Response to 'Review of the 2024 Stock Assessment for King Threadfin (*Polydactylus macrochir*) in the Gulf of Carpentaria, Queensland'

June 2024

Queensland Government

This publication has been compiled by A.B. Campbell, M. Tanimoto, O.J. Whybird and A.R. Fox of Fisheries Queensland, Department of Agriculture and Fisheries

Painting on cover page by Alise Fox.

Enquiries and feedback regarding this document can be made as follows:

Email:	info@daf.qld.gov.au
Telephone:	13 25 23 (Queensland callers only)
	(07) 3404 6999 (outside Queensland)
	Monday, Tuesday, Wednesday and Friday: 8 am to 5 pm, Thursday: 9 am to 5 pm
Post:	Department of Agriculture and Fisheries GPO Box 46 BRISBANE QLD 4001 AUSTRALIA
Website:	daf.qld.gov.au

Contents

1 Preliminary remarks

We thank the reviewers for their thorough examination of the technical content of the assessment and for the interactive approach they took throughout the review process. We would especially like to acknowledge the large quantity of background material they had to quickly absorb, and their engagement with our fisher stakeholder project team representatives. The combination of these factors has served to significantly strengthen the final published version of the assessment.

The additional analyses conducted during and after the review, either in direct response to reviewer suggestions, or as lines of enquiry initiated by the review discussions, have served to reinforce from our perspective the reviewer's principal finding: "The estimated biomass trend and stock status appear generally consistent with the catch history and the current age structure, but stock status is likely considerably more uncertain than the current estimate."

The 'current estimate' referred to is the estimate (with uncertainty intervals) provided in the final submission to review. The final report now incorporates additional work stemming from the review which has widened these intervals. However, the updated diagnostics suggest that uncertainty is still underestimated.

During the review process additional modelling work was done, and not all the outputs from this work are included in the updated final report. Some of them were included in an unpublished 'Companion' document, provided to the reviewers. In this response we reference a 'Karumba-only' model which was suggested by the reviewers and for which more than ten scenarios were explored. While this exploration was instructive the models were ultimately considered too preliminary to include in final reporting. The reviewer's report contains images of diagnostic plots provided to the reviewers during this exploration phase.

We start with the review recommendations and a brief comment on changes that have been made to address them. Subsequent sections include more detail. We invited the reviewers to follow up with feedback on the updated report and a draft of this response, and this 'response to response' is included here as the final section.

Most references in this document are references to figures and sections in Campbell et al. (2024).

2 Review recommendations

The reviewers' recommendations (Hoyle and Dunn 2024, Section 8) are reproduced in full here, with our respones *in italics* below each recommendation.

1. Expand data characterisation analyses as much as possible, including but not limited to the following.

   (a) Investigate monitoring data for spatial and temporal patterns in age and size structure.

   Done and added to the KTF Stock assessment 'Companion' document.

   (b) Show representativeness of sampling through time with respect both to the catch and to locations.

   Done and added to the KTF Stock assessment 'Companion' document.

2. Revise CPUE series.

   (a) Use an error distribution model that will allow full diagnostics.

   Done.

   (b) Provide enough diagnostics to determine whether the model is appropriate for the data.

   Done.

   (c) Explore alternative targeting analyses based on cluster analysis.

   Will consider for next assessment.

   (d) Conduct separate analyses by region.

   Will consider for next assessment, noting the FRDC project may provide relevant information. Exploratory Karumba-only model was useful in building understanding.

   (e) Apply models that allow for spatial and spatiotemporal variation in catch rates within regions.

   Done.

   (f) Consider using a revised river flow dataset.

   River flow removed.

   (g) Consider running models with the R package mgcv.
   Done.

3. Data preparation

   (a) Allow for spatial structure in size and age comps when developing composition datasets, since the model assumes that catch is sampled uniformly from the whole population.

   This has been explored.

4. Biological inputs

   (a) Use a range of steepness values from 0.55 to 0.95, and sigmaR of 0.6.

   The mid-point of this range (0.75) was adopted for the mid-point of the final ensemble. Computational/time constraints prevented exploring more than 0.7 and 0.8 in the final models.

   (b) Consider Lorenzen M.

   Left to for consideration next assessment.

5. Modelling

   (a) Apply a standard Francis data weighting method.

   Done.

   (b) Run a separate assessment model for each region.

   Done for exploratory Karumba and whole-Gulf, additional regionalisation for future work.

   (c) Identify hypotheses to explain why the model cannot fit the ages-at-length in the early period, and develop alternative models based on these

   With the Francis weighting and other changes this discrepancy has improved, however future work needs to examine this further.

   (d) Ensure that issues with incomplete convergence, no Hessian, and poor MCMC diagnostics are resolved

   MLE convergence is good. MCMC switched to parallel NUTS, Rhat statistics generated.

   (e) Allocate composition data to the region where it was sampled

   Done - singled fleet model now.

   (f) Consider non-asymptotic selectivity

   Explored, but not included in final ensemble.

   (g) Include month of sampling in the composition data

Not done, see comment in section 4.4.

3 Review of spatial aspects of the stock assessment

3.1 Spatial structure of the king threadfin stock

The reviewers' literature study on this, critique of the views expressed in the draft report and canvassing of the views of project team fisher stakeholders on this issue has been very valuable. We note the remark by Cadrin et al. (2023) that 'A summary of stock identification and how well it matches the current assessment or management unit should be updated in every stock assessment report (e.g., an updated summary of information should be a generic term of reference for all assessments)" and have updated our standard stock assessment report template to explictly include this. Section 2.2 of Campbell et al. (2024) has been added to focus on what is known about the spatial structure of the stock. It now references this contribution by the reviewers and updates the language (originally in the introduction) to more appropriately caveat our current level of understanding.

3.2 Assumptions of stock structure

It may be helpful to distinguish two broad levels at which this can be examined:

1. the appropriateness of the total spatial scale of the population model that underpins the assessment with regard to likely level of reproductive connectivity (or 'demographic dependence on ecologically relevant timescales' (Haugen et al. 2022)) at that spatial scale.
2. the handling of spatial complexity in the case that below this scale complexity remains (e.g. via explicitly spatial population model structure, 'fleets as areas', spatial data weighting or other some other means).

The reviewers critique focuses primarily on the second level but we will start with a brief comment on the first. This is where project team discussion in meeting 1 (Campbell et al. (2024, Appendix A)) were mainly focussed. The decision to proceed with a total spatial scale that matches the current management unit (whole of Gulf) was difficult particularly for the western 'boundary' and was utimately a practical one in light of our current state of knowledge. This is a major driver for Robins (2024).

In terms of Cadrin et al. (2023)'s ontology we are for now assuming a population or meta-population at the larger scale and then dealing with this situation within it: 'If population structure is too complex to define distinct spatial stocks, stock assessment may require spatial stratification, spatiotemporal analysis, or stock composition analysis to account for heterogeneity.'

In the draft models, this sub-Gulf scale spatial complexity was handled by a kind of 'psuedo fleets-as-areas' approach (with different fleets driven by regional catch rate predictions but mirrored selectivity). This has been replaced with:

1. a simple single fleet population model
2. spatial catch rate predictions that are spatially weighted (in four different ways, more on this below) for input as a single time series to drive the single fleet
3. a re-anlaysis of the full age-length data set (ultimately not spatially re-weighted, more on why below)

We retained the original decision to exclude data for catch rate purposes north of 13 degrees south.

The different spatial weighting approaches led to minor differences in the final time series. A comparison of method 2 (equally weighted intervals) and method 4 (number of 6 min grids fished) can be seen in Figure 3.10. Two reasons are suggested for the small effect. The first is that at smaller spatial scales, as in the exploratory 'Karumba only' analyses conducted during review, it appeared that there was a trend 'consistent with significantly greater depletion of areas with higher fishing pressure (Figure 20)'. However at larger scales, of the 'interval' order constructed for the final analyses you see quite a different pattern emerge (Figure B.11). The Flinders and Norman river basin high catch grids (we prefer to call them high catch rather than high fishing pressure, as fishing pressure assumes the depletion levels we are trying to infer) actually have higher catch rates on average. The Karumba-only pattern noticed by the reviewers can still be discerned in the uptick moving westwards from AD to AE in Figure B.11, but at larger scales the trend if anything is the opposite. This highlights that there is more to be done in terms of properly accounting for spatial residuals (and properly weighting over space) but it also points to a lack of simple spatial trends that might otherwise drive the kind of bias to which the reviewers refer. While the effect plots do not have interaction terms integrated out, we believe the general inference here (that there is no clear inverse relationship between highly fished areas and catch rates) to be robust.

The second is that different approaches to spatial weighting are going to have the greatest impact when spatial imputation is being done to account for spatial expansion or contraction of the fishery (e.g. as highlighted in Walters (2003)). Such large-scale transition appears to pre-date the major catch rate data set and fits with the written accounts (oral histories and Eldorado) of the development of the inshore net fishery in the Gulf of Carpentaria. It is possible a hierarchical statistical analysis of (a) historical Gulf logbook, (b) modern compulsory logbook and (c) modern vessel tracking could be a valuable exercise in capturing some of this phenomenon and thereby contributing more information to the population model.

Regarding length-age data, several exploratory spatial re-weightings were attempted along with a re-analysis of the full length-age data sets. Low sample sizes outside of the Flinders and Norman river basins limited this effort and a satisfactory alternative spatial weighting was not identified. River vs fore-shore weighting was also experimented with but again limited knowledge of the origin of most samples in this respect limited what could be achieved. The final models used this re-analysis which incorporated samples over the the full extent of the area considered by the catch rate analysis and where the initial sampling weight was driven by a unit of sampling akin to a trip/haul (Punt 2023, Table 2).

Overall, we agree with the reviewers that spatial stratification (river-basin dependent, accessibility dependent, river-vs-foreshore dependent etc) and consequent spatial weighting issues are important and have explored several aspects of this. A key resource for future spatial modelling of this kind will be Robins (2024) which we discuss more below.

4 Data inputs

4.1 Data characterisation

Data characterisation is a key part of the team's process however we agree with the reviewers that it could be approached in a more structured way, and better documented. As noted, as part of the review, a 'companion' document was generated contaning a large quantity of additional model diagnostic and data characterisation content. This was continually added to throughout the review, and forms a good basis for characterisation content to be generated for all assessments.

4.2 CPUE

As noted above, the catch per unit effort standardisation was significantly reworked and is discussed in more detail elsewhere. In terms of data characterisation and diagnostics, all outputs requested by the reviewers were generated and provided either in the companion document or as figures in the updated stock assessment report.

4.2.1 Targeting

The 'bespoke' method of filtering for targeting has now been better described (Section 2.4). Cluster analysis can be tried in future analyses as well as incorporting price information. Most powerful may be the use of vessel tracking to distinguish river vs foreshore sets as this is a highly relevant factor and this will also be explored.

4.2.2 Covariates

As discussed above, the catch rate model and in partcular the spatial handling of regions was reworked, predominantly in line with reviewer recommendations. See Figures 3.6 and 3.7 for separate spatial catch rate predictions across 'intervals' (30 min grid projections onto longitude travelling west from Karumba, and latitude travelling north). The intervals are then weighted according to four different metrics:

1. The original spatial prediction approach based on sample size
2. Equal weighting of each interval
3. A rough approximation to habitat area for each interval, constructed by calculating each interval's length of rivers and coastline (based on shapefiles)
4. A different approximation to habitat area by considering the number of distinct 6 min grids that were ever fished within that interval over the duration of the logbook period

The latter weighting was the reviewers preferred option (Hoyle 2024) and was the method used for the final models.

One aspect of this additional work that has not made the final cut for the published report is the construction of entirely separate population models for spatial regions below that of the Gulf in scale (the intervals of the current catch rate model are brought together into a single time series using different spatial weighting approaches as discusssed above). A 'Karumba only' model was developed for exploratory purposes as suggested by the reviewers, but the outputs are preliminary and have not been included in the final report. This is important work to continue in the next assessment.

River flow—this has been removed from the catch rate analysis for the final report. We support the reviewer's recommendation in this regard of 'future research to develop a better understanding of the mechanisms through which flow may affect threadfin populations' and considered that at this stage not enough is known about these mechanisms to include it only here as as catchability adjustment.

4.2.3 Model fitting and index development

The quasi-negative binomial approach may be a useful catch rate modelling tool but we agree with the reviewers that this needs to be better demonstrated. The quasi-negative binomial modelling approach was removed and replaced with a hurdle delta-lognormal model, implemented using the *mgcv* R package as recommended by the reviewers. We also used DHARMa package to generate residual plots and other diagnostics.

4.2.4 Fishing power

As discussed above, there is now no longer any distinction between regions for q_{inc}, and it is no longer estimated inside the model in any sense. It is an important source of variation and we appreciate the reviewers feedback elicited from industry members and documented in their review. Future work should consider this further and a component 3.1 of Robins (2024) is dedicated to this.

In the final models dependence on this parameter was sensitivity tested by fixing it at 0% and 1% per annum.

4.3 Catch estimates

The fit to the 1980s and 1990s age-length data is improved. Karumba-only models were trialled with lower catch during the period ten years earlier without any noticeable improvement observed / any clear link between early age-length fit and catch history.

4.4 Age and length data

As discussed above this has been reworked. Significant time and fishery monitoring expertise was dedicated to the task of understanding and extracting maximal information on fishing location, and then sensitivity testing different spatial weightings in the population model, at both Karumba and Gulf-wide scales. This did not lead to any clearly preferable model. The exercise was limited by how much location information could be extracted.

Regarding seasonality and how the data is entered into the population model, this is a bias-variance trade-off and given our sample sizes annual is probably the correct choice here.

5 Biological parameters

5.1 Length-weight

Nothing to add.

5.2 Recruitment

σ_R of 0.6 was the sole value adopted for maximum likelihood and Monte carlo Markov Chain recruitment deviation estimation.

5.3 Natural mortality

For the final eight scenarios we had:

- four scenarios estimating M using a strong prior with median 5.4/Amax, where Amax was set at 20 years
- four scenarios with M fixed at 5.4/Amax

5.4 Steepness

For the final eight scenarios we had:

- four scenarios with steepness fixed at 0.7
- four scenarios with steepness fixed at 0.8

We agree with the reviewers that the data in this fishery appear to be insufficiently informative about either natural mortality or steepness for them to be cleanly estimated. We have adopted Hamel and Cope (2022) / reviewer recommendation as a reference point in the case of the former, and the (Leigh et al. 2021) steepness value / reviewer recommendation as a reference point in the case of the latter.

We do not consider Moore et al. (2017) to provide 'clear evidence of king threadfin's density dependent response to depletion' due to the likely confounding of this signal with the environment. Mixed evidence perhaps.

6 Model structure

6.1 Growth

Following the age-length data reanalysis described above, only 1988–1990 have been retained for age-at-length composition purposes. Pearson residuals can be seen in Figure E.5 for Scenario 1, Figure E.16 for Scenario 2, etc., but here we display the comparable plot to that from Figure 4 in the review and make a few remarks about how these plots differ.

- During the early period, the model no longer displays ‘subtantial and consistent’ under-prediction of age across the majority of the size classes which have observations. On average, some under-prediction remains but it is significantly reduced, partiularly in 1989 and 1990.
- The model scenarios 1 through 8 differ in several respects from those in the draft report, as documented elsewhere in this response, and we have not yet identified any clear primary cause for the improvment.
- Sampling from the early period is notable for its variety of mesh sizes being used (where known, often this is unknown), and the banded structure across size classes seen in 1988 (also, previously seen in 1991, 1992 and 1993) is likely a reflection of mesh size dependence.
- This is particularly the case for the few observations in 1988 at 95cm+ which likely came from 10-inch mesh and/or heavier ply.
- For ‘biological sampling’, and focussed as these studies were on growth and age-at-length, in theory this mesh-size dependence should not matter. That is, if the assumption is that the age of the fish is independent of the mesh used, conditional on size at capture, is met.
- In reality there could be many reasons for lack of independence.
- The reviewers mention catchment dependence, as well as somatic growth being dependent on long-time scale (decadal) fluctations in environmental conditions and/or fishing pressure.
- The early data contain a greater number of males, whereas for the later data females are more heavily represented, and we are not sure why.
- Improved age-at-length fit in 1988-1990 notwithstanding, the filtering out of the other years needs to be addressed.
- Further inspection and spatial re-weighting (or spatial/statistical modelling, given the sample sizes and categories of unknown) of this early data set is probably sensible before external covariates like flow are considered in any part of the model.
- Francis scalars for the age-at-length data set were routinely around 0.3, without convergence, whereas length composition Francis scalars converged.
- Despite the biological sampling focus of the early data, length composition data was retained for 1986–1990, and the implied age-composition fits for 1988-1990 are provided in Figure E.4, Figure E.15 etc.
- On balance, we judge that the early data are providing useful information to the model on growth and beyond (four scenarios now estimate natural mortality), but much remains to be resolved here.

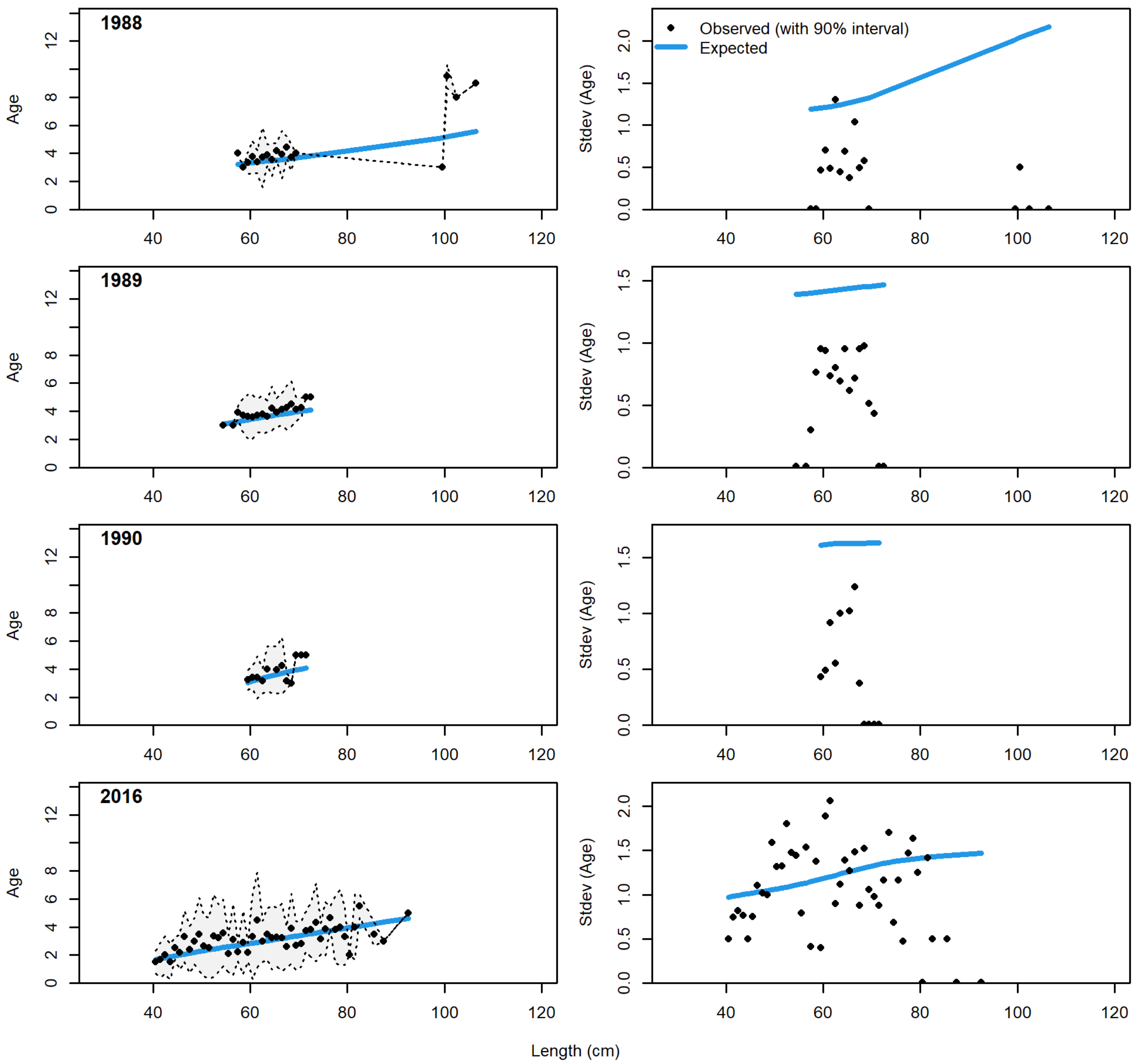

Figure 6.1: Conditional age-at-length plot for data from 1988–1990, and 2016. Plots show mean age-at-length by size-class (obs. and exp.) with 90% CIs based on adding 1.64 SE of mean to the data

6.2 Recruitment

Lognormal bias correction has been applied for maximum likeihood estimation. The bias correction ramp suggestion immediately converged.

One exploratory Karumba-only model did include a link between river flow and recruitment, using a Stock Synthesis stock-recruitment 'regime' parameter. Model convergence was achieved and the effect on biomass and the recruitment deviation trend was significant. However, it was unclear how to properly guage the value of this exploration given the challenges discussed elsewhere in this response. A structured approach to hypothesis testing will be required given the many possible mechanisms at work.

It is possible that the introduction of river flow as a survey fleet rather than a direct model covariate is a better way to go as this allows the environmental data to provide evidence for the recruitment deviations (Methot and Wetzel 2013; Schirripa et al. 2009). Note this would only be addressing flow impacts on biology, not catchability.

6.3 Fleet definitions

These have been reworked and the model is now a single fleet model.

6.4 Selectivity

Dome-shaped selectivity was trialled in the Karumba-only model. Project team discussion following the review confirmed reviewer perspectives on this, we agree non-asymptotic selectivity should be considered in the next assessment.

6.5 Data weighting strategy

This method has been removed and replaced with Francis weighting. Also note the length-age data reanalysis constructed intial ('stage 1') sample sizes using Fishery Monitoring 'catches' as the sampling unit.

6.6 Ensemble scenarios

Plausibility was a condition on the diagnostics, e.g. recruitment deviation trend. It is not a term used to define the ensemble grid in the updated report. The ensemble grid is defined in Section 2.7.4 'Sensitivity tests'.

7 Model outputs

7.1 Model diagnostics

Fits to CPUE (Figures E.1, E.12, etc.), length composition (Figures E.3, E.14, etc.) and age-at-length (Figures E.5, E.16, etc.) are improved. Likelihood profiles are significantly improved (Section D 2.1), likely due to the reduction of the inital sample weights as noted. All these figure references are to figures in the final report (Campbell et al. 2024).

7.2 MCMC and convergence

MCMC diagnostics were generated by running parallel chains using the No U-Turn Sampler and then R packages *adnuts* and *shinystan* to examine effective sample sizes and Rh-at values. These are provided in Setion D.1. The majority of parameters had R-hat values less than 1.05 and effective sample sizes over 300. A useful reference in this regard was Monnahan et al. (2019).

All 8 scenarios converged with positive definite hessians to final gradients of 1e-3 or less.

7.3 Abundance and recruitment trends

Recruitment trends across the final ensemble are provided in Figure D 12, and for each individual scenario in Figures E.2, E.13, E.24 etc. As noted, statistical weighting of age-length is greatly reduced but patterns in deviation parameters remain. The dynamic B0 plot for Scenario 1 now looks like this:

The blue line now does not trend down over the whole time series to the same degree, but does have clear decadal scale periods of fluctuation. One intepretation is that the environment is indeed behaving like the blue line would suggest. Another interpretation is that some form of model misspecification is driving the deviation patterns. Estuarine-based fisheries in the Gulf of Carpentaria (e.g., banana prawn, barramundi, mud crabs and by association king threadfin) are strongly environmentally influenced system and until we better understand the mechanisms behind those influences and build them into our models it will be hard to know.

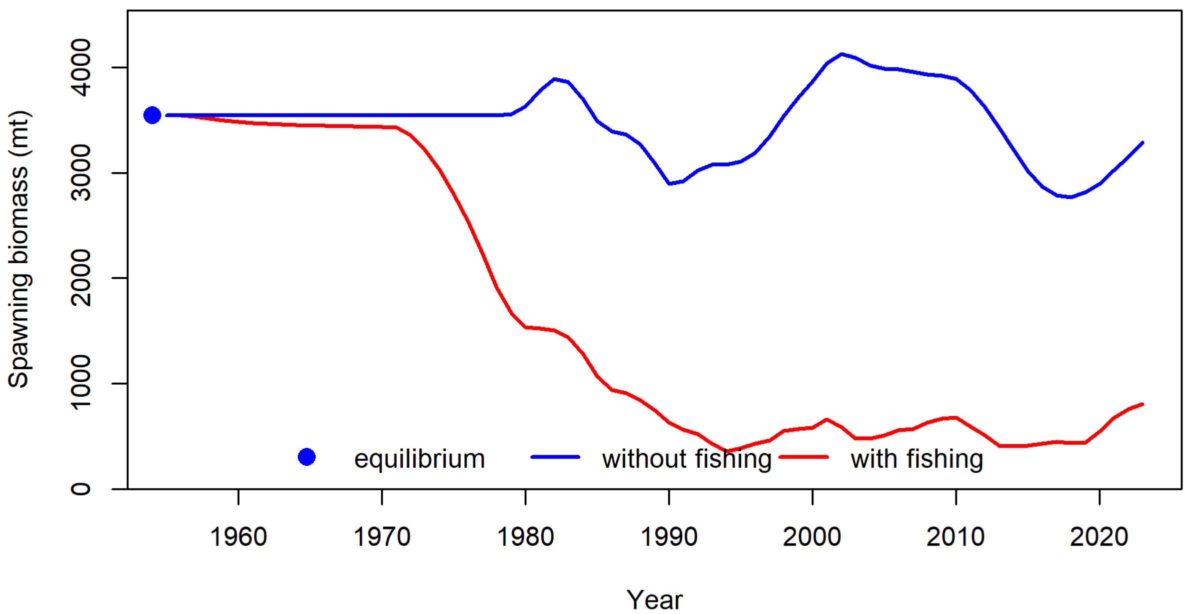

Figure 7.1: Dynamic B0 plot. The lower line shows the time series of estimated Spawning biomass (mt) in the presence of fishing mortality. The upper line shows the time series that could occur under the same dynamics (including deviations in recruitment), but without fishing. The point at the left represents the unfished equilibrium

8 Discussion

The review put a lot of focus on spatial weighting issues, and in particular on the idea that the draft models were over-weighting more heavily depleted regions and thus biasing the biomass results downwards. Having worked through several exploratory models and spatial re-weighting exercises, and having considered the spatial catch rate residuals at interval-scale, we think this is a less likely source of bias than several other key issues. These include, but are not limited to:

- targeting
- fishing power related catchablity increases
- river flow dependence of recruitment and/or natural mortality and/or growth
- river-foreshore / catchment dependent demographics
- functional reproductive connectivity extent / isolation by distance
- life history parameters
- catch history
- improved use of 1986–1994 data to inform mortality related parameters given the sampling program was designed for biological purposes rather than to be population-representative

An FRDC research project proposal, Robins (2024), will address several of these points. The objectives of this proposal are provided in Section 4.4.2.

9 Reviewer feedback on response to review

After applying the changes to the stock assessment model and inputs, the authors of Hoyle and Dunn (2024) provided the following feedback in response to the prompt questions which are denoted in bold text.

To what extent are these changes an improvement?

We appreciate the substantial effort in a short time by the assessment team to respond to our recommendations in the review and incorporate them in an updated assessment. The updates involve significant improvements, both for this assessment, and in approaches used which will help future assessments. The assessment is linked more closely to an understanding of the biology, and to information provided by fishing industry representatives.

1. Data preparation
   (a) More characterisation of data inputs (companion report).
   (b) Substantial reduction of the internal conflict between datasets.
   (c) Include more of a spatial component in the CPUE - habitat-based.
   (d) Remove the river flow component from the CPUE modeling.
   (e) Use of standard distributions for the CPUE analysis, which allowed more testing of whether they fitted the data.
   (f) Provision of a wider range of CPUE diagnostics.
   (g) Improvement of the approach for steepness, natural mortality and sigmaR.
2. Modeling methods
   (a) Use of Francis weighting rather than an unvalidated method.
   (b) Lognormal bias correction ramp for recruitment.
   (c) Models appear better converged and with less internal conflict, and somewhat better MCMC diagnostics, although there are still indications of unstable convergence.

What further things should be considered for the next assessment, especially as material from the FRDC project becomes available?

There are a number of areas where further work is likely to be helpful and improve both the estimates and understanding of the system. This work is likely best carried out once more information is available about the spatial structure of the stock and its dynamics.

Data preparation

- Catch
  - The extent of the early depletion depends on the catch before the late 1980s, but this catch remains very uncertain. Consider including catch scenarios as part of an ensemble approach.
  - The assessment assumes that the population starts at virgin biomass in 1955, but some catch would have been taken by indigenous communities. This should be considered along with ongoing Indigenous catch.
- Selectivity
  - Develop an approach for applying non-asymptotic selectivity.
- Age composition data

- The current analysis removes ageing data for 1986, 1991–1993, and 2008–2011, and 2015. Removal of 2015 is explained and is reasonable. However, the exclusion of the years 1988 and 1991–93 appears somewhat ad hoc, which the report acknowledges. Removal of the mid-2000s data is not explained in any depth.
- In general, I'm not convinced by the argument that age at length data may be affected by mesh selectivity. I agree that selectivity will affect sizes caught, but it should not affect observations of age at length, which are already conditioned on fish size. However, I agree there are some plausible possible explanations–see below.
- Nevertheless, it is appropriate to avoid combining, in a single model, datasets that simply don't fit together. Excluding at least the early data may be justified, though alternatives should be considered when time permits. The 1986 and 1991–3 data were inconsistent with the 1988–1990 data as well as more recent data. Removing these data represents a hypothesis that the inconsistent samples came from a different part of the population that had experienced a different fishing history, and that they were not representative of the part of the population being assessed.
- Including all the early data may lead to their effective exclusion from influence, because once Francis weighting converges (it should be allowed to—see below), all the early age data may be completely downweighted.
- The report acknowledges that the age-at-length data need further work, and that the model needs more work to develop hypotheses and alternative scenarios. The various possible explanations for the poor fits of the age-at-length data should be listed, researched, and then explored with models and either included in an ensemble or rejected.
- Possible hypotheses include:
  * There are local subpopulations, and the stock has undergone serial depletion, as boats and freezer technology improved. Perhaps the samples with old fish were from locations that hadn't been fished for a while, perhaps even from the Northern Territory. Improving freezers, or adding freezers to boats with low draught, may have allowed vessels to start fishing in more remote areas and still bring fish to market.
  * The 1970s catch estimates may be much too high, and depletion in the late 1980s was less than estimated.
  * The estimated CPUE trend may be very wrong due to some unknown feature, such as a higher degree of effort creep than expected, so that depletion in the late 1980s was less than estimated.
  * Some form of ageing bias may have occurred with the early otolith samples.
  * Large changes may have occurred in productivity/recruitment through time.
  * Growth in the early period may have been slower than recent estimates, e.g., due to density-dependence, leading to higher estimates of age at length.

• Length composition data
  - Data from the Pormpuraaw and Aurukun regions have been included, although they are believed to come from incomplete samples with large fish not sampled. If still considered to be unrepresentative, they should not be included in the model.
  - Include sampling month (also for age data).

• CPUE
  - Exploration of patterns in the CPUE data is always helpful. I think the progress recorded here is a good example of this. CPUE is one of the most influential parts of the model, so further work in this area is encouraged.

  - Model diagnostics are helpful. Note the lack of fit of some fisher effects in the binomial model indicating a very high level of non-zero catch rates for these fishers (Figure B.2). This may reflect a targeting or reporting issue. The sample sizes for these fishers may be low.
  - The 'response to the review' notes that balancing the CPUE weights so as not to over-emphasize trends in the more heavily depleted regions had less impact on the CPUE trend than the reviewers expected. In general, we agree. However, we note that the time-area interactions in the CPUE analysis model were applied at relatively large spatial scales. At smaller spatial scales, analyses of Karumba-area CPUE (Campbell 2024) identified lower catch rates closer to Karumba, which may be consistent with fine-scale depletion (more fishing leading to lower density). The larger scale analyses are unable to pick this up. At larger scales, the spatial patterns may reflect variation in habitat quality affecting densities of fish, and lower effort is likely where fish density is lower.
  - It was hard to diagnose the CPUE models because the result tables were missing. Additional diagnostics such as influence plots would be very helpful.
  - Effect plots for the lognormal model component show year, month, and interval effects, but each of these estimates is at the base level of the other parameters, without allowing for the effects of interaction terms. It would be more informative to provide results for each parameter after using prediction from the model to integrate across the values of other parameters. The patterns of the interaction effects must have been important for the lognormal component because the indices vary significantly by interval.
  - Also note that the delta model component omits year-interval and month-interval interactions, probably due to lack of sufficient data to estimate them. However, they are likely to occur and may be important, and could probably be estimated using different approaches, e.g., by using smoothers rather than categorical variables for interaction terms (in a GAM), and/or by using a spatiotemporal model.
  - The models were fitted entirely with categorical variables, and base values for some variables were associated with sparse data instead of using the most common value (e.g. Figure B.6). Using smoothers for variables like month, mesh, net length, and spatial effects is likely to provide more precision and accuracy.
  - Spatial aspects should include riverine versus coastal effort, once that information is available.
  - Targeting remains an important issue, and we agree with this comment in the report. The current approach to targeting is ad hoc, and better methods are available. Information from species composition and vessel tracking can be used to infer both targeting and fishing location (coast versus estuary).
- Biological parameters
  - A wider range of natural mortality scenarios might be considered for the ensemble. M is very influential and there is a high degree of uncertainty about this parameter.

Modeling methods

- Spatial
  - Spatial factors remain important. There is a need to understand the degree to which there are local (river catchment) dynamics, and how much mixing there is between locations at different spatial scales and age classes. The proposed FRDC project has potential to contribute significantly to understanding that will help structure the assessment. Movement information from acoustic tagging could be informative about mixing of older age classes. Otolith mi-

crochemistry could be informative about movements across the life span, as well as about relationships between environment / flow, movement, and recruitment.

  - Running separate assessments by region would provide insight into the information currently available by region.
  - There are intermediate steps between the current approach (assuming a single pool of fish) and independent models: 1) an areas-as-fleets approach, or 2) a model with a shared breeding stock but independent substocks in separate regions. Option 1 would (like the current model) not allow for different abundance trends by region, but it would allow for any spatially varying selectivity and size distribution so is likely to fit the composition data better, which is important. Option 2 may be a better approach if biomass trends do indeed vary spatially, but it can be very difficult to estimate movement parameters, and a convincing converged model may not be achievable.
  - Option 3, using a separate model for each region, has some important similarities to option 2 above but assumes no movement between regions. It is easier to work with one area at a time, particularly in a low-data situation where there is very little information about some areas. This approach may mainly be useful for developing understanding, and then used to fit a model based on areas as fleets, or option 2 with fixed or highly constrained movement parameters.
  - Ultimately the objective is to estimate the dynamics and spawning potential of the stock. Currently it is unclear how the parts of the metapopulation link together, and research such as the proposed FRDC project is needed to help inform this. The habitat near Karumba may be key to the spawning dynamics, or the stock may require a certain amount of spawning in each region, or there may be important contributions from spawners in every individual catchment. Assessments for these cases probably need different structure, and different information.

- Convergence
  - There are still apparently some problems fitting to the early data - see the unusual variation pattern in the MCMC medians (Figure 3.17).
- Data weighting
  - The Francis weighting process should be repeated until the weight estimates converge. Lack of convergence is a sign of model assumption/structure/data issues - say between early and late data, in which case you could deal with it via alternate scenarios.
  - Francis weighting should be applied independently for age data series with different sampling approaches. I suggest that there should really be at least 3 groups 1980s/90s, 2008-2011, and since 2016. The early sampling may not have used a sampling ‘design’ anyway – the sampling methods are unclear.
  - Note – check that better approaches for Francis weighting of age-at-length data (Punt 2017) are included in the current version of SS.
- Fleet structure
  - Including river flow as a survey fleet would involve fitting it to an age class. This is appropriate for an index of recruitment, but it would not work as a covariate for CPUE. It therefore addresses a different issue from the previous method of including it in the CPUE model.

References

Cadrin, Steven X, Daniel R Goethel, Aaron Berger, and Ernesto Jardim (2023). “Best practices for defining spatial boundaries and spatial structure in stock assessment”. In: *Fisheries Research* 262.106650.

Campbell, Alex (2024). personal communication.

Campbell, Alexander B., Mai Tanimoto, Olivia J. Whybird, and Alise R. Fox (May 2024). *Stock assessment of king threadfin* (Polydactylus macrochir) *in the Gulf of Carpentaria, Queensland, Australia, with data to December 2022*. Fisheries Queensland, Department of Agricluture and Fisheries.

Hamel, Owen S and Jason M Cope (2022). “Development and considerations for application of a longevity-based prior for the natural mortality rate”. In: *Fisheries Research* 256, p. 106477.

Haugen, JB, GB Skomal, TH Curtis, and SX Cadrin (2022). “Interdisciplinary stock identification of North Atlantic porbeagle *(Lamna nasus)*”. In: *Journal of Northwest Atlantic Fishery Science* 53.

Hoyle, Simon D. (Apr. 7, 2024). personal communication.

Hoyle, Simon D. and Alistair Dunn (Mar. 2024). *Review of the 2024 stock assessment for king threadfin* (Polydactylus macrochir) *in the Gulf of Carpentaria, Queensland*. Nelson, NZ: Hoyle Consulting Ltd.

Leigh, George M., Mai Tanimoto, and Olivia J. Whybird (2021). *Stock assessment of king threadfin (*Polydactylus macrochir*) in Queensland, Australia*. Brisbane: Department of Agriculture and Fisheries.

Methot R. D., Jr. and C.R. Wetzel (2013). “Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management”. In: *Fisheries Research* 142, pp. 86–99.

Monnahan, Cole C, Trevor A Branch, James T Thorson, Ian J Stewart, and Cody S Szuwalski (Apr. 2019). “Overcoming long Bayesian run times in integrated fisheries stock assessments”. In: *ICES Journal of Marine Science* 76.6, pp. 1477–1488. eprint: `https://academic.oup.com/icesjms/article-pdf/76/6/1477/31247349/fsz059.pdf`.

Moore, Bradley R., Jason M. Stapley, Ashley J. Williams, and David. J. Welch (Mar. 2017). “Overexploitation causes profound demographic changes to the protandrous hermaphrodite king threadfin (Polydactylus macrochir) in Queensland’s Gulf of Carpentaria, Australia”. In: *Fisheries Research* 187, pp. 199–208.

Punt, André E (2017). “Some insights into data weighting in integrated stock assessments”. In: *Fisheries research* 192, pp. 52–65.

— (2023). “Those who fail to learn from history are condemned to repeat it: A perspective on current stock assessment good practices and the consequences of not following them”. In: *Fisheries Research* 261, p. 106642.

Robins, Julie B. (2024). *Gulf of Carpentaria King Threadfin (Polydactylus macrochir) - addressing the knowledge gaps to support assessment, management and sustainable harvest*. FRDC Funding Application, Project Number: 2023-199. Queensland Department of Agriculture and Fisheries.

Schirripa, Michael J, C Phillip Goodyear, and Jr. Methot R. D. (2009). “Testing different methods of incorporating climate data into the assessment of US West Coast sablefish”. In: *ICES Journal of Marine Science* 66.7, pp. 1605–1613.

Walters, Carl (2003). “Folly and fantasy in the analysis of spatial catch rate data”. In: *Canadian Journal of Fisheries and Aquatic Sciences* 60.12, pp. 1433–1436.