Tiny Assessment Model (tinyAM) :: CHEATSHEET

Formula-driven age-structured stock assessment

1 Setup tidy inputs

Four tables: `catch`, `index`, `weight`, `maturity`. Keys: year–age. Indices also require `survey`; `fleet` is optional for catch. Optional `weight$type`: stock (biomass) or catch (yield); catch weights may be shared or fleet specific. Stock weights are required and also used for yield if catch weights are absent.

`catch`

year	age	obs	unit
1983	2	18	thousand fish
1983	3	2585	thousand fish
⋮	⋮	⋮	⋮
2025	19	0	thousand fish
2025	20	0	thousand fish

Predictors: F in `catch`; q in `index`; recruitment/N/M in optional `pop_covariates`.

`index` (thousand fish)

year	age	survey	samp_time	obs	q_block
1983	2	Fall 2J3KL	0.8	698822	2
1983	3	Fall 2J3KL	0.8	517528	3
⋮	⋮	⋮	⋮	⋮	⋮
2025	19	Fall 2J3KL	0.8	0.0	6-20
2025	20	Fall 2J3KL	0.8	0.0	6-20

`weight`

year	age	obs	unit
1983	2	0.0704	kg
1983	3	0.201	kg
⋮	⋮	⋮	⋮
2025	19	12.8	kg
2025	20	13.7	kg

`maturity`

year	age	obs
1983	2	0.00003
1983	3	0.00486
⋮	⋮	⋮
2025	19	1.00000
2025	20	1.00000

Example: `cod_obs`. Assemble and validate tables:

```
dat <- make_tam_data(
  catch = catch_data,
  index = index_data,
  weight = weight_data,
  maturity = maturity_data
)
```

2 Specify model hypotheses

These examples resemble statistical catch-at-age (SCAA), state-space assessment model (SAM), and bespoke Northern cod assessment model (NCAM) structures. They illustrate alternative assumptions using the same observations.

SCAA: fixed recruitment and mortality effects

Deterministic cohort survival; annual recruitment parameters; separable fishing mortality; fixed natural mortality.

```
scaa <- fit_tam(
  dat, years = 1983:2025, ages = 2:14,
  N_settings = list(process = "off", init = "exp"),
  rec_settings = list(process = "off", mu_form = ~ factor(year)),
  F_settings = list(process = "off",
    mu_form = ~ factor(year) + factor(age), mean_ages = 5:14),
  M_settings = list(process = "off",
    mu_supplied = ~ I(0.3), mean_ages = 5:14),
  catch_settings = list(sd_form = ~ 1),
  index_settings = list(sd_form = ~ 1, q_form = ~ mono(q_block))
)
```

SAM: stochastic abundance and fishing mortality

Abundance deviations, random-walk recruitment, and correlated random walks in fishing mortality; natural mortality stays fixed.

```
sam <- update(scaa,
  N_settings = list(process = "iid"),
  rec_settings = list(process = "rw", mu_form = NULL),
  F_settings = list(process = "cor_rw", mu_form = NULL)
)
```

NCAM: variation in fishing and natural mortality

Deterministic cohort survival, random-walk recruitment, and AR1 variation in fishing mortality and age-blocked natural mortality.

```
ncam <- update(scaa,
  rec_settings = list(process = "rw", mu_form = NULL),
  F_settings = list(process = "ar1", mu_form = ~ 1),
  M_settings = list(process = "ar1", age_breaks = c(2, 5, 8, 14))
)
```

Use `check_tam()` to inspect numerical diagnostics before interpreting estimates.

3 Visualize and compare

Explore model fits

```
vis_tam(scaa, sam, ncam)
```

Local dashboard / online: Explore these three example fits and their diagnostics.

Test survey forecast skill

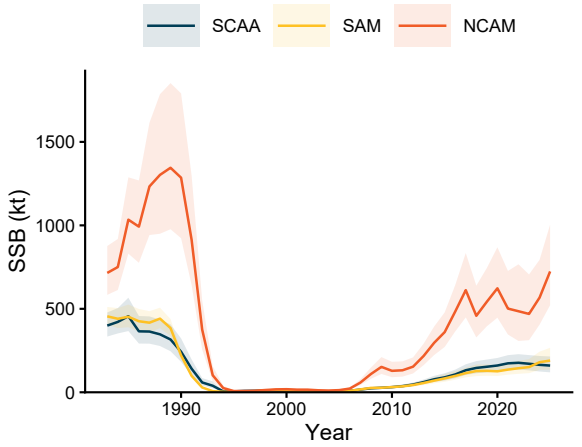
```
hind <- fit_hindcast(scaa, sam, ncam)
```

Assess survey forecast skill conditional on observed catches, and retrospective stability. Inspect `hind$fold_diagnostics` before comparing scores.

Access tables for report figures

```
out <- tidy_tam(scaa, sam, ncam)
ssb <- out$pop$ssb
index <- out$obs_pred$index
```

Population estimates and observation diagnostics are tidy tables, ready for static plots and report summaries.



Spawning stock biomass (kt). Lines show model estimates; shaded bands show 95% confidence intervals.

tinyAM :: MODEL REFERENCE

Settings, equations and the population assumptions they represent

Abundance and survival

N is start-of-year abundance. Each year, survivors move forward one age; survivors in the terminal age remain in the plus group. Let $Z = F + M$; recruitment is $R_t = N_{t,1}$.

$$\begin{aligned}\widehat{N}_{t,a} &= N_{t-1,a-1}e^{-Z_{t-1,a-1}},\\ \widehat{N}_{t,A} &= N_{t-1,A-1}e^{-Z_{t-1,A-1}}\\ &\quad + N_{t-1,A}e^{-Z_{t-1,A}},\end{aligned}$$

$$\log N_{t,a} = \log \widehat{N}_{t,a} + X^N_{t,a}\beta_N + \eta^N_{t,a}.$$

Optional `N_settings$mu_form` adds proportional apparent gains/losses; η^N is the process deviation. `init = "exp"` initializes older ages from exponential survivorship; alternatives: `"free"`, `"latent"`, `"mono"`, `"capped"`. Indices $1, \dots, A$ run from the youngest modeled age to the plus group.
For $t \geq 2$, non-plus ages use the first survival equation.

Recruitment

`rec_settings`: `process = "off"` estimates a mean without process deviations; `mu_form = ~ factor(year)` gives annual recruitment. `"rw"` gives a random walk. With IID/AR1 residuals, `mu_form = ~ bh(ssb)` or `~ ricker(ssb)` adds a stock–recruit curve.

$$\begin{aligned}\text{RW: } \log R_t - \log R_{t-1} &= \epsilon_t,\\ \log R_t &= \log g(SSB_{t-L}) + u_t,\\ g_{BH}(S) &= \alpha S/(1+\beta S),\\ g_{Ricker}(S) &= \alpha S e^{-\beta S}.\end{aligned}$$

$\alpha, \beta > 0$; $\epsilon_t \sim \mathcal{N}(0, \sigma_R^2)$. IID: $u_t = \epsilon_t$; AR1: $u_t = \phi_R u_{t-1} + \epsilon_t$, with a stationary start. Curves describe medians. Lag L defaults to the youngest biological age. Boundary R is free by default; `init = "latent"` integrates it. No RW + curve.

Biomass and spawning biomass

Spawning biomass sums mature biomass across biological ages after mortality up to spawning.

$$SSB_t = \sum_b n_{t,b} w_{t,b} p_{t,b} e^{-\tau_{t,b} Z_{t,b}}.$$

Biomass is $\sum_b n_{t,b} w_{t,b}$. b: retained biological age; n: its abundance; w: stock weight; p: mature fraction; τ: year fraction from year start to spawning (default 0; set `maturity$spawn_time` for age/year-specific values). Hidden ages retain biology/timing and share plus-age mortality.

Mean fishing mortality

Specify mean fishing mortality with `F_settings$mu_form`. For example, `mu_form = ~ factor(year) + factor(age)` gives separable year and age effects. With `process = "off"`, F follows this mean exactly. See `?for_fleet` for fleet-specific formulas.

$$\begin{aligned}S(a) &= \{1 + e^{-k(a-a_{50})}\}^{-1}, \quad k > 0,\\ \log F_{t,a,f} &= X_{t,a,f}\beta_F + \log S_f(a) + \eta^F_{t,a,f}.\end{aligned}$$

Total F is $\sum_f F_{t,a,f}$. Different shapes can be specified with `logistic()`, `double_logistic()`, or smooths such as `s()` and `t2(age, year)`. Check mean/process separation.

Natural mortality

`M_settings`: `mu_supplied = ~ M0` supplies the baseline mortality surface.

$$\log M_{t,a} = \log M^0_{t,a} + \eta^M_{t,a}.$$

`M0` comes from `pop_covariates`. M states share age blocks and start in year 2 by default; other cells retain the mean. `mu_form` adds log-mean effects.

Variation in N, F and M

N/F/M: `"off"`, `"iid"`, `"rw"`, `"ar1"`; F also `"cor_rw"`. Off means no deviations; RW starts are unpenalized. Curve/smooth F means exclude intrinsic RW. Fleets have independent processes; `cor_form = ~ fleet` separates their correlations. `sd_form` groups process SDs.

$$\begin{aligned}\text{IID: } \eta_{t,a} &\sim \mathcal{N}(0, \sigma_a^2),\\ \text{RW: } \eta_{t,a} - \eta_{t-1,a} &\sim \mathcal{N}(0, \sigma_a^2).\end{aligned}$$

AR1 stationary covariance (state/block distances):

$$\text{Cov}(\eta_{t,a}, \eta_{s,b}) = \frac{\sigma_a \sigma_b \phi_t^{|t-s|} \phi_a^{|a-b|}}{(1-\phi_t^2)(1-\phi_a^2)}.$$

F `"cor_rw"`: independent annual increment vectors with

$$\Sigma_{ab} = \sigma_a \sigma_b \rho^{|a-b|}, \quad -1 < \rho < 1.$$

Survey catchability

`q_block` leaves age patterns unrestricted; `mono(q_block)` imposes non-decreasing catchability under default log link.

$$\eta_{q,k} = \beta_q + \sum_{j<k} d_j, \quad d_j \geq 0, \qquad q_k = e^{\eta_{q,k}}.$$

Zero steps give exact plateaus. With `q_link = "logit"`:

$$q_k = \{1 + e^{-\eta_{q,k}}\}^{-1}.$$

`~ logistic(age)` gives a logistic age pattern. Terms `iid()`, `rw()` and `ar1()` add Gaussian variation; q covariates come from `index`. See `?for_survey` for survey-specific formulas.

Catch and survey observations

Catch and survey observations are linked to population quantities by the predictions below; observed log values have Gaussian errors.

$$\begin{aligned}\widetilde{C}_{t,a,f} &= N_{t,a}\frac{F_{t,a,f}}{Z_{t,a}}(1 - e^{-Z_{t,a}}),\\ \tilde{I}_i &= q_i N_{t_i,a_i} e^{-s_i Z_{t_i,a_i}},\\ \log Y_i &\sim \mathcal{N}(\log \widetilde{Y}_i, \sigma_i^2),\\ \sigma_i &= \sigma_i^0 \exp(X_{sd,i}\beta_{sd}).\end{aligned}$$

Y_i, $\widetilde{Y}_i$: observed and predicted catch/index; s_i: survey fraction. Catch/index `sd_form = ~ fleet` or `~ survey` gives group SDs. Errors are independent. Predictions are medians; zero/NA observations are excluded. Yield is $\sum_{a,f} C_{t,a,f} w^{catch}_{t,a,f}$; biomass/SSB use stock weights. Forecasts average each weight type/fleet separately.

Fit and prediction checks

`check_tam()` checks numerical and statistical diagnostics; `fit_hindcast()` evaluates survey forecasts given observed removals and retrospective stability. The joint likelihood includes active process, formula-effect and observation densities. RTMB integrates random effects by Laplace approximation; intervals use local curvature. Standardized residuals $(\log Y_i - \log \widetilde{Y}_i)/\sigma_i$ depend on fitted, shrunken states. Convergence does not establish identifiability. Hindcasts report survey log-RMSE, MASE, scored/failed folds and Mohn's ρ.