# Appendix 1: Description of data

Appendix 1

The integrated population model provides a framework for synthesizing multiple demographic datasets to mutually inform shared parameters. In Appendix 1.1, we further describe the Drayton Harbor time series data (D1). In Appendix 1.2, we describe additional information about the size-at-age data (D2), the model fitted to size-at-age data (D2), code used to fit the model, and posterior summaries. In Appendix 1.3, we describe the mark-recapture data (D3).

Appendix 1.1

The time series data (D1) was collected in Drayton Harbor, an enclosed bay in the Salish Sea in Washington State, USA.

Figure A1.1: Bathymetric map of Drayton Harbor. Colors indicate depth and black line indicates coast line.
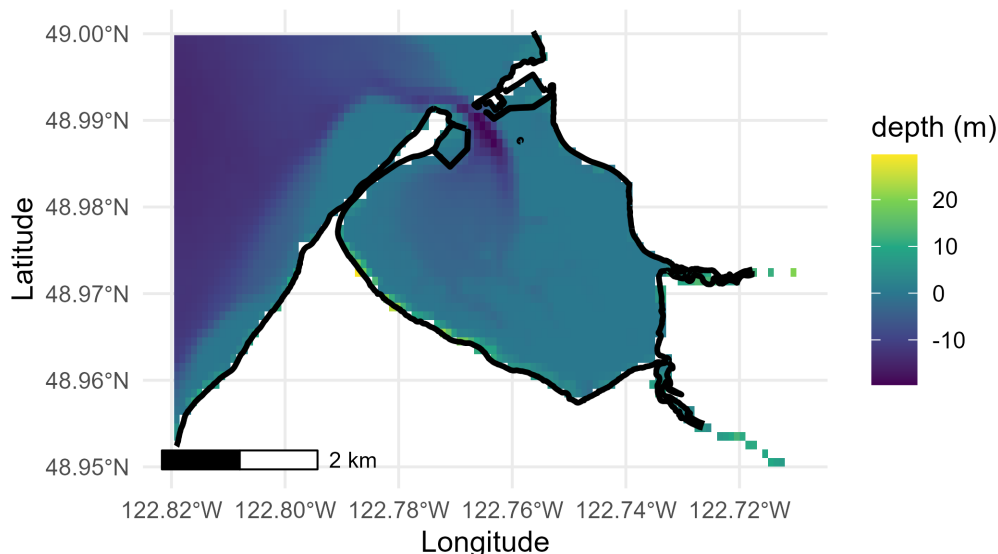
Below we describe the total removal effort and catch of European green crab (EGC, *C. maenas*) in Drayton Harbor (D1). While the removal count data is collected at many time points throughout the trapping season (~ April - October), we discretize the time intervals into biweekly periods in the model.

Drayton Harbor data were subset to only include recruits (age < 1) that appear in a single cohort in the fall of each year (Figure 1B). This subset only removed three recruit individuals in 2022 from the dataset, which were from an off-cycle recruitment cohort.

Three different trap types were used with different mesh sizes and entrance openings: Fukui, Minnow, and Shrimp. Fukui traps have 12 mm mesh with approximately 20 cm entrance openings. Gee-brand minnow traps have 6 mm mesh with approximately 5 cm entrances. Promar-brand 4-way shrimp traps have 1/2"-1" mesh size and approximately 10 cm entrance opening

Table A1.1: Number of crabs caught and total removal effort (number of traps) during each biweek and year. Julian day signifies the first Julian day of the discretized biweek period. F, M, and S correspond to Fukui, Minnow, and Shrimp traps, respectively.

Julian day	year	F (catch)	F (effort)	M (catch)	M (effort)	S (catch)	S (effort)
137	2020	12	48	22	43	0	0
152	2020	16	67	10	62	0	0
167	2020	8	62	5	57	0	0
182	2020	4	100	0	96	0	0
198	2020	1	64	0	44	0	0
213	2020	12	61	0	60	0	0
229	2020	7	82	0	71	0	0
244	2020	8	97	4	87	0	0
259	2020	4	15	0	10	0	0
259	2020	15	57	23	52	0	0
91	2021	1	180	4	180	0	0
106	2021	7	168	1	168	0	0
121	2021	8	130	2	130	0	0
137	2021	9	207	2	205	0	0
152	2021	4	180	0	180	0	0
167	2021	6	198	0	198	0	0
182	2021	2	120	0	120	0	0
198	2021	0	123	0	123	0	0
213	2021	18	200	0	203	0	0
229	2021	1	251	1	251	12	27
244	2021	8	296	3	297	27	39
259	2021	0	0	0	60	23	24
106	2022	0	135	0	135	3	12
121	2022	2	195	0	195	13	17
137	2022	0	90	0	90	4	12
152	2022	0	127	1	125	4	16
167	2022	2	120	0	120	22	27
182	2022	0	0	0	0	4	6
198	2022	0	62	1	62	15	23
213	2022	0	207	0	207	20	71
229	2022	0	30	0	30	6	50
244	2022	0	0	0	0	5	26
259	2022	0	0	2	26	23	25
259	2022	7	27	29	118	37	50
290	2022	0	0	0	0	6	12
305	2022	8	30	16	30	16	10
106	2023	0	120	2	120	3	20
121	2023	2	163	1	163	20	31
137	2023	2	95	2	95	23	34
152	2023	0	0	0	0	12	26
167	2023	5	60	2	60	8	28
182	2023	0	80	2	80	18	14
198	2023	0	60	0	60	2	26
213	2023	0	43	2	45	2	22
229	2023	0	72	0	149	6	31
244	2023	0	15	0	45	0	18
259	2023	0	30	1	105	10	30
259	2023	0	0	0	0	0	6

Julian day	year	F (catch)	F (effort)	M (catch)	M (effort)	S (catch)	S (effort)
290	2023	1	45	1	90	14	112

Appendix 1.2

Below we provide additional information about the size-at-age data (D2), the model fitted to size-at-age data (D2), code used to fit data to the model, and posterior summaries.

Appendix 1.2.1

Size-at-age data (D2) are collected from two sources: 1) records from crab removal observations in northeastern Pacific estuaries from Yamada et al. 2005, when the somatic growth of a strong recruitment class was tracked over time (Yamada et al. 2005), and 2) crabs with an assigned year class from Drayton Harbor that were easily identifiable based on size and carapace color. While data from these Drayton Harbor recruits enter the integrated population model likelihood twice (D1 and D2), extensive simulation-based research has revealed that IPMs are robust to dependent data (Abadi et al. 2010). Ages are assigned assuming recruitment settlement in March.

Table A1.2: Number of size-at-age records during each year and waterbody in the northeastern Pacific.

Year	Waterbody	n
1995	Humboldt Bay	2
1997	Coos Bay	2
1998	Yaquina Bay	1
1999	Barkley Sound	2
1999	Price Bay	1
2000	Clayoquot Sound	1
2000	Nootka Sound	1
2001	Little Espinosa Inlet	3
2007	Klaskino Inlet	1
2011	Gale Passage	2
2015	Kildidt Inlet	2
2016	Higgins Passage	1
2016	Westcott Bay	1
2017	Becher Bay	3
2017	Padilla Bay	1
2017	Whidbey Island	2
2018	Port Townshend Bay	1
2018	Scow Bay	1
2018	Westcott Bay	2
2018	Whidbey Island	1
2018	Witty's Lagoon	1
2019	Bellingham	3
2019	Boundary Bay	3
2019	Chuckanut Bay	1
2019	Esquimalt Lagoon	2
2019	Lummi Bay	2
2019	Salt Spring Island	4
2019	Samish Bay	1
2019	Sequim Bay	3
2019	Westcott Bay	1
2019	Witty's Lagoon	2
2020	Chuckanut Bay	1
2020	Daajing Giids	1
2020	Drayton Harbor	253
2021	Drayton Harbor	30

Year	Waterbody	n
2022	Drayton Harbor	125

Appendix 1.2.2

The integrated population model (IPM) is fitted sequentially, such that the seasonal growth parameters are fit with the size-at-age records, and the summarized posteriors are used to develop prior distributions in the overall IPM. Below, we detail the model used to fit the seasonal growth parameters with the size-at-age data. The size-at-age data, $W_{a,i}$, indicating carapace width at age a in year i are fit to the same seasonal growth model as in the main text (Somers et al. 1988). Except here, we account for non-independence among data collected in the same year, since growth rate is likely affected by water temperature, which varies from year-to-year. The expected carapace width, $\widetilde{W}_a$, of a crab at age a is:

$$\widetilde{W}_{a,i} = y_\infty(1 - exp(-k(a - t_0) - s(a) + s(t0))) + \epsilon_i$$

$$s(a) = \frac{Ck}{2\pi} sin(2\pi(a - t_s))$$

$$s(t_0) = \frac{Ck}{2\pi} sin(2\pi(t_0 - t_s))$$

To account for non-independence among data collected within the same year, we add a normally distributed error term with standard deviation, τ_y.

$$\epsilon_i \sim Normal(0, \tau_y)$$

To account for individual variation in growth rate, the observed size-at-age data, $W_{a,i}$, follows a normal distribution, with the expected carapace width $\widetilde{W}_{a,i}$, and standard deviation, τ_w.

$$W_{a,i} \sim Normal(\widetilde{W}_{a,i}, \tau_w)$$

We use uninformative priors for the parameters:

$$k \sim Uniform(0, 2)$$

$$C \sim Uniform(0, 2)$$

$$t_s \sim Uniform(-1, 0)$$

$$t_0 \sim Uniform(-1, 1)$$

$$\tau_y \sim Uniform(0, 100)$$

$$\tau_w \sim Uniform(0, 100)$$

$$y_\infty \sim Uniform(80, 140)$$

Appendix 1.2.3

Below is the code used to fit the seasonal growth model to the size-at-age data. First, we will write the model code with the R package, *nimble* (de Valpine et al. 2017).

```r
library(nimble)
# model code
seasonal_growth <- nimbleCode({
  # prior distributions
  size_inf ~ dunif(70, 140) # asymptotic size
  k ~ dunif(0, 2) # exponential rate of approach to asymptotic size
  C ~ dunif(0, 2) # size of seasonal variation
  ts ~ dunif(-1, 0) # time between t=0 and start of growth oscillation
  t0 ~ dunif(-1, 1) # age organism has 0 size
  tau_y ~ dunif(0, 100) # process error standard deviation
  tau_ranef ~ dunif(0, 100) # year random effect standard deviation
  for (i in 1:nsizes) {
    # process error
    size[i] ~ dnorm(y_hat[i], tau_y)
    # deterministic portion
    y_hat[i] <- (
      size_inf * (1 - exp(-k * (t[i] - t0) - s_t[i] + s_t0)) + ranef[year[i]]
    )
    # sinusoidal growth after inflection point
    s_t[i] <- (C * k / (2 * pi)) * sin(2 * pi * (t[i] - ts))
  }
  # seasonal growth before inflection point
  s_t0 <- (C * k / (2 * pi)) * sin(2 * pi * (t0 - ts))
  # year random effect
  for (y in 1:nyears) {
    ranef[y] ~ dnorm(0, tau_ranef)
  }
})
```

Next we will bundle up the data, constants, and initial values for the MCMC.

```r
# bundle up data and constants
constants <- list(
  nsizes = length(data$CW), # number of crabs
  year = data$year_index, # year index
  nyears = length(unique(data$year_index)), # number of years
  pi = pi # pi
)
data <- list(
  size = data$CW, # size captured
  t = data$age # ages
)
# set initial values
inits_season <- list(
  size_inf = 100,
  k = 0.6,
  C = 0.7,
  ts = -0.6,
  t0 = 0,
```

```
  tau_y = 0.05,
  tau_ranef = 0.05,
  ranef = rep(0, length(unique(data$year_index)))
)
```

Finally, we will run the MCMC in parallel:

```
# run MCMC chains in parallel
cl <- makeCluster(4)
set.seed(10120)
clusterExport(cl, c("seasonal_growth", "inits_season", "data", "constants"))
# Create a function with all the needed code
out <- clusterEvalQ(cl, {
  library(nimble)
  library(coda)
  # build the model
  model_season <- nimbleModel(seasonal_growth, constants, data, inits_season)
  # build the MCMC
  mcmcConf_season <- configureMCMC(
    model_season,
    monitors = c("size_inf", "k", "C", "ts", "t0", "tau_y", "tau_ranef")
  )
  model_mcmc_season <- buildMCMC(mcmcConf_season)
  # compile the model and MCMC
  # model
  cmodel_season <- compileNimble(model_season)
  # MCMC
  cmodel_mcmc_season <- compileNimble(model_mcmc_season, project = model_season)
  # run
  cmodel_mcmc_season$run(100000, thin = 10,
                          reset = FALSE)
  return(as.mcmc(as.matrix(cmodel_mcmc_season$mvSamples)))
})
stopCluster(cl)
```
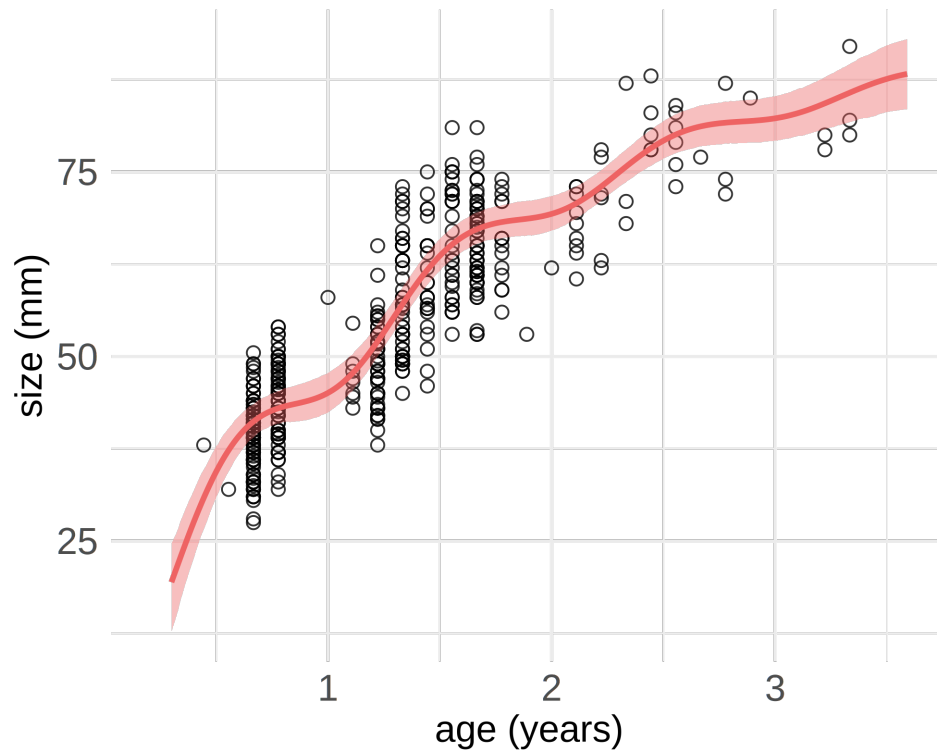
Appendix 1.2.4

Here we provide posterior summaries of the seasonal growth parameters fit with the size-at-age data.

Table A1.2: Posterior summaries, including the mean, standard deviation, 95% credibility interval (highest density interval), Rhat, and effective sample size. An X in 'IPM prior' indicates that the posterior is used in the overall IPM. Descriptions of parameters can be found in Appendix 1.2.2.

parameter	mean	sd	95 CI	Rhat	ESS	IPM prior
C	0.816	0.13	(0.559, 1.067)	1.00	7883	X
k	0.635	0.096	(0.457, 0.83)	1.02	189	X
y_∞	97.58	5.57	(86.81, 109.07)	1.02	205	X
t_0	-0.033	0.145	(-0.328, 0.199)	1.01	268	
τ_w	0.445	0.605	(0.02, 1.535)	1.02	2165	
τ_y	0.029	0.002	(0.025, 0.033)	1.00	31490	
t_s	-0.645	0.028	(-0.697, -0.589)	1.00	1982	X

Figure A1.2: Size-at-age data and model fit. Each posterior sample was used to predict the age-size relationship. The solid line corresponds to the median prediction, and the shaded area corresponds to the 95% credibility interval.

Figure A1.3: Trace plots of posterior samples of size-at-age model. Colors refer to separate chains. Descriptions of parameters can be found in Appendix 1.2.2.
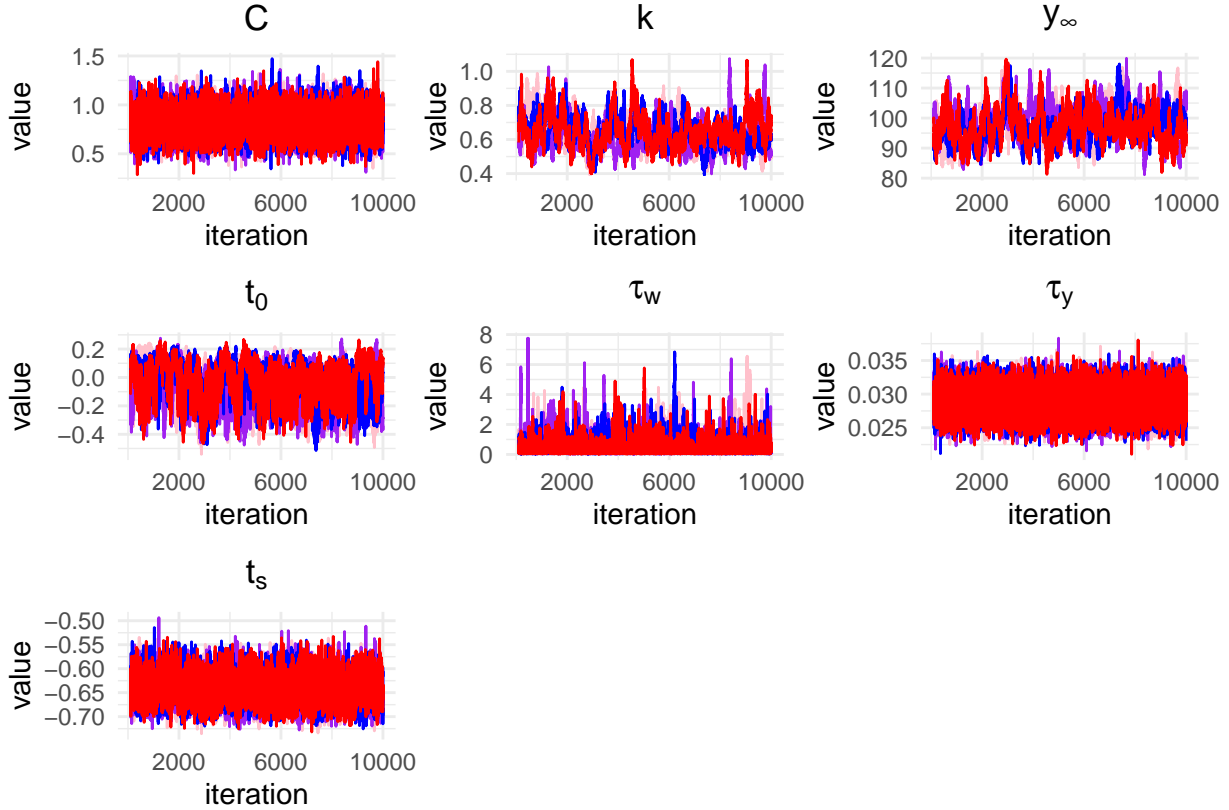
Appendix 1.3

Here we describe the batch mark-recapture data (D3). The data was taken from Grosholz et al. 2021, where the co-authors conducted an extensive capture-mark-recapture experiment to estimate population size at Seadrift Lagoon in California, USA (Grosholz et al. 2021).

The authors deployed 15 Fukui traps in June within the lagoon. Over the succeeding three days, they marked all crabs caught by clipping two spines on the carapace, recording their size, and releasing them back to their original location. Within successive weeks as part of an eradication program, crabs were removed, and the number and carapace width of retrieved (marked and unmarked) crabs were recorded.

Data was retrieved from NSF's Biological and Chemical Oceanography Data Management Office at https: //www.bco-dmo.org/person/699768. While the mark-recapture experiment took place from 2011-2018, we only used data from one year, 2011, since the marking data was unambiguous.

References

Abadi, Fitsum, Olivier Gimenez, Raphaël Arlettaz, and Michael Schaub. 2010. "An Assessment of Integrated Population Models: Bias, Accuracy, and Violation of the Assumption of Independence." *Ecology* 91 (1): 7–14.

de Valpine, Perry, Daniel Turek, Christopher J Paciorek, Clifford Anderson-Bergman, Duncan Temple Lang, and Rastislav Bodik. 2017. "Programming with Models: Writing Statistical Algorithms for General Model Structures with NIMBLE." *Journal of Computational and Graphical Statistics* 26 (2): 403–13.

Somers, IF et al. 1988. "On a Seasonally Oscillating Growth Function." *Fishbyte* 6 (1): 8–11.

Yamada, Sylvia Behrens, Brett R Dumbauld, Alex Kalin, Christopher E Hunt, Ron Figlar-Barnes, and Andrea Randall. 2005. "Growth and Persistence of a Recent Invader Carcinus Maenas in Estuaries of the Northeastern Pacific." *Biological Invasions* 7: 309–21.