Supplementary materials

Supplement 1: Survey information

Table S1. Compiled information from monthly effort standardised beached bird survey programs used in this manuscript. Presented are the total number of surveys and beach segments surveyed by each program, where BCBBS and COASST are also split into regions (outer coast of Washington, Oregon, California, vs inside waters of the Salish Sea, and Alaska). In addition, total carcass counts, and gross taxonomic composition are presented for major marine bird families included in our analyses.

	BCOMB	BWATCH	COASST (outer: CA-WA)	BCBBS (outer: VI)	COASST (Salish Sea)	BCBBS (Salish Sea)	COASST (AK)	**TOT**
Surveys	5045	22214	25105	807	27424	6578	6588	**93761**
Beaches	28	68	316	19	363	134	159	**1087**
Est.	May-97	Sep-93	Dec-99	Aug-02	Jan-99	Aug-02	Aug-05	
Bird finds								
Alcidae	17057	20411	38954	92	718	202	6832	**84266**
	[34.2]	[39.5]	[52.4]	[22.5]	[29.3]	[20.0]	[67.5]	**[44.4]**
Procellaridae	11319	8303	12997	186	83	8	1498	**34394**
	[22.7]	[16.1]	[17.5]	[45.5]	[3.4]	[0.8]	[14.8]	**[18.1]**
Diomedeidae	25	24	267	1	1		6	**324**
	[0.1]	[0.0]	[0.4]	[0.2]	[0.0]		[0.1]	**[0.2]**
Hydrobatidae	64	56	379	2	3		131	**635**
	[0.1]	[0.1]	[0.5]	[0.5]	[0.1]		[1.3]	**[0.3]**
Laridae	5537	9861	10974	75	950	498	1481	**29376**
	[11.1]	[19.1]	[14.7]	[18.3]	[38.8]	[49.4]	[14.6]	**[15.5]**
Phalacrocoracidae	7183	5387	3860	20	189	62	68	**16769**
	[14.4]	[10.4]	[5.2]	[4.9]	[7.7]	[6.2]	[0.7]	**[8.8]**
Sulidae	1	5	5		1			**12**
	[0.0]	[0.0]	[0.0]		[0.0]			**[0.0]**
Pelecanidae	1031	932	489	1	1			**2454**
	[2.1]	[1.8]	[0.7]	[0.2]	[0.0]			**[1.3]**
Podicipedidae	4938	4298	2874	2	148	32	9	**12301**
	[9.9]	[8.3]	[3.9]	[0.5]	[6.0]	[3.2]	[0.1]	**[6.5]**
Gaviidae	1419	990	819	5	142	33	15	**3423**
	[2.8]	[1.9]	[1.1]	[1.2]	[5.8]	[3.3]	[0.1]	**[1.8]**
Anatidae (Tribe:	1154	1289	2344	5	209	171	81	**5253**
Mergini)	[2.3]	[2.5]	[3.2]	[1.2]	[8.5]	[17.0]	[0.8]	**[2.8]**
Scolopacidae	95	124	440	20	4	2	2	**687**
(Genus:	[0.2]	[0.2]	[0.6]	[4.9]	[0.2]	[0.2]	[0.0]	**[0.4]**
Phalaropus)								
All marine birds	**49823**	**51680**	**74402**	**409**	**2449**	**1008**	**10123**	**189894**
Species count (marine)	**64**	**80**	**83**	**27**	**56**	**38**	**57**	**106**

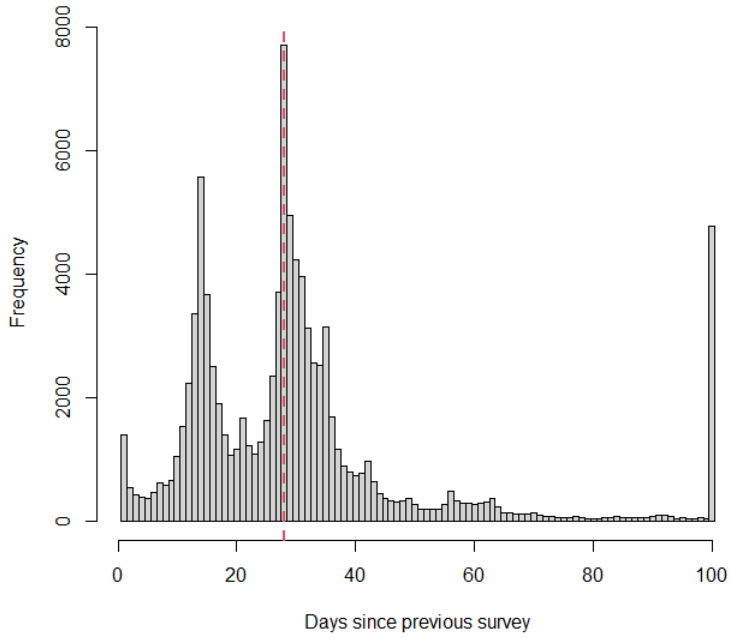

Figure S1. Histogram illustrating the distribution of survey intervals (number of days since the previous survey on that particular beach segment) present in the beached bird dataset. Survey intervals greater than 100 days, along with first surveys per beach segment, are assigned a value of 100 days such that all surveys can be shown on a reasonable scale. The vertical red-dashed line shows a survey interval of 28-days, which was the most prevalent interval present in the combined dataset.

CCLME regional divisions

Table S2. Beach counts by year among California Current regional divisions, along with information on carcass encounter rate (ER, birds per km), timing of peak ER and gross taxonomic composition in each region.

Year	Monterey	Farallones	Mendocino	Humboldt	S Oregon	Columbia River Plume	Juan de Fuca Eddy	Strait of Juan de Fuca	Puget Sound	Strait of Georgia
Beach count										
1993		26								
1994		37								
1995		44								
1996		48								
1997	7	49								
1998	9	45								
1999	7	41				4		2		
2000	7	44				9	1	5		2
2001	15	41			6	17	18	16	1	10
2002	12	43			4	28	25	31	8	32
2003	19	41			4	21	23	39	13	51
2004	18	40			3	23	24	41	35	56
2005	20	40			11	30	25	47	47	48
2006	23	40		6	13	33	31	55	58	49
2007	23	40		16	15	45	28	50	56	41
2008	22	40		16	17	43	27	45	49	35
2009	21	40		26	20	56	34	57	57	65
2010	22	40		29	21	50	30	64	82	84
2011	23	39		28	25	59	28	63	113	87
2012	23	40		37	33	66	35	74	117	121
2013	24	40		33	45	80	32	76	107	120
2014	24	48	4	43	48	85	36	74	106	115
2015	25	52	9	51	46	74	38	76	95	112
2016	26	53	10	43	47	70	34	77	82	109
2017	26	53	9	33	40	72	35	64	59	92
2018	25	54	13	35	45	63	27	54	56	98
2019	24	54	13	36	38	64	23	56	66	94
2020	22	55	11	35	34	54	18	53	52	94
2021	16	53	11	41	43	59	17	49	54	84
Bulk beached bird information										
Mean ER	2.38	1.18	2.25	2.18	2.71	2.32	0.98	0.11	0.06	0.10
Median ER	1.30	0.31	0.00	1.00	1.08	0.68	0.00	0.00	0.00	0.00
Peak Month	Nov	Sep	May	Aug	Sep	Dec	Sep	Aug	Oct	Aug
	(3.0)	(2.1)	(5.2)	(4.3)	(4.8)	(4.8)	(3.5)	(0.3)	(0.2)	(0.4)
Species 1	comu	comu	comu	comu	comu	comu	comu	rhau	gull	gull
Species 2	nofu	nofu	gull	gull	nofu	nofu	nofu	gull	comu	comu
Species 3	brac	wegu	nofu	wegr	caau	caau	caau	comu	rhau	wwsc

* Species codes: comu = common murre, nofu = northern fulmar, brac = Brandt's cormorant, wegu = western gull, gull = large gull complex, wegr = western/Clark's grebe, caau = Cassin's auklet, rhau = rhinoceros auklet, wwsc = white-winged scoter

To ascertain whether neighboring California Current regions should be combined, we fitted two models to the combined data across neighboring regions. The first model was similar in form to that described in the main text:

$$\ln(\mu_{i,j}) = \beta_0 + s(d_i) + \delta_{t_i} + \epsilon_j + \ln(L_{i,j}) \quad \text{[eqn. S1]}$$

$$C_{i,j} \sim NB(\mu_{i,j}, \theta) \quad \text{[eqn. S2]}$$

Where $C_{i,j}$ is the observed carcass count on survey i made on beach segment *j*, $\mu_{i,j}$ is the expected mean count, β_0 is the overall intercept, $s(d_i)$ is the seasonal smooth (cyclic-cubic) term of day of year of survey *i*, d_i, δ_{t_i} is the random effect of time-period t_i (i.e. anomaly) within which the survey was performed, ϵ_j is the random effect of beach segment, *j*, $L_{i,j}$ is the length of beach surveyed, and θ is the negative binomial dispersion parameter. The second model had the same structure, modified such that terms handling seasonality, $s(d_i)$, and anomalies through time, δ_{t_i}, were estimated separately for each of the two regions being compared, whereas in the first model these terms were the same for each region. The shared effects, and region-specific effects models were fitted using mixed effects generalised additive models (GAMMs) in the mgcv package in R (Wood 2017), and were subsequently compared based on AICc. We conclude that splitting neighboring regions was supported if the region-specific model had a lower AICc value than the shared effects model. In all cases within the California Current LME, the region-specific model was an improvement over the shared effects model based on AICc model comparisons, whereas regional sub-division was not supported in the Gulf of Alaska (**Table S3**).

Table S3. Statistics comparing models with shared effects (seasonality and departures from baseline) among regions to a model containing region-specific effects. Each row corresponds to a neighboring pair of regions that were combined for a specific test of whether regional-effects were supported compared to a shared set of effects.

Region 1	Region 2	Model AICc		Δ_{AICc}
		Shared	Regional	
California Current				
Monterey	Farallones	99134.9	98764.8	-370.1
Farallones	Mendocino	73470.2	73412.0	-58.2
Mendocino	Humboldt	17392.0	17368.3	-23.7
Humboldt	S Oregon	34994.4	34295.6	-698.8
S Oregon	Columbia Plume	59466.6	59102.1	-364.5
Columbia Plume	JDF Eddy	52370.5	52062.1	-308.4
Strait of JDF	Strait of Georgia	13670.9	13596.7	-74.2
Puget Sound	Strait of Georgia	10769.6	10750.2	-19.4
Gulf of Alaska				
Kenai Fiords	Cook Inlet	4212.4	4251.6	39.2
Kenai Fiords	Kodiak	2401.0	2416.7	15.7
Cook Inlet	Kodiak	4020.3	4028.3	8
All GoA regions		5293.8	5348.4	54.6

For each region, data were then split into constituent regions, and each region was then tested to identify whether additional model components of beached bird program and survey interval were supported. Both terms were included as fixed effects. Because survey interval was undefined for first surveys on a beach segment, and because some surveys had recommenced on a beach segment that had not been surveyed for 40 days or more, it was not possible to include interval as a continuous predictor without omitting many potentially informative surveys. Therefore, survey interval was coded as a factorial term, with levels

- interval: ≤ 7-days
- interval: 8-20 days
- interval: 21-39 days
- interval: ≥ 40 days

informed by the relative frequency distribution of survey interval (**Figure S1**) where first/recommenced surveys after a break in survey coverage of more than 40 days were grouped into the final category.

Each model consisted of the core model components (outlined in the main text), in addition to these additional factors of program, interval, or both. The best model among those trialled was identified based on minimizing AICc (**Table S4**).

Table S4. Comparison among model structures for baseline model fitting.

Region	**Model AICc values**			
	Base model	**+ interval**	**+ program**	**+ interval + program**
Monterey	26388.6	26362.8		
Farallones	71988.6	71925.6		
Mendocino	1469.7	1471.5		
Humboldt	15902.3	15903.8		
S Oregon	18413.0	18386.9		
Columbia Plume	40948.1	40701.7		
JDF Eddy	11324.5	11319.8	11319.7	11315.0
Strait of JDF	7181.2	7184.7	7181.8	7185.3
Puget Sound	4332.0	4325.6		
Strait of Georgia	6412.1	6416.7	6411.9	6416.5
Gulf of Alaska	5291.6	5293.5		
S Bering Sea	1812.2	1812.8		
N Bering Sea	1133.3	1138.4		

Supplement 2: Baseline model definition and diagnostic plots

Below are presented the full equations, and parameter definitions (**Table S5**) for the models fitted to beached bird monitoring data. Each model was a negative binomial 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 within a region (i.e., equations 1-2 in Supplement 1). Within this model, random effects are normally distributed on log-scale (log-link), such that

$$\epsilon_j \sim N(0, \sigma_\epsilon) \quad \text{[eqn. S3]}$$

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

$$\delta_t = \rho\delta_{t-1} + e_t \quad \text{[eqn. S4]}$$

$$e_t \sim N(0, \sigma_t) \quad \text{[eqn. S5]}$$

where ρ is the lag-1 autocorrelation strength, and e_t are the residuals of this process, assumed to be normally distributed with standard deviation σ_t. In order to estimate temporal random effects, δ_t, along with autocorrelation (ρ) and variability (σ_t), we defined the set of random effects as deriving from a multivariate normal distribution centered on zero, and variance covariance matrix Σ

$$\delta_t \sim MVN(0, \Sigma) \quad \text{[eqn. S6]}$$

$$\Sigma_{i,j} = \rho^{|i-j|}\frac{\sigma_t^2}{1-\rho^2} \quad \text{[eqn. S7]}$$

where $\rho^{|i-j|}$ models the decay in correlation strength between observations that are $i - j$ time periods apart. Specifying the random effects in this way allowed for anomalies to persist through time, as opposed to the assumption that neighbouring time periods were independent.

Table S5. Model parameter descriptions, symbols, and priors used when fitting baseline models using MCMC estimation in STAN.

Parameter	Symbol	Prior	Constraints
Core parameters			
Intercept	β_0	~ N(0, 100)	
Seasonal smooth parameters	β_{1-12}	~ N(0, 100)	
Inter-beach standard deviation	σ_ϵ	~ N(0, 100)	lower: 0
Inter-period standard deviation	σ_t	~ N(0, 100)	lower: 0
Inter-period correlation	ρ	~ N(0, 100)	lower: -1, upper: 1
Negative binomial dispersion	θ	~ N(0, 100)	lower: 0
Optional parameters			
Program-specific intercepts	$\beta_{0,P}$	~ N(0, 100)	
Survey-interval effects	$\beta_{0,I}$	~ N(0, 100)	

Monterey

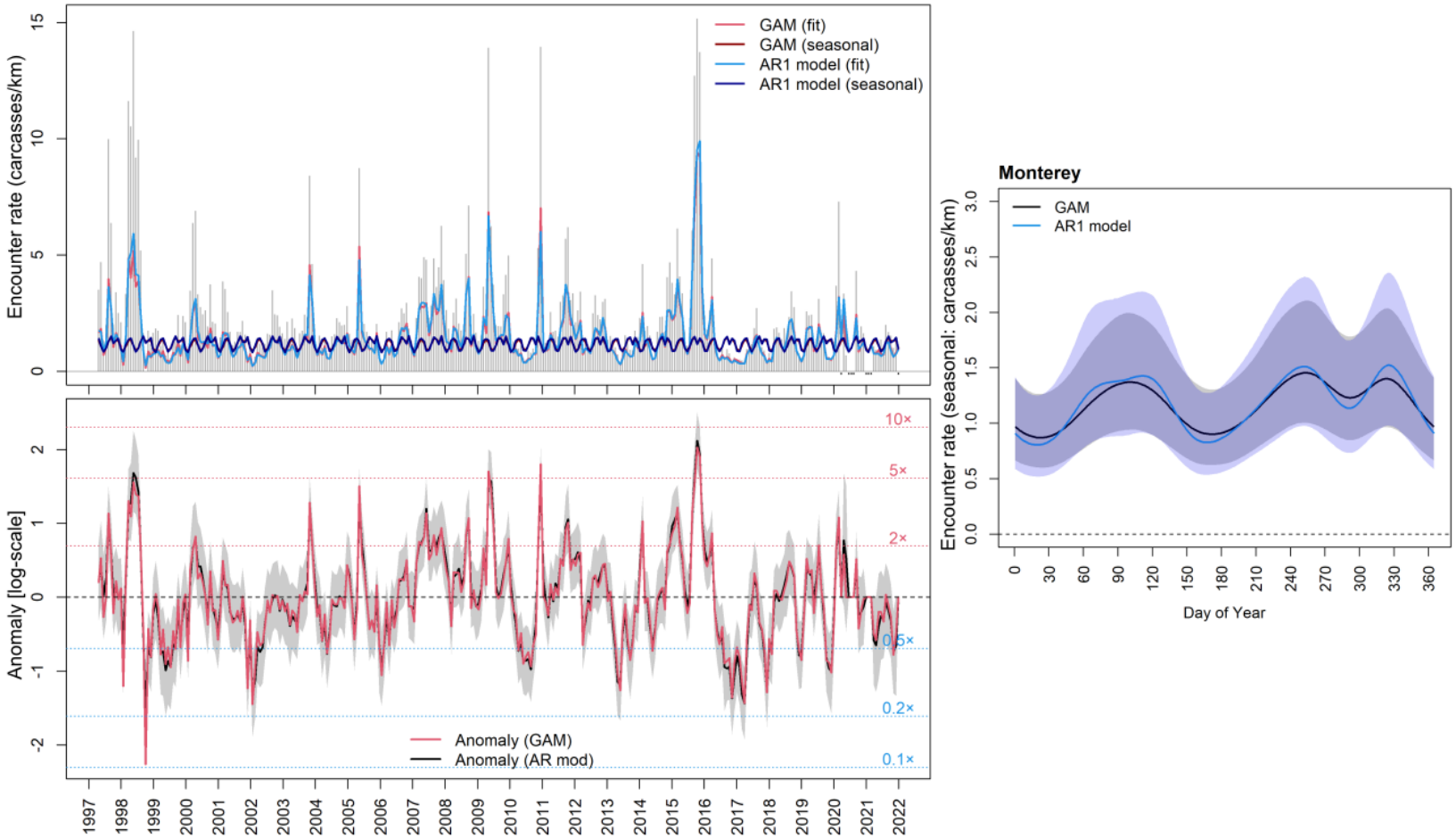

Figure S2. Average encounter rate (ER) for the Monterey region, 1997-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Farallones

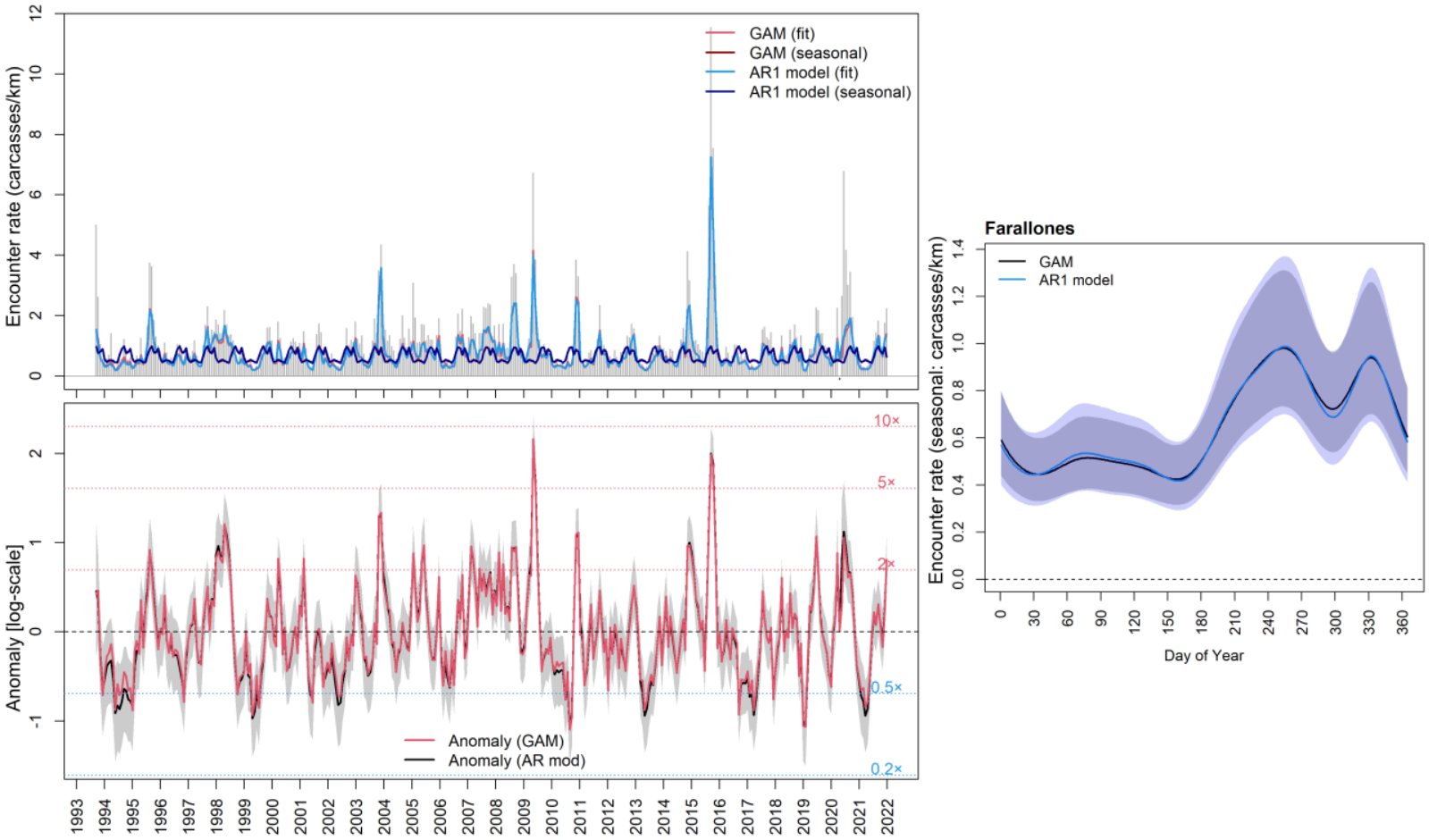

Figure S3. Average encounter rate (ER) for the Farallones region, 1993-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Mendocino

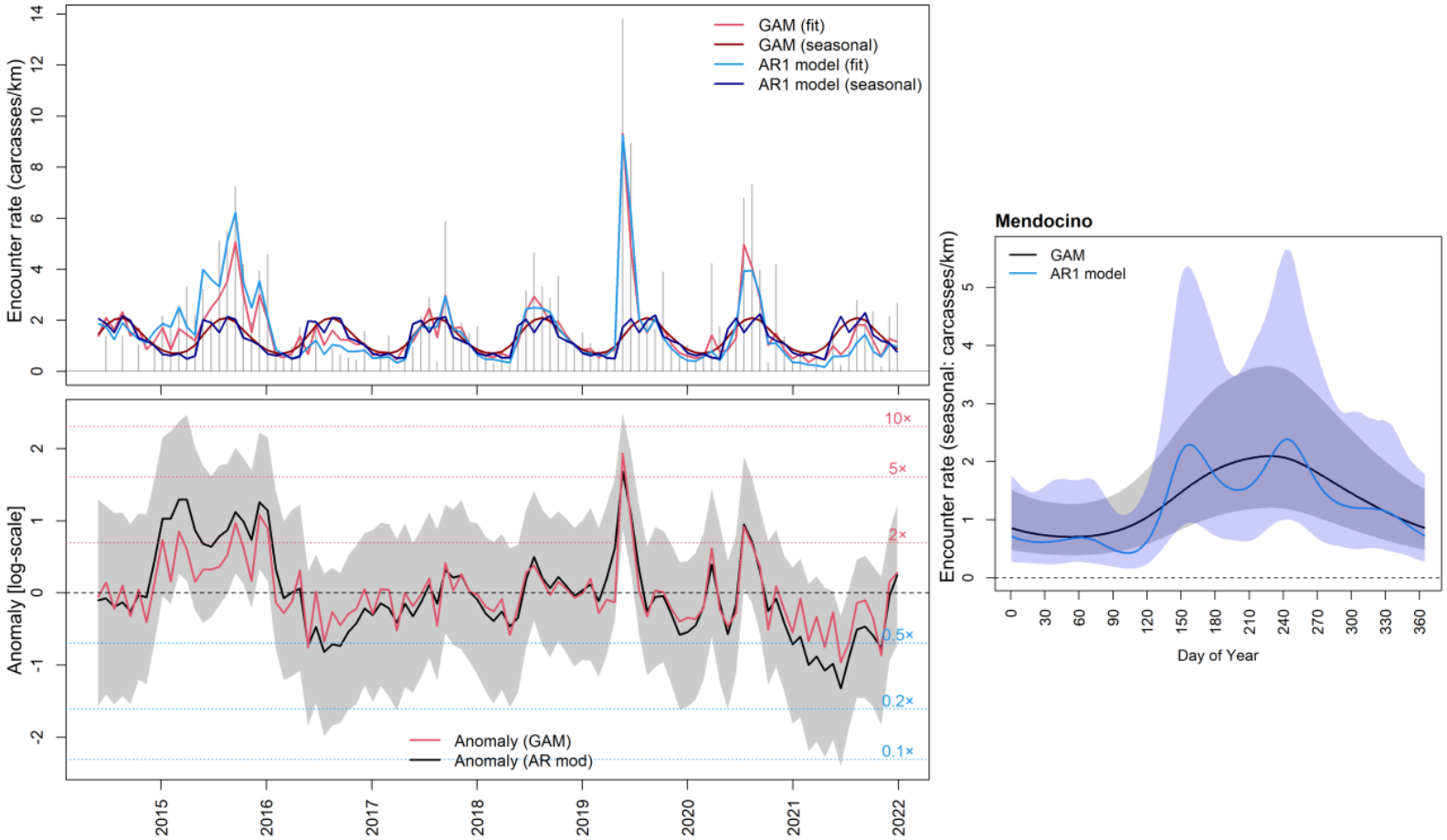

Figure S4. Average encounter rate (ER) for the Mendocino region, 2014-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Humboldt

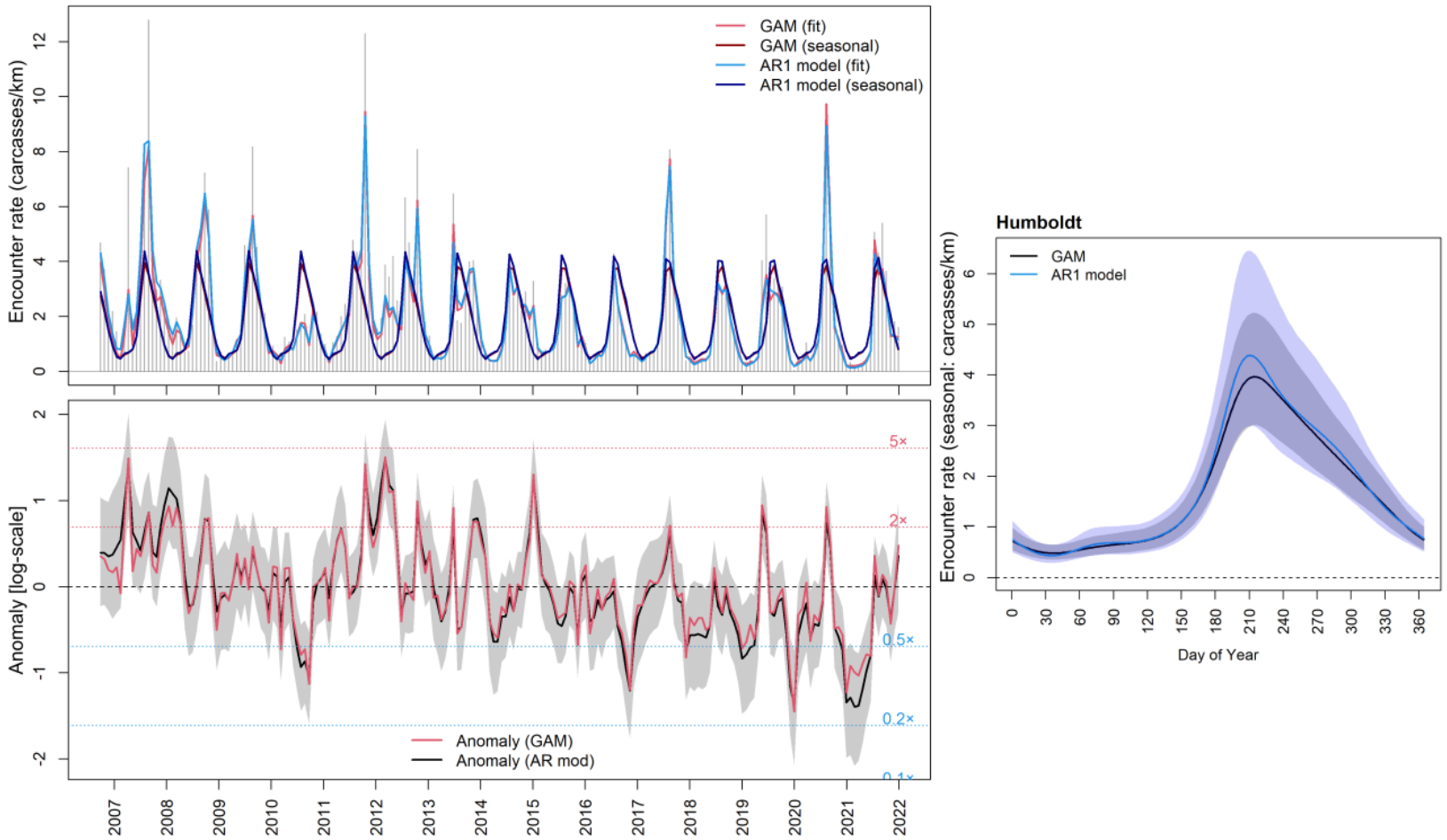

Figure S5. Average encounter rate (ER) for the Humboldt region, 2007-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

S Oregon

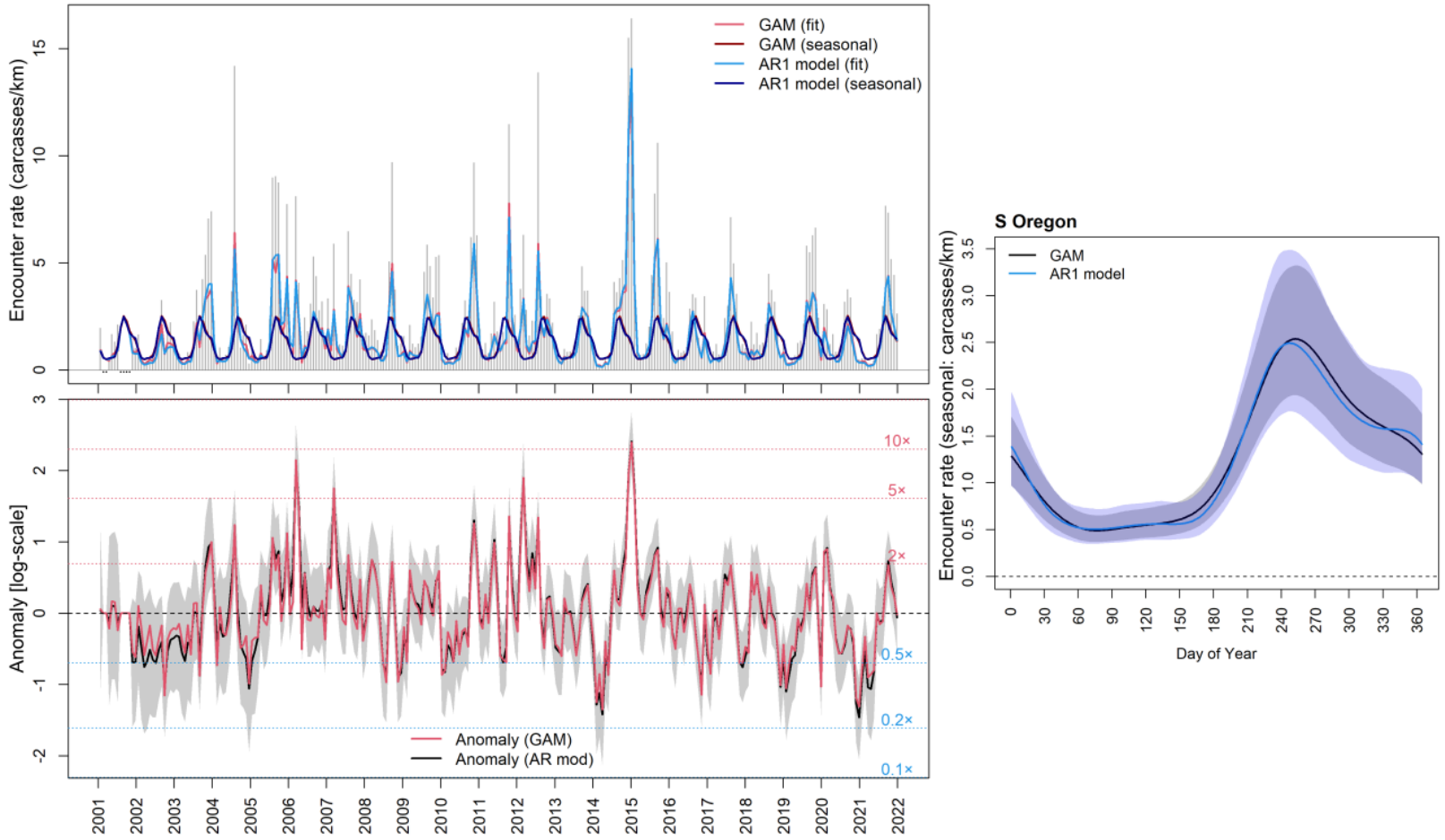

Figure S6. Average encounter rate (ER) for the Southern Oregon, 2001-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Columbia River Plume

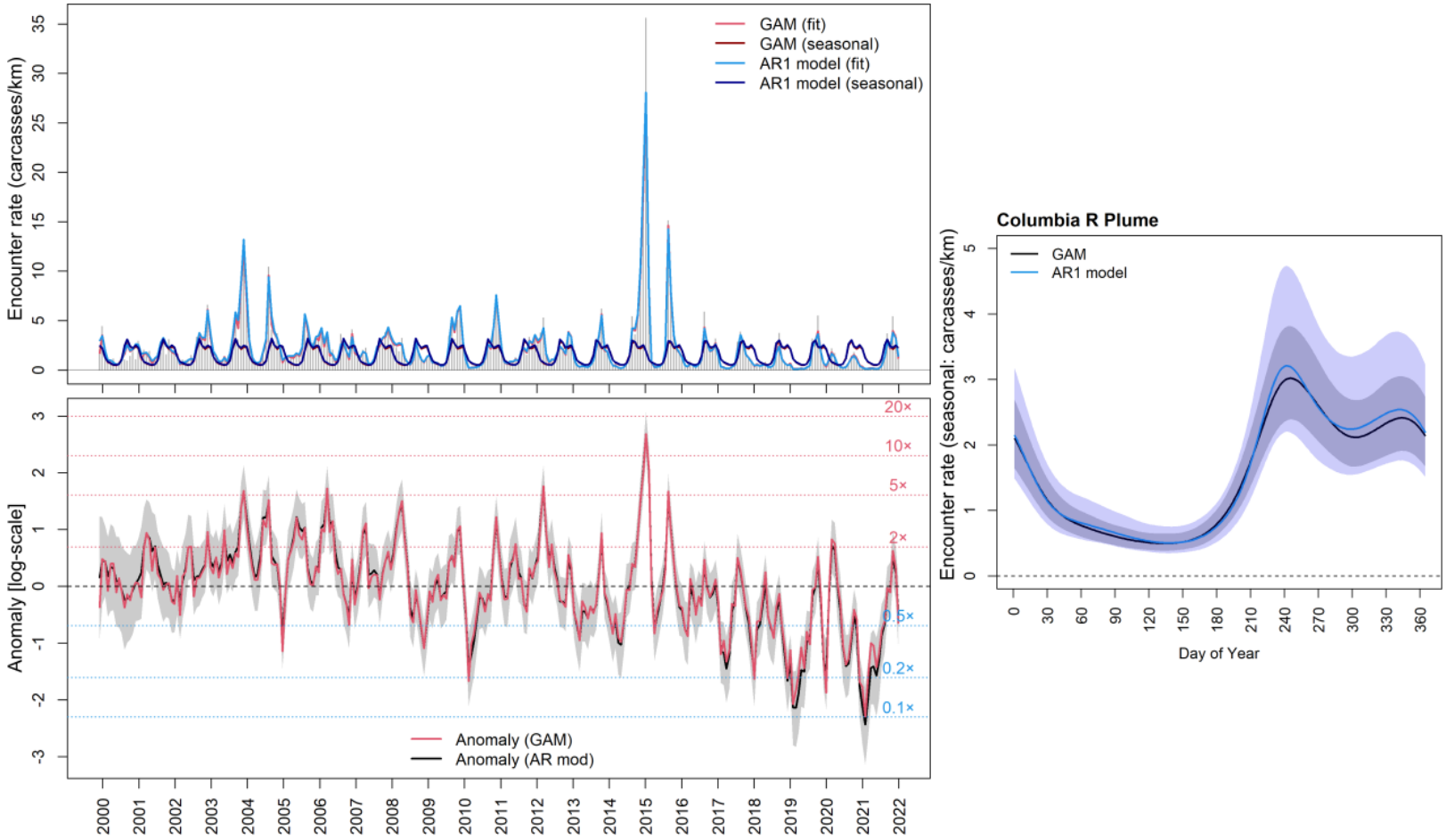

Figure S7. Average encounter rate (ER) for the Columbia River Plume region, 2000-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Juan de Fuca Eddy

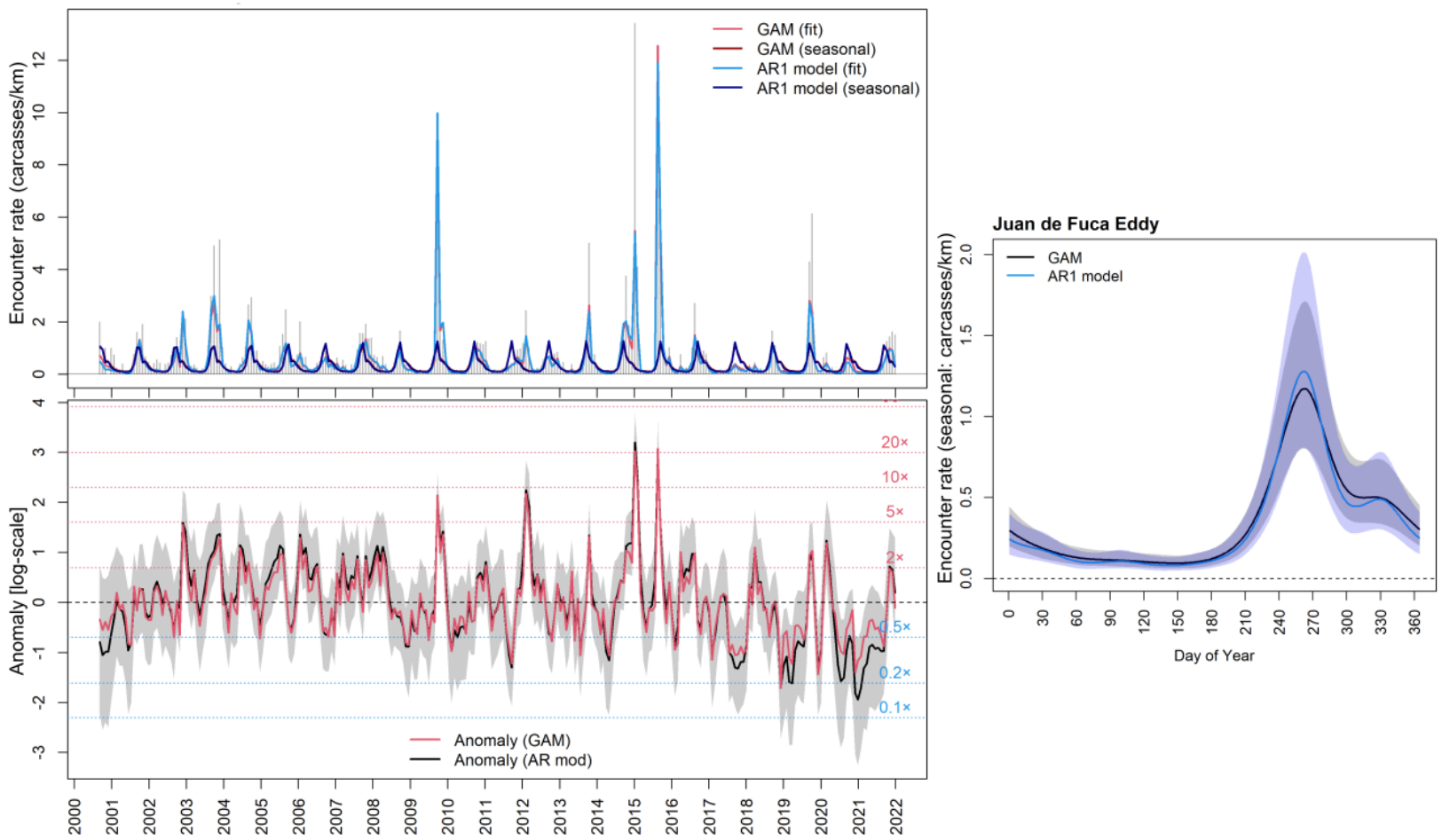

Figure S8. Average encounter rate (ER) for the Juan de Fuca Eddy region, 2000-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Strait of Juan de Fuca

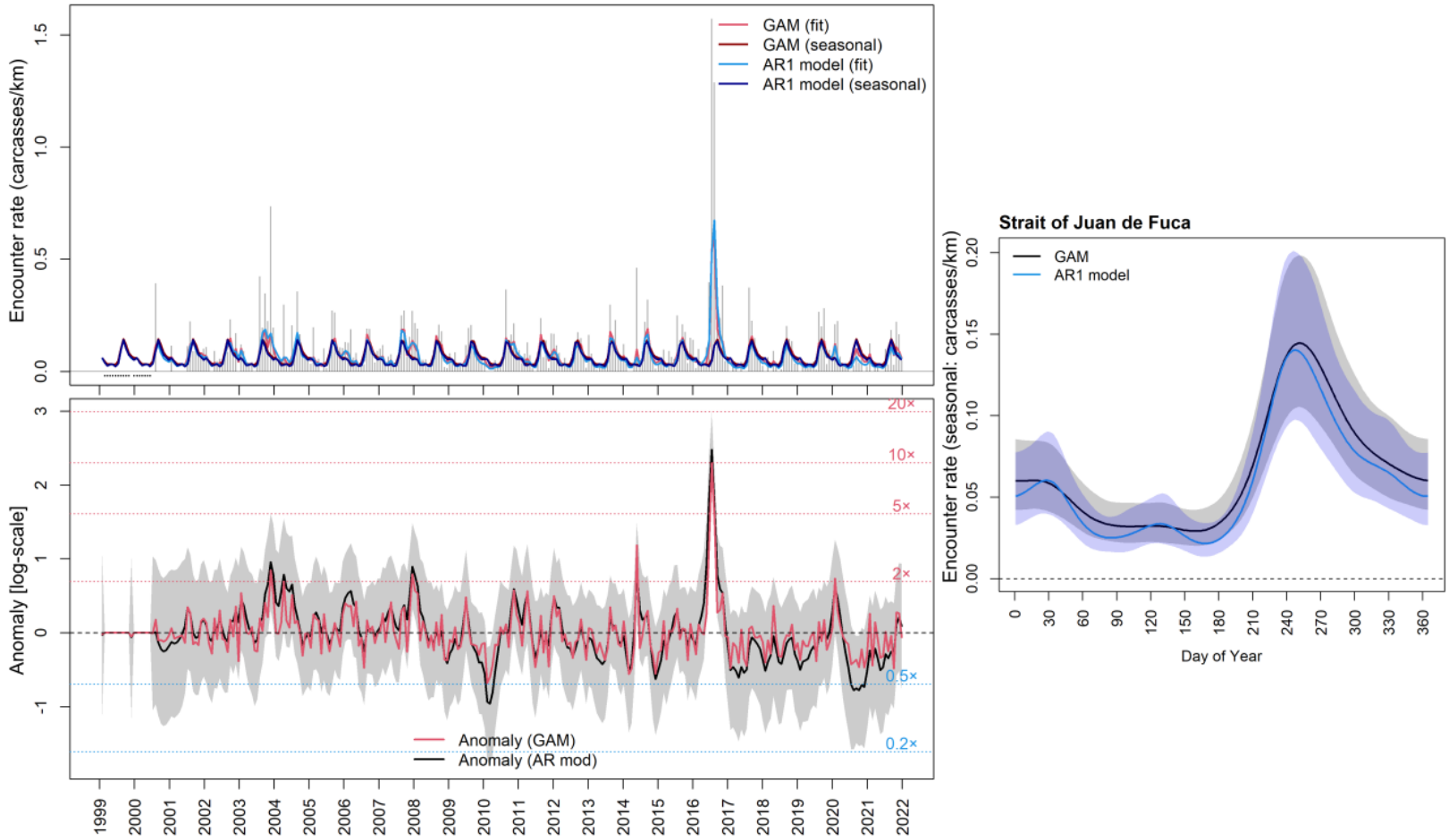

Figure S9. Average encounter rate (ER) for the Strait of Juan de Fuca region, 2000-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Puget Sound

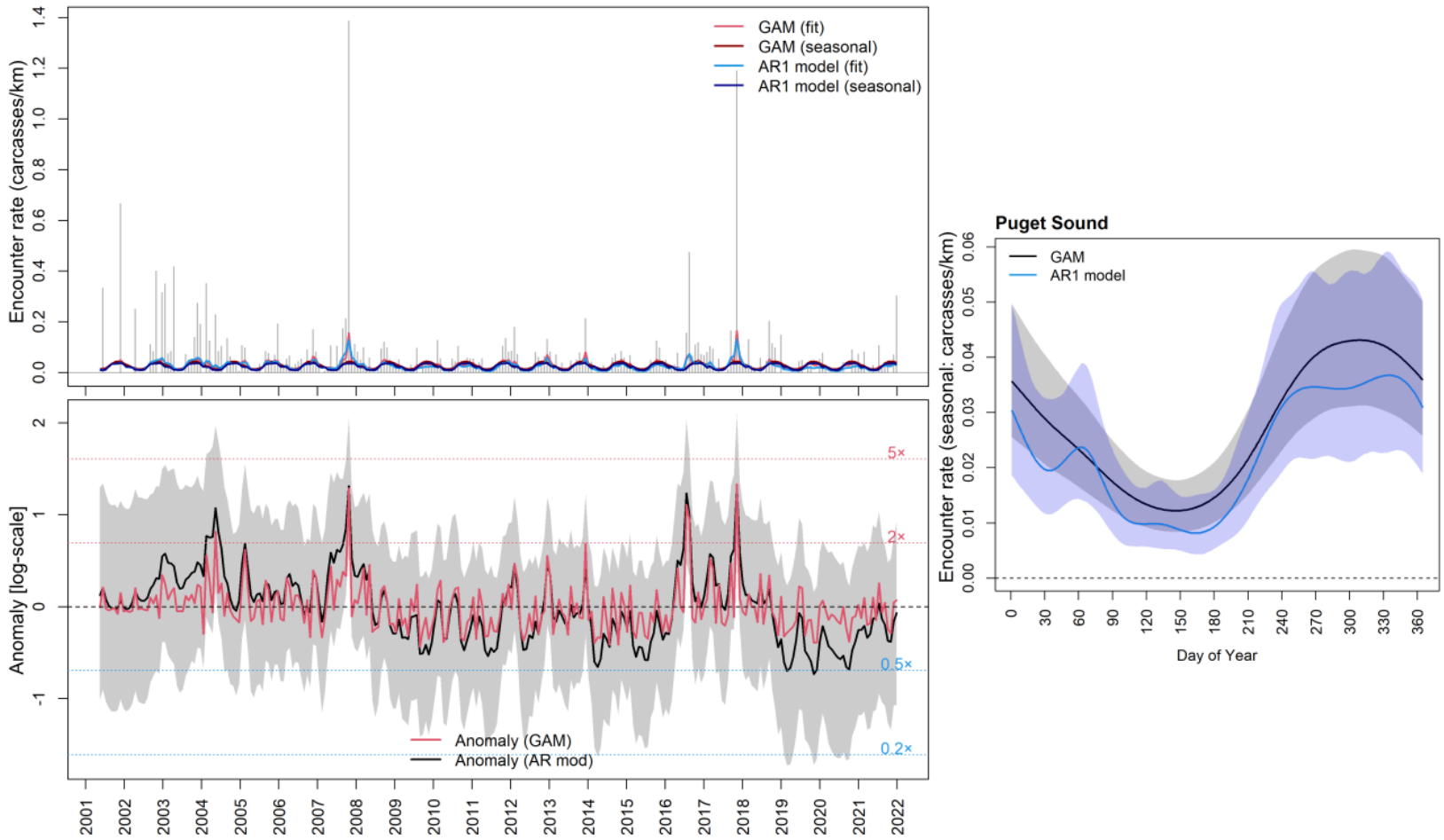

Figure S10. Average encounter rate (ER) for the Puget Sound region, 2000-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Strait of Georgia

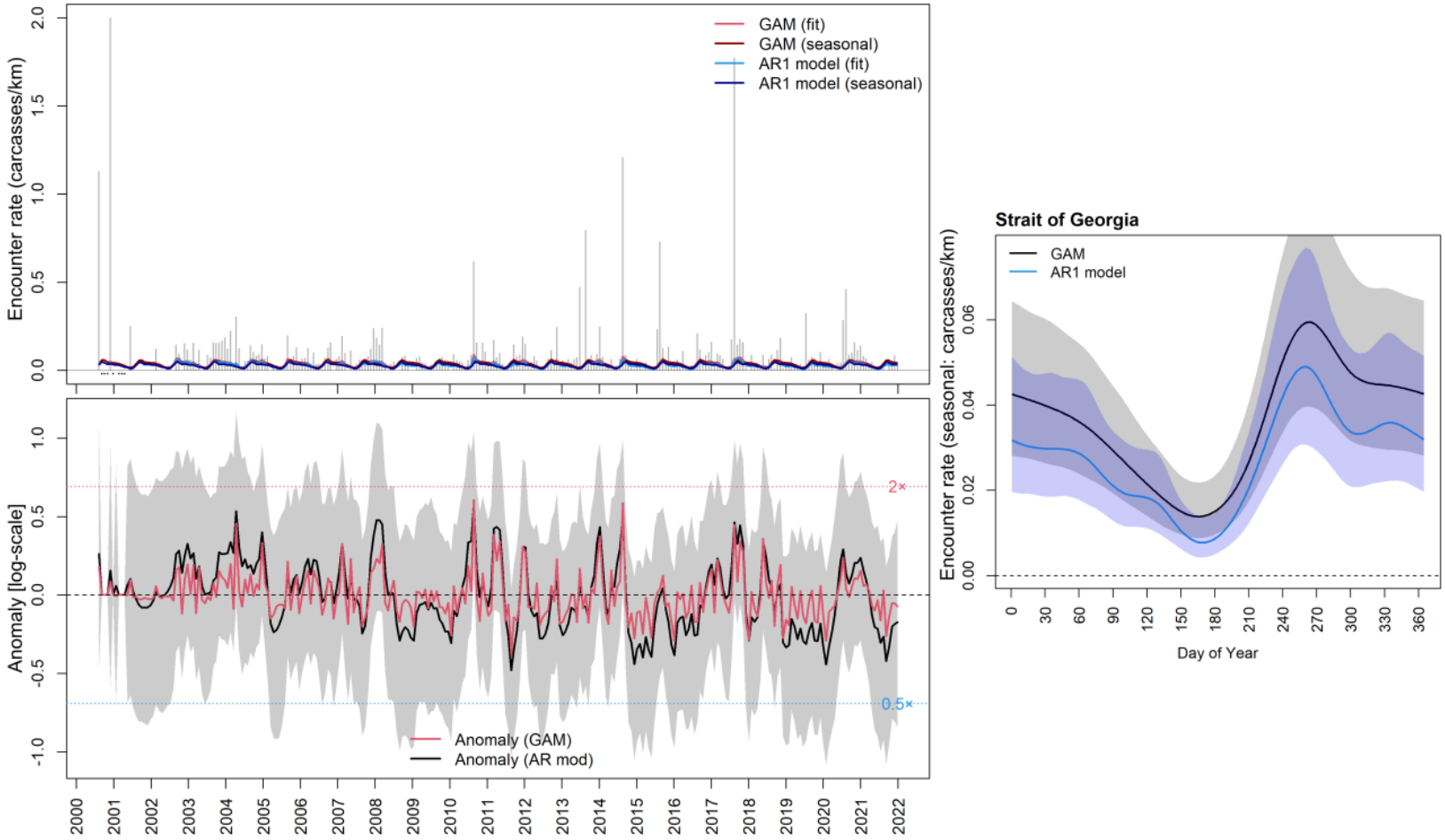

Figure S11. Average encounter rate (ER) for the Strait of Georgia region, 2000-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Gulf of Alaska

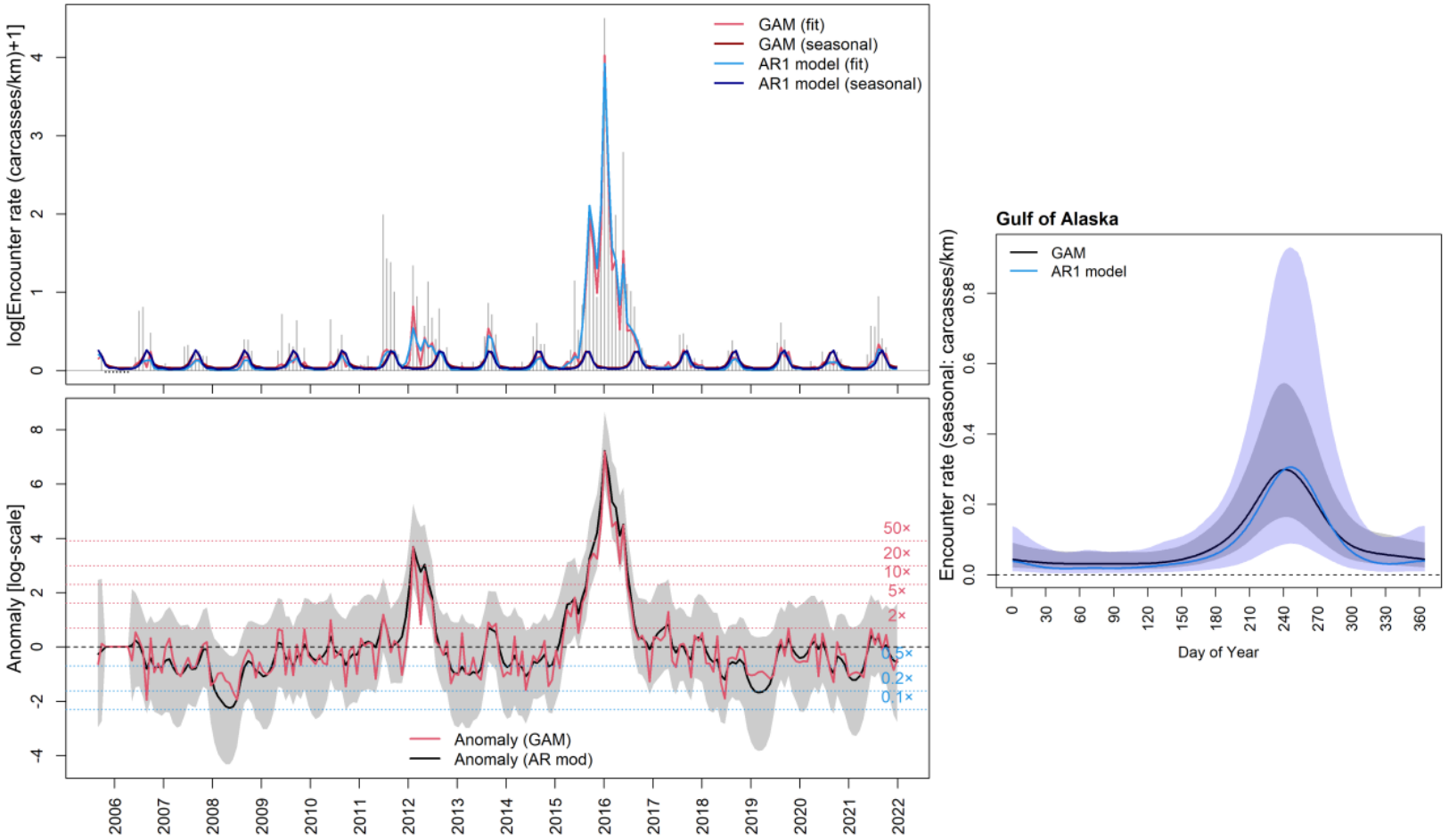

Figure S12. Average encounter rate (ER) for the Gulf of Alaska, 2006-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year. Note: the y-axis on the upper plot is log(x+1) of the encounter rate for the purposes of visualization.

Southern Bering Sea

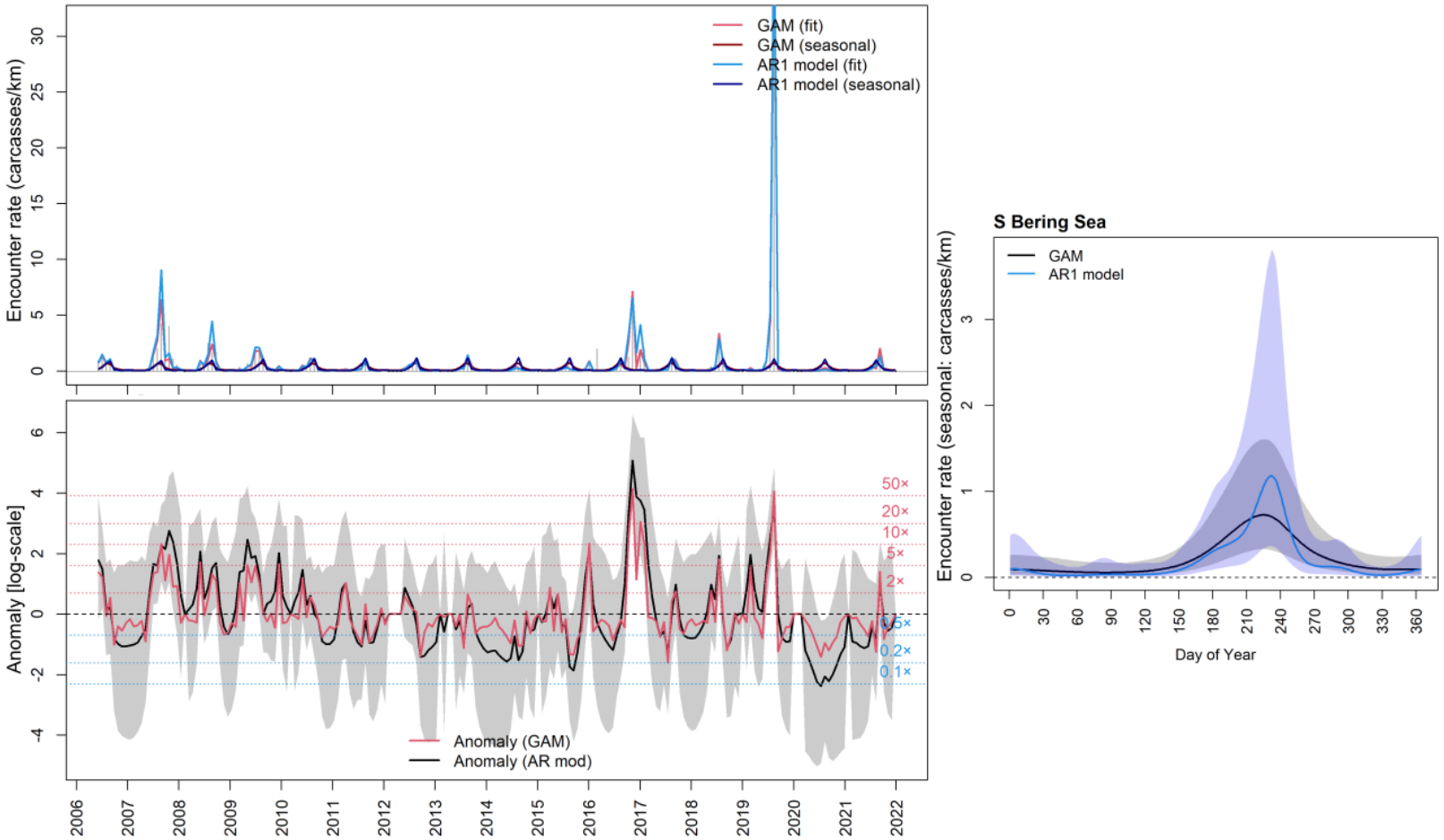

Figure S13. Average encounter rate (ER) for the Southern Bering Sea, 2006-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Northern Bering Sea

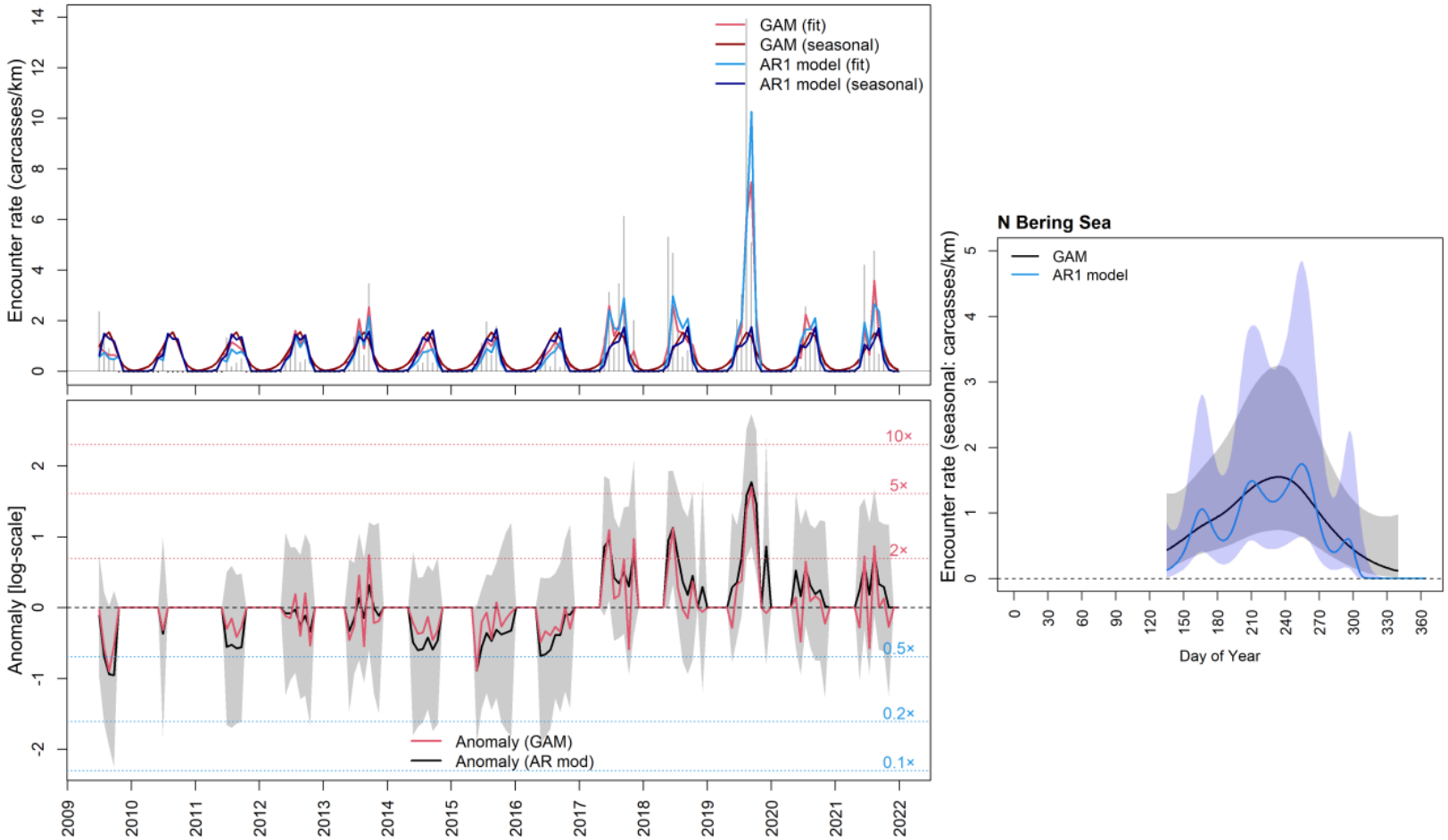

Figure S14. Average encounter rate (ER) for the Northern Bering Sea, 2009-2021 (upper), along with model estimated time-series of 28-day anomalies (lower). In each panel, fitted values are shown for models with temporally autocorrelated anomalies (AR1) versus uncorrelated anomalies (GAM).The upper plot shows observed mean ER (grey bars) averaged across all surveys within each 28-day period, overlaid by model-fitted values including only seasonal terms, or seasonal plus 28-day anomalies (fit). The lower panel shows estimated 28-day ER anomalies, indicating departure from seasonal baseline through time (grey shading: 95% CI from the AR1 model). Given the log-link function, anomaly values are additive on the log-scale (as plotted), but multiplicative on the response (encounter rate) scale. Equivalent multiplier rates are shown by horizontal dashed lines. The right panel shows estimated seasonal (baseline) patterns in ER as a function of day of year.

Supplement 3: Opportunistic information

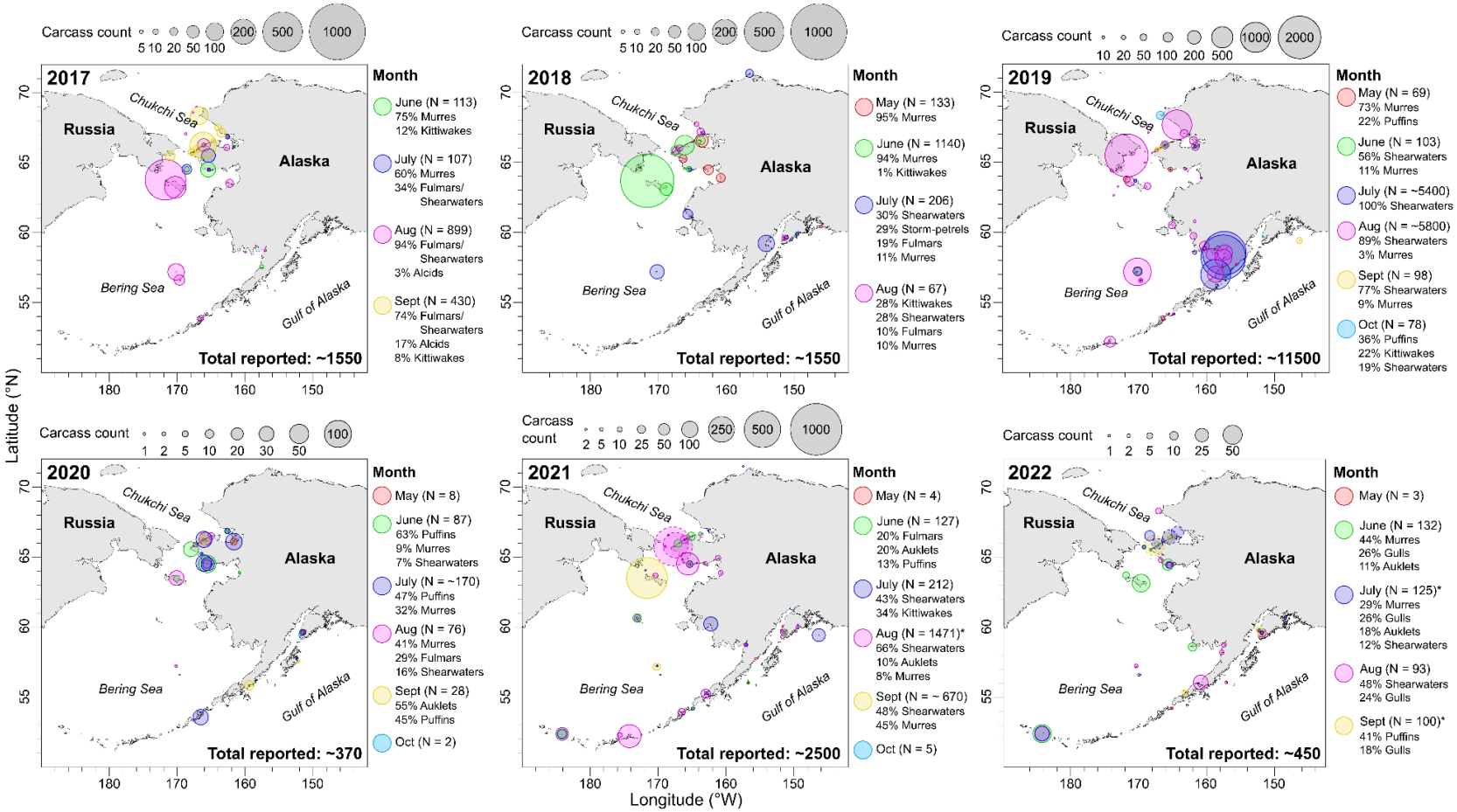

Figure S15. Mosaic of annual seabird mortality reports. These maps were produced annually in collaboration with US Fish and Wildlife Service as a tool for communicating where, and when, seabird mortality had been reported, with circle size indicating summed carcass counts reported in 100×100 km grid cells each month. Contributing data primarily consists of opportunistic reports from coastal communities, as well as a small number of at-sea reports, and counts derived from National Parks Service aerial surveys in the Bering Strait region, which are shown as dashed circles to distinguish them from on the ground reports. Total carcass counts, and approximate taxonomic composition are also shown per month in the map legend.

Supplement 4: Event characteristics and aggregation

Table S6. Unusual mortality event characteristic definitions.

Name	Description	Calculation
n_{bird}	Observed number of marine bird carcasses across all surveys within event bounds	
start date	Start date of the earliest 28-day period within event bounds	
end date	End date of the latest 28-day period within event bounds	
duration	Number of days between event **start** and **end** dates	
average encounter rate	Mean beached bird encounter rate (birds km^{-1}) estimated by the model evaluated across 28-day time-periods within event bounds.	$\frac{1}{n_e}\sum_{i}^{n_e}\widehat{\mu_i}$
maximum encounter rate	Maximum beached bird encounter rate (birds km^{-1}) estimated by the model evaluated across 28-day time-periods within event bounds.	$\exp\left(\max\left(\widehat{\mu_{i:j}}\right)\right)$
average anomaly	Ratio of estimated beached bird encounter rate relative to baseline averaged across 28-day time-periods within event bounds.	$\exp\left(\frac{1}{n_e}\sum_{i}^{n_e}\delta_i\right)$
maximum anomaly	Peak ratio of estimated beached bird encounter rate relative to baseline across 28-day time-periods within event bounds.	$\exp\left(\max\left(\delta_{i:j}\right)\right)$
extent	Event extent/coverage of beaches where anomalously high carcass counts were observed.	Region extent (km) multiplied by proportion of beaches where observed carcass count ≥ 3 × expected count anytime within event bounds
magnitude	Proxy measure representative of cumulative carcass count, taking into account average encounter rate per survey (birds km^{-1}), event extent (i.e. km affected) and event duration.	Average encounter rate × event duration × event extent

a: δ_i: anomaly values within event

b: n_e: Number of 28-day periods within event

c: $\widehat{\mu_i}$: expected carcass count evaluated at the mid-point of period *i* on a generic 1 km beach segment (i.e. beach effect = 0)

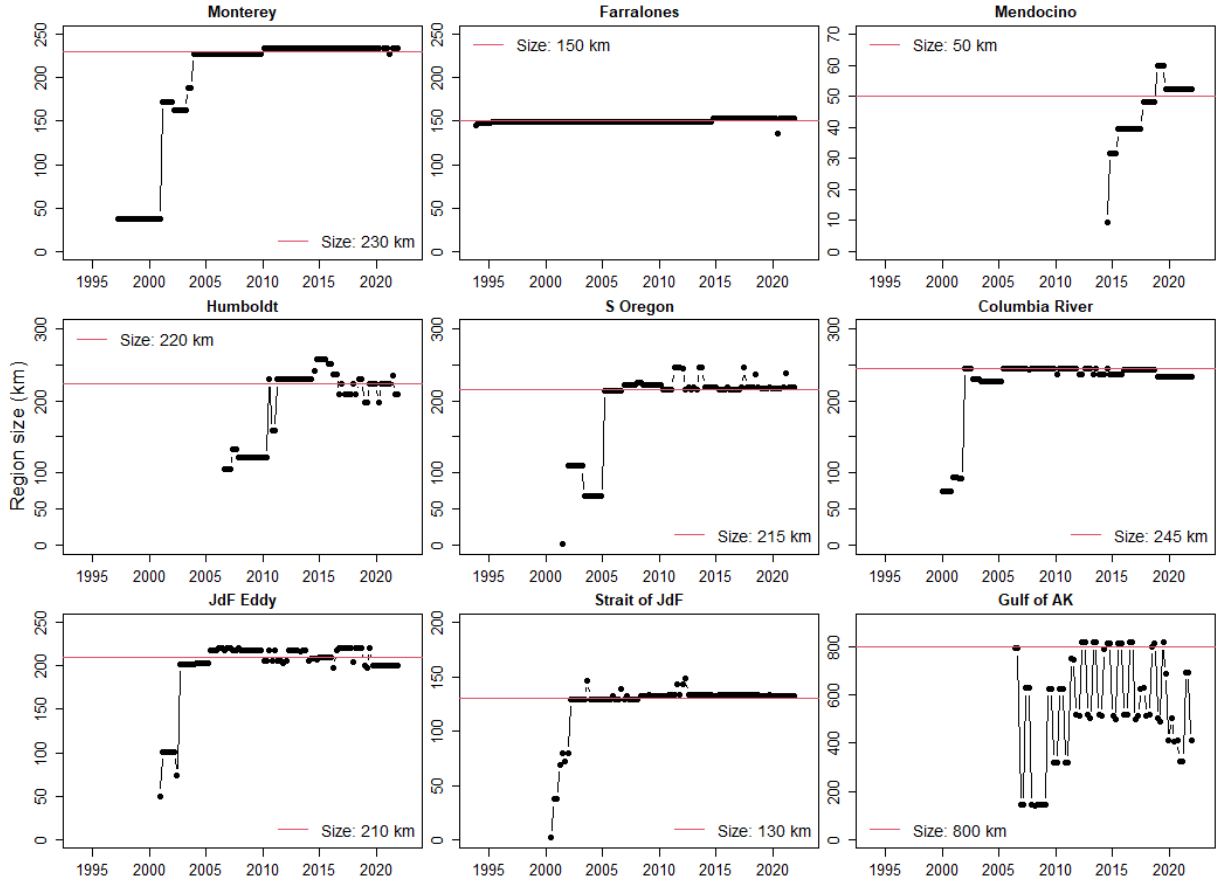

Figure S16. Regional extents through time for each of the California Current regions, excluding the inside waters of Puget Sound and the Strait of Georgia, and the Gulf of Alaska. Extents represent the maximum point to point distance between beaches monitored in that region evaluated based on a 90 day moving time-window centred on each plotted point. Horizontal red lines indicate the assigned extent for each region, which is also given in the figure panel legends.

Event aggregation

Mortality events identified at the regional level (n = 149) were then screened to remove events constituting fewer than 50 carcass observations that likely occurred due to relatively low survey effort, and/or low baseline seasonality. Of the 149 regional events, 38 were excluded based on this threshold, and we then aggregated the remaining 111 events based on:

- location: requiring they occur in the same large marine ecosystem
- timing: requiring events occur within the same seasonal window (event mid-points separated by < 90 days)
- taxonomic composition: requiring they be categorised in the same event typology based on a hierarchical cluster analysis of taxonomic composition among events

The last step was achieved by calculating taxonomic composition for each event, by tallying the number of seabird carcasses observed within event boundaries (in region and time-window) into major taxonomic groupings (**Table S7**). This was then converted to percent composition of each group in each event. Percent composition was then used to calculate a distance matrix (Euclidean distance) representative of the similarity/dissimilarity in percent taxonomic composition between event pairs. This distance matrix was then used to construct a hierarchical scheme of event similarities using the hclust function in R. This function performs a hierarchical cluster analysis of a given distance matrix, grouping similar multivariate observations by iteratively combining the two most similar clusters, starting with each observation in its own cluster, until all events are combined into the same node. Based on the resulting dendrogram of event similarities, we then identified event clusters by cutting the tree at a certain height. The optimum number of clusters was identified by calculating the mean silhouette width across events, where silhouette width is a measure of how similar an observation is to its own cluster compared to other clusters within that scheme (Rousseeuw 1987). Thus the mean silhouette width provides a method for comparing the strength/cohesiveness of a given clustering scheme, and therefore can be used as a method of identifying an optimum number of clusters that does not over-aggregate/-divide events.

Analysis of mean silhouette widths among competing cluster schemes suggested that a clustering scheme resulting from applying the Average linkage criterion with 9 clusters performed best of those investigated (**Figure S17**). We then used this clustering scheme to investigate individual event membership within each cluster according to Silhouette width, and also the general pattern of taxonomic composition within and among clusters (**Figure S18**). Using this clustering scheme, the majority of events had relatively high group membership, with only 3 of the 111 events showing a negative silhouette width value (JDF Eddy Sep 2004; Monterey Mar 1998; Humboldt Mar 2007), indicating that they were a potential outlier in their assigned group. Each of the nine clusters are characterized by their dominant species (left to right in **Figure S18**): kittiwakes (n = 1), puffins (n = 9), fulmar (n = 25), small alcids (n = 5), seaducks (n = 1), murre (n = 49), gulls (n = 2), grebes (n = 10), and shearwaters (n = 9). However, the grebe dominated cluster was considerably more mixed than the other clusters, and the murre cluster seemed also to show a division between events overwhelmingly dominated by murres, and those that represented a more even mix of locally breeding birds, primarily murres, cormorants, gulls (**Figure S18**).

These 9 clusters were then used to guide aggregation of regional events as described in the main text.

Table S7. Taxonomic groups of species used in hierarchical cluster analyses of regional events.

Name	Family	Genera	Gross %
murre	Alcidae	Uria	33.01
puffin	Alcidae	Cerorhinca, Fratercula	3.39
small alcid	Alcidae	Aethia, Brachyramphus, Synthliboramphus	6.55
other alcid	Alcidae	Cepphus	1.38
shearwater	Procellariidae	Ardenna, Puffinus	4.06
fulmar	Procellariidae	Fulmarus	14.03
albatross	Diomedeidae	Phoebastria	0.17
storm-petrel	Hydrobatidae	Hydrobates	0.30
other procellariid	Procellariidae	Pterodroma	0.10
gull	Laridae	Larus, Chroicocephalus, Xema, Leucophaeus	14.51
kittiwake	Laridae	Rissa	0.66
other larid	Laridae/ Stercorariidae	Hydroprogne, Sterna, Onychoprion, Rynchops, Thalasseus, Stercorarius	0.19
cormorant	Phalacrocoracidae	Urile, Nannopterum	8.87
booby	Sulidae	Sula	0.01
pelican	Pelicanidae	Pelecanus	1.30
grebe	Podicipedidae	Aechmophorus, Podilymbus, Podiceps	6.53
loon	Gaviidae	Gavia	1.81
scoter	Anatidae	Melanitta	2.44
eider	Anatidae	Somateria	0.01
other diving duck	Anatidae	Bucephala, Aythya, Mergus, Histrionicus, Lophodytes, Clangula, Oxyura	0.31
phalarope	Scolopacidae	Phalaropus	0.36

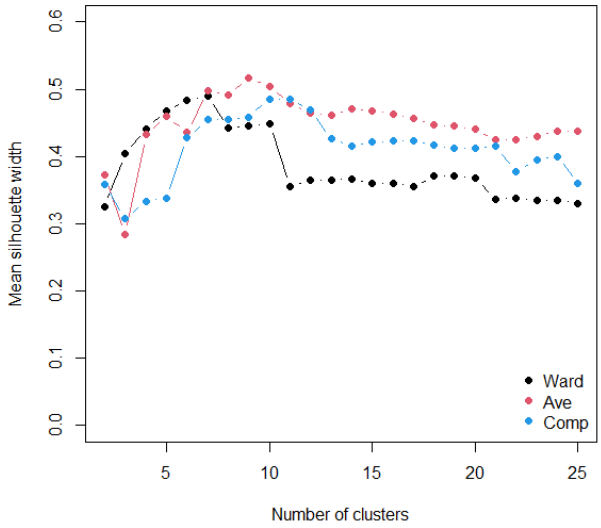

Figure S17. Mean silhouette width as a measure of the compactness/membership of regional mortality event clusters according to taxonomic composition for clustering schemes ranging from 2 to 25 clusters. Three different hierarchical linkage criterion were investigated (Average, Wards, Complete), and the mean silhouette width, averaged across mortality event-specific (n = 111) silhouette width in their assigned cluster is shown for each linkage criteria.

Taxonomic clustering of regional mortality events

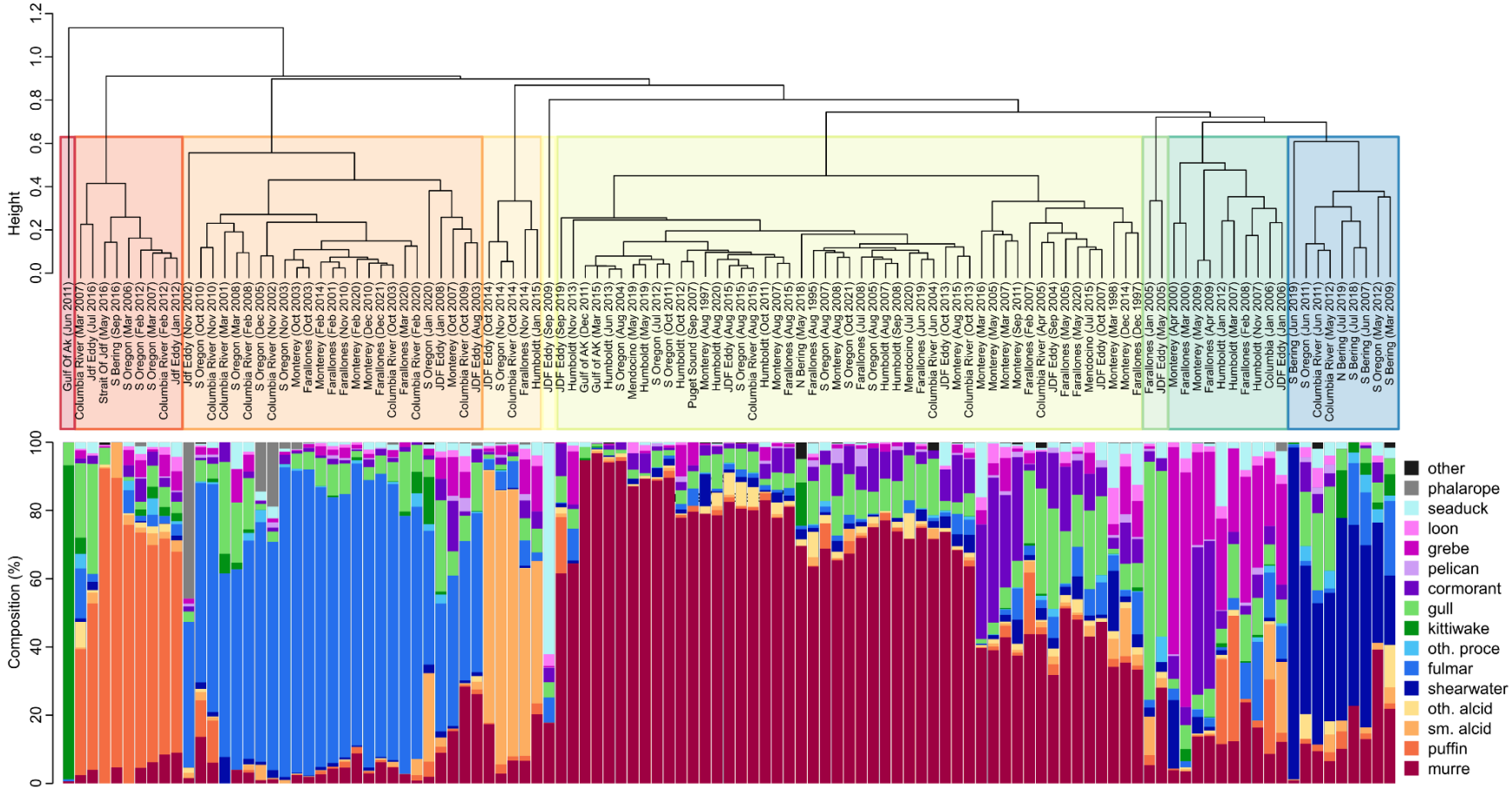

Figure S18. Hierarchical clustering scheme of regional mortality events (n = 111). The upper panel shows the dendrogram of event similarities based on proportional taxonomic composition, and is constructed using the average linkage criteria. Shaded polygons indicate the 9 clusters that came out as the optimal number of clusters based on mean silhouette width across events among clustering schemes with 2-25 clusters. The lower panel shows individual event proportional composition among major taxonomic groups, which we used to identify event typologies.

Supplement 5: Event characteristic analyses

Linear models: Event characteristics ~ f(time)

Aggregated event statistics (average anomaly, average encounter rate, extent, duration, magnitude) were used to identify whether event characteristics had consistently 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 gls; Pinheiro and Bates 2022). The additional variance function was required to control for increasing residual variance (initial model attempts displayed varying degrees of heteroscedasticity) through time, which was modelled according to

$$\sigma^2(t) = \sigma_0^2 \mathrm{e}^{2\theta t} \qquad \text{[eqn. S8]}$$

where $\sigma^2(t)$ is the residual variance function of time, t, σ_0^2 is the residual variance at t = 0, and θ controls the rate at which variance increases with t. Including this variance function resulted in model residuals of approximately constant magnitude (after integrating out the exponential modification), meeting model variance assumptions. We also investigated whether model residuals were autocorrelated through time by specifying a continuous autoregressive structure, but this was dropped from final models as in all cases the corresponding correlation coefficient was estimated to be zero. Bootstrap resampling (n = 1000 permutations) was used to estimate the 95% confidence intervals of model coefficients regarding event characteristic trends through time.

Table S8. Model statistics and trend estimates from models fitted to event characteristics through time. Models were fitted via weighted least squares, including an exponential variance function of time to account for heteroscedasticity of model residuals. Mean trend estimates of time from the corresponding models are given, along with their bootstrapped 95% confidence intervals.

		All events			Event category: 2-4		
Response variable	**Model**	**AICc**	**R2**	**Mean trend estimate [95% CI]**	**AICc**	**R2**	**Mean trend estimate [95% CI]**
Average anomaly	ln(y) ~ 1	72.3			39.9		
	ln(y) ~ time	66.9	3	0.021 [0.003, 0.039]	41.9	0	0.016 [-0.015, 0.061]
Average encounter rate	ln(y) ~ 1	142.3			67.7		
	ln(y) ~ time	144.0	0	0.016 [-0.022, 0.051]	69.8	0	0.024 [-0.019, 0.088]
Duration	ln(y) ~ 1	113.1			64.8		
	ln(y) ~ time	115.4	0	0.005 [-0.031, 0.048]	67.3	0	-0.006 [-0.054, 0.058]
Extent	ln(y) ~ 1	92.4			34.3		
	ln(y) ~ time	93.9	0	-0.048 [-0.132, 0.030]	37.1	0	0.031 [-0.057, 0.129]
Event magnitude	ln(y) ~ 1	133.9			55.9		
	ln(y) ~ time	136.2	0	-0.050 [-0.190, 0.099]	55.6	3	0.136 [-0.028, 0.322]

Linear models: Event characteristics ~ f(SSTa)

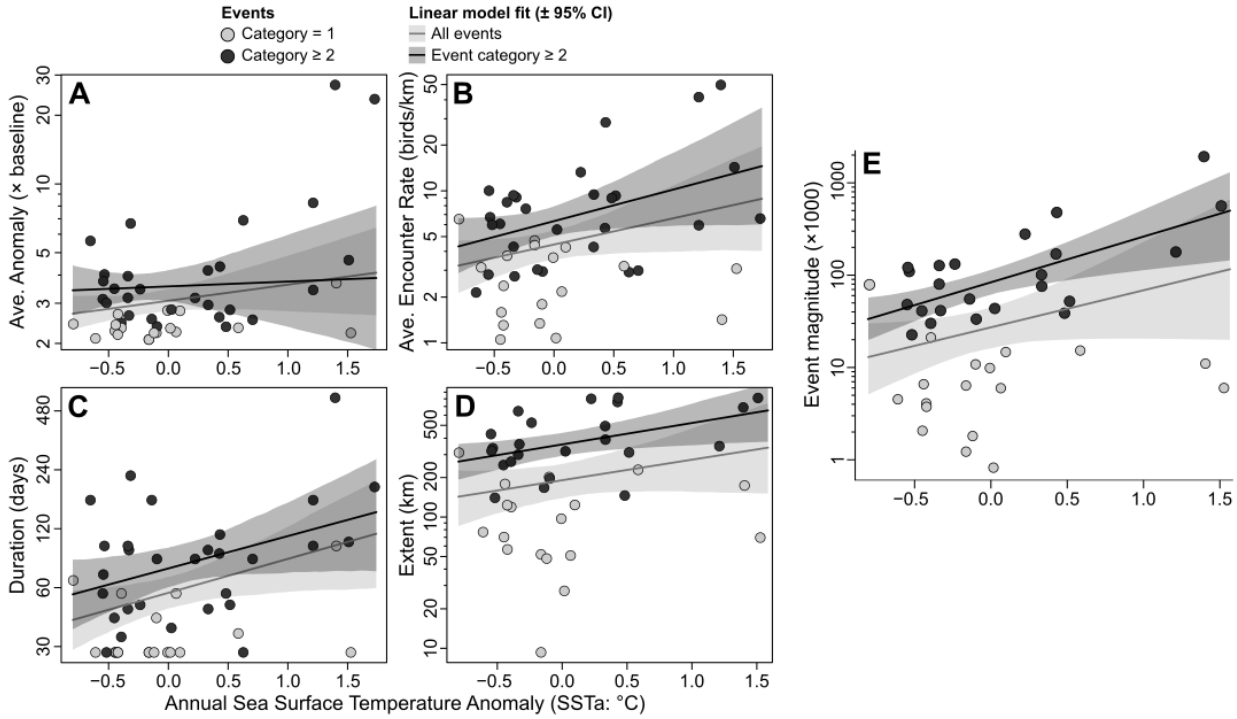

Figure S19. Aggregated unusual mortality event characteristics as a function of annual (year prior to event onset) sea surface temperature anomaly (SSTa). Event characteristics shown are: (**A**) average anomaly, defined as the multiplicative difference between observed and expected baseline carcass encounter rate (birds per km); (**B**) average carcass encounter rate per km of beach surveyed; (**C**) event duration, calculated as the number of days where encounter rate anomaly was above 2; (**D**) event extent, estimated as the km of shoreline where beached bird encounter rate was elevated based on the proportion of beaches registering above baseline encounter rates; and (**E**) event magnitude, calculated as the product of average carcass encounter rate, duration, and extent as a proportional measure of total deposition. In panels (**A**-**E**), events are delineated by event magnitude category. Lines and shaded polygons indicate fitted linear regression models (mean and 95% confidence interval) of event characteristics as a function of SSTa fitted to all data points (grey), and only magnitude category events 2 or above (black).

Table S9. Full linear model parameters of event characteristics as a function of annual SSTa and season. In each case, season effects are presented as the difference in response relative to autumn (intercept), and parameter values are given as mean and 95% confidence interval in []. All models were fitted to log-transformed response variables.

Response	SSTa	Season
All event categories		
Average event anomaly	0.15 [-0.18, 0.39]	Spring: -0.04 [-0.36, 0.34] Summer: -0.06 [-0.30, 0.26] Winter: -0.02 [-0.28, 0.27]
Average encounter rate	0.38 [-0.05, 0.84]	Spring: -0.53 [-1.14, 0.12] Summer: -0.14 [-0.73, 0.36] Winter: -1.01 [-1.46, -0.60]
Duration	0.45 [0.05, 0.85]	Spring: -0.21 [-0.75, 0.32] Summer: -0.30 [-0.83, 0.17] Winter: -0.01 [-0.51, 0.51]
Extent	0.40 [-0.13, 0.87]	Spring: -0.39 [-1.14, 0.40] Summer: -0.58 [-1.77, 0.50] Winter: -0.38 [-1.10, 0.40]
Event magnitude	1.11 [-0.15, 2.21]	Spring: -1.44 [-2.70, -0.09] Summer: -0.91 [-2.35, 0.48] Winter: -1.30 [-2.47, -0.11]
Event categories: 2 - 4		
Average event anomaly	-0.07 [-0.85, 1.00]	Spring: -0.13 [-0.68, 0.72] Summer: 0.04 [-0.32, -1.27] Winter: 0.10 [-0.24, 0.53]
Average encounter rate	0.32 [-0.15, 0.80]	Spring: -0.06 [-0.74, 0.43] Summer: 0.00 [-0.70, 0.57] Winter: -0.86 [-1.25, -0.34]
Duration	0.49 [0.06, 0.82]	Spring: 0.00 [-0.71, 0.72] Summer: -0.42 [-1.02, 0.19] Winter: 0.51 [0.03, 1.03]
Extent	0.37 [-0.04, 0.76]	Spring: -0.16 [-0.89, 0.46] Summer: -0.01 [-0.86, 0.90] Winter: 0.07 [-0.49, 0.56]
Event magnitude	1.24 [0.33, 1.92]	Spring: -0.68 [-1.78, 0.36] Summer: -0.35 [-1.28, 0.97] Winter: -0.28 [-1.09, 0.57]

Supplement 6: Marine heatwave association analyses

The randomization of mortality events proceeded via the following steps

1. Event start dates, and duration, are tied to the baseline model resolution (28 day time-windows), therefore event start dates could only be one of the possible 28-day windows, which formed our sample pool from which to draw dates at random
2. For each event, a start date was sampled at random from the available 28-day time-windows, restricted to those where monitoring had been performed in that region
3. If that event crossed region-boundaries, then the shared regional events were also shifted in time such that the time difference among event onset dates remained constant (i.e. they move as a block)
4. This was repeated until all regional events had been assigned a randomized onset date
5. The candidate assortment was then tested for validity, with the following fail conditions
    a. Condition 1: if an event went past the end of 2021 (i.e. a 112 day event was randomly placed at 1-Dec-2021, such that while the start of the event was within our data time-window, the end was not)
    b. Condition 2: if an event was assigned to a time when no surveys were performed in that region. i.e. an event (1) that was linked to another (2) was placed at a time that was valid for (1), but due to variable coverage was not valid for (2) given differences in program establishment in different regions
    c. Condition 3: if events overlapped in time within a region, or were butted against one another
6. If the assortment failed any one of these conditions, it was thrown out and another assortment generated in its stead
    a. Our initial attempts at this process involved using all regional events, but due to the relatively larger number of category 1 events, very few valid assortments were found due to the increased likelihood of event overlap
7. If the assortment was valid, then the time to most recent MHW was calculated for each event, and a set of null sample statistics (events within 12 months, mean event delay) was calculated from that information

The observed and null distributions of test statistics are presented in **Table S10**.

Table S10. Results of marine heatwave association analyses. Values represent observed test statistic values, and the mean/95% range of statistic values generated via permutation analyses (n = 1, 000 permutations) representing the null hypothesis that mortality events occur at random as a test of the association between mortality events and marine heatwaves. Test statistics are the mean time gap between the onset dates of a mortality event and the most recent marine heatwave for each event, as well as the number of mortality events within 12-months of a marine heatwave. Results are presented for 2 different scenarios based on mortality event encounter rate category (2 or above, 3 or above). Presented p-values represent the proportion of permutations with test statistics that were ≤ (mean time gap), or ≥ (event count) the observed value.

Mortality event category	Mean time-gap				Events within 12 months of MHW			
		Null distribution				Null distribution		
	Obs.	Mean	95% range	p-val	Obs.	Mean	95% range	p-val
2-4	340	489	[317, 679]	0.044	35	28.9	[21, 36]	0.068
3-4	295	570	[316, 895]	0.035	16	12.2	[7, 17]	0.120

References

Wood, S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

Rousseeuw, P.J. (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math., 20, 53–65.