



The 2018 stock assessment of pāua (*Haliotis iris*) for PAU 5D

New Zealand Fisheries Assessment Report 2019/39

P. Neubauer
L. Tremblay-Boyer

ISSN 1179-5352 (online)
ISBN 978-1-99-000825-2 (online)

September 2019



Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
<http://www.mpi.govt.nz/news-and-resources/publications>
<http://fs.fish.govt.nz> go to Document library/Research reports

© Crown Copyright - Fisheries New Zealand

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
1 INTRODUCTION	2
2 METHODS	2
2.1 Assessment model	2
2.1.1 Summary of changes from recent assessments	2
2.1.2 Population dynamics	3
2.2 Data models	4
2.2.1 Prior distributions	5
2.2.2 Data weighting	6
2.2.3 Technical model details	7
2.3 Stock assessment	7
3 RESULTS	13
3.1 Comparing mid-year biomass formulations	13
3.2 From ADMB to Stan	13
3.3 Growth scenarios: from modified growth assumptions to re-formulated growth	14
3.4 Towards a new base case: reasoning about data weight	14
3.5 Base case results and agreed sensitivity runs	15
4 DISCUSSION	18
5 ACKNOWLEDGMENTS	19
6 REFERENCES	19
APPENDIX A MODEL COMPARISON	21
A.1 Growth model: using data versus informed priors	21
A.2 Reducing growth: to what effect?	23
A.3 Re-formulating growth	29
A.4 Towards a new base-case: data weights	35
A.5 Towards a new base-case: additional constraints	41
A.6 Towards a new base-case: structural assumptions	47
APPENDIX B INDIVIDUAL MODEL RUNS	53
B.1 Base case: Model dwCSLF0.1 uwCPUE2	53
B.1.1 Markov Chain Monte Carlo and posteriors	53
B.2 Main sensitivity run 1: Model dwCSLF0.1 uwCPUE2 low MCV	55
B.2.1 Markov Chain Monte Carlo and posteriors	55
B.3 Main sensitivity run 2: Model dwCSLF0.1	57
B.3.1 Markov Chain Monte Carlo and posteriors	57

EXECUTIVE SUMMARY

Neubauer, P.; Tremblay-Boyer, L. (2019). The 2018 stock assessment of pāua (*Haliotis iris*) for PAU 5D. *New Zealand Fishery Assessment Report 2019/39*. 58 p.

We conducted a Bayesian length-based stock assessment for pāua (*Haliotis iris*) in pāua quota management area PAU 5D. The assessment model used the same population dynamics model as previous assessments, but the data models that linked the population dynamics (process) model with different data types were significantly updated. Unlike previous assessments, the present model used only model-derived inputs and did not fit directly to data. It, therefore, represents a Bayesian synthesis of available information rather than an integrated model that fits to data directly. This development is most significant for the growth process as represented in the model, which previously relied on fitting directly to available data from particular quota management areas. In this model, growth information was given in the form of a rather vague prior, and the model used this information with other inputs to estimate stock-level growth. This approach led to slower estimated growth across all models that were considered to be robust (by the Fisheries New Zealand Shellfish Working Group), which in turn resulted in higher estimated spawning stock biomass.

Other changes in the current stock assessment included a new Dirichlet-Multinomial/Multivariate-Normal formulation for catch sampling length frequency (CSLF) data, and estimation of process error for both CSLF and catch-per-unit-effort (CPUE) inputs.

Relative weighting between these inputs was achieved via explicit constraints on information loss in terms of Kullback-Leibler divergence (KLD) for the two main input sources (CPUE and CSLF). The KLD was then used to calculate posterior information loss in the model for both input types, and to adjust weights in favour of one or the other input.

Two sets of relative weights and two sets of priors for natural mortality M were selected as the main sensitivity runs alongside a base case that gave slightly more weight to CPUE data. Both the base case and the two main sensitivity runs suggested that the stock status was near or at 40% of the unfished spawning stock. For all models, however, the relative available biomass was relatively low (near 25%). Projections at current catch levels suggest that rebuilding towards higher levels of available biomass will be likely to be relatively slow despite current voluntary shelving of 35% of the allowable commercial catch.

1. INTRODUCTION

New Zealand abalone, pāua (*Haliotis iris*), is commercially fished throughout New Zealand, with its management based on different quota management areas (QMAs). The management of pāua fisheries includes regular stock assessments that determine the stock status of a particular QMA. These stock assessments are based on statistical models that estimate the current and projected stock status, as well as the exploitation rate of the portion of the population that is impacted by fishing.

The southern South Island pāua quota management area 5 (PAU 5) consists of three sub-areas, PAU 5A (Fiordland), PAU 5B (Stewart Island), and PAU 5D (Otago and Southland east coast). An earlier stock assessment for sub-area PAU 5D estimated that the fishery was below 40% of virgin biomass (Fu 2013), prompting fishers to shelve (i.e., not fish) 20% of the total allowable commercial catch (TACC) of 89 t for the 2014–15 fishing year, 30% for the 2015–16 and 2016–17 fishing years, and 35% for the 2017–18. A subsequent assessment up to 2016 (Marsh & Fu 2017) estimated that stock biomass was still below the 40% soft limit, but would likely increase subsequently under constant catch levels. Key uncertainties included the growth model, the treatment and representativeness of the catch-per-unit-effort (CPUE) time series and assumptions of natural mortality.

The current report presents the most recent stock assessment for pāua in PAU 5D, including data up to the 2017–2018 fishing year (the New Zealand fishing year spans from 1 October to 30 September the following year). The present assessment is based on the length-based population dynamics model for pāua initially described in Breen et al. (2003), relying on annual catch statistics from commercial, recreational, customary and illegal fishing, length composition of landings, standardised catch-per-unit effort indices, and maturity and growth data from tag-recapture studies to fit the model to data. The current assessment built on previous iterations by implementing a number of novel developments, especially in the treatment of input data with a fully Bayesian approach. Key changes also included a revised growth model that allowed for more representative overall growth of pāua in PAU 5D, and the development of an explicit framework to assess the information contributed to the model by CPUE versus length composition data. Data inputs and processing prior to inclusion in the model are described in a companion report (Neubauer & Tremblay-Boyer 2019).

2. METHODS

2.1 Assessment model

2.1.1 Summary of changes from recent assessments

The stock assessment model employed in this assessment used the general length-based population dynamics model developed by Breen et al. (2003) and employed in subsequent stock assessments (e.g., Marsh & Fu 2017). However, many of the data models have been updated to reflect current knowledge of pāua demographics and developments in statistical practice.

Key developments in the current analysis included:

- Previous assessments used a mixture of direct fitting to input data (e.g., tag-recapture data for estimating growth and maturity data to estimate length-at-maturity) and fitting to model or pre-processed outputs (e.g., catch-per-unit-effort (CPUE) index, assuming error as an output of a generalised linear standardisation model and post-processing e.g., Fu et al. (2017), and proportions-at-length data processed from catch sampling programmes). In contrast, the current model was not directly fitted to data and we used a consistent Bayesian methodology throughout. This included the processing of input data using Bayesian models as necessary, with outputs subsequently used as priors in the assessment.
- Default down-weighting of commercial length-frequency (LF) data was replaced by explicit weighting of CPUE and LF likelihoods in terms of penalties on information loss, and inspection of posterior information loss for both datasets. This update led to an explicit weighting of CPUE vs LF

datasets that increased transparency in regards to data weighting by analysts. Sensitivity analyses informing decisions on data weighting are also reported.

- Growth parameters were estimated using a vague prior derived from a meta-analytic model for all pāua tag-recapture data available across Quota Management Areas (QMAs). The small subset of pāua growth data from PAU 5D alone was deemed to be unrepresentative of growth of the overall stock in this management area, as samples were predominantly from sites with fast-growing, large pāua.
- Length-at-maturity was explicitly linked to growth in the meta-analytic growth model; it was informed by estimated growth *via* correlations of size-at-maturity with growth rate data.
- Except for the recruitment variability standard deviation, all parameter values were estimated in the current iteration of the assessment model. Previous assessments fixed a range of parameter values such as steepness, the standard deviation of recruitment variability and process error.
- All model runs were carried out in Stan (Stan Development Team 2018) to take advantage of improved Markov Chain Monte Carlo (MCMC) methods and diagnostics.

2.1.2 Population dynamics

The main pāua population dynamics are described by Breen et al. (2003), but some changes were recently implemented following recommendations by an international expert review panel for the stock assessment (Butterworth et al. 2015). Specifically, the review panel suggested that CPUE data be fitted to mid-year instead of beginning-of-year available biomass. However, during the working group process for the present assessment, it was found that this change had little to no effect when correctly specified (see below). For this reason, it was agreed to subsequently discard this change to prioritise model parsimony.

Population dynamics are written as beginning-of-year values N at length l , with $l \in [1, L]$ in year y as:

$$N_y = (SN_{y-1} \circ SF_{y-1})G + R_y, \quad (1)$$

where N_y is used to denote the vector of numbers at lengths $1 \dots L$ (i.e., omitting the subscript denotes a vector); $S = \exp(-M)$ is survival from natural mortality, SF is the length-specific survival after fishing (\circ is the element-wise multiplication), G is a $L \times L$ growth transition matrix and R is recruitment, which is evenly distributed among the first five length classes in the model. Element $G_{i,j}$ of G is then the proportion of fish transitioning from length class i to length class j in a given year.

Survival S is derived from M , an estimated parameter in the model. Survival from fishing, SF_y , is calculated by applying selectivity V_y in year y to the overall exploitation rate U_y such that $SF_y = 1 - U_y V_y$, with U_y the ratio of catch TC_y in year y to biomass B_y , or $U_y = TC_y/B_y$. Biomass is obtained by multiplying numbers at length N_y by a vector w of weight-at-length.

Selectivity was assumed to be logistic or inverse log-log. The logistic selectivity describes a smooth increase in selectivity that is symmetric about the size at 50% selectivity (D_{50}). With D_{50} and size at 95% selectivity (D_{95} and D_{95}), selectivity was estimated using

$$V_{y,l} = \frac{1}{1 + \exp\left(\frac{-\log(19)(l - D_{50})}{D_{95}}\right)}, \quad (2)$$

where D_s is a specified offset from the baseline selectivity, usually reflecting voluntary increases in minimum harvest size (MHS), and D_a is an estimated parameter that accounts for often only partial

implementation of increased MHS across a QMA (e.g., only some statistical areas will be fished at higher MHS).

Recruitment R_y was assumed to follow a Beverton-Holt stock-recruit relationship whereby steepness h , equilibrium recruitment R_0 , and annual recruitment deviations R_{dev} determine recruitment from a given spawning stock biomass; spawning stock biomass (SSB) was determined from the weight-at-length relationship w and the proportion mature at length y . The latter was estimated using a prior derived from the growth-maturity model and was adjusted in the model *via* estimated growth and its correlation with maturity (see Neubauer & Tremblay-Boyer 2019 for details).

Growth, G , was calculated from the estimated mean growth in PAU 5D, growth variability (standard deviation), and the proportion p_l of the PAU 5D population that does not grow at any length l . G is then:

$$G_{i,i} = p_i + (1 - p_i) \int_0^{i+k/2} \text{CLN}(\mu_i, \sigma_i), \quad (3)$$

$$G_{i,j} = (1 - p_i) \int_{i-k/2}^{i+k/2} \text{CLN}(\mu_i, \sigma_i) \quad \text{for } j = i + 1, L - 1, \quad (4)$$

$$G_{i,L} = (1 - p_i) \left(1 - \int_0^{L-k/2} \text{CLN}(\mu_i, \sigma_i)\right), \quad (5)$$

where $\text{CLN}(\mu_i, \sigma_i)$ is the cumulative distribution function of the log-normal distribution with mean μ and standard deviation σ on the log-scale.

2.2 Data models

The assessment model was fitted to three main data sources: the CPUE index, length-frequency distributions derived from commercial catch sampling, and growth and maturity priors from tag-recapture and maturation sampling programmes.

CPUE The CPUE index was included on the log-scale and modelled as a normally distributed variable with:

$$\text{CPUE}_y \sim N\left(\text{CPUE}_y^M, \sqrt{\text{OE}_{\text{CPUE}}^2 + \text{PE}_{\text{CPUE}}^2}\right) \quad (6)$$

with PE_{CPUE} the CPUE process error; CPUE_y^M the model-predicted CPUE in year y , calculated as the log of the proportion q of the available biomass in year y , $B_y^{\text{avail}} = (V_y \circ N_y)w$, i.e., $\text{CPUE}_y^M = \log(q) + \log(B_y^{\text{avail}}) * \beta$, with log catchability $\log(q)$ estimated in the model using a normal distribution with prior mean and standard deviation. The parameter β modulates the relation between CPUE and available biomass, and introduces hyperstability for $\beta < 1$.

Previous pāua stock assessments fixed $\sqrt{\text{OE}_{\text{CPUE}}^2 + \text{PE}_{\text{CPUE}}^2}$, whereas we chose to fix only OE_{CPUE} and estimate PE_{CPUE} . This change obviated the need to prescribe a desired fit *a priori* through the fitting of a smoother, and the calculation of expected $\sqrt{\text{OE}_{\text{CPUE}}^2 + \text{PE}_{\text{CPUE}}^2}$ (which was previously assumed the same for all observations). In this version of the model, we instead used the yearly estimated error from the Bayesian CPUE standardisation as OE_{CPUE} , which led to more uncertain CPUE early in the time series (e.g., Catch Effort Landing Return (CELR) data, especially year 1), and increased precision in later years (cf. figure 13 in Neubauer & Tremblay-Boyer 2019).

Catch sampling length frequencies (CSLF) Analogous to CPUE inputs, yearly CSLF proportions at length were included as the mean of centralised log-ratio (CLR) transformed estimated mean proportions from the Dirichlet-multinomial model for raw CSLF data. The model derived estimated length frequencies for each year with associated error, which was considered to be the observation error for CSLF data (analogous to the OE for uncertainty in the CPUE index). The CLR transformation moved the data from an L dimensional simplex (i.e., $\sum_{i=1}^L p_{y,i} = 1$) to an unconstrained L dimensional space. The observed mean CSLF values for year y are thus specified as multivariate-normal (MN) distributed with uncertainty, and correlations specified by a $L \times L$ dimensional covariance matrix of observation error OE_{CSLF_y} . Due to the strong correlations (positive and negative), we used a multiplicative process error formulation, giving:

$$CLR(CSLF_y) \sim MN\left(CSLF_y^M, (1 + PE_{CSLF_y})OE_{CSLF_y}\right), \quad \text{with} \quad (7)$$

$$CSLF_y^M = CLR((V_y \circ N_y)U_y), \quad (8)$$

where $CLR((V_y \circ N_y)U_y)$ are the CLR-transformed predicted selected proportions at length in the model, and $PE_{CLR(CSLF_y)} > 0$ the process error that is additional to OE.

Growth and length-at-maturity The present iteration of the assessment model was exclusively fitted to model outputs from pre-processing models on the input data. This approach was mainly taken for computational convenience; for example, the growth-maturation model takes considerable time to fit growth and length-at-maturity data for pāua, given long-range correlations in the model, the explicit solving of the differential equation for growth and the dataset have been expanded to include measurements across all QMAs.

Instead of fitting to growth data, we specified a relatively uninformative joint prior distribution for mean (log-scale) growth increments $\mu = \mu_1, \dots, \mu_L$, the log of the (log-scale) growth standard deviations $\sigma = \sigma_1, \dots, \sigma_L$, the logit of the proportions z of the total population that exhibits zero growth at each length class, and the logit of the proportion of the total population that are mature in each length class (y). The functional form of the relationship between elements of μ, σ, z , and y is determined by an overall covariance matrix that encodes correlations both within and between these variables. We, therefore, have:

$$\langle \mu, \sigma, z, y \rangle \sim MN(\langle \tilde{\mu}, \tilde{\sigma}, \tilde{z}, \tilde{y} \rangle, \text{cov}(\langle \tilde{\mu}, \tilde{\sigma}, \tilde{z}, \tilde{y} \rangle)), \quad (9)$$

where the tilde designates samples from the posterior distribution of the growth-maturation model.

2.2.1 Prior distributions

The CPUE process error was estimated in the model using a half-normal prior distribution (N^0), with prior standard deviation $\tau_{PE_{CPUE}}$.

$$PE_{CPUE} \sim N^0(\tau_{PE_{CPUE}})$$

Similarly, the CSLF process error was estimated in the model using a half-normal prior distribution, with prior standard deviation $\tau_{PE_{CSLF}}$.

Recruitment deviations (R_{dev}), equilibrium recruitment (R_0), natural mortality (M), catchability ($\log(q)$), length at 50% selectivity (D_{50}) and 95% selectivity offset (D_{95}) were assigned log-normal priors, parameterised in terms of mean and standard deviation (sd; on the log-scale), with the sample mean for R_{dev} forced to one.

Table 1: Default priors used in the pāua stock assessment model (LN=Lognormal, N=Normal, N⁰=half-normal), with prior standard deviation (SD) shown on the log-scale and on the positive scale (CPUE, catch-per-unit-effort; CSLF, catch sampling length frequency).

Parameter	Symbol	Prior	Mean	SD	SD (pos)
Equilibrium recruitment	R_0	LN	13.5	0.5	4.4×10^5
Recruitment deviations	R_{dev}	LN	0	2	54.1
Natural mortality	M	LN	$\log(0.12)$	0.2	0.02
Catchability	q	LN	-13	2	0
Length at 50% selectivity	D_{50}	LN	$\log(123)$	0.03	3.69
95% selectivity offset	D_{95}	LN	$\log(5)$	0.5	3.02
Selectivity increase	D_a	LN	0	1	2.16
Steepness	h	Beta	0.71	0.12	
CPUE process error	PE_{CPUE}	N ⁰ (0.05)	0.04	0.03	
CSLF process error	PE_{CSLF}	N ⁰ (2)	0.80	0.6	

Steepness h was estimated in this iteration of the assessment model and was assigned a beta distribution prior with parameters a and b , with $a = 10$ and $b = 4$ the default prior, leading to a wide prior that put most of the weight at $h > 0.5$ (see Table 1 for other default priors).

2.2.2 Data weighting

In this assessment, we used the Kullback-Leibler divergence (KLD) as an alternative method for data weighting *via* a measure of information loss. The method relies on the premise that there should be no *a priori* preference for any one dataset, and that relative weight should emerge as part of the analysis and model refinement process. In addition, it makes use of the total distribution for the compositional data rather than just the first moment (e.g., mean length).

The KLD measures the information loss from a signal like the distribution of input data (e.g., the CPUE index and observation error) to the distribution needed to fit the model (e.g., predicted CPUE and observation plus process error). We can minimize the information loss from a data source by using a scaled KLD as weight, which is mathematically equivalent to multiplying the log-likelihood by a constant. While this practice is generally equivalent to re-weighting using an abstract “sample size” measure for the multinomial distribution, for example, it has a clear interpretation in terms of information loss. In addition to the increased transparency from data weighing decisions, we can directly compare the KLD between CPUE and compositional data (CSLF) to assess the information loss for each dataset for a particular weighting. This comparison provides direct knowledge of the relative information gained by the model from CSLF and CPUE data.

For the model above, the mean KLD across Y years for the normally-distributed log CPUE index is:

$$KLD_{CPUE} = 1/Y \sum_y 0.5 \left(\log \left(\frac{OE_y^2 + PE^2}{OE_y^2} \right) + \frac{OE_y^2 + (CPUE_y - CPUE_y^M)^2}{OE_y^2 + PE^2} - 1 \right) \quad (10)$$

Similarly, for the CSLF inputs, the multivariate equivalent (scaled to match the univariate KLD) is:

$$KLD_{CSLF} = 1/Y \sum_y 0.5 L^{-1} \left(\log \left(\frac{|\Sigma_y^M|}{|OE_y|} \right) + \text{tr}((\Sigma_y^M)^{-1} OE_y) - L + \right. \\ \left. (\text{CSLF}_y - \text{CSLF}_y^M)(\Sigma_y^M)^{-1}(\text{CSLF}_y - \text{CSLF}_y^M) \right) \quad (11)$$

with $\Sigma_y^M = \text{OE}_y(1 + \text{PE})$ the covariance matrix including OE and PE. The latter terms refer to observation and process error for each data source, respectively.

2.2.3 Technical model details

Initialisation The model was initialised for a period of 60 years with constant recruitment at R_0 and no fishing.

MCMC configuration All MCMCs were run using the no-u-turn-sampler (NUTS) implemented in Stan (Stan Development Team 2018). The Stan language is more efficient than conventional Metropolis Hastings or Gibbs sampling for MCMC, and also provides diagnostics that can signal biased MCMC transitions (divergences) and potential bias in estimated quantities from these transitions. All MCMC chains were, therefore, monitored for divergent transitions to ensure that MCMCs were not biased. Eight independent chains were run over 6000 iterations, with the first 1000 samples discarded for each chain, and a further 2000 samples saved for inference and post-processing.

Computing environment The model was implemented in Stan (Stan Development Team 2018) and launched in R *via* RStan (R Core Team 2018, Stan Development Team 2018), using a version of the previous assessments' ADMB (AD model builder) code ported to Stan (by Chris Nottingham, University of Auckland) as a basis for development. We modified the code to ensure that we could exactly reproduce past assessments using Stan for Bayesian computations.

We favoured Stan for continuing the development of the assessment model for several reasons: first, Stan has robust and efficient MCMC routines that are straight-forward to run in parallel over many cores and on remote (cloud-based) computing platforms; in addition, Stan's MCMC provides valuable diagnostics about the MCMC that can help diagnose potential bias or poor performance of the MCMC; lastly, Stan's considerable library of standard probability densities and available post-processing libraries in R mean that integration with pre-and post-processing in R is more straightforward.

We made extensive use of built-in functions in Stan, especially formulations of multivariate normal densities based on Cholesky-factorised covariance matrices to gain efficiency. All post-processing was performed in R, using R Markdown (Allaire et al. 2018) and knitr (Xie 2014). All code was run within Docker containers for full reproducibility; all code was version controlled on Github.

Automation All input models, assessment runs and post-processing were launched on a cloud-based reproducible reporting system (at Dragonfly Data Science), and implemented as a data processing framework from input data to final reports. This framework enabled assessments to be re-run with updated data with a few simple commands. We intend to use this feature in the future to automatically update management procedures for pāua on the basis of biomass assessments.

2.3 Stock assessment

The stock assessment involved a progression from the ADMB model used in previous assessments to the Stan model used currently (see Appendix A, Figures A-1 to A-32 for details). The progression included the following steps:

1. Comparison of mid-year biomass for fitting CPUE The previous assessment model was compared using three formulations for fitting CPUE. The first formulation relied on previously used fitting of CPUE to beginning-of-year biomass. The second formulation used a comparison with mid-year biomass – i.e., the biomass calculated mid-year by applying half of natural mortality ($S^{\text{mid}} = 0.5S$) and by applying

half the year's growth via $G^{\text{mid}} = 0.5G$. This formulation was used in the previous ADMB model. This formulation can be expressed as

$$N_t^{\text{mid}} = 0.5 * N_t + 0.5 * N_t * S * SF_t * G_t. \quad (12)$$

(Note that this formulation is different from the first iteration of calculating mid-year biomass in the ADMB model, which was $N_t^{\text{mid}} = N_t * S^{\text{mid}} * SF_t^{\text{mid}}$ and did not consider growth). The most recent formulation, however, still fails to consider that growth is a non-linear process; the latter means that the growth matrix resulting from applying the growth model for $t = 0.5$ is not the same as taking half of the yearly growth ($G^{\text{mid}} \neq 0.5G$). Nevertheless, as the growth model is not process based, it is unclear how the population standard deviation for smaller-than-yearly time-steps ought to be calculated.

A third alternative is:

$$N_t^{\text{mid}} = N_t * S^{\text{mid}} * SF_t^{\text{mid}} G^{\text{mid}}, \quad (13)$$

with G^{mid} calculated from the growth model given $t = 0.5$ and assuming the standard deviation evolves linearly with time.

2. From ADMB to Stan A set of model runs was performed to compare outputs from the previous stock assessment for PAU 5D to the transferred assessment in Stan. These model runs were mainly performed to ensure that previous outputs could be reproduced with the new code base.

3. Re-formulating PAU 5D growth The previous assessment was sensitive to input assumptions about growth and natural mortality (Marsh & Fu 2017): with a negative-exponential growth model, the estimated stock status was more than 15% lower than the status estimated using an inverse logistic model. There was little information to base the model selection on, with one model performing better with available growth data (inverse logistic), whereas the other model produced closer fits to LF data (negative exponential). Comparison with growth data from other QMAs and with additional growth data collected since 2016 revealed that the available growth data for PAU 5D was from sites with some of the fastest and largest growing pāua in the country – although the new data, represented by relatively few recaptures, also showed that sites with slow growing pāua are present in PAU 5D (see Neubauer & Tremblay-Boyer 2019). In addition, the practice of fitting directly to these data introduced considerable constraints on growth (Figure 1), with little uncertainty about mean growth and realised growth variability.

Consultation with pāua fishers suggested that there are a number of areas with slow-growing pāua and substantial biomass in PAU 5D. These areas are occasionally fished, but annual catch is relatively low, due to their low productivity. By comparison, areas with fast-growing pāua are more regularly fished, but these areas may have substantially lower standing biomass as pāua rapidly recruit into the fishery. This difference is a challenge for the stock assessment model, as the latter attributes growth over the entire population, not just fast-growing sites. This aspect, in turn, may lead to a mismatch between growth estimated in the model from the tag-recapture data, and the actual growth pattern across the entire population. To assess the degree to which this bias might influence the model, we ran different model scenarios in which growth was reduced by a factor of 0.9, 0.8, and 0.6 by modifying the mean growth parameters accordingly.

One possibility to address the potential conflict between growth data and QMA-wide growth is to re-formulate the growth process in the model and to give less weight to growth data from sites with fast-growing pāua. As the proportion of fishing grounds that grow particular rates is unknown, we opted for an informed prior on growth that was estimated from all available growth data for South Island pāua. This model choice means that we assumed that i) growth rates in PAU 5D are *a priori* comparable to those in all other regions of the South Island, and ii) that to some extent, the mean and variability of growth rates across all sampled sites are informative about growth rates across all of PAU 5D. With these assumptions,

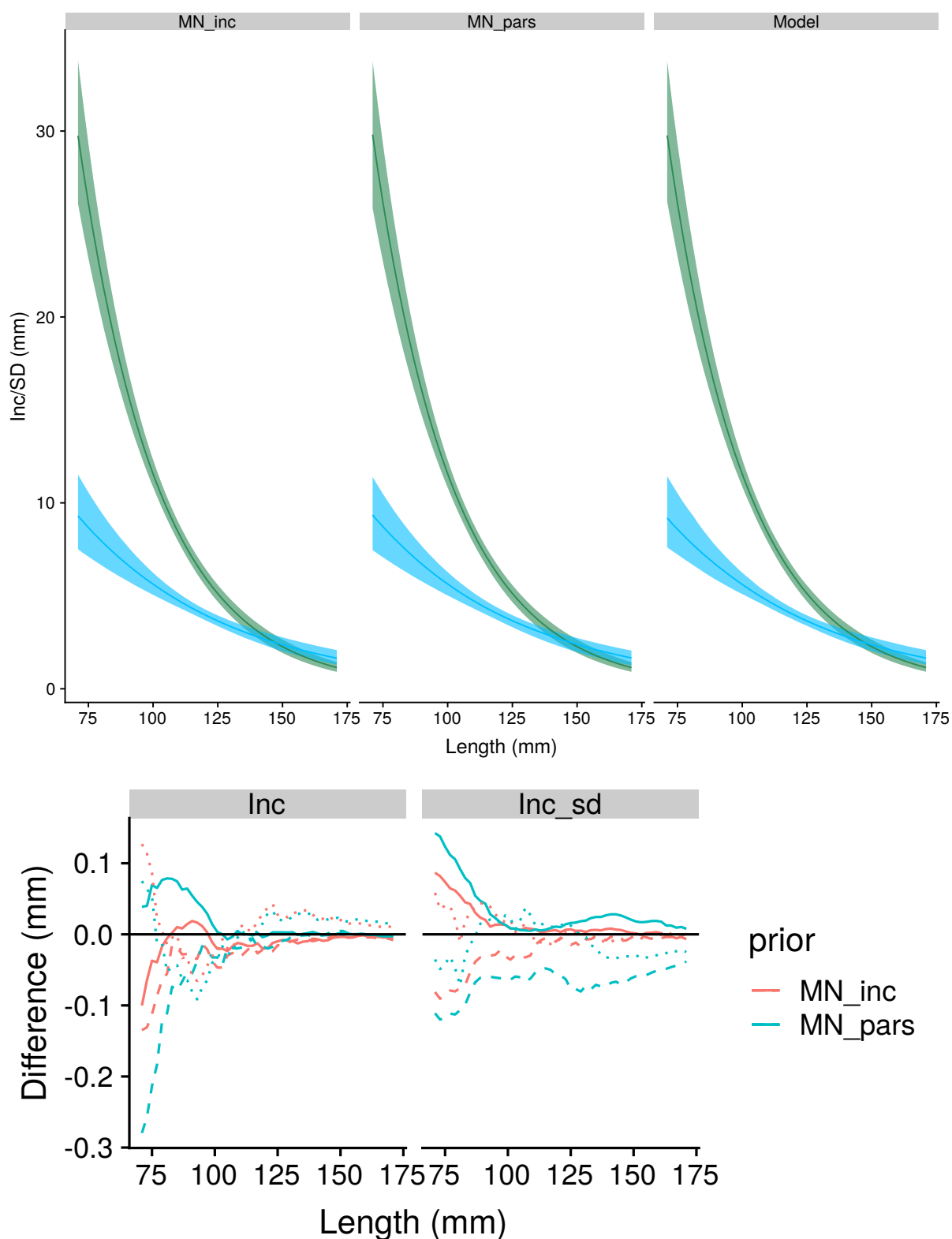


Figure 1: Modelling of pāua growth in management area PAU 5D. Top row: Posterior distribution for mean growth (green) and growth standard deviation (blue), with priors implied by a multivariate normal (MN) distribution estimated from fitted Markov Chain Monte Carlo output for growth increments (inc) and growth parameters (pars), and posterior distributions for a negative exponential growth model fitted to data from PAU 5D. Bottom row: Difference from the fit for simulations from the MN priors for growth.

we constructed an informative prior on growth that enforced some structure on growth in the model, but left it to the model to find the actual growth with the closest fit given all other data sources.

We also made the assumption that for each length l , we can describe growth by a proportion of growing individuals at sites with $L_\infty > l$, and a proportion of individuals with zero growth beyond l (i.e., at sites with $L_\infty \leq l$). This assumption leads to a hurdle model for growth, where an increasingly small proportion of pāua still growing at large sizes. This formulation avoids assumptions about large proportions of negative growth. Lastly, the growth model estimated the connection between the growth process and maturation (i.e., energy invested in gonad development as opposed to somatic growth). This connection means that the growth model implied a joint prior on growth and maturity (Figure 2).

The new growth/maturation formulation makes the growth/maturity process equal to other “data” sources in the model—for neither CPUE nor CSLFs the model is fitted to raw data. Nevertheless, in a Bayesian model, there is a mathematical equivalence between fitting these components outside of the model and within the model, as long as the posterior distribution of the parameters for any dataset are adequately described, based on the raw data and priors. In practice, however, we relied on a parametric representation of this posterior, which necessarily introduced some error. Nevertheless, for the growth data, we showed that by using simulations, only a small error was introduced between fitting the growth model directly and describing the growth process as a MN prior (calculated from the MCMC outputs of the fitted model) on growth parameters or increments and (log) standard deviations (see Figure 1).

4. Towards a new base case Once more flexibility was given to the model to determine QMA-wide growth patterns, we aimed to improve model fits by adjusting dataset weights in the model. The default weighting gave most weight to CPUE and comparatively little weight to CSLF data. Modifying this status quo, however, raised the question of what constitutes a reasonable weight for either dataset. In addition, we were uncertain of the effect of fixing different parameters in the assessment model and, therefore, chose to free up all parameters before sequentially constraining them towards a reasonable solution. This approach meant fitting steepness (previously fixed at $h = 0.75$), recruitment variability (previously $\sigma_R = 0.4$), CPUE and CSLF process error (previously fixed via fitting a smoother through CPUE and calculating the SD and iterative-re-weighting, respectively) in the model. In addition, the new formulation for CSLF data allowed us to represent PE and OE for CSLF data explicitly, and to derive realised weights via the KLD. We assigned a relatively wide prior to M , but chose to make use of domain knowledge to constrain priors for R_0 , selectivity and q . We realised after a number of trials that estimating recruitment variability introduced considerable instability. For this reason, we explored scenarios of fixing σ_R to different values.

We subsequently explored a grid of models with different degrees of constraints on process error terms, M , and relative weights for CPUE and CSLF. The realised weight was calculated using KLD (as described above; see Table 2 for the grid of scenarios).

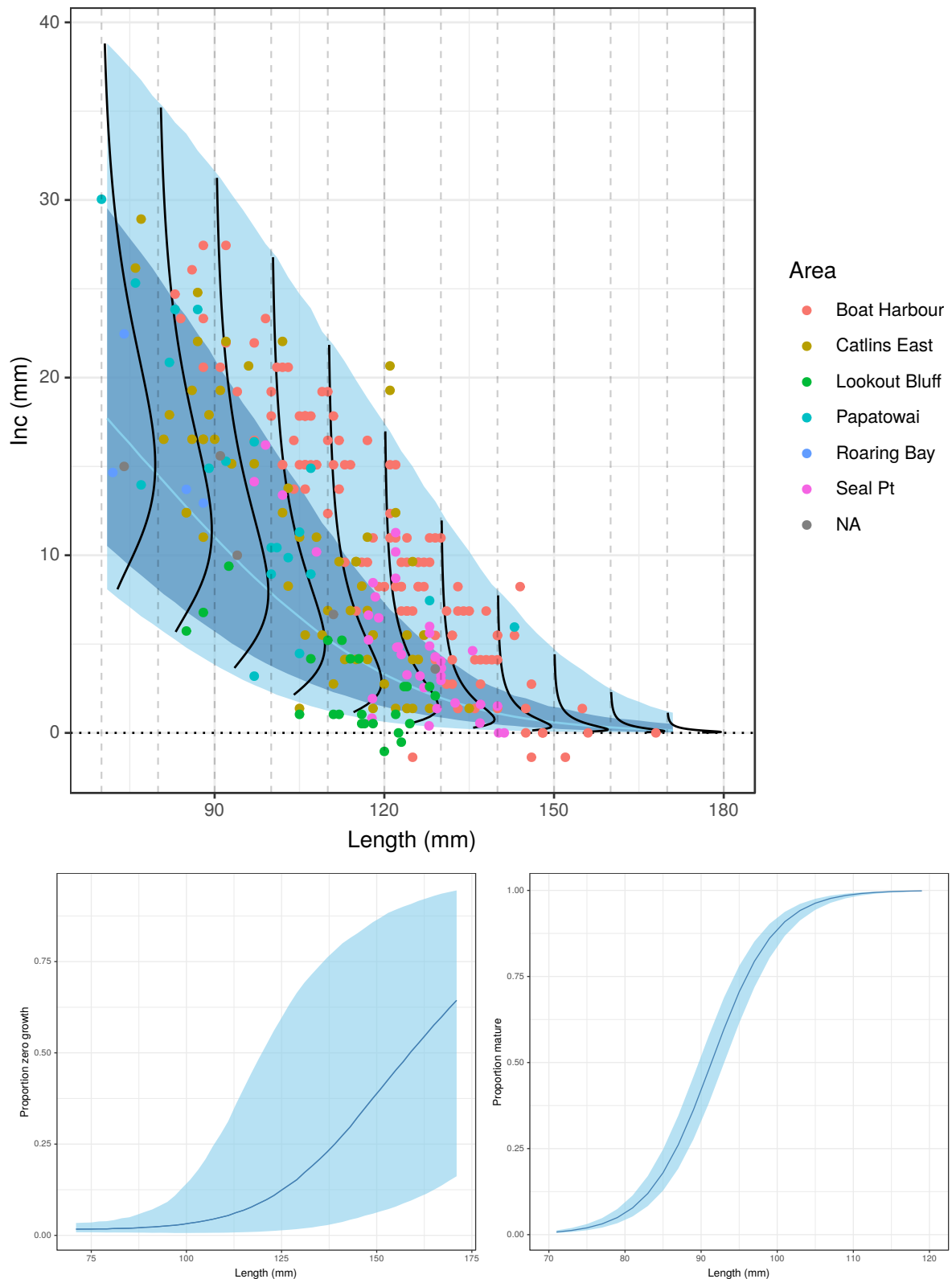


Figure 2: Priors implied from a meta-analysis of growth of South Island pāua, based on model-fitting to all tag-increment data from different areas (NA, no available area data). Shown is the joint prior for positive growth increments (top), proportion of local populations not growing at a particular size l (bottom left), and population level maturity (bottom right). For positive increments, dark blue shading shows uncertainty about mean growth, light blue line indicates posterior median for mean growth; light blue area shows the posterior median for the population standard deviation applied to mean growth; black line indicates the implied distribution of growth at the median of the prior.

Table 2: Different model runs for the stock assessment of pāua in management area PAU 5D. Stock assessment scenarios and associated priors for M , σ_R and catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) process error (PE). Likelihood weights for CSLF and CPUE data are multipliers on the log likelihood (equivalent to penalising the Kullback-Leibler divergence). dw refers to down-weighting, uw refers to up-weighting, low MCV indicated lower prior SD for M , high sigmaR suggests stronger constraint on σ_R . PMS: constrained process error, M and σ_R priors; RWQ: random walk q; 2 CPUE: Split CPUE time series; CPUEpow0.75: hyper-stable CPUE.

Run	Label	M prior sd	σ_R	CPUE pe	CPUE weight	CSLF pe	CSLF weight
1	growth vB rfx	0.10	0.40				
2	growth vB LLWx1	0.10	0.40				
3	growth vB LLWx0.1	0.10	0.40				
4	penalty model noconstr	0.20	2.00	1.00	1.00	5.00	1.00
5	base low PE	0.20	2.00	0.05	1.00	2.00	1.00
6	base low PMS	0.10	0.40	0.05	1.00	2.00	1.00
7	dwCSLF0.1	0.20	2.00	0.05	1.00	2.00	0.10
8	dwCSLF0.01	0.20	2.00	0.05	1.00	2.00	0.01
9	dwCSLF0.1 uwCPUE2	0.20	2.00	0.05	2.00	2.00	0.10
10	base low PE lowMCV	0.10	2.00	0.05	1.00	2.00	1.00
11	dwCSLF0.1 low MCV	0.10	2.00	0.05	1.00	2.00	0.10
12	dwCSLF0.01 low MCV	0.10	2.00	0.05	1.00	2.00	0.01
13	dwCSLF0.1 uwCPUE 2 low MCV	0.10	2.00	0.05	2.00	2.00	0.10
14	base low PE high sigmaR	0.20	0.40	0.05	1.00	2.00	1.00
15	dwCSLF0.1 high sigmaR	0.20	0.40	0.05	1.00	2.00	0.10
16	dwCSLF0.01 high sigmaR	0.20	0.40	0.05	1.00	2.00	0.01
17	dwCSLF0.1 uwCPUE2 high sigmaR	0.20	0.40	0.05	2.00	2.00	0.10
18	dwCSLF0.1 uwCPUE2 high sigmaRM	0.10	0.40	0.05	2.00	2.00	0.10
19	dwCSLF0.1 uwCPUE 2 RWQ	0.20	2.00	0.05	2.00	2.00	0.10
20	dwCSLF0.1 uwCPUE2 2CPUE	0.20	2.00	0.05	2.00	2.00	0.10
21	CPUEpow0.75	0.20	2.00	0.05	2.00	2.00	0.10

3. RESULTS

3.1 Comparing mid-year biomass formulations

The exact formulation of the mid-year biomass had little impact on biomass trends or stock status estimates for PAU 5D (Figure 3), with only small differences evident in trends and status estimates. For PAU 5B, no detectable differences were found between the formulations (Figure 4), and we attributed the slight changes in PAU 5D to the overall sensitivity of the model to assumptions in this area (subsequent discussions in the Shellfish Working Group (SFWG) led to the omission of the mid-year formulations to ensure model parsimony).

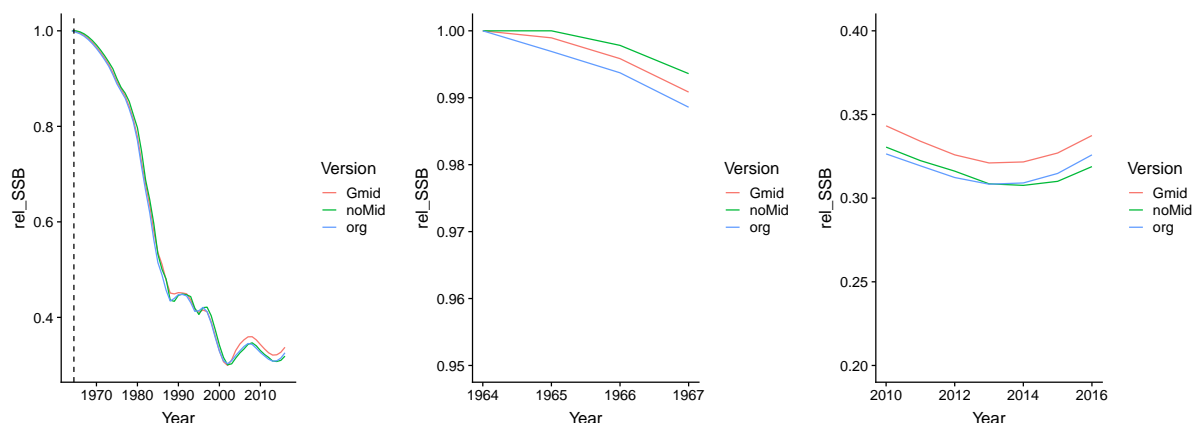


Figure 3: Comparison of the original formulation used in previous pāua stock assessments to calculate mid-year biomass (org) with an updated formulation based on corrected mid-year growth (Gmid) and a formulation where catch-per-unit-effort was fitted to beginning-of-year biomass in PAU 5D (noMid)(rel_SSB, relative spawning stock biomass).

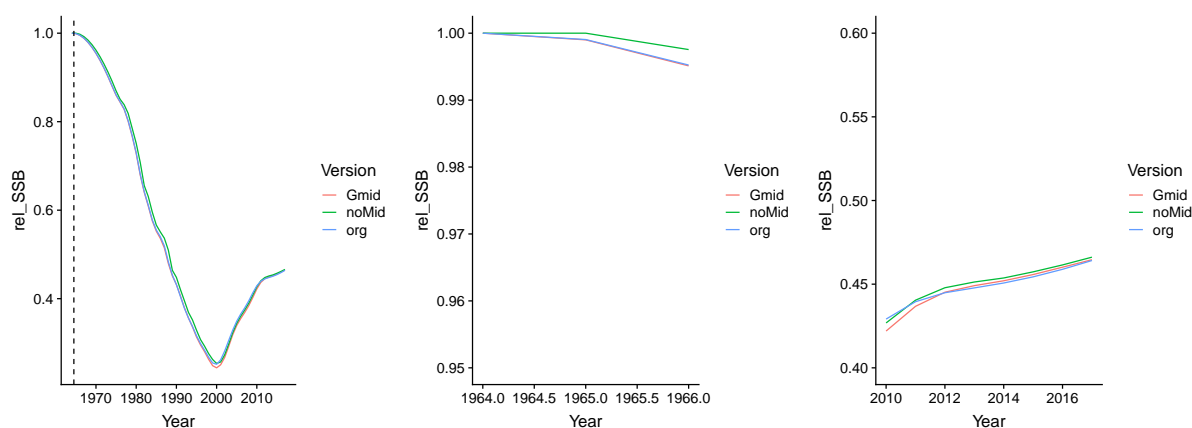


Figure 4: Comparison of the original formulation used in previous pāua stock assessments to calculate mid-year biomass (org) with an updated formulation based on corrected mid-year growth (Gmid) and a formulation where catch-per-unit-effort was fitted to beginning-of-year biomass in PAU 5B (noMid)(rel_SSB, relative spawning stock biomass).

3.2 From ADMB to Stan

We compared Stan outputs with the previous ADMB model using the relative spawning biomass as a metric that aggregates all parameters in the model in a quantity that is sensitive to change, especially in PAU 5D. We found that ADMB results were replicated to at least four significant digits for MPD (MAP) runs, leading to nearly identical posterior distributions.

3.3 Growth scenarios: from modified growth assumptions to re-formulated growth

Setting an informed prior on growth parameters compared with fitting growth directly in the model led to only small differences in estimated growth (Figure A-1), with slightly slower growth for models with the multivariate normal formulation. Due to the strong sensitivity of the assessment model to growth assumptions, however, estimated stock status in the models with multivariate normal priors shifted slightly towards higher status (Figure A-2).

By modifying the prior for growth parameters (i.e., by forcing growth to be slower (Figure A-3), we produced a strong response in model outcomes (Figure A-4): slower growth led to lower estimates of M , but higher estimates for equilibrium recruitment (R_0) and relative spawning stock biomass (Figure A-8). This outcome was largely due to larger amounts of biomass being estimated at sizes below the selected size (Figure A-5). Nevertheless, this modification did not markedly improve LF fits (Figure A-7), and led to progressively poorer CPUE fits (Figure A-6).

Freeing up growth (vB growth scenario) led to lower estimates of growth rates and natural mortality (Figure A-9), but did not greatly improve fits to CSLF data, and led to a smoothed fit through CPUE (Figures A-12, A-13). It also led to higher stock status estimates than the previous negative-exponential model (Figures A-9, A-14), owing to slower growth and larger standing biomass (Figures A-10, A-11). Progressively up-weighting CSLF data in the previous multinomial CSLF model (by using arbitrary weights of 0.1 and 1 – the latter meaning no down-weighting) led to improved fits to CPUE and CSLF data. It also led to fast growth and higher M estimates, but halved the estimate for stock status (for a CSLF weight of 1) owing to a much smaller estimate of R_0 and, by consequence, B_0 (Figure A-9). Nevertheless, these model fits were considered dubious, with a bimodal distribution for population length-frequencies (Figure A-11) and growth estimates that were beyond the observed data (Figure A-10).

3.4 Towards a new base case: reasoning about data weight

This section describes the new base case on the basis of model fits and relative information loss for the two key datasets that determine abundance estimates (CPUE and CSLF). We first developed a model with the Dirichlet Multinomial (DP)/ Multivariate Normal prior for CSLF data, with no external forcing (e.g., data weighting). Without prior weighting and with all parameters free (except σ_R fixed at a large value), we found that the model produced divergent MCMC iterations and poor convergence, indicating potential bias. This model was, therefore, not retained. By constraining process error priors and priors on M and σ_R , we could fit this model without divergent iterations; however, this model produced poor fits to CPUE data with high CPUE KLD, but close fits to CSLF data with a KLD for CSLF of about 0.56 (0.53;0.56;0.60) versus 1.41 for CPUE (Figures A-15, A-18, A-19). This model showed many characteristics of the previous model with CSLF not downweighted (cf. Figure A-11), suggesting that, without input weight adjustments, the DPM-MN prior for proportion data provides similar results to the previous multinomial prior.

Progressively down-weighting CSLF by a factor of 0.1 and 0.01 led to progressively improving CPUE fits and lower estimates for growth rates, without markedly impacting on CSLF fits (Figures A-16, A-18, A-19). At a weight of 0.1 for CSLF, the KLD for both CPUE and CSLF data became comparable (CPUE KLD 0.85 vs CSLF KLD 0.73 (0.65;0.73;0.83);), with a slightly higher information loss for CPUE inputs (Figure A-15). This difference was inverted by adding weight to CPUE data, which forced a better CPUE fit (KLD 0.67), with little impact on the CSLF fit. The Shellfish Working Group selected the latter case as the base case, noting that other models that did not upweight CPUE provided similar outcomes in terms of stock status.

Constraining M and σ_R in this base-case model had opposing effects: constraining M via a tighter prior led to lower estimates of M (Figure A-21), with lower stock status and estimated available biomass (Figure A-26). Constraining σ_R led to higher M estimates and considerably closer fits to CPUE data (Figure A-24); however, it also led to lower relative spawning biomass (stock status) than the base case with no constraints (Figures A-21, A-26). Constraining both M and σ_R resulted in similar fits to the

base-case model, with slightly higher M , but slightly lower stock status (Figure A-21). Overall, models with constrained σ_R led to unrealistic growth patterns in the form of fast growth beyond observed tag-increments (Figure A-22) that were not considered credible.

We tested more structural sensitivities in the form of models with two CPUE series (i.e., allowing for catchability to change by splitting the CPUE index between years with CELR and with Pāua Catch Effort Landing Return (PCELR) forms), adding a random walk on catchability; we also introduced hyper-stability in CPUE by setting $\beta = 0.75$. Splitting the time series led to a higher status estimate without markedly affecting any other quantity (Figure A-27). This increase in estimated stock status resulted from estimates of slightly lower recent catchability. This outcome differed to the model with time-varying catchability q estimated in a random walk (Figure A-21). The latter model attributed a significant part of CPUE process error to changes in q , with a probability of 0.83 that q exhibited a recent increase (Figure 5). This scenario also led to lower estimates of M . The overall trend was markedly different from the base case, with a much more recent decline in biomass (Figure A-32). This decline was to a stock level that was estimated to be higher than in the base case, but rapidly declining since the early 2000s. Estimated growth was similar in all scenarios except for the q random walk, which displayed lower growth and a higher proportion of stock with zero growth for larger length classes (Figure A-28).

Considering that the random-walk model for q resulted in high estimates for biomass in the early 2000s, the SFWG concluded that this model was likely to be unrealistic and a poor reflection of reality. Nevertheless, there was a general expectation that catchability probably increased to some extent over the past decades. In spite of this expectation, the model with two CPUE time series estimated almost equal catchability for both periods and, therefore, also contradicted the random-walk model for q .

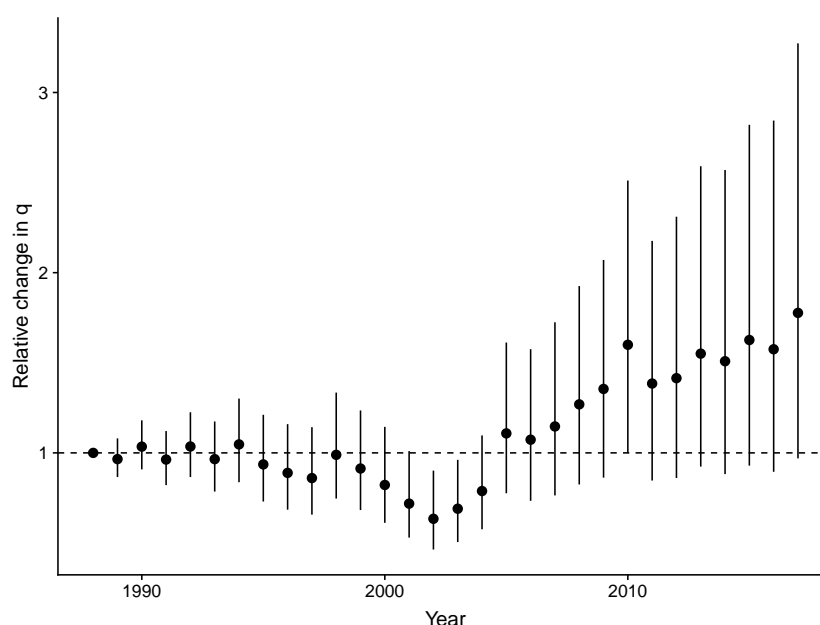


Figure 5: Change in estimated catchability q of pāua (using a random walk model) from the first year with catch-per-unit-effort data.

3.5 Base case results and agreed sensitivity runs

The proposed base case (see Appendix B.1) suggested a relatively flat trend in spawning stock biomass over the past seven years, following a slow downwards trend from 2005 to 2011. The base case also indicated a high probability that the stock is currently near target spawning stock biomass (Table 3, Figure ??), with little to no probability that it is below the soft limit of $0.2SSB_0$. This inference was supported by all sensitivity runs (Table 3). Nevertheless, relative available biomass was markedly lower than the spawning stock biomass (Figure ??), meaning that a considerable part of the spawning biomass was

below the minimum harvest size (Figure A-29). The latter aspect was supported by assessing the same quantity in terms of numbers relative to total number of pāua in the model: only around 20% of pāua were available for harvest at any time (Figure ??). Projections suggested relatively stable SSB for scenarios of current catch and 10% or 20% increased or decreased catch (Table ??). For all catch scenarios, available biomass was projected to slowly increase, although this increase is somewhat uncertain (i.e., 60% likelihood on an increase in three years over current available biomass at current catch).

Two scenarios were agreed as main sensitivity runs that bracketed estimated stock status in the base-case run. The first scenario was the base case with a more restrictive prior for M (log-normal SD of 0.1 instead of 0.2, which forced M to a lower point in the assessment (see Appendix B.2); it also led to lower recent stock status, all else being equal (Table 3; Figure ??). Nevertheless, this scenario also suggested a recent upturn in the fishery, despite a lower stock status estimate. This model run suggested a potentially stronger impact from recent shelving measures than the base case would suggest. Projections from this scenario largely agreed with those from the base case (Table ??).

A second main sensitivity scenario did not upweight CPUE and, therefore, only downweighted CSLF data (see Appendix B.3). This sensitivity run resulted in declining recent spawning stock biomass trends (Figure ??), despite resulting in slightly higher estimates for current stock status (Table 3). The declining trend continued for projections in this scenario regardless of the applied catch (i.e., even at lower-than-current catch; Table ??). For both main sensitivity runs, the probability of stock status being or falling below the soft limit was close to zero.

Table 3: Model runs for the stock assessment of pāua in management area PAU 5D. The base case is indicated with an asterisk, the two main sensitivity runs are indicated with plus signs; section references for detailed results are given for these runs. Posterior quantities for data weight (catch-per-unit-effort (CPUE) and catch sampling length frequency (CSLF) Kullback-Leibler divergence (KLD)), natural mortality M, stock status (relative spawning stock biomass), relative available biomass and probability of the stock status being below the soft limit ($0.2SSB_0$). Numbers are posterior means, with the 0.025, 0.500 and 0.975 posterior quantiles in parentheses.

Run	KLD CPUE	KLD CSLF	M	Stock status	Available	P(SSB>0.2)	Figure section
growth vB rfx			0.12 (0.10;0.12;0.14)	0.56 (0.42;0.56;0.71)		1.00	
base low PE lowMCV	1.41 (1.16;1.41;1.66)	0.56 (0.53;0.56;0.60)	0.17 (0.14;0.17;0.19)	0.32 (0.22;0.32;0.47)	0.24 (0.17;0.24;0.33)	0.99	
dwCSLF0.1 ⁺	0.85 (0.70;0.84;1.05)	0.73 (0.65;0.73;0.83)	0.15 (0.12;0.14;0.19)	0.46 (0.28;0.44;0.71)	0.30 (0.19;0.29;0.46)	1.00	B.3
dwCSLF0.01	0.84 (0.68;0.83;1.03)	1.16 (0.87;1.13;1.64)	0.11 (0.08;0.11;0.15)	0.58 (0.38;0.57;0.81)	0.47 (0.29;0.47;0.68)	1.00	
dwCSLF0.1 uwCPUE2*	0.67 (0.53;0.67;0.82)	0.74 (0.66;0.73;0.84)	0.15 (0.12;0.15;0.19)	0.42 (0.25;0.40;0.65)	0.26 (0.17;0.25;0.39)	1.00	B.1
dwCSLF0.1 uwCPUE2 low MCV ⁺	0.69 (0.53;0.68;0.92)	0.74 (0.66;0.74;0.84)	0.13 (0.11;0.13;0.16)	0.38 (0.24;0.36;0.56)	0.24 (0.16;0.23;0.35)	1.00	B.2
dwCSLF0.1 uwCPUE2 high sigmaR	0.42 (0.21;0.39;0.72)	0.70 (0.63;0.69;0.79)	0.24 (0.18;0.24;0.32)	0.39 (0.27;0.38;0.54)	0.27 (0.20;0.26;0.36)	1.00	
dwCSLF0.1 uwCPUE2 high sigmaRM	0.61 (0.31;0.63;0.78)	0.71 (0.64;0.71;0.82)	0.16 (0.14;0.16;0.19)	0.36 (0.25;0.36;0.49)	0.24 (0.18;0.24;0.32)	1.00	
dwCSLF0.1 uwCPUE2 2CPUE	0.70 (0.57;0.71;0.83)	0.74 (0.66;0.74;0.85)	0.15 (0.12;0.14;0.19)	0.48 (0.29;0.47;0.74)	0.33 (0.20;0.32;0.52)	1.00	
CPUEpow0.75	0.67 (0.55;0.67;0.81)	0.73 (0.65;0.72;0.83)	0.15 (0.12;0.15;0.20)	0.40 (0.24;0.38;0.65)	0.23 (0.15;0.23;0.35)	1.00	
dwCSLF0.1 uwCPUE2 RWQ	0.26 (0.14;0.25;0.41)	0.71 (0.63;0.70;0.81)	0.11 (0.08;0.11;0.16)	0.49 (0.27;0.48;0.75)	0.35 (0.17;0.35;0.56)	1.00	

4. DISCUSSION

This assessment of the pāua stock in management area PAU 5D presents some significant changes from previous models. Of these changes, changing models for CPUE and growth data had the greatest impact.

The new CPUE standardisation model integrated both CELR and PCELR data types in a single index that made use of all available variables for standardisation. This aspect led to a correction of CPUE from the raw geometric-mean index, determined by changes in catch distribution among divers of different abilities and increased diving in good conditions. Both changes may be a consequence of recent shelving of 35% of the total allowable catch, which may make it less attractive for less efficient operators to participate in the fishery, and which may allow more able operators to fish during better weather conditions. Although the standardisation corrected the raw CPUE downwards, spatial resource use can introduce downward bias in CPUE, especially after catch reductions (Neubauer 2017).

Changing the growth representation in the model led to the greatest change in the assessment outcome: the model estimated growth to be substantially lower than suggested by available data for PAU 5D. This leads to slower growth from recruit size to the first fishery-selected size-classes. As a consequence, the model estimated higher spawning stock biomass at sizes below the MLS than previous models for the same area. Although this change led to more optimistic estimates of stock status, the relative available biomass was largely unaffected; this aspect meant that, although the model estimated a larger spawning stock, the part of the stock that was fished had not changed markedly from previous model estimates, and remained near 25% of un-fished available biomass.

Other changes, such as changing the CSLF model to a DM-MN model for standardisation and estimation of process error, only led to minor changes in assessment outcomes. Data weighting in recent assessments followed Francis (2011) and Francis (2017), suggesting that relative abundance indices in the form of CPUE and survey indices should, in general, weigh more strongly in integrated stock assessment models than catch composition data. As a consequence, LF data received relatively little weight in previous assessments, as it was downweighted by the formulation based on Francis (2011). This formulation (see equation T1.8 in Francis 2011) compared mean length predicted in the assessment model to observed mean length, downweighting composition data according to the discrepancy between observed and predicted values in an iterative procedure. This procedure is based on the assumption that any mis-fit is due to process error, meaning the model formulation does not adequately represent the processes leading to composition data.

Although there are sound general arguments in favour of downweighting compositional data, this approach is not universally adopted. Data weighting in integrated stock assessments is currently recognised as a key challenge worldwide, and ongoing discussions by assessment scientists elsewhere highlight a number of strategies to address it (Wang & Maunder 2017, Minto-Vera et al. 2017, Maunder & Piner 2017). Arguments for downweighting composition data usually invoke the complexity of processes that lead to composition data: recruitment, growth, natural mortality and fishery selectivity all interact to determine catch composition (Maunder & Piner 2017). Together, these processes have the potential to induce large discrepancies (process error) between model compositions and actual composition data. Nevertheless, similar arguments could be made for CPUE indices, which often determine the abundance trends in pāua assessments given severe downweighting of compositional data. The CPUE indices have been criticised as measures of relative abundance for a number of reasons (Hilborn & Walters 1987, Harley et al. 2001), especially in abalone fisheries (e.g., Prince & Hilborn 1998); the indices only relate to actual abundance through catchability, which in itself carries the risk of considerable process error. In addition, downweighting composition data discards potentially valuable information about recruitment, growth and natural mortality (Wang & Maunder 2017, Minto-Vera et al. 2017) and can, therefore, lead to biased assessment outcomes. Lastly, default downweighting of composition data makes it difficult to appraise the amount of information that is retained (or lost) from each dataset, and carries the risk of under-representing model uncertainty associated with alternative weight scenarios.

We suggest that the combination of KLD weighting of likelihoods and description of posterior information loss through the KLD provide a consistent way to explore data weighting alternatives beyond the

default down-weighting of CSLF data. In this context, the models agreed on by the Shellfish Working Group led to nearly equal information loss from both data types. As nearly all parameters in the model were currently estimated, the main uncertainties were in the data weighting and prior weight for production-related parameters such as M and σ_R . We showed that constraining σ_R consistently led to unrealistic growth and M patterns that suggest an overly constrained σ_R that destabilises the model. The main sensitivity runs were, therefore, taken to be variations on data-weighting and the strength of the M prior. All scenarios give relatively optimistic estimates of stock status with little to no probability that the stock is currently overfished. Although this outcome may be viewed as conflicting with recent shelving of quota by the commercial fishery stakeholders, the available biomass continues to be low in relative terms, meaning that from a commercial point of view, the status of the fishery may not be as good as the stock status with respect to maturation (i.e., SSB) would suggest.

5. ACKNOWLEDGMENTS

We thank Julie Hills, Tom McCowan, Storm Stanley, Jeremy Cooper and the members of the Shellfish Working Group for stimulating discussion and ideas that led to various developments in this assessment.

This research was funded by Ministry for Primary Industries project PAU2018-01.

6. REFERENCES

- Allaire, J.; Xie, Y.; McPherson, J.; Luraschi, J.; Ushey, K.; Atkins, A.; Wickham, H.; Cheng, J.; Chang, W. (2018). Rmarkdown: Dynamic documents for r. R package version 1.10. Retrieved from <https://CRAN.R-project.org/package=rmarkdown>.
- Breen, P.A.; Kim, S.W.; Andrew, N.L. (2003). A length-based Bayesian stock assessment model for the New Zealand abalone *Haliotis iris*. *Marine and Freshwater Research* 54 (5): 619–634.
- Butterworth, D.S.; Haddon, M.; Haist, V.; Helidoniotis, F. (2015). Report on the New Zealand paua stock assessment model; 2015. *New Zealand Fisheries Science Review* 2015/4: 31 p. Retrieved from https://fs.fish.govt.nz/Doc/23947/FSR_2015_04_2015_Paua_review.pdf.ashx.
- Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* 68 (6): 1124–1138. doi:10.1139/f2011-025.
- Francis, R.I.C.C. (2017). Revisiting data weighting in fisheries stock assessment models. *Fisheries Research* 192: 5–15. doi:10.1016/j.fishres.2016.06.006.
- Fu, D. (2013). The 2012 stock assessment of paua (*Haliotis iris*) for PAU 5D. *New Zealand Fisheries Assessment Report* 2013/57. 51 p.
- Fu, D.; McKenzie, A.; Marsh, C. (2017). Summary of input data for the 2016 PAU 5D stock assessment. *New Zealand Fisheries Assessment Report* 2017/32. 79 p.
- Harley, S.J.; Myers, R.A.; Dunn, A. (2001). Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences* 58 (9): 1760–1772.
- Hilborn, R.; Walters, C.J. (1987). A general model for simulation of stock and fleet dynamics in spatially heterogeneous fisheries. *Canadian Journal of Fisheries and Aquatic Sciences* 44 (7): 1366–1369.
- Marsh, C.; Fu, D. (2017). The 2016 stock assessment of paua (*Haliotis iris*) for PAU 5D. *New Zealand Fisheries Assessment Report* 2017/33. 48 p.
- Maunder, M.N.; Piner, K.R. (2017). Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. *Fisheries Research* 192: 16–27. doi:10.1016/j.fishres.2016.04.022.
- Minte-Vera, C.V.; Maunder, M.N.; Aires-da-Silva, A.M.; Satoh, K.; Uosaki, K. (2017). Get the biology right, or use size-composition data at your own risk. *Fisheries Research* 192: 114–125. doi:10.1016/j.fishres.2017.01.014.
- Neubauer, P. (2017). Spatial bias in pāua *Haliotis iris* catch-per-unit-effort. *New Zealand Fisheries Assessment Report* 2017/57. Retrieved from <http://www.mpi.govt.nz/dmsdocument/23272-far-201757-spatial-bias-in-paua-cpue>.

- Neubauer, P.; Tremblay-Boyer, L. (2019). Input data for the 2018 stock assessment of pāua (*Haliotis iris*) for PAU 5D. *New Zealand Fisheries Assessment Report 2019/38*. 40 p.
- Prince, J.; Hilborn, R. (1998). Concentration profiles and invertebrate fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences* 125: 187–196.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.17.3. Retrieved from <http://mc-stan.org/>.
- Wang, S.-P.; Maunder, M.N. (2017). Is down-weighting composition data adequate for dealing with model misspecification, or do we need to fix the model? *Fisheries Research* 192: 41–51.
- Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in r. *In*: V. Stodden; F. Leisch; R.D. Peng (Eds.), *Implementing reproducible computational research*. ISBN 978-1466561595. Chapman. Retrieved from <http://www.crcpress.com/product/isbn/9781466561595>.

APPENDIX A: MODEL COMPARISON

A.1 Growth model: using data versus informed priors

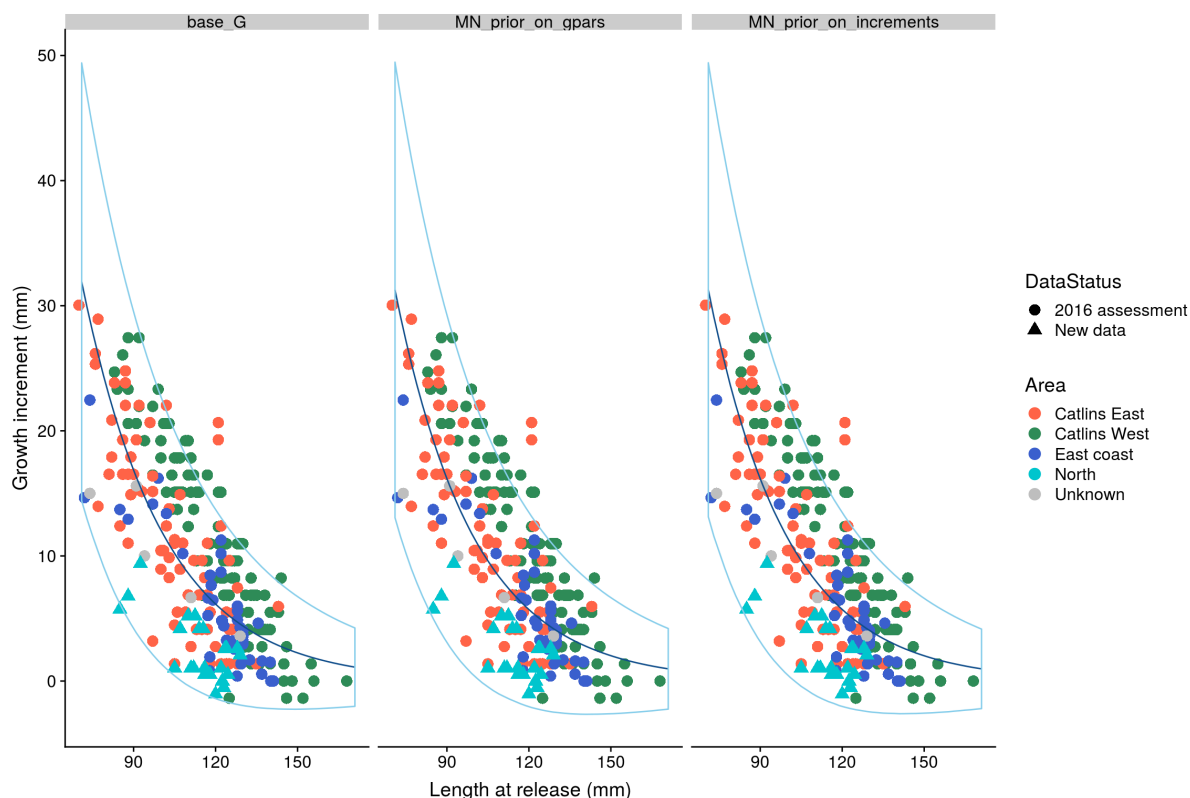


Figure A-1: Comparison of posterior mean growth of pāua (population mean (dark blue line) and standard deviation (light blue ribbon)) from assessment model runs in Stan, using a direct translation of the ADMB model (base-G) with modified growth priors (multivariate normal (MN) on growth parameters (gpars) and MN prior on growth increments). Data were from different South Island areas, including data used in the previous (2016) assessment and new data.

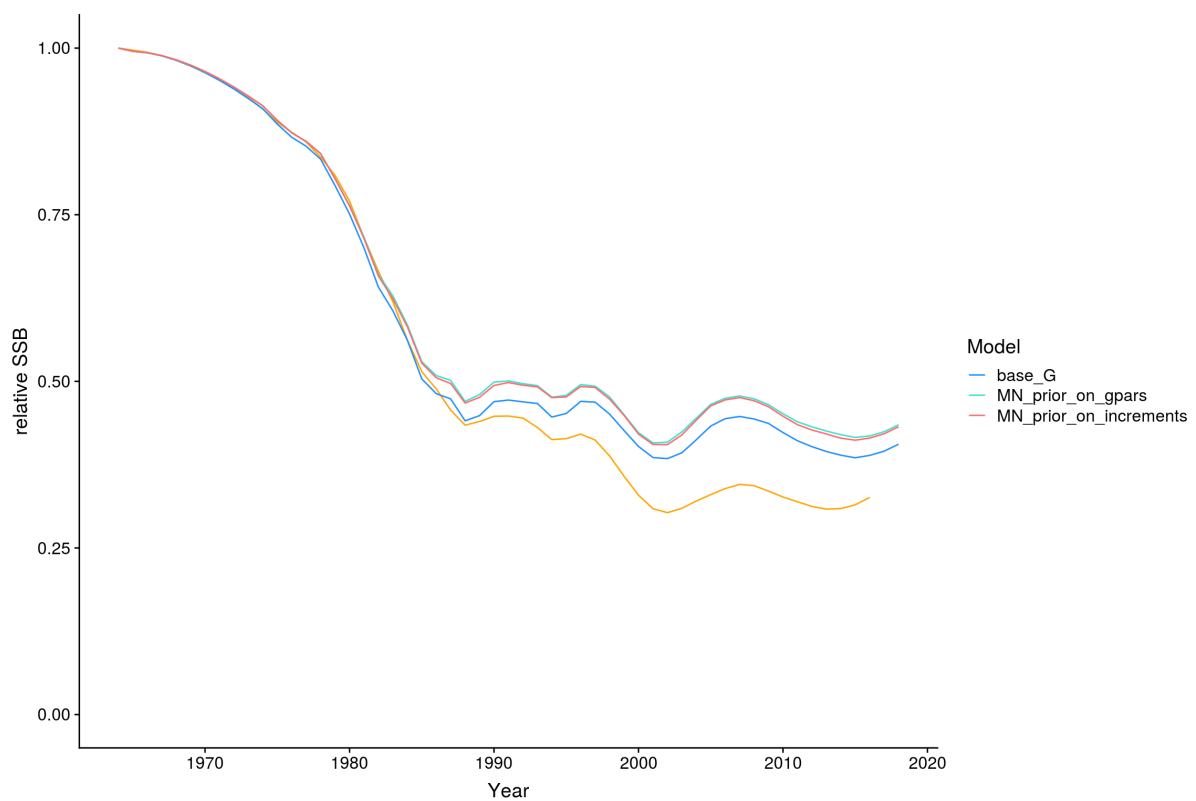


Figure A-2: Comparison of posterior mean pāua stock status (relative spawning stock biomass, SSB) from assessment model runs in Stan, using a direct translation of the ADMB model (base-G) with modified growth priors (multivariate normal (MN) on growth parameters (gpars) and MN prior on growth increments). The mode of joint posterior distribution (MPD) for the base case from the previous assessment is shown in orange for reference (note that the Markov Chain Monte Carlo of this model run provided a higher stock status estimate, comparable to the base-G run in this graph).

A.2 Reducing growth: to what effect?

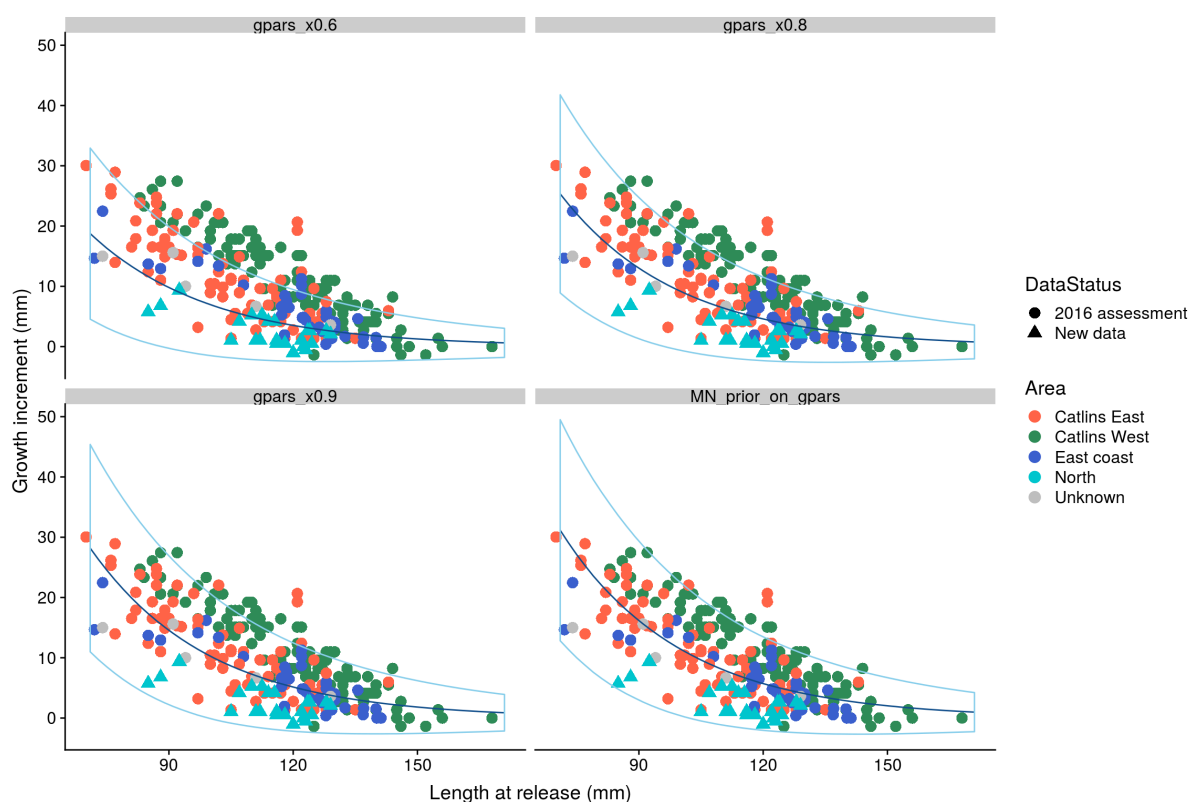


Figure A-3: Comparison of posterior mean growth of pāua (population mean (dark blue line) and standard deviation (light blue ribbon)) from model runs in Stan, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9. Data were from different South Island areas, including data to the previous (2016) assessment and new data.

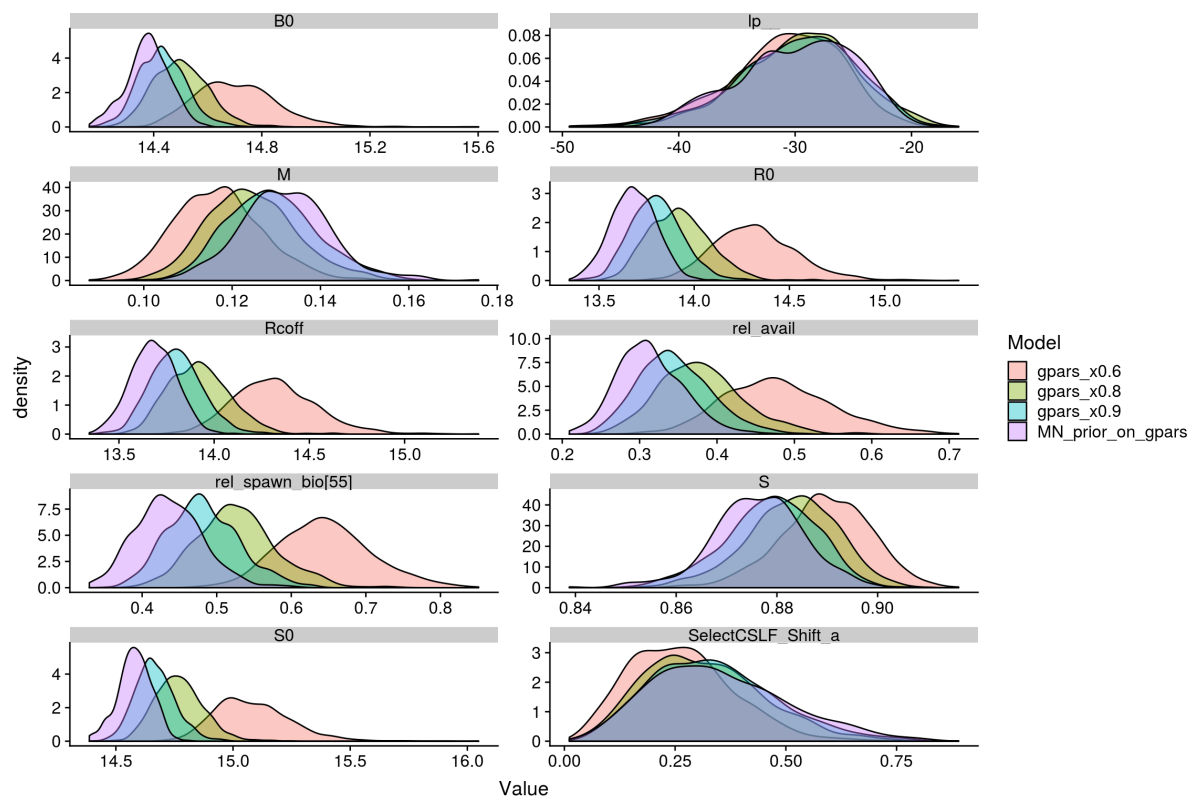


Figure A-4: Comparison of posterior densities for parameters from model runs in Stan, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9.

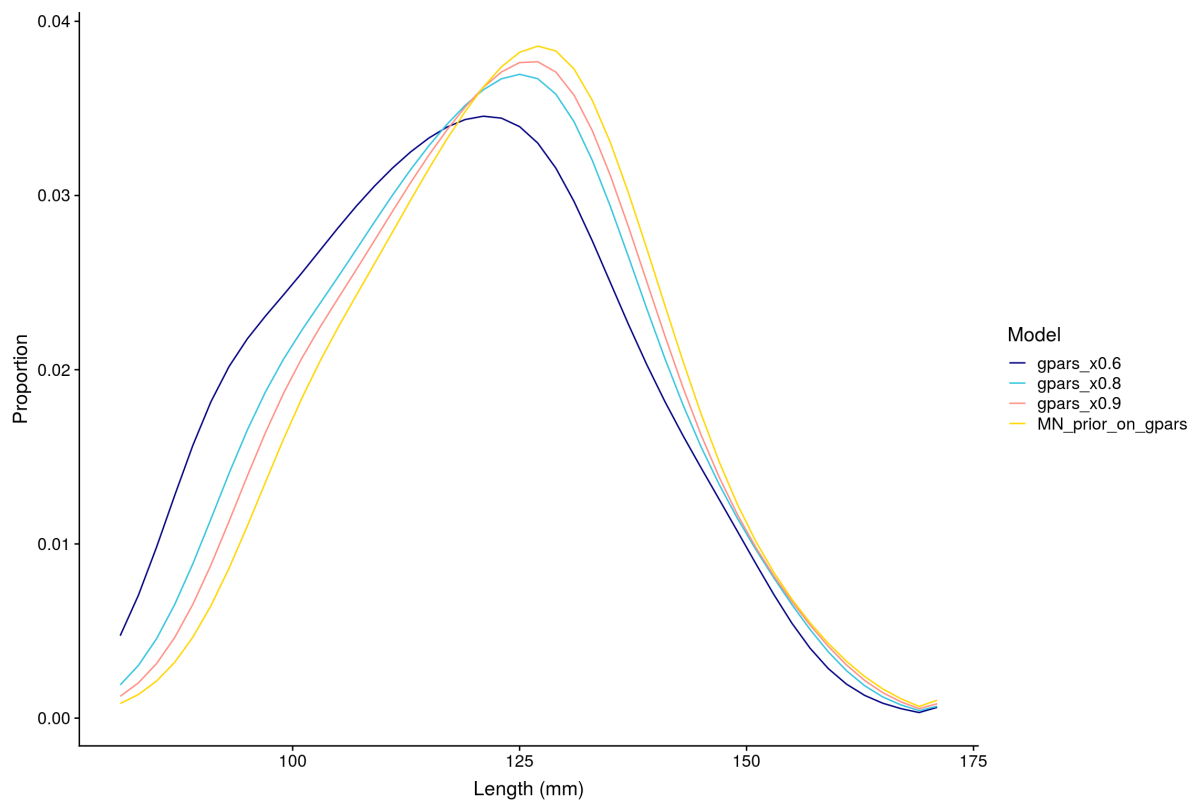


Figure A-5: Comparison of posterior mean proportions-at-length of pāua from model runs in Stan, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9.

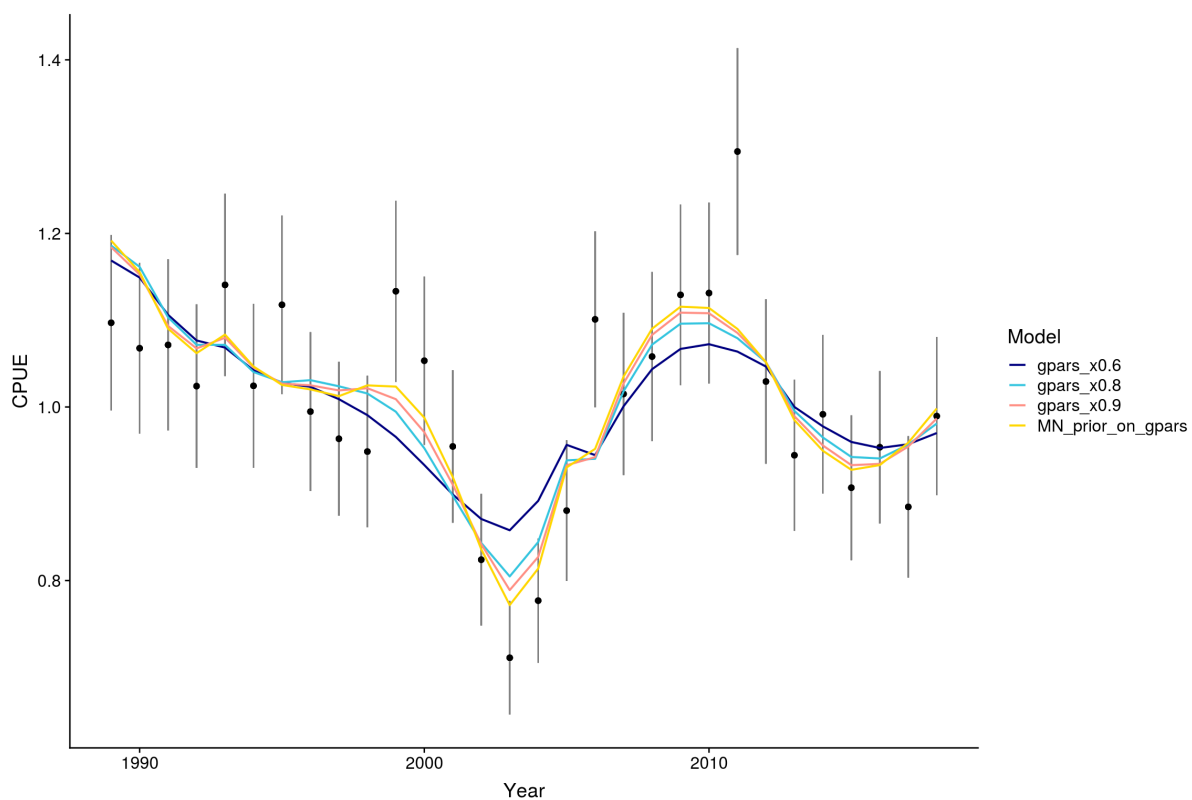


Figure A-6: Comparison of posterior mean predicted catch-per-unit-effort (CPUE) of pāua with estimated CPUE index and assumed process and observation errors (black points and error bars) from model runs in Stan, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9.

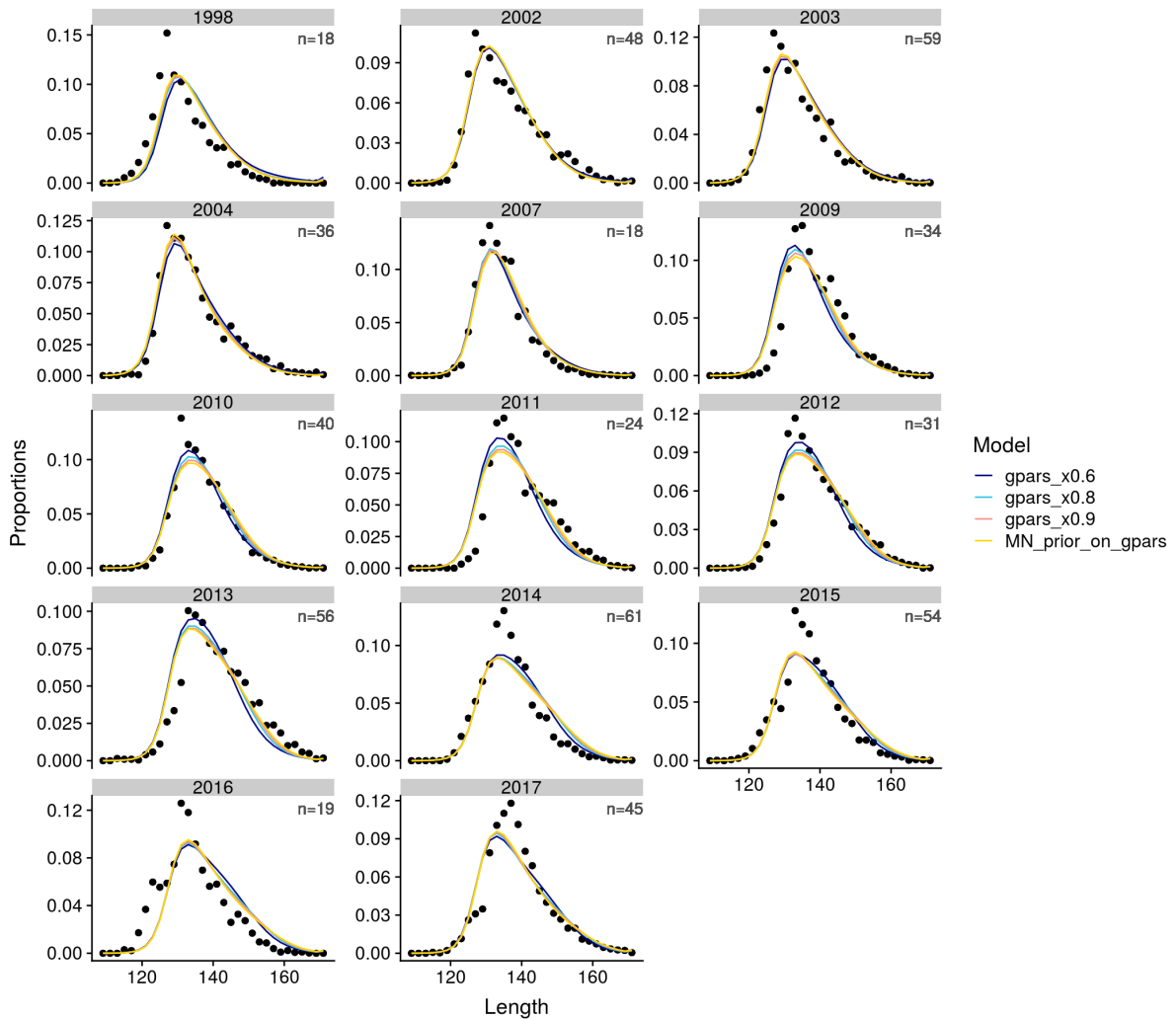


Figure A-7: Comparison of posterior mean predicted catch sampling length frequency (CSLF) of pāua with estimated CSLF proportions from model runs in Stan, using a direct translation of the ADMB model with modified growth priors multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9.

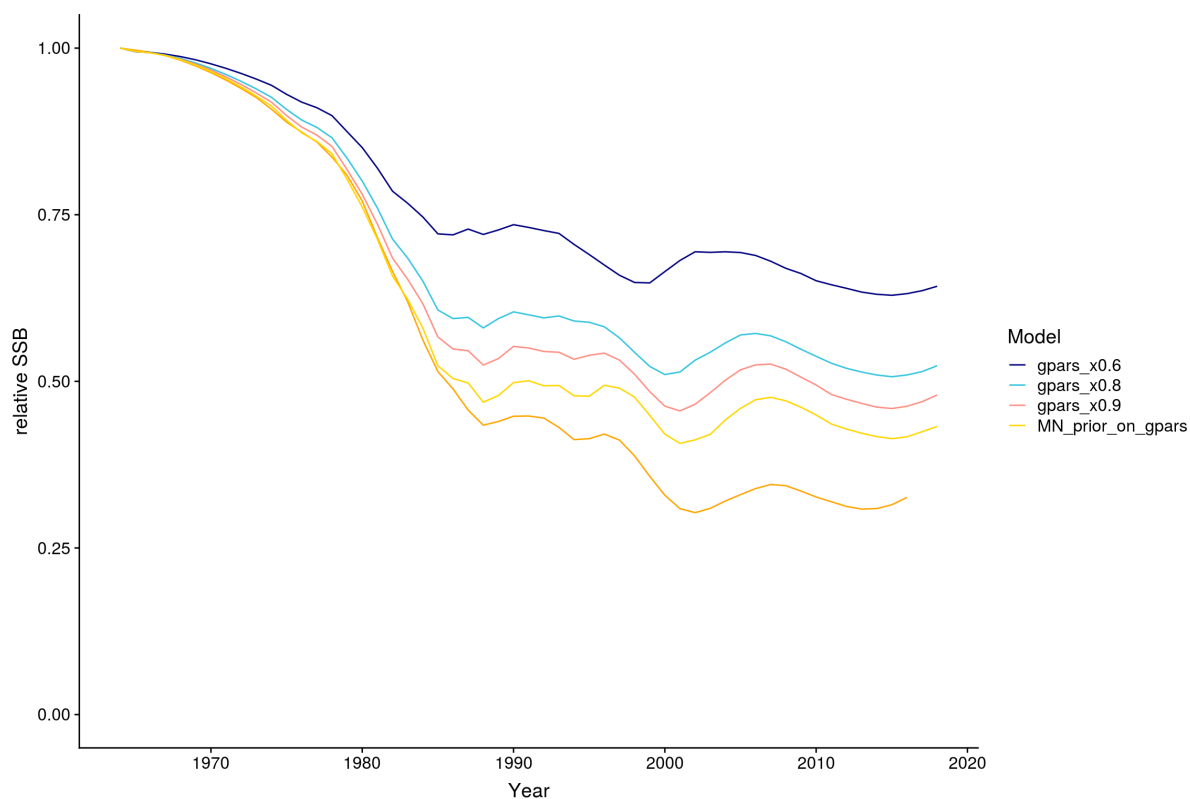


Figure A-8: Comparison of posterior mean stock status of pāua from assessment model runs in Stan, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth parameters (gpars)) and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9. The mode of joint posterior distribution (MPD) for the base case from the previous assessment (Marsh; Fu 2017) is shown in orange for reference (note that the Markov Chain Monte Carlo of this model run provided a higher stock status estimate, comparable to the base-G run in this graph).

A.3 Re-formulating growth

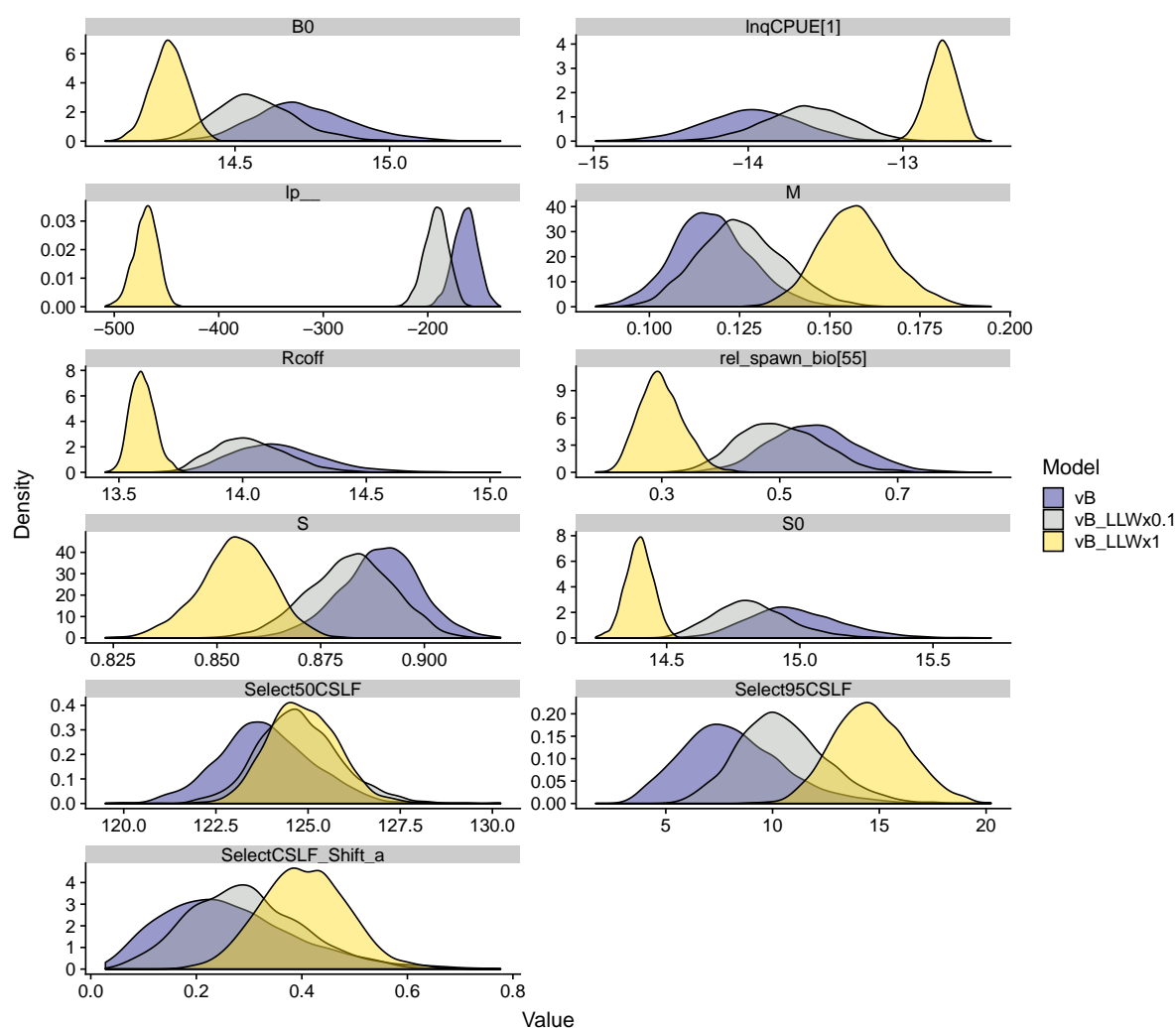


Figure A-9: Comparison of posterior densities for parameters, using a direct translation of the ADMB model with modified growth priors (multivariate normal (MN) on growth increments) estimated using all South Island pāua tag increment data, with varying weights for the catch sampling length frequency (CSLF) dataset. Model runs with modified weights are labelled as: LLW1 = no downweighting of CSLF data; LLW0.1 = CSLF inputs downweighted by a factor of 0.1 on the multinomial sample size. The base case was re-weighting using the method by Francis (2011).

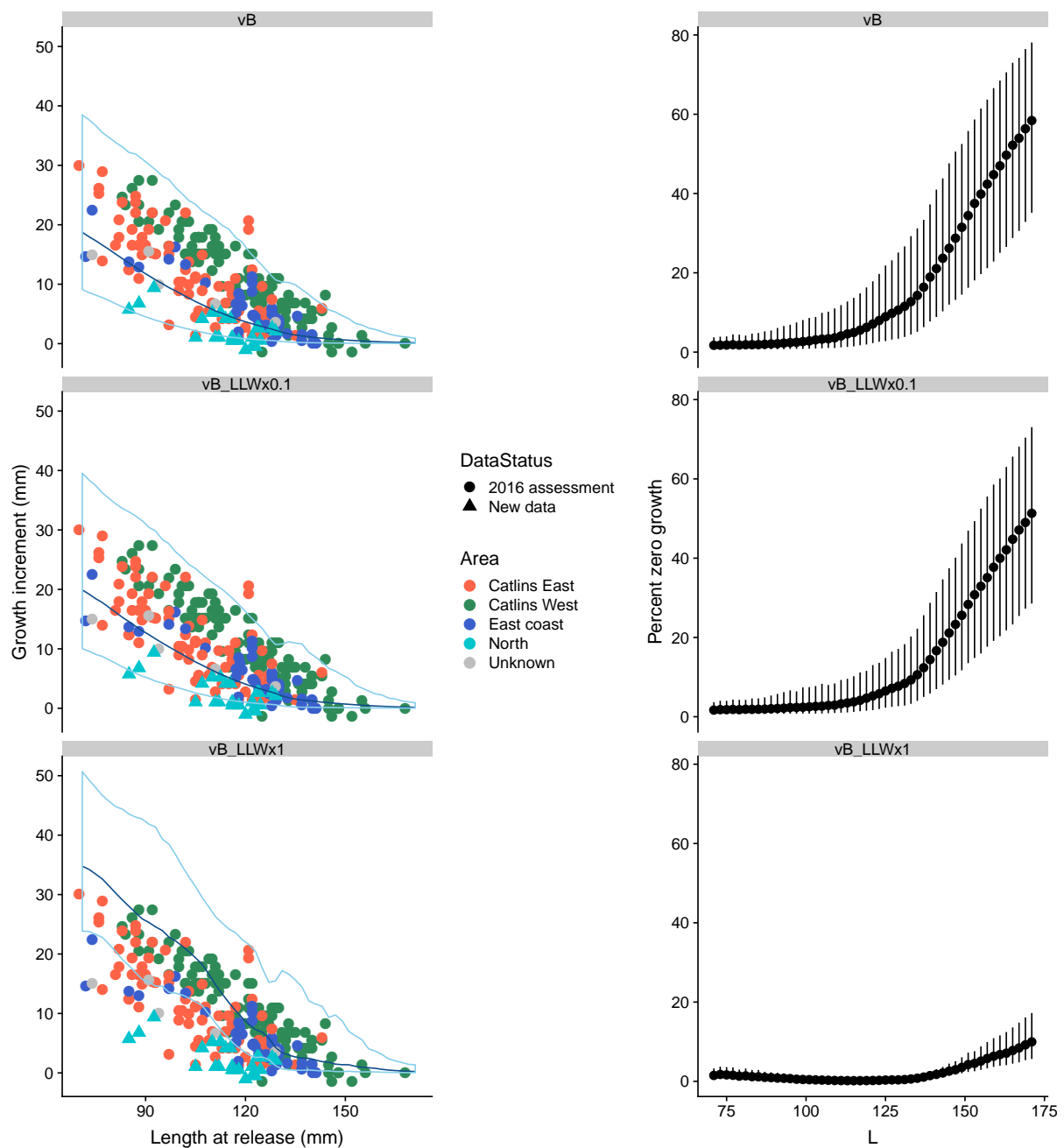


Figure A-10: Comparison of posterior mean growth (population mean (dark blue line) and standard deviation (light blue ribbon)), using a direct translation of the ADB model with modified growth priors (multivariate normal on growth parameters (gpars)), estimated using all South Island pāua tag increment data, with varying weights for the catch sampling length frequency (CSLF) dataset. Model runs with modified weights were: LLW1 = no downweighting of CS LF data; LLW0.1 = CS LF inputs downweighted by a factor of 0.1 on the multinomial sample size. The base case was re-weighting using the method by Francis (2011).

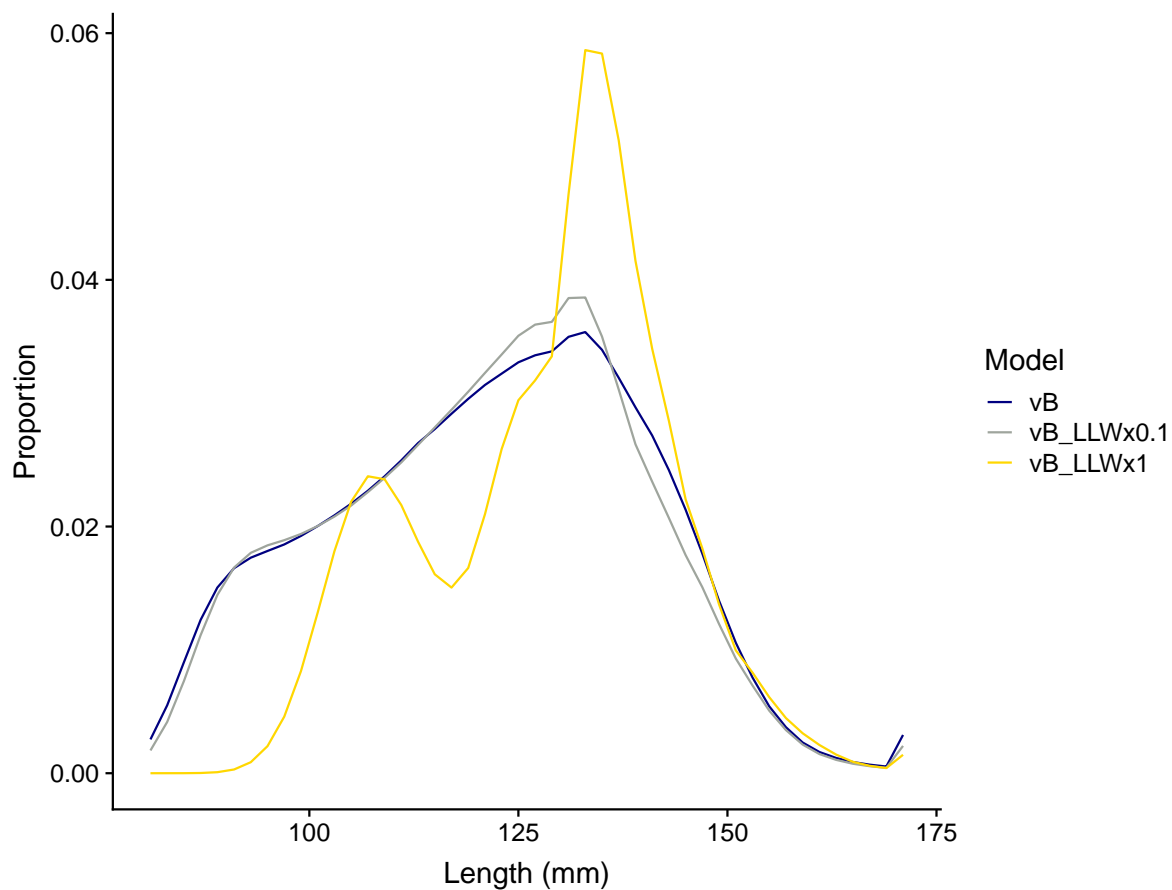


Figure A-11: Comparison of posterior mean proportions-at-length, using a direct translation of the ADBM model with modified growth priors (multivariate normal on growth increments) estimated using all South Island pāua tag increment data, with varying weights for the catch sampling length frequency (CSLF) dataset. Model runs with modified weights are labelled as: LLW1 = no downweighting of CSLF data; LLW0.1 = CSLF inputs downweighted by a factor of 0.1 on the multinomial sample size. The base case was re-weighting using the method by Francis (2011).

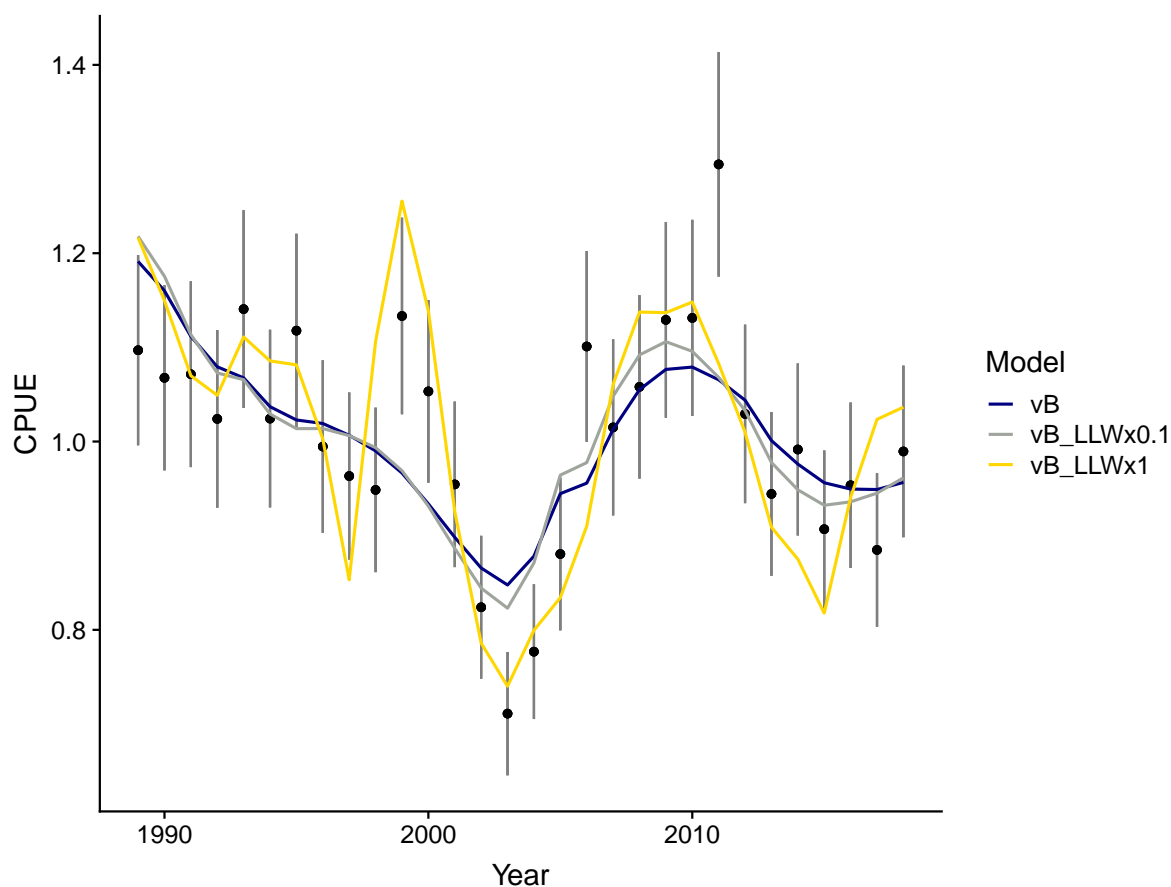


Figure A-12: Comparison of posterior median predicted catch-per-unit-effort (CPUE) with estimated CPUE index and assumed process and observation errors (black points and error bars), using a direct translation of the ADMB model with modified growth priors (multivariate normal on growth increments) estimated using all South Island pāua tag increment data, with varying weights for the catch sampling length frequency (CSLF) dataset. Model runs with modified weights are labelled as: LLW1 = no downweighting of CSLF data; LLW0.1 = CSLF inputs downweighted by a factor of 0.1 on the multinomial sample size. The base case was re-weighting using the method by Francis (2011).

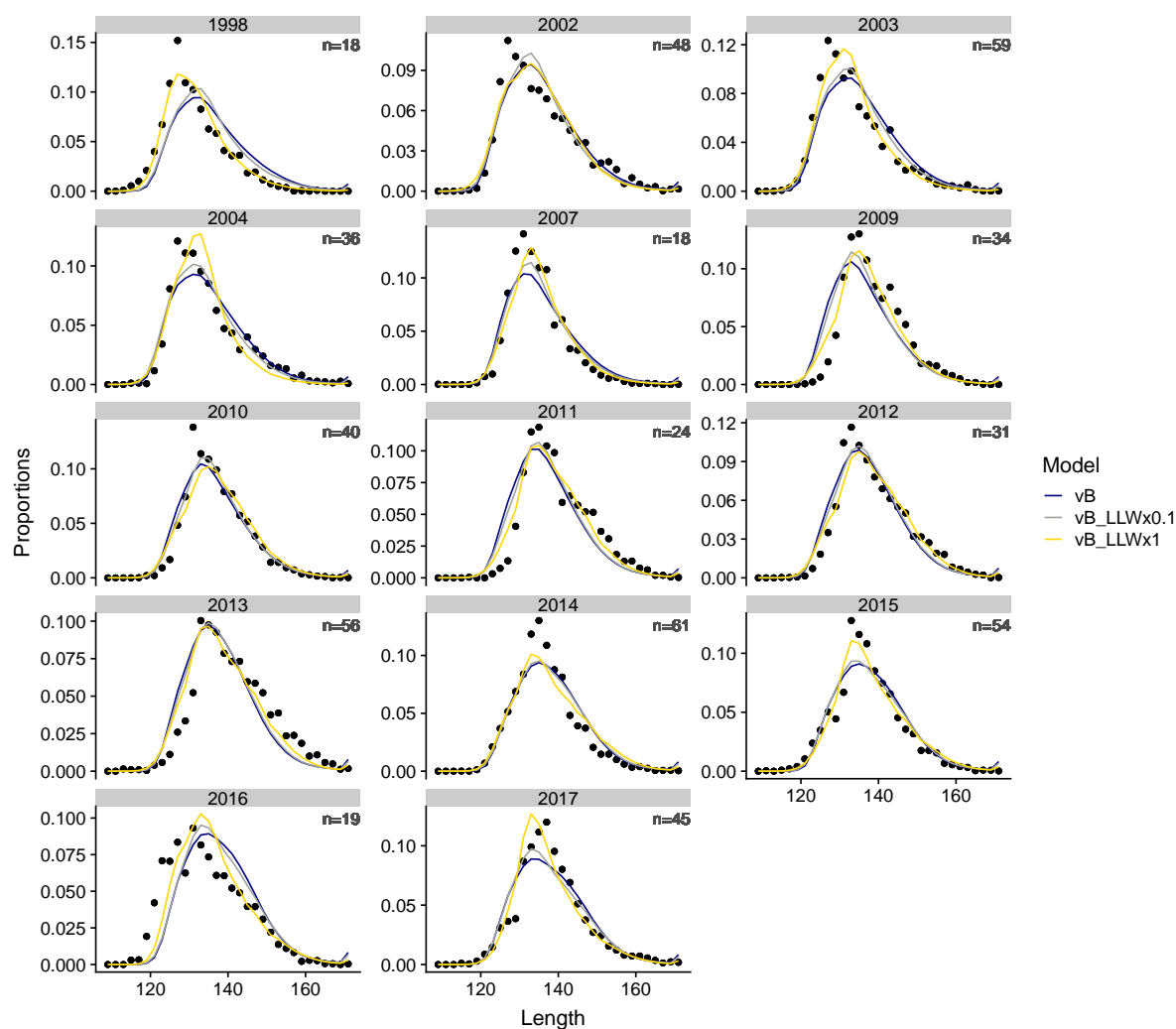


Figure A-13: Comparison of posterior mean predicted catch sampling length frequency (CSLF) with estimated CSLF proportions (using the model described in Section 2.1), with varying weights for CSLF and catch-per-unit-effort datasets. Model runs are labelled with the weights as: downweighted (dw(Dataset)<amount>), upweighted=uw, or base (no weights).



Figure A-14: Comparison of posterior median predicted relative spawning stock biomass trend, using a direct translation of the ADMB model with modified growth priors (multivariate normal on growth increments) estimated using all South Island pāua tag increment data, with varying weights for the catch sampling length frequency (CSLF) dataset. Model runs with modified weights are labelled as: LLW1 = no downweighting of CSLF data; LLW0.1 = CSLF inputs downweighted by a factor of 0.1 on the multinomial sample size. The base case was re-weighted using the method by Francis (2011). The mode of joint posterior distribution (MPD) for the base case from the previous assessment is shown in orange for reference (note that the Markov Chain Monte Carlo of this model run provided a higher stock status estimate, comparable to the base-G run in this graph).

A.4 Towards a new base-case: data weights

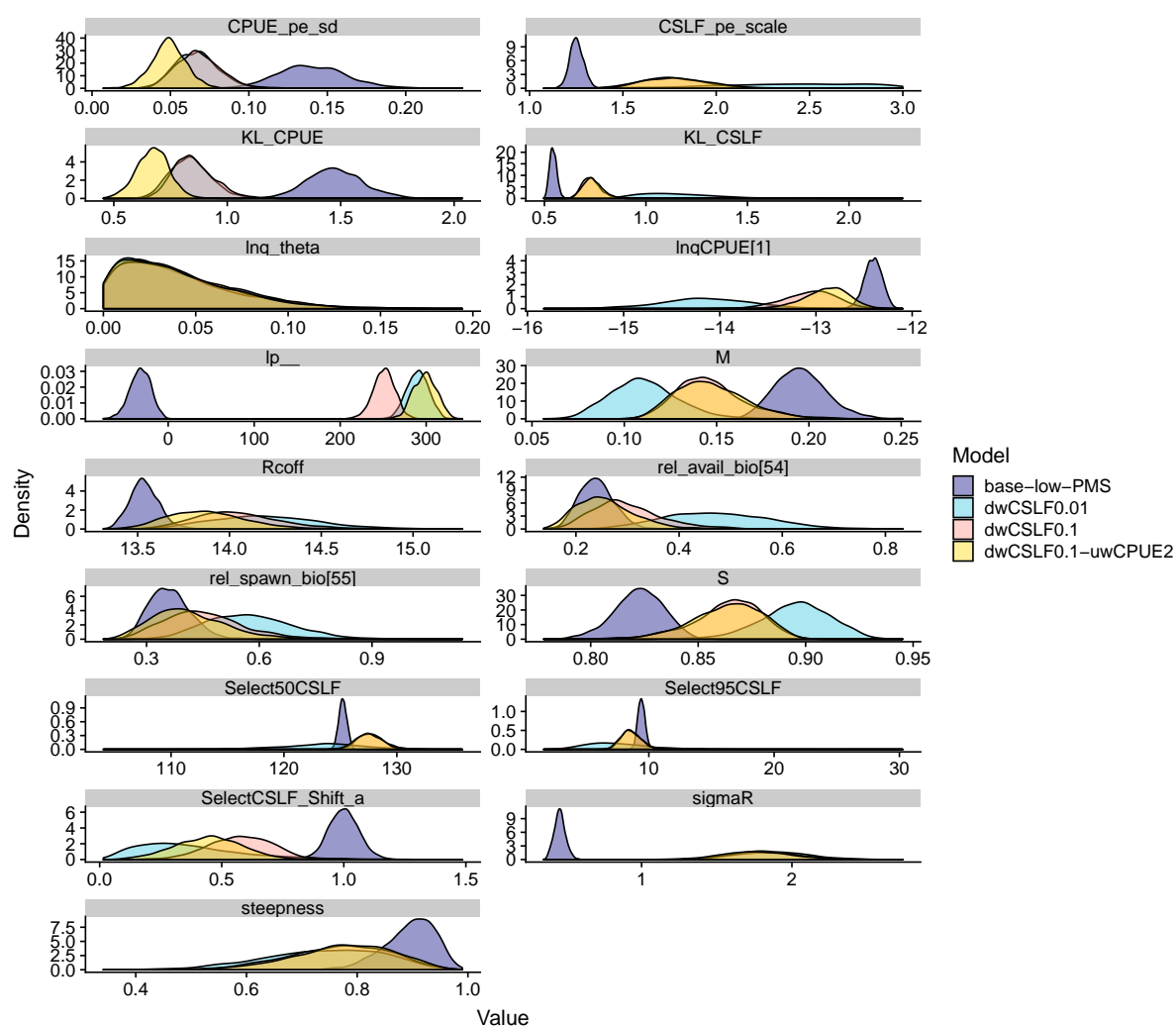


Figure A-15: Comparison of posterior densities for parameters, using the model described in 2.1, with varying weights for catch sampling length frequency (CSLF) CSLF and catch-per-unit-effort (CPUE) datasets. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, or base (no weights).

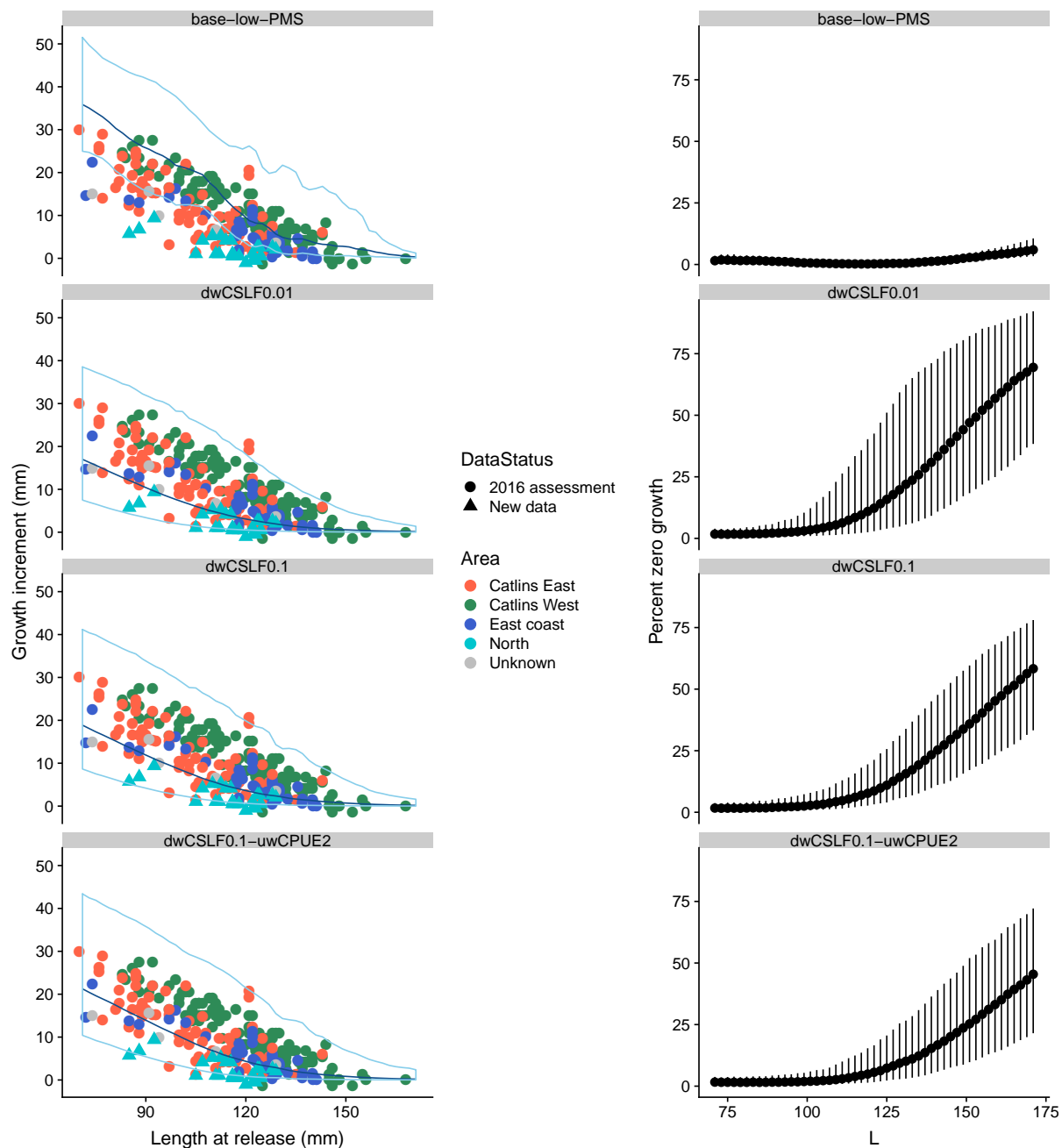


Figure A-16: Comparison of posterior mean growth (population mean (dark blue line) and standard deviation (light blue ribbon)(using the model described in Section 2.1), with varying weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets. Model runs are labelled with the weights as: downweighted (dw(Dataset)<amount>), upweighted=uw, or base (no weights).

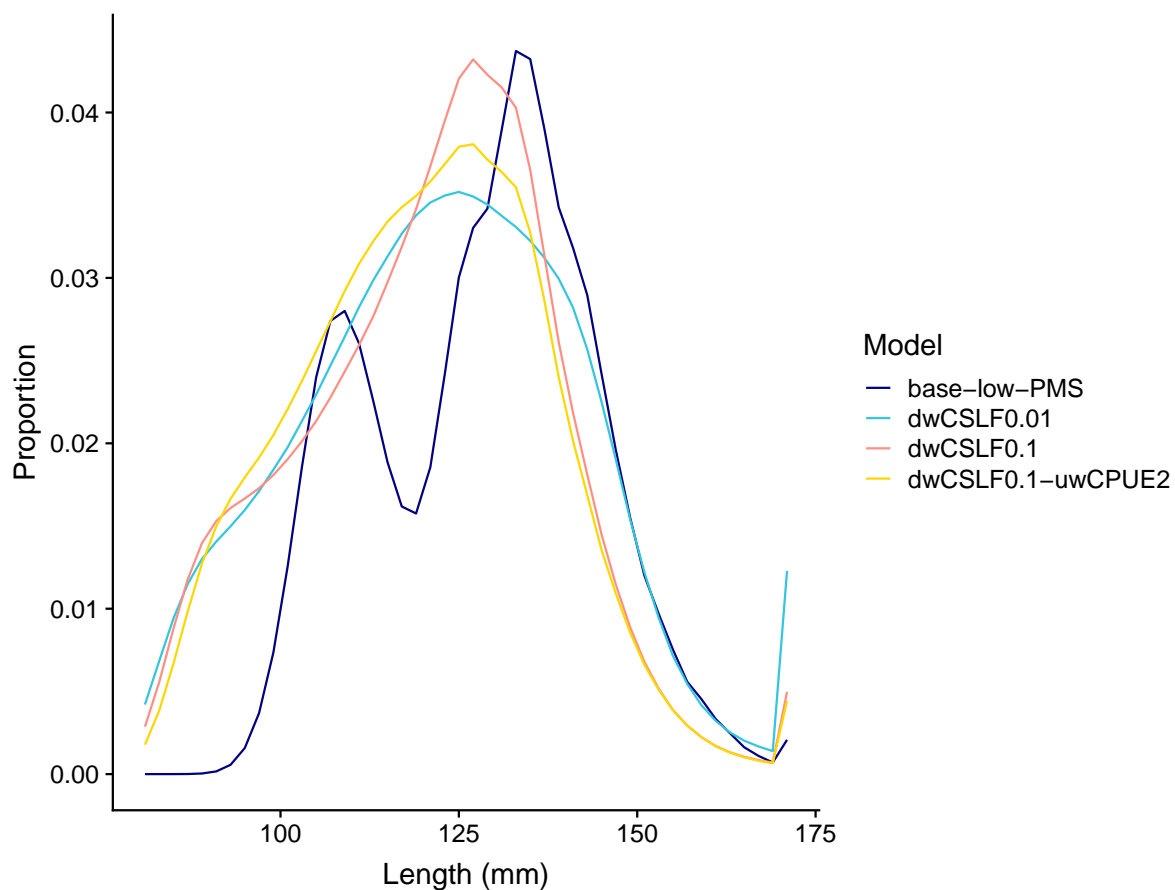


Figure A-17: Comparison of posterior mean proportions-at-length (using the model described in Section 2.1), with varying weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets. Model runs are labelled with the weights as: downweighted (dw<Dataset>(amount)), upweighted=uw, or base (no weights).

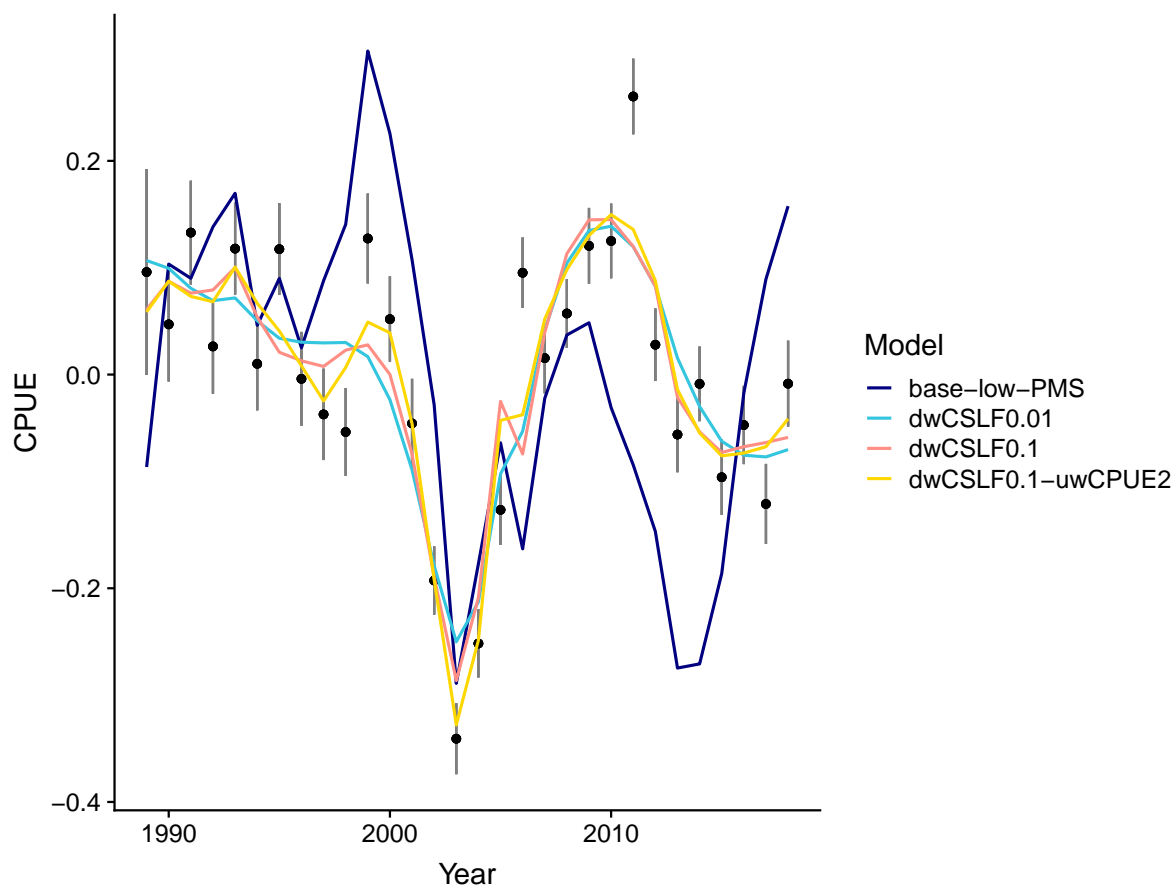


Figure A-18: Comparison of posterior median predicted catch-per-unit-effort (CPUE) with estimated CPUE index and assumed process and observation errors (black points and error bars) (using the model described in Section 2.1), with varying weights for catch sampling length frequency (CSLF) and CPUE datasets. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, or base (no weights).

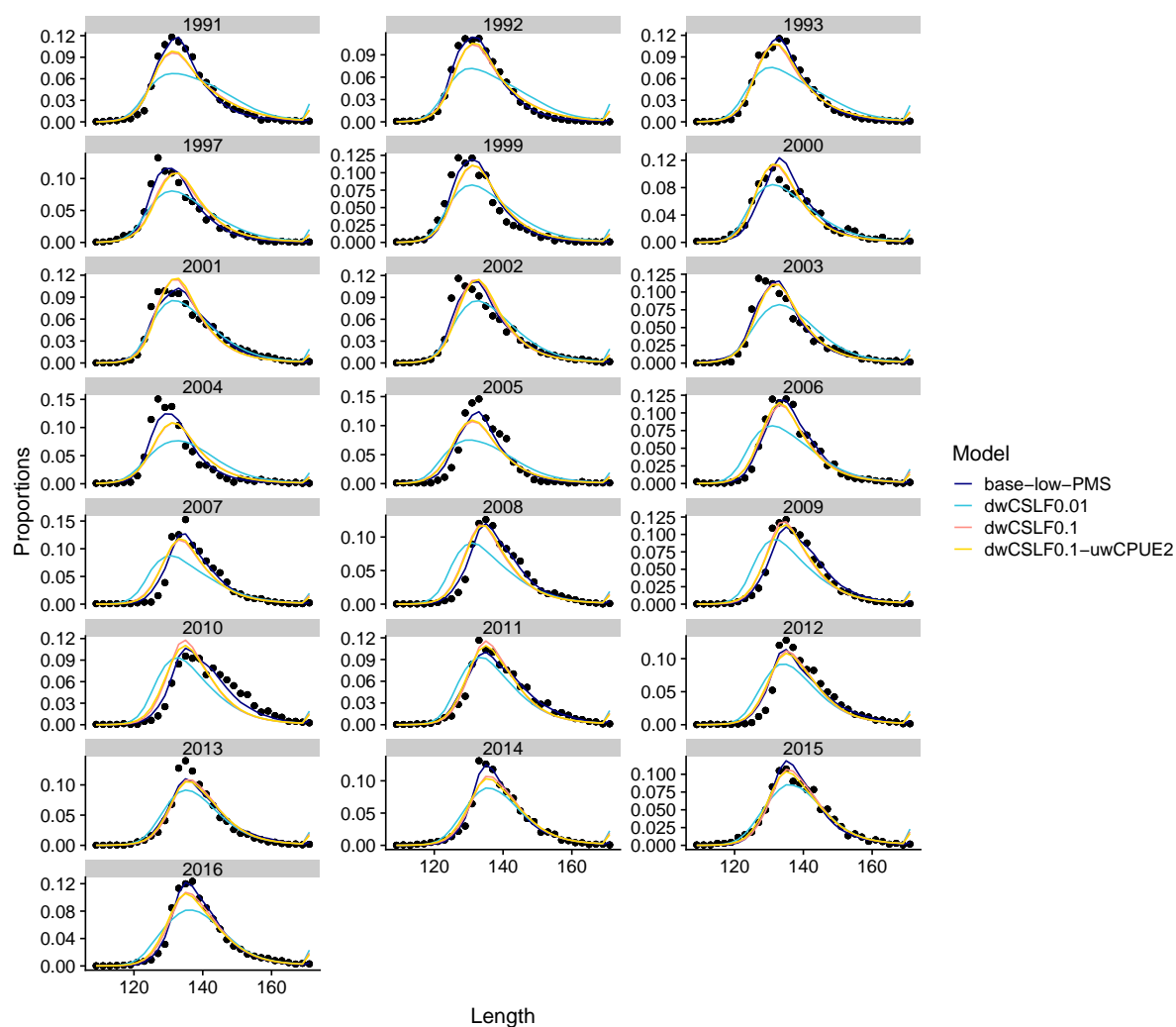


Figure A-19: Comparison of posterior mean predicted catch sampling length frequency (CSLF) with estimated CSLF proportions (using the model described in Section 2.1), with varying weights for CSLF and catch-per-unit-effort (CPUE) datasets. Model runs are labelled with the weights as: downweighted (dw(Dataset)<amount>), upweighted=uw, or base (no weights).

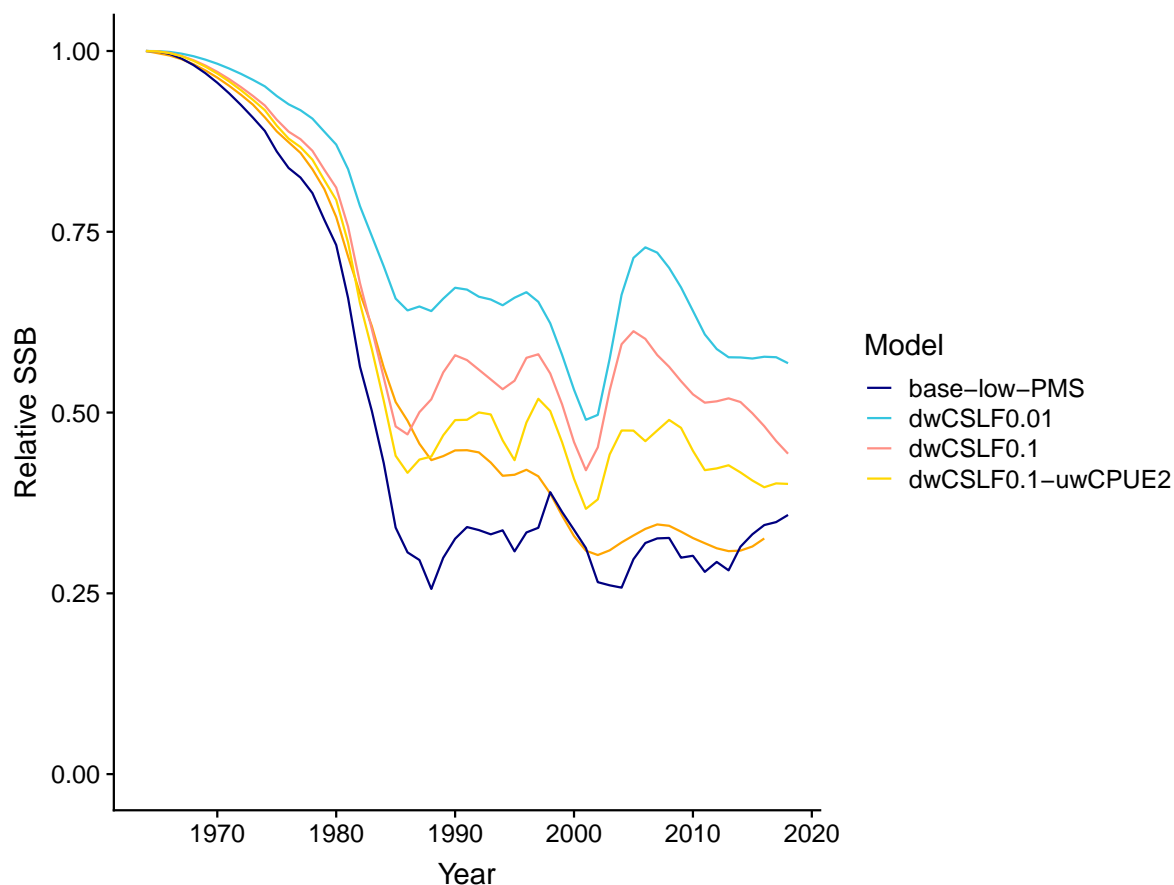


Figure A-20: Comparison of posterior median predicted relative spawning stock biomass trend (using the model described in Section 2.1), with varying weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, or base (no weights).

A.5 Towards a new base-case: additional constraints

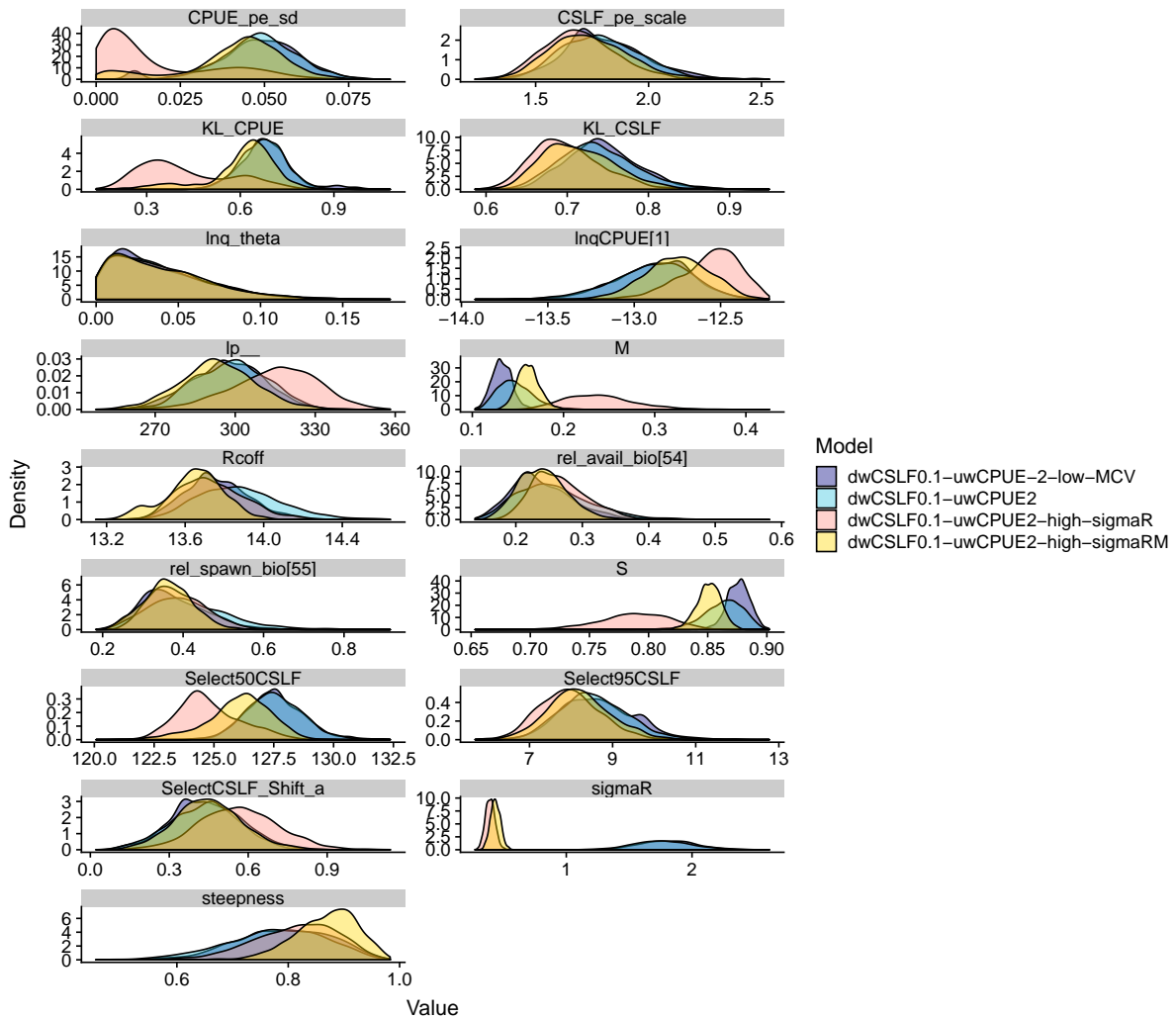


Figure A-21: Comparison of posterior densities for parameters (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) data-sets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with constraints/priors imposed on parameters: low MC = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

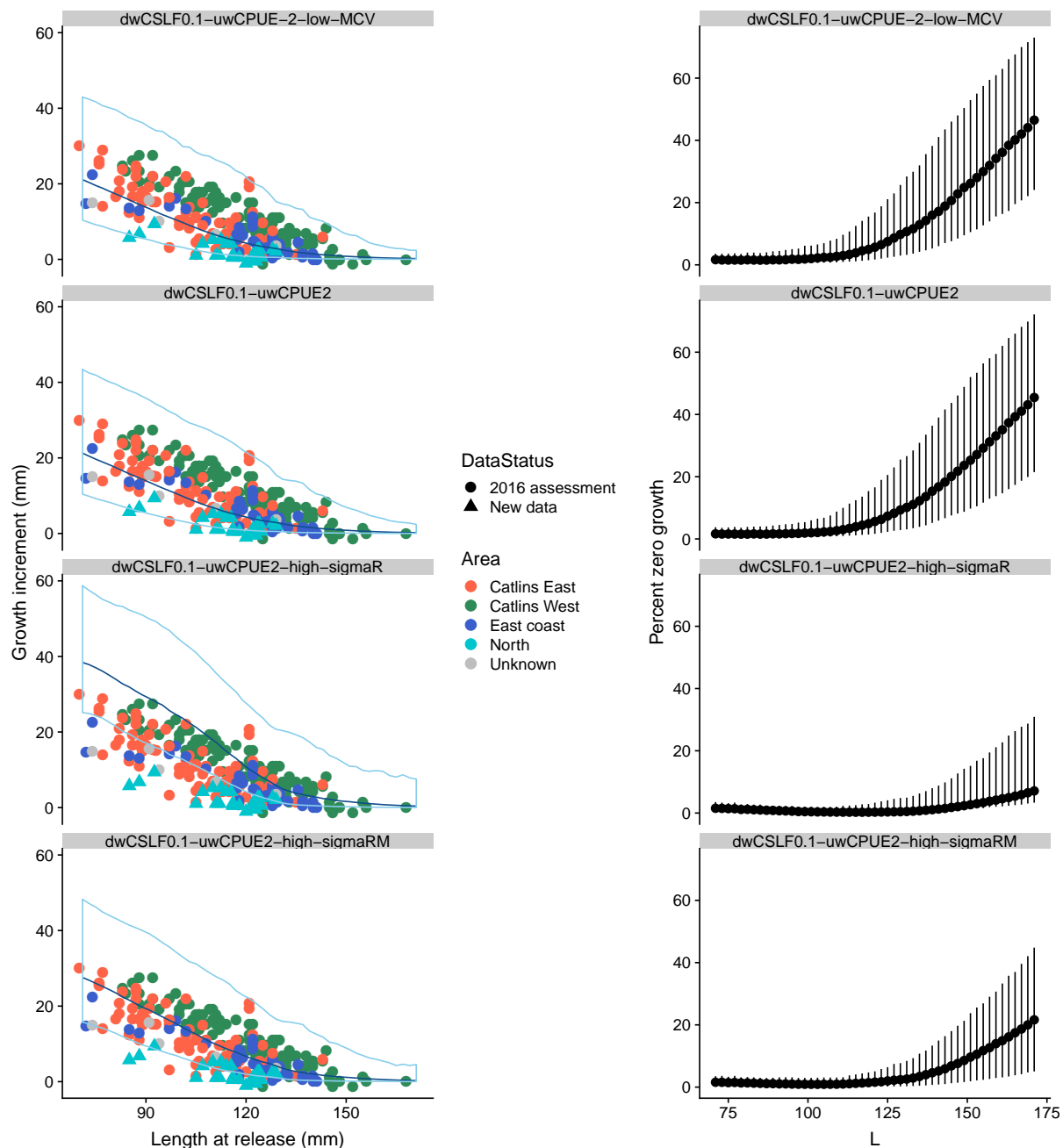


Figure A-22: Comparison of posterior mean growth (population mean (dark blue line) and standard deviation (light blue ribbon)) (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with constraints/priors imposed on parameters: low MCV = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

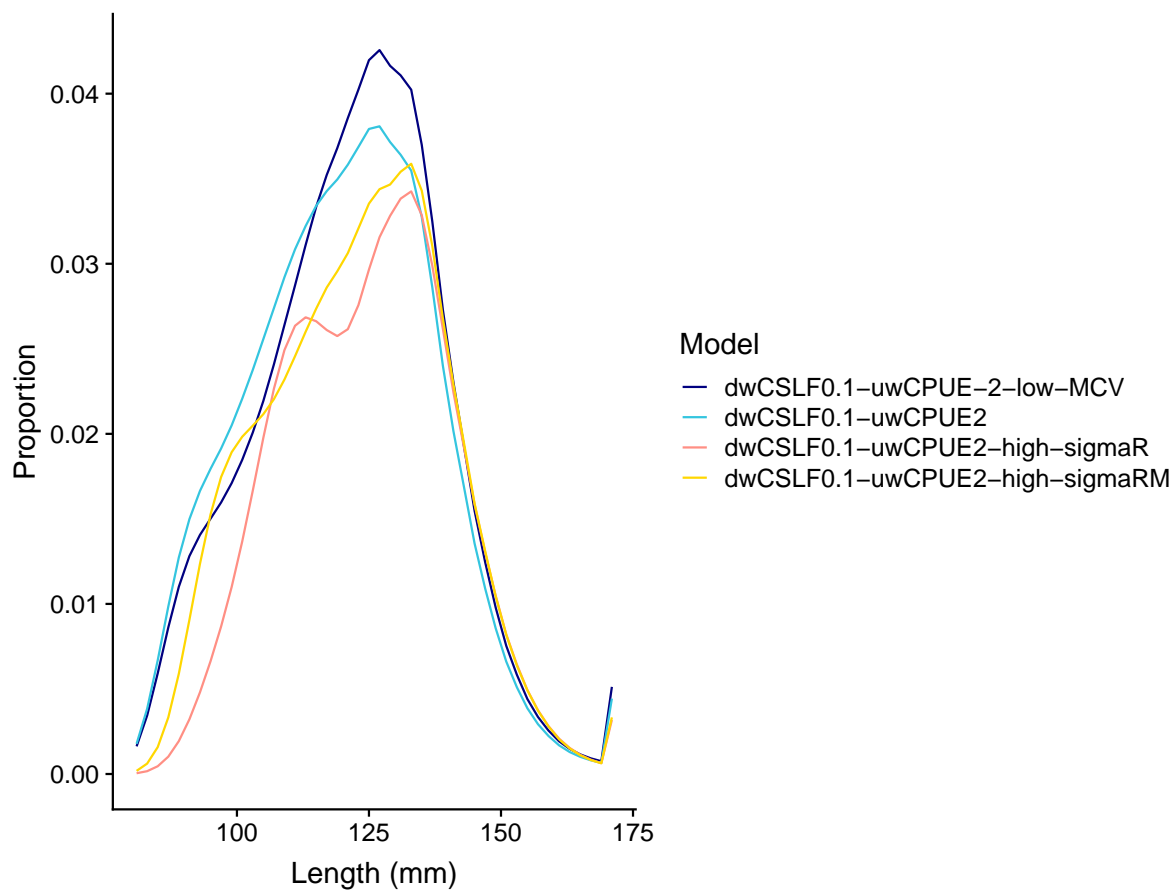


Figure A-23: Comparison of posterior mean proportions-at-length (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw<Dataset>(amount)), upweighted=uw, and with constraints/priors imposed on parameters: low MCV = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

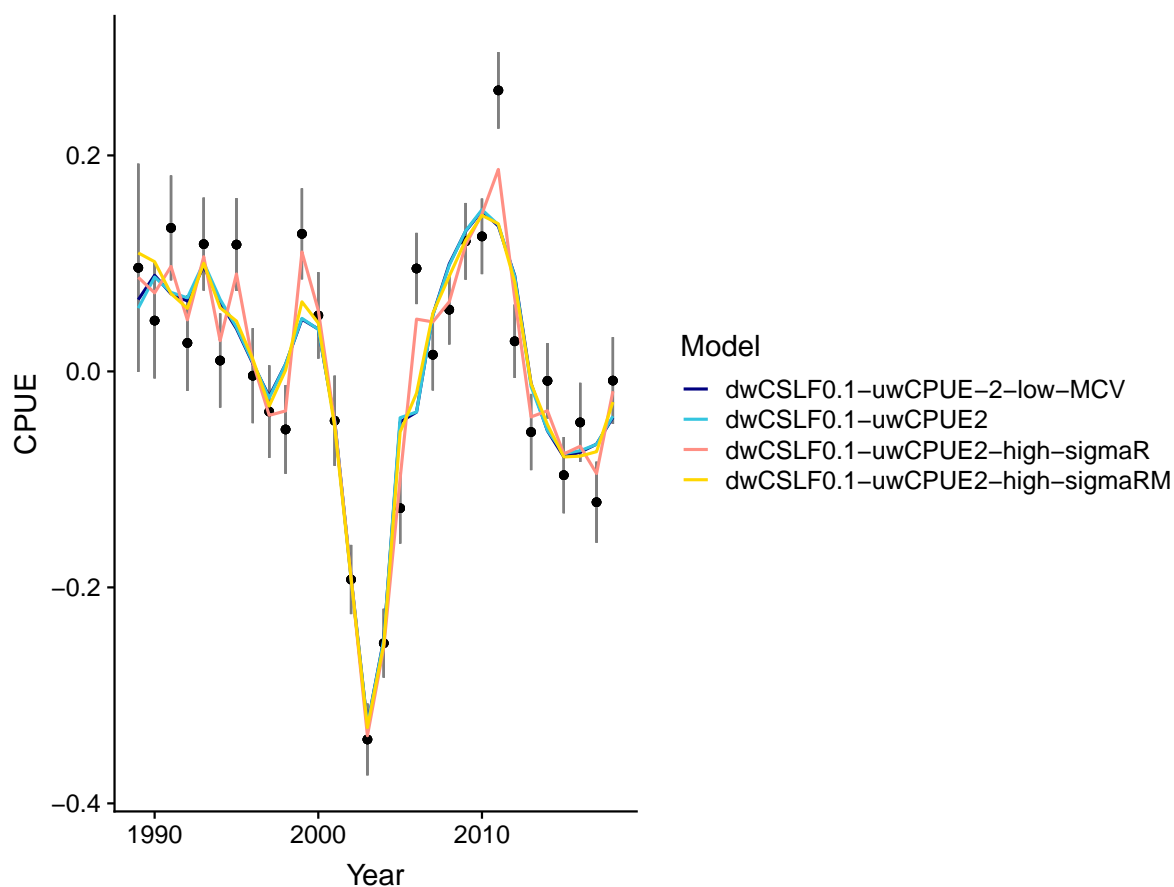


Figure A-24: Comparison of posterior mean predicted catch-per-unit-effort (CPUE) with estimated CPUE index and assumed process and observation errors (black points and error bars) (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and CPUE datasets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with constraints/priors imposed on parameters: low MCV = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

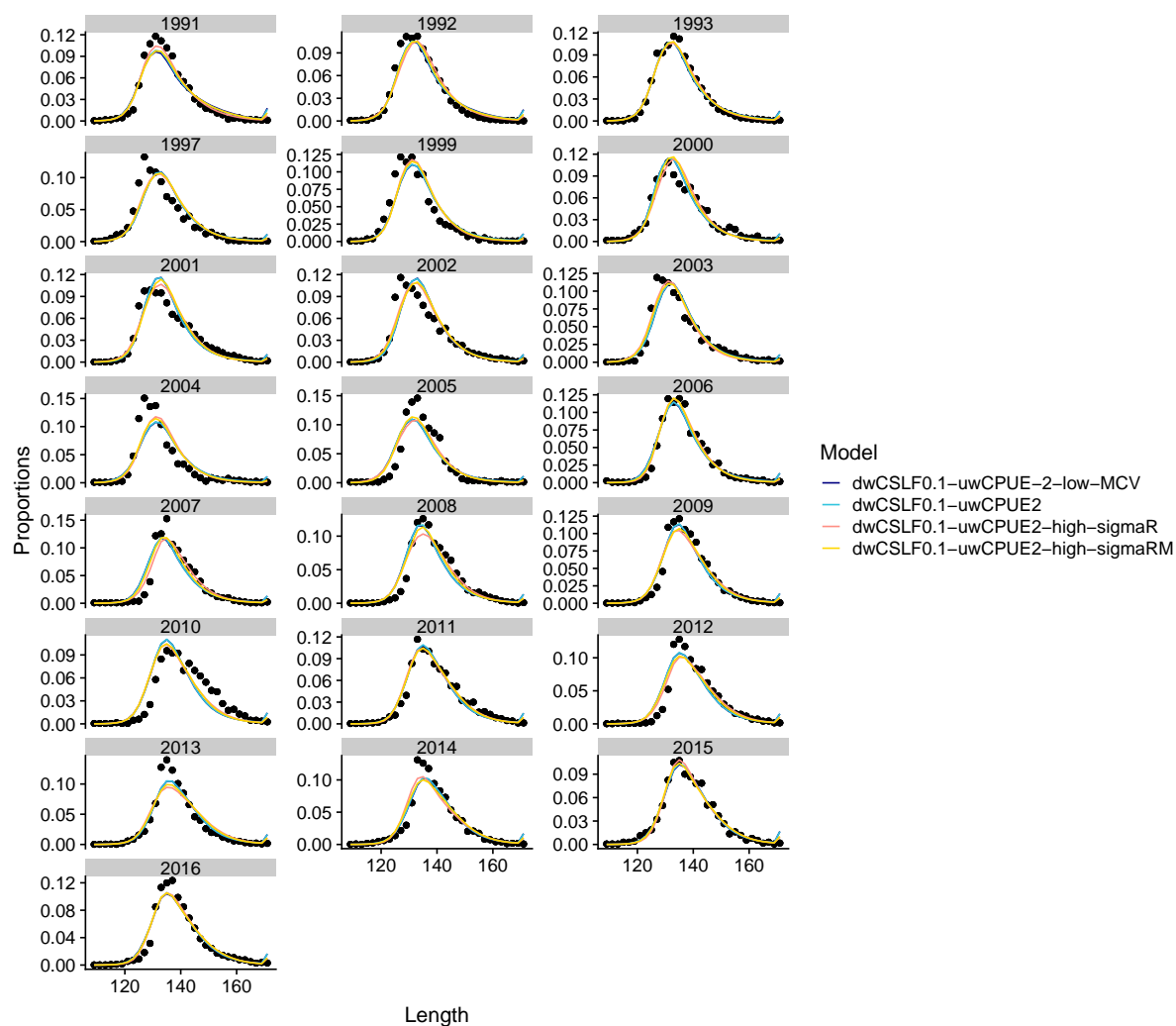


Figure A-25: Comparison of posterior mean predicted catch sampling length frequency (CSLF) with estimated CSLF proportions, using the model described in 2.1, with constant weights for CSLF and catch-per-unit-effort (CPUE) datasets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw/Dataset)⟨amount⟩, upweighted=uw, and with constraints/priors imposed on parameters: low MCV = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

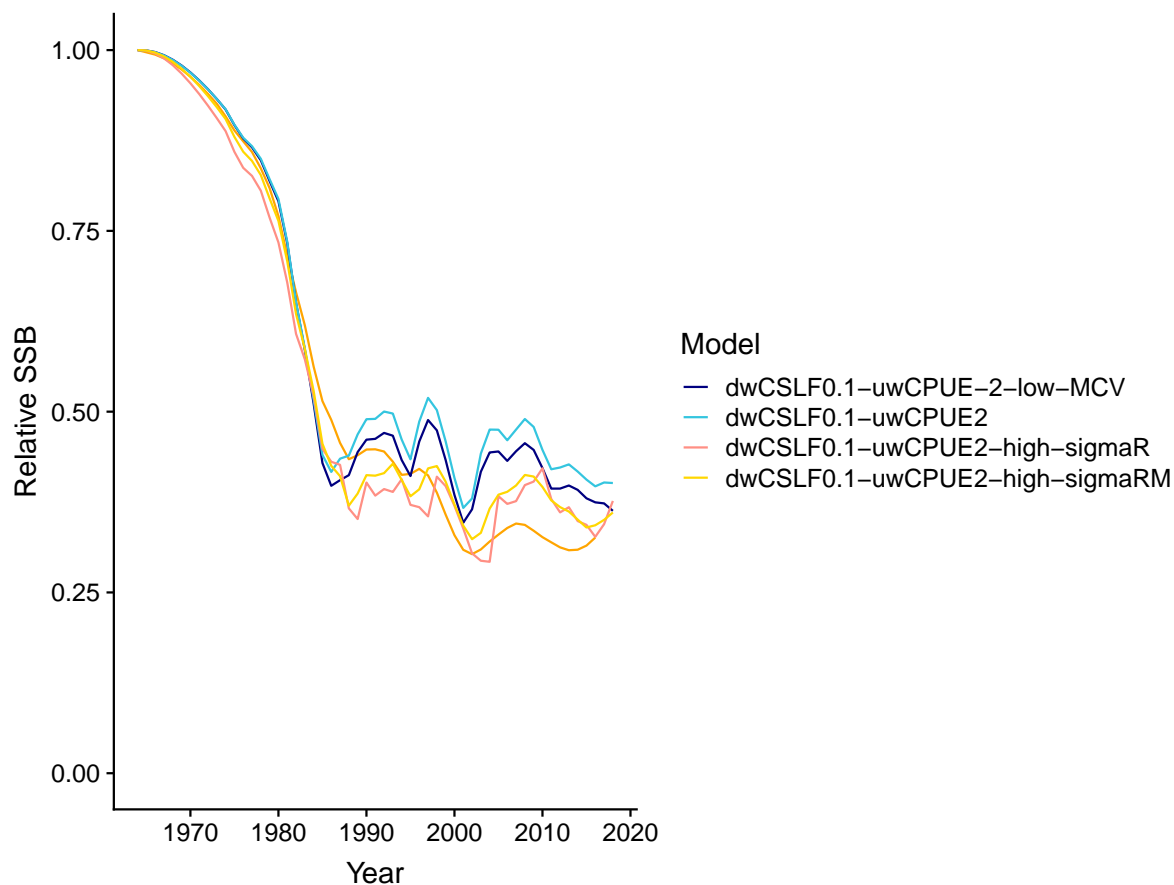


Figure A-26: Comparison of posterior median predicted relative spawning stock biomass trend (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying assumptions for M and σ_R . Model runs are labelled with the weights as: downweighted (dw<Dataset>(amount)), upweighted=uw, and with constraints/priors imposed on parameters: low MCV = log-normal standard deviation reduced to 0.1 (from 0.2), high sigmaR = constraint on σ_R (0.4 instead of 2). For high sigmaRM, both constraints were applied.

A.6 Towards a new base-case: structural assumptions

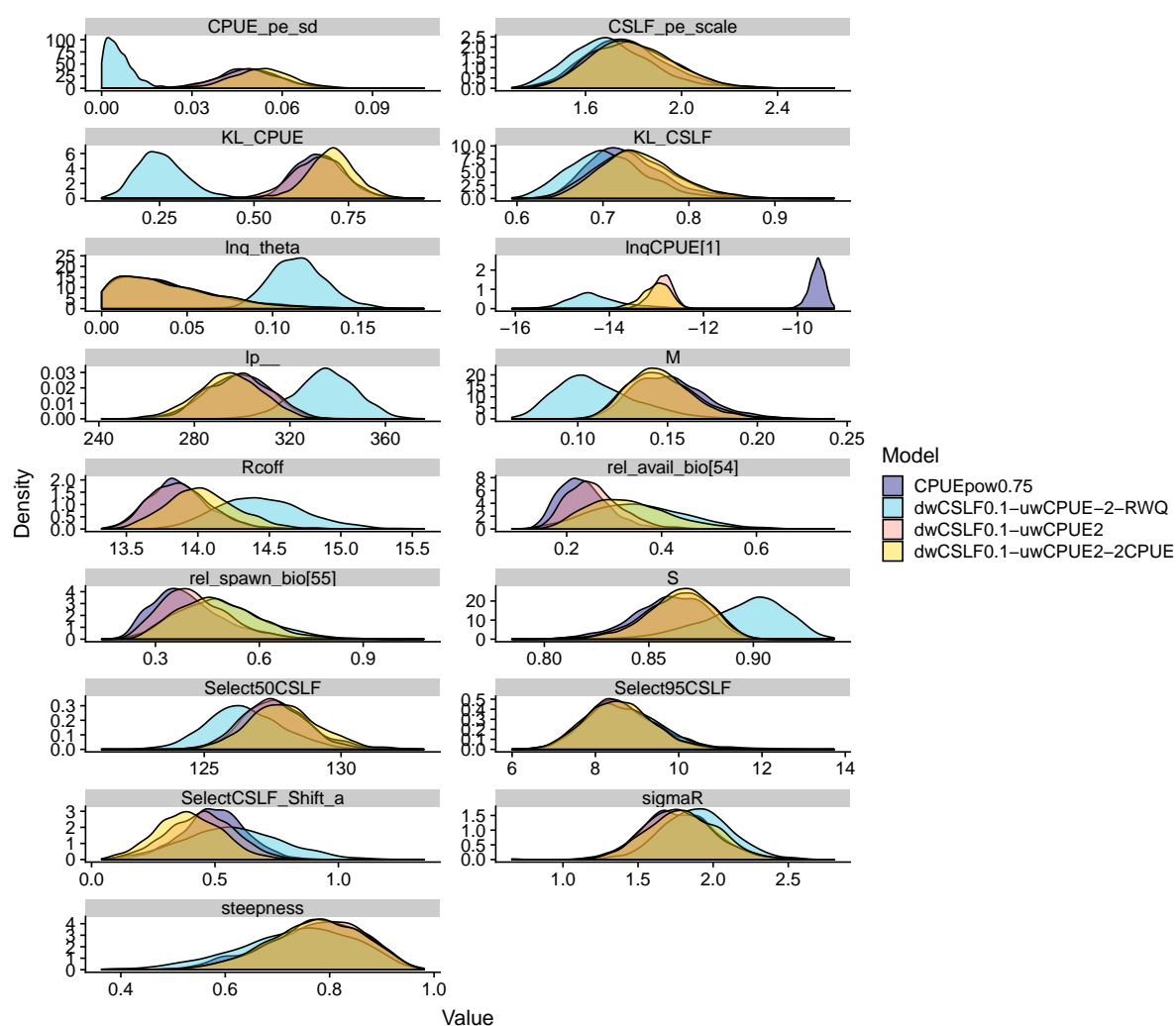


Figure A-27: Comparison of posterior densities for parameters (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) data-sets, but varying structural assumptions. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with structural assumptions labelled as: RWQ= random walk on q ; 2CPUE= split CPUE (reflecting a change in catch reporting from catch effort landing return (CELR) to paua catch effort landing return (PCELR) forms); CPUEpow0.75 = hyper-stable CPUE with $\beta = 0.75$.

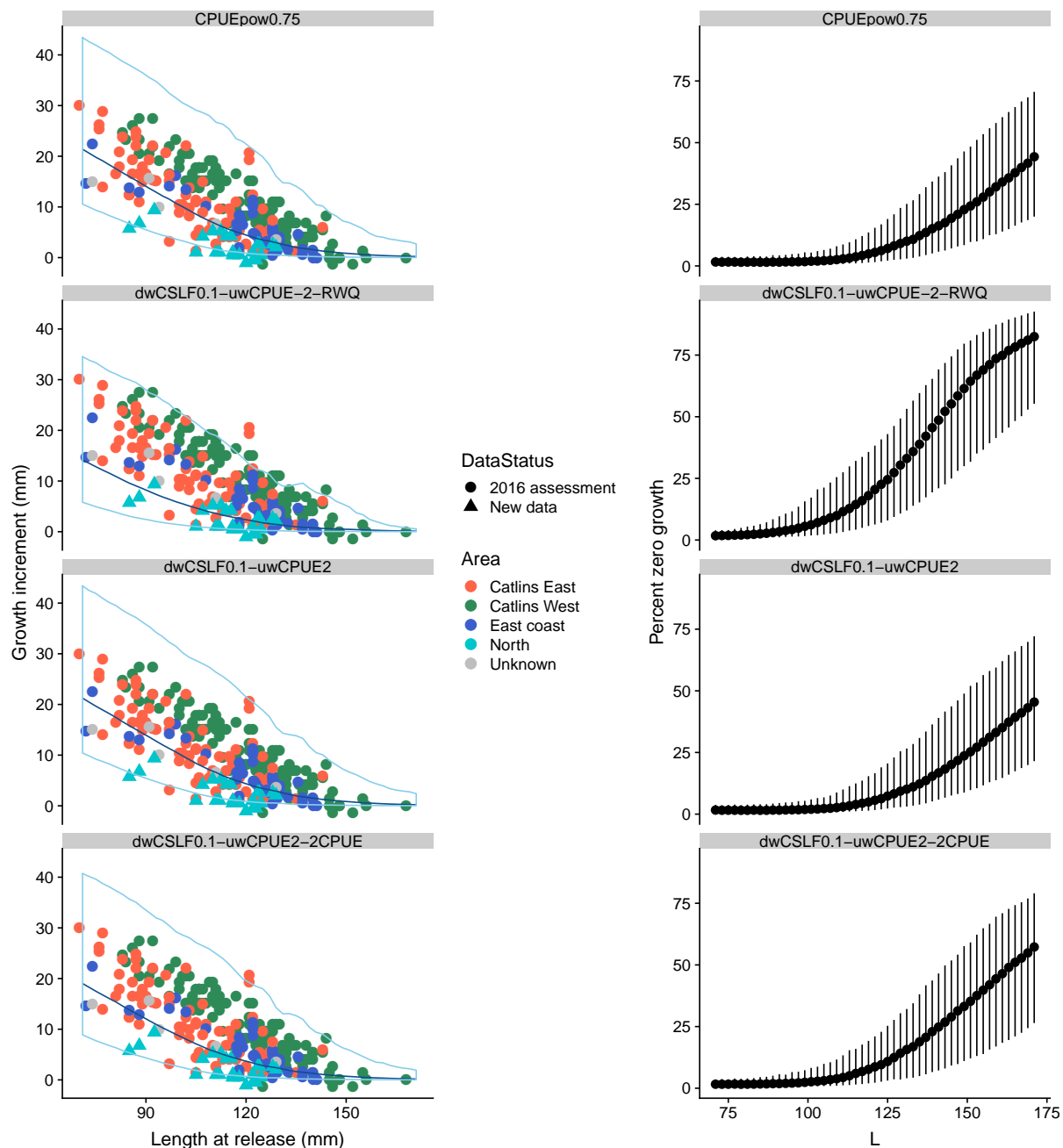


Figure A-28: Comparison of posterior mean growth (population mean (dark blue line) and standard deviation (light blue ribbon))(using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying structural assumptions. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with structural assumptions labelled as: RWQ= random walk on q; 2CPUE= split CPUE (reflecting a change in catch reporting from catch effort landing return (CELR) to paua catch effort landing return (PCELR) forms); CPUE_{pow0.75} = hyper-stable CPUE with $\beta = 0.75$.

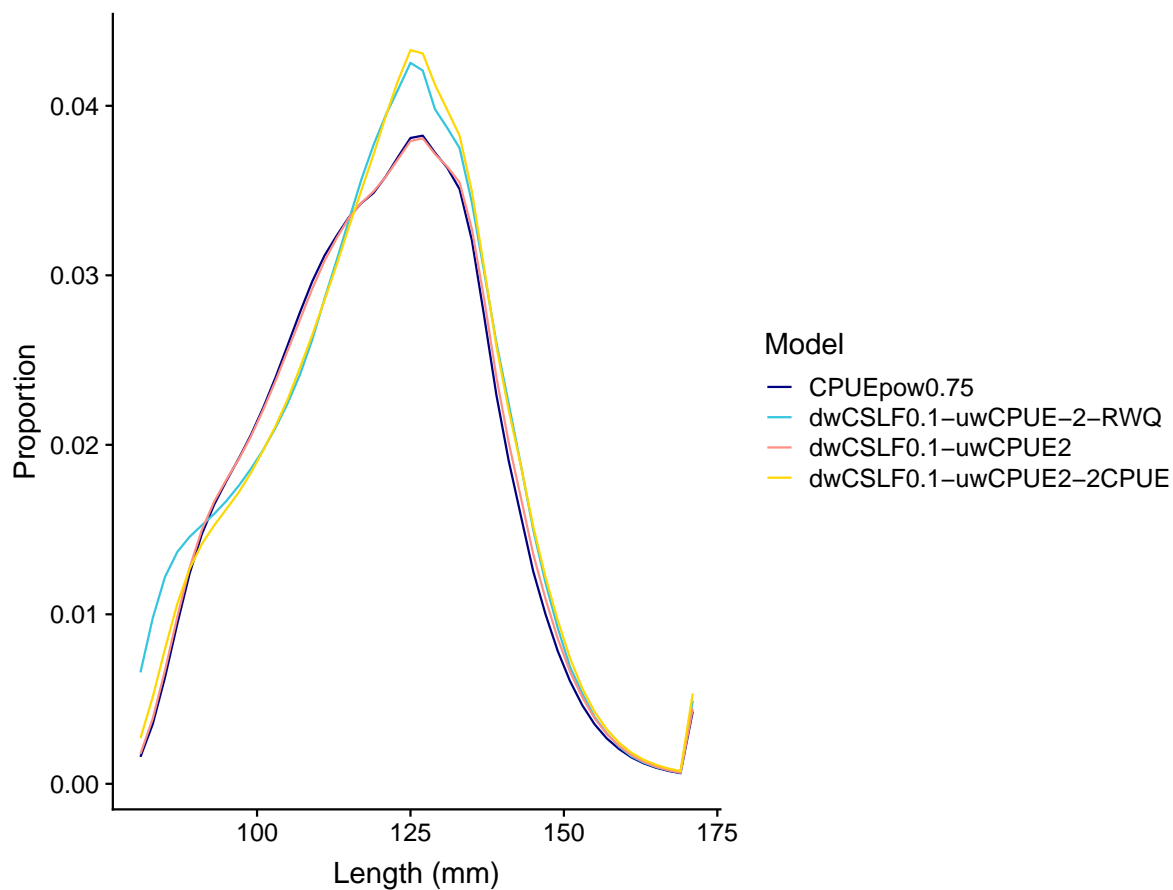


Figure A-29: Comparison of posterior mean proportions-at-length (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying structural assumptions. Model runs are labelled with the weights as: downweighted (dw<Dataset>(amount)), upweighted=uw, and with structural assumptions labelled as: RWQ= random walk on q ; 2CPUE= split CPUE (reflecting a change in catch reporting from catch effort landing return (CELR) to paua catch effort landing return (PCELR) forms); CPUEpow0.75 = hyper-stable CPUE with $\beta = 0.75$.

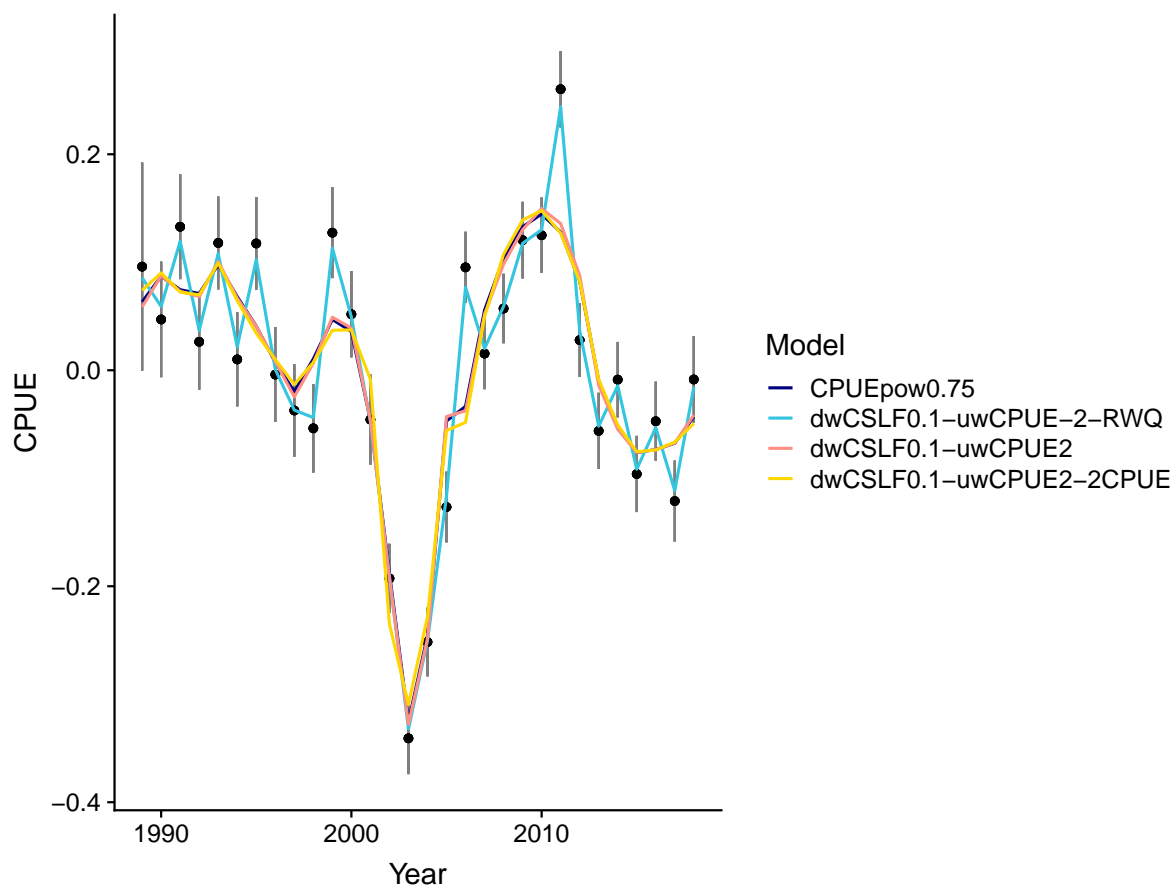


Figure A-30: Comparison of posterior mean predicted catch-per-unit-effort (CPUE) with estimated CPUE index and assumed process and observation errors (black points and error bars), using a direct translation of the ADMB model with modified growth priors (multivariate normal on growth parameters), and modified parameters for mean growth lowered by a factor of 0.6, 0.8 and 0.9.

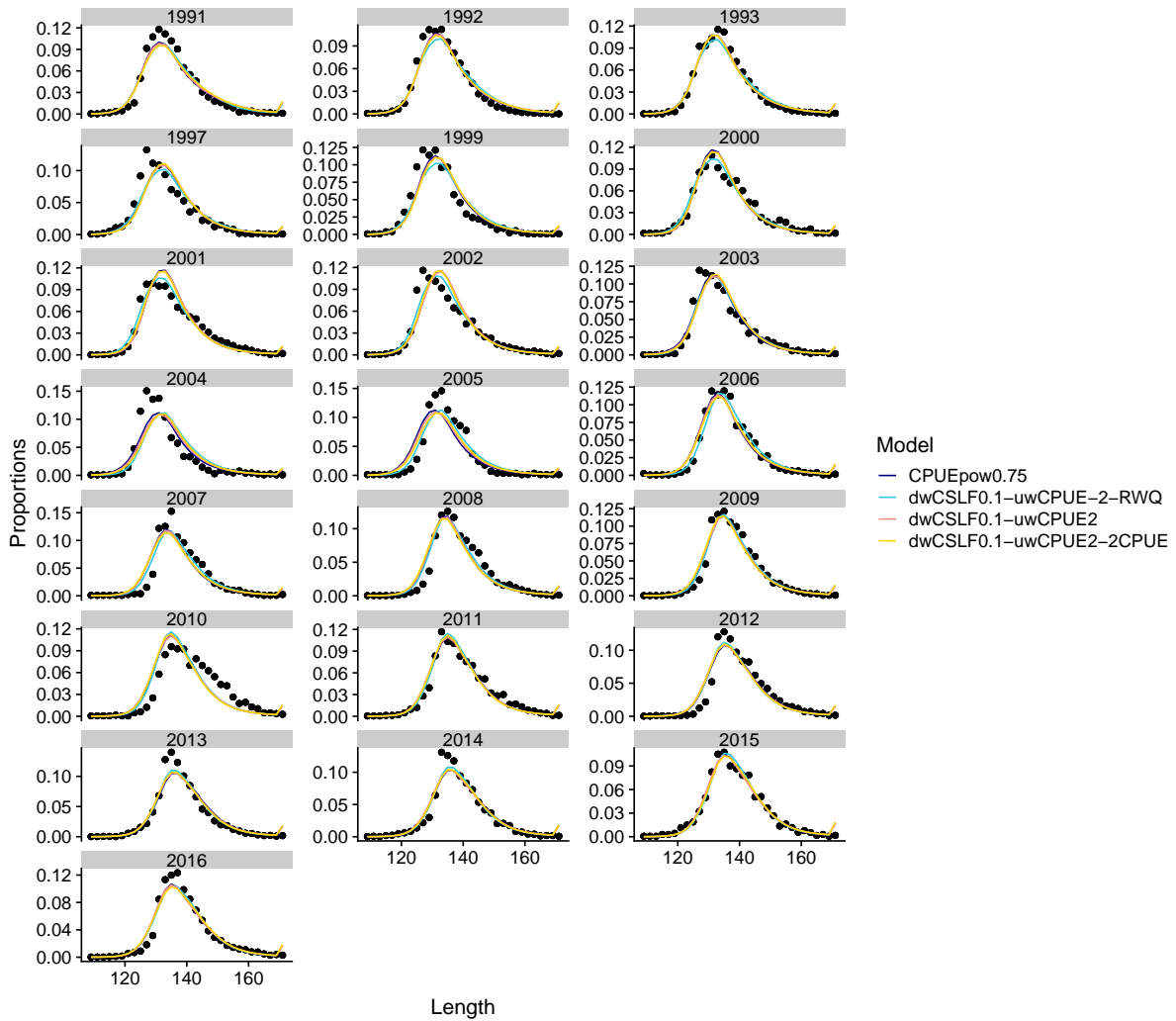


Figure A-31: Comparison of posterior mean predicted catch sampling length frequency (CSLF) with estimated CSLF proportions (using the model described in Section 2.1), with constant weights for CSLF and catch-per-unit-effort (CPUE) datasets, but varying structural assumptions. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with structural assumptions labelled as: RWQ= random walk on q ; 2CPUE= split CPUE (reflecting a change in catch reporting from catch effort landing return (CELRL) to paua catch effort landing return (PCELRL) forms); CPUEpow0.75 = hyper-stable CPUE with $\beta = 0.75$.

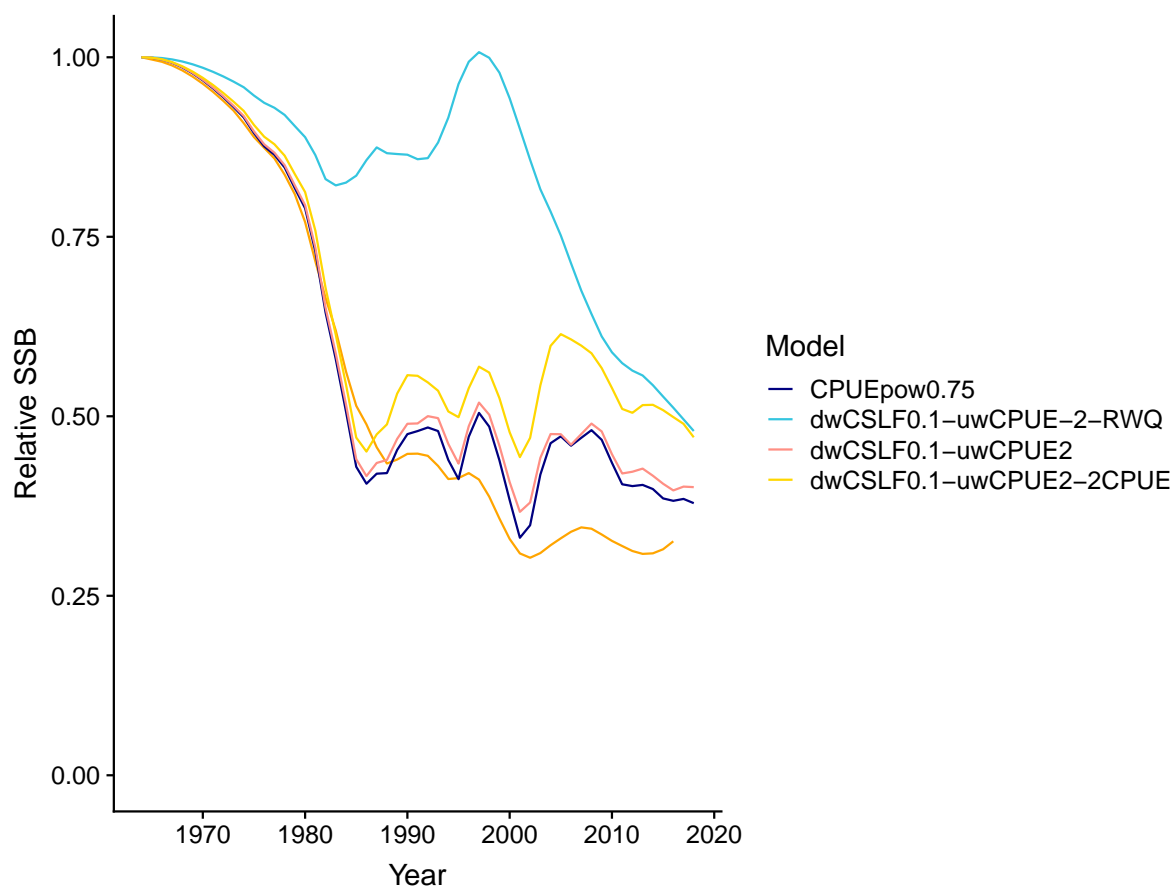


Figure A-32: Comparison of posterior median predicted relative spawning-stock biomass trend (using the model described in Section 2.1), with constant weights for catch sampling length frequency (CSLF) and catch-per-unit-effort (CPUE) datasets, but varying structural assumptions. Model runs are labelled with the weights as: downweighted (dw<Dataset><amount>), upweighted=uw, and with structural assumptions labelled as: RWQ= random walk on q ; 2CPUE= split CPUE (reflecting a change in catch reporting from catch effort landing return (CELR) to paua catch effort landing return (PCELR) forms); CPUEpow0.75 = hyper-stable CPUE with $\beta = 0.75$.

APPENDIX B: INDIVIDUAL MODEL RUNS

B.1 Base case: Model dwCSLF0.1 uwCPUE2

B.1.1 Markov Chain Monte Carlo and posteriors

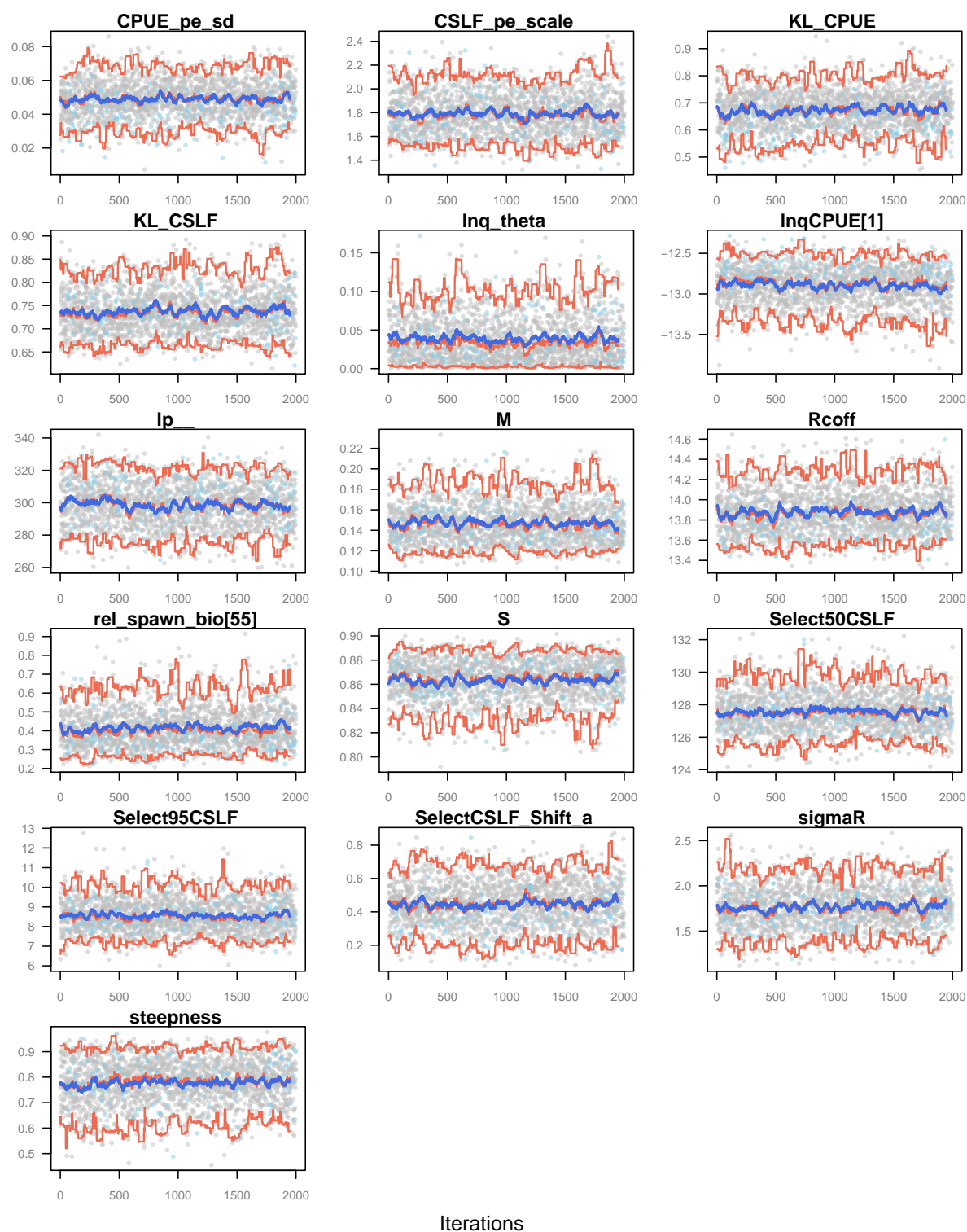


Figure B-33: Traces of Markov Chain Monte Carlo estimation for the marginal posterior distribution of key model parameters for the base case stock assessment model of pāua in PAU 5D.

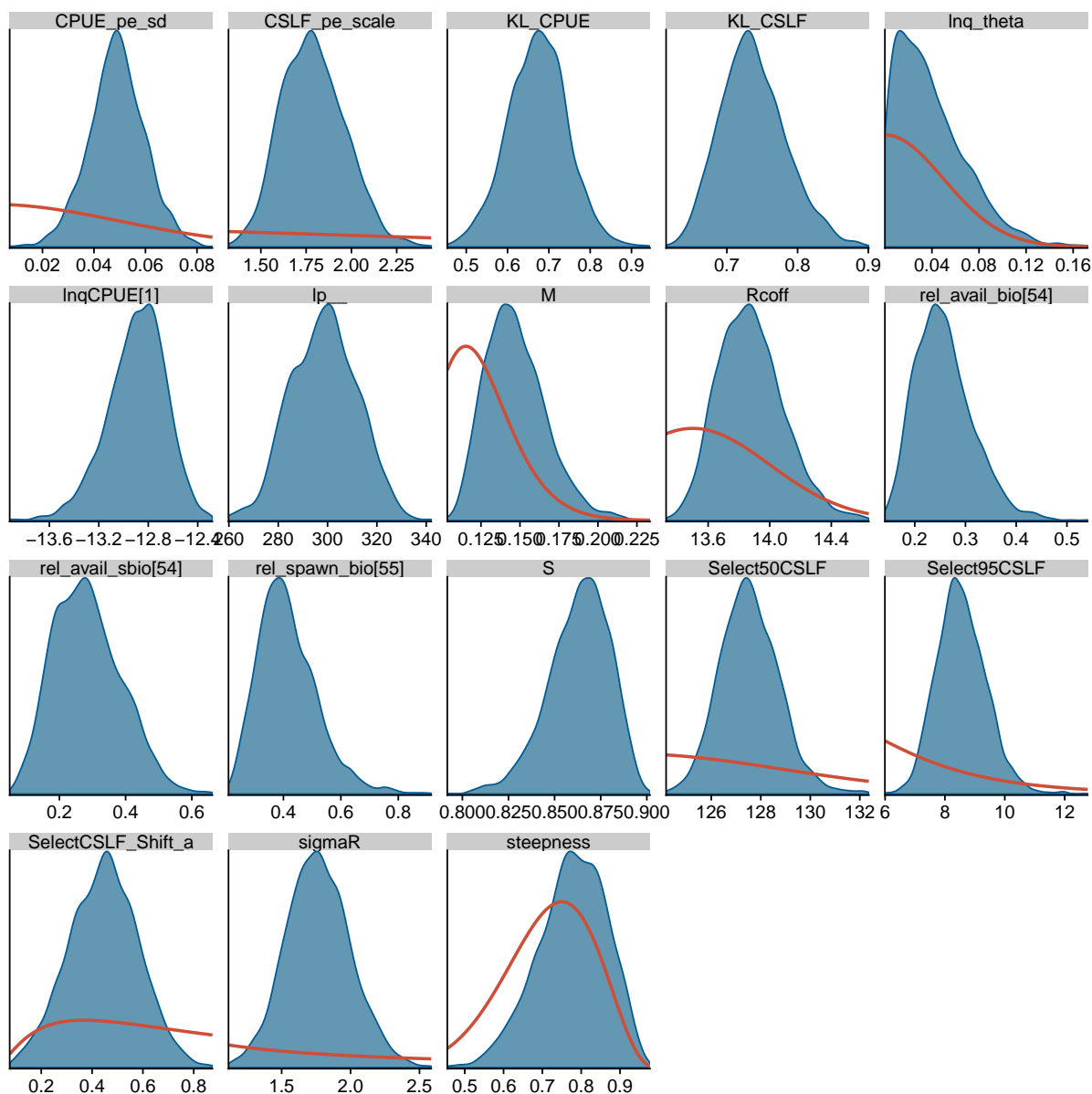


Figure B-34: Marginal posterior densities of key model parameters for the base case stock assessment model of pāua in PAU 5D, with prior densities outlined in red.

B.2 Main sensitivity run 1: Model dwCSLF0.1 uwCPUE2 low MCV

B.2.1 Markov Chain Monte Carlo and posteriors

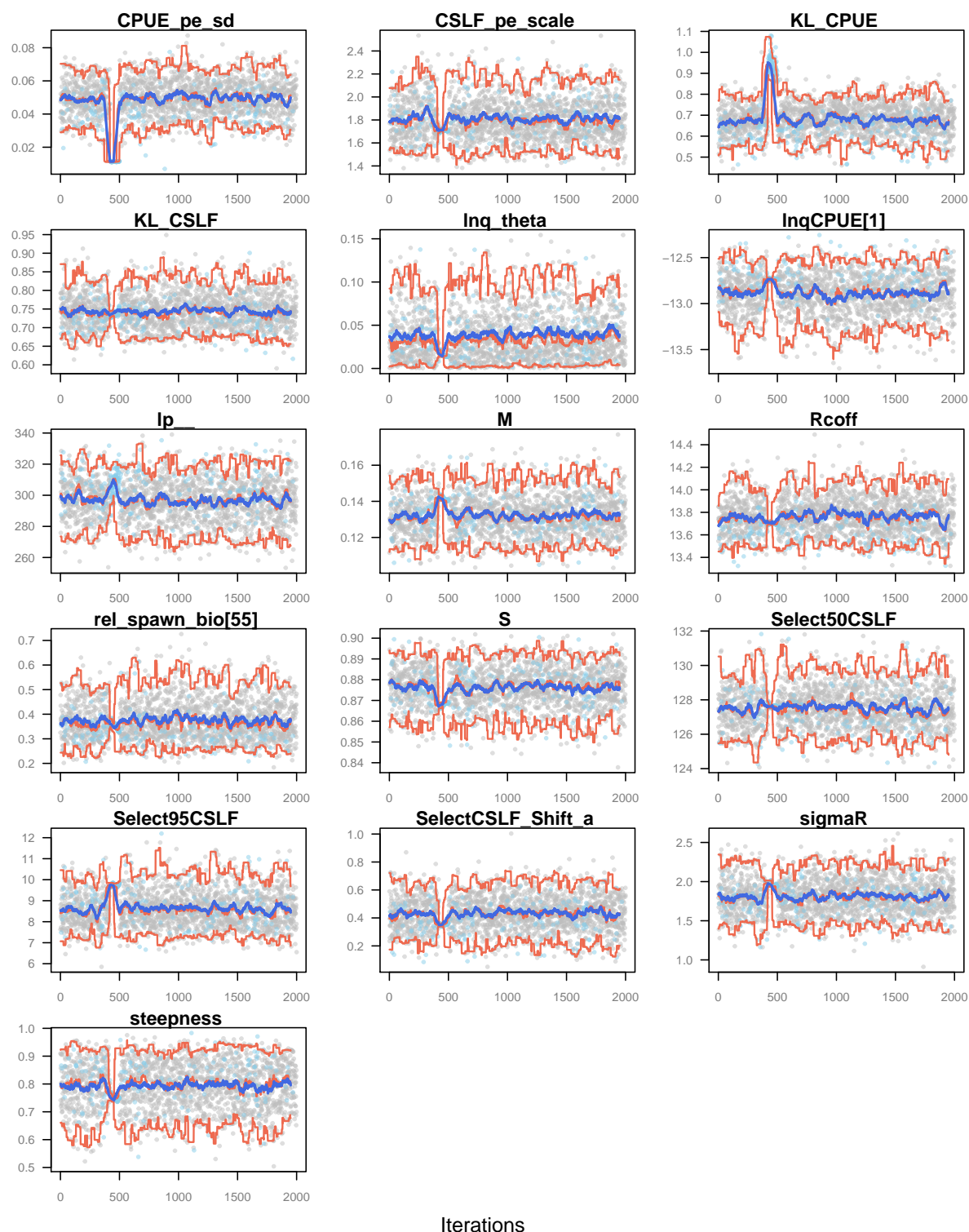


Figure B-35: Traces of Markov Chain Monte Carlo estimation for the marginal posterior distribution of key model parameters for the main sensitivity run 1 of the stock assessment model of pāua in PAU 5D.

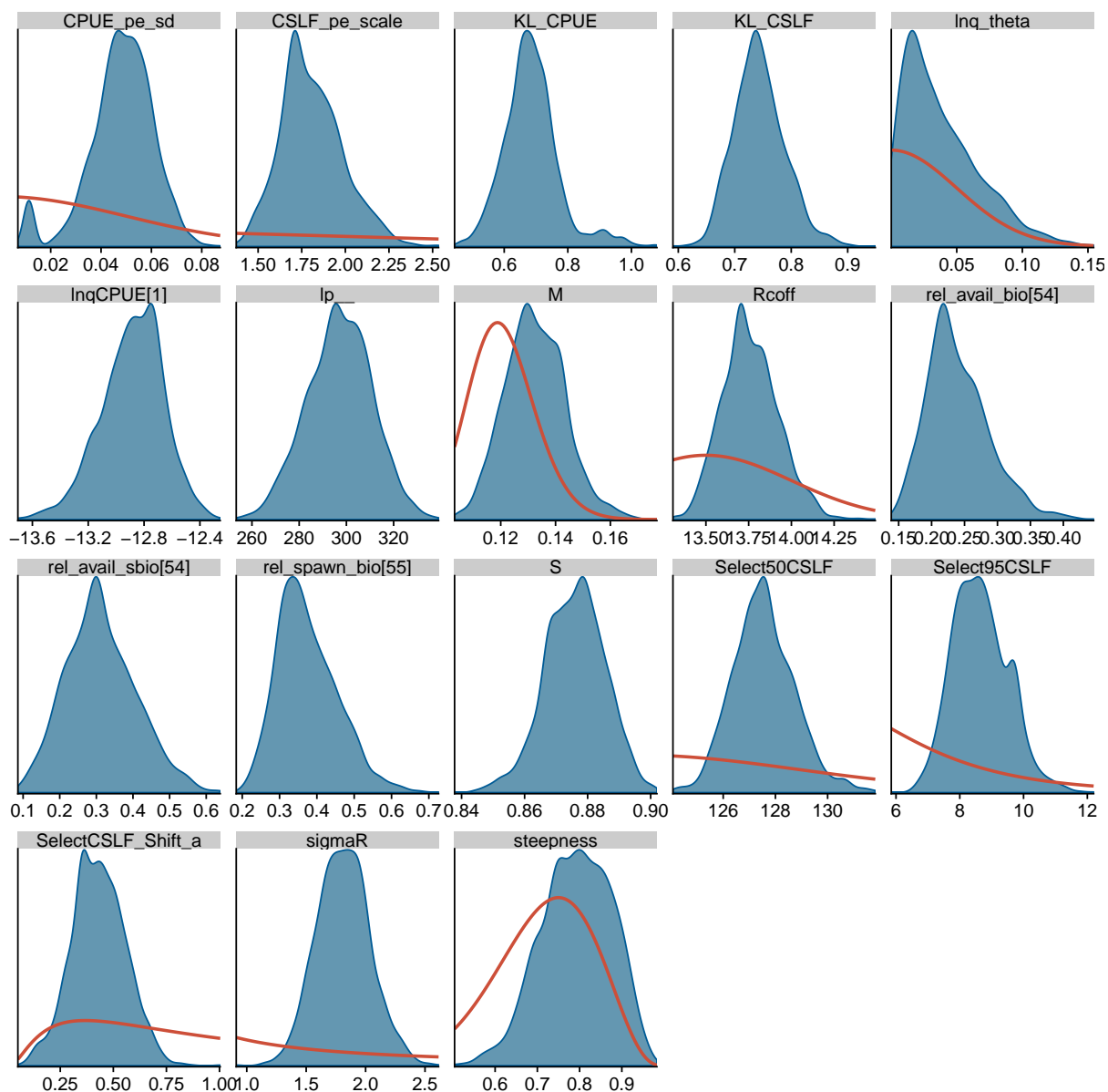


Figure B-36: Marginal posterior densities of key model parameters for the main sensitivity run 1 of the stock assessment model of pāua in PAU 5D, with prior densities outlined in red.

B.3 Main sensitivity run 2: Model dwCSLF0.1

B.3.1 Markov Chain Monte Carlo and posteriors

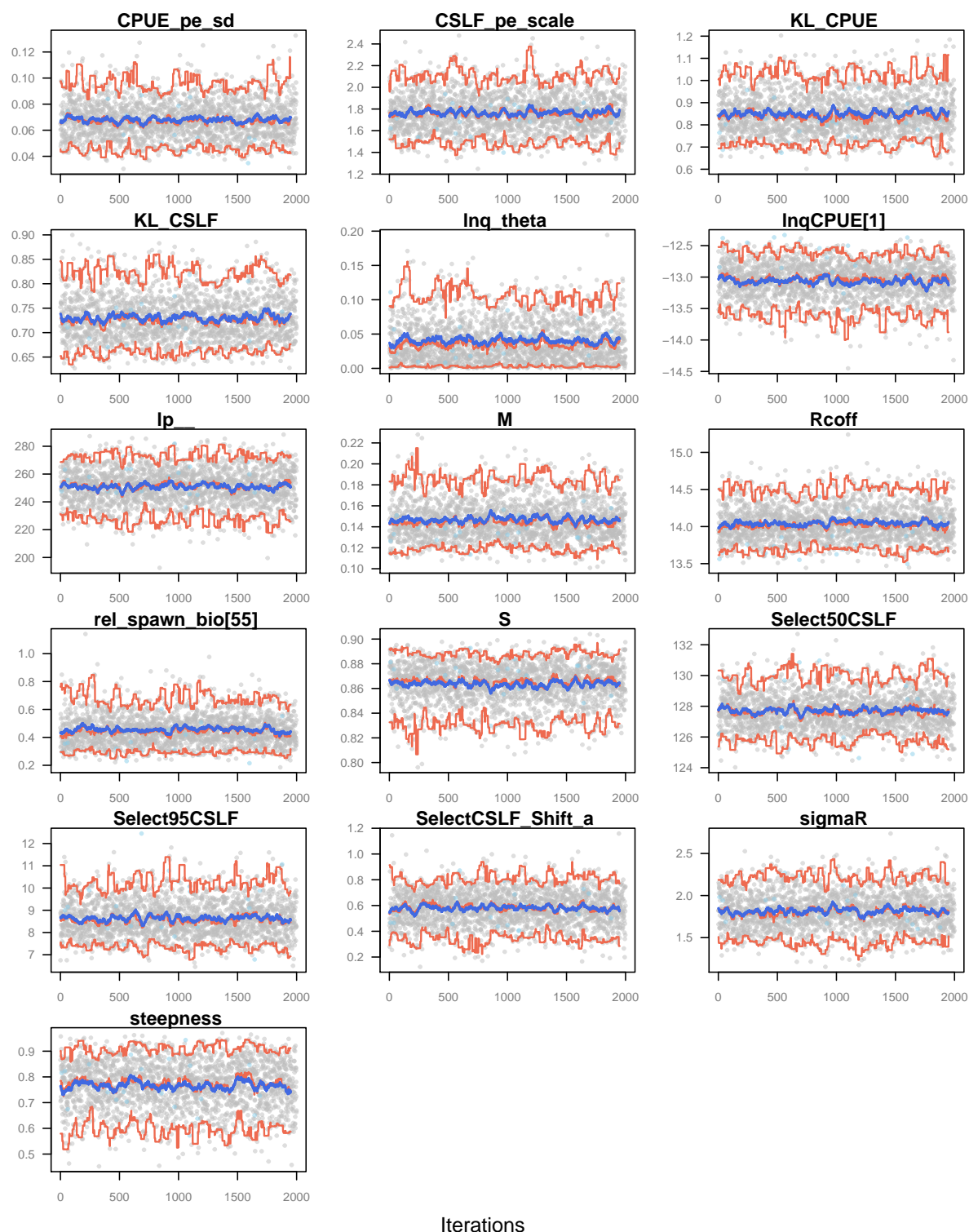


Figure B-37: Traces of Markov Chain Monte Carlo estimation for the marginal posterior distribution of key model parameters for the main sensitivity run 1 of the stock assessment model of pāua in PAU 5D.

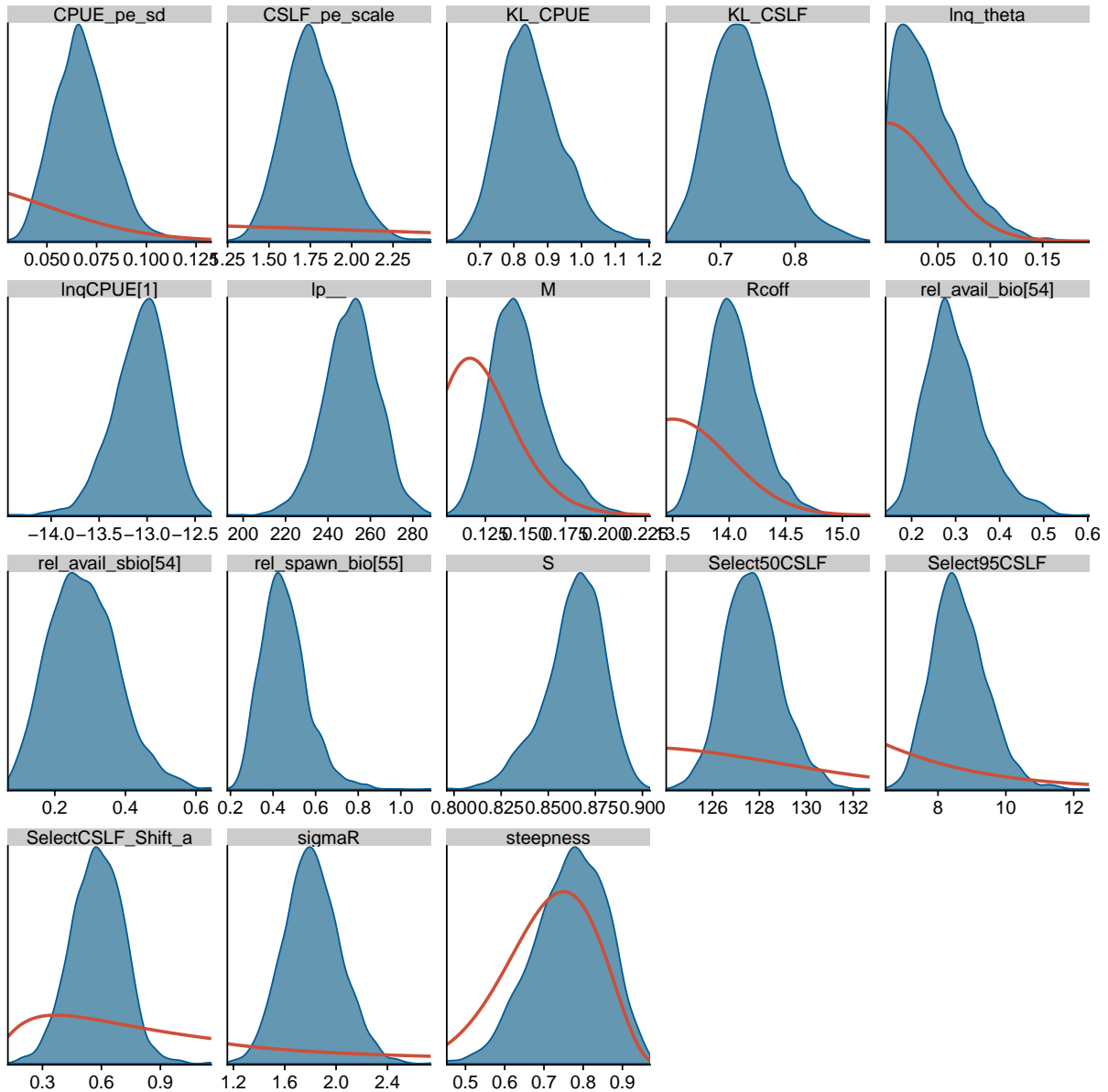


Figure B-38: Marginal posterior densities of key model parameters for the main sensitivity run 1 of the stock assessment model of pāua in PAU 5D, with prior densities outlined in red.