SCRS/2013/191 Collect. Vol. Sci. Pap. ICCAT, 70(2): 639-645 (2014)

ON THE PROBABILITY OF DETECTING CHANGES IN THE RECRUITMENT OF WESTERN BLUEFIN TUNA WITH INCREASED SPAWNING BIOMASS

Clay E. Porch

SUMMARY

This paper examines the statistical power to determine if the recruitment of western Atlantic bluefin tuna increases with spawning biomass when the spawning stock is allowed to rebuild under various catch scenarios. Stochastic projections were conducted using the bootstrap methodology employed by the 2012 SCRS assessment of western Atlantic bluefin tuna. Tables of statistical power are generated by tallying the bootstrap observation of recruitment under Beverton and Holt relationships with various degrees of compensation, TAC levels and years elapsed. The statistical power to discriminate between the HRS and LRS is very low even with little compensation (K=infinity) given the expected increases in SSB with a TAC of 2,500 t. The current TAC of 1,750 t could allow the spawning biomass to rebuild enough to afford moderate power to discriminate between the HRS and LRS by the year 2024 except for the combination of high compensation (K<=20,000) and low alpha. A TAC of 1,000 t or less should allow the spawning biomass to rebuild enough to afford moderate power to discriminate between the HRS and LRS by the end of the rebuilding period (2018) and high power (>0.8) by 2025.

RÉSUMÉ

Le présent document examine la puissance statistique pour déterminer si le recrutement du thon rouge de l'Atlantique Ouest augmente avec la biomasse reproductrice lorsqu'on permet au stock reproducteur de se rétablir en vertu de divers scénarios de capture. Des projections stochastiques ont été réalisées à l'aide de la méthodologie du bootstrap utilisée dans l'évaluation du thon rouge de l'Ouest effectuée en 2012 par le SCRS. Les tableaux de puissance statistique sont créés en récapitulant l'observation par bootstrap du recrutement en vertu des relations de Beverton et Holt avec divers degrés de compensation, niveaux de TAC et années écoulées. La puissance statistique pour distinguer entre HRS et LRS est très faible avec peu de compensation (K=infinité) compte tenu des augmentations escomptées de la SSB avec un TAC de 2.500 t. Le TAC actuel de 1.750 t pourrait permettre à la biomasse reproductrice de se rétablir suffisamment pour fournir une puissance modérée pour distinguer entre HRS et LRS avant 2024, sauf pour la combinaison de forte compensation (K<=20.000) et de faible alpha. Un TAC de 1.000 t ou moins devrait permettre à la biomasse reproductrice de se rétablir suffisamment pour fournir une puissance modérée pour distinguer entre HRS et LRS avant la fin de la période de rétablissement (2018) et une forte puissance (>0,8%) avant 2025.

RESUMEN

En este documento se examina el poder estadístico para determinar si el reclutamiento del atún rojo del Atlántico oeste se incrementa con la biomasa reproductora cuando se permite al stock reproductor recuperarse bajo diversos escenarios de captura. Se llevaron a cabo proyecciones estocásticas utilizando la metodología de bootstrap usada en la evaluación del SCRS de 2012 del atún rojo del Atlántico occidental. Las tablas de poder estadístico se generan resumiendo la observación del reclutamiento en el marco de una relación Beverton y Holt, con varios grados de depensación, niveles de TAC y años transcurridos. El poder estadístico de discriminar entre HRS y LRS es muy bajo incluso con poca compensación (K = infinito), dados los incrementos previstos en la SSB con un TAC de 2.500 t. El TAC actual de 1.750 t permitirá a la biomasa reproductora recuperarse lo suficiente para permitir un poder moderado de discriminar entre HRS y LRS, desde ahora hasta 2024, excepto para la combinación de alta depensación (K= 20.000) y un valor bajo de alfa. Se prevé que un TAC de 1.000 t o menos permitirá a la biomasa reproductora recuperarse lo suficiente para permitir un poder moderado de discriminar entre HRS y LRS antes del final del período de recuperación (2018) y un poder alto (>0,8) antes de 2025.

KEYWORDS

Stock assessment, Stock-recruitment relationships, Projections, Management measures

1. Introduction

The nature of the relationship between the spawning biomass of western bluefin tuna and the subsequent recruitment of yearlings has been a topic of long-standing debate among scientists and managers alike (Rosenberg *et al.* 2012). Currently, the International Commission for the Conservation of Atlantic Tuna's Standing Committee on Research and Statistics (SCRS) bases its management advice on two competing scenarios: one that assumes future recruitment will increase with spawning biomass (the high recruitment potential scenario; HRS) and one that assumes future recruitment will remain at recent levels (the low recruitment potential scenario; LRS). The SCRS has indicated that it "has no strong evidence to favor either scenario over the other and notes that both are plausible (but not extreme) lower and upper bounds on rebuilding potential."

The ability to statistically discriminate between the LRS and HRS recruitment hypotheses depends on the degree to which recruitment increases with spawning biomass in contrast to the degree that it fluctuates owing to other unpredictable factors. In the case of WBFT, the magnitude of recruitment appears to vary considerably from one year to the next, but there has been little change in the spawning biomass over the last 30 years (**Figure 1**). Accordingly, there is insufficient contrast in SSB to distinguish the true nature of the stock-recruitment relationship unless the data from the 1960s and early 1970s are used (which are arguably less reliable) or the spawning biomass is allowed to increase in the future. The latter alternative of course begs the question how much contrast in spawning biomass must there be.

The 2012 Executive Summary for Western Bluefin Tuna (SCRS 2012) identified a potential opportunity to test the LRS and HRS scenarios by capitalizing on the apparently large 2003 yearclass: "Maintaining catch at current levels (1,750 t) is expected to allow the spawning biomass to increase, which may help resolve the issue of low and high recruitment potential. For example, should the high recruitment hypothesis be correct, allowing substantial increases in spawning biomass should lead to higher recruitment." However, no analyses were conducted to suggest exactly how much the spawning biomass would need to be allowed to increase in order to distinguish the two hypotheses with any statistical certainty. To this end, Commissioners attending the First Meeting of the Working Group of Fisheries Managers and Scientists in Support of the Western Atlantic Bluefin Tuna Stock Assessment (ICCAT, 2013) recently requested the SCRS to conduct an analysis to "Determine how long it would take the western Atlantic bluefin tuna stock to reach spawning stock biomass levels that would allow for the testing of the stock-recruit relationship (under different levels of TAC)."

This paper uses simulations to determine the power to detect an increasing relationship between recruitment and spawning biomass when the spawning stock is allowed to rebuild under various catch scenarios.

2. Methods

The so-called HRS typically expresses recruitment R as a Beverton and Holt function of the spawning biomass S in the preceding year, which may be written

$$(1) \quad R = \frac{R_0}{1+K/S}$$

Values for the parameters describing the maximum expected recruitment R_0 and compensation K are typically estimated by fitting equation (1) to a time series of S and R using maximum likelihood techniques. Note that, above a certain threshold level of spawning biomass, the LRS is equivalent to the case where K=0 (i.e., constant recruitment expected at R_0). Thus, in an ideal circumstance the problem would boil down to testing whether the maximum likelihood estimate for K ($\hat{k}$) is statistically greater than zero. That is, an apparent trend in recruitment with spawning biomass may be deemed to be statistically significant if the probability p is low that an estimate as large or larger than a value $\hat{k}$ would be obtained if in fact the LRS ($K=0$) were true. The non-parametric equivalent to testing if an estimate $\hat{k}$ is statistically greater than zero is to test is if the average observed recruitment $\bar{r}$ is statistically greater than the expectation of the LRS (R_0^{LRS}), i.e., that the probability of observing an average recruitment of $\bar{r}$ or greater under the LRS is low.

The p-level that defines 'low', often denoted by the symbol α, is typically set at 0.05 or less. In cases where $p> \alpha$, it can be fairly stated that procedure failed to reject the LRS. However, the failure to reject the LRS should not necessarily be interpreted as acceptance of it. As Peterman (1990) points out, this jump in logic is not justified unless the probability β of committing a type 2 error (i.e., rejecting the HRS when it is true), is sufficiently low. The quantity 1-β is commonly referred to as the power of the statistical procedure. In this context it is the probability of correctly identifying the existence of a trend in recruitment with spawning biomass.

In the particular case of WBFT, there has been little change in the estimated spawning biomass since the late 1970s. As a result, almost any value of K provides an equally good fit to the data and one fails to reject the LRS. However, the statistical power of this hypothesis test is very low. The question then becomes by how much must the spawning biomass increase and for how many years in order for the hypothesis test to have reasonable power. Inasmuch as the primary tool for allowing spawning biomass to increase has been to reduce the total allowable catch (TAC), the problem reduces to what size TAC must be implemented for how many years.

The power of any statistical test depends on the number of observations, the contrast in the covariate (here S) and the standard deviation of observed recruitment relative to the expectation σ. In the present example we assume that recruitment varies from the expected value as lognormal-distributed random variable with mean 0 and σ equal to the sample standard deviation in observed log-scale recruitments since 1976 (as spawning biomass is estimated to have changed little over this period). Stochastic projections of recruitment were then made through the year 2025 under a range of values of K and TAC trajectories using the bootstrap projection methods adopted by the 2012 SCRS Bluefin tuna species group (SCRS 2012b). The catch in 2013 was assumed to be equal to the current TAC of 1,750 mt.

Average recruitments were computed from the vector of annual recruitments projected in each bootstrap ($r_{b,i}$), beginning in projection year 2014 and using n years of projected recruitments:

$$\bar{r}_b(n) = \frac{1}{n}\sum_{i=2014}^{2014+n-1} r_{b,i} \qquad\qquad n = 1,2,\ldots 12$$

The critical value for the mean recruitment r_c where the LRS would be rejected was calculated as the (1-α)-percentile of the collection of annual recruitments derived from each bootstrap when K=0 (**Figure 2**). The probability of a type 2 error, β, was calculated as fraction of average recruitments $\bar{r}_b(n)$ that exceeded the critical value r_c (see **Figure 3**).

The statistical power of the hypothesis test that $\bar{r} > r_c$, 1-β, was computed for three levels of α (0.05, 0.1, and 0.2), seven levels of TAC (0, 500, 1000, 1500, 1750, 2000 and 2500 mt), and three levels of K (the value estimated during the last assessment, a lower value of 20,000 and the upper limit of infinity). Note that the estimates of R_0 and K derived from a time series of S and R values are typically correlated. Accordingly, the statistical power of the hypothesis test should depend on the magnitude of the true values of both R_0 and K. In that case one would need to run the aforementioned projections for a range of K and R_0 values. However, given that the 1981-2008 data[1] includes many observations of R for similar values of S, it can be expected that refitting equation (1) using new data will result in similar predictions of R for the 1981-2008 average spawning biomass (as shown in **Figure 1**). Accordingly, the value of R_0 for the HRS (R_0^{HRS}) will be approximately equal to

$$(2)\quad R_0^{HRS} = \frac{R_0^{LRS}}{1+K/S_{1981-2008}}$$

and it is sufficient for this exercise to vary K alone.

3. Results and discussion

The expected increase in spawning biomass under the various combinations of K and TAC are shown in **Figure 4**. The corresponding results of the power analysis are summarized in **Table 1** and **Figure 5** for the various levels of α.

The statistical power to discriminate between the HRS and LRS is estimated to be very low with a TAC of 2,500 mt even with little compensation (K=infinity) because the spawning biomass is not expected to grow substantially. The current TAC of 1750 mt could allow the spawning biomass to rebuild enough to afford moderate power to discriminate between the HRS and LRS by the year 2024 except for the combination of high compensation (K<=20,000) and low alpha. A TAC of 1000 mt or less should allow the spawning biomass to rebuild enough to afford moderate power to discriminate between the HRS and LRS by the end of the rebuilding period (2018) and high power (>0.8) by 2025.

---

[1] Consistent with the 2012 SCRS, the last three years of recruitment estimates from the stock assessment (2009-2011) were not used in the calculations.

It was pointed out during the 2013 bluefin species group meeting that it may take several years before realized recruitments show up in the fishery and can be reliably estimated in the VPA. Indeed, the SCRS typically discounts the estimates of recruitment from the last three years precisely for this reason. An implication of this is that it may take some additional time to detect a statistically powerful change in recruitment. On the other hand, the present analysis excludes the recruitment estimates for the last three years in the assessment (2009-2011) during which the quota was somewhat reduced. As these recruitments would be better estimated in subsequent assessments, they could be incorporated into the power analysis, which might somewhat offset the aforementioned delay. The implementation of a fishery-independent survey of yearling bluefin tuna could also mitigate such a delay.

Acknowledgments

Shannon Cass-Calay provided the projection files used in the 2012 assessment. M. Lauretta read the manuscript and provided helpful comments.

References

SCRS. 2012. Report of the Standing Committee on Research and Statistics (SCRS). PLE-104 / 2012. 300 p.

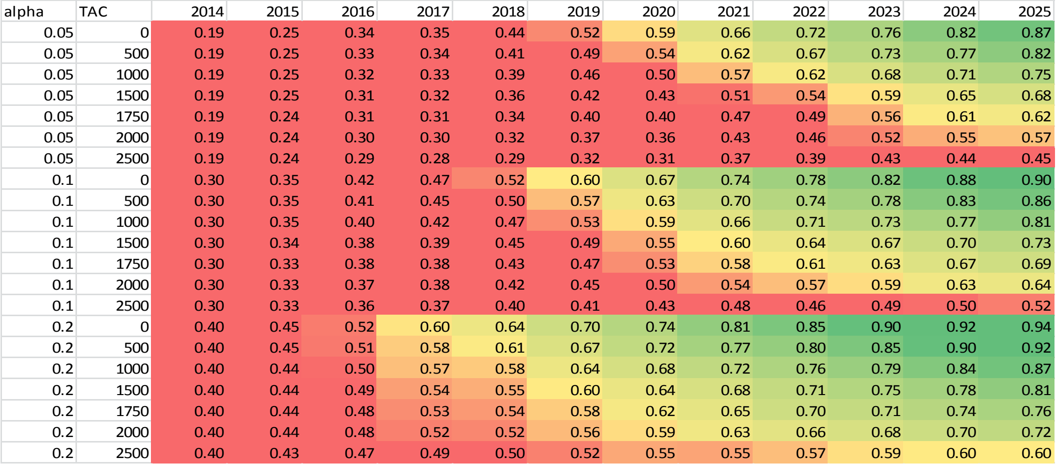
SCRS. 2012b. Report of the 2012 Atlantic Bluefin Tuna Stock Assessment Session, Madrid, Spain, September 4-11, 2012. SCRS/2012/015. 124 p.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquat. Sci. 47: 2-15.

Rosenberg, A., A. Cooper, M. Maunder, M. Mcallister, R. Methot, S. Miller, C. Porch, J. Powers, T. Quinn, V. Restrepo, G. Scott, J.C. Seijo, G. Stefansson, and J. Walter. Scientific Examination of Western Atlantic Bluefin Tuna Stock Recruitment Relationships. SCRS/2012/162.

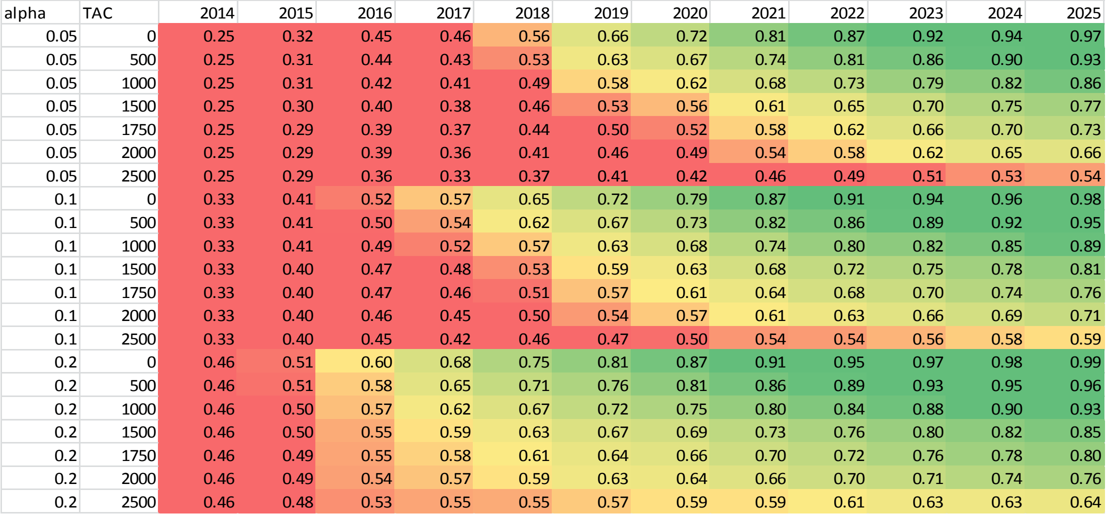
Table 1. Statistical power of tests to reject the LRS as a function of the extent of compensation *K*, rejection criteria α, TAC and number of years.

K=20,000



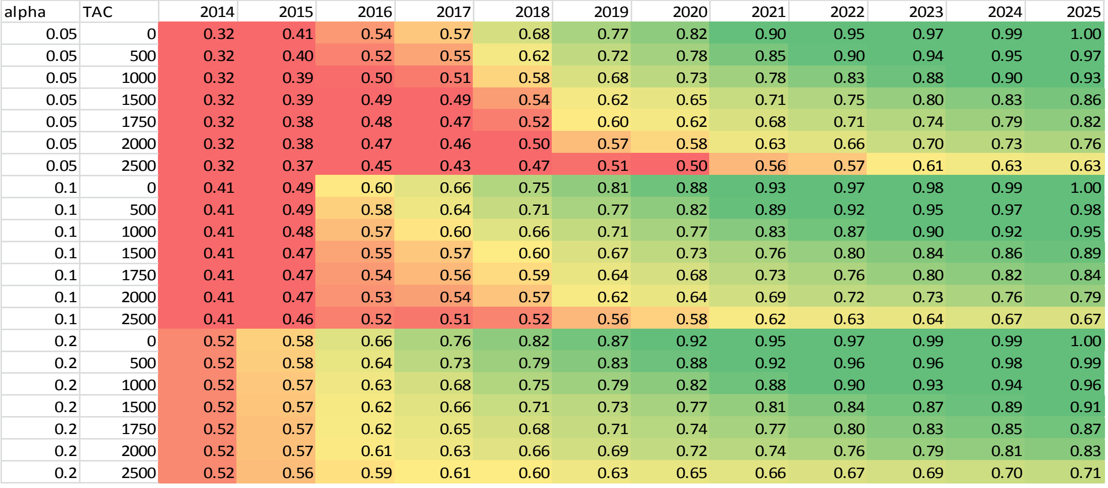
alpha	TAC	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
0.05	0	0.19	0.25	0.34	0.35	0.44	0.52	0.59	0.66	0.72	0.76	0.82	0.87
0.05	500	0.19	0.25	0.33	0.34	0.41	0.49	0.54	0.62	0.67	0.73	0.77	0.82
0.05	1000	0.19	0.25	0.32	0.33	0.39	0.46	0.50	0.57	0.62	0.68	0.71	0.75
0.05	1500	0.19	0.25	0.31	0.32	0.36	0.42	0.43	0.51	0.54	0.59	0.65	0.68
0.05	1750	0.19	0.24	0.31	0.31	0.34	0.40	0.40	0.47	0.49	0.56	0.61	0.62
0.05	2000	0.19	0.24	0.30	0.30	0.32	0.37	0.36	0.43	0.46	0.52	0.55	0.57
0.05	2500	0.19	0.24	0.29	0.28	0.29	0.32	0.31	0.37	0.39	0.43	0.44	0.45
0.1	0	0.30	0.35	0.42	0.47	0.52	0.60	0.67	0.74	0.78	0.82	0.88	0.90
0.1	500	0.30	0.35	0.41	0.45	0.50	0.57	0.63	0.70	0.74	0.78	0.83	0.86
0.1	1000	0.30	0.35	0.40	0.42	0.47	0.53	0.59	0.66	0.71	0.73	0.77	0.81
0.1	1500	0.30	0.34	0.38	0.39	0.45	0.49	0.55	0.60	0.64	0.67	0.70	0.73
0.1	1750	0.30	0.33	0.38	0.38	0.43	0.47	0.53	0.58	0.61	0.63	0.67	0.69
0.1	2000	0.30	0.33	0.37	0.38	0.42	0.45	0.50	0.54	0.57	0.59	0.63	0.64
0.1	2500	0.30	0.33	0.36	0.37	0.40	0.41	0.43	0.48	0.46	0.49	0.50	0.52
0.2	0	0.40	0.45	0.52	0.60	0.64	0.70	0.74	0.81	0.85	0.90	0.92	0.94
0.2	500	0.40	0.45	0.51	0.58	0.61	0.67	0.72	0.77	0.80	0.85	0.90	0.92
0.2	1000	0.40	0.44	0.50	0.57	0.58	0.64	0.68	0.72	0.76	0.79	0.84	0.87
0.2	1500	0.40	0.44	0.49	0.54	0.55	0.60	0.64	0.68	0.71	0.75	0.78	0.81
0.2	1750	0.40	0.44	0.48	0.53	0.54	0.58	0.62	0.65	0.70	0.71	0.74	0.76
0.2	2000	0.40	0.44	0.48	0.52	0.52	0.56	0.59	0.63	0.66	0.68	0.70	0.72
0.2	2500	0.40	0.43	0.47	0.49	0.50	0.52	0.55	0.55	0.57	0.59	0.60	0.60

K=estimated (75,000)



alpha	TAC	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
0.05	0	0.25	0.32	0.45	0.46	0.56	0.66	0.72	0.81	0.87	0.92	0.94	0.97
0.05	500	0.25	0.31	0.44	0.43	0.53	0.63	0.67	0.74	0.81	0.86	0.90	0.93
0.05	1000	0.25	0.31	0.42	0.41	0.49	0.58	0.62	0.68	0.73	0.79	0.82	0.86
0.05	1500	0.25	0.30	0.40	0.38	0.46	0.53	0.56	0.61	0.65	0.70	0.75	0.77
0.05	1750	0.25	0.29	0.39	0.37	0.44	0.50	0.52	0.58	0.62	0.66	0.70	0.73
0.05	2000	0.25	0.29	0.39	0.36	0.41	0.46	0.49	0.54	0.58	0.62	0.65	0.66
0.05	2500	0.25	0.29	0.36	0.33	0.37	0.41	0.42	0.46	0.49	0.51	0.53	0.54
0.1	0	0.33	0.41	0.52	0.57	0.65	0.72	0.79	0.87	0.91	0.94	0.96	0.98
0.1	500	0.33	0.41	0.50	0.54	0.62	0.67	0.73	0.82	0.86	0.89	0.92	0.95
0.1	1000	0.33	0.41	0.49	0.52	0.57	0.63	0.68	0.74	0.80	0.82	0.85	0.89
0.1	1500	0.33	0.40	0.47	0.48	0.53	0.59	0.63	0.68	0.72	0.75	0.78	0.81
0.1	1750	0.33	0.40	0.47	0.46	0.51	0.57	0.61	0.64	0.68	0.70	0.74	0.76
0.1	2000	0.33	0.40	0.46	0.45	0.50	0.54	0.57	0.61	0.63	0.66	0.69	0.71
0.1	2500	0.33	0.40	0.45	0.42	0.46	0.47	0.50	0.54	0.54	0.56	0.58	0.59
0.2	0	0.46	0.51	0.60	0.68	0.75	0.81	0.87	0.91	0.95	0.97	0.98	0.99
0.2	500	0.46	0.51	0.58	0.65	0.71	0.76	0.81	0.86	0.89	0.93	0.95	0.96
0.2	1000	0.46	0.50	0.57	0.62	0.67	0.72	0.75	0.80	0.84	0.88	0.90	0.93
0.2	1500	0.46	0.50	0.55	0.59	0.63	0.67	0.69	0.73	0.76	0.80	0.82	0.85
0.2	1750	0.46	0.49	0.55	0.58	0.61	0.64	0.66	0.70	0.72	0.76	0.78	0.80
0.2	2000	0.46	0.49	0.54	0.57	0.59	0.63	0.64	0.66	0.70	0.71	0.74	0.76
0.2	2500	0.46	0.48	0.53	0.55	0.55	0.57	0.59	0.59	0.61	0.63	0.63	0.64

K=infinity



alpha	TAC	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
0.05	0	0.32	0.41	0.54	0.57	0.68	0.77	0.82	0.90	0.95	0.97	0.99	1.00
0.05	500	0.32	0.40	0.52	0.55	0.62	0.72	0.78	0.85	0.90	0.94	0.95	0.97
0.05	1000	0.32	0.39	0.50	0.51	0.58	0.68	0.73	0.78	0.83	0.88	0.90	0.93
0.05	1500	0.32	0.39	0.49	0.49	0.54	0.62	0.65	0.71	0.75	0.80	0.83	0.86
0.05	1750	0.32	0.38	0.48	0.47	0.52	0.60	0.62	0.68	0.71	0.74	0.79	0.82
0.05	2000	0.32	0.38	0.47	0.46	0.50	0.57	0.58	0.63	0.66	0.70	0.73	0.76
0.05	2500	0.32	0.37	0.45	0.43	0.47	0.51	0.50	0.56	0.57	0.61	0.63	0.63
0.1	0	0.41	0.49	0.60	0.66	0.75	0.81	0.88	0.93	0.97	0.98	0.99	1.00
0.1	500	0.41	0.49	0.58	0.64	0.71	0.77	0.82	0.89	0.92	0.95	0.97	0.98
0.1	1000	0.41	0.48	0.57	0.60	0.66	0.71	0.77	0.83	0.87	0.90	0.92	0.95
0.1	1500	0.41	0.47	0.55	0.57	0.60	0.67	0.73	0.76	0.80	0.84	0.86	0.89
0.1	1750	0.41	0.47	0.54	0.56	0.59	0.64	0.68	0.73	0.76	0.80	0.82	0.84
0.1	2000	0.41	0.47	0.53	0.54	0.57	0.62	0.64	0.69	0.72	0.73	0.76	0.79
0.1	2500	0.41	0.46	0.52	0.51	0.52	0.56	0.58	0.62	0.63	0.64	0.67	0.67
0.2	0	0.52	0.58	0.66	0.76	0.82	0.87	0.92	0.95	0.97	0.99	0.99	1.00
0.2	500	0.52	0.58	0.64	0.73	0.79	0.83	0.88	0.92	0.96	0.96	0.98	0.99
0.2	1000	0.52	0.57	0.63	0.68	0.75	0.79	0.82	0.88	0.90	0.93	0.94	0.96
0.2	1500	0.52	0.57	0.62	0.66	0.71	0.73	0.77	0.81	0.84	0.87	0.89	0.91
0.2	1750	0.52	0.57	0.62	0.65	0.68	0.71	0.74	0.77	0.80	0.83	0.85	0.87
0.2	2000	0.52	0.57	0.61	0.63	0.66	0.69	0.72	0.74	0.76	0.79	0.81	0.83
0.2	2500	0.52	0.56	0.59	0.61	0.60	0.63	0.65	0.66	0.67	0.69	0.70	0.71

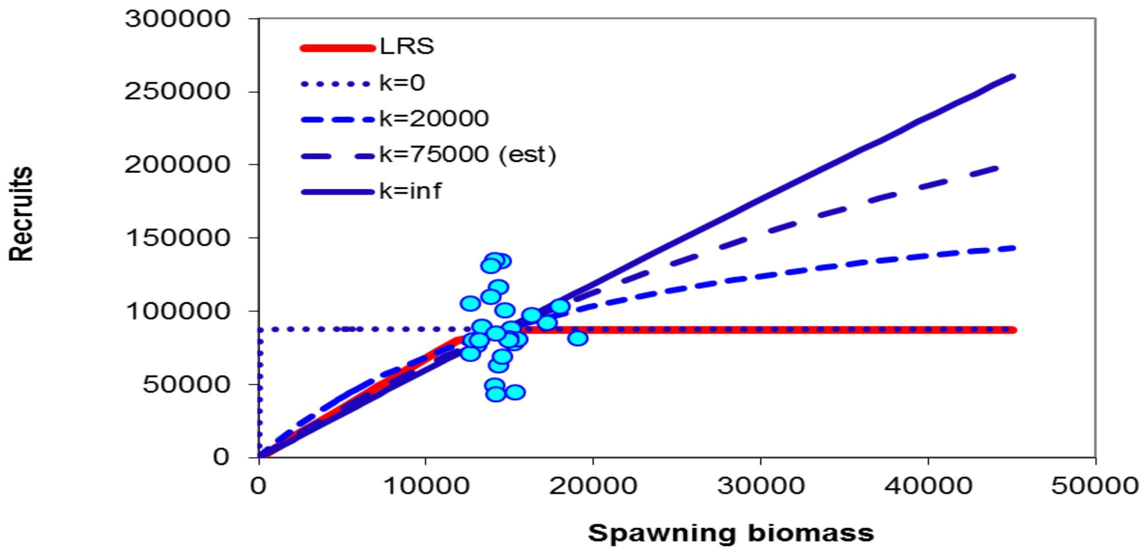


Figure 1. Comparison of LRS and HRS with various levels of compensation (*K*).

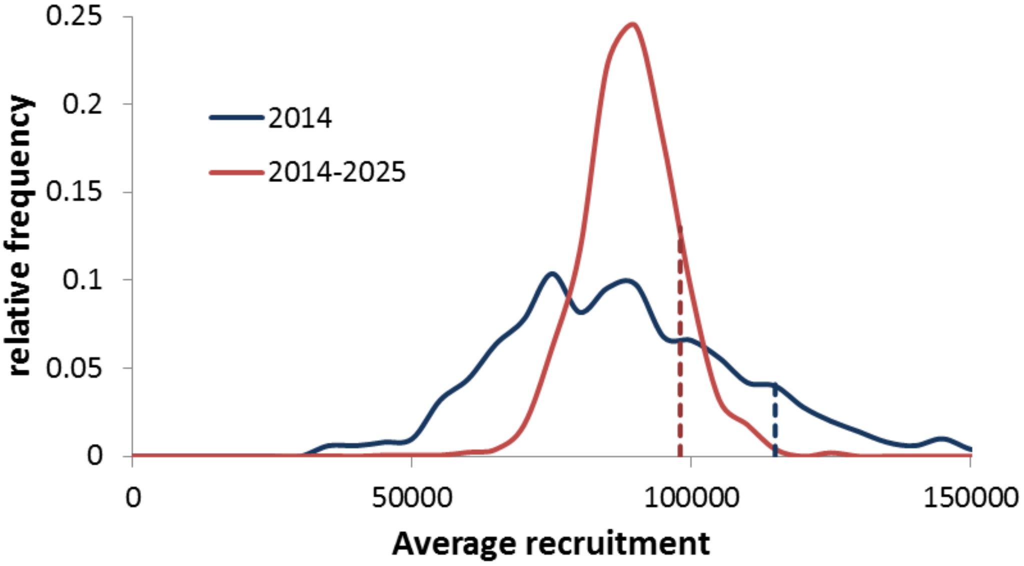


Figure 2. Distribution of bootstrapped average recruitment under the LRS when the average is calculated with one year of data (2014, blue curve) and 12 years of data (2014-2025, red curve). The dashed lines indicated the corresponding critical values (r_c) obtained with α = 10%.

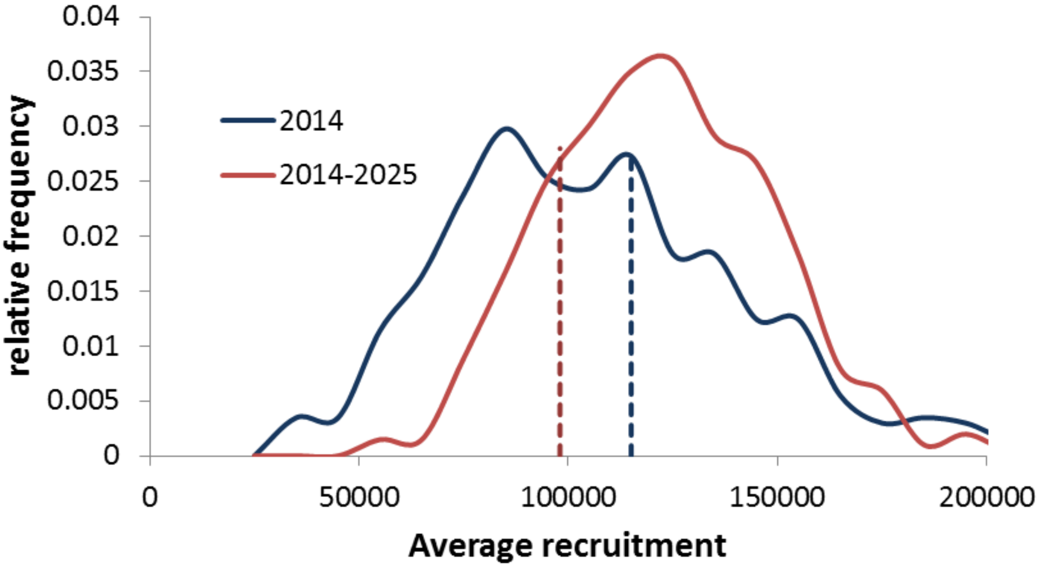


Figure 3. Distribution of bootstrapped average recruitment under the HRS estimated by the 2012 SCRS (*K*=75,000) and TAC=1750 mt when the average is calculated with one year of data (2014, blue curve) and 12 years of data (2014-2025, red curve). The dashed lines indicate the critical values (r_c) obtained with α = 10%. The cumulative proportion to the right of r_c is the power (1-β).

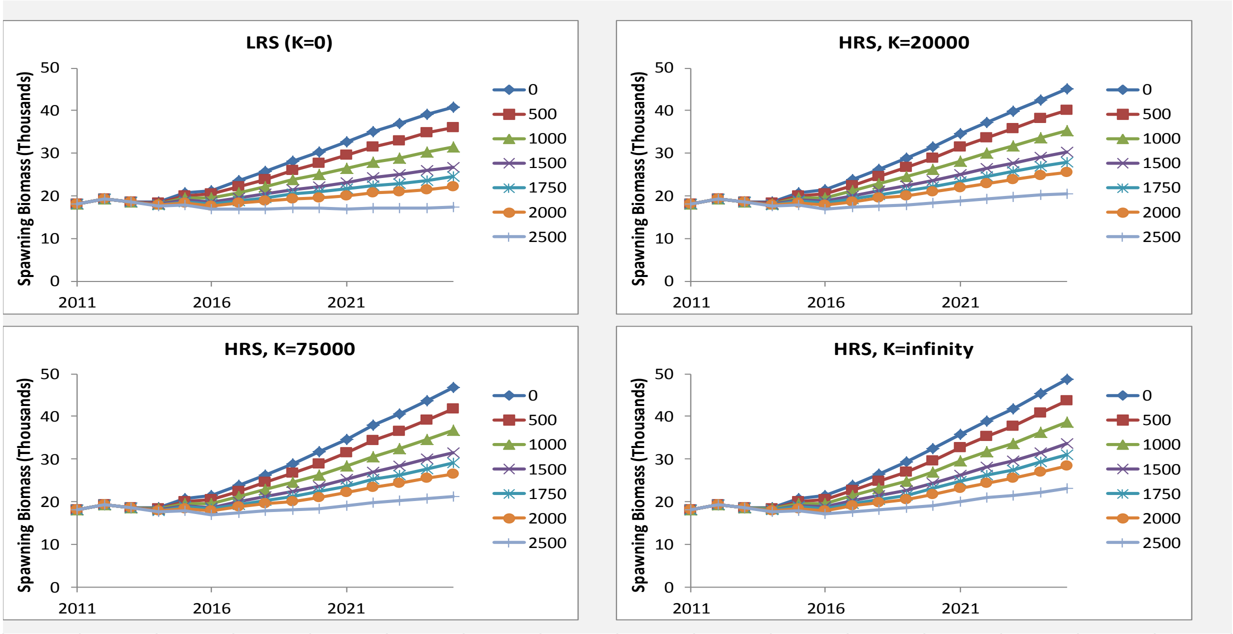


Figure 4. Projected trends in spawning biomass from the 2012 base-case assessment under various combinations of *K* and TAC.

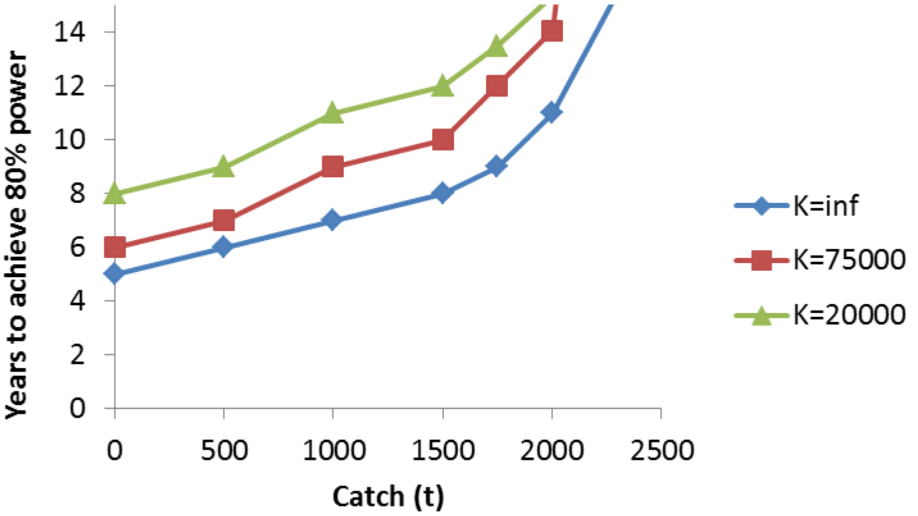


Figure 5. Time required to reach 80% power under the different levels of catch and *K* scenarios for α = 20%.