

Single-region YFT2023 model

Nick Davies

1 July 2025

Summary

As part of the P123 project “Scoping the next Stock Assessment software”, a step includes the preparation of a simplified tuna stock assessment model and input data set. This report presents the preliminary results of this step using the yellowfin tuna stock assessment model as was developed in 2023. The original model was spatially stratified into 5 defined regions. A simplified version was developed using MULTIFAN-CL having no spatial stratification (a single region), while retaining the original fisheries definitions for the 32 capture fisheries, and their fishery-specific input data structures. Instead of the five region-specific survey fisheries for which CPUE indices were available, the simplified model defined a single survey fishery representative of the entire model domain. The simplified model was fitted and achieved convergence, and a comparison with the original multi-region model is presented here.

Method

The existing YFT2023 diagnostic case model that employs the catch-conditioned method was used for preparing the input data, and to demonstrate the application of MULTIFAN-CL to a spatially unstratified tuna data set.

De-stratification of the model spatial configuration and observations

The input files that specify the spatial stratification, recruitment and movement parameterisations (*.ini and *.frq) were duly modified to define a single region with no movement diffusion coefficients or spatial recruitments. All fisheries and all tagging release events were defined to occur in the same single region. All other fishery-specific observations (size compositions, catch, effort, conditional age-length data) for the capture fisheries (1 to 32) were retained without modification; as these fisheries were re-defined as occurring in the single region. A single standardized CPUE time series was available for the survey fisheries combined over all 5 regions, and was taken as being representative of the entire model domain, i.e., a single-region index. Similarly, the size composition data among all five survey fisheries were aggregated into a single data set. This single survey fishery (33) was defined as such for the simplified model. Estimates of temporal precision for the CPUE indices were available, allowing a concentrated CPUE likelihood to be estimated for the simplified model.

The initial MULTIFAN-CL “-makepar” operation was completed, and the resultant 00.par file structure was assessed for correctly excluding all spatial stratification or movement parameterisation. All phases of the doital minimization that originally included spatial parameter estimation, were modified to de-activate the estimation of these parameters. The phase 1 operation was successfully completed, and then all subsequent phases of the minimization run to convergence.

The models included in this comparison were:

mult_regs	- original YFT2023 diagnostic case model, having multiple (5) regional strata
sngl_reg_cpue1	- simplified single-region model, employs non-concentrated CPUE likelihood
sngl_reg_conc_cpue1	- simplified single-region model, employs concentrated CPUE likelihood

Results

Both the single-region models produced stable minimisations with converged solutions obtained (maximum gradients $\sim 2.0 \times 10^{-4}$). Whereas the multi-region model produced a positive definite Hessian (PDH) solution, neither of the single-region models were PDH solutions (Table 1).

In respect of the size-composition data, comparable fits to the length-frequencies were obtained among the models, but markedly worse weight-frequency fits were obtained for the single-region models (21% worse negative-log likelihood) (Table 1). Similarly, worse fits were obtained to the tagging data (9.6% worse negative-log likelihood). Slightly better (1.1%) fits were obtained to the conditional age-length data by the single-region models. There was negligible difference in the fit to the CPUE indices among the single-region models (Figure 8), with that using the concentrated likelihood differing only slightly in respect of the indices in the early time periods for which temporal precision was low.

There were notable differences in the recruitment and natural mortality estimates of the single-region models compared to the multi-region model. Absolute recruitments were approximately 50% lower in all time periods (Figure 2), but with similar temporal variation. The Lorenzen natural mortality function was lower overall by on average approximately 10%, which accorded lower mortality, particularly for age classes < 12 quarters (Figure 5). In contrast, the growth function was almost identical (Figure 4). A number of the estimated selectivity-at-age functions were more dome-shaped, or with lower right-hand limbs (Figure 6).

Despite these differences in model fit to the observations and estimated parameters, overall absolute biomass was generally similar among the multi- and single-region models (Figure 1), and showed similar temporal variation. Noticeable differences were however evident in the recent periods (the past 15 years), with absolute total biomass of the single-region models being 23% lower (Table 1). Despite this similarity, the estimated depletion of adult biomass is at a 35% lower level compared to the multi-region model (Table 1 and Figure 7). Equilibrium quantities also differed, with MSY being 9.5% lower, and unfished total biomass (B_0) being 8.6% higher (Table 1).

These differences in the model derived variables can be attributed to the various parameter estimates that differ among the multi- and single-region models. It seems surprising that a model having substantially lower productivity (50% of the recruitments), yet sustains the equivalent total removals and produces a comparable absolute abundance, with only 10% lower equilibrium yield. Notably, the lower Lorenzen natural mortality over all age classes substantially reduces the total mortality – hence, B_0 is higher. The more dome-shaped selectivities and being lower for older age classes, reduces fishing mortality for the older age classes and allows for “cryptic” biomass, and increases MSY. Also, the spatial dynamics of the multi-region model in respect of recruitment, movement and fishing mortality on sub-populations may increase the overall fishing mortality estimates. Whereas, the fishing mortality of all fisheries acting upon a single population serves to exclude the effects of high impacts on sub-populations. The combined effects

of the above parameter differences can accord higher resilience of the single-region model to fishing mortality. However, the lower natural mortality rate contributes to greater estimated depletion, indicating the relative fishing impact is higher. A closer examination of the estimated adult depletion within the equatorial regions of the multi-region model (regions 2, 3 and 4) reveals substantial regional differences, with that of region 2 being very similar to the estimated level of the single-region models. This might indicate a dominant effect of the region 2 observations upon the single-region model estimates.

Tables and Figures

Table 1. Comparison of selected model parameters, negative log-likelihood terms and derived quantities, for the multi-region (**mult_reg**s) and single-region models (**sn**gl_reg_cpue1, **sn**gl_reg_conc_cpue1).

Model quantity	mult_reg	sn	gl_reg_cpue1	sn	gl_reg_conc_cpue1	%diff_nonconc	%diff_conc
MSY	169700	153500	153600	-9.55	-9.49		
Ccurr.MSY	4.224	4.665	4.662	10.44	10.37		
Fmsy	0.073	0.060	0.060	-18.08	-18.08		
Fmult	1.676	1.209	1.206	-27.86	-28.04		
Fcurr.Fmsy	0.597	0.827	0.829	38.63	38.97		
B0	8510000	9241000	9244000	8.59	8.63		
Bmsy	2336000	2580000	2581000	10.45	10.49		
Bcurr	4648625	3587013	3577813	-22.84	-23.04		
SB0	5347000	6573000	6576000	22.93	22.98		
SBmsy	1072000	1440000	1440000	34.33	34.33		
SBcurr	2404755	1996569	1991219	-16.97	-17.20		
Bcurr.Bmsy	1.990	1.390	1.386	-30.13	-30.34		
SBcurr.SBmsy	2.243	1.387	1.383	-38.19	-38.36		
SBcurr.SBcurrFO	0.454	0.296	0.295	-34.92	-35.08		
SBlatest.SBlatestFO	0.429	0.292	0.291	-31.95	-32.15		
obj_bhsteep	0.311	0.415	0.415	33.51	33.31		
obj_lencomp	-154969.914	-153345.580	-153345.611	-1.05	-1.05		
obj_wtcomp	-610342.171	-483274.416	-483273.668	-20.82	-20.82		
obj_tagdata	13217.201	14479.783	14479.435	9.55	9.55		
obj_ageIngdta	2480.443	2452.738	2452.727	-1.12	-1.12		
obj_cpue	-1157.645	-314.031	-313.991	-72.87	-72.88		
Obj	-750515.800	-619859.458	-619858.965	-17.41	-17.41		
No. parameters	1901	445	445	-76.59	-76.59		
gradient	0.0001268	0.0001774	0.0001527	39.90	20.46		
Lmin	19.800	19.800	19.800	0.00	0.00		
Lmax	141.958	141.924	141.923	-0.02	-0.02		
K	0.132	0.132	0.132	0.03	0.03		
PDH	Yes	No	No	-	-		

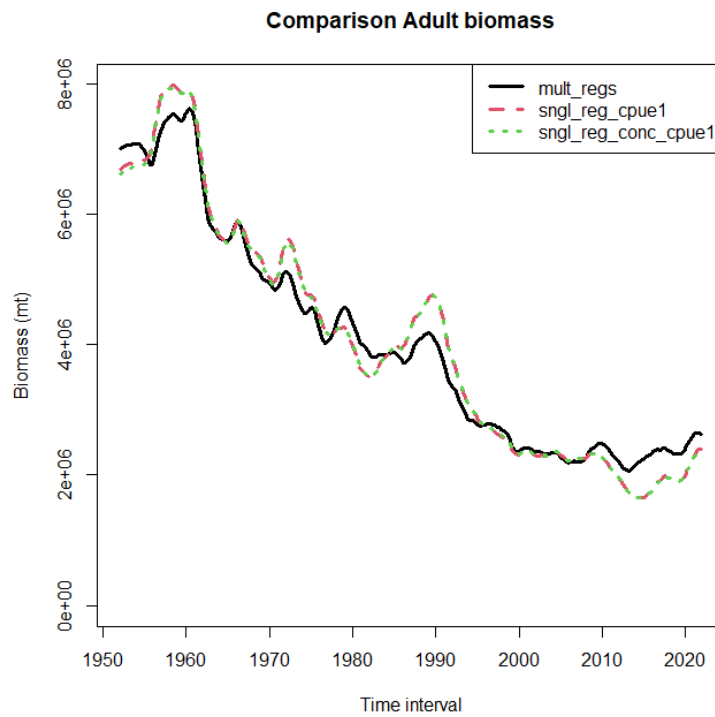


Figure 1. Comparison of adult biomass between the multi-region (mult_regs) and single-region (sngl_reg_cpue1, sngl_reg_conc_cpue1) models.

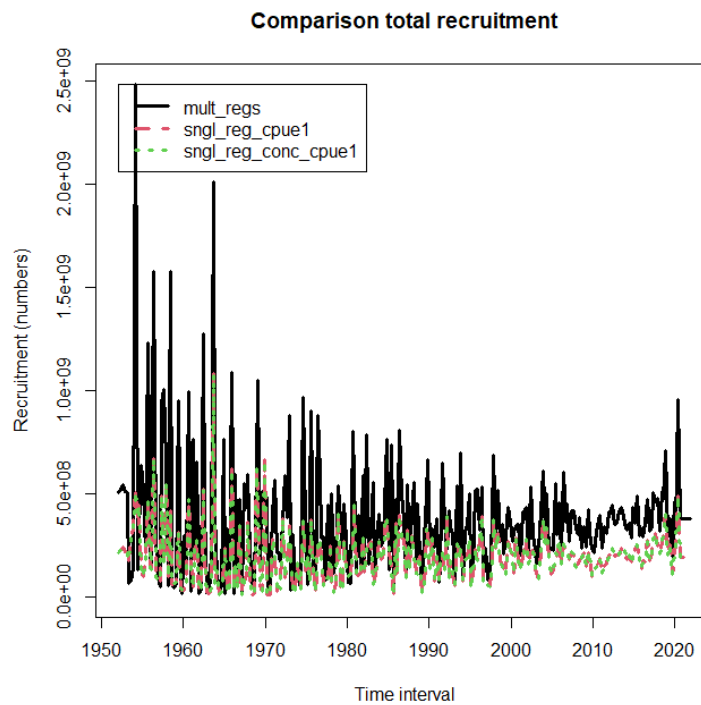


Figure 2. Comparison of estimated absolute recruitment between the multi-region (mult_regs) and single-region (sngl_reg_cpue1, sngl_reg_conc_cpue1) models.

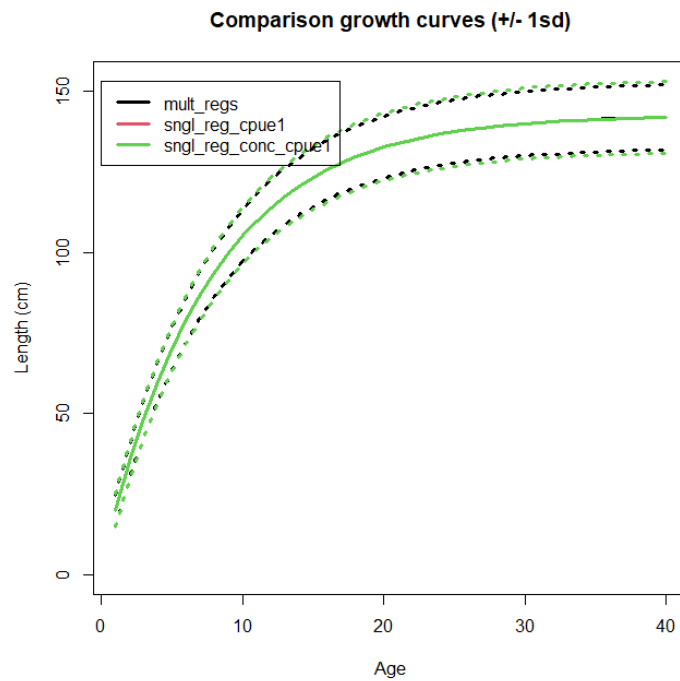


Figure 4. Comparison of estimated von Bertalanffy growth between the multi-region (mult_regs) and single-region (sngl_reg_cpue1, sngl_reg_conc_cpue1) models.

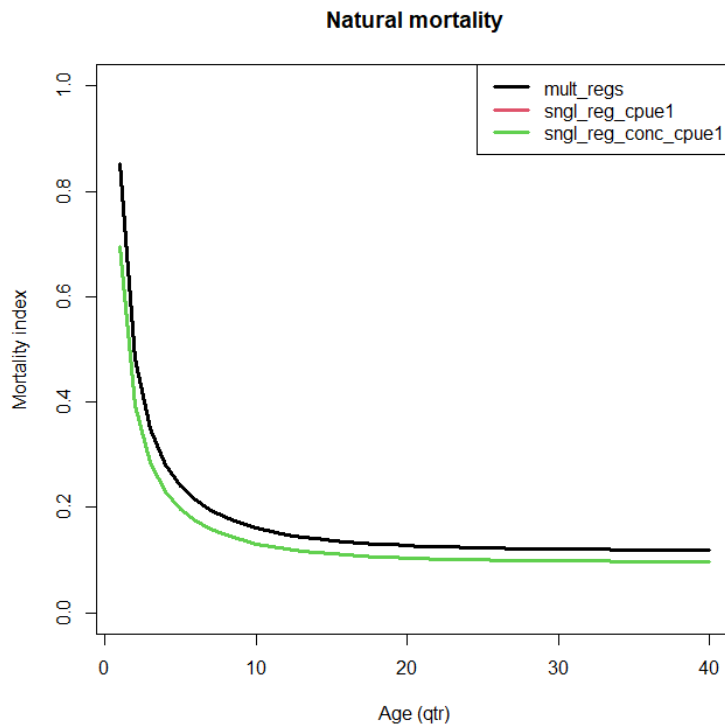


Figure 5. Comparison of estimated Lorenzen natural mortality functions between the multi-region (mult_regs) and single-region (sngl_reg_cpue1, sngl_reg_conc_cpue1) models.

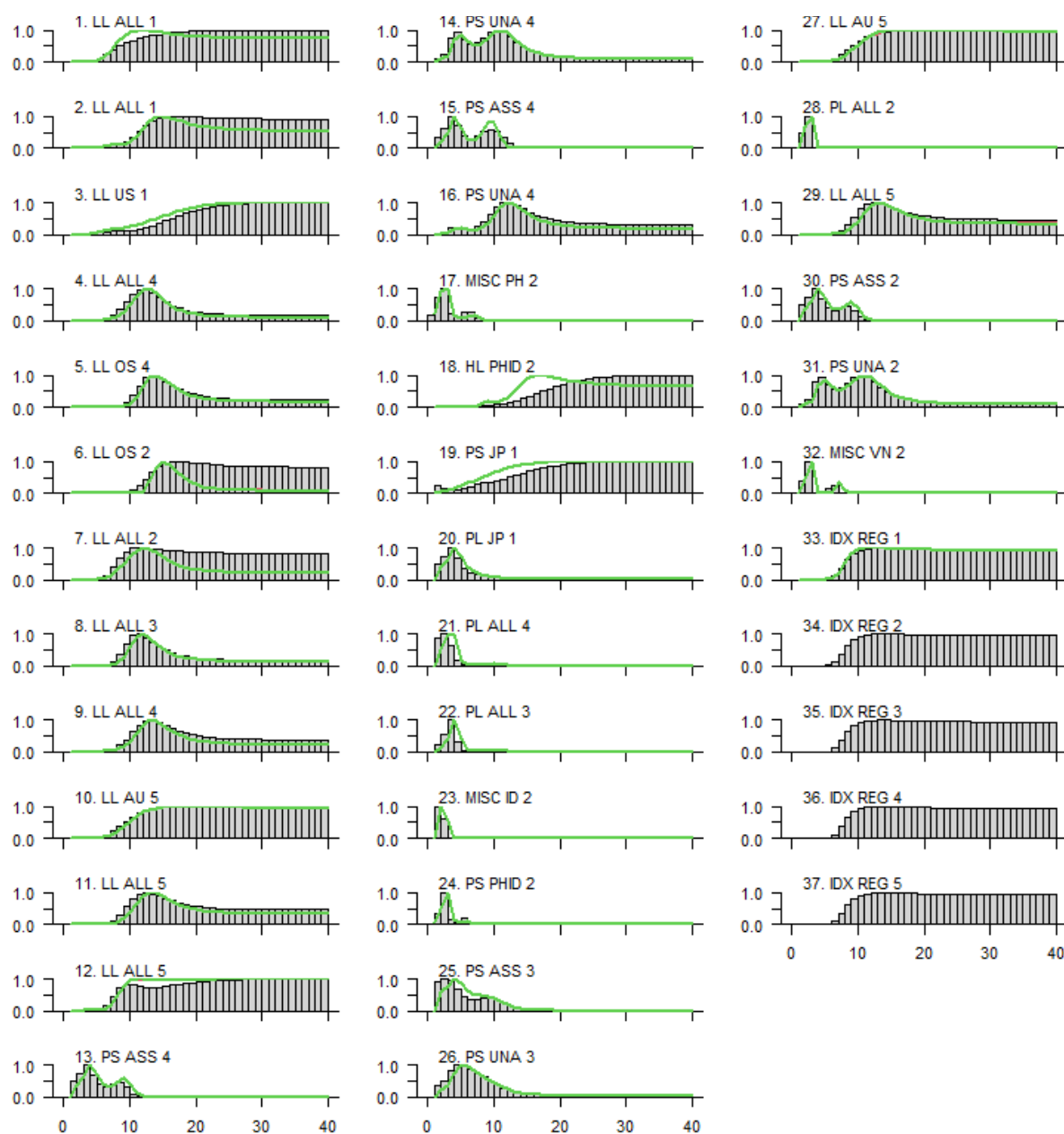


Figure 6. Comparison of estimated selectivity-at-age between the multi-region (mult_regs, histograms) and single-region (sngl_reg_cpue1, sngl_reg_conc_cpue1, red and green lines) models.

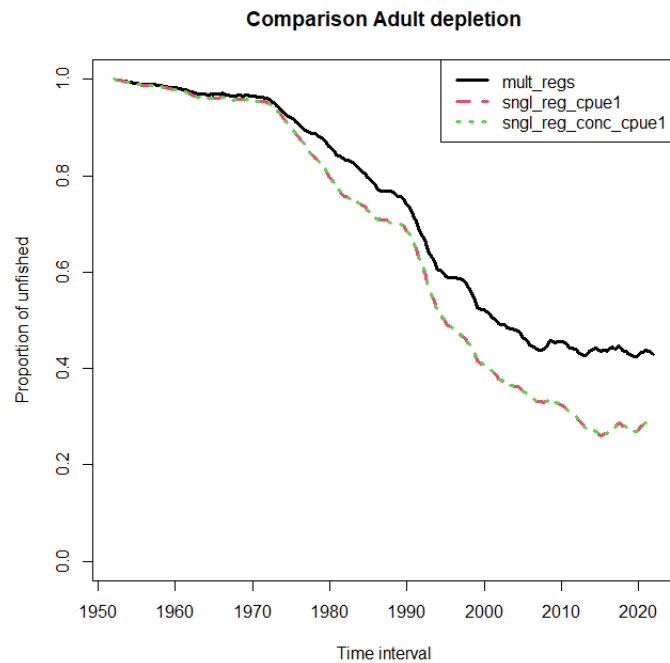


Figure 7. Comparison of estimated spawning biomass depletion between the multi-region (**mult_regs**) and single-region (**sngl_reg_cpue1**, **sngl_reg_conc_cpue1**) models.

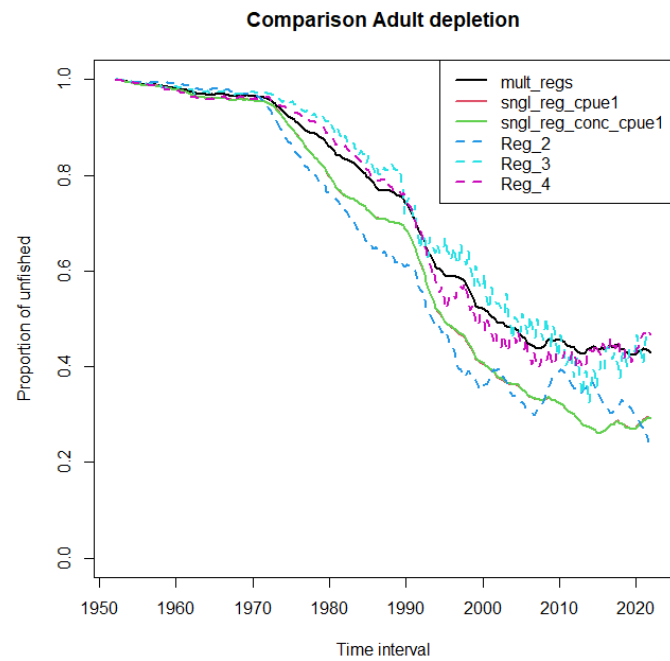


Figure 8. Comparison of estimated spawning biomass depletion between component regions of the multi-region model (**Reg_2**, **Reg_3**, **Reg_4**) and single-region (**sngl_reg_cpue1**, **sngl_reg_conc_cpue1**) models.

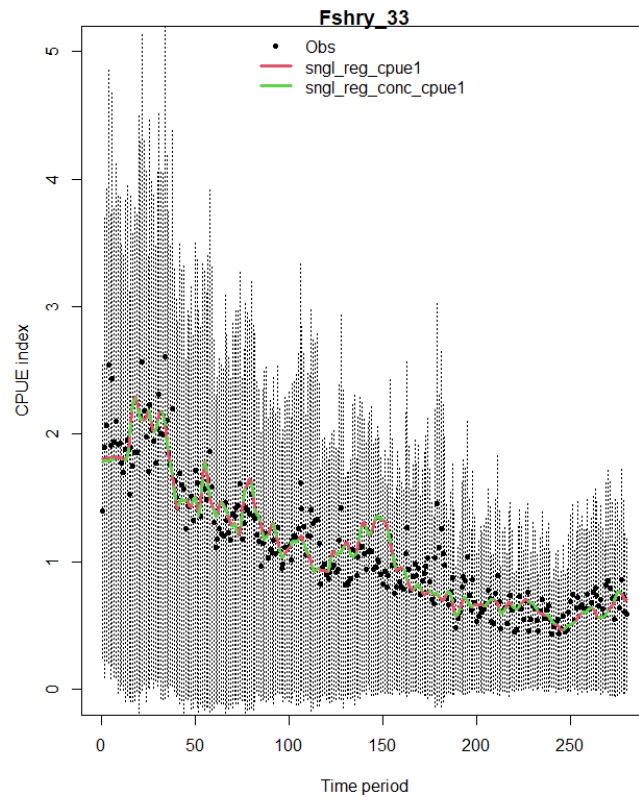


Figure 9. Comparison of the fit to the observed CPUE indices between the single-region models employing the non-concentrated (sngl_reg_cpue1), and concentrated (sngl_reg_conc_cpue1) likelihood formulations. Vertical dashed lines are the index + or - CV.