Adapting benchmarks of biological status for persistent changes in productivity and variability in exploitation history with a focus on data-limited populations (Conservation Units) of Pacific salmon in north and central BC

Draft Final Report

Carrie Holt[1], Brooke Davis[1], Steve Cox-Rogers[2]

[1]Pacific Biological Station, DFO, Nanaimo, BC; [2]DFO, Prince Rupert, BC

Abstract

Canada's Wild Salmon Policy requires the biological assessment of conservation units (CUs) of Pacific salmon to ensure their conservation for future generations. A "stop light" approach has been adopted that uses stock-recruit information and benchmarks to assign green, amber, or red status to CUs. Data limitations for many CUs require the exploration of revised benchmarks to ensure conservation objectives are achieved when stock-recruitment data are not available. In this study we compared the performance of revised benchmarks that consider variation in productivity and exploitation rates using prospective simulation modelling and retrospective analyses of empirical data for sockeye salmon in northern and central BC. In retrospective analyses, we found that benchmarks based on percentiles of escapement time series were generally more precautionary than previously adopted stock-recruit based benchmarks. The simulation study yielded similar results. In most cases percentile benchmarks were a precautionary choice to reach conservation objectives. However, when CU productivity was low and harvest rates high, neither benchmark type was precautionary and percentile-based benchmarks were especially risky. Data truncation methods that adjust benchmarks to current productivity regimes were variably effective depending on the benchmark type and current population dynamics, and should therefore be considered carefully on a case-by-case basis.

Introduction

The aim of the Pacific Salmon Treaty is to prevent overfishing and provide for optimum production of both Canadian and US salmon, which requires information on the biological status of component Conservation Units (CUs). The first goal of the Northern Fund is to improve stock assessment, for example, by developing assessment methods that can be applied across populations where data are limited. Data limitations are an ongoing issue for salmon population assessments in both northern BC and southeast Alaska. In particular, assessments that do not account for exploitation history and persistent changes in productivity may be associated with increased risk of population depletion when exploitation has been high and/or productivity is low. Alternatively, these methods may be too precautionary when exploitation has been low and/or productivity is high. Although assessment methods that account for persistent changes in productivity have been developed for data-rich CUs such as Fraser River sockeye (e.g.Grant et al. 2011), similar methods have not been developed for data-limited CUs.

To address this gap, our original research objectives were to (1) identify candidate revised benchmark(s) for data-limited CUs (having time-series on relative abundances of spawners only) that remain effective under persistent changes in productivity and/or variability in exploitation rate history, and (2) compare performance of revised benchmarks that account for changes in productivity and variability in exploitation rate history for data-limited CUs against benchmarks derived from data-intensive methods using prospective, simulation modelling, and retrospective analyses of empirical data.

However, the methods proposed to address these objectives include comparisons of benchmarks derived for data-limited with data-rich scenarios, and require assumptions about data quality. Given pervasive observation errors, inconsistent time-series, and uncertainties in assessments, estimates of benchmarks differ, sometimes greatly, from underlying "true" values. Previous evaluations of the relative performance of benchmarks have not fully accounted for these data uncertainties. Therefore, we revised the first objective to evaluate benchmarks given uncertainties in underlying data. In addition, we focused our evaluation of methods that account for changes in productivity to a data truncation approach (objective 2). This approach uses subsets of time-series data to estimate benchmarks when productivity or capacity has changed over time.

Revised objectives:

(1) Identify relative performance of benchmarks for data-limited CUs that use either spawner-recruitment data or time-series of spawner abundances alone (percentile benchmarks) by assessing how well they track benchmarks derived from the "true" underlying stock-recruitment parameters assuming perfect knowledge. Performance was evaluated in (a) retrospective analyses of empirical data and (b) a simulation model of a hypothetical CU.
(2) Use simulation analyses to compare percentile benchmarks that have been revised to account for changes in capacity and/or productivity against benchmarks derived from data-intensive methods, where revised benchmarks are calculated by limiting time-series to either historical high-production (productivity) or current low-production (productivity) regime.

Canada's Wild Salmon Policy and biological benchmarks

Canada's Wild Salmon Policy (2005) outlines strategies to ensure the conservation of wild Pacific Salmon for future generations. The policy requires the biological assessment of CUs into one of three status zones: green, amber and red. The lower benchmark, delineating red and amber zones, is to be established at a level ensuring the CUs is buffered from being considered at risk of extinction under COSEWIC, the Committee on the Status of Endangered Wildlife in Canada, taking into account data uncertainties and harvest management. The upper benchmark, delineating amber and green zones is the escapement level associated with the maximum average annual catch, under current environmental conditions. While this policy lays out a basic framework for the assessment of conservation status of CUs, it does not require a single set of benchmarks for all CUs. Rather, it states that benchmarks will be determined on a "case-by case basis, and depend on available information and the risk tolerance applied" (DFO 2005).

For populations with time-series of stock-recruit data, benchmarks were identified by Holt et al. (2009) to be robust to uncertainties in underlying stock productivity. These benchmarks are based on the Ricker stock-recruit relationship, which is widely used for Pacific salmon populations (Ricker 1975). The lower benchmark, S_{gen}, is the number of spawners required to rebound to S_{MSY} within one generation under equilibrium conditions, in the absence of fishing. The upper benchmark is 80% of S_{MSY}, the number of spawners required to achieve maximum sustainable yield (MSY). For those CUs with limited or uncertain stock-recruit data, alternative benchmarks are being developed. Percentile-based approaches have been proposed for determining sustainable escapement goals (SEGs) or conservation benchmarks under Canada's Wild Salmon Policy (Clark et al. 2014, Holt and Folkes 2015). These methods require escapement data only, and simply compare current escapement levels with the percentiles of historical observations. The Alaska Department of Fish and Game (ADF&G) compared various percentiles as a basis for Salmon Escapement Goals, SEGs (intended to be equivalent to S_{MSY}) in a simulation evaluation and retrospective analysis (Clark et al. 2014). Based on this study, a multi-tiered system was recommended where percentile values for SEGs are chosen based on data contrast, data uncertainty, and harvest rates. In particular, Clark et al. (2015) recommended that percentile-based SEGs not be used when harvest rates are high (> 40%), or spawner and recruitment data show little contrast over time and measurement error is high. Our evaluation differs from Clark et al. (2015) in that we evaluated the extent to which percentile benchmarks were consistent with biological benchmarks already identified under the Wild Salmon Policy, instead of applying (and evaluating) them as escapement goals for management at MSY levels. For data-limited chum salmon CUs in southern BC, percentile benchmarks at the 25th and 75th percentiles have been proposed and provisionally implemented as lower and upper benchmarks, respectively (Hilborn et al. 2013). These percentiles are higher, and therefore more conservative, or precautionary, than SEGs proposed by the ADF&G (Clark et al. 2014).

We first provide Methods and Results for Objective 1(a), evaluating data-limited benchmarks using retrospective analyses, followed by Objective 1(b), evaluating data-limited benchmarks in simulation. We then describe Methods and Result for Objective 2, evaluating benchmarks that use truncated time-series data to account for changes in productivity, and provide a synthesized Discussion for both objectives.

Objective 1a. Evaluating benchmarks using retrospective analyses

In retrospective analyses, we compared biological status of 49 CUs of north and central coast sockeye salmon using data-rich and data-limited benchmarks derived from historical time-series. Escapement data were available for these CUs for varying time intervals, ranging from 1954-2014. Reconstructed recruitment data range from (brood years) 1954-2009 (Table 1). For some CUs, available data were continuous, for other there were missing data within the time series. Run reconstructions used aggregate catch data and migration timing and patterns of fish from specific CUs through different fisheries to estimate the number returning to each CU. However, CU-specific return estimates are very sensitive to uncertainty in migration timing and patterns (see Korman et al. 2013). Also, fitting the Ricker model to uncertain data can lead to biased parameter estimates because of observation errors-in-variables and time-series biases (Walters and Martell 2004). Long time-series across highly variable spawning stocks can help to reduce these biases. However, caution in the interpretation of results is warranted. These results should be considered with those from the simulation model that incorporates numerous sources of data uncertainties (Objective 1b).

For the data-rich scenario, we compared benchmarks derived from two different forms of the Ricker model, the standard model which estimates parameters independently for each CU, and a hierarchical model where CUs are assumed to have similar productivities. Hierarchical models may reduce uncertainties and biases in parameter estimation mentioned above by sharing information on productivity across CUs. CU-specific productivities were drawn from a global hyper-distribution given evidence for spatial covariation in productivity among populations within regions (Pyper et al. 2002). Following the results of Holt and Folkes (2015) who investigated the impacts of temporal changes in productivity, we further identified temporal trends in productivity over time in these CUs using a recursive Bayes modelling approach.

Methods

Spawner and recruitment data were available for 55 CUs of sockeye salmon, but we limited our analysis to those with >10 years of spawner and recruitment data, which are not necessarily contiguous (50 CUs, n=12-56 years). Estimated returns by brood year were prepared by English (2013) and English et al. (2012) with updated escapements until 2014 for most sites (Cox-Rogers, S. pers. comm.).

Estimates of freshwater capacity were available for 11 CUs, which are used to provide improved estimates of S_{max}, the spawner abundances at which maximum recruits are produced. We omitted available capacity information for two CUs (Morice and Bear) because those data did not adequately capture rearing capacity in those systems. Freshwater capacity estimates are derived from photosynthetic rates in nursery lakes as a proxy for food availability for juveniles (Cox-Rogers et al. 2010, Shortreed et al. 1999). This method assumes that the main factor limiting recruitment is freshwater rearing capacity.

We first identified benchmarks and assessed status in the most recent year using all available data. Retrospective analyses were then carried out by sequentially calculating benchmarks and assessing status using all available data up to a given year. For both percentile and stock-recruitment based benchmarks, we assumed that 10 years of data were required to estimate the first benchmark, and benchmarks were re-estimated every year for which new data were available. Since recruitment information was required for stock-recruitment based benchmarks, and recruitment from a given brood year cannot be calculated until the oldest age class fishery (or 95% of the age distribution) has recruited to the fishery. Ricker benchmarks lag behind percentile-based benchmarks by 5 or 6 years, depending on the highest age class needing to return for 95% of the age distribution to be accounted for (Table 1). Benchmarks were compared to generational mean escapement to determine status. Generational mean escapement was estimated as the four-year running geometric average. When there were gaps in recruitment data but escapement data were available, the most recent benchmarks were compared to that year's escapement value to estimate status.

Standard Ricker Model

For each year with sufficient data, a Ricker model (Eqn. 1) was fit in a Bayesian context using Markov Chain Monte Carlo (MCMC) methods.

(1) $R = \alpha S e^{-\beta S}$

R is the abundance of adult recruits from a given spawning event and *S* is the number of spawners that generated those recruits (also referred to as escapement). The parameter α is productivity, the recruits per spawner at low spawner abundances, and β is the reciprocal of S_{max}. We linearized the equation and incorporated normally distributed process error:

(2) $R = \log(\alpha) + \log(S) - \beta S + \nu,\ \nu \sim normal(0,\ \tau_\nu)$,

where τ_ν is the precision in process error and precision is the reciprocal of variance. We applied a weakly informative prior on α to ensure values greater than zero and within the bounds of observed productivity values for sockeye salmon (Dorner et al. 2008, Fig. 1).

(3) $\log(\alpha) \sim normal(1,1)$

The prior for β was set indirectly by applying a prior on its reciprocal, S_{max}. We applied an informative lognormal prior for those sites by converting rearing capacity estimates to spawner capacities, S_{PR} (Cox-Rogers et al. 2010) (Eqn. 4a). The mean of the informative, lognormal prior on S_{max} was set to S_{PR}. When prior information was not available, the mean of the lognormal prior was set to the mean observed escapement with a variance that was wide enough to be uninformative over the range from zero to twice the maximum spawner abundances (Eqn 4b).

(4a) $S_{max} \sim lognormal(S_{PR},\ \tau_S),\ \ \tau_S = 1/\log(CV^2 + 1)$, and
(4b) $S_{max} \sim lognormal(log(mean(S_{obs}))\,,\ \tau_S),\ \ \tau_S = 1/\log(CV^2 + 1)$,

τ_S is the variance of the log-normal prior, calculated using a standard transformation of the coefficient of variation, CV, in normal space to log-normal space. We set the CV of the informative lognormal prior to 0.3 based on Korman et al. (2013), and for the uninformative prior to 5, which covered the range of observed escapement values.

In a sensitivity analysis for CUs without S_{PR} estimates, we fit models using uninformative uniform instead of log-normal priors (Eqn. 4c). We set the lower bound of the uniform distribution to zero and upper bound to twice the maximum observed spawner value, given the assumption that CUs had not been depleted below half of S_{max} over the historical record.

(4c) $S_{max} \sim uniform(1, max(S_{obs}) * 2)$

Uninformative gamma priors were used for the τ parameters,

(5) $\tau_{v,}\tau_{S} \sim gamma(0.01, 0.001)$.

Hierarchical Ricker Model

We also estimated Ricker parameters using a hierarchical version of the standard Ricker model (Eqns. 1 and 2), where parameters for each CU, i, were estimated simultaneously. CUs were grouped into three categories: coastal, interior, and high interior, based on Freshwater Adaptive Zones (FAZ) described by Holtby and Ciruna (2007) (Table 1). Parameters from each CU within a group were estimated simultaneously, where CU-specific α_i values are drawn from a common, normal distribution:

(6a) $R = \alpha_i S e^{-\beta_i S}, v \sim normal(0, \tau_v)$,

(6b) $\alpha_i \sim normal(\mu_\alpha, \tau_\alpha)$,

where parameter μ_α is the mean of the normal distribution of a given group and τ_α is precision. This model assumes that productivities are similar among CUs within a group, a reasonable assumption given their geographic proximity and similar biogeoclimatic conditions.

The same prior distributions were used as for the standard Ricker model (Eqn. 4), with the addition of a prior on the global mean and variance of alpha, μ_α:

(7) $\log(\mu_\alpha) \sim normal(1,1)$

Additionally, to impose an uninformative prior on τ_α we put an more intuitive uninformative prior on variance, σ_α, where $\sigma_\alpha = 1/\tau_\alpha$:

(8) $\sigma_\alpha \sim Uniform(0, 100)$.

Models were fit using MCMC using JAGS (Plummer 2003), interfaced through R version 3.2.0 (R Development Core Team 2016) using package "R2jags" (Su and Yajima 2012). Model convergence was assessed using Gelman-Rubin statistics and visual inspection of trace plots.

Benchmarks

For stock-recruitment based benchmarks, the lower benchmark, S_{gen}, was calculated numerically, according to the following equation (Holt 2009):

(9) $S_{MSY} = S_{gen}\, \alpha\, e^{-\beta S_{gen}}$

The upper benchmark was calculated using an approximation developed by Hilborn and Walters (1992):

(10) $0.8\,S_{MSY} = \; 0.8\frac{\log(\alpha)}{\beta}(0.5 - 0.07\log(\alpha))$

Percentile-based benchmarks were calculated as the 25th and 75th percentile of observed spawner abundances ranked from lowest to highest, for the lower and upper benchmarks respectively (S_{25th}, S_{75th}).

Changes in productivity

To identify changes in productivity over time for north and central coast sockeye CUs and assess how these changes affect benchmark performance, we fit a recursive Bayes model to stock-recruitment data which allowed α to vary over time for each CU (Malick and Cox 2016). We fit this model using all available data for each site. It follows the basic Ricker form with a time-varying α parameter,

(11) $R = \alpha_t S e^{-\beta S} e^{v}, v \sim normal(0, \tau_v)$,

where α_i is productivity in brood year t. The model assumes that α changes over time following a Gaussian random walk:

(12) $\log(\alpha_t) = \log(\alpha_{t-1}) + w,\; w \sim normal\big(0,\, \tau_w\big)$

The same prior distributions are applied as for the standard Ricker model (equation 5), with the addition of a normally distributed prior on α in year 1, and a uniform prior on the variance associated with the Gaussian random walk, σ_w, where $\sigma_w = 1/\tau_w$:

(13a) $\log(\alpha_1) \sim normal(1,1)$, and

(13b) $\sigma_w \sim Uniform(0, 100)$.

Results

Effects of priors on parameter estimates

In retrospective analyses of the standard Ricker model, varying the prior distribution on $S_{\max}$ resulted in benchmark estimates that generally fell within the range of uncertainty for a given parameter. For 98% of CU-year combinations, the estimated benchmarks from log-normal and uniform priors fell within the 95% credible bounds of each other. Where the estimates did not fall within those bounds, the lognormal prior resulted in benchmarks that were significantly higher than those estimated using the uniform prior, which tended to occur when log-normal priors on $S_{\max}$ were informative, i.e., capacity information was available. These discrepancies occurred early in the retrospective analyses where few years of data were available and parameters were highly uncertain. Since these discrepancies were rare and did not persist throughout the whole time series, results for the lognormal prior model are presented here, which takes advantage of available capacity data.

Current benchmarks and status

Lower percentile benchmarks (S_{25th}) tended to be similar in value to lower Ricker-based benchmarks (S_{gen}), whereas upper percentile benchmarks (S_{75th}) were generally much higher than the Ricker-based upper benchmarks (80% S_{MSY}, Fig. 2). The posterior densities of the upper and lower benchmarks overlapped and, in some cases, were nearly indistinguishable, e.g., Owikeno (Fig. 2al). Stock-recruitment benchmarks varied between the standard and hierarchical Ricker models (comparing Fig. 2(i) vs .(ii) for each CU, a-aw), but these differences were small compared with large uncertainties in benchmark estimates. As expected, the uncertainty in benchmarks was reduced for the hierarchical Ricker model compared with standard Ricker for most CUs. However, for some CUs, benchmark uncertainty increased for the hierarchical vs. the standard Ricker model (e.g., Kainet Creek, Fig. 2s). In these cases, CU-specific productivities estimated from the standard Ricker model differed significantly from neighbouring CUs in their FAZ grouping. Therefore, the additional information provided in the hierarchical model increased the uncertainty in productivity compared with the standard model. All benchmark values are provided in the Appendix, Tables A1 and A2.

Retrospective analyses

In retrospective analyses, percentile benchmarks tended to vary over time more than Ricker-based benchmarks, but overall, percentile benchmarks were more precautionary (Fig. 3, Tables 2-4). Percentile-based benchmarks were the same or higher than Ricker-based benchmarks for 95% of CU-year combinations. The few exceptions where percentile-based benchmarks were lower than Ricker-based benchmarks were associated with either long periods of low escapement, long-term declines in percentile benchmarks, and relatively consistent Ricker-based benchmark (e.g., Kitlope and Mary Cove, Fig.3u & ad), or short time-series of fragmented data, percentile benchmarks that were sensitive to occasional low escapement values, and relatively constant Ricker-based benchmarks (e.g., Motase and Port John, Fig.3aj & am).

In addition, percentile benchmarks were found to provide the same status, or more precautionary status compared to Ricker-based benchmarks for most CU-year combinations (Fig. 4, Tables 2-4). Specifically, Figure 4 shows time-varying statuses derived from each set of lower and upper benchmarks. Uncertainty in status is shown by comparing escapement to the lower and upper credible intervals of each benchmark, for Ricker-based benchmarks only (Fig. 4, shaded regions above and below coloured bars). Large uncertainties in Ricker-based benchmarks were associated with variability in status derived from the best-estimate and lower and upper credible intervals. For example, for the CU, Alistar, in 1978, status using the standard Ricker model was amber, but was green or red when using the lower or upper credible intervals for those benchmarks, respectively (Fig. 4a).

Time-varying productivity

Time-varying productivities estimated using a recursive Bayes model for each CU, showed unique patterns among CUs (Fig. 5). For many CUs, there was considerable uncertainty in productivity over time, indicated by wide error bounds compared with the magnitude of inter-annual variability. A thorough analysis of patterns and possible synchrony in trends among CUs was beyond the scope of this study. However, the decline observed for numerous CUs warrants further investigation (e.g., Babine Early Wild, Fig. 5d). Using long-term data to estimate stock-recruitment parameters and benchmarks that spans decades where productivity has changed may lead to poor Ricker-model fits, large uncertainty in parameters and benchmarks, and possible biases.

Objective 1b. Evaluating benchmarks using simulation analyses

Methods

We adapted the simulation model of Holt and Folkes (2015) to evaluate data-limited benchmarks. As in Holt and Folkes (2015), the model included five components representing population dynamics, observations of abundances, management (including the derivation of benchmarks), harvest, and performance evaluation (Fig. 6). In particular, the model included natural variability in adult recruitment based on a Ricker spawner-recruitment relationship with variable age at maturity, errors in observations of abundances, assessments of biological status relative to benchmarks, the application of a harvest control rule, and uncertainties in the outcomes from implementing management decisions. See Holt and Folkes (2015) for model equations.

Our model differed from that of Holt and Folkes (2015) in 9 ways:

- The population dynamics sub-model was parameterized for sockeye salmon instead of chum salmon (i.e., stock-recruitment model, age-at-maturity, and stray rates)
- The population dynamics sub-model included covariance in Ricker residuals among subpopulations within a CU, instead of assuming sub-populations varied independently.
- The observation sub-model was more realistic in that catches (or, alternately exploitation rates) were observed with observation errors and recruitments by brood year were then calculated using estimated ages-at-maturity, instead of applying observation error directly to “true” recruits by brood year. Estimates of ages-at-maturity used to calculate recruitment by brood year were constant over time, estimated from mean observed ages-at-maturity over an initialization period of 13 years.
- In the observation sub-model, we evaluated scenarios where spawner abundances were observed with a consistent negative (or positive) bias that was not corrected for in the assessment.
- In the assessment sub-model, we evaluated scenarios where only a portion of subpopulations were sampled within a CU and a constant expansion factor was applied to derive escapement estimates for the entire CU. The expansion factor was estimated from observed complete sampling in a 3-year initialization period.
- The assessment sub-model focused on percentile benchmarks and stock-recruitment benchmarks applied under the Wild Salmon Policy (S_{gen} and 80% of S_{MSY} for the lower and upper benchmarks, respectively)
- In the harvest sub-model, a constant low harvest rate (20%) was applied instead of a harvest control rule with limit and/or target reference points. The harvest rates applied over the initialization period were varied in a sensitivity analysis (but remained constant over that period) to reflect the varied harvest rate histories observed among CUs.
- In the performance module, benchmarks were evaluated based on the deviations between benchmark estimates and the “true” underlying values. In the case of percentile benchmarks, we evaluated deviations between 25th percentile (lower benchmark) and the “true” S_{gen} value, and between the 75th percentile (upper benchmark) and the “true” 80% of S_{MSY} value. Specifically, we evaluated mean percent error and mean raw error because we were interested in the direction of bias (i.e., if the estimated benchmark was above or below the “true” benchmark). We focused our results on mean percent error, MPE, as this

metric is scale independent, making comparisons in sensitivity across benchmarks more intuitive. Results for mean raw error are provided in the Appendix.

- The model was run over 50 years, instead of 100 to provide a more realistic time-series length for estimating benchmarks. The model was run over 5000 MC trials, the number of trials required to stabilize output metrics at (standard error <=3% in performance metrics). The model was initialized for 20 years after a 5-year pre-initialization period necessary to generate the first recruitment by brood year.

Parameterization

The population dynamics sub-model was parameterized based on previous empirical studies in the primary literature and governmental reports on sockeye salmon. See Holt and Folkes (2015; Appendix) for model equations. The productivity parameter of the spawner-recruitment relationship, a (defined as $\log_e$(recruits/spawner) at low spawner abundances, and referred to simply as productivity here) and the range considered in sensitivity analyses (Table 4, see below for more details) were chosen to bound productivities commonly observed for sockeye salmon stocks in BC, 0.5-2.0 (Holt and Bradford 2011, Korman et al. 2013). Productivity and spawner abundances at equilibrium abundance, S_{eq} (set at 10 000 fish) were assumed to be equal among subpopulations.

We assumed an autocorrelation coefficient of 0.6, based on coefficients estimated for 8 Bristol Bay, Alaska stocks of sockeye salmon (ranging from 0.17-0.72; Peterman et al. 2003) and 30 sockeye salmon stocks in BC and Alaska (ranging fron -0.41-0.84; Korman Josh et al. 1995), and considered a range of plausible autocorrelation coefficients (0 and 0.9) in sensitivity analyses (Table 4). The standard deviation in recruitment residuals (in log-space) was set to 0.75, within the range of values estimated from the same Bristol Bay stocks (0.39-0.76; Peterman et al. 2013), and within the range estimated for sockeye salmon in BC and Alaska (0.49-1.63; Korman et al. 1995). The average proportions of mature adults at ages 3, 4, and 5 were estimated for the Babine Early Wild sockeye salmon CU, 1960-2005 (0.08, 0.42, and 0.50, respectively, (S. Cox-Rogers pers. comm. 2016), as an example age distribution for sockeye. The variance in the proportions of ages at maturity was estimated from empirical time-series data for proportions of age-at-maturity for Babine Early Wild sockeye salmon (S. Cox-Rogers pers. comm. 2016). The probability of straying among adult recruits was assumed to be relatively low (1%) because of weak evidence for dispersal among natal spawning sites (Quinn 1993).

In the observation sub-model, we assumed the standard deviation in estimates of spawner abundance around the true values,(observation errors) was equal to 0.5 (in log-space), which corresponds to an upper estimate of the uncertainty in spawner abundances derived from various visual surveys of Pacific salmon (Cousens 1982, Szerlong and Rundio 2008). On the north and central coast of BC, sockeye salmon abundance is largely estimated from visual surveys, which typically produces relatively imprecise estimates. We also considered a lower estimate of observation errors (0.2) in a sensitivity analysis. In the absence of quantitative estimates of uncertainty in total exploitation rate estimates, we assumed a standard deviation in normally distributed errors of 0.03 based on deviations between two independent time-series of exploitation rates (English et al. 2013), with a sensitivity analysis using an upper estimate of that standard deviation (0.1).

The standard deviation of outcome uncertainty was estimated at 0.3, an upper estimate of outcome uncertainty estimated by Holt Carrie A and Peterman (2006) for Fraser River sockeye salmon. Because the standard deviation of outcome uncertainty is not widely estimated for Pacific salmon, and likely varies widely among stocks and management approaches, we also considered an upper value of 0.5 in a sensitivity analysis.

Sensitivity analyses

To assess the strength and direction of effects of input parameters on benchmark performance (measured as deviations between estimated lower benchmarks and "true" lower benchmarks) we performed a sensitivity analysis where input parameters were varied individually while all others were held constant. These sensitivity analyses did not assess the sensitivity of performance to interactions among input variables. For two parameters that had a relatively large effect on performance: productivity and initial harvest rates, a bivariate sensitivity analysis was performed to assess their combined effect on benchmark performance. To further consider interactions among all input variables, we performed a global sensitivity analysis using the Morris method. Similar to univariate analyses, the Morris method varies each input parameter one at time, but in contrast to univariate analyses, this is done at different points of the factor input space (i.e., at different combinations of other variables) (Morris et al. 2014). The mean elemental effect of an input parameter from the Morris method is an index of the sensitivity of benchmark performance to uncertainty in that parameter. The standard deviation of the elemental effects is an index of sensitivity of benchmark performance to interactions of that variable with other variables. The Morris method was run using the R package, sensitivity v.1.11.1. We focused univariate and global sensitivity analyses on lower benchmarks (25th percentile and S_{gen}), but also considered sensitivity of upper benchmarks (75th percentile and 80% S_{MSY}) in our bivariate sensitivity analysis.

Results

Simulation model outputs for an example CU are presented in Fig. 7. Mean percent error between estimated and "true" benchmarks was generally greater than zero, especially for percentile benchmarks (Fig 7, right panel). Percentile benchmarks tended to be more precautionary than stock-recruitment based benchmarks (i.e., positive deviations were greater), but both were precautionary. Estimates of stock-recruitment benchmarks differed from the "true" values because estimates were based on observed data (black line in Fig 7, left panel, and solid dots in Fig. 7, middle panel) rather than "true" data (grey line in Fig 7, left panel, and hollow dots in Fig. 7, middle panel). The assessed stock-recruitment model (black curve, Fig. 7, middle panel) differed from the "true" underlying model (grey curve, Fig. 7, middle panel), due to those errors in spawner abundances and time-series biases (Walters and Martell 2004).

Overall, performance of lower benchmarks (both S_{gen} and S_{25th}) was more sensitive to uncertainty in productivity than to other input parameters (Fig. 8 a & b, respectively). Low productivity values (leftmost black bar) were associated with negative deviations from base case (i.e., benchmarks that were less precautionary than the base case); high productivities (leftmost white bar) were associated with positive deviations (i.e., benchmarks that were less precautionary).For the lower benchmark, S_{gen}, Ricker autocorrelation had moderate impacts on performance and the remaining input parameters had relatively weak effects on performance (<30%). For the lower percentile benchmark, S_{25th}, initial harvest rates had a moderate effect on benchmark performance, and the remaining input parameters had relatively weak effects on performance (<50%). Similar patterns of results were found for mean raw error of estimated benchmark from the "true" value (Appendix, Fig. A1).

Our model assumed spawner abundances at equilibrium, S_{eq}, remained constant as productivity varied in sensitivity analyses (as in Holt and Bradford 2011). When we considered an alternate assumption where S_{max} remained constant, but S_{eq} declined with productivity, we found similar results (within 5%). This alternate assumption represents a scenario of simultaneous declines capacity and productivity.

We further explored bivariate sensitivity analyses of the effects of variability in productivity and initial harvest rates on benchmark MPEs. At moderate to high productivity and low initial harvest rates, both S_{25th} and S_{gen} benchmarks were precautionary (i.e., benchmarks were equal to or higher than "true" S_{gen} lower benchmark) (Fig. 9, top left portions of panels). At low productivity and high initial harvest rates (Fig. 9, bottom right corner) neither benchmark was precautionary, but S_{gen} performed slightly better (i.e., was slightly closer to "true" value than S_{25th}).

When we superimposed CU productivities and harvest rates for north and central coast sockeye, the associated mean percent error for estimated lower benchmarks were greater than zero for all CUs except one, for the benchmark S_{gen} (Fig. 8, symbols lie above zero contour line). For upper benchmarks, the estimated values were always greater than the true values, and these positive deviations were greatest for the S_{75th} benchmark (Fig. 10). The curvilinear relationship between the estimated 80% S_{MSY} benchmark and initial harvest rates and productivity were due to the covariance between productivity and carrying capacity when estimating the stock-recruitment relationship and the resulting confounding effect on estimates of S_{MSY}.

The global sensitivity analyses showed similar patterns to univariate and bivariate sensitivity analyses. The mean elemental effects (magnitude of sensitivity) were greatest for productivity for both S_{25th} and S_{gen} benchmarks (Fig. 11). Initial harvest rates were secondarily important for S_{25th} benchmark (Fig. 11b), and the Ricker autocorrelation coefficient and observation errors were also important for S_{gen}. The Ricker autocorrelation coefficient was especially important in combination with other input variables, resulting in relatively high standard deviation in elemental effects.

Objective 2. Adapting benchmarks to changes in productivity

For Objective 2, we simulated an abrupt decline in productivity in the population dynamics sub-model, and evaluated benchmarks that used truncated data sets that represented either historical baseline conditions (high productivity period) or recent conditions (low productivity period). This data-truncation approach was adopted by Grant et al. (2011) to evaluate biological status of sockeye salmon on the Fraser River using biological benchmarks under the Wild Salmon Policy. Specifically, Grant et al. (2011) found that benchmarks estimated using spawner and recruitment data from only the recent low-productivity period were larger (i.e., more precautionary) than those that were estimated from the entire time-series. This approach has not been applied to other species or regions, where the data required for assessing changes in productivity are often lacking.

Methods

We adapted the simulation model from Objective 1 in 4 ways:

- The population dynamics sub-model included time-varying productivity modelled as an abrupt decline from $\log_e$(recruits/spawner)= 2 to 1 at year 35. The magnitude of this change in productivity was chosen to reflect observed changes in the productivity for sockeye salmon (Dorner et al. 2008). Evaluating the effects of different temporal different patterns in productivity was outside the scope of this study. We considered both the scenario where S_{eq} remains constant as productivity changes, and where S_{max} remains constant, and S_{eq} declines with productivity. The latter reflects simultaneous changes in capacity and productivity.
- The population model was run over 70 years (35 years prior to and after the productivity shift).
- In the assessment sub-model, benchmark estimation occurred in the final year of the simulation only.
- Assessments used either the entire time-series, the first 30 years (base line, high-productivity period), or the final 30 years (recent, low-productivity period) to estimate benchmarks.

Results

As expected, regime shifts from high to low productivity were associated with increases in "true" (i.e., deterministic) S_{gen}, 80% of S_{MSY}, and S_{max} when S_{eq} was assumed constant (Fig. 12a-e). In contrast, when S_{max} was assumed constant, declines in productivity were associated with declines in "true" 80% of S_{MSY}, and increases in "true" S_{gen} and S_{eq} (Fig. 12f-j). The latter assumption incorporates a decline in total capacity of the CU to sustain a population, as well as decline in recruits/spawner at low spawner abundances (see Fig. 3 of Holt and Folkes 2015).

Truncating time-series data used to estimate benchmarks to the recent low-productivity period resulted in lower estimates of productivity (Fig. 12f) and higher estimates of S_{gen} (i.e., more precautionary) (Fig. 13c) under constant S_{eq}, as expected from the previous deterministic analyses. The opposite was true when the historical period was used (higher productivity and lower S_{gen} estimates). Median estimates of S_{gen} were consistently below the "true" value (dashed line Fig. 13c), though the confidence intervals covered the “true” value in all three scenarios. The upper benchmark, 80% of S_{MSY}, did not change consistently with data truncation (Fig. 13e). Both S_{25th} and S_{75th} percentile benchmarks declined (i.e., became less precautionary) when data were truncated to the recent period, but these values were consistently higher than the “true” S_{gen} and 80% of S_{MSY} benchmarks, respectively. Although S_{gen} became more precautionary as data were truncated to the recent period and percentile benchmarks became less precautionary, percentile benchmarks were still consistently greater than true values (mean percent errors were >> zero) (Fig. 13a).

Similar patterns were observed under the assumption of constant $S_{max,}$ with an abrupt decline in productivity (Fig. 14b and 15), with two exceptions. When only the recent data were used, the estimate of productivity did not decline compared with using the entire data set (Fig. 15a), but the estimate of S_{eq} did decline (Fig. 15b). Confounding between estimates of productivity and S_{eq} resulted in a relatively low (instead of high) values for S_{gen} when only recent data were used (Fig. 15c).

Discussion and Conclusions

Our retrospective analysis of 49 sockeye salmon CUs indicates that 25th and 75th percentile benchmarks are generally more precautionary than stock-recruitment based benchmarks adopted under the Wild Salmon Policy, and are therefore a precautionary choice in data-limited situations. The few exceptions where percentile benchmarks were less precautionary than Ricker-based benchmarks did not occur in the most recent year (i.e., they occurred in retrospective assessments that used shorter time series), and were associated with either long-term, persistent declines in escapement, or highly fragmented and variable data. In particular, CUs with short time series of escapement data need to be approached with caution, as percentile benchmarks will change drastically with each new observation. Upper percentile benchmarks (S_{75th}) were considerably higher than upper stock-recruitment based benchmarks (>>80% S_{MSY}), and, by definition, green status occurred in only a quarter of observed years. If upper percentile benchmarks are used to inform fisheries management targets, these will likely lie above S_{MSY} levels, resulting in harvests below MSY.

For the 49 CUs analyzed here, benchmarks derived from hierarchical Ricker models were virtually indistinguishable from those estimated using standard Ricker models. In retrospective analyses, the standard Ricker model and hierarchical Ricker model gave the same status for 95% of CU-year combinations. However, benchmarks derived from the hierarchical Ricker model were generally less uncertain than those from the standard model. Given large uncertainties in the stock-recruitment data and inconsistent time-series, a hierarchical approach is recommended over standard Ricker model when there is support for similar productivities among CUs within regions.

Our retrospective analysis illustrates that most north and central coast CUs of sockeye salmon have shown considerable change in conservation status over time and these changes depend on the benchmarks used. Percentile-based benchmarks were associated with higher variability in status over time than stock-recruitment benchmarks. One caveat on the application of percentile-based benchmarks is that uncertainties in benchmarks are not provided. Nevertheless, status determined using the percentile benchmarks were consistently more precautionary than Ricker-based benchmarks, and are therefore a viable choice where Ricker-based benchmarks cannot be calculated or are highly uncertain.

Similarly, our simulation model suggests that the lower percentile benchmark, S_{25th}, tends to be more precautionary than the corresponding stock-recruitment based benchmark, S_{gen} when historical harvest rates are moderate to low, current harvest rates are low, and productivity is moderate to low. At low productivity and high exploitation rates, neither benchmark is precautionary, but the percentile benchmark is especially negatively biased. Although most CUs in our study had historical exploitation rates and productivities associated with relatively precautionary estimates of benchmarks, uncertainties in the remaining input parameters may affect benchmark performance. Specifically, S_{gen} performance was also sensitive to Ricker autocorrelation coefficient, Ricker sigma (residual variance), and observation errors in abundances.

Our retrospective analyses assumed constant productivity over time, but observed temporal variability in productivity for most sockeye CUs suggests that the Ricker stock-recruit relationships may vary through time (here and documented in Malick and Cox 2016). Indeed, there is widespread evidence for abrupt regime shifts in salmon productivity followed by relatively constant productivities (Beamish et al. 1999, Hare et al. 1999). Given widespread changes in productivity, a data-truncation approach that uses data from either the recent period or a historical baseline may result in benchmarks that are more precise and in some cases, more precautionary. Although truncation of time-series to the most recent period has been suggested as a method to account for declines in productivity in a precautionary manner (Grant et al. 2011), this may result in a shifting baseline for percentile benchmarks since they tend to decline with abundances. In other words, declines in productivity affect S_{gen} and percentile benchmarks in opposite directions.

Our simulation model suggests that truncating data to a recent low-productivity period results in S_{gen} values that are more precautionary and percentile benchmarks that are less precautionary than when the entire time-series is used, only under the scenario of constant S_{eq}. When S_{eq} also declines with productivity (simultaneous declines in production), then both S_{gen} and S_{25th} benchmarks become less precautionary, but this change is greatest for the S_{25th} benchmark. Despite these divergent trends with data truncation, percentile benchmarks remain more precautionary overall (with caveats from Objective 1 described above).

Any changes to benchmarks in response to changes in productivity (i.e., data truncation) will require careful consideration of strength of evidence, causal mechanisms, and reversibility of changes, among other factors, and should be implemented on a case-by-case basis, If the observed decline in abundances is due to density-independent changes in productivity that are reversible, then precautionary (i.e., relatively high) benchmarks are warranted to maintain resilience of the CU. If the observed decline is due to density-dependent changes in capacity that are well understood and irreversible (i.e., a persistent shift to a low-production regime), reductions in benchmarks may be warranted in rare cases to reflect this decline in production. Duplisea and Cadigan (2012) provide recommendations on the conditions that would be required to make such adjustments.

One limitation in our data truncation analysis is that it assumes that shifts in productivity can been detected accurately, and data are truncated to within a specific regime. The results of our Ricker model with time-varying productivity in Objective 1a demonstrate that temporal estimates of productivity are highly uncertainty, making regime shifts challenging to identify. Developing improved methods for detecting regime shifts within these data sets is an area for further investigation.

References

Beamish R, Noakes D, McFarlane G, Klyashtorin L, Ivanov V, Kurashov V. 1999. The regime concept and natural trends in the production of Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences 56:516-526.

Clark RA, Eggers DM, Munro AR, Fleischman SJ, Bue BG, Hasbrouck JJ. 2014. An Evaluation of the Percentile Approach for Establishing Sustainable Escapement Goals in lieu of Stock Productivity Information. Fishery Manuscript No. 14-06, Alaska Department of Fish and Game.

Cousens NBF. 1982. A review of salmon escapement estimation techniques. Can. Tech. Rep. Fish. Aquat. Sci. 1982/1108.

Cox-Rogers S, Hume JMB, Shortreed KS, Spilsted B. 2010. A Risk Assessment Model for Skeena River Sockeye Salmon. Can. Manuscr. Rep. Fish. Aquat. Sci. 2010/2920.

DFO (Fisheries and Oceans Canada). 2005. Canada's Policy For Conservation of Wild Pacific Salmon. Cat. No. Fs23-476/2005E, DFO (Fisheries and Oceans Canada).

Dorner B, Peterman RM, Haeseker SL. 2008. Historical trends in productivity of 120 Pacific pink, chum, and sockeye salmon stocks reconstructed by using a Kalman filter. Canadian Journal of Fisheries and Aquatic Science 65: 1842–1866.

Duplisea D, Cadigan N. 2012. Proceedings of the national workshop for technical expertise in stock assessment (TESA): maximum sustainable yield (MSY) reference points and the precautionary approach when productivity varies. Can. Sci. Adv. Sec. Proc. Ser. 2012/055.

English K. 2013. Extend the time-series of catch and escapement estimates for Skeena sockeye, pink, chum and coho salmon stocks. Report prepared by LGL limited for the Pacific Salmon Foundation.

English KK, Mochizuk T, Robichaud D. 2012. Review of North and Central Coast Salmon Indicator Streams and Estimating Escapement, Catch and Run Size for each Salmon Conservation Unit. Prepared by LGL Limited Environmental Research Associates for Pacific Salmon Foundation, and Fisheries and Oceans Canada.

Grant SCH, MacDonald BL, Cone TE, Holt CA, Cass A, Porszt EJ, Hume JMB, Pon LB. 2011. Evaluation of Uncertainty in Fraser Sockeye (*Oncorhynchus nerka*) Wild Salmon Policy Status using Abundance and Trends in Abundance Metrics. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/087:viii + 183 p.

Hare SR, Mantua NJ, Francis RC. 1999. Inverse production regimes: Alaska and west coast Pacific salmon. Fisheries 24:6-14.

Hilborn R, Schmidt D, English K, Devitt S. 2013. British Columbia Chum Salmon (*Oncorhynchus keta*) Fisheries; British Columbia Coastal and Adjacent Canadian Pacific EEZ Waters. Ref: (82506), Intertek Moody marine.

Hilborn R, Walters CJE. 1992. Quantitative Fisheries Stock Assessment; Choice Dynamics & Uncertainty. Springer US.

Holt C, Cass A, Holtby B, Riddell B. 2009. Indicators of status and benchmarks for conservation units in Canada's Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/058.

Holt CA, Bradford MJ. 2011. Evaluating benchmarks of population status for Pacific salmon. North American Journal of Fisheries Management 31:363-378.

Holt CA, Folkes MJP. 2015. Cautions on using percentile-based benchmarks of status for data-limited populations of Pacific salmon under persistent trends in productivity and uncertain outcomes from harvest management. Fisheries Research 171:188-200.

Holt CA, Peterman RM. 2006. Missing the target: uncertainties in achieving management goals in fisheries on Fraser River, British Columbia, sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences 63:2722-2733.

Holtby LB, Ciruna KA. 2007. Conservation Units for Pacific Salmon under the Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Res. Doc. 2007/070.

Korman J, Cox-Rogers S, English K. 2013. Benchmark Analysis for Pacific Salmon Conservation Units in the Skeena Watershed. Pacific Salmon Foundation.

Korman J, Peterman RM, Walters CJ. 1995. Empirical and theoretical analyses of correction of time-series bias in stock-recruitment relationships of sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences 52:2174-2189.

Malick MJ, Cox SP. 2016. Regional-Scale Declines in productivity of pink and chum salmon stocks in western North America. PLoS One 11:e0146009.

Morris DJ, Speirs DC, Cameron AI, Heath MR. 2014. Global sensitivity analysis of an end-to-end marine ecosystem model of the North Sea: Factors affecting the biomass of fish and benthos. Ecological Modelling 273:251-263.

Peterman RM, Pyper BJ, MacGregor BW. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences 60:809-824.

Plummer M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing. Vienna, Austria.

Pyper BJ, Mueter FJ, Peterman RM, Blackbourn DJ, Wood CC. 2002. Spatial covariation in survival rates of northeast Pacific chum salmon. Transactions of the American Fisheries Society 131:343-363.

Quinn TP. 1993. A review of homing and straying of wild and hatchery-produced salmon. Fisheries Research 18:29-44.

R Development Core Team. 2016. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ricker WE. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bulletin of the Fisheries Research Board of Canada, No 119., Ottawa.

Shortreed KS, Hume JMB, Stockner JG. 1999. Using photosynthetic rates to estimate the juvenile sockeye rearing capacity of British Columbia lakes. Pages 502-521 in Knudsen EE, Stewart CR, MacDonald DD, Williams JE, Reiser DW, eds. Sustainable Fisheries Management: Pacific Salmon, CRC Press LLC, Boca Raton, New York.

Su Y-S, Yajima M. 2012. R2jags: A Package for Running jags from R. R package version 0.03-08, URL http://CRAN. R-project. org/package= R2jags.

Szerlong RG, Rundio DE. 2008. A statistical modeling method for estimating mortality and abundance of spawning salmon from a time series of counts. Canadian Journal of Fisheries and Aquatic Sciences 65:17-26.

Walters CJ, Martell SJD. 2004. Fisheries Ecology and Management. Princeton University Press.

Figure captions

Fig. 1. Prior distribution of Ricker α parameter, productivity.

Fig. 2. Observed spawner-recruit data over time, with fitted Ricker curves and associated benchmarks for the standard Bayesian Ricker model (i), and the Bayesian hierarchical Ricker model (ii). Shaded regions indicate 95% credible intervals, delineated by 2.5th and 97.5th posterior densities. Red and green circles on x-axis identify percentile-based benchmarks (S_{25th} and S_{75th}, respectively). Cross indicates most recent data point. Colours of points increase in darkness as years progress towards the most current year.

Fig. 3. Raw and generational average escapement over time, with retrospective conservation benchmarks overlaid for three benchmarks types: (i) standard Ricker model; (ii) hierarchical Ricker model; (iii) percentile. Shaded regions indicate 95% credible intervals, delineated by 2.5th and 97.5th posterior densities. Retrospective benchmarks use all available data up to that year.

Fig. 4. Standardized raw and generational average escapements across CUs, with conservation status indicated by coloured bars below. Transparent bars indicate upper and lower credible interval bounds, based on 2.5th and 97.5th posterior densities of estimated parameters. . Gaps exist either because of missing recruitment data, or because status was not assessed when α values were < 1.5, as suggested by Holt and Ogden (2013).

Fig. 5. Estimated Ricker αvalues using a recursive Bayes model, which allows alpha to vary over time, within a given CU. Grey shaded polygons indicate 95% credible intervals based on posterior densities of estimated α values.

Fig.6. Schematic of simulation model used to evaluate benchmark performance.

Fig 7(a). Time-series of observed spawner abundances (black line) and "true" spawner abundances (grey line) and benchmarks for one Monte Carlo trial. Benchmarks are estimated annually base on all data up until that year: annual estimates of 80% S_{MSY} (upper benchmark, green dashed line), annual estimates of S_{gen} (lower benchmark, red dashed line), the 75th percentile benchmark (green dotted line), and the 25th percentile (red dotted line). 95% confidence intervals are shown for estimates of stock-recruitment based benchmarks (green and red shading for the upper and lower benchmarks, respectively). (b) Observed spawner and recruitment data (solid black dots) and "true" data (grey hollow dots) for the final year of one Monte Carlo trial. The "true" underlying stock-recruitment relationship is shown with the grey curve and the estimate based on observed data is shown with the black curve. (c) Mean percent error, MPE, between estimated and "true" benchmark averaged over all Monte Carlo trials. Red bars are the mean percent error from the "true" S_{gen} (lower benchmark), and green bars the mean percent error from the "true" 80% S_{MSY}. Asterisk in panel (c) is MPE beyond the limit of the y-axis, 229%.

Fig. 8. Difference in Mean Percent Error, MPE, of estimated lower benchmark (S_{gen}, (a), and S_{25th} (b)) and the "true" lower benchmarks ("true" S_{gen}), between sensitivity analyses listed on the x-

axis and the base case scenario. Black bars are analyses where the input parameter was increased relative to the base case (see Table 1); white bars are analyses where the input parameter was reduced relative to the base case. Positive values indicate sensitivity analyses where MPE increased under that change in input parameter from the base case; negative values indicate analyses where the MPE declined under that change in input parameter. Asterisks denote values above the limit of the *y*-axis, 156% (a) and 884% (b).

Fig. 9. Mean percent error, MPE, of the estimated lower benchmark (S_{25th} (a), and S_{gen} (b)) from the "true" S_{gen} value along a gradient in initial harvest rates (*x*-axis) and productivities (*y*-axis) derived from a simulation model of a hypothetical salmon CU. Dots indicate MPE of CUs assuming productivities estimated from hierarchical Ricker models and mean harvest rate over available time-series for each CU.

Fig. 10. Mean percent error of the estimated upper benchmark (S_{75th} percentile (a), and 80% S_{MSY} (b)) from the "true" 80% S_{MSY} value along a gradient in initial harvest rates (*x*-axis) and productivities (*y*-axis) derived from a simulation model of a hypothetical salmon CU. Dots indicate MPE of CUs assuming productivities estimated from hierarchical Ricker models and mean harvest rate over available time-series for each CU.

Fig. 11. Sensitivity indices of the effects of individual variables (mean elemental effect, *x*-axis) and interactions among variables (standard deviation in elemental effects, *y*-axis). Indices were derived from the Morris method, a global sensitivity analyses for the mean percent error of estimated lower benchmarks (S_{gen} (a), and S_{25th} (b)) from "true" benchmarks. Input variables with values >100 on either axis are labelled.

Fig. 12. "True" values of Ricker parameters and benchmarks under assumption of constant spawners at equilibrium, S_{eq}, (a-e) or constant spawner abundances at maximum recruitment S_{max} (f-j), with abrupt changes in productivity in year 35 of simulation.

Fig. 13. Box plots of parameter and benchmark values in final year of the simulation averaged over all Monte Carlo trials, using either the first 30 years of data (dark grey boxes), all 70 years of data (light grey boxes), or the most recent 30 years of data (white boxes). Boxes represent the lower quartile, median, and upper quartiles of the parameter distribution. Whiskers are the 95% confidence intervals. Dashed lines represent the "true" value for each parameter. For the percentile benchmarks, the dashed lines represent the "true" S_{gen} (e) and "true" 80% of S_{MSY} (f). S_{eq} is held constant in simulations as productivity varies.

Fig. 14. Mean percent error of estimated Ricker parameters and lower and upper benchmarks from the "true" values, using only the first 30 years of data (black bars), all 70 years (grey bars), and the most recent 30 years (white bars) assuming either S_{eq} remains constant over time (a), or S_{max} remains constant (b).

Fig. 15. Caption as for Fig. 13, except S_{max} is held constant and S_{eq} declines in simulations as productivity varies.

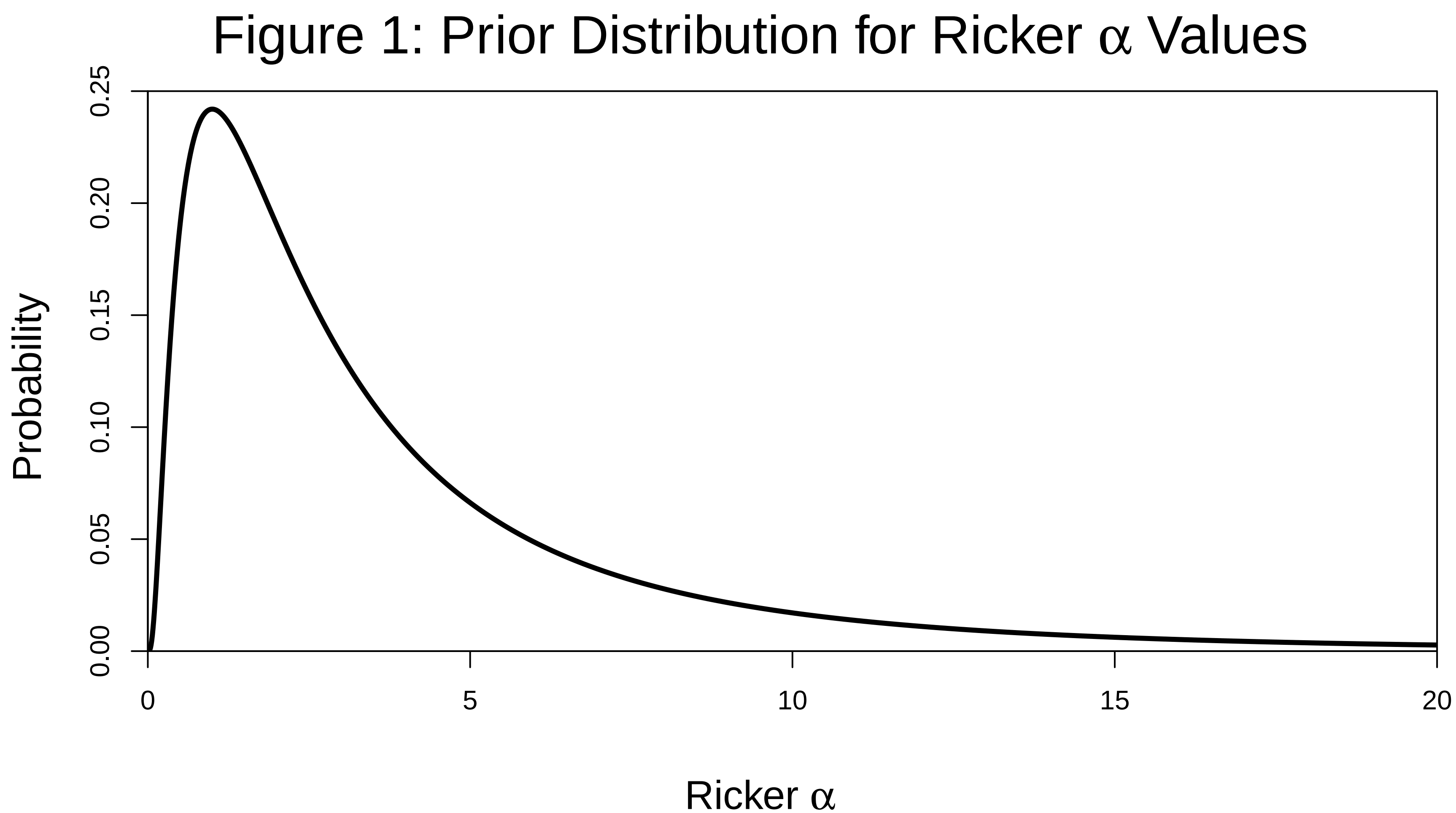


Figure 1: Prior Distribution for Ricker α Values
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0
5
10
15
20
Ricker α

Figure 2: Final Year Ricker Curves and Benchmarks

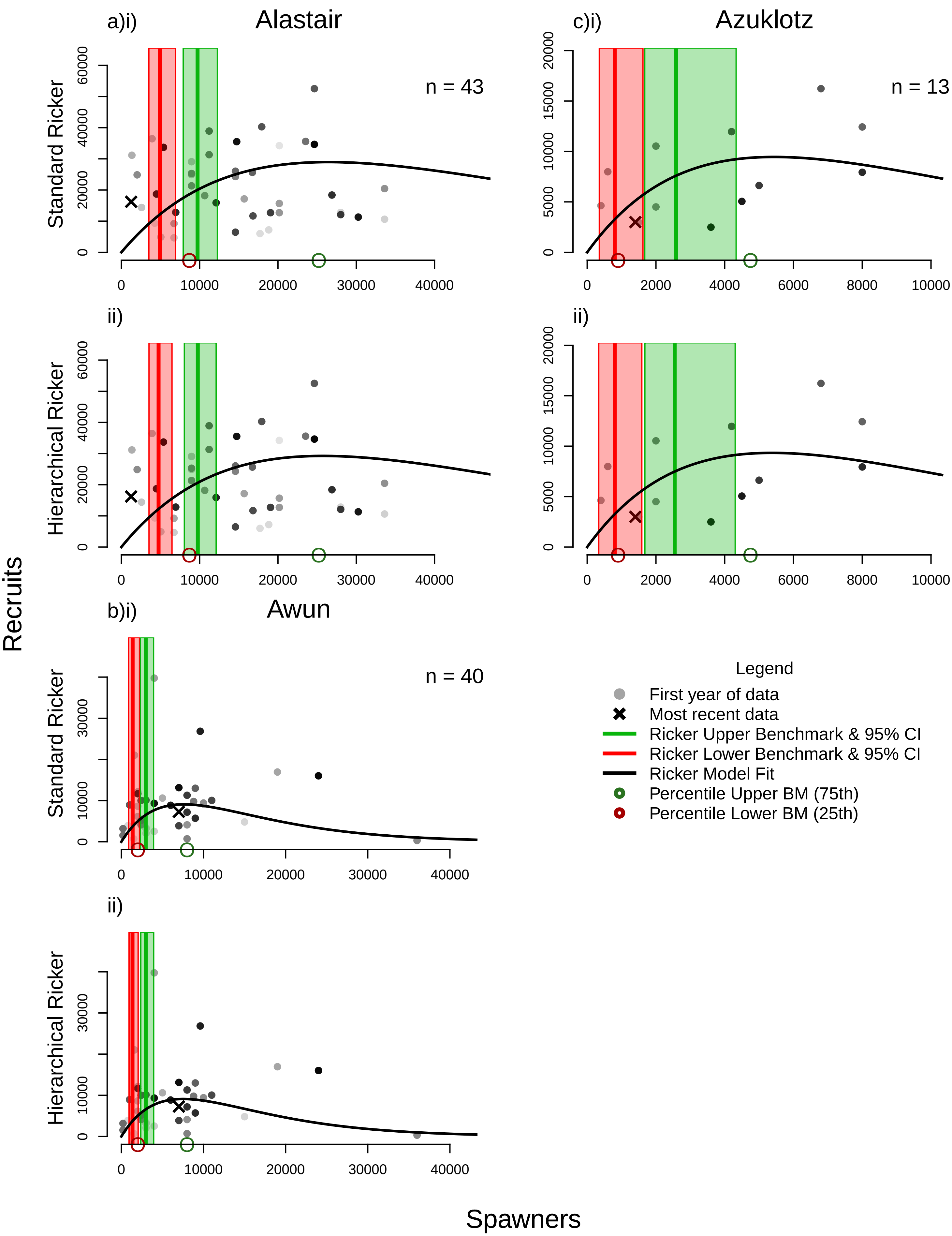


Figure 2: Final Year Ricker Curves and Benchmarks

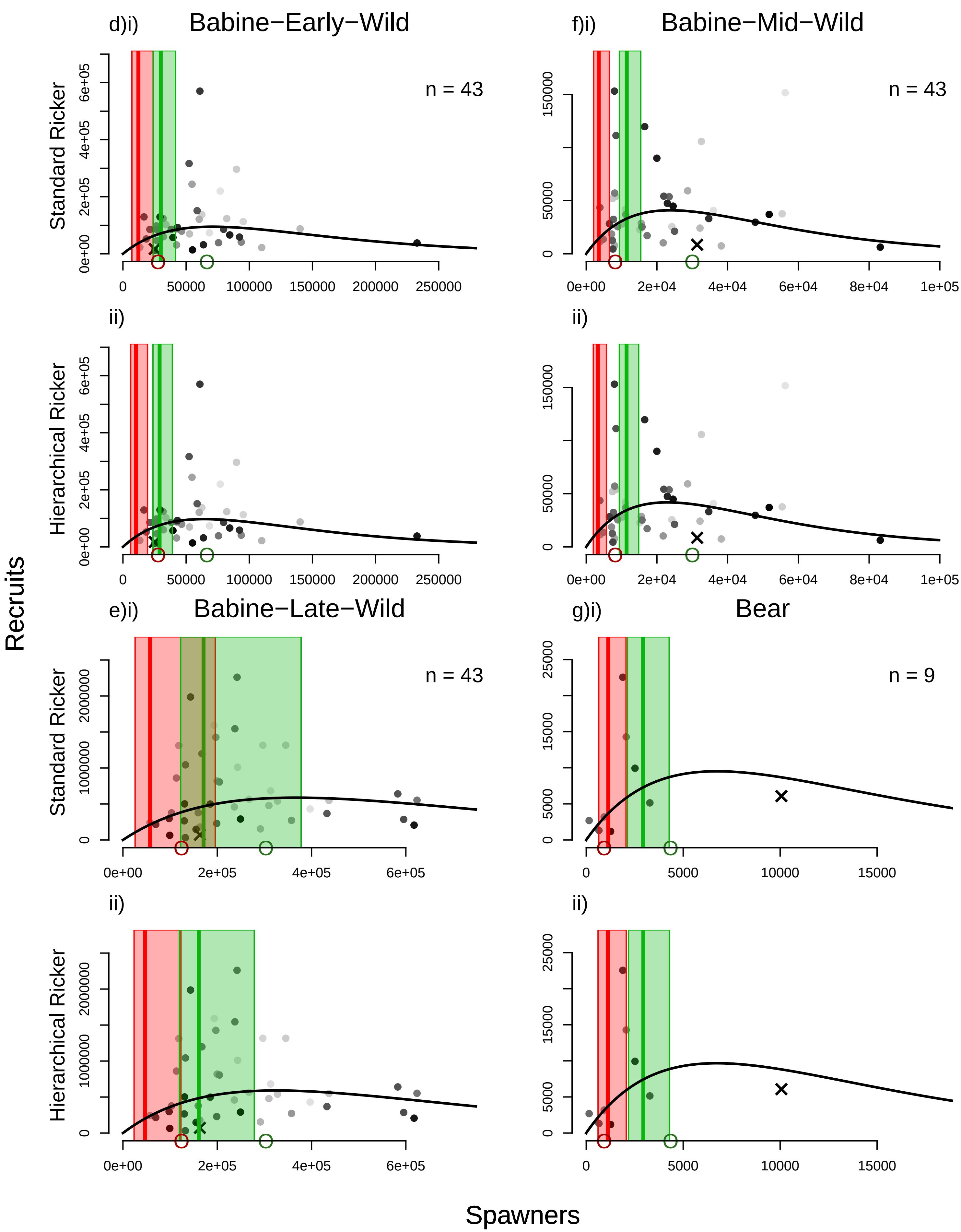


Figure 2: Final Year Ricker Curves and Benchmarks

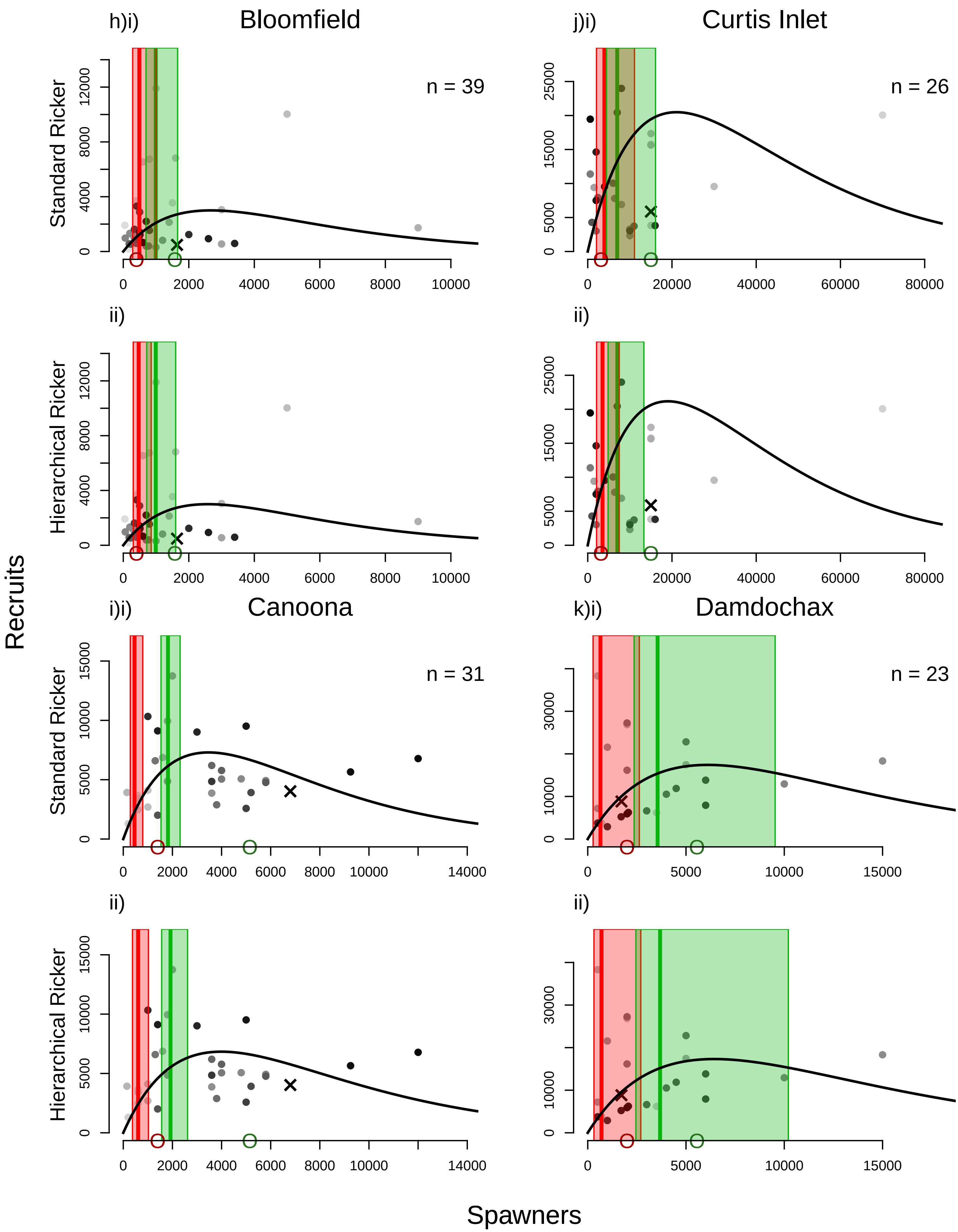


Figure 2: Final Year Ricker Curves and Benchmarks

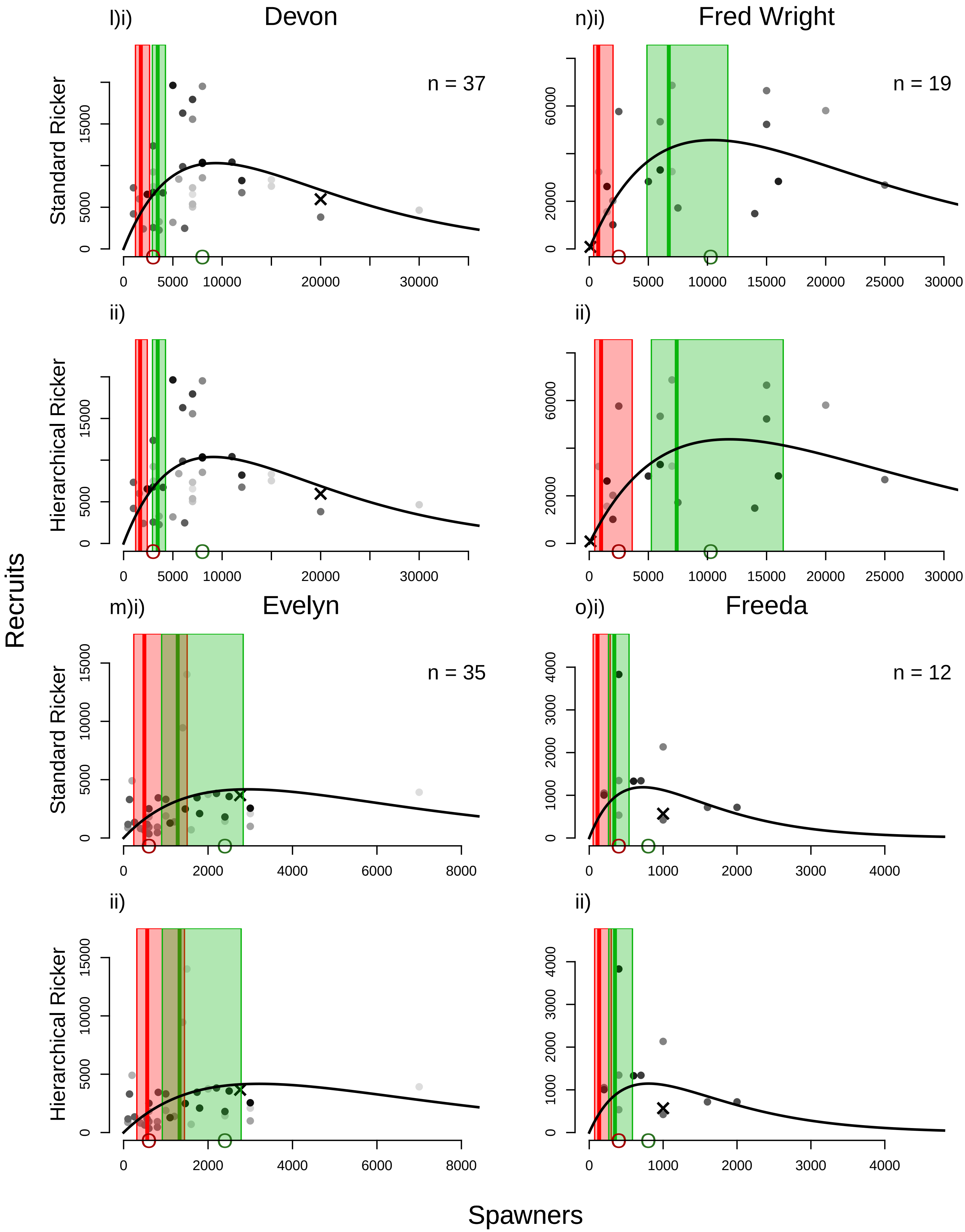


Figure 2: Final Year Ricker Curves and Benchmarks

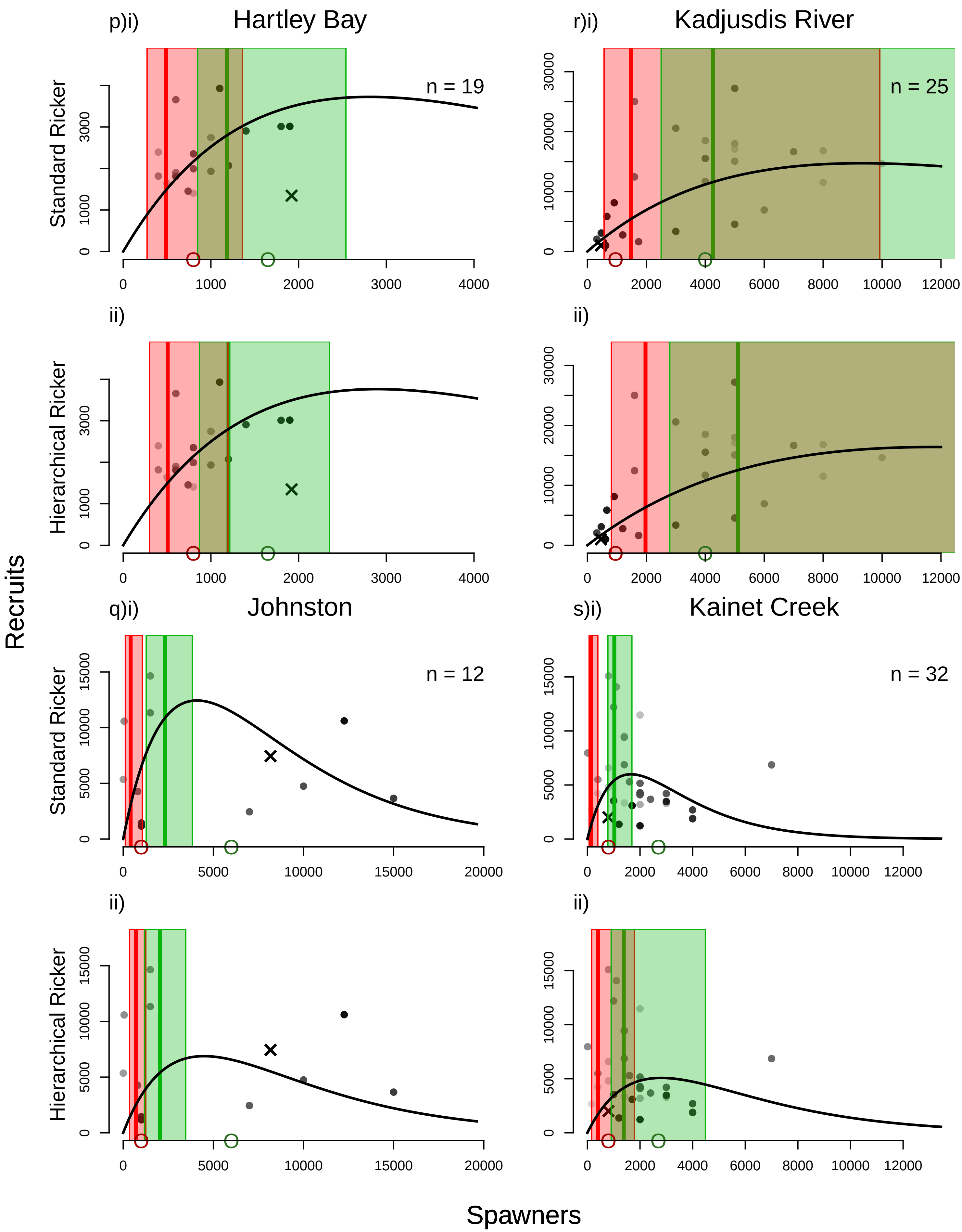


Figure 2: Final Year Ricker Curves and Benchmarks

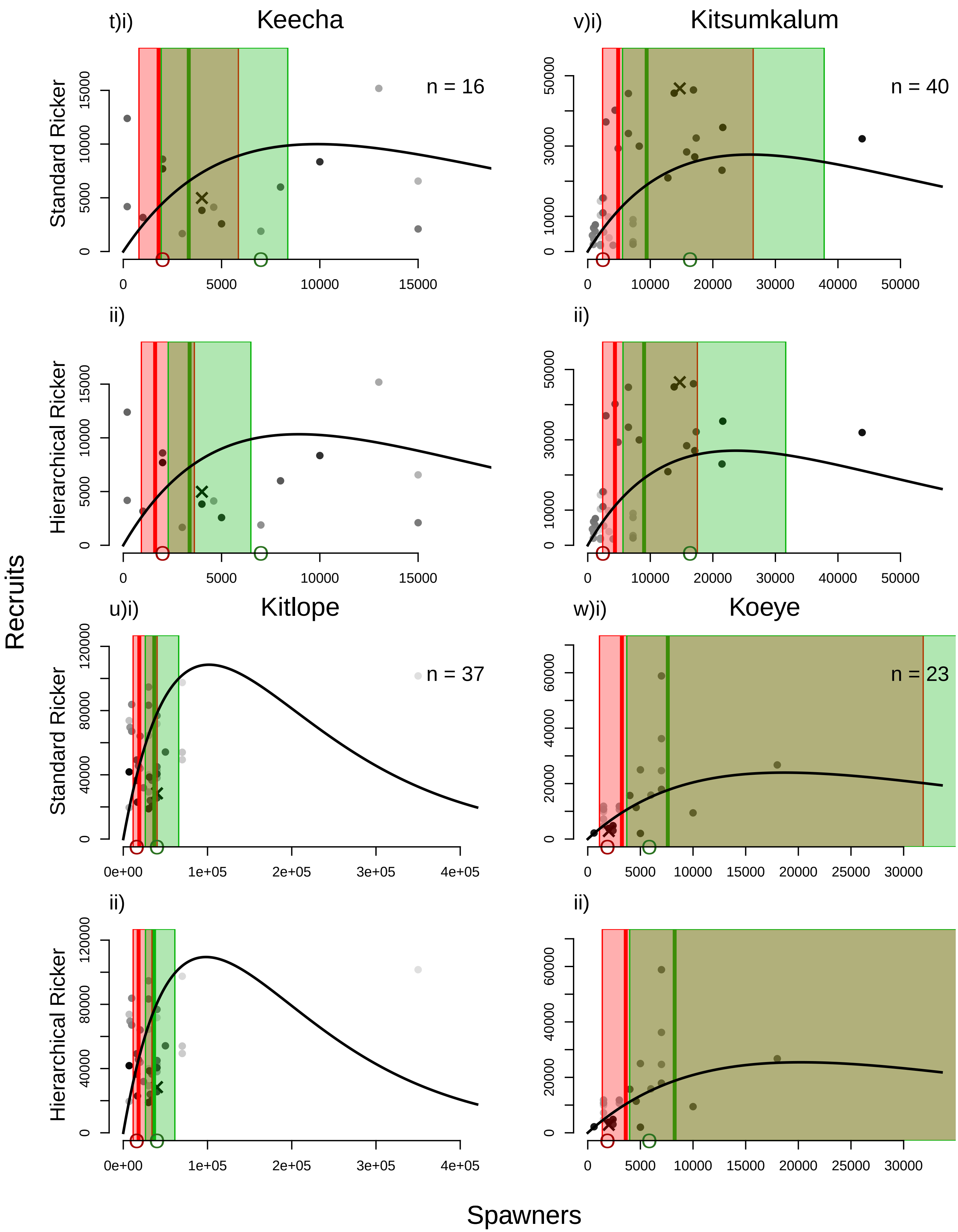


Figure 2: Final Year Ricker Curves and Benchmarks

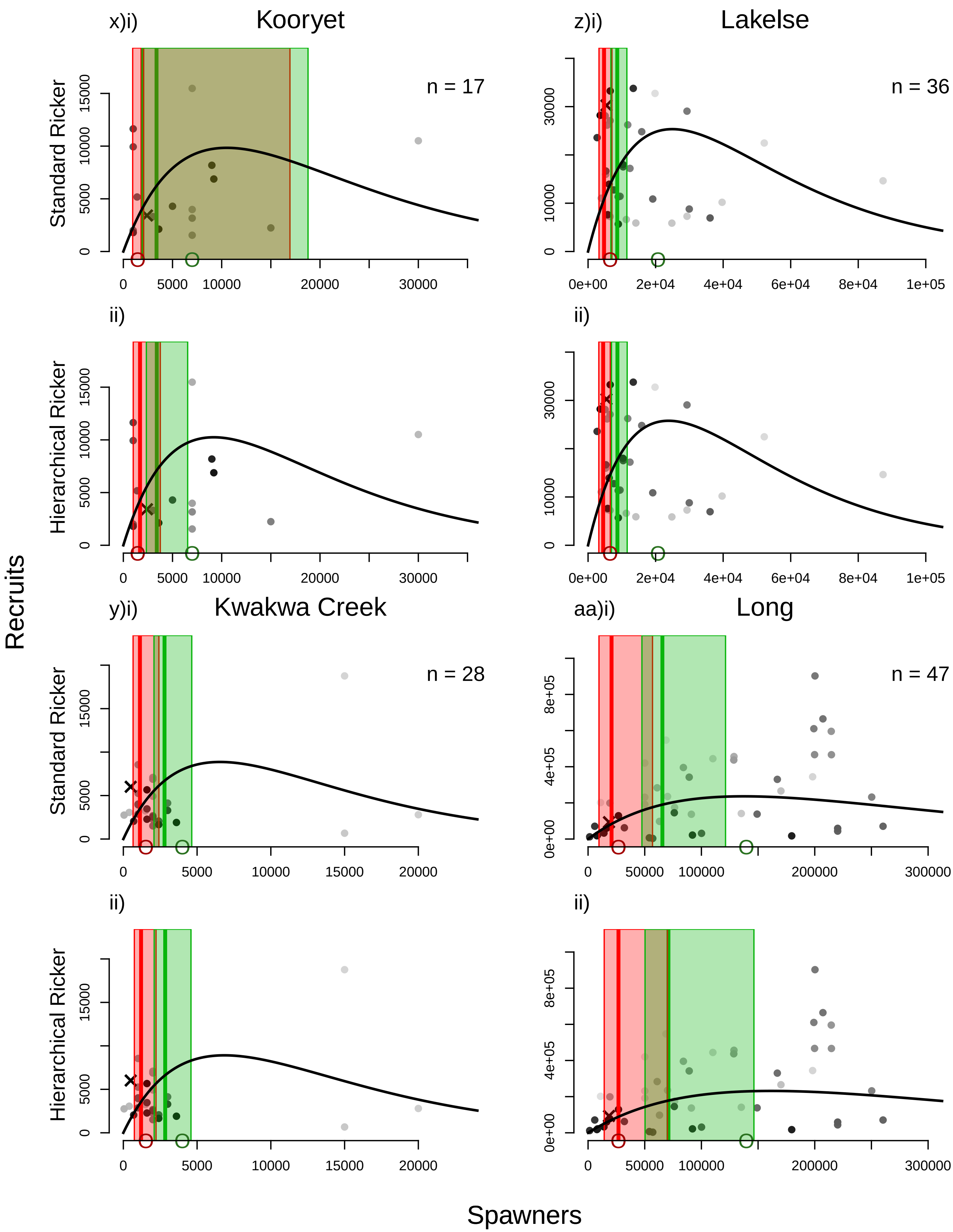


Figure 2: Final Year Ricker Curves and Benchmarks

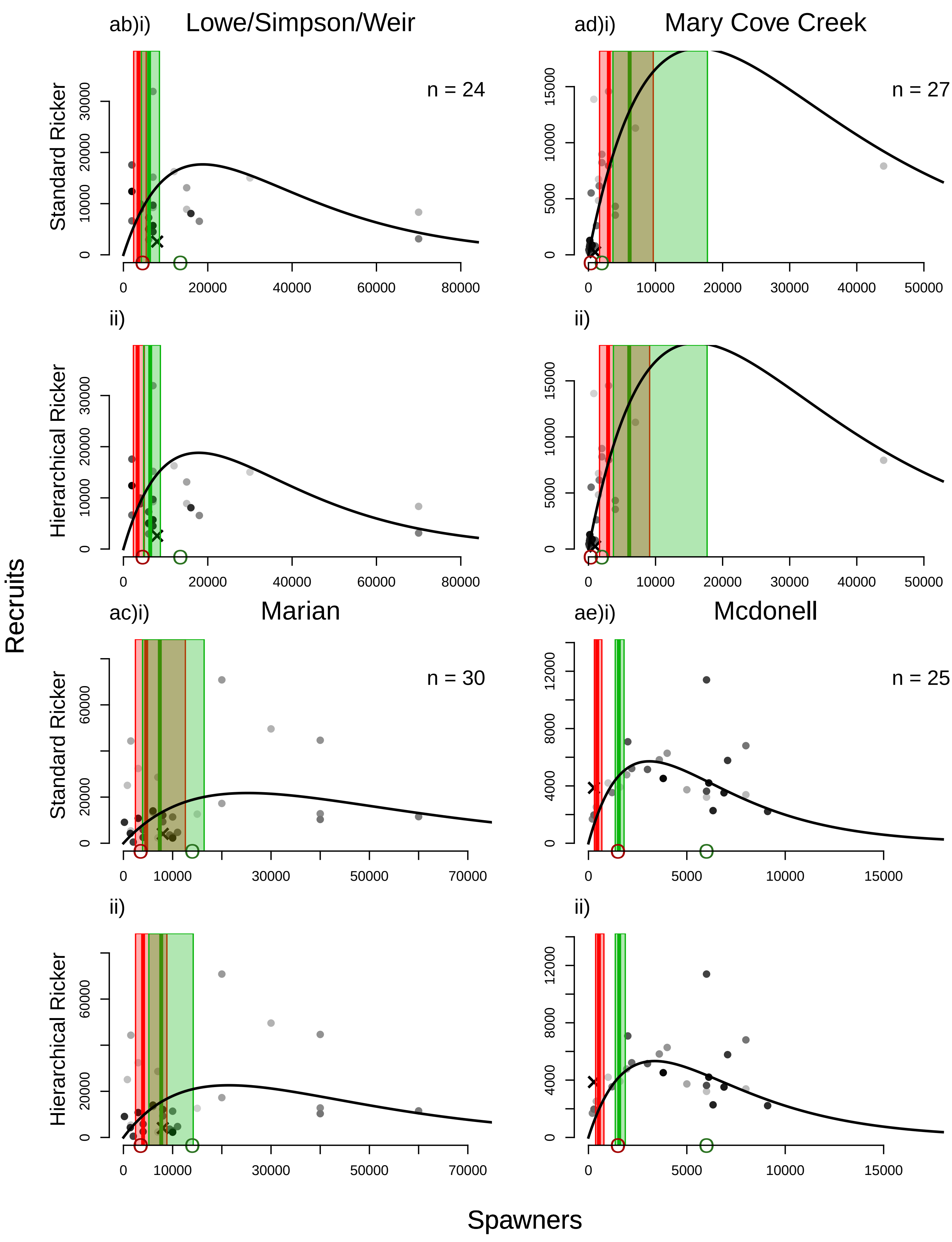


Figure 2: Final Year Ricker Curves and Benchmarks

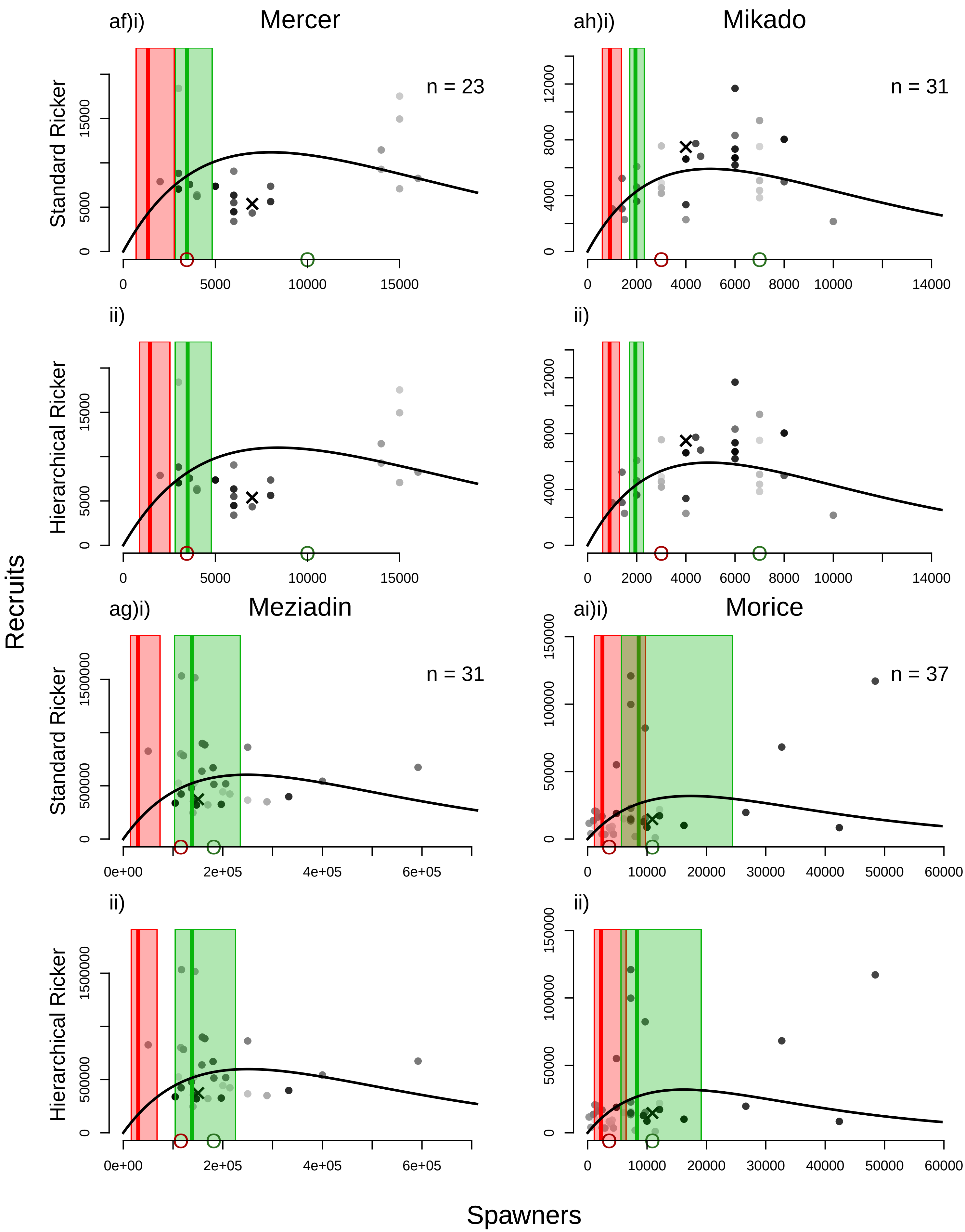


Figure 2: Final Year Ricker Curves and Benchmarks

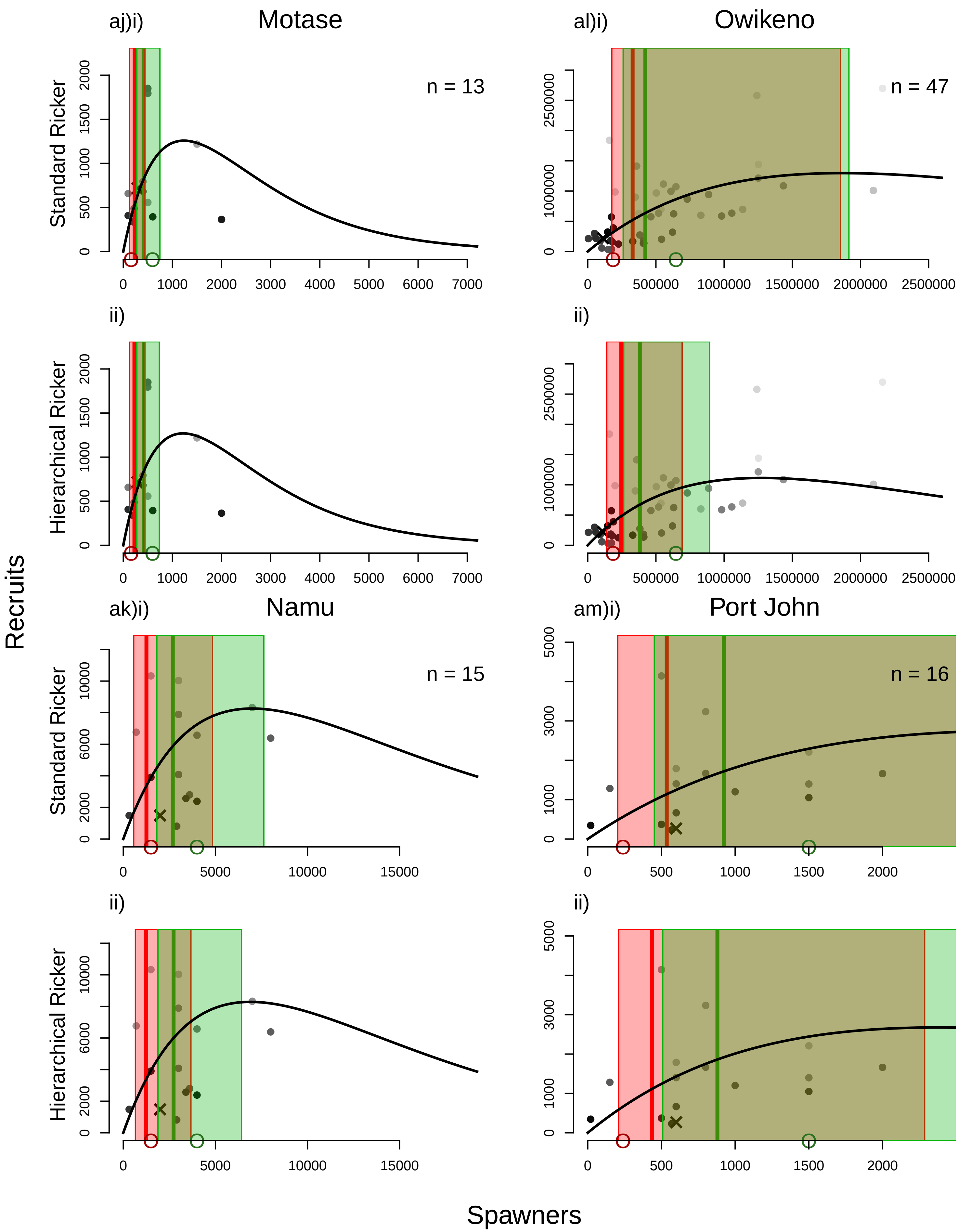


Figure 2: Final Year Ricker Curves and Benchmarks

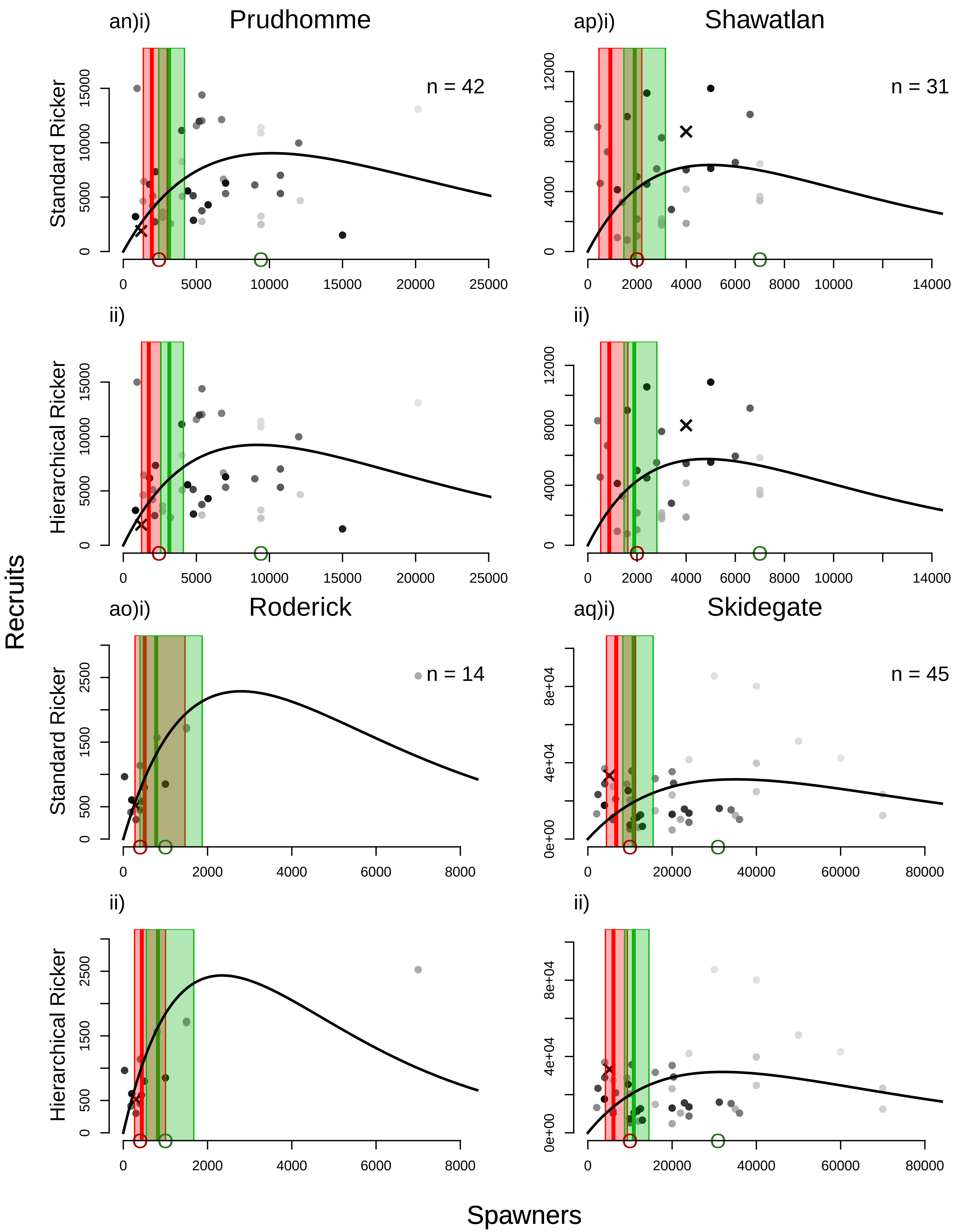


Figure 2: Final Year Ricker Curves and Benchmarks

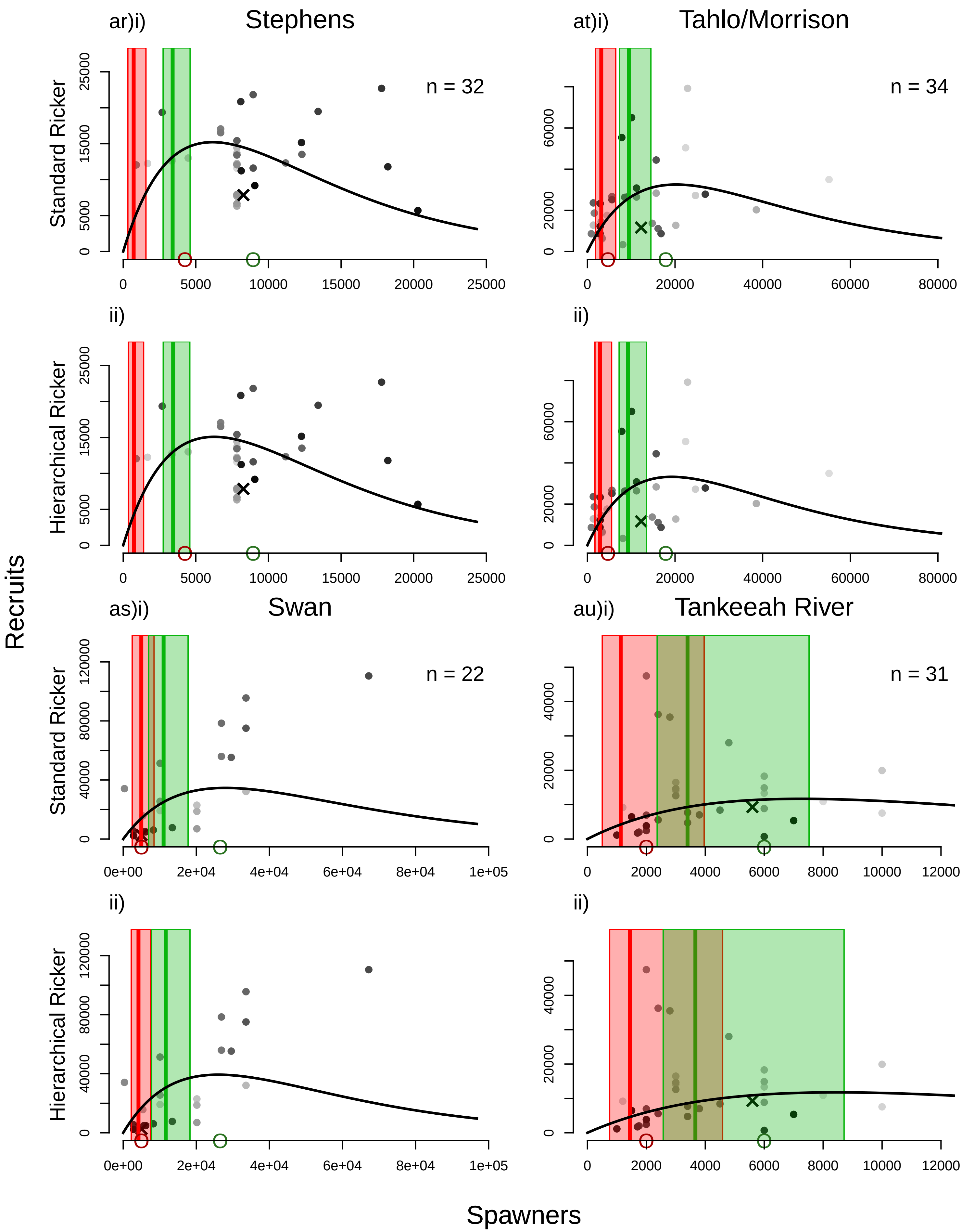


Figure 2: Final Year Ricker Curves and Benchmarks

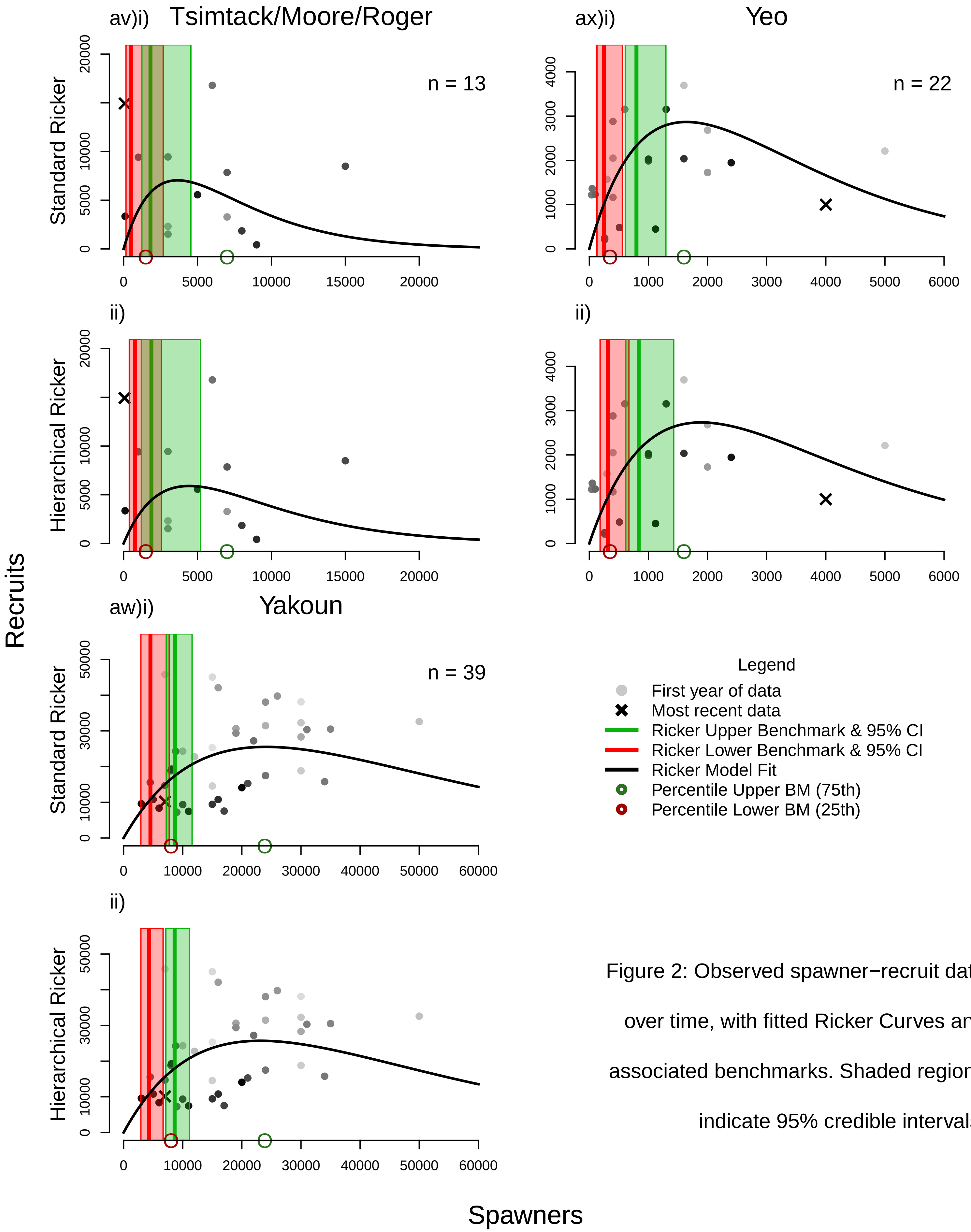


Figure 2: Observed spawner−recruit data over time, with fitted Ricker Curves and associated benchmarks. Shaded regions indicate 95% credible intervals.

Figure 3: Conservation Benchmarks Over Time

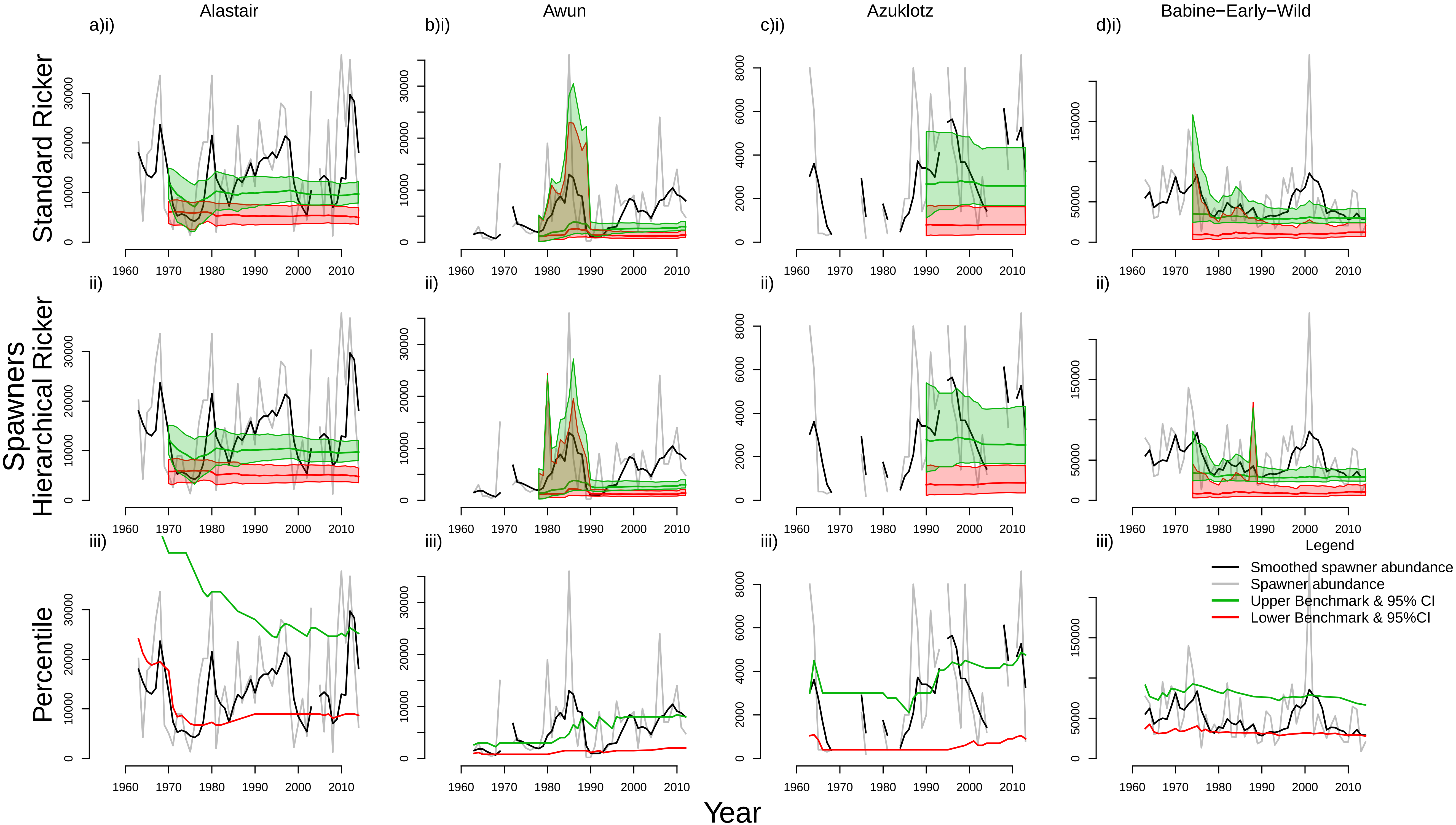
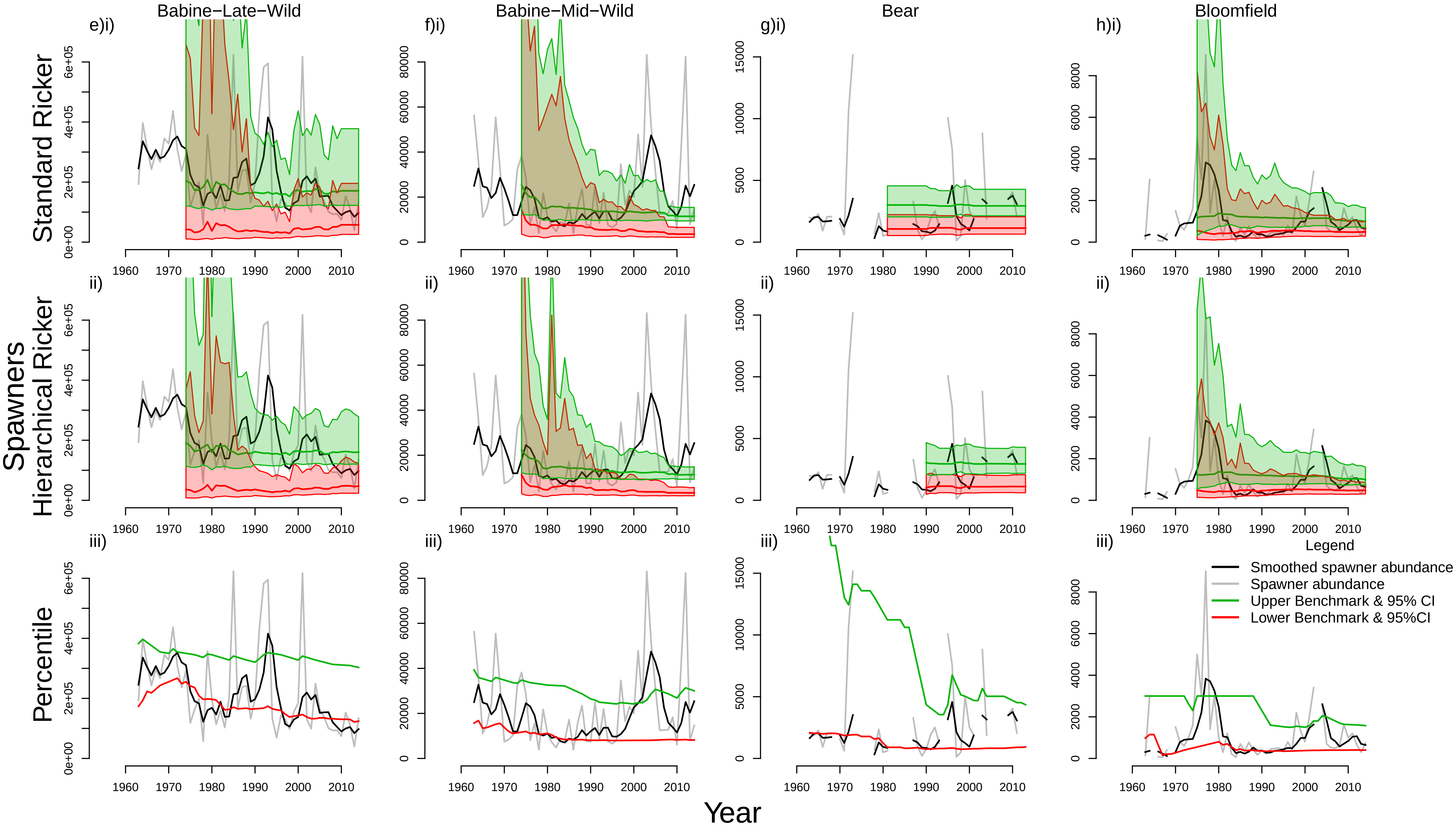


Figure 3: Conservation Benchmarks Over Time



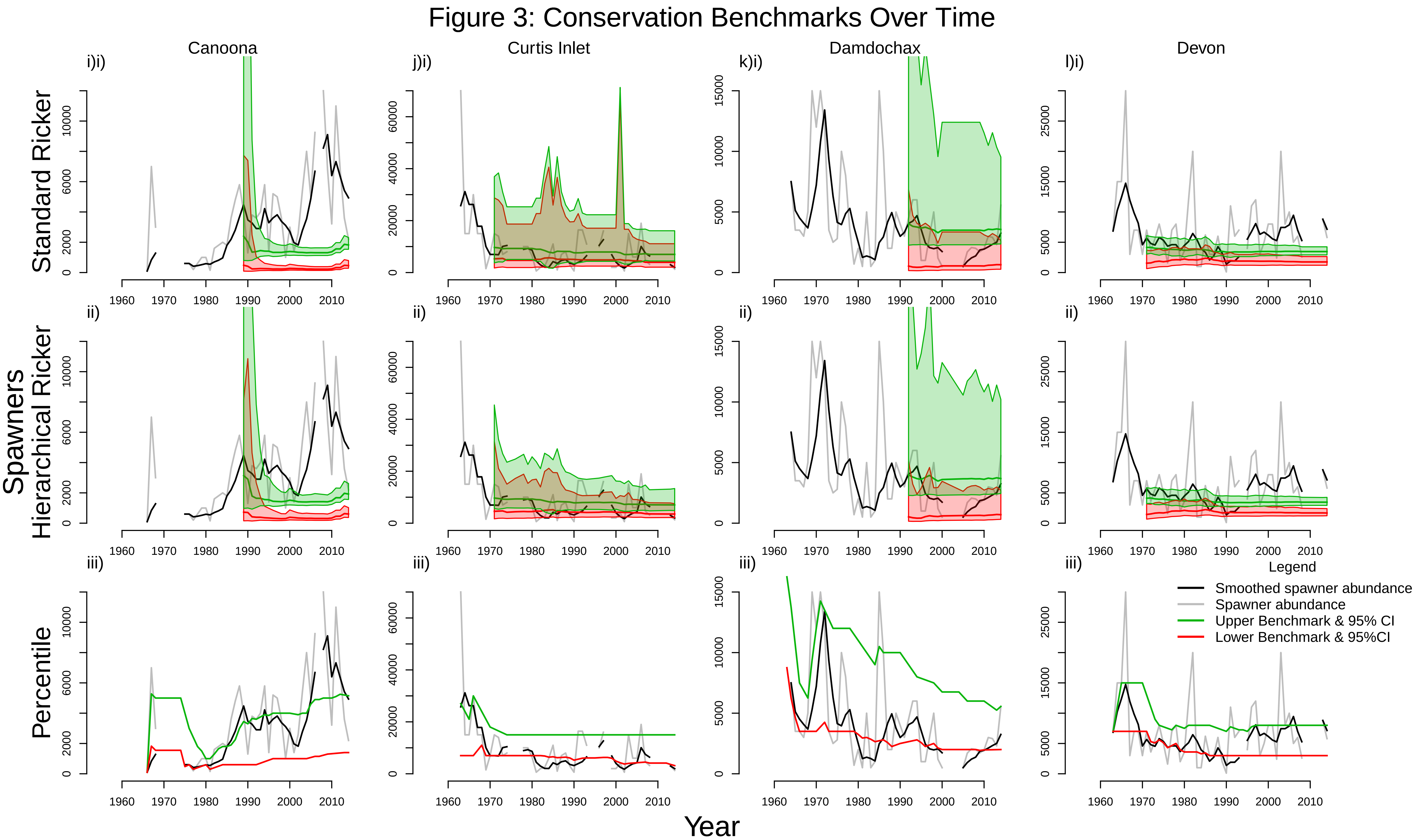
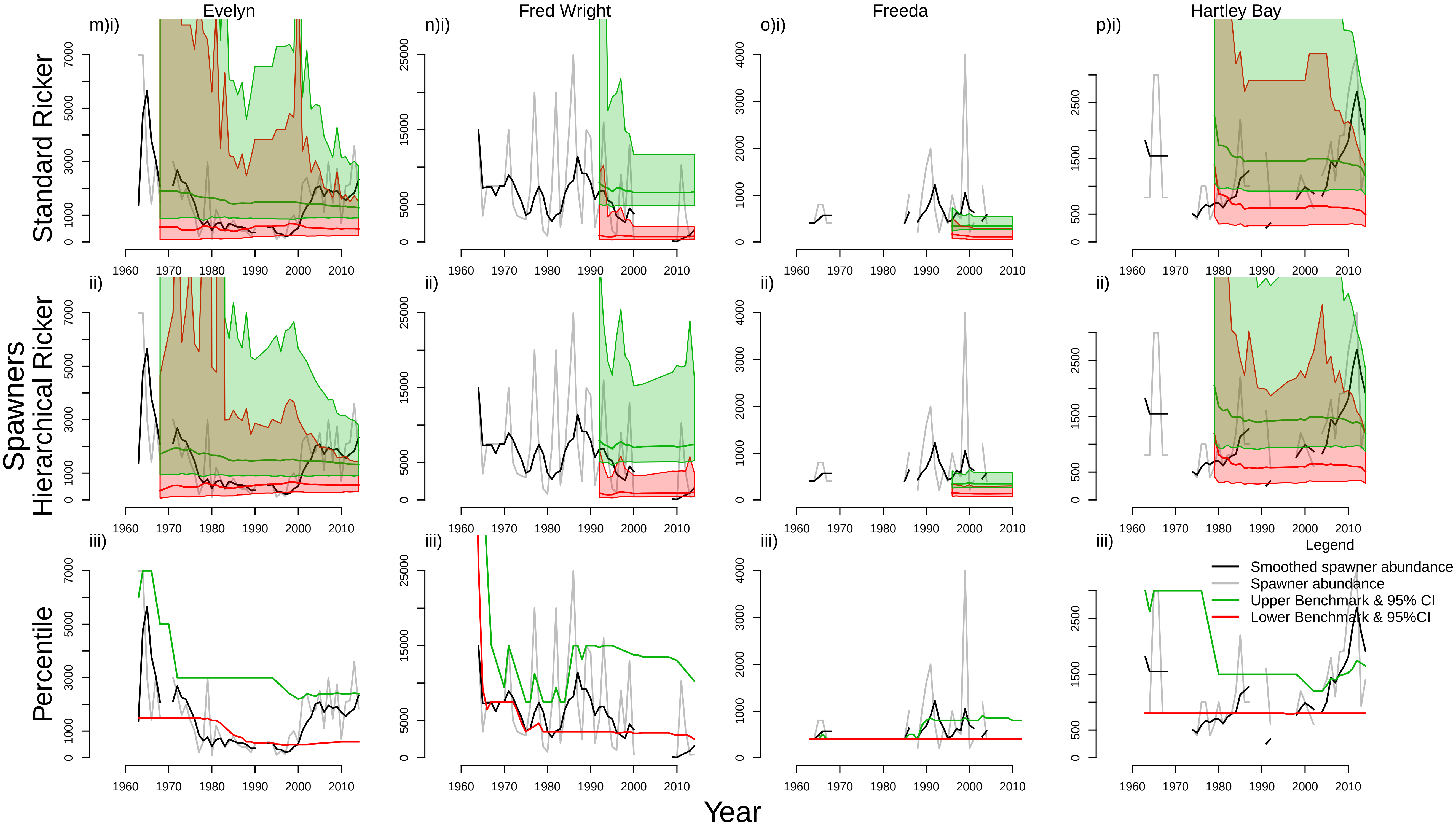


Figure 3: Conservation Benchmarks Over Time
Canoona
Curtis Inlet
Damdochax
Devon
i)i)
j)i)
k)i)
l)i)
ii)
iii)
Standard Ricker
Hierarchical Ricker
Percentile
Spawners
Year
Legend
Smoothed spawner abundance
Spawner abundance
Upper Benchmark & 95% CI
Lower Benchmark & 95%CI

Figure 3: Conservation Benchmarks Over Time



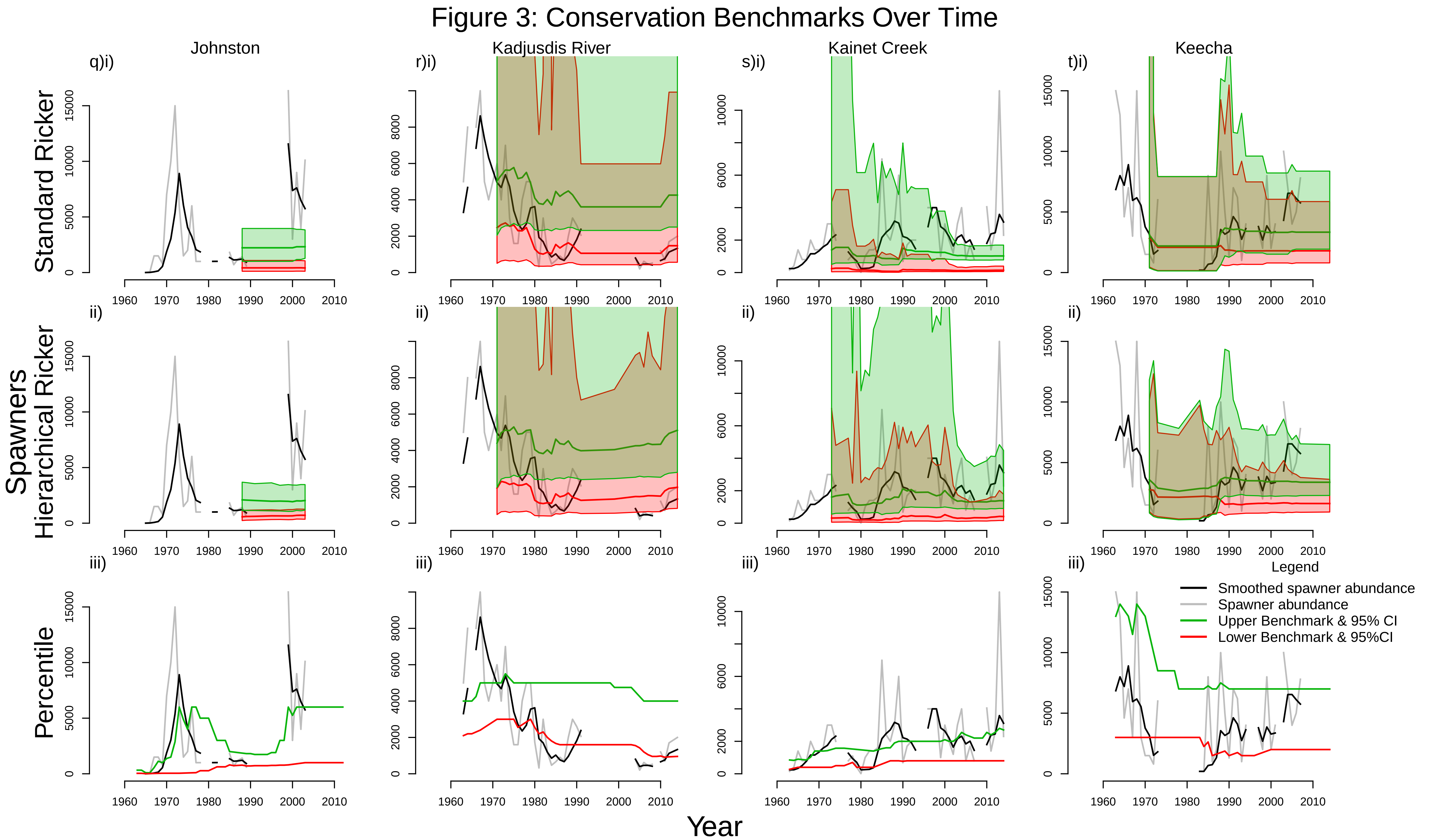


Figure 3: Conservation Benchmarks Over Time
Johnston
Kadjusdis River
Kainet Creek
Keecha
q)i)
r)i)
s)i)
t)i)
ii)
iii)
Standard Ricker
Hierarchical Ricker
Percentile
Spawners
Year
Legend
Smoothed spawner abundance
Spawner abundance
Upper Benchmark & 95% CI
Lower Benchmark & 95%CI

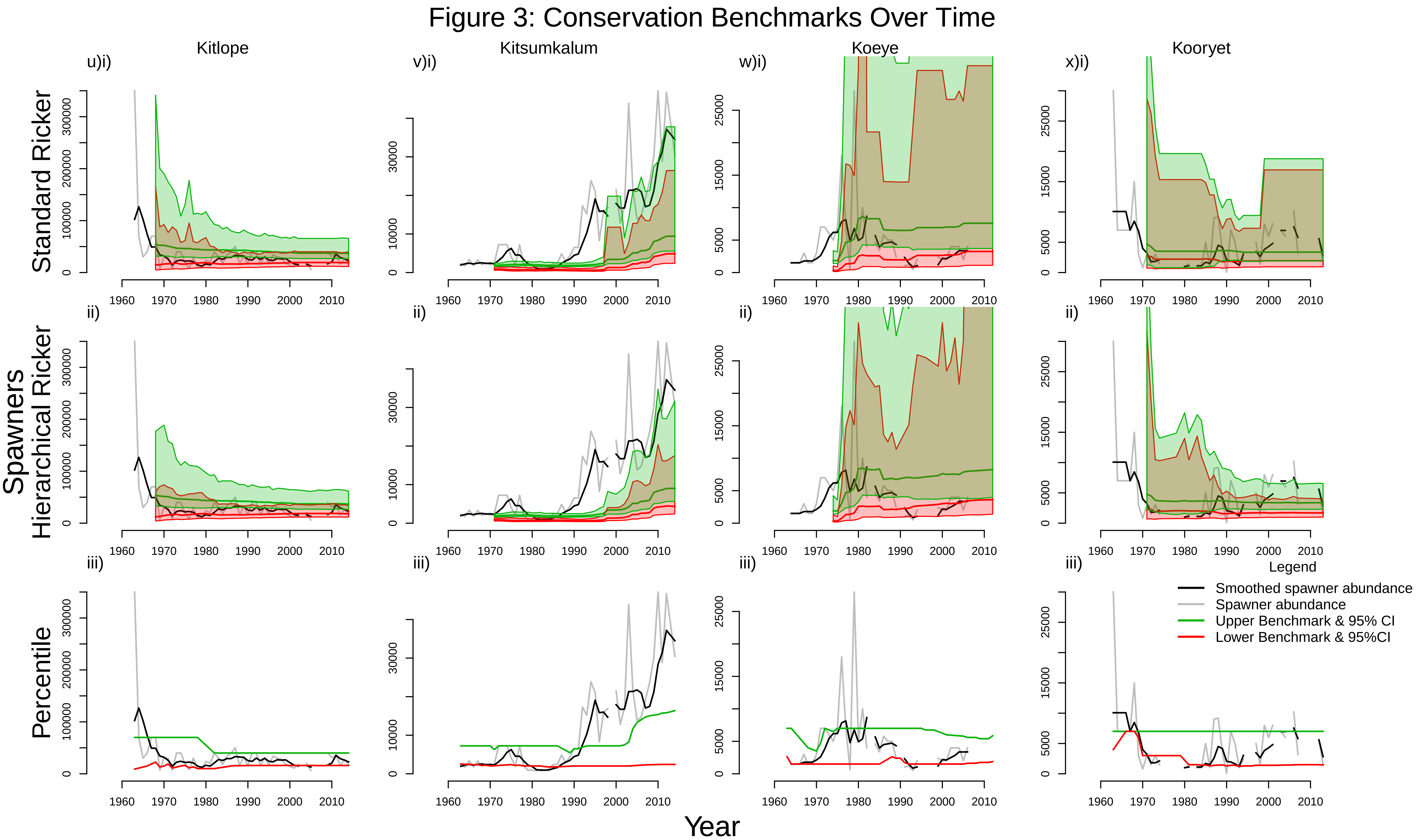


Figure 3: Conservation Benchmarks Over Time
Kitlope
Kitsumkalum
Koeye
Kooryet
u)i)
v)i)
w)i)
x)i)
ii)
iii)
Standard Ricker
Hierarchical Ricker
Percentile
Spawners
Year
Legend
Smoothed spawner abundance
Spawner abundance
Upper Benchmark & 95% CI
Lower Benchmark & 95%CI

Figure 3: Conservation Benchmarks Over Time

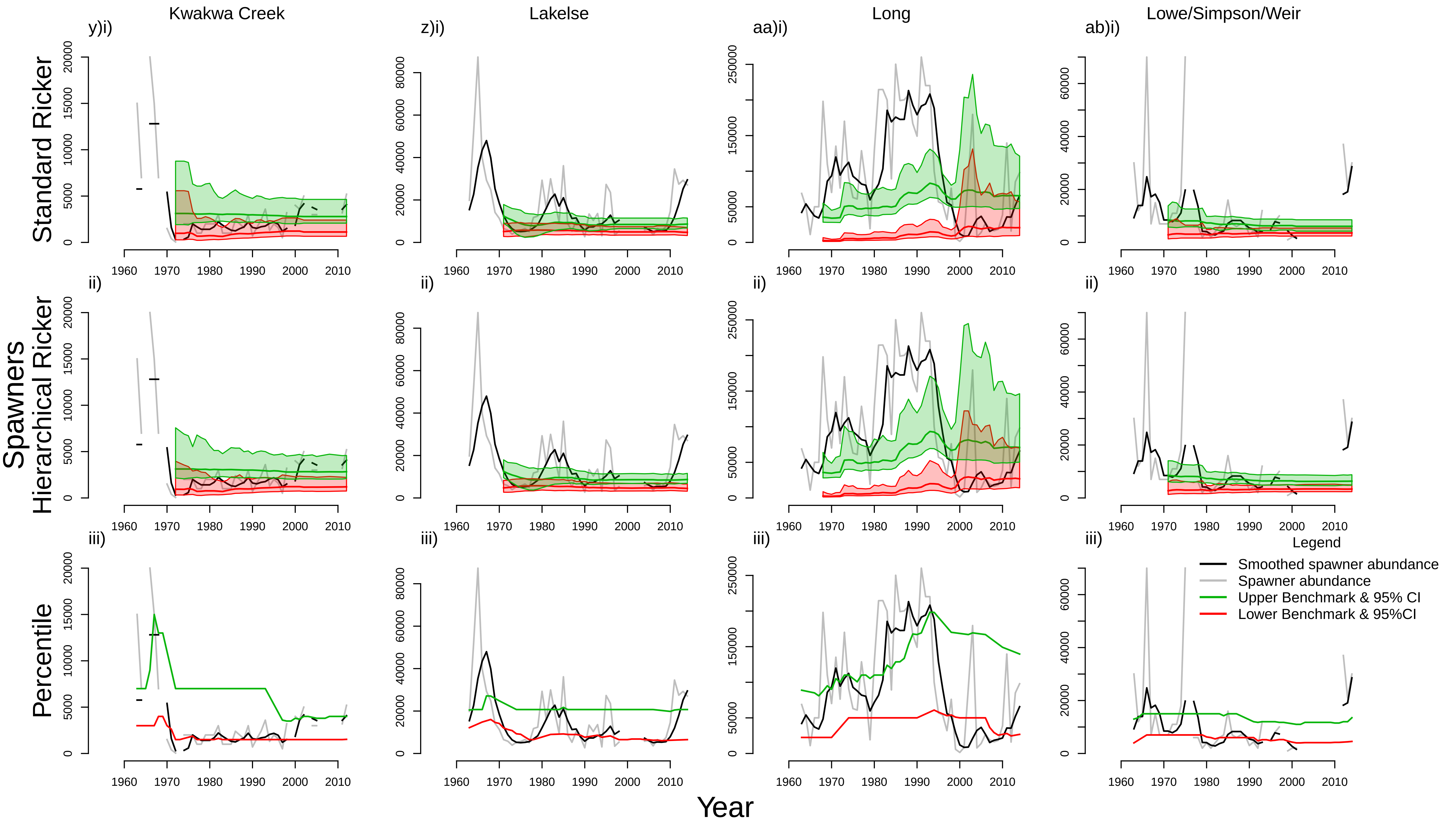


Figure 3: Conservation Benchmarks Over Time

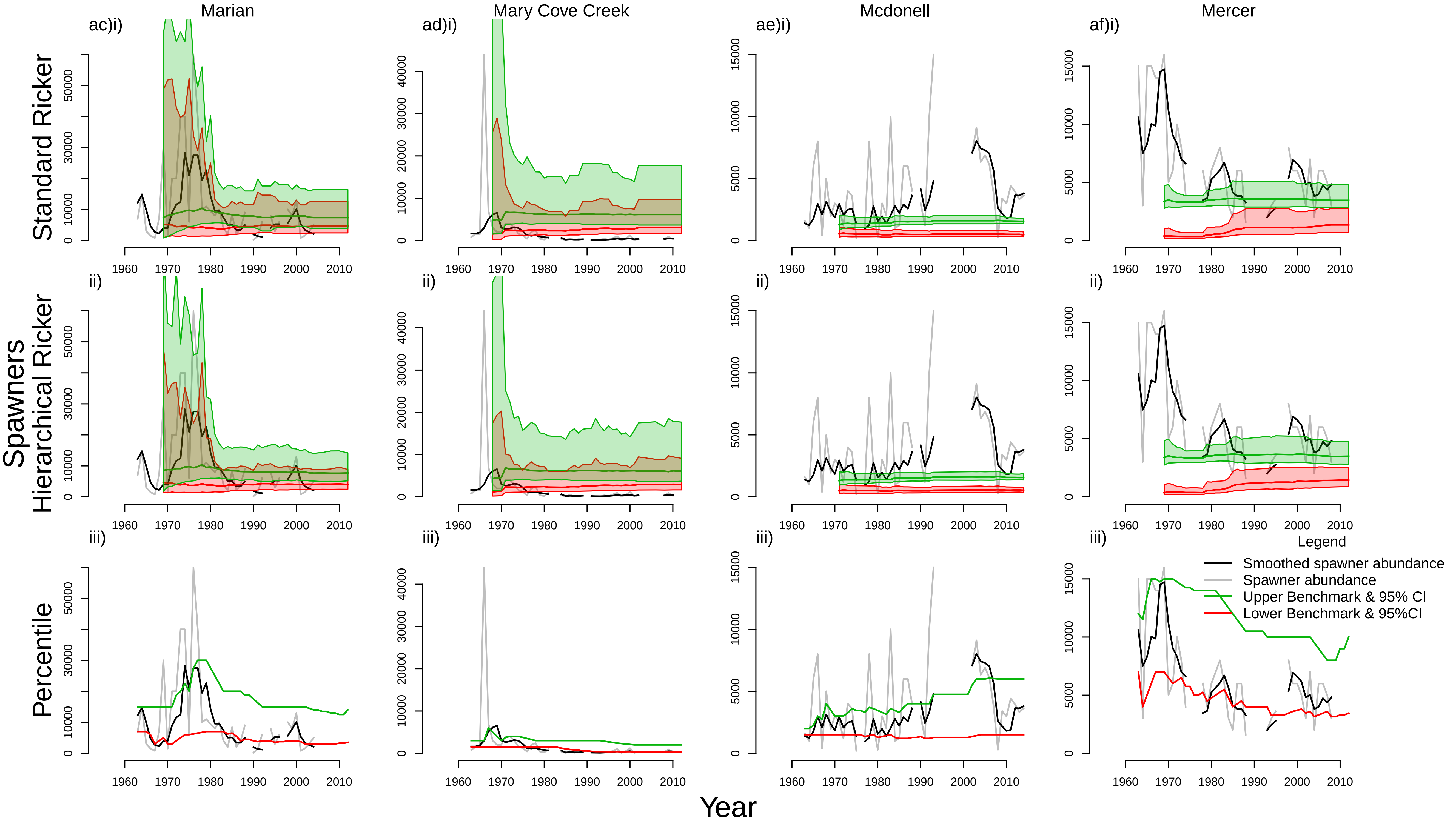


Figure 3: Conservation Benchmarks Over Time

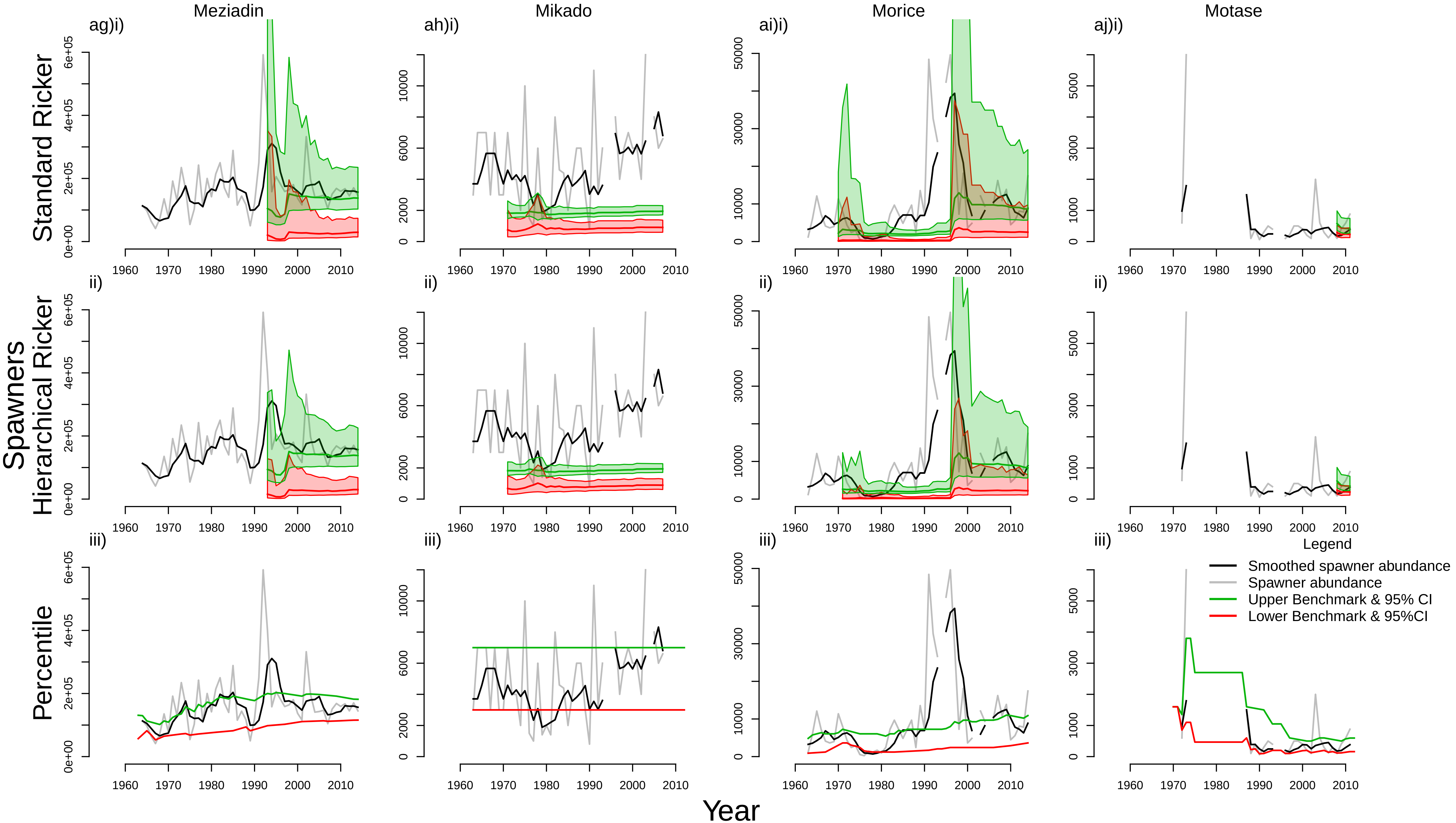


Figure 3: Conservation Benchmarks Over Time

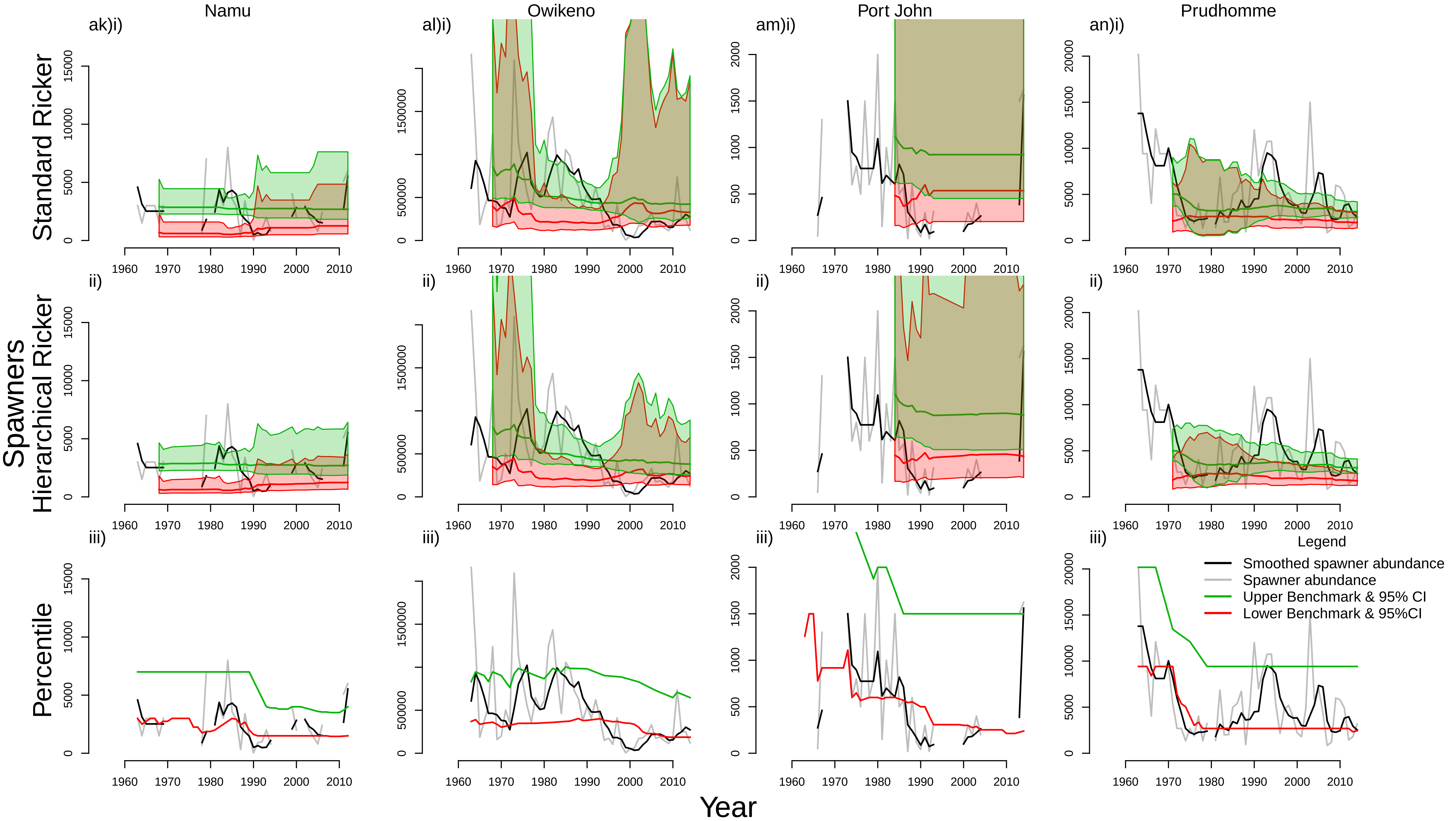


Figure 3: Conservation Benchmarks Over Time

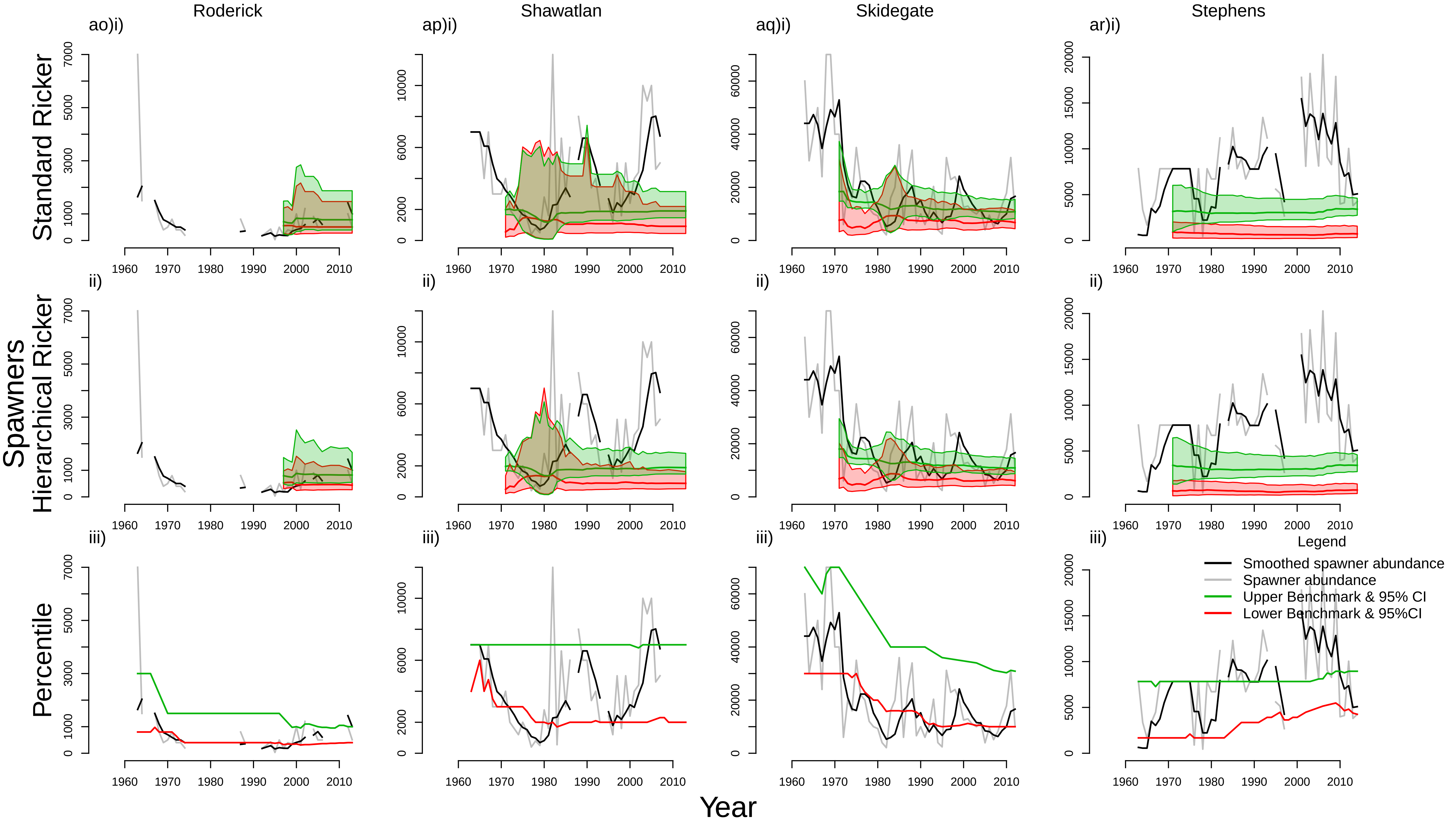


Figure 3: Conservation Benchmarks Over Time

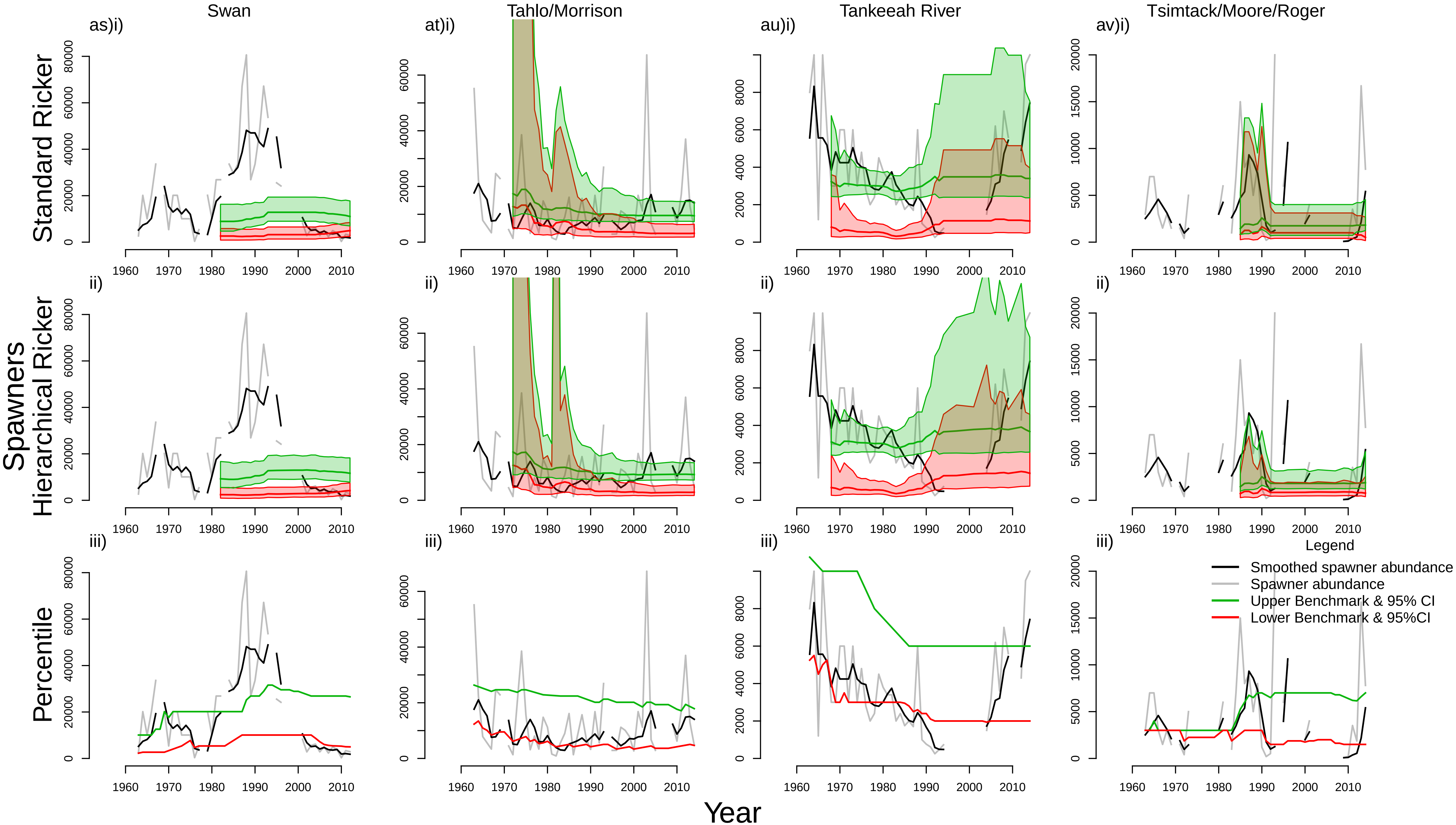


Figure 3: Conservation Benchmarks Over Time

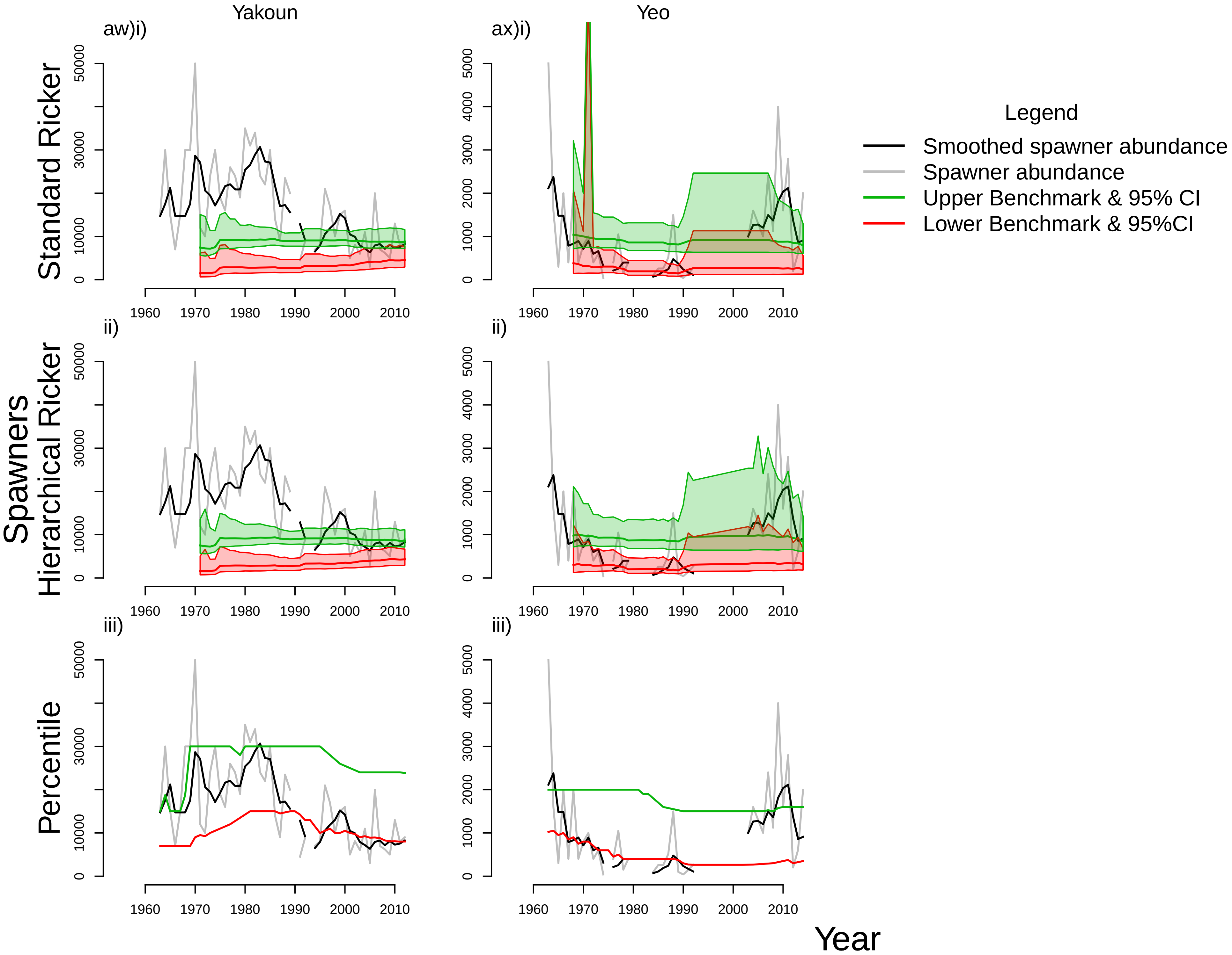


Figure 4: Status Over Time

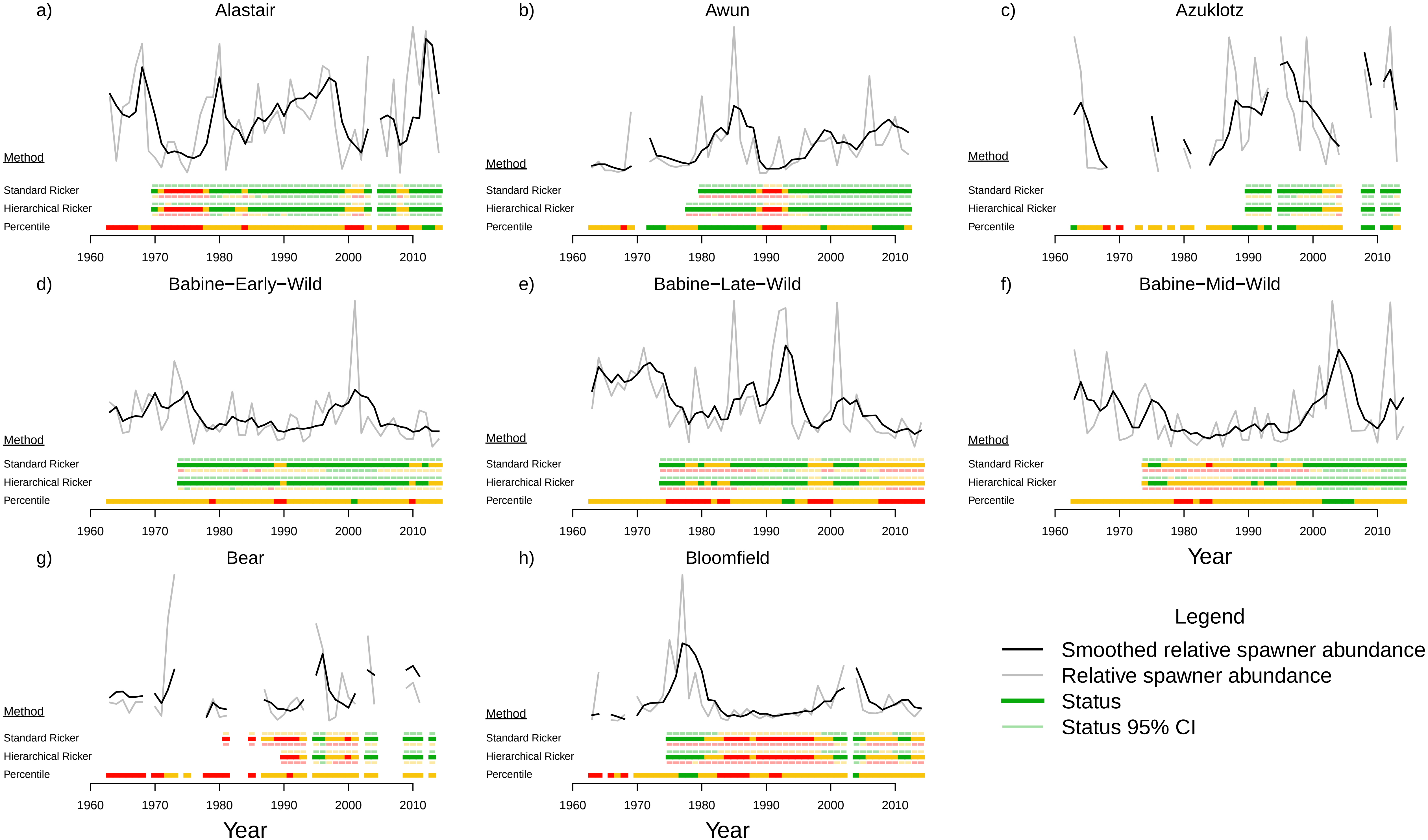
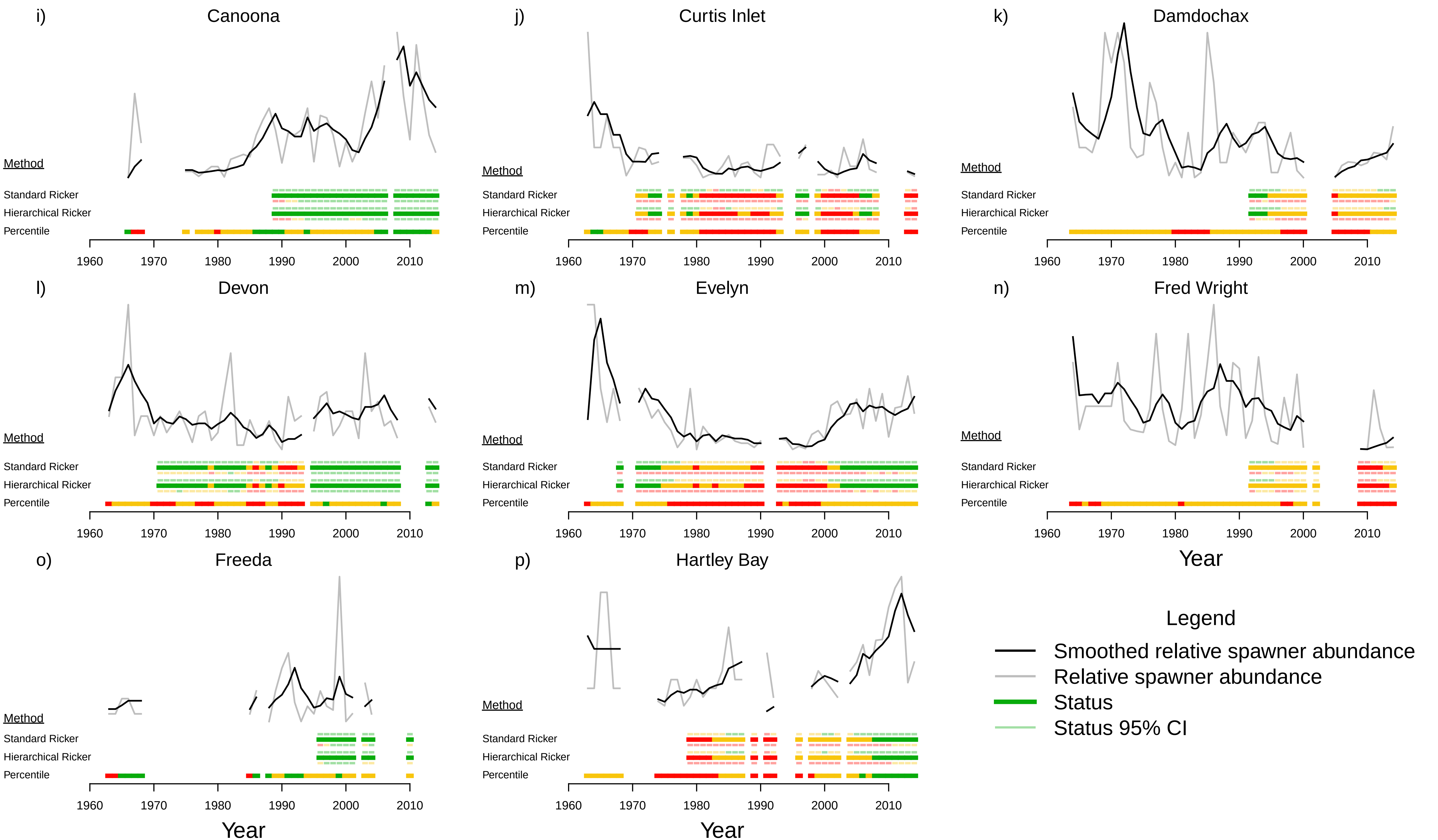


Figure 4: Status Over Time



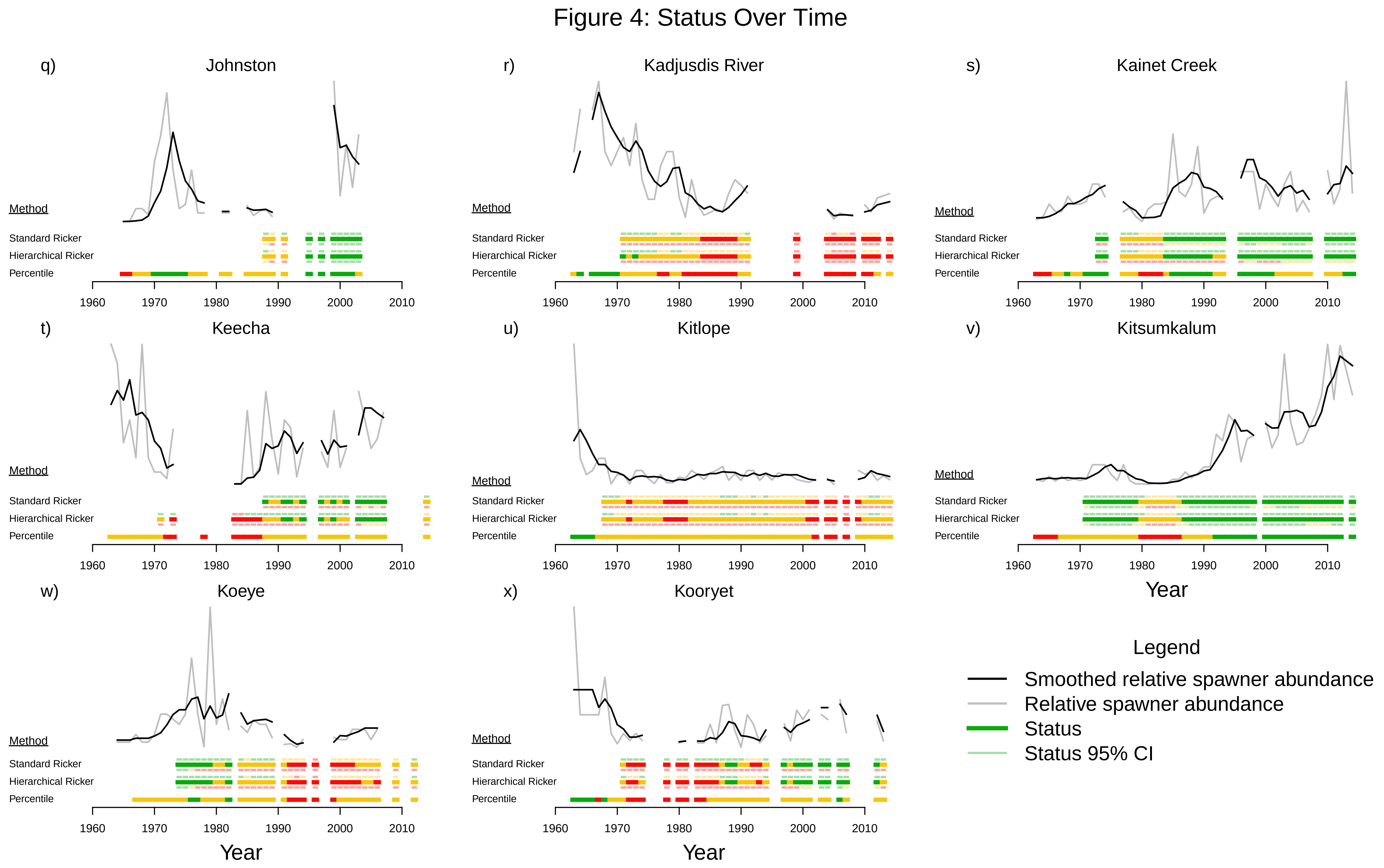


Figure 4: Status Over Time
q) Johnston
r) Kadjusdis River
s) Kainet Creek
t) Keecha
u) Kitlope
v) Kitsumkalum
w) Koeye
x) Kooryet
Method
Standard Ricker
Hierarchical Ricker
Percentile
1960
1970
1980
1990
2000
2010
Year
Legend
Smoothed relative spawner abundance
Relative spawner abundance
Status
Status 95% CI

Figure 4: Status Over Time

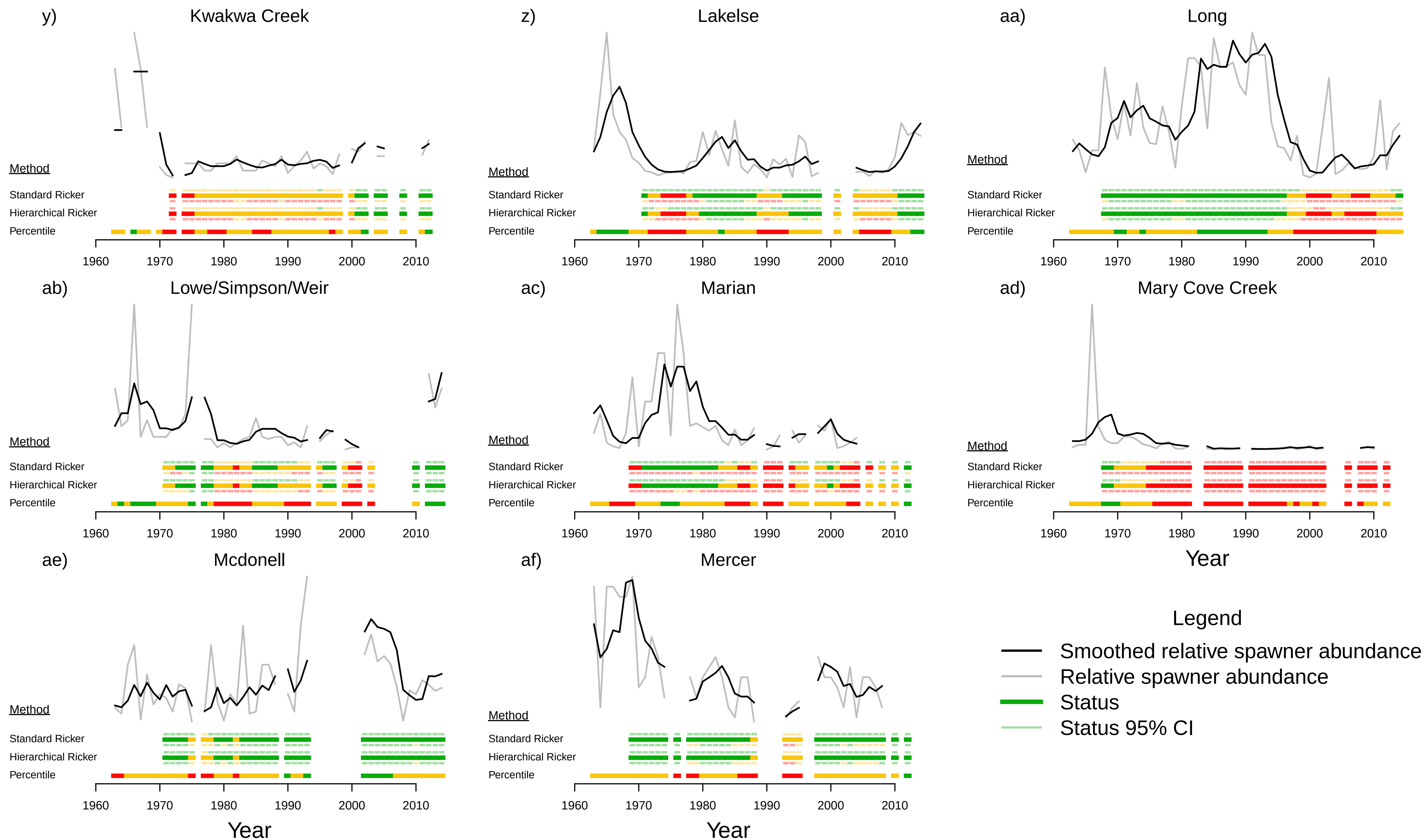


Figure 4: Status Over Time

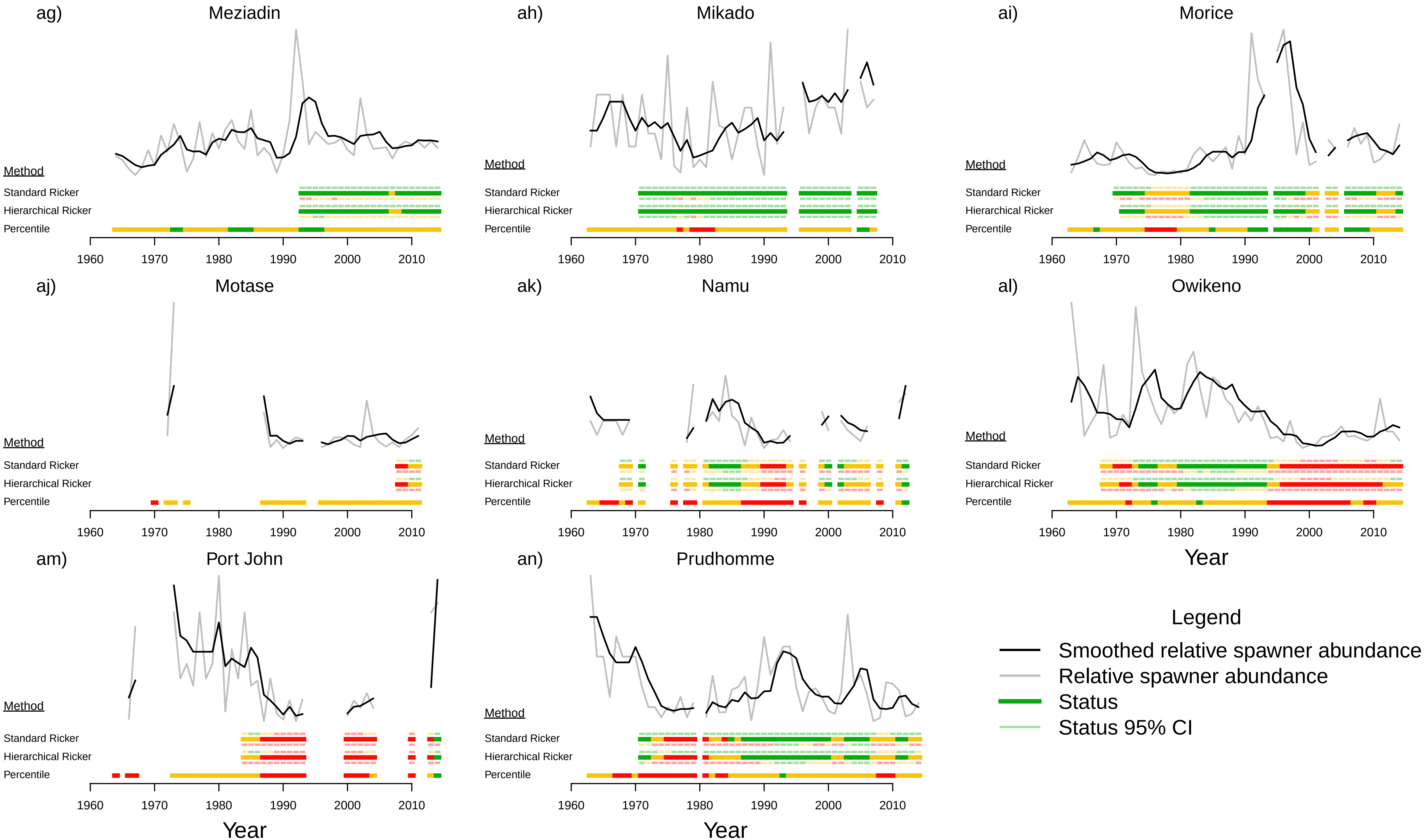


Figure 4: Status Over Time

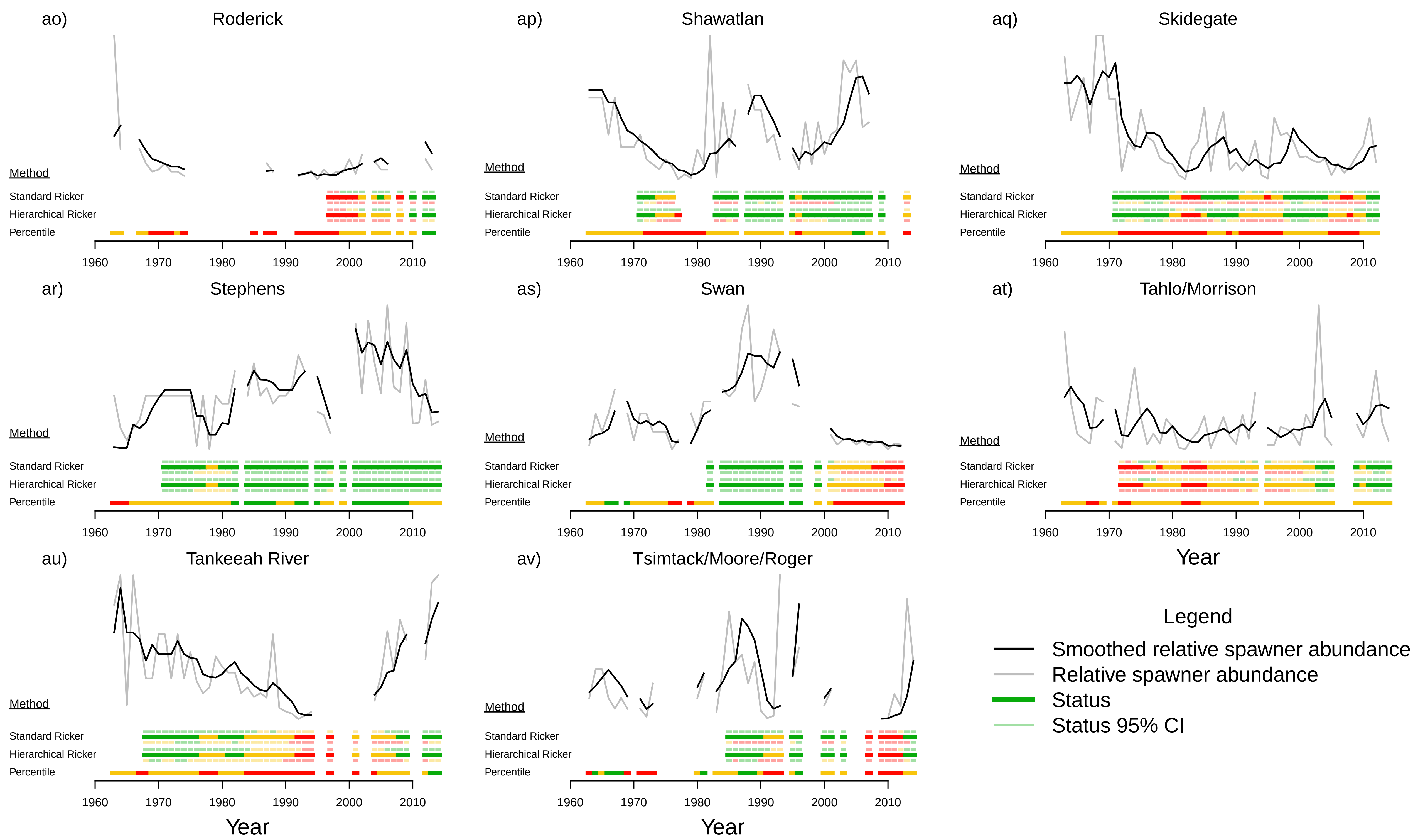


Figure 4: Status Over Time

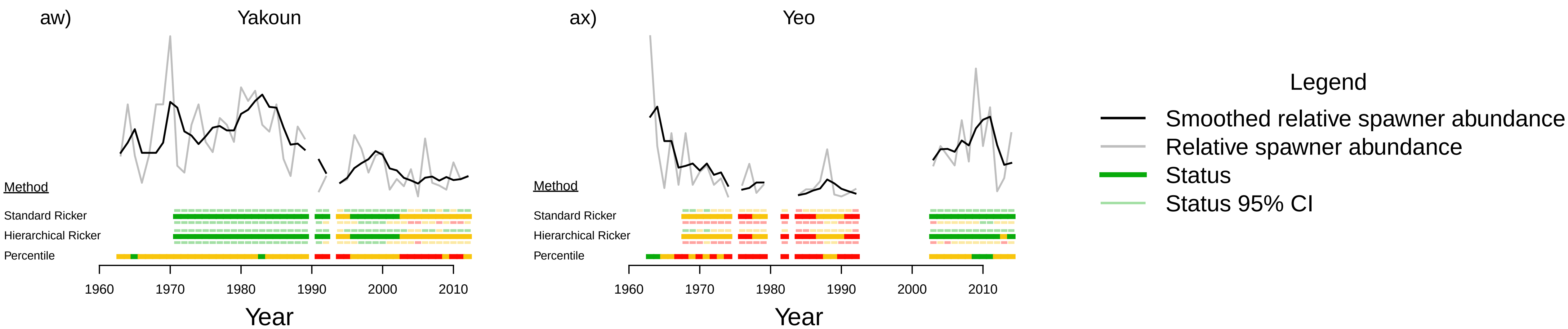


Figure 5: Ricker α Estimated by Recursive Bayes

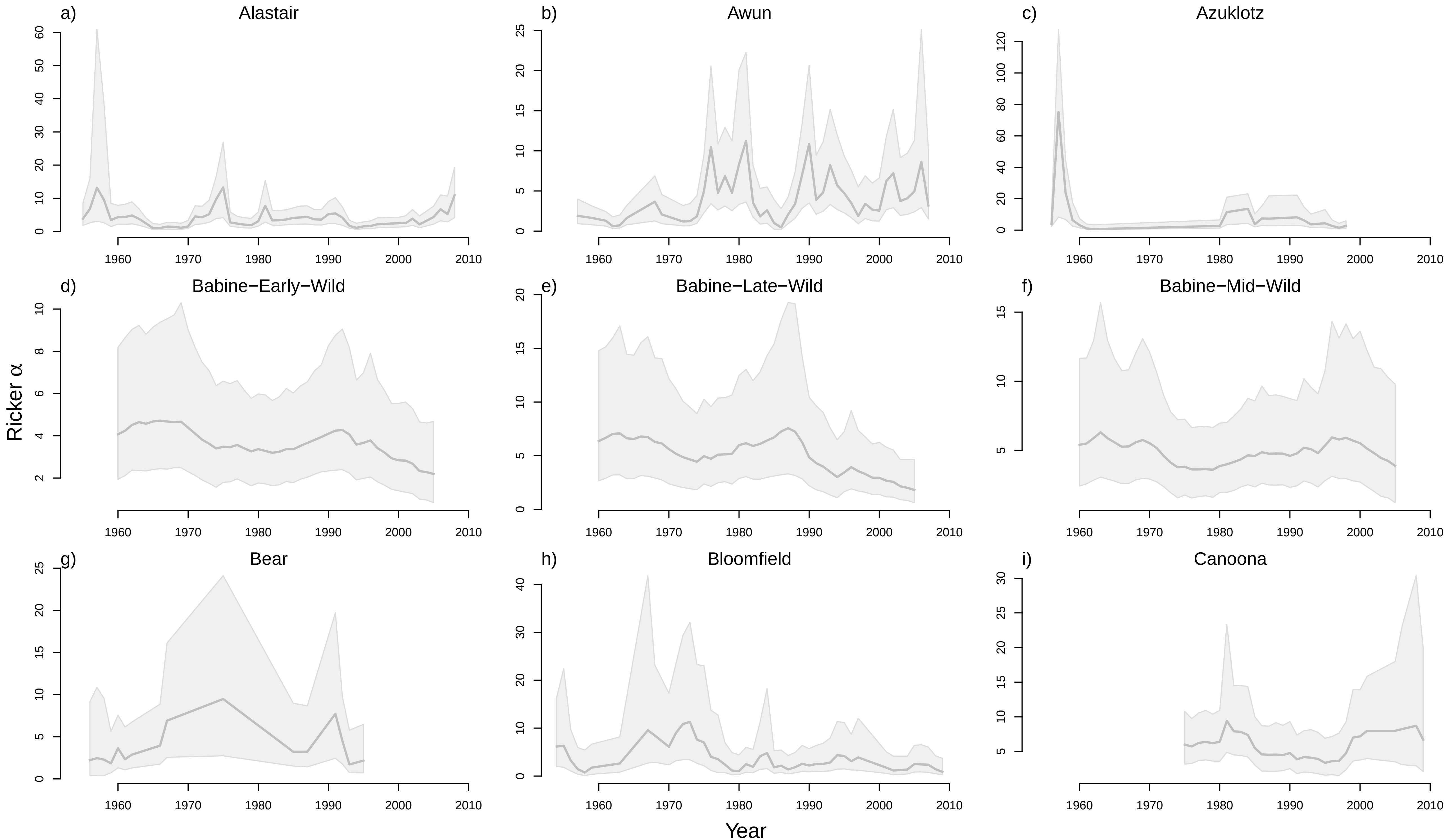


Figure 5: Ricker α Estimated by Recursive Bayes

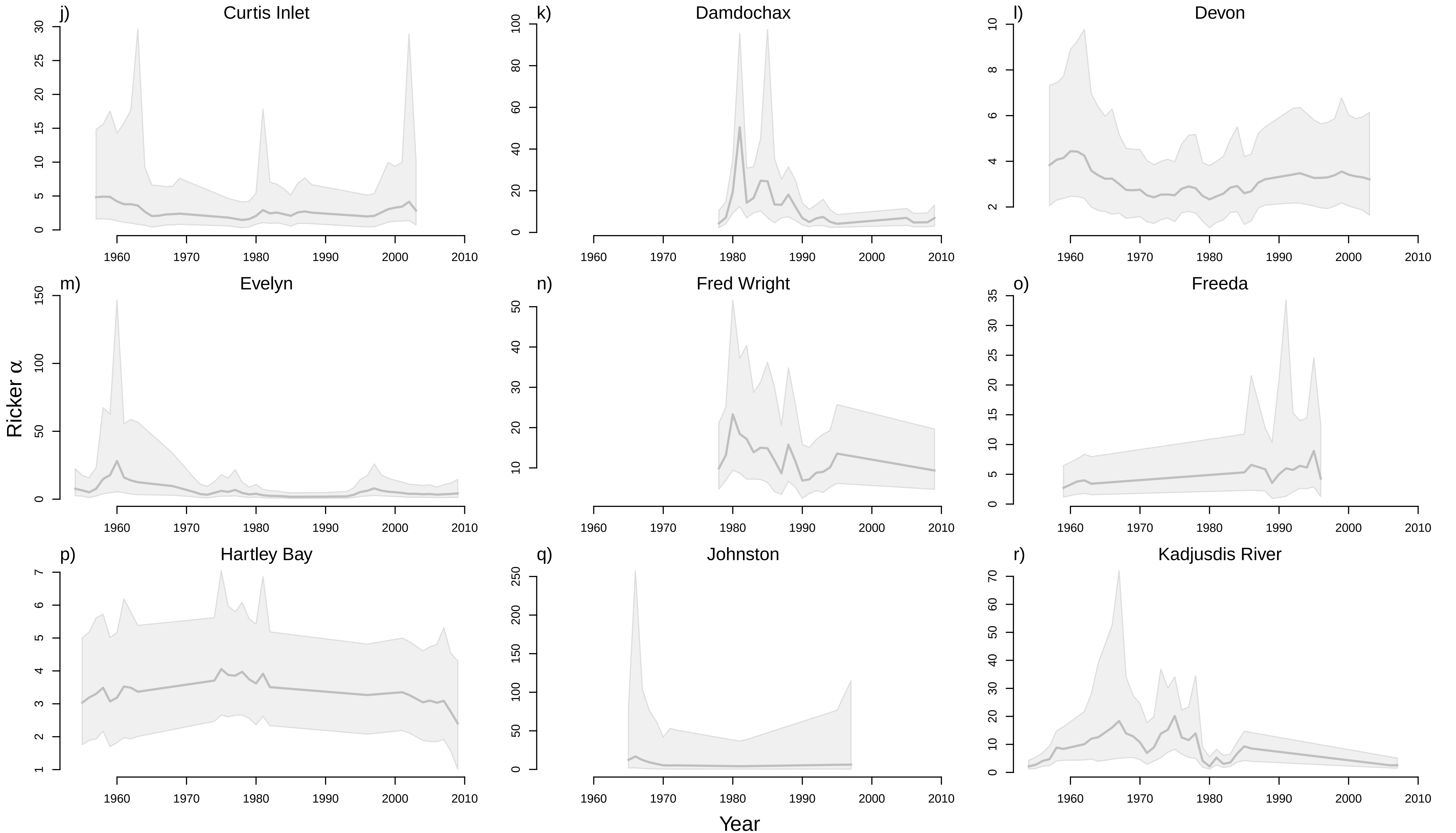


Figure 5: Ricker α Estimated by Recursive Bayes

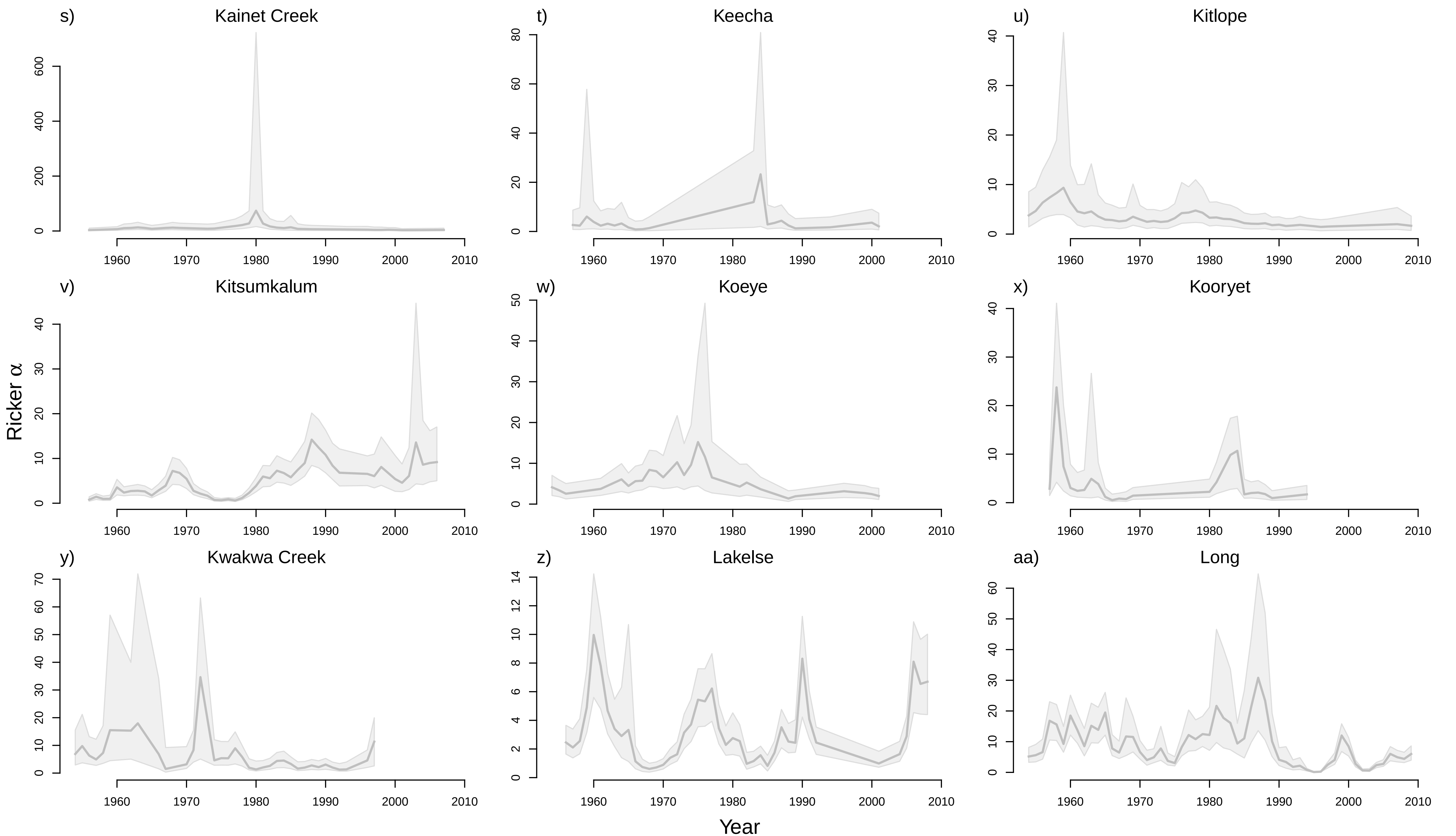


Figure 5: Ricker α Estimated by Recursive Bayes

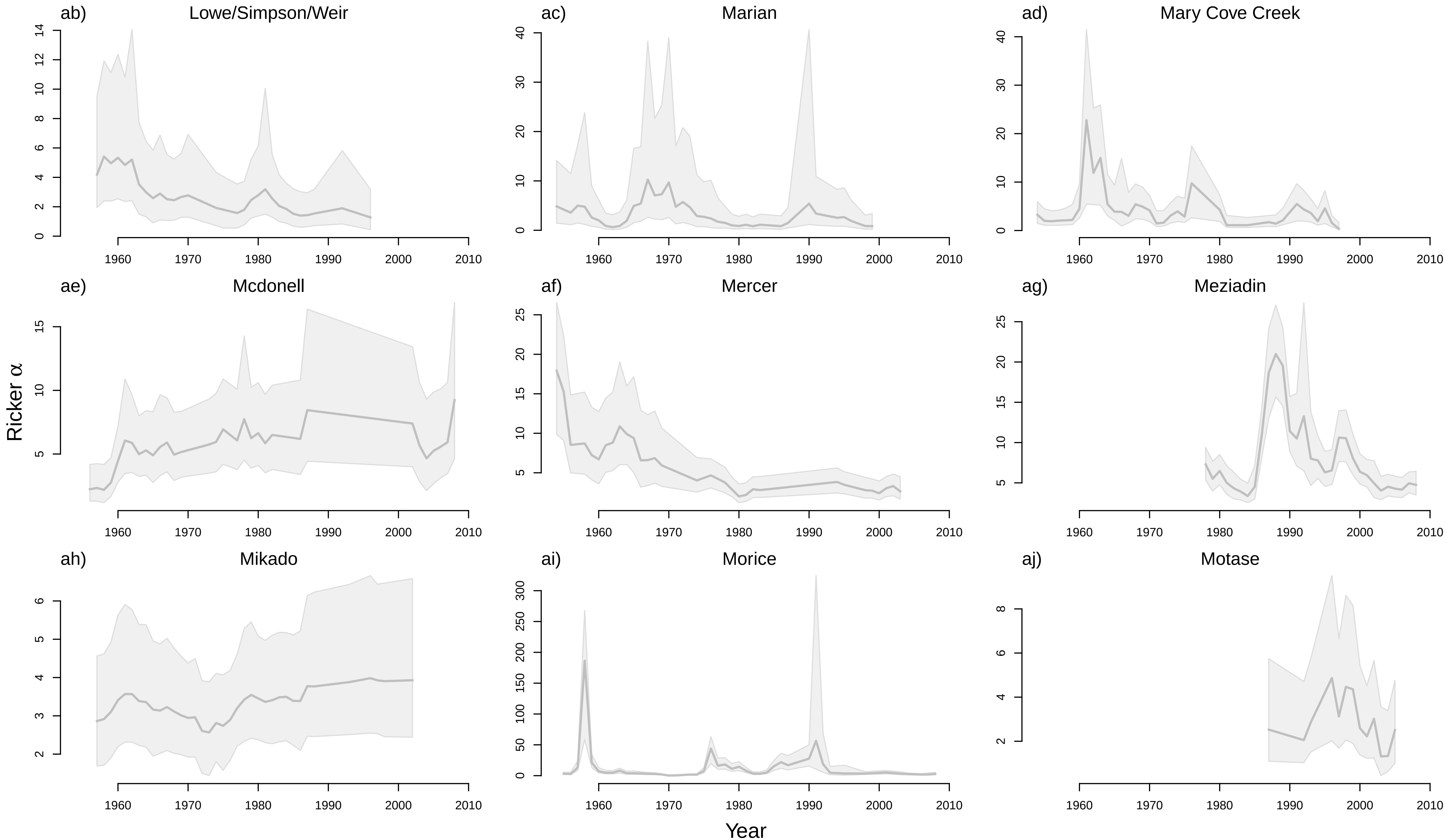


Figure 5: Ricker α Estimated by Recursive Bayes

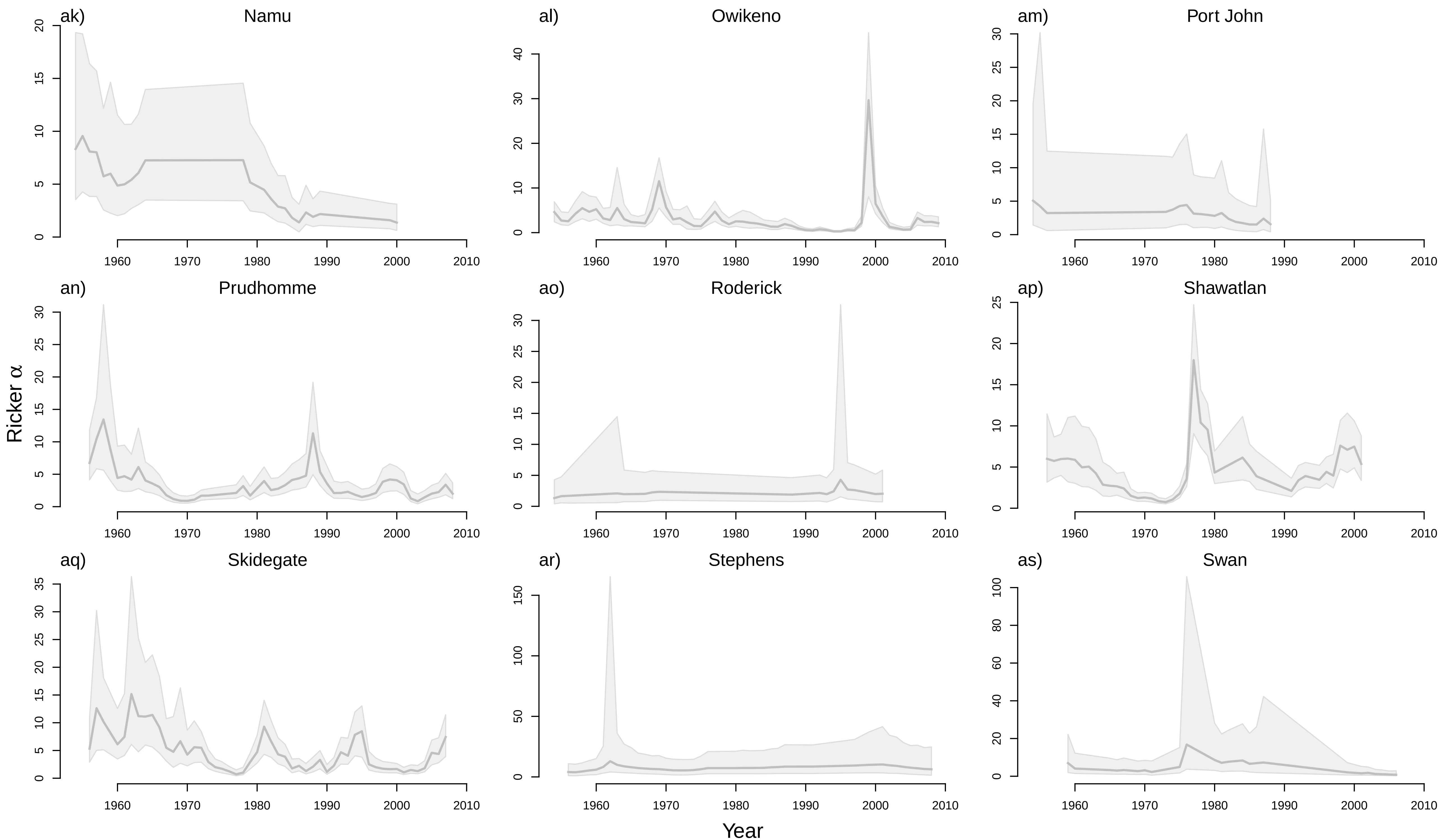


Figure 5: Ricker α Estimated by Recursive Bayes

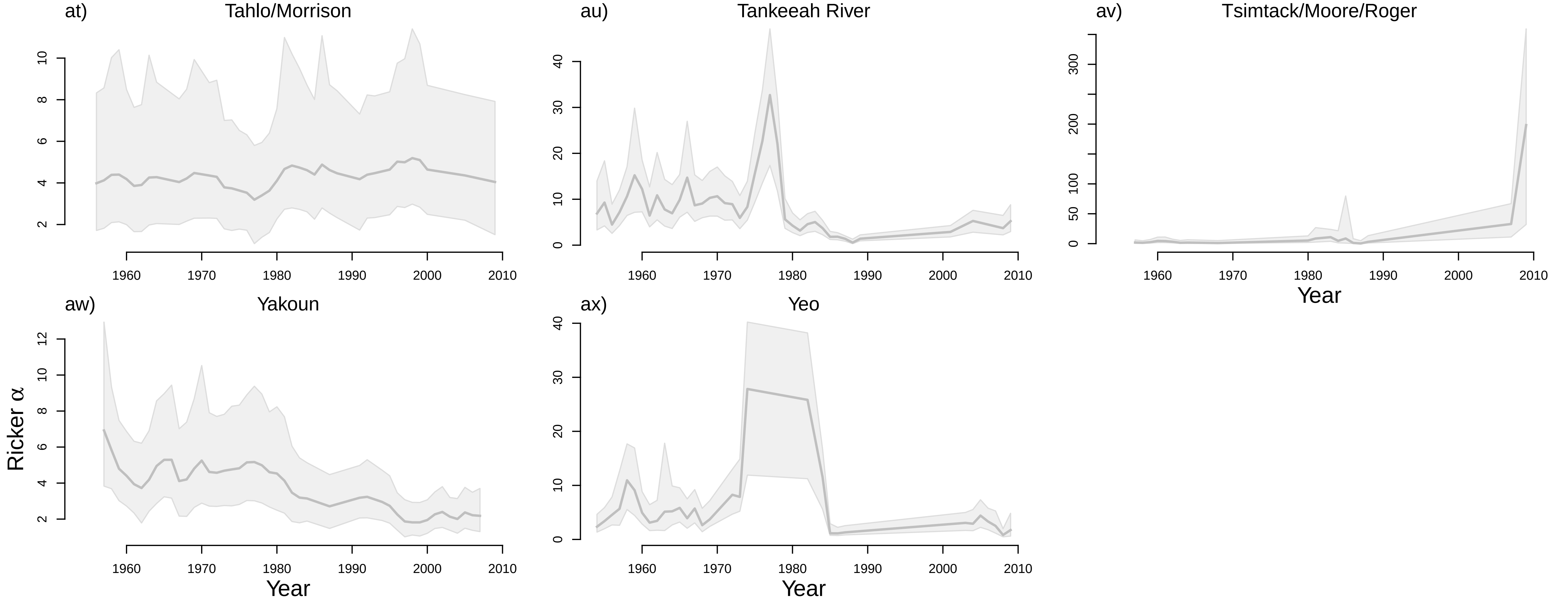


Fig.6. Schematic of simulation model used to evaluate benchmark performance.

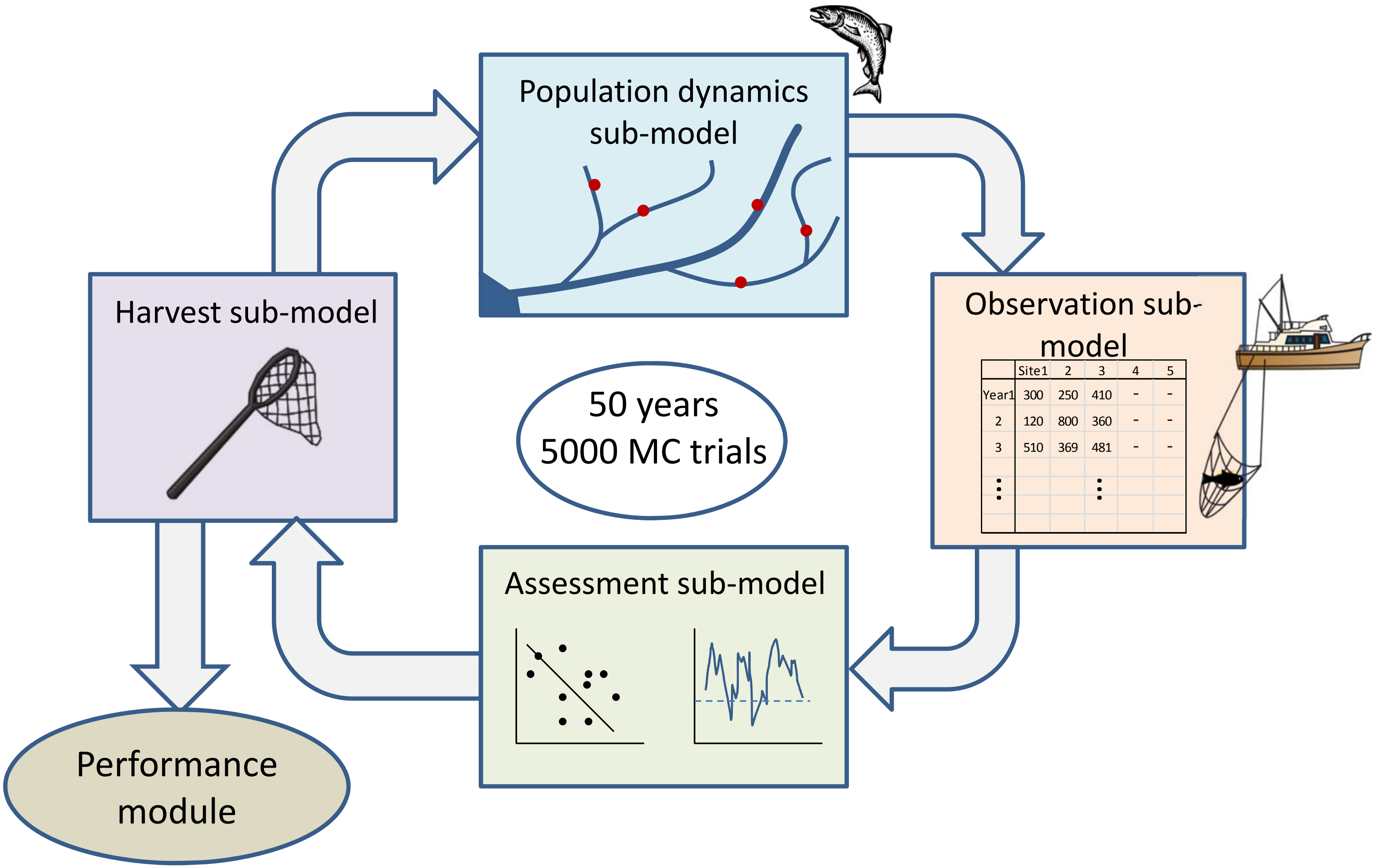


Fig 7(a). Time-series of observed spawner abundances (black line) and “true” spawner abundances (grey line) and benchmarks for one Monte Carlo trial. Benchmarks are estimated annually base on all data up until that year: annual estimates of 80% S_{MSY} (upper benchmark, green dashed line), annual estimates of S_{gen} (lower benchmark, red dashed line), the 75th percentile benchmark (green dotted line), and the 25th percentile (red dotted line). 95% confidence intervals are shown for estimates of stock-recruitment based benchmarks (green and red shading for the upper and lower benchmarks, respectively). (b) Observed spawner and recruitment data (solid black dots) and “true” data (grey hollow dots) for the final year of one Monte Carlo trial. The “true” underlying stock-recruitment relationship is shown with the grey curve and the estimate based on observed data is shown with the black curve. (c) Mean percent error, MPE, between estimated and “true” benchmark averaged over all Monte Carlo trials. Red bars are the mean percent error from the “true” S_{gen} (lower benchmark), and green bars the mean percent error from the “true” 80% S_{MSY}. Asterisk in panel (c) is MPE beyond the limit of the *y*-axis, 229%.

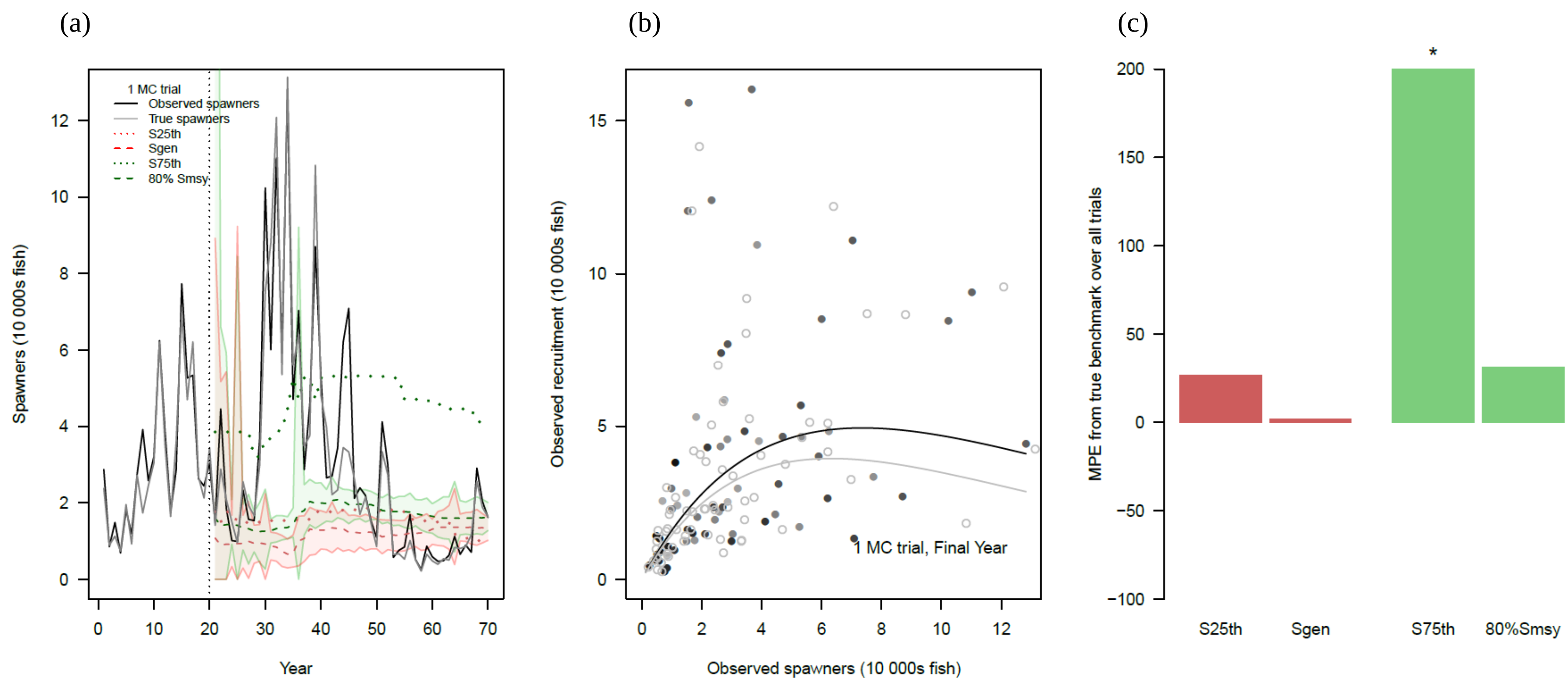


Fig. 8. Difference in Mean Percent Error, MPE, of estimated lower benchmark (S_{gen}, (a), and S_{25th} (b)) and the "true" lower benchmarks ("true" S_{gen}), between sensitivity analyses listed on the *x*-axis and the base case scenario. Black bars are analyses where the input parameter was increased relative to the base case (see Table 1); white bars are analyses where the input parameter was reduced relative to the base case. Positive values indicate sensitivity analyses where MPE increased under that change in input parameter from the base case; negative values indicate analyses where the MPE declined under that change in input parameter. Asterisks denote values above the limit of the *y*-axis, 156% (a) and 884% (b).

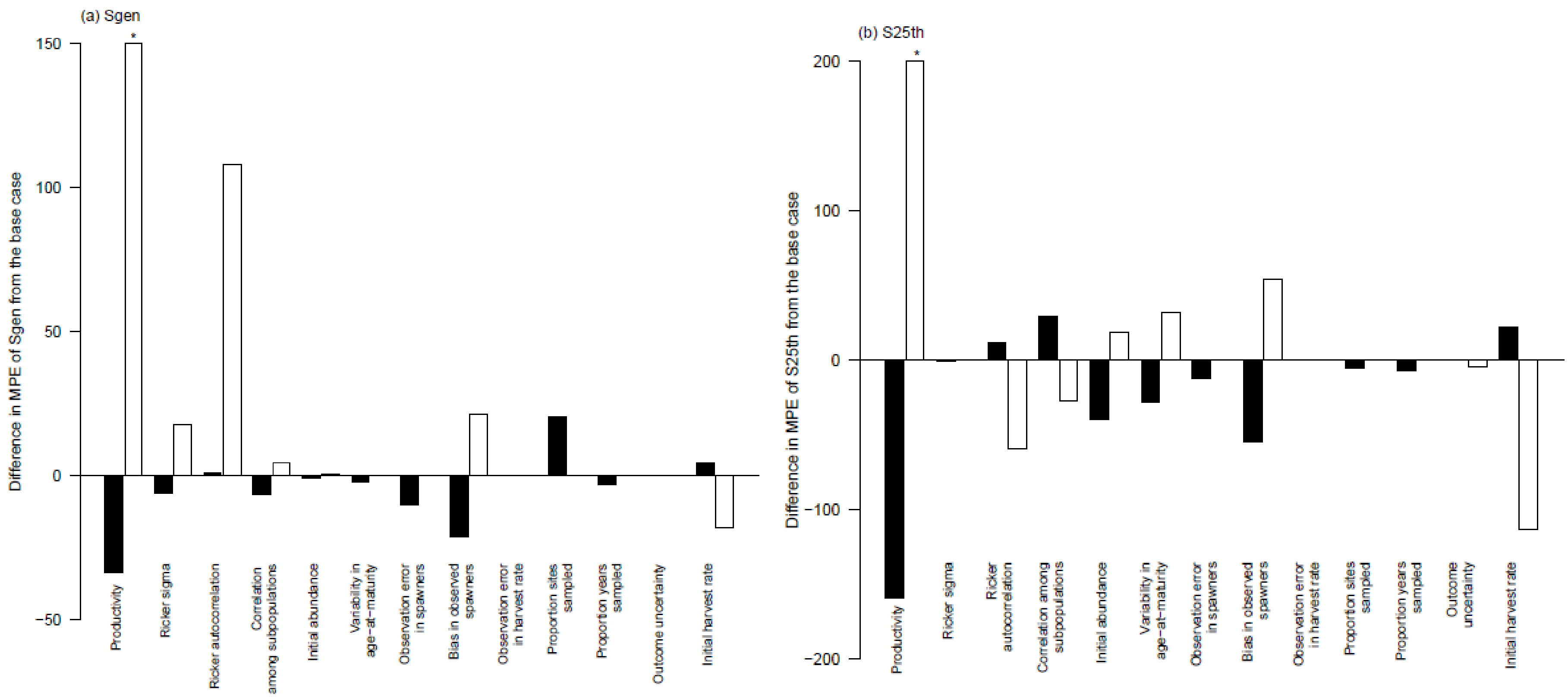


Fig. 9. Mean percent error, MPE, of the estimated lower benchmark (S_{25th} (a), and S_{gen} (b)) from the "true" S_{gen} value along a gradient in initial harvest rates (*x*-axis) and productivities (*y*-axis) derived from a simulation model of a hypothetical salmon CU. Dots indicate MPE of CUs assuming productivities estimated from hierarchical Ricker models and mean harvest rate over available time-series for each CU.

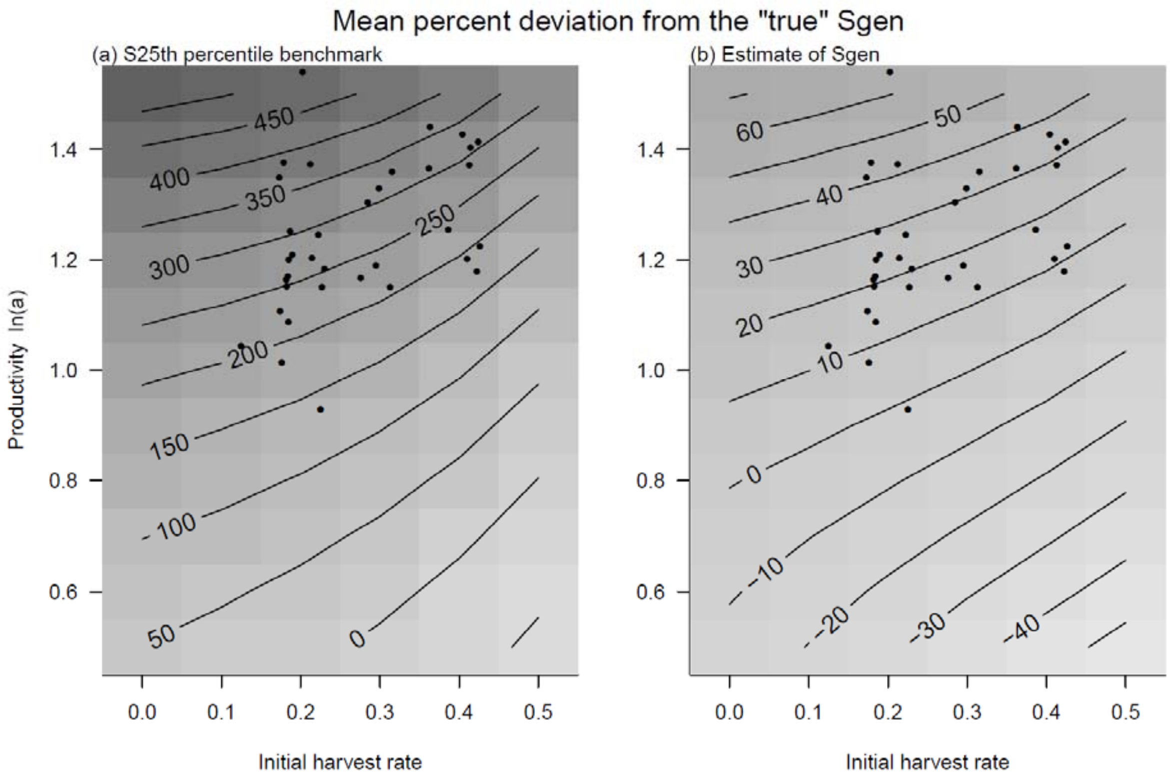


Fig. 10. Mean percent error of the estimated upper benchmark (S_{75th} percentile (a), and 80% S_{MSY} (b)) from the "true" 80% S_{MSY} value along a gradient in initial harvest rates (*x*-axis) and productivities (*y*-axis) derived from a simulation model of a hypothetical salmon CU. Dots indicate MPE of CUs assuming productivities estimated from hierarchical Ricker models and mean harvest rate over available time-series for each CU.

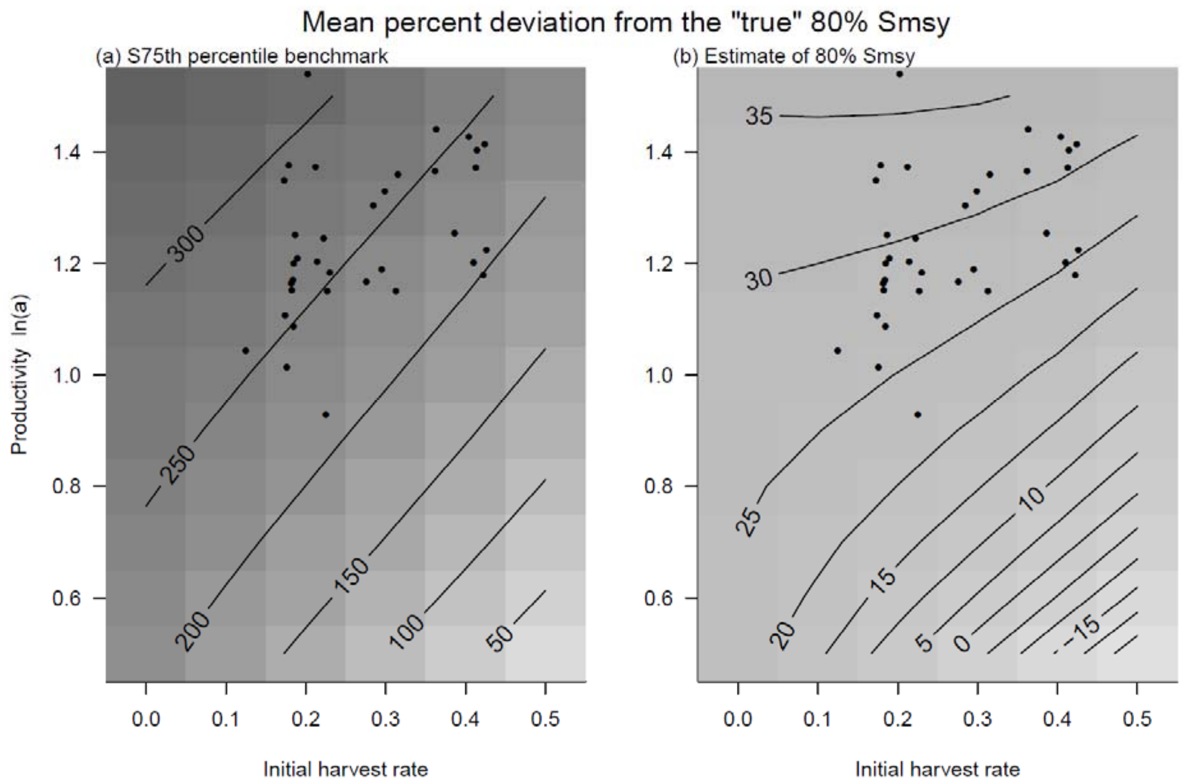


Fig. 11. Sensitivity indices of the effects of individual variables (mean elemental effect, *x*-axis) and interactions among variables (standard deviation in elemental effects, *y*-axis). Indices were derived from the Morris method, a global sensitivity analyses for the mean percent error of estimated lower benchmarks (S_{gen} (a), and S_{25th} (b)) from "true" benchmarks. Input variables with values >100 on either axis are labelled.

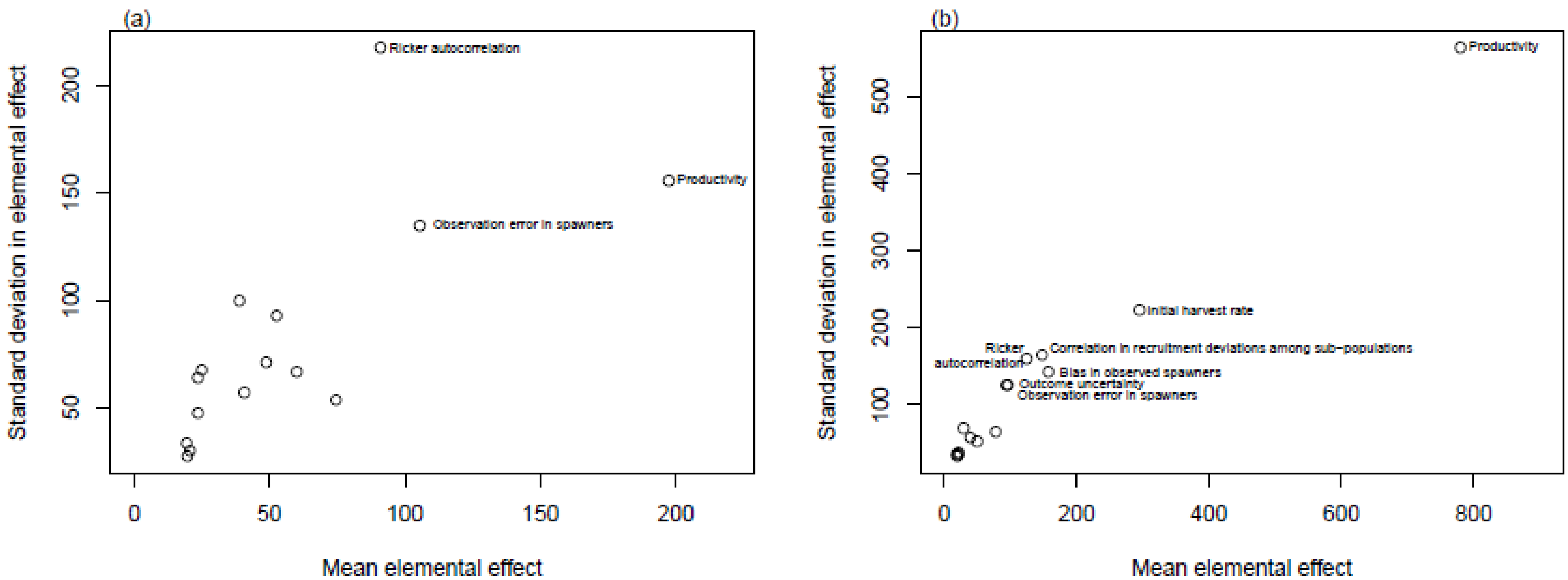


Fig. 12. “True” values of Ricker parameters and benchmarks under assumption of constant spawners at equilibrium, S_{eq}, (a-e) or constant spawner abundances at maximum recruitment S_{max} (f-j), with abrupt changes in productivity in year 35 of simulation.

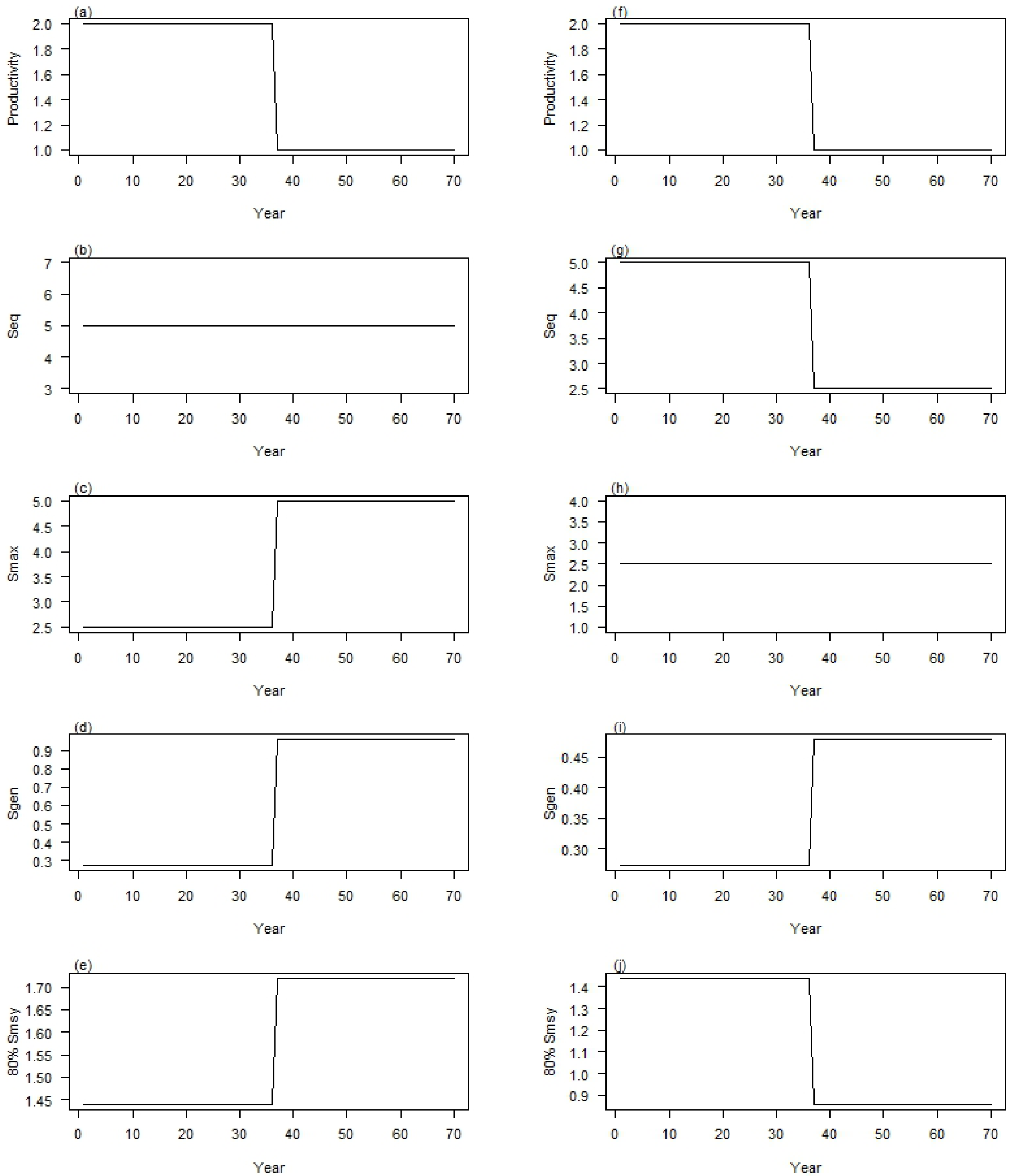


Fig. 13. Box plots of parameter and benchmark values in final year of the simulation averaged over all Monte Carlo trials, using either the first 30 years of data (dark grey boxes), all 70 years of data (light grey boxes), or the most recent 30 years of data (white boxes). Boxes represent the lower quartile, median, and upper quartiles of the parameter distribution. Whiskers are the 95% confidence intervals. Dashed lines represent the "true" value for each parameter. For the percentile benchmarks, the dashed lines represent the "true" S_{gen} (e) and "true" 80% of S_{MSY} (f). S_{eq} is held constant in simulations as productivity varies.

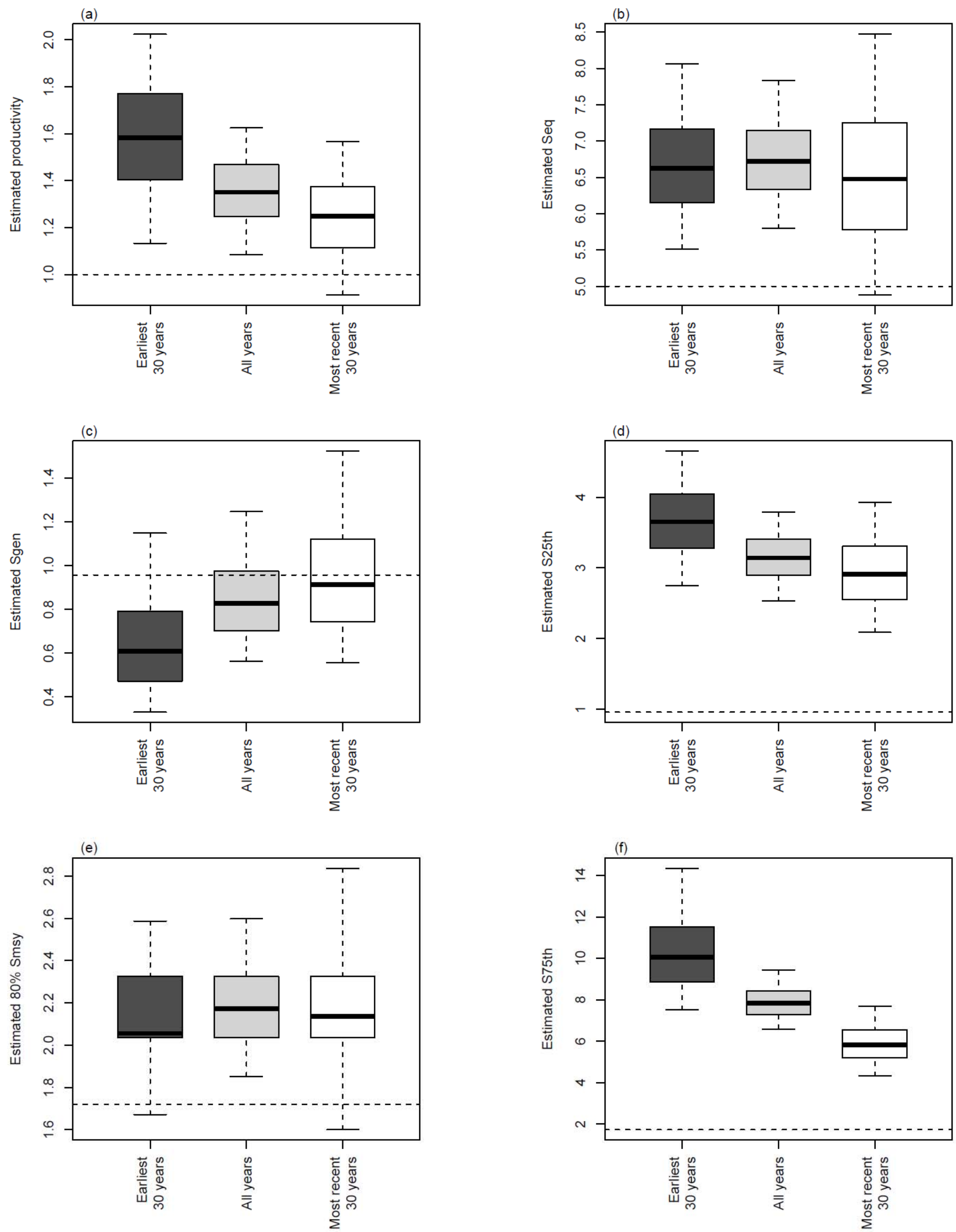


Fig. 14. Mean percent error of estimated Ricker parameters and lower and upper benchmarks from the “true” values, using only the first 30 years of data (black bars), all 70 years (grey bars), and the most recent 30 years (white bars) assuming either S_{eq} remains constant over time (a), or S_{max} remains constant (b).

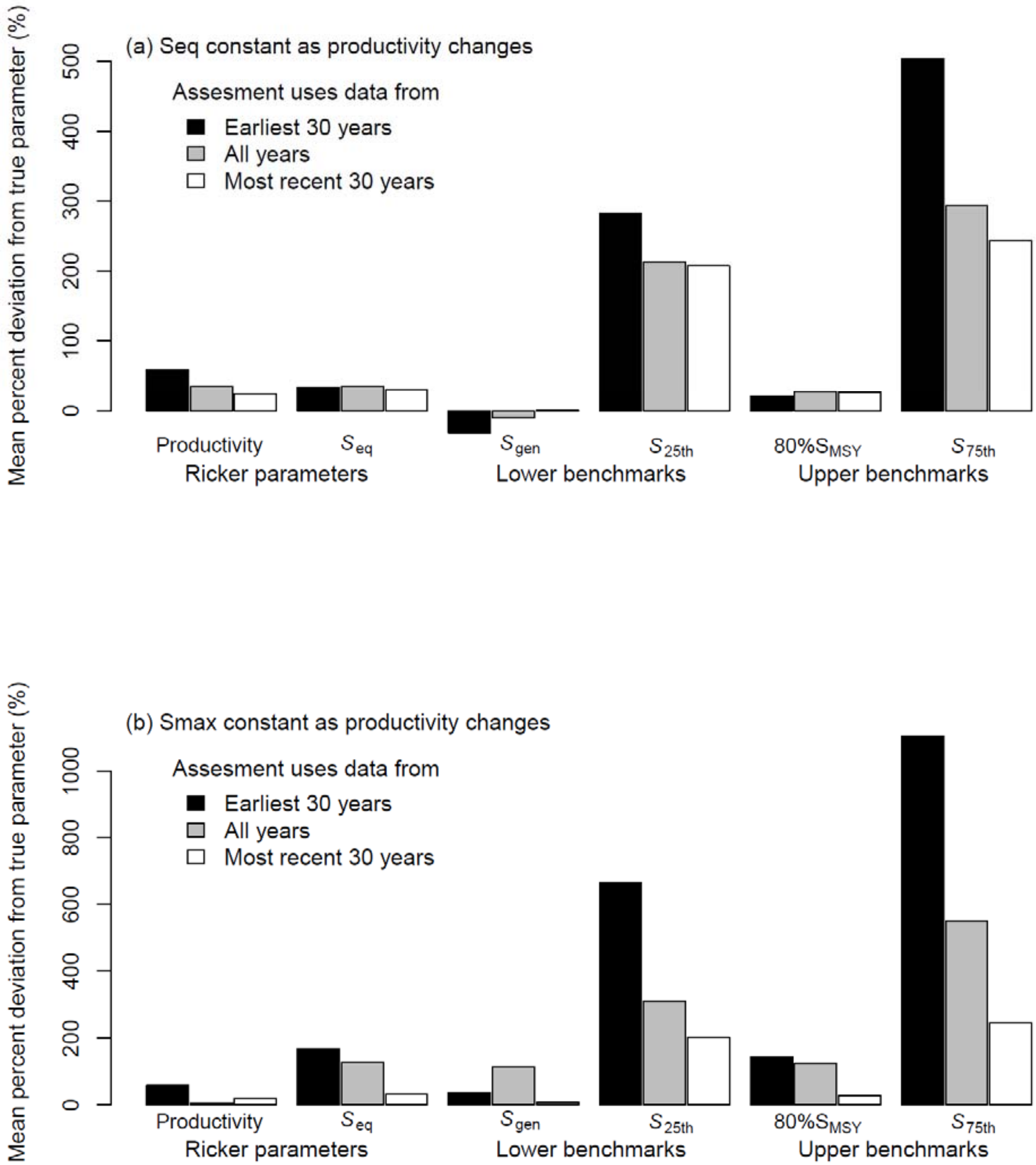
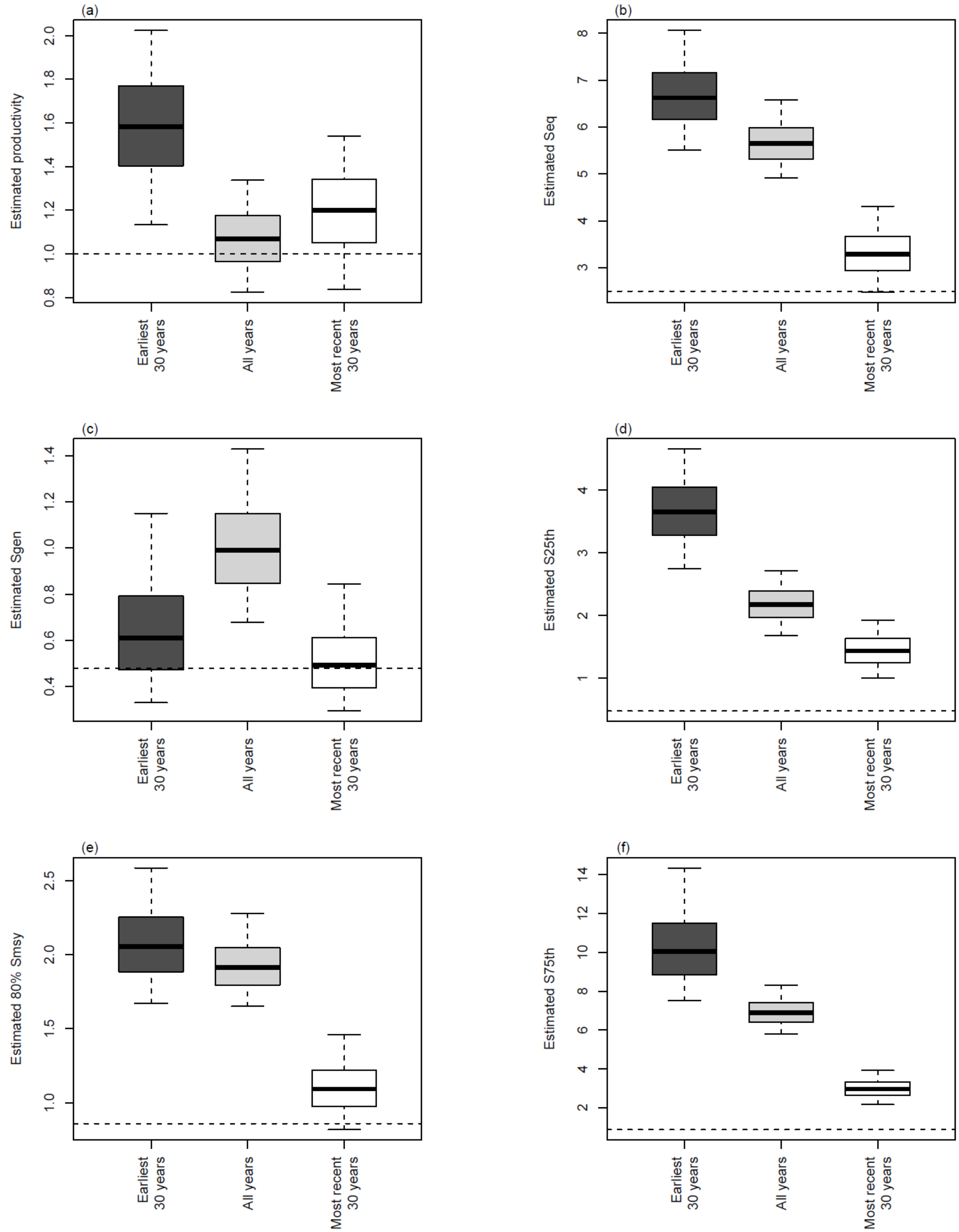


Fig. 15. Caption as for Fig. 13, except S_{max} is held constant and S_{eq} declines in simulations as productivity varies.



Tables

Table 1. CU groupings and data availability across CUs. *Some Capacity estimates were not used in the final models, because they were far outside the range of observed spawner levels.

Conservation Unit	Years of Data	Average Exploitation Rate	FAZ	Type	Max Age	S_{max} Estimate
Alastair	51	0.30	Lower Skeena	Coastal	6	23437
Awun	44	0.20	Queen Charlottes	Coastal	5	NA
Azuklotz	20	0.53	Upper Skeena	High Interior	6	5933
Babine-Early-Wild	46	0.45	Middle Skeena	Interior	5	NA
Babine-Late-Wild	46	0.62	Middle Skeena	Interior	5	NA
Babine-Mid-Wild	46	0.52	Middle Skeena	Interior	5	NA
Bear	16	0.53	Upper Skeena	High Interior	6	NA
Bloomfield	45	0.40	Hecate Lowlands	Coastal	5	NA
Canoona	31	0.37	N Coastal Streams	Coastal	5	NA
Curtis Inlet	32	0.19	Hecate Lowlands	Coastal	5	NA
Damdochax	23	0.70	Upper Nass	Interior	5	NA
Devon	43	0.19	Hecate Lowlands	Coastal	5	NA
Evelyn	44	0.39	N Coastal Streams	Coastal	5	NA
Fred Wright	19	0.73	Upper Nass	Interior	5	NA
Freeda	15	0.24	Hecate Lowlands	Coastal	5	NA
Hartley Bay	27	0.40	Hecate Lowlands	Coastal	5	NA
Johnston	12	0.19	Lower Skeena	Coastal	6	4125
Kadjusdis River	32	0.56	Hecate Lowlands	Coastal	5	NA
Kainet Creek	36	0.48	N Coastal Streams	Coastal	5	NA
Keecha	22	0.17	Hecate Lowlands	Coastal	5	NA
Kitlope	46	0.43	N Coastal Streams	Coastal	5	NA
Kitsumkalum	47	0.45	Lower Skeena	Coastal	6	NA
Koeye	27	0.44	Hecate Lowlands	Coastal	5	NA
Kooryet	23	0.17	Hecate Lowlands	Coastal	5	NA
Kwakwa Creek	35	0.44	Hecate Lowlands	Coastal	5	NA
Lakelse	43	0.17	Lower Skeena	Coastal	6	35916
Long	56	0.46	Rivers-Smith Inlets	Coastal	5	NA
Lowe/Simpson/Weir	30	0.18	Hecate Lowlands	Coastal	5	NA
Marian	38	0.20	Queen Charlottes	Coastal	5	NA
Mary Cove Creek	36	0.52	Hecate Lowlands	Coastal	5	NA
Mcdonell	32	0.20	Lower Skeena	Coastal	6	4072
Mercer	31	0.20	Queen Charlottes	Coastal	5	NA
Meziadin	31	0.68	Upper Nass	Interior	6	NA
Mikado	37	0.19	Hecate Lowlands	Coastal	5	NA
Morice	45	0.54	Middle Skeena	Interior	6	NA
Motase	13	0.50	Upper Skeena	High Interior	6	1764
Namu	24	0.45	Hecate Lowlands	Coastal	5	NA
Owikeno	56	0.25	Rivers-Smith Inlets	Coastal	5	NA
Port John	19	0.51	Hecate Lowlands	Coastal	5	NA
Prudhomme	49	0.19	Hecate Lowlands	Coastal	6	NA
Roderick	17	0.35	Hecate Lowlands	Coastal	5	NA
Shawatlan	38	0.20	Hecate Lowlands	Coastal	6	NA
Skidegate	51	0.20	Queen Charlottes	Coastal	5	NA
Stephens	39	0.36	Middle Skeena	Interior	6	7069
Swan	24	0.35	Middle Skeena	Interior	6	21432
Tahlo/Morrison	41	0.52	Middle Skeena	Interior	5	NA
Tankeeah River	40	0.53	Hecate Lowlands	Coastal	5	NA
Tsimtack/Moore/Roger	19	0.15	Hecate Lowlands	Coastal	5	NA
Yakoun	45	0.20	Queen Charlottes	Coastal	5	NA
Yeo	31	0.42	Hecate Lowlands	Coastal	5	NA

Table 2. Proportion of years where Ricker-based status and percentile-based status match, by CU and Ricker Model. Additionally, the proportion of years where the Percentile-based status match OR are more precautionary.

Conservation Unit	Basic Ricker match	Hierarchical match	Basic Ricker: match or more precautionary	Hierarchical Ricker: match or more precautionary
Alastair	0.20	0.23	1.00	1.00
Awun	0.61	0.57	1.00	1.00
Azuklotz	0.68	0.68	1.00	1.00
Babine-Early-Wild	0.10	0.07	1.00	1.00
Babine-Late-Wild	0.15	0.12	1.00	1.00
Babine-Mid-Wild	0.56	0.44	1.00	1.00
Bear	0.45	0.35	0.82	0.82
Bloomfield	0.59	0.59	0.82	0.82
Canoona	0.44	0.44	1.00	1.00
Curtis Inlet	0.74	0.63	1.00	1.00
Damdochax	0.37	0.37	1.00	1.00
Devon	0.21	0.15	1.00	1.00
Evelyn	0.28	0.33	0.95	0.95
Fred Wright	0.75	0.81	1.00	1.00
Freeda	0.11	0.11	1.00	1.00
Hartley Bay	0.86	0.86	0.97	0.97
Johnston	0.90	0.90	1.00	1.00
Kadjusdis River	0.77	0.71	0.94	0.94
Kainet Creek	0.56	0.61	1.00	1.00
Keecha	0.33	0.60	1.00	1.00
Kitlope	0.84	0.84	0.84	0.84
Kitsumkalum	0.50	0.50	1.00	1.00
Koeye	0.77	0.70	0.90	0.83
Kooryet	0.53	0.53	0.87	0.90
Kwakwa Creek	0.65	0.65	1.00	1.00
Lakelse	0.28	0.30	1.00	1.00
Long	0.53	0.60	1.00	1.00
Lowe/Simpson/Weir	0.36	0.36	1.00	1.00
Marian	0.57	0.59	0.89	0.92
Mary Cove Creek	0.76	0.76	0.76	0.76
Mcdonell	0.21	0.21	1.00	1.00
Mercer	0.03	0.03	1.00	1.00
Meziadin	0.23	0.27	1.00	1.00
Mikado	0.06	0.06	1.00	1.00
Morice	0.50	0.51	0.98	0.98
Motase	0.50	0.50	0.50	0.50
Namu	0.42	0.42	1.00	1.00
Owikeno	0.47	0.55	0.83	0.91
Port John	0.89	0.89	0.89	0.89
Prudhomme	0.35	0.35	1.00	1.00
Roderick	0.54	0.69	0.69	0.77
Shawatlan	0.06	0.09	1.00	1.00
Skidegate	0.17	0.12	1.00	1.00
Stephens	0.50	0.50	1.00	1.00
Swan	0.69	0.62	1.00	1.00
Tahlo/Morrison	0.69	0.72	0.90	0.92
Tankeeah River	0.24	0.26	1.00	1.00
Tsimtack/Moore/Roger	0.43	0.43	1.00	1.00
Yakoun	0.08	0.08	1.00	1.00
Yeo	0.48	0.52	1.00	1.00

Table 3 - Estimated status based on benchmarks using all available data, compared to the final year of escapement data, as listed below.

Conservation Unit	Year	Percentile Status	Basic Ricker Status	Hierarchical Ricker Status
Alastair	2014	Amber	Green	Green
Awun	2012	Amber	Green	Green
Azuklotz	2013	Amber	Green	Green
Babine-Early-Wild	2014	Amber	Amber	Amber
Babine-Late-Wild	2014	Red	Amber	Amber
Babine-Mid-Wild	2014	Amber	Green	Green
Bear	2013	Amber	Green	Green
Bloomfield	2014	Amber	Amber	Amber
Canoona	2014	Amber	Green	Green
Curtis Inlet	2014	Red	Red	Red
Damdochax	2014	Amber	Amber	Amber
Devon	2014	Amber	Green	Green
Evelyn	2014	Amber	Green	Green
Fred Wright	2014	Red	Amber	Amber
Freeda	2010	Amber	Green	Green
Hartley Bay	2014	Green	Green	Green
Johnston	2003	Amber	Green	Green
Kadjusdis River	2014	Amber	Red	Red
Kainet Creek	2014	Green	Green	Green
Keecha	2014	Amber	Amber	Amber
Kitlope	2014	Amber	Amber	Amber
Kitsumkalum	2014	Green	Green	Green
Koeye	2012	Amber	Amber	Amber
Kooryet	2013	Amber	Amber	Amber
Kwakwa Creek	2012	Green	Green	Green
Lakelse	2014	Green	Green	Green
Long	2014	Amber	Green	Amber
Lowe/Simpson/Weir	2014	Green	Green	Green
Marian	2012	Green	Green	Green
Mary Cove Creek	2012	Amber	Red	Red
Mcdonell	2014	Amber	Green	Green
Mercer	2012	Green	Green	Green
Meziadin	2014	Amber	Green	Green
Mikado	2007	Amber	Green	Green
Morice	2014	Amber	Green	Green
Motase	2011	Amber	Amber	Amber
Namu	2012	Green	Green	Green
Owikeno	2014	Amber	Red	Amber
Port John	2014	Green	Green	Green
Prudhomme	2014	Amber	Amber	Amber
Roderick	2013	Green	Green	Green
Shawatlan	2013	Red	Amber	Amber
Skidegate	2012	Amber	Green	Green
Stephens	2014	Amber	Green	Green
Swan	2012	Red	Red	Red
Tahlo/Morrison	2014	Amber	Green	Green
Tankeeah River	2014	Green	Green	Green
Tsimtack/Moore/Roger	2014	Amber	Green	Green
Yakoun	2012	Amber	Amber	Amber
Yeo	2014	Amber	Green	Green

Table 4. Parameters used as base case, univariate sensitivity analyses, and global sensitivity analyses of simulation model to evaluate lower benchmarks.

Sub-model	**Parameter**	**Base-case Value**	**Values considered in univariate sensitivity analyses**	**Range considered in global sensitivity analyses**
Population dynamics sub-model	Ricker productivity parameter	1	0.5 (low) and 2.0 (high)	0.5-2.0
	Ricker autocorrelation coefficient	0.6	0 (low) and 0.9 (high)	0-1.0
	Standard deviation in Ricker residuals	0.75	0.6 (low) and 1.0 (high)	0.6-1.0
	Average proportions at age-of-maturity	Age 3= 8% Age 4=42% Age 5=50%		
	Variability in age-at-maturity, ϖ, specified in multivariate logistic distribution	0.6	0.1 (low) and 0.9 (high)	0.1-0.9
	Correlation in recruitment residuals among subpopulations within a CU	0.4	0 (low) and 1.0 (high)	0-1.0
	Initial spawner abundances	$0.2\times S_{eq}$, where S_{eq} is spawner abundances at equilibrium	$0.1\times S_{eq}$ (low) and $0.3\times S_{eq}$ (high)	$0.1\times S_{eq}$-$0.3\times S_{eq}$
	Stray rate	0.01		
Observation sub-model	Standard deviation in observation errors of spawner abundances	0.5	0.2 (low)	0-1.0
	Standard deviation in normally distributed observation errors of harvest rates (constrained to between 0 and 1)	0.03	0 (low) and 0.1 (high)	0-1.0
	Multiplicative bias in observed spawner abundances not accounted for in assessment	1	0.8 (negative bias) and 1.2 (positive bias)	0.8-1.2
Assessment sub-model	Proportion of subpopulations sampled within a CU	100%	50% (low)	50%-100%
	Proportion of years that CU is sampled	100%	60% (low)	60%-100%
Harvest sub-model	Harvest rate during initialization period	20%	10% (low) and 50% (high)	10%-60%
	Outcome uncertainty (standard deviation in differences between target and realized harvest rates)	0.3	0.5 (high)	0-0.9

Appendix

Table A.1. Benchmark values across three methods used: standard Ricker model, hierarchical Ricker model (S_{gen} and 80% S_{MSY}) and percentiles (25th and 75th).

	Standard Ricker	Hierarchical Ricker	Percentile	Standard Ricker	Hierarchical Ricker	Percentile
	Alastair			Awun		
Upper Benchmark	9731	9759	25214	2964	2977	8000
Lower Benchmark	4936	4753	8685	1370	1357	2000
	Azuklotz			Babine-Early-Wild		
Upper Benchmark	2585	2545	4750	29906	29008	66515
Lower Benchmark	802	802	900	12237	10442	27816
	Babine-Late-Wild			Babine-Mid-Wild		
Upper Benchmark	170894	160797	303079	11470	11417	30029
Lower Benchmark	57635	47117	124230	3562	3302	8223
	Bear			Bloomfield		
Upper Benchmark	2935	2945	4350	992	992	1575
Lower Benchmark	1136	1113	926	492	470	403
	Canoona			Curtis Inlet		
Upper Benchmark	1820	1921	5150	6996	7034	15000
Lower Benchmark	458	605	1400	3952	3516	3150
	Damdochax			Devon		
Upper Benchmark	3554	3680	5557	3463	3462	8000
Lower Benchmark	647	701	2000	1751	1678	3000
	Evelyn			Fred Wright		
Upper Benchmark	1282	1326	2400	6730	7398	10273
Lower Benchmark	492	560	600	761	1005	2500
	Freeda			Hartley Bay		
Upper Benchmark	340	352	800	1183	1200	1650
Lower Benchmark	112	138	400	489	508	800
	Johnston			Kadjusdis River		
Upper Benchmark	2322	1995	6000	4260	5110	4000
Lower Benchmark	413	750	1000	1476	1972	955
	Kainet Creek			Keecha		
Upper Benchmark	1023	1383	2700	3331	3375	7000
Lower Benchmark	142	411	800	1796	1620	2000
	Kitlope			Kitsumkalum		
Upper Benchmark	36797	36604	40000	9419	9011	16360
Lower Benchmark	19025	18173	16000	4864	4350	2426
	Koeye			Kooryet		
Upper Benchmark	7603	8255	5850	3375	3401	7000
Lower Benchmark	3243	3610	1875	1936	1698	1460
	Kwakwa Creek			Lakelse		
Upper Benchmark	2792	2836	4000	8645	8693	20706
Lower Benchmark	1125	1201	1525	4735	4474	6530

	Standard Ricker	Hierarchical Ricker	Percentile	Standard Ricker	Hierarchical Ricker	Percentile
	Long			Lowe/Simpson/Weir		
Upper Benchmark	65575	70898	139642	6082	6352	13500
Lower Benchmark	20713	26833	26827	3576	3362	4550
	Marian			Mary Cove Creek		
Upper Benchmark	7393	7679	14000	6143	6077	2000
Lower Benchmark	4635	3994	3500	3039	2926	350
	Mcdonell			Mercer		
Upper Benchmark	1545	1562	6000	3451	3494	10000
Lower Benchmark	445	532	1500	1347	1457	3450
	Meziadin			Mikado		
Upper Benchmark	137790	138152	181649	1947	1941	7000
Lower Benchmark	29388	29998	115705	906	894	3000
	Morice			Motase		
Upper Benchmark	8593	8286	10897	414	418	596
Lower Benchmark	2488	2220	3632	231	226	160
	Namu			Owikeno		
Upper Benchmark	2686	2729	4000	422542	381728	646872
Lower Benchmark	1256	1246	1500	329015	244892	185785
	Port John			Prudhomme		
Upper Benchmark	923	880	1500	3136	3150	9416
Lower Benchmark	536	436	239	1953	1748	2445
	Roderick			Shawatlan		
Upper Benchmark	780	822	1000	1905	1888	7000
Lower Benchmark	509	440	400	914	872	2000
	Skidegate			Stephens		
Upper Benchmark	10849	10905	30900	3402	3439	8947
Lower Benchmark	6735	6079	10000	719	747	4250
	Swan			Tahlo/Morrison		
Upper Benchmark	11045	11633	26540	9474	9250	17884
Lower Benchmark	4957	4179	4985	3169	2891	4685
	Tankeeah River			Tsimtack/Moore/Roger		
Upper Benchmark	3398	3665	6000	1816	1879	7000
Lower Benchmark	1131	1442	2000	508	762	1500
	Yakoun			Yeo		
Upper Benchmark	8685	8626	23875	797	837	1600
Lower Benchmark	4535	4311	8025	244	313	350

Table A.2. Parameter and benchmark estimates and upper/lower credible interval bounds delineated as 2.5th and 97.5th posterior densities for most recent year of model estimation for each CU.

	Alastair				Awun				Azuklotz			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.98	4.00	3.09	3.91	3.24	4.75	3.30	4.48	4.73	8.75	4.72	8.91
		2.22		2.39		2.19		2.43		2.60		2.65
S_{max}	26,432	36,057	25,779	33,625	7,606	11,479	7,528	10,637	5,439	9,592	5,375	9,315
		20,647		20,689		5,687		5,812		3,305		3,284
S_{gen}	4,936	6,934	4,770	6,422	1,370	2,200	1,353	2,032	802	1,622	802	1,589
		3,512		3,546		887		931		352		337
80% S_{MSY}	9,731	12,259	9,745	12,113	2,964	3,926	2,987	3,892	2,585	4,332	2,545	4,305
		7,898		8,069		2,320		2,377		1,675		1,678

	Babine-Early-Wild				Babine-Late-Wild				Babine-Mid-Wild			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.64	5.43	4.11	5.96	4.36	7.22	4.93	7.65	4.68	6.87	4.97	7.11
		2.40		2.76	NA	2.64		3.14		3.12		3.50
S_{max}	70,971	124,097	64,418	103,537	366,314	1,052,898	325,342	679,750	23,790	36,924	22,910	33,497
		50,319		47,215		224,212		216,826		17,708		17,482
S_{gen}	12,237	23,900	10,442	19,460	57,635	195,682	47,117	123,007	3,562	6,548	3,302	5,715
		7,070		6,124		25,719		23,534		2,123		2,007
80% S_{MSY}	29,906	41,564	29,008	39,098	170,894	377,800	160,797	278,457	11,470	15,426	11,417	14,844
		24,044		23,787		122,433		119,608		9,415		9,365

	Bear				Bloomfield				Canoona			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.83	6.48	3.91	6.60	3.05	4.68	3.16	4.39	5.71	7.95	4.65	6.51
		2.22		2.32		1.93		2.26		3.92		3.39
S_{max}	6,750	11,052	6,734	11,076	2,677	5,286	2,582	4,539	3,473	4,907	3,993	6,014
		4,926		4,922		1,787		1,816		2,727		3,013
S_{gen}	1,136	2,054	1,113	2,067	492	1,006	470	852	458	793	605	1,025
		647		611		287		304		286		374
80% S_{MSY}	2,935	4,278	2,945	4,292	992	1,658	992	1,599	1,820	2,313	1,921	2,617
		2,132		2,183		694		715		1,536		1,562

	Curtis Inlet				Damdochax				Devon			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.64	4.45	3.02	4.27	7.74	12.67	7.31	11.51	2.98	4.05	3.10	3.99
		1.58		2.07		4.57		4.70		2.16		2.37
S_{max}	21,090	62,782	19,036	40,155	6,111	18,867	6,437	20,173	9,381	13,721	9,101	12,449
		12,656		12,329		3,744		3,990		7,164		7,287
S_{gen}	3,952	11,099	3,516	7,478	647	2,624	701	2,700	1,751	2,639	1,678	2,403
		2,076		2,095		272		316		1,203		1,225
80% S_{MSY}	6,996	16,096	7,034	13,336	3,554	9,535	3,680	10,206	3,463	4,248	3,462	4,255
		4,461		4,835		2,360		2,449		2,914		2,931

	Evelyn				Fred Wright				Freeda			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.86	6.23	3.53	5.09	11.92	18.43	10.03	15.90	4.45	7.72	3.75	5.65
		2.37		2.53		7.17		5.86		2.38		2.58
S_{max}	2,939	8,014	3,211	7,695	10,424	19,759	11,847	29,358	726	1,555	822	1,570
		1,844		2,053		7,251		7,880		486		564
S_{gen}	492	1,505	560	1,442	761	2,014	1,005	3,635	112	284	140	294
		241		314		373		465		53		79
80% S_{MSY}	1,282	2,833	1,326	2,785	6,730	11,723	7,398	16,407	340	539	351	576
		902		916		4,881		5,254		260		264

	Hartley Bay				Johnston				Kadjusdis River			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.60	4.77	3.53	4.50	8.27	24.65	3.72	6.60	4.30	6.66	3.79	5.42
		2.63		2.73		2.30		2.41		2.86		2.82
S_{max}	2,811	7,168	2,897	6,334	4,090	6,648	4,601	7,451	9,319	56,854	11,756	79,916
		1,809		1,913		2,626		2,939		4,622		5,685
S_{gen}	489	1,361	508	1,194	413	1,060	769	1,320	1,476	9,924	1,972	14,288
		272		299		117		432		566		813
80% S_{MSY}	1,183	2,540	1,200	2,352	2,322	3,837	1,969	3,336	4,260	22,753	5,110	31,005
		848		869		1,276		1,152		2,503		2,796

	Kainet Creek				Keecha				Kitlope			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	9.85	16.79	4.92	8.82	2.75	5.62	3.14	4.73	2.90	3.97	3.03	3.91
		5.46		3.22		1.33		2.08		2.13		2.31
S_{max}	1,653	3,104	2,806	10,392	9,873	37,922	8,932	19,386	101,800	208,672	98,217	178,560
		1,170		1,619		5,682		5,701		67,186		67,801
S_{gen}	142	396	411	1,782	1,796	5,857	1,620	3,616	19,025	40,204	18,173	34,035
		65		162		797		920		11,663		11,693
80% S_{MSY}	1,023	1,691	1,383	4,484	3,331	8,369	3,375	6,492	36,797	65,796	36,604	61,250
		779		903		1,932		2,286		26,098		26,347

	Kitsumkalum				Koeye				Kooryet			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.89	4.09	3.06	4.04	3.49	5.34	3.41	4.72	2.54	4.72	3.01	4.36
		2.06		2.26		2.42		2.61		1.27		2.02
S_{max}	25,969	138,227	23,876	105,023	18,674	172,687	19,572	154,877	10,519	172,937	9,120	19,878
		13,834		13,875		7,654		8,733		6,092		6,043
S_{gen}	4,864	26,450	4,403	20,140	3,243	31,858	3,491	28,936	1,936	16,943	1,685	3,801
		2,376		2,358		1,114		1,378		951		1,010
80% S_{MSY}	9,419	37,802	8,979	33,416	7,603	61,967	7,931	58,511	3,375	18,783	3,368	6,615
		5,561		5,618		3,710		3,937		1,906		2,325

	Kwakwa Creek				Lakelse				Long			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.68	5.47	3.52	4.88	2.76	3.75	2.93	3.82	4.68	7.79	3.90	5.69
		2.40		2.57		2.03		2.25		2.87		2.77
S_{max}	6,554	13,024	6,816	11,996	24,945	37,068	23,935	34,009	137,290	313,191	161,465	378,128
		4,592		4,757		18,551		18,327		87,868		104,051
S_{gen}	1,125	2,411	1,192	2,210	4,735	7,070	4,482	6,518	20,713	56,878	26,833	70,117
		660		738		3,242		3,186		9,638		14,249
80% S_{MSY}	2,792	4,639	2,814	4,565	8,645	11,493	8,671	11,608	65,575	121,290	70,898	146,340
		2,077		2,061		6,753		6,941		47,520		50,278

	Lowe/Simpson/Weir				Marian				Mary Cove Creek			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.55	3.82	2.86	3.85	2.34	4.20	2.89	4.17	3.04	4.27	3.14	4.16
		1.71		2.11		1.34		1.92		2.15		2.38
S_{max}	18,818	29,191	17,860	25,845	25,286	81,194	21,425	44,393	16,457	51,506	15,784	47,427
		13,812		13,368		14,691		14,525		9,547		9,206
S_{gen}	3,576	5,415	3,362	4,951	4,635	12,586	3,997	8,350	3,039	9,660	2,883	8,906
		2,426		2,362		2,432		2,495		1,655		1,639
80% S_{MSY}	6,082	8,542	6,352	8,764	7,393	16,406	7,656	14,226	6,143	17,737	6,039	17,380
		4,251		4,790		3,881		5,195		3,653		3,604

	Mcdonell				Mercer				Meziadin			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	5.02	6.85	4.39	5.97	3.80	6.23	3.53	5.10	6.61	9.97	6.48	9.38
		3.58		3.30		2.32		2.55		4.31		4.45
S_{max}	3,096	3,995	3,341	4,306	8,015	14,479	8,435	12,862	248,576	482,065	251,649	477,821
		2,567		2,742		5,539		6,149		167,636		174,602
S_{gen}	445	681	522	754	1,347	2,766	1,479	2,429	29,388	73,875	30,121	73,462
		302		359		697		892		14,621		15,885
80% S_{MSY}	1,545	1,810	1,557	1,841	3,451	4,825	3,494	4,662	137,790	235,158	139,031	230,450
		1,361		1,364		2,832		2,843		102,959		104,598

	Mikado				Morice				Motase			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.23	4.46	3.26	4.23	5.01	8.20	5.38	8.27	2.77	4.28	2.84	4.34
		2.36		2.51		3.07		3.51		1.74		1.78
S_{max}	4,990	7,116	4,935	6,714	17,323	57,609	16,196	40,346	1,232	2,343	1,215	2,249
		3,818		3,927		10,470		10,252		760		751
S_{gen}	906	1,375	895	1,291	2,488	9,758	2,219	6,368	231	440	226	418
		597		637		1,131		1,112		127		126
80% S_{MSY}	1,947	2,306	1,944	2,294	8,593	24,424	8,326	18,453	414	745	418	734
		1,714		1,709		5,700		5,621		276		276

	Namu				Owikeno				Port John			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	3.21	5.54	3.29	4.75	1.89	2.78	2.37	3.33	2.55	4.92	3.07	4.49
		1.83		2.28		1.32		1.66		1.38		1.98
S_{max}	6,989	26,886	6,844	16,769	1,865,768	########	1,274,512	3,859,755	2,932	24,817	2,370	11,839
		3,961		4,257		925,618		756,019		1,307		1,260
S_{gen}	1,256	4,844	1,230	3,158	329,015	1,853,066	244,892	692,274	536	3,876	436	2,286
		566		665		176,395		139,389		203		209
80% S_{MSY}	2,686	7,627	2,695	5,952	422,542	1,915,239	381,728	893,347	923	5,253	880	3,759
		1,822		1,861		259,761		263,934		452		509

	Prudhomme				Roderick				Shawatlan			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.42	3.32	2.71	3.58	2.22	3.86	2.78	3.93	3.14	5.16	3.22	4.54
		1.72		1.99		1.29		1.86		1.91		2.29
S_{max}	10,160	16,771	9,281	13,951	2,796	9,735	2,421	5,514	4,990	11,839	4,878	8,517
		7,442		7,063		1,617		1,541		3,151		3,430
S_{gen}	1,953	3,003	1,773	2,654	509	1,465	456	1,039	914	2,195	884	1,629
		1,367		1,255		278		275		450		543
80% S_{MSY}	3,136	4,187	3,171	4,086	780	1,872	836	1,773	1,905	3,157	1,896	2,867
		2,427		2,585		395		539		1,465		1,494

	Skidegate				Stephens				Swan			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.43	3.48	2.76	3.65	6.72	12.47	6.58	11.59	3.36	6.24	4.18	6.84
		1.63		2.01		3.70		4.02		1.94		2.53
S_{max}	35,069	65,204	31,675	48,465	6,155	9,631	6,238	9,162	27,982	44,942	26,105	41,271
		24,503		23,441		4,418		4,489		18,109		17,137
S_{gen}	6,735	11,323	6,032	9,227	719	1,565	744	1,418	4,957	8,434	4,176	7,462
		4,413		4,151		319		352		2,456		2,185
80% S_{MSY}	10,849	15,500	10,934	14,596	3,402	4,606	3,435	4,626	11,045	17,753	11,688	17,999
		8,305		8,778		2,747		2,725		6,964		7,843

	Tahlo/Morrison				Tankeeah River				Tsimtack/Moore/Roger			
Model	Basic		Hierarchical		Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	4.36	6.22	4.70	6.45	4.37	7.20	3.77	5.49	5.24	12.92	3.61	6.13
		3.04		3.32		2.60		2.67		1.91		2.34
S_{max}	20,253	36,175	19,224	31,789	7,270	20,627	8,467	24,101	3,650	15,670	4,439	13,822
		14,267		13,870		4,404		5,345		2,251		2,782
S_{gen}	3,169	6,478	2,891	5,533	1,131	3,961	1,442	4,590	508	2,676	762	2,554
		1,839		1,725		504		755		161		385
80% S_{MSY}	9,474	14,495	9,250	13,498	3,398	7,523	3,665	8,710	1,816	4,551	1,879	5,197
		7,324		7,239		2,367		2,571		1,234		1,199

	Yakoun				Yeo			
Model	Basic		Hierarchical		Basic		Hierarchical	
Statistic	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**	**Estimate**	**UCL/LCL**
Ricker α	2.88	3.99	3.04	3.98	4.75	7.38	3.93	5.61
		2.08		2.32		2.92		2.77
S_{max}	24,090	40,130	22,927	33,504	1,642	3,085	1,891	3,673
		17,549		17,190		1,142		1,316
S_{gen}	4,535	7,712	4,251	6,472	244	558	313	670
		2,922		2,878		129		182
80% S_{MSY}	8,685	11,582	8,590	10,921	797	1,295	837	1,428
		7,205		7,179		608		617

Fig. A1. Difference in Mean Raw Error, MRE, of estimated lower benchmark (S_{gen}, (a), and S_{25th} (b)) and the "true" lower benchmarks ("true" S_{gen}), between sensitivity analyses listed on the X-axis and the base case scenario. Black bars are analyses where the input parameter was increased relative to the base case (see Table 1); white bars are analyses where the input parameter was reduced relative to the base case. Positive values indicate sensitivity analyses where MRE increased under that change in input parameter from the base case; negative values indicate analyses where the MRE declined under that change in input parameter.