

Guidelines for the use of the stochastic production model in continuous time (SPiCT)

T.K. Mildenerberger, A. Kokkalis, C.W. Berg

27 September, 2019

This is a living document, be sure to check for the latest update ([link](#)). The SPiCT package is actively developed, check for the most recent package version and report problems at ([link](#)).

Document version	SPiCT version	Comments
v 1.0	v 1.2.8	Initial version (2019-09-23)

This document provides specific guidelines for the use of the stochastic production model in continuous time (SPiCT; Pedersen and Berg, 2017). The following is divided into three parts, containing: (i) the main assumptions and data requirements of SPiCT, (ii) a checklist for the acceptance of a SPiCT assessment, and (iii) other helpful tips.

Most of the following guidelines and tips are accompanied by sample code that assumes

- input data is contained in a list called ‘inp’,
- the results are contained in a list called ‘fit’ after fitting SPiCT to the input data with `fit <- fit.spict(inp)`, and
- calculating the one-step-ahead residuals with `fit <- calc.osa.resid(fit)`.

Main assumptions and input data for SPiCT

- Catch data should be representative of both landings and bycatch. It is also possible to use a time series of landings, but then the interpretation of the results is changed. If available, seasonal catches should be used as input. Catches are assumed to be taken over a period of time (e.g. years or quarters). The time vector in SPiCT should reflect the beginning of each catch interval (e.g. 2002.25 and 2002.75 for the second and fourth quarter catches, respectively). Additionally, the vector `inp$dte` should reflect the length of each time interval (e.g. 1 for annual and 0.25 for quarterly catches, respectively).
- Stock size indices should be in terms of biomass (not numbers) and representative of the exploitable stock biomass (ESB). In many cases, the gear selectivity of the commercial and scientific fleets do not coincide and thus the stock size indices have to be corrected to exclude smaller individuals that are not recruited to the commercial fleet.
- Biomass indices are assumed to be snapshots on given time points. Therefore, the timing of survey indices has to be given as decimal years reflecting the timing of the survey in the vector `inp$timeI` (e.g. 1995.5 for the middle of the year). Commercial CPUE indices should be associated with the midpoint of the interval of the corresponding catches, i.e. middle of the year if they are based on yearly aggregated catches and effort.
- If input data spans long periods and there is ecological evidence for potential changes in productivity, it is possible to fit productivity regime shifts or long-term gradual changes in productivity.
- The Euler discretisation has to be finer than the data, i.e. for yearly data the Euler time step (`inp$dteuler`) has to be smaller than 1. A value of 1 changes the model to a discrete time production model with different model assumptions. The default value is 1/16 year, which is sufficient in most cases. The function `check.euler(fit)` allows to check if the discretisation is fine enough.

Checklist for the acceptance of a SPiCT assessment

ICES category 3 stocks can be managed using the official advice rule based on SPiCT (3.1.1 and 3.1.2 in ICES, 2018), which requires an accepted SPiCT assessment.

1. The assessment **converged** (`fitoptconvergence` equals 0).
2. All **variance parameters** of the model parameters are **finite** (`all(is.finite(fit$sd))` should be TRUE).
3. **No violation of model assumptions** based on one-step-ahead residuals (bias, auto-correlation, normality). This means, that p-values are insignificant (≤ 0.05), indicated by green titles in the graphs of `spictplo.diagnostics(fit)`. Slight violations of these assumptions do not necessarily invalidate model results.
4. **Consistent patterns in the retrospective analysis** (`fit <- fit.retro(fit)`). This means that there is no tendency of consistent under- or overestimation of the relative fishing mortality ($\frac{F}{F_{MSY}}$) and relative biomass ($\frac{B}{B_{MSY}}$) in successive assessment. The retrospective trajectories of those two quantities should be inside the confidence intervals of the base run.
5. **Realistic production curve**. The shape of the production curve should not be too skewed ($\frac{B_{MSY}}{K}$ should be between 0.1 and 0.9). Low values of $\frac{B_{MSY}}{K}$ allow for an infinite population growth rate (`calc.bmsyk(fit)`).
6. **High assessment uncertainty** can indicate a lack of contrast in the input data or violation of the ecological model assumptions. The main variance parameters (`logsdb`, `logsdc`, `logsdi`, `logsdf`) should not be unrealistically high. Confidence intervals for $\frac{B}{B_{MSY}}$ and $\frac{F}{F_{MSY}}$ should not span more than 1 order of magnitude (`calc.om(fit)`).
7. **Initial values do not influence the parameter estimates** (`fit <- check.ini(fit)`). The estimates should be the same for all initial values (`fit$ckeck.ini$resmat`). Runs which did not converge should not be considered in this regard.

Optional model and assessment tuning in case of assessment rejection

Disclaimer: An uncertain assessment is okay; the model should not be tweaked and parameters should not be fixed unless there is sufficient information and evidence to make such assumptions. Rather than avoiding uncertainty, it should be accounted for by means of stochastic harvest control rules.

- Increase **iterations of optimisation** Special error code (8) exceeded max.iter.. `inp$optimiser.control = list(iter.max = 1e4, eval.max = 1e4)`
- If the catch time series is longer than the survey(s), **shortening the catch time series** to cover only the period where there is an available biomass index may help model convergence. `inp = shorten.inp(inp, 2005, 2018)`
- **Adjust** the parameter determining the **shape of the production curve** ‘logn’ or impose a prior (be aware that this modification might cause smaller confidence bands).
- A prior for n based on a meta analysis, e.g. for all species pooled based on Thorson et al. (2011):
`inp$priors$logn <- c(log(log(1.478)), 0.6, 1)`
 - Tighter Schaefer prior for n (or any other prior e.g. from meta studies): `inp$priors$logn <- c(log(2), 0.5, 1)`
 - Fixing n to resemble the Schaefer production model (or the meta study, alternatively):
`inpinilogn <- log(2); inp$phases$logn <- -1`

- Use a **prior for the initial depletion level** based on available information, FOR EXAMPLE:
 - If evidence or expert knowledge allows to infer that there was low or no exploitation before the beginning of the available data: initial depletion level could be assumed to be close to the carrying capacity (e.g. `inp$priors$logbkfrac <- c(log(0.8),0.5,1)`)
 - If evidence or expert knowledge allows to infer that there was low or no exploitation before the beginning of the available data: initial depletion level could be assumed to be a certain fraction of the carrying capacity (e.g. `inp$priors$logbkfrac <- c(log(0.2),0.5,1)`)
- If information on the level of uncertainty in the biomass index and/or the uncertainty of the catch is available prior distribution could be used for the observation error term of the indices ('`logsdi`') and catches ('`logsdci`'), respectively. This requires to remove the priors for the ratios of process to observation errors ('`logalpha`') and ('`logbeta`'), respectively.