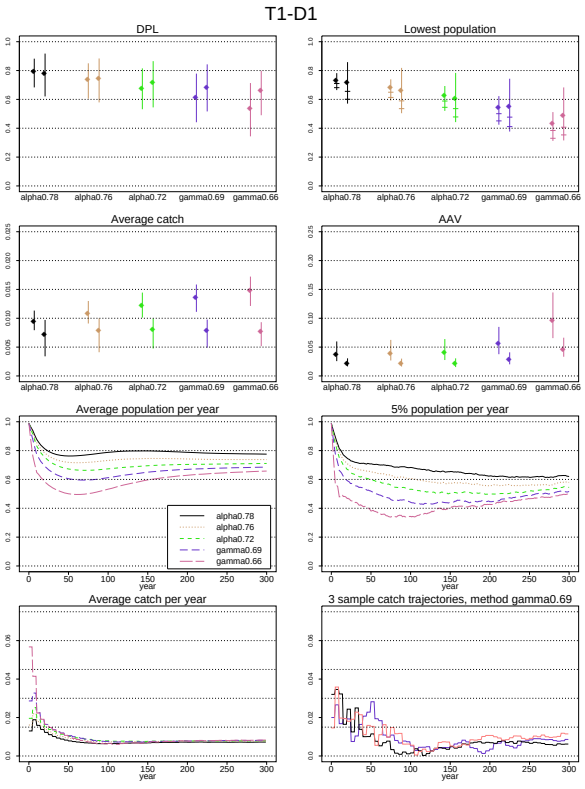
Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Simulation trials for a re-tuned Catch Limit Algorithm



Note no SAMBA/10/06

Authors Magne Aldrin
Ragnar Bang Huseby
Tore Schweder

Date April 26, 2006

The authors

Magne Aldrin is chief research scientist at the Norwegian Computing Center and has a Ph.D. in statistics from the University of Oslo. Ragnar Bang Huseby is senior research scientist at the Norwegian Computing Center and has a M.Sc. in mathematics from the University of Oslo. Tore Schweder is professor in statistics at the University of Oslo and works part-time at the Norwegian Computing Center

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Simulation trials for a re-tuned Catch Limit Algorithm**
Authors	**Magne Aldrin** <magne.aldrin@nr.no> **Ragnar Bang Huseby** <ragnar.bang.huseby@nr.no> **Tore Schweder** <tore.schweder@econ.uio.no>
Date	April 26, 2006
Publication number	SAMBA/10/06

Abstract

The catch limit algorithm (CLA) is a central part of the revised management procedure (RMP) for whaling developed by the International Whaling Commission (IWC), and used by Norway to manage its minke whaling. CLA has previously been tuned to a specified final depletion after 100 years of management, based on simulations from a population model with maximum sustainable yield (MSY) at 1% of the mature component of the stock. According to the "old" tuning procedure, median depletion after 100 years is set to 60%, 66% or 72% of carrying capacity.

In this paper we change the horizon from 100 to 300 years to allow the managed population to come closer to a stable level. We also measure productivity of the stock (MSYR) in terms of the total stock, excluding calves, and we re-tune the catch limit algorithm.

Our "new" tuning procedure assumes $MSYR^{1+}$=1%, and target a given level of depletion after 300 years of management. The traditional tunings lead to target depletions of 0.72, 0.74 and 0.78 according to the "new" procedure. These three versions of the CLA together with two new versions tuned to 0.66 and 0.69, are investigated by simulation with respect to yield and stock conservation properties in a series of scenarios including the base case trials and the more taxing robustness trials that have been considered in previous studies and implementation reviews.

Keywords

Target group

Availability

Project

Project number

Research field

Number of pages 59

Contents

1 Introduction

The revised management procedure RMP and its rule for setting catch quotas, the catch limit algorithm CLA, were developed around 1990. Its catch performance and robustness were tested on simulation trials based on an age- and sex-structured population dynamics model for whales. The productivity of the stock was specified in terms of its mature component through the maximum sustainable yield relative to the number of sexually mature whales $MSYR^{mature}$. All simulation trials were carried out for a period of 100 years of management.

One of the so-called base case trials (T1-D1) acts as a tuning case. Tuning means to set the parameters of the CLA to values making the median final depletion reach a specified level after a certain number of years of management, conventionally 100 years. Due to randomness in the survey part of the model, final depletion has a distribution of which its median shall reach the target. We refer to this as the "old" tuning procedure. Final depletions between 0.60 and 0.72 were typical tuning levels with this "old" tuning procedure.

Figure 1 shows that the traditional tuning does not meet the target median depletion level in the long run. Population dynamics in baleen whales is simply to slow in the tuning trial to have the population process reach anywhere near its stationary state in 100 years, which traditionally has been taken as the management horizon. The population is still on average increasing until about 300 years from a low slightly below 100 years. For all the three levels of "old" tuning, the long term median depletion level is around 74% of carrying capacity (Table 1).

$NYEAR$=100	$NYEAR$=300
0.72	0.76
0.66	0.74
0.60	0.73

Table 1. Median depletion level after 100 and 300 years of management for "old" tuning levels 0.72, 0.66 and 0.60, based on 100 simulationas from the "old" tuning trial with $MSYR^{mature}$=1%.

For these reasons we investigate a "new" method of tuning the RMP (IWC/56/22), we take 300 years as the management horizon, and we also measure productivity in terms of the total population excluding calves, denoted the $1+$ population component. At least for minke whales, it is the $1+$ component which is subject to surveying and abundance estimation. Even if the productivity now is specified in terms of the $1+$ component of the population, the degree of stock depletion is still defined in terms of the sexually mature stock component.

Our "new" tuning procedure is based on simulations from a population model

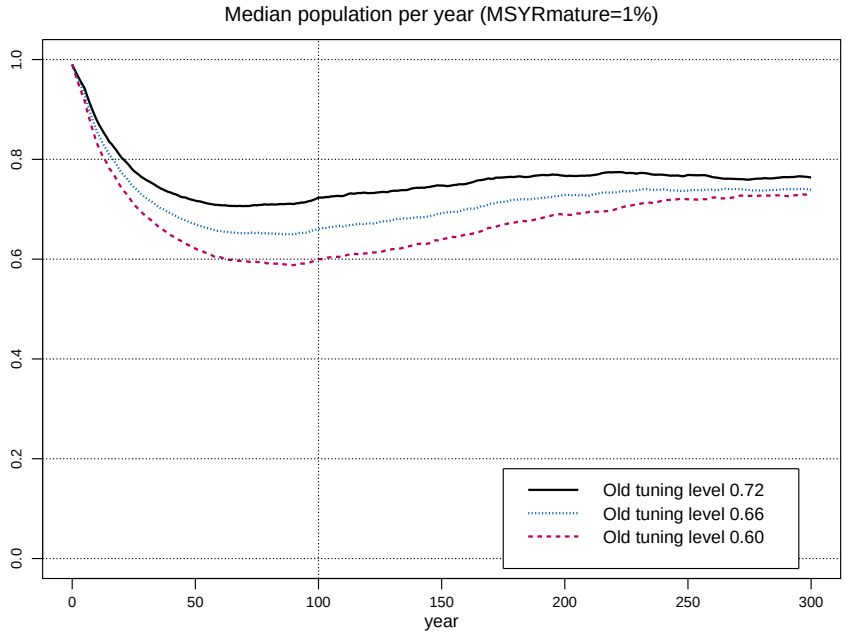


Figure 1. Median yearly population size based on 100 simulations from the "old" tuning trial with $MSYR^{mature}$=1%.

with $MSYR^{1+} = 1\%$. The target, final depletion, is median population level after 300 years of management. The "old" tuning levels 0.60, 0.66 and 0.72 correspond to "new" tuning levels 0.72, 0.74 and 0.78, respectively (Note that these numbers assume $MSYR^{1+} = 1\%$ and therefore differ slightly from the numbers in Table 1 which assume $MSYR^{mature} = 1\%$). We further re-tune the CLA to "new" tuning levels 0.69 and 0.66, but with another tuning parameter than the traditional quantile of the posterior distribution for the internal catch limit.

We use $MSYR^{1+} = 1\%$ as the lower limit of plausibility for productivity in baleen whales. For gray whales, bowhead whales and minke whales in West Greenland, the lower limit for productivity has been taken at least this high. We return briefly to this issue in the Discussion.

The performance of variants of the CLA, tuned to "new" levels between 0.66 and 0.78 is investigated on several simulation trials. The productivity in these population models are measured in terms of $MSYR^{1+}$ rather than $MSYR^{mature}$. Yield- and conservation statistics are calculated both over the first 100 and over the next 200 years of management. The aim of the robustness trials is to investigate the various versions of the CLA with respect to productivity for the whalers, but more importantly, conservation properties in various marginally plausible scenarios that might pose a risk of severe depletion.

In Section 2 we first review the CLA, then present the simulation trials, and finally describe the tuning process. The results are given in Section 3, and they

are briefly discussed in Section 4.

2 Methods

2.1 The Catch Limit Algorithm (CLA)

The input data to the catch limit algorithm (IWC, 1999, p. 251-258) consists of a time series of historic annual catches and a time series of absolute abundance estimates along with their standard errors and correlations on the logarithmic scale.

The internal population model of the catch limit algorithm is defined by the following dynamics

$$\begin{aligned} P_0 &= \frac{P_T}{D_T}, \\ P_{t+1} &= P_t - C_t + 1.4184\,\mu\,P_t(1-(\frac{P_t}{P_0})^2) \quad (0 \le t < T), \end{aligned} \tag{1}$$

where

- 0 is the first year of recorded catch, and T is the current year of management (i.e. the first year of an assessment cycle). P_0 is regarded as pristine population size, and P_t is the population size in numbers at the beginning of year t,
- C_t is the catch in numbers in year t,
- $D_T = P_T/P_0$ is the ratio of the population size at the beginning of year T to the population size at the beginning of year zero, measuring stock depletion,
- μ is a parameter describing the productivity,
- the historic catch series used in assessments covers years 0 to $T-1$.

The abundance estimates are assumed to be log-normally distributed with a given (estimated) information matrix for the on the log scale. The likelihood based on the abundance data is

$$\text{Likelihood}(\mu, D_T, b) \propto \exp\left(-0.5(\mathbf{a}-\mathbf{p}-\beta\mathbf{1})'H(\mathbf{a}-\mathbf{p}-\beta\mathbf{1})\right) \tag{2}$$

where the symbol $\propto$ means proportional to, and where

- $\mathbf{a}$ is the vector of logarithms of the estimates of population size by year,
- $\mathbf{p}$ is the vector of logarithms of the modelled annual population sizes for the years with population estimates, $p_t = \ln(P_t)$,

- β is the logarithm of the bias parameter, thus $b = \exp(\beta)$;
- H is the information matrix of the a vector. H is assumed nonsingular, and $V = H^{-1}$ is the covariance matrix of the vector a.

The parameters μ, D_T, and b are assigned independent uniform prior distributions making their joint prior distribution uniform over the region

$$[\mu_{\min}, \mu_{\max}] \times [D_{T,\min}, D_{T,\max}] \times [b_{\min}, b_{\max}], \tag{3}$$

where $\mu_{\min}$, $\mu_{\max}$, $D_{T,\min}$, $D_{T,\max}$, $b_{\min}$, and $b_{\max}$ are chosen constants. We will use $\mu_{\min} = 0.0$, $\mu_{\max} = 0.05$, $D_{T,\min} = 0.0$, $D_{T,\max} = 1.0$, $b_{\min} = 0.0$, and $b_{\max} = 1.6667$, which are the values used in the current implementation of the CLA.

A distinctive feature of the CLA is that abundance data are strongly down-weighted to obtain desired robustness properties. In the internal model, all variances and covariances of logarithmic abundance estimates are actually multiplied by 16. The historic catch data are furthermore assumed to be accurate, without any measurement errors. The posterior density function of the parameters μ, D_T, and b is therefore

$$\text{Posterior}(\mu, D_T, b) \propto \text{Prior}(\mu, D_T, b) \cdot \text{Likelihood}(\mu, D_T, b)^s, \quad s = 1/16 \tag{4}$$

The presence of a deflation parameter $0 < s < 1$ down-weights the survey information relative to a strict Bayesian approach.

The internal catch limit is the following function of μ, D_T, and P_T:

$$L_T = \begin{cases} 0 & \text{if } D_T \leq IPL \\ \gamma\mu(D_T - IPL)P_T & \text{if } D_T > IPL \end{cases} \tag{5}$$

where the internal protection level IPL is a control parameter. In this work IPL is fixed to 0.54. Traditionally, $\gamma = 3$. We will use this parameter as a tuning parameter in 2.3.

The internal catch limit can be regarded as the catch limit in the hypothetical case of perfect knowledge of population parameters and size. However, in the Bayesian formalism, L_T is regarded as a random variable, with marginal posterior distribution obtained from the joint posterior distribution of (μ, D_T, b). The actual catch limit z is defined as a certain percentile of its distribution,

$$P(L_T < z|data) \leq \alpha \leq P(L_T \leq z|data) \tag{6}$$

for a given value of the tuning parameter α.

We use the implementation of the algorithm of the Norwegian Computing Center (Huseby and Aldrin 2006). This implementation is available from the IWC secretariat. It is a FORTRAN subroutine called CATCHLIMIT, and the present

version is from January 2006. The parameters α and γ above are input parameters to the subroutine, and are called IN_PPROB and IN_PSLOPE, respectively. The accuracy of the algorithm depends on parameters set by the user. Details regarding the accuracy parameters used in this work are given in Appendix B.

2.2 Simulation trials

The CLA is tuned to five different target levels, as described in Section 2.3. The performance of each of these five specifications of the CLA is tested in various scenarios. The purpose is to investigate their performance with respect to yield- and conservation properties under various assumptions regarding productivity and other aspects of the population dynamics of the stock to be managed, and assumptions regarding the statistical properties of the abundance estimates used as input to the CLA. The scenarios, or trials, vary from optimistic to pessimistic assumptions on recruitment, initial population size, bias in sighting surveys etcetera. These simulation trials are performed using the FORTRAN program MANTST which is available from the IWC secretariat. Our version of MANTST is based on version 11 (received from the IWC secretariat in January 2005), but modified by Andre Punt to allow for projections more than 100 year.

The trials are based on an age- and sex-structured population dynamics model for whales with density-dependent fertility. For given model parameters, and given catches, the model generates deterministic population trajectories. Based on the series of previous catches and the stochastically generated series of abundance estimates with associated uncertainty, the CLA calculates the catch limit in the current year. The population vector at the beginning of next year is then calculated by the model, and a new abundance estimate is randomly generated if that year is assigned for abundance estimation. The CLA is then applied, and the whole process is moved forwards. The abundance estimates are generated according to the distributional assumptions in the trial, and according to current status of the simulated stock. Due to the stochasticity in the abundance estimates, the population trajectory is stochastic. Thus, replicate simulations are performed in each trial.

The age-specific fertility rates in the population dynamics model depend on the number of sexually mature whales in the stock. At carrying capacity K measured in number of mature individuals, and with a stable age distribution, the schedule of density independent natural mortalities balances the fertility and there is no net recruitment.

The CLA calculates the catch limit in number of whales. Removals are assumed equal to catch limit, and the yield is recorded as a fraction of K. The degree of stock depletion is defined in terms of the sexually mature stock component, and is recorded as a fraction of K. As mentioned above, and further touched

upon in Section 4, we measure productivity by $MSYR^{1+}$, and use 1% and 4% as reference values.

All the simulation trials have a period of 30 years with historical catches before year 0, which is the first year where catch quotas are set by the CLA. The quotas are based on simulated abundance estimates, thought of as being based on sighting surveys. Sighting surveys with corresponding abundance estimates are then in the simulations usually repeated every five years.

Two parameters are varied systematically from trial to trial:

- The initial depletion, e.g. the population size in year 0 relative to K. This parameter is usually either $0.99K$ (trials coded by D), $0.60K$ (coded by S) or $0.30K$ (coded by R).

- The maximum sustainable yield rate ($MSYR^{1+}$), which usually is either 1% (denoted D1, S1 or R1) or 4% (denoted D4 etcetera).

When these two parameters are fixed, the carrying capacity is given implicitly. Bias and variability in abundance estimates will also be specified in the various trials, along with other aspects of the population dynamics, the environment or the observational scheme.

A comprehensive list of earlier trials is found in Allison (2002). We have rerun a selected sample of these, with the two important differences that productivity is now measured by $MSYR^{1+}$ rather than by $MSYR^{mature}$, but using the same levels (1% and 4%), and we use horizon 300 rather than 100 years. Figure 2 shows average (over 100 replicate simulations) depletion by year of management for three different trials. The same version of the CLA was used in these three trials. The population size has not stabilized after 100 years of management, but is close to a stable level after 300 years. Compared to Figure 1, where the CLA tuned by the "old" method to the three traditional targets are displayed, the convergence is somewhat slower, at least for trial T3-R1. Although for this trial and for a few additional trials where the population size still changes appreciably after 300 years, we found this time horizon to be sufficient to see “almost the whole story” within an acceptable computer time. Note that computing demand grows faster than linear with number of years of management since data gets more and more complex.

Table 2 gives a list of the trials we have run. 400 replicate simulations are used in the T1 trials and 100 replicate simulation in the others. We use the same seeds for the random number generator as was used in the original trials. The trials are described in more detail in IWC (1992, p. 317-318). Some trials need some additional comments:

- The trial T1-D1 is of special interest. It is used to tune the CLA, i.e. to set its parameters such that the median final population size after a given number

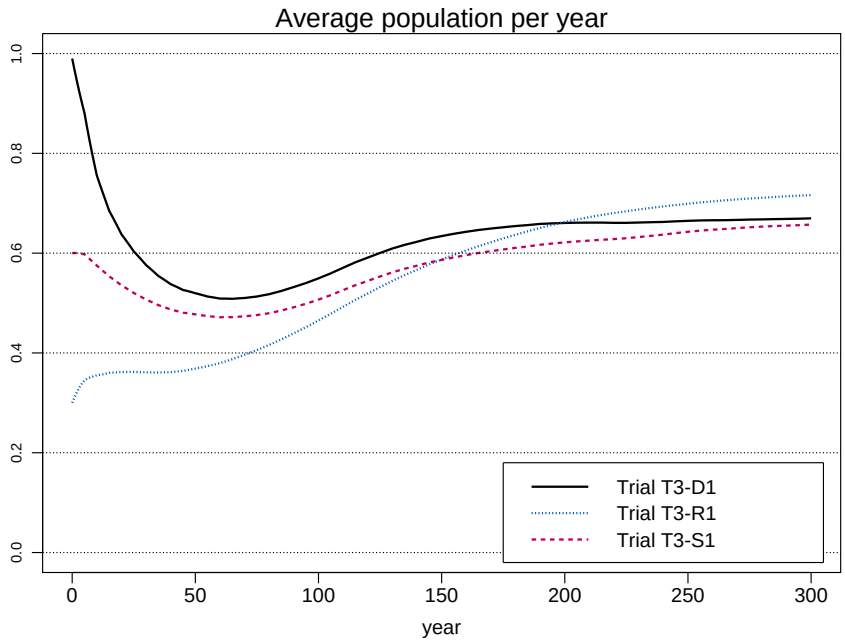


Figure 2. Average yearly population size based on 100 simulations from trials T3-D1, T3-R1 and T3-S1 with $\alpha = 0.5222$ and $\gamma = 3$.

of years matches a specified level. This is treated in more detail in Section 2.3.

- The trials T1-D0.6, T1-R0.6 and T1-S0.6 are additional to the original set. Here the recruitment parameter $MSYR^{1+}$ is lower than 1% to investigate what may happen if $MSYR^{1+}$ is less than what is regarded as the minimum plausible level. In these trials $MSYR^{mature}$ is set to 1%, which corresponds to $MSYR^{1+}\approx$0.6% according to Punt and Allison (2004). These trials are identical to the original T1-D1, T1-R1 and T1-S1 trials.

- The T9 trials have episodic events where the population size is halved. The events occur at rate 0.02 (every fifty years on average), but the times of events are stochastic and independent within and between simulations.

- In the T12A and T12B trials, the carrying capacity K first changes linearly from year 0 to year 100 as in the original trials. After 100 years K is kept constant, since further change at the same rate seems biological implausible.

For completeness, Table 3 lists those original trials (according to Allison 2002) as we regard to be of minor interests and have chosen not to run.

Trial name	Initial size	$MSYR^{1+}$ (%)	Description
T1-D1	0.99	1	Base case
T1-D4	0.99	4	Base case
T1-R1	0.30	1	Base case
T1-R4	0.30	4	Base case
T1-S1	0.60	1	Base case
T1-S4	0.60	4	Base case
T1-D0.6	0.99	0.6	$MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%
T1-S0.6	0.60	0.6	$MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%
T1-R0.6	0.30	0.6	$MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%
T2-D1	0.99	1	50% negative bias in abundance estimates
T2-R1	0.30	1	50% negative bias in abundance estimates
T3-D1	0.99	1	50% positive bias in abundance estimates
T3-R1	0.30	1	50% positive bias in abundance estimates
T3-S1	0.60	1	50% positive bias in abundance estimates
T4-X1	0.05	1	Initial depletion = $0.05K$
T6-R1	0.30	1	Reported historic catch = 50% of true catch
T6-R4	0.30	4	Reported historic catch = 50% of true catch
T9-D1	0.99	1	Episodic events; (Rate =0.02)
T9-D4	0.99	4	Episodic events; (Rate =0.02)
T9-R1	0.30	1	Episodic events; (Rate =0.02)
T9-R4	0.30	4	Episodic events; (Rate =0.02)
T12A-D1-2	0.99	1	Linear increase in K from K_0 to $2 * K_0$ after 100 years, then constant.
T12A-D4-2	0.99	4	As above
T12A-R1-2	0.30	1	As above
T12A-R4-2	0.30	4	As above
T12B-D1-05	0.99	1	Linear decrease in K from K_0 to $0.5 * K_0$ after 100 years, then constant
T12B-D4-05	0.99	4	As above
T12B-R1-05	0.30	1	As above
T12B-R4-05	0.30	4	As above
T15-D1-1	0.99	1	Surveys every 10 years
T15-D4-1	0.99	4	Surveys every 10 years
T15-R1-1	0.30	1	Surveys every 10 years
T15-R4-1	0.30	4	Surveys every 10 years

Table 2. Trials performed.

Trial name	Initial size	$MSYR^{1+}$ (%)	Description
T5-R1	0.30	1	25yrs protection before management
T5-R4	0.30	4	25yrs protection before management
T7-D1	0.99	1	Age at maturity: 10.0
T7-D4	0.99	4	Age at maturity: 10.0
T7-R1	0.30	1	Age at maturity: 10.0
T7-R4	0.30	4	Age at maturity: 10.0
T10-D1-4	0.99	1	Tent model; MSYL= $.40K$
T10-D4-4	0.99	4	Tent model; MSYL= $.40K$
T10-R1-4	0.30	1	Tent model; MSYL= $.40K$
T10-R4-4	0.30	4	Tent model; MSYL= $.40K$
T11-D1-8	0.99	1	Tent model; MSYL= $.80K$
T11-D4-8	0.99	4	Tent model; MSYL= $.80K$
T11-R1-8	0.30	1	Tent model; MSYL= $.80K$
T11-R4-8	0.30	4	Tent model; MSYL= $.80K$
T13-D1-33	0.99	1	MSYR steps up or down every 33 yrs
T13-D4-33	0.99	4	MSYR steps up or down every 33 yrs
T13-R1-33	0.30	1	MSYR steps up or down every 33 yrs
T13-R4-33	0.30	4	MSYR steps up or down every 33 yrs
T14-D1-	0.99	1	No surveys after year -1
T14-R1-	0.30	1	No surveys after year -1
T15-D1-1	0.99	1	Surveys every 10 years
T15-D4-1	0.99	4	Surveys every 10 years
T15-R1-1	0.30	1	Surveys every 10 years
T15-R4-1	0.30	4	Surveys every 10 years

Table 3. Trials not performed.

2.3 Tuning CLA

The CLA is tuned by finding values of its parameters that makes the median (over replicate runs) final depletion meet a target value in the T1-D1 trial.

Traditionally, the management horizon has been $NYEAR$=100, and $MSYR^{mature}$=1%. Furthermore, α was varied and used as tuning parameter to meet the required tuning level, whereas γ was fixed to the value 3. Table 4 shows the values of α corresponding to "old" tuning levels 0.72, 0.66 and 0.60 (from Huseby and Aldrin, 2000)

Instead we use $NYEAR$=300 and $MSYR^{1+}$=1%, as argued above and further discussed in Section 4. Table 4 also shows the corresponding "new" tuning levels for α varying from 0.4015 to 0.8, whereas γ is still fixed at 3. For this value of γ, α is clearly not usable as a tuning parameter when the target is 0.7 or lower. Even high values of α will actually in the long run produce catch limits that are too low to make median depletion after 300 years meet a target lower than 0.7 when $MSYR^{1+} = 1\%$.

In addition, α-values higher than 0.5 are conceptually problematic. Increased uncertainty in the posterior distribution of the internal catch limit L_T will in fact push percentiles above the median out towards higher values. Catch limits might therefore increase with increasing uncertainties in abundance estimates if $\alpha > 0.5$.

For these reasons Aldrin, Huseby and Schweder (2006) investigated various other methods of tuning the CLA,. They found that γ is a suitable tuning parameter together with fixing $\alpha = 0.5$. Table 5 shows values of γ found to achieve target levels of 0.69 and 0.66 for median depletion after 300 years in trial T1-D1. More details on the tuning are found in appendix A.

α	γ	"Old" tuning level with $NYEAR$=100 and $MSYR^{mature}$=1%	"New" tuning level with $NYEAR$=300 and $MSYR^{1+}$=1%
0.4015	3	0.72	0.78
0.4629	3	0.66	0.76
0.5222	3	0.60	0.72
0.6000	3		0.70
0.8000	3		0.70

Table 4. Values of α for fixed γ with corresponding "old" and "new" tuning levels. "Old" tuning level is based on >50000 simulations., "New" tuning level is based on 400 simulations with the standard common set of seeds.

α	γ	"New" tuning level with $NYEAR$=300 and $MSYR^{1+}$=1%
0.5	4.7157	0.69
0.5	9.3443	0.66

Table 5. Values of γ for fixed α with corresponding "new" tuning levels. "New" tuning level is based on >50000 simulations, but calculations on 400 simulations with the standard common sets of seeds give the same numbers within the accuracy of two decimals.

3 Results of simulation trials

We investigate the performance properties of five variants of the CLA. There are two variants with γ as tuning parameter with "new" tuning levels 0.69 and 0.66, from now on called gamma0.69 and gamma0.66. There are also three variants with α as tuning parameter with "new" tuning levels 0.78, 0.76 and 0.72, here called alpha0.78, alpha0.76 and alpha0.72 respectively. The alpha0.72 variant is currently used to set catch limits for northeastern Atlantic minke whales.

These five variants differ primarily in their tuning levels. The difference in tuning method, i.e. by γ or α is of secondary importance. Table 6 gives a summary of the five variants.

name	α	γ	"New" tuning level with $NYEAR$=300 and $MSYR^{1+}$=1%
alpha0.78	0.4015	3	0.78
alpha0.76	0.4629	3	0.76
alpha0.72	0.5222	3	0.72
gamma0.69	0.5	4.7157	0.69
gamma0.66	0.5	9.3443	0.66

Table 6. Combinations of α and γ for five tunings of the CLA. "New" tuning level is based on >50000 simulations, but calculations on 400 simulations with the standard common sets of seeds give the same numbers within the accuracy of two decimals.

All population and catch quantities reported are scaled by the carrying capacity K in year 0. For each trial and for each variant of CLA, four main quantities are calculated based on the first 100 and the next 200 years of each single simulation:

· Final depletion (DPL) at the end of the period in question.

- Lowest population over the period.
- Average catch.
- Average annual catch variation (AAV), defined as $ave|C_t - C_{t-1}|/ave(C_t)$, where C_t is the catch in year t and ave means average over the period in question.

For each of these quantities, the following summary statistics are calculated based on the 100 or 400 simulations:

- The median.
- The 5% and 95% values. The 5%-values are calculated as the 5th lowest value if they are based on 100 simulations or the 20th lowest if they are based on 400 simulations, according to the description in IWC (1992, p. 317-318). The 95% values are similarly calculated as the 96th or 381st value.
- The 10% and 25% values, for the lowest population only.

The results from each trial separately are shown in the upper four panels of Figures 3 to 35 (placed after the references in this paper), and the numerical values are found at
`http://www.nr.no/~aldrin/whales/allresults.txt` . In each panel of the figures, the results are given pairwise for each variant of CLA, with the results for the first 100 years to the left and for the next 200 years to the right. The various summary statistics are symbolised by:

- The median is shown as a diamond.
- The 90% interval between the 5% and 95% values is shown as a vertical line.
- The 10% and 25% values for the lowest population are shown as two short horizontal lines.

The next three panels in each figure show time series of year-specific summary statistics for each variant of CLA. The three statistics are average population, 5% population level, and average catch. The final panel in each figure shows catch trajectories for three individual simulations with catch quotas set by the gamma0.69 variant, from the first three of the standard common sets of seeds.

Note that the scales on the y-axes for the population panels are usually between 0 and 1, except for the T12A trials, where the upper limits are 2. In the figures for the T12A and T12B trials, the year-specific K is marked by a separate line. The scales for AAV are the same in all figures. The scales on the y-axes for the catch panels may vary within and between figures.

4 Discussion

The schedules of fertility and natural mortality determines the productivity of the stock which is summarized in the maximum sustainable yield rate $MSYR$. In the RMP context, $MSYR$ has traditionally been defined in terms of the sexually mature stock (IWC, 1992, p. 317-318). In the context of management procedures for aboriginal subsistence whaling, $MSYR$ has however consistently been measured as the ratio between the maximum sustainable yield and the total number of whales excluding calves. We have also summarized stock productivity in this way, denoted $MSYR^{1+}$, and we have used 1% and 4% as reference values for $MSYR^{1+}$ as these have been considered reasonable in the context of aboriginal subsistence whaling (AWMP).

For minke whales in off West Greenland, and for bowhead whales, the range of plausible values for $MSYR^{1+}$ has been from 1% to 4%, while the range has been 1.5% to 5.5% for gray whales (IWC, 1998 p.209-10; 1999 p.130; 2000 p.130; 2002 p.20; 2003 p.27; 2004 p.188; 2005 p.16).

Our context is that of investigating the conservation properties of various proposed catch limit algorithms in a number of plausible scenarios, which also has been the context for simulation testing of proposed strike limit algorithms aboriginal subsistence whaling. For consistency with respect to plausible productivity in baleen whales, we have therefore settled for $MSYR^{1+} = 1\%$ as the lower limit (except in trials T1-D0.6, T1-R0.6 and T1-S0.6).

In addition to use $MSYR^{1+} = 1\%$ as the lower limit of plausible productivity, we have extended the horizon to 300 years of management. We have actually demonstrated that the population dynamics for baleen whales, as modelled in the population dynamics model , is too slow to allow a population subject to management by the proposed catch limit algorithms to come anywhere near stability in 100 years. It is the continued performance of the algorithm that is at issue, both with respect to yield and conservation, and we find that 100 years of management is not a long run. For most scenarios the population is close to stability after 300 years of management.

The five variants of the CLA are broadly ordered by their target depletion, also with respect to average catch and lowest population level. As expected, catches tend to go up, and lowest population tends to go down, as target depletion level goes down.

The gamma0.66 variants might be regarded as giving too high yield on the cost of depressing the population to low with appreciable probability in several trials: T1-S0.6, T3-D1, T3-R1, T3-S1, T6-R4, T12A-R4.

The variants gamma0.69 and alpha0.72 show quite similar performance. The yield is somewhat better in the former, while the latter in some trials depress

the population slightly less. In the trial T12A-R4 the two variants show quite different performance. Note that depletion is measured relative to initial K. When gamma0.69 drives the 5% relative population level down towards 0.7 after 300 years, the 5% level will tend to 0.35 relative to current carrying capacity. If this is not regarded dangerously low, gamma0.69 behaves better than alpha0.72 in this trial.

The performance of the alpha variants of the CLA were also investigated in the original simulation trials in the nineties, when the population model were based on $MSYR^{mature}$. In trials T1-D0.6, T1-R0.6 and T1-S0.6 productivity is set at $MSYR^{mature} = 1\%$ which corresponds to $MSYR^{1+}$ somewhat above 0.6, and thus outside the plausible range. In these old base case trials, all the five variants considered perform satisfactorily.

Acknowledgements

The authors are grateful to Andre Punt, Doug Butterworth and Lars Walløe for good help when designing this study and when presenting its results. We also thank Andre Punt for modifying the MANTST program, version 11, to allow for projections more than 100 years.

This work has been funded by the Norwegian Ministry of Fisheries and Coastal Affairs and the Norwegian Research Council by the project 169046/S40.

References

Allison, C. (2002). Adjunct 2 - Report on tuning the Norwegian CLA. J. Cetacean Res. Manage., 4 (Suppl), 108-113.

Allison, C. and Punt, A.E. (2004). Appendix 2 - MSYL, MSYR and Density-Dependence Trials. J. Cetacean Res. Manage., 6 (Suppl), 92-95.

Aldrin, M. and Huseby, R.B. and Schweder, T. (2006). Variants of the Catch Limit Algorithm. NR-note SAMBA/11/06, Norwegian Computing Center.

Huseby, R.B. and Aldrin, M (2000). Tuning of the catch limit algorithm. NR-note SAMBA/26/00, Norwegian Computing Center.

Huseby, R.B. and Aldrin, M (2006). Updated documentation of a Fortran 77 subroutine implementing the catch limit algorithm - Version January 2006. NR-note SAMBA/06/06, Norwegian Computing Center.

IWC (1992). Rep. Int. Whal. Commn 42.

IWC (1998). Rep. Int. Whal. Commn 48.

IWC (1999 - 2005). J. Cetacean Res. Manage. 1-7 (Suppl.)

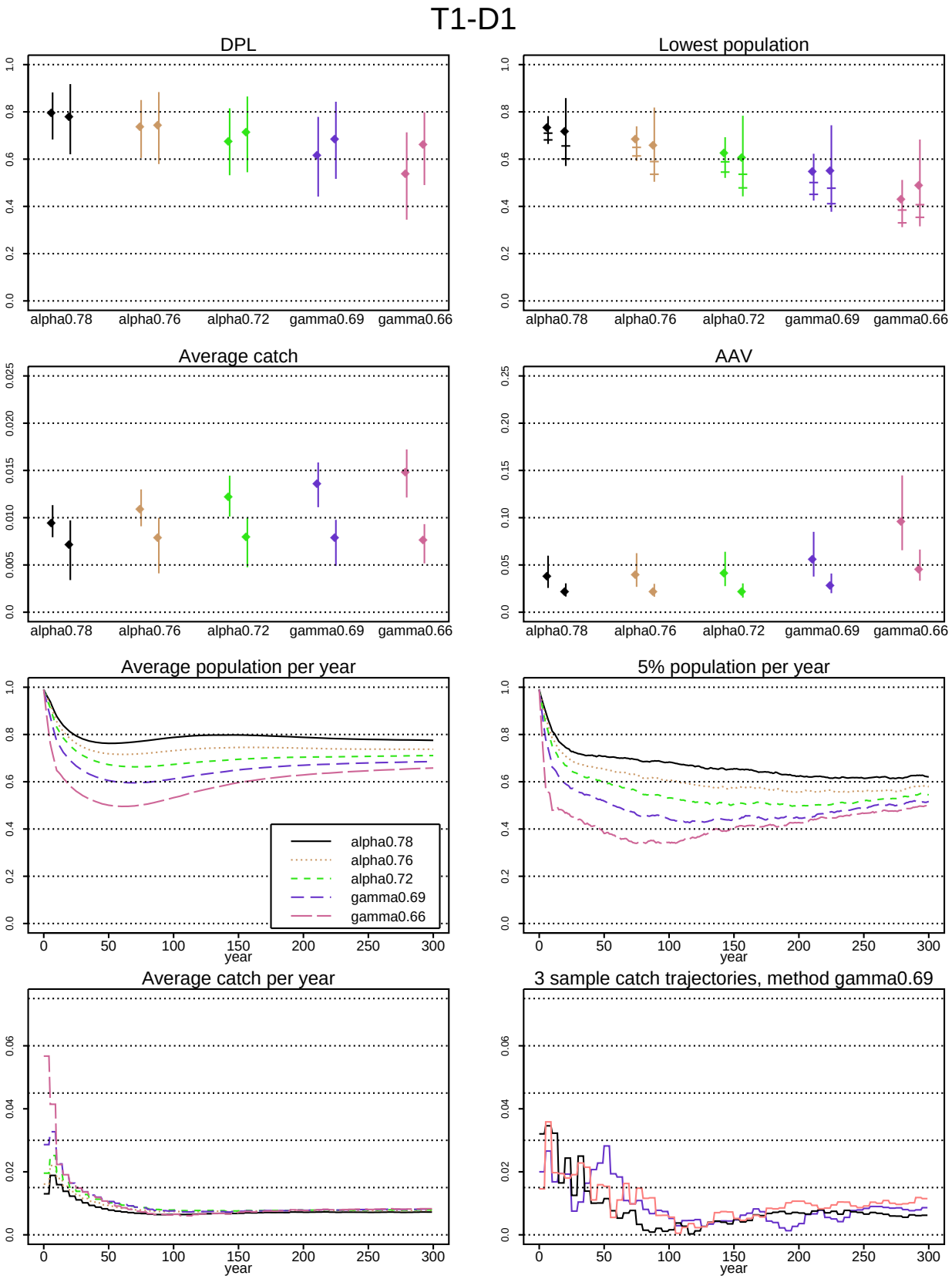


Figure 3. Trial T1-D1. Base case. The pairs of results by CLA variant in the upper four panels represent the first 100 years (left) and the last 200 years (right).

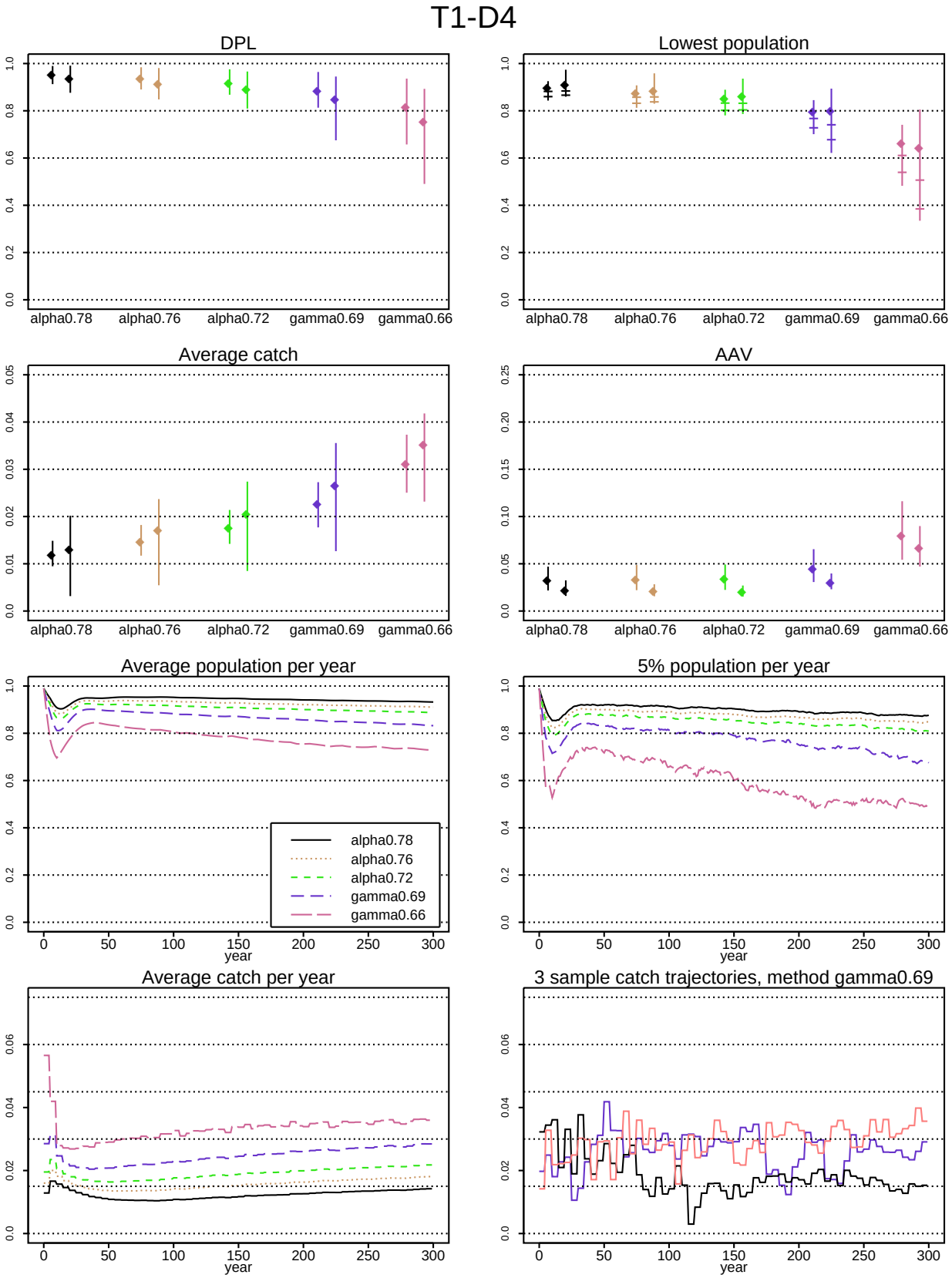


Figure 4. Trial T1-D4. Base case.

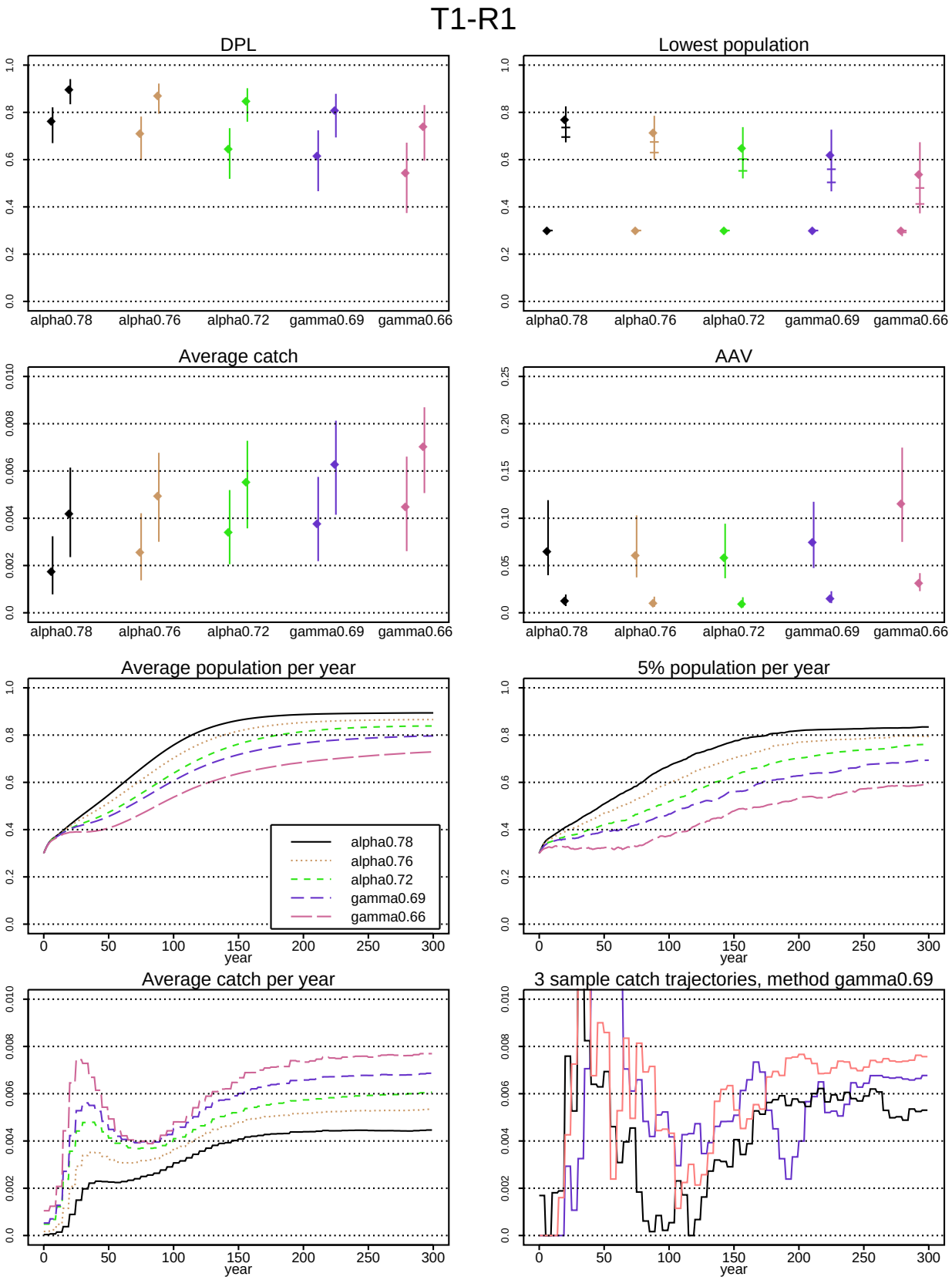


Figure 5. Trial T1-R1. Base case.

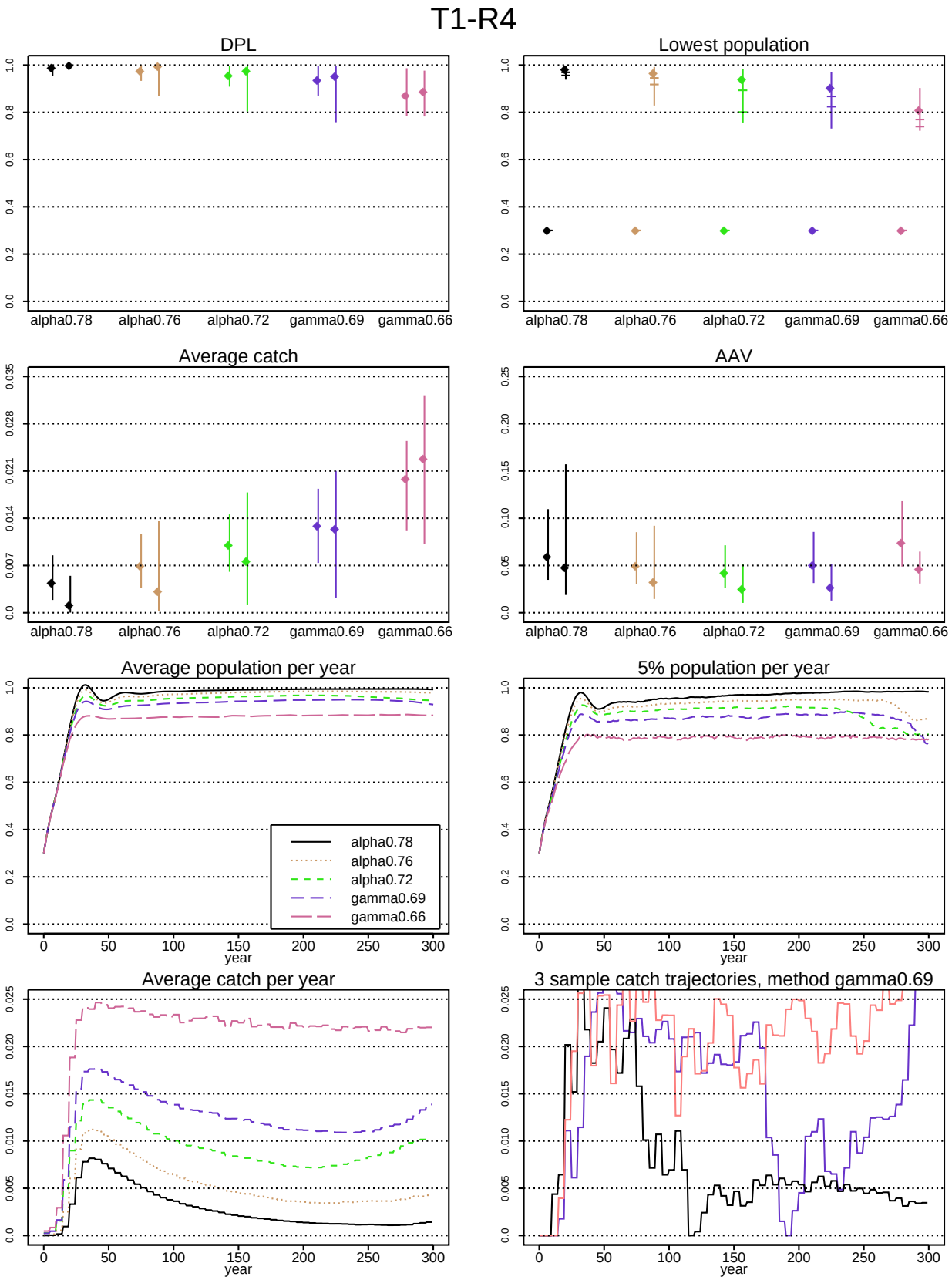


Figure 6. Trial T1-R4. Base case.

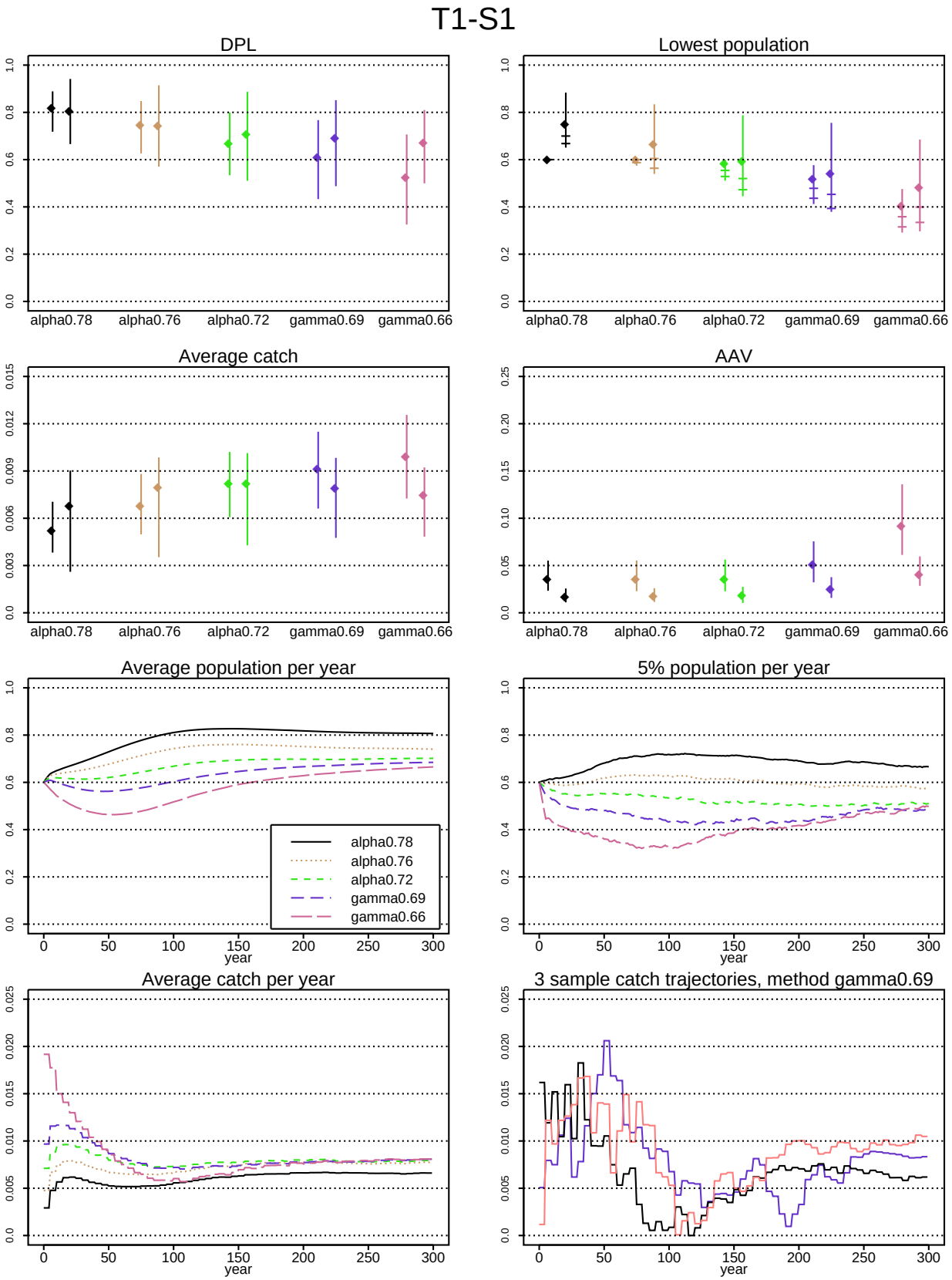


Figure 7. Trial T1-S1. Base case.

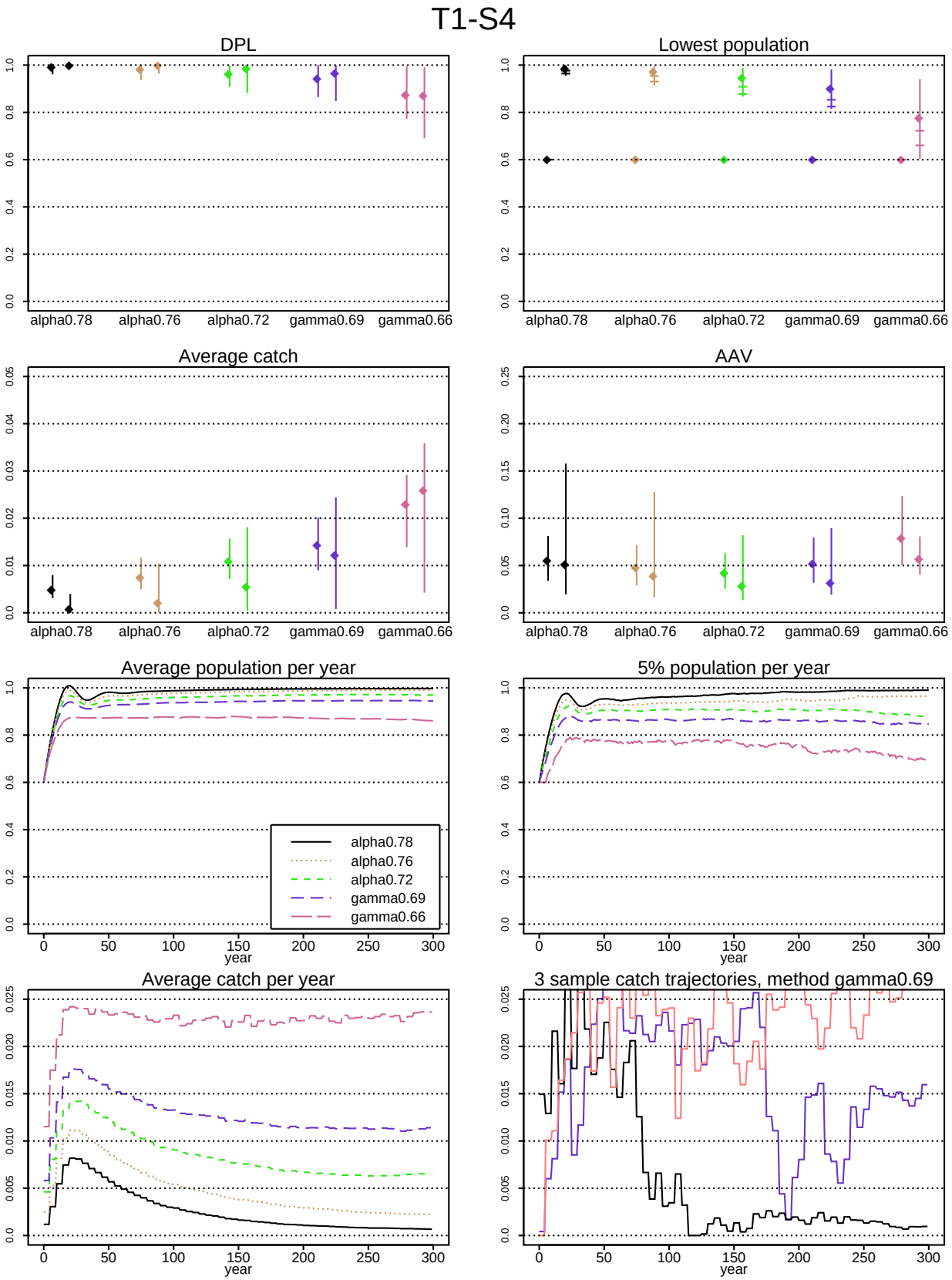


Figure 8. Trial T1-S4. Base case.

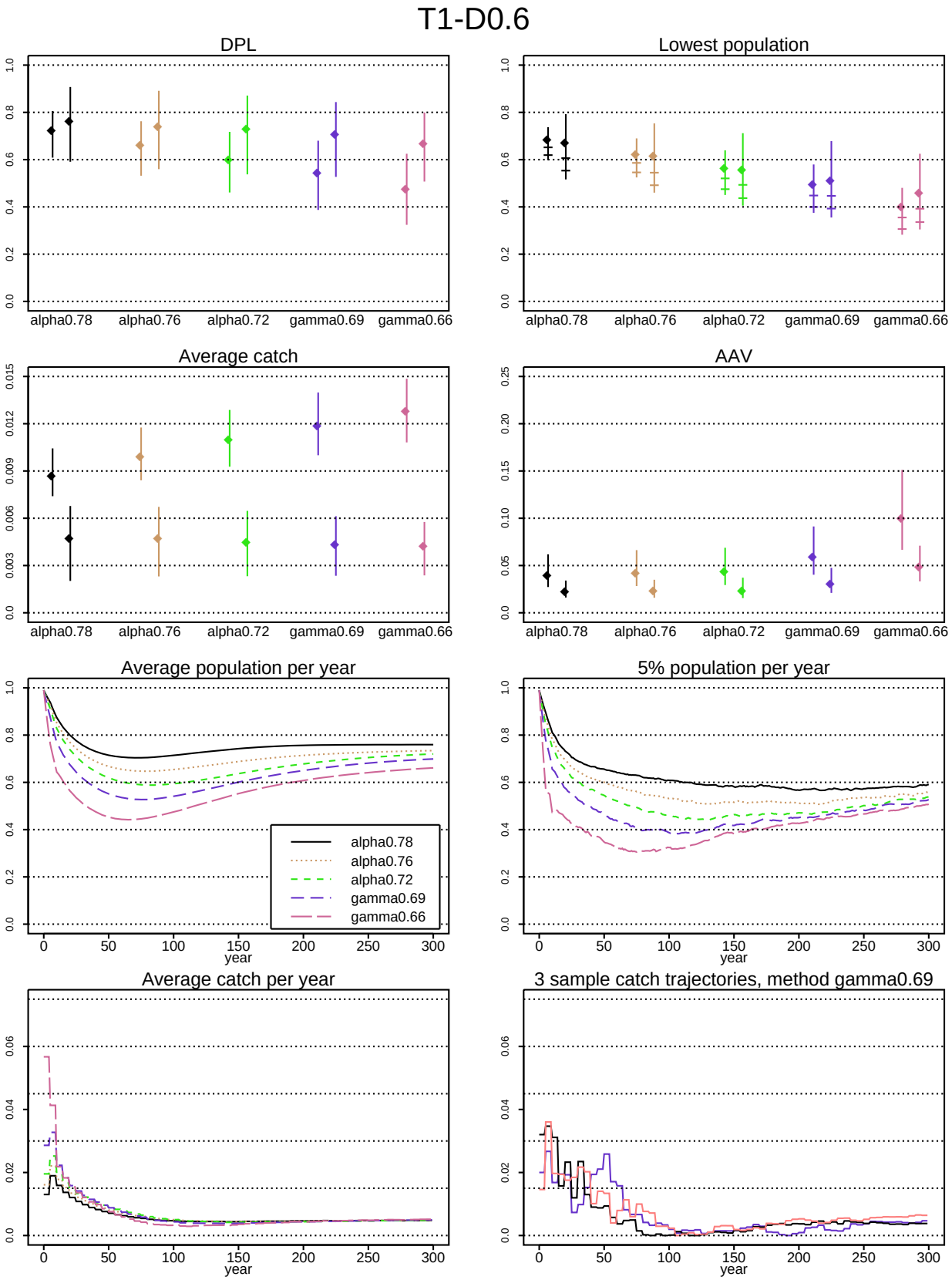


Figure 9. Trial T1-D0.6. $MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%.

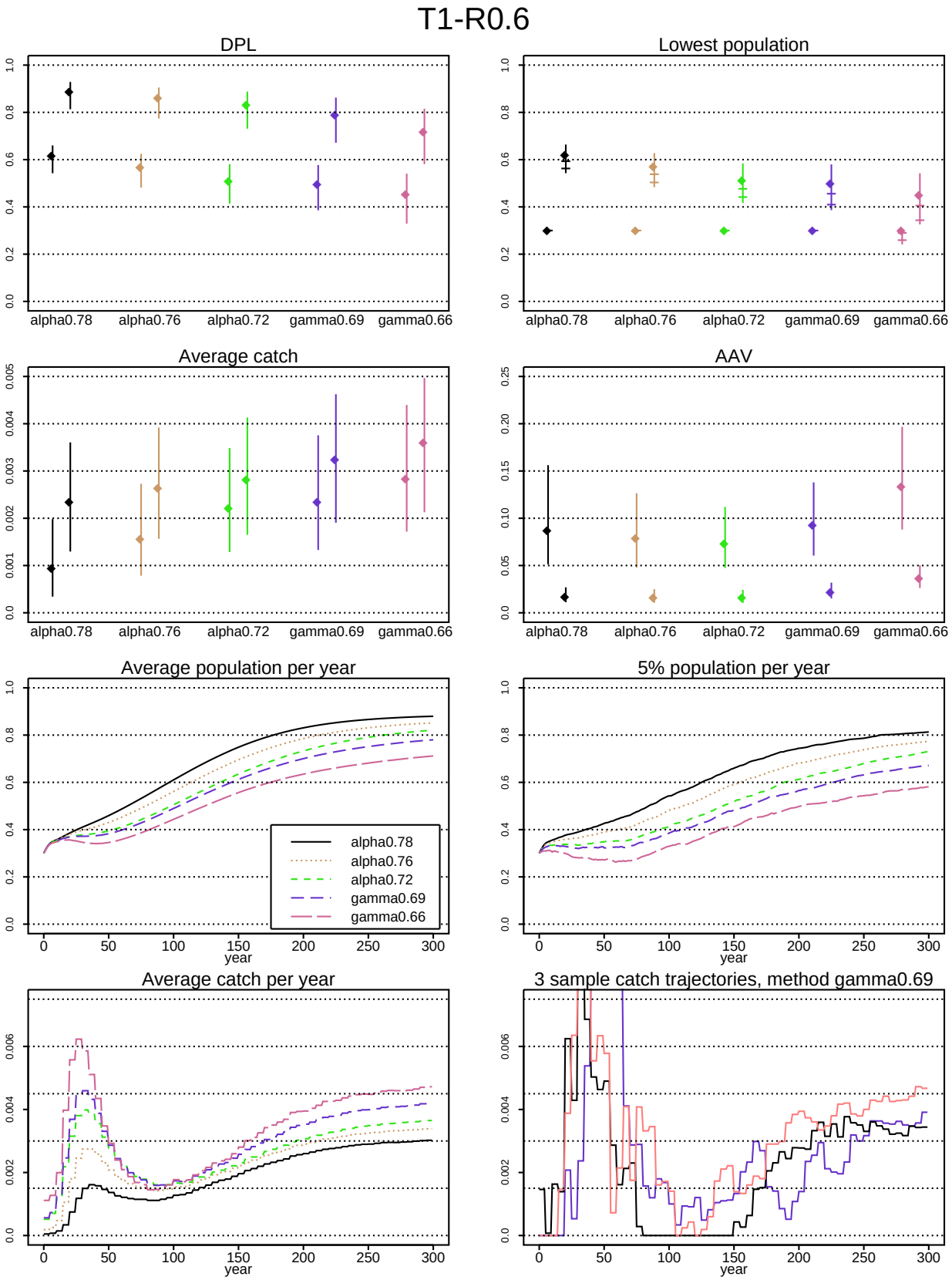


Figure 10. Trial T1-R0.6. $MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%.

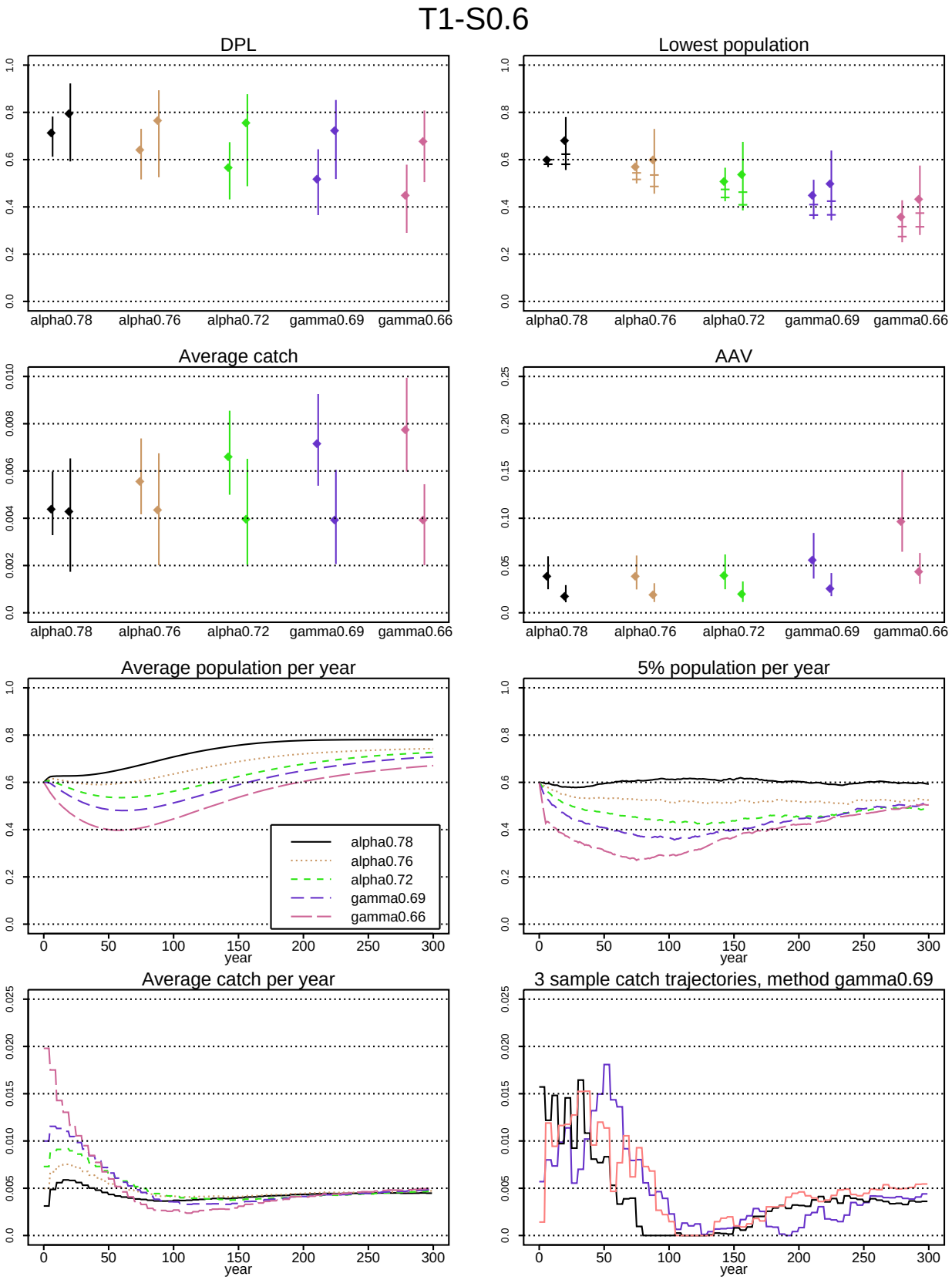


Figure 11. Trial T1-S0.6. $MSYR^{mature}$=1% => $MSYR^{1+}$≈0.6%.

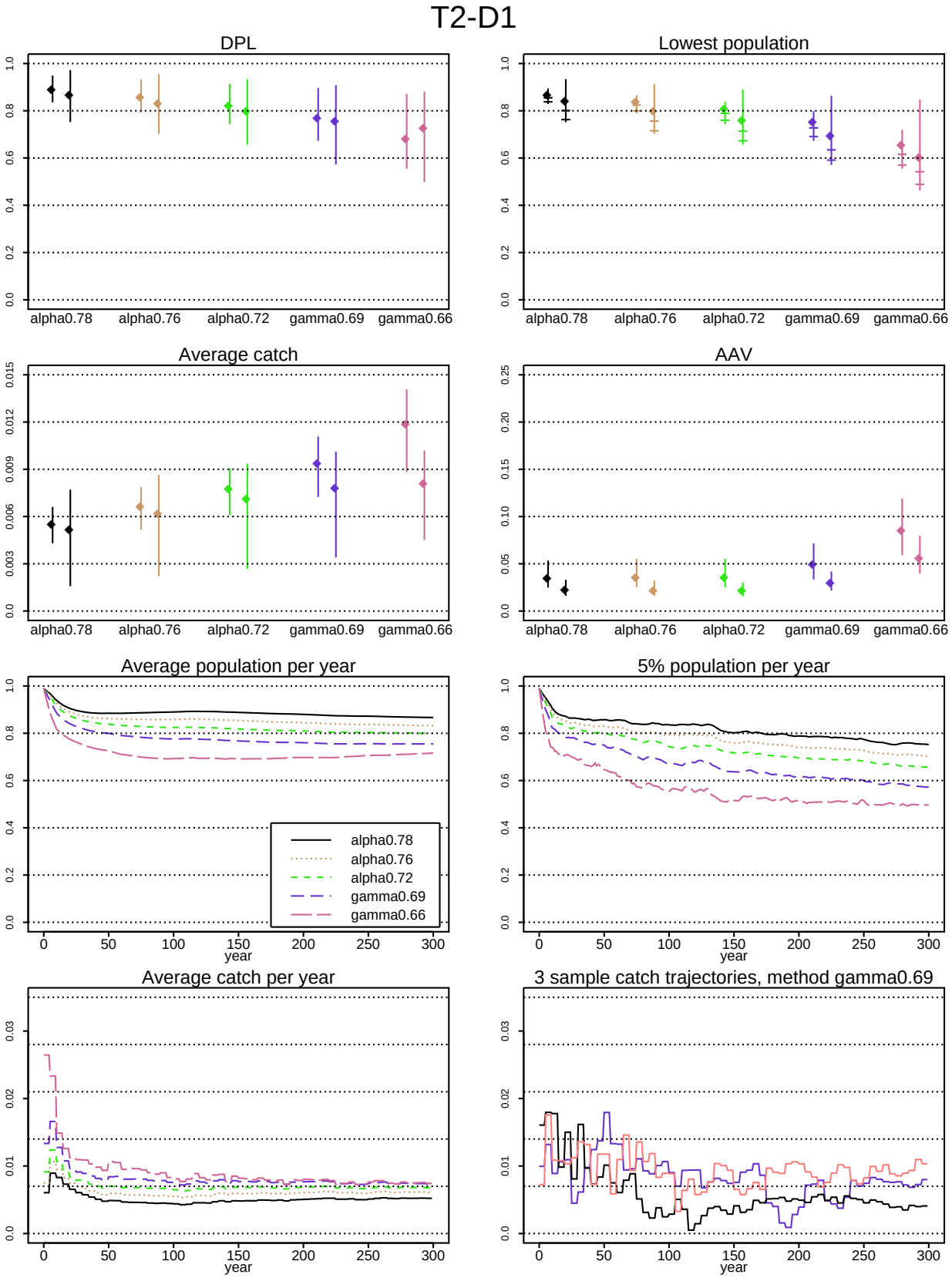


Figure 12. Trial T2-D1. 50% negative bias in abundance estimates.

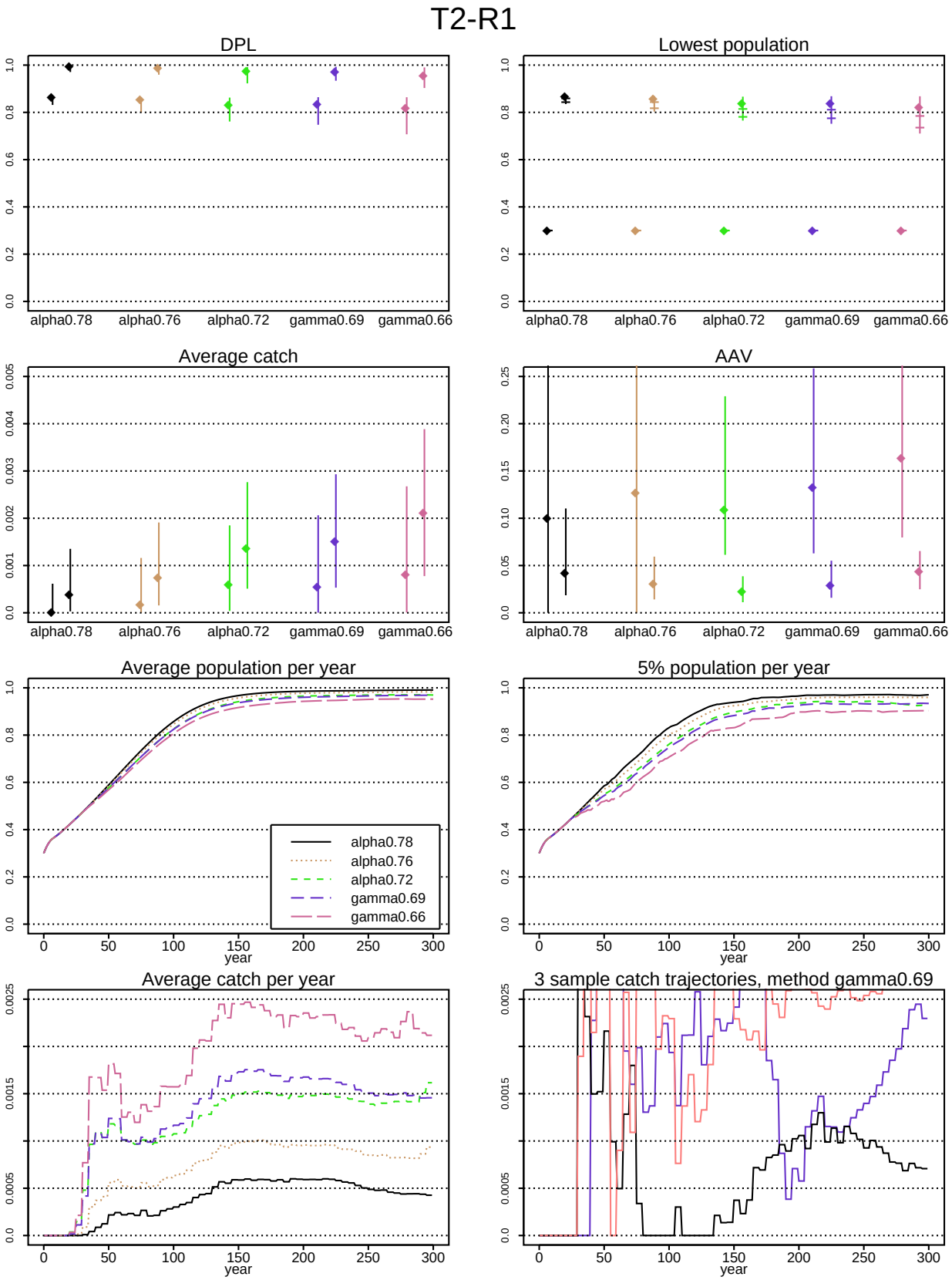


Figure 13. Trial T2-R1. 50% negative bias in abundance estimates.

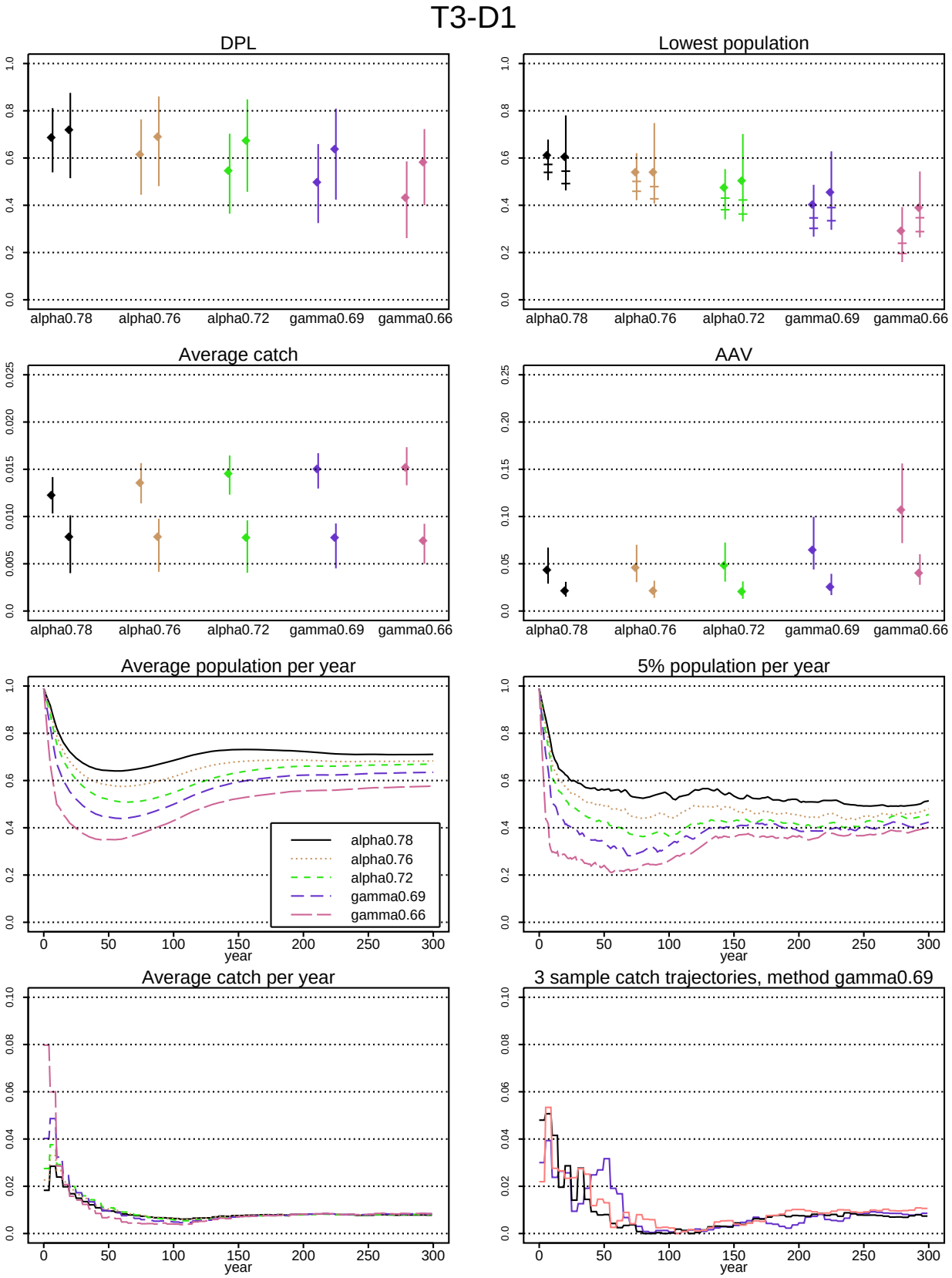


Figure 14. Trial T3-D1. 50% positive bias in abundance estimates.

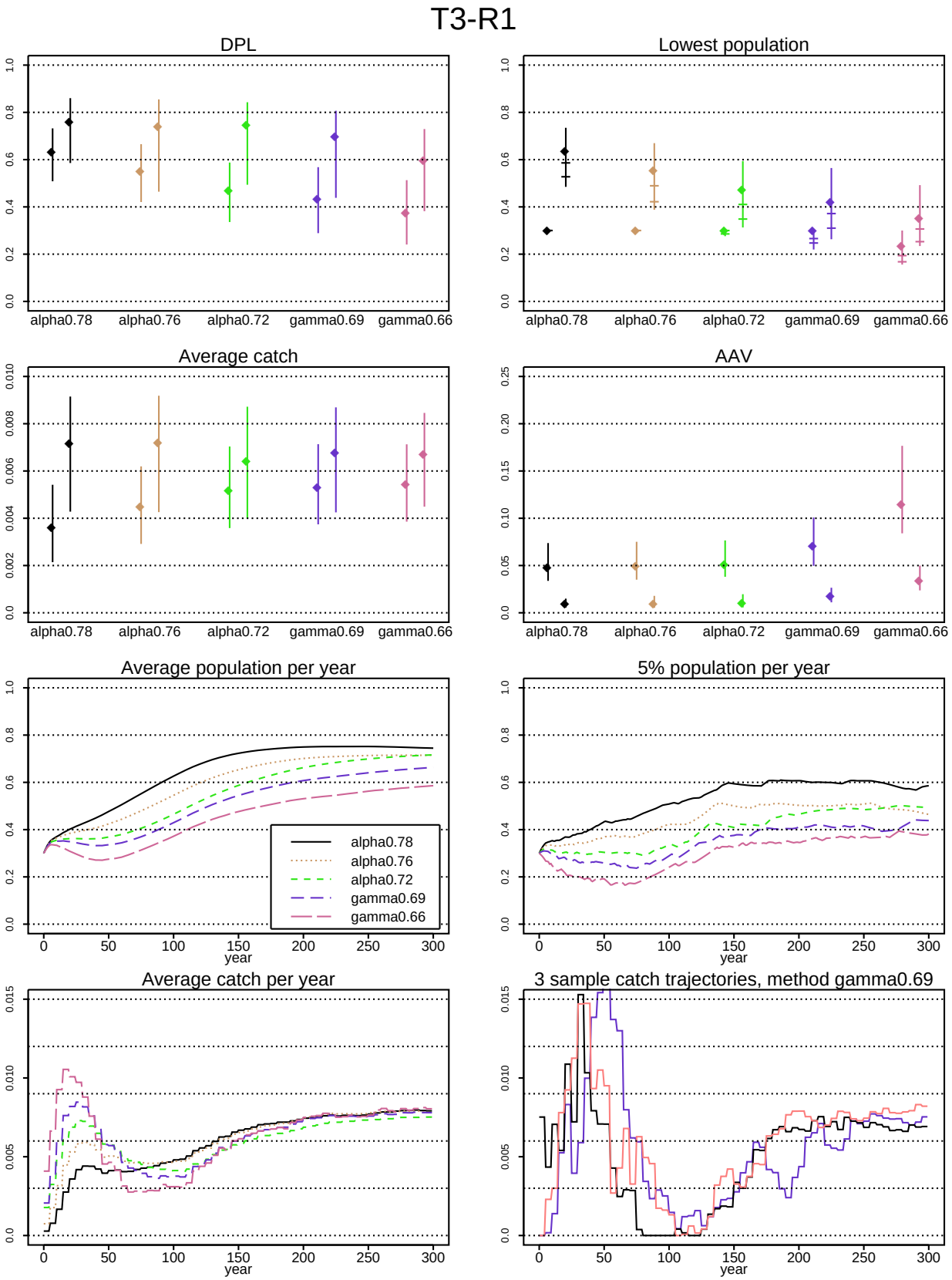


Figure 15. Trial T3-R1. 50% positive bias in abundance estimates.

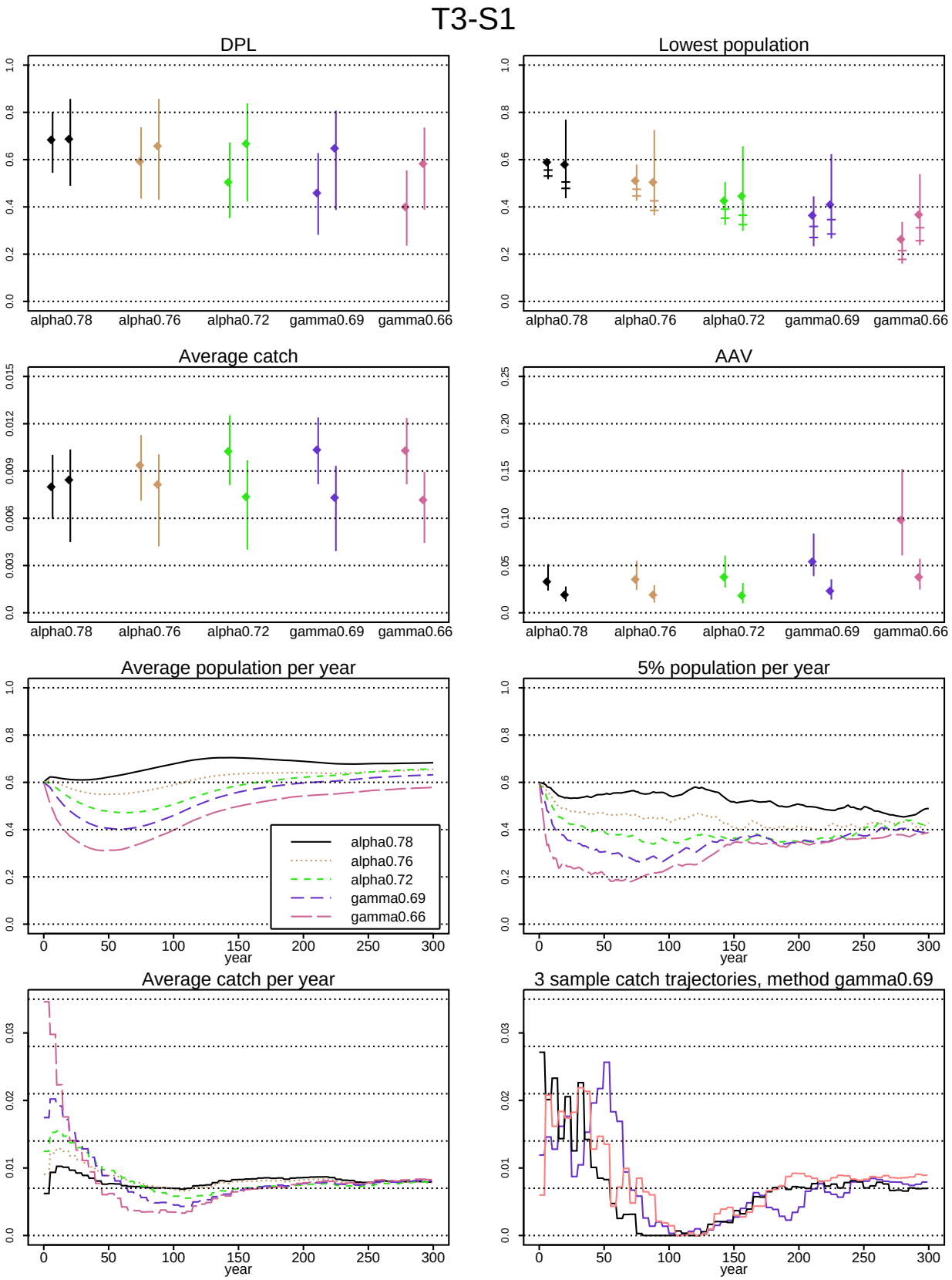


Figure 16. Trial T3-S1. 50% positive bias in abundance estimates.

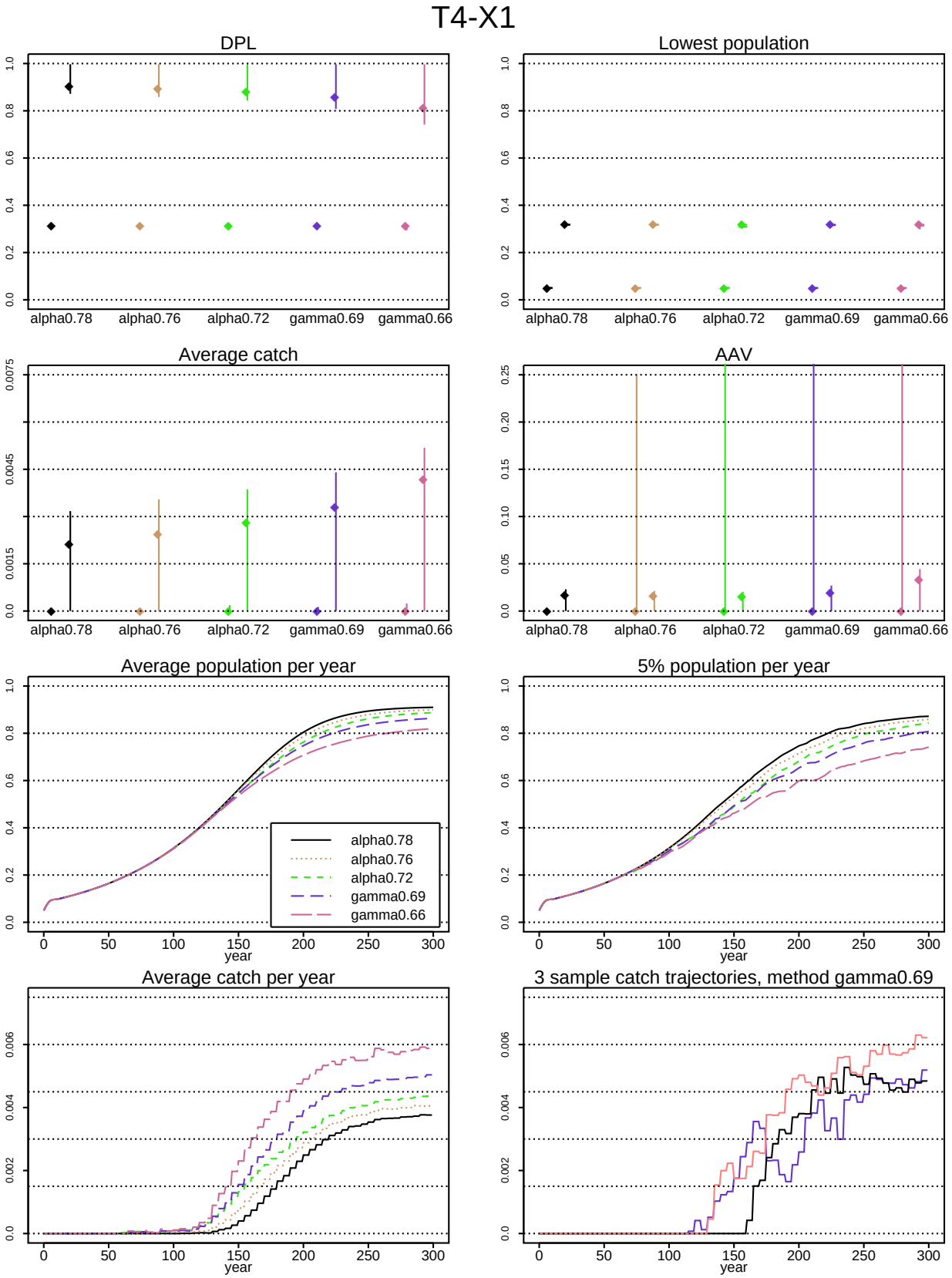


Figure 17. Trial T4-X1. Initial depletion 0.05K.

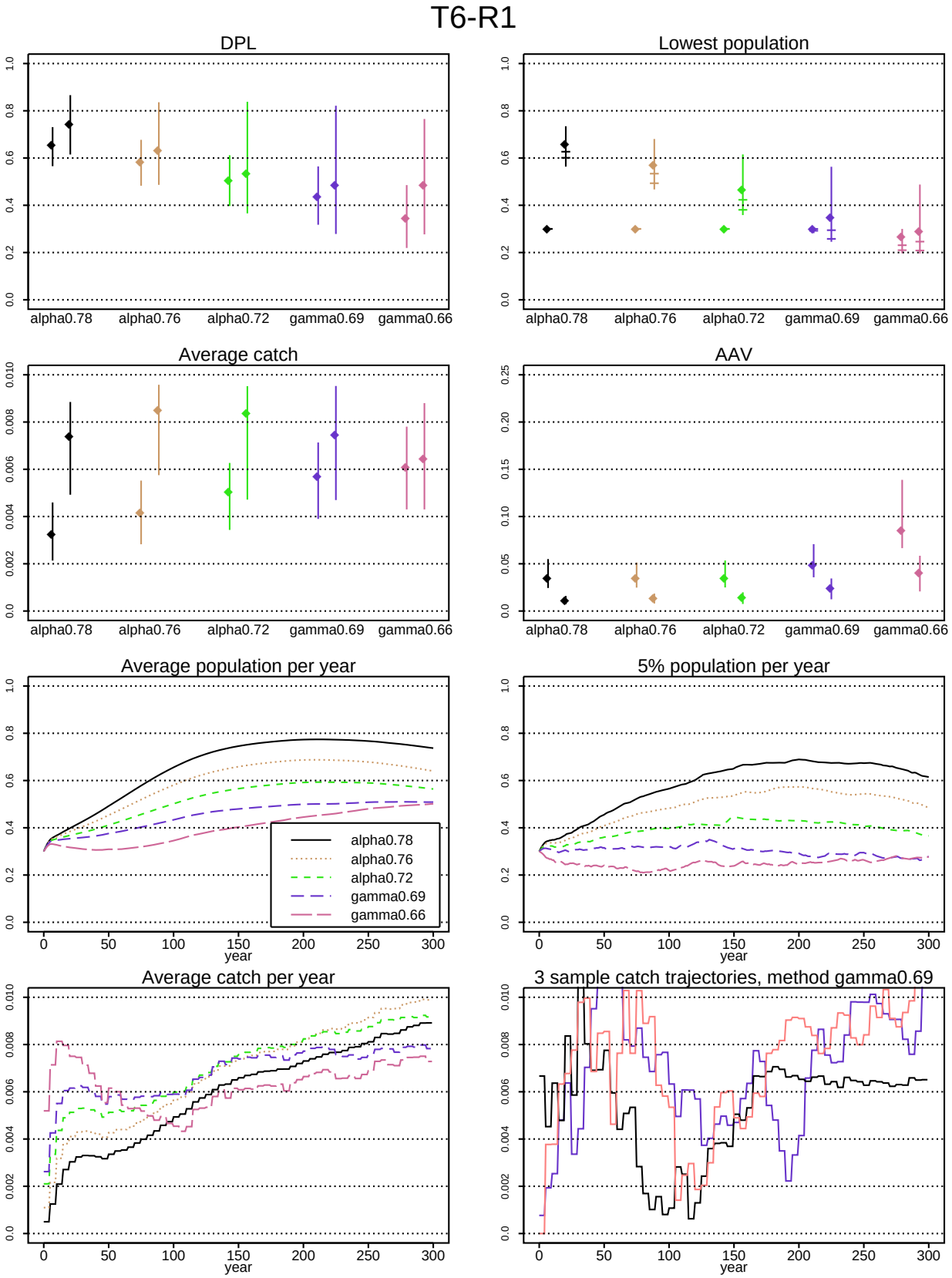


Figure 18. Trial T6-R1. Reported historic catch = 50% of true catch.

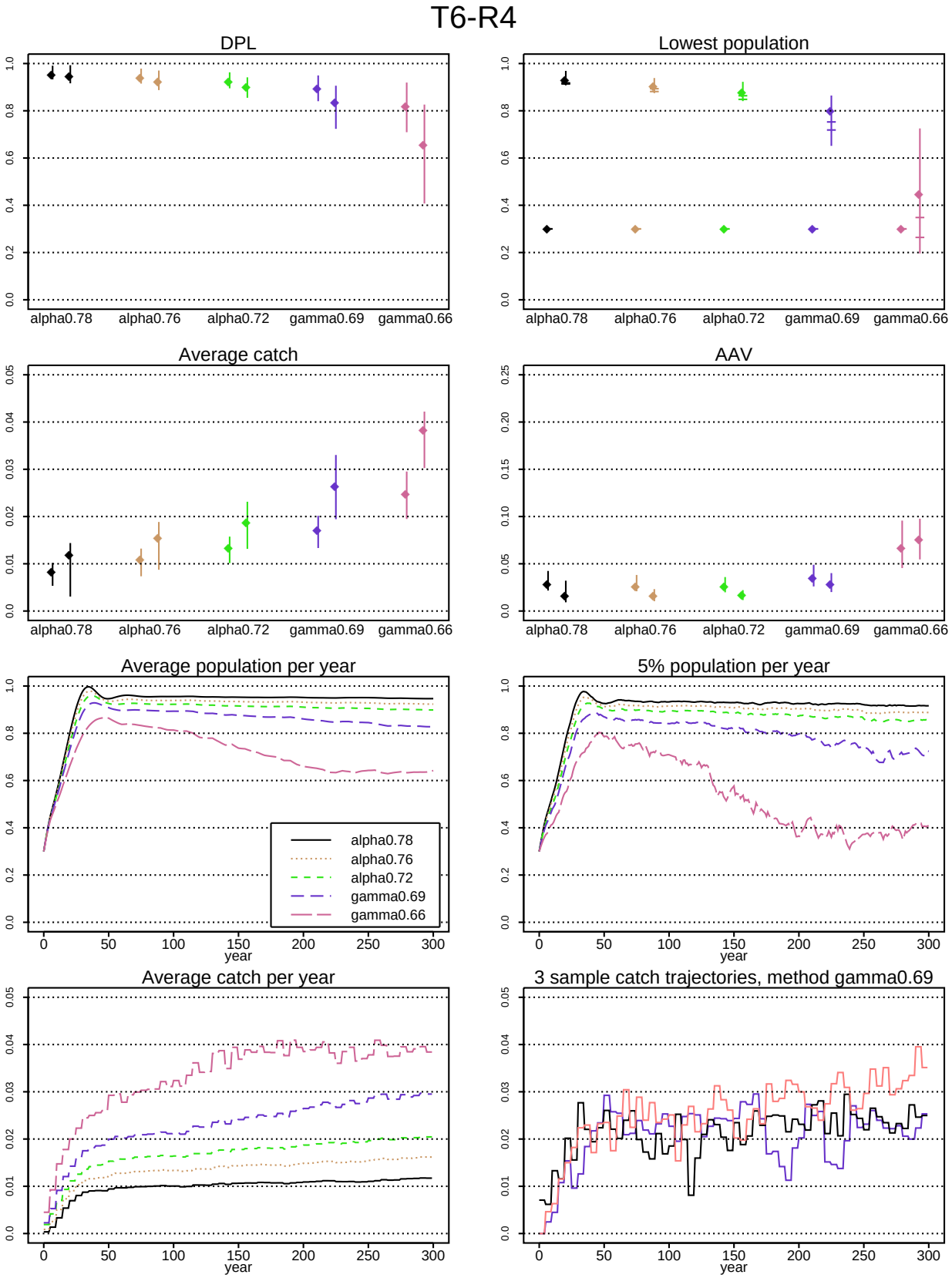


Figure 19. Trial T6-R4. Reported historic catch = 50% of true catch.

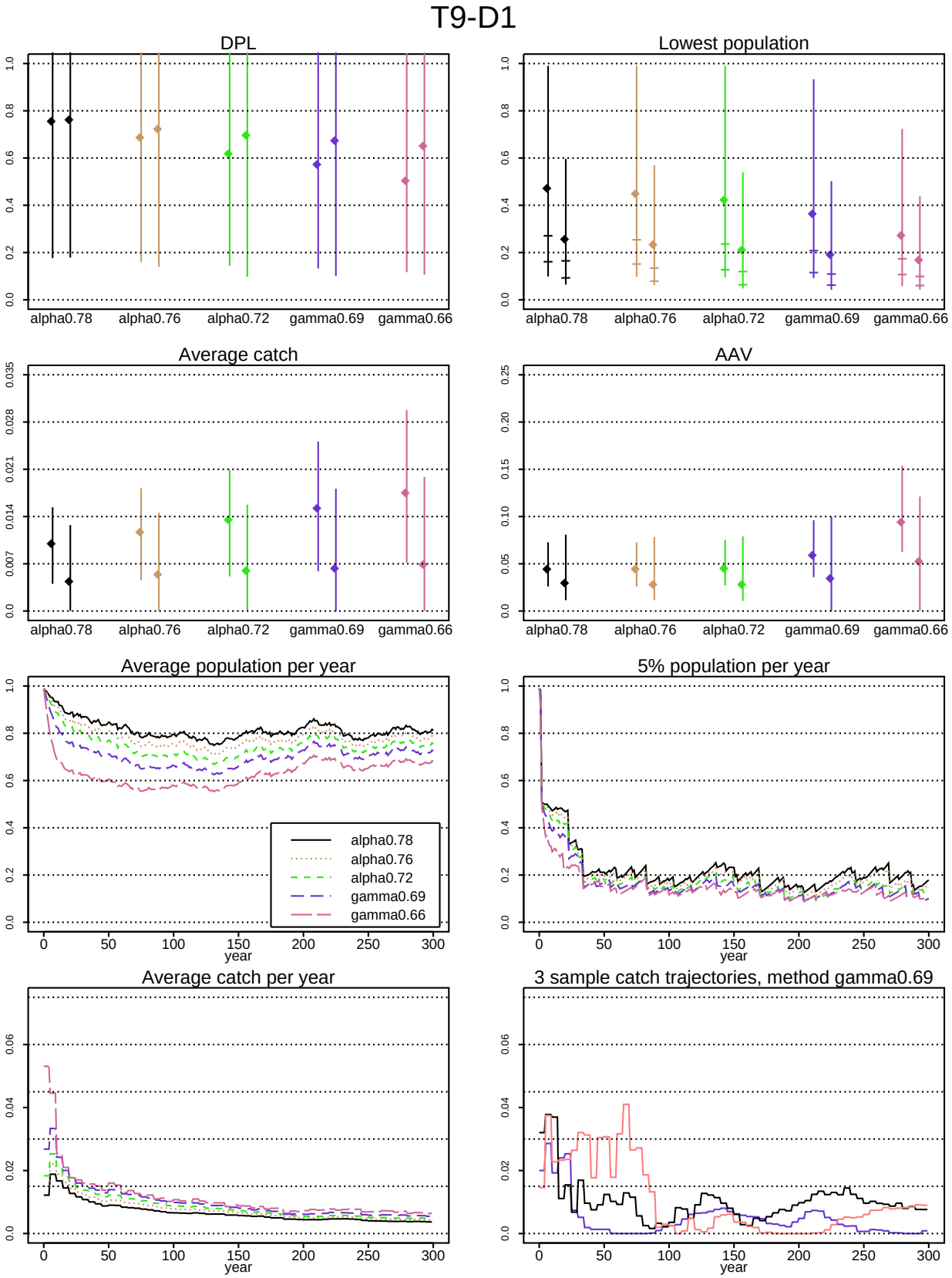


Figure 20. Trial T9-D1. Episodic events; (Rate = 0.02).

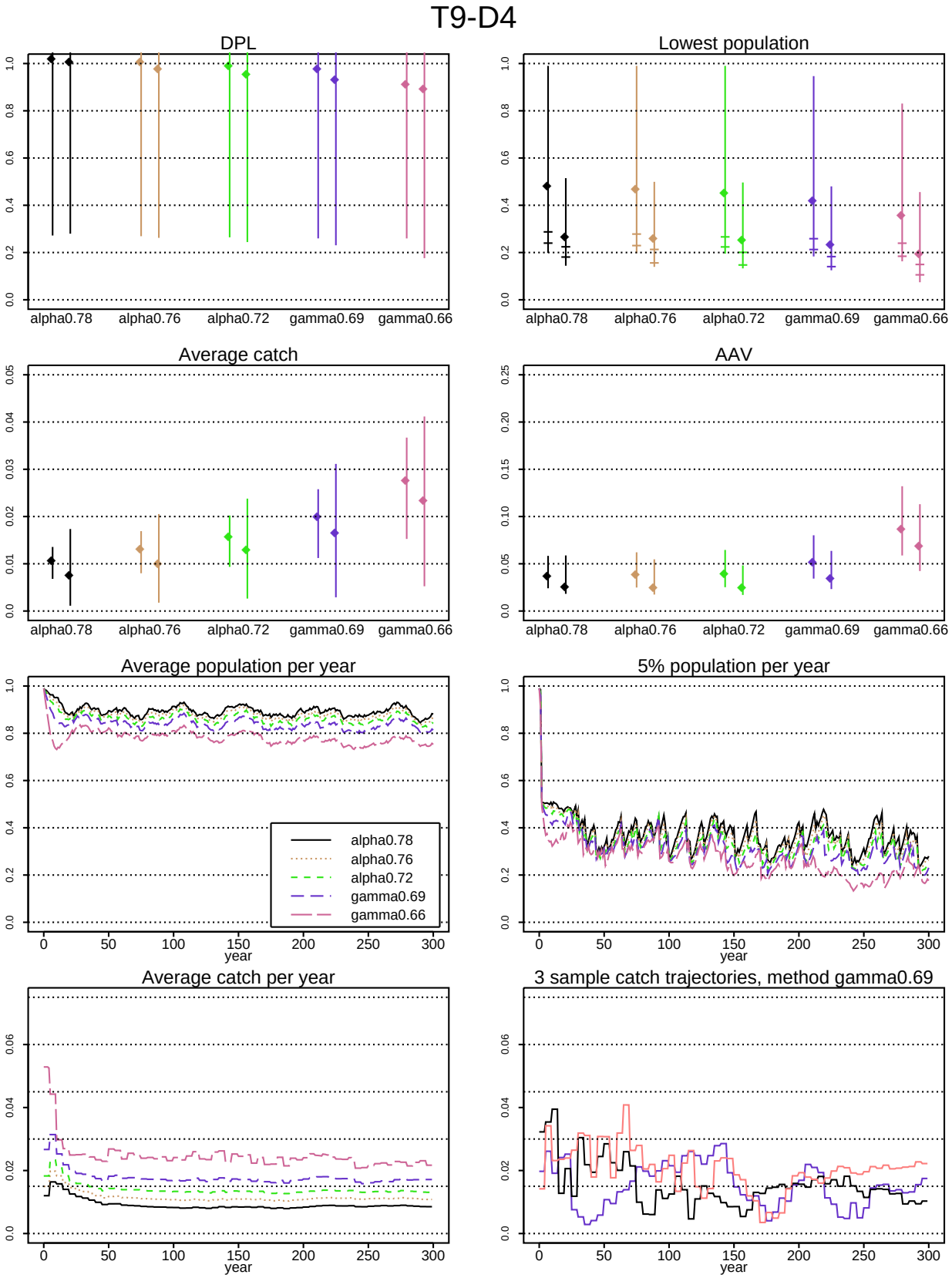


Figure 21. Trial T9-D4. Episodic events; (Rate = 0.02).

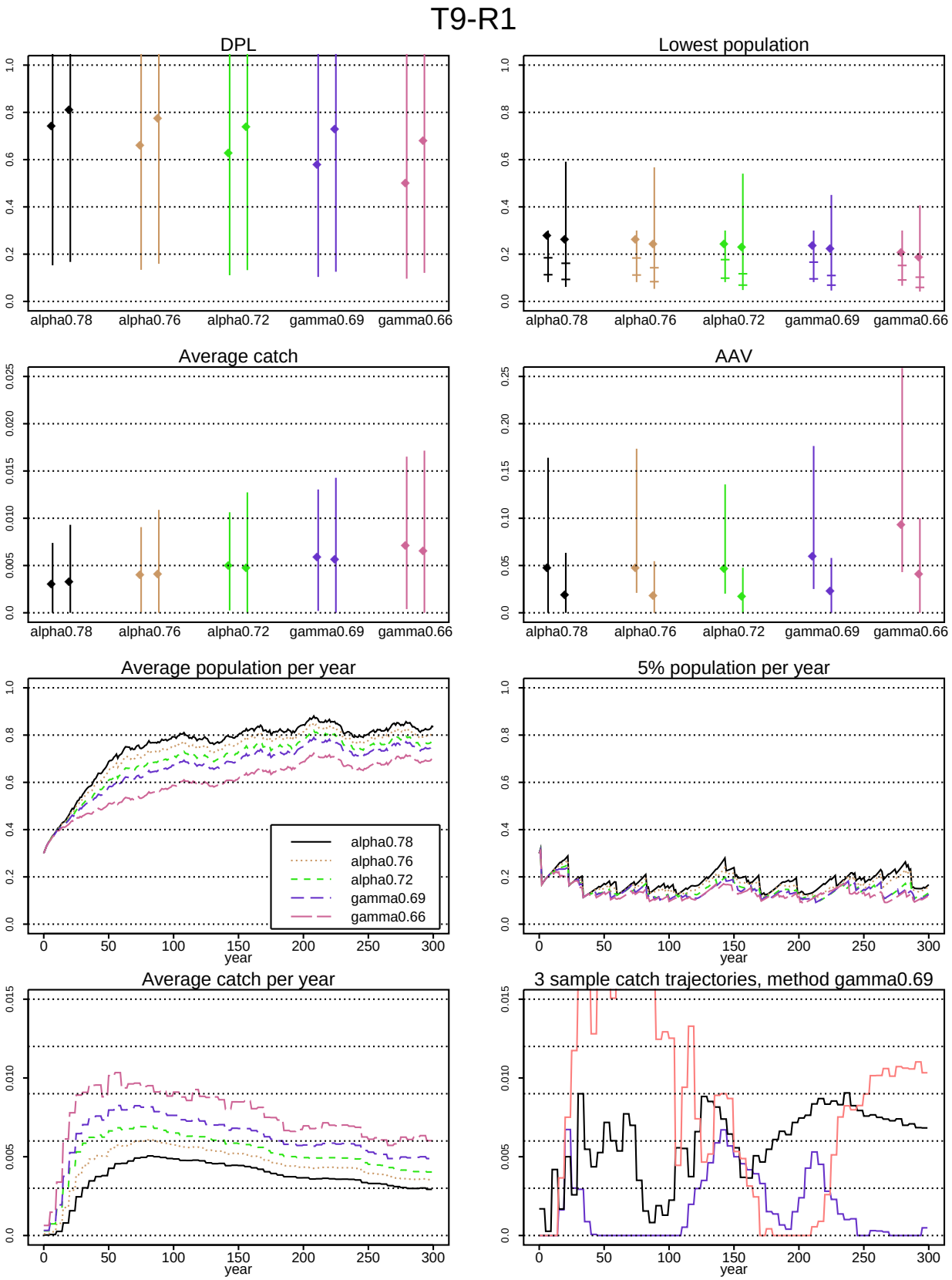


Figure 22. Trial T9-R1. Episodic events; (Rate = 0.02).

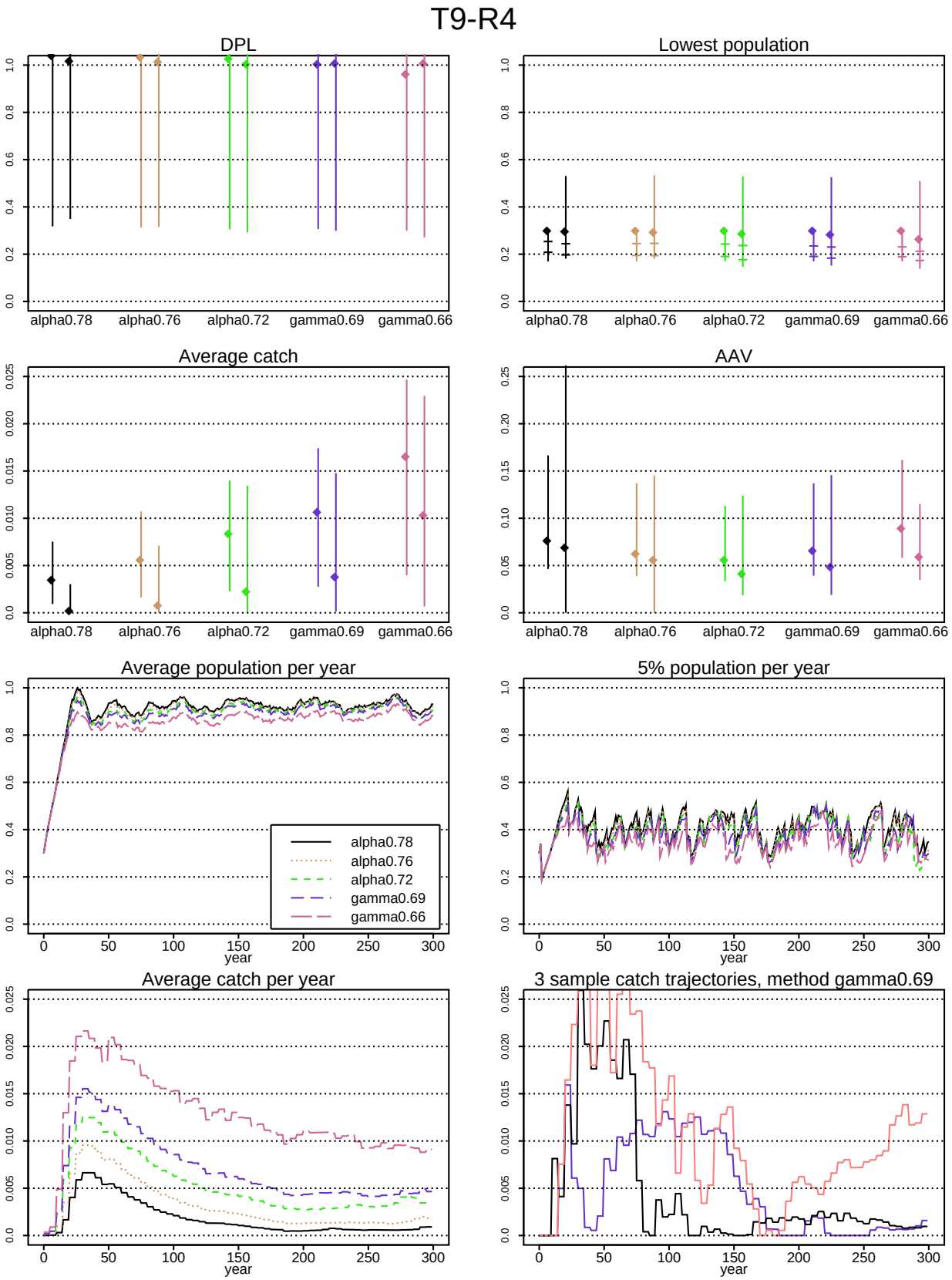


Figure 23. Trial T9-R4. Episodic events; (Rate = 0.02).

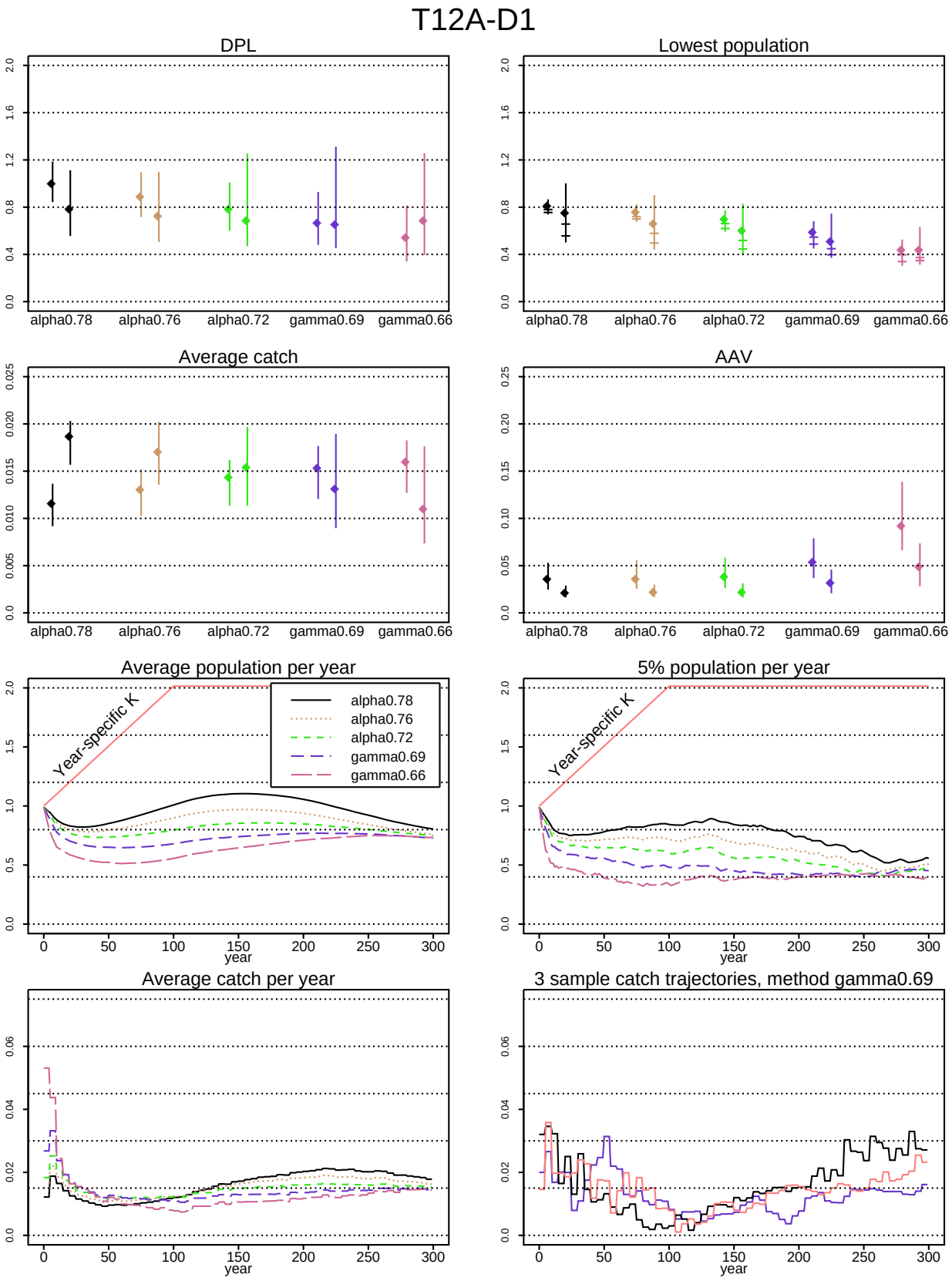


Figure 24. Trial T12A-D1. Linear increase in K from K_0 to $2 \cdot K_0$ after 100 years, then constant.

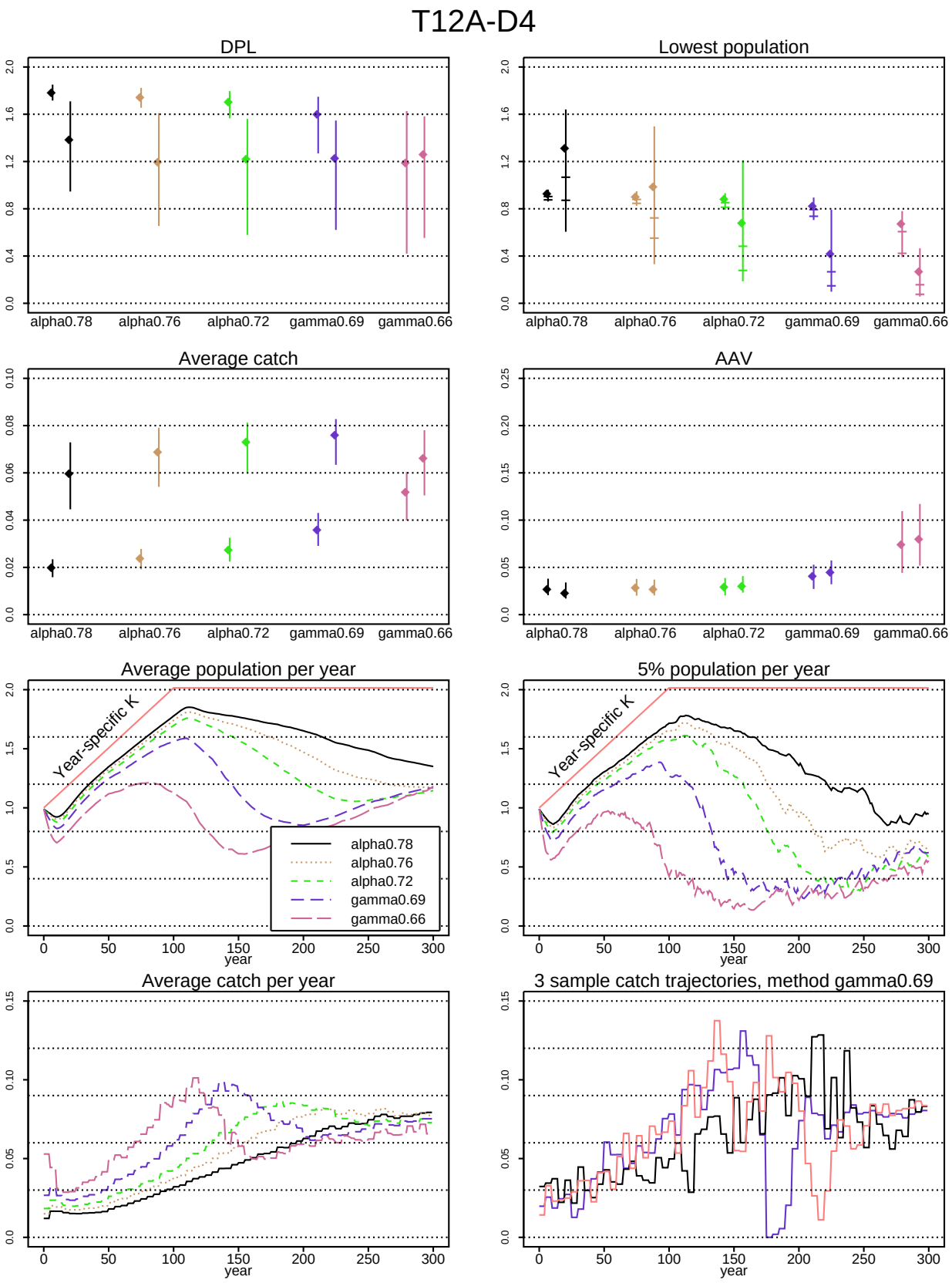


Figure 25. Trial T12A-D4. Linear increase in K from K_0 to $2 \cdot K_0$ after 100 years, then constant.

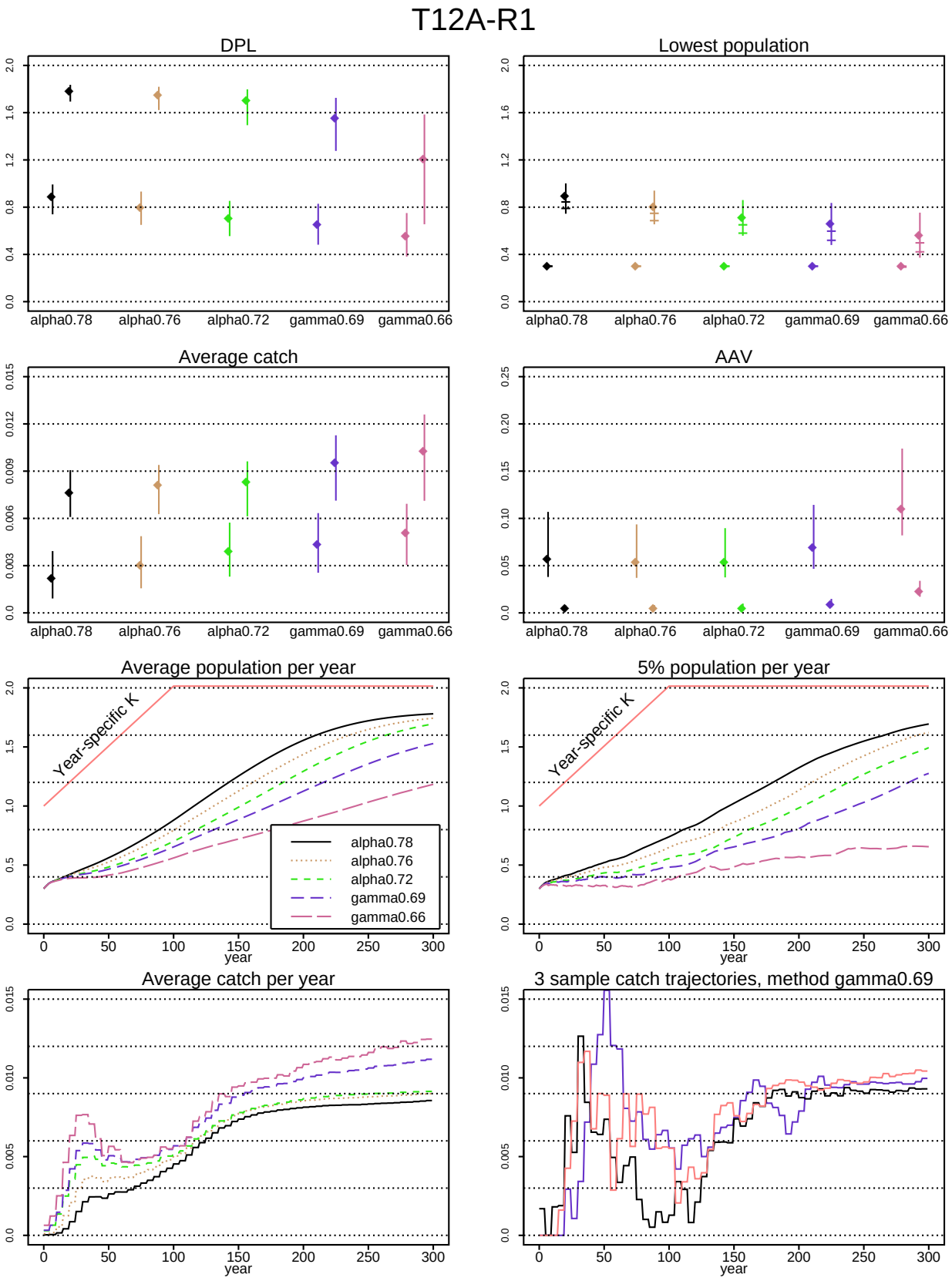


Figure 26. Trial T12A-R1. Linear increase in K from K_0 to $2 \cdot K_0$ after 100 years, then constant.

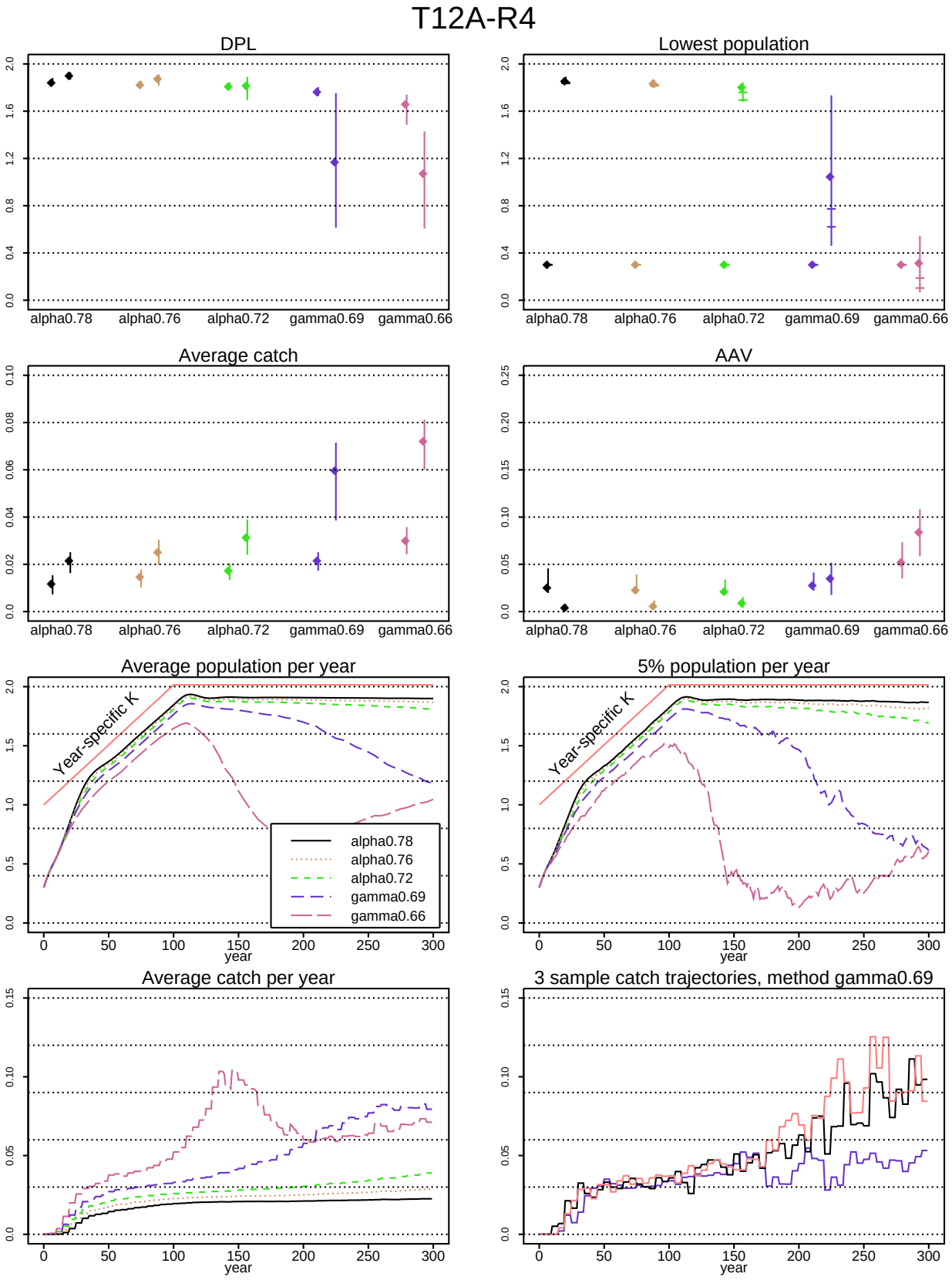


Figure 27. Trial T12A-R4. Linear increase in K from K_0 to $2 \cdot K_0$ after 100 years, then constant.

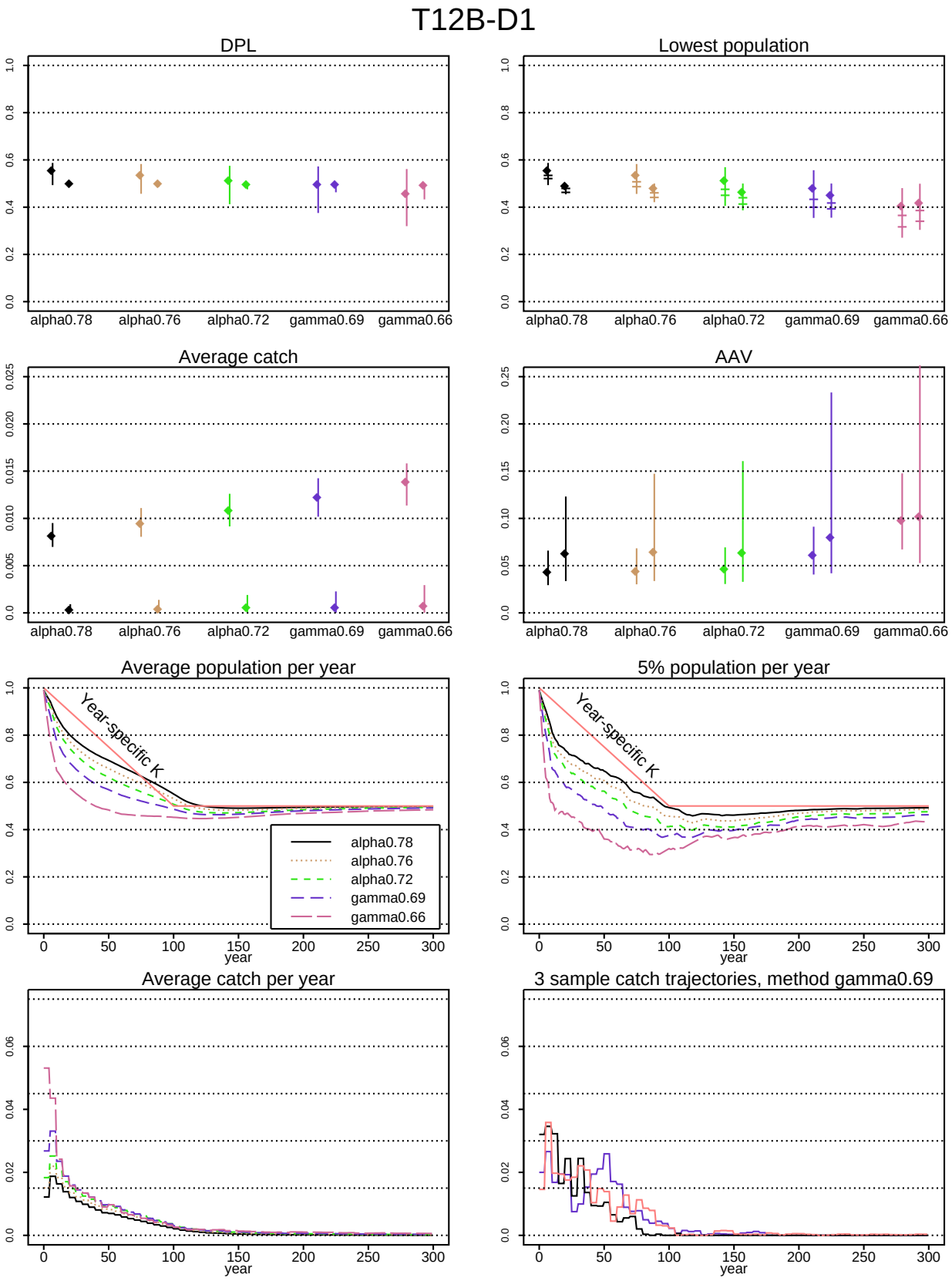


Figure 28. Trial T12B-D1. Linear decrease in K from K_0 to $0.5 \cdot K_0$ after 100 years, then constant.

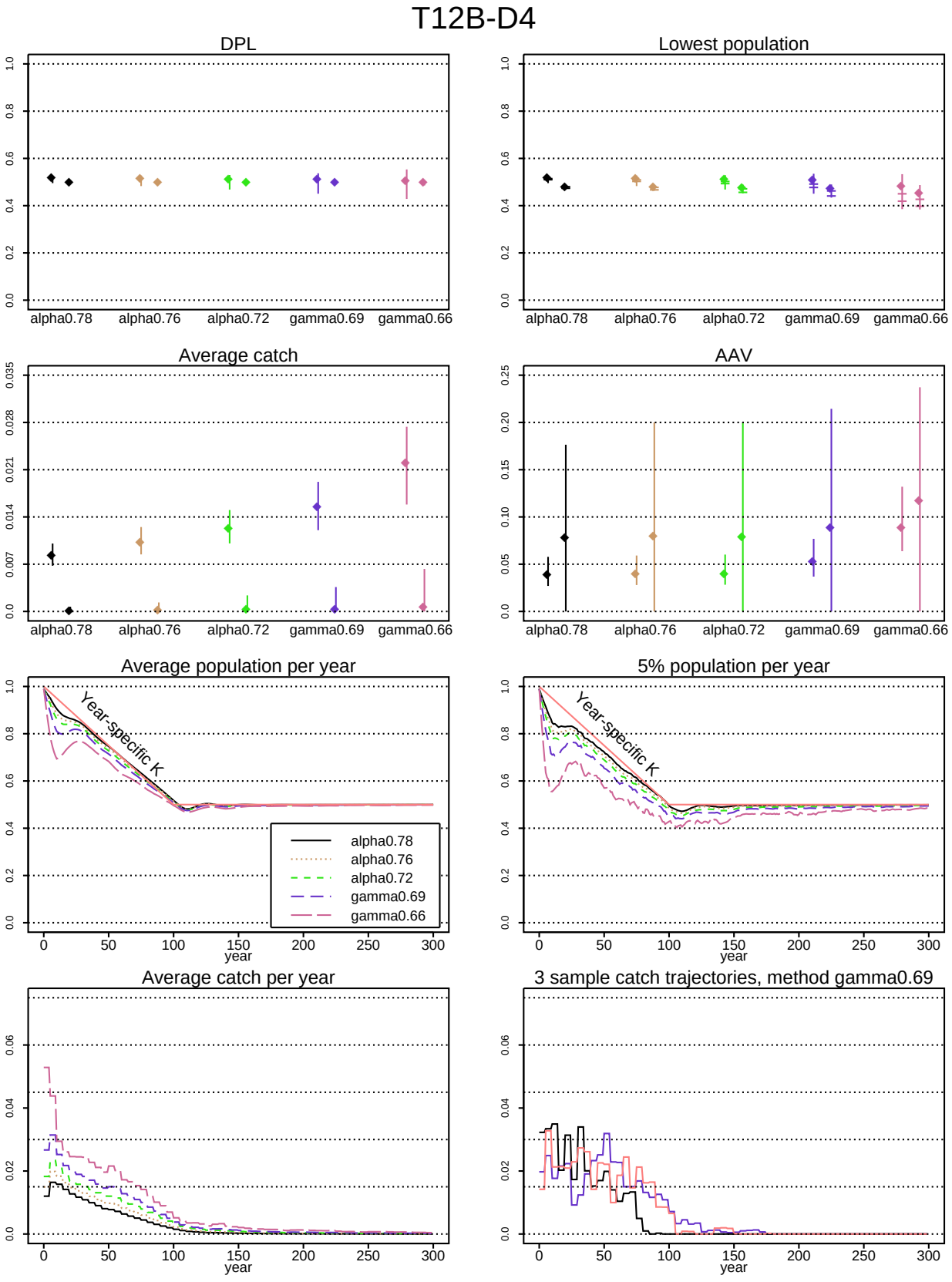


Figure 29. Trial T12B-D4. Linear decrease in K from K_0 to $0.5 \cdot K_0$ after 100 years, then constant.

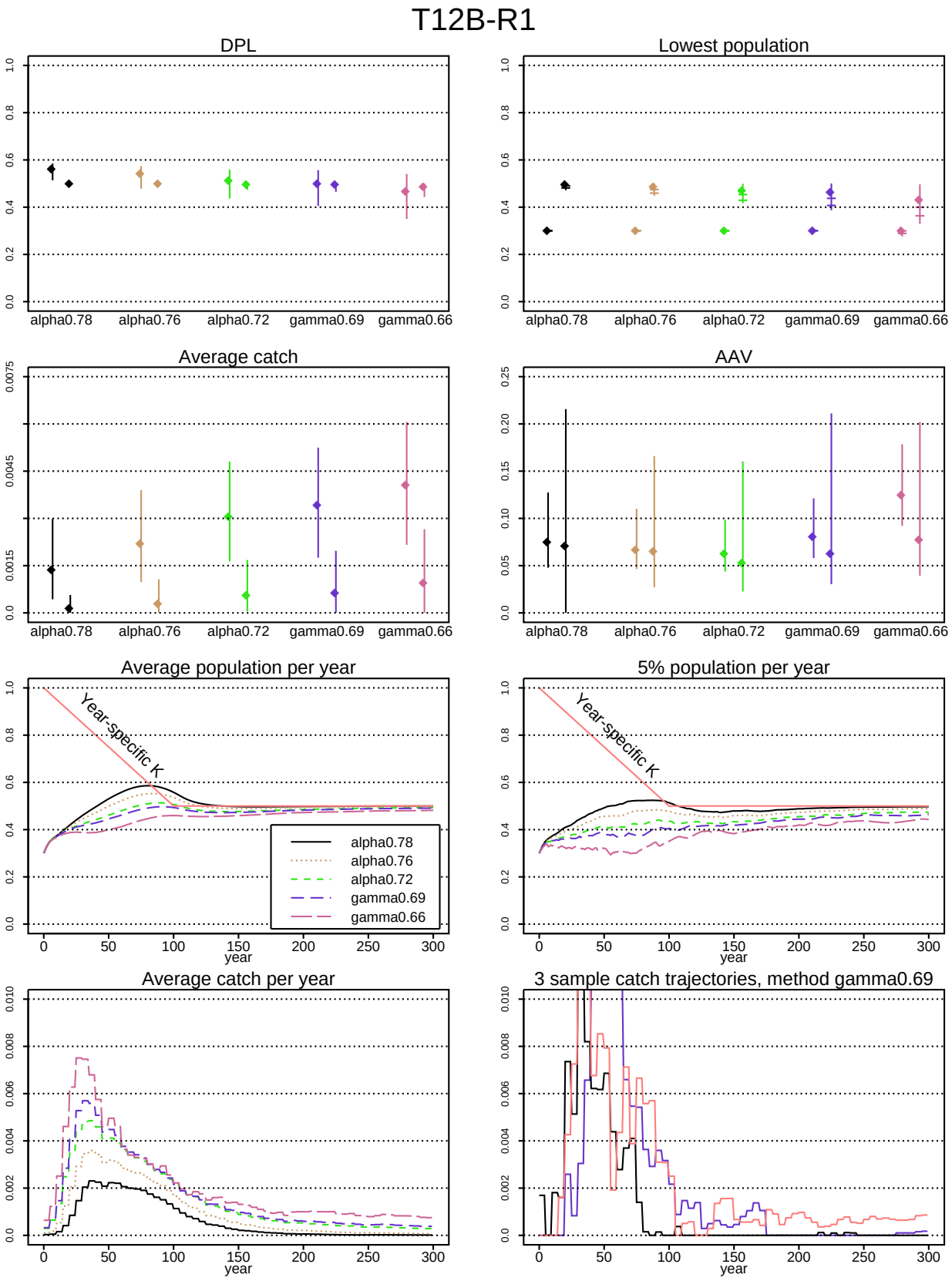


Figure 30. Trial T12B-R1. Linear decrease in K from K_0 to $0.5 \cdot K_0$ after 100 years, then constant.

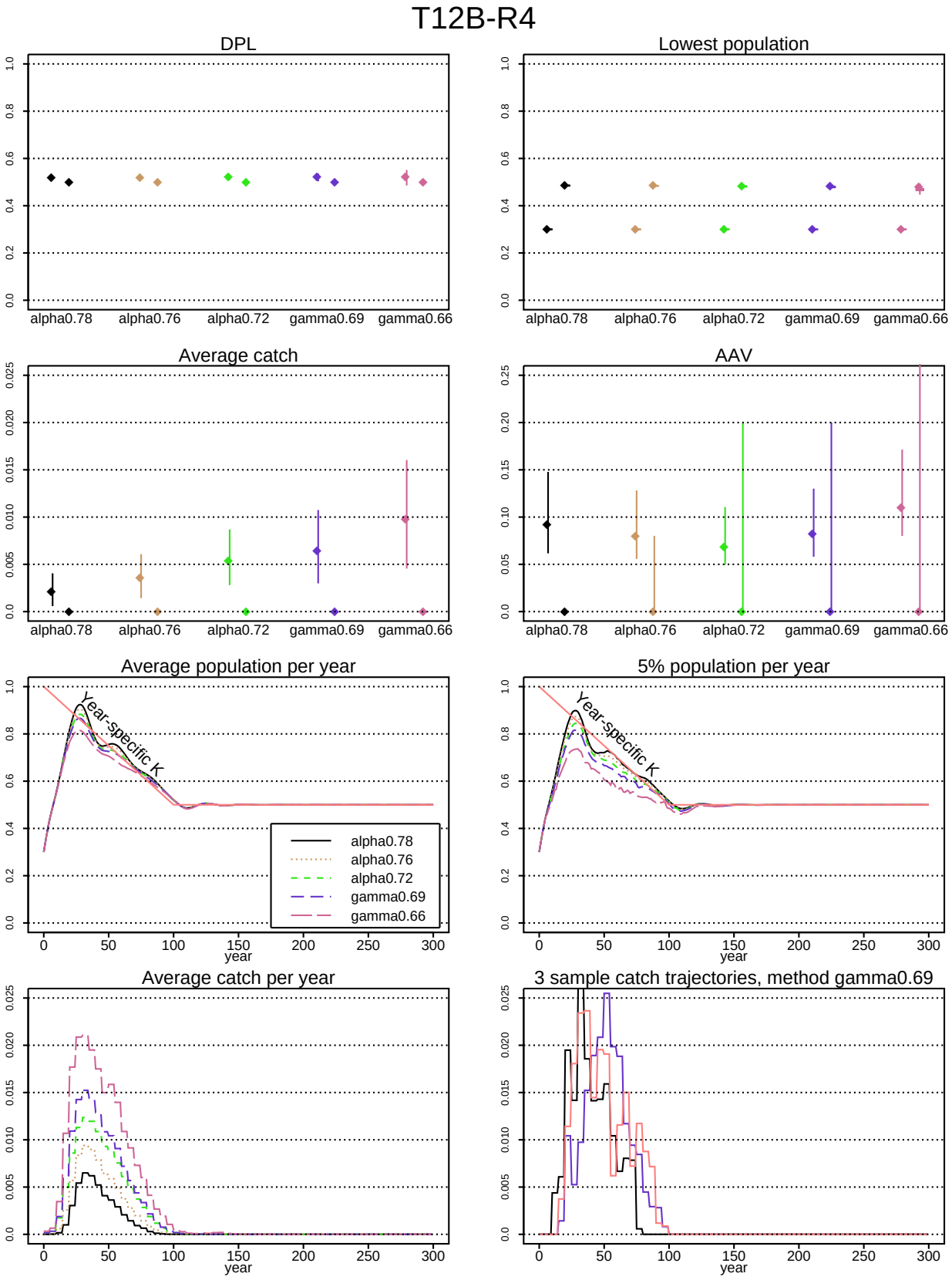


Figure 31. Trial T12B-R4. Linear decrease in K from K_0 to $0.5 \cdot K_0$ after 100 years, then constant.

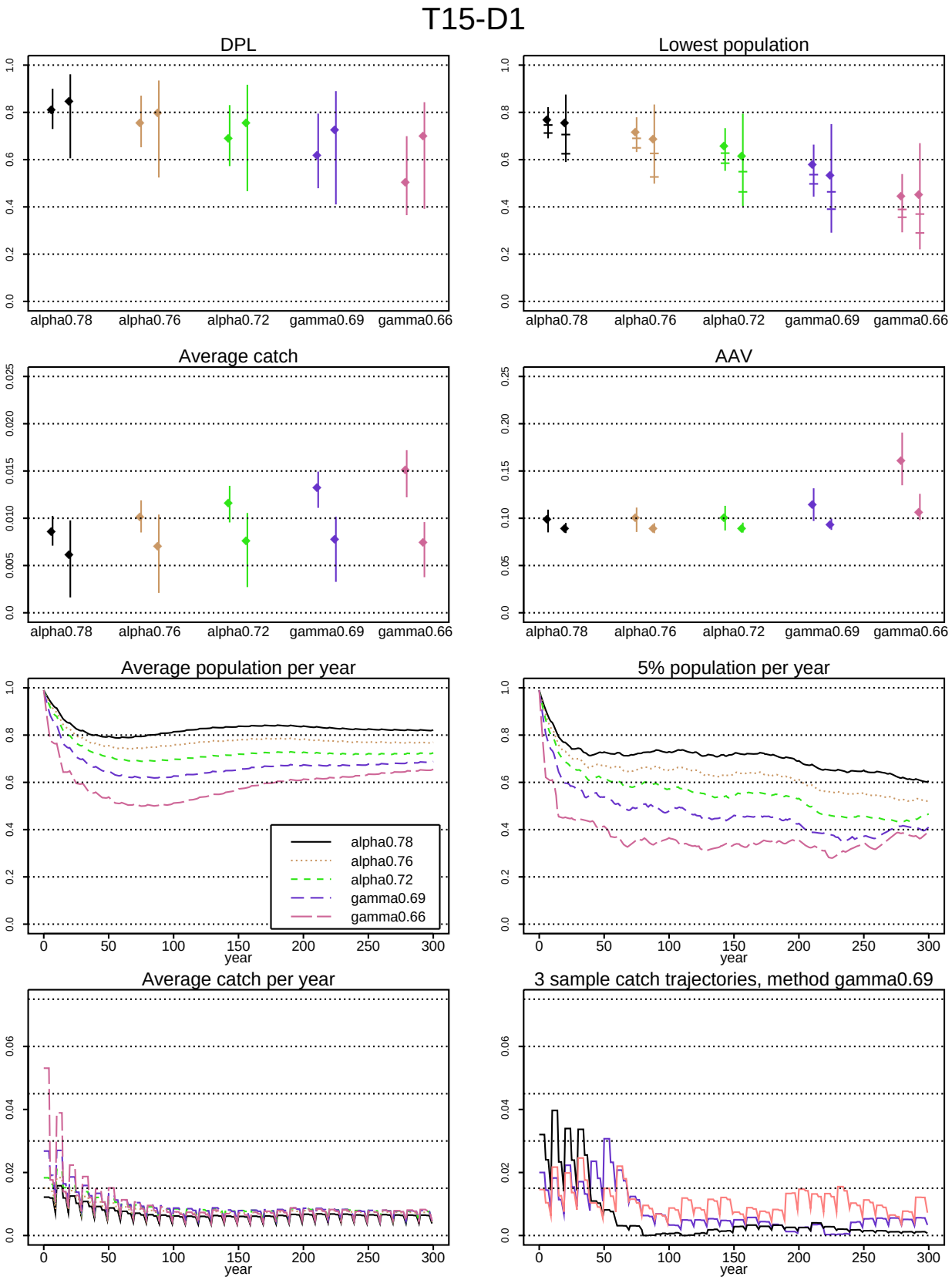


Figure 32. Trial T15-D1. Surveys every 10 years.

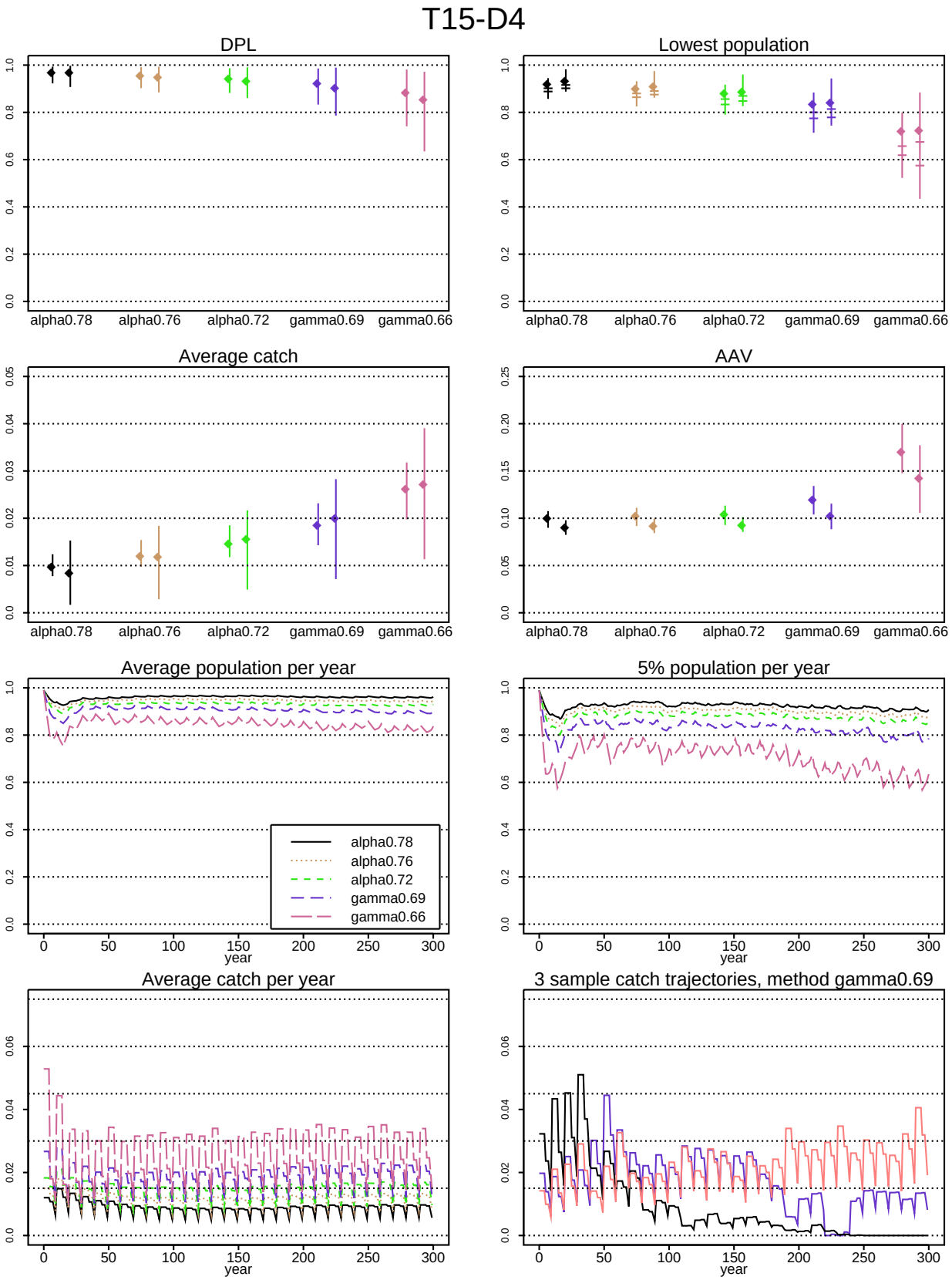


Figure 33. Trial T15-D4. Surveys every 10 years.

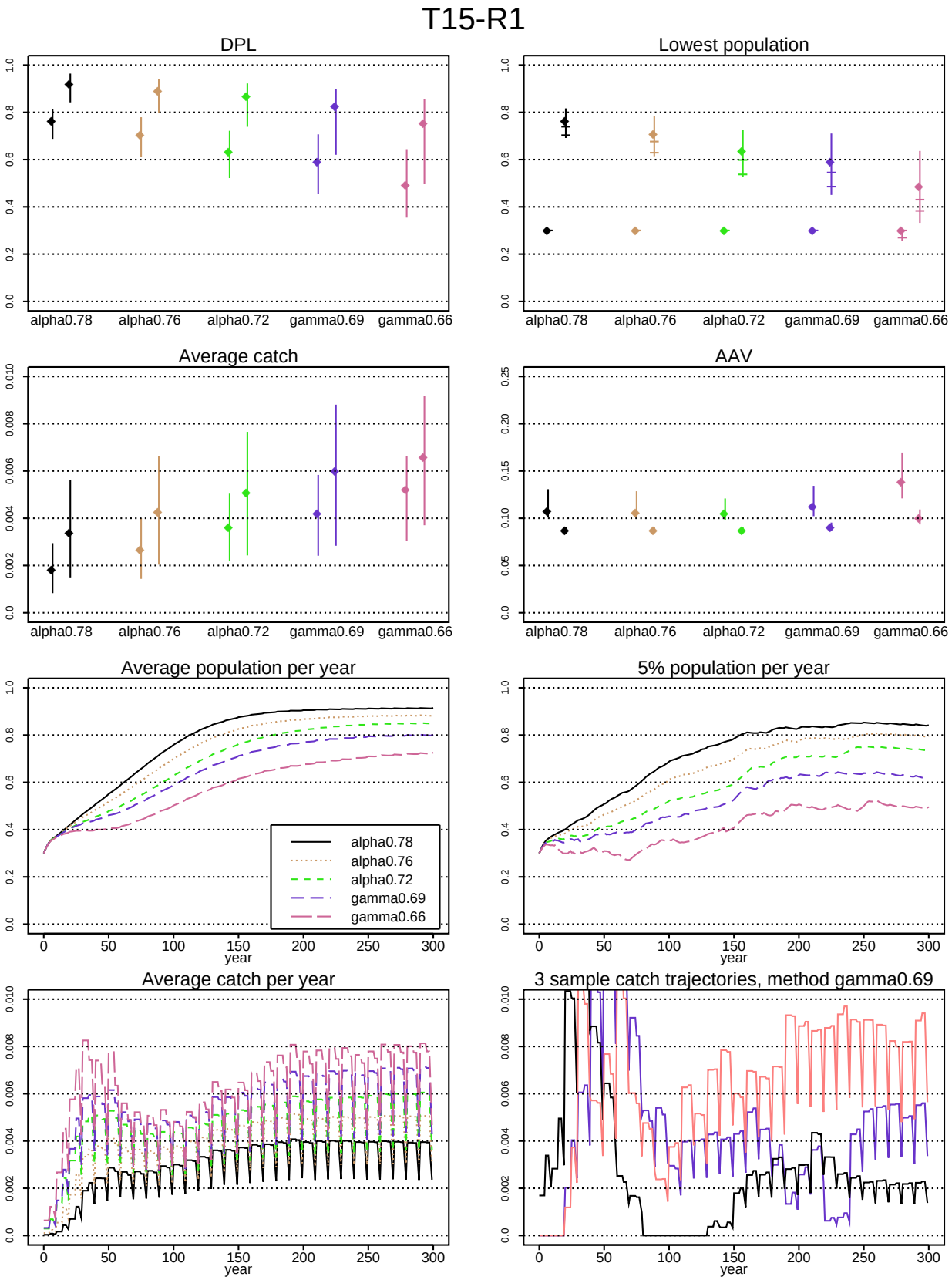


Figure 34. Trial T15-R1. Surveys every 10 years.

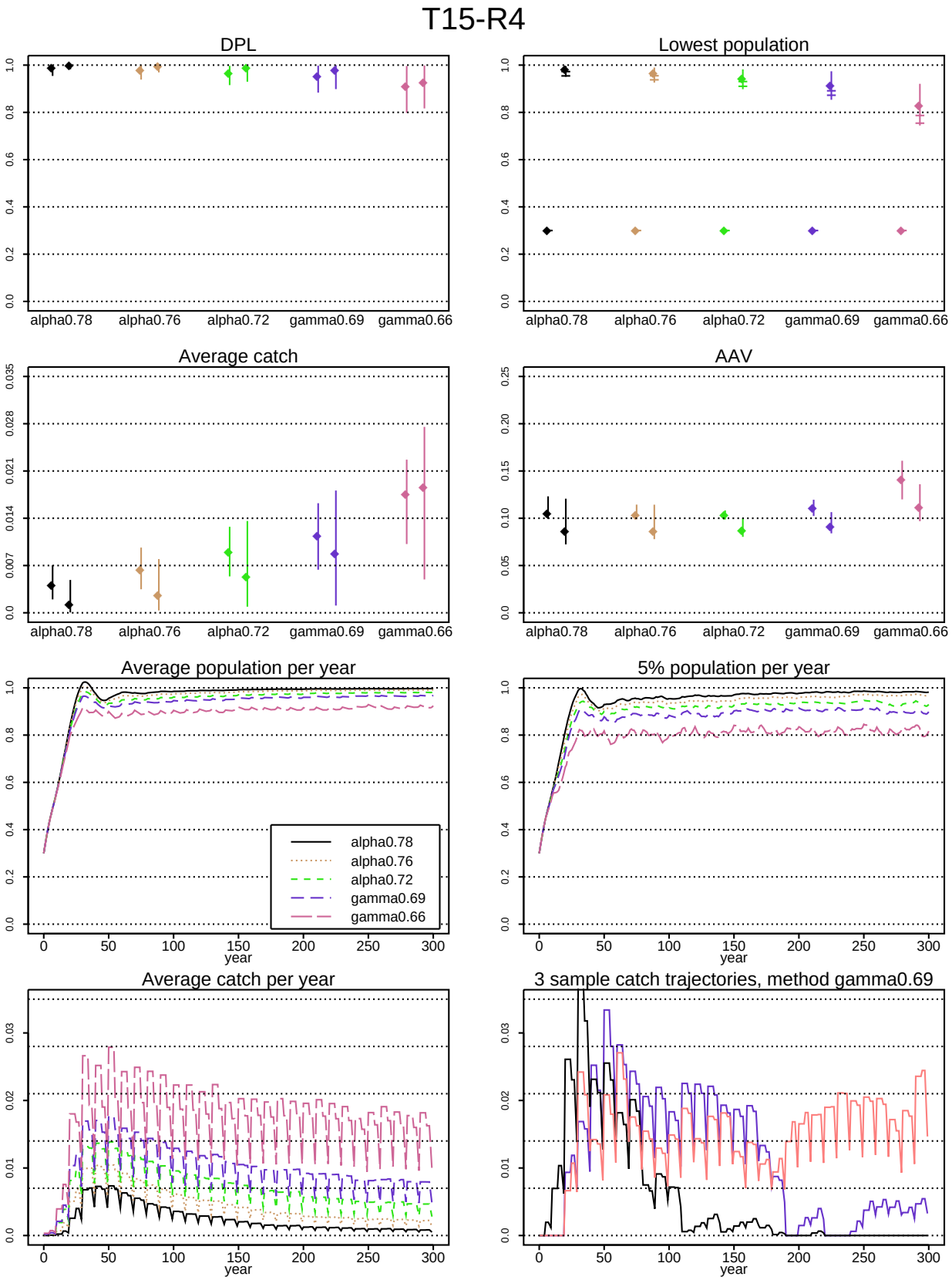


Figure 35. Trial T15-R4. Surveys every 10 years.

A Appendix Tuning details

By the "new" tuning process we mean the process that determines the parameters of the CLA to yield a specified median final depletion level in the T1-D1 trial with $MSYR^{1+}$=1%. The final depletion is equal to the final population size after $NYEAR$=300 years of management divided by the pristine population size. In the tuning of CLA (see 2.3), γ has been varied, whereas $\alpha = 0.5$ has been fixed.

Let DPL denote the final depletion. The population model is stochastic. Accordingly, DPL is a random variable. The median of DPL, which is denoted $med(DPL)$, is a function of γ. On a short interval this function is approximately linear. Thus

$$med(DPL) = a + b\,\gamma$$

where a and b are constants. The sample median of DPL based on N independent realizations of DPL is approximately equal to

$$med(DPL) + \epsilon$$

where ϵ is a random variable with expectation zero if N is sufficiently large. Let $y_1, y_2, \ldots, y_M$ be sample medians from M independent simulation trials with γ-values $\gamma_1, \gamma_2, \ldots, \gamma_M$; with N simulations in each trial. We have chosen $N = 1000$. Then we have approximately

$$y_i = a + b\,\gamma_i + \epsilon_i; i = 1, 2, \ldots, M$$

where the ϵ_is are independent with zero expectation with equal variance provided that the γ_is vary in a short interval. It follows that the parameters a and b can be estimated by ordinary least squares regression. Let $\hat{a}$ and $\hat{b}$ denote the estimates. The estimated median final depletion level is then given by

$$\hat{med(DPL)}(\gamma) = \hat{a} + \hat{b}\,\gamma \tag{A.1}$$

for a given γ. Having specified the median final depletion level we can solve for γ. Let γ^* be the solution. Then

$$\hat{med(DPL)}(\gamma^*) = \hat{a} + \hat{b}\,\gamma^*$$

is an estimator of $med(DPL)(\gamma^*)$. The estimated standard error of this estimator is

$$\hat{SE}(med(\hat{DPL})) = s\sqrt{\frac{1}{M} + \frac{(\gamma^* - \bar{\gamma})^2}{S_\gamma^2}} \tag{A.2}$$

where $\bar{\gamma}$ is the average taken over the γ_is, and s is defined by

$$s^2 = \frac{1}{M-2}\sum_{i=1}^{M}(y_i - \hat{a} - \hat{b}\,\gamma_i)^2 \tag{A.3}$$

and S_γ^2 is defined by

$$S_\gamma^2 = \sum_{i=1}^{M} (\gamma_i - \bar{\gamma})^2. \tag{A.4}$$

An approximate 95% confidence interval for $med(DPL)(\gamma^*)$ is then

$$(med(\hat{D}PL)(\gamma^*) - 2\,\hat{SE}(med(\hat{D}PL)), med(\hat{D}PL)(\gamma^*) + 2\,\hat{SE}(med(\hat{D}PL))).$$

The endpoints of the corresponding confidence interval for γ^* are found by replacing the left-hand side of (A.1) with the endpoints of the confidence interval for $med(DPL)(\gamma^*)$ and then solving for γ.

We have tuned the catch limit algorithm for the two cases where the median final depletion level is either 0.66 or 0.69. The results are based on independent sample medians of DPL corresponding to different values of γ. Each sample median is based on 1000 independent trials.

Figures A.1 and A.2 show the plots relevant to tuning to median depletion levels of 0.66 and 0.69, respectively. The ordinary least squares regression line is plotted along with the plot of sample medians of depletion vs. γ. The results are summarised in Table A.1.

α	γ	"New" tuning level with $NYEAR$=300 and $MSYR^{1+}$=1% $med(DPL)$	95% conf.int. for $med(DPL)$
0.5	4.7157	0.66	(0.659, 0.661)
0.5	9.3443	0.69	(0.689, 0.691)

Table A.1. Tuning results. The columns of the table represent the values of α and γ, the tuning level, and the 95% confidence interval for the median final depletion for the given value of γ.

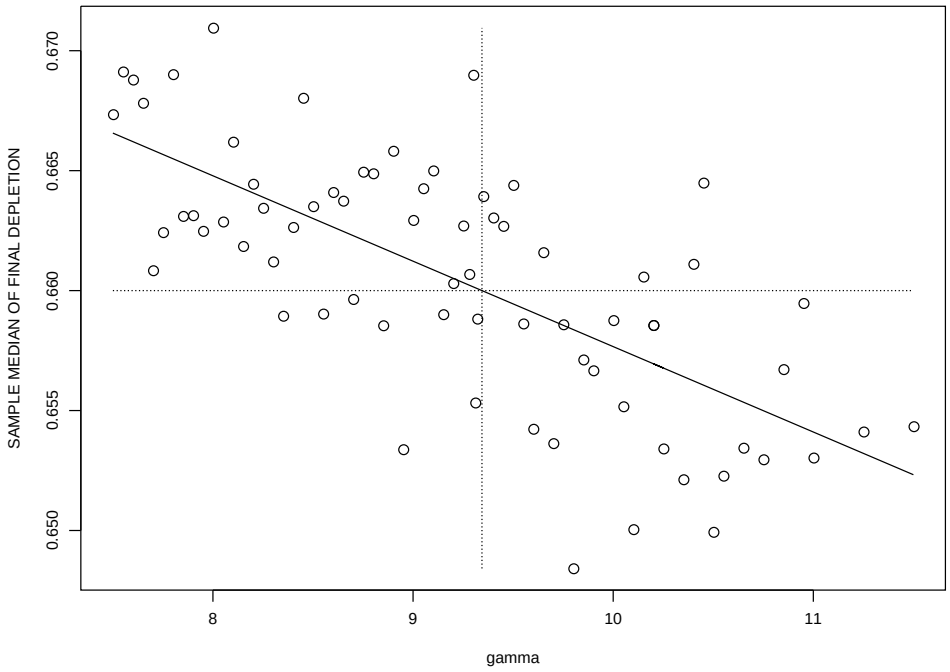


Figure A.1. Tuning to “new” tuning level 0.66.

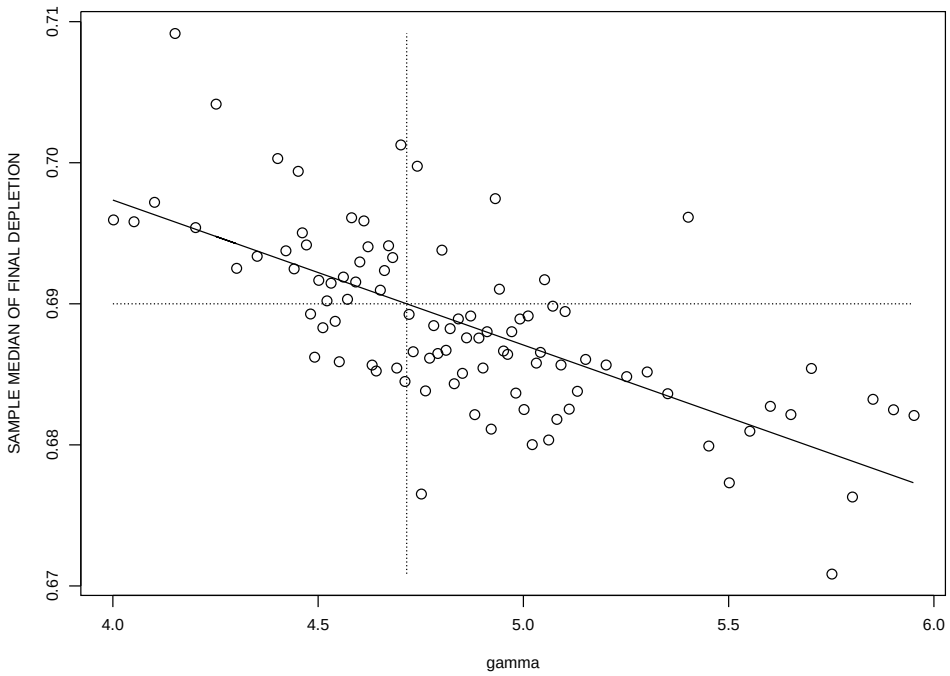


Figure A.2. Tuning to “new” tuning level 0.69.

B Appendix Accuracy adjustment

Details concerning the computation of the catch limit are described in section 3 of (Huseby and Aldrin, 2006). In the algorithm, approximations of the catch limit are calculated using n-point Gaussian-Legendre integration rules. The procedure is carried out for $n = 8, 16, \ldots, 2^J$ where J is a positive integer, or until the difference between successive approximations becomes less than ϵ where ϵ is a positive number. The constants J and ϵ need to be specified. The corresponding input parameters to the FORTRAN subroutine CATCHLIMIT (Huseby and Aldrin, 2006) are IN_NOF_RULE and ACCQUOTA, which are related to J and ϵ by IN_NOF_RULE=$J - 2$ and ACCQUOTA=ϵ. The computational burden increases as J increases or ϵ decreases. In addition, the computational burden increases as the number of abundance estimates increases.

It is reasonable to let ϵ be proportional to K_{1+}, the pristine $1+$ population size. The proportionality factor is 0.0001 if the number of abundance estimates does not exceed 20, i.e. until 100 years of management if new abundance estimates becomes available every five years. If the number of abundance estimates is greater than 20 the proportionality factor is 0.0004.

In the simulation trials performed here, $J = 10$ as far as the number of abundance estimates is less or equal 20, and $J = 8$ when there are more than 20 abundance estimates. However, in the tuning of CLA (Appendix A), $J = 10$ always.