Shifts in the size of fish from a culturally important recreational fishery

Mark Scheuerell

03 June 2021

Data

The original Tengu Derby data were provided to me by Tom Quinn on 16 June 2020 in the form of an MS Excel file titled `Tengu_derby_leaders through 2019 derby.xls`. I exported one worksheet of interest (`data in kg`) as `~/data/tengu_derby_data.csv`. The same Excel file also included a worksheet with information from WDFW on the number of natural- and hatchery-origin Chinook, and the mean mass of Chinook (`Losee Chiook data`), which I exported as `~/data/wdfw_data.csv`

```
## set data dir
datadir <- here::here("data")
## import raw Tengu data
tengu_data <- readr::read_csv(file.path(datadir, "tengu_derby_data.csv"))
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   derby = col_double(),
##   year = col_double(),
##   month = col_character(),
##   days = col_double(),
##   members = col_double(),
##   total_catch = col_double(),
##   n_over_10 = col_double(),
##   n_over_5 = col_double(),
##   size_1 = col_double(),
##   size_2 = col_double(),
##   size_3 = col_double(),
##   size_4 = col_double(),
##   size_5 = col_double()
## )
```

```
## import raw WDFW data
wdfw_data <- readr::read_csv(file.path(datadir, "wdfw_data.csv"))
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   year = col_double(),
##   NOR = col_double(),
##   HOR = col_double(),
##   total = col_double(),
##   size = col_double()
## )
```

Changes in fish size over time

The data set contains three different indicators of fish size over time:

1) the total number of fish over 10 pounds (~4.55 kg);
2) the total number of fish over 5 pounds (~2.27 kg); and
3) the masses (kg) of the 5 largest fish.

Clearly (1) and (2) will be correlated, as (1) is a subset of (2). Furthermore, the probability of catching a fish greater than 5 or 10 pounds clearly increases as the total number of fish caught also increases. Thus, I modeled the proportion of fish caught in a given year that were greater than the 2 size thresholds.

```
## proportion of fish >10 lbs
p10 <- tengu_data$n_over_10 / tengu_data$total_catch
## screen for p = 0 & change to p = 0.001
p10[p10 == 0] <- 0.001

## proportion of fish >5 lbs
p5 <- tengu_data$n_over_5 / tengu_data$total_catch
## screen for p = 1 & change to p = 0.999
p5[p5 == 1] <- 0.999

## combine proportional data
fish_sizes <- cbind("Prop. over 10 lbs" = p10, "Prop. over 5 lbs" = p5) %>%
  ts(start = min(tengu_data$year))
```

Here are plots of the two size metrics over time.

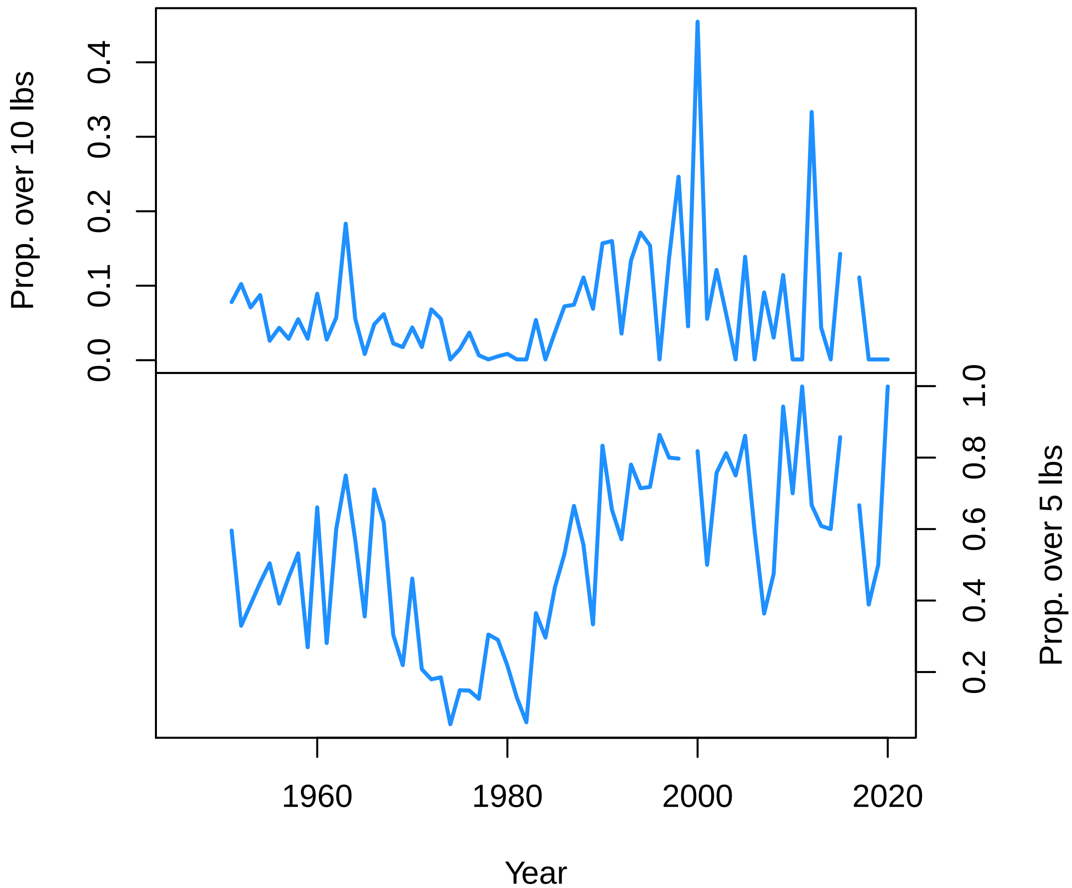

Figure 1. Time series of the proportion of fish over 10 pounds (top) and over 5 pounds (bottom).

There appears to be an overall decline in fish size from the mid 1940s until the early 1980s, when fish sizes increased rapidly before declining again until present.

Check annual size data

size_check <- tengu_data %>% select(starts_with("size")) %>% apply(1, is.na) %>% t()

Fit biased random walk

Just as I did with the CPUE data, I fit both biased and unbiased forms of random walk models to the fish size data, with the response being the logit-transformed proportions of fish over 10 and 5 pounds.

```
## response for fish >10 lbs
l_size_10 <- matrix(qlogis(p10), nrow = 1)

## response for fish >5 lbs
```

```r
l_size_5 <- matrix(qlogis(p5), nrow = 1)

## model setup
mod_list <- list(
  B = matrix(1),
  U = matrix("u"),
  Q = matrix("q"),
  Z = matrix(1),
  A = matrix(0),
  R = matrix("r")
)

## over 10
## fit model with bias (Eqn 1)
size_brw_10 <- MARSS(l_size_10, model = mod_list)

## Success! abstol and log-log tests passed at 61 iterations.
## Alert: conv.test.slope.tol is 0.5.
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 61 iterations.
## Log-likelihood: -140.8205
## AIC: 289.641   AICc: 290.266
##
##       Estimate
## R.r     2.8030
## U.u    -0.0401
## Q.q     0.1307
## x0.x0  -2.3235
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.

## 95% CI for bias
size_brw_10 <- MARSSparamCIs(size_brw_10)


## over 5
## fit model with bias (Eqn 1)
size_brw_5 <- MARSS(l_size_5, model = mod_list)

## Success! abstol and log-log tests passed at 65 iterations.
## Alert: conv.test.slope.tol is 0.5.
```

```
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 65 iterations.
## Log-likelihood: -118.5948
## AIC: 245.1895   AICc: 245.8244
##
##       Estimate
## R.r     1.5330
## U.u     0.0355
## Q.q     0.0769
## x0.x0  -0.5255
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

```
## 95% CI for bias
size_brw_5 <- MARSSparamCIs(size_brw_5)
```

Fit unbiased random walk

```
## set bias to 0
mod_list$U <- matrix(0)

## over 10
## fit model without bias (Eqn 2)
size_rw_10 <- MARSS(l_size_10, model = mod_list)
```

```
## Success! abstol and log-log tests passed at 57 iterations.
## Alert: conv.test.slope.tol is 0.5.
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 57 iterations.
## Log-likelihood: -141.1965
## AIC: 288.3929   AICc: 288.7621
##
##       Estimate
## R.r      2.797
## Q.q      0.146
## x0.x0   -2.715
## Initial states (x0) defined at t=0
```

```
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

```
## over 5
## fit model without bias (Eqn 2)
size_rw_5 <- MARSS(l_size_5, model = mod_list)
```

```
## Success! abstol and log-log tests passed at 57 iterations.
## Alert: conv.test.slope.tol is 0.5.
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 57 iterations.
## Log-likelihood: -119.0947
## AIC: 244.1894   AICc: 244.5644
##
##       Estimate
## R.r     1.5172
## Q.q     0.0932
## x0.x0  -0.1698
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

For the model based upon the proportion of fish over 10 pounds, the 95% confidence interval for the bias term (u) is (-0.132, 0.052), and the model with a bias term has an AIC value that is only ~-1.5 units lower than the model without a bias term, suggesting there is essentially no data support for a systematic downward trend in the log-transformed size over time. This is not much of a surprise, however, given the apparent temporal patterns in the data.

For the model based upon the proportion of fish over 5 pounds, the 95% confidence interval for the bias term (u) is (-0.035, 0.106), and the model with a bias term has an AIC value that is only ~-1.3 units lower than the model without a bias term, suggesting there is essentially no data support for a systematic downward trend in the log-transformed size over time.

Model fit to size data

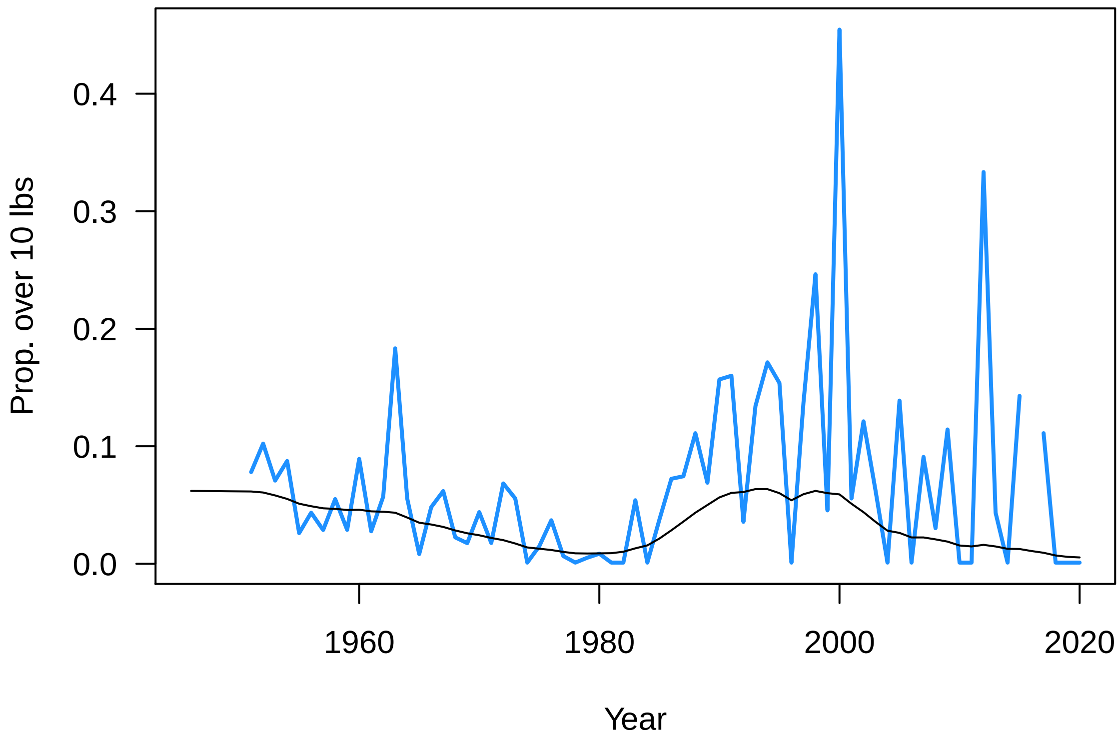

Figure 2. Time series of observed size (blue) and the fit for the random walk model (Eqn 3; black).

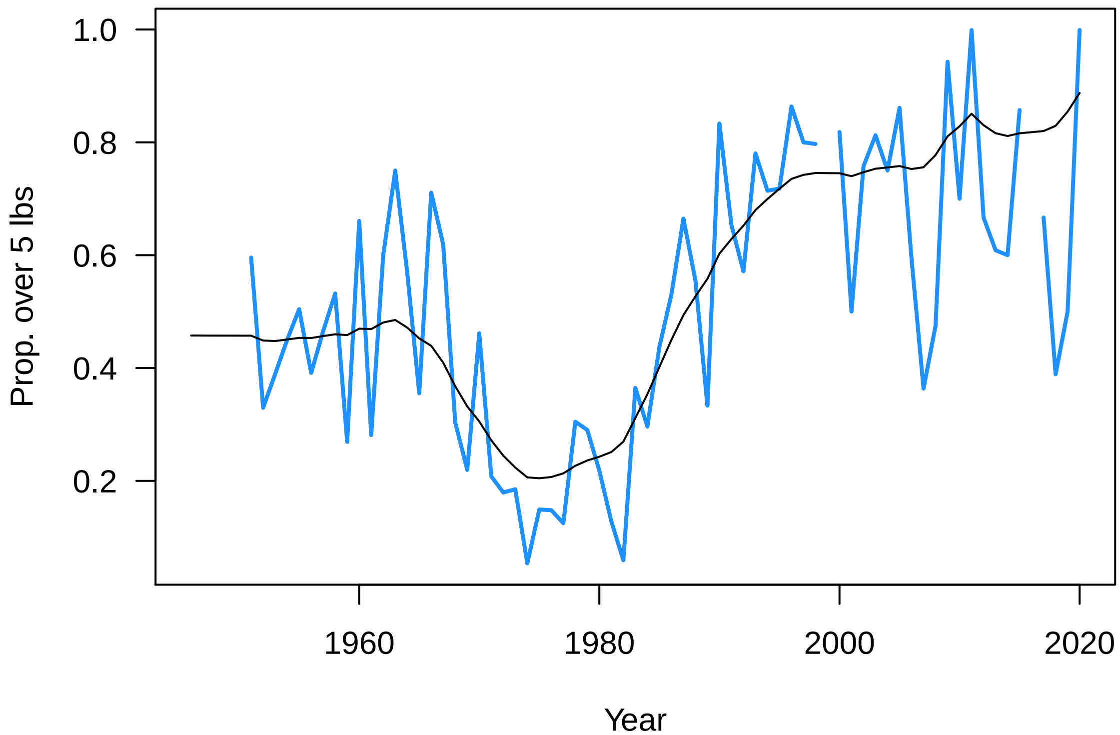

Figure 3. Time series of observed size (blue) and the fit for the random walk model (Eqn 3; black).

Comparisons between the Tengu Derby & WDFW

Although the two data sources come from different times, places, and gear types, they both contain information on the temporal changes in size and CPUE over time. I investigated whether or not the temporal trends in the two data sources track one another (i.e., are representative of one "state of nature"). To do so, I used a multivariate state-space model of the general form

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \tag{1}$$

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t \tag{2}$$

For both forms of the model, $\mathbf{y}_t$ is a $[2 \times 1]$ vector of the observed data from both sources, $\mathbf{a}$ is a $[2 \times 1]$ vector of offsets (intercepts), and $\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$. For both models, I assumed that the observation errors at time t ($\mathbf{v}_t$) are independent and differently distributed, such that

$$\mathbf{R} = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}$$

Fish size

The time series from WDFW begins in 1970 and runs through 2015, but the Tengy Derby data is missing size information for 2015, so I restricted my analysis to the 45 years from 1970-2014. Again I fit models to the log-transformed size data.

One pattern over time

For the model with only one state of nature, $\mathbf{Z}$ is a $[2 \times 1]$ vector of 1's, $\mathbf{x}_t$ is a $[1 \times 1]$ scalar of the true state, and $\mathbf{w}_t \sim \mathrm{N}(0, q)$, such that

$$\begin{bmatrix} y_{\text{Tengu}} \\ y_{\text{WDFW}} \end{bmatrix}_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_t + \begin{bmatrix} a_{\text{Tengu}} \\ a_{\text{WDFW}} \end{bmatrix} + \begin{bmatrix} v_{\text{Tengu}} \\ v_{\text{WDFW}} \end{bmatrix}_t \tag{3}$$

$$x_t = x_{t-1} + w_t \tag{4}$$

```
## mean mass of 2 largest fish
mean_of_top_2 <- apply(tengu_data[,c(9:10)], 1, mean, na.rm = TRUE)
mean_of_top_2[is.nan(mean_of_top_2)] <- NA

## select only years when 5 fish were weighed
tengu_sizes <- tengu_data %>%
  select(starts_with("size")) %>%
  apply(1, mean)


## get ts of Tengu catch
# catch <- tengu_data$total_catch[tengu_data$year >= 1970 & tengu_data$year <= 2014]

## select common data
yy <- cbind(tengu = tengu_sizes[tengu_data$year >= 1970 & tengu_data$year <= 2014],
            wdfw = wdfw_data$size[wdfw_data$year >= 1970 & wdfw_data$year <= 2014])

## model defn for Eqns 6 & 7
mod_list <- list(
  B = matrix(1),
  U = matrix(0),
  Q = matrix("q"),
  Z = matrix(1, nrow = 2, ncol = 1),
  A = matrix(c("T", "W"), nrow = 2, ncol = 1),
  # R = matrix(list("T", 0, 0, "W"), 2, 2)
  R = matrix(list("r", 0, 0, "r"), 2, 2)
)

## fit Eqns 6 & 7
size_both_1 <- MARSS(t(log(yy)), model = mod_list)
```

```
## Success! abstol and log-log tests passed at 62 iterations.
## Alert: conv.test.slope.tol is 0.5.
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 62 iterations.
## Log-likelihood: 15.98604
## AIC: -21.97209   AICc: -21.24038
##
##        Estimate
## A.T   -0.245864
## A.W    0.223262
## R.r    0.036988
## Q.q    0.000818
## x0.x0  1.811714
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

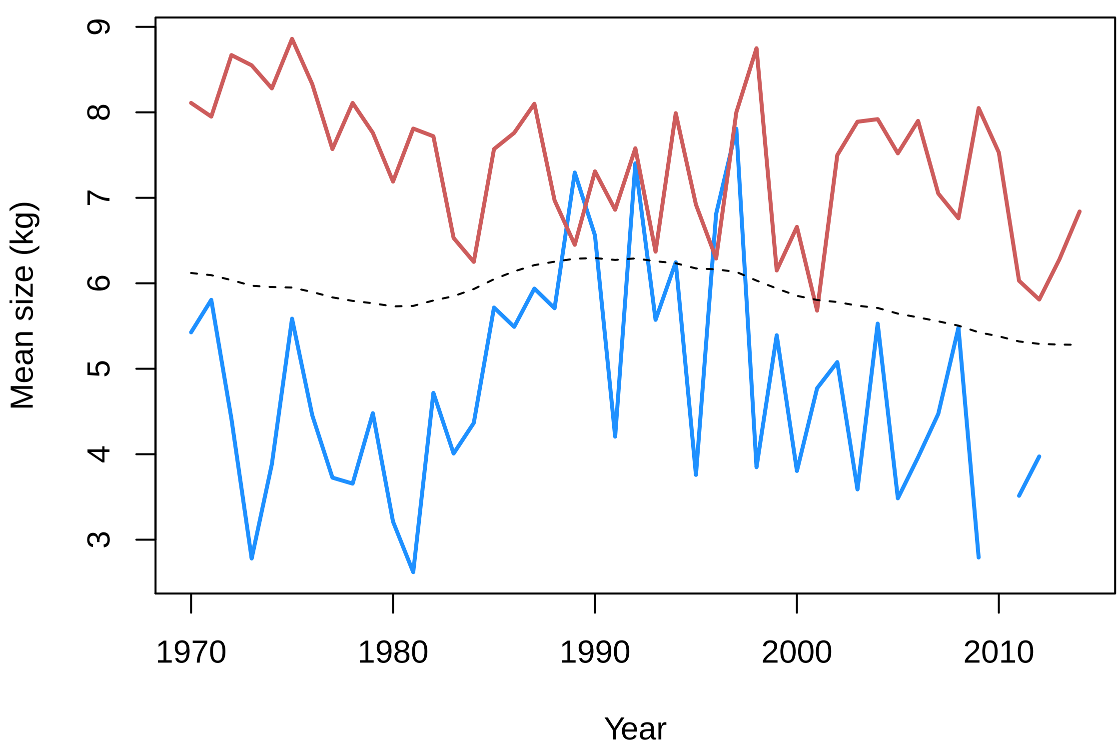

Figure 3. Time series of observed size (blue) and the fit for the random walk model (Eqn 3; black).

Two patterns over time

For the model with two different states of nature, $\mathbf{Z}$ is a $[2 \times 2]$ identity matrix, $\mathbf{x}_t$ is a $[2 \times 1]$ vector of the true states, and $\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$, such that

$$\begin{bmatrix} y_{\text{Tengu}} \\ y_{\text{WDFW}} \end{bmatrix}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_t + \begin{bmatrix} a_{\text{Tengu}} \\ a_{\text{WDFW}} \end{bmatrix} + \begin{bmatrix} v_{\text{Tengu}} \\ v_{\text{WDFW}} \end{bmatrix}_t \tag{5}$$

$$\begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_t = \begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_{t-1} + \begin{bmatrix} w_{\text{Tengu}} \\ w_{\text{WDFW}} \end{bmatrix}_t \tag{6}$$

```
## model defn for Eqns 8 & 9
mod_list <- list(
  B = diag(2),
  U = matrix("u", nrow = 2, ncol = 1),
  # Q = matrix(list("T", 0, 0, "W"), 2, 2),
  Q = matrix(list("q", 0, 0, "q"), 2, 2),
  Z = diag(2),
  A = matrix(c(0, 0), nrow = 2, ncol = 1),
  # R = matrix(list("T", 0, 0, "W"), 2, 2)
  R = matrix(list("r", 0, 0, "r"), 2, 2)
)

## fit Eqns 8 & 9
size_both_2 <- MARSS(t(log(yy)), model = mod_list)

## Success! abstol and log-log tests passed at 53 iterations.
## Alert: conv.test.slope.tol is 0.5.
## Test with smaller values (<0.1) to ensure convergence.
##
## MARSS fit is
## Estimation method: kem
## Convergence test: conv.test.slope.tol = 0.5, abstol = 0.001
## Estimation converged in 53 iterations.
## Log-likelihood: 19.86981
## AIC: -29.73961   AICc: -29.0079
##
##             Estimate
## R.r          0.02822
## U.u         -0.00494
## Q.q          0.00244
## x0.X.tengu  1.54875
## x0.X.wdfw   2.12301
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

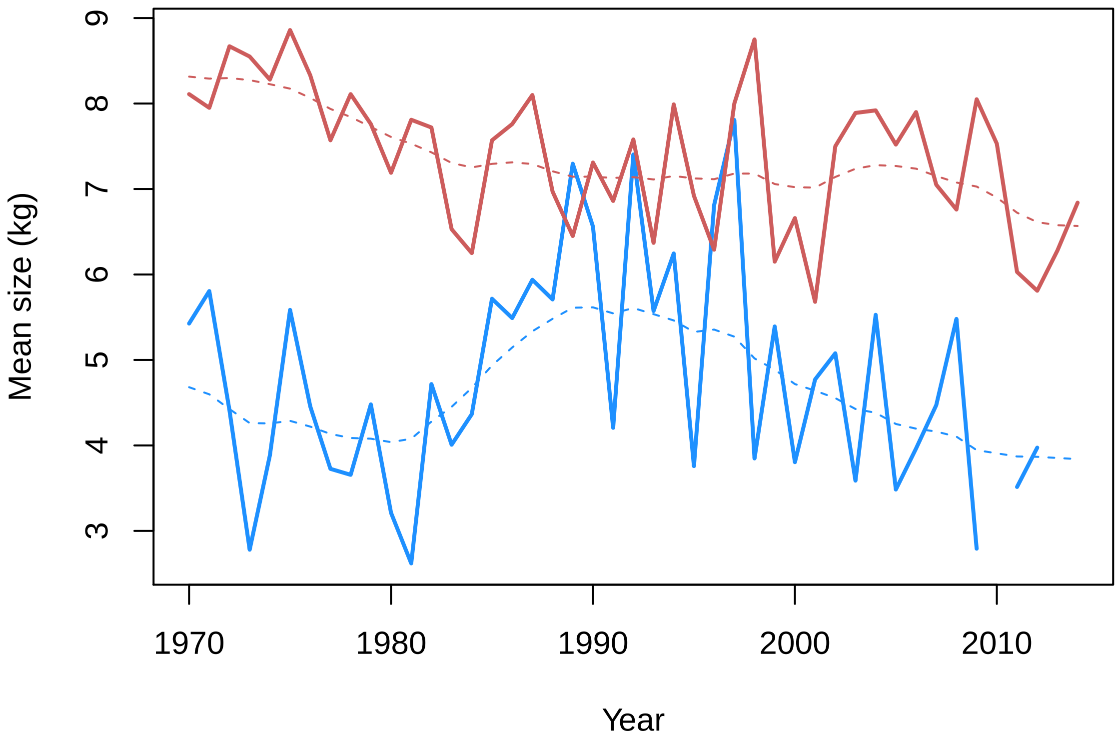

Figure 4. Time series of observed fish size from the Tengu derby (blue) and WDFW surveys (red), including fits from the multivariate random walk model (Eqns 8 & 9; dashed).

Summary of size comparison

The model with one common state has an AICc value of -21.2 and the model with two unique states has an AICc value of -29, which indicates rather modest support for two unique temporal patterns in the data.