing

We sourced data from 4 beached bird monitoring programs: BeachCOMBERS (BCOMBERS: Central California, established 1997), Beach Watch (BWATCH: Central California, established 1993), the Coastal Observation and Seabird Survey Team (COASST: Northern California to Washington, and Alaska, USA, established in 1999), and the British Columbia Beached Bird Survey (BCBBS: Canada, established in 2002; Fig. 1). All 4 programs employ an effort-standardized survey protocol, whereby participants survey a fixed beach segment of known length for bird carcasses. Although participants are free to survey more frequently, they are asked to perform surveys at least monthly (BCOMBERS, COASST, BCBBS), or every 2 wk (BWATCH). As a result, the majority (82%) of surveys completed within 100 d of the previous survey on that beach segment (95% of all surveys) were done so within 10–40 d (Fig. S1 in Supplement 1). All participants are trained by relevant program staff.

A survey consists of a single-pass search of the beach, from the water’s edge to the leading edge of vegetation. For each carcass encountered, species identity is determined via morphological (foot type), morphometric (tarsus, culmen, and wing chord measurements), and plumage characteristics, with the aid of a dichotomous beached bird key (Ainley et al. 1994, Hass & Parrish 2002). Following carcass identification, carcasses are photographed and either uniquely tagged (COASST) or marked (BCBBS, BCOMBERS, BWATCH) to identify them as a previously found carcass on subsequent surveys. Following each survey, data sheets and photographs are sent to program staff for verification and archiving.

Of the data available to us, we limited our analyses to surveys completed through 2021. We also excluded surveys from 3 regions: (1) south of 35.3° N in south-central California (BCOMBERS) due to the lack of long-term data (surveys began in 2013); (2) all beaches in Southeast Alaska (COASST) due to the exceptionally low carcass encounter rate in that region (42 birds across 1617 surveys, or 0.017 birds km⁻¹); and (3) the outer coast of British Columbia north of Tofino (49.16° N), as these locations were surveyed infrequently (Fig. 1). The remaining data encompassed 93 761 surveys across 1087 beach segments (Table S1 in Supplement 1). Data were screened to remove carcasses found and marked in a previous survey to avoid double counting. We retained carcasses independent of age group (adult, juvenile, fledgling, unknown age), as age was not recorded for the majority of species. Finally, we limited our analyses to species that are either entirely or seasonally dependent on the marine environment (hereafter marine birds), excluding all species that solely depend on terrestrial or freshwater environments and carcasses that could not be identified to any taxonomic level. These data were processed to create our base data set, which consisted
of survey-level counts of all newly found marine birds, beach length surveyed (km), survey interval (time since last survey in days), location (beach ID, latitude, longitude), and date.

In addition to regular beached bird monitoring data, we compiled information from opportunistic reports between 2017 and 2021 in the Bering and Chukchi Seas to bolster information from regions where regular monthly monitoring is scarce or absent (see Supplement 3). The majority of these data were communicated to the US Fish and Wildlife Service by coastal community members and/or Alaska Sea Grant Marine Advisors located in remote regions of northwest Alaska. While there was a concerted effort to record opportunistic reports following the mortality events in Alaska in 2015–2016, the reports were submitted voluntarily by individuals encountering marine bird carcasses. Therefore, opportunistic reports represent presence-only reports and are likely biased in space toward population centers that were aware of avenues for submitting reports. Reports were highly variable in the type(s) of information communicated; however, they always consisted of carcass counts or estimates at a given time and location. Due to concerns over spatial biases and lack of carcass-absence data, we included these data in event documentation and description but not statistical analysis.

### 2.2. Baseline model fitting

To model patterns in effort-standardized carcass abundance, we assigned the available data to one of 4 large marine ecosystems (LMEs): California Current (CCLME), Gulf of Alaska, Southern Bering Sea, and Northern Bering Sea/Southern Chukchi Sea (hereafter Northern Bering), based on survey location (Fig. 1). For the CCLME and Gulf of Alaska, we explored whether smaller regional subdivisions (~200 km) were necessary to capture seasonality and departures from baseline in beached bird abundance while avoiding spatio-temporal biases due to changes in survey coverage and extent within each LME. Because Southern and Northern Bering Sea surveys were concentrated in few locations, regional delineation was not attempted. Regional subdivisions were created based on major oceanographic or geographic features, such as river plumes and headlands (Fig. 1), beached bird abundance patterns, and
when surveys were established (Table S2 in Supplement 1). By comparing models with and without regional subdivisions, we found support for a subdivision within the CCLME but not for the Gulf of Alaska (Table S3 in Supplement 1). Therefore, models in the CCLME were fitted at the regional scale (Fig. 1E–G), whereas models for the Gulf of Alaska and Southern and Northern Bering Sea were fitted at the LME scale. For ease of language, we refer to the spatial scale of all models as ‘region’.

For each region, we fitted generalized additive mixed-effects models (GAMMs) to the beached bird data in order to estimate baseline seasonality as a function of day-of-year as well as departures from baseline through time. Our models contained 2 major components. The first was a seasonal effect, modeled as a cyclic cubic spline of survey day-of-year. This accounted for seasonality that was consistently observed across years, with fitted values that were continuous (i.e. cyclic) across the year end-start boundary. Next, we divided time (1993–2021) into discrete 28 d periods and labeled surveys according to the period (starting from the earliest surveyed month; 1 Sep 1993) in which they occurred. A 28 d period was used as it matches the most frequently observed survey interval in our data set (Fig. S1 in Supplement 1). Departures from baseline were modeled by including 28 d period as a random effect, allowing the fitted values for each period to fluctuate about long-term seasonality. The magnitude and direction of these random effects indicate to what extent carcass encounter rates (carcasses km−1) were above or below baseline for each period. Given that, we refer to this term as ‘encounter rate anomaly’.

We used a negative binomial (NB) GAMM (log-link function) with an offset term to account for differing beach lengths and a random effect of beach to account for consistent variability among beaches:

\[
\ln(\mu_{ij}) = \beta_0 + s(d_j) + \delta_t + \gamma_j + \ln(L_{ij}) \tag{1}
\]

\[
C_{ij} \sim \text{NB}(\mu_{ij}, \theta) \tag{2}
\]

where \(C_{ij}\) is the observed carcass count on survey \(i\) on beach segment \(j\), \(\mu_{ij}\) is the expected count, \(\beta_0\) is the intercept, \(s(d_j)\) is the seasonal smooth term of day-of-year, \(d_j\) is the random effect of time-period \(t\) (i.e. anomaly) within which the survey occurred, \(\gamma_j\) is the beach random effect, \(L_{ij}\) is beach length surveyed, and \(\theta\) is the negative binomial dispersion parameter. Within this model, random effects are normally distributed on the log-scale (log-link), such that:

\[
\gamma_j \sim \text{N}(0, \sigma_j) \tag{3}
\]

where \(\sigma_j\) is the standard deviation of beach random effects. However, temporal random effects were assumed to be autocorrelated through time, which we model via an AR1 process:

\[
d_t = \rho d_{t-1} + e_t \tag{4}
\]

\[
e_t \sim \text{N}(0, \sigma_e) \tag{5}
\]

where \(\rho\) is the lag-1 autocorrelation strength and \(e_t\) are the residuals of this process, assumed to be normally distributed with standard deviation \(\sigma_e\). Specifying autocorrelated random effects allowed for anomalies to persist through time and addresses the assumption of independence among random effects that would otherwise be violated.

Given the log-link function, random effects are log-normally distributed on the response (encounter rate) scale, allowing for positive deviations (i.e. elevated encounter rates) that are naturally more extreme than negative deviations (i.e. depressed encounter rates). For the same reason, random effects are multiplicative, such that anomaly estimates of \(\pm 1\) represent encounter rates that are \(e^{\pm 1} = 0.37\) (−) and 2.71 (+) times baseline, providing a natural way to classify departures from baseline.

In addition to these core model components, we also considered the fixed effects of beached bird program and survey interval to control for differences in survey protocol and carcass accumulation time between surveys, respectively. For each region, we compared models with and without these additional terms (program effects were only suitable in regions with overlapping programs; Fig. 1) and chose the most parsimonious model (i.e. that with the lowest Akaike’s information criterion adjusted for small sample sizes, AICc) for each region to perform all subsequent analyses with (Table S4 in Supplement 1).

Models were fitted using Stan (Stan Development Team 2022a) and RStan (Stan Development Team 2022b), the R interface to Stan. Model parameters were defined via vague priors and estimated using 4 Markov chain Monte Carlo chains, each running for 12,000 iterations split equally between warm-up and estimation phases. Model parameter convergence was evaluated based on traceplots and convergence criteria (R-hat \(\leq 1.05\); Vehtari et al. 2021). Models were further evaluated by checking that random effects distributions met the assumptions of normality, and model fit was checked through comparisons between observed and fitted values before proceeding with subsequent analyses (Supplement 2).
2.3. Event identification and classification

Mortality events were first identified at the region level based on model-estimated encounter rate anomalies. Regional events were identified as one or more contiguous time periods (i.e. 28 d) where the encounter rate was estimated to be at least twice baseline. Following that, we excluded all matching incidences where fewer than 50 carcasses were observed, removing cases that were a reflection of low and/or localized survey effort.

For each regional mortality event, a series of characteristics were calculated by either summing or averaging across time period (28 d) specific model estimates within the event time window (Table S6 in Supplement 4). These were (1) ‘duration’, calculated as time-period count multiplied by 28; (2) ‘average encounter rate’, calculated as mean across time-period-specific encounter rate estimates; (3) ‘average anomaly’, defined as mean across-encounter-rate anomalies, and converted into relative units:

$$\exp\left(\frac{1}{n_e} \sum_i^n \delta_i \right)$$

(9)

where \(n_e\) is the number of time periods within the event, and therefore represents the average relative departure from baseline; and (4) ‘extent’, defined as the event spatial coverage. Extent was calculated by first estimating the proportion of beaches within the event region where the observed count on any survey throughout the event duration was at least 3 times higher than the baseline expected count for that survey. This provided a measure of the coverage or ubiquity of elevated abundances, separating localized events from region-wide events. Due to differences in region extent, we then multiplied this proportion by the maximum point-to-point distance (km) between beach segments within each region as a measure of overall region size. We examined how this point-to-point distance varied through time due to the addition or removal of sites, and in most regions, it reached an asymptotic value as survey coverage became more comprehensive (Fig. S16 in Supplement 4). For each region, we chose the asymptotic value as a constant value to be applied in all extent calculations to avoid a perceived increase due to expanding survey coverage. Extents were not calculated for the Salish Sea regions due to the complex nature of the associated coastlines, or for the Northern or Southern Bering Sea due to the lack of comprehensive coverage (Fig. 1).

Following extent calculations, event ‘magnitude’ was calculated as the product of average carcass encounter rate, extent, and duration. This measure integrates carcass encounter rate over relevant space and time and can therefore be used to compare collective relative magnitudes among events. However, it should not be interpreted as cumulative deposition (i.e. the product of carcasses \(km^{-1}\) survey\(^{-1}\) × km × days), as it omits factors necessary to convert carcass counts to deposition (Jones et al. 2017).

Mortality events were also sorted into 4 encounter rate categories. This was based on average encounter rate and anomaly relative to baseline (Table 1). Category 1 events were just above our minimum threshold of twice baseline. By contrast, Category 4 events represented only the most severe carcass encounter rates: anomalies greater than 5 times the baseline and averaging at least 8 carcasses \(km^{-1}\).

Given the separation of US/Canadian coastlines outside of Alaska into distinct regions, single expansive events were documented multiple times in different regions if the event crossed regional boundaries. We therefore combined regional events into a smaller number of distinct events. This was based on location, requiring that they occur within the same LME; timing, requiring that event mid-points be separated by <90 d; and taxonomic composition, requiring that they be categorized in the same event typology based on a hierarchical cluster analysis of taxonomic composition among regional events (see Supplement 4). Aggregated event characteristics were then calculated by either averaging (encounter rate, duration) or summing (extent, magnitude) across constituent regional event values.

2.4. SST data and MHW identification

SST data were obtained from the NOAA OI SST V2 High Resolution Dataset (daily, spatial resolution: 0.25°; Reynolds et al. 2002) for 1981–2021, inclusive. For each region, we defined a bounding box from the corresponding shoreline out to 100 km offshore and calculated the average SST within that area for each

<table>
<thead>
<tr>
<th>Category</th>
<th>Average encounter rate (carcasses km(^{-1}))</th>
<th>Average anomaly (× baseline)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&gt;1</td>
<td>&gt;2</td>
</tr>
<tr>
<td>2</td>
<td>&gt;2</td>
<td>&gt;3</td>
</tr>
<tr>
<td>3</td>
<td>&gt;4</td>
<td>&gt;4</td>
</tr>
<tr>
<td>4</td>
<td>&gt;8</td>
<td>&gt;5</td>
</tr>
</tbody>
</table>
The regional SST values were then smoothed using a 30-d Gaussian kernel smooth after Hobday et al. (2016). From this, we calculated the region-averaged 30 yr climatology (1981–2010) by taking the mean regional SST for each calendar day-of-year. We then converted the regional SST time-series to SST anomaly (SSTa) by subtracting the climatological average. For each mortality event identified from beached bird data, we calculated the mean SSTa value for the corresponding region(s) averaged over the year prior to mortality event onset (onset date − 360 to onset date, inclusive) providing a measure of temperature conditions prior to the event. The resolution of SST data prevented calculations within the Salish Sea. For the single mortality event identified in the Salish Sea, we used SST data from the Juan de Fuca Eddy region (Fig. 1), as that mortality event extended from the Strait of Juan de Fuca to the outer coast of northern Washington.

In addition, we used SST data at the regional and LME scale (<100 km offshore) to identify MHW days (SST > 90th climatological percentile for that day-of-year) and MHW episodes (contiguous period of at least 6 MHW days, allowing for a maximum 2 d break) following the methods of Hobday et al. (2016). For each MHW episode, we calculated onset and termination dates as well as average intensity (mean SSTa).

In all instances, associated SST and MHW results should be interpreted within the context of large-scale (200–1000 km) temperature perturbations given the spatial scale of our analyses.

2.5. Aggregated event characteristics and connection to SSTa

Aggregated event statistics (average anomaly, average encounter rate, extent, duration, magnitude; averaged or summed across constituent regional events) were used to identify whether event characteristics had changed throughout the study period. To do this, we fit linear models with time as the single predictor variable, using weighted least squares with an exponential variance function ('varExp' in the 'gls' function in R; Pinheiro & Bates 2022) to control for heteroscedasticity that was evident upon examining the data. We then used bootstrap resampling (n = 1000 permutations) to estimate the 95% confidence intervals of model coefficients regarding event characteristic trends through time.

We also investigated whether there was a relationship between aggregated event characteristics and prior-year averaged SSTa. We used the same modeling process as described for trends through time but also included season (winter: Dec–Feb, spring: Mar–May, summer: Jun–Aug, autumn: Sep–Nov) as a factorial variable based on event onset date. For both time and SSTa analyses, we fitted models to characteristics data from all aggregated events as well as to a reduced data set that did not contain Category 1 events (Table 1) to identify whether patterns were more evident for higher encounter rate mortality events.

To determine whether there were thresholds in SSTa associated with mortality event severity, non-linear relationships between aggregated event magnitude and SSTa were explored. We fitted shape-constrained (SC-GAM; Pya 2022) and ordinary generalized additive models (GAMs; Wood 2017), using the ‘scam’ and ‘gam’ functions in the ‘scam’ and ‘mgcv’ packages in R, respectively (Wood 2017). The former approach allows for specification of monotonic functions between SSTa and event magnitude that are likely to be more representative of real-world functional relationships. We compared AICc between 4 different models for event magnitude: a null model, an unconstrained GAM with SSTa as a smooth term, and 2 SC-GAMs, each containing SSTa effects via a smooth term but differing with regard to functional constraints (monotonic increase, monotonic decrease). The best-fitting model (lowest AICc) was used to estimate the difference in event magnitude between SSTa = 0°C and SSTa = 1°C. Furthermore, functional thresholds were also calculated from the fitted function by calculating the SSTa value where model-estimated magnitude approached within 10% of its minimum (lower threshold) and maximum (upper) values. These thresholds provide measures of SSTa below or above which we would expect lower or larger magnitude events, respectively. Confidence intervals for these thresholds were determined via posterior simulation (Simpson 2022).

2.6. MHW association analyses

If mortality events are associated with MHWs, they should occur more frequently following MHW onset than would be expected by chance. To test this hypothesis, regional mortality events were linked to the most recent MHW. We calculated 2 test statistics from these data: the number of mortality events occurring within 12 mo of a MHW and the mean time delay, both based on time between MHW onset and mortality event onset dates. We then created a procedure where mortality events were randomly re-
arranged in time to represent the occurrence of mortality events de-coupled from MHWs and then calculated the corresponding statistics for each realization of this randomization process (for details see Supplement 6). Null distributions of our test statistics were constructed based on 1000 repetitions of this procedure, and we used this to calculate p-values and the 95% range of values expected under the null hypothesis that mortality events occurred at random with respect to MHWs. We performed this test procedure using encounter rate Category 2–4 events as well as only using Category 3–4 events to determine whether results differed when considering only the most extreme events.

2.7. Comparison of impacts between prolonged MHWs

To explore the degree to which there is a system-level pattern of response to MHWs, we focused on the progression of beached bird encounter rates, before, during, and after 3 prolonged MHWs in the Northeast Pacific: the 1997–1998 ENSO event (Humphries et al. 2015), the 2014–2016 Northeast Pacific MHW (Di Lorenzo & Mantua 2016), and the 2018–2019 ENSO event (Chen et al. 2021).

We examined the time series of encounter rate anomalies from 6 mo before through 4 yr after the onset of each MHW to identify whether, and for how long, beached bird abundances and phenology were disrupted following MHW occurrence. We limited these exploratory analyses to the central California Current regions, as monitoring elsewhere was established after the 1997–1998 ENSO event.

All analyses were performed in R version 4.2.1 (R Core Team 2022).

3. RESULTS

3.1. Event identification and aggregation

Across the 13 regions examined, 28 d encounter rate anomalies (hereafter anomalies) exceeded the 2× baseline threshold for 11.7% of the 3119 region × time periods examined (Fig. 2). These were not equally spread among regions and ranged from 0–4% in the sheltered waters of the Salish Sea to 26% in the Southern Bering Sea (Fig. 2A). Grouping contiguous 28 d periods resulted in the identification of 111 region-specific events (Fig. 2A). The majority (82%) of regional mortality events identified were constituents of broader-extent events spanning multiple regions (Fig. 2A). Grouping regional mortality events based on timing, location, and taxonomic composition (Fig. S18 in Supplement 4) resulted in the identification of 49 distinct events (Fig. 2B). The CCLME dominated event counts, with only 9 of the 49 regionally aggregated events occurring in the Gulf of Alaska and either the Northern or Southern Bering Sea. However, survey coverage was considerably higher in the CCLME, such that event detection was much higher. Although event counts per year were higher from 2005 onwards (Fig. 2B), this is largely a reflection of expansion of survey effort (i.e. Fig. 2A). Furthermore, prior to 2001, survey coverage in the CCLME was limited to central California (Beach Watch and BeachCOMBERS; Fig. 1), likely limiting event detection and giving rise to the sparsity of events before the early 2000s (Fig. 2B). Only 2 events were identified in the Salish Sea (2007 and 2016), with the 2016 event extending to the northern outer coast of Washington (Fig. 2A).

Category 4 events, corresponding to the highest carcass encounter rates, rarely occurred (n = 8), with 5 occurring between 2014 and 2019 (Fig. 2B–D). These events were dominated by alcids between 2014 (Cassin’s auklets Ptychoramphus aleuticus) and 2019 (common murres), and by procellariids prior to 2014 (northern fulmars Fulmarus glacialis in 2003) and from 2019 onwards (short-tailed shearwaters Ardenna tenuirostris; Table 2). The one exception was a 2009 mortality event that primarily affected scoters Melanitta sp. (Table 2). Events with the highest encounter rates did not necessarily coincide with those that were extensive in space (Table 2). Mortality events spanning more than 500 km were identified in 2005 (murres), 2010 (fulmars), and 2012 ( puffins, primarily rhinoceros auklets Cerorhinca monocerata), but these were all associated with relatively lower average encounter rates (<10 birds km⁻¹; Table 2). Collectively, mortality events prior to 2014 were predominantly either spatially extensive but relatively low average encounter rate (2005, 2010, 2012) or high average encounter rate but more localized (2009; Table 2). Only one high encounter rate, large-extent event was identified prior to 2014 (2003), whereas 4 were identified from 2014–2019 (Table 2). However, categorization of the 1997–1998 mortality event in the CCLME and the 2007 mortality event in the Southern Bering Sea (Table 2) as either large-scale and/or intense is uncertain given that they occurred when survey coverage was relatively lower in these regions.

The lack of coverage in Alaska prior to 2006 combined with sparse coverage relative to the CCLME
Fig. 2. Regional patterns of carcass encounter rate (carcasses km⁻¹) anomaly relative to seasonal baselines, and summarized information regarding mortality event occurrence. (A) Region-specific encounter rate anomalies, defined as the multiplicative difference between observed carcass encounter rate and seasonal baselines, from 1993–2021. Grey bars: beached bird monitoring data were unavailable. (B) Aggregated event counts by year, shown according to event onset year and color-coded according to encounter rate category (see Table 1). (C) Opportunistic reports of seabird mortality in the Southern and Northern Bering Seas, shown in annual-summarized form as cumulative carcass count (black) and mean carcass count per report (grey). Given the nature of opportunistic reporting, these data are not effort-standardized. (D) Graphical representation of mortality event occurrence through time (circle size is scaled to relative event magnitude), calculated as the product of event-averaged encounter rate, duration, and spatial extent (calculated values are equivalent to values shown in D × 10⁴). Events are color-coded according to dominant taxonomic group (Gr: grebe *Aechmophorus* spp.; Co: cormorants *Uria* spp.; Sc: scoters *Melanitta* spp.; Ki: kittiwakes *Rissa* spp.). For events primarily documented in opportunistic reports (dashed circles), spatial extent was approximately estimated (500 km in 2017–2018, 800 km in 2019, 100 km at other times). Major events documented in the literature (1993: Piatt & Van Pelt 1997; 1997: Baduini et al. 2001; 1998: Piatt et al. 1999; 2013: Bodenstein et al. 2015) but not captured in our analyses are also shown (no outline) as equal-sized circles, as equivalent magnitude measures were not comparable among published events. Events documented via opportunistic reports and published literature in (D) are shown for illustration purposes only and are not included in further statistical analyses.
certainly biased our results towards the south. Despite the sparsity of regular survey coverage (i.e. Fig. 1), our primary analyses indicated mortality events in the Southern Bering Sea in 2007 and 2016 and throughout the Southern and Northern Bering Sea in 2018 and 2019. However, opportunistic reports were also suggestive of mortality events in the Northern Bering Sea in 2017 and 2021 (Fig. 2C), primarily concentrated in the Bering Strait, north of St. Lawrence Island (Fig. S15 in Supplement 3). The total number of carcasses reported opportunistically along with the mean counts per report suggest similar-magnitude mortality events in 2017 (approximate total carcass count, N = 1200), 2018 (N = 1200), and 2021 (N = 1300), in addition to the much larger and more expansive (Southern and Northern Bering Sea) event in 2019 (N = 10 000; Fig. 2C). Opportunistic reports were also made in 2020; however, the total number of carcasses reported (N = 250) and the average count per report were seemingly not reflective of unusually high mortality in that year (Fig. 2C).

### 3.2. Event characteristic analyses

Aggregated event (n = 49) statistics varied considerably among events, with each individual characteristic spanning one or more orders of magnitude (Fig. 3). Relative to long-term baselines of carcass encounter rate, 57% of all regionally aggregated events had encounter rate anomalies 2–3 times above seasonal baselines, with only 6 events exceeding 5 times the baseline (Fig. 3A). However, all of these events were from LMEs in Alaska, where baseline encounter rates are very low at certain times of year, potentially inflating these measures. Duration and spatial extent were also skewed towards shorter, smaller events, with only 14% of all events lasting longer than 4 mo and 16% (n = 43; excluding Southern and Northern Bering Sea events where extent calculations were not performed) spanning more than 500 km (Fig. 3).

Aggregated event characteristics displayed no consistent linear trend through time (Table S8 in

---

### Table 2. Large-scale and/or high encounter rate marine bird mortality event characteristics and taxonomic composition. Events are aggregated and represent those that were either in the highest encounter rate category or extended over more than 500 km of coastline (bold). Summaries for events in the Northern and Southern Bering Sea that were primarily documented by opportunistic reports are given for years where the opportunistic count exceeded 1000 birds. Max encounter rates represent the peak encounter rate observed in a particular region and month throughout the mortality event duration. Counts and encounter rates represent all marine bird carcass finds within event bounds. CC: California Current Large Marine Ecosystem (LME); GA: Gulf of Alaska LME; SBS: Southern Bering Sea LME; NBS: Northern Bering Sea LME

<table>
<thead>
<tr>
<th>LME</th>
<th>Timing</th>
<th>Encounter rate</th>
<th>Extent (km)</th>
<th>Magn. (×1000)</th>
<th>Count</th>
<th>Primary</th>
<th>Secondary</th>
<th>Tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cat. Mean (max.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCa</td>
<td>Dec–Jul 1997–1998</td>
<td>3 3.0 (5.9)</td>
<td>348</td>
<td>178</td>
<td>2996</td>
<td>Murre (34%)</td>
<td>Scoter (12%)</td>
<td>Shearwater (11%)</td>
</tr>
<tr>
<td>CC</td>
<td>Aug–Feb 2003–2004</td>
<td>4 4.2 (13.2)</td>
<td>798</td>
<td>279</td>
<td>3776</td>
<td>Fulmar (79%)</td>
<td>Murre (7%)</td>
<td>Gull (6%)</td>
</tr>
<tr>
<td>CC</td>
<td>Apr–Oct 2005</td>
<td>3 2.5 (5.7)</td>
<td>755</td>
<td>170</td>
<td>2328</td>
<td>Murre (48%)</td>
<td>Cormorant (24%)</td>
<td>Gull (15%)</td>
</tr>
<tr>
<td>SBSb</td>
<td>Jun–Nov 2007</td>
<td>2.2 (9.0)</td>
<td>320</td>
<td>121</td>
<td>2612</td>
<td>Scoter (36%)</td>
<td>Murre (22%)</td>
<td>Fulmar (22%)</td>
</tr>
<tr>
<td>CC</td>
<td>Sep–Dec 2009</td>
<td>5.4 (10.0)</td>
<td>524</td>
<td>132</td>
<td>3379</td>
<td>Fulmar (76%)</td>
<td>Gull (6%)</td>
<td>Murre (5%)</td>
</tr>
<tr>
<td>CC</td>
<td>Oct–Dec 2010</td>
<td>3 5.1 (7.6)</td>
<td>641</td>
<td>127</td>
<td>1631</td>
<td>Puffin (53%)</td>
<td>Grebe (12%)</td>
<td>Murre (9%)</td>
</tr>
<tr>
<td>CC</td>
<td>Jan–May 2012</td>
<td>3 2.0 (4.3)</td>
<td>813</td>
<td>478</td>
<td>13238</td>
<td>Auklet (66%)</td>
<td>Murre (17%)</td>
<td>Fulmar (6%)</td>
</tr>
<tr>
<td>CC</td>
<td>Oct–Mar 2014–2015</td>
<td>5.2 (28.3)</td>
<td>810</td>
<td>562</td>
<td>11312</td>
<td>Murre (76%)</td>
<td>Cormorant (8%)</td>
<td>Gull (5%)</td>
</tr>
<tr>
<td>GA</td>
<td>May–Mar 2015–2016</td>
<td>5.0 (49.5)</td>
<td>850</td>
<td>1924</td>
<td>5109</td>
<td>Murre (97%)</td>
<td>Gull (1%)</td>
<td>Auklet (1%)</td>
</tr>
<tr>
<td>CC</td>
<td>May–Oct 2019</td>
<td>3 3.4 (9.3)</td>
<td>311</td>
<td>52</td>
<td>1322</td>
<td>Murre (75%)</td>
<td>Cormorant (5%)</td>
<td>Puffin (4%)</td>
</tr>
<tr>
<td>SBSc</td>
<td>Jun–Oct 2019</td>
<td>4 10.3 (41.2)</td>
<td>960</td>
<td></td>
<td>1318</td>
<td>Shearwater (60%)</td>
<td>Murre (27%)</td>
<td>Kittiwake (7%)</td>
</tr>
</tbody>
</table>

Opportunistic reports (years with >1000 birds reported)

<table>
<thead>
<tr>
<th>LME</th>
<th>Timing</th>
<th>Extent (km)</th>
<th>Magn. (×1000)</th>
<th>Count</th>
<th>Primary</th>
<th>Secondary</th>
<th>Tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBS</td>
<td>Jun–Sep 2017</td>
<td>−600</td>
<td></td>
<td>1177</td>
<td>Fulmar (44%)</td>
<td>Shearwater (30%)</td>
<td>Murre (15%)</td>
</tr>
<tr>
<td>NBS</td>
<td>May–Jul 2018</td>
<td>−500</td>
<td></td>
<td>1232</td>
<td>Murre (92%)</td>
<td>Shearwater (3%)</td>
<td>Fulmar (1%)</td>
</tr>
<tr>
<td>NBS</td>
<td>Jun–Oct 2019</td>
<td>−800</td>
<td></td>
<td>10507</td>
<td>Shearwater (95%)</td>
<td>Murre (2%)</td>
<td>Kittiwake (1%)</td>
</tr>
<tr>
<td>NBS</td>
<td>Jun–Sep 2021</td>
<td>−300</td>
<td></td>
<td>1318</td>
<td>Shearwater (60%)</td>
<td>Murre (27%)</td>
<td>Kittiwake (7%)</td>
</tr>
</tbody>
</table>

--

*aThis event occurred prior to widespread monitoring throughout the CCLME, and therefore the event extent and magnitude are likely underestimates

*bThis event occurred early after monitoring commenced in the Southern Bering Sea, and few sites were monitored at that time

This event is also noted in the Opportunistic reports category.
For the majority of characteristics, models with time as a predictor had higher AICc values than corresponding null models and bootstrapped 95% confidence intervals of trend estimates overlapped with zero (Table S8 in Supplement 5). We found moderate to weak support for linearly increasing trends of encounter rate anomaly (trend = 0.021, 95% CI = [0.003, 0.039], ΔAICc = 5.4 compared to null) and event magnitude (trend = 0.136, 95% CI = [−0.028, 0.322], ΔAICc = 0.3), albeit the latter only when excluding the lowest event category (Table S8 in Supplement 5). However, these results were driven almost exclusively by events in 2015–2016, and trend estimates were not robust to their removal. In summary, there was weak-to-no evidence that either individual event characteristics or cumulative event magnitude increased (linearly) throughout the study period.

However, we found that aggregated mortality event characteristics were correlated with SST anomalies averaged over the year prior to event onset date, tending toward more extreme events following periods of elevated SST (Fig. 3). Not all mortality events, however, were associated with warmer conditions. Year-averaged SST anomalies varied widely among mortality events, ranging from cooler (SST0 = −0.8°C) to considerably warmer (SST0 = +1.7°C) than normal conditions (Fig. 3). Events following warm (SST0 > 0) conditions included a greater proportion of higher encounter rate category (2−4) events (cool: 54%, n = 27; warm: 71%, n = 22), and averaged longer durations (cool: mean = 65 d; warm: mean = 90 d), and
larger spatial extents (cool: mean = 233 km; warm: mean = 408 km) (Fig. 3).

With the exception of encounter rate anomaly, models of event characteristics with SSTa as the sole predictor were supported over null models based on AICc (Table 3). For event extent and average carcass encounter rate, evidence in favor of the SSTa model over the null model varied from weak to moderate (ΔAICc < 2; Table 3). However, there was relatively stronger evidence of SSTa effects for event duration and magnitude response variables (ΔAICc > 2; Table 3). Despite that, SSTa only explained a small proportion of event characteristic variation (3–15%; Table 3), and while SSTa trend coefficient point estimates were uniformly positive, confidence intervals overlapped with zero for all but event duration (Table 3). However, excluding Category 1 events strengthened SSTa relationships for extent, duration, and event magnitude (Table 3).

The inclusion of season as a categorical predictor was only supported for models of average carcass encounter rate and for event duration when Category 1 events were excluded (Table 3). Autumn had the highest carcass encounter rates compared to other seasons, possibly resulting from elevated post-breeding mortality of juveniles and adults, whereas event durations were more prolonged in winter, and shorter during summer (Table S9 in Supplement 5).

Non-linear models of event magnitude as a function of prior-year averaged SSTa revealed a sigmoidal (step-function) relationship for both shape-constrained (SC-GAM: monotonically increasing function) and unconstrained GAMs (Fig. 3E). Both model types were supported over the null model (lower AICc) and were functionally equivalent in shape. Given that, and along with the lower AICc value of the SC-GAM compared to the unconstrained GAM (ΔAICc = 1; Table 3), we based all

Table 3. Model statistics and trend estimates from models fitted to event characteristics as a function of year-averaged sea surface temperature anomaly (SSTa). Model statistics for non-linear generalized additive models of event magnitude are presented for unconstrained (GAM) and shape-constrained (SC-GAM) models. Mean trend estimates of SSTa coefficients are given for linear models, along with their bootstrapped 95% confidence intervals. AICc: Akaike’s information criterion adjusted for small sample size.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Model</th>
<th>All events</th>
<th>Event Category: 2–4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AICc</td>
<td>R²</td>
<td>Trend estimate (95% CI)</td>
</tr>
<tr>
<td>Average event anomaly</td>
<td>ln(y) ~ 1</td>
<td>51.6</td>
<td>3 0.16 (−0.09, 0.38)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa</td>
<td>52.5</td>
<td>1 0.16 (−0.09, 0.38)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ season</td>
<td>58.6</td>
<td>4 0.15 (−0.18, 0.39)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa + season</td>
<td>60.0</td>
<td>4 0.15 (−0.18, 0.39)</td>
</tr>
<tr>
<td>Average carcass encounter rate</td>
<td>ln(y) ~ 1</td>
<td>119.6</td>
<td>1 0.16 (−0.09, 0.38)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa</td>
<td>118.4</td>
<td>7 0.41 (−0.01, 0.90)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ season</td>
<td>110.2</td>
<td>30 0.41 (−0.01, 0.90)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa + season</td>
<td>109.5</td>
<td>35 0.38 (−0.05, 0.84)</td>
</tr>
<tr>
<td>Duration</td>
<td>ln(y) ~ 1</td>
<td>104.7</td>
<td>15 0.40 (0.02, 0.74)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa</td>
<td>99.7</td>
<td>15 0.40 (0.02, 0.74)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ season</td>
<td>111.5</td>
<td>1 0.40 (0.02, 0.74)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa + season</td>
<td>105.3</td>
<td>18 0.45 (0.05, 0.85)</td>
</tr>
<tr>
<td>Extent</td>
<td>ln(y) ~ 1</td>
<td>120.2</td>
<td>5 0.37 (−0.10, 0.78)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa</td>
<td>120.6</td>
<td>5 0.37 (−0.10, 0.78)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ season</td>
<td>126.7</td>
<td>2 0.37 (−0.10, 0.78)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa + season</td>
<td>126.9</td>
<td>8 0.37 (−0.13, 0.87)</td>
</tr>
<tr>
<td>Event magnitude</td>
<td>ln(y) ~ 1</td>
<td>165.6</td>
<td>10 0.90 (−0.10, 1.87)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa</td>
<td>163.5</td>
<td>10 0.90 (−0.10, 1.87)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ season</td>
<td>170.8</td>
<td>5 0.80 (−0.01, 1.62)</td>
</tr>
<tr>
<td></td>
<td>ln(y) ~ SSTa + season</td>
<td>166.6</td>
<td>20 1.11 (−0.15, 2.21)</td>
</tr>
<tr>
<td></td>
<td>GAM: ln(y) ~ s(SSTa)</td>
<td>163.9</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>SC-GAM: ln(y) ~ s(SSTa)</td>
<td>162.9</td>
<td>16</td>
</tr>
</tbody>
</table>

*These models did not contain the variance function as there was no indication of heteroscedasticity of model residuals

*Presented model statistics are for the SC-GAM with a monotonically increasing functional form. The equivalent model, specified to include a monotonically decreasing function, was equal to the null model, as no decreasing relationship was plausible given the data.
further results on the SC-GAM. For colder than normal conditions (i.e. \(\text{SST}_a < 0^\circ\text{C}\)), the model estimated a near constant and relatively lower average event magnitude, indicating no impact of ocean temperature on marine bird mortality event magnitude within this temperature range. However, above \(\text{SST}_a = 0^\circ\text{C}\) (lower threshold estimate, 95% CI = \([-0.45, 0.28]\)°C; Fig. 3E). Furthermore, between \(\text{SST}_a = 0^\circ\text{C}\) and \(\text{SST}_a = +1^\circ\text{C}\), the model estimated a mean difference of 1.67 (95% CI = \([0.14, 3.28]\)) in log-magnitude. This suggests that mortality event magnitudes averaged 5.3 times higher (95% CI = \([1.1, 26.6]\)) when preceded by prolonged warming above +1°C than when following near climatological average (i.e. \(\text{SST}_a \sim 0^\circ\text{C}\)) conditions.

### 3.3. MHW association analyses

From 1993 to 2021, MHWs, operationally defined as SST assessed at the LME scale exceeding the 90th climatological percentile for a minimum of 6 consecutive days, occurred in clusters of years (Fig. 4A). During this period, MHW days occurred 13.3% of the time throughout the CCLME but were more common to
the north (Gulf of Alaska: 23.6%; Southern Bering Sea: 22.1%; Northern Bering Sea: 18.2%; Fig. 4A). Several MHW events were prolonged, lasting months to years, specifically in 1997–1998, 2003–2005, and 2014–2020 (Fig. 4B).

While multiple mortality events were observed following MHWs (Fig. 4B), they also occurred between MHW episodes (Fig. 4B). However, assessed at the regional scale, the time between mortality events and the most recent MHW was skewed, with half (24 of the 48 Category 2–4 mortality events) occurring within 120 d of a MHW, rising to almost three-quarters (35 of 48) occurring within 1 yr (Fig. 4C). Permutation testing via random re-assortment of mortality events at the regional level suggested that, while not a necessary requirement, mortality events occur more frequently following MHWs, suggestive of an association between these 2 phenomena. There were more regional mortality events within 1 yr of a MHW (observed: 35, 95% range from permutations = [21, 36], p = 0.068), as well as a shorter mean time gap (observed: 340 d, 95% range = [317, 679], p = 0.044) between MHWs and mortality events than expected by chance. This pattern was also true when limiting permutation testing to only the more extreme regional mortality events (Category 3–4; Table S10 in Supplement 6).

3.4. Post MHW impacts on beached bird abundance

To explore potential similarities in ecosystem response to prolonged MHWs, we examined 3 focal MHWs (1997–1998 ENSO, 2014–2016 NE Pacific MHW, 2019 ENSO) within the central California Current (Farallones and Monterey regions) where our data set is most comprehensive. The progression of regional encounter rate anomalies, aligned relative to MHW onset date, bore strong similarities in both the timing of mortality event occurrence and in the following period of depressed carcass encounter rates (negative anomaly values; Fig. 5). Regional mortality events appeared to cluster into 2 distinct time windows following MHW onset: approximately 1–6 mo later, with a second mode at approximately

![Fig. 5. Anomaly in beached bird encounter rate (carcasses km⁻¹) in the central California Current large marine ecosystem following major marine heatwave (MHW) events in (A) 1997–1998, (B) 2014–2015, and (C) 2019 plotted according to time since MHW onset. Plotted points are the model-derived anomaly values for each 28 d time window estimated for the Monterey and Farallones regions; line and shaded regions: anomaly averaged across these 2 regions. All panels are aligned according to MHW onset date (vertical dashed line); dark red and black polygons: MHW and mortality event durations, respectively]
10–16 mo (Fig. 5). Following that, an approximate 16 mo period (~450–500 d) of depressed carcass encounter rates was also consistently observed (Fig. 5). Although the magnitude and exact timing of mortality events and the subsequent depression differed among the 3 MHW episodes, the emergent pattern and durations appeared consistent among the 3 large-scale MHWs examined (Fig. 5). These comparisons suggest that, at least within the CCLME, ecosystem response to a prolonged MHW may include a cascade of marine bird mortality events followed by subsequent reductions in beaching rates collectively lasting nearly 3 yr.

4. DISCUSSION

Our analysis suggests that low-level marine bird mortality events, indicated by significant elevations in carcass beaching rate, are fairly common (Fig. 2A). However, mass mortality events, with order-of-magnitude increases in carcass encounter rate spanning multiple months and regions (i.e. Jones et al. 2018, Piatt et al. 2020; Fig. 2) are rare. We demonstrate that these extreme events are more likely following MHWs (Fig. 4), and that event magnitudes are on average 5 times larger following sustained (annual) SSTs of 0.5–1°C above climatology (1981–2010; Fig. 3E). Our results also suggest a systemic response to MHWs in the CCLME (Fig. 5), consisting of repeated mortality event occurrences followed by an extended period of depressed carcass deposition; a pattern that seemingly spans 3 yr post-MHW.

Our study is also the first to present a synoptic view of the mortality events that occurred in the Northeast Pacific from 2014 through 2019, starting in the CCLME in 2014 (Jones et al. 2018), propagating northwards through the Gulf of Alaska in 2015–2016 (Piatt et al. 2020) and culminating in a succession of events in the Southern and Northern Bering Sea in 2016 through 2019 (Jones et al. 2019, Romano et al. 2020, Van Hemert et al. 2020, Will et al. 2020a; Fig. 2). We believe that this sequence of mortality events is unprecedented in terms of their frequency, extent, and magnitude. Furthermore, these events coincided with multiple ocean-heating phenomena; specifically, the prolonged MHW (Di Lorenzo & Mantua 2016, Lim et al. 2017) of 2014–2016 and transition to a warm stanza accompanied by massive retraction of sea ice in the Bering and Chukchi Seas (Stabeno & Bell 2019). The sequence of mass mortality events from 2014–2019 likely resulted in excess mortality of several million birds.

4.1. Methodological considerations

Prior to 2009 in the Northern Bering Sea and 2006 elsewhere in Alaska, we have no effort-controlled data for Alaska and rely on published accounts (i.e. Piatt & Van Pelt 1997, Baduini et al. 2001) to infer event occurrence (Fig. 2). A similar temporal bias also exists prior to 2001 within the CCLME, with effort-controlled data south of Mendocino only (Fig. 1). Consequently, interpretation of mortality event occurrence and/or scale prior to these epochs should be treated with caution. In particular, the spatial scale of the 1997–1998 mortality event in the CCLME was likely underestimated due to the lack of concurrent effort-controlled monitoring outside central California. Because the 1997–1998 ENSO-driven event coincided with mortality of shearwaters in the Southern Bering Sea (Baduini et al. 2001) as well as common murres in the Gulf of Alaska (Piatt et al. 1999), this may well have been a multi-ecosystem mortality event sequence similar to those observed following the 2014–2016 Northeast Pacific MHW, albeit of a lower magnitude.

Even with the inclusion of opportunistic data from the Northern Bering Sea (Fig. S15 in Supplement 3), lower coverage in this region inhibited our ability to identify mortality events other than those that were large-scale and/or coincident with beached bird monitoring (e.g. the localized puffin mortality event documented on St Paul Island; Jones et al. 2019). Given the temporal extent, geographic breadth, and grain of survey coverage throughout the CCLME relative to Alaska, many more localized events were detected in the CCLME, potentially biasing our results towards over-representing the former. To some degree, regionalization of analyses into approximately 200 km segments of coastline within the CCLME addressed this issue, as very local events (~10s of km) would not register according to our event criteria. While data gaps likely induced spatial biases and differences in spatial scale, we do not believe that they impact the inferred relationships between event occurrence or characteristics and SSTs. We maintain that our main results — the occurrence of an unprecedented sequence of mass mortality events between 2014 and 2019 and the relationships with SSTa and MHWs — are robust to these data gaps.

4.2. Population impact

Marine bird populations can withstand significant annual mortality without declining (Frederiksen et al. 2018).
2008), especially when elevated mortality is limited to pre-reproductive age classes and does not persist over multiple years (Tuck et al. 2015, Johns et al. 2022). The degree to which mass beaching events are representative of increased mortality versus changes in distribution (e.g. Jones et al. 2018) is debatable, although both are a necessary requirement for mortality events with carcass encounter rates that are orders of magnitude above normal. Furthermore, whether mortality observed during any large mortality event is additive versus compensatory is also likely to fall along a continuum. Thus, while sudden, distinct episodes of marine bird mortality can be highly visible, the associated population impacts may in some instances be offset by reduced mortality overall (Frederiksen et al. 2008).

For most marine bird populations, assessing population impacts is only possible on the breeding colonies. Even in high single-event mortality situations, changes in colony size may be obscured by surviving older pre-reproductives and adult age classes otherwise excluded from securing breeding positions (i.e. the floating population) ‘filling up’ newly vacated space (Johns et al. 2022). Despite this obscuring factor, there is evidence that several of the mass mortality events described herein depressed breeding population size and subsequent vital rates.

Following the Cassin’s auklet mortality event in the winter of 2014–2015, Jones et al. (2018) reported a 15% decline in burrow occupancy rates on Triangle Island, British Columbia, the largest Cassin’s auklet colony worldwide. To the south on Southeast Farallon Island, California, Johns et al. (2022) estimated that annual apparent survival in 2014 was 51%, down from 80% on average. While burrow occupancy in 2015 (following the mortality event) was not depressed, auklets shifted their breeding strategy from ‘record high’ double-clutching (40–80%) prior to the event (2010–2014) to less than 10% of pairs engaging in this strategy over the following years (2015–2019). These authors hypothesized that higher than average productivity from 2010–2014 likely increased the floating population, which subsequently occupied available burrows following elevated mortality in 2014.

Subsequent to the short-tailed shearwater mass mortality event in the Southern and Northern Bering Sea in 2019, Glencross et al. (2021) reported delayed arrival of adults to Bruny Island in Southeast Tasmania, Australia, and a curtailed ‘pre-laying exodus’ to the Southern Ocean to forage prior to breeding. Although burrow occupancy was not depressed, breeding success was low (33% compared to 50–75% typically) and nest abandonment was widespread. These authors suggest that foraging stress in the Bering Sea forced adults farther north and left survivors in poorer body condition facing a longer migration.

Finally, and most seriously, leading into and following the extreme die-off of common murres in Alaska in 2015–2016, Piatt et al. (2020) reported severe depression of reproductive success throughout monitored colonies in Alaska (n = 13), with all colonies failing (i.e. colony abandonment) at least once during 2015–2017 and multi-year failures at 8 colonies. Indeed, total breeding failures continued into 2019 at some colonies in the Gulf of Alaska (Schoen et al. 2022, this Theme Section). These changes were concurrent with, or subsequent to, the loss of upwards of one-quarter of all breeders in Alaska during the mortality event, constituting at least 1 million individuals. In 2018, a second mortality event of common murres also occurred, centered in the Northern Bering Sea (Romano et al. 2020). We suggest that when multiple mass mortality events involve a single species, even the most robust populations can become threatened.

4.3. Mechanisms

Our work reinforces studies worldwide indicating that a warmer ocean will be different in terms of community composition (Brown et al. 2010, Gurgel et al. 2020) and ecosystem function (Smale et al. 2019, Smith et al. 2023). The mechanisms by which those changes occur are diverse and complex (Gissi et al. 2021), including both direct and indirect effects (e.g. trophic-mediated changes to grazing pressure; Batten et al. 2022). The range of mortality events contained in this synthesis are indicative of the myriad of mechanistic pathways ending in sudden and dramatic, albeit short-term, increases in marine bird mortality. Bottom-up trophic factors appear to include lower productivity (Beardall et al. 2009, Gao et al. 2021), in some cases associated with phenological shifts in seasonal upwelling (e.g. Parrish et al. 2007), and alterations to zooplankton (e.g. copepods, euphausiids) community composition (Hipfner et al. 2020, Batten et al. 2022), with a concurrent shift in size distribution towards smaller, relatively energy-poor species (Arimitsu et al. 2021, Killeen et al. 2022). However, a host of non-trophic impacts have also been associated with warmer Northeast Pacific and Bering Sea waters, including increases in disease (Bodenstein et al. 2015, Knowles et al. 2019) and the occur-

It is worth noting that beached bird abundance, although certainly an indicator of increased mortality, is not linearly related to population vital rates. This is because a bird that becomes moribund or dies at sea will be scavenged or sink before washing ashore unless the body is already within a narrow strip of coastal ocean (Wiese 2003, O’Hara & Morgan 2006). Therefore, sudden and dramatic increases in carcass deposition rate imply that victims had already made decisions to stay within, or migrate to, the coastal zone. Given that many of the species experiencing mass mortality events in the Northeast Pacific since 2014 were alcids and/or procellariids (Table 2), which have a distinctly offshore distribution post-breeding (e.g. Hunt 1990, Hatch et al. 2000, 2010, Hedd et al. 2012), one interpretation is that deteriorating offshore conditions drove birds onshore. Following the 2014–2016 Northeast Pacific MHW, changes in forage fish abundance (Santora et al. 2020) and quality (Arimitsu et al. 2021) of nearshore species have been documented, which may have provoked inshore movement (Jones et al. 2018). It remains unknown, however, how or even whether offshore prey fields changed.

Finally, the under-representation in these mortality events of several abundant nearshore marine bird taxa, including larids and phalacrocorids, suggests that ecosystem impacts that precipitate mortality events are not uniformly experienced among marine bird taxa. Some taxonomic groups may be relatively buffered against ecosystem shifts due to a combination of foraging behavior (Osborne et al. 2020), metabolic adaptations (Burke & Montevecchi 2018), and/or dietary breadth (Baltz & Morejohn 1977, Will et al. 2020b, Tate et al. 2021). For instance, larids (Blight et al. 2015) and phalacrocorids (Watanuki et al. 2004, Kotzerka et al. 2011) are generalist foragers, which may buffer them against losses of individual prey sources. Will et al. (2020a) found that reductions in the availability of pelagic prey in the Northern Bering Sea from 2016–2019 resulted in higher nutritional stress in common murres relative to congeneric thick-billed murres Uria lomvia, which they attributed to a more diverse diet and benthic foraging habits in the latter species. In summary, different rates of mortality among seabird species may be related to their greater ability to exploit alternate prey than the pelagic species that were severely impacted during recent MHWs (Arimitsu et al. 2021).

4.4. Indicators of change

Several studies have suggested that the North Pacific will be less productive as it continues to warm (Ainsworth et al. 2011, Bryndum-Buchholz et al. 2019), resulting in a lower carrying capacity for epipelagic upper trophics (Woodworth-Jelfcoats et al. 2017), including marine birds. Piatt et al. (2020) proposed that ocean warming creates an ‘ectothermic vise’ for marine birds. Warming reduces bottom-up production but increases metabolism and food demands of forage fish, thereby reducing their size and energy content. Simultaneously, warming also increases the metabolism and food demands of large predatory fish (Holsman & Aydin 2015), increasing top-down competition for forage fish. At the broadest scale, our data suggest that recent warming in the Northeast Pacific has provoked an increase in the frequency and magnitude of marine bird mass mortality events. However, we hypothesize that it is not the absolute value of SST, but rather the abrupt increase in ocean temperature (i.e. following MHW onset, Fig. 5) with concomitant bottom-up impacts on trophic dynamics that provokes a mortality event (Piatt et al. 2020). In some instances, these changes may occur over time scales of weeks (Montevecchi et al. 2021).

Our models hint that following a prolonged MHW, marine bird mortality events spaced at 1–6 mo, and again at 10–16 mo, preceded a relatively quiescent period of less-than-normal deposition in the CCLME. While we are confident that mortality events are an indication of excess mortality in the short term, there are multiple competing hypotheses for lower deposition over subsequent months to years. These include depressed breeding productivity and colonystay at sea distribution, compensatory mortality in adult birds, and/or reductions in population size before affected populations adjust back towards carrying capacity. Several regions in the CCLME appear to display long-term reductions in carcass encounter rate that did not fully recover following the 2014–2016 Northeast Pacific MHW (i.e. Columbia River [Fig. S7] and Juan de Fuca Eddy [Fig. S8 in Supplement 2]), potentially indicative of population decline. As MHW frequency increases (Benedetti-Cecchi 2021), we posit that the chance of compounding effects will increase as marine bird populations fail to fully recover between MHW episodes.
4.5. Conclusions

While marine bird mortality events are one manifestation of ecosystem change, there is still much uncertainty regarding causal mechanisms and the degree of population impact. Nevertheless, the occurrence of multiple sequential mass mortality events in the Northeast Pacific and Bering Sea between 2014 and 2019 represents an order of magnitude increase in their frequency, given the occurrence of 1–2 events of a similar magnitude in the 20 yr prior (Table 2). Given that these events likely include some of the largest marine bird mortality events that have ever been documented, and that each was associated with changes in population vital rates, it is likely that the abundance of North Pacific marine birds was reduced by millions of birds as a result. Thus, the question now becomes whether carrying capacities will be absolutely lower given chronic forcing towards warmer conditions as a baseline rather than an anomaly (Benedetti-Cecchi 2021); and secondarily, whether the frequency of marine bird mortality events will slow should the rate of warming decelerate—an admittedly optimistic view given the continued increase in greenhouse gas emissions globally.

Acknowledgements. A data set of this scale would be impossible to collect and maintain without the efforts of thousands of volunteer participants and interns over 2 decades and across 4 citizen-science programs: COASST (University of Washington), BeachCOMBERS (Moss Landing Marine Laboratories), Beach Watch (Greater Farallones Association), and BC Beached Bird Survey (Birds Canada). Special mention to Mireya Bejarano and Kathya Argueta (US Fish and Wildlife Service, USFWS) for their efforts to prepare the BeachCOMBERS data set for these analyses. Funding for the Beach Watch project was provided by NOAA–Greater Farallones National Marine Sanctuary and Greater Farallones Association. We thank the National Parks Service in Alaska for their contributions to COASST annual data sets, with special thanks to contributors David Payer, Letty Hughes, Raime Fronstin, Mary Hake, and Pam Sousanes for their support of seabird surveys in the Arctic parks. Similar thanks go to annual summer surveyors across remote areas within the Alaska Maritime National Wildlife Refuge. Annual contributions of data and survey development in the Pribilof Islands would not have been possible without the efforts of Pamela Lestenko, Paul Melovidov, and Aaron Lestenko of the Ecosystem Conservation Office (ECO). In addition, the opportunistic reports collected in Alaska from 2017–2021 by the USFWS are the result of dedicated efforts by a large group of federal, state, community, and tribal partners, including Brandon Almasuk (Kawerak Inc.), Gay Sheffield and Melissa Good (AK Sea Grant), Sarah Schoen (USGS), and Elizabeth Labunska (USFWS). We thank the 3 anonymous reviewers, and content editor (J. Piatt) for taking the time and effort necessary to review and provide constructive comments on this work. The overall quality of the manuscript was much improved following their feedback. The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the USFWS.

LITERATURE CITED


Barkhordarian A, Nielsen DM, Baehr J (2022) Recent marine heatwaves in the North Pacific warming pool can be attributed to rising atmospheric levels of greenhouse gases. Commun Earth Environ 3:131


Jones et al.: Ocean warming and seabird mortality events


Tate HM, Studholme KR, Domalik AD, Drever MC and others (2021) Interannual measures of nutritional stress during a marine heatwave (the Blob) differ between two North Pacific seabird species. Conserv Physiol 9:coab090


Editorial responsibility: John F. Piatt (Guest Editor), Anchorage, Alaska, USA
Reviewed by: W. Montevecchi and 2 anonymous referees

Submitted: November 16, 2022
Accepted: May 15, 2023
Proofs received from author(s): June 30, 2023