OCEANS AND ATMOSPHERE FLAGSHIP
www.csiro.au

Targeting and CPUE definition in the SESSF trawl fishery through auxiliary data

Mark V. Bravington and Scott D. Foster

FRDC Project No. 2008/002
June 2015

FISHERIES RESEARCH & DEVELOPMENT CORPORATION

ISBN 978-1-4863-0572-8 (Online version)
ISBN 978-1-4863-0573-5 (Print version)

Title: Targeting and CPUE definition in the SESSF trawl fishery through auxiliary data
Authors: Mark V. Bravington and Scott D. Foster

FRDC Project Number: 2008/002

Published by: CSIRO's Oceans and Atmosphere Flagship

The Fisheries Research and Development Corporation plans, invests in and manages fisheries research and development throughout Australia. It is a statutory authority within the portfolio of the federal Minister for Agriculture, Fisheries and Forestry, jointly funded by the Australian Government and the fishing industry.

Researcher Contact Details

Name: Scott Foster
Address: CSIRO Marine Labs
Castray Esplanade, Hobart, TAS 7001
Phone: 03 6232 5178
Fax: 03 6232 5000
Email: scott.foster@csiro.au

FRDC Contact Details

Address: 25 Geils Court
Deakin ACT 2600
Phone: 02 6285 0400
Fax: 02 6285 0499
Email: frdc@frdc.com.au
Web: www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

Contents

List of Tables

List of Figures

Acknowledgments

We would like to thank those who attended the Hobart workshop SESSF CPUE workshop in April 2007. This collection of RAG chairs, fisheries managers and stock assessment scientists highlighted to us the importance of targeting, and what is done about it during analysis, in the conduct and management of the SESSF fisheries.

Many thanks go to Malcolm Haddon, Geoff Tuck and Dave Smith who supported this work through its long journey; to Ian Knuckey for obtaining the elusive Sydney market data for us; and to FRDC for their patience. Thanks also to André Punt and David Sampson for the insightful comments on an earlier version of this report.

Most of all, our thanks go to John Jarvis and his colleagues for taking the time to enlighten us about the real factors behind the numbers.

Executive Summary

Management of the SESSF is still heavily driven by CTS trawl CPUE data, despite the acknowledged interpretational problems and consequent risks (of failing to detect either overexploitation or recovery). CPUE is currently standardized species-by-species, but this raises questions at a species level about appropriate data subsetting, and whether long-term changes in fishing incentives have distorted the CPUE trend. The answers assumed do have an impact on TACs, and there is no purely statistical criterion that can be used at a species-by-species level to resolve them.

This project explores whether it is possible to answer these questions more systematically by taking a multispecies viewpoint on each shot, to somehow take account of "targeting", i.e. fisher-controlled specifics of each trawl shot that affect what it is likely to catch, but which are not recorded in logbook data. While there are several published approaches to "multispecies CPUE", none are statistically satisfactory in a setting as complex as the CTS, so we needed a new approach. It is clear from discussions with industry that some degree of targeting is possible in the CTS, and also that the economic drivers to catch particular species have changed significantly since the introduction of ITQs in the early 1990s. We developed a model of economic drivers to generate prior probabilities that each shot would be targeted in the various different ways, and we linked this model to a CPUE standardization carried out simultaneously for all the main quota species. Overall, the model estimates what the targeting types are (in terms of how they affect catch rates of different species), the shifts in targeting over the years, and various parameters related to the economic drivers, as well as the usual abundance index series from a "standard standardization" of CPUE. The model is complicated, but at least it should be statistically coherent and less susceptible to the pitfalls which plague simpler ad hoc approaches to multispecies CPUE, in particular because the choice-of-type for each shot has an explicit underlying model rather than relying merely on a criterion of statistical tidiness. How good that underlying model actually is, is another question in itself.

In practice, the new model does not suggest markedly different trends in CPUE compared with a standard species-by-species standardization, even though the estimated effects of the various types on catch are substantial, and even though there are apparent trends in targeting at least in some seasons and depths. Either the changes in targeting have had little impact on *overall* CPUE across all shots in the CTS, or there are strong trends but the model cannot detect them. While our model is not perfect— it is exploratory, and because of time and computational constraints we could not incorporate every single bell-and-whistle— we are doubtful that a more complicated model (if feasible at all) would discover a stronger signal. We are also skeptical that simpler approaches (e.g. without economic drivers) to multispecies CPUE in the CTS can avoid the pitfalls outlined in the Introduction; any such proposal should be considered beforehand in that light, in mathematical terms and/or through simple simulation testing (for which we have provided some software).

Given that (i) our inferred abundance indices are not much different from a standardization (at least when zero-catch shots are included), and (ii) that our model is complex and consequently fragile to run

and would require substantial further work to bring to "production quality", we do not recommend that our model should replace species-by-species CPUE standardization. That is not at all to say that the latter is fine— rather that the issues seem irresolvable, and that from a management perspective there is little point in putting much further work into improving the handling of CTS CPUE. In particular, and regardless of whether one takes a single-species or multi-species approach, "effort creep" (improvements in fishing efficiency that are not captured in logbook or other data, and thus cannot be accounted for in standardization) remains a major concern in CTS CPUE, especially given CPUE's strong influence on the SESSF's top-tier Harvest Control Rules, which in generic simulation tests have been shown vulnerable to effort creep. The need for more reliable fishery-independent abundance indices in the SESSF, whether from surveys or some new approach, is as strong as ever.

Introduction

1.1 Characteristics of the Commonwealth Trawl Sector (CTS)

The CTS is the Commonwealth-managed (i.e. offshore) demersal trawl component of what is now called the "Southern and Eastern Scalefish and Shark Fishery" (SESSF)[1]. The CTS is the largest component of the SESSF, currently catching around 10,000t annually (formerly much higher) with GVP of $51M in 2011 (Woodhams et al., 2012). It stretches south from Sydney, around Tasmania, to Cape Jervis near Adelaide in South Australia, in a narrow band about 8-15km wide that starts 3nmi offshore (on the continental shelf, shallower than about 200m) and extends down the continental slope to about 700m. Some form of trawl fishery has existed in parts of this region for over 100 years, and there have been numerous jurisdictional and operational changes. For further details, see the annual status reports (e.g. Larcombe and Begg, 2008; Woodhams et al., 2012) from which much of this material is sourced, and Klaer, 2006 for a historical perspective.

The CTS catches over 100 species, but about 80% of the catch comprised of 20-25 quota-managed "stocks" (species, populations, and/or species-groups); in 2014, 66% of the quota-related revenue came from just three species, though historically the spread has been wider. Some of the species are caught to a lesser extent in other components of the SESFF. Overfishing has historically impacted many of the quota species, leading to the progressive introduction of total allowable catches (TACs) and individual transferable quotas (ITQs) from the late 1980s, and buyouts ("structural adjustments") in 2006. Management is now by ITQ with limited entry, plus some restrictions on gears, times, and areas. By the mid-2000s, TACs had become sufficiently constraining to impact fishing behaviour, and the second buyout achieved substantial reductions in active fleet size and in real effort. Also, TACs on some species have impacts on catches of species with similar habitat preferences (companion species, see Klaer and Smith, 2012). Nevertheless, in defiance of economic theory there is still substantial "latent quota" for some species (i.e. quota that is held, but not leased or used). TACs are set based on single-species stock assessments which are mostly driven by standardized trawl CPUE for that species, assumed directly proportional to abundance.

Most trawl shots in the CTS catch several species. Species-level CPUE is derived from shot-level logbook data which include a number of variables relevant to CPUE standardization (e.g. duration of trawl, time of day, average depth, start and end location), but not all (e.g. tweaks to the way the gear is set, proximity to seabed features, entire track of the shot). CPUE standardization is currently done separately for each species using generalised linear models (GLMs), with some filtering to remove shots deemed "irrelevant" for that species— e.g. the vexed question of whether to remove zeros Haddon, 2012. For most of the CTS quota species, the standardized CPUE index is one of the primary driver of the stock assessment, from which Total Allowable Catch (TAC) is set by a Harvest Control Rule (HCR) based on estimated overfished/overfishing

[1]The official nomenclature has changed over the years

status; different HCRs (“tiers 1-4”) are in use for different species, depending on the perceived reliability of “the assessment” (in a broad sense). The details are complicated (see e.g. AFMA, 2014) and irrelevant to this report; the point to note is that CPUE remains one of the main determinants of TAC, despite numerous acknowledged shortcomings. Section 1.3 discusses some pitfalls in general (non-CTS) settings of CPUE dependence, both single- and multi-species.

Experienced fishers in the CTS have considerable, though not limitless, scope to “target” their trawl shots to increase the probability of catching particular species. Part of this targeting is captured in logbook data (e.g. average depth, approximate location) and can be “standardized out”, but part cannot (e.g. fine-scale positioning of trawls). In addition, the incentives to catch, or avoid catching, particular species have changed substantially as TACs have been reduced. This potentially creates an extra problem for CPUE: once a species has been overfished to the point where a really low TAC is set to allow recovery, targeting will deliberately avoid that species in future (since it cannot be profitably caught) so that the apparent CPUE will never recover. This has sometimes been proffered as an explanation for the apparent non-recovery of several once-overfished SESSF species where quota restrictions were expected to lead to recovery, including blue warehou and school sharks.

1.1.1 Questionnaire data

To get some insight into (i) the operational issues of targeting, and (ii) the changing incentives to fish in different ways, we prepared a questionnaire for CTS fishers. The responses we received come from only a limited part of the present and past CTS (fishers involved in the RAGs and their immediate colleagues), so cannot be considered comprehensive. Nevertheless, the replies— and subsequent discussion with fishers and scientists— were extremely informative and led us to develop a model structure that was quite different from our preliminary expectations, with a much stronger economic focus.

The questions are shown in Appendix A. For confidentiality reasons, this report does not include the actual responses. Below is our summary of the salient points from the point-of-view of handling unrecorded targeting in a CPUE standardization; this omits a lot of detail in the responses that is important operationally, but not in our view for modelling— e.g. fine-scale positioning of trawls. We have also included some points that are general background to the CTS. Many of these points have exceptions, which we have not listed in detail.

- Boats typically operate in 3-5 day trips, with several shots a day.
  - Logbooks do not provide enough information to reliably separate one trip from the next.
  - Catches are aggregated within a trip, and sold on return to port mostly through one of two main sources, The Sydney Fish Market or one of theMelbourne markets.
  - Market price information is available to fishers during a trip, and thus plays a role in deciding where/how to fish. However, prices can fluctuate suddenly and unpredictably, due e.g. to imports, between making a shot and landing the catch. There are strong (and more predictable) seasonal variations in the price of particular species.
- At the time of writing, vessel-level catches are reconciled by AFMA against vessel-level quota currently 4 times a year. It is not necessary to hold quota for a species at the time it is landed; quota can be leased retrospectively, up until the next reconciliation.
  - If quota is unused by the end of the 12 month TAC year (late April), there is limited rollover

  - The cost of leasing quota on the "open market" (often from holding companies) can nowadays sometimes approach the market price for a species; in the early days of quota management, before TAC became seriously restrictive, 10% was usual.
  - Open-market quota becomes more difficult to lease (thus, more expensive) towards the end of the annual TAC period (late April).
  - Informal agreements are widespread, to lease future quota if necessary at low cost between vessels and companies.
  - Discarding can occur, when a vessel does not expect to be able to retrospectively lease enough quota affordably.
  - Despite TACs sufficient to restrict effort and an ITQ system, there is often still substantial uncaught TAC for many species, even the handful of "iconic" ones that dominate earnings and drive fishing decisions (noting that the "iconic" list has changed over time).
  - There is now a disincentive to fish a species when it is at its most seasonally available, because the market will be swamped then and prices low; the best annual return from a given quota is theoretically obtained by catching when the species is least easily caught. This is a shift away from the cultural incentive to visibly bring in as much fish as possible.
  - "Mixed-bag" shots (and especially trips) have become economically much more desirable, though they are operationally somewhat inefficient per kilo precisely because they tend to avoid the easiest-to-catch species at that time.

- Operationally, the likely mix of species caught can be controlled in several ways:
  - via the depth profile of the shot (mixed-bag shots in particular travelling more up and down slope, rather than along contours)
  - by staying closer to or further from seabed features, on a fine scale (i.e. distances of a few hundred metres) so as to match species' habitat preferences;
  - by (limited) gear adjustments, e.g. footline tightness.
  - by changing direction, perhaps in response to echosounder observations (e.g. returning to a depth where fish appeared dense)

- Not all relevant data is recorded in logbooks. For example, the logbook records only average depth rather than full depth profile, has limited information on gear settings, and has no track information between start and end locations.

- Some shots also have be spent learning about current distribution, because of within- and between-season movements
  - Big reductions in fleet size have reduced the amount of within-fleet information-sharing.

Overall, it is very apparent that economic drivers, specifically quota tightening and enforcement, *have* substantially changed behaviour in the fishery: an avoidance of fish at times/places when they are plentiful. Quite likely, this has had an some impact on CPUE, or more accurately on how CPUE relates to abundance—but given that CPUE is effectively averaged across all shots in a season, even the direction of the impact is hard to guess.

1.2 Terminology

There is no universal agreement on what the term "targeting" means. When CTS fishers make a trawl shot, they make many decisions which affect the likely catch and species-mix: where along the coast, depth, time of day, gear settings, track direction (along or across contours), etc. From the point-of-view of the fisher, "targeting" may encompass all of these decisions, which could in principle be captured numerically as covariates and included in standardization.

From the point-of-view of CPUE analysis, the important point is not so much what label to attach to a specific "type" of targeting, but crucial distinction is this:

(i) *recorded* factors, whether inside fishers' control (such as depth of shot) or outside it which are obtainable from logbook/vessel/general data and therefore can included in standardizations; and

(ii) *unrecorded* factors under the control of fishers which do affect the likely catch, but which cannot be included in a conventional CPUE standardization. (See next chapter for what this might mean in the CTS.)

There is also a third category— unrecorded factors known to a fisher but *not* under their control, such as wind strength and direction. While such factors are very important in deciding how to make the next shot, omitting them is no problem for CPUE analysis as long as they have no strong long-term trend; they simply contribute to the general "noise" in CPUE data, which is allowed for and would still be substantial even if every single conceivable variable was measured.

In this report, unless specifically stated otherwise we use "targeting" to refer *only* to the second class of *unrecorded-but-controlled* covariates. This is not what everyone would mean by "targeting", which can sometimes be used to describe other phenomena (such as what species the fisher *hoped* to catch), or to include controlled covariates which are recorded (at least partly, like average depth of a shot). But despite the potential for confusion, it is essential to find *some* short term for unrecorded-but-controlled-covariates, and we have opted for brevity.

We have also assumed that targeting, even though it combines several covariates some of which can be varied continuously rather than stepwise, can be summarized into a small number of discrete categories. This is essential both to make the modelling feasible, and to give some chance of interpretability to the results— though just because an approximation is unavoidable, it does not follow that it is adequate.

Our model for "targeting" embodies it as a multiplier on average catch rates per species, over and above the effect of recorded covariates that already included in standardization. Since in practice a fishers' decision is not split into "recorded and unrecorded", some care is needed when trying to interpret the magnitude of our "targeting" estimates— they pertain only to aspects not captured elsewhere in CPUE standardization.

1.3 Review: CPUE and MS CPUE

There are well-known problems with relying on commercial CPUE as an index of relative abundance, even in single-species fisheries. The summary is that CPUE (even after statistical standardization to account for factors other than abundance that are known to affect catch rates) can change systematically but undetectably over time for reasons unconnected with abundance. This can sometimes mean, for example, that there is no *measured* drop in CPUE despite a genuine drop in abundance ("hyperstability"; Hilborn and Walters, 1991), until it's too late and the stock, and the fishery, collapse. Importantly— and largely independent of any multispecies/targeting aspects— insufficient information is available in the CTS (and in most fisheries that we are aware of) to retrospectively assess "effort creep"[2], i.e. the increases in fishing efficiency arising from improvements in fishing equipment such as advent of GPS, improved depth sounders, etc. This is liable to make trends inferred from CPUE over-optimistic, the more so the longer the series. For overviews of CPUE and its pitfalls of CPUE, see Hilborn and Walters, 1991 and Maunder et al., 2006. For an exampleof a fishery that has gone to considerable lengths to measure effort-creep, one should look at the Northern Prawn Fishery (Cartwright, 2005).

Before use as an abundance index, CPUE is usually standardized for the reason given above, by fitting a statistical model such as a generalised linear mixed model (GLMM) or a generalised additive model (GAM) using recorded covariates (e.g. Maunder and Punt, 2004). This is not necessarily simple, because of the large number of recorded covariates typically available and the consequent range of possible models that could be fitted. In principle, though, if the available covariates do capture all the relevant information about each fishing operation, including drivers of any effort-creep, then CPUE standardization is a well-understood statistical procedure that is not intrinsically problematic. This report focusses on *unrecorded* covariates, so for simplicity we therefore assume that standardization with respect to recorded covariates is "done right"; our own model is in effect a standardization procedure that allows for unrecorded as well as recorded covariates.

Allowing for targeting, and dealing with MS CPUE generally, is an especially complicated business. There are a number of published approaches for specific and/or generic applications (He et al., 1997; Pelletier and Ferraris, 2000; Punzón et al., 2010; Iriondo et al., 2010; Punzón et al., 2011; Winker et al., 2013; Winker et al., 2014), but no clearly "right answer" and no clear review[3]. And despite the work we put into this report, we ourselves have no great confidence that we have arrived at the "right answer", either. However, we have at least become aware of many pitfalls. To provide some clarity in a complicated subject with a confusing literature where *ad hoc* proposals are common, we have organized this review by listing those pitfalls, with conceptual examples and literature references as appropriate. We have also written an accompanying R package that can be used to explore some of the points in simple illustrative examples (but not for analysing real data): see Chapter 5.1.2.

The examples here are not meant to be realistic— they are deliberately extreme, each to highlight a specific point which may be just as relevant in more complex realistic settings, but will be harder to see. Some of the issues around MS CPUE can also be seen in single-species examples, though they may have extra bite in MS CPUE settings; for clarity, we have used single-species examples where possible.

1.3.1 "Automatic" analyses are risky

Consider analysing a hypothetical single-species fishery where each "shot" either catches 1 or 0 fish from that species. Over time, there is an increase in the proportion of zeros: what does this mean for abundance? In the absence of other insight, this is statistically *impossible* to resolve, because the data can be explained equally well by two very different models:

1. Abundance has declined, so each shot is less likely to catch a fish;

2. Abundance has stayed the same, but shots are being done differently over time, in a way that reduces the probability of catching a fish (e.g. because fuel costs have increased).

Actually, these models are just part of a continuum; it's possible that efficiency has actually *increased* while abundance has decreased *faster* than the observed trend.

[2] also known as "technological creep"

[3] At the time of writing this report, we are aware of plans to produce such a review, but it is some way off completion.

The real issue here is nothing to do with zeros *per se*— it's about the importance of understanding enough about the "motive and opportunity" of fishing to pick the right model, and the fundamental inability of any statistical method *on its own* to choose between phenomena that are statistically confounded[4].

Clearly, this issue is important when modelling targeting, because the range of plausible models with different implications may be particularly large. However, it applies in simpler settings too, e.g. to "shot selection" as currently used in single-species CPUE standardizations for the CTS— for example, should shots that catch nearly-zero of a species be included in the standardization dataset for that species?

Points to ponder: would this problem suddenly go away just because more species are included in the analysis? If so, why? and would it matter whether the species truly have similar, or different, abundance trends? If it is impossible to distinguish between two models statistically, and there is not enough "insight" to make the call, can Management Strategy Evaluation (MSE) be used instead?

1.3.2 Two-step analyses

Some approaches to targeting split the problem into two steps. First, they use (only) the catch-composition of each shot to classify the targeting, either by assigning to one of a fixed number of discrete "types" (e.g. He et al., 1997; Pelletier and Ferraris, 2000; Stephens and MacCall, 2004; Iriondo et al., 2010; Punzón et al., 2010; Carvalho et al., 2010; Castro et al., 2010; Punzón et al., 2011; Deporte et al., 2012), or by converting the composition into a low-dimensional continuous variable which is assumed to serve as a proxy for the controlled-but-unrecorded covariates (Winker et al., 2013). In either case, the result is treated as an "exact covariate" in the second step, a "standard" species-by-species CPUE standardization[5].

If the first step (classification) always worked perfectly, the two-step approach would be statistically fine. That may indeed be the case in some fisheries where there are limited options for targeting and good *a priori* understanding of "motive and opportunity" (see section 1.4). But the two-step approach seems often to be applied automatically (i.e. relying on statistical devices such as clustering algorithms to do the first step) and can go wrong in several ways that may invalidate the statistical assumptions of the second step (standardization).

1.3.2.1 Changing composition due to changing relative abundance

Different species will have different abundance trends over time, but a classifying algorithm will not be able to account for this. In some cases, the change can be extreme: for example, in the eastern CTS region, dogsharks and skates made up about 50% of biomass in deepwater trawls in the mid-1970s but under 5% in the mid-1990s[6]. There is no way to "tell" a standard classification algorithm about this, and the upshot will be muddling of abundance trends with targeting changes arising from misclassification.

Points to ponder: how will changing composition affect the classifier algorithm? If the classifier somehow allows "cluster mix" to change over time, how does it avoid confounding changes-in-abundance with changes-in-cluster-composition? Does a two-stage model really need to be embedded inside a full stock assessment (an "Integrated Assessment") to constrain the real changes in abundance?

[4]Both authors of this report are statisticians, so we are hardly "anti-statistics". But our experience is that, while statistical analysis can be enormously powerful if done right, (i) it can easily be done wrong so that the conclusions aren't trustworthy, and (ii) "a statistical model" is not a magical cure for ignorance about the underlying processes.

[5]That is, the shot-by-shot covariates will be the same for all species, but of course there will be differences in the the catches and the estimated standardization coefficients, including the coefficient(s) associated with the "targeting classifier".

[6]From research shots made with the same vessel and comparable gear: Graham et al., 2001, Table 1.

1.3.2.2 Circular arguments

The classification step is applied to shot-by-shot catch data from all species jointly, and then the resulting "targeting classifier" is used as a covariate in standardizing each species in turn. This means that the "Left Hand Side" (response) variable in the standardization— the catch of that species— also appears in the "Right Hand Side" explanatory covariates. This violates all standard statistical assumptions, and we are not aware of any standard statistical approach that allows for such a circular argument. As a simple example of the possible consequences, suppose you are trying to estimate a time trend, i.e. the effect of a covariate T on a response Y (e.g. raw CPUE), by regression; if you also include Y or some close transform of it as another covariate in the regression, then the statistical model will be very happy to let Y-on-the-right do a perfect job of explaining Y-on-the-left, and there will be nothing left for T to do.

The offense may not be entirely as heinous as it sounds, in that the inputs to the classifier are often normalized, so that only the *relative* catch of one species to another appears on the RHS; still, noise that increases the LHS will also increase the RHS. Also, if the transformation applied by the classifier is sufficiently brutal (e.g. discretizing to a small number of types), then the noise component of the RHS may be largely suppressed. But unless that can be demonstrated somehow in a particular application (which would, again, presumably require really good understanding of a limited range of fishing options), the circularity remains a statistical minefield.

For an example, suppose there are just two species A & B, which occur patchily but independently of each other, and that the fishery does not in fact target at all. A classifier working off proportional composition could decide there are three "types" of targeting: A-wanted, B-wanted, or mixed (both big, or both small). If A's abundance declines but B's does not, then the proportion of shots classified as "mixed" will increase at the expense of "A-wanted" shots; in the standardization, the estimated catchability of A in mixed shots will be lower than in A-wanted shots, so the abundance decline in A will be misinterpreted as a change in targeting. This harks back to the problems of "automatism" and changing composition, but by a somewhat different route.

Note that there is one potential diagnostic: in a two-step approach where the first step goes wrong in the way just illustrated, the trends in catch rate of A *within* "targeting type" will be different. We are unaware of this having been systematically checked in practice— perhaps because the classifier step is sometimes used for "shot selection", with standardization conducted only on shots deemed "targeted" to the species in question, so the comparison is never done. But even if the different-trends diagnostic was checked and found wanting, there is no obvious remedy within the two-step format.

There is one obvious cure for circularity, but it may be worse than the disease: if you are about to standardize species A, then just run the classifier step *without* including the catch of A. The downside is that the catch of A may be highly informative about targeting; the temptation of using this extra piece of information seems generally to get the better of more abstruse-seeming concerns about circularity.

Point to ponder: is adding more species guaranteed to dilute the circularity problem into irrelevance? (Note the catches of species which co-occur by chance, will not be statistically independent even within shots of a given targeting type.)

Point to ponder: is it a good/bad idea to use only one targeting-type of shot in standardization? If each targeting-type genuinely describes a consistent set of operational characteristics of a shot, why would you not expect to see parallel abundance indiex series from each type? (But see also "Aggregation" below).

1.3.2.3 Errors-in-variables

Economic and medical statistical practice is very familiar with the problem of "errors-in-variables" (EinV[7]), whereby a covariate used on the RHS of a statistical model is measured with error. The best-known consequence of ignoring EinV is attenuation bias; if you are trying to study the effect of covariate (RHS) X on a response (LHS) Y, but you are forced to use a noisy measurement $\tilde{X}$ instead of the true X, then a standard regression (or GLM) will systematically *underestimate* the effect of X; the estimated coefficient of X will be shrunk towards zero. While this theory can be extended to categorical variables, the algebra and conclusions are more messy. Attentuation still occurs, but its effect is dependent on more conditions: such as the type of misclassification.

In settings with multiple covariates, some of which are measured exactly while others aren't, the implications of ignoring EinV are just as bad but more complicated. The presence of covariates-with-error (conventionally called X before error and W as measured, i.e. with error included) leads to bias in the estimated effects of the covariates-without-error (Z), usually attenuation (towards zero) but sometimes exaggeration depending on the relationship between X and Z. For two-step MS CPUE, one can think of W as the classifier output (with X being the "true type" of a shot), and the without-error covariate Z as year. Our goal would be to estimate the time trend in abundance, i.e. the effect of Z on Y, which in general will be biased if a non-EinV approach is used. [The exception is if X and Z are uncorrelated— i.e. that targeting has not changed, which would rather undermine the purpose of looking at it in the first place.][8]

Even though the EinV literature is sometimes forbidding, it is undeniably vast, so it is disconcerting not to see it referenced in this area of fisheries. A good technical introduction is a set of lecture notes (Carroll, 2011), which on p37 has a formula for bias very relevant to the MSCPUE setup above. Note that the circularity issue is an additional complication not covered in the lecture notes; the simple example in the previous section illustrates how circularity can lead to an extreme form of attenuation bias.

Even though EinV appears to be a useful framework for considering two-step MSCPUE models, in particular for highlighting some pitfalls, the solutions most often in econometrics or medical statistics (instrumental variables etc.; see Carroll et al., 2006) do not seem particularly useful for MSCPUE, in particular because they do not easily accommodate the "non-automatic insights" that we feel are crucial to any successful treatment of MSCPUE. Our own model does not follow the two-step paradigm (and the reasons should be clear by now), so we did not follow up further on EinV.

Point to ponder: when considering the likely performance of any proposed two-step approach to MSCPUE, how would the approach fit within the EinV canon?

[7] Also known as "measurement error". Both terms are unfortunate, since of course the LHS (response) of the model is also a variable and usually measured with error (or noise), but that is already allowed for standard statistical models. "Measurement-errors-in-covariates" would be better, since covariates by definition only appear on the RHS of the model, but here we have stuck with the commonest nomenclature, which comes from medical statistics, where the response (e.g. life or death) is unambiguous, albeit subject to random variability. .

[8] Winker et al., 2013 include simulation results, which at first sight do not show evidence of attenuation. This is puzzling, since the two-step setup in that paper should certainly be susceptible to EinV-style attenuation and/or exaggeration. We conjecture that the reason lies in the way the simulations are organized, and the results reported: each simulation corresponds to a randomly-chosen scenario where positive and negative trends are equally and independently likely for each species, so overall bias would be zero however bad the EinV phenomena (because attenuation/exaggeration are always towards/away-from zero, irrespective of sign). To explore this, the results would need to be categorized by the signs of the trends in the simulation; we have suggested this to the author.

1.3.3 Aggregate indices

Behind any attempt to construct a single stand-alone CPUE index for a species— whether as a "single species" or as part of MSCPUE — is the notion that the index should be proportional to abundance (usually, summed over some set of age classes). One easy way for this to fail, is if there are shifts in depth preference with age[9]. The "depth effect" in a standardization will then partly reflect the species' age composition, as well as any genuine variation in individual catchability with depth. Changes-over-time in age composition will then violate the assumption of proportionality[10]. The age breakdown in the catch may well be used elsewhere in the stock assessment, but is still out-of-step with a CPUE-based abundance index.

There is an additional consequence for MSCPUE, because (for example) targeting type may also vary with depth[11]. Since no two species have the same intrinsic depth preferences, there may be differential trends in species' CPUE even *within* a targeting type (if we could somehow know the latter exactly). This could be a particular problem for non-two-step models (like ours) that expect to see similar trends (but at different levels) for each species across targeting types. Having said that, if targeting types really do represent different operational characteristics that remain fixed over time, then the assumption seems reasonable, at least *within* age-groups as well as within species.

Unlike some of the issues above, this problem at least has an in-principle solution: don't over-aggregate. If "the index" is broken into several separate indices, for example by depth and/or fish length, then the interpretational uncertainty about "depth effects" is retained until later in the assessment, where there is more hope of a resolution because of the constraints imposed by population dynamics[12]. Indeed, one might go further and argue that all the shot-by-shot data should be retained into the assessment, so that the assessment includes CPUE standardization; and then, given the existence of targeting, maybe all the species-level assessments should be done simultaneously. That would be the culmination of Integrated Assessment, as advocated e.g. in Maunder et al., 2006:

> "Integrated models use all available information, so they can be used to find inconsistencies in the data."

This may or may not be correct philosophically, but in our experience the seemingly-innocuous "so" can be a substantial overstatement; the daunting complexity of an Integrated Assessment can act as a massive deterrent to thorough model-and-data-checking, and the task can become inhumanly difficult— an aspect that is somewhat acknowledged in Maunder and Punt, 2013.

Point to ponder: does this mean that Integrated Assessment is required? In reality, how practically feasible would it be to diagnose the reliability of the MSCPUE (or single species) submodel within an Integrated Assessment?

[9]The argument applies not just to depth and age, but to any covariate (recorded or not) and individual characteristic where catchability may vary systematically with that characteristic within the population (i.e. the entity whose abundance we are hoping to measure), and where the composition of the population with that covariate may vary systematically over time. It's simpler to say "depth and age".

[10]Essentially the same point is made by Maunder et al., 2006, but in the context of community composition in a mixed-species CPUE, rather than age composition.

[11]This gets unavoidably confusing. For a fisher, "depth" is of course a *key* part of "targeting"; but in this report, where targeting covariates are by definition unrecorded, depth is *not* part of targeting.

[12]For the CTS, age-class standardisations are unlikely to be possible due to the small amount of age-stratified catch data (only available through observed data)

1.4 Rationale for our modelling approach

Because of the pitfalls just noted, we elected not to follow any existing approach to MSCPUE, but rather to develop a new one from scratch, the aims being statistical coherence, minimization of confounding, and allowance for economic drivers.

Our thinking was as follows:

- "Automatic" approaches to inferring targeting, i.e. those that rely on purely statistical/numerical criteria without a clear underlying set of assumptions about the fishery operations, are to be avoided because, as "Thought experiment #1" indicates, the same data may be equally well fitted by quite different scenarios about abundance change. Relying on *implicit* unexamined assumptions to somehow choose between these scenarios is too risky.
- A two-step approach (classify each shot, then use classifier output as a factor in standardization, and/or as a basis for shot selection) may work *if* the fishery operation is very well understood, with a small number of targeting options and very clear-cut outcomes [13]. But if not, then there are fundamental statistical problems ("Thought experiment #3") which are not addressed in existing approaches. In our view, the CTS does *not* fall into the "well-understood and clear-cut" class.
- While EinV approaches might alleviate some problems of the two-step approach, and are well-known in applied statistics, they do not seem naturally applicable to targeting, in particular being geared to an "automatic" worldview.
- A Bayesian approach could in principle alleviate both the "naivety" and "two-step" problems. It would need to include:
  - a prior distribution which describes the probabilities that a shot would be targeted in each possible way, given some characteristics of the shot (vessel, location, date, etc.) but not on what actually got caught;
  - a likelihood which describes the probability distribution of the species-by-species catches of that shot, under each of the ways the shot could have been targeted;
  - rather than ascribing a *definite* targeting type to a particular shot, the model can compute the posterior probabilities of the shot being "targeting type 1", "targeting type 2", etc., , given the characteristics of the shot *and* what was caught

There is nothing magic in the word "Bayesian" to guarantee success. In this case particularly, it is essential to structure the "prior" part appropriately. For example, one "inappropriate prior" would be a model that said: every year, the average prior probability (across vessels) of each "type" of shot can change, and those averages can vary freely from year-to-year. This model would fall straight into the "automatic" trap; a trend in abundance would be confounded with a trend in average prior probability. Changes in economic considerations are claimed to be the reason for systematic changes in targeting in the CTS, so it makes sense to build the prior accordingly.

An overview of the Bayesian model, which we propose for conducting a MS-CPUE standardisation in the CTS, is given in Chapter 1.4, and the equations are supplied in Appendix B. The models and results

[13] For example: the Indonesian longline fishery in the North Australian Basin catches either yellowfin tuna (with shallow sets) or bigeye / albacore tuna (with deep sets). Type of set can be inferred quite reliably from catch composition, even if hook depth is not recorded.

presented in this report should be seen as exploratory; the problem we addressed is difficult, and in order to produce any solution it was necessary to make more simplifying assumptions than one would want to use in a finished product. For further comment on what could and what should not be attempted, see Chapter 4.5.

Objectives

The primary objective of this project was exploratory development of a statistically rigorous framework (including software) for conducting CPUE analyses in the heavily multi-species setting of the CTS, specifically:

1. Develop mixture models for log-book data that deal appropriately with “zeros” and that incorporates auxiliary data (e.g. catch composition, market price, fine scale habitat and environmental data) to help account for targeting.
2. Use models from (1) to develop predictors of fishing effort type using only the log-book and auxiliary data.
3. Make our software available to fishery scientists involved in CPUE standardization.

Methodology

Here we describe the broad features of our model; equations are in Appendix B. Most focus is on the "choice submodel", i.e. the shot-level prior for targeting type, since that is the least-familiar and most-difficult part.

Some aspects of our implementation may seem unintuitive, especially concerning expected profit, but they are driven by the need to have a statistical model where parameters can actually be estimated— in particular, to keep the number of parameters as low as possible, and to avoid parameter redundancy/aliassing. Nevertheless, we think that the model should reflect *qualitatively* the phenomena described earlier.

- We split the CTS into four parts: summer/winter[1], and shallow/deep. The species mix differs greatly across the four combinations, and separate analyses were done for each. (There are some subtleties around splitting by depth in this way, discussed in Section 3.3.1.)

- We assumed there were a small number of discrete targeting types in each part, and that the available "pool" of possible types was fixed over time. It's not obvious from the questionnaires and discussions how many types there "should" be, nor indeed how sharp the distinctions between types really are[2]. However, from a statistical perspective, fixed discrete types are quite hard enough.
  We explored models with from 1-5 types, 1-type corresponding to a "standard standardization" (no targeting possible) and 5-types being the computational limit. Each fishery will have its own number of targetting types. However, for comparison, we note that 3-9 metiers have been suggested in other fisheries around the world with the higher numbers corresponding to more complex fisheries (numbers taken from He et al., 1997; Pelletier and Ferraris, 2000; Iriondo et al., 2010; Punzón et al., 2010; Punzón et al., 2011). In the analysis here, the types are in no way pre-specified, but are allowed to emerge from the statistical estimation step; this is not because of statistical puritanism, but simply because the CTS seems too complicated to make reliable *a priori* quantitative assumptions on what the "types" should be like.
  The main unrecorded operational factors that we have in mind as separating the types are:

  – depth profile of the shot (mostly, whether it is along or across contours). Only average depth is recorded in CTS logbooks, and the start and end locations of the shot are usually inadequate to infer much about the path between. See below for further remarks on depth;

  – fine-scale tracks relative to bottom features;

[1] Each a 3-month period, coinciding with the choices made in designing the Fishery-Independent Survey (Knuckey et al., 2013) mainly so that we could borrow that standardization model; catch rates are reputedly more stable from year-to-year during these months, being less susceptible to timing variations in the arrival of warm/cold conditions.

[2] This is partly a function of needing to separate the statistical problem (accounting for unrecorded covariates) from the on-the-ground targeting decision faced by fishers (not just the unrecorded covariates, but most importantly where and how deep).

  - gear adjustments[3].

- Each vessel has its own "preference", drawn from a pre-specified hyperprior distribution that describes how that vessel would differ from an average vessel in the shots it tends to make, *if* economic conditions made all types of shot equally lucrative across the fishery[4]. The hidden mechanism which we assume drives that preference, is the effective quota available to each boat (see next point). Preferences are allowed to change at 5-year intervals[5]. To keep an already-complex model from getting out of hand, the preference applies to *type* rather than to individual species.

- The "effective quota" for each vessel is not just the quota it has leased at the time a shot is made, but also the quota that it knows it *could* obtain inexpensively before the next reconciliation, via a complex network of informal arrangements. While data on currently-held quota is *potentially* available (AFMA holds the transaction records, but there are confidentiality requirements), the thing that matters is effective quota, which cannot be directly measured.

- For each shot, the prior probability that is of each of the possible types, given the vessel and the recorded covariates but not the actual catch data, is assumed to be affected by the fisher's expectation of profit for that shot under the different types. That depends on the expected catch of all species by type, weighted by both (i) current market prices (known to both the fisher and the model), and (ii) either the price that would have to be paid to lease matching quota if quota is not currently held, or if it is held then the price at which that quota might subsequently be leased out to others if not unused. Item (ii) is something like a "shadow price" which is known or guesstimated by the fisher, but is not directly available to the model. Instead, it is proxied by combining the vessel preference (which reflects quota holdings) with a measure of "TAC tightness", as explained next.

  - Several other factors strongly influence a fisher's expectation of likely profit when deciding how to fish: for example, water current strength, market knowledge of incoming catches, and short-term information on fish distribution. They are omitted from our model in the hope that they will average out over time but at the expense of potential covariance amongst shots. Including these data might— if the right data even exist— reduce noise, but should not substantially affect conclusions.
  - One extra factor that we have tried to include, is the price-depression (elasticity) effect of accumulated catch of each species in the rest of the trip[6]. The inability to separate trips in the logbook data adds another unwelcome layer of difficulty here, but in any case we suspect this has minor impact on the overall model because it will either tend to average out across a season, or be captured in the vessel preference.

- When the TAC for a species becomes "tighter" relative to the species' abundance, then it will become less attractive for the fleet as a whole to aim to catch that species, on average and over the course of an

[3]The logbooks record some information on gear, but do not capture all adjustments— and probably never could do so systematically.

[4]The fixed hyperprior is to prevent average preferences from drifting over time (to avoid the risk of confounding with abundance trends). However, that applies only to "ideal preferences". In practice, quota constraints and species price changes may drive the entire fleet away/towards certain types— but there is assumed to be no other reason why that would happen.

[5]The fixed-for-5-years approach is not ideal statistically; a more flexible hidden-Markov-type model would be neater, but was much too complicated for this project.

[6]This is a small fishery (by world standards) where an individual vessel's landings is a substantial part of the day's landings, perhaps even the days total for some species. In larger fisheries the price depression from a single boat is likely to be irrelevant compared to fleetwide landings. The model should be changed accordingly, if applied to those fisheries.

entire year[7]. The mechanism is that lease price will be higher, so the marginal revenue will be lower (on average across vessels). The model tries to accommodate this by assuming that the "fleetwise marginal price" for a species[8] is a multiple of the actual market price, where the multiplier will vary from 1 (if the TAC does not constrain the fishery at all) down to 0 if the TAC is zero (when no-one can make any money from catching that species). The multiplier is a nonlinear function of the ratio between the species' relative abundance that year— already the key term in the standardization— and its TAC, so no new variables or mechanisms are required, though extra parameters do need to be estimated.

- To compute the prior probabilities that a specific vessel will adopt each of the different possible targeting types for a specific shot, the model combines the "fleetwise marginal revenue"[9] with the vessel-specific preferences. Of course, not every shot will then correspond to the model's somewhat naive notion of the economically-best choice, so there is also a "rationality" or "perfection-of-knowledge" parameter to be estimated; larger values correspond to better ability to predict catches and forecast future quota prices, and/or to smaller effects of unrecorded *and* uncontrollable shot-specific factors such as weather conditions, known to the fishers but not to the model, which would diminish the model's accuracy. That last point is important: from the questionnaire and interview data, it is unsurprisingly clear that these unmeasured/uncontrollable short-term factors are often dominant in fishing decisions. But on the assumption that they will largely cancel out over the course of a fishing season (there is a strong incentive to use, or re-lease, held quota at *some* point within its 12-month lifespan; the decisions are only about when to do so), we feel reasonably justified in omitting them from the model.

- Having finally computed the prior probabilities of each targeting-type for each shot for each vessel for whatever set of parameters is currently being considered, "standardization" is conceptually carried out in the statistical setting of a "finite mixture model" (McLachlan and Peel, 2004), where the joint probability of the observed catches of all species (given recorded covariates such as trawl duration ("effort") and depth) is summed across the different possible targeting types, using their prior probabilities as weights. This comparatively straightforward step is easier to follow in an equation than in words (Section B.1).

- All parameters are estimated simultaneously from the joint likelihood, driven mostly by the catches per shot, but also using some characteristics of each shot which are indicative of that vessel's preferences. Estimation is completely separate for the summer/winter deep/shallow data sets.

- There are many operationally-important factors which are not included in the model, and which aren't statistically feasible to include: e.g., weather, current strength, water temperature, recent knowledge and feedback from other vessels. However, these omissions do not of themselves make the model "wrong" or useless[10], provided the factors left out are sources of *noise* (random variability over time) rather than systematic trend.

3.1 Depth and location

Shot depth is perhaps the single most important factor in determining what gets caught. As such, it should play a pivotal rôle in determining the targetting type of a shot. Also, it affects the catch rate of the shot. In

[7] Of course, some vessels may still retain enough quota that they do not need to shift effort elsewhere; the model attempts to capture that via vessel preferences, as explained *next*.

[8] IE the profit per kilo that would be made by an average boat catching that species at that time

[9] FMR = anticipated catch for the species and type, times fleetwise marginal price for the species, summed over species

[10] "All models are wrong, but some are useful": George Box, *Empirical model-building and response surfaces*, 1987

terms of our model, there are four key points about shot depth in the CTS:

1. The logbooks record only average depth, and there is no reliable way of determining whether a shot was along- or across-contours. A shot that runs down-slope from 50m to 400m with a mean depth of 250m will tend to catch a wider range of species— but less of any individual species— than a shot along the 250m contour.

2. The model has to learn about each vessel's preferences. Besides the actual catch composition of a shot, the depth chosen for it is an important piece of information; a hypothetical vessel that held mostly silver trevally quota would not waste time fishing deeper than 100m, so seeing a high number of shallow shots for a vessel is indicative of a type-preference that favour silver trevally (or other shallower-water species).

3. There is a dramatic change in species compositions between the shallowest (50m) and deepest (700m) average depths in our dataset. In principle, our model formulation ought to take this in its stride. In practice, though, we ran into numerical and interpretational problems trying to fit all species and all depths together[11]. The pragmatic resolution was to split the analysis completely by depth (at the 200m average depth line), using a different range of species in each. This discards some information that is potentially useful for inferring preferences, but was computationally unavoidable. The 200m isobath was chosen as it seemed to naturally occur as a break point for many of the species under consideration.

4. Decisions about the unrecorded covariates that drive "targeting type" in our constrained definition, are in practice entirely linked to decisions about recorded covariates, especially (average) shot depth; it would be senseless to rig the gear for an off-the-bottom species but then trawl at a depth where it does not occur. When computing expected revenue across different options for the shot, we therefore considered not just targeting type, but rather the combination of targeting type and average depth (split into a small number of sub-bands), given that the decision was already made to make a "basically shallow" or "basically deep" shot as just described. We treated the latitude of the shot[12] as a "given" (i.e. not another option like type or depth), and assumed that the choice would then be made about depth and type; this seems reasonable since the continental slope in Eastern Australia is very long North-South but just a few kilometers wide, especially within each of the two depth bands we used.

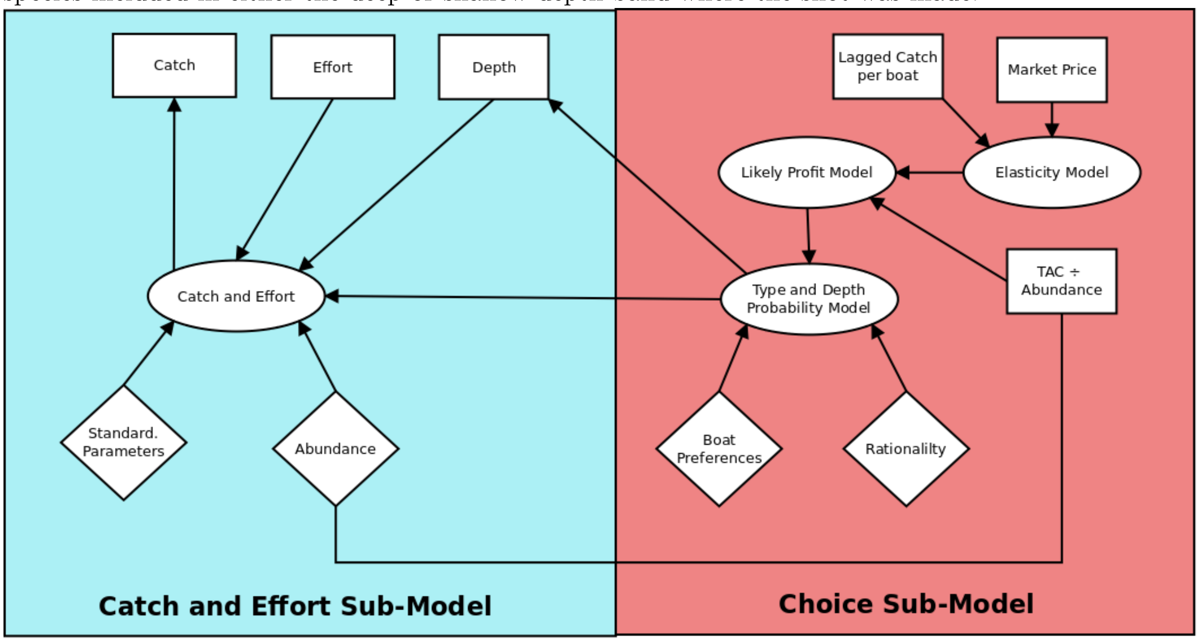
5. In reality, the species mix also varies (much more gradually) with latitude, so choice of location is also somewhat informative about vessel preferences. However, location is really a trip-level decision influenced by many factors (travel time, fuel costs) as well as anticipated catches, and within a trip the position along the coast of one shot is quite constrained by the latitude of the previous shot (not the case with depth), so it would not be easy to extract the information. Thus we did not try to include location in the targeting prior[13], though it is included in the standardization part of the model because most species do exhibit systematic spatial variation in catch rates (Knuckey et al., 2013).

[11] We do have various theories post hoc to explain why problems might occur, but they are hard to summarize and have little bearing on the outcome, since we chose to bypass the problems using a depth split.

[12] Actually, its position when projected onto an axis that roughly tracks a shallow contour running close to the coastline, running mostly North-South; see Knuckey et al., 2013.

[13] Individual-vessel-based fleet simulation models that do try to model choice of fishing grounds have been developed for some fisheries worldwide, and may exist for the CTS. However, while such models may yield interesting qualitative insights, we see no prospect that they could (or should) be ever be used for *quantitative inference* such as CPUE standardization, basically because the range of implicit or explicit modelling options is too high-dimensional for well-founded statistical approaches to work, whether Bayesian or otherwise.

Figure 3.1.1: Components of our model for CPUE standardization with targeting
“Effort” comprises the covariates of any normal standardization. “Catch” is multi-species across all quota species included in either the deep or shallow depth-band where the shot was made.



3.2 Catch and Effort Sub-Model

The catch and effort sub-model closely resembles a GLMM, such as would be used in a modern single-species CPUE standardisation; the only difference is targetting type is a latent variable rather than a measured covariate. We used the same basic formulation as in the FIS (Peel et al., 2013)— a log link with a Tweedie error distribution, and keeping the FIS estimates of smoothing and Tweedie-power parameters— since this gave satisfactory model performance. The covariates we used were: trawl duration (as an offset), average depth (continuous), time-of-day (factor), year (factor), and position along coast (continuous). Penalised regression splines (see Wood, 2006) were used for depth and position along coast. Both had 9 basis functions and penalty taken from Peel et al., 2013. Time of day was entered as a factor, levels taken from the logbooks themselves, with 3 levels (XXXXXXXXXFIND OUT LEVELSXXXXXX). All terms were added as main-effects only (no interactions). Trawl duration was considered to be an offset only. In particular, we assume that the length of trawl has a multiplicative effect on catch— a double-lengthed trawl is expected to catch twice the amount as a single-lengthed trawl. This is a common assumption in this fishery (e.g. Peel et al., 2011). See Appendix B for details and see Jørgensen, 1997; Smyth, 1996; Foster and Bravington, 2013 for background on Tweedie GLMs and see Appendix Dfor more model details of how the catch and effort model was applied to the CTS.

3.3 Data

3.3.1 Logbook Data

Our analysis covered 1994-2008, a time of major changes in the fishery (and major reductions in CPUE for many species): from the early days of ITQs, through widespread adoption of GPS and the tightening of TACs, to a substantial reduction in fleet-size at the 2006 buyout ("structural adjustment").

The logbook data contain numerous errors and outliers that can be removed prior to analyses. To filter the data, we followed Darbyshire et al. (2008) and Klaer and Smith (2008), specifically excluding shots that: (1) lack key variables, such as catch, effort, depth; or (2) have a reported position on land; or (3) are unusually short or long; or (4) are out of the usual depth range. We also excluded all shots below 700m since these are almost exclusively directed at one particular species, orange roughy— one case where targeting is easily determined. See Table C.1.1 for a complete description of the filtering rules.

Since this is an exploratory analysis, we simplified the problem slightly by subsetting in space and time. In particular, we considered only the south-east coast between Sydney and Hobart excluding Bass Strait (CTS zones 10/20/30) and only two seasons: winter (July–September; 48,864 shots) and summer (January–March; 45,612 shots). The seasons match those used for the Fishery Independent Survey (Knuckey et al., 2013); the idea is that sea conditions (especially temperature) are more stable from year-to-year in those seasons than in the spring and autumn "shoulders" when the East Australia current arrives/departs. The periods we used are also well within the TAC-year, which starts in May. This approach makes the assumption that there has been no shift in relative effort into, or out of, these seasons.

Of the many species recorded, we restricted attention to the nine that make up the bulk of the landings and revenue: tiger flathead, spotted warehou, pink ling, blue grenadier, jackass morwong, redfish, john dory, silver trevally and mirror dory.

The number of shots made by different vessels is highly skewed. The bulk of shots come from a small number of vessels; since each vessel included in the model entails the estimation of extra parameters, to keep computational burden down we considered only the 39 vessels with more than 1000 reported shots, which account for 74% of shots overall. Full details of the data, and how it was obtained by subsetting, is given in Appendix C.

3.3.2 Environmental Data

The region of interest is a narrow strip 8-15km wide but over 1500km long. The coast and the shelf-break, run roughly but not exactly North-South. It is more biologically meaningful to model species' spatial distribution not via (latitude, longitude) but rather by (depth,position-along-coast) where the latter is a projection of location onto a smooth curved axis that roughly follows the 50m contour line; it is also more statistically parsimonious, since the depth-effect on a species is roughly orthogonal to the coast-effect. We used the same axis as in the Fishery Independent Survey Peel et al. (2013) and Knuckey et al. (2013), which also meant we could borrow the estimates of within- and between-year spatial variability estimated in designing that survey, for use in our own "standardization". This variable is illustrated in Figure 3.3.1.

3.3.3 Market Data

The price data for 1994-2008 was purchased from Sydney fish market, the sole wholesaler market in Sydney. Data from the other main markets, in Melbourne, proved harder to obtain, so for this exploratory exercise we did not pursue it— it remains unclear whether those data still exist.

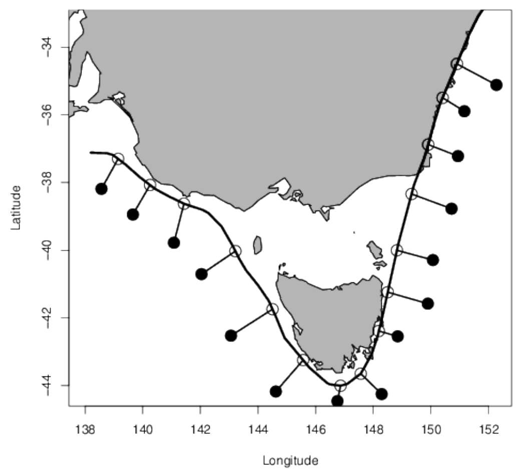


Figure 3.3.1: Illustration of the position-along-coast variable. It is the curve that (roughly) tracks a bathymetry contour around the southern part of the continent. Only the south-east coast data were considered in this analysis (so the curve west of Tasmania is irrelevant).

The data record the species and amount of fish sold in each transaction, and the price paid. Even the commonest species are not sold every day, so there are gaps in the species-level price data. From these data, we were also able to reconstruct the "running cumulative catch" over a few days (typically 3-5 days are required for incoming landings to clear the market), for use in the elasticity model.

3.3.4 TACs and ITQs

Annual TACs and total catches are available on the AFMA website. Vessel catches are constrained to a proportion of the TAC, via the vessel's quota-holding. Although AFMA keeps a record of quota transactions and in principle could reconstruct the quota holding of any vessel at any time, we did not try to obtain those data, partly because of confidentiality issues but mostly because we did not think it would be much use for modelling, as explained earlier in this chapter.

Results

4.1 Elasticity from Market Data

The first step in the analysis was to determine the effect of landing an amount of fish into the market, to help later on with predicting a vessel's likely shot decision based on *its* catches so far that trip. See Section 3.3.3 for a description of the market data and Appendix C.2 for a full description of the model and its parameters. Briefly, the expected price of a species is modelled as a function of the amount of that species currently in the market via a single estimated "elasticity" parameter, with additional trend terms to absorb long-term effects.

For illustration, we show results for tiger flathead and redfish (Figure 4.1.1), which represent the extremes among the 8 species we eventually focussed on. Tiger flathead has little seasonal variation, and only a gradual increase in price. In contrast, redfish have a strong seasonal pattern, and a more dramatic long-term increase in price. Most species, but not all, showed some long-term increase in price. This is due, in part, to inflation but it is also confounded with other time-varying drivers (such as consumer trends). For the purposes of this analysis, the cause is immaterial and it is the effect that is important.

The estimated parameters are given in Table 4.1.1. For this part of the modelling, we included all species with market data, not just the species we later focussed on.

The signs on all estimated coefficients are negative (except for roughskin dogfish), showing that there is some degree of saturation is possible for all species. The price response seems least for orange roughy and gemfish. In the choice sub-model, the expected price is used. After all, fishers can only make decisions about what they expect to catch. This involves using the predicted market price and also the elasticity parameters in the prior for shot type (the choice sub-model).

In principle, the elasticity analysis could be extended to use multispecies landings as covariates (not just the species in question), since there is presumably some degree of substitution between some species. There is no fundamental statistical obstacle to doing so, just extra work e.g. in model selection, so we did not pursue it for this project.

4.2 Targeting Analysis

For brevity, we present here the results from just the two "winter" subsets (deep and shallow); the same details for "summer" are in Appendix E.

For each subset, we experimented with varying the number of targeting types between one (i.e. our version of a classical CPUE standardization, without targeting) and five (the computational limit). Even after fitting the five models, it is not obvious which number of types is best; in the statistical field of "finite mixture models", of which our model is an elaborate extension, there is still no agreed and reliable criterion

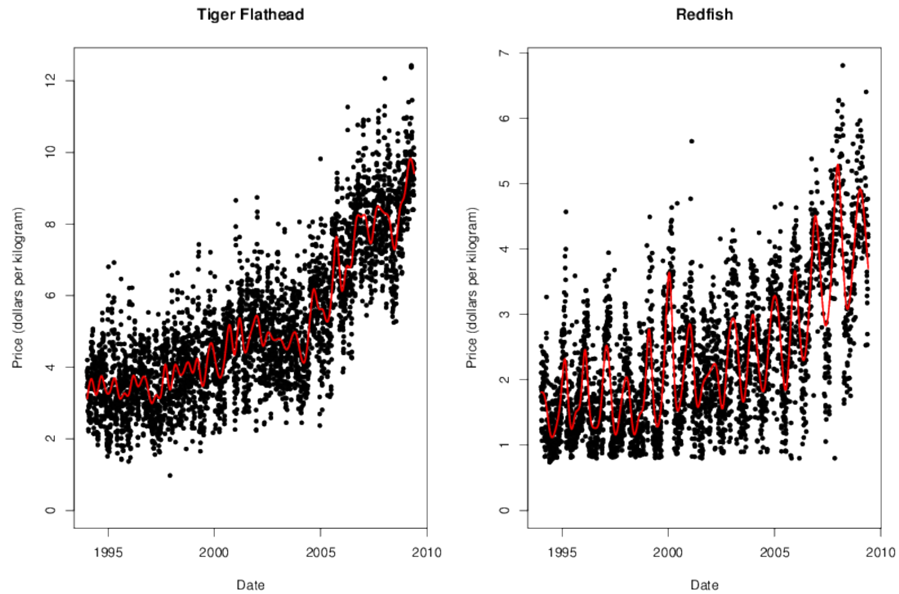


Figure 4.1.1: Observed (black dots) and expected (red line) price for tiger flathead and redfish. The Statistical model is defined in Appendix C.2.

for choosing how many categories on purely statistical grounds, and our model poses a number of extra challanges than the simpler regression-like settings (Hui et al., in press). Here we have presented results for the number-of-types that yielded the highest "rationality parameter" estimates, i.e. where the model's assessment of prior economic incentives across types for each shot is best aligned with the apparent type of the shot (see Appendix B for equation and definition). However, we make no claim that this particular choice is "the truth", and later we present the overall implications for CPUE for all 5 variants.

The estimated rationality parameters for the different subsets and numbers-of-types are given in Table 4.2.1. There is no obvious scale for interpreting these numbers, but they are comparable across numbers-of-types. Negative values would be nonsensical (implying that fishers tend to avoid the shot-type that the model thinks would earn the most revenue), but all the estimates are comfortably positive.

The final piece of evidence reported is how the targeting types might have changed over time. This is inferred from calculating the yearly sum of the (posterior) probabilities for each targeting type, where the sum is over each of the observed shots. The posterior probabilities are denoted by $\Pr(g_i|d_i, \{C_{si}\})$ where g_i is the targeting type for shot i, d_i is the shot's measured depth and $\{C_{si}\}$ is the set of species catches. The posterior probabilities are defined as

$$\begin{aligned}\Pr(g_i|d_i, \{C_{si}\}) &= \frac{\Pr(\{C_{si}\}, d_i, g_i)}{\sum_{g_i'=1}^{G} \Pr(\{C_{si}\}, d_i, g_i')} \\ &= \frac{\Pr(\{C_{si}\}|d_i, g_i)\Pr(d_i, g_i)}{\sum_{g_i'=1}^{G} \Pr(\{C_{si}\}|d_i, g_i')\Pr(d_i, g_i')}\end{aligned} \tag{4.2.1}$$

where all probabilities are calculated during the fitting process and are a result of the model. Note that $\Pr(d_i, g_i)$ is the joint probability of the depth and targetting type (see Section 3.1).

	Elasticity	Which Data Set
Bigeye Ocean Perch	-0.14	-
Blueeye Trevalla	-0.05	-
Blue Grenadier	-0.10	Deep
Blue Warehou	-0.06	-
Gemfish	-0.01	-
Jackass Morwong	-0.14	Shallow & Deep
John Dory	-0.07	Shallow
Ling	-0.08	Deep
Mirror Dory	-0.12	Shallow & Deep
Orange Roughy	-0.01	-
Redfish	-0.17	Shallow & Deep
Roughskin Dogfish	0.00	-
Royal Red Prawn	-0.04	-
School Whiting	-0.08	-
Silver Trevally	-0.12	Shallow
Spikey Oreo	-0.05	-
Spotted Warehou	-0.13	Shallow & Deep
Tiger Flathead	-0.15	Shallow & Deep

Table 4.1.1: Estimated elasticity coefficients β_s for each species s. The more negative the value, the more "saturatable" the market for that species. A sudden doubling of the amount on the market would reduce the price of that species by a factor 2^{β_s} , a tripling by 3^{β_s}, etc.

4.3 Winter Deep

The four-group model has the highest rationality parameter. Estimates of typewise catchability effects are in Table 4.3.1, and time-trends in Figure 4.3.1. The variation between types in implied catch rates is very large for some species, particularly tiger flathead and blue grenadier— which are the two highest-value species. The numbers in the table should not be over-interpreted, though, because decisions about "type" are in practice made jointly with decisions about depth, so in practice one type might never be made at the depth typical for another type, so that the implied catch ratios would never actually be seen.

A qualitative interpretation of the types and their trends might be as follows:

Type 1: Redfish, morwong, flathead, ling favoured; spotted warehou unlikely— mixed bag? Proportion stable over time.

Type 2: Blue grenadier mainly; unlikely to catch morwong or redfish, and definitely not flathead. Decreasing proportion over time.

Type 3: Flathead; poor for ling, no chance of catching grenadier: increasing proportion over time.

Type 4: Spotted warehou and some blue grenadier; stable proportion.

The inferred relative abundance series for the targeting models[1] is presented in Figure 4.3.2, along with two

[1]Including the 1-type model, which is the "*any color you like as long as it's black*" version of targeting (attrib. Henry Ford).

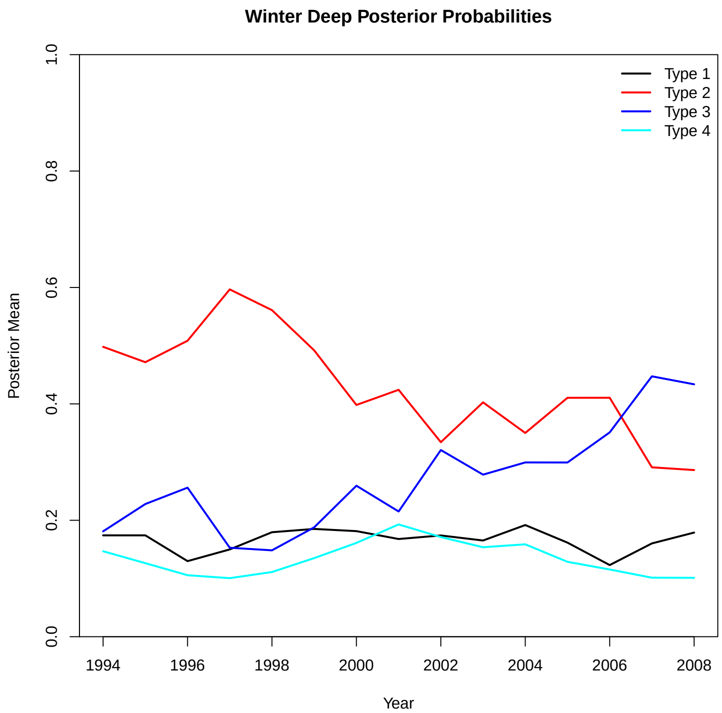


Figure 4.3.1: Proportion of **Winter Deep** shots performed under each of the targeting types over time. Calculated as the mean of the posterior probabilites of the observed shots.

	G=1	G=2	G=3	G=4	G=5
Winter Deep	-	0.93	1.16	1.45	1.38
Winter Shallow	-	13.98	15.36	3.06	2.69
Summer Deep	-	6.80	4.04	5.85	4.30
Summer Shallow	-	9.51	17.58	5.72	5.18

Table 4.2.1: Estimated "rationality/perfection-of-knowledge" parameters κ for different data subsets and assumed numbers of types. The larger the value, the more likely a shot is to correspond to the model's notion of best likely revenue.

	Blue Grenadier	Jackass Morwong	Ling	Mirror Dory	Redfish	Spotted Warehou	Tiger Flathead
Type 1	2.12	1.12	0.58	-0.24	1.60	-1.50	2.18
Type 2	3.22	-1.61	0.46	0.36	-1.51	-0.34	-5.86
Type 3	-7.80	0.26	-1.07	-0.10	-0.31	-0.53	2.70
Type 4	2.46	0.23	0.02	-0.02	0.21	2.36	0.98

Table 4.3.1: **Winter deep** catchability parameters (γ_{sg} in the model). Positive values mean higher expected catches for a species than an "average shot", all else being equal— which, as the text points out, it may not be. Estimates are on a natural-log scale so, for any species, a difference of 3 units between two types corresponds to a ratio of $e^3 \approx 20$ in expected catches. Columns sum to one to enforce estimability (sum-to-zero constraints).

simple reference unstandardized series[2]. We specifically do not include the actual standardisation results as these are based on a different subset of the log-book data. Therefore, we *a priori* do not expect it to agree with our results. For the most part, there is little difference between the series, although our tentatively-preferred 4-types model (solid black) is perhaps less spiky than the others, especially the unstandardized cases and especially at the start of some series. The 4-types model is generally very similar to the equivalent standardized-with-no-targeting model (dotted turquoise).

Of course, if a statistical model is told to find 4 types, or 40, then that is exactly what it will come up with, regardless of whether all the types are clearly discriminated in the data. One post hoc check on overfitting is to look at the shot-by-shot posterior probabilities; if the model is "inventing" types, then these will tend to be scattered between 0 and 1 for each type, but if the types are meaningful then one would expect the posterior probabilities to be clustered near the ends. Figure 4.3.3 shows histograms by type (to avoid a 3D or 4D display). There is no objective measure of goodness here— fisheries shot-by-shot data are notoriously variable, and a shot which catches very little of anything is going to be difficult to classify even for an "accurate" model— but the results are certainly not uniform. The spikes near zero and the scatter elsewhere imply that, while the model is confident at deciding that a shot does *not* belong to some type(s), it is less clear about precisely which type it does belong to. This may suggest overfitting (too many types); similar plots for the other three data subsets (where in all cases the preferred model has fewer than 4 types) show much sharper clustering. Other measures of prior-posterior concordance could be developed, but given the intrinsic noisiness of catch-per-shot data, we are not sure that much insight would be gained.

[2]The "nozeros" reference case (dashed red) is a log-normal GLM with the same covariates as the targetting model; the "geometric" case (dashed blue) is a geometric mean CPUE per year (obtained by fitting a log-normal GLM with only year as a covariate).

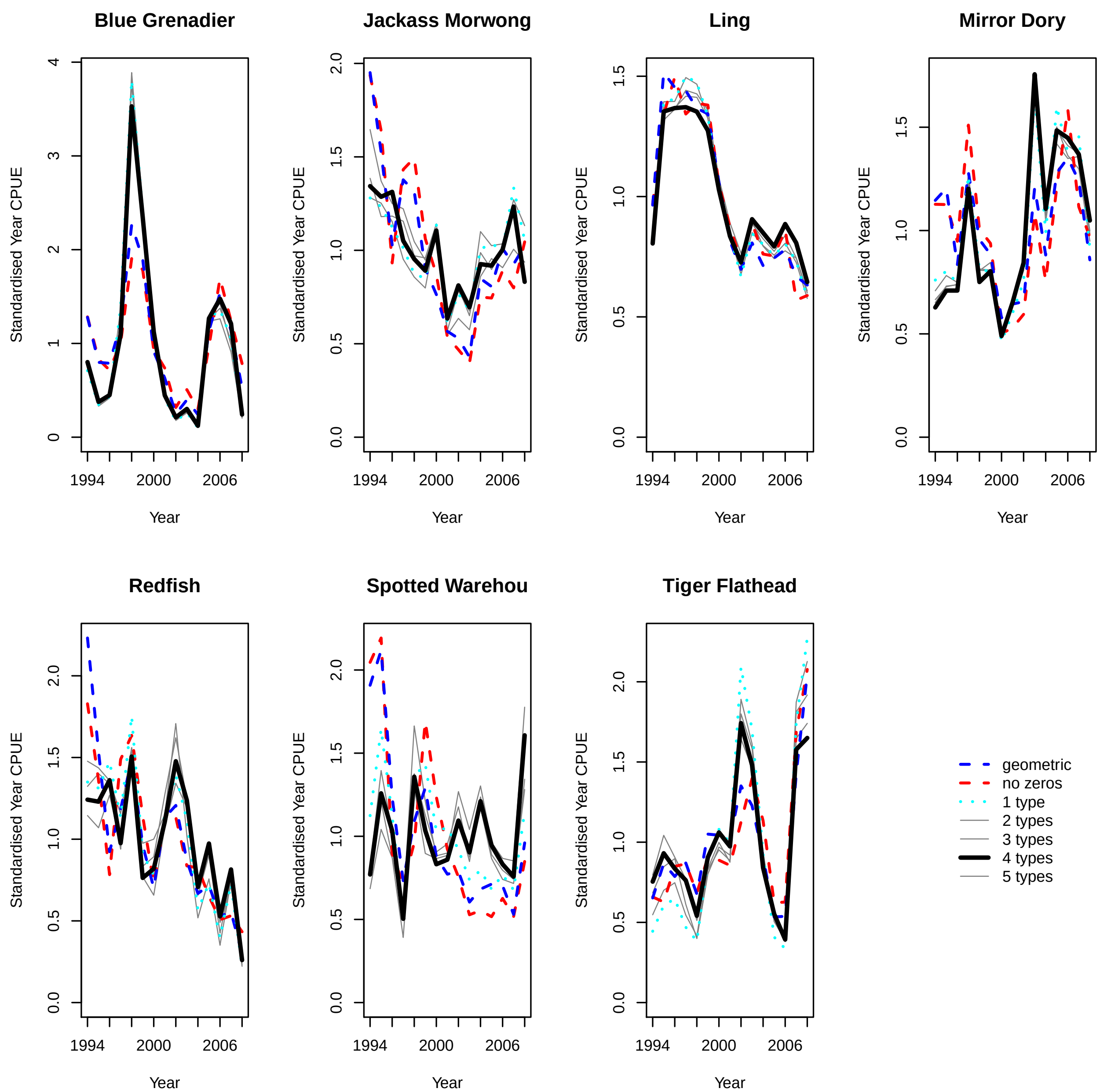


Figure 4.3.2: **Winter Deep** Relative abundance series estimated by the different targeting models. Dotted-green is our version of a standardized CPUE without targeting (a Tweedie GLM with covariates); solid-black is the 4-type model, which had the highest "rationality estimate". Different numbers of targetting type have been intnetionally given the same line type. This reduces graph clutter whilst retaining the sense of variation. Dashed lines are CPUE series that remove zeros, without covariates (blue dash) and with covariates (red dash). Both series without zeros were obtained using a log-normal GLM.

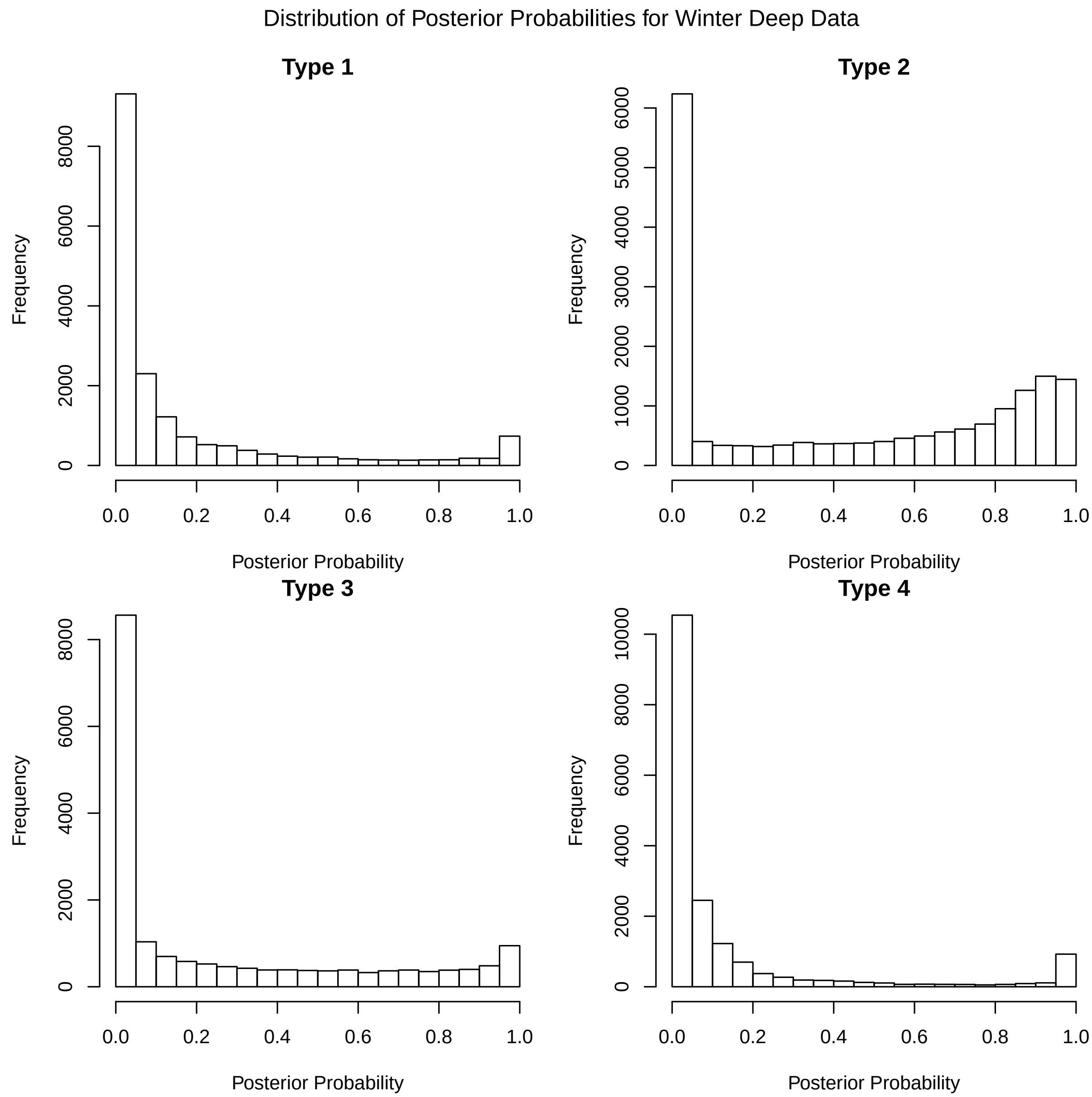


Figure 4.3.3: Distribution of posterior probabilities of each shot in the **winter deep** data set. The distribution is taken over all the observed shots.

With 4 types in this model, vessel preferences are hard to summarize, so consideration is deferred to the winter-shallow results next, where the chosen model has fewer types.

4.4 Winter Shallow

Here the 3-types model has “highest rationality”. Typewise catchabilities and trends are shown in Table 4.4.1 and Figure 4.4.1.

Type 1: Spotted warehou, redfish, and mirror dory favoured; the worst way to catch flathead. Declining trend, and now uncommon.

Type 2: Silver trevally especially, intermediate for most other species. Declining trend.

Type 3: Flathead good, but the worst type for all other species. Strong increasing trend.

The estimated effect of type on catch rates varies considerably between species: there is almost no difference for john dory, but around 100-fold differences for spotted warehou, redfish, and mirror dory. As for winter-deep subset, the “all else being equal” caveat applies. The inferred effect for tiger flathead is modest— about 50% better in Type 3 than Type 1— but important because flathead drives a high proportion of revenue. Flathead catches also vary less from shot-to-shot than many other species’, so will generally show less noise in estimated parameters, including type-specific catchability effects; when comparing best-to-worst catchabilities in these tables, estimation noise will tend to systematically inflate the range.

	Jackass Morwong	John Dory	Mirror Dory	Redfish	Silver Trevally	Spotted Warehou	Tiger Flathead
Type 1	0.40	-0.01	2.32	2.18	-0.69	2.73	-0.19
Type 2	0.75	0.07	-0.41	-0.36	1.64	-0.11	0.03
Type 3	-1.15	-0.06	-1.91	-1.82	-0.95	-2.62	0.16

Table 4.4.1: **Winter shallow** catchability parameters for the four targeting type model, on a log scale; see Table 4.3.1 for explanation.

Figure 4.4.2 shows the estimated winter-shallow abundance trends under the various targeting models (i.e. different numbers of types). Flathead CPUE is remarkably insensitive to the model used. For the other species, standardization clearly does have an effect (the reference dashed lines look different to the rest), but beyond that the number of targeting types assumed seems to make little overall difference for most species except silver trevally. Model choice does seem to have an effect at the start of the series, where there is a sharp peak in abundance for morwong, mirror dory, redfish, and spotted warehou according to models without targeting (dotted turquoise), but not for the 3-types model. The start of the series is, unsurprisingly, exactly when the 3-types model estimates that the fastest changes in targeting occurred (Figure 4.4.1).

Shotwise posterior probability histograms for winter-shallow are shown in Figure 4.4.3. Clustering is clear, and much sharper than for winter-deep; this is not surprising since in effect only two types are seen over most of the time period. With only 3 types in the winter-shallow model, it is feasible to examine vessel preferences.We tried to set the model up so that there should, so to speak, be no “average intrinsic preference for any specific preference”, *unless* economic drivers dictate otherwise— the whole point being, of course, that economic drivers *do* dictate otherwise. If vessel preferences turned out nevertheless to be clustered towards particular types, that might suggest the model was missing something of significance. Figure 4.4.4A shows that they are not, at least for winter-shallow subset (also true for summer-shallow: Appendix E).

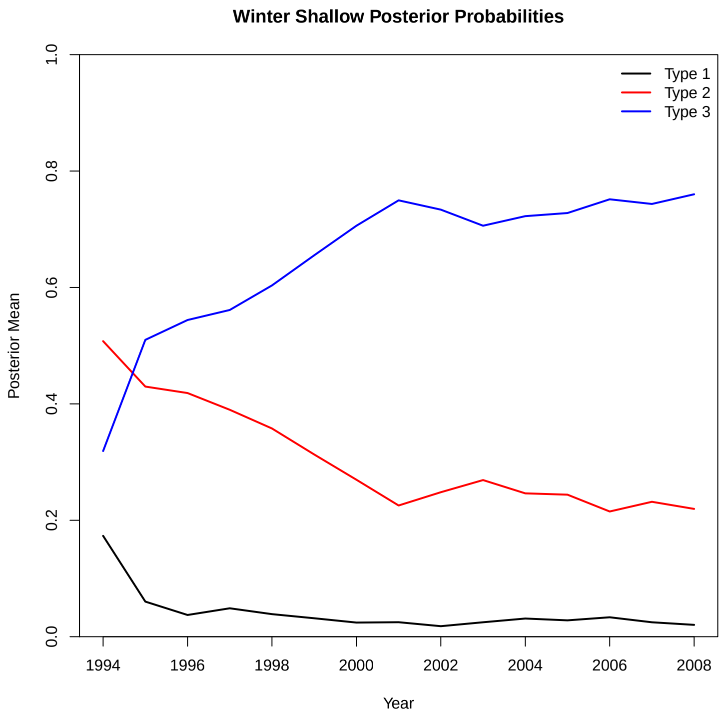


Figure 4.4.1: Proportion of **Winter Shallow** shots performed under each of the targeting types over time. Calculated as the mean of the posterior probabilites of the observed shots.

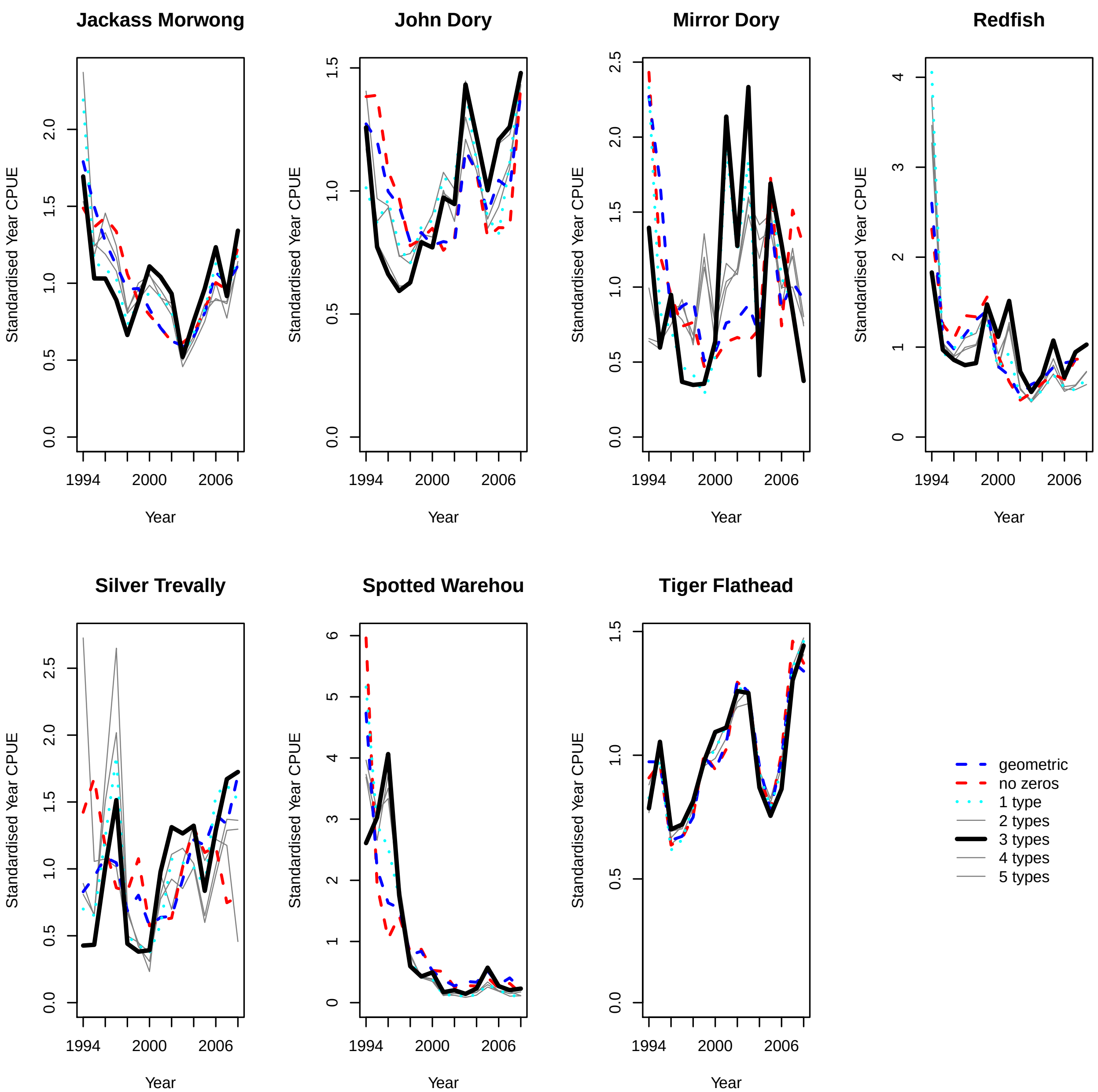


Figure 4.4.2: **Winter Shallow** CPUE series under the different targeting models. See caption of Figure 4.3.2.

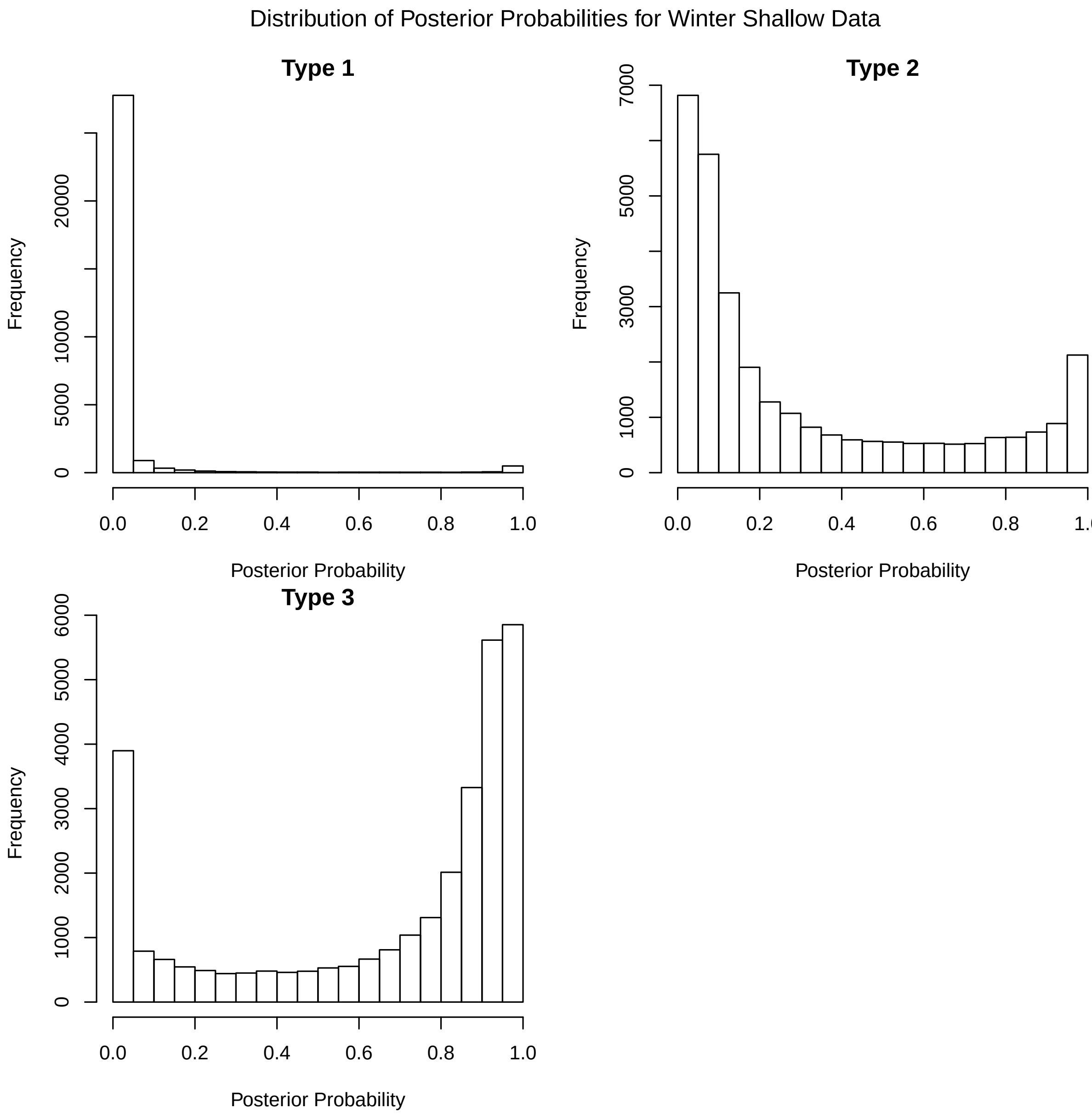


Figure 4.4.3: Distribution of posterior probabilities for the **winter shallow** data. The distribution is taken over all the observed shots.

Vessel preferences can change over time, and the 5-year fixed blocks we were forced to use, for computational reasons, might in principle be too long to adequately reflect this. To check, we plotted the preference-traces for each boat from one block to the next (Figure 4.4.4B-F, split arbitrarily into sets of boats for clarity). The impression is that within-vessel changes over 5 years (arrow-chains) are typically smaller than between-vessel differences (different colours). This suggests that the 5-year-block approximation is not seriously distorting our results. (If this model ever was to be taken further, though, a more sophisticated model for changing preferences would be necessary; but the statistical and computational challenges should not be underestimated.)

4.5 Concluding notes on results

The summer shallow and deep results are shown in Appendix E; in both cases, the preferred model had in effect only 2 types[3]. Of course there are differences in detail between the four winter/summer shallow/deep subsets, but our overall impressions are quite similar: the inferred types do imply quite large differences in expected catch rates of some species due to targeting, but there is rather little effect overall on inferred abundance trends when targeting is incorporated into standardization.

We did not see any particularly concerning diagnostics, although the loose clustering in the winter-deep results may suggest that 4 types is too many. In particular, the results we have presented do not contain any "rubbish" types— where one type has lower catchability for every species than another. However, we did find some instances in other models, at least among the 4-types and 5-types models, and in some of the preliminary results that we presented at RAG meetings [4]. Rubbish types are a conundrum, because it seems hard to imagine why a fisher might *choose* to make a shot one way when they would be likely to catch more of *everything* by doing it another way. (Admittedly, there might be some explanation in terms of exploratory shots and not "burning" quota when prices are low.) For diagnostic purposes, we view rubbish types as a warning sign: they are statistically tempting because of the improvement in statistical fit for shots which didn't catch much, but since rubbish shots do not (presumably) make sense in terms of the model's narrow view of economics, they *may* signify internal tension in the estimation. Since the inferred abundance series seem fairly insensitive to the number of targeting types assumed even though some of the models imply rubbish types, the latter is probably not a major concern for our conclusions.

If our exploratory model was ever adapted for serious use on CPUE standardization, a number of issues would require substantial further thought and work. A by-no-means exhaustive list would be:

- model selection (number of types);
- model selection (covariates included into the model, functional form and interactions);
- what to do if the selected model is applied one year and suddenly decides that a rubbish type is needed— or more generally if the inferred types change dramatically from one run to the next;
- quantify uncertainty, especially how to estimate covariances of the time-series (easy if just one model is selected, but hard if results need to take model uncertainty/selection into account)
- getting price data from the Melbourne markets as well as the Sydney market.

[3] The summer-shallow preferred model technically has 3 types, but the 3rd is uncommon and— unlike winter-shallow— its proportion is stable over time.

[4] Resource Assessment Group: the industry and science forum where stock assessments and research outcomes are discussed.

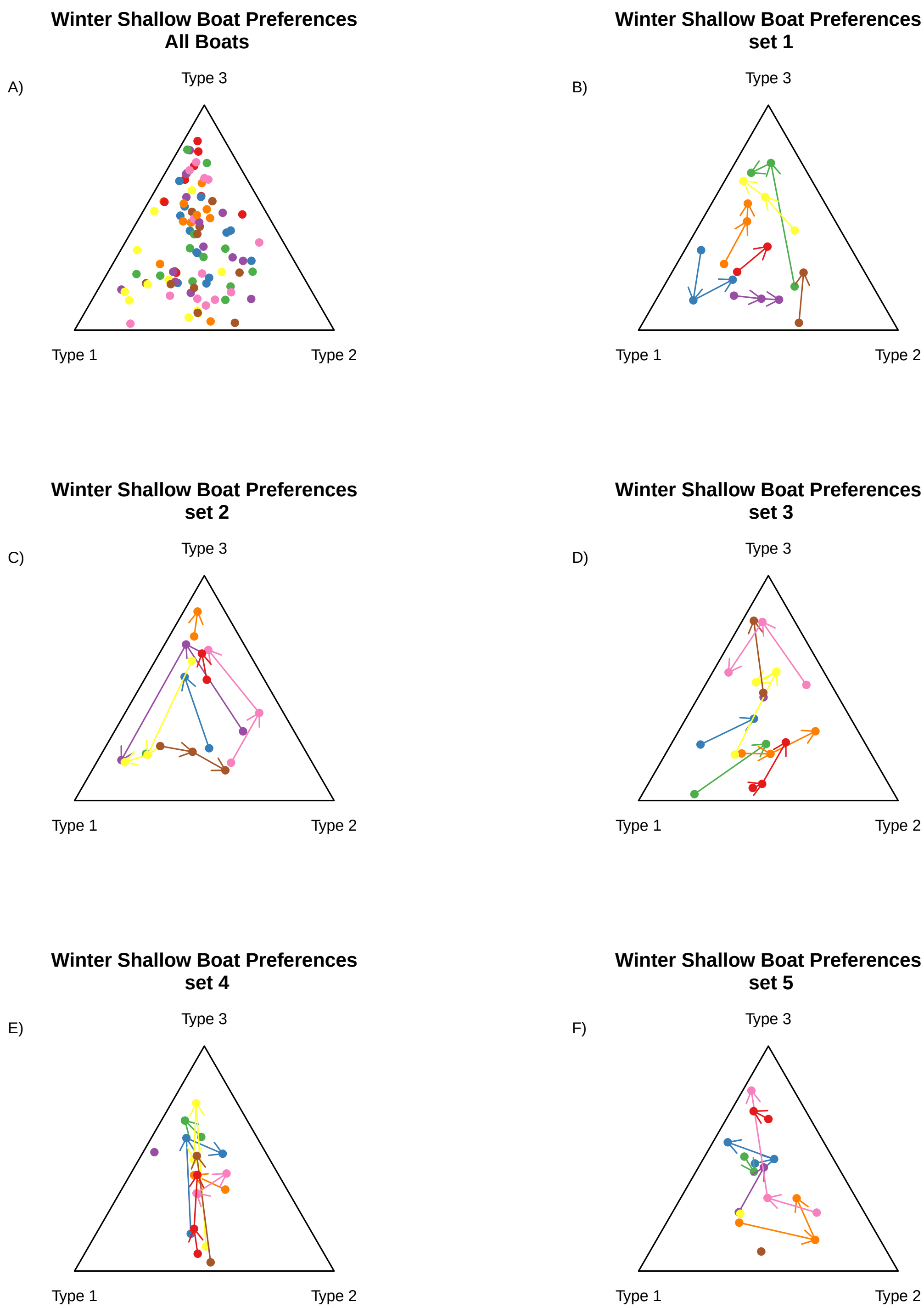


Figure 4.4.4: **Winter Shallow** vessel preferences. Vessels are allowed to "choose" a new preference every 5 year block (block boundaries are staggered between vessels). Each vessel in a 5 year period (a dot) has a preference for each of the three types (the vertices of the triangle). Panel A shows all vessels-and-blocks, coloured by vessel.. Panels B–F each show a subset of vessels (colours) and how their preferences changed between blocks (arrows).

Conclusions and Discussion

Targeting[1] is certainly possible in the CTS, and there is good reason to expect that it has changed over the last 20 years in response to economic drivers and to changing stock abundances. The likely effect on CPUE for individual species is not necessarily clear *a priori*, though, given that CPUE (after current standardization is applied) is averaged over shots. Since there is no direct measure of targeting in the CTS logbook data, the usual approaches to CPUE standardization are not able to account for targeting.

We derived a new statistical approach, with the two aims of avoiding the statistical pitfalls, and reflecting the economic drivers. The model includes a "typical CPUE standardization", plus a prior distribution on likely targeting "type" for each shot. That prior is driven mainly by quota tightness and individual vessel preferences, which are estimated within the model and assumed to reflect the amount of quota readily available to that vessel. There is no need to treat zero catches specially, since the usual implicit reason for excluding them (targeting) is explictly allowed for in our model. Some compromises were required to make statistical estimation possible, but in the end we were able to fit the model with several different assumed numbers of "targeting types" (2-5). There are no glaring inconsistencies in the results, taken at face value: the main inferred types (i.e. the mix of species) do not seem ridiculous, the inferred choice of shot-type does appear to be broadly consistent with inferred economic drivers[2], and there is no obvious "arbitrary trend" in preferences (i.e. aside from trends driven by economics) that could be confounded with real trends in abundance.

Model results did show evidence of changed targeting over time, though only in winter. However, there was little effect on inferred abundance trends; the with-targeting results seem a little less spiky, but for most species the overall picture was largely unchanged. In other words, the standardized CPUE trend *with* targeting was pretty similar to *without*, which is how CPUE is currently handled. The good news is that the results are not sensitive to choosing one particular model (e.g. number of targeting types); the either-good-or-bad-news is that targeting changes do not appear to have affected overall CPUE much.

5.1 How reliable are these results?

To incorporate economic drivers, it seems to us that a complicated statistical model is unavoidable, even though we kept it as simple as we could, so there is a lot that could go wrong. Below we consider the two main potential problem areas: model details and assumptions; and data limitations.

[1]taken in this report to mean unrecorded factors, under control of fishers, that influence catches (see Section 1.2).
[2]Insofar as the estimated "rationality parameters" are positive, as one would expect.

5.1.1 Model details

Any statistical model involves some approximations and assumptions; however, these are not important as long as the resultant "model error" is small relative either to data noise, or to the consequences arising from the way the model is used. This project is exploratory (there was no existing statistical approach to allowing for economic incentives in targeting) and the aim is to investigate whether a useful targeting model *could* be developed, not to develop a perfect model straightaway— a much bigger task.

Here we have concentrated on the least-familiar aspects of our model; of course, the usual statistical considerations also apply, just as in single-species standardization (e.g. choosing reasonable functional forms, appropriate mean-variance relationships, avoidance of over-parametrization).

5.1.1.1 Targeting can be adequately represented by a few discrete types

Finite-mixture models, including our model in this project, can be quite challenging to make fit, and there is a distinct upper limit on the number of types that can be included: in this project, certainly no more than 5 in each of the two "deepish" and "shallowish" bands. It is not obvious in advance how many "types" there should be. As noted in the Introduction, there are good reasons why fishers do not think about "targeting" in exactly the same way that we have been forced to (by data constraints), and the questionnaire responses did not consistently describe any one clear set of "types". However, the comment was made that "*the iconic species drive the decision*", and only 3-6 iconic species were proposed, fewer in a single depth band. Since we have split our analysis by broad depth band anyway, our suspicion is that 5 types per band should be more than adequate.

A harder question is to what extent "discrete types" actually exist; individual shots last for several hours, vessels get some feedback on fish densities during a shot and can somewhat change direction if they wish, and increasing incentives for "mixed bag" shots may blur some of the types. Our impression, though, is that the posterior distributions of shot-type (e.g. Figure 4.4.3) do show some evidence of concentration (clearly not driven just by vessel preferences, since the latter are quite diffuse: Figure 4.4.4), rather than the more uniform scatter that might be expected if the model was just "making things up because it was told to".

Failure to represent the variety of real types properly within the model would presumably weaken the inferred effects of economic drivers. However, even if that really is a problem, we do not see any realistic prospect of building an even-more-complex statistical model that could accommodate it.

5.1.1.2 Handle for "shadow price"

We have used a convenient but rather indirect approach to reflect anticipated lease costs for unheld quote and/or "opportunity costs" for leasing held quota to others. Some approximation is inevitable; it does not seem remotely possible to develop a statistically-estimable model that really matches a fisher's process of revenue-forecasting, and we can never expect to have all the relevant data. Also, we have necessarily concentrated on modelling "average" behaviour across a fishing year; this is (according to the questionnaire responses) quite different from how fishers decide about shots, where shorter-term factors (immediate issues of fishing conditions and market state, and certainly no further than the next quota reconciliation point) are necessarily dominant.

Nevertheless, we expect our approach to show qualitatively correct behaviour: the economic incentive to pursue particular species varies from vessel to vessel, and tighter TACs on a species will tend to deter the fleet as a whole from making shots likely to catch that species. For simplicity, though, we did make two possibly-avoidable assumptions that may somewhat distort our results, both in the "TAC tightness" part of the model:

- that the proportion of "unavailable" TAC is constant over time;
- that TAC tightness is independent of "available effort" in the fleet.

The point about quota tightness is that, if there was just one boat left in the fishery, then it wouldn't matter what the TAC was. No one boat could catch the whole TAC whatever type of shots it made, so there would be no deterrent effect of "tightness" on individual shots. There was a major reduction in CTS fishing capacity in the 2006 buyout, which may have reduced the deterrent effect of tightness; it is not obvious what the implications for CPUE would be.

With both of those assumptions, there might be enough aggregate data available to produce a better model, although at the cost of doing considerably more development work.

5.1.1.3 Vessel fishing-power effects are omitted

CPUE standardizations often include a "vessel effect" on overall fishing power. Our model formulation is more parsimonious than a set of single species standardisations, in the sense that our vessel effects are concentrated into a small number of types rather than one for each of a large number of species. Much of the vessel effects in single-species analysis may reflect targeting, rather than fishing power. We could have included one more vessel-specific parameter to describe "fishing power", but we doubt that conclusions would change much.

5.1.1.4 Vessel preferences: fixed for 5yr blocks, and controlled by prior

Vessels can change their preferred species over time. Constraining each vessel to have the same preference for the whole 12-year period would risk blurring the economic drivers. However, allowing each vessel to change its preference every year would leave an enormous number of parameters to be estimated, and might hamper the ability of the model to discern each set of preferences. As so often in statistical modelling, there is a bias-variance trade-off.

Our choice of 5yr blocks reflects computational limitations— we had originally planned to model vessel preference as a modified random walk, whereby the fleetwise rate-of-change in preferences would also be estimated, but that proved too complicated. The 5yr "fixed effect" blocks were the most that could be handled. However, it seems from Figure 4.4.4B–F that the within-vessel changes from one 5yr block to the next are generally smaller than the between-vessel differences, suggesting that some consistency is likely over 5yr timescales. We do not see this as a major problem in an exploratory model, although a more flexible approach would be needed if this model were ever fully developed.

To avoid confounding between real changes in abundance and any "accidental" (unmotivated by economics) trend in preferences, we used a fixed prior over time to shrink[3] the vessel preferences. Again, our original intention was to estimate this prior's variance (an Empirical Bayes approach) but this proved too complicated, so a somewhat arbitrary choice was made. This still leaves some potential for confounding, although it is not obvious in Figure 4.4.4This is one area where the model could be improved, perhaps by adding fleetwise constraints so that not the prior does not treat vessels as independent (on the basis that non-latent quota has to go somewhere).

[3]In the technical statistical sense, i.e. that in the absence of information to the contrary, most boats will look somewhat like a composite-average-boat.

5.1.1.5 Age-aggregation and size-depth effects

Some CTS species show a strong size-depth relationship, e.g. spotted warehou (Bax and Williams, 2000). The Introduction notes the problems this can cause for CPUE interpretation, in a single-species as well as multi-species setting (where the consequences might be slightly worse. Further, market price per kilo in the SESSF often depends on size as well as species; we just used an average price to avoid excessive complication. In principle, targeting decisions might take size-depth relationship into account. Somehow this seems unlikely to be the "straw that broke the model's back", but it is worth remembering that this problem exists for any CPUE analysis that aggregates across size/age where it should not. The question raised in the Introduction remains open: whether/how far CPUE standardization should be embedded within an Integrated Assessment, rather than run as a preliminary step and then used as a stand-alone input into an assessment. But the complexity of any multi-species approach to that is quite daunting.

5.1.1.6 Omitted covariates

The problem with statistical modelling in fisheries is often not lack of data *per se*, but lack of *informative* data. In CPUE standardization, there can be an inexhaustible supply of covariates that "might help a bit". Omitting them is not usually a problem, except insofar as the inferred trends in abundance become more precise when more of the variance is explained by covariates (and subject to the usual statistical proviso, that overfitted models can have worse predictive power). While there could no doubt be some improvements in the model if extra covariates were used, we would be surprised[4] to see much change in the qualitative conclusions.

One covariate worthy of specific mention, is seafloor habitat data. We originally intended to explore using the mapping data in Williams et al., 2010, because choice of habitat for trawling is sometimes an operational decision that is informative about targeting. In the end we did not pursue this, for four reasons:

1. Time constraints on model development, and the complexity of the mapping data (many different classifications at different hierarchical levels), which would create major complications in model-selection.

2. There is no fine-scale track data for shots to reflect e.g. proximity to outcrops— a known factor in species distribution and an operational consideration in targeting.

3. The mapping data is still too coarse for such features anyway (finest level recorded is 1km^2).

4. Large-scale spatial effects on species distribution are absorbed anyway in the spatial distribution part of the standardization.

5.1.1.7 Drivers really are economic

The economic considerations that we have attempted to model are not necessarily the only reason why fishers choose to do one thing rather than another. Anecdotally at least, social considerations also used to be important in the CTS, for example demonstrating fishing prowess by landing a lot of fish even if it means reduced revenue from a quota, or in terms of information-sharing about recent and future shots. It is unlikely that such factors have vanished altogether under the icy glare of economics. Also, the quota/revenue considerations that were explained to us are very intricate; it must take time to learn how to "play the game", and for the game itself to evolve, especially given the major structural changes in the fishery in still-recent

[4]That is: among covariates that we know to be available now. If different data were collected in future, this might not remain true; see next section.

years. We see no prospect of building a data-driven statistical model for CPUE standardization in the CTS that can magically account for such things.

5.1.1.8 Effort creep

We did not try to model changes in fishing power, e.g. due to GPS or other technology, or conceivably to increasing average skipper expertise in the face of fleet reductions and the march of time. It is possible that such effort creep may not affect all species' catchability in the same way; for example, when really accurate location data are available, species that stay close to outcrops may be easier to target, whereas there may be less effect for more widely-distributed species. However, we suspect that any multi-species aspects of effort creep are dwarfed by the general problem it raises for primarily-CPUE-driven management; it is a serious matter for the CTS that is not accounted for in the assessment or in the Harvest Control Rules, and we are unaware of any effective remedy while CPUE remains so prominent.

5.1.2 Would more work or better data change the story?

Many statistical models could be improved with more work, but diminishing-returns quickly set in because of intrinsic limits in the information content of a dataset about a quantity of interest. Ours could certainly be improved, and we have noted a couple of specific examples above. In fact, to bring the model to "production quality" (i.e. for routine use as a standardization tool), substantially more work would be required, including at a minimum: preference prior variance estimation, a better model for changing preferences, allowance for vessel effects on catch rates not just preferences, and much more on model selection.

This begs the question: if not the model, then what about the data? Allowing for targeting in standardization means allowing for unrecorded but fisher-controlled covariates, so in principle targeting might be handled easily if the right covariates could be recorded. In practice, though, there are two problems. First, extra covariates in future would not help with retrospective interpretation of CPUE. Historical trends in CPUE during a period of fast depletion in the 1990s-early-2000s (noting that effort has now been cut back substantially) will continue to be a major driving force in the kind of Harvest Control Rules currently used to manage the SESSF; given the Rules, considerable importance attaches to the question of whether biomass is assessed *right now* (2015) to be above or below 20% of pre-exploitation levels, and no future CPUE data is going to change that assessment by much. Second, it is not obvious that the right data could be collected— though this is speculative, and somewhat beyond the scope of this report. In principle, we suspect that modelling might be somewhat improved (perhaps to the point of not having to worry about unrecorded covariates) by knowing more about the entire track of a shot than just average depth and start/end locations, since the across/along-contour distinction is important for expected catch. However, it is not obvious how any standardization could blend two halves of a dataset with and without that information. Beyond that aspect of the data, it is not clear whether the ideal information could be recorded in any consistent way, e.g. points of detail about gear settings. And one proposed extension, of asking fishers to record "intended targeting" in advance of a shot— presumably in terms of what species they *hoped* to catch— is not necessarily going to translate consistently to operational details about what they actually *did*; the latter is what is really needed for better standardization. We note that asking fishers for targetting information has, in theorey, already been tried. The resulting data has been universally disregarded as unreliable (M Haddon *pers. comm.*).

Recommendations, Implications, and Further Development

Targeting is a real phenomenon in the CTS that has changed over time, which may have led to changes in CPUE independent of abundance, but that cannot be measured directly. We have shown that it is (just about) possible to construct a statistically-defensible CPUE standardization that (in principle) makes allowance for changes in targeting driven by (some) economic factors. However, we do not recommend that our model be adopted for CPUE standardization in the CTS, even after adaptation. It is much more complicated than existing approaches, likely less robust than an off-the-shelf statistical model for "standard standardization" such as a GLMM, and would require substantial further work to get the details right; and the answers we got are not different enough to make the extra work worthwhile. Any minor improvements would be dwarfed by the elephant-in-the-CPUE-room— effort creep— which is unaccounted both in our model and in "standard standardization".

There are two possible conclusions from our results: either long-term changes in targeting have not much affected overall CPUE (after standardization) in which case there is not much point in taking the work further; or the model we developed was inadequate to handle targeting properly. We certainly acknowledge that the latter is possible; we tried our best but the data, our time, and our abilities are all limited. Even though the economic representation in the model is crude, from a statistical and computational point of view the model is already so complicated that it is approaching the limits of feasibility. Would it be worth the effort to try developing a better model? While one can never rule out the possibility that a better model might somehow change the answer, our feeling is that, if *any* model could ever pick up a strong effect of targeting on CPUE trend, then our exploratory attempt should have shown a much larger effect than it did. This is not to say that there really is no strong effect— but if there is, then we see little chance of ever capturing it in a model. One reviewer of this report suggested that it might be worthwhile unravelling why our model didn't substantially alter trends. This has academic merit, but it must be based on: 1) pretending that we have data that is not available for the CTS, or 2) simplifying the model by removing sources of information. There is little practical utility in pretending to have data that is not available— it will be applicable to any real standardisation. Also, removing information can only have the effect of blurring the targetting effect even more. Hence, we cannot recommend this line of research for the CTS.

There is always a temptation to try squeezing more blood out of the CPUE stone, especially when pickings are so slim in the rest of the stock-assessment data. However, we are skeptical of claims for simpler approaches to standardization-with-targeting in a complex fishery like the CTS (e.g. omitting economic drivers) because of the statistical pitfalls listed near the end of the Introduction: in particular, the danger of reliance on purely statistical classification criteria without an underlying mechanism, and the various problems with a two-step classify-then-standardize approach. Therefore, we recommend that any further proposal to improve CPUE standardization in the CTS by addressing targeting should *first* demonstrate— by argument and/or simulation— why it will avoid those pitfalls, and (to avoid duplication of substantial effort) why it might be expected to yield different results from ours. With respect to the pitfalls, we have found simple simulations to be more informative than complicated ones, and we have provided a small R package that might be useful for exploration in simple and generic (non-CTS) settings. The code for fitting our own targeting model to CTS data is also available as an R (and C) package.

As to the implications for assessment and management of the SESSF: we do not see much prospect of ever getting much more out of CPUE in the CTS than is currently achieved by standardization[1]. If nothing else, perhaps this project will dispel any illusions on that score. Worldwide, in fisheries which have the opportunity, CPUE has been abandoned as the index-of-choice because of all the interpretational issues for which there is no guarantee that any statistical fix will work, and because of the serious implications of making the wrong interpretation or failing to span the range of possible interpretations. There are certainly problems with CTS CPUE, especially effort creep which to the best of our knowledge there is no realistic hope of quantifying in the CTS; and the problems are not mitigated by the current SESSF HCRs, which are considerably affected by the estimated depletion (i.e. current biomass relative to unexploited) which for many species is largely determined by CPUE trends[2] (see Punt et al., 2002 for a CTS specific study of this issue). When the top-tier HCRs have been simulation-tested in generic settings for the impact of unmodelled effort creep, they have unsurprisingly not done well (Andre Punt, *pers.comm.*)..

One of the motivations behind this study was the lengthy and unresolved debate about whether zero-catch shots of a species should be included when doing single-species CPUE standardization in the CTS (and indeed, when was zero really zero?). This "much ado about nothing" is a shot selection question that is really just the tip of the iceberg of the wider issue of targeting. One of our remits in this project was to find a coherent way to resolve the question and, in a technical sense, we have succeeded; our multispecies standardization uses all the shot data and zeros should definitely be included. However, given that we are not recommending routine use of our model, the question of what to do about zeros remains. This is somewhat tied up with the minutiae of single-species standardization[3], which are beyond the scope of this project, but the following general points apply:

- There is no purely statistical criterion (such as a goodness-of-fit criterion) that can be relied on to make the decision;
- The operational details of the CTS are too complex, and the logbook data too incomplete, to reliably decide shot-by-shot on "operational" grounds, except in special case
- The only safe method is to check what might happen in management terms when zeros are/are not included (but should not/should have been), i.e. by MSE. Intuitively, omitting zeros when they are meaningful could lead to inadvertent overexploitation, whereas including them when they are not meaningful (e.g. because of changed targeting) could lead to overly conservative management. The SESSF Harvest Strategy is presumably set up to give more emphasis to one or other of these situations, so simulations may not actually be needed, or very simple ones may suffice.

[1]The question of "what to do about zeros" is still important and worth reconsidering in single-species CPUE standardization, though it is really linked to targeting.

[2]Albeit via a stock assessment at least for the "top tier" HCRs, but the assessments are currently still at the mercy of CPUE.

[3]For example, in a well-chosen GLMM, it should not matter if shots outside of a species' depth range are included, since they will be "standardized out" anyway. The more difficult question concerns shots that fail to catch a species inside its usual depth range, season, etc., where failure-to-catch may or may not be indicative of an operational decision about targeting.

The desirability of an alternative, more reliable, abundance index for the SESSF is obvious. The Fishery Independent Survey (Knuckey et al., 2013) should at least be immune to effort creep, although it has higher short-term variability than CPUE because of the limited number of shots; we note that there has been no investigation of how often or how thoroughly the FIS would need to be done in order to provide adequate inputs for the Harvest Control Rules. We also note that there has been a recent success with a completely new fishery-independent abundance estimation method— Close-Kin Mark-Recapture— for Southern Bluefin Tuna (Bravington et al., 2014), and that the CKMR approach is being developed into a long-term monitoring index for that species. CKMR is not totally free, but it is far cheaper than surveys[4], and does not suffer from the problems of CPUE; although it has not yet been scoped specifically for any SESSF species (except school shark, where a new project began in 2015) and the specifics would be quite different to SBT, there is no obvious reason why it would not work for many of the key SESSF species.

[4] And is most cost-effective for the highest-value most-abundant species.

Extension and Adoption

Several presentations of our results were made to SESSF RAGs and other meetings. Most presented only intermediate results, but a complete overview is available (CSIRO MSCPUE workshop, Hobart, 2014). These presentations to RAG meetings is intenionally the primary source of extension as it is the scientists and managers who have the prime interest in this work.

In addition, some of the early results have been presented to scientific audiences through a journal article (Foster and Bravington, 2013) and through invitations to conferences/workshops (Foster and Bravington, 2010a; Foster and Bravington, 2010b).

Project Materials Developed

fishMod_0.25.tar.gz An R package to estimate and summarise a number of different types of single-species zero-inclusive models. The statistical and numerical methods employed in this package are necessary pre-cursors to the full multi-species standardization. This package is available from CRAN.

fishModFULL_0.2.tar.gz An R package to estimate the full targetting-inclusive multi-species CPUE model.

mscpue_1.0.7.tar.gz An R package to illustrate some of the pitfalls, highlighted in Section , of various approaches to multi-species CPUE standardisation

Single Species Paper (Foster and Bravington, 2013) A formal description of the methods in the single species (with zeros) package fishMod_0.25.tar.gz.

Appendices

Appendix A

Questionnaire on Targeting Practices

1. Say you're already out on a trip and are thinking about where/how to shoot next.
   (a) Which species are important in your decisions?
   (b) What would you actually do, to go after the species you decide on? E.G. depth, depth profile thru shot; duration; net height; proximity to reefs...
2. Before shooting, how good an idea do you have about likely prices for different species?
   (a) How fast do prices fluctuate?
3. What influences your decision about where/how to shoot? E.G. current quota holdings, lease-ability
4. Is the total catch from your own trip enough to affect market price in its own right?
   (a) If so, do you have to go for mixed bags, in order not to swamp the market with one species?
   (b) How much does supply of one species affect the price of that and other species?
5. If you don't already hold quota for a species, how hard is it to lease retrospectively?
   (a) Does this change through the quota year (e.g. late, when there's not much TAC left for a species)?
   (b) How good an idea would you have beforehand about the likely cost?
6. If you hold unused quota for a species late in the year, do you deliberately lease it out, rather than try to use it yourself?

Appendix B

Model Details

Details of the model are now given. The different parts of the model, and the intuition of how they fit together, are given in Chapter 1.4.

B.1 Sub-model for catch-and-effort data

The basic model used to describe catch and effort data is a simple multiplicative model – close to a generalised linear model (GLM) with a log-link. For the i^{th} trawl and the s^{th} species, the model is (ignoring random components, for now)

$$C_{si} = q_{si} E_i B_{si},$$

where C_{si} is a catch record, q_{is} is the catchability, E_i is the effort and B_{si} is the biomass for the species. The catchability, q_{is} is the probability of catching the species if it is actually at the location of trawling. The biomass B_{si} is the biomass available to the shot, not the population biomass. This model can be linearised using a log transformation

$$\begin{aligned} \log C_{si} &= \log q_{si} + \log E_i + \log B_{si} && \text{(B.1.1)} \\ &= \log E_i + \log\left(q_{si} B_{si}\right) \\ &= \log E_i + \log\left(B^*_{si}\right) \text{ say,} && \text{(B.1.2)} \end{aligned}$$

where we still have not taken any consideration of any random variation. While illustrative, the first form of the linearised model (B.1.1) is not practical as the biomass and catchability are not separable, not without replicate trawls anyway. This leaves us with the final form for practical uses, (B.1.2).

Random variation can be incorporated using a log link function, and an explicit model for the mean response is defined at the same time. That is,

$$\begin{aligned} \log \mathbb{E}\left(C_{si}\right) &= \log E_i + \log\left(B^*_{si}\right) \\ &= h_s(\boldsymbol{x}_i) + \log E_i, \text{ say,} && \text{(B.1.3)} \end{aligned}$$

where $h_s(\cdot)$ is a function of shot-specific covariates $\boldsymbol{x}_i$. The variables in $\boldsymbol{x}_i$ are things that are thought to be related to catch, either through altering biomass B_{si}, catchability q_{si}, or both. We choose the function $h_s(\cdot)$ to be a penalised smoothing spline, a generalised additive model (e.g. Wood, 2006). As such, the flexible function can be expressed as a linear combination, that is $h_s(\cdot) = \boldsymbol{x}_i^{*\top}\boldsymbol{\tau}_s$ where $\boldsymbol{\tau}_s = (\boldsymbol{\tau}_{s1}, \boldsymbol{\tau}_{s0})$ is a vector of parameters for spline terms ($\boldsymbol{\tau}_{s1}$), and other terms ($\boldsymbol{\tau}_{s0}$). The vector $\boldsymbol{x}_i^*$ is a transformed version (basis expanded) of the original observations $\boldsymbol{x}_i$.

We prefer to model C_{si} directly, rather than catch-per-unit-effort (CPUE; C_{si}/E_i), as CPUE varies with both catch and effort – not just catch. Hence, its relationship with catch can be obscured. Typical examples of variable to include in $\boldsymbol{x}_i$ are: year, depth, spatial location, and time of day. The model is completed with the specification of the distribution of departures from the expectation, as defined in (B.1.3). This is not a simple choice as the chance of not catching a species is quite high, and these zeros can cause complications for standard exponential family distributions. We favour the use of Tweedie departures as these allow for zeros in a principled manner (see Smyth, 1996; Candy, 2004; Shono, 2008; Tascheria et al., 2010; Foster and Bravington, 2013).

To account for targeting in the catch and effort analyses we assume a small number, G, of different targeting practices and allow an extra targeting-specific term in the model, depending on the targeting type. For targeting type $g \in [1, \dots, G]$ the model is

$$\log \mathbb{E}\left(C_{si} | \text{targetting type } g\right) = h_s(\boldsymbol{x}_i) + \gamma_{sg} + \log E_i, \tag{B.1.4}$$

where γ_{sg} can be interpreted as an adjustment to the catchability of the species (the intercept), although another interpretation is an adjustment to the effort. This is just a GLM when the targeting type is known. Since targeting type is not recorded however, we introduce a latent (random) variable $\boldsymbol{z}_i$. This latent $G \times 1$ vector has zeros everywhere except for a single one at the g^{th} location. The conditional model then becomes

$$\log \mathbb{E}\left(C_{si} | \boldsymbol{z}_i\right) = h_s(\boldsymbol{x}_i) + \boldsymbol{z}_i^\top \boldsymbol{\gamma}_s + \log E_i,$$

where $\boldsymbol{\gamma}_s$ is the vector of all G adjustment terms for species s. The distribution of the $\boldsymbol{z}_i$ is required, we assume a multinomial with a single draw and the vector of probabilities $\boldsymbol{\pi}_i$. That is $\boldsymbol{z}_i \sim \text{mult}(1, \boldsymbol{\pi}_i)$. This type of model is commonly known as a finite mixture model (McLachlan and Peel, 2004). Our model has marginal expectation

$$\mathbb{E}\left(C_{si}\right) = \sum_{g=1}^{G} \pi_{ig} \mathbb{E}\left(C_{si} | \text{targetting type } g\right).$$

We refer to the model for catch C_{si}, that takes into account the different targeting types, as the targeted catch-and-effort model. Its simple structure is intuitive, simple and appealing. However, the full model is not complete yet. It is completed with specification of the $\boldsymbol{\pi}_i$ vector, which we refer to as the choice model.

B.2 The choice sub-model

This part of the model assumes that the probability of making a shot of some particular type depends on the Anticipated Net Revenue. The shot-type with the highest ANR (according to the model) is the most likely to get chosen, but any type is possible, for two reasons:

- the model does not have access to all the information that the fisher does;
- fishers do not always make perfect decisions.

In practice, the decision about shot-type (as we classify them) is intimately linked to the decision about average depth of the shot, so we compute ANRs across all combinations of shot-type and average depth[1].

[1] We assumed for simplicity that operational costs are about the same for all fishing options, so that profits are determined by shot outcome not by shot expenditure.

If there was enough data to calculate expected revenue specifically for the vessel b that is about to make the shot, then the prior probability of deciding to make the shot of type g at average depth[2] d would be calculated as:

$$\mathbb{P}\left[d, g|b, x, t\right] = \frac{\mathrm{ANR}^{\kappa}_{dgbxt}}{\sum_{d^* g^*} \mathrm{ANR}^{\kappa}_{d^* g^* bxt}}$$

where t is date and x stands for the other "conditioning data" for the shot: location-along-coast (which affect species density), whether the shot is to be basically deep or basically shallow, plus current economic factors explained below. The parameter κ is a "rationality" or "perfection-of-knowledge" parameter; values near zero correspond to shots being chosen randomly without any economic motive, and as the value of κ increases, the more likely a fisher is to choose the type-and-depth that the model reckons would be best.

Since we cannot hope to properly model vessel-specific net revenues, we approximate the effect by instead applying a vessel-specific preference weighting to the different types:

$$\mathbb{P}\left[d, g|b, x, t\right] = \frac{P_{bg}\mathrm{FANR}^{\kappa}_{dgbxt}}{\sum_{d^* g^*} P_{bg}\mathrm{FANR}^{\kappa}_{d^* g^* bxt}}$$

where the FANR is now calculated as a "Fleetwise average"— though there is still a small vessel-specific adjustment to allow for trip-level effects, explained below.

The term FANR_{dgbxt} is calculated as the expected difference of the trip revenue with and without the shot, where the "trip" comprises all shots in the three days prior (note that trips cannot be identified from logbook data). This takes into account that landings in the *next* shot can affect revenue from the *previous* shot in a trip, because of elasticity, as well as vice versa. The formula is

$$\mathrm{FANR}_{dgbxt} = \sum_{s=1}^{S} \left[\mathbb{E}\left(C_{si}|dgxt\right) + c^{+}_{stb}\right] p_{st} - \sum_{s=1}^{S} c^{+}_{stb} p^{+}_{st},$$

where:

- s is species
- p_{st} is the anticipated "net price" (see below) of s at landing, *including* the anticipated catch from the imminent shot
  - p^{+}_{st} is the anticipated "net price" without the imminent catch
- C_s is the random variable for the catch of species s in the imminent shot
  - c^{+}_{stb} is the known sum of the catches for this vessel over the three preceding days (typical trip duration; unfortunately, actual trips cannot be distinguished from the logbook data).

The anticipated net prices should reflect

[2] Here "average depth" is treated as discrete rather than continuous, being split into 25 narrow sub-bands within the "basically deep" or "basically shallow" range.

1. across the fleet on average: the difference between market price p_{st}^{mkt} (known to the model and to the fisher) and the fleetwise-average cost of having to lease enough unheld quota (unknown, but depends on TAC-tightness), and

2. for the vessel making the shot: the elasticity adjustment from catches earlier in the trip.

We dealt with the TAC issue via:

$$p_{st}^{\text{net}} = p_{st}^{\text{mkt}} \times \left(1 - \exp\left(-\theta_s \frac{T_s}{B_s^*}\right)\right) \tag{B.2.1}$$

where T_s is the quota for that species s that year (year subscript omitted for brevity), B_s^* is the abundance index for the species (i.e. the goal of the standardization, and represented by parameters already in the model), and $\theta_s > 0$ is a scaling parameter that allows for overall catchability of the species and the proportion of the TAC that is unavailable for lease.

The trip-specific elasticity adjustment is given by

$$\begin{aligned} p_{st} &= p_{st}^{\text{net}} \exp\left(\beta_s \log\left(c_{stb}^{+} + \mathbb{E}\left(C_s | dgxt\right) + 1\right)\right) \\ p_{st}^{+} &= p_{st}^{\text{net}} \exp\left(\beta_s \log\left(c_{stb}^{+} + 1\right)\right) \end{aligned}$$

Estimates of β_s come from the preliminary analysis described in Section C.2.

Vessel-specific type preferences, notated by $\boldsymbol{P}_b$ and discussed in Chapter 1.4, are assumed to be Dirichlet-distributed random effects with equal means $1/G$ (i.e. no average fleetwise preference for any type, except as economics dictates). The concentration parameter of the Dirichlet was chosen to make the distribution concave (to avoid numerical difficulties), but only mildly (to avoid biasing the results). In principle, this parameter could be estimated (e.g. by Laplace approximation), but the computational task was too great.

B.3 Log-likelihood

For estimation purposes we use, as a log-likelihood, the log of the joint probability density for the observations (C_{si}, $i = 1 \ldots n$ and $s = 1 \ldots S$), the (random) smoothing spline effects ($\boldsymbol{\tau}_{sj}$, $s = 1 \ldots S$, $j = 1 \ldots J$ where J is the number of smooth terms in the model), the boat preferences ($\boldsymbol{P}_b$, $b = 1 \ldots B$ where B is the number of boats), and the trawl depth (d_i). The log-likelihood is marginal to the targeting groups, obtained by summation – the usual approach for mixture models.

The conditional log-likelihood contribution for the catch of the i^{th} trawl at depth d_i is given by

$$\ell_i^{(1)} = \log\left(\sum_{g=1}^{G} \pi_{ig}(\kappa, \boldsymbol{\theta}, \{\boldsymbol{\tau}\}_{s=1}^{S}) \prod_{s=1}^{S} f(C_{si} | \text{targetting type } g \text{ depth } d_i)\right) \tag{B.3.1}$$

where $f(C_{si}|\text{targetting type } g \text{ depth } d_i)$ is the density from a Tweedie distribution (with mean given by (B.1.4) and some known power parameter), and $\pi_{ig}(\cdot)$ is the joint probability of targeting type g and depth d_i. Note that this is not marginal to depth, rather it is a true joint distribution according to the model in Chapter 1.4 and this appendix. The (joint) log-likelihood is simply the sum of these values, penalised for the other random effects. That is

$$\ell^{(2)} = \sum_{i=1}^{n} \ell_i^{(1)} - \sum_{s=1}^{S} \sum_{j=1}^{J} \frac{1}{2\sigma_{sj}^2} \boldsymbol{\tau}_{sj}^{\top} \boldsymbol{S}_j \boldsymbol{\tau}_{sj} - \sum_{b=1}^{B} D(\boldsymbol{P}_b), \tag{B.3.2}$$

where σ^2_{sj}, $\boldsymbol{\tau}_{sj}$, $\boldsymbol{S}_j$ are the variance components, the random effects, the smoothing (inverse variance) matrix for the j^{th} penalised smoothing spline in the model respectively, and $D(\cdot)$ is the log-density for a Dirchlet distribution. We parameterise the Dirichlet distribution using a concentration parameter, α^* and a set of G scaled parameters that sum to one, $\boldsymbol{\alpha}$. To ensure concavity, we assume that each element of $\boldsymbol{\alpha}$ is at least 1. This is performed by assuming that at least one observation is in each group. This is identical to using a posterior density where one observation from each group is observed, and the observation comes from a multinomial. Under these considerations, the function $D(\cdot)$ is

$$D(\boldsymbol{x}) = \log \Gamma(\alpha^* + G) - \sum_{g=1}^{G} \Gamma(\alpha_g + 1) + \sum_{g-1}^{G} (\alpha_g) \log x_g.$$

B.3.1 Estimate-taming penalties

The smoothing spline terms $\{\boldsymbol{\tau}_j\}$ have been introduced as random effects. This has been done to penalise the flexibility of the model and the size of the penalty is inversely proportional to the variance of the random effects. This has well-known benefits, chief amongst them is that it smooths the likelihood surface and makes estimation easier. There are costs too, such as biased estimates with the amount of bias depending on the severity of the penalty. We have introduced the Dirichlet boat effects with a similar motivation and we introduce some further penalties in this section. The prime motivation is to smooth the log-likelihood to make estimation easier.

The location parameters in the catch and effort sub-model ($\{\boldsymbol{\tau}_{so}\}_{s=1}^{S}$ in Section B.1) are assumed to have a mild (quadratic) penalty. This implies that they are assumed to be normal random effects with mean zero. The group and species adjustments in the targeted catch and effort model ($\{\gamma_{sg}\}$ in Section B.1) are also assumed to be random normal effects, with mean zero. Likewise the log of the TAC scaling parameters (θ_s in B.2.1) and the log of the confusion parameter (κ in Section B.2) all are assumed to be normal with mean zero. The penalised log-likelihood is the one used for estimation, it is

$$\ell^{(3)} = \ell^{(2)} - \sum_{s=1}^{S} \frac{1}{2\sigma_o^2} \boldsymbol{\tau}_{so}^\top \boldsymbol{\tau}_{so} - \sum_{s=1}^{S}\sum_{g=1}^{G} \frac{1}{2\sigma_\gamma^2} \gamma_{sg}^2 - \sum_{s=1}^{S} \frac{1}{2\sigma_\theta^2} (\log \theta_s)^2 - \frac{1}{2\sigma_\kappa^2} (\log \kappa)^2 \tag{B.3.3}$$

The variances of these random effects, introduced for computational reasons, are all extra tuning parameters. As such, their choice is important. However, the SETF data are plentiful and the effect of these tuning parameters is likely to be small – except for parameter values very far from that expected. For this reason, we try to choose mild tuning parameters.

B.4 Estimation

Estimation is performed by maximising the penalised log-likelihood in (B.3.3). Maximisation is carried out using a quasi-Newton optimisation routine Nash and Sofer, 1996, which provides super-linear convergence. The quasi-Newton method requires the first derivatives of the penalised log-likelihood. We obtain these using automatic differentiation Griewank, 2001 as implemented in the CppAD C++ library Bell, 2011.

The calculation of the penalised log-likelihood requires evaluation of the Tweedie density, for fixed power parameter. We use the method of Dunn and Smyth, 2005, also see Appendix A of Foster and Bravington, 2013. We pre-specify the power parameter based on the results from Peel et al., 2013, who used a univariate Tweedie GLM and estimated the power parameter by choosing it to make the residuals homogeneous.

Limited testing of these estimates, based on the methods in Foster and Bravington, 2013, suggest that they should be adequate. For computational thriftiness, it is important to pre-specify the power parameters as this enables pre-calculation of many of the terms required by the series evaluation method of Dunn and Smyth, 2005.

The calculation of the penalised log-likelihood is an involved process that involves many calculations. For this reason, the optimisation procedure is slow – for the summer data (three targeting groups and eight species), the model will take up to 18 hours to converge. Unfortunately, this limits the breadth of the scenarios that can be investigated.

We originally planned a fully Bayesian implementation, in which hyperparameters such as the variance of vessel-effects would also be estimable parameters. Such models need special-purpose software to fit, e.g. able to handle automatic Laplace Approximation (Skaug and Fournier, 2006), and in 2008 there were only two options: ADMB (which was at that time causing great problems to one of us in a different project), and CppAD (able to handle Laplace Approximation in principle, but not previously tested). In the end, though, we had to opt for a simpler approach of fixing the hyperparameters manually, adequate for an exploratory model; but we were by then committed to CppAD.

Appendix C

Details of Data Used

C.1 Log book data

The details of the subsetting procedure are given in Table C.1.1. These filters, when applied to the entire log-book data deliver the data used in this study. They are based on the approaches described in Darbyshire et al., 2008 and Klaer and Smith, 2008.

Table C.1.1: Filtering rules to remove erroneous data and to homogenise the remaining data. All named variables must not have missing values.

	Condition	Number of Trawls
No Filtering		660,184
Longitude & Latitude	–	660,184
Vessel (Boat)	–	660,184
Catch & Effort	–	633,628
SETF Zone	–	633,628
Depth	Average depth < 700m	547,972
Period of Day	Not 'unkown'	538,215
Start and End in Ocean	–	529,739
Trawl Distance	0 < Trawl distance < 30 nautical miles	501,261
Trawl Time	0.5 < Trawl time < 8 hours	495,222
Depth	Average depth > 10m	494,983
Roughy Trawls	Trawl time == 0 & Depth > 500	494,976
Years	1994–2008	349,854
SETF Zones	10, 20, and 30	247,874
Season	Winter and Summer	123,937
Active Boats	>1000 trawls in entire period	94,689
Boats active in boat.time	>100 trawls for each boat.times	94,476
Number of trawls to analyse: 94,726 (winter 48,864; summer 45,612)		

C.2 Fish market data

We do not require the observed price for the model. Rather, it is the expected price that is important and the elasticity (substitution) effects for lagged catches. We start development by considering that expected price is obtained from prediction from a generalised additive model (Wood, 2006, for example) that contains a smooth term for time only. It is simply

$$\log\left(p_s(t_i)\right) = \log \mathbb{E}\left(y_{st}\right) = f(t),$$

where y_{st} is the observed price for the s^{th} species on the t^{th} day and $f(t)$ is the penalised spline function value for day t. This is a multiplicative model, which is preferred here as it provided a better fit to an additive model. The additive model failed to account for the varying size of the deviations due to inflation, especially so when elasticity effects are included.

The time-only model ignores any effects arising from elasticity, substitution and other market-based drivers of price. In essence, this models describes the expected price *marginal* to these market-based drivers. That is, it predicts the expected price *irrespective* of what the market based effects are. It is needed in the full mixture model.

The mixture model also needs the elasticity effects – the amount of price reduction for having a large amount of a species already on the market. These effects are obtained from an extension of the time-only GAM. This approach models some of the fine scale temporal effects that are not identified by the spline. The model is more economically sensible as it accounts for the two major sources of price variation: time trend and elasticity, where time is both inter- and intra-annual. In addition to the smooth term, another is added for the amount (kg) sold in the last *three* days. That is

$$\log \mathbb{E}\left(y_{st}\right) = f_2(t) + \beta_s \log\left(C_{st}^{+} + 1\right),$$

where C_{st}^{+} is the cumulative amount of species s sold in the last 3 days and β_s is the elasticity parameter. The lagged catch is included as $\log\left(C_{st}^{+} + 1\right)$ as this scale both (approximately) linearised the relationship effectively, and it down weighted the effect of particularly large catches. Note that this equation could, in principle, be extended to account for substitution effects (amount of other species on the market affect price). This has not been done – it would require a more detailed analysis of the fish market data, which the data themselves are unlikely to support.

This analysis was performed for all species that had more than 500kg market turnover throughout the entire 1994-2008 time period. Figure 4.1.1 illustrates the fit of this model for two species.

Appendix D

Model Specifics for the Logbook Data Analysis

We fitted separate targeting models to the summer/winter and shallow/deep data, see Section C.1. This was done as it made computation tractable and also as many species behave like different stocks over the different seasons/depths. The sub-model for the catch and effort data (see Section B.1) was specified as

$$\log \mathbb{E}\left(C_{si} | \text{targetting type } g\right) = \tau_{s,\text{year}}(i) + h_s(\text{along}_i) + h_s(\text{depth}_i) + \tau_{s,\text{period}}(i) + \gamma_{sg} + \log E_i,$$

where C_{si} is the catch of the s^{th} species in the i^{th} shot, $\tau_{s,\text{year}}(\cdot)$ and $\tau_{s,\text{period}}(\cdot)$ are indicator functions that produce discrete values for each year and day period (night, day and a combination of night and day (mixed)) respectively, γ_{sg} is the targeting type factor, E_i is the effort for shot i, and $h_s(\text{along}_i)$ and $h_s(\text{depth}_i)$ are smoothing splines functions for the along-coast and depth variables respectively. All effects in this catch-and-effort model are species dependent, except for the offset (of course).

The smoothing spline functions for depth and along-coast were defined as cubic regression smoothing splines, each with 9 bases functions. The variance components for the random effect formulation of the smoothing splines were taken from non-targeted fits to each species. Ideally, these components would have been estimated in the full model. However, computational limitations prevented it. Previous work (Peel et al., 2013) used a similar approach to single-species analyses ans showd that these two spline effects are, perhaps surprisingly, effectively independent for a large number of species.

There are a large number of possible models that we could have fitted. Some may have even been more sensible than the one used. However, we considered this to be a plausible model that was computationally feasible. Based on previous work (Peel et al., 2013) we suspect that any refinements would produce only modest improvements to the model's fit.

The (joint) sub-model for targeting type and for depth requires the specification of the number of depth classes to use in the approximation of the integral. Here we use $D = 25$ depth ranges. This was chosen as a compromise between accuracy and computational feasibility. In a similar spirit the number and locations of reference points for estimating the total abundance for the TAC-adjusted revenue need to be specified, see B.2.1. We use a set of 125 locations that are spread evenly through the along-coast variable. These points all follow the contour defined by the average trawl depth.

Vessel effects are included in the model using the methods described in Section B.2. We allow for time-varying boat effects by categorising time for each boat into 5 year (maximum) blocks. The minimum time block is 5 years as smaller time periods were thought to be too variable. The change-points for the vessels are

Table D.0.1: Penalties used for the SETF analyses.

Terms	Type	Penalty
$\{\delta_{sg}\}_{g=1}^{G}$	Variance	1
$\{\boldsymbol{\tau}_o\}_{s=1}^{S}$	Variance	3
$\{\log\theta_s\}_{s=1}^{S}$	Variance	1
$\log\kappa$	Variance	2
$\{\boldsymbol{p}_b\}_{b=1}^{B}$	Concentration	10

staggered so that changes in the fleet behaviour does not affect the catch trend. Each vessel is assigned into one of 5 groups, each of which has change points at different years. Any boat and time block combination that has less than 100 observations was removed. This leaves a total of 108 boat by time block combinations.

The penalties used in an effort to smooth the log-likelihood surface are listed in Table D.0.1. In practice, the choice of penalties do not make much difference as the information in the data is strong enough to over-power the penalties. However, the penalties do 'smooth out' some of the more irregular features of improbable combinations of parameter values. Sensitivity to these choices is shown to be low in Appendix F.

There are a large number of parameters in this model. We give the complete list in Table XXX, along with their type and their number.

Table D.0.2: Listing of parameters for the model applied to the CTS fishery. Note that the spline wiggliness term is taken from the literature, the elasticity parameters are pre-estimated from the market data and the penalties are specified (not estimated). All other parameters are estimated from the log-book data. The number of parameters is for a G targetting group model with B vessels (recall that vessel in this analysis is a boat within a 5 year time block).

	Parameter	Type	No. per Species	No. per Model	Notes
Catch and Effort Sub-Model	$\boldsymbol{\tau}_{s,year}$	factor	23	161	
	$\boldsymbol{\tau}_{s,period}$	factor	2	14	
	$h_s(along)$	spline	9	63	
	$h_s(depth)$	spline	9	63	
	γ_{sg}	tagetting term	G	7G	
Choice Sub-model	κ	rationality	–	1	
	θ_s	TAC scaling	1	7	
	β_s	elasticity	1	7	Pre-estimated from market data
	$\boldsymbol{P}_b$	boat preferences	G per boat	BG	B=boat-block combinations
Penalties and Priors	σ^2_{sj}	spline wiggliness	–	1	Taken from Peel et al., 2013
	σ^2_0	penalty	–	1	for $\tau_{s,year}$ and $\tau_{s,period}$
	σ^2_γ	penalty	–	1	for γ_{sg}
	σ^2_θ	penalty	–	1	for $\{\theta_s\}$
	σ^2_κ	penalty	–	1	for κ
	α^*	penalty	–	1	for $\{\boldsymbol{P}_b\}$

Appendix E

Summer Results

The winter results, deep and shallow, are presented in the body of the document. Here, we present the results of the summer analyses. We do not provide the same level of interpretation, as most of it would be repetition of the ideas presented previously.

E.1 Summer Deep

A 2-type model has the highest "rationality parameter"; the catchability parameters are given in Table E.1.1. The types are clearly separated into favouring flathead/redfish/morwong (Type 1), or blue grenadier (Type 2), which is more common. There is not much change in proportions of each type over time (Figure E.1.1). Posterior probabilities of shot-type show good clustering (Figure E.1.2). Clustering is more distinct here than for the other subsets where more types are involved.

	Blue Grenadier	Jackass Morwong	Ling	Mirror Dory	Redfish	Spotted Warehou	Tiger Flathead
Type 1	-1.21	1.02	-0.33	-0.39	2.62	-0.31	6.15
Type 2	1.21	-1.02	0.33	0.39	-2.62	0.31	-6.15

Table E.1.1: **Summer deep** catchability parameters for the chosen model; see main report for explanation. With only two types, it is mathematically inevitable that these log-catchability effects will have equal magnitude but opposite signs in the two types.

In terms of abundance, the inferred series under different models are given in Figure E.1.3. On the whole there is again no *major* difference when targeting is included (though basic standardization does matter), though the chosen 2-type model (solid black) does give a less spiky series than the 1-type model (standardization but no targeting; dotted turquoise) especially for redfish and grenadier. While this is also true for spotted warehou, one of the (non-chosen) targeting models with *more* than two types suggests a different story to the 2-type model; there is some sensitivity to model-choice here.

E.2 Summer Shallow

Here a 3-type model gave considerably the highest rationality parameter, although one type is rare (~4% of all shots). The two main types could be described as pro- or anti-flathead ("Type 3" and "Type 2" in

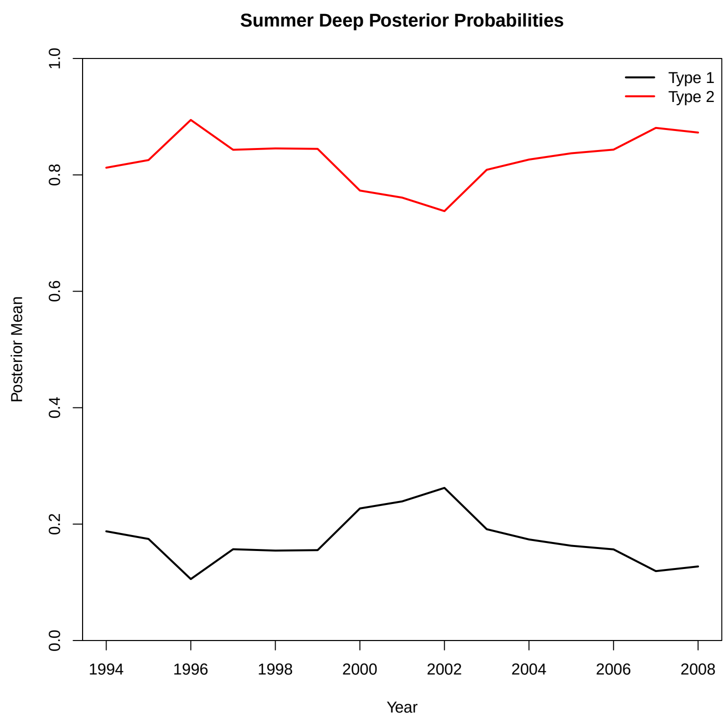


Figure E.1.1: Proportion of **Summer deep** shot-types over time

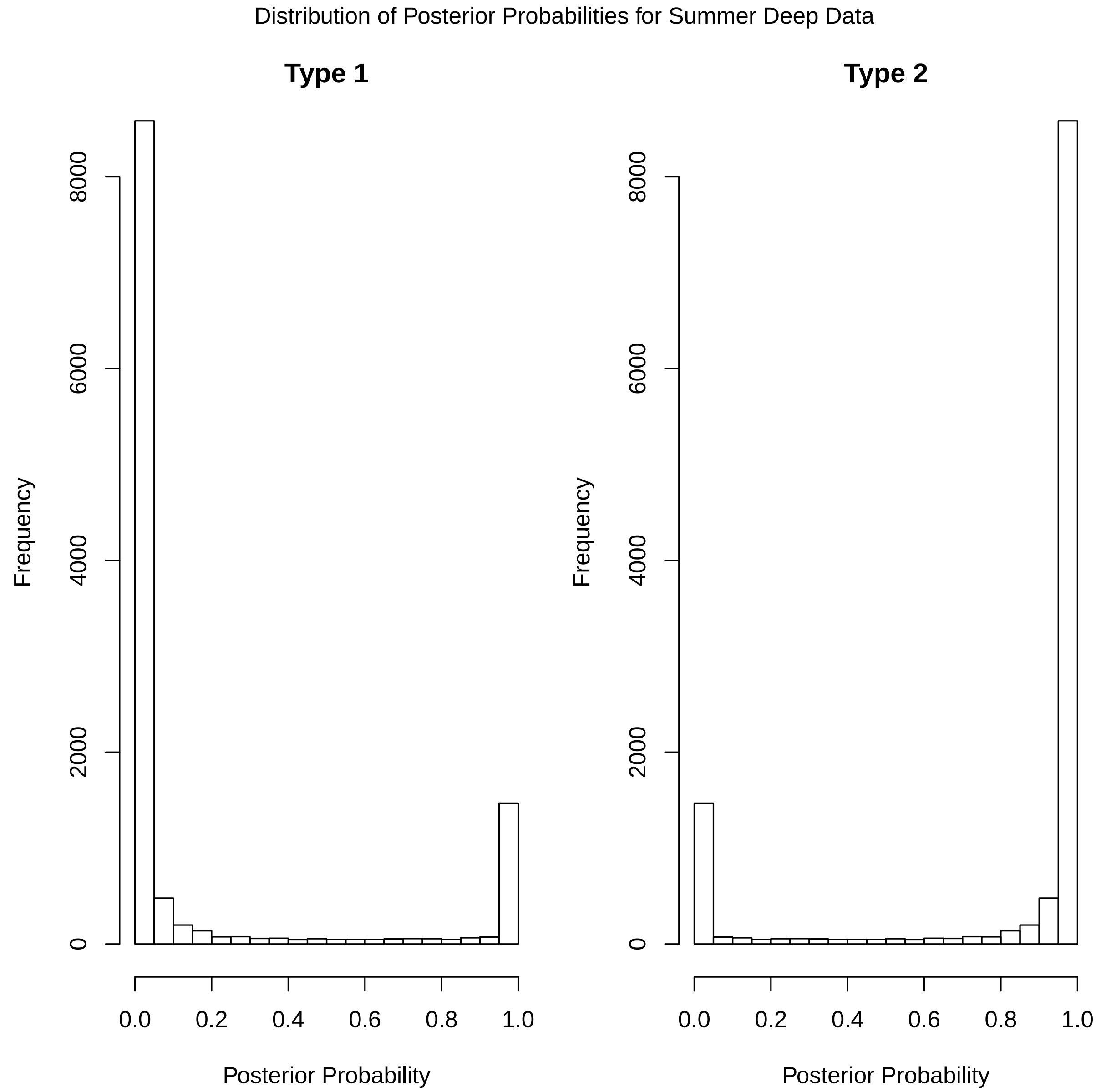


Figure E.1.2: Posterior probabilities of shot-type for the **summer deep** data. The mirror-image is inevitable since there are only two types in this model.

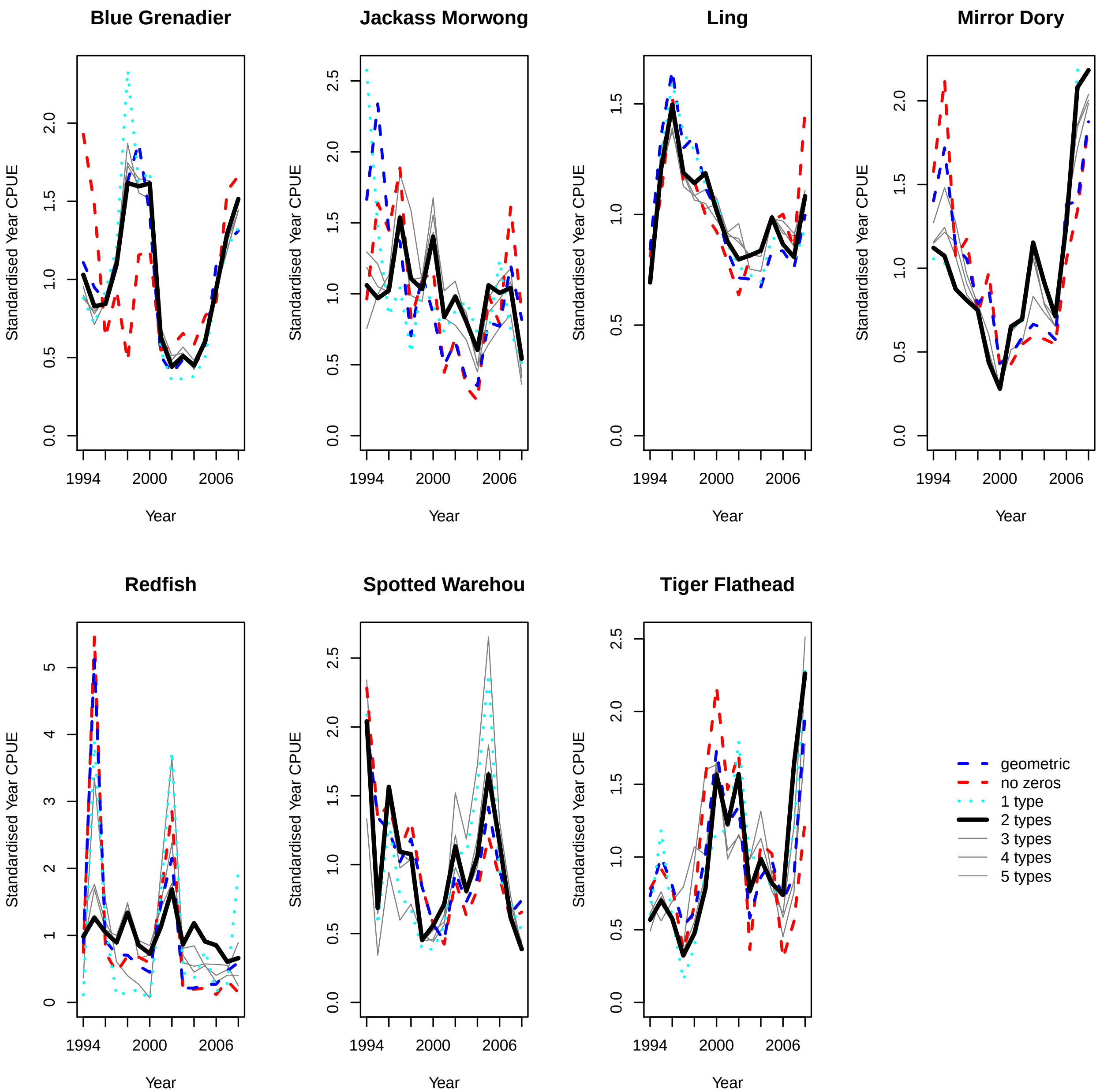


Figure E.1.3: **Summer Deep** abundance indices under different targeting models. See main report for explanation of the different lines, in particular Figure 4.3.2.

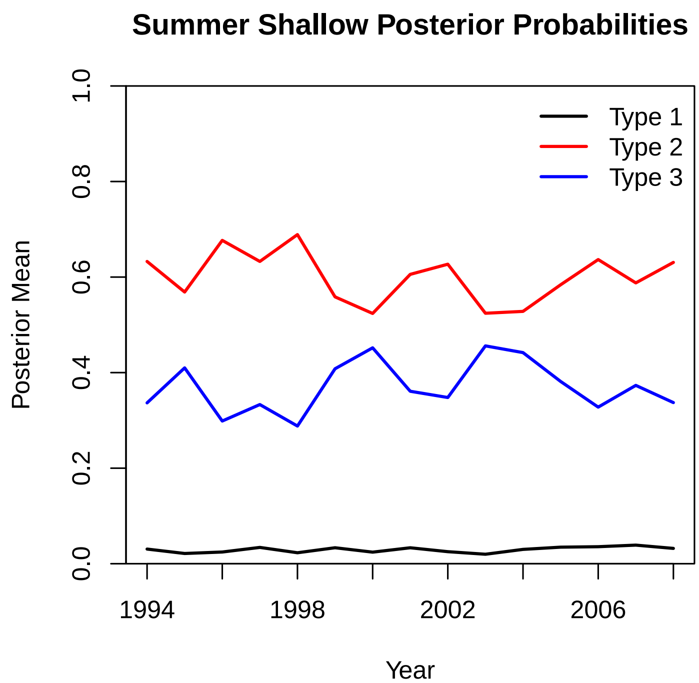


Figure E.2.1: Proportion of **Summer shallow** shots of each type over time.

Table E.2.1 respectively), although flathead is widespread and the estimated "pro-flathead" effect is only about 20%. However, the pro-flathead type has effectively no chance of catching morwong, mirror dory, or spotted warehou. The rare third type ("Type 1") apparently favours mirror dory, redfish and silver trevally against flathead. There is not much trend in the proportions over time (E.2.1). Posterior probabilities are fairly well separated.

	Jackass Morwong	John Dory	Mirror Dory	Redfish	Silver Trevally	Spotted Warehou	Tiger Flathead
Type 1	4.74	-0.18	5.42	2.89	2.70	2.23	-0.14
Type 2	4.89	0.06	1.88	-0.32	-0.67	2.58	-0.02
Type 3	-9.63	0.12	-7.30	-2.57	-2.03	-4.82	0.16

Table E.2.1: **Summer shallow** log-catchability effects. Note that Type 1 is rare (˜4% of shots).

The prevalence of each targeting type does not appear to change over time, see Figure E.2.1. The prevalence of targeting type 2 has a higher prevalence than type 3 and a much higher prevalence of type 3 throughout the study period.

The various abundance indices for summer-shallow are given in Figure E.2.3. Again, some species are insensitive to model choice, for most others standardization has an effect but targeting does not, and for a few (morwong, john dory, mirror dory, spotted warehou) the with-targeting results are a little less spiky, though not much different in overall trend.

Vessel preferences are again well-scattered (Figure E.2.4A) and within-vessel changes over time are gen-

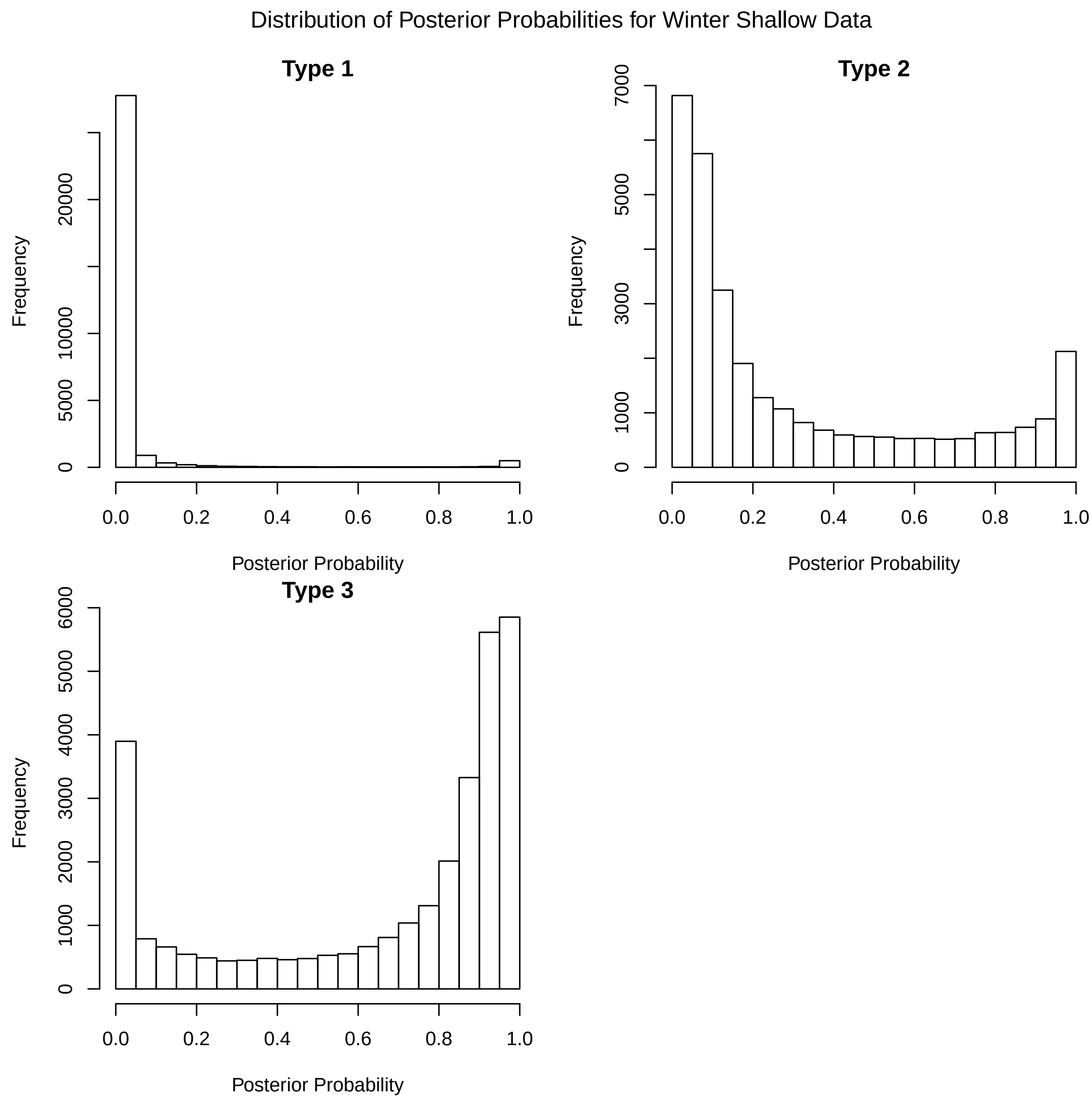


Figure E.2.2: Distribution of posterior probabilities for the **summer shallow** data.

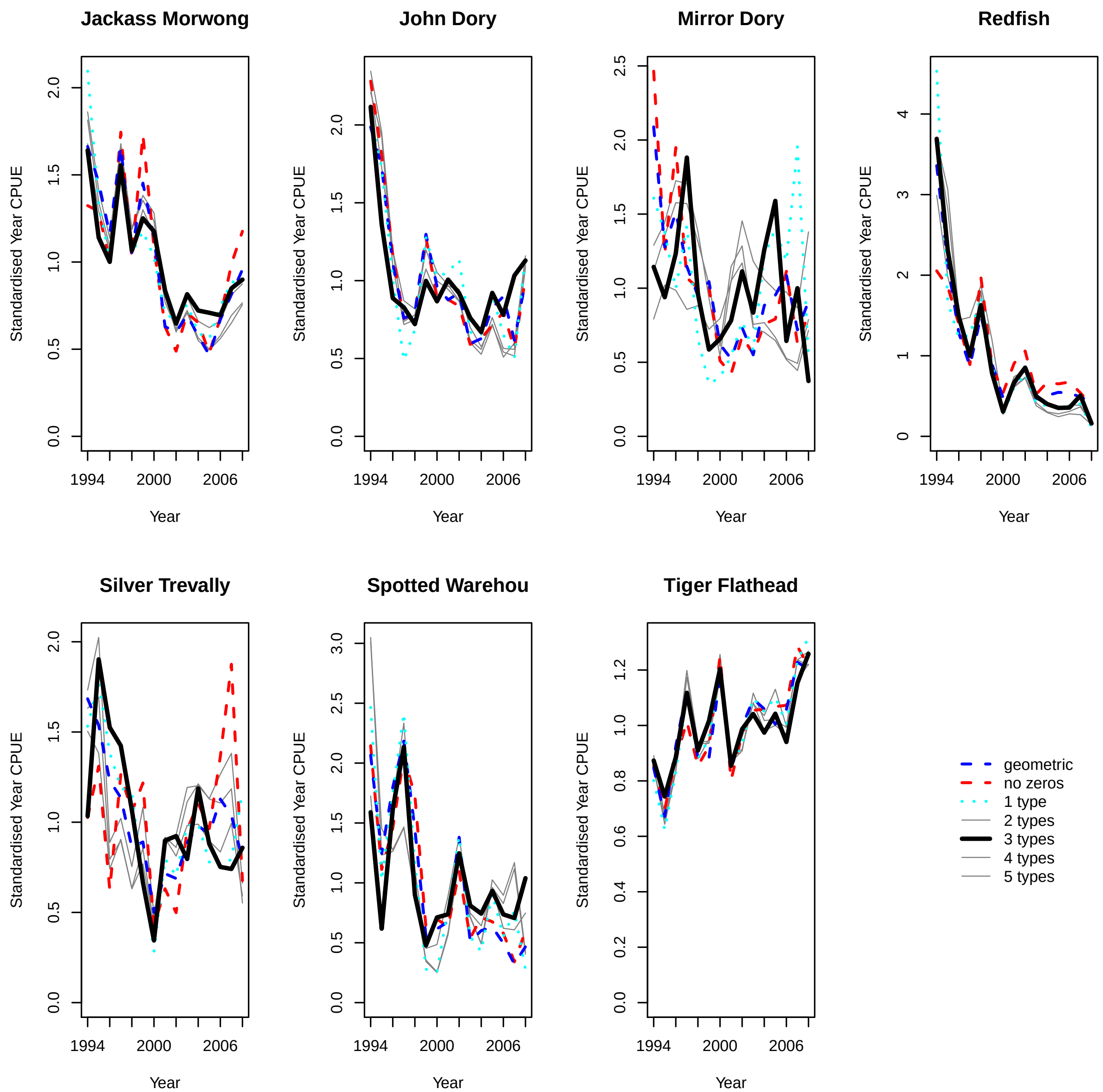


Figure E.2.3: **Summer Shallow** abundance indices under different targeting and standardization models. Solid black vs dotted-turquoise correspond to 3-type vs 1-type targeting. See main report for explanation of the different lines, in particular Figure 4.3.2.

erally smaller than between-vessel differences (Figure E.2.4B–F), suggesting no major problem with the 5-year-fixed-block approximation. Some of the vessel preferences are quite close to the rare "Type 1" vertex, so those shots may come from a particular subset of the fishery.

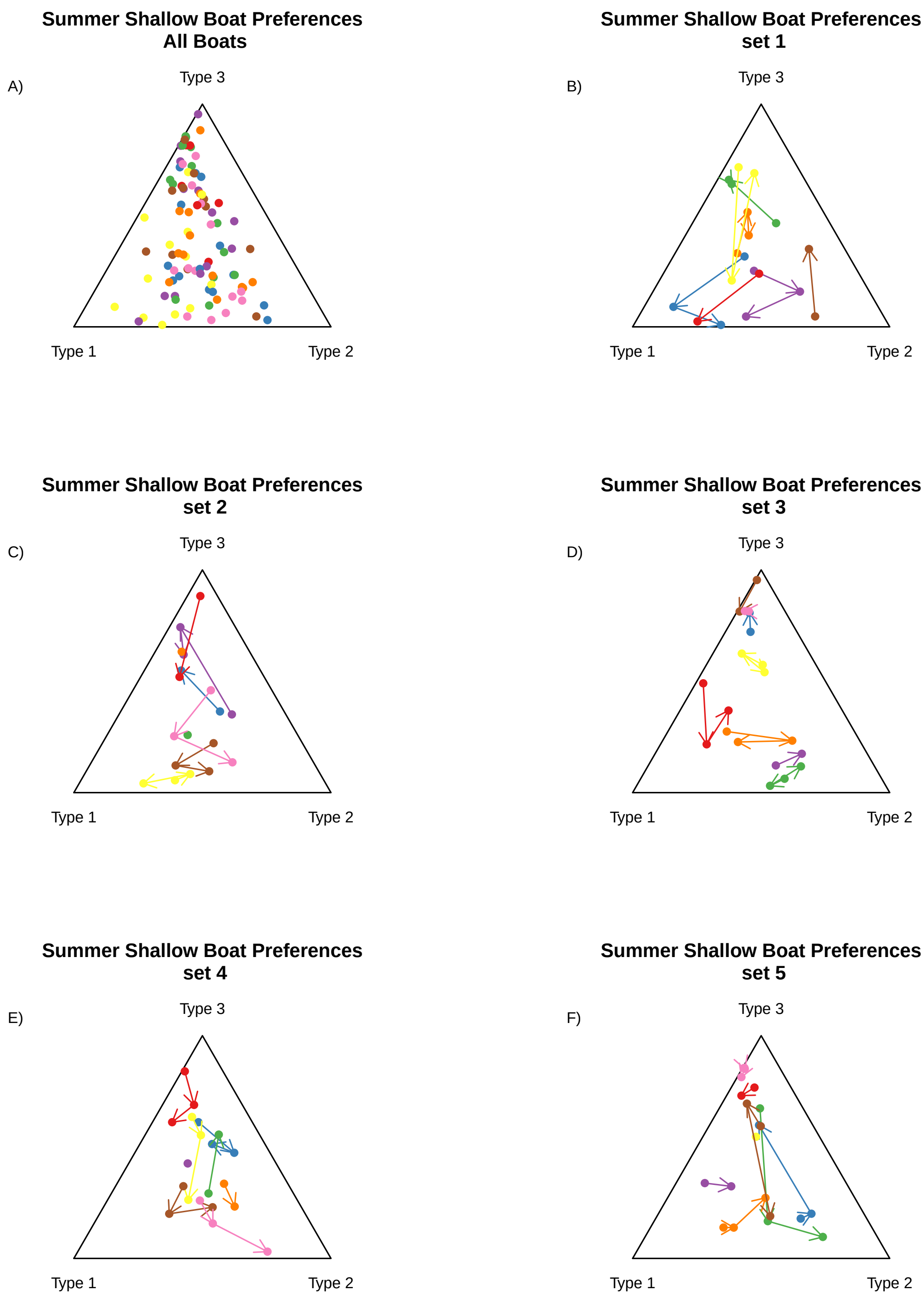


Figure E.2.4: **Summer Shallow** boat preferences on ternary diagrams. Each boat (a dot) has a preference for each of the three different targeting types (vertices of triangle). A boats proximity to a vertice measures its preference to that targeting type. Panel A) presents all boat effects (vessels and periods), coloured by vessel. Panels B) through to F) give a subset of vessels, whose preferences have been tracked over time (between periods). Sequential periods have been joined by arrows. Subsetting was performed to reduce the amount of information per panel to an interpretable amount.

Appendix F

Sensitivity Analysis

F.1 Penalisation

The penalties on the parameters are arbitrary and their effect on the resulting conclusions should be checked. The sensitivity of the inference from the model on the amount of penalisation can be checked by fitting models with more, and less, penalisation. To this end, we use the winter data (combined shallow and deep) and the three component model to inspect how sensitive the resulting CPUE series is to the amount of penalisation in the log-likelihood. Note that this is a different data set than previous analysis. The difference is an unfortunate accident of history, but not an overly large one. The choice of data should not affect the sensitivity of the model because the amount of information in the data will be comparable and so the effect of penalisation will also be comparable. We fitted two additional models, one with more penalisations and one with less penalisation than the model with penalties defined in Table D.0.1. See Table F.1.1 for altered penalties in the sensitivity analysis.

The resulting CPUE series do differ, but not by much (see Figure F.1.1). This amount of variation is small, especially when compared to the amount of variation exhibited between the models of different numbers of targetting types. We do not feel that the choice of penalties invalidates the inferences drawn from the model.

F.2 Time of Day and Length of Trawls

There has been some discussion as to whether the period of day that the trawl was undertaken (day, night, or mixed) is adequately accounted for in the model. Likewise, it is possible that filtering out shots that are

Table F.1.1: Penalties used for the sensitivity to penalties analyses. Compare with Table D.0.1.

Terms	Type	Less Penalty	More Penalty
$\{\delta_{sg}\}_{g=1}^{G}$	Variance	10	0.5
$\{\boldsymbol{\tau}_o\}_{s=1}^{S}$	Variance	10	1
$\{\log\theta_s\}_{s=1}^{S}$	Variance	10	0.5
$\log\kappa$	Variance	2	2
$\{\boldsymbol{p}_b\}_{b=1}^{B}$	Concentration	0.3	30

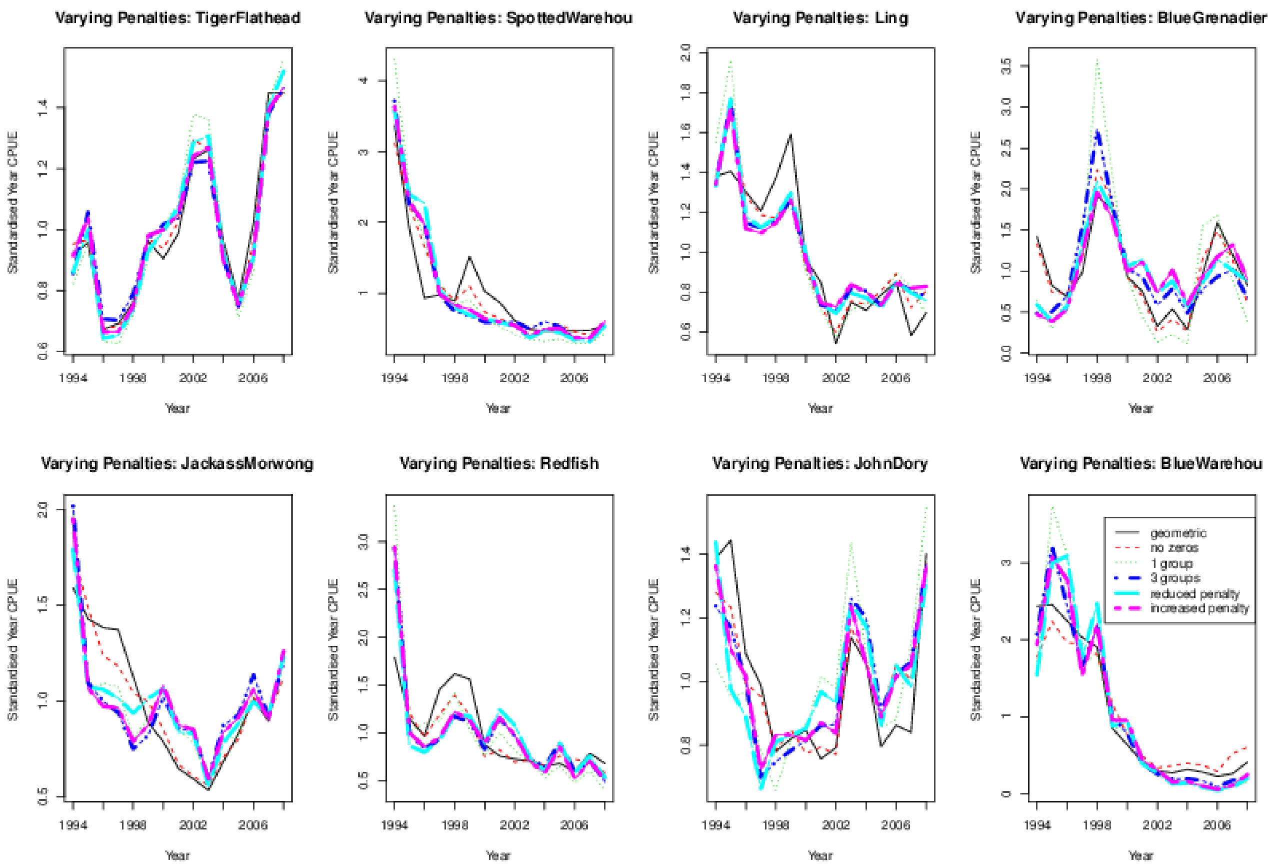


Figure F.1.1: Sensitivity to penalisation. Winter catch rates for different targetting models (all depths).

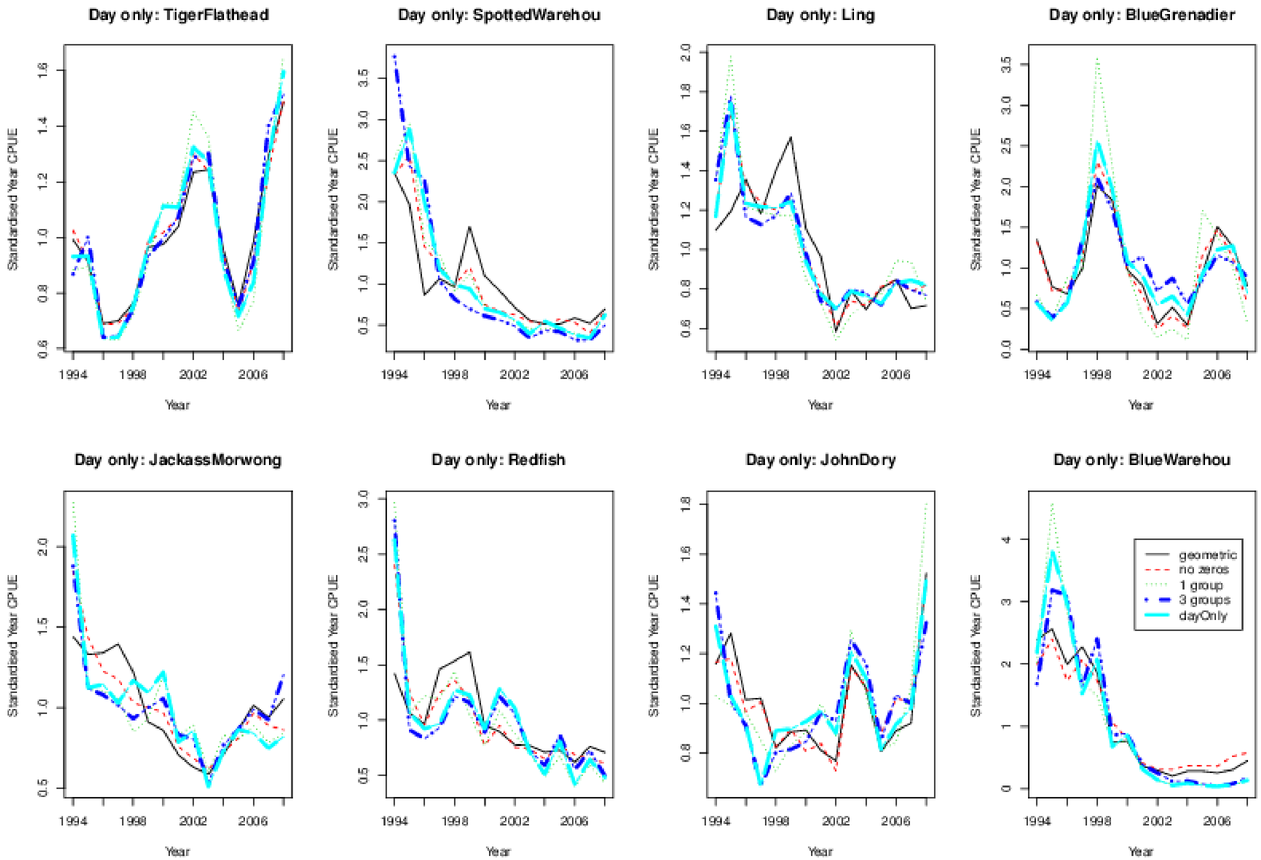


Figure F.2.1: Comparison of model fitted to data with and without night shots (winter data, all depths).

too long or too short will make a substantial difference – it will obtain a more homogeneous set of shots. To investigate these two issues we fitted an extra two models; one to those shots that were undertaken during the daytime and one to those whose trawl time was limited to be between 2.5 and 4.5 hours. The number of trawls analysed in either case was 36,648 for the daytime shots and 33,685 for the effort-limited shots. The resulting CPUE series are presented in Figures F.2.1 and F.2.2. The effect on the CPUE series of using only the daytime shots is largely insignificant as is the effect of using a limited effort range. This implies that the inferences from the model are robust to subsetting data for time of day and for shot length.

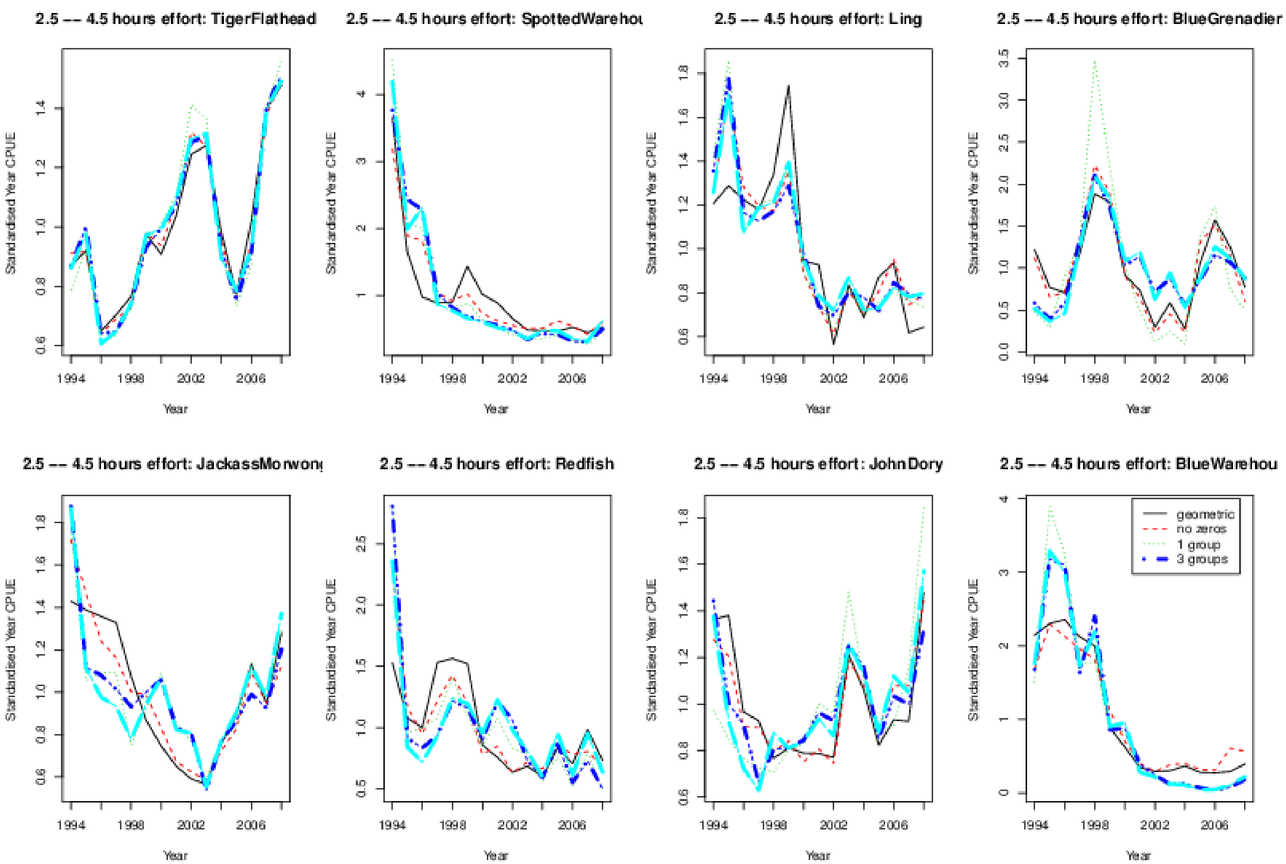


Figure F.2.2: Comparison of model fitted with restricted trawl duration (2.5 to 4.5 hours) to that will winter trawls (all dpeths).

Bibliography

AFMA (2014). *Harvest Strategy Framework for the Southern and Eastern Scalefish and Shark Fishery.* Tech. rep.

Bax, N and A Williams (2000). *Habitat and fisheries production in the Southeast Fishery ecosystem.* Tech. rep. FRDC Report 94/040. Fisheries Research and Development Corporation of Australia.

Bell, B (2011). *CppAD: a package for C++ algorithmic differentiation, COIN-OR.* http://www.coin-or.org/CppAD/, version 2011/01/27.

Bravington, M, P Grewe, and C Davies (2014). *Fishery-independent estimate of spawning biomass of Southern Bluefin Tuna through identification of close-kin using genetic markers.* FRDC Report 2007/034. CSIRO, Australia.

Candy, SG (2004). "Modelling catch and effort data using generalised linear models, the Tweedie distribution, random vessel effects and random stratum-by-year effects". In: *CCAMLR Science* 11, pp. 59–80.

Carroll, R (2011). *Measurement error in covariates.*

Carroll, R, D Ruppert, L Stefanski, and C Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

Cartwright, I (2005). "The Australian Northern Prawn Fishery". In: *Successful Fisheries Management Issues, Case Studies and Perspectives. SIFAR/World Bank Study of Good Management Practice in Sustainable Fisheries.* Ed. by S Cunningham and T Bostock. The Netherlands: Eburon Academic Publishers, pp. 197–231.

Carvalho, FC, DJ Murie, FH Hazin, HG Hazin, B Leite-Mourato, P Travassos, and GH Burgess (2010). "Catch rates and size composition of blue sharks (Prionace glauca) caught by the Brazilian pelagic longline fleet in the southwestern Atlantic Ocean". In: *Aquatic Living Resources* 23 (04), pp. 373–385.

Castro, J, A Punzón, GJ Pierce, M Marín, and E Abad (2010). "Identification of métiers of the Northern Spanish coastal bottom pair trawl fleet by using the partitioning method {CLARA}". In: *Fisheries Research* 102.1–2, pp. 184–190.

Darbyshire, R, A Sampaklis, P Hobsbawn, and A Morison (2008). *Spatial analysis of catch per unit effort data in the Southern and Eastern Scalefish and Shark Fishery: a case study for jackass morwong.* Tech. rep. Canberra: Australian Government's Bureau of Rural Sciences.

Deporte, N, C Ulrich, M Stéphaniem, and S Demanèche (2012). "Regional Métier definition: a comparatice investigation of statstical methods using a workflow applied to international otter trawl fisheries in the North Sea". In: *ICES Journal of Marine Science* 69.2, pp. 331–342.

Dunn, P and G Smyth (2005). "Series evaluation of Tweedie exponential dispersion model densities". In: *Statistics and Computing* 15, pp. 267–280.

Foster, S and M Bravington (2010a). "Accounting for targetting and exact zeros in fisheries catch and effort data". In: *Invited presentation at: Workshop on Data Science.* Keio University, Japan.

Foster, S and M Bravington (2010b). “Fisheries Catch and Effort Analyses: Dealing with exact zeros through extensions of Tweedie models.” In: *Invited talk at: the Japanese Joint Statistical Meeting*. Waseda University, Japan.

— (2013). “A Poisson-Gamma Model for the Analysis of Non-Negative Continuous Data”. In: *Journal of Environmental and Ecological Statistics*, pages.

Graham, K, N Andrew, and K Hodgson (2001). “Changes in relative abundance of sharks and rays on Australian South East Fishery trawl grounds after twenty years of fishing”. In: *Marine and Freshwater Research* 52.4, pp. 549–561.

Griewank, A (2001). *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. SIAM.

Haddon, M (2012). *Using innovative techniques to analyse trends in abundance for non-target species*. Tech. rep. FRDC Report 2010/057. Fisheries Research and Development Corporation of Australia.

He, X, KA Bigelow, and CH Boggs (1997). “Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery”. In: *Fisheries Research* 31.1–2, pp. 147–158.

Hilborn, R and CJ Walters (1991). *Quantitative fisheries stock assessment: choice, dynamics, and uncertainty*. Chapman and Hall.

Hui, F, D Warton, and S Foster (in press). “Order Selection in Finite Mixture Models: Complete or Observed Likelihood Information Criteria?” In: *Biometrika*.

Iriondo, A, R Prellezo, m Santurtún, D Garcia, I Quincoces, and E Mugerza (2010). “A multivariate approach for metier definition: A case study of Basque Country Trawlers”. In: *Revista de Investigación Marina* 17, pp. 140–148.

Jørgensen, B (1997). *The Theory of Dispersion Models*. Chapman & Hall.

Klaer, NL (2006). “Changes in the structure of demersal fish communities of the south east Australian continental shelf from 1915 to 1961”. PhD thesis. University of Canberra.

Klaer, NL and DC Smith (2012). “Determining primary and companion species in a multi-species fishery: Implications for {TAC} setting”. In: *Marine Policy* 36.3, pp. 606–612.

Klaer, N and D Smith (2008). *Species associations and companion TACs in the SESSF*. Tech. rep. Canberra: Australian Fisheries Management Authority.

Knuckey, I, M Bravington, D Peel, M Koopman, M Fuller, N Klaer, J Day, J Upston, and R Hudson (2013). *Implementation of a fishery independent survey for the Southern and Eastern Scalefish and Shark Fishery*. Tech. rep. FRDC Report 2006/028. Fisheries Research and Development Corporation of Australia, p. 168.

Larcombe, J and G Begg (2008). *Fishery status reports 2007: status of fish stocks managed by the Australian Government*. Tech. rep. Bureau of Rural Sciences, Canberra.

Maunder, MN and AE Punt (2004). “Standardizing catch and effort data: A review of recent approaches”. In: *Fisheries Research* 70.2-3, pp. 141–159.

Maunder, MN and AE Punt (2013). “A review of integrated analysis in fisheries stock assessment”. In: *Fisheries Research* 142, pp. 61–74.

Maunder, MN, JR Sibert, A Fonteneau, J Hampton, P Kleiber, and SJ Harley (2006). “Interpreting catch per unit effort data to assess the status of individual stocks and communities”. In: *ICES Journal of Marine Science: Journal du Conseil* 63.8, pp. 1373–1385.

McLachlan, G and D Peel (2004). *Finite mixture models*. John Wiley and Sons.

Nash, S and A Sofer (1996). *Linear and Nonlinear Programming*. Singapore: McGraw-Hill.

Peel, D, M Bravington, N Kelly, SN Wood, and I Knuckey (2013). “A Model-Based Approach to Designing a Fishery-Independent Survey”. In: *Journal of Agricultural, Biological, and Environmental Statistics* 18.1, pp. 1–21.

Peel, D, M Bravington, N Kelly, S Wood, and I Knuckey (2011). "A model-based approach to designing a fishery independent survey". In: *submitted*.

Pelletier, D and J Ferraris (2000). "A multivariate approach for defining fishing tactics from commercial catch and effort data". In: *Canadian Journal of Fisheries and Aquatic Sciences* 57.1, pp. 51–65.

Punt, A, A Smith, and C Gurong (2002). "Evaluation of management tools for Australia's South East Fishery. Towards selecting appropriate harvest strategies". In: *Marine and Freshwater Research* 53, pp. 645–660.

Punzón, A, A Serrano, J Castro, E Abad, J Gil, and P Pereda (2011). "Deep-water fishing tactics of the Spanish fleet in the Northeast Atlantic. Seasonal and spatial distribution". In: *Scientia Marina* 75, pp. 465–476.

Punzón, A, C Hernández, E Abad, J Castro, N Pérez, and V Trujillo (2010). "Spanish otter trawl fisheries in the Cantabrian Sea". In: *ICES Journal of Marine Science: Journal du Conseil* 67.8, pp. 1604–1616.

Shono, H (2008). "Application of the Tweedie distribution to zero-catch data in CPUE analysis". In: *Fisheries Research* 93, pp. 154–162.

Skaug, HJ and D Fournier (2006). "Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models". In: *Computational Statistics and Data Analysis* 51, pp. 699–709.

Smyth, G (1996). "Regression modelling of quantity data with exact zeros". In: *Proceedings of the Second Australia-Japan Workshop on Stochastic Models in Engineering, Technology and Management*. Technology Management Centre, University of Queensland, pp. 572–580.

Stephens, A and A MacCall (2004). "A multispecies approach to subsetting logbook data for purposes of estimating CPUE". In: *Fisheries Research* 70.2–3. Models in Fisheries Research: GLMs, GAMS and GLMMs, pp. 299–310.

Tascheria, R, J Saavedra-Nievasb, and R Roa-Ureta (2010). "Statistical models to standardize catch rates in the multi-species trawl fishery for Patagonian grenadier (*Macruronus magellanicus*) off Southern Chile". In: *Fisheries Science* 105, pp. 200–214.

Williams, A, N Bax, and B Barker (2010). *Integrating fishing industry knowledge of fishing grounds with scientific data on seabed habitats for informed spatial management and ESD evaluation in the SEF*. Tech. rep. FRDC Report 2000/153. Fisheries Research and Development Corporation of Australia.

Winker, H, SE Kerwath, and CG Attwood (2013). "Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery". In: *Fisheries Research* 139, pp. 118–131.

— (2014). "Proof of concept for a novel procedure to standardize multispecies catch and effort data". In: *Fisheries Research* 155, pp. 149–159.

Wood, S (2006). *Generalized Additive Models*. Texts in Statistical Science. Chapman & Hall/CRC.

Woodhams, J, S Viera, and I Stobutzki (2012). *Fishery status reports 2011*. Tech. rep. Australian Bureau of Agricultural, Resource Economics, and Sciences, Canberra.

CONTACT US
t 1300 363 400
+61 3 9545 2176
e enquiries@csiro.au
w www.csiro.au

AT CSIRO WE SHAPE THE FUTURE

We do this by using science to solve real issues. Our research makes a difference to industry, people and the planet.

As Australia's national science agency we've been pushing the edge of what's possible for over 85 years. Today we have more than 5,000 talented people working out of 50-plus centres in Australia and internationally. Our people work closely with industry and communities to leave a lasting legacy. Collectively, our innovation and excellence places us in the top ten applied research agencies in the world.

WE ASK, WE SEEK AND WE SOLVE

FOR FURTHER INFORMATION

Oceans and Atmosphere Flagship
Scott D. Foster
t +61 3 6232 5178
e scott.foster@csiro.au
w http://www.csiro.au/en/Research/OandA

Oceans and Atmosphere Flagship
Mark V. Bravington
t +61 3 6232 5118
e mark.bravington@csiro.au
w http://www.csiro.au/en/Research/OandA