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

Supplementary material to accompany Quinn et al.

Mark Scheuerell

10 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 decided to model the proportion of fish caught in a given year that were greater than the 2 size thresholds. To prevent numerical errors from the necessary logit transform I replaced any `0`'s or `1`'s with `0.001` and `0.999`, respectively.

```
## proportion of fish >5 lbs & >10 lbs
size_props <- tengu_data %>%
  mutate(
    ## proportion of fish >5 lbs
    p5 = case_when(
      n_over_5 == total_catch ~ 0.9999,
      TRUE ~ n_over_5 / total_catch),
    ## proportion of fish >10 lbs
    p10 = case_when(
      n_over_10 == 0 ~ 0.0001,
      TRUE ~ n_over_10 / total_catch)) %>%
  ## select cols of interest
  select(year, p5, p10) %>%
  ## remove years prior to 1950 with all NA's
  filter(year > 1949)
```

Here are plots of the two proportional size metrics over time.

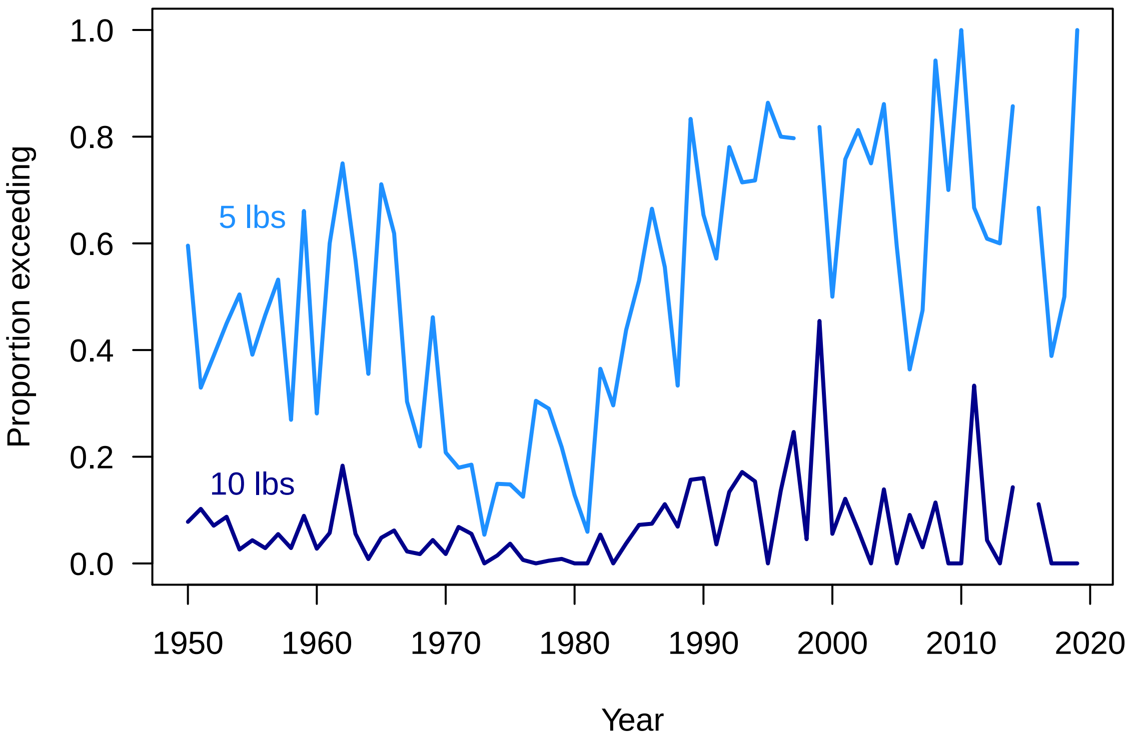

Figure 1. Time series of the proportion of fish caught that exceeded 5 pounds (light blue) and 10 pounds (dark blue).

There appears to be an overall decline in fish size from 1950 until the early 1980s, when fish sizes increased rapidly before declining again until present.

Univariate random walk models

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 those over 5 pounds. The models take the general form

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

$$y_t = x_t + a + v_t \tag{2}$$

where x_t is the true, but unobserved size of fish in year t, u is the bias term, y_t is the observed proportion of fish exceeding a size threshold, a is an offset, and w_t and v_t are normally distributed process and observation errors, respectively.

By fitting models with and without the bias (u) term, we can evaluate the data support for a trend over time. For models with a bias term, we can also obtain a confidence interval around u.

Fit biased random walk

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

## response for fish >5 lbs
l_size_5 <- matrix(qlogis(size_props$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 69 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 69 iterations.
## Log-likelihood: -165.9049
## AIC: 339.8099   AICc: 340.4349
##
##       Estimate
## R.r     6.0223
## U.u    -0.0611
## Q.q     0.2135
## x0.x0  -2.4454
## 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 84 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 84 iterations.
## Log-likelihood: -132.3186
## AIC: 272.6371   AICc: 273.272
##
##       Estimate
## R.r     2.4572
## U.u     0.0425
## Q.q     0.0684
## x0.x0  -0.4689
## 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 59 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 59 iterations.
## Log-likelihood: -166.4277
```

```
## AIC: 338.8555   AICc: 339.2247
##
##       Estimate
## R.r      5.980
## Q.q      0.265
## x0.x0   -2.758
## 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 66 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 66 iterations.
## Log-likelihood: -133.0377
## AIC: 272.0754   AICc: 272.4504
##
##       Estimate
## R.r      2.421
## Q.q      0.099
## x0.x0   -0.182
## 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 (-0.181, 0.058), and the model with a bias term has an AIC value that is actually -1.2 units greater 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 (-0.026, 0.111), and the model with a bias term has an AIC value that is -0.8 units greater than the model without a bias term, again suggesting there is essentially no data support for a systematic downward trend in the log-transformed size over time.

Model fit to proportional size data

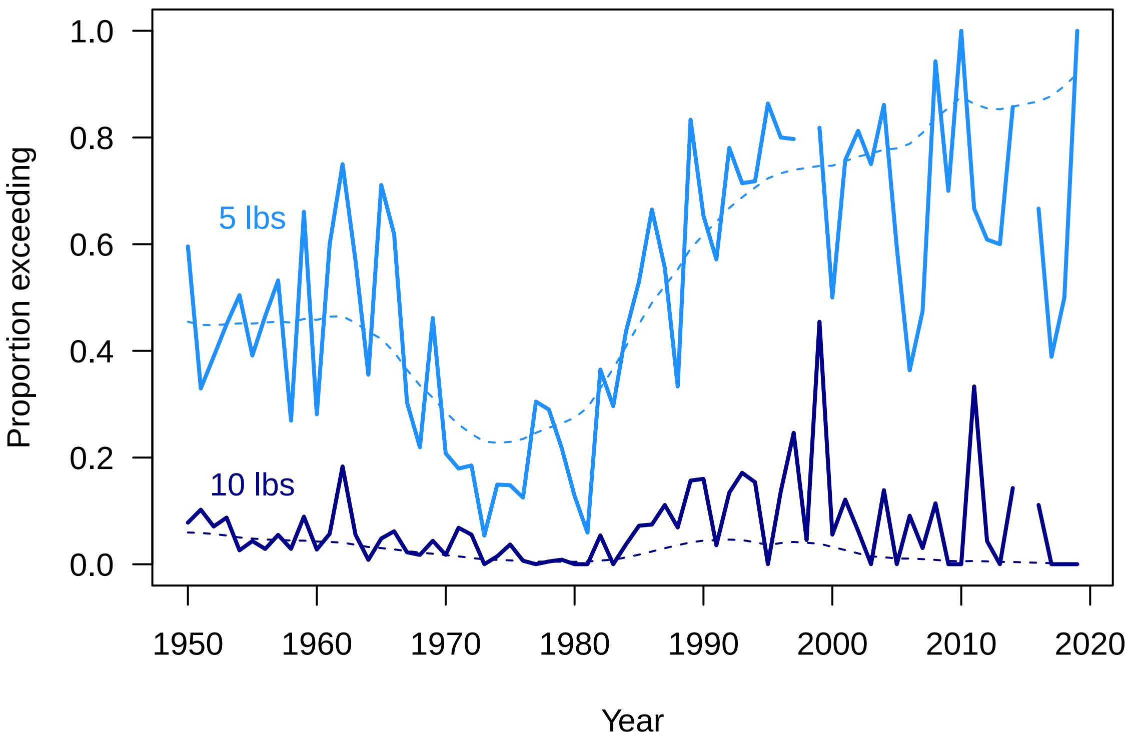

Figure 2. Time series of the proportion of fish caught that exceeded 5 pounds (solid, light blue) and 10 pounds (solid, dark blue), and the fitted values from their respective models (dashed lines).

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

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \tag{3}$$

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t \tag{4}$$

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 identically distributed, such that

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

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

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

$$x_t = x_{t-1} + w_t \tag{6}$$

```r
## get mean of top-5; code as NA if <5 fish were weighed
tengu_sizes <- tengu_data %>%
  filter(year >= 1970 & year <= 2014) %>%
  select(starts_with("size")) %>%
  apply(1, mean)

## get corresponding size data from WDFW
wdfw_sizes <- wdfw_data %>%
  filter(year >= 1970 & year <= 2014) %>%
  select(size)

## combined lengths into matrix for MARSS
yy <- cbind(tengu = tengu_sizes, wdfw = wdfw_sizes)

## 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("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.
```

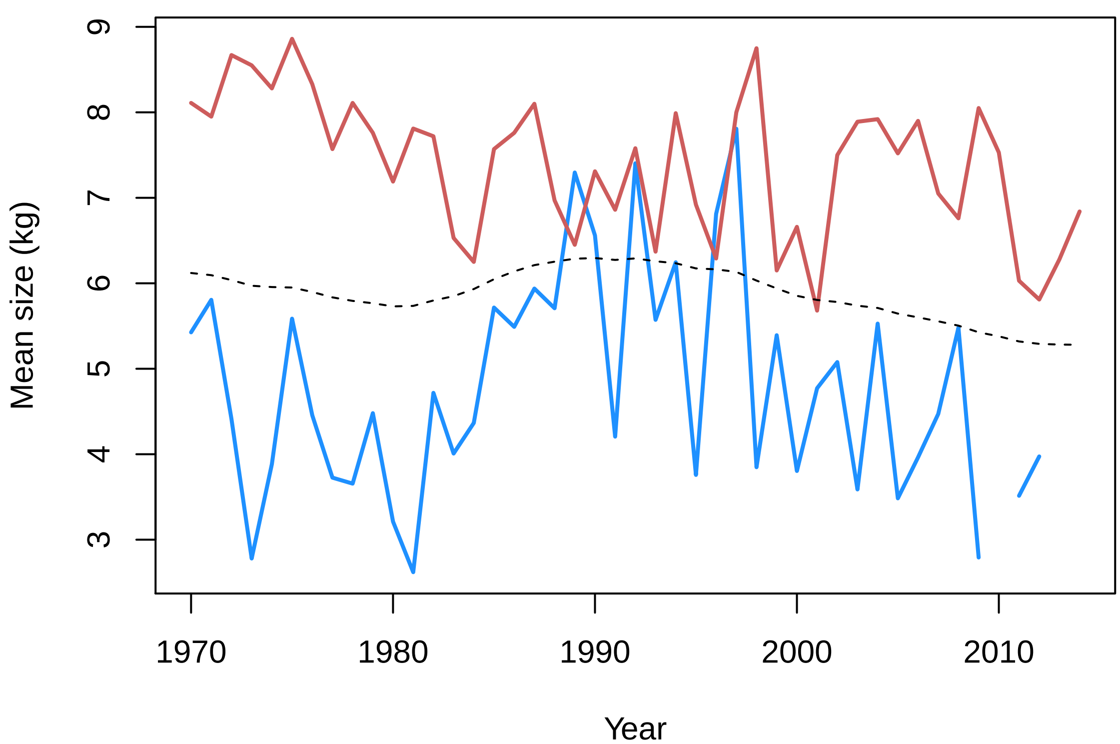

Figure 3. Time series of observed fish size from the Tengu derby (blue) and WDFW surveys (red), including fits from the multivariate random walk model (Eqns 6 & 7; dashed).

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

$$\begin{bmatrix} y_{\text{Tengu}} \\ y_{\text{WDFW}} \end{bmatrix}_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_t + \begin{bmatrix} a_{\text{Tengu}} \\ a_{\text{WDFW}} \end{bmatrix} + \begin{bmatrix} v_{\text{Tengu}} \\ v_{\text{WDFW}} \end{bmatrix}_t \tag{7}$$

$$\begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_t = \begin{bmatrix} x_{\text{Tengu}} \\ x_{\text{WDFW}} \end{bmatrix}_{t-1} + \begin{bmatrix} w_{\text{Tengu}} \\ w_{\text{WDFW}} \end{bmatrix}_t \tag{8}$$

```
## 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.size    2.12301
## Initial states (x0) defined at t=0
##
## Standard errors have not been calculated.
## Use MARSSparamCIs to compute CIs and bias estimates.
```

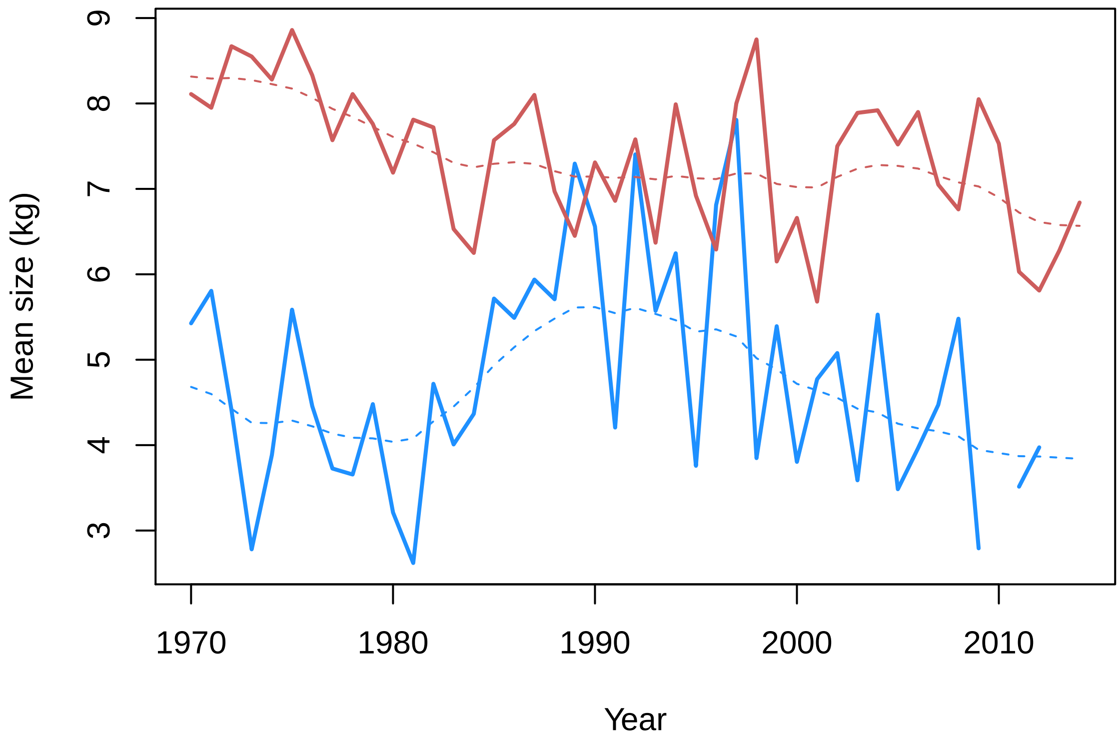

Figure 4. Time series of observed fish size from the Tengu derby (blue) and WDFW surveys (red), including fits from the multivariate random walk model (Eqns 8 & 9; dashed).

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.