**GeMS: A generalized management strategy evaluation framework for fisheries**

**Main Message:** GeMS is a general simulation framework that evaluates the performance of fisheries management strategies under a range of underlying population dynamics.

**Abstract**

Generalized Management Strategy Evaluation (GeMS) provides a flexible simulation platform to quantitatively answer questions in fisheries management while incorporating the uncertainty of the management process itself. GeMS allows for time-varying parameters in all population processes, using a two-box, age- and length-structured operating model. Environmental change, shifts in species distribution, and exploitation can change population processes over time. Time-variation in many population processes can have dramatic impacts on sustainable yields, but this variability is often not incorporated into management. Here, we demonstrate the current capabilities of the GeMS framework using two examples. We conclude by discussing potential uses for GeMS and how it complements the existing tools used for management strategy evaluations.

**Introduction**

One of the main goals of natural resource management is to ensure the sustainable exploitation of a resource, with clearly described trade-offs and acceptable outcomes to all stakeholders. Effectively managing an exploited population requires an understanding of the way a population will respond to changes in exploitation. This can be a difficult task because of the uncertainties inherent in the understanding of an exploited population’s dynamics and implementation of management. In recent years, management strategy evaluation (MSE; Smith, 1994) has been proposed as the gold standard for the testing of management strategies under uncertainty (Punt, 2008; Punt et al., 2016; Rademeyer et al., 2007). MSE is a process of creating simulated fisheries (the ‘operating model’) and testing different assessment tools (the ‘estimation model’) and harvest control rules to evaluate their performance in metrics like long term yield, variability in yield, and probability of overfishing (A’mar et al., 2009; Punt et al., 2014; Rademeyer et al., 2007). In doing so, it allows for quantitatively comparing the performance of a suite of management strategies given some assumed underlying dynamics.

MSE has been used for many fisheries around the globe to compare and select management strategies. MSE was pioneered by de la Mare (1986) for the International Whaling Commission, and more recently, and has been used by Hurtado-Ferro et al. (2010) for Japanese sardine (*Sardinops melanostictus*), Wetzel and Punt (2011) for US west coast flatfish and groundfish, and Smith et al. (2014) for Australian fisheries (see Punt et al. (2016) for more examples). Butterworth and Punt (1999) recommended the development of “case-specific” procedures over “generic” ones, yet they also acknowledged that a major hurdle to conducting MSEs is a general lack of modelling skills and of software packages that can implement “generalized” operating models. In response, the Fisheries Library for R (FLR; Kell et al., 2007), an open-source generic MSE framework, was developed to allow scientists to conduct MSEs. However, FLR does not allow for time-varying components or spatial structure in its operating models, which will likely be increasingly important in management given a changing climate and shifting stocks (Szuwalski and Hollowed, 2016). Operating models, often used to estimate quantities used in management, are a simplification of the natural world, although it is often unclear how well these simplifications represent nature. As such, operating models can take on a “hypothesis-oriented approach” (Kell et al., 2006; Punt et al., 2014), where alternative operating models incorporating expert beliefs and different assumptions about population drivers are evaluated.

Ecological processes can vary over time due to external factors such as climate change (e.g. Hollowed et al., 2013), fishing behavior (e.g. Szuwalski et al., 2017), regime shifts (e.g. Conners et al., 2002; Szuwalski and Punt, 2013a), and these processes will likely continue to fluctuate in the future. However, parameters in assessment models used in management (such as natural mortality *M* and growth parameters *Κ* and *L∞*) are often either fixed or estimated as time-invariant despite evidence to the contrary (e.g. Thorson et al., 2015; Webber and Thorson, 2016). Furthermore, not considering temporal variation in creating harvest control rules could potentially result in failed management of the stock (e.g. Hurtado-Ferro et al., 2010; Lee et al., 2018). Concurrently, the effects of temporal variation on management vary (see A’mar et al., 2009; Mueter et al., 2011 for studies on walleye pollock *Theragra chalcogramma*), and a mechanistic understanding of how a population fluctuates over time is often unavailable (Thorson, 2015). As such, management strategies robust to fluctuations in the dynamics of the population are desirable (Holt and Punt, 2009; Punt et al., 2014; Szuwalski and Punt, 2013b).

The true impact of time-varying processes on a population remains unknown, but simulation allows managers to evaluate management alternatives over different projected ‘states of nature’. MSEs conducted on walleye pollock evaluated the performance of management strategies under scenarios of climate change (A’mar et al., 2009; Ianelli et al., 2011). These studies found that under changing environmental conditions, keeping the management strategy at status quo had a high probability of resulting in an overfished stock. In the Bering Sea, status quo harvest control rules for snow crab (*Chionoecetes opilio*) were found to result in the stock being potentially falsely declared overfished due to regime shifts in the system dynamics (Szuwalski and Punt, 2013b). Based on these studies, it is clear that management strategies robust to temporal variation are important for the sustainable management of fisheries. Spatial variation can also play an important role in determining appropriate management strategies. Ying et al. (2011) showed that managing three subpopulations as a single unit was likely to result in overexploitation at a local level for small yellow croaker (*Larimichthys polyactis*). These case-specific evaluations in systems that are often data-rich represent significant investment of time and resources that may not be possible in many fishery systems across the world.

Here we present a Generalized Management Strategy Evaluation framework (GeMS) that allows for the testing of management strategy performance over a range of scenarios. GeMS allows for time-variation in every population process in the operating model, and also can accommodate movement between two spatially-distinct populations. GeMS only requires inputs of life history parameters to begin an MSE, but time series of historical fishing effort and recruitment from an existing stock assessment can also be incorporated. From this starting point, fisheries managers can quickly begin the process of a ‘generalized’ management strategy evaluation. This may help stakeholders to make a decision on whether effort and resources should be invested in a more exhaustive MSE for their system. GeMS is written in R (R Core Team, 2017) and Auto Differential Model Builder (ADMB; Fournier et al., 2012), two open-source tools, in the hope of ensuring ease of use, reproducibility, and transparency (Anderson et al., 2014; Kell et al., 2007).

**Framework**

*General Overview*

MSE is accomplished by simulating populations with characteristics similar to the target population, drawing data with error from the simulated populations required to assess the stock using a given management strategy, applying an assessment method to estimate quantities used in management (e.g. the current biomass and biological reference points), and finally using a specified harvest control rule to determine the allowable catches in a given year (Figure 1). The calculated allowable catches are removed from the simulated population and process is repeated into the ‘future’ to mimic the feedback that occurs in the management of an actual fishery (A’mar et al., 2009; Punt et al., 2016; Rademeyer et al., 2007). GeMS uses this general framework, broken into three components:

1. Operating model: The operating model for GeMS is a two-area, age- and length-structured population dynamics model with capacity for time-varying parameters in all population and management processes.
2. Estimation model: Currently, two estimation models exist: a Schaeffer model, and an age- and length-structured model with the capacity to allow growth, natural mortality, and selectivity to vary over time.
3. Harvest control rule (HCR): Currently, available HCRs are a fixed exploitation rate, a fixed catch, and a sloped control rule based on proxies of target biomasses and fishing mortalities.

A full description of the equations defining each component can be found in the appendix.

*Basic steps to GeMS*

The package can be forked from Github at https://github.com/szuwalski/GeMS. After it is downloaded, all that is needed to begin an MSE is a control file. The control file dictates the entire MSE process and includes parameters of the operating model, the properties of the estimation model, and the type of HCR that should be applied. GeMS comes with a master control file (Master\_CTL.csv) to serve as a template, as well as several examples based on a cod-like stock from on Hurtado et al. (2015). A function call run\_GeMS() initiates the MSE, and produces basic diagnostics and output, which can then be analyzed by the user according to their needs.

*Example 1: Performance of a production model*

The surplus production model (Ricker, 1975) is one of the simplest models of fish population dynamics, only considering changes in exploitable biomass of the fishable stock (Polacheck et al., 1993). Written in the Schaefer form (Schaefer, 1954), this model is computationally simple, and requires few data sources (only catch and an index of relative abundance) to compute management reference points (BMSY and FMSY). However, the model also makes assumptions that may over-simplify population characteristics, such as assuming all individuals are identical within a population, which might result in inappropriate management advice.

An example MSE that examines the management performance of a surplus production model for a population with cod-like characteristics with three different levels of productivity (as seen through the ‘steepness’ parameter of the stock-recruit relationship) is included in the folder ‘inst/extdata/Cod\_1\_Production’ of the GeMS github repository. A production model was selected in the control file for each scenario (i.e. in “Cod\_Base\_CTL.csv”, “Cod\_HighProd\_CTL.csv”, and “Cod\_LowProd\_CTL.csv”) as the desired assessment method and the MSE was executed by running code similar to this R code:

library(GeMS)

CurrentDirectory<-“C:/GeMS/MyDir”

OMNames <-c(“Cod\_Base\_CTL”, “Cod\_HighProd\_CTL”, “Cod\_LowProd\_CTL”)

run\_GeMS(MSEdir=CurrentDirectory,CTLNameList=OMNames)

In general, production models underestimate the fishing mortality that would produce maximum sustainable yield in this example (figure 2). Consequently, the stock is underexploited (figure 3). The extent to which the optimal catches are underestimated declines as the productivity of the stock decreases. Figures 2 and 3 were produced from the output of “ProductionModelOutput”, which writes figures in .PNG format to a folder named ‘Plots’ in the working directory. A step by step walkthrough of how to manipulate steepness to explore changes in management performance can be found in the wiki. This simple example illustrates how to evaluate the impact of an estimation model on management performance by changing one parameter at a time in the underlying population dynamics and comparing results. With this information, an investigator will have a better understanding of the circumstances under which a particular estimation model would provide appropriate estimates of management-related quantities and can focus research priorities and data collection endeavors. Production models did not provide unbiased management advice for these scenarios, but age-structured models may provide better estimates.

*Example 2: Retrospective patterns in age-structured models*

Age-structured integrated assessment models, the general form of which is described in Walters (1969), include variables such has natural mortality rates, vulnerability to fishing, and stock-recruitment relationships (Lawson and Hilborn, 1985). Integrated analysis models combine a variety of data sources into a single analysis (Fournier and Archibald, 1982; Maunder and Punt, 2013), and provide advice on stock status and management quantities (ICES, 2012). These assessment models often estimate management quantities better than other types of models (Butterworth and Rademeyer, 2008; Punt et al., 2002; Radomski et al., 2005).

However, time-variation in population processes can hamper the ability of assessments to estimate quantities important in management without bias (e.g. Johnson et al., 2014; Lee et al., 2018; Szuwalski et al., 2017b). Retrospective patterns are defined as systematic biases in estimates of derived quantities from a model, given increasing years of data (Mohn, 1999). When retrospective patterns are strong, perhaps due to a misspecified stock assessment model, the model could be rendered unsuitable for management purposes (Hurtado-Ferro et al., 2015). For example, a Pacific halibut (*Hippoglossus stenolepis*) stock assessment was found to have consistently overestimated biomass and underestimated harvest rates (Clark and Hare, 2008; Valero, 2012). Valero (2012) found that this resulted in inappropriate harvest strategies being set for managing the stock. One of the ways in which this could be mitigated is by allowing a process (such as growth, natural mortality, or selectivity) to vary over the time, even if the “true” time-varying process is unknown (Szuwalski et al., 2017b). However, reducing retrospective patterns does not necessarily translate to unbiased management reference points, and GeMS can be used to investigate such a question.

GeMS can simulate circumstances that give rise to retrospective biases by allowing time-variation in population processes like growth, natural mortality, and selectivity in the operating model. Using the cod-like population again as an example, we present an example MSE with three scenarios. The first scenario applied the age-structured assessment method to data drawn from an operating model in which the assumptions of the assessment matched the operating model’s dynamics. The second and third scenarios were run with control files in which the parameter governing natural mortality in the operating model (“NatMn”) was input as a vector, with a sudden change of the parameter value in the middle of the time series. The performance of an assessment method in which natural mortality varied was then compared to an assessment in which natural mortality was static. The 10 replicates of the MSE was conducted for each scenario, and each replicate was projected 20 years, beginning with 50 years of historical data for assessment. GeMS can be ran in parallel for larger jobs by changing a few optional switches in ‘run\_GeMS’ (see code below.The example files can be found in the folder inst/extdata/Cod\_5\_AgeStructure of the Github repository.

run\_GeMS(MSEdir = CurrentDirectory, # directory in which the control files reside

CTLNameList = OMNames, # list of control files

runparallel = T, # use parallel processing

cores = 2, # use 2 cores in parallel

ADoptions = “-gbs 2000000000” # memory management.in the ADMB model)

First and foremost, the scenario in which the assumptions in the assessment matched the operating model dynamics returned unbiased estimates of quantities used in management (figure 5). This sets an important benchmark and verifies that the structure of the assessment method is correct. However, when applying the same assessment method to data drawn from the operating model in which natural mortality varied over time, retrospective patterns and biased management quantities resulted (as expected, figure 5). Retrospective patterns and biases (to some extent) were corrected in the assessment in which natural mortality was allowed to vary (figure 5). The way in which the time-invariant assessment methods ‘accommodate’ time-variation can be seen by the plots of estimated population process (figure 6). In this case, the assessment method in which natural mortality was fixed ‘accommodated’ variation in natural mortality by negative biases in recruitment before the change in M, positive biases in estimated F, and negative biases in fishery selectivity. A more detailed analysis of retrospective patterns using GeMS can be found in Szuwalski et al. (2017b).

**Discussion**

GeMS provides a simple method to quickly and quantitatively evaluate the performance of management strategies under different states of nature. It goes beyond the capabilities of most other MSE packages by allowing for variation in all processes in the operating model and providing the option of spatial dynamics, both of which will likely be important considerations under a changing climate. GeMS is not meant as a replacement for stock assessment and will not be able to capture all the idiosyncrasies of existing assessments. Therefore, GeMS is meant as a starting point for stakeholders to assess their needs for a full MSE using other software such as FLR (Kell et al., 2007). In addition to the questions described above, GeMS could help to answer questions including (but not limited to) the following:

* What is the value of improved assessment methods or data for assessment?
* How would implementing an MPA or changing gear types impact maximum sustainable yield?
* How does the performance of a management strategy using a production model change (in terms of long term yield) under scenarios in which climate change influences somatic growth?
* How does movement between two populations impact the performance of age-structured assessment methods?
* How would changes in species interactions (via range shifts, for example) influence the potential yield of a given species?

The code for GeMS is open source and published on Github so that users can provide feedback continuously, and an open dialogue between the developers and the users can be established. We hope that GeMS can efficiently evolve to meet additional needs and improve its capabilities through a transparent development and feedback process.

Acknowledgements

Packard Foundation, SFG nerds, Santa Barbara City for the nice weather to work in.

**References**

A’mar, Z.T., Punt, A.E., Dorn, M.W., 2009. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. ICES J. Mar. Sci. J. Cons. 66, 1614–1632. https://doi.org/10.1093/icesjms/fsp044

Anderson, S.C., Monnahan, C.C., Johnson, K.F., Ono, K., Valero, J.L., 2014. ss3sim: An R Package for Fisheries Stock Assessment Simulation with Stock Synthesis. PLOS ONE 9, e92725. https://doi.org/10.1371/journal.pone.0092725

Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985–998. https://doi.org/10.1006/jmsc.1999.0532

Butterworth, D.S., Rademeyer, R.A., 2008. Statistical catch-at-age analysis vs. ADAPT-VPA: the case of Gulf of Maine cod. ICES J. Mar. Sci. 65, 1717–1732. https://doi.org/10.1093/icesjms/fsn178

Clark, W.G., Hare, S.R., 2008. Assessment of the Pacific halibut stock at the end of 2007. Int Pac Halibut Comm Rep. Assess. Res. Act. 2007, 117–203.

Conners, M.E., Hollowed, A.B., Brown, E., 2002. Retrospective analysis of Bering Sea bottom trawl surveys: regime shift and ecosystem reorganization. Prog. Oceanogr. 55, 209–222. https://doi.org/10.1016/S0079-6611(02)00079-4

de la Mare, W.K., 1986. Simulation Studies on Management Procedures. Rep. Int. Whal. Comm. 36, 429–450.

Fournier, D., Archibald, C.P., 1982. A General Theory for Analyzing Catch at Age Data. Can. J. Fish. Aquat. Sci. 39, 1195–1207. https://doi.org/10.1139/f82-157

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., Sibert, J., 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27, 233–249. https://doi.org/10.1080/10556788.2011.597854

Hollowed, A.B., Barange, M., Beamish, R.J., Brander, K., Cochrane, K., Drinkwater, K., Foreman, M.G.G., Hare, J.A., Holt, J., Ito, S., Kim, S., King, J.R., Loeng, H., MacKenzie, B.R., Mueter, F.J., Okey, T.A., Peck, M.A., Radchenko, V.I., Rice, J.C., Schirripa, M.J., Yatsu, A., Yamanaka, Y., 2013. Projected impacts of climate change on marine fish and fisheries. ICES J. Mar. Sci. 70, 1023–1037. https://doi.org/10.1093/icesjms/fst081

Holt, C.A., Punt, A.E., 2009. Incorporating climate information into rebuilding plans for overfished groundfish species of the U.S. west coast. Fish. Res., Ecosystem Approach to Fisheries: Improvements on Traditional Management for Declining and Depleted StocksAnnual Meeting of the North Pacific Marine Science Organization 100, 57–67. https://doi.org/10.1016/j.fishres.2009.03.002

Hurtado-Ferro, F., Hiramatsu, K., Shirakihara, K., 2010. Allowing for environmental effects in a management strategy evaluation for Japanese sardine. ICES J. Mar. Sci. 67, 2012–2017. https://doi.org/10.1093/icesjms/fsq126

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. 72, 99–110. https://doi.org/10.1093/icesjms/fsu198

Ianelli, J.N., Hollowed, A.B., Haynie, A.C., Mueter, F.J., Bond, N.A., 2011. Evaluating management strategies for eastern Bering Sea walleye pollock (Theragra chalcogramma) in a changing environment. ICES J. Mar. Sci. J. Cons. 68, 1297–1304. https://doi.org/10.1093/icesjms/fsr010

ICES, 2012. Report on the Classification of Stock Assessment methods developed by SISAM, ICES CM 2012/ACOM/SCICOM:01.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-pre, K.A., Anderson, S.C., Cunningham, C.J., Hurtado-Ferro, F., Licandeo, R.R., Muradian, M.L., Ono, K., Szuwalski, C.S., Valero, J.L., Whitten, A.R., Punt, A.E., 2014. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES J. Mar. Sci. J. Cons. fsu055. https://doi.org/10.1093/icesjms/fsu055

Kell, L.T., De Oliveira, J.A.A., Punt, A.E., McAllister, M.K., Kuikka, S., 2006. Operational management procedures: an introduction to the use of management strategy evaluation frameworks., in: The Knowledge Base for Fisheries Management, Developments in Aquaculture and FIsheries Science. Elsevier, Amsterdam, pp. 379–407.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D., 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci. 64, 640–646. https://doi.org/10.1093/icesjms/fsm012

Lawson, T.A., Hilborn, R., 1985. Equilibrium Yields and Yield Isopleths from a General Age-Structured Model of Harvested Populations. Can. J. Fish. Aquat. Sci. 42, 1766–1771. https://doi.org/10.1139/f85-221

Lee, Q., Thorson, J.T., Gertseva, V.V., Punt, A.E., 2018. The benefits and risks of incorporating climate-driven growth variation into stock assessment models, with application to Splitnose Rockfish (Sebastes diploproa). ICES J. Mar. Sci. 75, 245–256. https://doi.org/10.1093/icesjms/fsx147

Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61–74. https://doi.org/10.1016/j.fishres.2012.07.025

Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488. https://doi.org/10.1006/jmsc.1999.0481

Mueter, F.J., Bond, N.A., Ianelli, J.N., Hollowed, A.B., 2011. Expected declines in recruitment of walleye pollock (Theragra chalcogramma) in the eastern Bering Sea under future climate change. ICES J. Mar. Sci. 68, 1284–1296. https://doi.org/10.1093/icesjms/fsr022

Polacheck, T., Hilborn, R., Punt, A.E., 1993. Fitting Surplus Production Models: Comparing Methods and Measuring Uncertainty. Can. J. Fish. Aquat. Sci. 50, 2597–2607. https://doi.org/10.1139/f93-284

Punt, A.E., 2008. Refocusing Stock Assessment in Support of Policy Evaluation, in: Tsukamoto, K., Kawamura, T., Takeuchi, T., Beard Jr, T.D., Kaiser, M.J. (Eds.), Fisheries for Global Welfare and Environment: Memorial Book of the 5th World Fisheries Congress 2008. pp. 139–152.

Punt, A.E., A’mar, T., Bond, N.A., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., Haltuch, M.A., Hollowed, A.B., Szuwalski, C., 2014. Fisheries management under climate and environmental uncertainty: control rules and performance simulation. ICES J. Mar. Sci. J. Cons. 71, 2208–2220. https://doi.org/10.1093/icesjms/fst057

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., Haddon, M., 2016. Management strategy evaluation: best practices. Fish Fish. 17, 303–334. https://doi.org/10.1111/faf.12104

Punt, A.E., Smith, A.D.M., Cui, G., 2002. Evaluation of management tools for Australia’s South East Fishery.2. How well can management quantities be estimated? Mar. Freshw. Res. 53, 631–644. https://doi.org/10.1071/mf01008

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rademeyer, R.A., Plagányi, É.E., Butterworth, D.S., 2007. Tips and tricks in designing management procedures. ICES J. Mar. Sci. 64, 618–625. https://doi.org/10.1093/icesjms/fsm050

Radomski, P., Bence, J.R., Quinn II, T.J., 2005. Comparison of virtual population analysis and statistical kill-at-age analysis for a recreational, kill-dominated fishery. Can. J. Fish. Aquat. Sci. 62, 436–452. https://doi.org/10.1139/f04-228

Ricker, W.E., 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bull. Fish. Res. Board Can. 191, 382.

Schaefer, M.B., 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm. Bull. 1, 23–56.

Smith, A.D.M., 1994. Management strategy evaluation – the light on the hill, in: Hancock, D.A. (Ed.), Population Dynamics for Fisheries Management: Australian Society for Fish Biology Workshop Proceedings, Perth, 24-25 August 1993. Australian Society for Fish Biology, Perth, Western Australia, pp. 249–253.

Smith, A.D.M., Smith, D.C., Haddon, M., Knuckey, I.A., Sainsbury, K.J., Sloan, S.R., 2014. Implementing harvest strategies in Australia: 5 years on. ICES J. Mar. Sci. 71, 195–203. https://doi.org/10.1093/icesjms/fst158

Szuwalski, C.S., Burgess, M.G., Costello, C., Gaines, S.D., 2017a. High fishery catches through trophic cascades in China. Proc. Natl. Acad. Sci. 114, 717–721. https://doi.org/10.1073/pnas.1612722114

Szuwalski, C.S., Hollowed, A.B., 2016. Climate change and non-stationary population processes in fisheries management. ICES J. Mar. Sci. 73, 1297–1305. https://doi.org/10.1093/icesjms/fsv229

Szuwalski, C.S., Ianelli, J.N., Punt, A.E., 2017b. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES J. Mar. Sci. https://doi.org/10.1093/icesjms/fsx159

Szuwalski, C.S., Punt, A.E., 2013a. Regime shifts and recruitment dynamics of snow crab, Chionoecetes opilio, in the eastern Bering Sea. Fish. Oceanogr. 22, 345–354. https://doi.org/10.1111/fog.12026

Szuwalski, C.S., Punt, A.E., 2013b. Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. ICES J. Mar. Sci. 70, 955–967. https://doi.org/10.1093/icesjms/fss182

Thorson, J.T., 2015. Spatio-temporal variation in fish condition is not consistently explained by density, temperature, or season for California Current groundfishes. Mar. Ecol. Prog. Ser. 526, 101–112.

Thorson, J.T., Hicks, A.C., Methot, R.D., 2015. Random effect estimation of time-varying factors in Stock Synthesis. ICES J. Mar. Sci. J. Cons. 72, 178–185. https://doi.org/10.1093/icesjms/fst211

Valero, J.L., 2012. Harvest policy considerations on retrospective bias and biomass projections. Int Pac Halibut Comm Rep. Assess. Res. Act. 2011 311–329.

Walters, C.J., 1969. A Generalized Computer Simulation Model for Fish Population Studies. Trans. Am. Fish. Soc. 98, 505–512. https://doi.org/10.1577/1548-8659(1969)98[505:AGCSMF]2.0.CO;2

Webber, D.N., Thorson, J.T., 2016. Variation in growth among individuals and over time: A case study and simulation experiment involving tagged Antarctic toothfish. Fish. Res., Growth: theory, estimation, and application in fishery stock assessment models 180, 67–76. https://doi.org/10.1016/j.fishres.2015.08.016

Wetzel, C.R., Punt, A.E., 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fish. Res. 110, 342–355. https://doi.org/10.1016/j.fishres.2011.04.024

Ying, Y., Chen, Y., Lin, L., Gao, T., 2011. Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. 68, 2101–2120. https://doi.org/10.1139/f2011-116

**Appendix**

*Operating model*

Population dynamics model

The population dynamics model is a single-sex, age-structured model that tracked the number of individuals in an age class by year, *N*a,y, and allowed natural mortality, *M*y to vary over time.

|  |  |  |
| --- | --- | --- |
|  |  | 1 |
|  |  | 2 |
|  |  | 3 |

Mean recruitment followed a Beverton-Holt stock-recruit relationship parameterized in terms of steepness (*h*; the expected proportion of virgin recruitment entering the population when spawning biomass is 20% of virgin levels), *R*0 (virgin recruitment) and *SB*0 (virgin spawning biomass). Recruitment variation was log-normally distributed with mean 0 and standard deviation σr.

|  |  |  |
| --- | --- | --- |
|  |  | 4 |
|  |  | 5 |

Annual fishing mortality, *F*y, was comprised of a separable annual selectivity at length curve in the estimation model as a logistic curve:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  | 6 |

where and determine the length *l* at which the probabilities of capture are respectively 50% and 95% in year l. Selectivity parameters were specified in the operating model in terms of length (equation 6) and converted to age conditional upon the length-at-age curve by rearranging equation 8 (which is conditional upon growth). Time variation in selectivity parameters in the operating model was specified by a vector input by the user. Maturity, *Mat*a, is also a logistic function of age and two estimated parameters (p50 and p95, which were the age at which the probability of maturing is 50% and 95%, respectively) and assumed constant over time:

|  |  |  |
| --- | --- | --- |
|  |  | 7 |

Length-at-age followed a growth increment von Bertalanffy form:

|  |  |  |
| --- | --- | --- |
|  |  | 8 |
|  |  | 8a |
|  |  | 8b |

where is the maximum length during year *y*, *ky* was the growth rate in year *y* and *t0* was the age corresponding to a predicted length of 0. Changes in *k*y were specified as a vector in the operating model. Weight was a constant function of length:

|  |  |  |
| --- | --- | --- |
|  |  | 9 |

Conditional length-at-age for both the catch and survey were calculated from the numbers at age matrix by specifying a single standard deviation σlen and the expected length-at-age determined using equation 9. The array, *LAa,z,y* with *a* rows representing the number of age classes and *z* columns representing the number of length bins during year *y* contained the probability of an individual of age *a* being length *z* (i.e. the proportion of each age class in each length bin). Length bins were pre-specified.

|  |  |  |
| --- | --- | --- |
|  |  | 10 |
|  |  | 11 |

where Binz was a vector of the midpoints of the specified length bins. Numbers at length were calculated from this matrix of (normalized) probabilities of length-at-age by multiplying each row by the number of individuals at age and then summing over rows (i.e. age).

|  |  |  |
| --- | --- | --- |
|  |  | 12 |

The catch weight during year *y* was calculated as:

|  |  |  |
| --- | --- | --- |
|  |  | 13 |

The survey-selected biomass at the time of the survey was calculated as:

|  |  |  |
| --- | --- | --- |
|  |  | 14 |

where was the survey selectivity defined as:

|  |  |  |
| --- | --- | --- |
|  |  | 15 |

As with fishery selectivity, the parameters associated with survey selectivity (and , the length at which the probability of being selected in the fishery is 50% and 95%, respectively) were specified according to length and then transformed to age within the operating model conditional upon growth. Survey selectivity at length was constant over time.

*Data simulation*

Catch biomass, catch length frequencies, fishery-independent survey indices of abundance and survey length frequencies were generated using the operating model with error to be used in the estimation models for each year in the simulation. Observed catch biomass and survey biomass were modeled as:

|  |  |  |
| --- | --- | --- |
|  |  | 17 |
|  |  | 18 |

where εyis a normally distributed random variable with a mean of 0 and a pre-specified standard deviation σε. Values for the coefficients of variation for all data sources are in Table 1. Observed length frequencies for both the catch and the survey are generated by sampling the true numbers at length calculated above using the ‘sample’ function in R a specified number of times (see Table 1 for sample sizes).

*Assessment method and harvest control rule*

The equations governing the population dynamics within the assessment matched those of the operating model. Table 1 lists the estimable parameters of the assessment method. Annual deviations from the mean were additional fixed-effects parameters with specified penalties (see ‘likelihood components’ below) when natural mortality, growth or selectivity were time-varying natural mortality. Within the estimation method, survey selectivity was estimated based on age and growth parameters were not estimated in the ‘base’ model. Average recruitment (μR) within the assessment method was estimated with annual deviations (Rdev,y); eq. 19).

|  |  |  |
| --- | --- | --- |
|  |  | 19 |

Fishery selectivity at age was estimated within the assessment as deviations around a mean.

|  |  |
| --- | --- |
|  | 20 |

*Likelihood components*

The assessment method was fit to the data generated from the operating model based on four likelihood components. The log-likelihoods (ignoring constants) for catch and the survey index of abundance were log-normal:

|  |  |  |
| --- | --- | --- |
|  |  | 20 |
|  |  | 21 |

Catch and survey length frequencies were fit under the assumption of multinomial sampling:

|  |  |  |
| --- | --- | --- |
|  |  | 22 |
|  |  | 23 |

where was the observed proportion at length of the catch in year *y*, was the predicted proportion at length in year *y* of the catch, was the observed proportion at length of the survey biomass in year *y*, and was the predicted proportion at length of the survey biomass in year *y*. The data were weighted with the same CVs and sample sizes with which they were generated (Table 1). Small penalties were added to the objective function to ensure the smoothness of estimated recruitment, fishing mortality, time-varying natural mortality, time-varying selectivity, and time-varying growth in the form of:

|  |  |  |
| --- | --- | --- |
|  |  | 24 |

where is the vector of the deviations associated with recruitment, fishing mortality, natural mortality, selectivity and/or growth and γx is the weight applied to each likelihood component (specified as a CV of ~0.41 for all quantities because it produced estimates that were not overly variable, but still contributed little to the objective function).

*Harvest control rule*

Management targets used in the harvest control rule were based on proxies for the biomass at which maximum sustainable yield occurs (*B*MSY) and the fishing mortality that produces that biomass at equilibrium (*F*MSY) using spawning-biomass-per-recruit methods (e.g. Clark, 1991; NPFMC, 2007). *F*35%, or the fishing mortality that reduces spawning biomass per recruit (SBPR) to 35% of virgin levels is used as a target fishing mortality for Alaskan crab fisheries (NPFMC, 2007). *B*35% was calculated as the SBPR corresponding to *F*35% multiplied by an average recruitment calculated from the entire time series of estimated recruitments. Calculated values of *F*35% and *B*35% were used in conjunction with a control rule to adjust the proportion of *F*35% that was applied to the population based on the status of the population relative to *B*35%. The fishing mortality derived from equation 24 was deemed the fishing mortality corresponding to the TAC (which coincides with the OFL), the *F*OFL, and was applied to the population to find the TAC using equation 14.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | 25 |
| Where, | |  | |
|  | | The currently estimated mature biomass in the projected year for TAC determination | |
|  | | Mature biomass resulting from fishing at | |
|  | | Fishing mortality that reduces the mature biomass per recruit to 35% of the unfished level | |
|  | | Determines the slope of the descending limb of the control rule (specified as 0.25 here) | |
|  | | Fraction of *B*35% below which fishing mortality is zero (specified as 0.05 here) | |

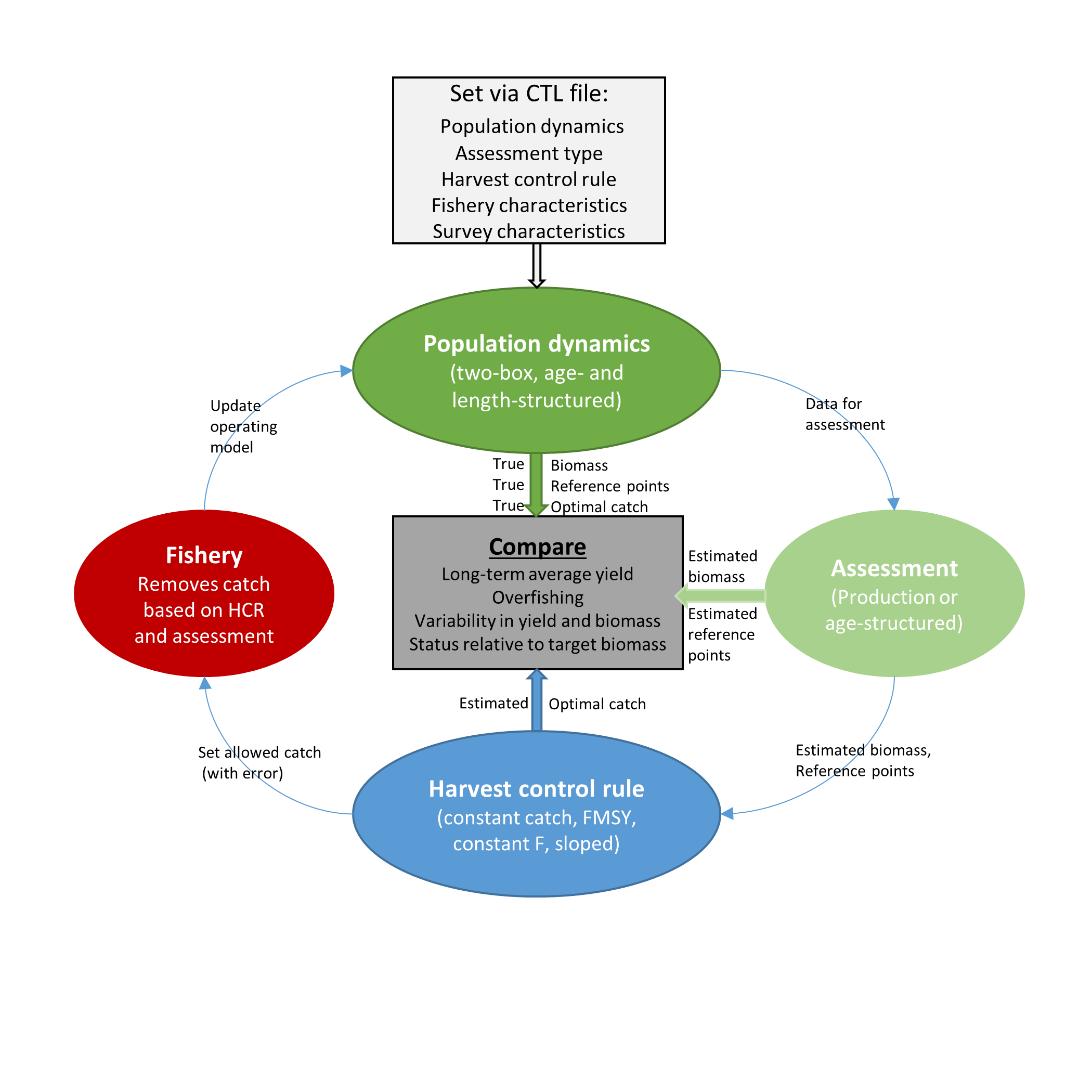


Figure 1. Flow chart representing the flow of information through a management strategy evaluation in GeMS.

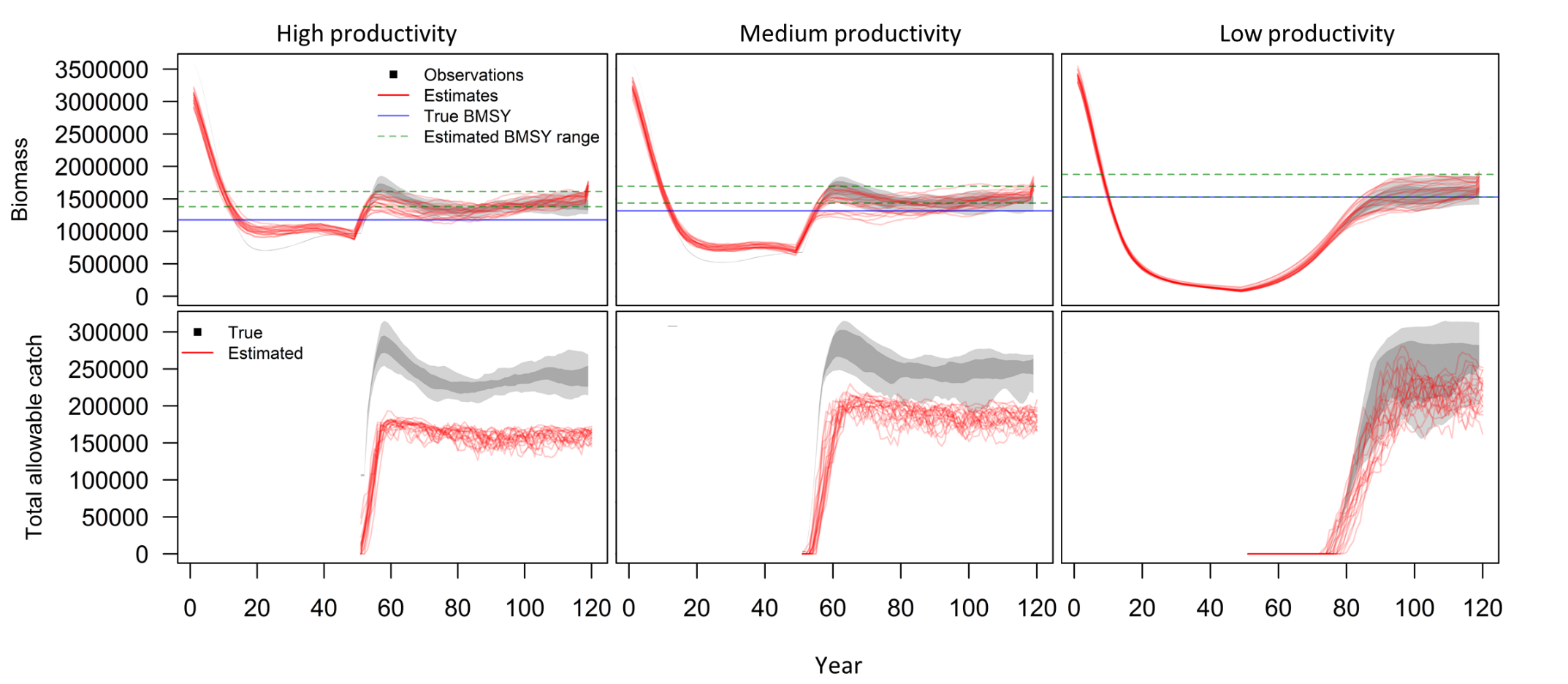


Figure 2. A comparison of the fits to biomass and estimated catch recommendations from a production model applied to simulated populations with high, medium, and low productivity (i.e. steepness equals 0.8, 0.65, 0.4 in the stock recruit relationship).

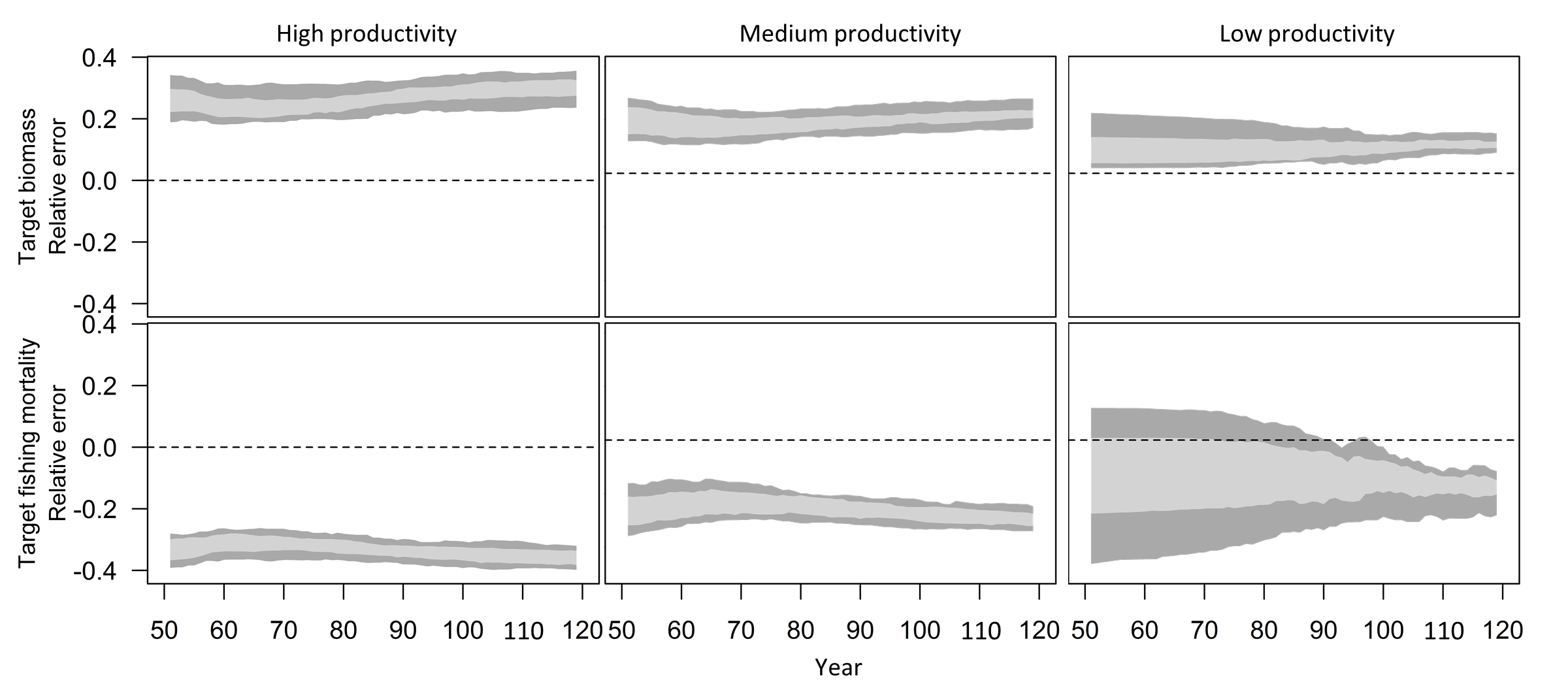


Figure 3. Relative error in estimated target biomasses and target fishing mortalities in each year of the projection period from a production model applied to simulated populations with high, medium, and low productivity (i.e. steepness equals 0.8, 0.65, 0.4 in the stock recruit relationship).

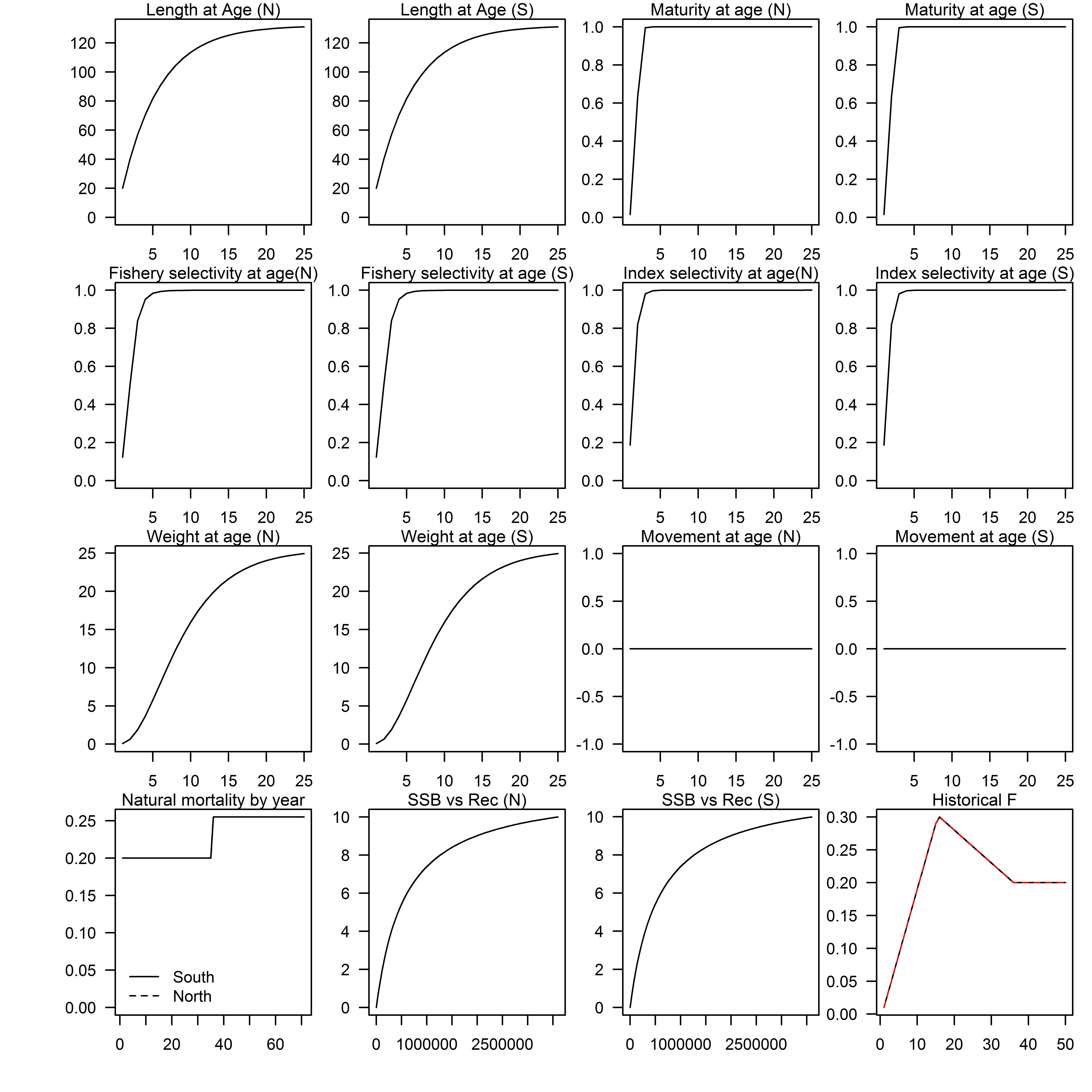


Figure 4. Standard output describing the life history processes in the operating model. In this case, only natural mortality varies over time, but all processes have the capacity to change.

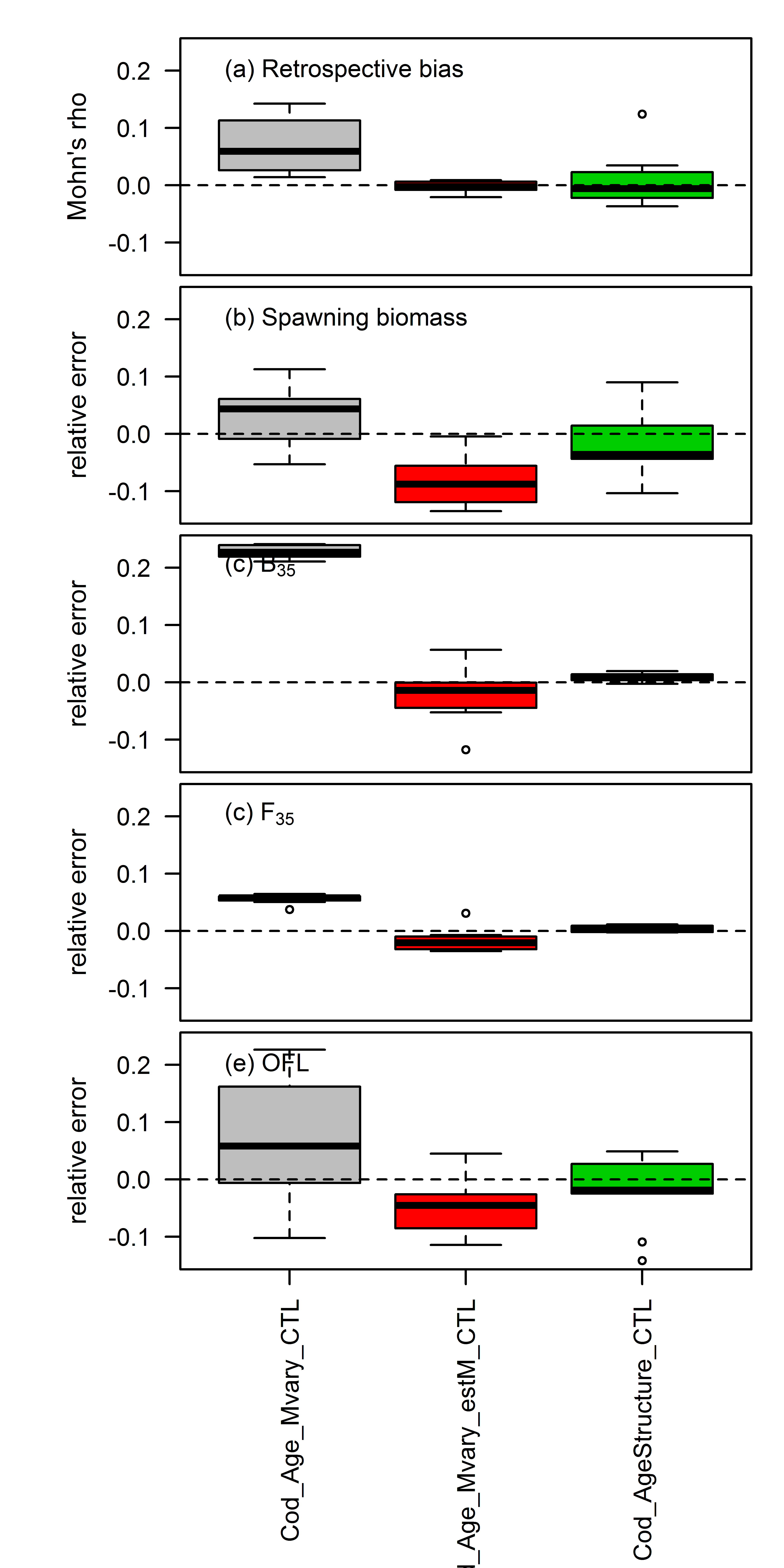


Figure 5. A comparison of the estimated management quantities from age-structured assessment methods. “Cod\_Age\_Mvary\_CTL” is an assessment method that fixes natural mortality at a single value applied to an operating model in which natural mortality varies over time. “Cod\_Age\_Mvary\_estM\_CTL” is an assessment method that estimates a time-varying natural mortality applied to an operating model in which natural mortality varies over time. “Cod\_AgeStructure\_CTL” is an assessment method that fixes natural mortality at a single value applied to an operating model in which natural mortality is constant over time.

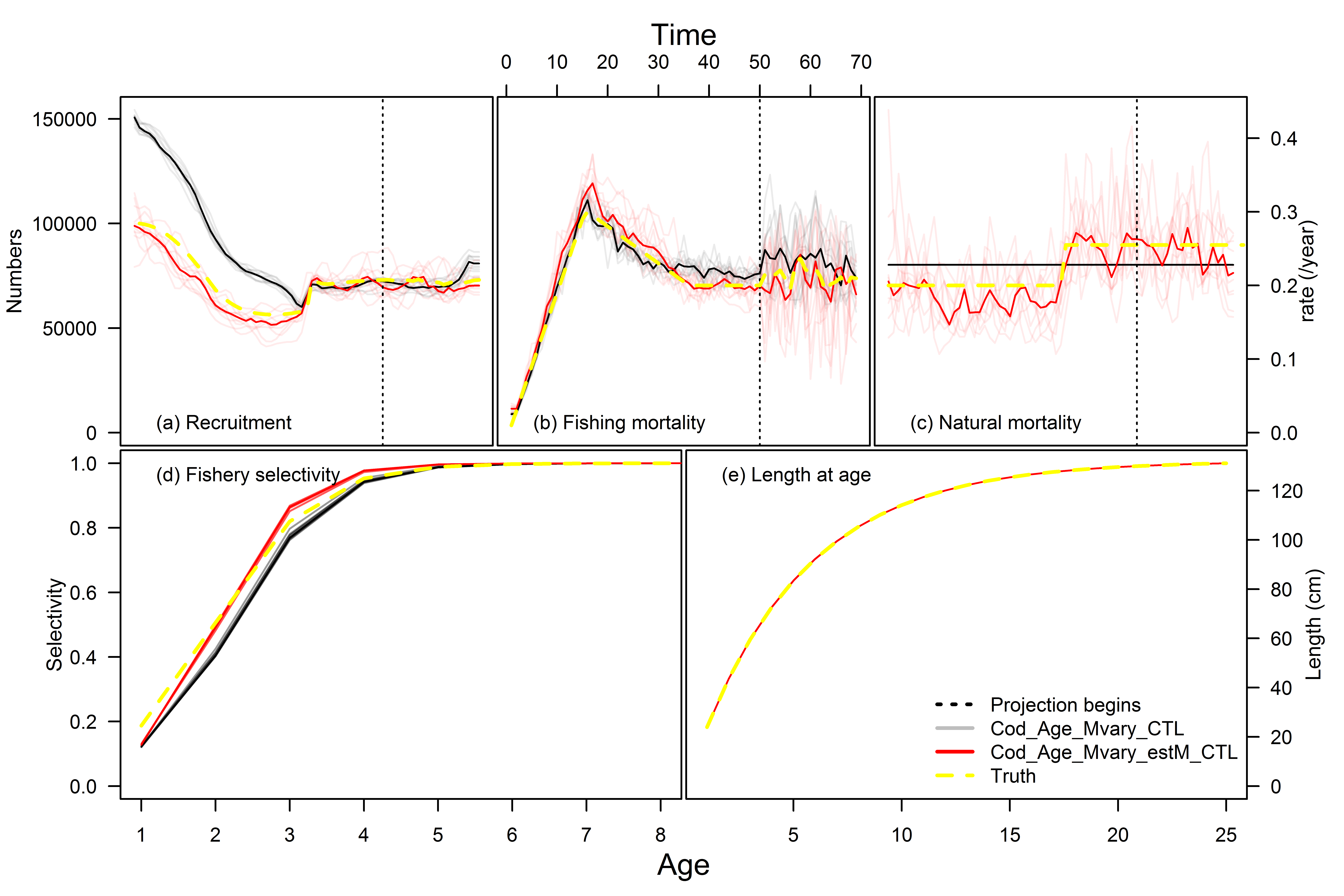


Figure 6. A comparison of the estimated population processes from 2 age-structured assessment methods applied to an operating model in which natural mortality varies over time. “Cod\_Age\_Mvary\_CTL” is an assessment method that fixes natural mortality at a single value; “Cod\_Age\_Mvary\_estM\_CTL” is an assessment method that estimates a time-varying natural mortality.