Fisheries New Zealand
Tini a Tangaroa

Input data for the 2018 stock assessment of pāua (*Haliotis iris*) for PAU 5D

New Zealand Fisheries Assessment Report 2019/38

P. Neubauer
L. Tremblay-Boyer

ISSN 1179-5352 (online)
ISBN 978-1-99-000824-5 (online)

September 2019

New Zealand Government

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

TABLE OF CONTENTS

EXECUTIVE SUMMARY

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.**

This report describes the data inputs to the 2018 stock assessment of New Zealand abalone, pāua (*Haliotis iris*), in quota management area (QMA) PAU 5D. The PAU 5D QMA covers the eastern coast of Southland and the Otago coast. Data inputs to the stock assessment were catch, catch-per-unit-effort (CPUE), catch sampling length frequency from commercial sampling (CSLF), growth data from tagging, and maturity. There were three years of additional data since the preceding assessment in 2016, e.g., catch data from 2015–16, 2016–17 and 2017–18.

The 2018 stock assessment also included a number of developments and updates to the modelling approach for CPUE, CSLF and growth and maturity data. For CPUE data, two previously separate time series with different resolutions for key covariates were merged by specifying the relationship between the two resolutions within the model structure. The accuracy of the effort variable was also improved by accounting for the probability that diver effort was reported individually or at the level of the crew. For the length-frequency composition data (CSLF), samples were standardised assuming a Dirichlet-Multinomial error distribution instead of rescaling them as in previous stock assessments. For the growth model, a new formulation based on a meta-analysis of growth across all pāua QMAs led to a new, less informative growth prior for population-wide growth in PAU 5D.

1. INTRODUCTION

The most recent stock assessment of New Zealand abalone, pauā (*Haliotis iris*), in quota management area (QMA) PAU 5D was conducted in 2018 (see Neubauer & Tremblay-Boyer 2019 and a map of the PAU 5D QMA in Figure 1). This report describes the data inputs to the 2018 stock assessment. The fishing year for this fishery is defined to cover the period from 1 October to 30 September the following year (here, we refer to a fishing year by using the latter year in the corresponding time period).

The PAU 5D pāua substock is part of the southern pāua QMA PAU 5, which was split into three subareas in the 1995 fishing year: the eastern Southland and Otago coast (PAU 5D), Fiordland (PAU 5A) and Stewart Island (PAU 5B; Figure 1). More than 90% of the catch in PAU 5D historically occurred in the Catlins region, in the southeast of South Island. The PAU 5 QMA itself was introduced as a QMA in 1986 and allocated a Total Allowable Commercial Catch (TACC) of 445 t, which was subsequently increased to 492 t for the 1992 fishing year. The TACC was split about evenly across substocks in 1995, with an allocation of 149 t for PAU 5D, reduced to 114 t in October 2002, and a further reduction to 89 t the following year.

In addition to the TACC allocated by Fisheries New Zealand, the commercial pāua industry undertakes voluntary management actions outlined in each Annual Operational Plan (AOP). These actions are agreed to and signed by quota and annual catch entitlement (ACE) holders. While the TACC has remained unchanged since 2003, the commercial pāua industry has undertaken voluntary shelving (i.e., non-fishing) of a portion of the quota for the last four years in response to a perceived mismatch between the allocated catch and stock status. In addition, the minimum legal size for fishing of 125 mm shell length was increased in 2010 to 130 mm and 128 mm shell length for key areas in PAU 5D. Also, a number of areas in this QMA are permanently closed to commercial fishing, but remain fished by recreational users. For this reason, it is considered that there is a distinct separation between commercial and recreational fisheries in PAU 5D.

The impact of changes in management areas on catch distribution is difficult to identify as new stock boundaries did not consistently align with previous reporting areas. Also, there was a delay in the adoption of new nomenclature in statutory reporting following each change. Starting in 1997, reporting for PAU 5D was split into 11 statistical reporting areas, further refined to 47 areas in 2002 when reporting at the level of fine-scale statistical areas became mandatory on the Pāua Catch Effort Landing Return (PCELR) forms. No changes to the reporting system have occurred since then (see Table 1 for a summary of reporting standards over the assessment period, and Figure 2 for the repartition of catch-reporting by area since the 1990s. More details on the fishery and its development are available in Fu et al. 2017).

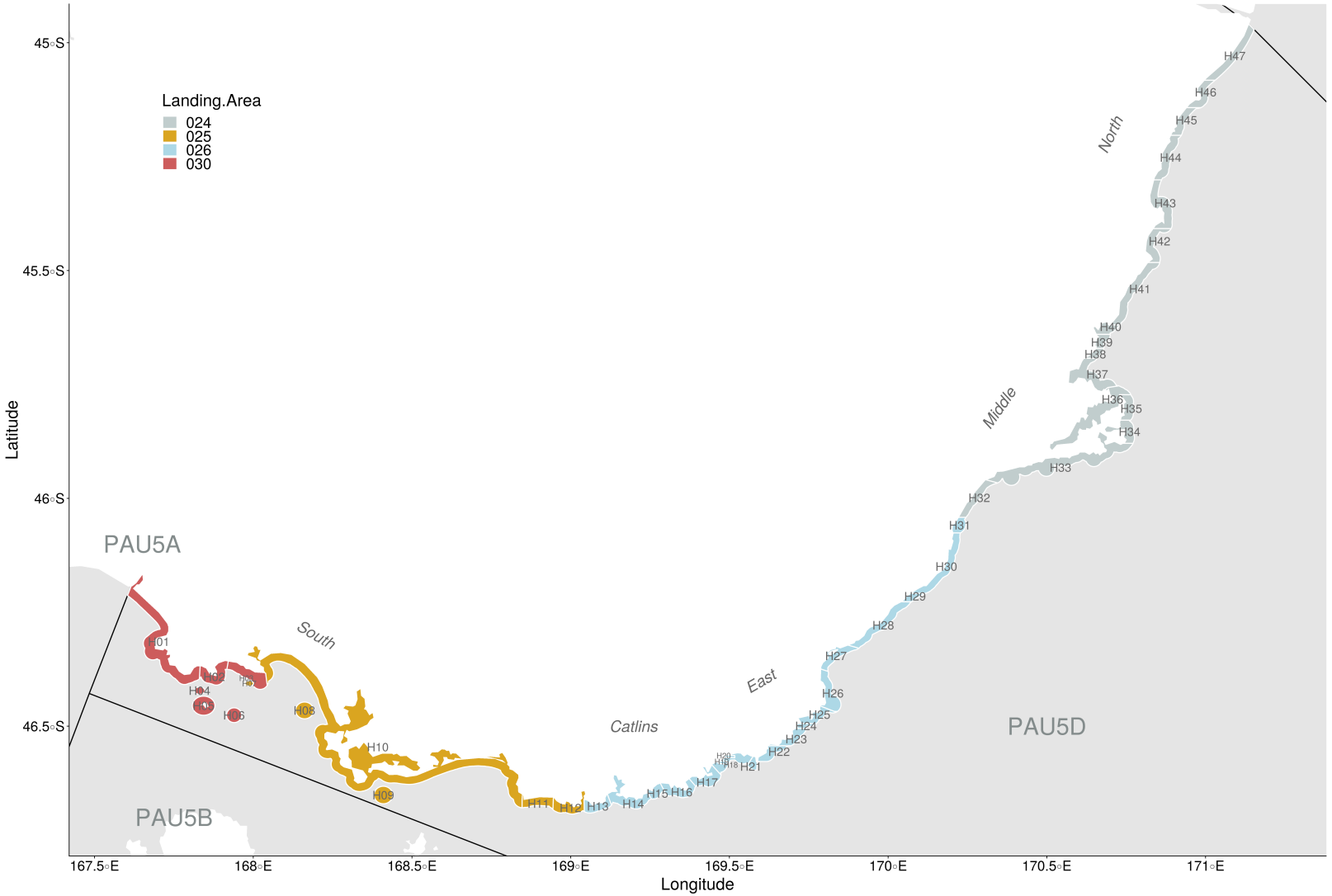

Figure 1: Pāua quota management area (QMA) PAU 5D, surrounding QMAs, and key spatial divisions within PAU 5D. Catch reporting was initially at a lower spatial resolution of General Statistical Areas (024, 025, 026 and 030) and subsequently (in 2002) changed to 47 Paua Statistical Areas (HXX labels). These fine-scale areas may be grouped within broader regions: South (H01 to H10), Catlins (H11 to H20), East (H21 to H25), Middle (H26 to H40) and North (H41 to H47).

Table 1: History of the spatial extent and resolution for pāua stock in sub-area PAU 5D since 1995, when quota management area (QMA) PAU 5 was split into three sub-areas including PAU 5D.

QMA		Statistical areas		
To Sep 1995	Oct 1995–present	1983–1997	1997–2001	2001–present
PAU5	PAU5D	024	D7–D11	P5DH32–P5DH47
		026	D4–D6	P5DH13-P5DH31
		025 (northern part)	D2–D3	P5DH03 P5DH06-P5DH12
		030 (north-eastern part)	D1	P5DH01-P5DH02 P5DH04-P5DH05

Data inputs to the stock assessment consisted of commercial, recreational, customary and illegal catch (reconstructed for part of the assessment period), catch-per-unit-effort (CPUE) from Catch Effort Landing Return (CELR) forms, catch sampling length frequency from commercial sampling (CSLF), growth measurements from tagging by research divers, and maturity (the distribution of key inputs to the assessment is shown in Figure 3).

These data inputs were updated with three years of additional data since the preceding assessment in 2016, i.e., data from 2015–16, 2016–17 and, for catch data, 2017–18. In addition, the most recent stock assessment included several developments and updates to the modelling approach for CPUE, CSLF and growth and maturity data. More specifically, for the CPUE, two previously separate time-series were merged using a single random-effects model that explicitly partitioned variance among nested factors for different reporting periods. In parallel, direct modelling of the probability that diver effort was reported individually or a the level of the crew meant that older CELR data did not have to be separated into reporting types on an empirical basis. For the commercial catch sampling length-frequency data, samples were standardised assuming a Dirichlet-Multinomial error distribution instead of the previously-applied catch-weighted rescaling; for the growth model, an expanded dataset was assembled to derive a growth prior for PAU 5D from growth data across all QMAs.

Detailed information of analytical approaches that were retained (i.e., not updated) from the previous assessment is provided by Fu et al. (2017), most notably the commercial catch reconstruction and assumptions underlying recreational, customary and illegal catch.

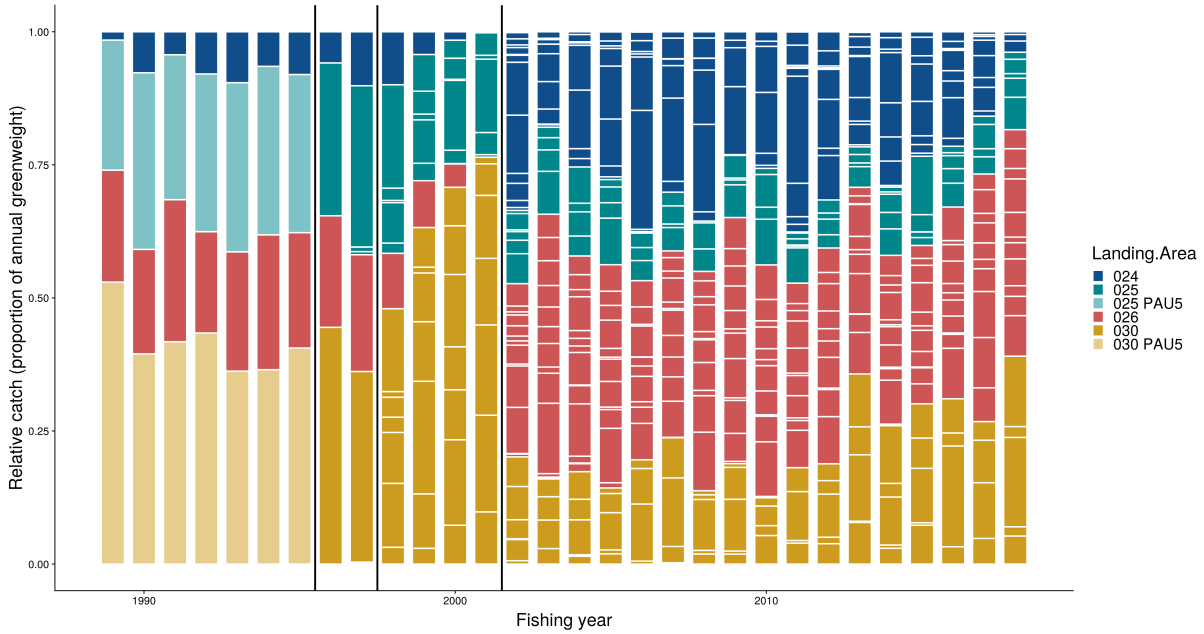

Figure 2: Relative distribution of pāua catch in PAU 5D amongst statistical areas as reporting system changed and the spatial resolution of reporting increased. Coloured bands indicate large-scale landing areas (in a lighter shade for 025 and 030 before 1995 as they straddled multiple quota management areas). Fine-scale areas within landing areas are outlined in white.

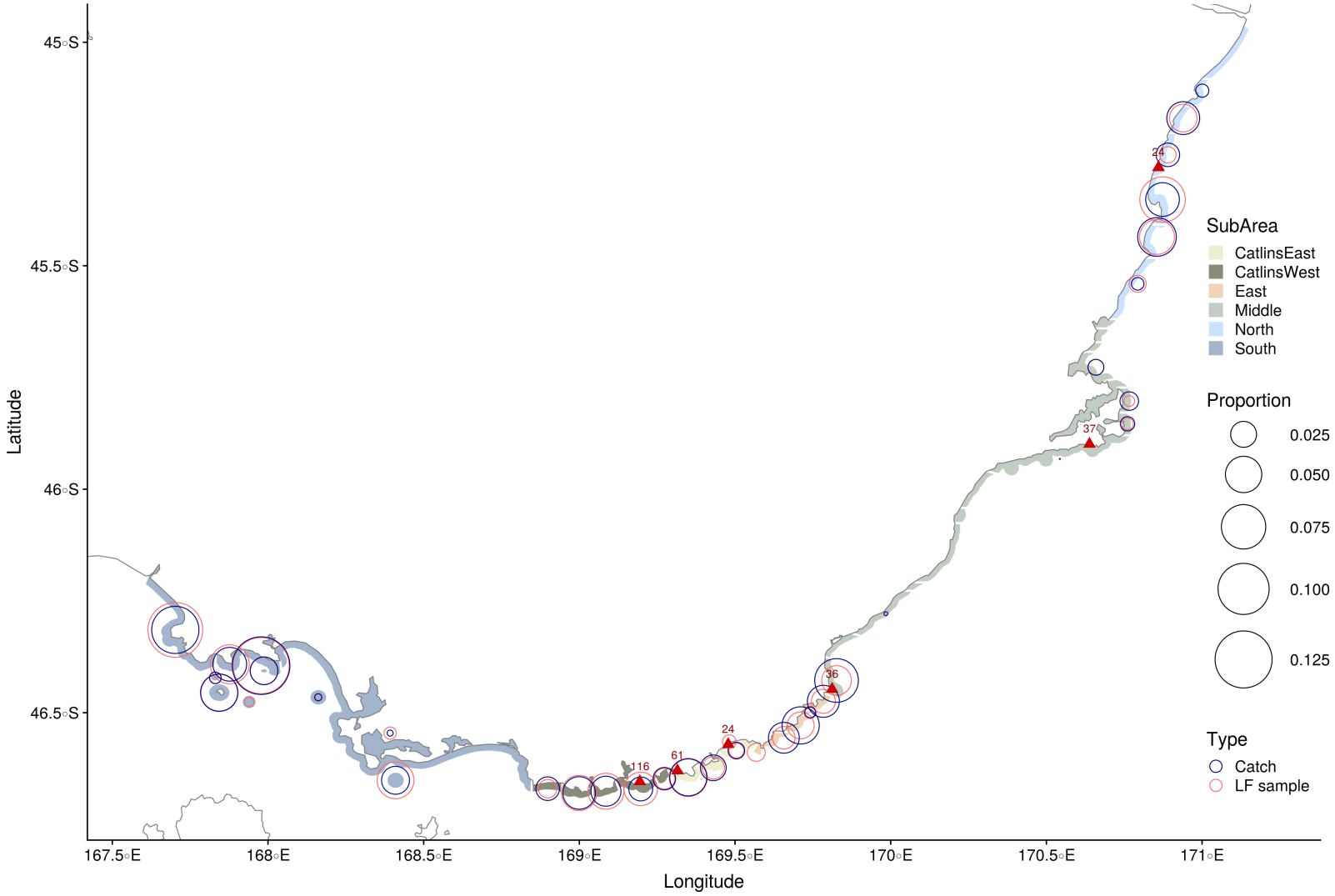

Figure 3: Distribution of key inputs used in the 2018 stock assessment of pāua in PAU 5D. Open circles show the relative proportion of catch (blue) and length-frequency (LF; red) samples from each statistical area (aggregated since 2012); red triangles show the location of growth samples and the number of tagged individuals recovered.

2. CATCH

2.1 Commercial catch

The commercial catch history for PAU 5D is difficult to follow owing to changes in the spatial resolution of the reporting framework (Table 1, Figure 2) and non-overlapping areas between reporting frameworks. More specifically, from 1974 to 1983, catch was reported at the unit of research strata within the PAU 5 QMA covering the south of South Island from Awarua Point on the west coast to Waitaki River on the east coast. Catches within this area could occur in any of the sub-areas now attributed to the PAU 5A, PAU 5B or PAU 5D substocks. Considerable effort by stock analysts in previous years, in consultation with the Fisheries New Zealand shellfish working group (SFWG), has led to a reconstructed pre-1995 catch history under the following assumptions:

- 25% of the catch in PAU 5 between 1974 and 1983 was from the PAU 5D sub-area;
- 7% and 25% of the catch attributed to Statistical Areas 030 and 025 between 1984 and 1995 was from PAU 5D (general Statistical Areas 024 and 026 are entirely within PAU 5D).

Different catch reconstruction scenarios were developed for PAU 5, notably for the breakdown of catch attributed to Statistical Area 030, but there was no effect on catch in PAU 5D. Since 2002, catch has been reported at the level of fine-scale Paua Statistical Areas (on PCELR forms; Figures 4 and 5 show the temporal and spatial distribution of recent commercial catch over fine-scale statistical areas for the period where those forms were implemented). The reconstructed catch history data used for the 2018 assessment was appended with new commercial data for 2016 and 2017 (as agreed by the SFWG). For the 2018 fishing year, the catch was set to 57.85 t in accordance with the 35% TACC shelving agreed by the industry (the reconstructed commercial catch history and recent catches collated from landings are shown in relation to PAU 5 commercial catch in Figure 6).

2.2 Recreational, customary and illegal catch

The recreational catch was assumed to increase linearly from 2 t in 1974 to 10 t in 2005 and subsequently (see Figure 7 for the PAU 5D catch history by catch category). This assumption was agreed upon by the SFWG based on two surveys in 1996 and 1999 (by the National Recreational Fishing Survey) of PAU 5 (no stock breakdown) and a third more recent survey (by the National Research Bureau Ltd). The latter survey was of PAU 5D, but considered to report catch mostly from areas closed to commercial fishing.

A previous sensitivity analysis with the recreational catch increasing to 20t in 2005 had minimal impact on assessment predictions and was not repeated here (Marsh & Fu 2017).

The customary and illegal catch were set to a constant 2 t and 10 t for PAU 5D, respectively, as agreed upon by the SFWG for previous assessments.

3. CATCH-PER-UNIT-EFFORT (CPUE)

3.1 Preparation of CELR and PCELR data

The stock assessment used data from both CELR (1989 to 2001) and PCELR (2002 to 2018) forms to construct a single CPUE index for the assessment. Data preparation procedures generally followed established protocols detailed in Fu et al. (2017). The steps can be summarised as follows:

1. Exclusion of Fisheries Statistical Unit (FSU) data as they represent a low proportion of catch and comprise a substantial number of missing fields.
2. Use only events with “diving” as method.

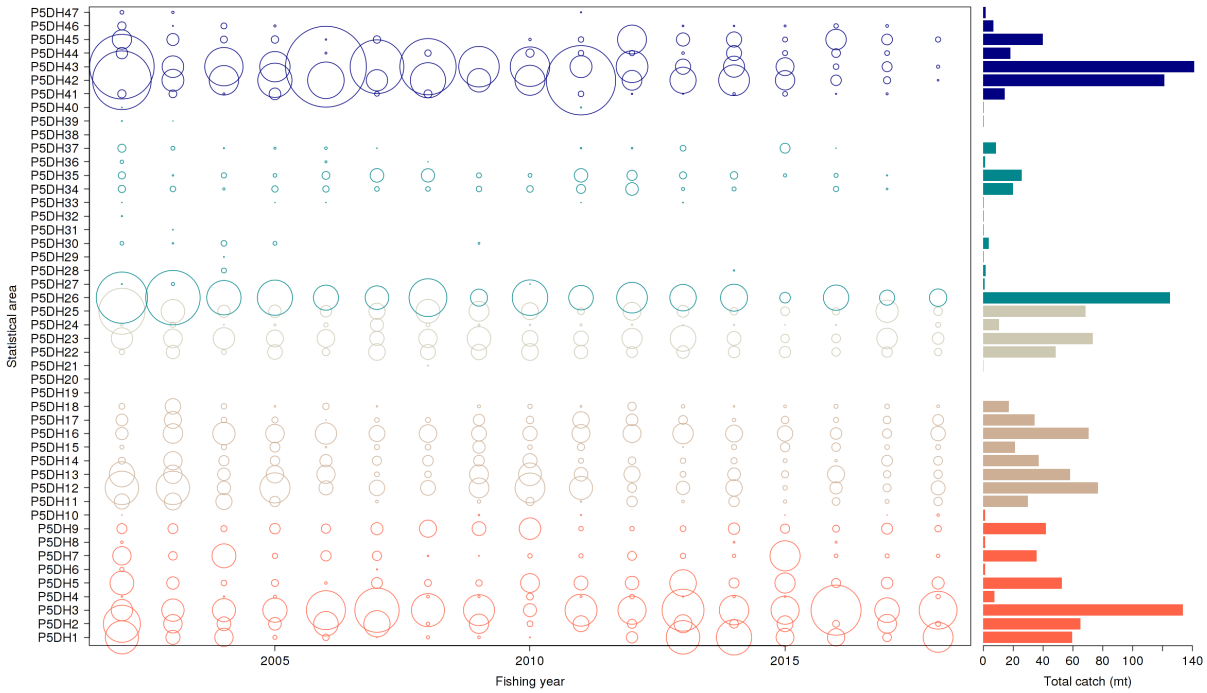

Figure 4: Relative trend in pāua catch (kg) over time by statistical areas in PAU 5D for the period from 2002 to 2018, with total catch over the same time period (right-hand side). Key regions within PAU 5D: South in orange, Catlins in tan, East in light yellow, Middle in turquoise and North in blue.

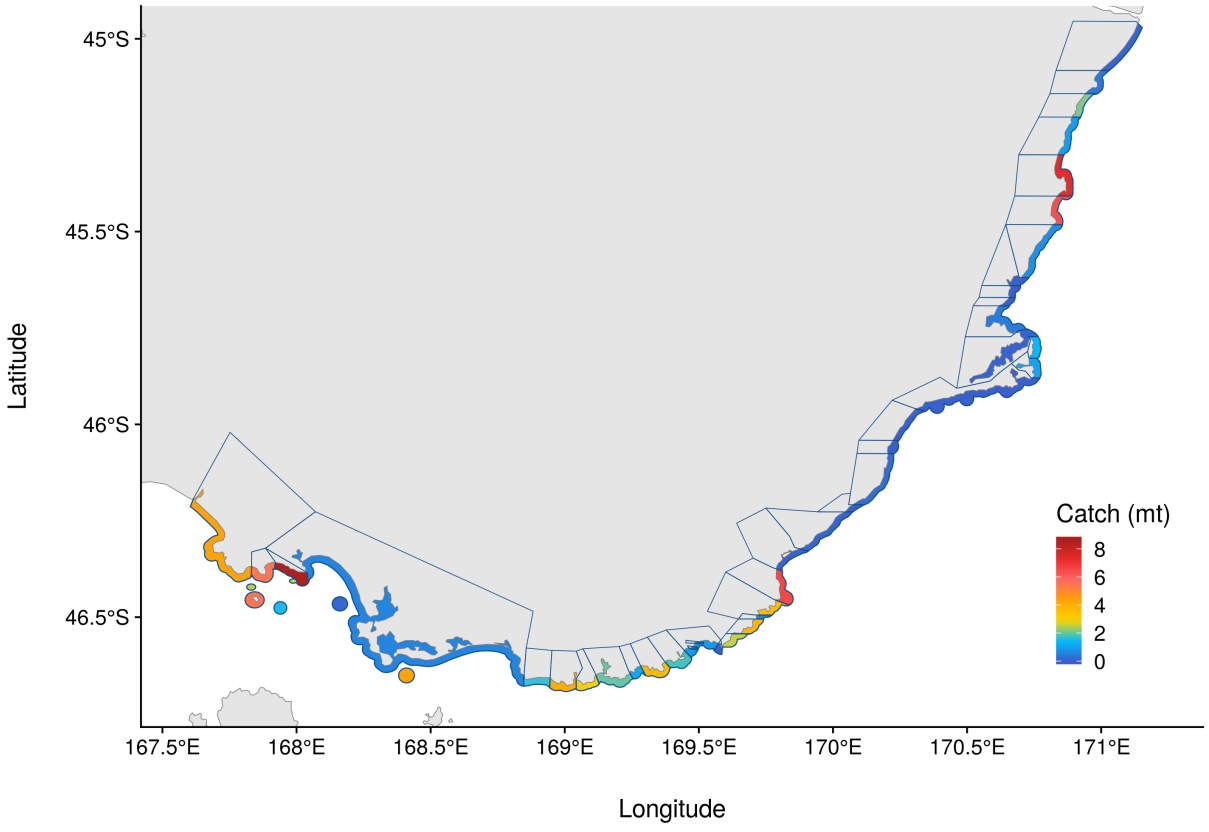

Figure 5: Average annual pāua catch (metric ton, mt) among fine-scale statistical areas in PAU 5D for the period 2002 to 2018.

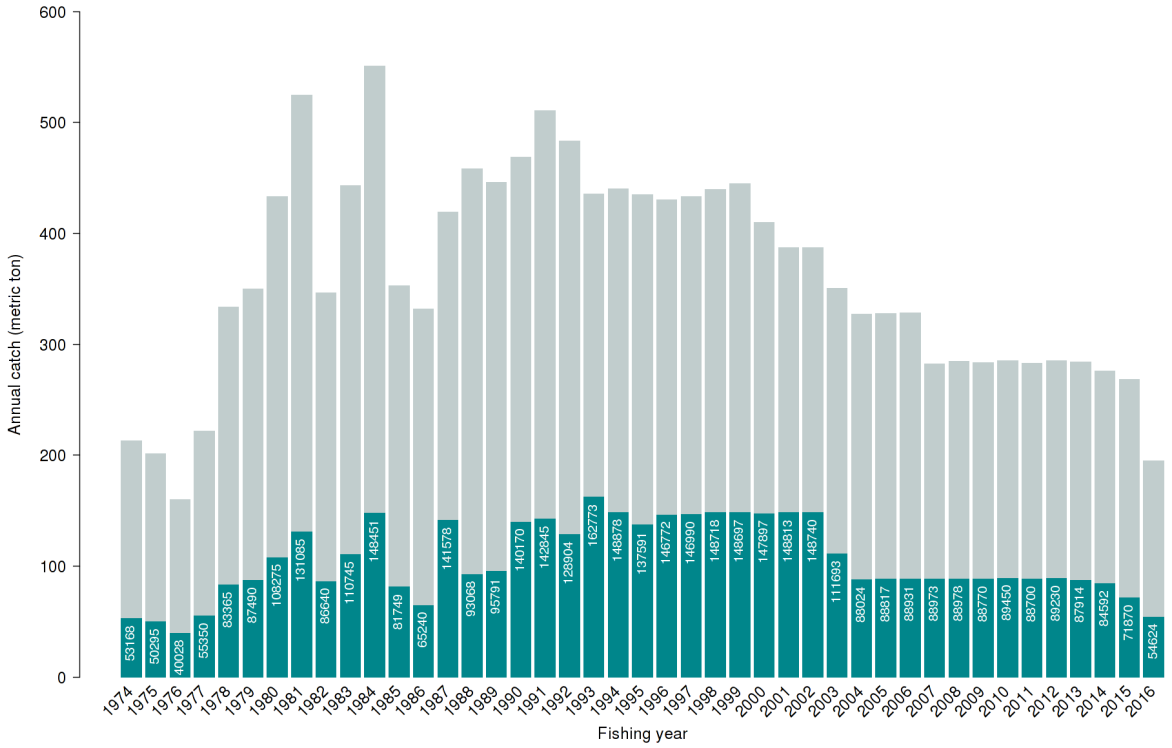

Figure 6: Estimated commercial pāua catch history for PAU 5D (blue-green) and total catch in quota management area PAU 5 (grey). Numbers on bars are exact values for the annual catch (kg) in sub-area PAU 5D.

Figure 7: Estimated pāua catch history for PAU 5D from 1974 to 2018 by fishery component. Fishery categories were commercial, customary, recreational, and illegal catch. Commercial catch was reconstructed up to 1995 when the QMA was created, and based on landing records thereafter.

3. Remove client/fisher identification numbers (FIN) and Statistical Areas that account for little diving (fewer than 50 events over all years). (Note, this step is for computational convenience only as the random-effects models employed in the 2018 stock assessment were not affected by these data.)

4. Allocate time on a per-diver basis in CELR data.

For CELR data, records were required to be recorded as combined effort, but were usually recorded as individual effort (i.e., hours per diver instead of hours for all divers). In the data preparation of recent assessments, this aspect was addressed by using a cut-off time for fishing duration. For fishing durations above this threshold time, data were considered to represent combined effort. The threshold value was set by the SFWG upon inspection of fishing duration times recorded for individual divers and comparison with fishing duration records for events with more than one diver (Figure 8). The cut-off time in the previous assessment was chosen as 10 h.

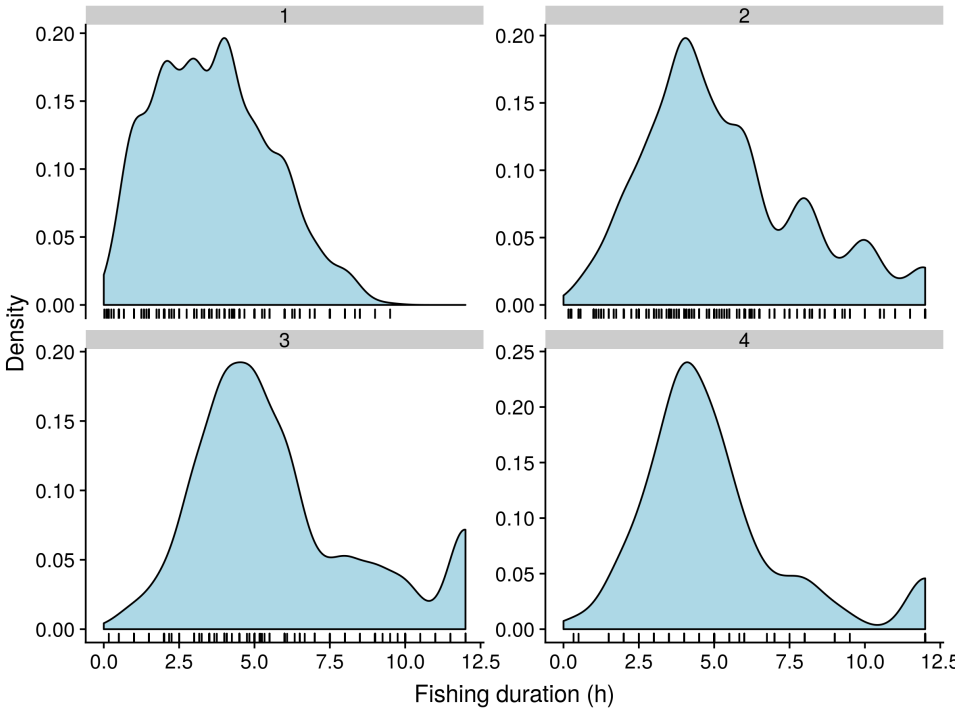

Figure 8: Reported fishing duration for Catch Effort Landing Return records with one to four divers (density of tick marks indicates the number of observations).

Choosing an arbitrary threshold for duration time may introduce bias – if some records were incorrectly classed as recorded by diver (i.e., not across the whole crew; Figure 9), then CPUE may be biased to some extent, especially if the proportion of these data included in the analysis changes over time (e.g., if reporting practices change across some or all crew members).

To mitigate this risk, we trialled a novel model-based classification procedure that assigned CELR records to per diver or per crew reporting types, based on records of single diver events (PCELR and CELR). The model built a representation of dive times T per diver for each crew as a truncated normal distribution $\mathrm{N}^0(\mu_c, \sigma_c)$ and $\mathrm{N}^0(\epsilon_i\mu_c, \epsilon_i\sigma_c)$, with effort ϵ_i and mean and variance for each crew c, and estimated probabilities $p_{c,y}$ for each crew and year that an event was recorded per diver. The resulting model for reported effort was a mixture-of-normals model, where μ_c and $p_{c,y}$ were estimated using random effects formulations, and vaguely informative priors on random effects variances:

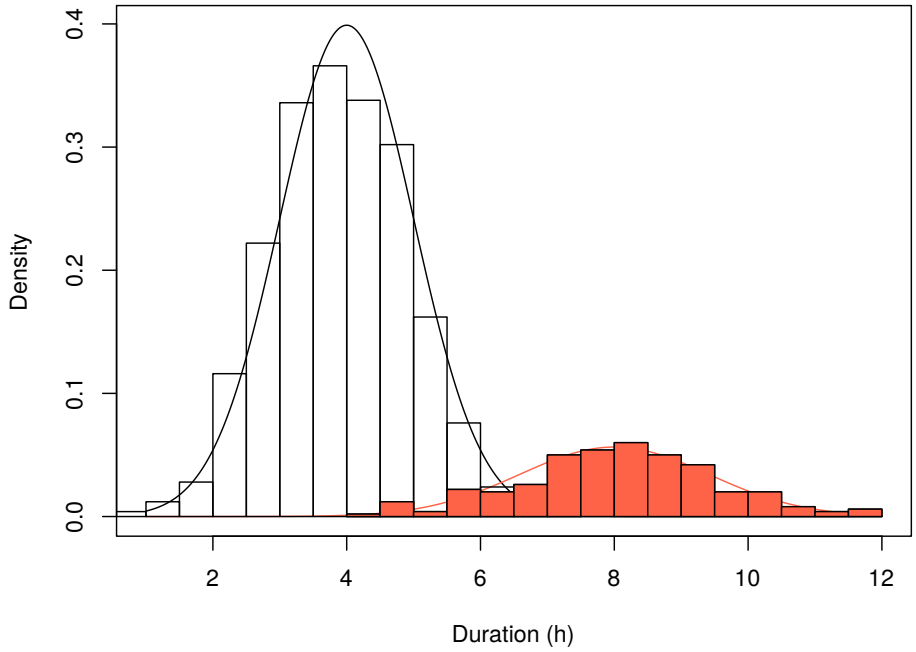

Figure 9: Illustration of potential risk of including data from crews recording combined effort when choosing an arbitrary cut-off time (e.g., 8h) to distinguish between records with duration per diver (black density) and records with duration per crew (i.e., combined effort; red density). Combined observed effort is shown as a histogram, illustrating the difficulty of choosing a definite cut-off time.

$$T_{i,c,y} \sim p_{c,y}\mathrm{N}^0(\mu_c, \sigma_c) + (1 - p_{c,y})\mathrm{N}^0(\epsilon_i\mu_c, \epsilon_i\sigma_c) \quad (1)$$

$$\mathrm{logit}(p_{c,y}) \sim \mathrm{N}(\delta_c, \sigma_p) \quad (2)$$

$$\delta_c \sim \mathrm{N}(\Delta, \sigma_\delta) \quad (3)$$

$$\mu_c \sim \mathrm{N}(\zeta, \sigma_\mu) \quad (4)$$

$$\sigma_c \sim \mathrm{N}(0, \sigma_\sigma), \quad (5)$$

with $\sigma_p, \sigma_\mu, \sigma_\delta$, and σ_σ random effects variance parameters. The model was implemented in Stan (Stan Development Team 2018) and estimated using Stan's Markov Chain Monte Carlo (MCMC) algorithms.

In the preparation of CPUE data, we applied one or the other reporting type (i.e., direct use of duration per diver or reported effort divided by the number of divers in a crew), depending on classification outcomes. For any crew/year combinations where $p_{c,y}$ was $p > 0.05$ and $p < 0.95$, data were not used for standardisation as the reporting type was considered too uncertain.

For most crew and year combinations, the estimated reporting regime had high certainty (Figure 10), and the data were, therefore, retained for CPUE analysis (Figure 11). For some crews, the reporting regime changed over time, and for some records, the classification was ambiguous – these records occurred most often for crews with few records overall and/or for single years for some crews.

3.2 CPUE model

Previous CPUE analyses typically separated CELR and PCELR indices, as the PCELR data provided a higher resolution of spatial fishing patterns and the identification (ID) of divers fishing (instead of

reporting client IDs). Separating the two datasets led to two separate, disjoint CPUE indices in the assessment, so that two separate catchability parameters had to be estimated. The alternative was to use a combined dataset for standardisation that discards PCELR information and produces a combined index using only variables available in the CELR data.

In this iteration of the assessment, we partitioned the variance and estimated how much of the residual variance in CELR data was due to factors accounted for in PCELR data. All PCELR variables commonly used in the standardisation were nested within CELR variables: divers were diving for a single client, fine-scale statistical areas were within larger research strata. These findings meant that residual variance in CELR data was due to variance explained by PCELR factors plus a common residual (PCELR + CELR). The model can be written as:

$$\mathbf{CPUE} \sim \mathrm{LN}(\boldsymbol{\mu}, \boldsymbol{\sigma}_{\mathbf{Obs}}) \tag{6}$$

$$\boldsymbol{\mu} = Z\boldsymbol{\zeta} \tag{7}$$

$$\zeta_i \sim \mathrm{N}(0, \tau_i) \tag{8}$$

$$\boldsymbol{\sigma}_{\mathbf{Obs}} = \sigma_{\mathrm{Res}} + \sum_i I^{\mathrm{CELR}} \tau_i \tag{9}$$

where CPUE is the classification-corrected catch-per-unit effort, μ is the mean vector that is determined by random effects (ζ applied to design matrix Z), with each random effect (ζ_i) drawn from a normal distribution with variance τ_i. Some of the random effects in ζ are variables available for PCELR data only – the diver and statistical area effects. As these are not available for CELR, the observation error for CELR data $\boldsymbol{\sigma}_{\mathbf{Obs}}$ is a combination of the variance attributed to PCELR effects ($\sum_i I^{CELR}\tau_i$), and common residual variance σ_{res}. The model was implemented in Stan and estimated using Stan's MCMC algorithms. Outputs from this model were compared to a combined index using only CELR level variables, and an index using the above model, but using an arbitrary time cut-off of 7 h instead of classification outputs.

3.3 CPUE indices

The log-normal CPUE index model fitted the data reasonably well (Figure 12), and produced an index that diverged relatively strongly from the raw CPUE in recent years (Figure 13). Despite a number of variables having comparable effect size in the standardisation model (Figure 14), the standardisation effect (i.e., lower standardised CPUE in recent years compared with raw CPUE) was mainly due to PCELR level variables, namely the diver effect: more catch was taken by more efficient divers in recent years (Figure 15), and in better dive conditions (Figure 16). Client number and statistical area also had an effect on the standardized CPUE, but the trend was less consistent over time (Figures 17 and 18).

3.4 CPUE outputs for the stock assessment

We trialled two approaches to include the estimated CPUE indices in the assessment model. For models that estimated CPUE process error, both CPUE and estimated observation error (σ_{OBS}) were used as direct inputs to the model without further modifications. For models that did not estimate process error, we followed procedures described in Fu et al. (2017): process error was assumed and estimated by fitting a smoother through the CPUE index and calculating a degree-of-freedom-corrected coefficient of variation (CV). This procedure was repeated for three to 12 degrees-of-freedom (df) for the smoother, and the CV corresponding to the smoother with the lowest df corrected CV was applied to the entire time series (9.98%) to represent both observation and process errors (Figure 19).

4. COMMERCIAL SAMPLING LENGTH FREQUENCY (CSLF)

4.1 General description

The pāua catch sampling data ("market sampling") consist of measurements of the basal shell length of pāua landed by the commercial fishery. These data are the longest measurement following the anterior-posterior axis of the shell lip. The basal length measurement does not include the spire (or hump) when overhanging, nor encrusting organisms. The market sampling was carried by staff of the National Institute of Water and Atmospheric Research (NIWA) until 2005–2006 and subsequently by the Pāua Industry Council.

An extract from the CSLF database resulted in 5694 new samples in addition to the 80 327 measurements that were available from the previous assessment. There were 28 sampling events in 2015–2016 and 54 sampling events in 2016–2017. Comparison of the raw ("pre-rescaling") distribution of length-frequency data with previous years found that dispersion for 2016 was greater compared with previous years, but for 2017 was similar to fishing years before 2016 (Figure 20). This finding might be due to the spatial distribution of sampled landings in 2016, which differed to some extent from that of other years (Figure 21); there was also a smaller sample size overall.

Considering median length data by fine-scale statistical areas from 2002 onwards, median lengths were higher in recent years than in the earlier 2002–2005 period; otherwise there were few notable patterns in length-frequency data. There was a clear west-to-east pattern of increasing median length within the South and the Catlins regions that was also noted in the previous assessment. In general, median lengths in these two regions were greater than elsewhere, especially in the recent period (2014 to 2017; Figure 22).

Two approaches were used to process the CSLF data prior to input to the assessment model. In the first instance, the approach used in the previous assessment was reproduced exactly for continuity (Fu et al. 2017). This approach reweights length-frequency samples to adjust for potential mismatches between the distribution of catch and length-frequency samples (Figure 23).

As such, samples were assigned to sub-areas within PAU 5D and length-frequency samples were scaled up to the sample's landing weight and then to the landings for the year in that sub-area using NIWA's "catch-at-age" software (Bull & Dunn 2002). Length-frequency samples were rescaled to add up to 1 (the reweighted length-compositions resulting from the first approach are shown in Figure 24). In the second approach, the length-frequency distributions were standardised by assuming a Dirichlet-Multinomial error distribution (see Section 4.2 for more details).

The size of the hump (spire) overhang has been recorded since 2006–2007. It is considered to be important in high-density regions, where some growth in shell height instead of basal length might occur. We examined the extent of the hump between years and regions for patterns indicative of growth unaccounted for in the basal shell length only (Figure 25). The current 2018 assessment model still used the basal-length data, but the importance of the hump as an alternative indicator of growth was raised with the SFWG, and potential changes to the sampling methodology are under discussion.

4.2 Standardising length frequency input data with the Dirichlet-Multinomial model

Samples of catch-at-length or catch-at-age from fisheries or surveys are often pre-processed outside of the assessment model (Stewart & Hamel 2014, Thorson 2014). The aim of this pre-processing is to account for known sources of variability when estimating stock-level composition data and to estimate appropriate weights for composition data that would result if the data had been independent. Model-based methods for standardising compositional data are relatively rarely used, but provide a means to address both of these objectives in a single step (Thorson 2014), similar to abundance index standardisation methods. In addition, these methods allow compositional data to be used in the same way as CPUE data, thus eliminating differences in data treatment that are largely due to the compositional nature of the catch-at-length/age data.

Thorson (2014) proposed two methods for standardising compositional data in a model-based framework: i) a normal approximation to individual multinomial proportions, and ii) a Dirichlet-Multinomial (DM) model. A simulation study found that the normal approximation was mainly unbiased due to accounting for known sources of variability in compositions, whereas the Dirichlet-Multinomial model was biased. Nevertheless, this bias was likely to be due to an uneven setup between the two modelling frameworks: unlike the setup for the normal approximation, the DM model was not structured in a way that could capture spatial variability in the sampled proportions and was, therefore, not adequately standardising for this variation. Here, we designed the DM model differently to Thorson (2014) to correct for this limitation: instead of using aggregated composition data by year and estimating a single overdispersion parameter, we used the raw composition samples and attributed overdispersion to known sources of variability (namely, spatial variability). With this approach, we standardised for sources of variability in much the same way that CPUE standardisation partials out variation from known sources in CPUE.

The conceptual model setup was as follows: we started with the assumption that for every year y, there is a true stock-level composition π of numbers-at-length. Nevertheless, there is spatial variability in the composition so that in any particular area one will find a true composition $\tilde{\pi}_a$, from which n_a pāua are sampled, giving a sample $\{\hat{\pi}_a, n_a\}$. We assumed that the $\tilde{\pi}_a$ are distributed according to a Dirichlet distribution with parameters β and π, meaning they are random samples of the true stock-level proportions π with β a concentration parameter. For lower β, the $\tilde{\pi}_a$ are more variable – β can be thought of as a sample size in a Bayesian context, and is inversely proportional to the variance. Crucially, we assumed that the realised β is not a fixed parameter globally or by year, but depends on which spatial areas were sampled. If a statistical area with a $\tilde{\pi}_a$ that is substantially different from the overall π is sampled disproportionately in a particular year, then β for that year ought to be low. To account for this contribution of spatial variability to the variability in the samples, we assumed that the realised variability $\tilde{\beta}$ depends on the true variability of $\tilde{\pi}_a$ around π, and a contribution of the actually sampled spatial areas, so that the realised variability is:

$$\tilde{\beta}_i = \exp(\beta + \omega_a),$$

where ω_a is a spatial random effect. This approach led to a model that effectively attributed variability to spatial sampling intensity, similar to the normal approximation method in Thorson (2014); however, it accounts for the compositional (i.e., non-independent, constrained) nature of the data. The model for observation i in area a can then be written as:

$$\hat{\pi}_{i,a,y} \sim \text{DMN}(n_i, \pi_y, \beta_a), \tag{10}$$
$$log(\beta_a) \sim \text{N}(\beta, \sigma^a), \tag{11}$$
$$\sigma^a \sim \text{N}^0(5), \tag{12}$$
$$\pi_y \sim \text{Dirichlet}(1/L), \tag{13}$$

where DMN is the Dirichlet-multinomial distribution. The DM model was implemented in Stan and parameters were estimated across all length-frequency data for 1992–1994, 1998 and 1999–2017 (please see Fu et al. (2017) for the rationale and reasons for missing data in the omitted years). The estimation used Stan's No U-turn sampler for full MCMC estimation of all model parameters.

4.3 From Dirichlet-Multinomial model outputs to assessment model inputs

The DM model suggested that spatial variation contributed significantly to between-sample variation in proportion-at-length data in PAU 5D (Figure 26). The posterior distributions for proportions-at-length were similar to raw inputs in most years (Figure 27), with overall uncertainty reflecting the amount of data available each year. In some years, the standardisation led to samples of largely small pāua to be downweighted (e.g., 2014–15). For most years, the model-based treatment of length-frequency data

gave similar yearly length-frequency distributions as the previous method, although for 2015–16, the DM standardisation results were more closely aligned with raw length frequencies and more similar to other years. In contrast, the previous re-weighting procedure resulted in a distribution that appeared different from raw length frequencies and from all other years.

Estimated posterior length-proportions for each year were extracted from the model for use in the stock assessment. The posterior co-variance estimated from clr transformed yearly proportions (π_y) from MCMC draws was used as an estimate of observation error for clr transformed proportions data in the final stock assessment model.

5. GROWTH AND MATURATION

5.1 Model setup

Growth was an important determinant of stock status in the previous assessment for PAU 5D (Marsh & Fu 2017), owing to a considerable difficulty in deciding which growth model was most appropriate for pāua in this QMA. Fits to tag-recapture data were considered to be superior using the inverse-logistic growth model (Helidoniotis et al. 2011) to fits with the negative-exponential model; however, fits to commercial length-frequency data were closer for the latter model. Without a clear preference for either model, the decision about model structure introduced a substantial uncertainty in the estimates of stock status, with estimates from the inverse-logistic model being about 15% higher than fits using the negative-exponential model.

Recent developments for pāua growth models suggest that flexible growth models based on energy balance equations (e.g., Ohnishi et al. 2012) can describe observed growth and maturation differences across pāua QMAs (Neubauer unpubl. data). Here, we fitted such a model to provide insights from a formal meta-analysis across all pāua growth data about likely growth patterns in pāua for the PAU 5D area. The model explicitly accounted for statistical area-specific variation in metabolic processes and, thereby, allowed us to gain an explicit understanding of spatial growth patterns in pāua.

The basic growth model is a variation of the von Bertalanffy growth model (vBM), which accounts for length-based allocation of energy to growth and reproduction. The von Bertalanffy growth model (vBM) is based on the assumption that growth is determined by the difference in energy intake (anabolism) and energy requirements to sustain biological functions (catabolsim). This balance is reflected in the equation:

$$\frac{\partial l}{\partial t} = (a - kl),$$

where l is the length in mm, a is anabolism and k is catabolism (also referred to as the Brody growth coefficient). Nevertheless, the von Bertalanffy growth model is considered an incomplete representation of energy allocation processes (Kozlowski et al. 2004, Lester et al. 2004, Quince et al. 2008, Minte-Vera et al. 2016). To address this short-coming, a number of models have been recently introduced that incorporate the cost of reproduction (Minte-Vera et al. 2016). Here, we followed the general model of Ohnishi et al. (2012) so that:

$$\text{vBM}(t) = \frac{\partial l}{\partial t} = (1 - y(l))(a - kl),$$

where $y(l)$ is a logistic allocation function that determines energy allocation to reproduction as a function of length l. In practice, this model can be fit to both length-at-maturity and growth data; however, $y(l)$ and length-at-maturity describe two separate processes. The former describes energy allocation to reproduction, whereas the latter describes the proportion of animals that are capable of releasing gametes (i.e., can contribute to recruitment). The latter process is a much more rapid transition, whereas the energy invested in reproduction tends to increase more gradually as a fraction of length. To convert from one

to the other, we introduced two parameters that scale from the probability of being classified as mature $p(l)$ to the reproductive allocation $y(l)$:

$$p(l) = \frac{1}{1 - \log(19)\exp(-(l - \text{L50})/(\text{L95} - \text{L50}))}, \tag{14}$$

$$y(l) = \frac{1}{1 - \log(19)\exp(-(l - \text{L50}/s)/(\text{L95} - \text{L50})/s^r)}, \tag{15}$$

where L50 and L95 are the lengths at which 50% and 95% of energy are allocated to reproduction, respectively, and s and r are positive scaling factors.

The statistical model correlated statistical-area-specific deviations in growth parameters a and k, and linked the maturation process to growth via a link between the statistical-area-specific L50 and k. The model further assumed a truncated observation process, whereby negative observed growth was ascribed to observation error (σ_{OE}), as negative weight growth is possible (i.e., animals shrink within the shell), but cannot be observed due to the measurement process (which measures the shell length). For the measurement and observation process we, therefore, have:

$$x_{i,s} \sim \text{Cauchy}(\hat{x}_{i,s}, \sigma_{OE}), \tag{16}$$

$$\hat{x}_{i,s} = \left\{ \begin{array}{ll} ,0, & \text{for } \dot{x}_{i,s} \leq 0, \\ \dot{x}_{i,s}, & \text{for } \dot{x}_{i,s} > 0, \end{array} \right\} \tag{17}$$

where $x_{i,s}$ is the observed increment for individual i in statistical area s and $\hat{x}$ is the increment that would have been observed if observation error was zero. The observation process is related to the growth process via a truncation process, which turns any predicted growth increment ($\dot{x}$) less than or equal to zero to a mean of zero growth in the measurement process. The predicted growth increment is calculated from the von Bertalanffy model process using numerical integration over the time at liberty T. The growth process itself is determined by statistical-area-specific variation in k and a, which are formulated as random effects drawn from a multivariate normal prior with estimated co-variance matrix Σ. In addition, we let a vary among individuals to capture some of the within-statistical area variability in growth (ζ). With this approach, we have:

$$\dot{x}_{i,s} = \text{vBM}(l_i, T, k_s, a_{i,s}, \text{L50}, \text{L95}, s, r) - l_i, \tag{18}$$

$$k_s = \mu_k + \omega_{k,s}, \tag{19}$$

$$a_{i,s} = \mu_a + \omega_{a,s} + \zeta_i, \tag{20}$$

$$\omega = \Sigma\delta, \tag{21}$$

$$\delta \sim \text{MN}(0, I), \tag{22}$$

$$\zeta \sim \text{N}(0, \sigma_i). \tag{23}$$

Here, we used MN for the multivariate normal distribution; delta is a $s \times 2$ matrix of un-scaled random effects for a and k, with ω the scaled effects (via the inner product with the estimated covariance matrix Σ). Lastly, we linked maturation to the statistical-area-specific Brody growth coefficient k_s and estimated the strength of the link σ_{km}:

$$p(l) = \frac{1}{1 - \log(19)\exp(-(l - \text{L50} - \sigma_{km}\delta_{k,s})/(\text{L95} - \text{L50}))}, \tag{24}$$

$$\text{l50} \sim \text{N}(90, 10), \tag{25}$$

$$\text{l95} - \text{l50} \sim \text{N}(15, 10), \tag{26}$$

$$\sigma_{km} \sim \text{N}(5), \tag{27}$$

$$s \sim \text{Beta}(2, 1), \tag{28}$$

$$r \sim \text{N}^0(\tau). \tag{29}$$

The model was implemented in Stan to take advantage of numerical ordinary differential equation solvers implemented in Stan. These were used within a MCMC procedure to estimate marginal posterior distributions for model parameters and derived quantities.

The model was fit to a combined dataset of 3313 tag-recapture observation in all South Island QMAs, and 1648 length-at-maturity samples across a combined 60 different statistical areas.

5.2 Model outputs

The model was slow to fit, but with a sufficiently high tree-depth setting, we were able to obtain reasonable MCMC performance; all parameters except the observation error variance were estimated with low error (Figure 28).

Model fits for all statistical areas in the dataset were considered adequate (Figure 29). Nevertheless, in some statistical areas, there were significant outliers that did not fit the overall growth pattern and were too extreme to be accounted for by general within-area variability. Estimated maturation also closely fitted the data (Figure 30).

Estimates from the von Bertalanaffy growth (and maturation) model provided insights into processes that determine growth and maturation variability in pāua. There was a strong negative relationship between L50 and estimated deviations in k, suggesting that animals in statistical areas with higher Brody growth coefficients mature at a smaller size (Figure 31). Nevertheless, this relationship was often offset by a concomitant increase in a for statistical areas with high k, so that the asymptotic size L_∞ was relatively unaffected by changes in k and energy allocation to reproduction (i.e., statistical areas with large animals may be early or late maturing).

For most QMAs, the variation between statistical areas was considerably greater than the variation between QMAs. This finding was confirmed by the initial exploration of this model, which suggested that little variability in the dataset can be attributed to differences between QMAs. For example, in PAU 5D, some statistical areas are amongst the fastest growing sites in the dataset (e.g. statistical area H14), whereas other statistical areas contain slow growing pāua (e.g. H44). In aggregate, however, it appears as though the PAU 5D data was dominated by samples with relatively fast growth. In contrast, data from other QMAs, such as PAU 3, were dominated by areas with slow growth. This aspect poses a challenge when using QMA-specific data in the assessment, as they are required to represent growth across the entire QMA. In PAU 5D, it is unclear what proportion of the total population grows at a rate that is as fast as that suggested by PAU 5D tag re-capture data.

5.3 Deriving a prior for growth and maturation

Using only PAU 5D data and fitting directly to this dataset either within or outside the assessment model provided a precise, but potentially unrepresentative, signal to the assessment model about QMA-wide growth. The resulting uncertainty about mean growth in the QMA was low based on the available tag-recapture data only (Figure 32), and the dataset suggested fast mean growth and a large proportion of pāua growing to large sizes.

Considering the tag-recapture meta-analysis overall, the SFWG agreed that this pattern of fast growth was unlikely to be representative of growth in all of PAU 5D. To provide a less informative but more accurate input to the model for growth, a prior was therefore constructed from the tag-recapture meta-analysis. The prior was derived by predicting growth across all statistical areas in the South Island wide dataset thought to be representative of PAU 5D, and by taking the posterior distribution of the (finite) population mean and standard deviation across statistical areas as a prior for growth in the assessment model.

As a process model, the von Bertalanffy growth model does not admit negative growth: negative growth can only occur if either a or k change, or if a large-sized pāua is released in an area where the estimated growth parameters imply a negative energy balance. Neither scenario is admissible in the current assessment model, which does not incorporate environmental variation or the necessary flexibility to lead to negative growth. Instead, we truncated the prior at zero and added a vector of zero growth proportions at length to the prior and assessment model; therefore, at each length l, a proportion $z(l)$ of the population grows according to a log-normal growth prior, and a proportion $(1 - z(l))$ of pāua is located in areas with no growth at length l (Figure 33).

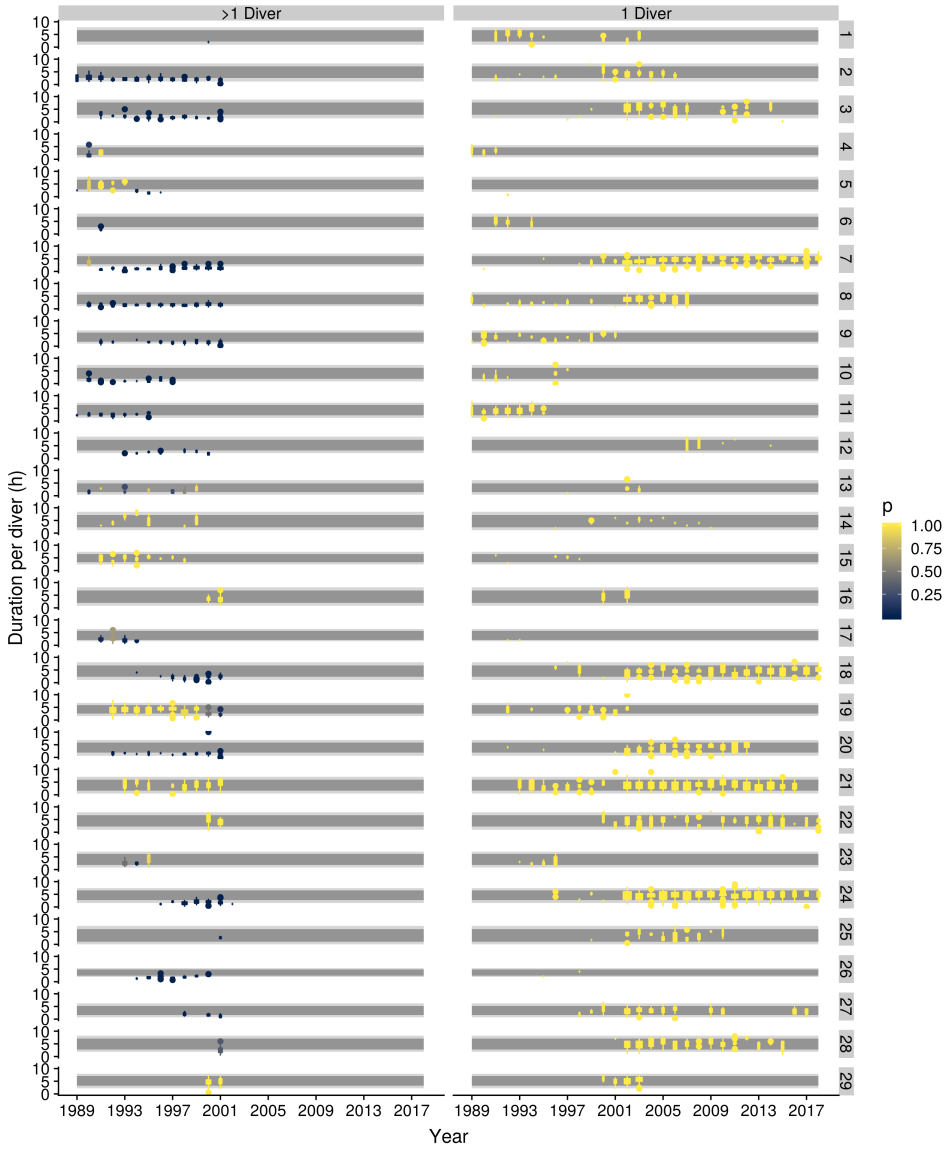

Figure 10: Classification of pāua fishing duration records into reporting by combined time in water ($p \approx 1$; yellow) and by diver for events with more than one diver ($p \approx 0$; blue). Each line corresponds to one crew (client identification), with the estimated distribution of fishing duration per diver in grey (dark grey: 80% confidence; light grey: 95% confidence). For events with more than one diver (left column), the reported fishing duration was divided by the crew number (boxplot), so that for records where reporting was by crew, the boxplot should be within the distribution indicated by the grey ribbon.

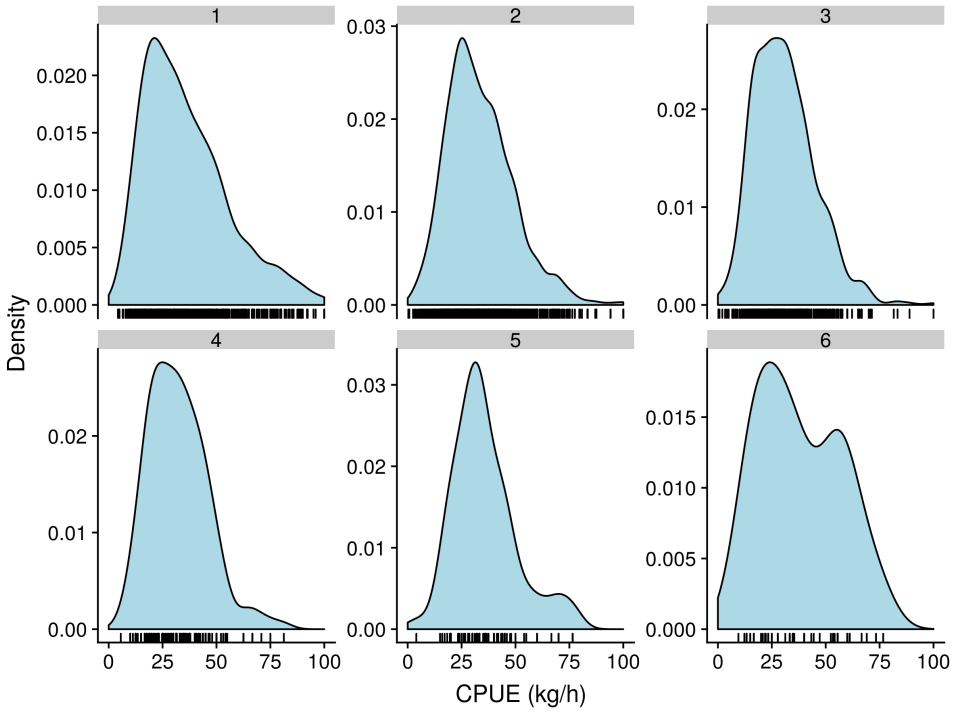

Figure 11: Densities of corrected catch-per-unit-effort (CPUE) by diver for records with one to six divers in the crew.

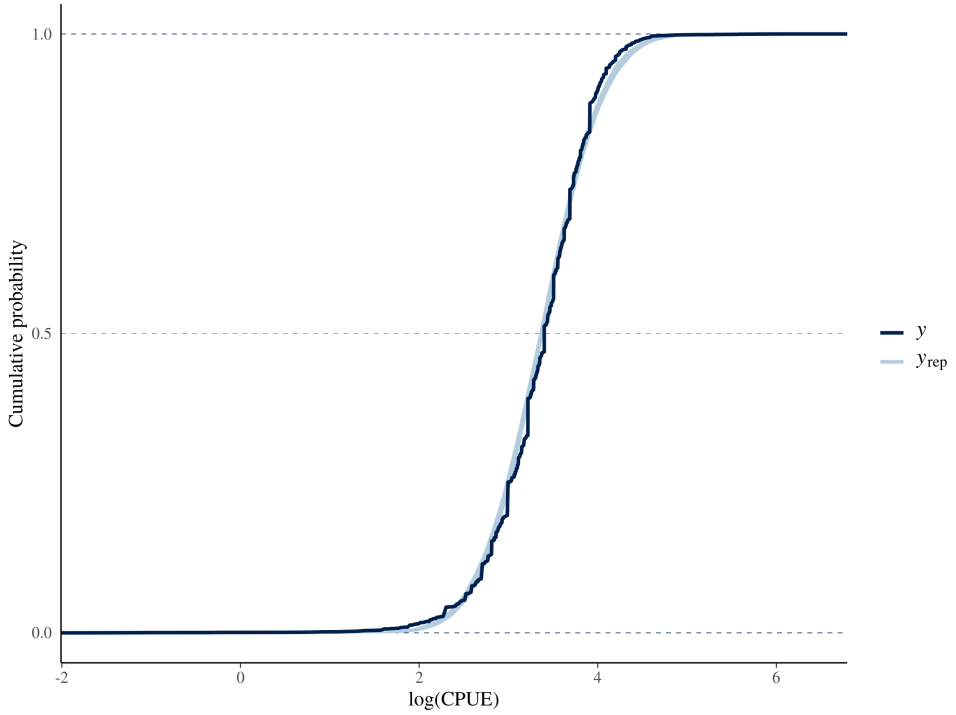

Figure 12: Fit of the log-normal generalised linear mixed model used for catch-per-unit-effort index standardisation. Shown is the cumulative distribution from posterior predictive draws from the model (i.e., predicting each data point; blue) compared with the empirical cumulative distribution (black line).

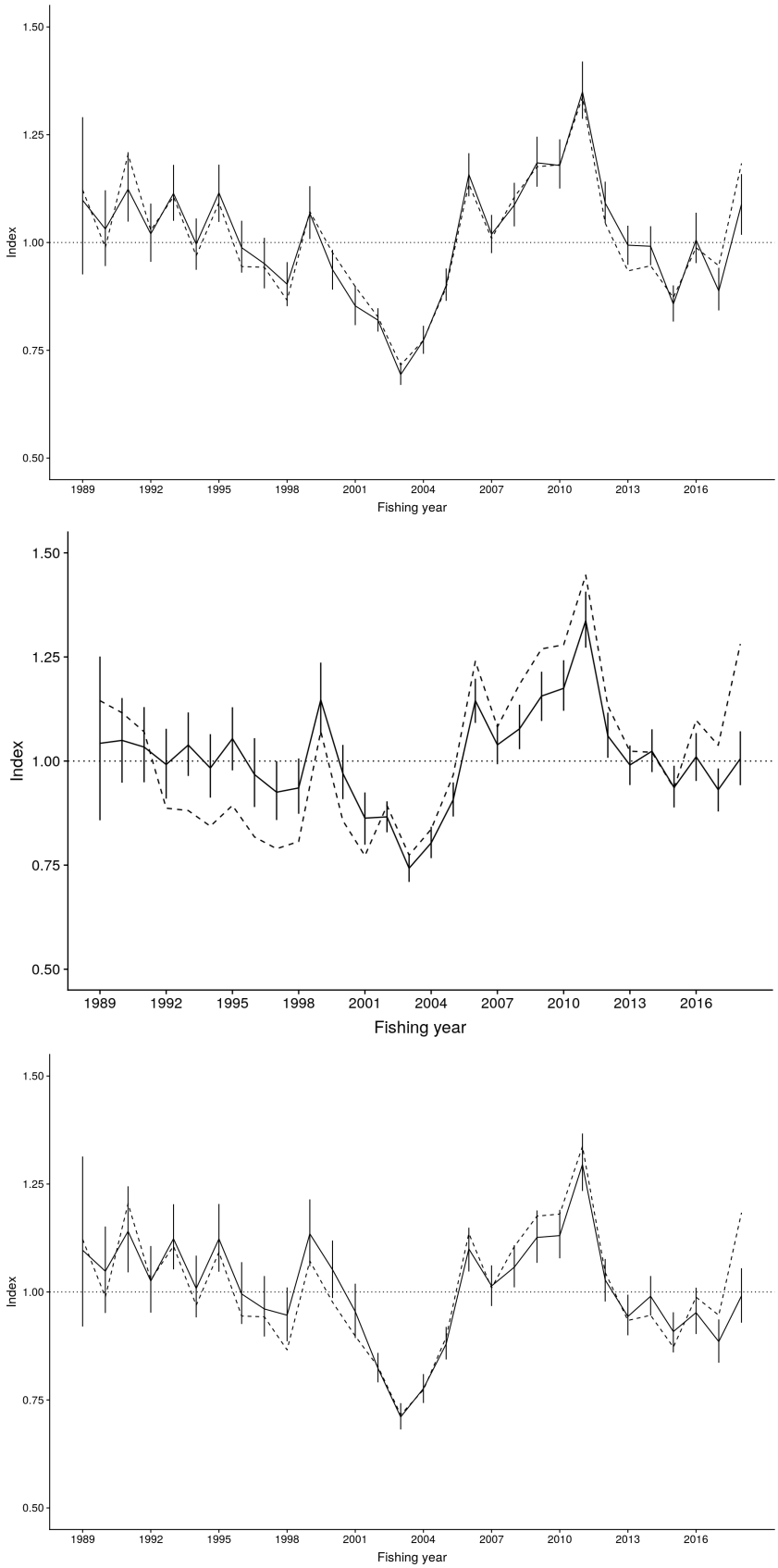

Figure 13: Standardisation of catch-per-unit-effort (CPUE) data using the generalised linear mixed model used for CPUE index standardisation. Shown are: (top) model applied to data with only the catch effort landing return (CELR) level variables, and truncation to include only data with fishing duration $<$7 h; (middle) above model with pāua CELR (PCELR) variables added; (bottom) full model with CELR and PCELR variables, applied to data corrected by the classification procedure. Un-standardised geometric mean CPUE included for reference (dashed line).

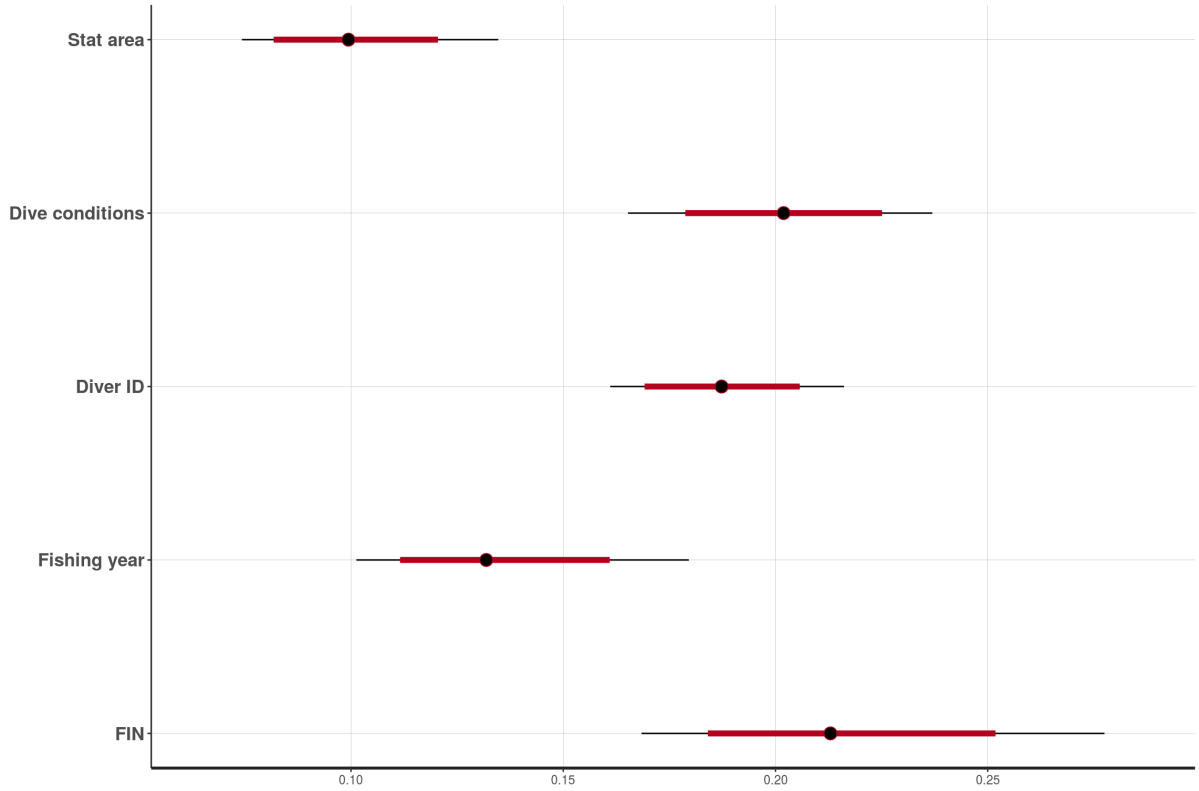

Figure 14: Effect size as variance explained for variables included in the random effects standardisation model.

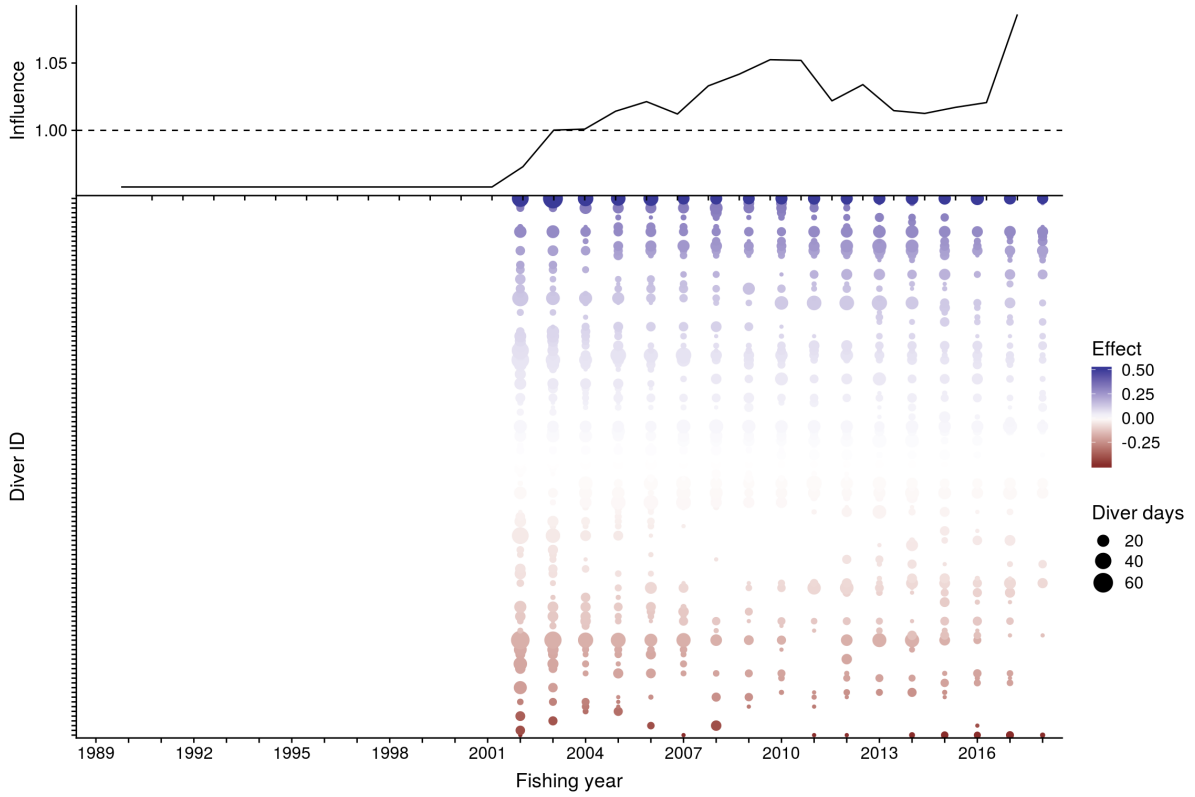

Figure 15: Influence plots for diver random effects, showing the effect on raw catch-per-unit-effort (CPUE; top) of applying estimated effect sizes (bottom) from the factor levels encountered each fishing year. A positive (>1) influence for a given time period means that the unstandardized index was increased due to the relative distribution of the covariate.

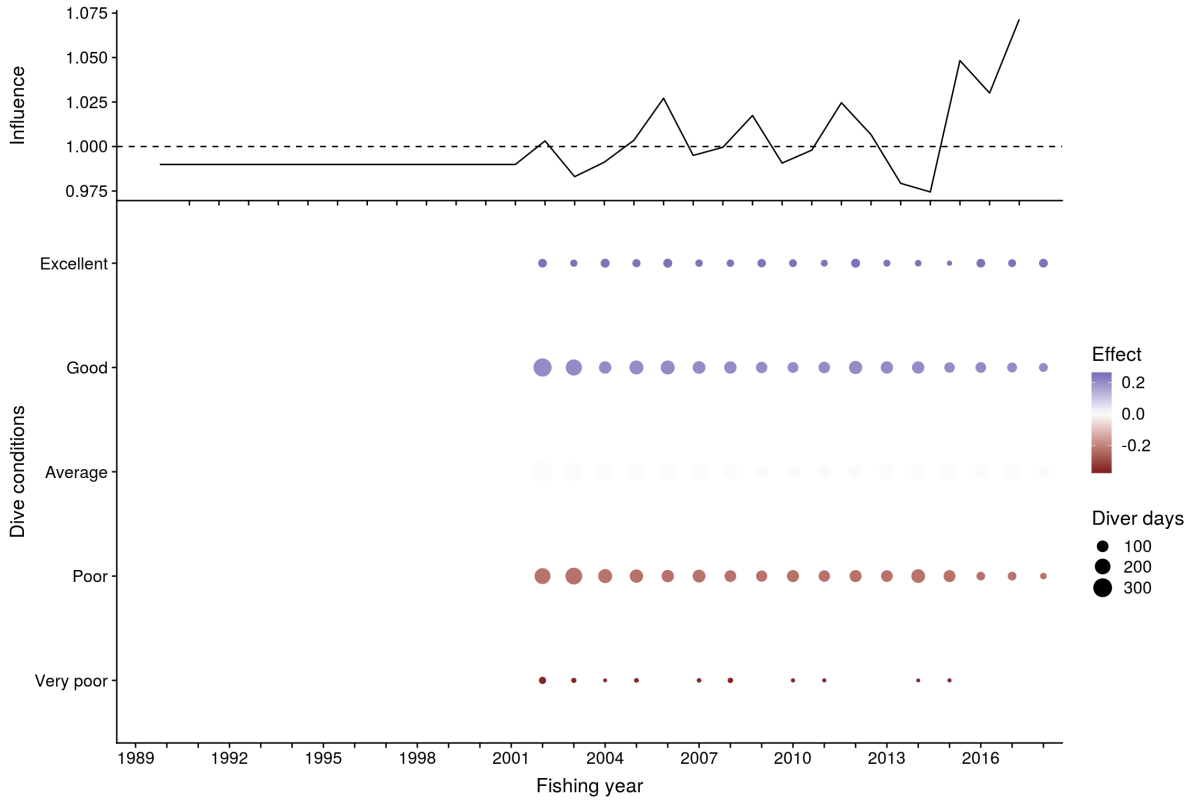

Figure 16: Influence plots for dive condition, showing the effect on raw catch-per-unit-effort (CPUE; top) of applying estimated effect sizes (bottom) from the factor levels encountered each fishing year. A positive (>1) influence for a given time period means that the unstandardized index was increased due to the relative distribution of the covariate.

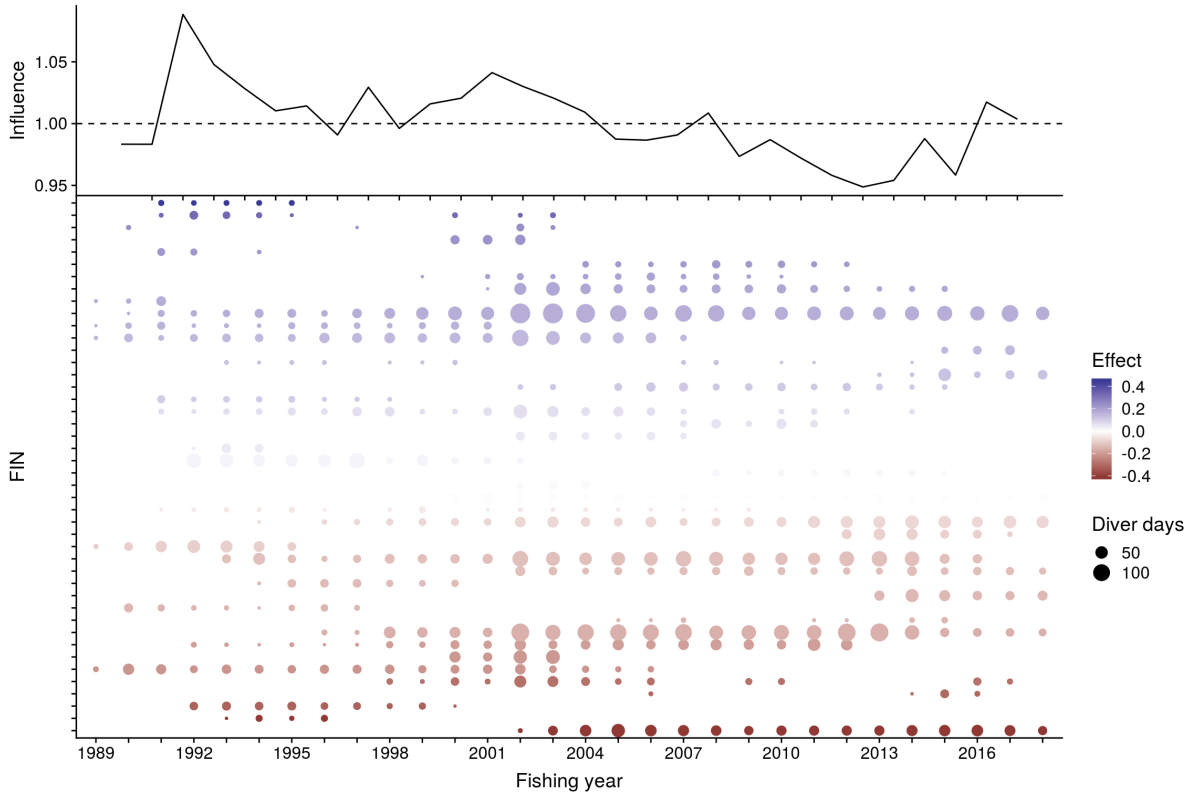

Figure 17: Influence plots for client number (FIN) effect, showing the effect on raw catch-per-unit-effort (CPUE; top) of applying estimated effect sizes (bottom) from the factor levels encountered each fishing year. A positive (>1) influence for a given time period means that the unstandardized index was increased due to the relative distribution of the covariate.

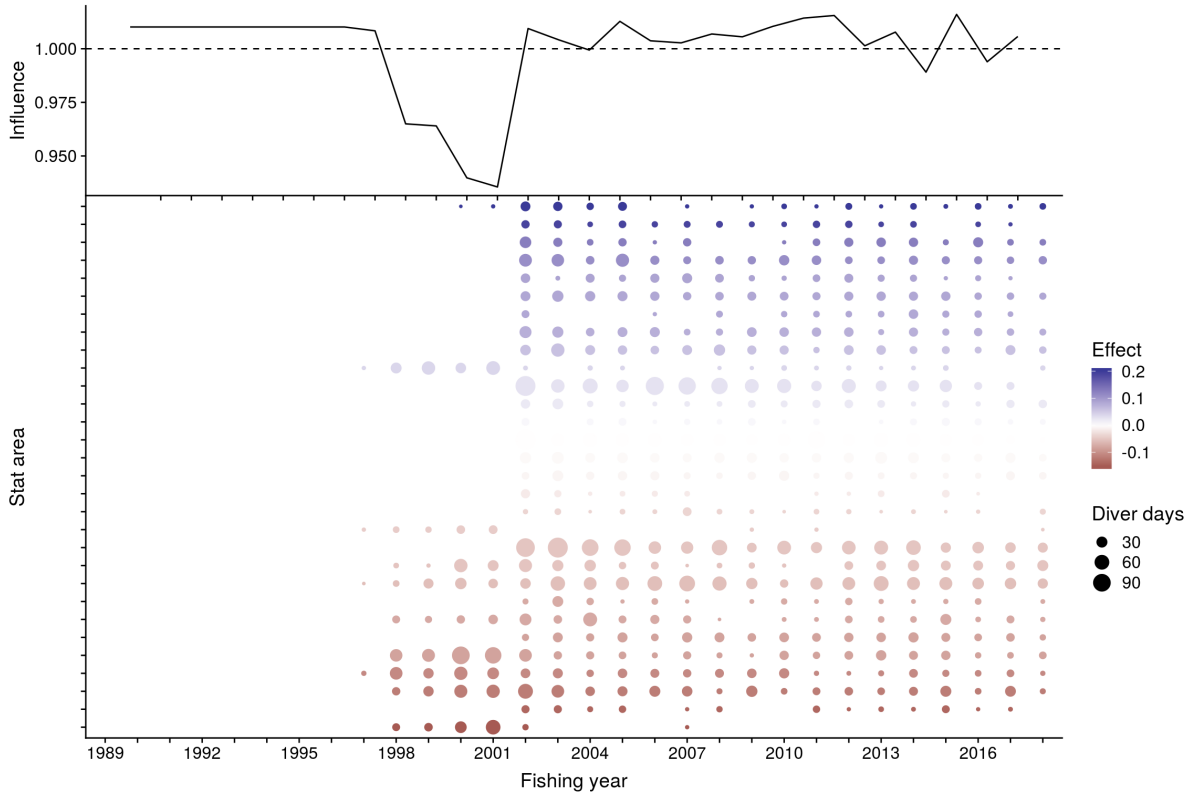

Figure 18: Influence plots for statistical (stat) area effect, showing the effect on raw catch-per-unit-effort (CPUE; top) of applying estimated effect sizes (bottom) from the factor levels encountered each fishing year. A positive (>1) influence for a given time period means that the unstandardized index was increased due to the relative distribution of the covariate.

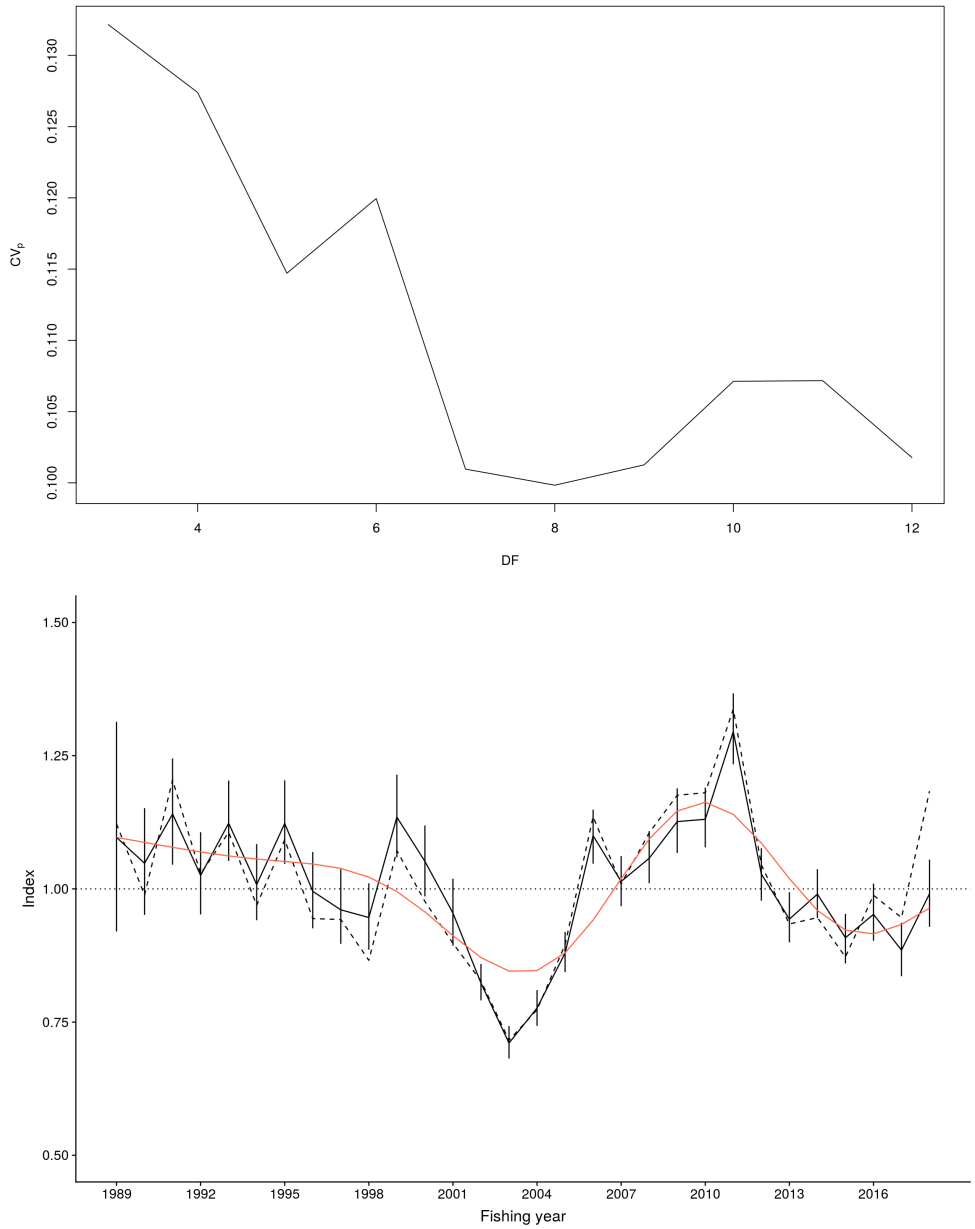

Figure 19: Estimation of process error by fitting a smoother and calculating a degree-of-freedom-corrected coefficient of variation (CV). Top: Degree-of-freedom-corrected CV as a function of the degrees-of-freedom of the smoother applied to the catch-per-unit-effort (CPUE) index. Bottom: Smoother with eight degrees of freedom (red) fitted through the CPUE index (black line with observation error bars). Un-standardised geometric mean CPUE included for reference (dashed line).

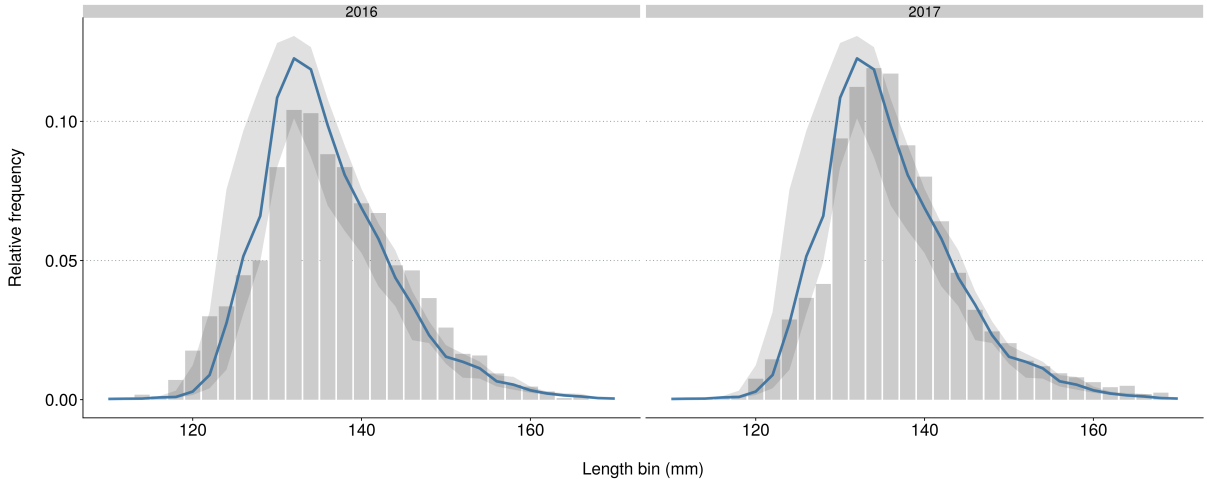

Figure 20: Comparison of length-frequency distributions for the two additional years of samples since the previous assessment in 2016 (grey bars, relative proportions, unscaled), compared with the distribution for the previous length-frequency samples. Median proportion by length class for fishing years 2002 to 2015 in blue, 25th to 75th quantiles indicated by the grey band.

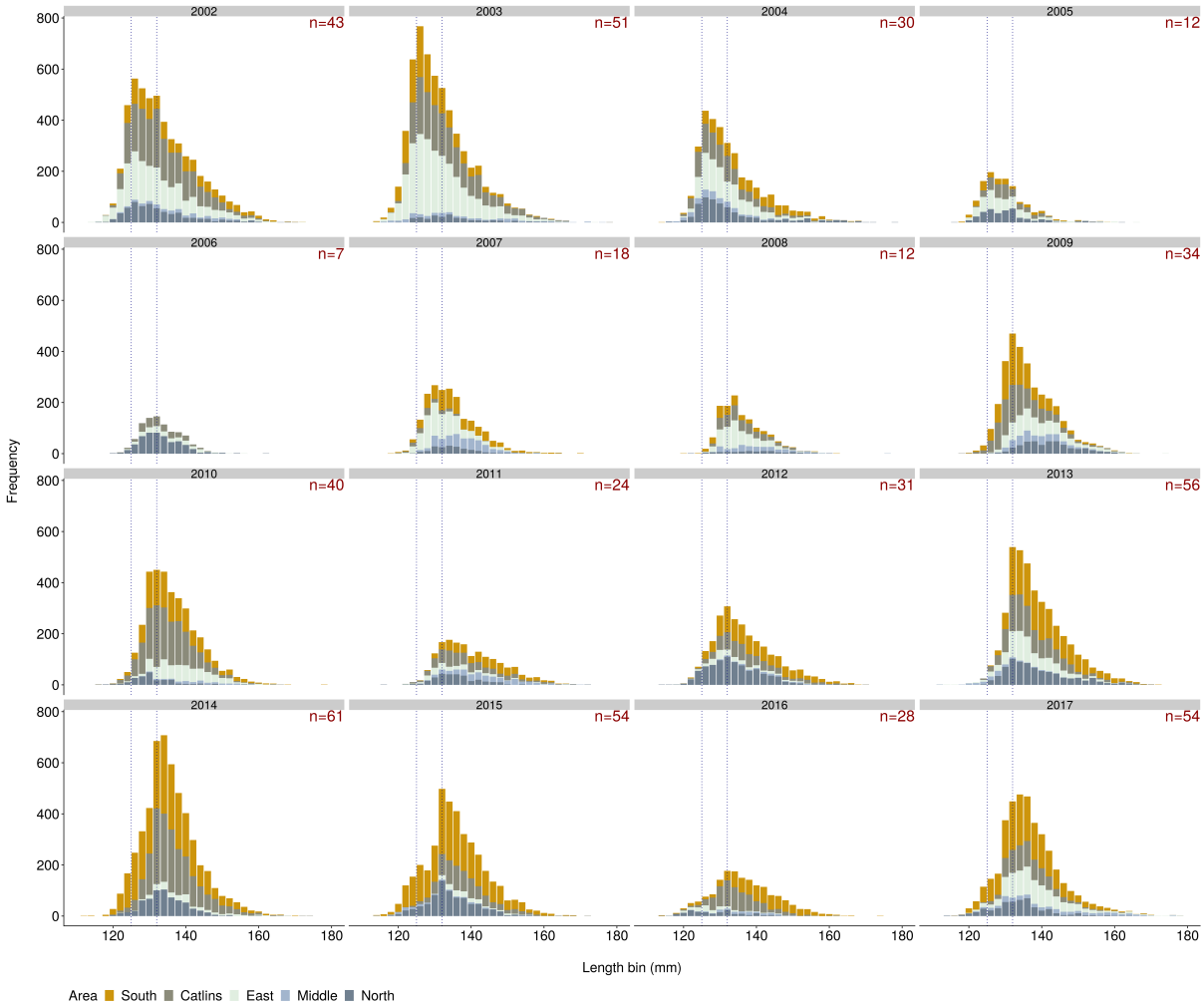

Figure 21: Unscaled pāua length-frequency distributions by fishing year from 2002 onwards, with the collection area highlighted. Number of landings sampled for each year indicated in red; dotted lines indicate the minimum legal size at 125 mm shell length and the more recent minimum harvest size at 132 mm shell length.

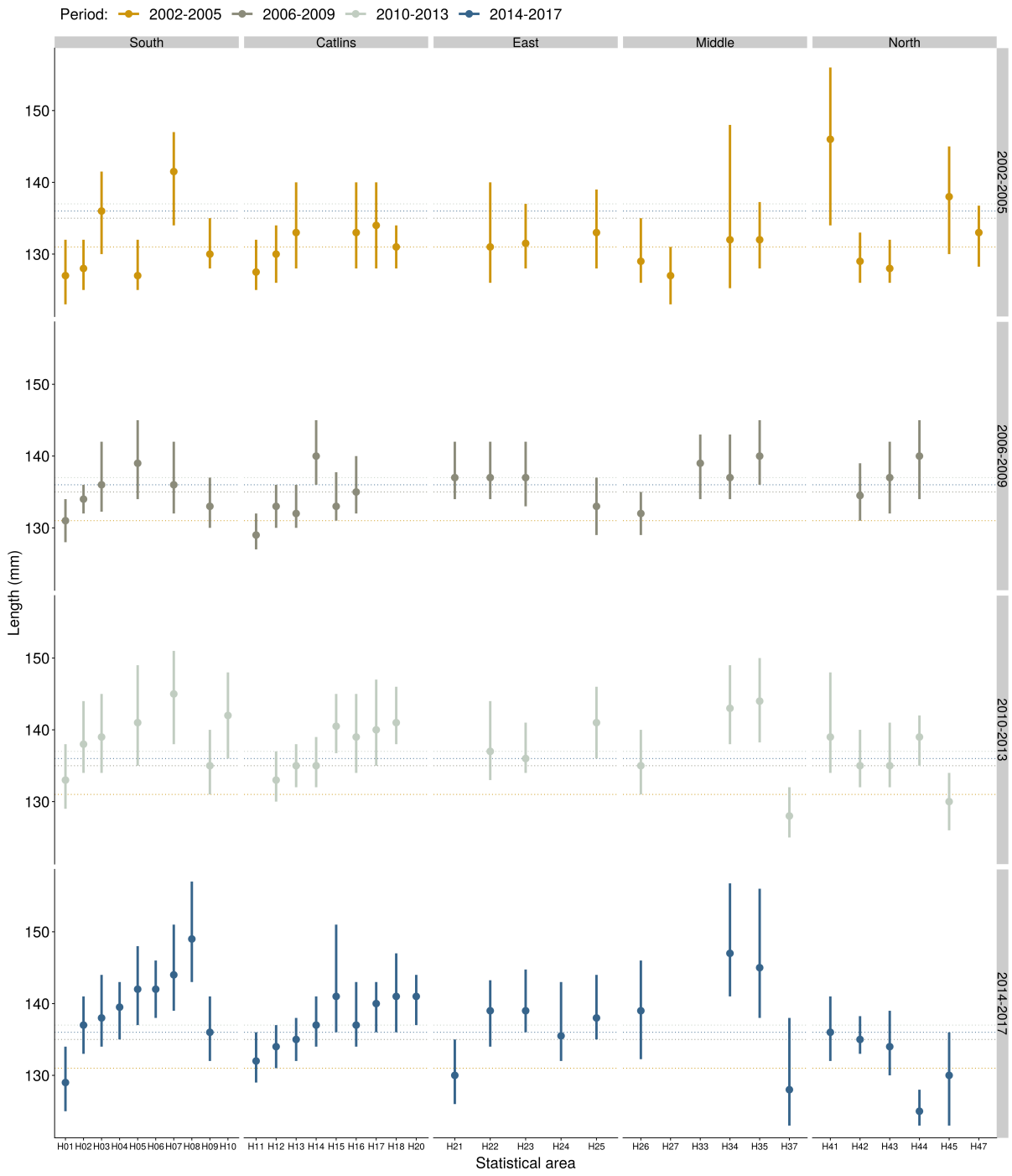

Figure 22: Median and 25th to 75th quantile bounds for pāua length samples by four-year period and statistical area. Summary statistics were calculated across all length samples within a four-year period in each statistical area, with dotted lines showing the median length over all statistical areas in each time period.

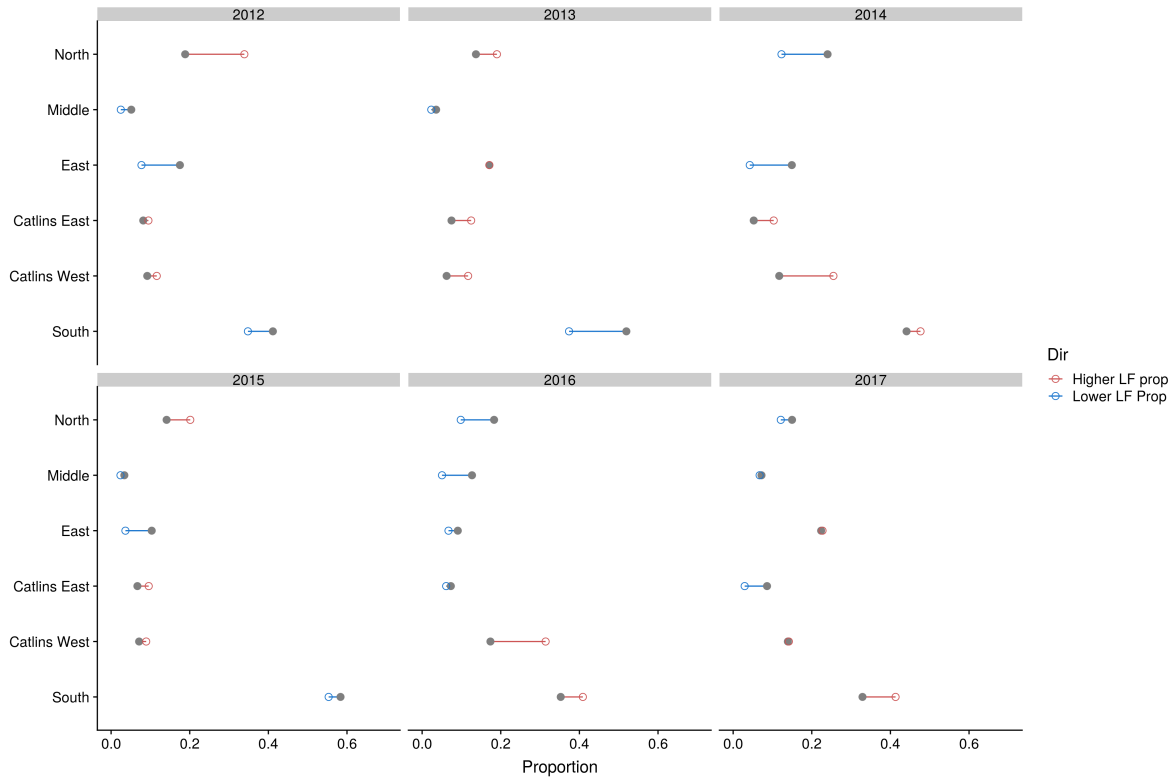

Figure 23: Relative distribution of pāua catch versus length-frequency (LF) samples by region for the six years of catch sampling length frequency data between 2012 and 2017. Proportion of total catch in each region for the year in grey, proportion of length samples in blue if it was less than the catch proportion for that year and in red if it was greater.

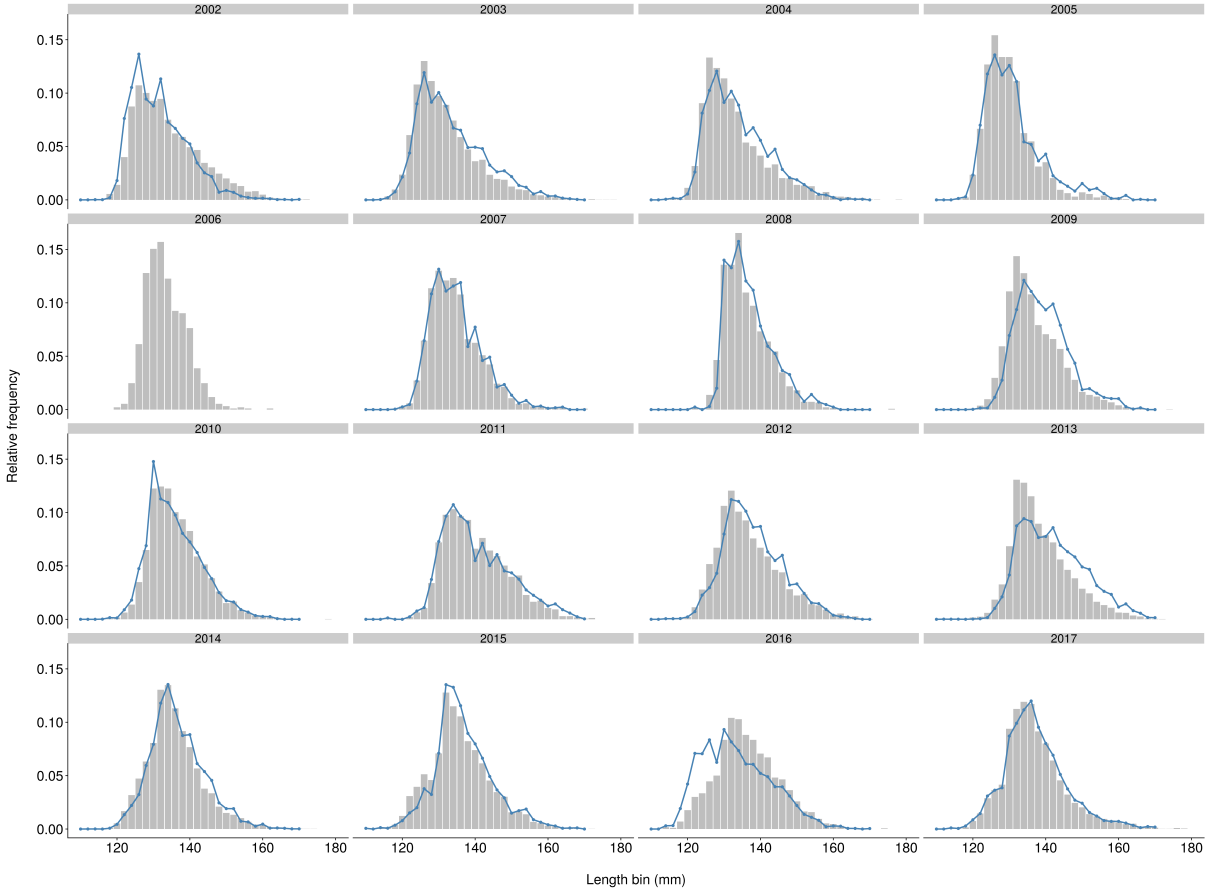

Figure 24: Unscaled relative pāua length-frequency distributions by fishing year from 2002 onwards, with the re-weighted distribution using the catch-at-age approach in blue.

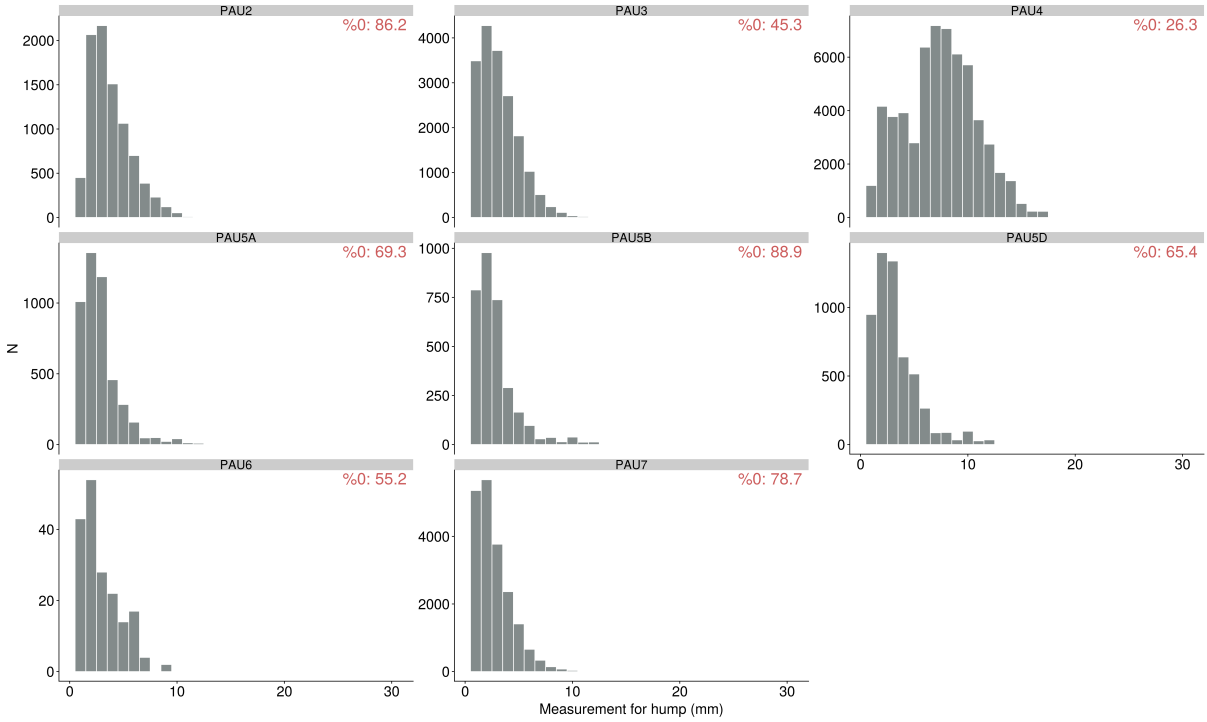

Figure 25: Prevalence (in red) and frequency distribution of hump (spire) size when positive by pāua Quota Management Area. Data were collected between 2006 and 2017.

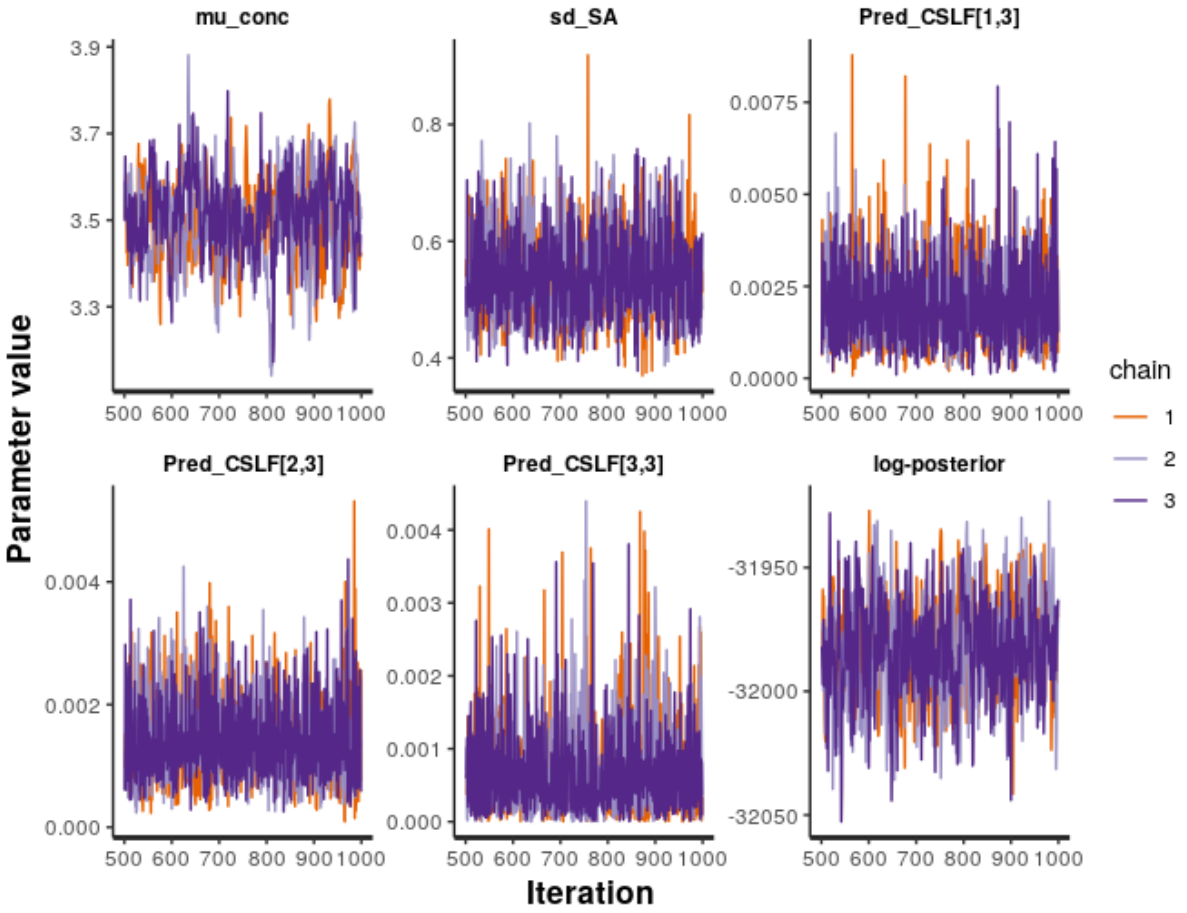

Figure 26: Markov Chain Monte Carlo traces for parameters in the Dirichlet-Multinomial DM length composition standardisation model. The parameter sd_SA in the trace plot shows the standard deviation (σ^a) of the statistical area random effect.

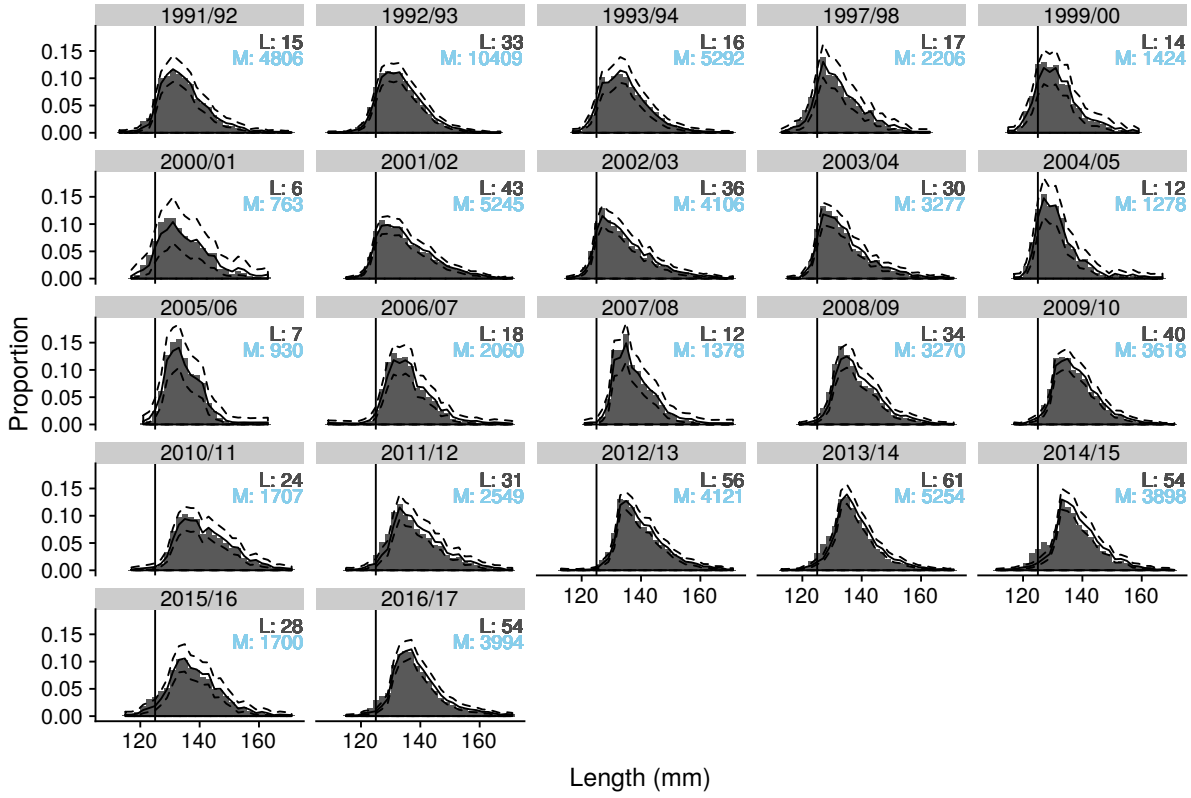

Figure 27: Dirichlet-Multinomial posterior distributions for yearly proportions π_y (black line) with 95% confidence intervals (dashed line). Raw catch sampling length frequency CSLF proportions are in grey; number of landings (L) in black; number of measurements (M) in blue.

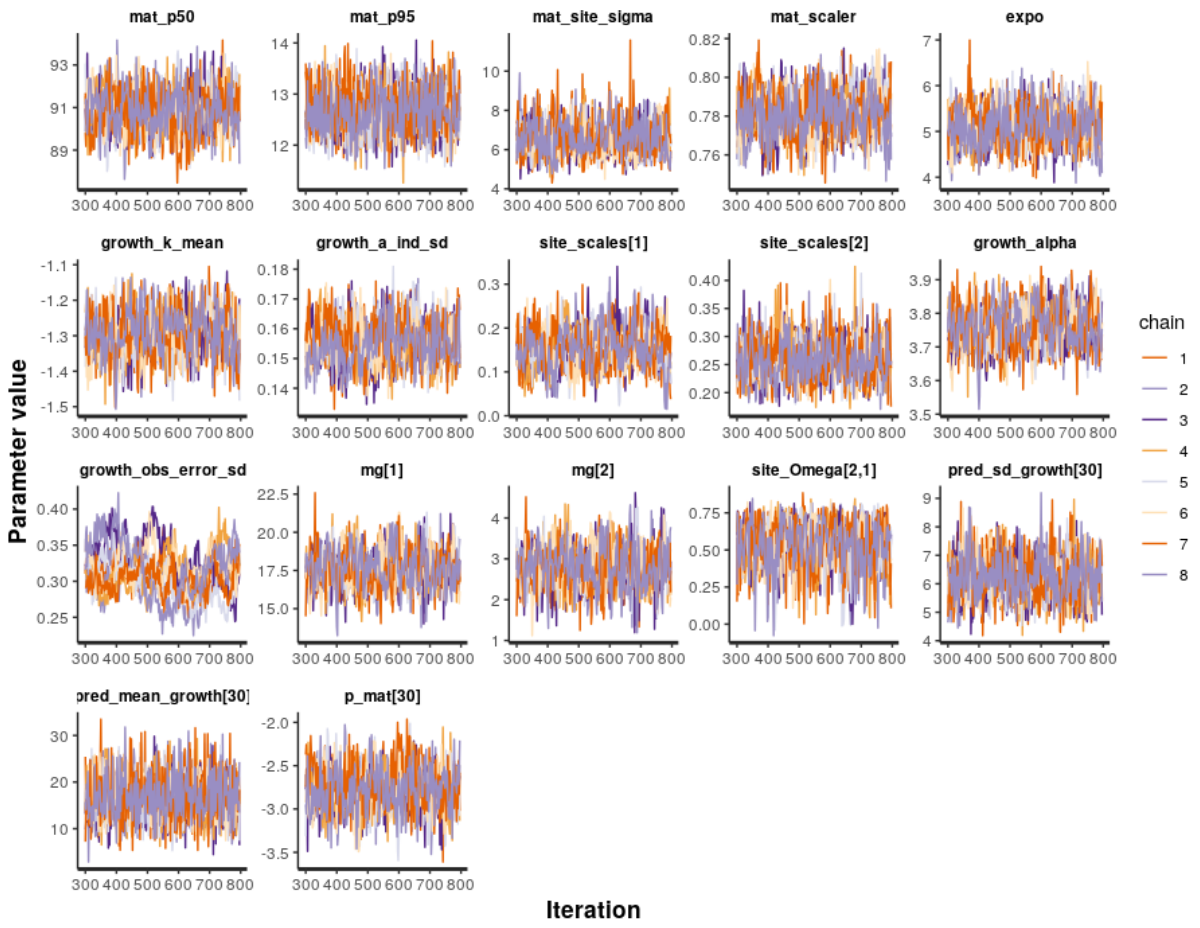

Figure 28: Markov Chain Monte Carlo traces for parameters in the von Bertalanffy growth model.

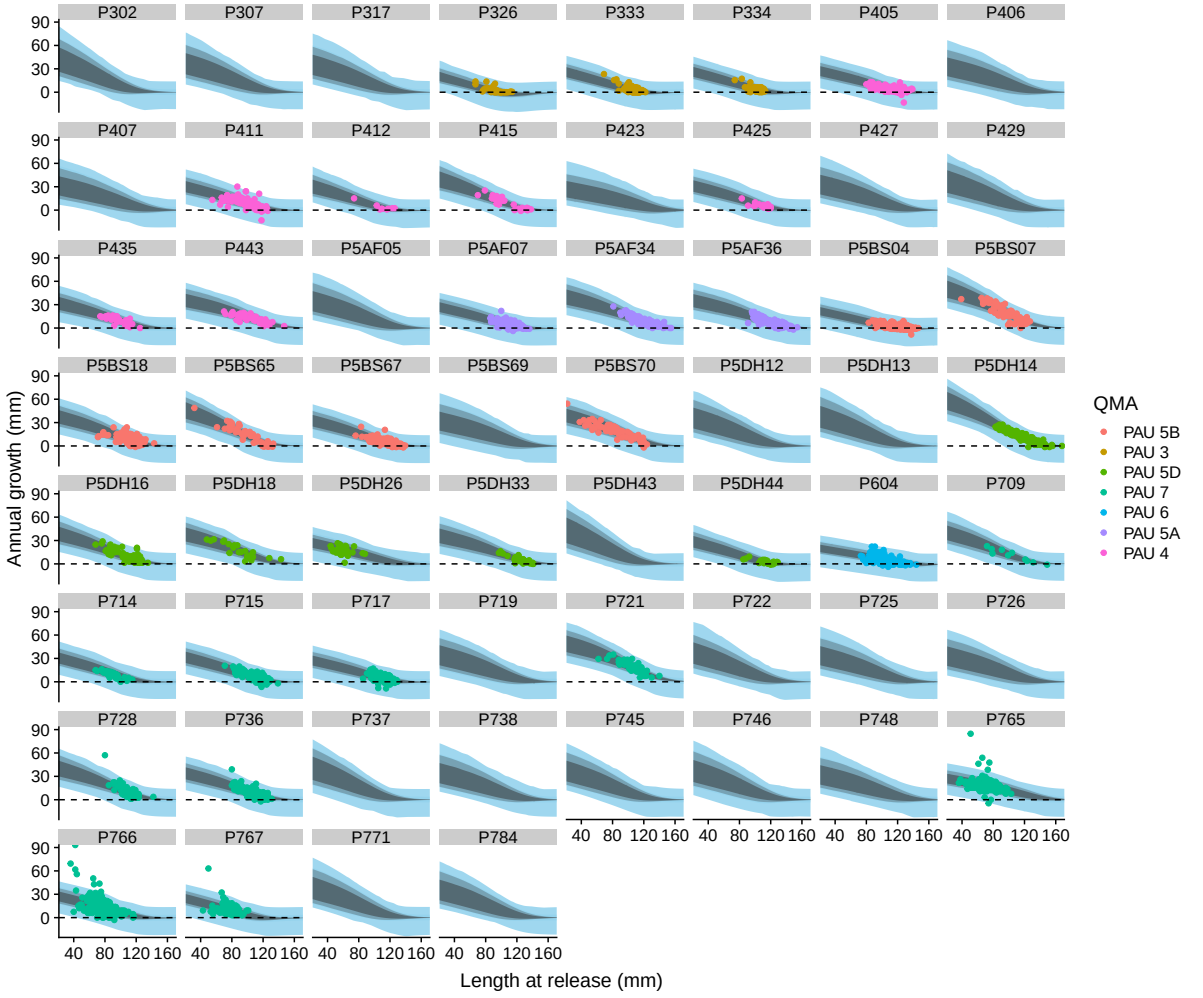

Figure 29: Confidence regions (dark shade: 80%; light shade: 95%; blue shade: 95% measurement error) for predicted true increment $x_{i,s}^{\cdot}$ at length l across all statistical areas in the pāua tag-recapture dataset (QMA, quota management area).

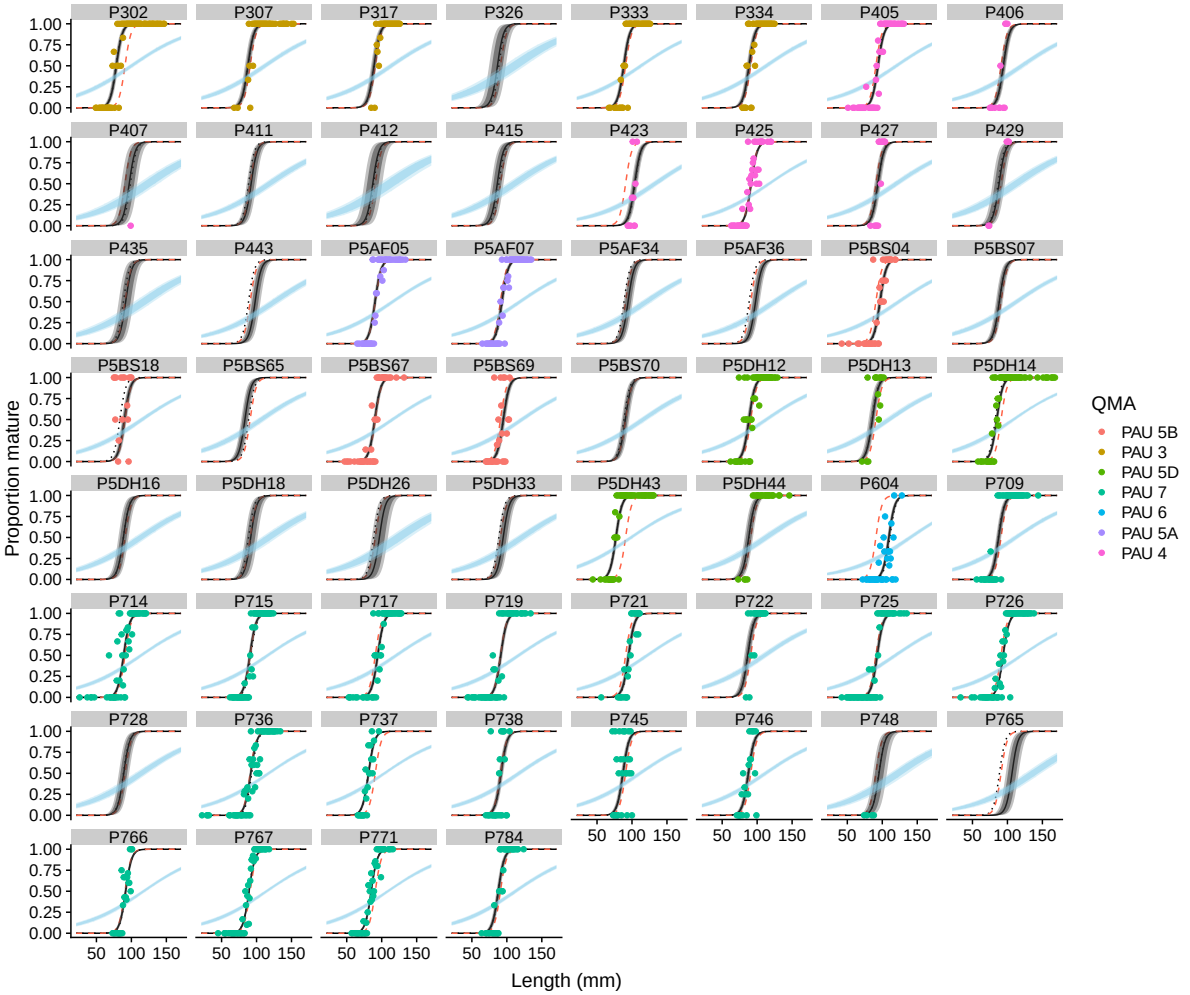

Figure 30: Confidence regions (dark shade: 80%; light shade: 95%) for predicted proportion labelled mature $p(l)$, with the corresponding proportion allocated to reproduction in light blue for all statistical areas in the pāua tag-recapture dataset (QMA, quota management area). Dark red dashed line indicates the population mean proportion mature at length, dark dotted line shows a generalised linear mixed model prediction for logistic length-at-maturity independent of growth.

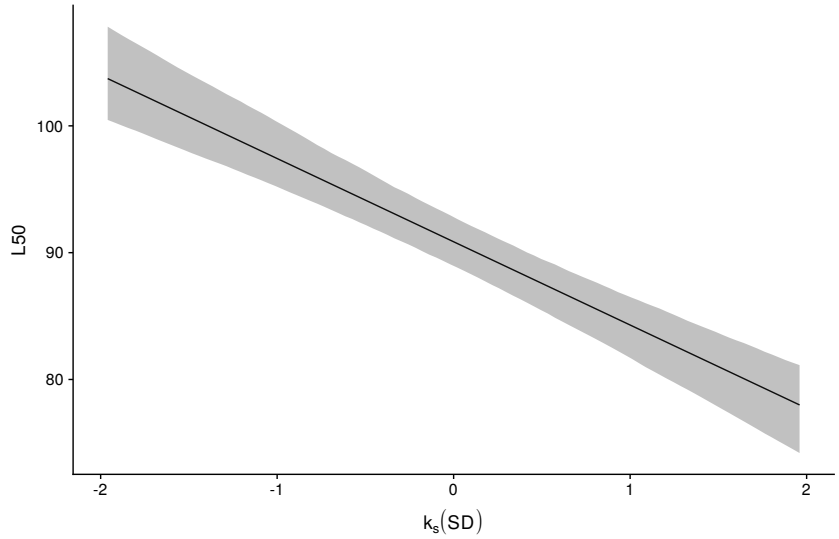

Figure 31: Predicted relationship between length-at-maturity parameter $L50$ and the deviation from the mean growth parameter k (SD, standard deviation). Confidence regions are shown in dark shade for the 95% confidence interval.

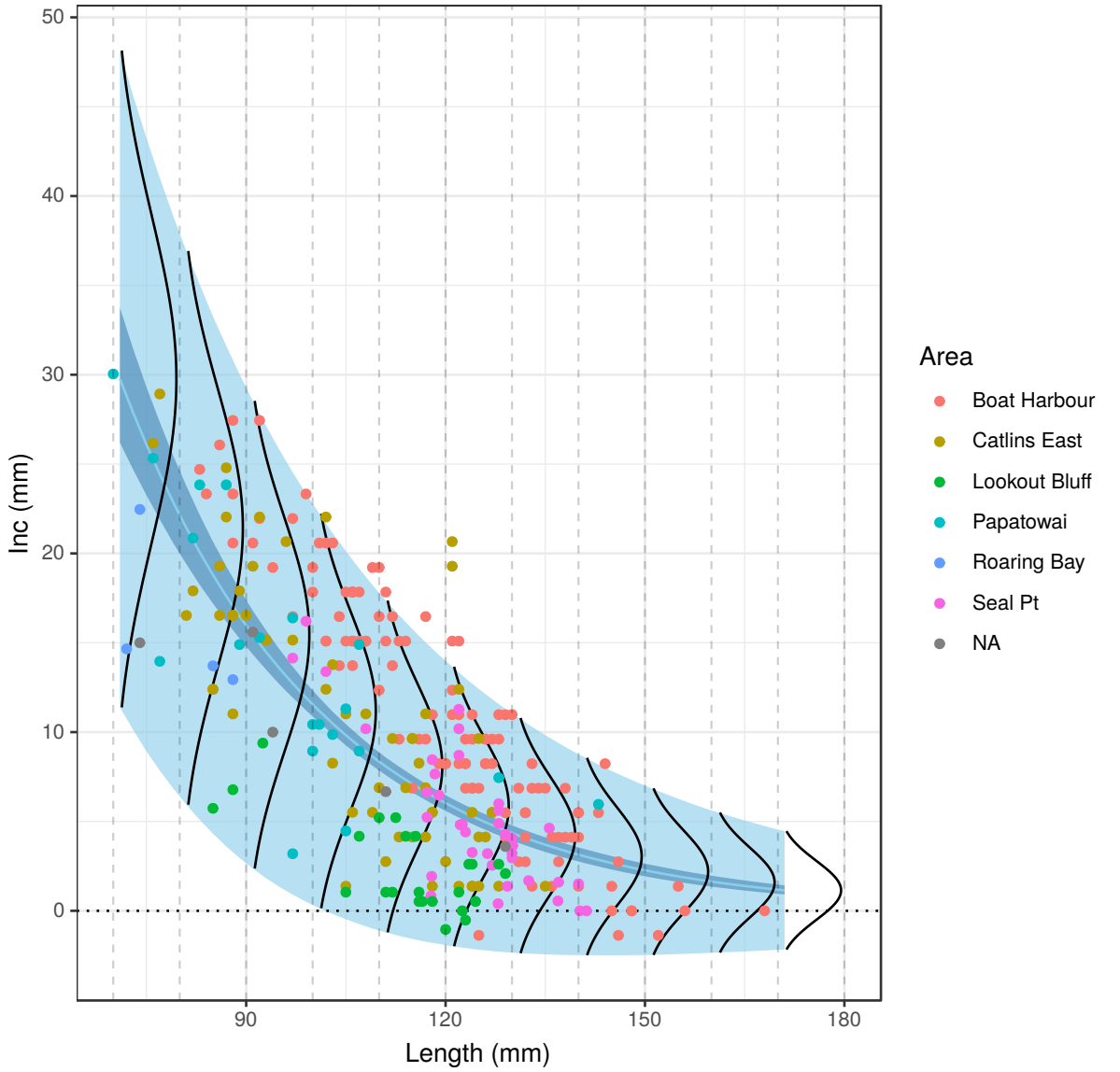

Figure 32: Prior implied by PAU 5D tag-recapture data for growth within the 2018 stock assessment of pāua. Dark blue area shows uncertainty about mean growth, light blue line the posterior median for mean growth; light blue area shows the posterior median for the population standard deviation applied to the mean growth; black line shows the implied distribution of growth at the median of the prior. Points show tag-recapture data from different statistical areas in PAU 5D (NA, no spatial information available).

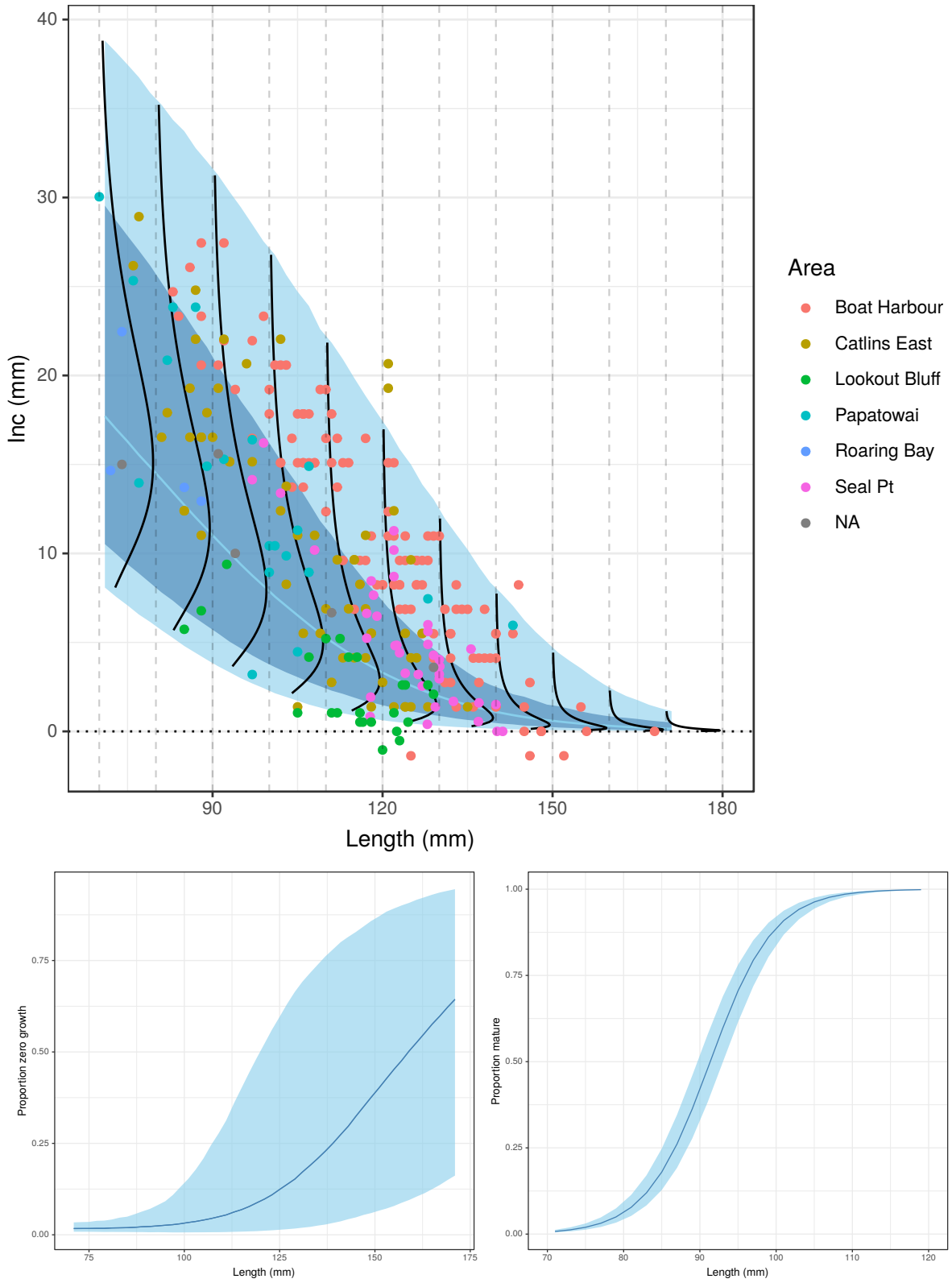

Figure 33: Priors implied from a meta-analysis of growth of South Island pāua, based on model-fitting to all tag-increment data from different areas (NA, no available area data). Shown is the joint prior for positive growth increments (top), proportion of local populations not growing at a particular 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.

6. ACKNOWLEDGMENTS

Many thanks to Tom McCowan for helpful feedback, Storm Stanley, Jeremy Cooper, Marine Pomarede and the FNZ Shellfish Working Group for discussions over the course of this project.

Funding for this research was provided by Fisheries New Zealand project PAU2018-01.

7. REFERENCES

Bull, B.; Dunn, A. (2002). Catch-at-age: User manual v1.06.2002/09/12. NIWA Technical Report 114.

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.

Helidoniotis, F.; Haddon, M.; Tuck, G.; Tarbath, D. (2011). The relative suitability of the von Bertalanffy, Gompertz and inverse logistic models for describing growth in blacklip abalone populations (*Haliotis rubra*) in Tasmania, Australia. *Fisheries Research 112 (1)*: 13–21.

Kozlowski, J.; Czarnoleski, M.; Dańko, M. (2004). Can optimal resource allocation models explain why ectotherms grow larger in cold? *Integrative and Comparative Biology 44 (6)*: 480–493.

Lester, N.P.; Shuter, B.J.; Abrams, P.A. (2004). Interpreting the von Bertalanffy model of somatic growth in fishes: the cost of reproduction. *Proceedings of the Royal Society of London B: Biological Sciences 271 (1548)*: 1625–1631. doi:10.1098/rspb.2004.2778.

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.

Minte-Vera, C.V.; Maunder, M.N.; Casselman, J.M.; Campana, S.E. (2016). Growth functions that incorporate the cost of reproduction. *Fisheries Research 80*: 31–54. doi:10.1016/j.fishres.2015.10.023.

Neubauer, P.; Tremblay-Boyer, L. (2019). The 2018 stock assessment of pāua (*Haliotis iris*) for PAU 5D. *Draft New Zealand Fisheries Assessment Report No. XX 2019/xx*.

Ohnishi, S.; Yamakawa, T.; Okamura, H.; Akamine, T. (2012). A note on the von Bertalanffy growth function concerning the allocation of surplus energy to reproduction. *Fishery Bulletin 110 (2)*: 223–229.

Quince, C.; Abrams, P.A.; Shuter, B.J.; Lester, N.P. (2008). Biphasic growth in fish I: theoretical foundations. *Journal of Theoretical Biology 254 (2)*: 197–206.

Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.17.3. Retrieved from http://mc-stan.org/.

Stewart, I.J.; Hamel, O.S. (2014). Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences 71 (4)*: 581–588. doi:10.1139/cjfas-2013-0289.

Thorson, J.T. (2014). Standardizing compositional data for stock assessment. *ICES Journal of Marine Science 71 (5)*: 1117–1128. doi:10.1093/icesjms/fst224.