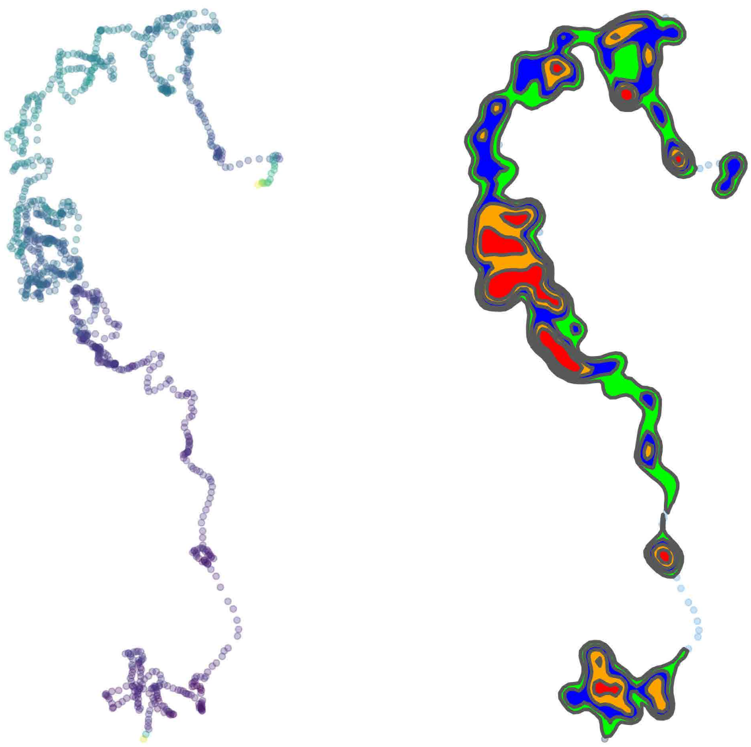
Can spatial fishery-dependent data be used to determine stock status in a spatially structured fishery?

Craig Mundy[1], Bill Venables[2], Malcolm Haddon[1,3], Catherine Dichmont[1,4]

[1]IMAS, University of Tasmania
[2]Venables Institute for Advanced Statistics
[3]Haddon.Net
[4]CDC-Consulting

November 7, 2024

FRDC Project 2017-026



Can spatial fishery-dependent data be used to determine stock status in a spatially structured fishery?

Craig Mundy, Malcolm Haddon, Catherine Dichmont, and Bill Venables

FRDC Project No **2017-026**

National Library of Australia Cataloguing-in-Publication entry

ISBN 978-1-922708-58-8

Ownership of Intellectual property rights

This publication (and any information sourced from it) should be attributed to Mundy, C.N., Venables, W.N., Haddon, M. and C. Dichmont (2024) *Can spatial fishery-dependent data be used to determine stock status in a spatially structured fishery?* IMAS, University of Tasmania, FRDC Final Report 2017-026 Hobart, 183 p.

Creative Commons licence

Disclaimer

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		FRDC Contact Details	
Name:	Dr Craig Mundy	Name:	Fisheries Research & Development Corporation
Insitution:	IMAS-FA		
Address:	15-21 Taroona Crescent	Address:	25 Geils Court
NA	Taroona, Tas, 7053		Deakin ACT 2600
Phone:	03 62268232	Phone:	02 6122 2100
Email:	Craig.Mundy@utas.edu.au	Email:	frdc@frdc.com.au
Web:	https://www.imas.utas.edu.au/	Web:	www.frdc.com.au

Table of contents

List of Figures

List of Tables

Acknowledgments

This project was supported by funding from the Fisheries Research Development Corporation on behalf of the Australian Government and was co-funded by the Institute of Marine and Antarctic Sciences (IMAS), University of Tasmania. This project received financial support from Natural Resources and Environment Tasmania (NRET), and from the Tasmanian Abalone Council Limited (TACL) to fund local development of GPS and depth data loggers to IMAS specifications. Luisa Forbes managed data logger allocation and retrieval to/from divers for the program from 2019. Luisa Forbes also provided critical support in the development, field testing, and finally deploying the next generation loggers. SciElex has accommodated and facilitated many of our requested changes in the firmware and helped to make this program a success.

Executive Summary

Capture of high resolution abalone fisheries data through the combination of passive GPS and depth data has been pioneered in Tasmania, with a 12 year time-series of mandatory use of the data logging hardware in Tasmania, covering around 90% of fishing activity. This project has exploited that time-series to address key challenges in assessment of abalone fisheries, particularly the global dilemma of quantifying catch rate hyperstability. The high-resolution spatial fisheries data have multiple applications, including the generation of new spatial performance measures, quantifying the spatial structure of our fisheries, as well as providing insights to the natural dynamics of dive fishery fleets.

Background

This project is the third in a series of projects aimed at developing the logic and data analysis workflow for electronic high resolution data in abalone fisheries. The first project (FRDC 2006/029) established a database backend/frontend for managing the deployment/retrieval of loggers to/from fishers, basic principals of handling and processing the electronic data stream in R, and structural components of two processed data formats. The second project (FRDC 2011/201) further refined the methods for processing the electronic data streams and examined basic properties of the spatial data (Chapters 5 and 6), as well as exploring a range of analytical applications to maximise the potential of the high-resolution data streams (Chapters 7 to 12).

Aims/objectives

Among the Australian states with a wild-harvest abalone fishery, Tasmania has the longest time-series of data collected from a GPS and depth data logger program (12 years at the completion of this project). Prior to incorporating this data in Fishery Assessments, there were two specific challenges to be addressed: 1. identify a statistical solution to quantify hyperstability in catch rates, and 2. develop a framework for incorporating spatial data into catch-rate standardisation. This project was focused on these two tasks, in addition to further exploration of questions that might be addressed using the electronic spatial data, in addition to refining our estimates of catch-rates.

There were four primary objectives identified: 1. Characterise the statistical properties, consistency, interpretability and assumptions of spatial and classic indicators of fishery performance. 2. Develop methods for inclusion of fine-scale spatial data in CPUE standardisations. 3. Identify methods for detecting hyperstability in CPUE. 4. Determine feasibility of spatial data based stock status determination in spatially structured fisheries.

Methodology

This project focused primarily on the dive polygon derived by extracting the 75th percentile isopleth from a bivariate Kernel Utilisation Distribution (KUD) applied to each discrete dive event. The area and maximum linear extent of the KUD polygon were extracted as candidate performance measures, while the centroid of each KUD was used to represent the spatial location of each dive event. The majority of work was focused around these two measures - polygon area, and polygon length. Data from 2012 to 2021 have been utilised in this report, with the most recent years excluded to protect time-sensitive recent fishing data.

Broader consideration of the spatial structure and dynamic of the fishery and the fleet, and possible new approaches, were examined using the Hexgrid product - where each hexagon contains the total catch and effort across all divers for each year in the time-series (see section 6.4 in Mundy 2012). For those unfamiliar with data products derived from the Tasmanian abalone GPS logger program, a detailed description is provided in Chapters 5 and 6 of Mundy, Jones, and Worthington (2018). For brevity, the detail of those basic methods and data descriptions are not repeated here.

This project spanned the COVID era, with work conducted through remote collaboration, and when permissible, face to face sessions. The project team was a collaboration between fisheries ecologists, statisticians and fisheries assessment scientists, combining local knowledge of the species, fisheries and data collection program with statistical and assessment expertise. We focused on a small number of regions within the Tasmanian Abalone fishery, particularly the North-West and central east coast as they provided contrasting temporal trends, and magnitude of change over time.

Selection of predictor variables was a key first step, with extensive investigation of environmental variables such as SST (sea surface temperature), wind speed and wave power, along with predictors that represent sub-units of statistical reporting areas. A further consideration once predictors are finalised is whether each term should be considered random or fixed, and the inclusion or not, of interaction terms (Bolker et al. 2009; Brown 2021; T. S. Clark and Linzer 2015; Silk, Harrison, and Hodgson 2020). Most abalone catch-rate data is log-normal, thus the question of whether to model the data as a gamma or log-normal with bias-correction on predicted means was an important consideration. While this made no difference to the outcome, the gamma approach while preferable in terms of simplicity, failed in several of the R packages available for mixed effect linear models.

All statistical analyses were conducted using R (R Core Team 2023) and several packages for mixed-effect models (lme4 Bates et al. 2015; glmmTMB Brooks et al. 2017; sdmTMB Anderson et al. 2022). This report is produced using Quarto within Rstudio (Allaire 2023).

Results/key findings

This project makes several important key findings. Firstly that, using passive GPS and depth data logging hardware, we are able to demonstrate that the effort component of the Tasmanian Abalone Fishery Docket returns are comparable to effort values determined from the GPS and depth data loggers. Given that daily catch is accurately weighed and reported, validation of the fishing effort component gives confidence in the long time-series of commercial catch and effort data currently utilised in the Tasmanian Abalone Fishery assessment. We were also able to demonstrate and quantify the existence of hyperstability in catch rate data, and, that hyperstability occurs when the stock is low and declining, but also when the stock is improving. While it has long been assumed that hyperstable catch rates are problematic in abalone fisheries, it has previously been unquantifiable and therefore difficult to address in Harvest Strategies. The current Tasmanian Empirical Harvest Strategy includes an asymmetric control rule as a precautionary approach to account for hyperstable catch rates when stocks are in decline. The findings from this project validate the inclusion of that measure and provide support for ongoing retention of an asymmetric control rule.

This project demonstrates that dive polygon area, rather than dive time, is a much more appropriate measure of effort. Consequently, future utilisation of the GPS program derived spatial fisheries data will utilise area in all calculations of catch rates, rather than as a predictor in the standardisation model. At the commencement of this project, we were intent on exploring how to incorporate spatial performance measure in catch rate standardisations. Other candidate spatial performance measures such as the number of dives per day were less useful, and are likely to have a non-linear relationship with stock abundance.

An exhaustive investigation of the value of including environmental predictors (appendices 13 - 15) in catch rate standardisations found little evidence at the scale of zones to support their inclusion. While sea temperature was not expected to be an important predictor of catch rates, wind speed and wave power were. There was however little direct evidence of an effect of these variables on mean annual catch rates, possibly because of the opportunity on any given day of fishing to find relatively sheltered locations to fish. Thus these two variables are likely to have an upper threshold, beyond which fishing is not practical (or safe) rather than as a direct determinant of catch rate. Further investigation of environmental influences on fleet dynamics is under way as part of a PhD project independent of this project.

Explicit inclusion of dive location in statistical catch rate standardisations appeared to have no material consequence on estimated annual mean catch rate. While spatial heterogeneity in abalone fisheries is well known, it was uncertain as to whether this mattered in the context of calculating catch rates. This finding simplifies the statistical modelling runs, as inclusion of spatial explicit terms either as a spline in a GAMM or as a spatial term in sdmTMB, has very high computing overheads. However as a secondary indicator, examination of the spatial distribution of residuals from spatially explicit Linear Mixed Effect Models provides valuable insights to the spatio-temporal dynamic within a Spatial Assessment Unit (SAU).

Inclusion of a rule-based process to identify distinct fishing grounds was trialled with some success. Abalone fisheries historically have used broad scale reporting areas (Blocks, Spatial Assessment Units, Spatial Management Units, Area). Over the past few decades, the realisation that there is considerable value in reporting at smaller scales has seen all abalone fisheries include finer scale statistical reporting areas, within which fishers must report daily catch and effort. These smaller areas (subblocks, mapcodes, reefcodes) are arbitrarily determined based on local knowledge, but often dissect contiguous rocky reef. An additional benefit of the GPS logger program is the ability to apply spatial clustering methods to identify discrete fishing grounds, resulting in a large number (mostly) of more appropriately bounded, small spatial units to capture structure within larger reporting areas.

Implications for relevant stakeholders

Development and validation of spatial performance measures from the Tasmanian GPS logger program has only been possible by the foresight of both the Tasmanian Government and the Tasmanian abalone fishery to establish GPS based data collection as a mandatory reporting requirement in 2012. This program, now in its 12th year has underpinned the success of this project. Other states wishing to capitalise on these findings are likely to be able to make use of a GPS data program within a shorter time-period, on the assumption that fishing operations and species are largely similar across states.

For Tasmania, this project provides the confidence to expand the Tasmanian Empirical Harvest Strategy to include catch rates based on area, and when using GPS logger data, to replace subblock terms with cluster ID to account for within SAU structure in catch rate standardisations. There remain challenges in terms of identifying Target reference Points (TRP) and Limit Reference Points (LRP) for catch rates calculated using *area* as the measure of effort. Proxy TRPs and LRPs may be defined by empirical relationships between dive time and dive area, and will be explored further as the time-series expands.

Docket book derived fishing effort for calculation of catch rates will remain a key component of Harvest Strategies across all jurisdictions, in order to capitalise on the long time-series of commercial Docket book data. It is essential that each jurisdiction include a mechanism to account for the presence of hyperstable catch rates in the mechanics of their harvest strategies, as has been done in Tasmania.

Keywords

Stock status, Blacklip Abalone, *Haliotis rubra rubra*, Tasmania, GPS, spatial regression, hyperstability, fleet dynamics

Objectives

1. Characterise the statistical properties, consistency, interpretability and assumptions of spatial and classic indicators of fishery performance.
2. Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.
3. Identify methods for detecting hyper-stability in CPUE.
4. Determine feasibility of spatial data based stock status determination in spatially structured fisheries.

Introduction

Abalone stock status assessment is intrinsically difficult because stocks are typically spatially structured, utilise small vessels, target sedentary species, and are prone to serial depletion. Fisheries with these characteristics have been termed S-fisheries by Orensanz et al (2005). Larval dispersal in the Blacklip Abalone *Haliotis rubra* is highly localised adding an additional spatial complexity. Obtaining representative biological data (growth, recruitment, mortality, emigration etc.) remains cost-prohibitive because abalone stocks are across extensive coastlines with substantial spatial structuring, requiring large numbers of sampling locations to capture the extent of variability. Local productivity of Blacklip is largely dependent on local abundance of reproductive individuals, and recovery from localised depletion is reliant on the remnant population as few external recruits will arrive to assist with population recovery (Miller, Maynard, and Mundy 2009). All of these features present significant challenges when assessing stock status.

Determining 'global' stock status requires some knowledge of the collective status of local populations within a stock. Accurately accounting for local dynamics (recruitment, growth, mortality) remains a challenge for model based abalone fishery assessments because it is unlikely that all local populations in a stock will share a global parameter set (see Mundy, Haddon, and McAllister 2023 for a more extensive discussion). Abalone assessments in many jurisdictions rely heavily on commercial catch-per-unit-effort (CPUE) data utlising fisher returns of catch and effort. Some jurisdictions have maintained a Fishery Independent Survey program, which can provide valuable performance measures in those jurisdictions. These programs are costly, and difficult to maintain representative coverage of the fished area. Given the micro-stock structure of abalone fisheries, use of fishery-dependent data to determine stock status is problematic because catch-rates are applicable only to the local population fished and are not indicative of 'global' stock levels (Parma et al. 2003). When there are sufficient fishing events, time-series of sample parameters (mean, distribution shape) appear to be a useful indicator of fishery performance sufficient to make Total Allowable Commercial Catch (TACC) decisions. However, without data on the spatial extent of the fishery and how the area of reef utilised by the fishery varies through time, it remains uncertain whether CPUE can provide a meaningful indicator of overall stock biomass, changes in fishing mortality that might indicate increasing levels of over fishing, or the effectiveness of management interventions. For this reason substantial effort has been invested in the development of spatial indicators, and theoretically, fine-scale fishery-dependent spatial data could be used in the development of a 'global' stock status indicator for abalone fisheries.

The introduction of the Status of Australian Fish Stocks (SAFS) process is generally considered to be an excellent initiative. However, the common reporting template is based on equilibrium concepts such as current biomass relative to unfished biomass (B_0), and relies on fisheries conforming to Dynamic Pool assumptions, which includes an assumption of a single homogeneous stock. Sedentary and sessile target species are unlikely to conform Dynamic Pool concepts (Seijo, Pérez, and Caddy 2004), and this is particularly applicable to abalone fisheries. The issue of lack of adherence to Dynamic Pool assumptions is an area both abalone scientists and abalone fishery managers have avoided, noting the challenges of factoring this understanding into both assessment and management of abalone fisheries. The lack of consideration in abalone and similar fisheries of just what a catch-rate indicator means in terms of stock status has created a significant stumbling block for scientists attempting to classify their fisheries as Sustainable, Depleting, Recovering, or Depleted. Collectively, our attempts to find a solution to the key issue - determining when a fishery has crossed a threshold from one category to another - have created tension, confusion, stress, and high levels of frustration across industry, management, research and FRDC.

Fishery-dependent data remain an important component of harvest control rules in Australian and New Zealand abalone fisheries (primarily catch-rates as Kg/Hr). However, with abalone, stable catch-rates may not indicate stable biomass and/or stable density. Catch-rates are frequently criticised because the effort (diving time) needed to take a quantity of catch may be influenced by density independent factors such as conditions at the time of fishing, experience, and the ability of fishers to adjust their fishing strategy to maintain catch rates (diver behaviour driven hyperstability). While there are many issues with the assumption that CPUE is a reliable proxy for abundance, it is assumed to be so despite the absence of robust data to validate use of CPUE in this way. In some jurisdictions CPUE is supplemented by sparse fishery-independent size and density data. Therefore, there is a need to review common assumptions, methods and interpretations of CPUE as a primary indicator, and to determine whether inclusion of spatial fishery data could provide a 'global' indicator of stock status for abalone fisheries.

This report examines the potential contribution to fishery assessments by using spatially explicit abalone fishing data. The project utilises spatial data collected by the Tasmanian abalone fishery GPS logger data program, along with catch and effort data from fisher docket returns. In particular we address the issue of hyperstability, which was the prime motivator of developing the passive GPS and depth data logger program in Tasmania in 2005. We also explore the potential of spatially explicit abalone fishing data to highlight spatio-temporal dynamics of fishing.

Part I

Performance Measure Properties

1 Properties of candidate spatial performance measures

1.1 Introduction

1.1.1 Review of classic and spatial effort variables

One of the primary motivations for pursuing the passive GPS and depth data logger concept was to improve the consistency and quality of how fishing effort is captured. Weather conditions are a strong determinant of where fishers choose to dive, and only rarely are conditions perfect. There may also be multiple dives over a section of coast, making the fishers task of estimating total diving effort at the end of the day challenging. In addition, different divers have different ideas of what counts as fishing effort. Thus, we expect effort recorded in divers' fishing docket books to be an approximation. The GPS logger program provides two forms of effort:

- Diving time, accurately measured by the automatic depth data logger
- Area, as measured by the area of the vessel footprint for each dive

Of interest is whether these variables are correlated with the measure of effort (hours fished) recorded by fishers on the docket returns. Understanding the degree of correlation will provide insights into whether earlier estimates of catch-rates might potentially be biased in a particular manner.

This chapter links to Objective 1: **Characterise the statistical properties, consistency, interpretability and assumptions of spatial and classic indicators of fishery performance.**

1.2 Methods

Two reporting blocks in the Tasmanian abalone fishery are used as contrasting case studies - Block 6 in the north west, and Block 21 in the south-east. The fishery in Block 6 has declined rapidly over the past 10 years, although with ~90% reductions in TACC, the fishery has recently shown increases in catch-rates indicating a degree of rebuilding. Block 21 in the south-east was rebuilding at the start of the time-series, however its fishery was substantially impacted by two external events in 2016; the first was the most significant marine heat wave recorded in recent times (Oliver et al. 2017), and the second was a significant storm generating the largest easterly swells on record. Nevertheless, from 2017, stocks appeared to continue rebuilding through to 2021. These two blocks provide contrasting trends in fishery performance as a way to examine the behaviour of the modelling approach under different circumstances.

A simple graphical comparison of the distribution of daily effort (dive time) derived from docket books and GPS/depth data loggers is used to illustrate the relationship, and examine how each form of effort estimate relates to other variables such as area and maximum extent of the dive polygon, and to daily catch.

1.3 Results

1.3.1 How accurate are Tasmanian abalone fisher Docketbook effort records?

Calculations of the percent difference between Docketbook estimates of effort recorded by divers and precise estimates of effort recorded by passive depth data loggers, suggests that Tasmanian commercial abalone divers provide surprisingly accurate estimates of actual fishing effort given the difficulties of working from small runabout fishing vessels (Figure 1.1). The proportion of docket records with ± 25% of the GPS estimates is over 80% in both blocks (Table 1.1), giving a degree of confidence in the Docketbook estimates of fishing effort. The discrepancy between docket and logger estimates of effort are used to filter data before final use for spatial Performance Measures. Where both GPS and depth data logger data are available and electronic measure of effort is consistent with the docket record of effort, those records are retained. Where GPS data is present and depth data are absent, those GPS data are processed with track classification algorithm to create a virtual depth data stream, and appended to the full data set. Where GPS data are missing but depth data are present, those data are removed from further analyses.

Data loss can exert a potential bias depending on whether the origins are human failure or hardware failure. Where there is a hardware failure, logger sets were replaced and the faulty units repaired and returned to a pool of spare units. Where there is human failure such as consistently forgetting to turn loggers on, or forgetting to charge units prior to fishing days there is the potential for bias in capture of fishing effort where we will achieve a lower coverage of data from fishers that are less diligent. There are multiple triggers for correct use of loggers, including reporting the serial ID of the logger hardware being used on the daily fisher returns, and a question in the pre-fishing phone report that asks the fishers to confirm that the loggers are charged and turned on. Hardware failure is the dominant cause of data loss, and we expect that to be unbiased with respect to fisher and region.

Marginal plots of the relationship between Docketbook effort and logger effort provide support for a relatively close relationship between Docket Book and logger forms of fishing effort (time) in both Block 6 and Block 21 (Figure 1.2, Figure 1.3). The scatter of points well below the line (in sub-plot 'a' in both cases) indicates days where there is proportionally short logger time compared to Docket Book time. These are primarily due to low battery or hardware failure, although are relatively minor occurrences (e.g. Figure 1.2 a).

Fishing effort recorded in docket returns is often rounded to the nearest half hour or hour, and occurs across the full range of daily fishing effort (Figure 1.2 (a), Figure 1.3 (a)), regardless of whether the fishing day was short or long. The relationship between area and time is non-linear, with increasing variance as area or effort increase (Figure 1.2 (c), Figure 1.3 (c)). The area of the dive polygon is linearly correlated with the maximum linear extent, indicating that either of these two variables would have a similar effect in a statistical model (Figure 1.2 (d), Figure 1.3 (d)).

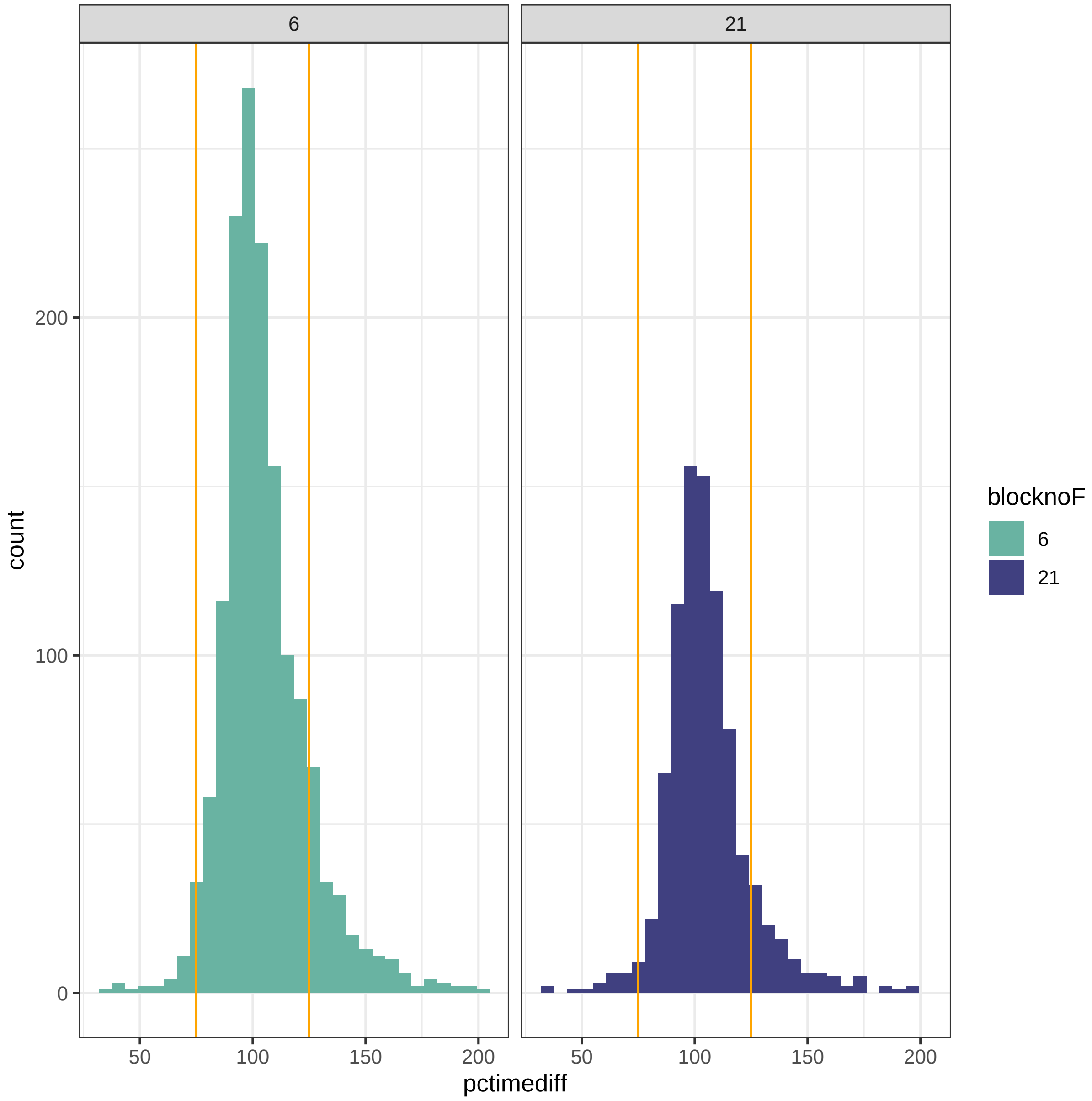


Figure 1.1: Distribution of percent time differences between daily effort reported from divers Docketbooks and the GPS/depth dataloggers. Orange vertical lines indicate region of 75% and 125% differences.

Table 1.1: Distribution of differences between daily effort reported from divers Docketbooks and the GPS/depth dataloggers. Notes: allrecs = all matching records; Recs25 = proportion of records where logger time was within 25% of the docket time; pc25 = Recs25 as a fraction of total records.

blockno	allRecs	Recs25	pc25
6	1560	1273	81.6%
21	914	766	83.8%

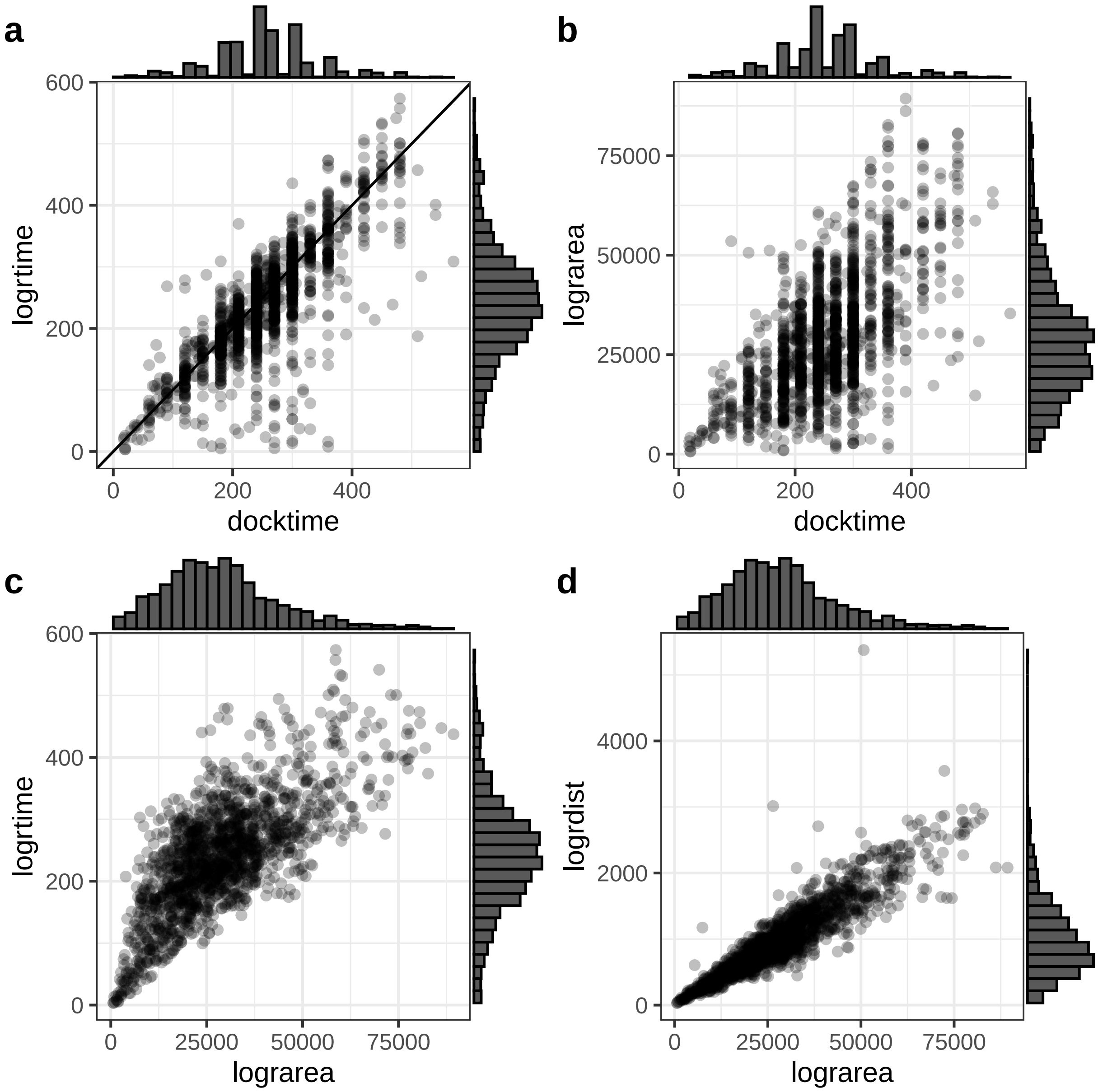


Figure 1.2: Relationship between different measures of effort for Block 6, Tasmanian Northern Zone. a) docktime vs logrtime - relationship between daily effort (minutes) from docket and logger; b) docktime vs lograrea - relationship between docket effort and dive polygon area; c) lograrea vs logrtime - relationship between logger area and logger time; d) lograrea vs logrdist - relationship between logger area and logger distance.

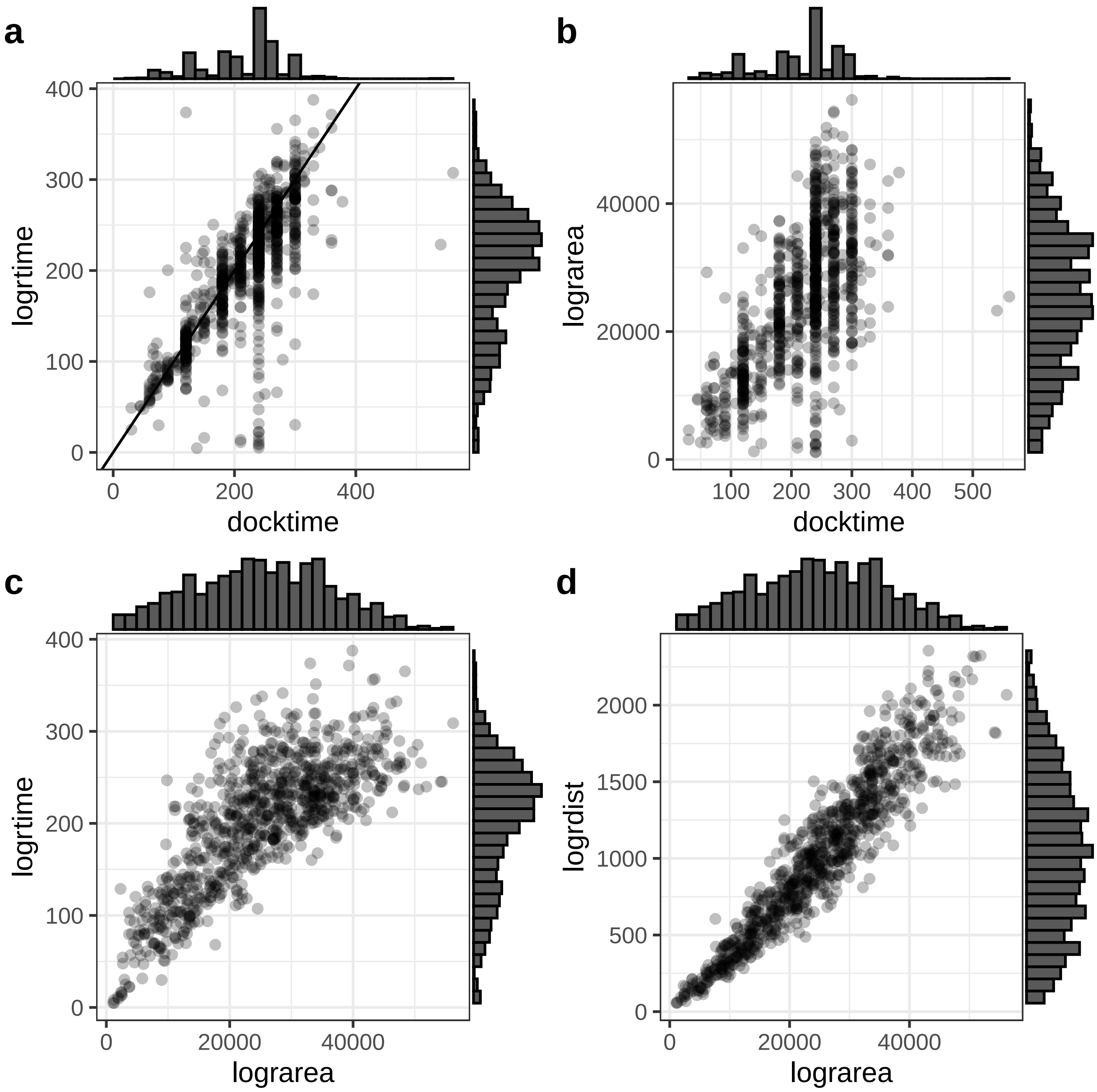


Figure 1.3: Relationship between different measures of effort for Block 21, Tasmanian Eastern Zone. a) docktime vs logrtime - relationship between daily effort (minutes) from docket and logger; b) docktime vs lograrea - relationship between docket effort and dive polygon area; c) lograrea vs logrtime - relationship between logger area and logger time; d) lograrea vs logrdist - relationship between logger area and logger distance.

1.3.1.1 Relationship between catch and time by data type

Catch broadly increases with fishing time (logger) in both blocks (Figure 1.4, Figure 1.5). Banding of the fishing time from Pocketbooks can be seen in plot b) again for both blocks (Figure 1.4, Figure 1.5), whereas catch appears as a continuous variable with no categorisation. This is expected given that catch is weighed to the nearest kilogram at the boat ramp (or jetty), during the transfer process from diver to processor. Although, all catch weight values in this study (and in the annual fishery assessment) are estimated weights, the fishers are much more adept and highly accurate at estimating weight of catch from each sub-block visited over the fishing day.

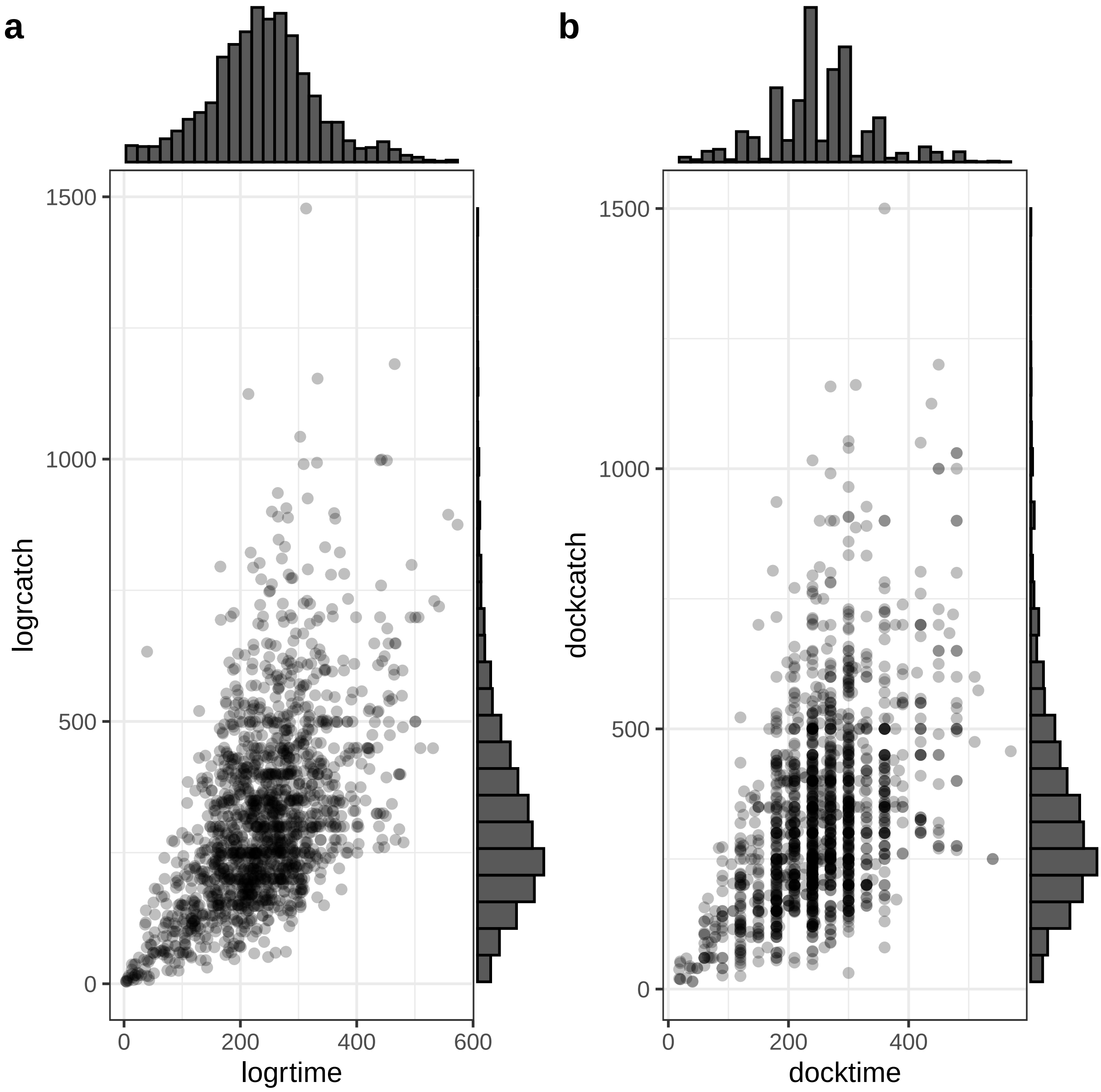


Figure 1.4: Relationship between catch and effort (logger time and docket time) for Block 6, Northern Zone; a) logger effort (minutes) and catch (kilograms), b) Docketbook effort (minutes) and catch (Kilograms). Note: daily catches > 1550Kg removed.

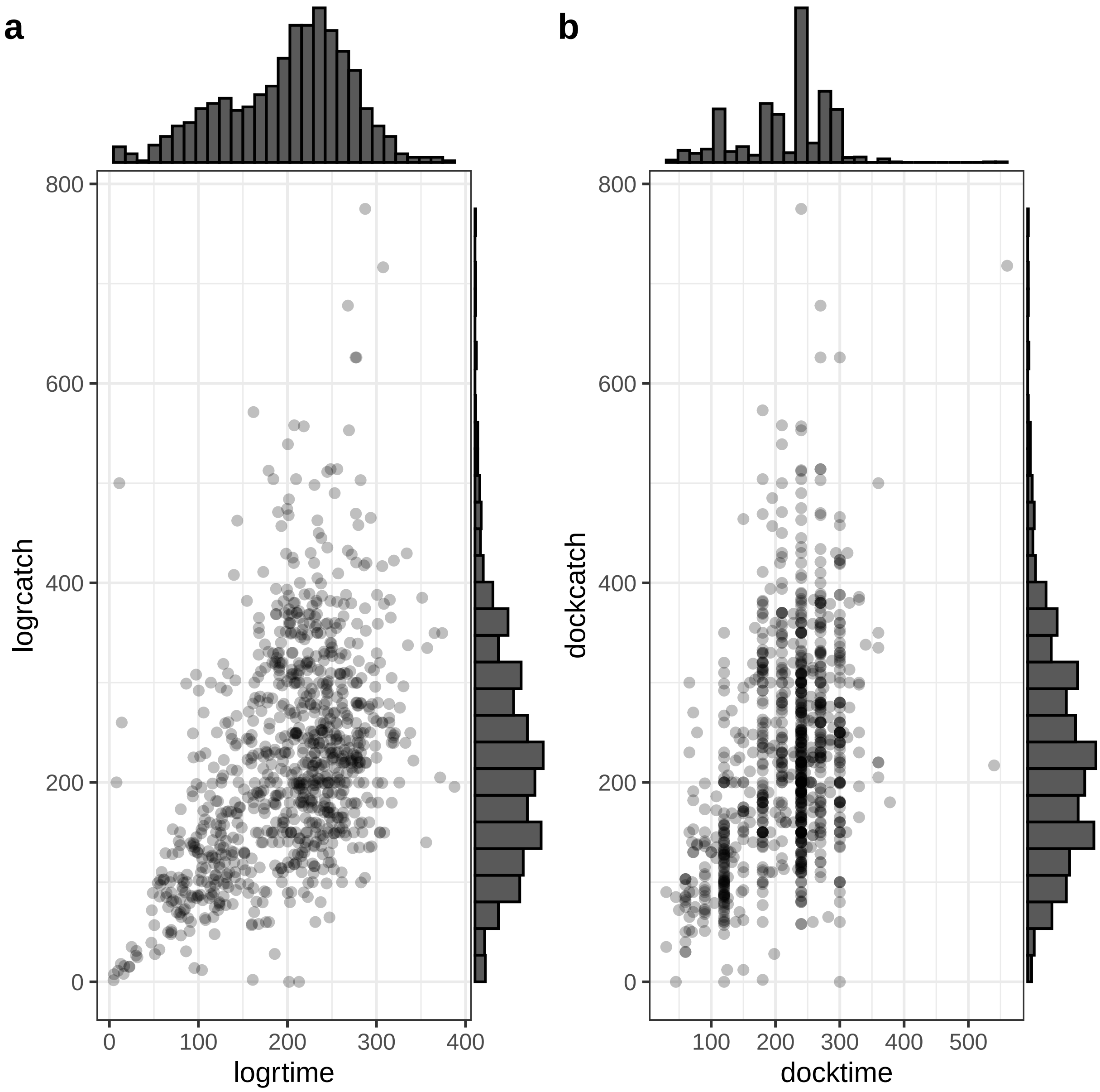


Figure 1.5: Relationship between catch and effort (logger time and docket time) for Block 21, Eastern Zone; a) logger effort (minutes) and catch (kilograms), b) Docketbook effort (minutes) and catch (Kilograms

1.3.1.2 Relationship between catch and area

There is considerable variation in the catch-dive area relationship in both block 6 and block 21 (Figure 1.6). This reflects high levels of variation among locations, fishing days, fishers, and state of the fishery at the time of fishing.

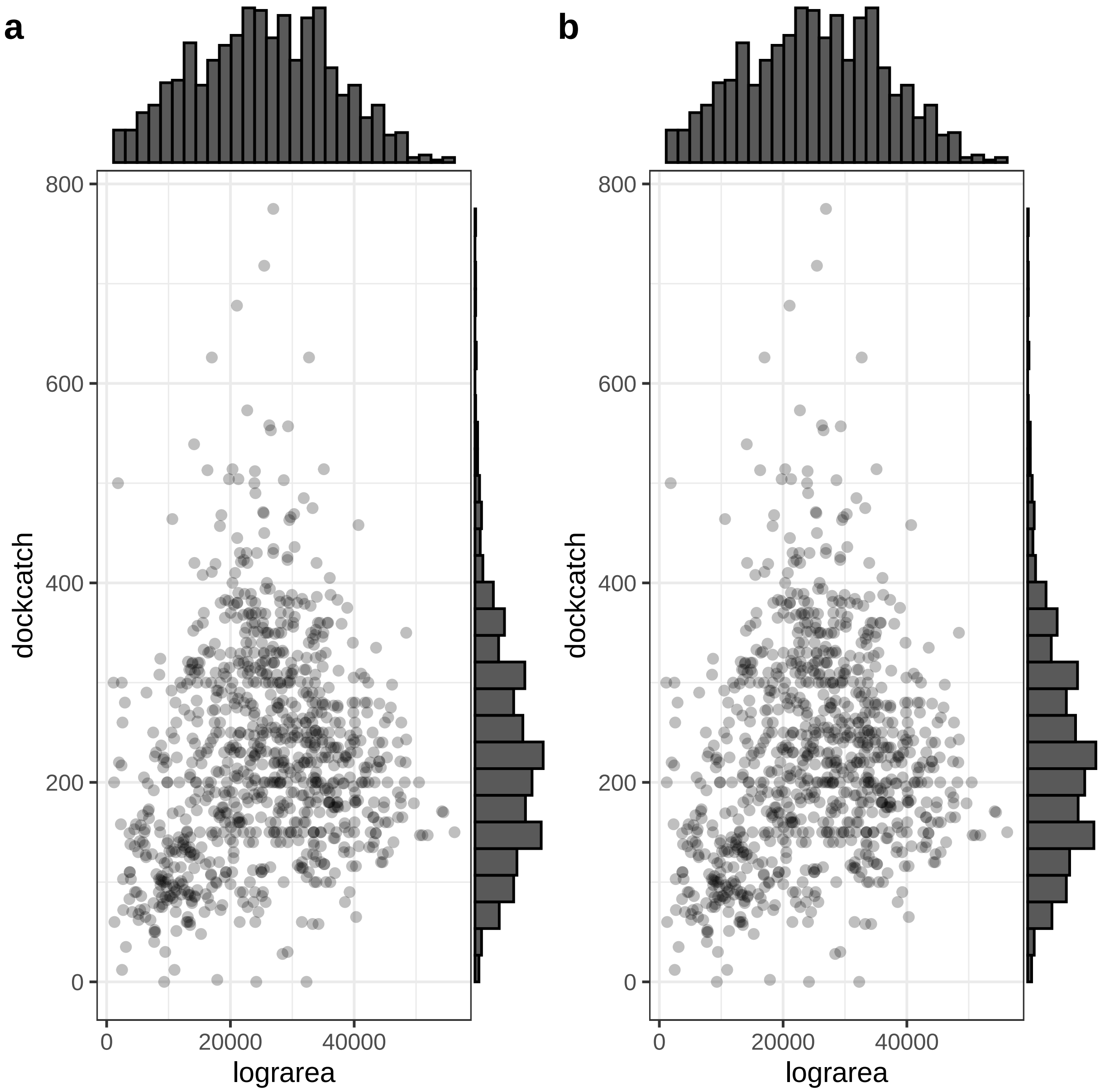


Figure 1.6: Relationship between reported Docketbook daily catch (dockcatch) and daily area from GPS loggers (lograrea) for Blocks 6 and 21; a) Block 6 logger area (square metres) and daily catch (kilograms), b) Block 21 logger area (square meters) and catch (Kilograms). Note: daily catches > 1550Kg removed.

1.4 Discussion

The Tasmanian abalone divers on the whole are providing relatively accurate estimates of fishing effort when recording hours fished on their daily Docket returns. Being mindful of the challenges of remembering how much time was spent underwater on each dive, and in each sub-block at the end of the day, the differences between the fisher memory (docket) and automatic logger (depth logger) estimates of fishing time were much less than thought likely. There is a risk that this high level of agreement (excluding biased records due to logger failure) is influenced by the knowledge that effort is being recorded electronically, leading fishers to be more careful with estimating there effort. Most Tasmanian abalone fishers have been using dive computers for the past two decades, and in most instances it assists them with estimating total hours in the water for their Docket returns. Thus, there is a level of confidence that we can infer a level of accuracy in effort reporting for the period prior to the introduction of GPS and depth datalogger hardware.

The agreement between diver and logger estimates of effort, despite clear patterns of rounding, provides confidence that the Docket returns from fishers contain meaningful estimates of their fishing effort. This is an important conclusion from this study in that there is now greater confidence in the longer time-series of catch and effort data available from the Docket records, while a time-series of logger based measures of effort is developed. It is also clear that relationships between several variables (fishing time, area, distance, catch) are either non-linear, and with increasing variance with increasing magnitude of the variables. This is a clear indication that there are other factors that influence daily catch in addition to effort. These will include human variables such as knowledge of the coast, experience of the deckhand, and also, physical conditions such as sea conditions, wind, etc). There is also likely to be a contribution from small- and regional-scale variation in abundance, and prior fishing activity at the dive site. These issues are explored further in the remaining chapters.

Part II

Optimising use of spatial dive event data

This section of the report is focused on identifying a base catch rate standardisation model, advantages of pursing different scales of aggregation, and finally on characterising and describing spatial performance measures derived from bivariate kernel utilisation distribution polygons of individual fishing events (KUD). Here we directly explore the potential to replace time as a unit of effort with dive polygon area.

2 Catch-rate standardisation strategy

2.1 Introduction

Central to any assessment of fishery performance is the availability of a valid time-series of an index of relative abundance. All Australian abalone fisheries assume that commercial catch-rates provide such an index, although in some cases, this is assumption should have caveats relating to hyperstability, a lack of contrast through time following the introduction of quota systems, and the inherent variation expected with abalone fishery statistics.

Before use in an assessment, ideally, catch-rates should be standardised. Statistical standardisation of catch-rates seek to identify and account for variation in the available data that is not a feature of the temporal trend in abalone abundance, and may include variables relating to gear types, fisher identity or spatial and environmental variables that may exert an influence on catch-rate trends (Hsu, Chang, and Ducharme-Barth 2022; Maunder and Punt 2004; Tanaka et al. 2022a; Thorson et al. 2023). Given the wildly different character of different commercial fisheries, there is no single best approach to such standardisation. For example, there are many different views on what predictor variables should be used, whether variables should be considered fixed, continuous or random, and the extent of any interactions between terms in the model. In the literature, that discussion appears never-ending and at least partly reflects the many different situations exhibited by different fisheries. Whether such discussion has value in practical terms is also debatable. For example, in many cases, providing the statistical model is constructed validly, the choice of whether to include a term as a random or fixed effect often makes relatively little difference.

This chapter establishes a common base statistical mixed effects model to be applied to both docket- and logger-based datasets. Additional predictor variables may be added to the base model to address questions of specific interest, or for specific regions, or even individual assessment areas within the fishery.

This chapter links to two objectives:

Objective 1: **Characterise the statistical properties, consistency, interpretability and assumptions of spatial and classic indicators of fishery performance.**

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.**

2.1.1 Statistical Models

The traditional performance measure in abalone fisheries is catch per unit effort (CPUE), with Effort represented by time (in water dive time) and Catch as kilograms of abalone landed, summed across all dives for a day.

"The Trouble with Ratios: an Overview

Unless there is reason to believe that the regression passes through the origin, the ratio is of dubious value. It may do for rough work, but careful experimentation deserves a more efficient statistical method."[1]

While we use the ratio CPUE religiously in abalone fisheries, it might be argued that the actual metric of interest is, how much catch can be achieved for a given unit of effort rather than the ratio of catch and effort. A more relevant question to ask of the data is, has the level of catch achieved per unit of effort changed over time. In the Tasmanian abalone fishery divers set out for the day with a known amount of catch to achieve and make decisions on where to fish based on their prior knowledge of where they are likely, or hopeful, of achieving that catch within the dive time allowed (or in sufficient time to meet the processor truck at the boat ramp at the pre-agreed time). Divers will also count as they fish, knowing that for every minute they spend in the water they need to achieve a certain level of catch. Divers, however, are familiar with and understand the ratio CPUE, perhaps largely because we have trained them over the past four decades to think this is the best way to examine performance of an abalone fishery.

The classic approach for statistical modelling of catch-rates is to use the ratio of catch per unit effort (CPUE) as the dependent (Mod0), with a number of explanatory or independent variables which may be categorical or continuous, and fixed or random depending on ones philosophical leaning. Abalone fisheries are well known for high levels of spatial variation in productivity and catch-rates, and this is addressed by including a within SAU predictor, in this case sub-block ('subblockno'). Fishers also have different strategies and diving behaviour, while in most locations we expect catch-rates to be influenced by season, for example higher levels of seaweed biomass during summer months affects catchability. We expect "among diver" and "among season" to contribute substantially to overall variation, and predictors are included for both these factors. Previously, the typical abalone standardisation model for example was:

```
Mod0 <- CPUE ~
        fishyear +
        subblockno +
        diver_id +
        fishmonth
```

Where all predictors are treated as fixed effects.

As most of our predictors are included to account for variability other than the effect of year, we should consider whether our additional predictors should be included as random effects. Where the number of levels in a random predictor is small (i.e. < 5) the practical difference between treating subbockno as either fixed or random will be small. The estimated values of a random effect are known as 'BLUP's (Best Linear Unbiased Predictions), and can be considered equivalent to parameter estimates for a fixed effect. The major difference between treating a predictor as a random or fixed effects is that the BLUPs are penalised, and provide some protection for levels of a factor that might be quite disparate to other levels, particularly where the number of samples in that level are marginal. Overall, treating a predictor such as subblockno as random takes advantage of a phenomenon known as 'borrowing strength', by assuming some similarity between individual levels (see for example, common trends among sub-blocks below in Figure 2.2). This provides a desirable feature for many of the SAUs where there is considerable heterogeneity in where fishers are active among seasons or among years, and often with small numbers of fishers with extensive local knowledge.

[1] George Snedecor (in: Curran-Everett (2013) p. 213)

Rather than use the CPUE as a ratio response variable, an alternative option is to model catch as the response variable and specify effort as an offset in the right hand side of the equation. This provides an identical outcome to using the catch-rate ratio as the response variable, but provides more flexibility when back-casting to predict annual mean catch-rate as part of the standardisation process. Rather than producing a time-series of annual mean CPUE, we can instead examine a time-series of the relative change in catch achieved for a given unit of effort (e.g. 1 hour), where log(effort) is included as an offset in the linear model. When we back-cast to estimate the catch achieved for 1 hour of effort, the output of this formula structure is identical to that from a model with the ratio CPUE as the dependent (i.e. Mod0). Combining these two alternative approaches (mixed effects, offset(log(catch))), we have a model form such as (Mod1).

```
Mod1 <- log(catch) ~ offset(log(time))  +
        fishyear +
        (1 | subblockno) +
        (1 | subblockno:fishyear) +
        (1 | diver_id) +
        (1 | fishmonth)
```

2.2 Methods

All mixed effect model and spatial analyses were conducted using R statistical software (R Core Team 2023). The sf package (Pebesma and Bivand 2023) was used to manage spatial objects within R throughout. A mixed-effects model (Mod1) was applied to the GPS logger time-series, with data aggregated by diver fishing day within the smallest scale reported on docket books (i.e. sub-blocks). A Log-normal model is used throughout, implemented using the r package glmmTMB (Brooks et al. 2017). The performance package (Lüdecke et al. 2021) was used to obtain diagnostics and Conditional and Marginal R^2 values, and Intra-class correlation coefficents for the fitted models. The mixedup package (M. Clark 2023) was used to extract variance components for random effects. Standardised catch-rates are plotted with the ggplot package (Wickham 2016), and multi-plot arrangements were composed with (Pedersen 2023). General functions within the tidyverse (Wickham et al. 2019) as well as base R were used to coerce and re-shape data as required. Tables were produced using the flextable package (Gohel and Skintzos 2023). Data were batch processed using custom functions and the purrr package (Wickham and Henry 2023).

An informal proportionality check (IPC) of catch and effort variables can be calculated with a minor re-arrangement of model terms, with offset(log(catch) replaced with log(catch) such that log(catch) becomes a continuous term in the model rather than as an offset. The IPC is the coefficient for log(catch), with an expected value of 1, which would confirm that a unit increase in effort achieves a unit increase in catch (Venables, pers comm). Marginal and conditional R^2 values are also calculated to provide an indicator of how well the model captures variation in the dataset. Marginal R^2 indicate the variance accounted for by the fixed effects in the model and the Conditional R^2 indicate the variance accounted for by the full model (fixed and random). Standard Residual and QQ plots are used to test whether assumptions of normality and variance are met adequately.

Statistical properties and diagnostics are explored in detail for a subset of Spatial Assessment Units (SAU). The IPC, R^2 values, and contribution of random effects are conducted on a selection of SAUs, while more detailed analyses are done for two SAUs - SAU5 and SAU11. The following base model formula was applied to each of the example SAUs:

log(catch_est) ~ offset(log(logrtime)) + fishyear + (1 | subblockno) + , (1 | subblockno:fishyear) + (1 | diver_id) + (1 | fishmonth)

A consideration of a range of standardisation approaches and the contribution of several environmental variables are detailed in the Appendix (Chapters 10-12).

2.3 Results

2.3.1 Informal proportionality check

IPC values for all SAUs examined here are close to the expected value of 1. Two notable exceptions are SAU21 and SAU49 where the IPC is 0.94 and 0.91 respectively (Table 2.1). This simple check provides confidence that the data conform to the expectation that catch is proportional to effort.

Table 2.1: Logger data: Informal proportionality check for three Spatial Assessment Units (SAUs) in the Tasmanian abalone fishery.

SAU	ipc
SAU11	0.98
SAU12IN	0.98
SAU13	1.00
SAU21	0.94
SAU49BS	0.91
SAU5	1.03
SAU53	1.00
SAU7	1.03

Table 2.2: Docketbook data: Informal proportionality check for three Spatial Assessment Units (SAUs) in the Tasmanian abalone fishery.

SAU	ipc
SAU11	0.97
SAU12IN	0.95
SAU13	0.91

SAU	ipc
SAU21	0.84
SAU49BS	0.65
SAU5	0.91
SAU53	0.93
SAU7	1.01

2.3.2 Marginal and Conditional R^2

Marginal R^2 (~ fixed effects) was low for all SAUs examined (Table 2.3). In the base model, the only fixed effect is Year (fishyear), with random effect terms accounting for much of the variation. The Conditional R^2 (both fixed and random effects) ranged from 0.30 to 0.70 for the SAUs examined. The relatively low level of variance explained suggests there are drivers not included in the model contributing to variation in the dataset. These may include spatial or environmental drivers, or, may reflect the high levels of variation commonly observed in commercial fisheries data (See(Figure 3.2 in Chapter 3 for an expansion of this issue).

Table 2.3: Marginal (R2_M) and Conditional(R2_C) R^2 for three Spatial Assessment Units (SAUs) in the Tasmanian abalone fishery.

SAU	R2_marginal	R2_conditional
SAU11	0.07	0.30
SAU12IN	0.06	0.35
SAU13	0.10	0.46
SAU21	0.15	0.55
SAU49BS	0.07	0.70
SAU5	0.21	0.45
SAU53	0.12	0.59
SAU7	0.09	0.47

2.3.3 Intra-class correlation coefficient

Of interest is the whether replicates within a grouping variable (random effect) resemble each other. The Intra-class correlation coefficient (ICC) provides a measure of within-group repeatability i.e. are records within a level within a group more similar to each other than to records in another level within the same group. Notably, our model term for seasonality (fishmonth) suggests there is very little repeatability within months (Table 2.4). ICC values close to zero suggest there might be little benefit in setting that model term as a random effect. Only two terms - diver_id and subblockno show moderate levels of repeatability, but the patterns are not consistent across SAUs.

Table 2.4: Intra-class correlation coefficient of individual random effects in the model.

SAU	diver_id	fishmonth	subblockno	subblockno:fishyear
SAU11	0.19	0.02	0.00	0.03
SAU12IN	0.20	0.01	0.08	0.02
SAU13	0.28	0.11	0.00	0.01
SAU21	0.40	0.01	0.03	0.03
SAU49BS	0.03	0.01	0.58	0.04
SAU5	0.13	0.02	0.11	0.05
SAU53	0.05	0.06	0.37	0.06
SAU7	0.14	0.01	0.17	0.09

2.3.4 Contribution of individual random effects to the model

For most SAUs diver identity was the most important random effect in the model, whereas the sub-block effect (representing areas within each SAU) was relatively unimportant. Whereas for SAU49 and SAU53, the relative importance of these two random effects were reversed (Table 2.5). Month of year was of negligible importance across all the SAUs examined here, with SAUs representing western, eastern, northern, and Bass Strait regions of the Tasmanian Abalone fishery. The relative contribution of random effects are almost identical to the ICC values.

Table 2.5: Contribution (% explained) of individual random effects to the model.

SAU	diver_id	subblockno:fishyear	fishmonth	subblockno	Residual
SAU11	0.192	0.027	0.024	0.000	0.756
SAU12IN	0.203	0.021	0.014	0.076	0.686
SAU13	0.277	0.009	0.110	0.001	0.603
SAU21	0.400	0.029	0.014	0.029	0.527
SAU49BS	0.033	0.041	0.013	0.585	0.328
SAU5	0.130	0.054	0.016	0.106	0.694
SAU53	0.047	0.057	0.064	0.367	0.465
SAU7	0.143	0.087	0.012	0.172	0.586

2.3.5 Case Study: SAU5 Northern Zone (Blocks 5 & 6)

Model diagnostics for the SAU5 dataset indicate reasonable adherence to assumptions required of data analysed with linear mixed-effect model. Normality of the residuals is perhaps the most concerning (Figure 2.1), although typical of fisheries data where the density of data points is sparse at the lower end and to a lesser extent at the upper end. Noting that a log-normal model is applied, the diagnostics for this linear mixed-effect model don't suggest any cause for concern that might require action.

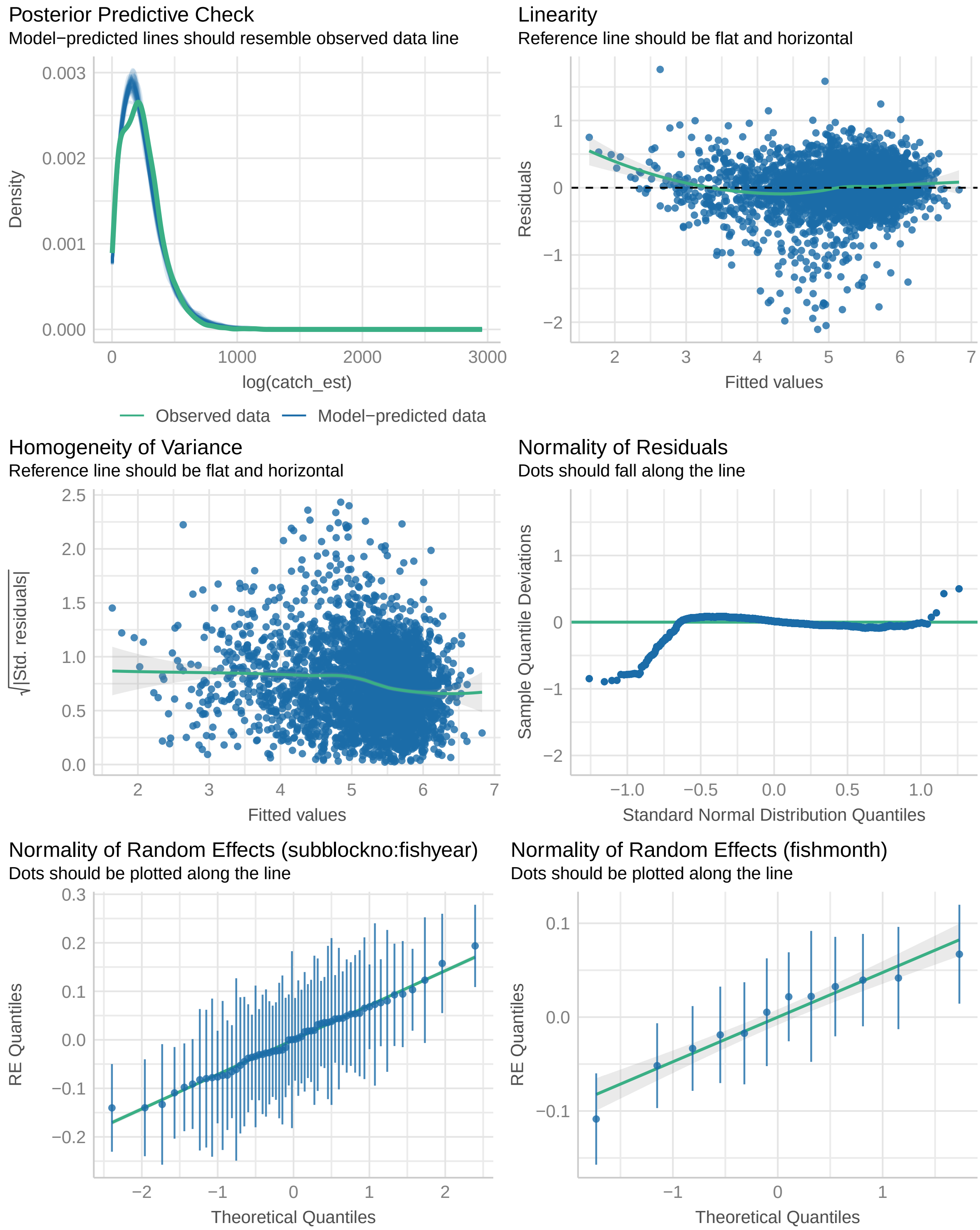


Figure 2.1: SAU5: Model diagnostics)

2.3.5.1 Standardised catch per 1 hour of diving effort

An advantage of the base statistical model is that the SAU wide trend can be presented, while accounting for variation among sub-blocks within the SAU. Additionally, via the partial interaction term (1 | subblock:fishyear) we can also examine the trends of individual sub-blocks across the same time-series (Figure 2.2). Sub-block 5A, while one of the top two sub-blocks in this SAU in terms of production, has always had a substantially lower catch-rate (Figure 2.2, Figure 2.3). Despite this, the trend in catch per 1 hour of effort is very similar across all subblocks.

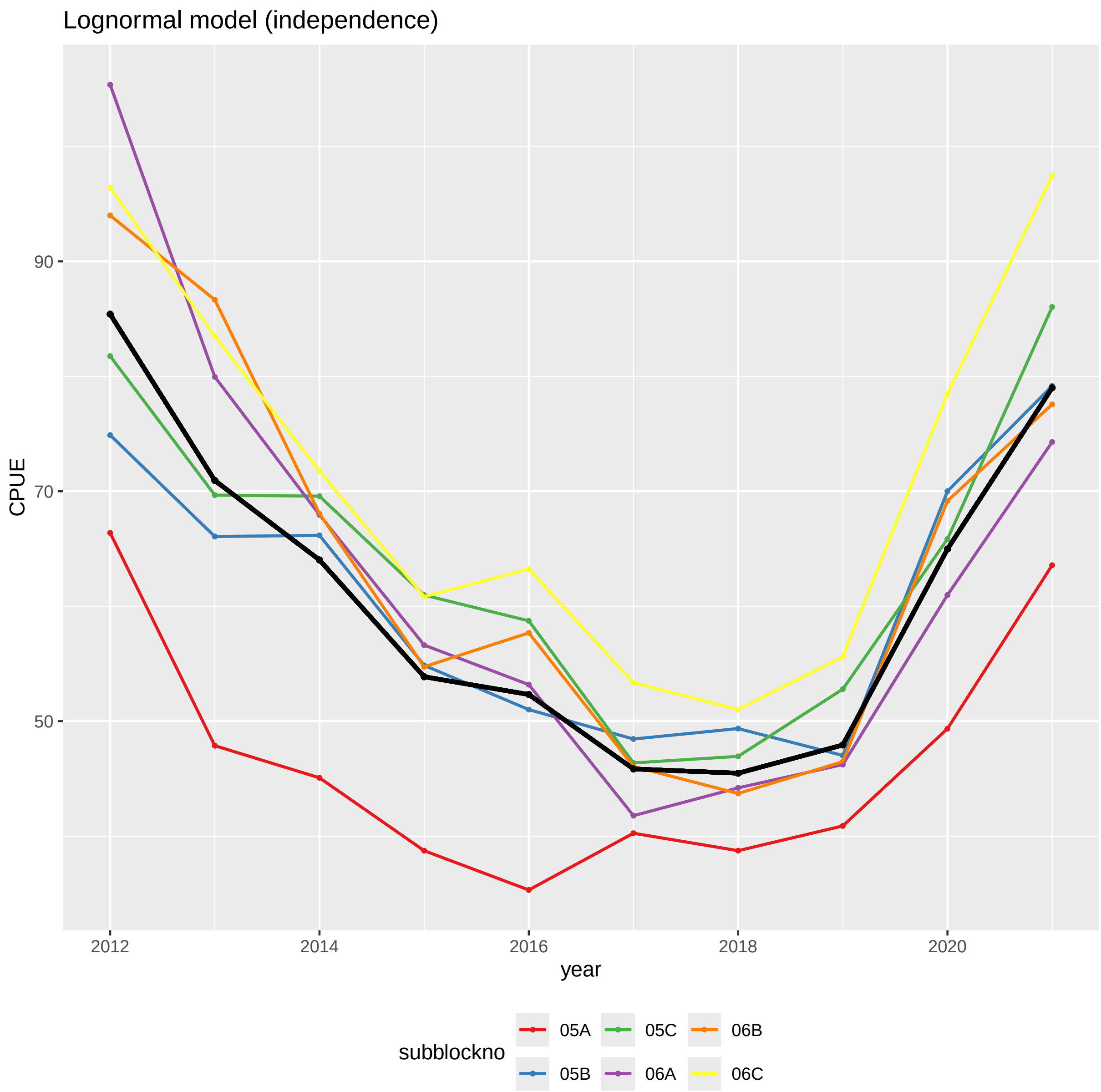


Figure 2.2: SAU5: glmmTMB - Log-normal CPUE standardisation (including bias correction). Response variable is log(catch), with log(effort) included as an offset. Black line indicates SAU mean catch-rate, and the coloured lines relate to the trend in each specific subblock, and the coloured lines relate to the trend in each specific subblock.

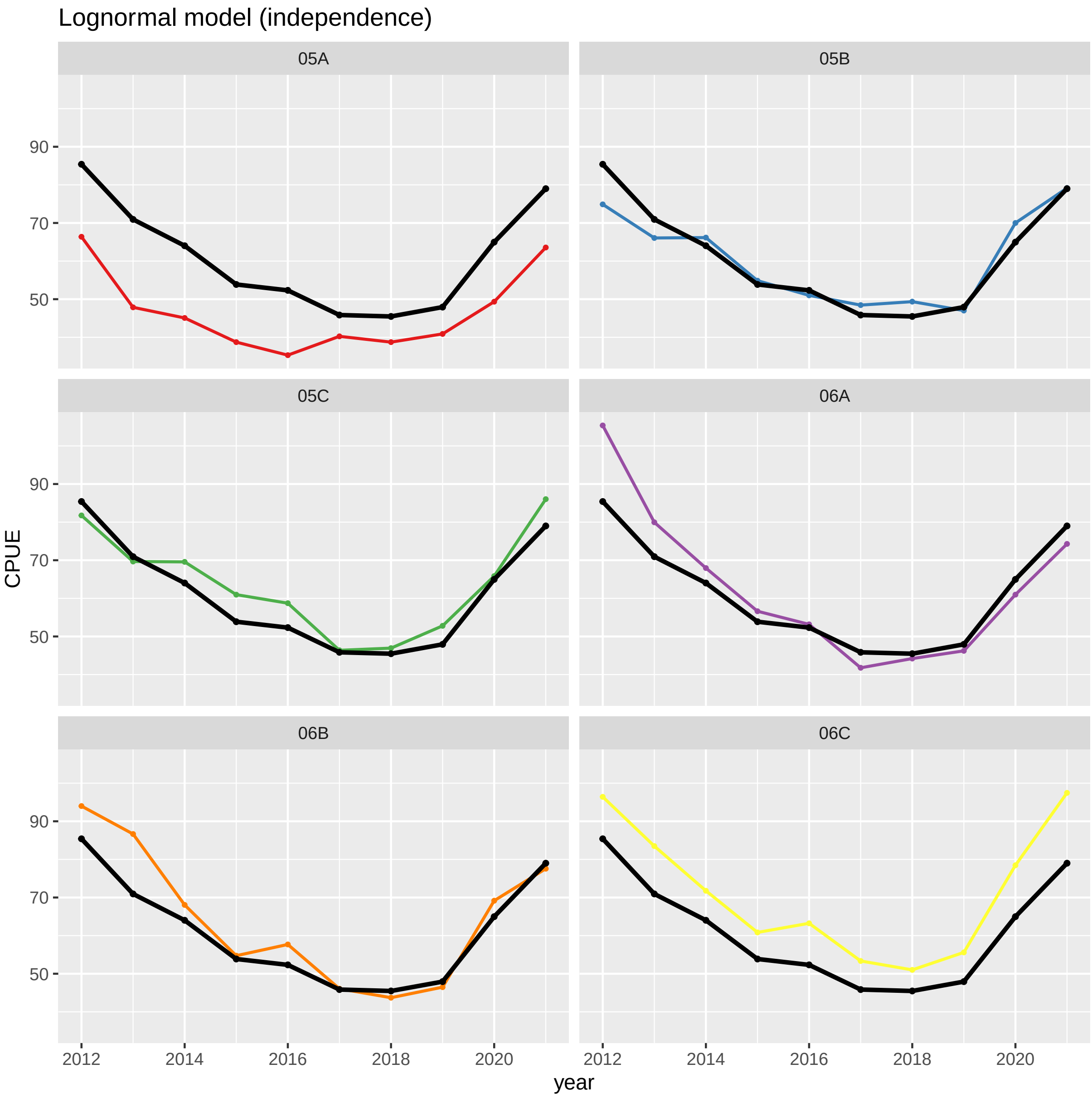


Figure 2.3: SAU5: glmmTMB - Log-normal CPUE standardisation (including bias correction) - facet plot. Black line indicates SAU mean catch-rate, and the coloured lines relate to the trend in each specific subblock, and the coloured lines relate to the trend in each specific subblock.

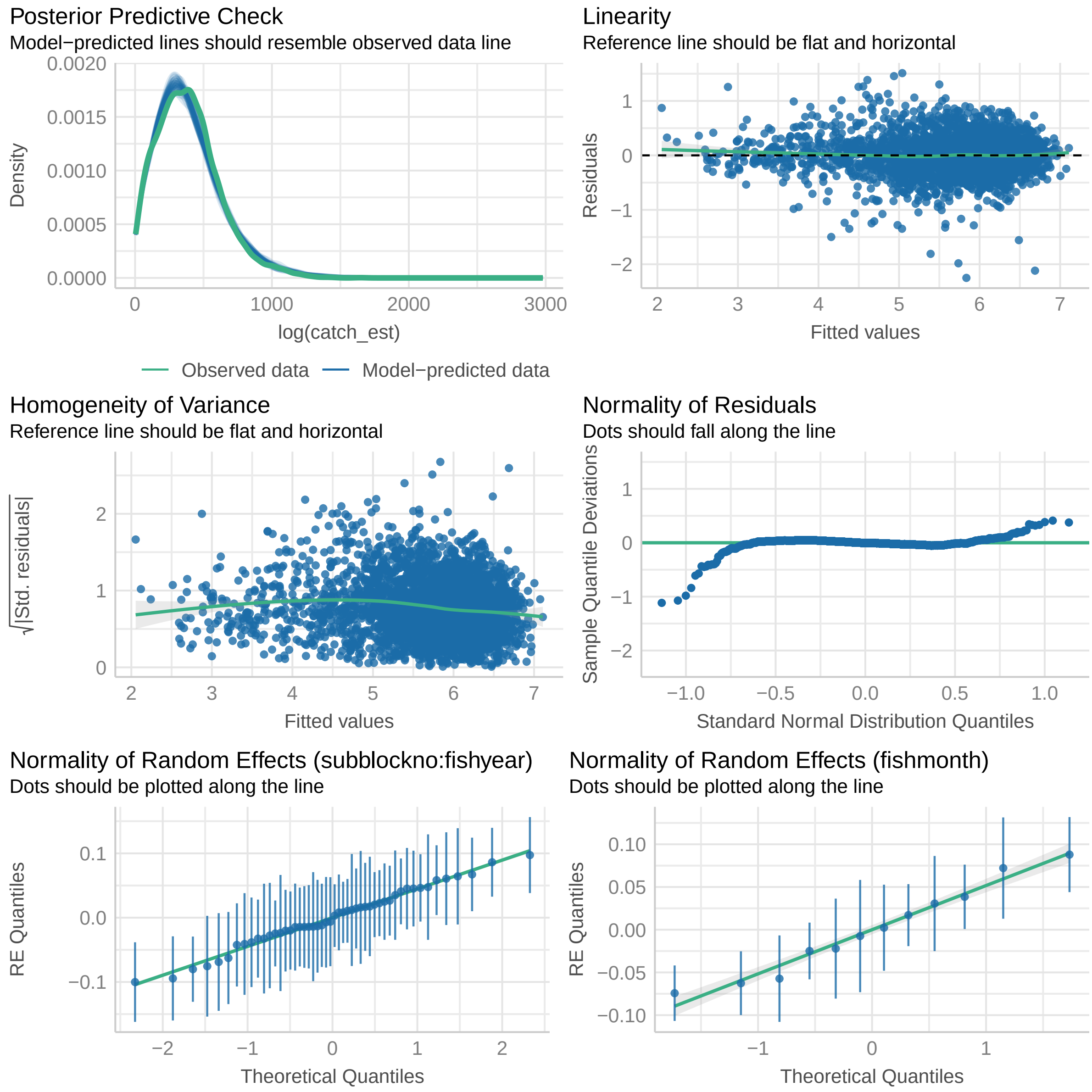


Figure 2.4: SAU11: Model diagnostics)

2.3.6 Case Study: SAU11 Western Zone (Block 11)

Diagnostics for SAU11 are similar to those for SAU5, even though SAU11 has a much greater level of fishing activity (SAU11 4116 records; SAU5 3237 records). Non-Normality of the residuals in SAU11 follows a similar pattern to SAU5. As for SAU5, most subblock catch rates follow a similar trend to the SAU wide trend. This may explain why the ICC and % variance explained for subblocks is consistently low - i.e the subblocks are essentially correlated. For SAU11 however, two subblocks however show different trends over the time series; subblock 11D had the lowest catch per unit effort at the start of the time-series, but the second highest by the end of the time-series, while subblock 11A had the highest catch per unit effort in 2012, and the lowest by 2019. While these two subblocks have slightly different trajectories, they remain largely in parallel.

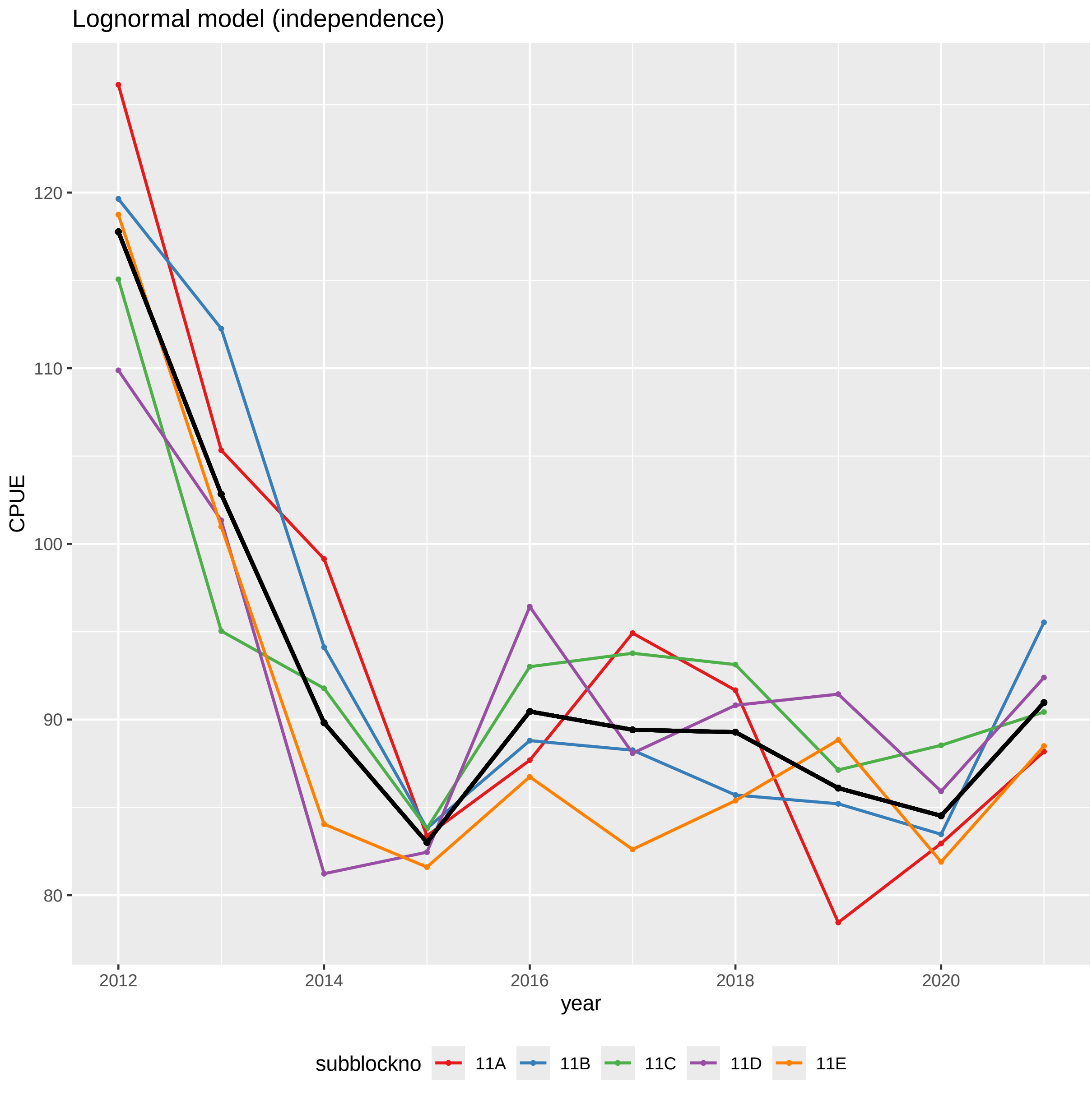


Figure 2.5: SAU11: glmmTMB - Log-normal CPUE standardisation (including bias correction). Response variable is log(catch), with log(effort) included as an offset. Black line indicates SAU mean catch-rate, and the coloured lines relate to the trend in each specific subblock, and the coloured lines relate to the trend in each specific subblock.

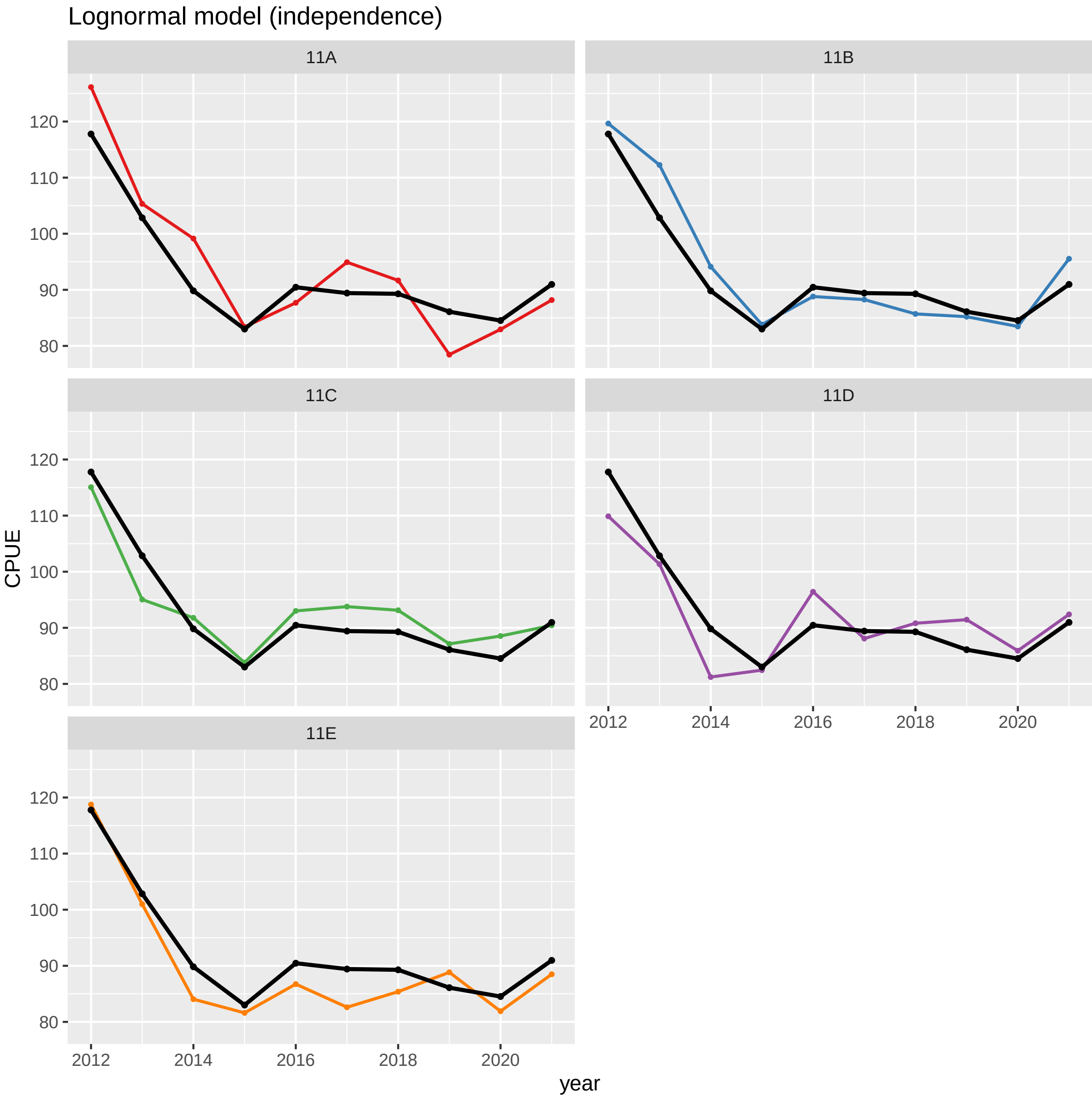


Figure 2.6: SAU11: glmmTMB - Log-normal CPUE standardisation(including bias correction) - facet plot. Black line indicates SAU mean catch-rate.

2.4 Discussion

A log-normal model was fitted using two R packages - glmmTMB (Brooks et al. 2017) and lme4 (Bates et al. 2015). The lme4 package has been chosen as for statistical standardisation of Tasmanian abalone catch-rates, largely due to the ease of prediction methods for back-casting to obtain a standardised catch rate. When attempting to fit a log-normal model using lme4 (Bates et al. 2015) or a gamma model using either lme4 or glmmTMB packages, convergence warnings were common, primarily for singularity, and always driven by the subblock factor having negligible or zero variance. An automated switch was applied such that for SAUs with less than four subblock levels, the subblockno and subblockno:fishyear terms were removed from the model. Implementation of the filter resolved the majority of convergence warnings when useing the lme4 package.

The base statistical model outlined here for statistical standardisation of Tasmanian abalone catch-rates provides a number of benefits in terms of flexibility, while accounting for key factors affecting fishing performance. The minimal contribution however, of the sub-block factor to model variance in most SAUs was unexpected (this is explored further in Chapter 3). When using the fishery docket data some of level of inaccuracy in reporting at the sub-block scale is expected, but that does not apply in the GPS logger program examples provided here. GPS logger program data maps sub-block onto the dataset using a spatial join between each dive polygon and a zone polygon with zone, region, SAU, block, and sub-block attributes.

An emerging characteristic of the Tasmanian GPS logger program data, is that different parts of the fishery have different drivers in terms of spatial structure. For this reason, a relatively simple base model, optimised for each SAU with additional variables where available is a better approach than attempting to apply a more complex model throughout. The model residual term accounted for the majority of data variance for all SAUs. Whether this simply that other factors are more important than those included in the base model, or there is a high level of noise in fisheries catch rate data is examined in the next chapter (Chapter 3).

3 The Importance of Scale

3.1 Introduction

Read (Barnett, Ward, and Anderson 2021)

Adapt and run this code https://pbs-assess.github.io/sdmTMB/articles/web_only/index-standardization.html

Catch and effort data in most fisheries are perceived as being noisy and variable. This perhaps unsurprising in that they often span large geographic distances, often with highly heterogeneous environments. Fishing takes place across a range of sea conditions, and effort is often rough guide rather than a precisely measured unit of time. We observed in Chapter 1 considerable levels of variation in the relationship between catch and effort. In Chapter 2 we observed moderate Conditional R^2 (~ 0.4 - 0.5) for many of the Spatial Assessment Units (SAUs) examined, indicating the terms included in the statistical model for catch rate standardisation are failing to capture a substantial proportion of the observed variance in the datasets. While we have included a within SAU spatial structure (subblock), this term contributed very little to the model outcomes (Table 2.5). This would suggest either a level of homogeneity among subblocks within an SAU, or, that high levels of variation occurs at a scale much smaller than the scale of a subblock. Alternatively, there are other factors not included in the model that contribute to high variability in daily diver catch rates.

A perceived advantage of the fine-scale fisheries collected through the GPS logger program is the potential to account for the high levels of variability in abalone fisheries catch rate data. Utilising the precise location of each fishing event, we an either dissect each SAU into smaller rule-based spatial units (rather than the arbitrarily imposed subblock boundary divisions), or go further and account for the exact spatial location in our linear model.

Abalone fishery catch rates are always perceived as highly variable and unreliable, yet illustrations of the extent of variability in catch rates within a Spatial Assessment Unit (SAU) are rarely provided. In order to better understand the challenges with obtaining a mean annual catch rate for an SAU, catch rates are calculated at a range of time steps, from daily (mean of all fishing events in an SAU per day) to mean catch rate within a moving window of 5, 10, 30 and 90 days.

This chapter links to two objectives:

Objective 1: **Characterise the statistical properties, consistency, interpretability and assumptions of spatial and classic indicators of fishery performance.**

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.** ## Methods

3.2 Methods

Input data are those spatial records that meet criteria outlined in Chapter 1. Noting that data capture is between 80% and 90% depending on the year, this opens a potential bias

3.2.1 Statistical Models

All lmer analyses here use the KUD dataset either pooled by diver/day, or by individual dive. The response is always log(catch) and modelled against log(effort) where effort is logger time in minutes.

3.2.1.1 Modelling response variable at different scales

While the GPS logger program records the duration (in minutes) for each fishing event, catch is reported as a daily total catch for each subblock visited. Catch per dive is achieved by pro-rata assignment of total catch to each dive based on duration of each dive. While we recognise this may have a smoothing effect on actual catch per dive, where catch rate for dives through the day might vary due to depth or other environmental drivers. By modelling catch at the dive level, rather than by day, we may be able to account for some of the unexplained variation in the model identified in Chapter 2. This is achieved by modelling catch achieved for one hour of effort at the day, and dive scales using the base mixed effects model outlined in Chapter 2.

3.2.1.2 Including within SAU spatial structure in standardisation models

Dividing an SAU into multiple spatial clusters as an alternative to arbitrary subblock divisions may capture more of the unexplained variation (residual variance) evident in the examination of the contribution of effects to total variance in Chapter 2. This is achieved by applying spatial scan methods to the dive centroid, to create an empirical rule-based SAU subarea. Spatial density-based clustering methods require two parameter settings; a distance value (epsilon) to be used as the radius of the neighbourhood search, and the minimum number of points required for a unique cluster. Here we use the dbscan package (Hahsler and Piekenbrock 2022) to assign a unique spatial cluster identifier to every dive record, with epsilon set at 300m and the minimum points per cluster set at 15. The new cluster variable is substituted for the subblock variable in the base standardisation model (Figure 3.1).

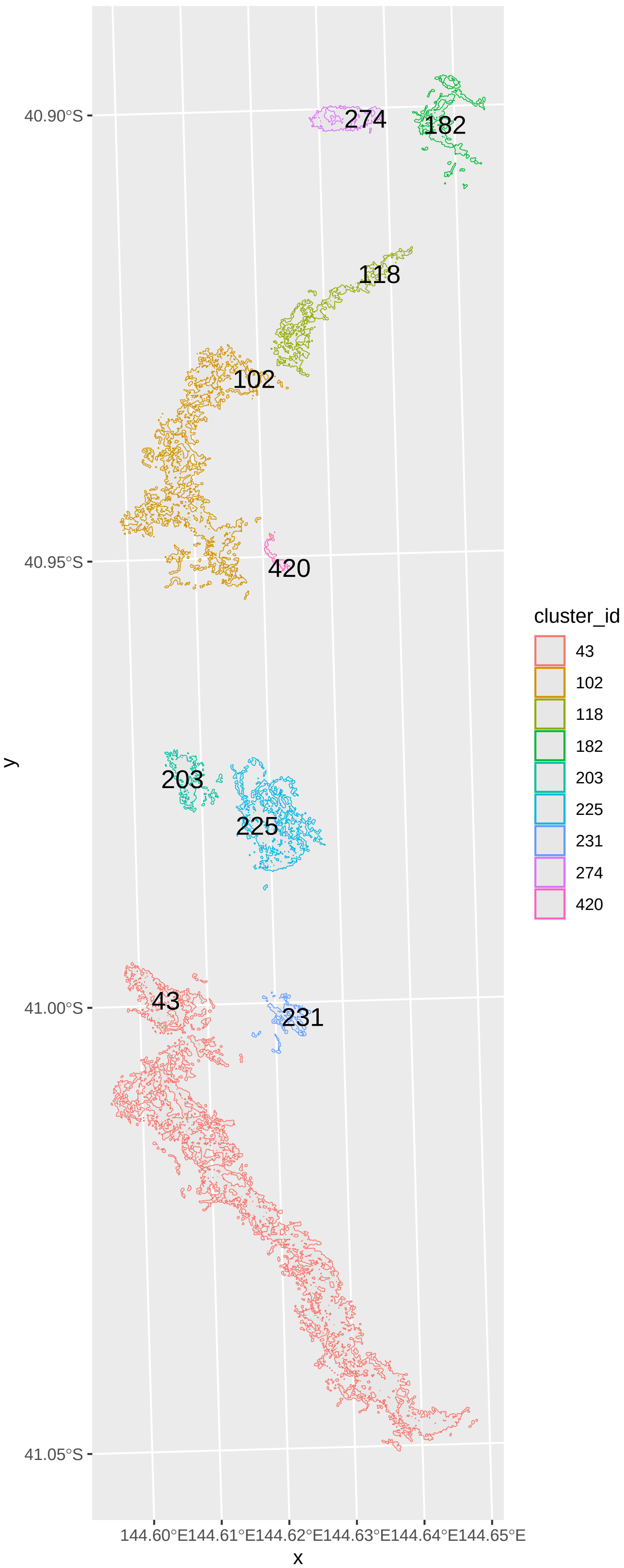


Figure 3.1: Spatial Clustering of all dive events within SAU5, Subblock 5C, Northern Zone (2012 - 2021) into discrete groups by distance. DBSCAN generated clusters using a setting of 300m separation between the centroid of fishing events, and a minimum of 15 points to form a cluster. This approach found nine distinct spatial clusters of fishing activity within the singel Subblock 5C.

3.2.1.3 Using high resolution information on location of fishing

Much of the variation in catch rates is thought to be associated with high levels of patchiness in resource abundance. By modelling fishing data at the scale of individual dive, the centroid of the dive polygon representing the location of the fishing event can be used in a spatial regression approach. While several packages offer fixed effect spatial auto-regression capability, very few enable direct use of the spatial information in a mixed effect model.

Two mixed effect spatial regression approaches are utilised here:

a) Coordinates (Easting and Northing) of the dive polygon centroid are included as a spline in a general additive mixed effect model (GAM) using the mgcv package (Wood 2011).
b) A spatio-temporal LMM was used to examine the influence of spatial structure on the model fit to the case study dataset using the sdmTMB package (Anderson et al. 2022).

The approach using GAMs involved fitting the base CPUE standardisation model as a GAM, and then repeating the model with a term for a spline of easting and northing of the dive centroid. Model fits were examined using deviance explained and AIC. Scaled prediction plots using predicted annual mean catch per 1 hour of effort were used to examine the impact of inclusion of the spatial term.

Spatially explicit GLMMs seek to address spatial and or temporal autocorrelation in a dataset, and provide insights into the space and time influences on response variables. The first step is to create an SPDE mesh using the centroid of the dive polygon, with a cut-off value of 5, where cut-off refers to the minimum distance between knots (vertices) of the resulting mesh (Anderson et al. 2022). This step builds a mesh across the spatial extent of the input dataset. For the GLMM component, the base standardisation model was used, except the (1| subblock:fishyear) term was dropped as sdmTMB does not yet support interactions of this type. Three models were applied, spatial switched off (non-spatial) which should be equivalent to a GLMM from non-spatial modelling packages, spatial switched on (space) and lastly with switching both spatial and spatio-temporal modes on (space-time). The temporal component was intentionally coarse for this exercise, with the time resolution set to fishing year. To better understand the spatial scale of variation in the datasets, this exercise was run firstly with the coarse scale subarea term subblock, and then again with subblock replaced with the finer scale subarea term cluster_id. Model performance was compared using AIC. If fishing performance is relatively homogeneous within clusters, we would expect relatively little difference between the non-spatial and spatial model fits where the finer scale cluster term is used.

3.3 Results

3.3.1 Variability in daily catch rate

Across all example SAUs, daily catch rates differ substantially over short time frames such as 10s of days (Figure 3.2). Daily catch rates appear to be lower and less variable during winter, but otherwise there is little evidence of a strong seasonal pattern. For example in SAU11, daily catch rate may vary by up to 150 Kg/Hr within as little as 30 days (Figure 3.2 SAU11 panel). Moving window mean catch rates of 5 and 10 days also show short-term oscillations of up to 80 Kg/Hr. Over longer time windows, the 30 and 90 day moving window mean catch rates continue to show substantial oscillations of up to 50 kg/Hr. There are many potential explanations for this level of variability such as weather on the day of fishing, diver identity and knowledge, the subblock or location within the SAU, or high variability in the local abundance across fishing events. The contribution of environmental variables was examined in detail in appendices 10-12, with little evidence that weather imparts a strong influence on fishing performance in the majority of cases. The contribution of heterogeneity abalone abundance at fin spatial scales is the focus of the remainder of this chapter.

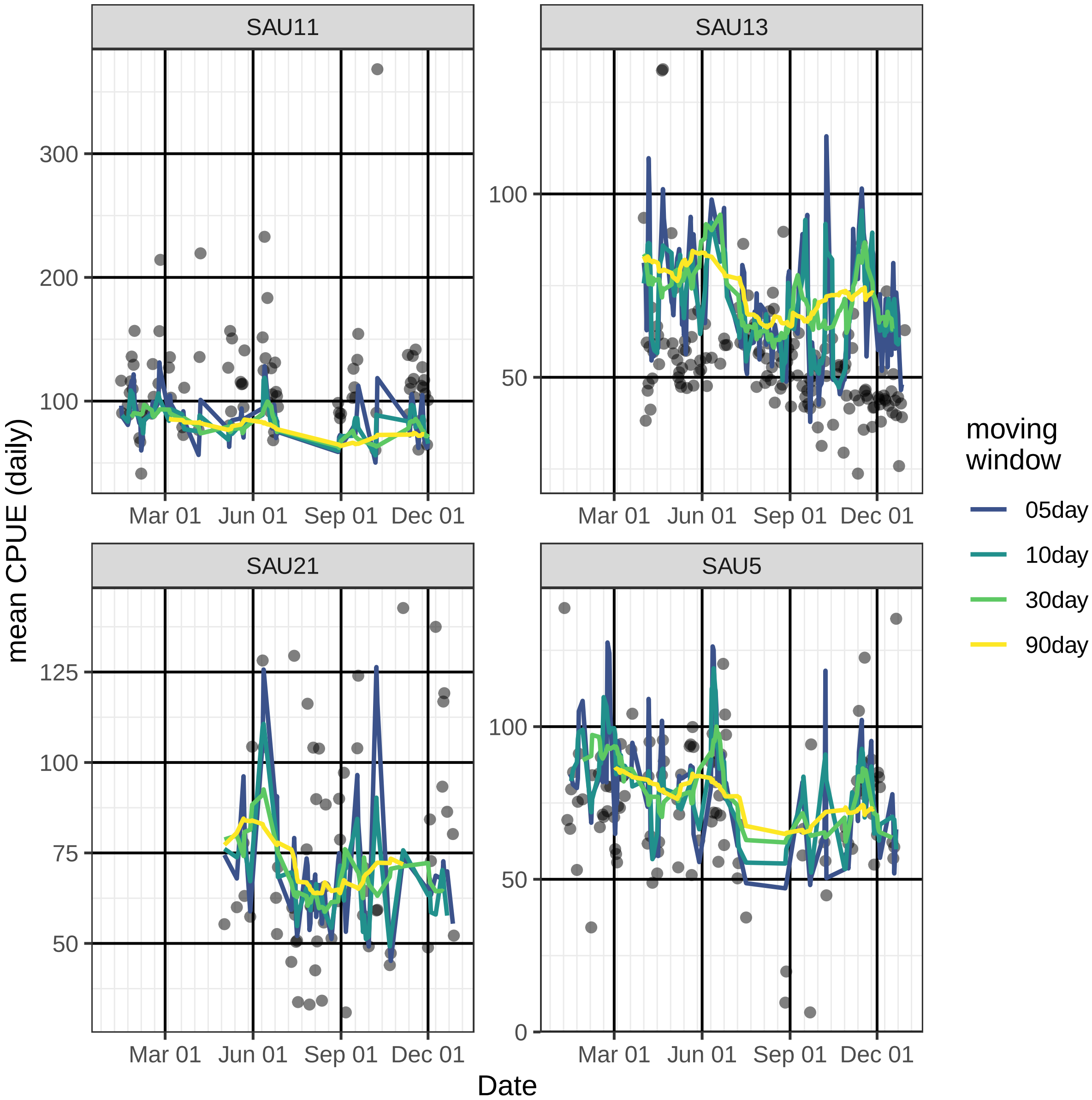


Figure 3.2: Mean daily catch rate (Kg/hr) by dive in 2013 for three case-study SAUs. Grey scale dots represent mean daily catch rate. Coloured lines represent a moving window mean catch rate calculated for window size: 5, 10, 30, and 90 days.

3.3.2 Modelling response variable at different scales: day vs dive

Modelling the response variable (estimated catch) at the spatial resolution of individual dives rather than aggregated up to each fishing day, provides a number of opportunities for quantifying scales of spatial structure in our datasets. There was almost no difference in predicted mean catch per hour of effort when data were modelled by day or disaggregated to the scale of each dive event (Figure 3.3). Mean expected catch only deviated in the final year of the time-series (2021), and is not likely due to the low sample size for 2021 in the study dataset.

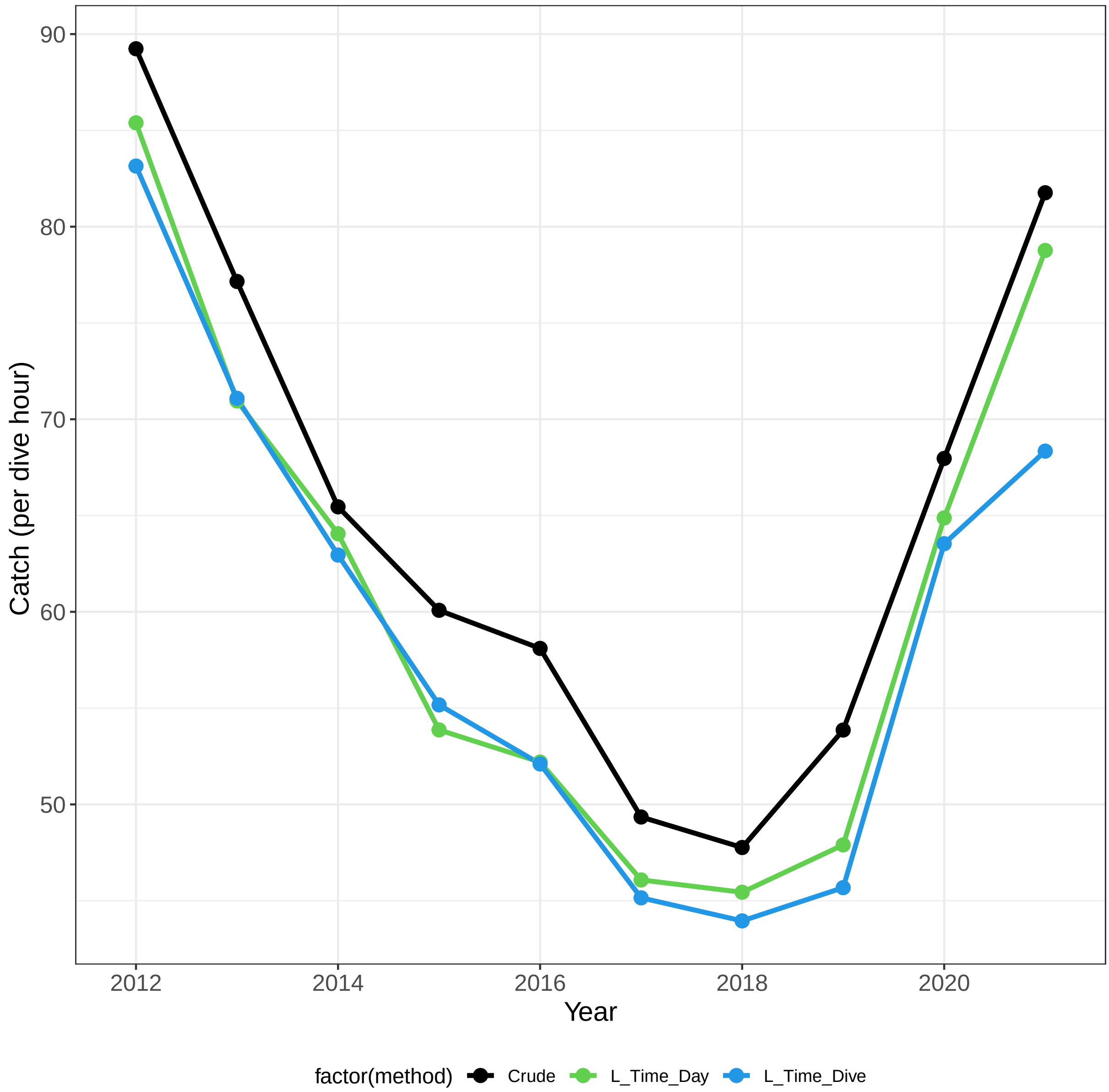


Figure 3.3: Modelling the response variable (catch) at day vs dive scale for Northern Zone, SAU 5(subblocks 5A, 5B, 5C, 6A, 6C, 6D). Catch achieved is predicted for 1 hour of fishing effort.

3.3.3 Accounting for within SAU spatial structure

Replacement of the coarse-scale (subblock) random effect predictor with a finer-scale (cluster_id) predictor to account for within SAU spatial structure returned mixed results. For some SAUs (SAU11, SAU21), replacing the coarse-scale subblock categorical predictor with the fine-scale cluster predictor, marginally increased the Conditional R^2 (Table 3.1), indicating a small improvement in capture of variation within the dataset. However, the inverse was observed for SAU49 where inclusion of cluster_id decreased Conditional R^2. This is supported by a corresponding increase in residual variance for SAU49 modelled with cluster_id (Table 3.2). While in most cases, the switch from coarse to fine for the sub-SAU term had little effect on the overall variance accounted for by the model, there was an effect on the proportion of variance accounted for by the model random effects. With the exception of SAU49 and SAU7, shifting to the finer cluster term decreased the residual variance, although there was no real change in SAU13, which had only three distinct spatial clusters of fishing events. The trend in estimated mean catch per 1 hour fishing event using subblock or cluster terms was also very similar (Figure 3.4). The absence of a strong effect of the within SAU spatial structure term (subblock vs cluster) on either Conditional R^2 or on estimated mean catch suggests, that in most SAUs they are not an effective mechanism to account for spatial structure in the dataset.

Table 3.1: Accounting for within SAU spatial structure with fine or coarse scale division. Marginal (R2_M) and Conditional(R2_C) R^2 for five Spatial Assessment Units (SAUs) in the Tasmanian abalone fishery.

structure	SAU	R2_marginal	R2_conditional
Sub-block			
	SAU11	0.07	0.41
	SAU13	0.10	0.52
	SAU21	0.17	0.58
	SAU49BS	0.07	0.75
	SAU5	0.25	0.53
	SAU53	0.12	0.78
	SAU7	0.07	0.57
Cluster			
	SAU11	0.07	0.50
	SAU13	0.12	0.56
	SAU21	0.19	0.62
	SAU49BS	0.12	0.68
	SAU5	0.22	0.55
	SAU53	0.14	0.76
	SAU7	0.06	0.57

Table 3.2: Accounting for within SAU spatial structure with fine or coarse scale division. Contribution (% explained) of individual random effects to the model.

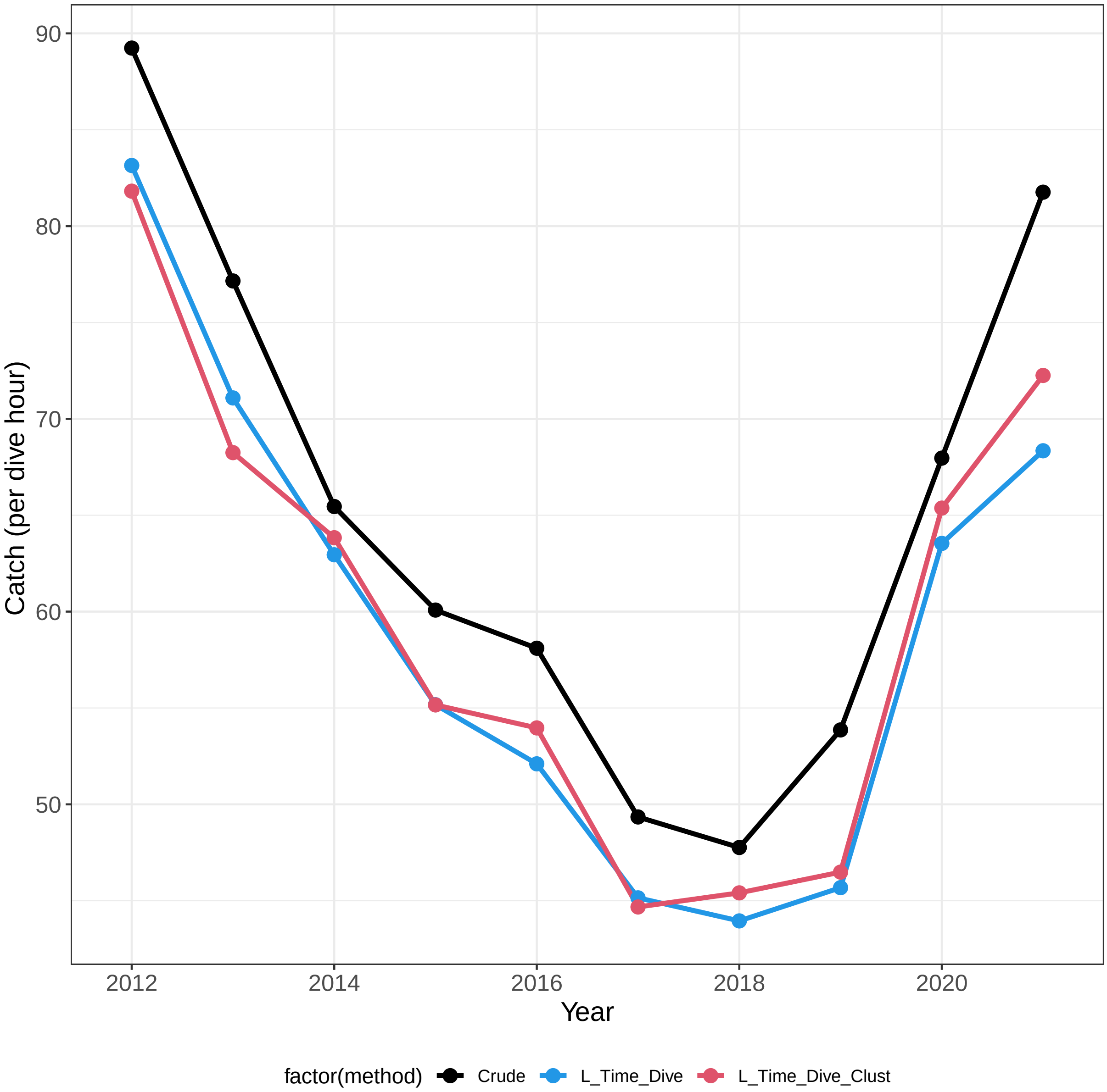


Figure 3.4: Modelling the response variable (catch) at the dive scale, with either subblock or cluster terms to account for within SAU spatial structure. Figure shows results for Northern Zone, SAU5 (subblocks 5A, 5B, 5C, 6A, 6C, 6D). Catch achieved is predicted for 1 hour of fishing effort.

3.3.4 Including precise location of dive events in the standardisation model

3.3.4.1 Using GAM with a spline to account for spatial structure

Inclusion of a spline term in a GAM model to account for precise location of each fishing event had minimal impact on the proportion of deviance explained by the model (Table 3.3), although there was a noticeable reduction in AIC with the inclusion of the spatial spline (Table 3.4). There appeared to be little corresponding impact on the temporal trend in the estimated annual mean catch rate (Figure 3.5), although the GAM model suggests the fishery reached a low point in 2016, whereas the non-spatial glmmTMB model suggests the low point in the time-series occurred in 2018 (Figure 3.4).

Exploration of the residuals by year and spatial location identify high levels of local variation in residuals, with substantial change among years (Figure 3.6).

Table 3.3: Change in model deviance explained by accounting for spatial location of each dive event by adding a spline on easting and northing of the dive centroid. Data used in this analysis were from SAU5 in the Tasmanian abalone fishery.

Scale	devExplained
non-spatial	0.84
spatial	0.87

Table 3.4: Change in model AIC with increasing spatial resolution.

model	df	AIC
dive	117.8	6,822.9
diveXY	145.9	4,369.0

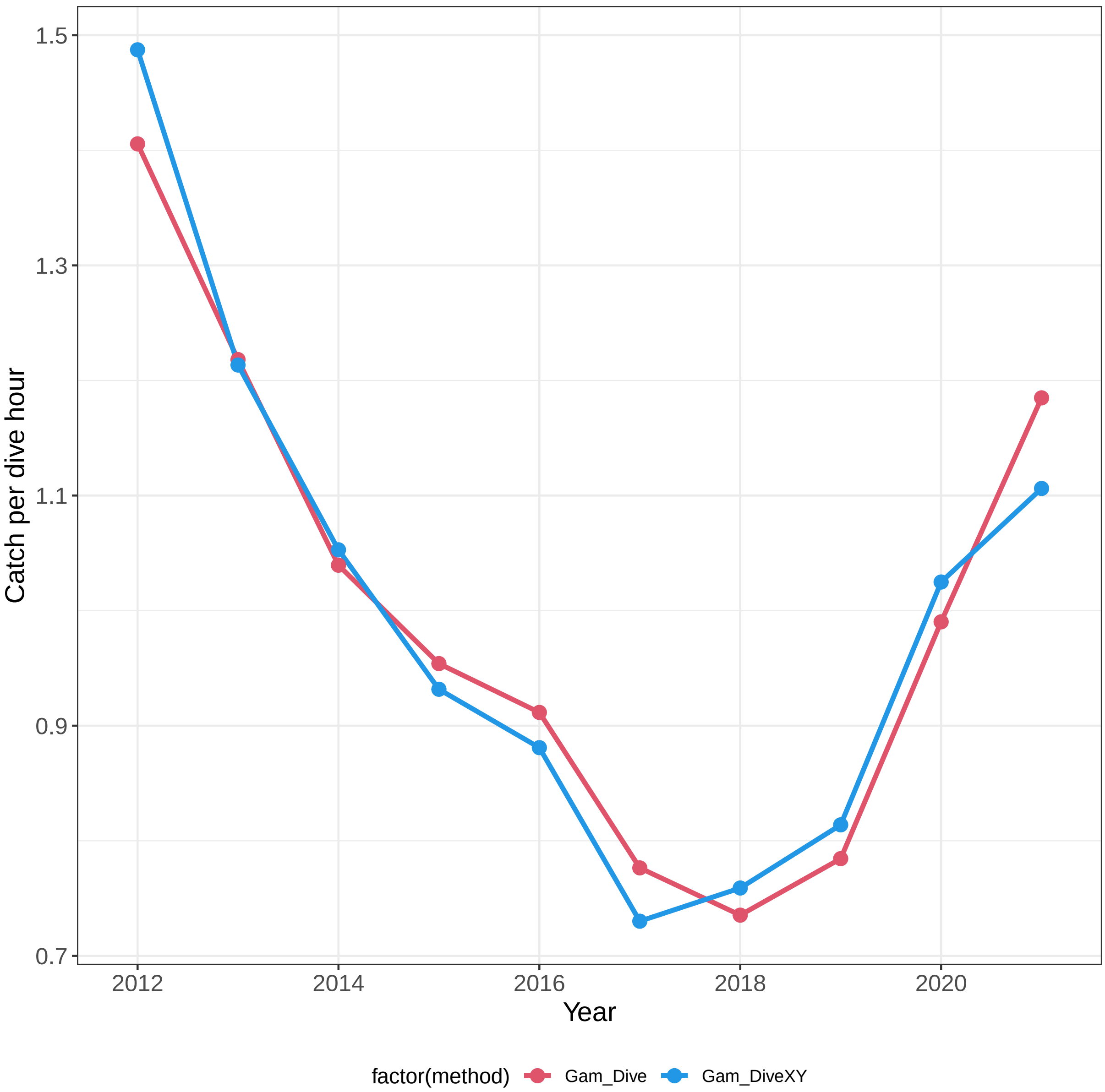


Figure 3.5: GAM Modelling day vs dive scale catch and effort for SAU5, Northern Zone. Standardised CPUA (catch achieved for 0.5Ha of fishing area).

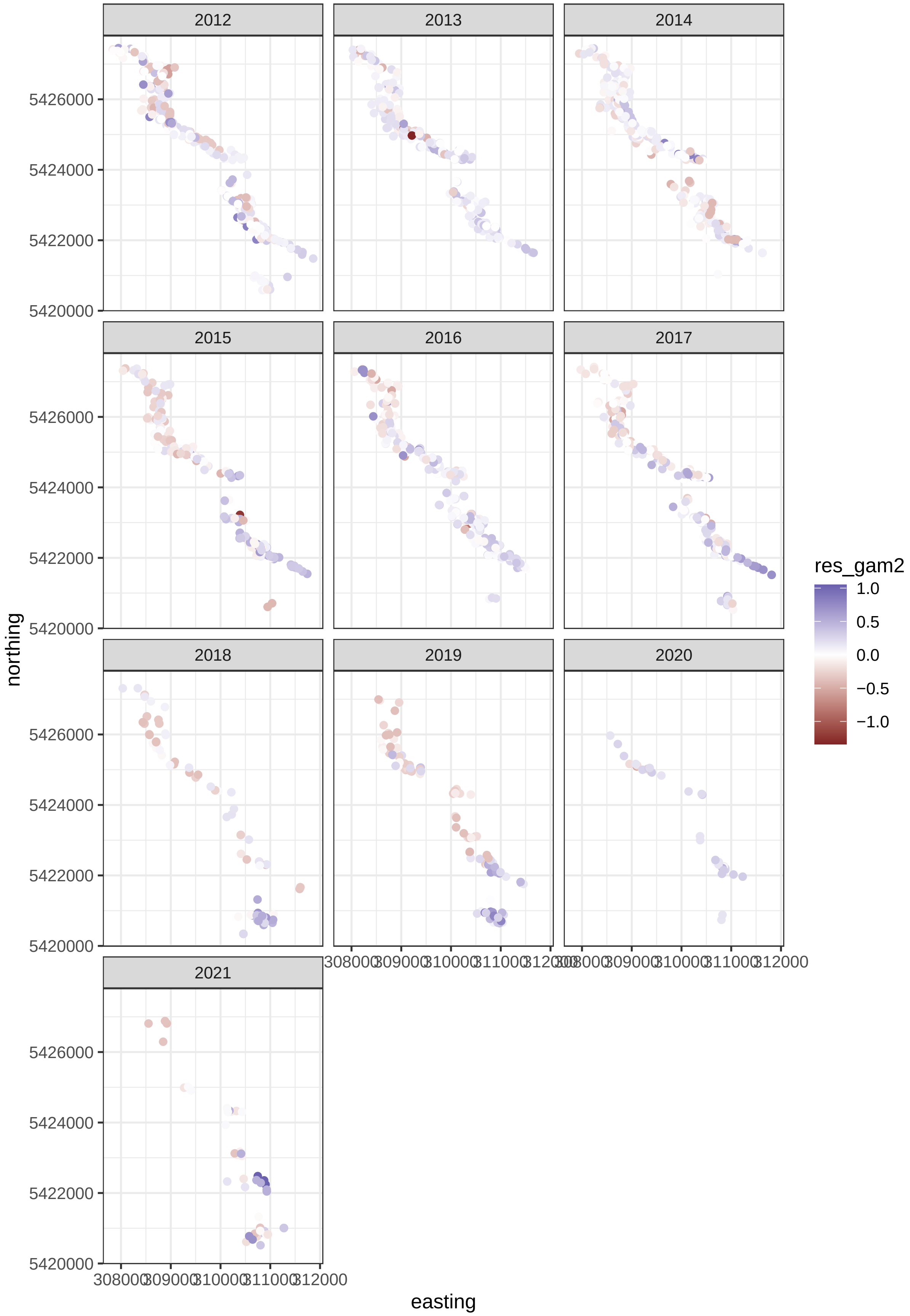


Figure 3.6: Residual plots by year for SAU5, Northern Zone from model fitted with subblock term. Only data from subblock 6C are presented here for clarity. Spatial location included as a spline smooth in the GAM model using easting and northing of the dive centroid.

3.3.4.2 Using spatially explicit GLMMs to account for spatial structure

The basis of the spatially explicit linear mixed effects model is a SPDE (stochastic partial differential equation) mesh Figure 3.7. The resolution of the mesh can be configured from fine to coarse, with obvious trade-offs between computation speed and spatial resolution.

Direct comparison between non-spatial and spatio-temporal model fits is not possible, although a comparison of the improved performance from using a different within-SAU term (subblock vs cluster) can provide insights into the improvements gained by addressing spatial structure directly in the model. For the non-spatial model, a marked reduction in model AIC was observed for the model using cluster_id as the within SAU term compared to using the coarser subblock term (Table 3.5). When spatial or spatio-temporal terms are included the improvement in AIC is negligible (Table 3.5).

Examination of the spatial structure in residuals through time suggests that the majority of variation in catch rates occurs at a very fine scale, although this does vary through time and some broad spatial structure is evident regardless of whether subblock or cluster is added as the within SAU term (Figure 3.8, Figure 3.9). The spatial map of residuals matches very closely to that returned from the GAM model (Figure 3.6) In 2012, there is considerable local variation in residuals across the fishing grounds in subblock 6C. In 2018 ad 2019, when fishery performance was lowest (Figure 3.4), the residuals are largely negative across all of Subblock 6C (Figure 3.8, Figure 3.9). While there is a small level of deviation between the non-spatial model residuals and the spatio-temporal model residuals, there are no outstanding differences, or differences that are of concern (Figure 3.10, Figure 3.11).

The difference between non-spatial, spatial, and spatio-temporal standardised annual mean catch rate was negligible, for both the subblock and cluster term models (Figure 3.12, Figure 3.13). This is consistent with a lack of major shift in the residuals (Figure 3.10, Figure 3.11), and also consistent with the spatial and non-spatial GAM models.

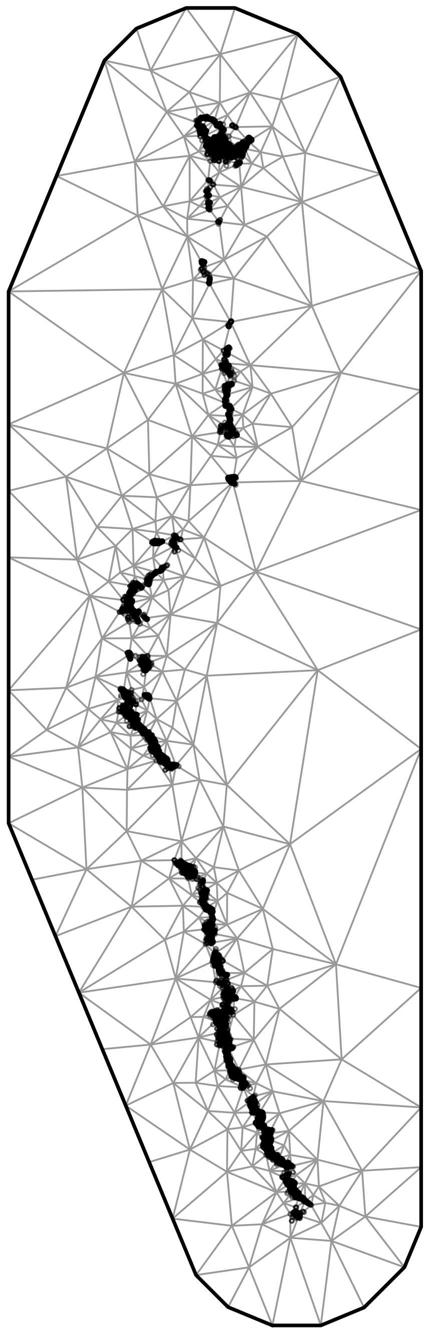


Figure 3.7: Spatial mesh for Block6 used as the basis for the spatially explicit lmm.

```
mesh5: 0.06 sec elapsed
```

Table 3.5: Spatially explicit GLM modelling of abalone abundance. GLM formula was identical for all three models, with **subblock** and **cluster** used as alternative within SAU random effect terms.

structure	withinSAU	df	AIC
non-spatial	subblock	14	5,893.1
	cluster	14	5,606.2
spatial	subblock	16	4,159.6
	cluster	16	4,160.3
space-time	subblock	17	2,888.2
	cluster	17	2,888.1

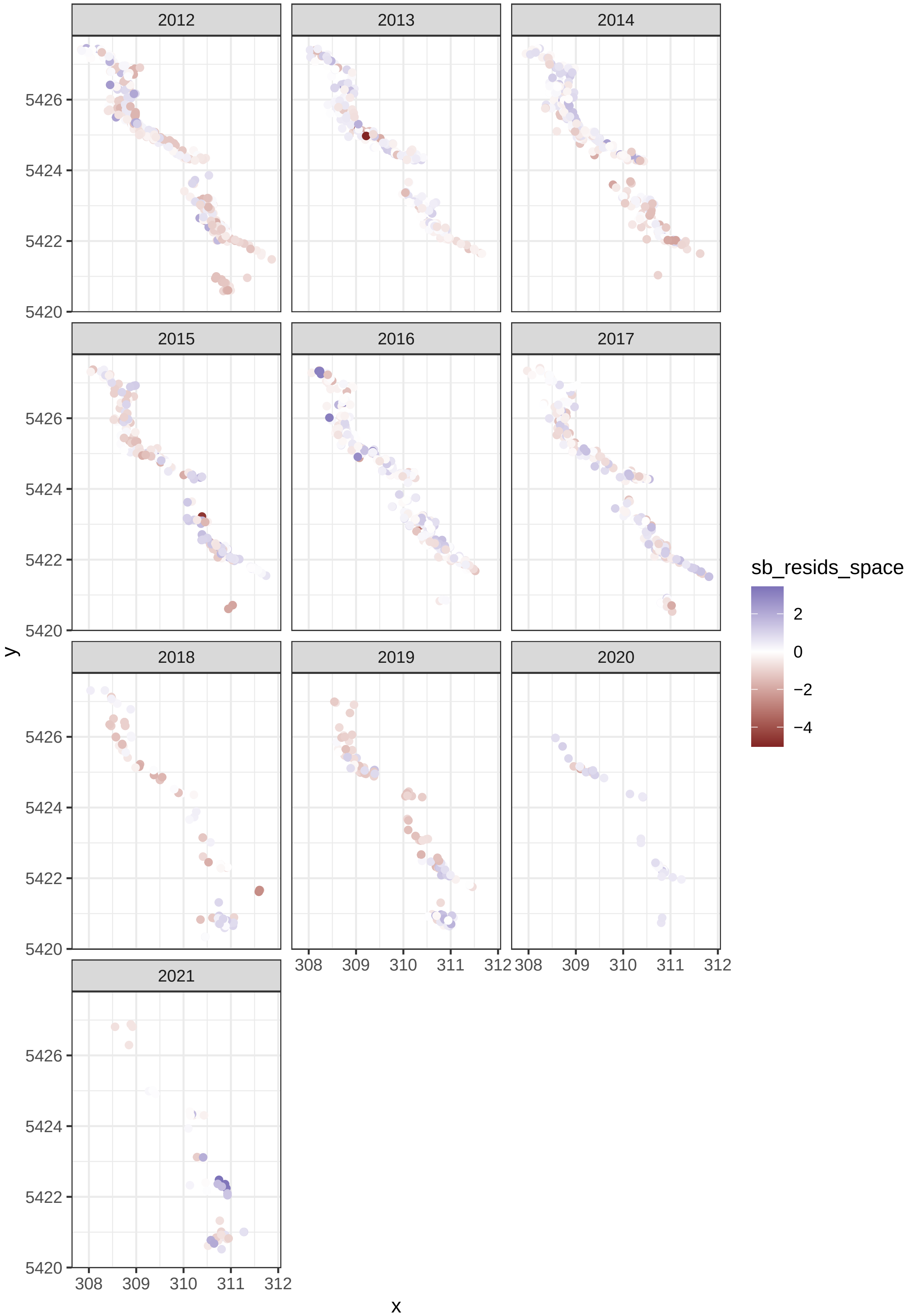


Figure 3.8: Residual plots by year for SAU5, Northern Zone from model fitted with subblock term.. Only data from subblock 6C are presented here for clarity. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid.

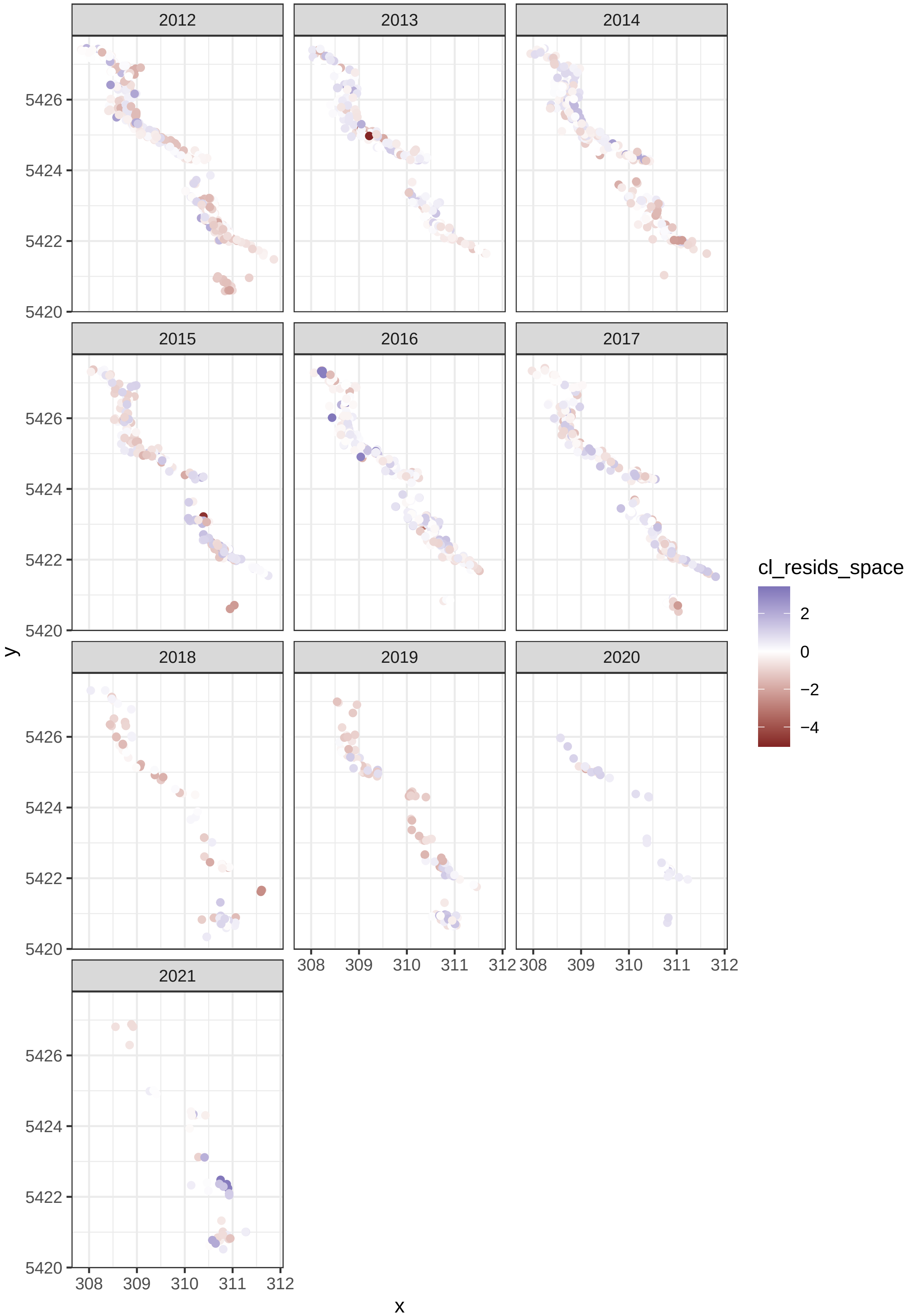


Figure 3.9: Residual plots by year for SAU5, Northern Zone from model fitted with cluster term. Only data from subblock 6C are presented here for clarity. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid.

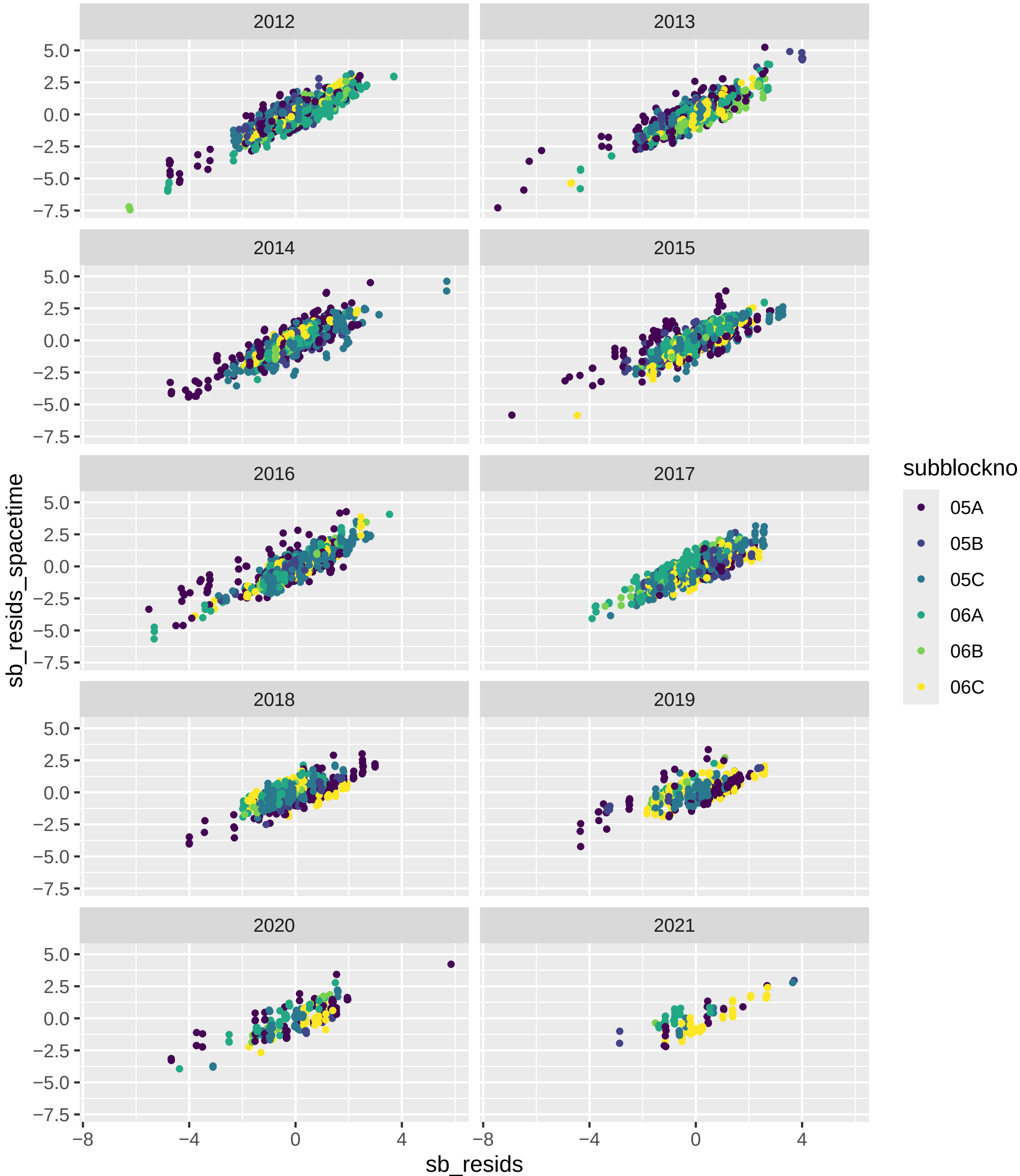


Figure 3.10: Plot of residuals for non-spatial vs spatiotemporal model for SAU5, Northern Zone from model fitted with subblock term. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid. Date of fishing is included as a Spatiotemporal term.

Figure 3.11: Plot of residuals for non-spatial vs spatiotemporal model for SAU5, Northern Zone from model fitted with subblock term. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid. Date of fishing is included as a Spatiotemporal term.

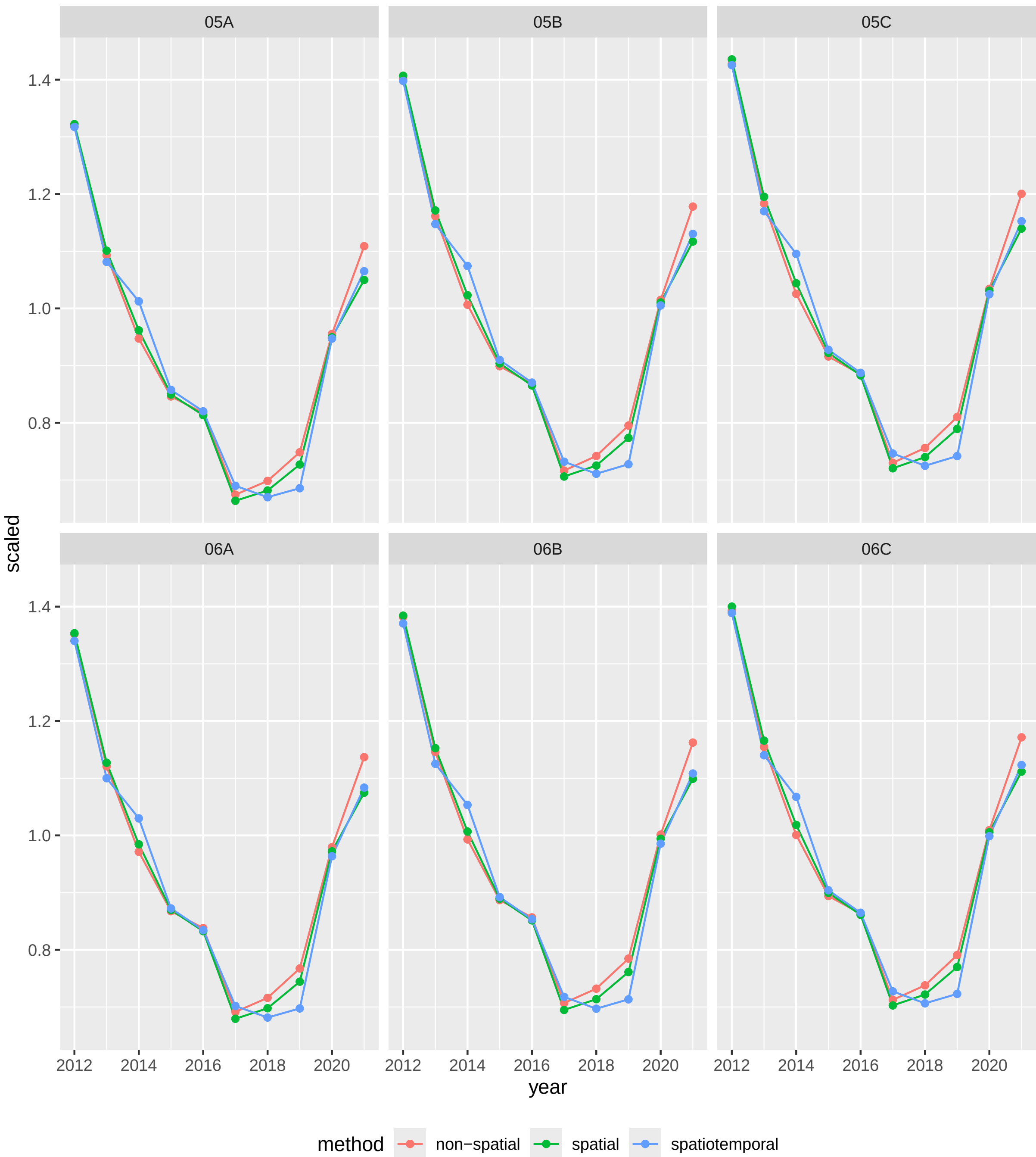


Figure 3.12: Space-time modelling of the response variable (catch) for SAU5, Northern Zone from model fitted with subblock term. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid. Date of fishing is included as a Spatiotemporal term.

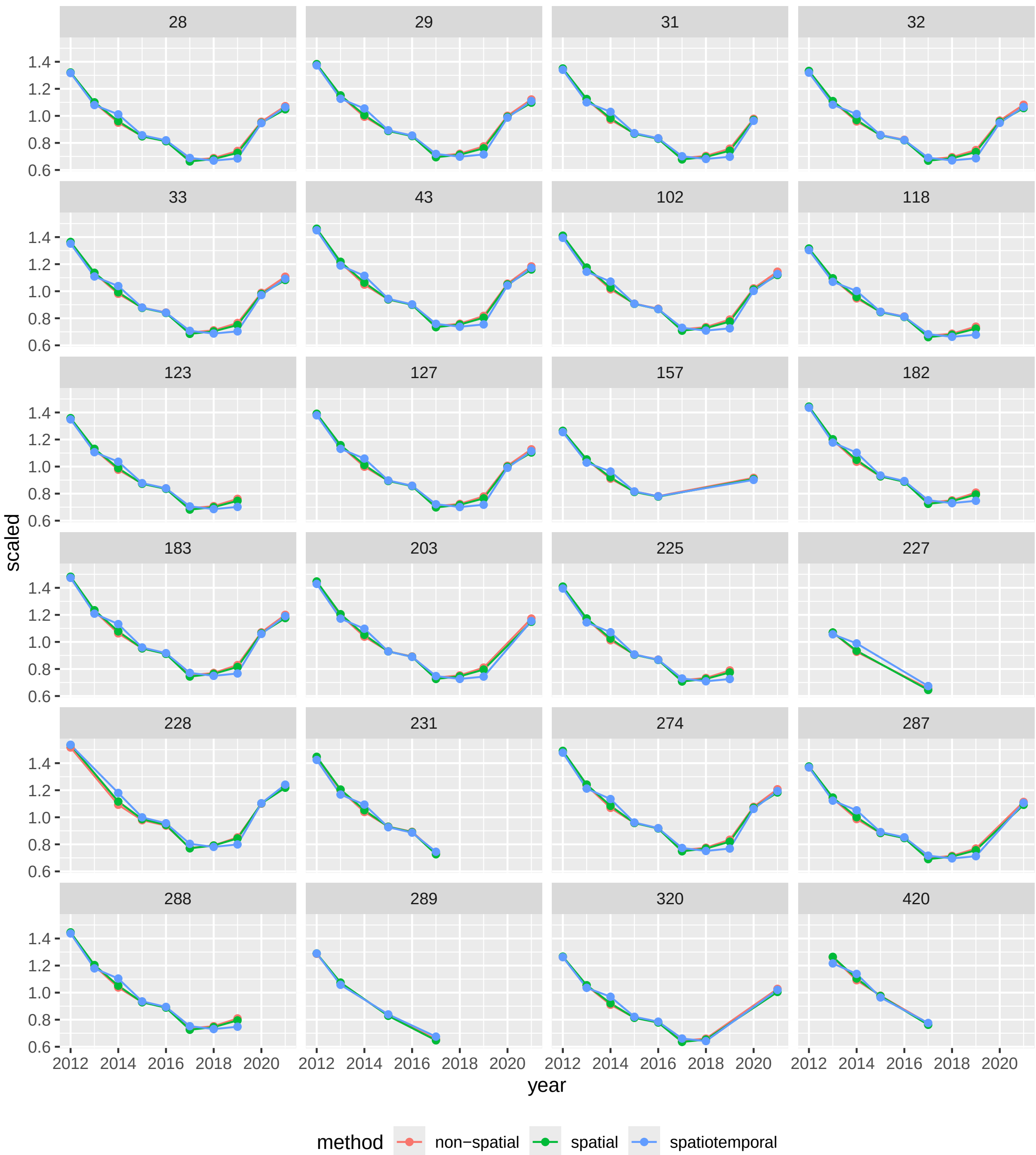


Figure 3.13: Space-time modelling of the response variable (catch) for SAU5, Northern Zone from model fitted with cluster term. Spatial location included as a mesh to the sdmTMB model using easting and northing of the dive centroid. Date of fishing is included as a Spatiotemporal term.

3.4 Discussion

The suite of results presented here confirm that geo-referenced fishery dependent data can assist with accounting for, and quantifying spatial structure in abalone fisheries data. That does not translate however, to major shifts in standardised mean annual catch rates. Mean catch per hour of effort was relatively robust to a range of modelling approaches with different spatial and temporal resolution of predictor terms. Surprisingly, aggregating response and predictor variables to the scale of fishing day did not have an adverse impact on the estimated annual mean catch rate. In terms of impact on catch rate trends, there appears limited direct benefit in applying spatial regression methods to catch rate standardisation given the negligible improvement found in this study. There are however other benefits of application of spatial linear models which are discussed below.

Utilising high resolution data on fishing activity to build maps of discrete reef systems, and adding a reef identifier to the model as a random effect does appear to be advantageous. This appears to be a valuable use of the fine scale fisheries data, regardless of whether a spatial term is added to the mixed effects model. While it is desirable to capture information on fisher Docket returns about geographic structure within an SAU, the formal subblock (mapcode, reefcode) reporting areas often dissect contiguous rocky reef. Whereas empirical rule-based determination of reef systems based on centroids of dive polygons is a more defensible approach to addressing within SAU spatial structure. Typically, the number of discrete reef areas within an SAU is higher than the number of formal subblock divisions, and meets the normally accepted requirement for random effects to have more than five levels. A greater difference is evident among clusters in annual catch rate trends (Figure 3.13) than when a coarse grouping level such as subblock is used (Figure 3.12).

The relatively new spatially explicit GLMM tools are providing capability to incorporate spatial terms and capture spatial structure in CPUE standardisation and/or research surveys (Evans et al. 2021; Tanaka et al. 2022b; Thompson et al. 2023; Thorson et al. 2023; E. J. Ward et al. 2022; Zhou, Campbell, and Hoyle 2019). Application of mixed effect models for CPUE standardisation require a relatively high level of expertise, and there are many different views on their efficacy and their implementation. Adding spatial terms to GLMMs necessarily adds complexity, with many more traps for young players (Hodges and Reich 2010). At this stage, mixed effects models with random spatial terms are being actively developed, with rapid advances made in terms of speed of computation. Implementation of these types of models as well as informative diagnostics and capacity to make comparisons with non-spatial models are still their infancy. It is very clear from the experimental implementation of spatially explicit mixed effect models in this chapter, that the capacity to explore temporal shifts in the spatial structure of residuals can provide great insight into underlying changes within SAUs, that would be impossible to capture without a GPS and depth data logger program.

This chapter has skimmed the surface of the potential use of spatial GLMMS. As the new tools mature, particularity those using TMB (Kristensen et al. 2016) and INLA (Krainski, Lindgren, and Rue 2023; Rue, Martino, and Chopin 2009) for the basis of CPU intensive spatial computations, the scope for application will broaden.

4 Quantifying hyperstability in abalone catch and effort data

4.1 Introduction

Hyperstable catch rates are widely assumed to feature in most wild fisheries around the world (Harley, Myers, and Dunn 2001), and is often cited as a justification for interpreting catch rates with caution, or relying heavily on fishery independent survey (FIS) data sets. FIS data are largely collected systematic along an experimental design and eliminate some of the biases in fishery-dependent data. FIS data collection can also be subject to hyperstability associated with density-dependence of the target species (Kotwicki, Ianelli, and Punt 2014), but avoid the potential effect of diver and fleet behaviour that might create a bias in a fishery-dependent catch rate signal. Behaviour of the target species, for example aggregating behaviour (Rose and Kulka 1999) is another key driver. Changes in the way fishers operate can be difficult to quantify, although Hamilton et al. (2016) were able to combine FIS with fishery dependent data to quantify hyperstability in a tropical reef fishery. Classical serial depletion is another driver of hyperstable catch rates, for example switching to a new species as abundance of one species declines has been a feature in early fisheries, including the Californian abalone fishery (Karpov et al. 2000). Travelling further as fishing grounds close to port become depleted is another form of spatial serial depletion resulting in hyperstable catch rates (Cardinale, Nugroho, and Jonson 2011), though is relatively uncommon in mature fisheries.

While hyperstable catch rates are often acknowledged, and occasionally described (see Figure 1 in Harley, Myers, and Dunn 2001), hyperstable catch rates are rarely quantified in a commercial wild harvest fishery. Documentation of hyperstable catch rates however, is becoming more common in recreational fisheries (e.g. H. G. M. Ward, Askey, and Post 2013). When hyperstable catch rates are described typically it is in the form of a power curve that describes a non-linear relationship between the index of abundance (catch rate) and biomass (e.g. Figure 4.1). When fishery-dependent indices of abundance such as CPUE is used to estimate biomass, the effect is that lower catch rates over-estimate remaining biomass. The challenge of confirming the existence of, and quantifying hyperstability in fishery catch rates is problematic whether using catch rate as an index of relative abundance directly in Empirical Harvest Strategies, or, as an index of relative abundance when conditioning integrated assessment models. The solution for Empirical Harvest Strategies is to be more precautionary, while assessment models can include a hyperstability coefficient and, examine the impact of that coefficient on model fits.

In well established abalone fisheries, rather than the typical form of hyperstability such as species switching, or serial depletion of fishing grounds, it is the capacity of divers to change their fishing behaviour to maintain a catch rate that is of greatest concern. Tasmanian abalone divers have long expressed frustration with reliance on time as a measure of effort, and frequently advised that they changed their fishing behaviour (usually swim speed) to maintain catch rates. Aggregation of adult abalone will be a factor in some abalone fisheries, for example fishers in the Tasmanian Western Zone blacklip fishery target aggregations, where has fishers in the Tasmanian Eastern Zone blacklip fishery are skimming and picking up individual animals as they traverse the reef. Whether this is a habitat driven pattern, or a function of the relative health of these two blacklip fisheries remains unclear.

Traditionally, we measure effort as time and model CPUE (ratio of catch and time) as the response variable against a number of likely predictors. Divers frequently comment that while catch per unit effort (CPUE) remains stable, they are covering more reef area in the same dive time to get their target catch. This common description of change in fishing behaviour in response to declining abalone abundance should be detectable by considering the spatial extent of each dive event. The spatial data collected in the Tasmanian passive GPS data logger program allows us to estimate two spatial forms of effort - the first being the reef area covered by the vessel, and the second being the maximum extent (linear distance) covered by the vessel during a dive as proxies for the area being fished by the diver. Dive polygon area and dive polygon length are highly correlated (Chapter 1), and in these analyses result in almost identical trend lines. Here we use dive polygon area as a measure of effort to contrast with dive time as a measure of effort from both the GPS logger program and dive time recorded on the Docket book data.

This chapter links to two objective:

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.** Objective 3: **Identify methods for detecting hyperstability in CPUE.**

4.2 Methods

Three SAUs in the Tasmanian abalone fishery are used as contrasting case studies - SAU5 in the north-west, SAU21 in the East, and SAU13, the most productive SAU in the south- east. SAU5 is highly spatially structured with variable growth, and has declined rapidly over the past 10 years. With ~90% reductions in TACC, this fishery has recently shown indications of rebuilding. SAU21 in the south-east was rebuilding at the start of the time-series, declined sharply due to two external events in 2016; the most significant marine heat wave recorded in recent times (Oliver et al. 2017), and a significant storm generating the largest easterly swells on record. From 2017, stocks appeared to keep rebuilding through to the end of the time-series used here (2021). SAU13 is the most productive region (Actaeons) in the Tasmanian blacklip abalone region, and currently accounts for more than 60% of the Eastern Zone TACC. This SAU has been slowly rebuilding over this time-series. These three SAUs provide contrasting trends in fishery performance as a way to examine consequences of using different forms of effort to estimate fishery performance.

In Chapter 2, we outlined our standardisation approach modelling catch as a function of effort (Mod1 below).

```
Mod1 <- log(catch) ~ offset(log(time))  +
        fishyear +
        (1 | subblockno) +
        (1 | subblockno:fishyear) +
        (1 | diver_id) +
        (1 | fishmonth)
```

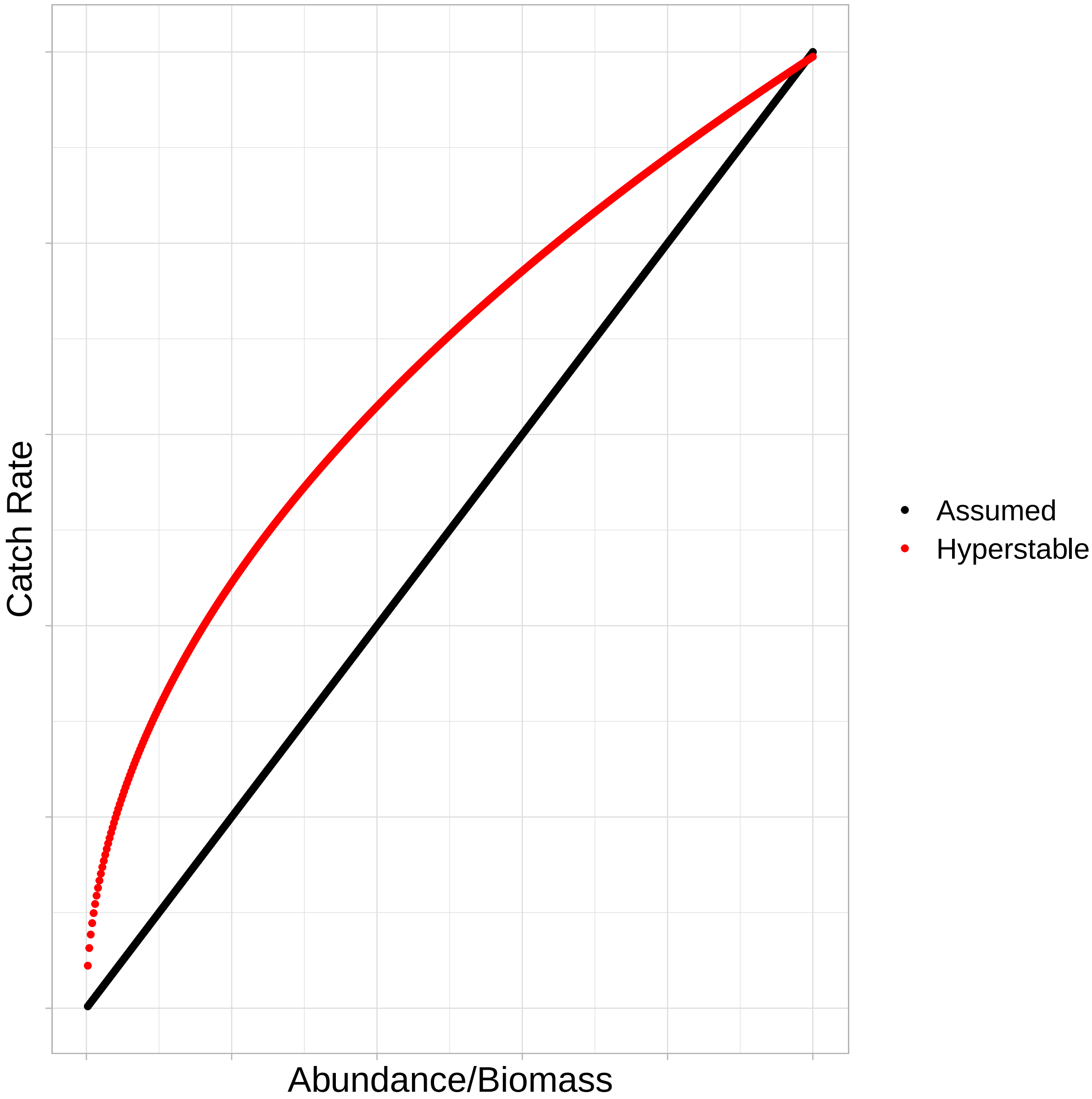


Figure 4.1: Hypothetical relationship between Biomass and CPUE, where CPUE is assumed to be an index of relative abundance (e.g. Harley, Myers, and Dunn 2001, fig. 1)). The black line (Assumed) represents the assumed relationship where the index is proportional to abundance, while the red line (Hyperstable) represents a non-linear hyperstable relationship between the index and actual abundance.

To contrast the effort required to achieve a nominal amount of catch (100kg), the effort and catch terms are reversed in the statistical model. Using this approach we are able to ask the question, what is the change in effort through time required to achieve 100kg of catch? The statistical model is then given by Mod2.

```
Mod2 <- log(time) ~ offset(log(catch))  +
        fishyear +
        (1 | subblockno) +
        (1 | subblockno:fishyear) +
        (1 | diver_id) +
        (1 | fishmonth)
```

To provide a comparison with traditional catch rate modelling, both Mod1 and Mod2 are applied to our case study SAUs. Mod1 is applied to both GPS logger and Docket book estimates of dive time (effort). Bias-corrected geometric mean catch-rates are usually presented when catch-rates are not standardised, but for simplicity here crude catch rate is calculated as total catch / total effort across all dives in each year, and included as a rough yardstick.

The first comparison is simply time as effort from two different sources:

- diver Docket book time (minutes)
- GPS program time (minutes)

Secondly we consider a comparison of effort represented as either dive area or dive time, and compared using mod2:

- GPS logger dive time (minutes)
- GPS dive polygon area (meters^2)

The datasets were restricted to the years 2012-21 for both GPS and Docket book datasets. To provide a reasonable comparison, both spatial and docket datasets operate on a per diver/day basis. For spatial data, individual dive metrics (area, dive time, catch) are aggregated to the scale of day.

A log-normal model was fitted using R (R Core Team 2023) and the glmmTMB package (Brooks et al. 2017), and summary figures produced using the tidyverse suite of R packages (Wickham et al. 2019). Data were batch processed using custom analysis functions and the purrr package (Wickham and Henry 2023). We use predict functions to back-cast the mean change in effort required to achieve 100kg of catch in each year of the time-series.

4.3 Results

4.3.1 Case study 1: SAU5

Fishery performance in SAU5 as assessed using the typical approach of catch as a function of effort (in this case 1 hour of dive time) identifies a sharp decline in catch rate from 2012 through 2018, and then a rebuilding catch rate from 2019 through 2021 (Figure 4.2A). Mean annual catch rates using time as a measure of effort from GPS logger and Docket book datasets are very similar (Figure 4.2A). Catch rates for both datasets decline by around

50% between 2012 and 2018. Crude catch rate follows a similar pattern to the standardised catch rates, with the greatest deviation when catch rates are low (2015-19).

Modelling effort as a function of catch (effort ~ catch) identifies an inverse pattern to mean catch rate trends (catch ~ offset(effort)). Substantial differences in estimated mean effort are required to achieve 100kg of catch by the two forms of effort - time, area (Figure 4.2B). As the fishery declines between 2012 and 2018, the dive time increases twofold, whereas over the same time-period the reef area required to catch 100kg is four times greater in 2018 compared to 2012 (Figure 4.2B). As the fishery improves between 2019 and 2021, the effort required to achieve 100kg of catch return to similar values observed in 2012, for both forms of effort.

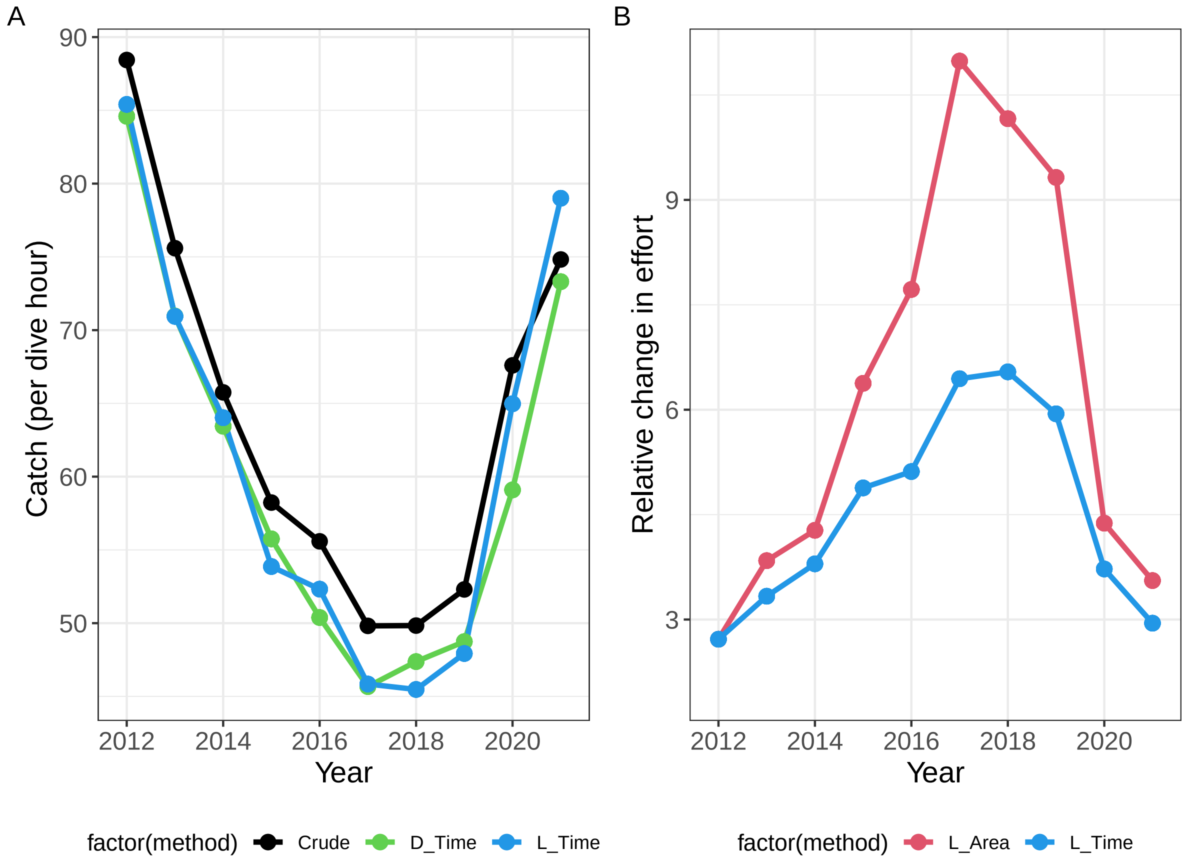


Figure 4.2: Modelling catch and effort for Northern Zone, SAU 5(subblocks 5A, 5B, 5C, 6A, 6C, 6D). A) Crude and standardised catch achieved for 1 hour of effort (docket and logger); B) Relative change in effort (logger area, logger time) required to achieve 100kg of catch.

4.3.2 Case study 2: SAU21

In contrast to SAU5, abalone abundance in SAU21 is rebuilding between 2012 and 2021, with a down turn in catch rate associated with the Marine Heat Wave in 2016 (Figure 4.3A). The typical approach in abalone fisheries of modelling catch as a function of effort (in this case 1 hour of dive time) identifies a sharp decline in catch rate in 2016, followed by an immediate increase in 2017 which continues through 2021 (Figure 4.3A). The GPS logger and Docket book estimated annual mean catch rates using time as a measure of effort are very similar (Figure 4.3A), with minor difference in mean catch rate and a common temporal trend. Catch rates for both datasets nearly double over the time-series, despite the impact of the Marine Heat Wave in 2016. Crude catch rate follows a similar pattern to the standardised catch rates, although there is some deviation in the first three years and again in the final year (Figure 4.3A). Crude catch rate is very similar to the catch ~ offset(effort) approach, although with some trend departure in the first (2012) and last (2021) years.

When modelling effort required to achieve 100kg of catch, a rather different pattern is apparent in SAU21 compared to SAU5. For SAU21, area as our measure of effort declines at a faster rate than when using time as our measure of effort (Figure 4.3B) i.e. the area required to achieve 100kg of catch reduces a greater rate with higher abalone abundance. This pattern is apparent in both rebuilding phases within the time-series - from 2012 to 2015, and then again from 2017 - 2021 (Figure 4.3B). The difference in the rate of relative change in effort required when catch rates are improving, is not as great as the rate of relative change when catch rates are declining (Figure 4.2B).

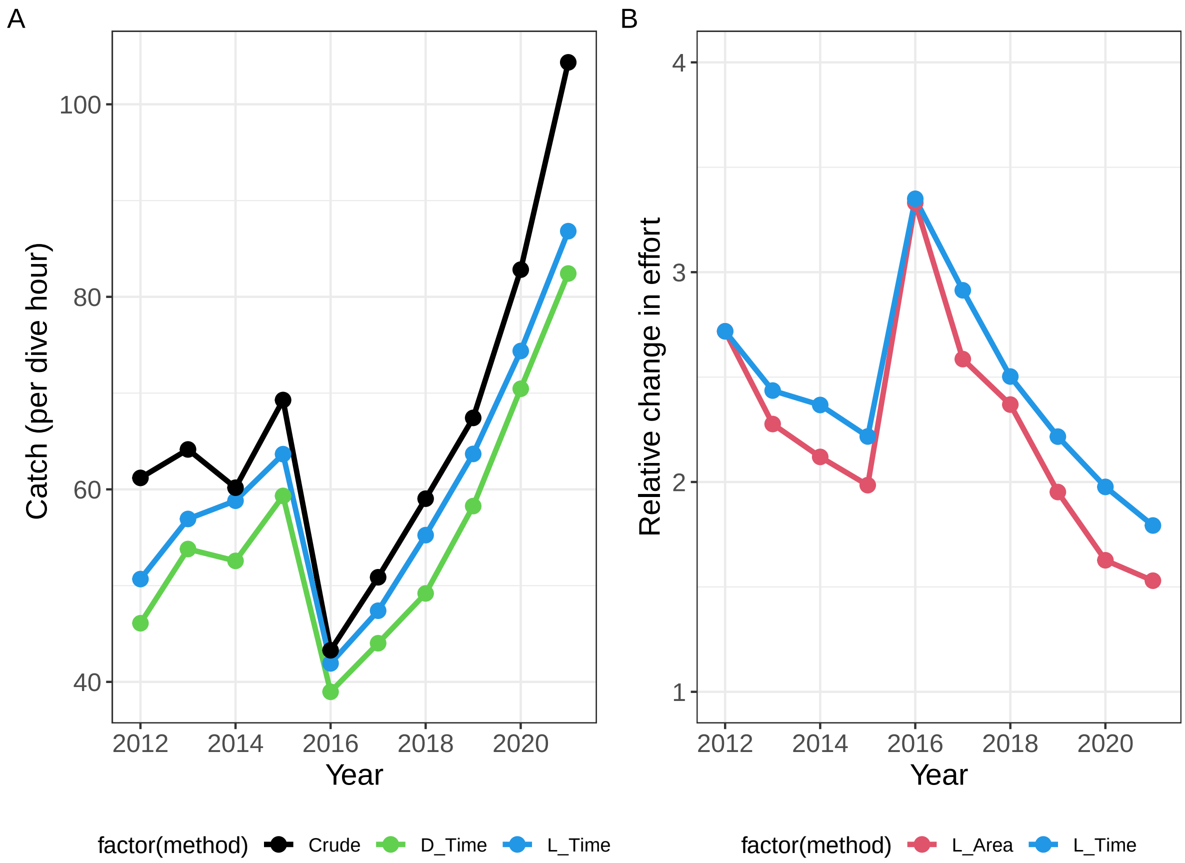


Figure 4.3: Modelling catch and effort for Eastern Zone, SAU21 (subblocks 21A, 21B, 21C). A) Crude and standardised catch achieved for 1 hour of effort; B) Relative change in effort required to achieve 100kg of catch.

4.3.3 Case study 3: SAU13

Block 13 Catch rates ((catch ~ offset(effort)) were stable in the first two years, with a dip in 2014, followed by improvements in catch rate through to 2021 ((Figure 4.4A)). In contrast to SAU5 and SAU21, the crude catch index also departs from the catch ~ offset(effort) standardised catch rate index for the middle phase of the time-series (2016-18) (Figure 4.4A). Mean catch rates for both Docket book and GPS logger estimates of time are almost identical for the entire time-series, using the catch ~ offset(effort) approach.

Estimates of mean effort required to achieve 100kg of catch over the time-series decline over the time-series, with the exception of 2014, where effort required increased sharply (Figure 4.4B). As observed for SAU21, when effort is measured as area of the dive polygon, the index of effort required declines faster than when time is our measure of effort.

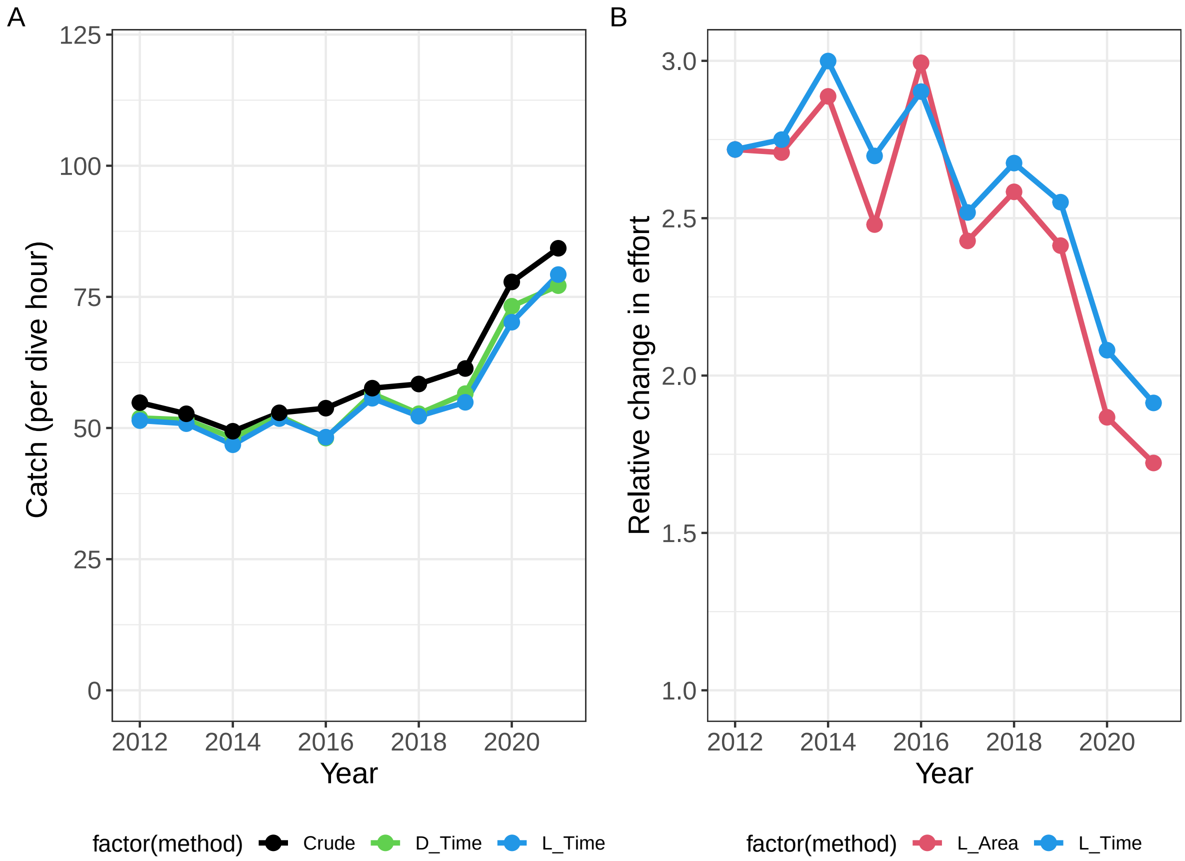


Figure 4.4: Modelling catch and effort for Eastern Zone, SAU13. A) Crude and standardised catch achieved for 1 hour of effort; B) Relative change in effort required to achieve 100kg of catch.

4.4 Discussion

4.4.1 Modelling relative change in effort

Divers routinely comment that they are covering much greater area per dive as abundance declines and by swimming faster as they search for abalone they are able to maintain catch rates (Kg/Hr). Through the GPS logger program, we have captured and quantified that spatial change in fishing activity. When area was used as our measure of effort, the relative change in effort required in SAU5 to harvest 100kg of catch increased markedly over time based measures of effort (Figure 4.2B). Similarly, the relative change in area declined faster than the relative change in time required to harvest 100kg from 2018 onwards as catch rates begin to improve (Figure 4.2A). These results clearly demonstrate that reliance on time as our measure of fishing effort under-estimates the fishing effort required to harvest 100kg of abalone (i.e. catch rate indices using time as a measure of effort are hyperstable). Hyperstable catch rates in an experimental Paua fishery was also observed by Abraham and Neubauer (2015), although this finding was assigned to serial depeltion over the course of the experiment. This is the first empirical confirmation of a change in fisher behaviour leading to hyperstable catch rates in an abalone fishery.

While the rate of change in area was greater than the change in time required to harvest 100kg of abalone, the magnitude of the change over the time-series (2012-21) varied greatly among SAUs. One possible explanation is that the magnitude of the relative change in effort is linked to the level of depletion. In terms of productivity and stock health, we can order the three SAU case studies as SAU13 in the best health, followed by SAU21, and then SAU5 which is recovering from severe depletion. Respectively, the magnitude of relative change was smallest at SAU13 (Figure 4.4B) and largest at SAU5 (Figure 4.2B). Alternatively, the observed patterns in rate of change of in effort required may be linked more strongly to the habitat type and reef complexity, with SAU5 having relatively low reef complexity in comparison to SAU21 and SAU13. However, SAU21 is much more complex in terms of reef structure than SAU13. The low topographic complexity of reef systems in SAU5 may have been a contributing factor to the extent of depletion, but is perhaps unlikely to be a key factor affecting the magnitude of change in relative effort observed in this study.

4.4.2 Hypothetical relationship between CPUE and biomass revisited

As well as differential rates of change expressed using time and area as measures of effort in a declining fishery, we also found evidence of hyperstability in catch rate indices when catch rates are above average and improving. Patterns of relative change in effort for area and time in SAU21 and SAU13 (improving fisheries), demonstrate that time based catch rate indices are also subject to hyperstability at the upper margins of catch rates when abalone abundance is improving. Thus hyperstability is not restricted to depleting stocks, and can affect catch rate indices at both ends of the catch rate range. In hypothetical terms, the non-linear relationship between catch rate index and biomass would be better described by Figure 4.5, than by Figure 4.1. Gaertner and Dreyfus-Leon (2004) alluded to the possibility of hyperstability occurring at upper and lower ranges of stock abundance in their simulations of a tuna purse-seine fishery (Gaertner and Dreyfus-Leon 2004, fig. 3). However, no empirical demonstration of hyperstability of an improving stock has been reported previously in the literature.

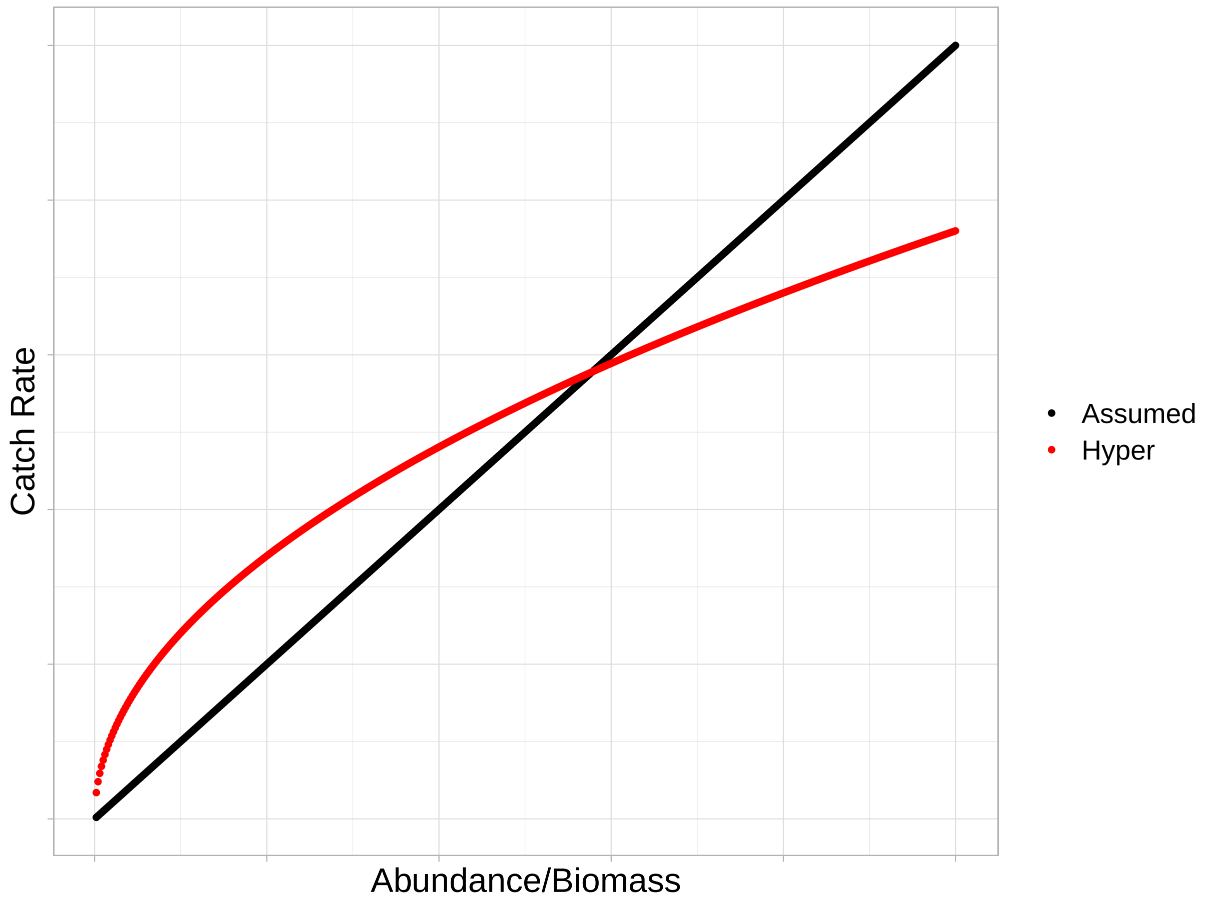


Figure 4.5: Revised hypothetical relationship between Biomass and CPUE (where CPUE is assumed to be an index of abundance). The non-linear relationship between CPUE and abundance is modified to reflect the finding that time based CPUE overestimate abundance when low, and underestiamte abundance when CPUE is high. The black line (Assumed) represents the assumed relationship where the index is proportional to abundance, while the red line (Hyperstable) represents a non-linear hyperstable relationship between the index and actual abundance.

Part III

Quantifying spatial dynamics of reef use: exploiting the grid data

The previous chapters focused on characterising and describing spatial performance measures derived from bivariate kernel utilisation distribution polygons of individual fishing events (KUD). Those chapters directly explore the potential to replace time as a unit of effort with dive polygon area. In Chapter 3 we created maps of residuals which clearly demonstrate a spatio-temporal dynamic of fishing. The chapters in this section focus on how we might use the geo-referenced fishing data to determine the total area of reef utilised, and identify metrics that could provide performance measures for patterns in reef utilisation.

5 Time integrated productivity index (TIP) data

5.1 Introduction

Fleet utilisation of fishing grounds in spatially structure fisheries is rarely described, and not well understood. In addition to choices made by fishers based on personal preference and restrictions imposed by weather conditions, patterns of exploitation may also be driven by fine-scale heterogeneity of stock productivity. Fishers often report a natural cycle of return times to specific fishing grounds, suggesting a strategy of exploiting a section of reef followed by a period of leaving that reef to recover. In a fishery with a large number of fishers, it is feasible that specific locations may be fished more frequently than individual fishers may be aware of. Most experienced fishers are also very adept at identifying sections of reef that are likely to support commercial quantities of catch. Thus the more attractive a section of reef, the greater the likely extent of visitation by fishers, and the greater the likelihood that section of reef will be fished each year. Whereas, some sections of reef may be unattractive, or thought to be unproductive because of past experience, or other location characteristics such as high frequency of white shark sightings. Abalone fishers identify and refer to sections of reef that are resilient to fishing as 'good recovery bottom' (GRB). As yet, we do not know specific locations of GRB, or whether they are permanent or temporary features of the fishing landscape depending on spatial and temporal patterns of exploitation. Habitat quality, food availability, recruitment dynamics along with local fishing pressure will all contribute to a high level of heterogeneity in local exploitation rates, and might be expected to impart a level of stability in relative productivity of local reef patches.

Here we examine spatial structure in fishing activity, as one measure of the spatial complexity of the Tasmanian abalone fishery. One mechanism for characterising and quantifying the dynamics of reef use is to classify the known fishable reef in terms of whether it is fished frequently (every year), occasionally (most years) or rarely (e.g. one year in every 10). In addition to the inter-annual frequency of fishing a given location, it is also informative to understand whether that location is productive in a relative sense, i.e, compared to other sections of reef fished in that year, does it support a higher level of total catch across all fishers fishing at that location. The combined periodicity in fishing and relative yield can be used as an index of productivity over the extent of the time-series available. We reviewed a similar metric developed and applied in an early project (see Chapter 9, Mundy, Jones, and Worthington 2018), and made some minor changes in order to make the metric more informative.

This chapter links to two objective:

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.** Objective 3: **Identify methods for detecting hyper-stability in CPUE.**

Objective 4: **Determine feasibility of spatial data based stock status determination in spatially structured fisheries**

5.2 Methods

5.2.1 Time Integrated Productivity (TIP) Index

In order to address some of the spatial heterogeneity questions raised above, we modified an earlier version of a Time Integrated Productivity Index, utilising the hexagon GRID data to better understand the importance of local reef areas to overall catch SAU. The TIP index may also be useful for identifying and characterising the resilience of patches of reef across the fishery, among other applications.

In this exercise we have used catch/cell (cell = 1 Hectare hexagonal grid cell), with the cells in descending order of catch (grouped by zone and block), separately for each year. Four new fields are calculated in order to plot the concentration area curve (CAC):

- RowNum = rank order of cells based on catch (1:n())
- Tz = rank order normalised to 1
- CumCatch = Cumulative catch started from the hex cell with the greatest catch
- Pt = Cumulative catch normalised to one

The TIP index is based on the quartiles of the Y axis (TIP) of the scaled concentration area curves produced from the above derived variables. A score from 4 to 1 is provided for each quartile, where the cells in the quartile of highest catch are assigned a score of 4, and the cells in the lowest quartile a score of 1 (Figure 5.1). Where a hex cell is not fished in a year, a Zero is assigned to the score for that year. The TIP index is created by summing the score across the available time-series - in this case eight years. The maximum value of the TIP index for an eight year time-series is therefore 32, and the minimum possible score is 1. A cell scores 32 only if it falls in the top 25% of cells in the Concentration Area Curve in all eight years.

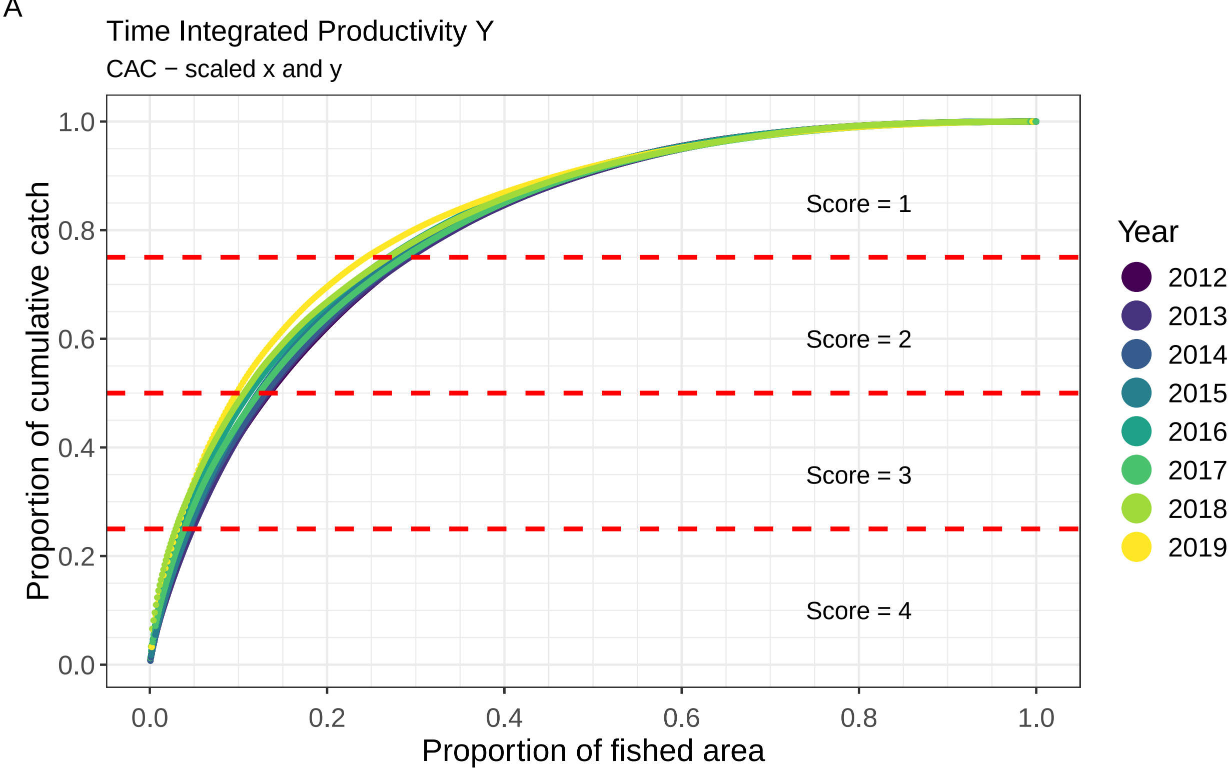
5.2.2 Spatial Structure in TIP index

The spatial structure of the TIP index was examined using Local Indicators of Spatial Autocorrelation (LISA) and Getis-Ord Hotspot analyses. LISA (Local Morans I) and G* (Getis-Ord) statistics were calculated from D Nearest Neighbour matrices with D set as a range of 0m to 110m from the centroid of each target cell. The inter-centroid distance of the Hexagonal grid is 105m, thus a maximum distance of 110m ensures that only the immediate cells are considered neighbours. The K-nearest neighbour approach will classify cells some distance from the target along the edge of known reefs systems as part of that reef system, creating a neighbour network that may not always be identical to a D-nearest neighbour approach. LISA (Local Morans I assigns a class to all observations depending on the relationship between each target cell and its neighbours. If a target cell has a high TIP index and its neighbours also have a high TIP index, it is classified as High-High. When the inverse is true, cells are classified as Low-Low. Where the Target cell has a high TIP index score, and is surround by low TIP index score cells, it is classified as High-Low, etc. Significance of Morans I is tested by permutation (999 reps), and cells with non-significant outcomes are Classed as 'Non-Significant'. Getis-Ord Hotspot analysis provides a continuous Z score, and critical values of the statistic for 90th and 95th percentiles are given. Significant scores that are positive are considered hotspots, while negative scores indicate cold spots.

5.2.2.1 Grid filtering

Most modern abalone fleets work 'live' while fishing. This means the skipper of the fishing boat will follow closely behind the diver, and will retrieve/deliver nets to/from the diver using a weighted drop line. The frequency of drop line throws will depend on catch rates, but is typically every five to ten minutes. Strong wind conditions, or days with higher swells can result in the fishing vessel drifting away from the diver for short periods. When processing spatial data from these days, it can result in minor and potentially spurious activity in grid cells adjacent to fished reef, often less than 5 minutes for a year, or for the entire time-series. Filtering of all grid cells where effort is greater than a minimum threshold is an important consideration. Removing grid cells with less than 5 minutes activity in any year results in a minor, but noticeable shift in the shape of the concentration area curves, and as a consequence, the TIP index (Figure 5.1).

A



B

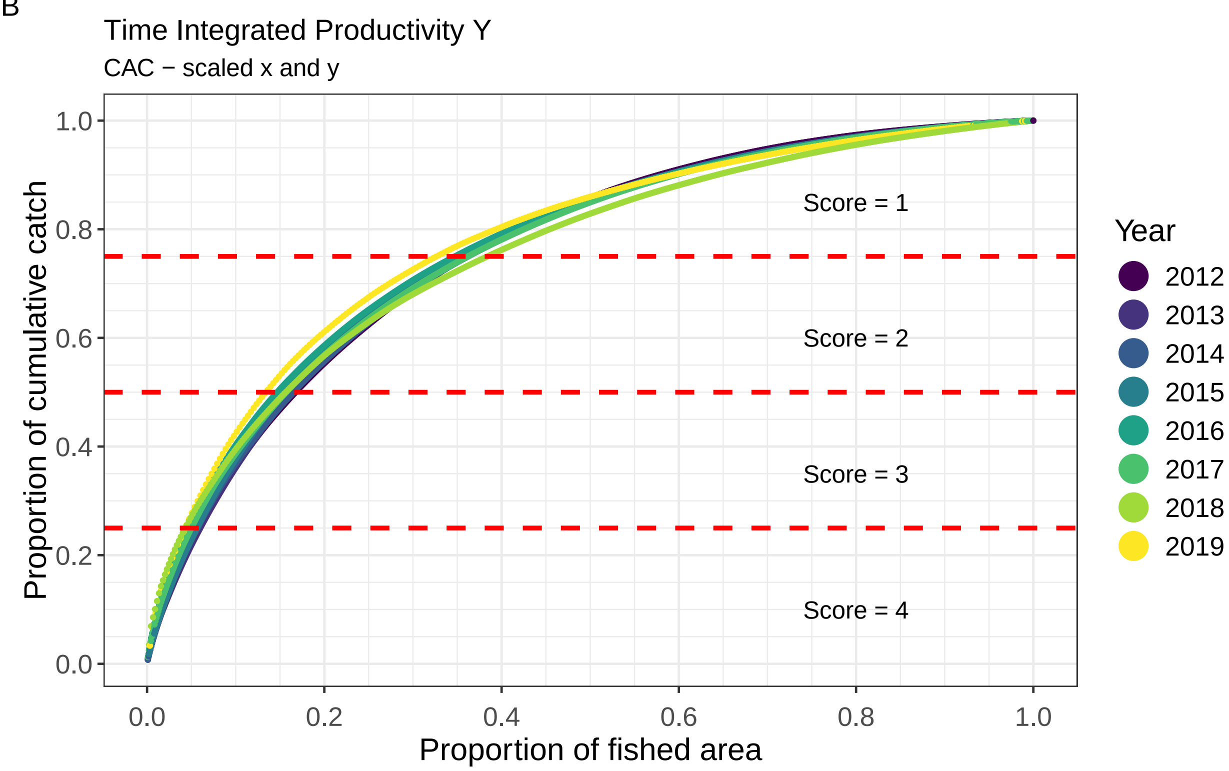


Figure 5.1: Time Integrated Productivity Y. Annual TIP score based on quartiles of cumalitve catch (vertical axis). *A*) full dataset wth no filtering, *B*) dataset filtered to remove cells with less than 5 minutes of effort (i.e. assumesthat cell was not fished in that year).

5.3 Results

Here we use Block 6, N to examine the properties of the TIP index. The frequency distribution of the TIP index is left skewed with relatively few grid cells achieving a TIP index of greater than 15 (Figure 5.2 A). There is a positive non-linear relationship between TIP index and total catch (pooled across the 8 year time series) (Figure 5.2 B). The majority of catch however is harvested from grid cells with a TIP index between 6 and 22, with cells with a very high TIP index delivering relatively little overall catch (Figure 5.2 C). In contrast, cells that are fished in at least six of the eight years account for the majority of the catch (Figure 5.2 D), and cells that are rarely visited (within our time series), support only relative small amounts of catch.

Examining the relationship between TIP and years fished indices is instructive for overall patterns, but masks any variation among individual grid cells within each index class. High TIP index cells, and/or cells that are fished in most years (six or more) have the highest median annual catch, but span a comparatively large range of annual catch values (Figure 5.3). Cells with a high TIP index are necessarily fished in most years, and have high catches relative to other cells in that year. Cells that are fished in most or all years, however show substantial variation, and do not always have high annual catches. While very high TIP index cells typically achieve higher total annual catch, they are relatively uncommon, and thus contribute less to the total catch of an SAU than might be assumed.

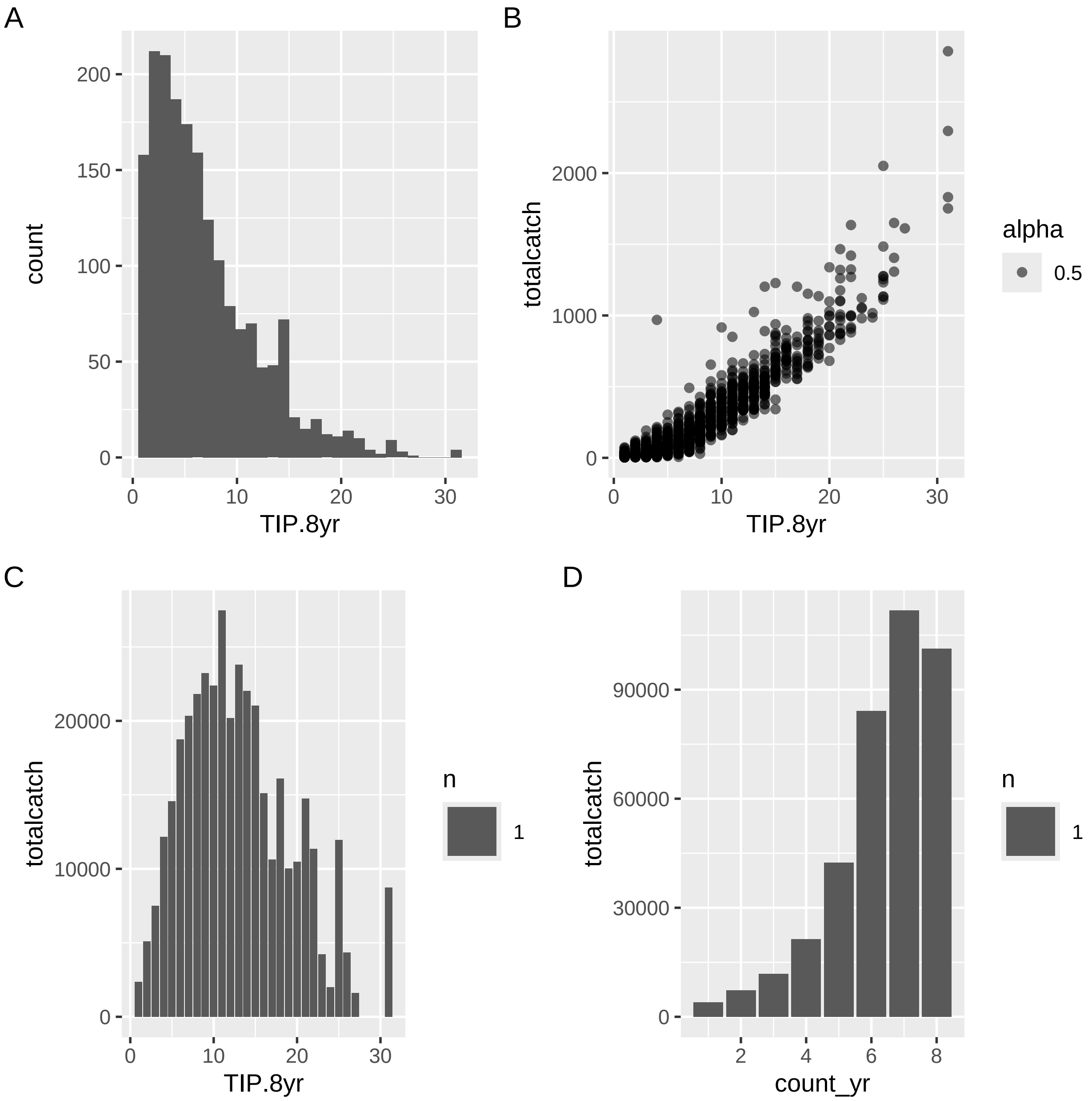


Figure 5.2: Summary information on TIP indices relative to frequency of occurence and catch. A) Frequency of TIP index; B) Relationship between TIP index and total catch (years pooled); C) Total catch by TIP Index (years pooled); D) Total catch by index of years fished (years pooled). Note: years fished is a count of the yrs when a given cell recorded more than the 5 minute effort threshold.

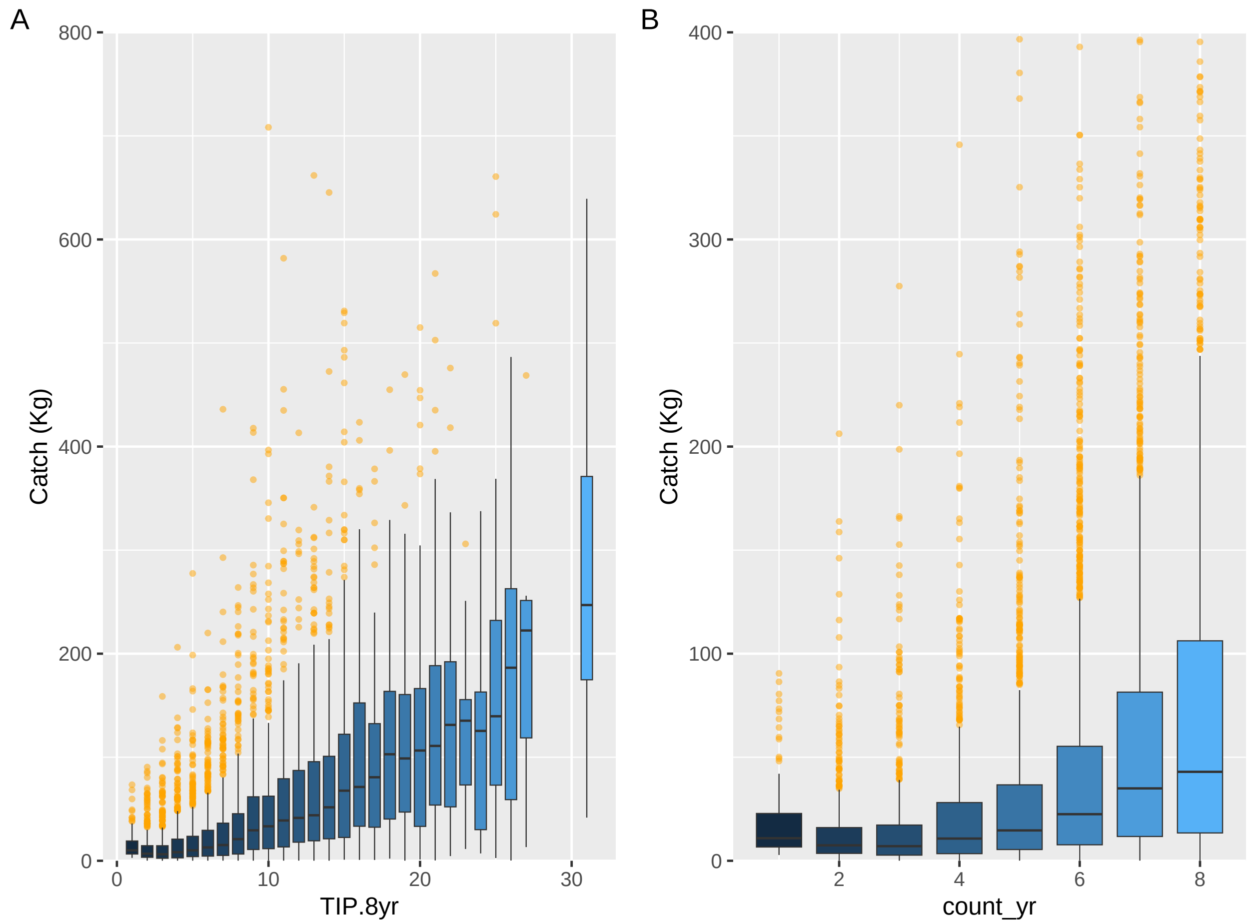


Figure 5.3: A) Boxplot of annual catch by TIP Index. B) Boxplot of annual catch by years fished. Note: crude catch rate calculated from the ratio of catch/effort within each grid cell.

5.3.0.1 Differences in annual catch among TIP and year fished indices

When the relationship between catch and TIP index is considered by year, we see a similar pattern to that observed for the dataset with years pooled, though there is evidence of a shift in the relationship over time. The abalone fishery in the example SAU used here declined over the time period of 2012-19, and we see a gradual lowering of the annual catch achieved by high TIP cells (Figure 5.4). In 2012 and 2013, high TIP cells were achieving a median annual catch over 200kg, whereas by 2018 and 2019, high TIP cells were typically achieving a total annual catch of less than 200kg (Figure 5.4). The shift in median annual catch through time was less marked when we simply examine cells by the number of years they had been fished (Figure 5.5).

Figure 5.4: Annual catch by TIP index in Block 6, Northern Zone, for each year in the time series (2012-19).

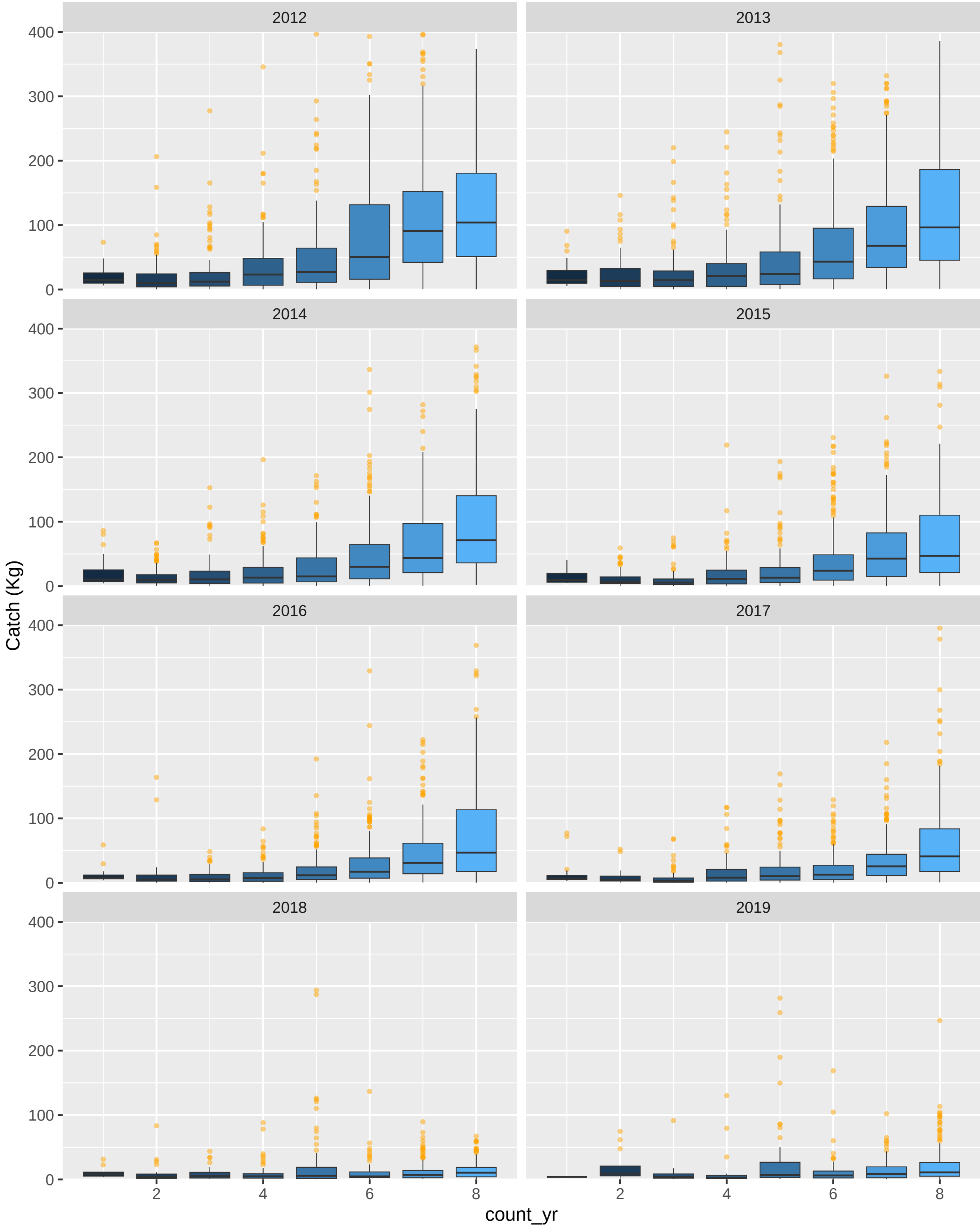


Figure 5.5: Annual catch by year index in Block 6, Northern Zone, for each year in the time series (2012-19).

5.3.0.2 Catch rate as a function of TIP and year fished indices

To simplify graphical presentations, the 32 level TIP index was categorised into 8 ordered classes, with TIP Class 1 containing TIP index scores of 1-4, Class 2 with 5-8, etc. A crude catch rate (CPUE) was calculated for each class by the ratio of annual catch and effort observed for each cell. There was no consistent ordering of CPUE and TIP Class in the example SAU (Figure 5.6 A), and all TIP classes displayed a common inter-annual trend. The relationship between years fished index and CPUE was less consistent. In the first year (2012), cells fished in only 1 year had markedly higher CPUE, while cells that were fished every year had the second lowest CPUE (Figure 5.6 B). At the end of the time series cells that had been fished in most years (7 or 8) had low to moderate CPUE, and the cells that had been fished between three and four years had markedly higher CPUE (Figure 5.6 B), and the cells that had only been fished in 2019 had the highest CPUE.

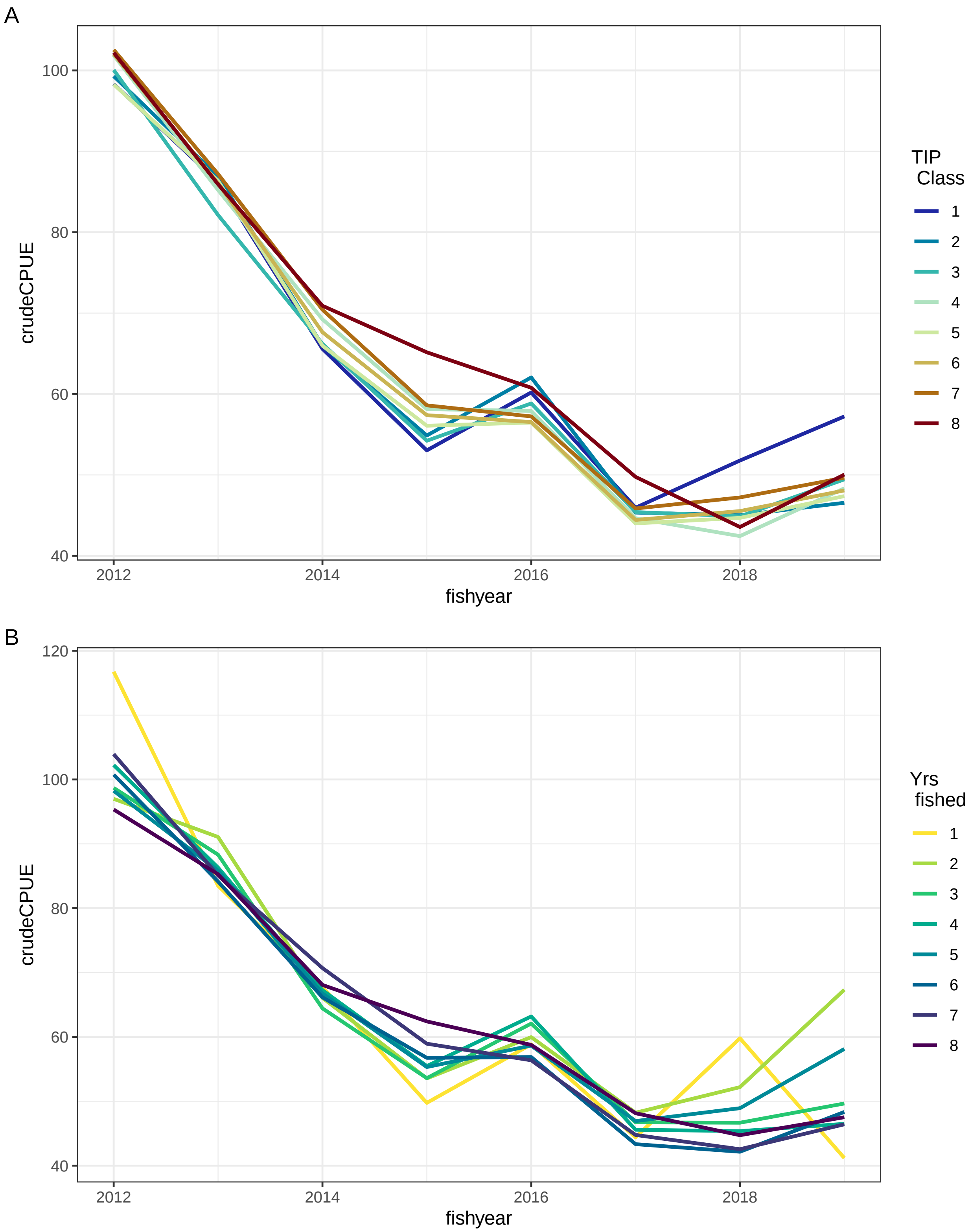


Figure 5.6: A) Crude mean cell CPUE by TIP index categorised into a bin size of 4 (i.e. TIP class of 8 includes cells with TIP score of 29-32 inclusive); B) Crude mean cell CPUE by number of years fished. Note: crude catch rate calculated from the ratio of catch/effort within each grid cell.

5.3.1 Spatial structure of TIP indices

The summaries above demonstrated that high TIP index cells were relatively uncommon, but supported relatively high annual catch, at a similar CPUE to moderate or low TIP cells. The spatial distribution of these high and low TIP index cells is also of interest. For example do they occur as individual cells within the fishing grounds, or as small or large clusters of cells that might identify reefs with underlying characteristics. Local Indices of Spatial Autocorrelation (LISA) and Getis-Ord HotSpot analyses provide a mechanism for examining spatial structure in the TIP index.

The spatial lag of the TIP score is positively correlated with the TIP score (Figure 5.7), confirming that there is spatial autocorrelation in the dataset. A map of Local Morans I colour coded by LISA class (High-High, Low-Low, etc,). The majority of cells are classified as Not-significant (i.e. no local autocorrelation). However cells classed as High-High are distributed through the fishing ground, either as individual cells are as small to large clusters of High-High cells (Figure 5.8). There were relatively few Low-Low cells in the example SAU. Most Low-Low cells were scattered across the fishing ground with only one cluster of any size evident (Figure 5.8). The Getis-Ord HotSpot analyses found a very similar pattern, particularly with the HotSpot cells, but provided a little more detail on ColdSpot cells than the LISA approach (Figure 5.9).

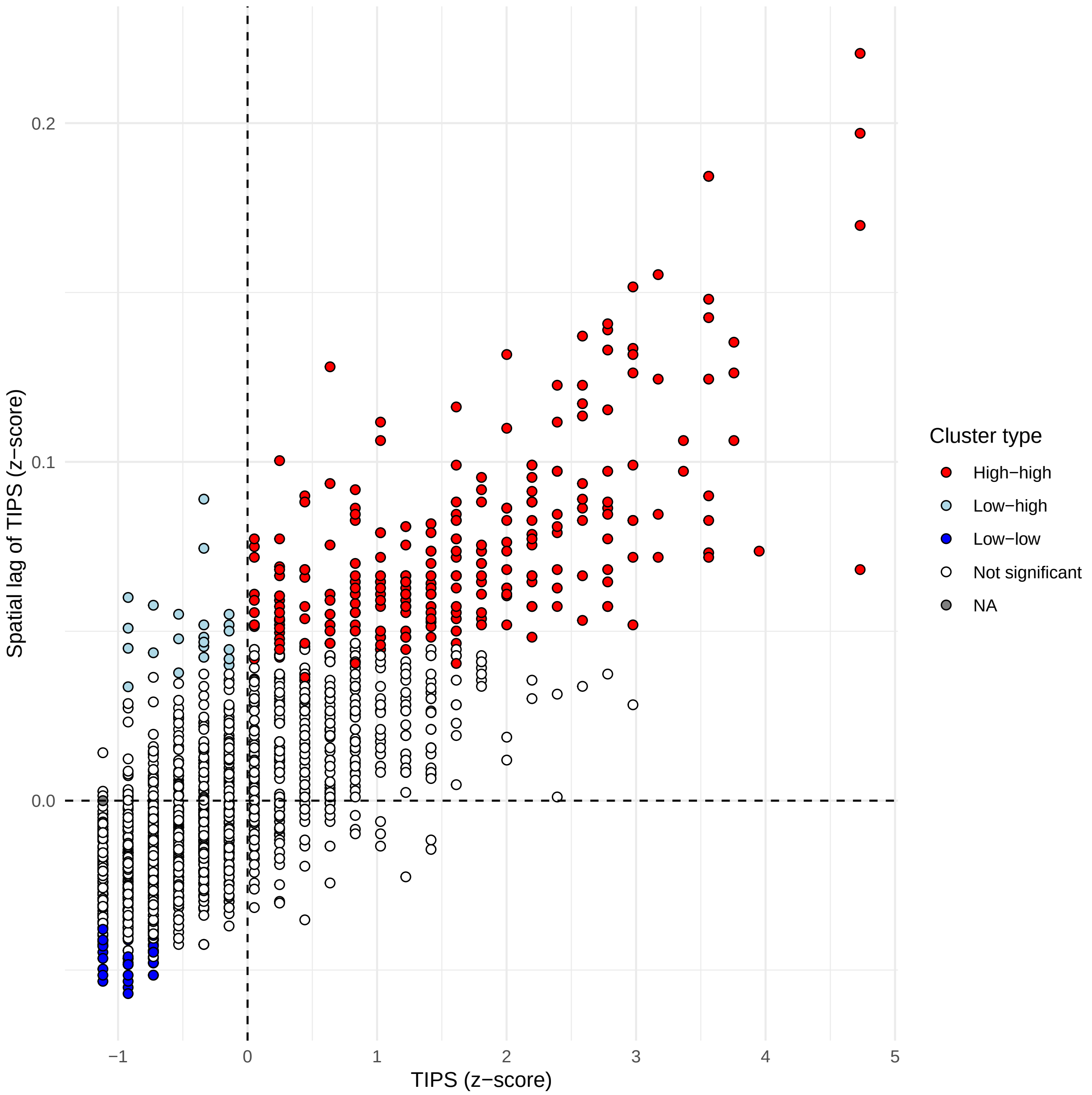


Figure 5.7: Morans I: Spatial lag of TIP against Z score. Cells colour coded by the Local Indicator of Spatial Autocorelation (LISA) index. The High-High class indicates the target cell has a high TIP index, and all neighbour cells also have a high TIP index (i.e. a 'hotspot'). Low-Low indicates the inverse with both Target cells and neighbour cells having low TIP indicies (i.e. a 'coldspot').

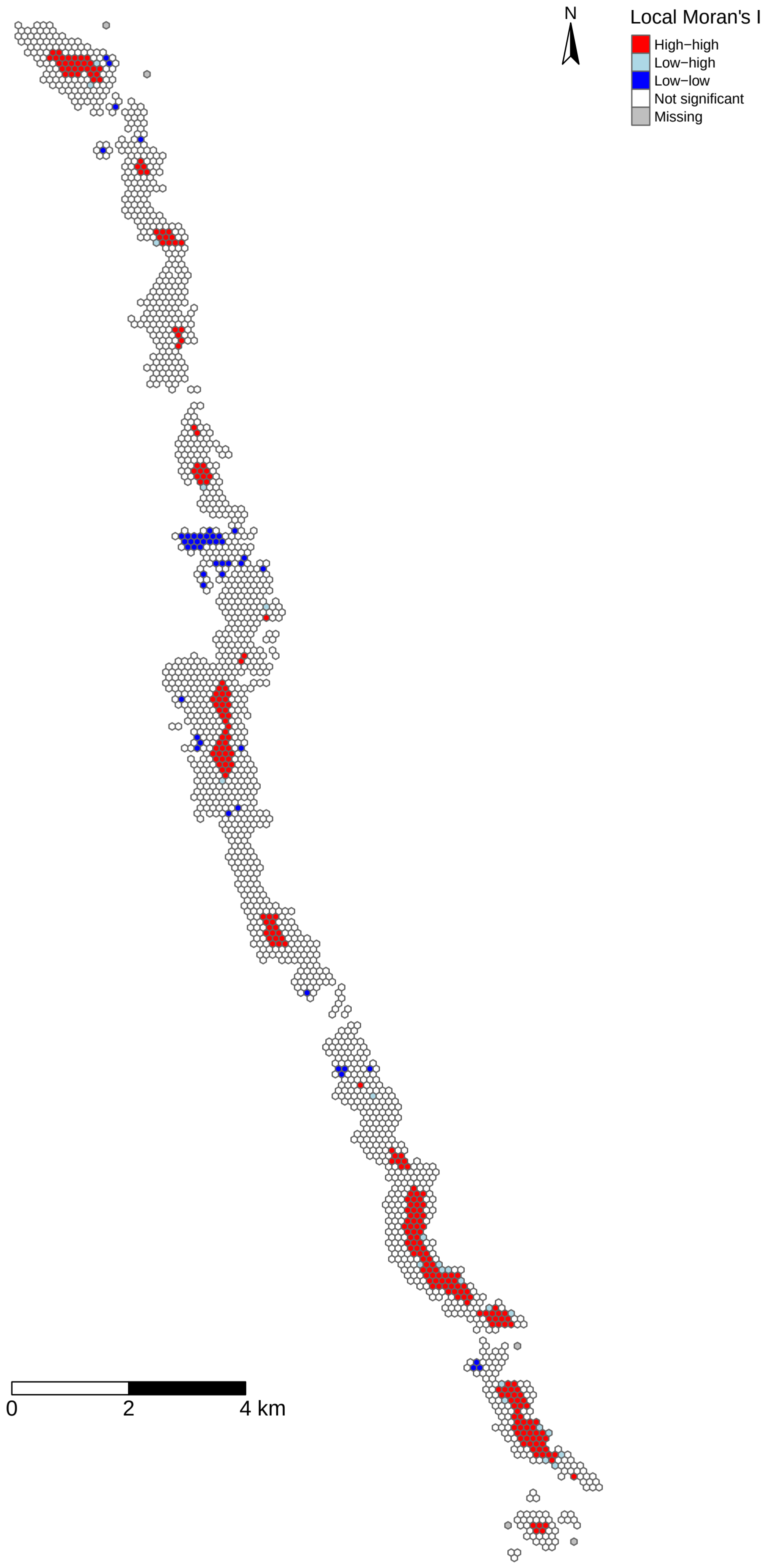


Figure 5.8: Spatial structue of Local Indicator of Spatial Auotcorelation (LISA) index. The High-High class indicates the target cell has a high TIP index, and all neighbour cells also have a high TIP index (i.e. a 'hotspot'). Low-Low indicates the inverse with both Target cells and neighbour cells having low TIP indicies (i.e. a 'coldspot').

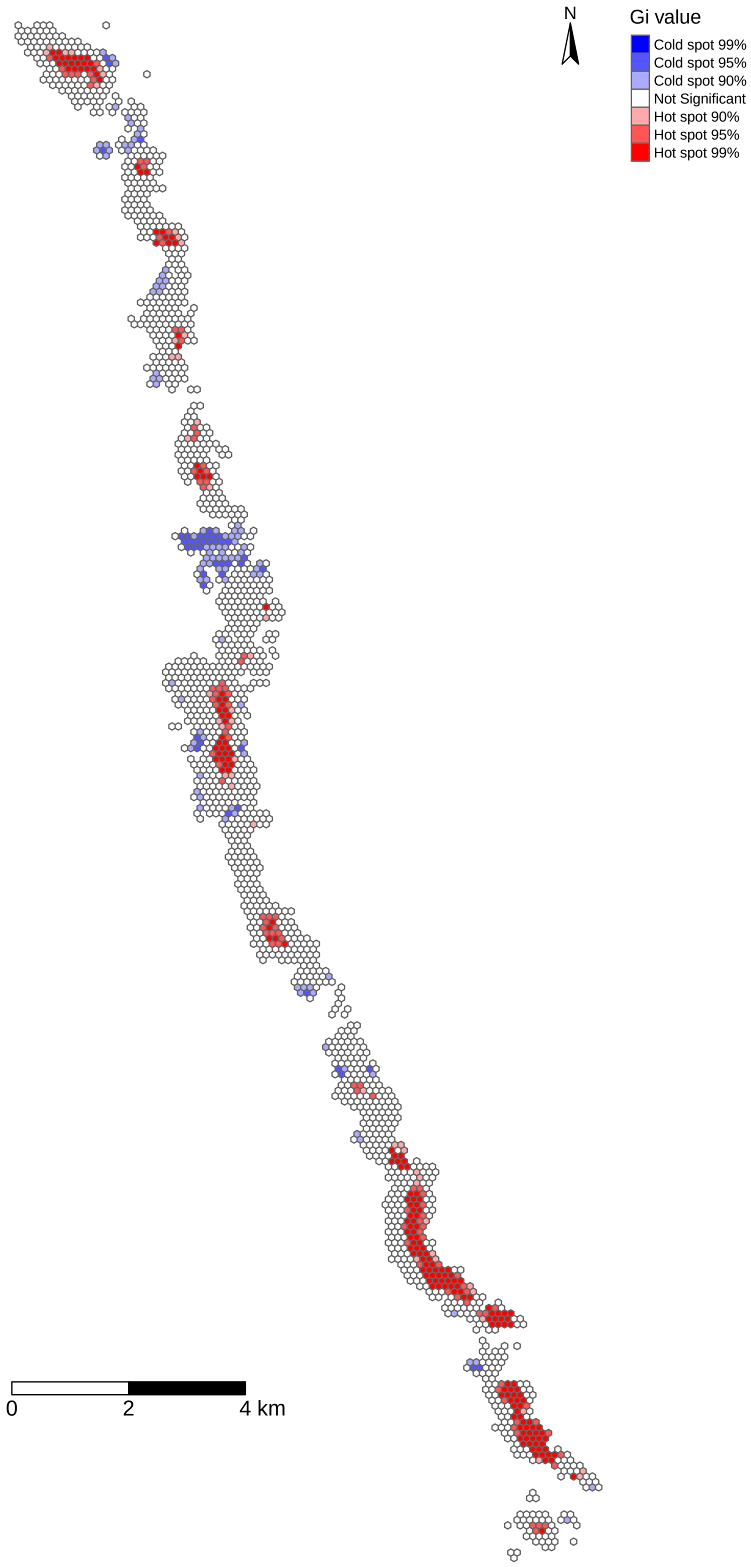


Figure 5.9: Getis-Ord G statistic for HotSpot analysis. When both Target and neighbour cells have high TIP indices, cells ae classified in three bands of 'HotSpot'. When Target and Neighbour cells have low TIP indices, they are classified in three bands indicating a 'ColdSpot'.

5.4 Discussion

The Time Integrated Productivity Index provides valuable insights on the spatial structure of productivity of fishing grounds. Total catch per cell was positively related to both the TIP index and years fished, although the overall contribution of these high TIP cells (or cells fished every year) to the SAU catch was relatively minor as these cells are relatively uncommon. Despite high TIP index cells supporting higher annual catch, there was no substantial difference in catch rates (crude CPUE) across TIP classes, or years fished.

High TIP index cells (hotspots) were not found to be clustered or grouped as a large block within the SAU fishing grounds used as an example in this study. Similarly low TIP index cells (coldspots) were distributed throughout the fishing ground, although the Getis-Ord HotSpot analysis identified an area of low-productivity in the central area of the case-study SAU (Figure 5.9). There was some evidence of a diminishing total annual catch from high TIP cells over time, but this likely to be a broader pattern of depletion across the SAU, as annual catch was simultaneously declining in low TIP index cells Figure 5.4.

The observation that individual areas of reef (identified here by one Hectare grid cells) show a temporal dynamic highlights an important consideration for how we use grid cell data in a fishery assessment. Critical, is that we do not assume that the spatial structure observed in any given year is static. While we did find that some reef areas consistently contribute high catch to the total SAU catch, the majority of reef was temporally variable in its annual contribution to observed catch.

6 Estimation of Total Fishable Area

6.1 Introduction

GPS logger equipment is required to be carried by all divers within the Tasmanian abalone fishery since 2012. This section will consider that GPS based data across the years 2012-19. Only the Western and Eastern zones will be considered, which avoids the complication of the mixed fishery for blacklip *Haliotis rubra*) and greenlip (*H. laevigata*) abalone that can occur in the Northern and Bass Strait zones. The GPS based data is much more spatially detailed than the more classical docket data and provides an opportunity to describe the empirical geographical extent of the the fishery in each of the statistical blocks used in Tasmania in the fishery assessments. The docket data reports at a subblock area at its finest whereas the GPS data is reported every 10 seconds. This provides the possibility of characterizing the actual extent of fishable reef around Tasmania. The question to be addressed here is:

What is the total fishable area of reef within each statistical block or Statistical Assessment Unit (SAU) in the eastern and western zones?

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.** Objective 3: **Identify methods for detecting hyper-stability in CPUE.**

Objective 4: **Determine feasibility of spatial data based stock status determination in spatially structured fisheries**

6.2 Methods

6.2.1 Identifying noise in the hexagon grid cell data

Mundy (2012) describes generating: "A hexagon grid cell network extending from the coast to a distance of two kilometres offshore was created for the entire coast of Tasmania and all offshore islands." (Mundy (2012), p42). Using software to count the number of GPS data points in each of these cells (each point represents 10 seconds of effort), allows the areal extent and real-time usage of the reefs to be monitored. The use of depth loggers by the divers enables any GPS data points obtained while the divers are not underwater, or when travelling by vessel to be eliminated from the dataset. However, because of the uncertainty added through the location of the dive vessel not being an exact reflection of the location of the diver there will be some cells which only appear to contain effort data, which, in reality, are only noise due to vessel drift. Thus, estimating the extent of the fishable reef is not a case of simply summing the number of cells within which effort is recorded. The intent here is to estimate the intentional fishing of the reef and to find a method that facilitates the removal of false positive cell records.

Table 6.1: The data fields used in the TIP database for the years 2012-19, showing their general properties. 'coastdist is the distance in m of the centroid to the coast, divers is the number of unique divers that have visted an cell across years, exposure is an index of wave exposure, count_yr = occur, totE, totC, and maxE reflect cell values across all years.

	varname	Index	isNA	Unique	Class	Min	Max	Example
1	cell	1	0	60532	integer	3756	1673900	3756
2	zone	2	0	4	character	0	0	W
3	block	3	0	56	integer	1	57	12
4	subblock	4	0	145	character	0	0	12C
5	coastdist	5	85	60482	numeric	7e-04	4989.223	1.851393
6	minutes	6	0	3391	numeric	0.1667	1687.3333	7.5
7	catch	7	0	212552	numeric	9e-04	3125.3839	26.75947
8	divers	8	0	34	integer	1	34	1
9	year	9	0	8	numeric	2012	2019	2016
10	exposure	10	0	17234	numeric	0	817934.8539	385243.6
11	tipx	11	0	32	numeric	1	32	2
12	count_yr	12	0	8	numeric	1	8	1
13	TIP	13	0	32	numeric	1	32	1
14	long	14	0	60532	numeric	143.7896	148.5072	146.3741
15	lat	15	0	60532	numeric	<NA>	<NA>	-43.74261
16	occur	16	0	8	numeric	1	8	1
17	totE	17	0	7635	numeric	0.1667	9881.5	7.5
18	totC	18	0	59096	numeric	0	11.2643	0.02675947
19	maxE	19	0	2674	numeric	0.1667	1687.3333	7.5

In addition, of course, when a diver or set of divers fishes in a particular cell, they will only fish on rocky reefs that potential hold abalone. The cells are just a grid arranged around the coast and will contain reef and sandy habitats to different degrees. What this means is, even if it is possible to find a way in which to exclude those cells that only accidentally contain effort records, any sum of the remaining 1 Ha cells will be biased high.

In this section the GPS logger data, summarized as 1 Ha cells, will be characterized and different approaches for selecting those cells that do not represent intentional abalone fishing will be compared to illustrate their effects. The objective is to identify the advantages and disadvantages of each approach and finally select the best compromise for identifying the total fishable area and subsequently for characterizing how the divers fish the reefs around Tasmania.

First the cell data will be characterized to illustrate the available information. The assumption is that there are two types of cell, those that experience true intentional fishing and false positives that only appear to experience fishing because of vessel movement. Those that experience true fishing will be defined as **true** cells and those that are only accidental will be referred to as **false** cells.

Within the GPS data from 2012-19 (see **(Table 6.1)**, for all blacklip abalone , without further filtering there are 220639 observations across a total of 0 separate 1 ha grid cells. These represent a total of 11021.899 tonnes of catch from 156340.2 hours of effort across the eight years.

The parts of **(Table 6.2)** that will be used are:

1) **minutes**, each cell sums up the number of GPS data points it contains. That number divided by 6 gives the number of minutes of effort.

2) **divers**, the number of separate dives that contribute to the effort in a cell is the number of divers it contains.
3) **catch** is obtained by dividing the reported catch for each dive by the number of 10 second GPS points in that dive and allocating that catch relative to the number of points in each cell visited. As the catch per dive (and per number of GPS points) differs between divers it is only the effort data that is invariant between divers.
4) **year** is the standard calendar year over which quotas are allocated.
5) **TIP** has been explained elsewhere. It is derived from the cumulative catch curve for the selected region of the fishery (in the cases considered here that is limited to statistical blocks), which is subdivided into four quartiles on the y-axis (25th, 50th and 75th). With eight years of data the maximum value possible is 32
6) **occur** is simply the number of years, across the eight available, in which effort is recorded in a cell.
7) **totE** is the total effort recorded within the same cell across the eight years.
8) **totC** is the total catch recorded within each cell across the eight years.
9) **maxE** is the maximum effort in any of the years for which effort is recorded within each cell.

6.2.2 Example Blocks

Four example blocks will be used to illustrate how the cell summary data can be used to describe the spatial dynamics of fishing; only blacklip abalone data (*Haliotis rubra*) will be used. Analyses will be conducted on all four blocks but the results for some blocks may be relegated to an appendix to allow the text to progress more fluidly. The example blocks are:

1) Block 6 in the Northern Zone is an example of a block that has been heavily exploited and in 2008 had its Legal Minimum Size reduced from 136mm to 132mm following an industry driven initiative.

2) Block 11 from the Western Zone is an example of a highly productive west coast block whose catches were increased following zonation in 2000 and experienced a long decline in CPUE following that increase.

3) Block 13 in the Eastern Zone is the primary fishing block on the east coast of Tasmania. Its productivity is remarkable with catches averaging 250 t with rather large changes through time. The productivity is largely based around the Actaeon Islands in subblock 13E.
4) Block 21 in the Eastern Zone had a long history of catches of around 50 tonnes from 1992 - 2011, followed by lower catches of about 20 t.

6.3 Results

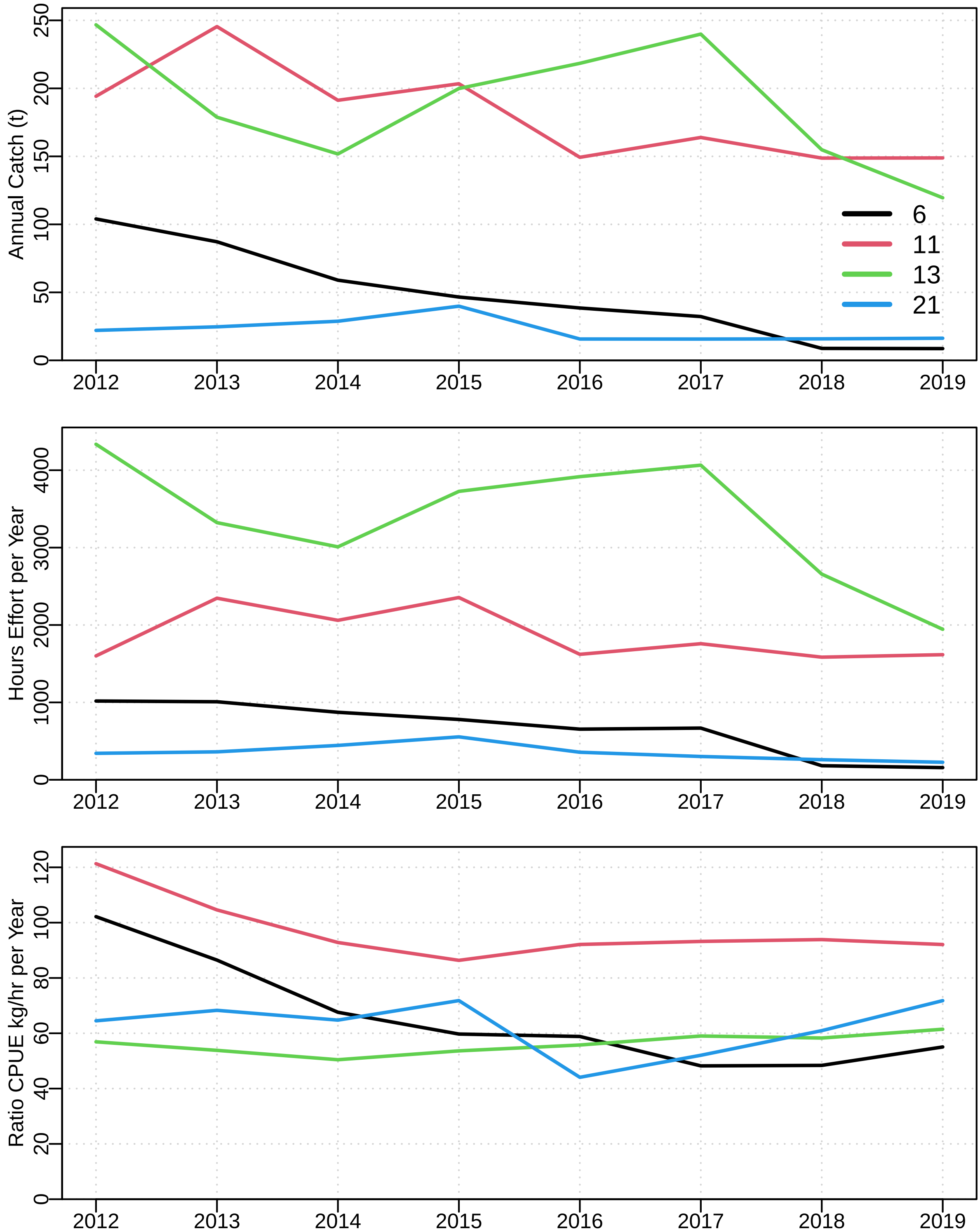


Figure 6.1: The annual catch, hours of effort, and ratio cpue (total catch/total effort) in the GPS data logger data set for each of the four example blocks.

While the catch levels in block 13 east approximately match those in block 11 the effort is much greater, which explains the higher CPUE in block 11. The results for block 6 illustrate its slow depletion over the 8 years and for block 21 demonstrate its relative stability.

Table 6.2: Summary information for each block across all years. nobs is the number of observations or records, ncells is the number of unique cells reporting effort, hours is the total effort, catch is the total catch, catchKg/Hex is the average catch in Kg per cell per block, and the RatioCE is the ratio CPUE for the block averaged across years.

	nobs	ncells	hours	catch t	catchKg/hex	RatioCE
6	8918	0	5339.1	384.836	Inf	72.079
11	24730	0	14942.6	1445.059	Inf	96.707
13	13615	0	26978.7	1509.916	Inf	55.967
21	4877	0	2844.4	178.727	Inf	62.834

Block 11 has more than double the apparent area of any other block but the averaged catch per cell in block 13 is more than double that in block 11. These differences become apparent once a strategy is implemented for removing false positive cells from consideration.

6.3.1 Frequency of Occurrence of Records

By tabulating the frequency of occurrence of records for each combination of occur and TIP it becomes apparent that most cells are visited in fewer than eight years. For example, by considering only block 11 (18178 / 24730 = 0.735; 73.5% visited in only 1 - 7 years)

There are only 5881 unique cells with reported effort in block 11 across the years 2012-19. There are 24730 records because many cells occur in multiple years. An alternative view would be to determine the number of unique cells found in each combination of occur and TIP. Note in the column of occur = 8 it is possible to divide each value by 8 to determine the number of unique cells, and the values in occur = 1 already equal the number of unique cells. However, in all other combinations the number of unique cells would need to be determined individually. The equivalent analyses for the other example blocks can be found in the appendix.

Table 6.3: The frequency of occurrence of each combination of occur and TIP within the 24730 observations within block 11's dataset.

	1	2	3	4	5	6	7	8
1	1229	0	0	0	0	0	0	0
2	35	1365	0	0	0	0	0	0
3	6	122	1419	0	0	0	0	0
4	2	40	186	1468	0	0	0	0
5	0	16	108	360	1430	0	0	0
6	0	6	42	180	510	1356	0	0
7	0	2	24	100	270	660	1050	0
8	0	0	18	40	200	456	728	424
9	0	0	3	24	110	468	497	488
10	0	0	0	12	65	234	672	488
11	0	0	0	4	45	174	462	440
12	0	0	3	4	40	120	378	456
13	0	0	0	0	10	78	336	392
14	0	0	0	0	15	42	196	456
15	0	0	0	0	0	54	182	520
16	0	0	0	0	0	24	119	464
17	0	0	0	0	5	18	91	208
18	0	0	0	0	0	0	105	328
19	0	0	0	0	0	6	56	328
20	0	0	0	0	0	0	35	256
21	0	0	0	0	0	0	28	232
22	0	0	0	0	0	0	28	248
23	0	0	0	0	0	0	0	240
24	0	0	0	0	0	0	7	120
25	0	0	0	0	0	0	0	96
26	0	0	0	0	0	0	0	80
27	0	0	0	0	0	0	0	104
28	0	0	0	0	0	0	0	40
29	0	0	0	0	0	0	0	56
30	0	0	0	0	0	0	0	32
31	0	0	0	0	0	0	0	40
32	0	0	0	0	0	0	0	16

Table 6.4: The frequency of occurrence of unique cells in each combination of occur and TIP within the 5881 unique cells within block 11's dataset.

	1	2	3	4	5	6	7	8
1	1229	0	0	0	0	0	0	0
2	35	683	0	0	0	0	0	0
3	6	61	473	0	0	0	0	0
4	2	20	62	367	0	0	0	0
5	0	8	36	90	286	0	0	0
6	0	3	14	45	102	226	0	0
7	0	1	8	25	54	110	150	0
8	0	0	6	10	40	76	104	53
9	0	0	1	6	22	78	71	61
10	0	0	0	3	13	39	96	61
11	0	0	0	1	9	29	66	55
12	0	0	1	1	8	20	54	57
13	0	0	0	0	2	13	48	49
14	0	0	0	0	3	7	28	57
15	0	0	0	0	0	9	26	65
16	0	0	0	0	0	4	17	58
17	0	0	0	0	1	3	13	26
18	0	0	0	0	0	0	15	41
19	0	0	0	0	0	1	8	41
20	0	0	0	0	0	0	5	32
21	0	0	0	0	0	0	4	29
22	0	0	0	0	0	0	4	31
23	0	0	0	0	0	0	0	30
24	0	0	0	0	0	0	1	15
25	0	0	0	0	0	0	0	12
26	0	0	0	0	0	0	0	10
27	0	0	0	0	0	0	0	13
28	0	0	0	0	0	0	0	5
29	0	0	0	0	0	0	0	7
30	0	0	0	0	0	0	0	4
31	0	0	0	0	0	0	0	5
32	0	0	0	0	0	0	0	2

If the proportions in each *occur* column are determined, **(Table 6.5)**, the proportion of unique cells experiencing less than 8 years of fishing is up to 86%, and there are 21.6% that are only fishing in single years. If, however, the distribution of minutes of effort is considered (**Figure 6.2**) it is clear that there are a large number of cells x year observations that contain only very small amounts of effort minutes. A value of less than around 3 minutes would not normally represent where a fisher has intentionally entered the water and then changed their minds and exit so as to fish elsewhere. The important point is whether the diver entered the water with the intention of fishing, which assumes they would only enter the water where they would expect to find abalone. It may be the case that they then almost immediately exit the water to try elsewhere but at least they entered the water intentionally. A threshold of 3 minutes appears to be a minimum period that would be highly unlikely to represent an intentional water entry. Somewhere between 3 - 5 minutes would be required to assess an area to be worthwhile spending any time searching.

Table 6.5: The proportion of cells experiencing different numbers of years of fishing.

	1	2	3	4	5	6	7	8
N	1272	776	601	548	540	615	710	819
Prop	0.216	0.132	0.102	0.093	0.092	0.105	0.121	0.139
CumProp	0.216	0.348	0.450	0.544	0.635	0.740	0.861	1.000

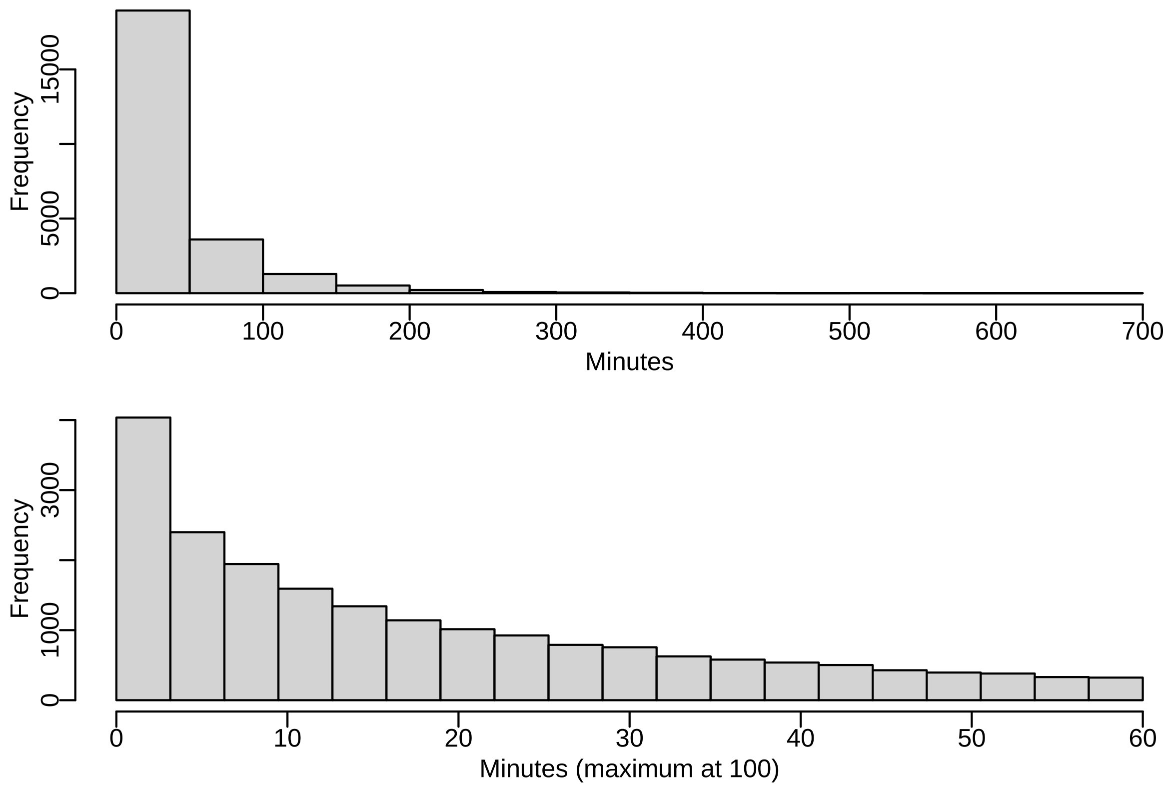


Figure 6.2: The distribution of effort as minutes in each cell within block 11. The lower plot is identical to the upper plot except with an upper bound of 60 minutes and breaks at 3 minutes so as to expand the visibility of the smaller values.

6.3.2 Catch by TIP Value

Now that the notion of TIP and occur have been introduced it is possible to plot out the frequency of records by TIP as well as the catch per TIP value.

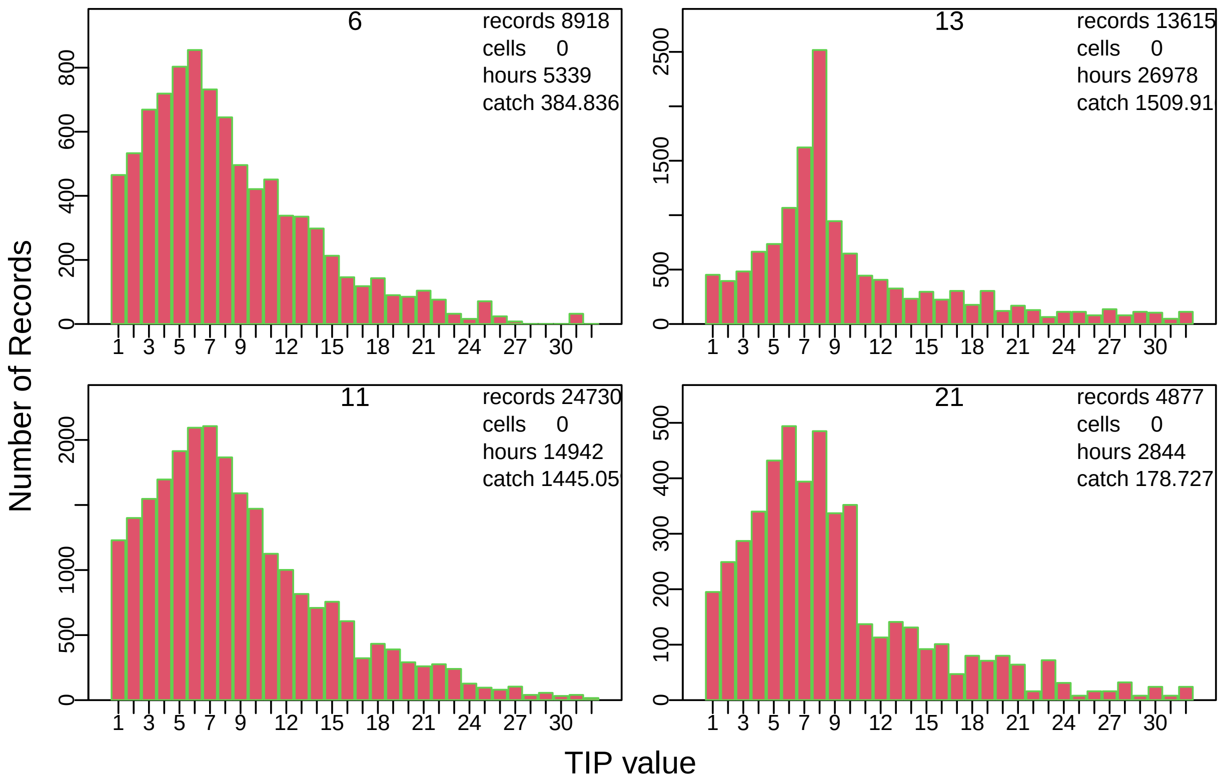


Figure 6.3: The frequency distribution of the TIP index for each example statistical block, 6N, 11W, 13E, 21E.

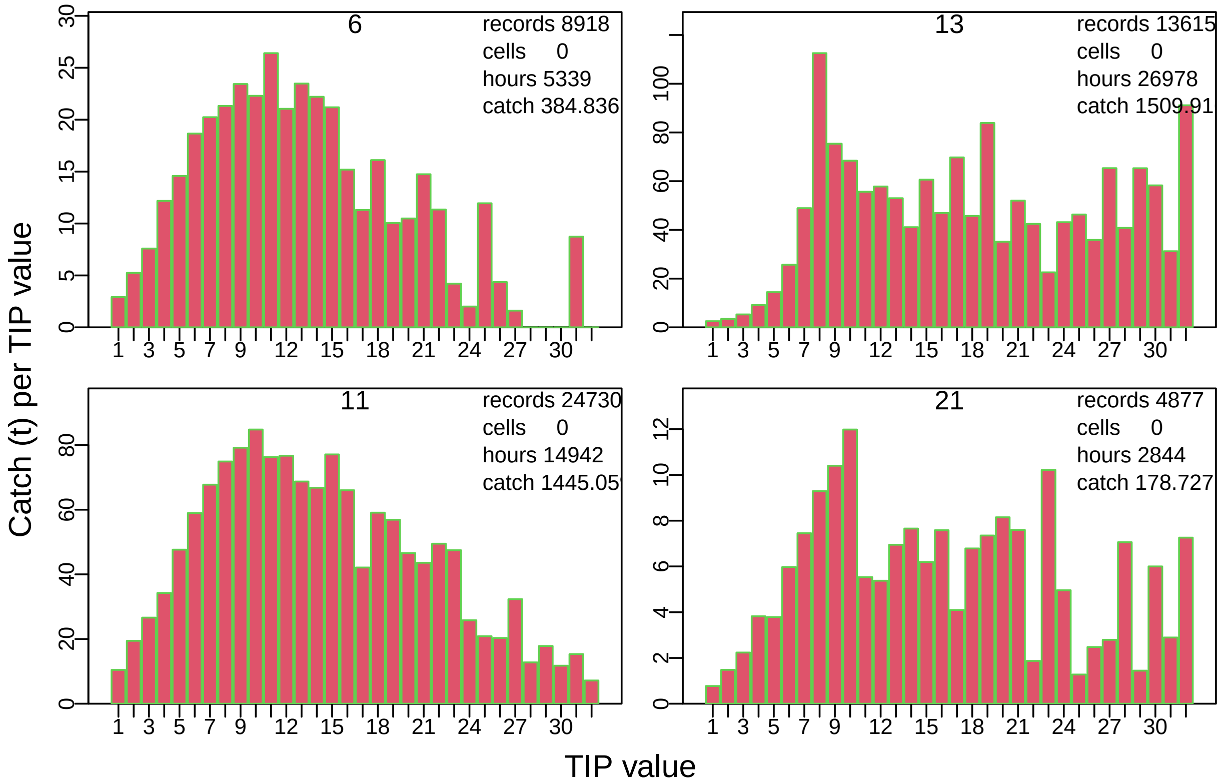


Figure 6.4: The distribution of catches for each TIP value for each example statistical block, 6N, 11W, 13E, 21E. The Eastern Zone blocks have more weight towards the higher TIP values.

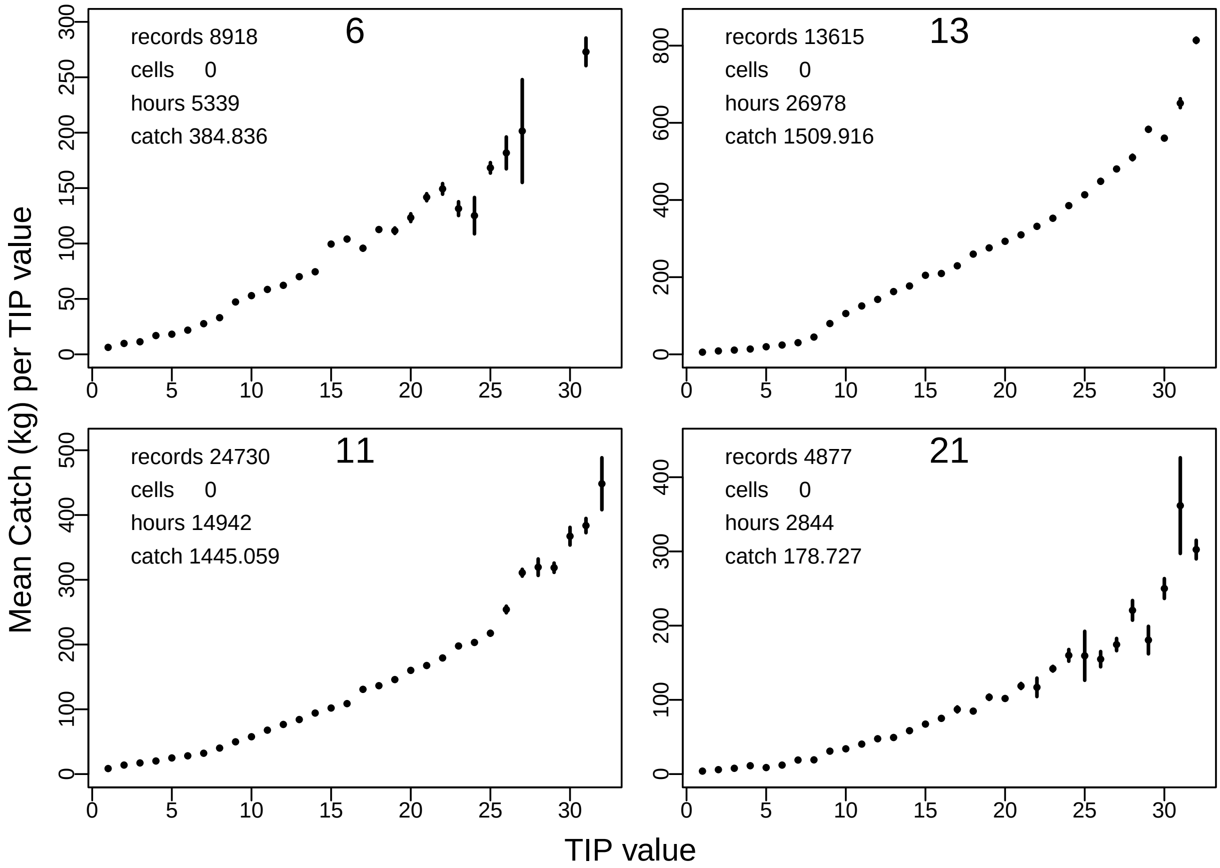


Figure 6.5: The mean catch per TIP value with 99% Normal CI for each example statistical block, 6N, 11W, 13E, 21E. The number of observations for each TIP value are so large that the CI are all small except for the TIP values with low representation.

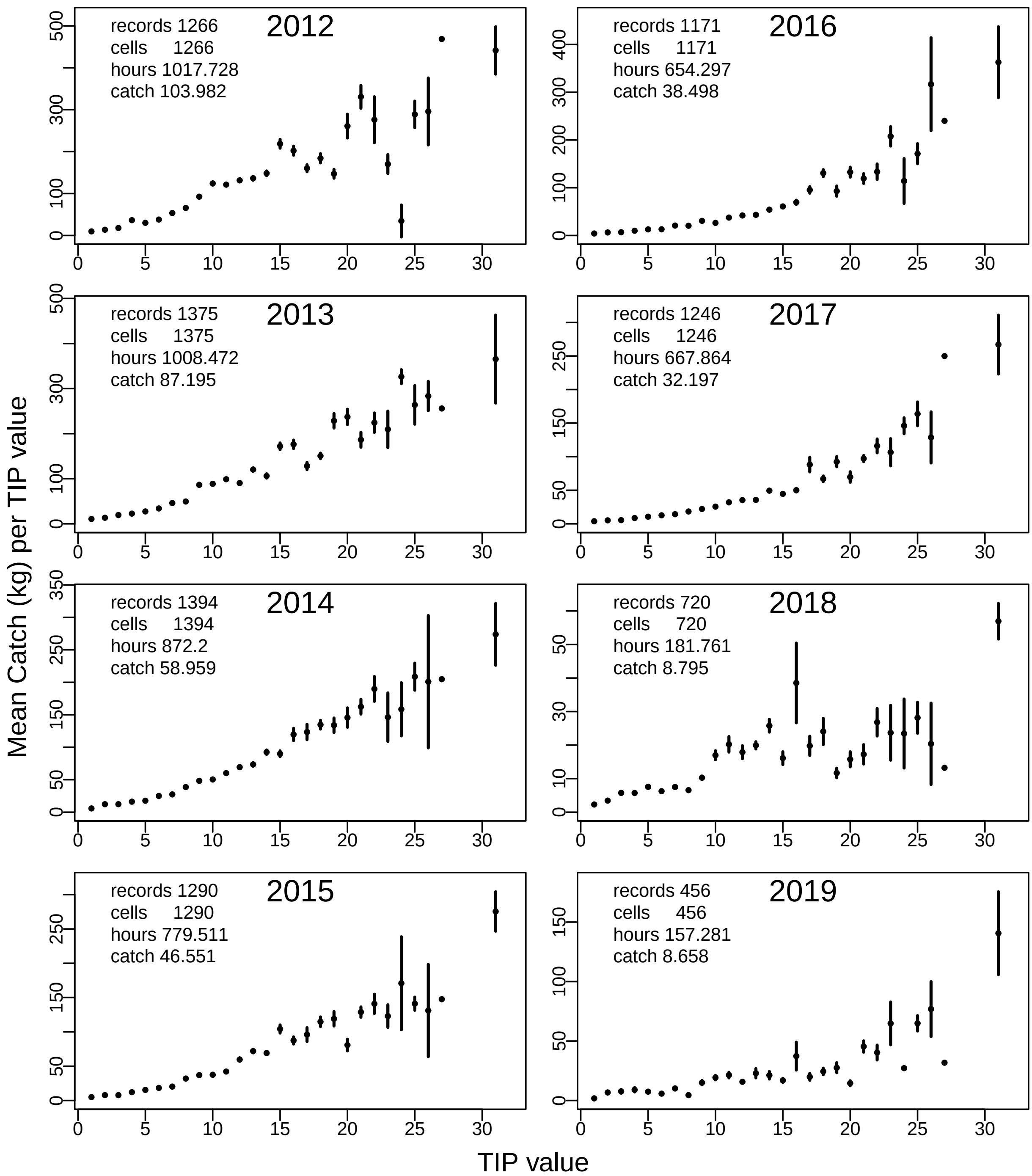


Figure 6.6: The mean catch per TIP value by year with 99% Normal CI for example statistical block 6N.

6.4 Alternative Approaches to Select for Effort

There are at least three different ways that one could select some chosen minimum level of effort:

1) Select X minutes from the **$minutes** field in the dataset.
2) Select Y minutes from the **$totE** field in the dataset.
3) Select Z minutes from the **$maxE** field in the dataset.

The use of X, Y, and Z minutes in the different approaches is merely to represent that because they are different approaches, different levels of effort might be required to gain the desired effect. While there are these alternatives there are some good reasons to exclude two of the options.

If one were to select against, for example, every cell by year combination where **$minutes** <= X minutes, this would certainly remove every instance of cells where only small amounts of effort were expressed. However, this only removes the individual records in each year in which such records occur. This does not appear to be what is needed when attempting to estimate the total fishable areas of reefs where intentional fishing occurs. While this approach would find all cells in each year that have <= 5 minutes, many of those may have much more effort in other years and so really contribute to the total fishable area (they are just not used in every year). This can be illustrated by selecting the unique cells that have at least one year where <= 3 minutes of effort are expressed. If all effort levels for those cells are plotted then much great levels of effort are present in at least some of them in other years (**Figure 6.7**).

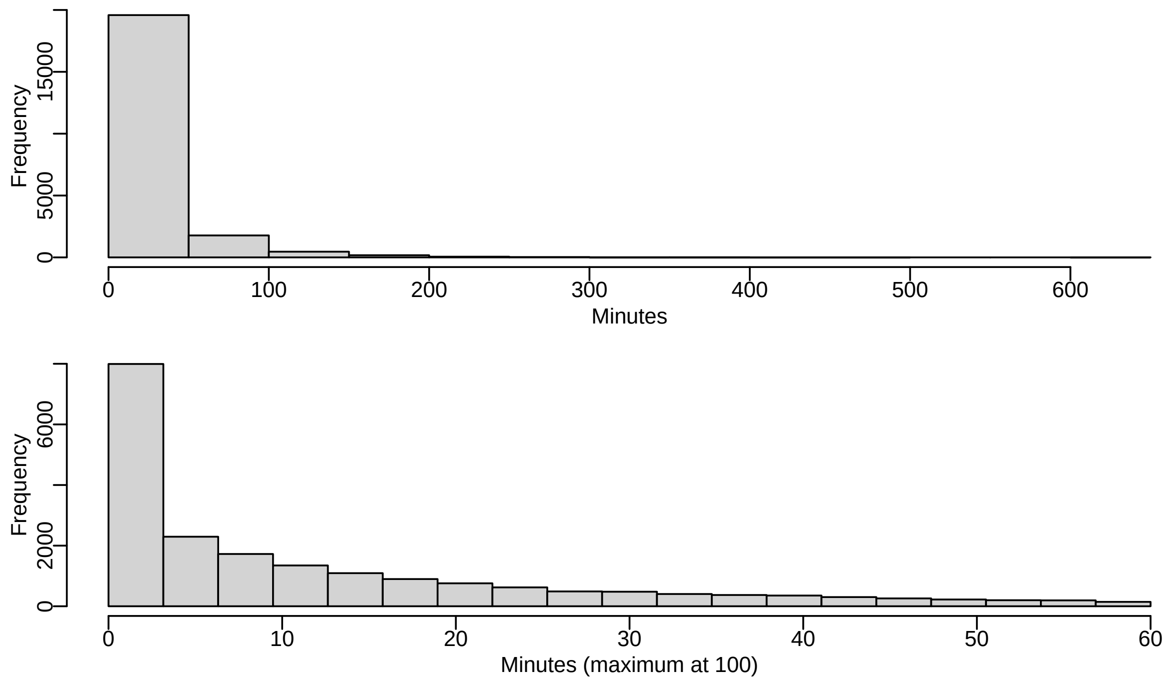


Figure 6.7: The distribution of effort as minutes in each unique cell within block 11 that has at least one year where effort was <= 3 minutes. The lower plot is identical to the upper plot except with an upper bound of 50 minutes so as to expand the visibility of the smaller values.

Using a selection criterion with the *totE* field has different but related problems. The range of effort expended within each cell can vary markedly between years and by using the total effort across years even if one were to use some low value, such as 5 minutes, this might arise from a series of seven very low values and one large value.

Following this train of thought it appears that the optimum approach would be to select some level of the maximum effort that occurs across all years. If a maximum effort (*maxE*) of 3 minutes or less is selected this ensures that there are no unexpectedly high years hidden within the selection and that the cells selected will be removed from all years. The question remains of what level of maximum effort across all years should be used to select against false positive cell records. By plotting up the effect of selecting for a range of different *maxE* values any discontinuities in the observed data can be made apparent. If there are no discontinuities then some value that would represent the time that would exclude any intentional fishing should be selected.

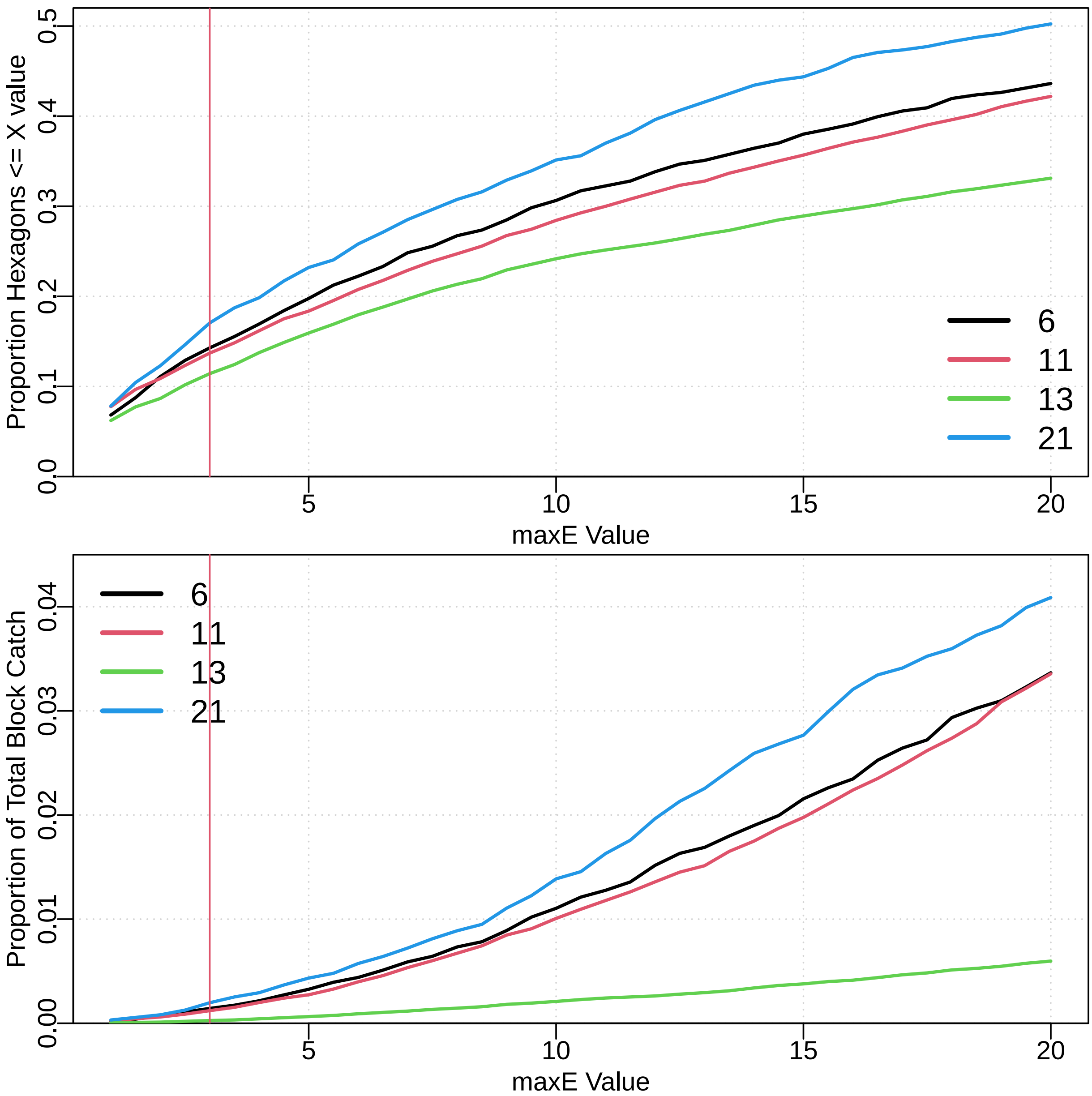


Figure 6.8: The proportion of the unique cells reported in each example block selected by different values of the maximum effort across years value (maxE). There are no obvious inflection points in any block and the vertical red line at 3 minutes is the nominally selected value of maxE that will be used leading to percentages of cells excluded of 14.3, 13.7, 11.4 and 17.1 in blocks 6, 11, 13, and 21, respectively.

6.4.0.1 Update Fishable Area

Table 6.6: Summary information for each block across year. ncells is the total number of unique cells identified in the observations for each block, area* is the number of cells selected with maximum effort levels of <= 3 minutes, and fishable Ha is the estimated fishable reef area in hectares after the false positive cells are removed.

	nobs	hours	catch t	ncells	area*	fishable Ha
6	8918	5339.1	384.836	0	0	0
11	24730	14942.6	1445.059	0	0	0
13	13615	26978.7	1509.916	0	0	0

Table 6.6: Summary information for each block across year. ncells is the total number of unique cells identified in the observations for each block, area* is the number of cells selected with maximum effort levels of <= 3 minutes, and fishable Ha is the estimated fishable reef area in hectares after the false positive cells are removed.

	nobs	hours	catch t	ncells	area*	fishable Ha
21	4877	2844.4	178.727	0	0	0

6.5 Discussion and Implications

The proportion of false positive cells is surprisingly high although most occur in the TIP = 1 by occur = 1 combination (i.e. the cell occurs in only one year, and is in the lowest quartile of catch for that year) . There were no occurrences of low maxE values in TIP values greater than 8.

Table 6.7: The count of cells in block 6 having a maxE value <= 3 in each combination of TIP and occur (see appendix for the other block results.

	1	2	3	4	5	6	7	8
1	244	0	0	0	0	0	0	0
2	0	50	0	0	0	0	0	0
3	0	0	19	0	0	0	0	0
4	0	0	0	3	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

The proportion of cells in each block with maxE values <= 3 minutes was always greater than 10%, but the proportion of catch lost in this manner was always less than 0.2% of the total, and in the case of block 13 was only 0.026%.

Table 6.8: A summary across blocks of the result of selecting on maxE <= 3 minutes.

	maxE	ncell	select	prop	catch	propcatch	totalcatch
6	3	2226	318	0.1429	548.203	0.00142	384835.6
11	3	5881	806	0.1371	1752.345	0.00121	1445059.5
13	3	2573	294	0.1143	395.136	0.00026	1509915.9
21	3	1073	183	0.1705	352.700	0.00197	178726.5

Despite the misplaced catch being such a small proportion of the total this selection against false positive cells implies that, in order to obtain accurate summary statistics relating to catch rates and related matters, the missing catches will need to be redistributed back into the cells in which most of the fishing from such events occurred. Given the relatively high number of individual cells affected in each block (**Table 6.8**), this is not a rapid or simple task. At a minimum the analyses should be restricted to those cells assumed to represent the true fishable reef area. This would entail removing those cell identified are false positives.

For example, if we consider one of the example blocks, we shall use block 13, the total number of unique cells used across the eight years is known. The number of cells visited in any single year is known (labelled 'records'). It is also possible to identify how many of the total unique cells were visited for the first time in each year (labelled 'new'). Similarly, how many of the cells visited in a year had been visited in the previous year ('prevyr'). By knowing the total unique cells it is also possible to tabulate how many of the total were not visited in a given year ('missed'). The catch and effort per year can also be added as can the proportion of the total known area that was used (**Table 6.9**).

Table 6.9: A summary of the grid cell use across years for block 6.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	1266	1266	1266	0	960	103.982	1017.728	0.569
2013	1375	400	1666	975	851	87.195	1008.472	0.618
2014	1394	250	1916	1016	832	58.959	872.200	0.626
2015	1290	126	2042	961	936	46.551	779.511	0.580
2016	1171	68	2110	845	1055	38.498	654.297	0.526
2017	1246	65	2175	847	980	32.197	667.864	0.560
2018	720	45	2220	539	1506	8.795	181.761	0.323
2019	456	6	2226	254	1770	8.658	157.281	0.205

But this outcome (**Table 6.9**) uses all known cells including those that we now assume are false positives. If those are removed by excluding all cells with maxE <= 3 minutes, rather different results are forthcoming.

Table 6.10: A summary of the grid cell use across years for block 6 after all cells with maximum effort levels <= 3 minutes are excluded. The final column 'propused' refers to the proportion of cells used in that year.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	1213	1213	1213	0	695	103.881	1016.675	0.636
2013	1308	345	1558	963	600	87.069	1006.947	0.686
2014	1301	173	1731	1009	607	58.827	870.250	0.682
2015	1228	80	1811	953	680	46.499	778.547	0.644
2016	1120	37	1848	842	788	38.441	653.358	0.587
2017	1199	41	1889	841	709	32.159	667.039	0.628
2018	678	18	1907	535	1230	8.762	181.161	0.355
2019	447	1	1908	253	1461	8.649	157.108	0.234

Comparing the two tables **Table 6.9** and **Table 6.10** the cumulative number of unique cells ('cumul') differs in 2019 by 318 (as predicted by **Table 6.8**). The 'known' area of reef has been empirically defined as those cells with intentional fishing. The effect or removing the false positive cells is to increase the proportion of the total used each year. However, even with this increase it is clear that the proportion of the known reef used in any single year can be only a fraction of the total.

In block 6, the proportion used in the last two years was unusually low, which reflected the very low catches. In the other years, however, the proportion of the known area used varied from 0.587 - 0.682 with an average of 0.643. Block 11 (see appendix) was similar varying from 0.51 - 0.658. The blocks from the east coast used a somewhat higher proportion with block 13 varying from 0.601 - 0.808, while block 21 varied between 0.542 - 0.766.

Table 6.11: The range of the total known reef area used across years for all blocks after all cells with effort levels <= 3 minutes are excluded. Block 6 should really not include the last two years as they were exceptionally low. The final column 'propused' refers to the proportion of cells used in that year.

	ReefArea	minimum	maximum	average
6	1908	0.234	0.686	0.556
11	5075	0.510	0.658	0.584
13	2279	0.601	0.808	0.726
21	890	0.542	0.766	0.646

6.6 Appendix - Analyses of Blocks not Included in Text

6.6.1 Frequency of Occurrence of Records

6.6.1.1 Block 6

Table 6.12: The frequency of occurrence of each combination of occur and TIP within block 6's dataset.

	1	2	3	4	5	6	7	8
1	465	0	0	0	0	0	0	0
2	7	526	0	0	0	0	0	0
3	3	36	630	0	0	0	0	0
4	1	14	72	632	0	0	0	0
5	0	6	27	180	590	0	0	0
6	0	0	12	48	285	510	0	0
7	0	0	6	32	125	324	245	0
8	0	0	3	8	75	264	175	120
9	0	0	0	28	40	192	196	40
10	0	0	0	12	50	132	147	80
11	0	0	0	0	50	114	175	112
12	0	0	0	0	5	96	133	104
13	0	0	0	0	15	42	182	96
14	0	0	0	0	15	42	105	136
15	0	0	0	0	10	48	91	64

Table 6.12: The frequency of occurrence of each combination of occur and TIP within block 6's dataset.

	1	2	3	4	5	6	7	8
16	0	0	0	0	5	12	105	24
17	0	0	0	0	0	12	42	64
18	0	0	0	0	0	12	91	40
19	0	0	0	0	0	6	28	56
20	0	0	0	0	0	0	21	64
21	0	0	0	0	0	6	42	56
22	0	0	0	0	0	6	14	56
23	0	0	0	0	0	0	0	32
24	0	0	0	0	0	0	0	16
25	0	0	0	0	0	0	7	64
26	0	0	0	0	0	0	0	24
27	0	0	0	0	0	0	0	8
31	0	0	0	0	0	0	0	32

Table 6.13: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 6's dataset.

	1	2	3	4	5	6	7	8
1	465	0	0	0	0	0	0	0
2	7	263	0	0	0	0	0	0
3	3	18	210	0	0	0	0	0
4	1	7	24	158	0	0	0	0
5	0	3	9	45	118	0	0	0
6	0	0	4	12	57	85	0	0
7	0	0	2	8	25	54	35	0
8	0	0	1	2	15	44	25	15
9	0	0	0	7	8	32	28	5
10	0	0	0	3	10	22	21	10
11	0	0	0	0	10	19	25	14
12	0	0	0	0	1	16	19	13
13	0	0	0	0	3	7	26	12
14	0	0	0	0	3	7	15	17
15	0	0	0	0	2	8	13	8
16	0	0	0	0	1	2	15	3
17	0	0	0	0	0	2	6	8
18	0	0	0	0	0	2	13	5
19	0	0	0	0	0	1	4	7
20	0	0	0	0	0	0	3	8
21	0	0	0	0	0	1	6	7

Table 6.13: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 6's dataset.

	1	2	3	4	5	6	7	8
22	0	0	0	0	0	1	2	7
23	0	0	0	0	0	0	0	4
24	0	0	0	0	0	0	0	2
25	0	0	0	0	0	0	1	8
26	0	0	0	0	0	0	0	3
27	0	0	0	0	0	0	0	1
28	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	4
32	0	0	0	0	0	0	0	0

6.6.1.2 Block 13

Table 6.14: The frequency of occurrence of each combination of occur and TIP within block 13's dataset.

	1	2	3	4	5	6	7	8
1	452	0	0	0	0	0	0	0
2	1	394	0	0	0	0	0	0
3	0	0	483	0	0	0	0	0
4	0	0	0	664	0	0	0	0
5	0	0	0	20	715	0	0	0
6	0	0	0	0	35	1032	0	0
7	0	0	0	0	5	84	1533	0
8	0	0	0	0	0	48	245	2224
9	0	0	0	0	0	0	105	840
10	0	0	0	0	0	0	63	584
11	0	0	0	0	0	0	28	416
12	0	0	0	0	0	0	14	392
13	0	0	0	0	0	0	14	312
14	0	0	0	0	0	0	0	232
15	0	0	0	0	0	0	0	296
16	0	0	0	0	0	0	0	224
17	0	0	0	0	0	0	0	304
18	0	0	0	0	0	0	0	176
19	0	0	0	0	0	0	0	304
20	0	0	0	0	0	0	0	120
21	0	0	0	0	0	0	0	168

Table 6.14: The frequency of occurrence of each combination of occur and TIP within block 13's dataset.

	1	2	3	4	5	6	7	8
22	0	0	0	0	0	0	0	128
23	0	0	0	0	0	0	0	64
24	0	0	0	0	0	0	0	112
25	0	0	0	0	0	0	0	112
26	0	0	0	0	0	0	0	80
27	0	0	0	0	0	0	0	136
28	0	0	0	0	0	0	0	80
29	0	0	0	0	0	0	0	112
30	0	0	0	0	0	0	0	104
31	0	0	0	0	0	0	0	48
32	0	0	0	0	0	0	0	112

Table 6.15: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 13's dataset.

	1	2	3	4	5	6	7	8
1	452	0	0	0	0	0	0	0
2	1	197	0	0	0	0	0	0
3	0	0	161	0	0	0	0	0
4	0	0	0	166	0	0	0	0
5	0	0	0	5	143	0	0	0
6	0	0	0	0	7	172	0	0
7	0	0	0	0	1	14	219	0
8	0	0	0	0	0	8	35	278
9	0	0	0	0	0	0	15	105
10	0	0	0	0	0	0	9	73
11	0	0	0	0	0	0	4	52
12	0	0	0	0	0	0	2	49
13	0	0	0	0	0	0	2	39
14	0	0	0	0	0	0	0	29
15	0	0	0	0	0	0	0	37
16	0	0	0	0	0	0	0	28
17	0	0	0	0	0	0	0	38
18	0	0	0	0	0	0	0	22
19	0	0	0	0	0	0	0	38
20	0	0	0	0	0	0	0	15
21	0	0	0	0	0	0	0	21
22	0	0	0	0	0	0	0	16
23	0	0	0	0	0	0	0	8

Table 6.15: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 13's dataset.

	1	2	3	4	5	6	7	8
24	0	0	0	0	0	0	0	14
25	0	0	0	0	0	0	0	14
26	0	0	0	0	0	0	0	10
27	0	0	0	0	0	0	0	17
28	0	0	0	0	0	0	0	10
29	0	0	0	0	0	0	0	14
30	0	0	0	0	0	0	0	13
31	0	0	0	0	0	0	0	6
32	0	0	0	0	0	0	0	14

6.6.1.3 Block 21

Table 6.16: The frequency of occurrence of each combination of occur and TIP within block 21's dataset.

	1	2	3	4	5	6	7	8
1	195	0	0	0	0	0	0	0
2	1	248	0	0	0	0	0	0
3	0	14	273	0	0	0	0	0
4	0	2	30	308	0	0	0	0
5	0	0	0	32	400	0	0	0
6	0	0	3	4	85	402	0	0
7	0	0	0	4	40	126	224	0
8	0	0	0	0	5	66	182	232
9	0	0	0	8	5	60	112	152
10	0	0	0	0	5	18	105	224
11	0	0	0	0	0	24	49	64
12	0	0	0	0	0	6	35	72
13	0	0	0	0	0	0	21	120
14	0	0	0	0	0	6	21	104
15	0	0	0	0	0	6	14	72
16	0	0	0	0	0	0	21	80
17	0	0	0	0	0	0	7	40
18	0	0	0	0	0	0	0	80
19	0	0	0	0	0	0	7	64
20	0	0	0	0	0	0	0	80
21	0	0	0	0	0	0	0	64
22	0	0	0	0	0	0	0	16
23	0	0	0	0	0	0	0	72

Table 6.16: The frequency of occurrence of each combination of occur and TIP within block 21's dataset.

	1	2	3	4	5	6	7	8
24	0	0	0	0	0	0	7	24
25	0	0	0	0	0	0	0	8
26	0	0	0	0	0	0	0	16
27	0	0	0	0	0	0	0	16
28	0	0	0	0	0	0	0	32
29	0	0	0	0	0	0	0	8
30	0	0	0	0	0	0	0	24
31	0	0	0	0	0	0	0	8
32	0	0	0	0	0	0	0	24

Table 6.17: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 21's dataset.

	1	2	3	4	5	6	7	8
1	195	0	0	0	0	0	0	0
2	1	124	0	0	0	0	0	0
3	0	7	91	0	0	0	0	0
4	0	1	10	77	0	0	0	0
5	0	0	0	8	80	0	0	0
6	0	0	1	1	17	67	0	0
7	0	0	0	1	8	21	32	0
8	0	0	0	0	1	11	26	29
9	0	0	0	2	1	10	16	19
10	0	0	0	0	1	3	15	28
11	0	0	0	0	0	4	7	8
12	0	0	0	0	0	1	5	9
13	0	0	0	0	0	0	3	15
14	0	0	0	0	0	1	3	13
15	0	0	0	0	0	1	2	9
16	0	0	0	0	0	0	3	10
17	0	0	0	0	0	0	1	5
18	0	0	0	0	0	0	0	10
19	0	0	0	0	0	0	1	8
20	0	0	0	0	0	0	0	10
21	0	0	0	0	0	0	0	8
22	0	0	0	0	0	0	0	2
23	0	0	0	0	0	0	0	9
24	0	0	0	0	0	0	1	3
25	0	0	0	0	0	0	0	1

Table 6.17: The frequency of occurrence of unique hexagons in each combination of occur and TIP within block 21's dataset.

	1	2	3	4	5	6	7	8
26	0	0	0	0	0	0	0	2
27	0	0	0	0	0	0	0	2
28	0	0	0	0	0	0	0	4
29	0	0	0	0	0	0	0	1
30	0	0	0	0	0	0	0	3
31	0	0	0	0	0	0	0	1
32	0	0	0	0	0	0	0	3

6.6.2 Catch by TIP Value

6.6.2.1 Block 11

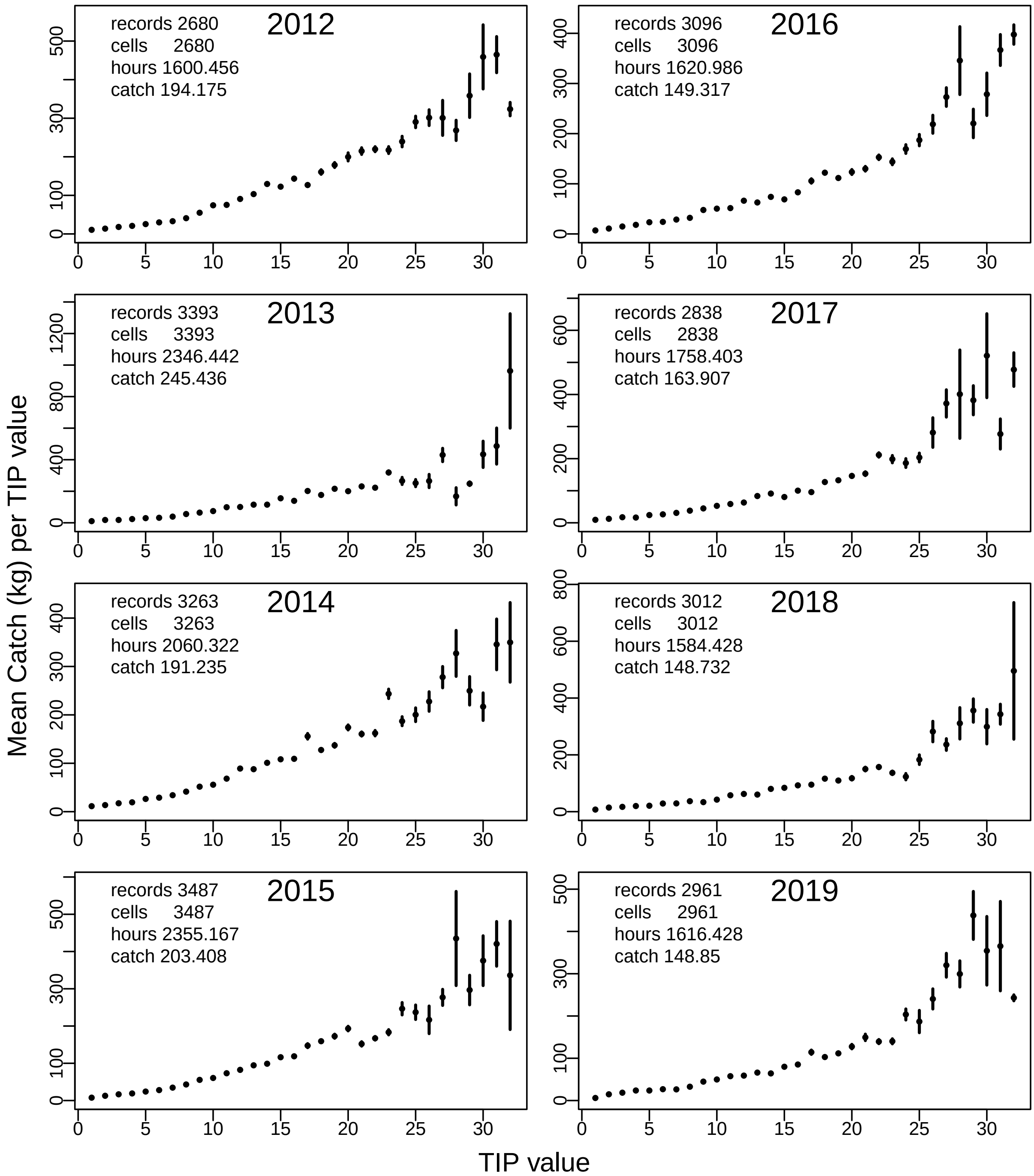


Figure 6.9: The mean catch per TIP value by year with 99% Normal CI for example statistical block 11.

6.6.2.2 Block 13

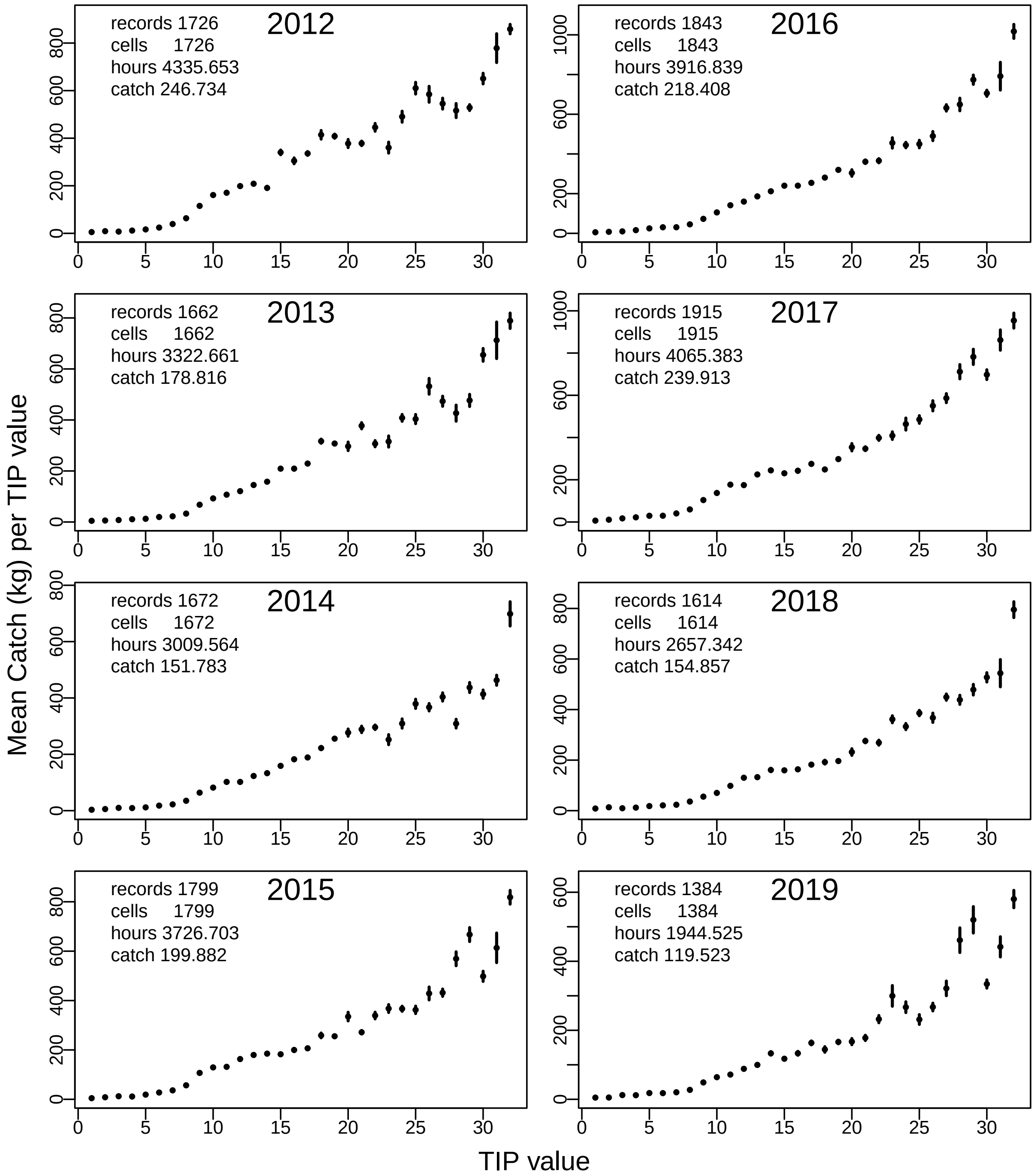


Figure 6.10: The mean catch per TIP value by year with 99% Normal CI for example statistical block 13.

6.6.2.3 Block 21

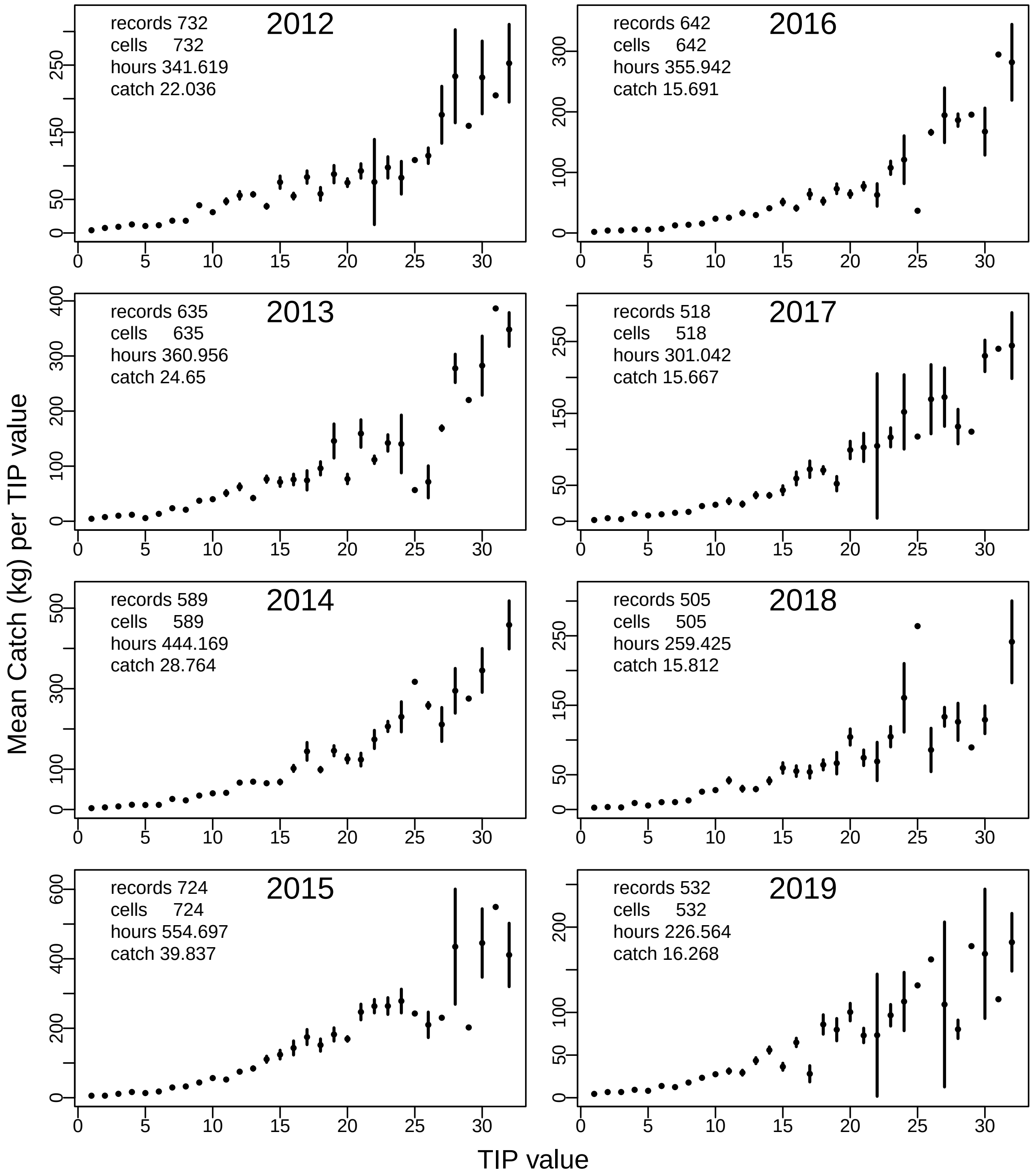


Figure 6.11: TThe mean catch per TIP value by year with 99% Normal CI for example statistical block 21.

6.6.3 Update Fishable Area

Most hexagons exhibiting maxE values <= 3 minutes occur in the occur = 1 and TIP = 1 combination. There were no occurrences of such hexagons in TIP values greater than 8.

Table 6.18: The count of hexagons in block 11 having a maxE value <= 3 in each combination of TIP and occur (see appendix for the other block results.

	1	2	3	4	5	6	7	8
1	638	0	0	0	0	0	0	0
2	0	120	0	0	0	0	0	0
3	0	0	38	0	0	0	0	0
4	0	0	0	6	0	0	0	0
5	0	0	0	0	3	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1

Table 6.19: The count of hexagons in block 13 having a maxE value <= 3 in each combination of TIP and occur (see appendix for the other block results.

	1	2	3	4	5	6	7	8
1	238	0	0	0	0	0	0	0
2	0	34	0	0	0	0	0	0
3	0	0	12	0	0	0	0	0
4	0	0	0	8	0	0	0	0
5	0	0	0	0	2	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

Table 6.20: The count of hexagons in block 21 having a maxE value <= 3 in each combination of TIP and occur (see appendix for the other block results.

	1	2	3	4	5	6	7	8
1	127	0	0	0	0	0	0	0
2	0	32	0	0	0	0	0	0
3	0	0	15	0	0	0	0	0
4	0	0	0	2	0	0	0	0
5	0	0	0	0	5	0	0	0
6	0	0	0	0	0	2	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

6.6.4 Implications

Table 6.21: A summary of the hexagon use across years for block 11.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	2680	2680	2680	0	3201	194.175	1600.456	0.456
2013	3393	1284	3964	2109	2488	245.436	2346.442	0.577
2014	3263	600	4564	2382	2618	191.235	2060.322	0.555
2015	3487	464	5028	2423	2394	203.408	2355.167	0.593
2016	3096	277	5305	2270	2785	149.317	1620.986	0.526
2017	2838	147	5452	1931	3043	163.907	1758.403	0.483
2018	3012	217	5669	1910	2869	148.732	1584.428	0.512
2019	2961	212	5881	1938	2920	148.850	1616.428	0.503

Table 6.22: A summary of the hexagon use across years for block 11 after all hexagons with effort levels <= 3 minutes are excluded.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	2587	2587	2587	0	2488	193.935	1598.536	0.510
2013	3237	1146	3733	2091	1838	245.119	2343.500	0.638
2014	3152	514	4247	2363	1923	191.003	2057.856	0.621
2015	3339	351	4598	2413	1736	203.168	2352.314	0.658
2016	2982	201	4799	2262	2093	149.114	1618.689	0.588
2017	2746	91	4890	1926	2329	163.770	1756.883	0.541
2018	2860	110	5000	1902	2215	148.530	1582.300	0.564
2019	2788	75	5075	1927	2287	148.669	1614.508	0.549

Table 6.23: A summary of the hexagon use across years for block 13.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	1726	1726	1726	0	847	246.734	4335.653	0.671
2013	1662	204	1930	1458	911	178.816	3322.661	0.646
2014	1672	163	2093	1380	901	151.783	3009.564	0.650
2015	1799	158	2251	1444	774	199.882	3726.703	0.699
2016	1843	112	2363	1526	730	218.408	3916.839	0.716
2017	1915	128	2491	1595	658	239.913	4065.383	0.744
2018	1614	69	2560	1434	959	154.857	2657.342	0.627
2019	1384	13	2573	1205	1189	119.523	1944.525	0.538

Table 6.24: A summary of the hexagon use across years for block 13 after all hexagons with effort levels <= 3 minutes are excluded.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	1681	1681	1681	0	598	246.689	4334.856	0.738
2013	1619	170	1851	1449	660	178.772	3321.847	0.710
2014	1621	123	1974	1372	658	151.731	3008.697	0.711
2015	1732	105	2079	1440	547	199.820	3725.511	0.760
2016	1798	82	2161	1521	481	218.345	3915.822	0.789
2017	1842	77	2238	1587	437	239.840	4064.136	0.808
2018	1568	36	2274	1430	711	154.812	2656.603	0.688
2019	1370	5	2279	1202	909	119.511	1944.319	0.601

Table 6.25: A summary of the hexagon use across years for block 21.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	732	732	732	0	341	22.036	341.619	0.682
2013	635	135	867	500	438	24.650	360.956	0.592
2014	589	58	925	438	484	28.764	444.169	0.549
2015	724	68	993	496	349	39.837	554.697	0.675
2016	642	27	1020	539	431	15.691	355.942	0.598
2017	518	21	1041	408	555	15.667	301.042	0.483
2018	505	20	1061	353	568	15.812	259.425	0.471
2019	532	12	1073	349	541	16.268	226.564	0.496

Table 6.26: A summary of the hexagon use across years for block 21 after all hexagons with effort levels <= 3 minutes are excluded.

	records	new	cumul	prevyr	missed	catch	hrs	propused
2012	676	676	676	0	214	21.966	340.519	0.760
2013	589	99	775	490	301	24.582	359.997	0.662
2014	557	39	814	431	333	28.715	443.453	0.626
2015	682	43	857	490	208	39.782	554.000	0.766
2016	611	11	868	533	279	15.667	355.400	0.687
2017	495	11	879	404	395	15.648	300.622	0.556
2018	482	8	887	351	408	15.788	258.972	0.542
2019	504	3	890	345	386	16.226	226.019	0.566

7 Grid turnover: temporal dynamics of reef use

7.1 Introduction

When preparing and interpreting fishery data (catch, catch rates), there is no consideration given to whether fishers utilise all or part of the known productive reef each year. This is fundamentally because very few fisheries have sufficient spatial data across the fleet to determine what proportion of the fishing grounds have been accessed in a given year. Without knowledge of what proportion of the fishers have accessed in a given year, there becomes and implicit assumption that either a) it doesn't matter, or b) most of the reef is utilised and therefore it doesn't matter.

Of specific interest were the following categories;

- *This Year* = total number number of cells fished this year
- *new* = new cells this year
- *Last year and this* = cells fished in two consecutive years (this year and the preceding year)
- *New this year* = cells fished this year, but not last year
- *Not this year, but last* = cells fished last year and not visited in the current year
- *Cumulative* = total known cells in the fishery

Objective 2: **Develop methods for inclusion of fine-scale spatial data in CPUE standardisations.** Objective 3: **Identify methods for detecting hyper-stability in CPUE.**

Objective 4: **Determine feasibility of spatial data based stock status determination in spatially structured fisheries**

7.1.1 Statewide patterns of grid cell turnover

A state-wide summary of the grid dynamics indicate that in the eight year of this time-series, areas are being utilised that haven't been utilised in the previous seven years (Figure 7.1). This approach to examining the dynamics of reef utilisation indicate that the total number of cells utilised annually declined from 2016 to 2019, which is consistent with wide-ranging TACC reductions in those years. In the final year of this time-series however, there will still almost 1000 new cells visited across the fishery (Table 7.1, Figure 7.1). By the end of the time-series, there were over 47,500 1Ha Hexagonal cells visited between 2012 and 2019.

Table 7.1: Annual turnover of grid cells across the Tasmanian abalone fishery, excluding cells with less than 5 minutes of fishing effort. Explanation of legend: This Year - total number number of cells fished this year; Last year and this - cells fished in two consecutive years (this year and preceding year); New this year - cells not previously fished; Not this year, but last - cells fished last year and not visited in the current year; Cumulative - total known cells in the fishery

	2012	2013	2014	2015	2016	2017	2018	2019
years	21937	23345	24260	24048	22603	20216	17611	15905
new	21937	8953	5608	3819	2966	2025	1457	928
both_years	0	14392	15570	15820	14857	13317	11292	10112
this_year	0	8953	8690	8228	7746	6899	6319	5793
last_year	0	7545	7775	8440	9191	9286	8924	7499
all_years	21937	30890	36498	40317	43283	45308	46765	47693

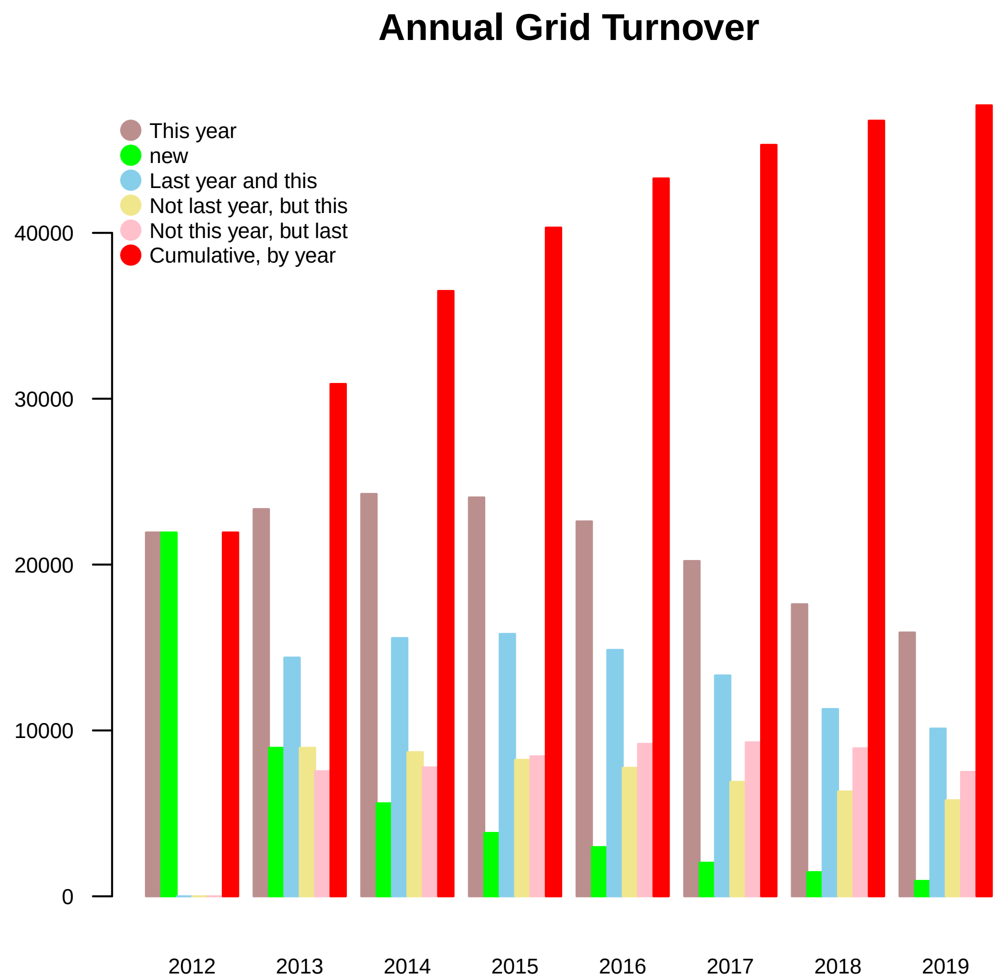


Figure 7.1: Dynamics of fishing. Annual turnover of grid cells across the Tasmanian abalone fishery, excluding cells with less than 5 minutes of fishing effort. Explanation of legend: This Year - total number number of cells fished this year; Last year and this - cells fished in two consecutive years (this year and preceding year); New this year - cells not previously fished; Not this year, but last - cells fished last year and not visted in the current year; Cumulative - total known cells in the fishery

Table 7.2: Dynamics of fishing in Block 6. Annual turnover of grid cells, excluding cells with less than 5 minutes of fishing effort. Explanation of legend: years - total number number of cells fished this year; New this year - cells not previously fished; both_years - cells fished in two consecutive years (this year and preceding year); this_year - cells fished this year and not visited in the previous year; last_year - cells fished last year and not visited in the current year; all_years - total known cells in the fishery

	2012	2013	2014	2015	2016	2017	2018	2019
years	983	1064	1040	961	831	917	418	297
new	983	353	182	95	47	58	18	5
both_years	0	711	745	689	583	575	320	139
this_year	0	353	295	272	248	342	98	158
last_year	0	272	319	351	378	256	597	279
all_years	983	1336	1518	1613	1660	1718	1736	1741

Table 7.3: Dynamics of fishing in Block 21. Annual turnover of grid cells, excluding cells with less than 5 minutes of fishing effort. Explanation of legend: years - total number number of cells fished this year; New this year - cells not previously fished; both_years - cells fished in two consecutive years (this year and preceding year); this_year - cells fished this year and not visited in the previous year; last_year - cells fished last year and not visited in the current year; all_years - total known cells in the fishery

	2012	2013	2014	2015	2016	2017	2018	2019
years	485	444	447	557	438	368	344	362
new	485	117	57	78	19	22	15	12
both_years	0	327	321	379	370	275	243	230
this_year	0	117	126	178	68	93	101	132
last_year	0	158	123	68	187	163	125	114
all_years	485	602	659	737	756	778	793	805

7.1.2 Block level dynamics of grid cell activity

At a smaller scale, the grid turnover metrics confirm that reef accessed to harvest abalone is not common among years, and, there may be as little as 50% overlap in reef utilised among consecutive years (Table 7.2, Table 7.3). In the depleting fishery (Block 6), the mean fishing grounds in common among consecutive years was 0.58 (Table 7.2; Years vs both_years), except for 2019, where there was only 0.47 reef used in common. A similar pattern was observed for the rebuilding Block 21, with the mean fishing grounds in common among consecutive years marginally higher at 0.63 (Table 7.2; Years vs both_years).

In overall terms, the proportion of the total known fishery utilised in anyone year averaged 0.47 in Block 6, and 0.53 in Block 21. In the final three years in Block 6 that fraction reduced to an average of 0.31, whereas in the rebuilding Block 21, that fraction was 0.44, and closer to the annual average.

7.2 Discussion

In the absence of spatially explicit fisheries data we are forced to ignore any underlying spatio-temporal dynamics within an SAU. The alternative is to make explicit assumptions about how we believe the fishery operates. In terms of reef utilised to achieve a harvest in a fishing year, and whether the amount of reef area changes through time, or, all fishing grounds are used every year is is a major gap in our understanding of our abalone fisheries. In this chapter we demonstrate that around two-thirds of reef utilised is common among consecutive years, with only half of the known fishing grounds utilised each year. Interestingly, we observed a decline in the average fraction of the total known reef area in the final three years of the eight year time-series. For the depleted Block 6, this is likely to reflect that fishers have targeted the few remaining productive fishing grounds in the fishery. Whereas for the rebuilding Block 21, with high catch rates and constraints on catch, the fishers will be able to obtain allocated catch from a smaller proportion of the total fishing grounds. Thus two different factors underpin a similar result - in Block the small level of reef used is a function of stock dynamics, whereas in Block 21 it is a management induced response.

Given the level of reef use in common among consecutive years is around two-thirds, we might conclude that fishers are not fishing at random each year, and experience in the previous year might encourage them to return if fishing was good, or find a new location if fishing was poor. This metric provides some insight into the fleet preference of where to fish each year, though does not account for changes in where individual divers may choose to fish across years. It is unlikely this new understanding of how fishers utilise reef areas will change existing data treatment of docket book, or even spatial data. It should however, influence how we think about recovery and depletion of fisheries, whether our management actions might be inducing trends in spatial utilisation of fishing grounds. As most of the Tasmanian abalone fishery has been either in a depleting phase or rebuilding phase over the period of this study, it is unclear whether the patterns we see are a subset or modified form of natural cycles of exploitation of fishing grounds (within and among years).

A question frequently asked in abalone fisheries is whether our fisheries are undergoing spatial contraction. This question can only be answered with geo-referenced fishery-dependent data, although is not necessarily as simple as looking at which grid cells are not being utilised. Declining TACC also trigger a reduction in total effort, with corresponding reduction in spatial area utilised, unless abundance is declining rapidly and effort and area used increases inversely to catch. Fishers also rotate through their known patches, which may create a signal in how the entire reef is used. This will be especially prevalent in SAUs which support small amounts of catch, and are visited by relatively few fishers often who have good local knowledge of that SAU. In this chapter we looked at cells that we fished last year, but not this year. This approach could be modified to examine those cells that have not been fished for a period of time after a particular year. For example, how many cells fished in 2012, were not fished again within the next five years, or indeed the entire remainder of the time series.

The level of annual turnover in fishing grounds identified here suggests there will be limited success in utilising the grid based datasets for the purpose of predicting either future yield. Previous attempts to utilise the same dataset explored in this chapter (hexagonal grid dataset) to predict future catch as a function of recent catch history was very poor for the Tasmanian Western Zone (Mundy, Jones, and Worthington 2018, chap. 8, Table 8). That finding may well be a result of the lack of a high level of overlap among consecutive years here.

Part IV

Interpretation and implications

8 Can spatial fisheries data improve assessment of stock status?

The Tasmanian high-resolution spatial fisheries data collection program is highly cost-effective given the scope and scale of the information the program delivers. Spatial fisheries data derived from GPS logger programs provide multiple benefits for determining stock status in spatially structured abalone fisheries:

1. provision of a more responsive measure of effort in the form of dive polygon area,
2. identification of discrete reef systems as an improvement on within Spatial Assessment Unit spatial structure,
3. through a temporal index of productivity (reported in an earlier milestone), and,
4. provision of capacity to quantify the extent of reef used each year, and among consecutive years.

8.1 Space is a better measure of effort than time

Area as a measure of effort in diver fisheries appears to offer greater sensitivity to stock changes and is less affected by hyperstability than time as a measure of effort. Passive GPS and Depth data logging technology capture dive time and dive area with greater precision than estimated daily dive effort by divers. Even when we have greater precision in capturing fishing effort in the form of dive time with depth data loggers, area however appears to be a better measure of effort to assess performance of the fishery, and therefore stock status. In essence, using both Area and Time as measure of effort, we are able to see the cross-over from when search time dominates the dive to when handling time is the limiting factor and search time is minimal.

The findings that catch rates based on diving time are subject to hyperstability are not evidence that CPUE (effort as *Time*) is a inaccurate measure of fishery performance as many would argue. This project merely confirms hyperstability is present at both high and low stock levels. Where *Time* is the only measure of effort available, Harvest Strategies can still effectively utilise CPUE as a performance measure, with simple adjustments such that as CPUE deviates from historical median catch rates, outcomes should be increasingly conservative, and when CPUE approaches historic high levels, outcomes can be more optimistic. However, it appears that the level of hyperstability is not entirely symmetrical, thus caution is advised when adjusting HS outcomes at the very high end of the scale.

8.2 Explicit inclusion of dive location in catch rate standardisation

Abalone fisheries are well known for being highly variable, both in terms of fishery dynamics, biology and productivity, at local scales. Largely this comes from 'Fisher Ecological Knowledge' or 'Local Knowledge' (Anuchiracheeva et al. 2003; Hall and Close 2007; Hall et al. 2009; Nguyen et al. 2019), and has rarely been quantified through a data collection and analysis process. In Chapter 3 we illustrated that daily catch rates, can be highly variable, either as mean daily catch rates, or smoothed with a moving window average of 5, 20, or 30 days. As most catch rate standardisations account for season and fisher identity, one remaining predictor that had not been previously examined was the dive location as a primary driver of high variation in mean daily catch rate within a reporting area. Inclusion of spatial location as a term in the catch rate standardisations however, had negligible impact on predicted mean annual catch rate. Comparisons of models with and without spatial terms are rather difficult to compare directly through common methods such as R^2 or AIC. The tools for spatial mixed effect models are still under rapid development, as are the diagnostic procedures and methods for comparison with non-spatial models. This project has touched on these methods, and is an area for future activity, although largely that will be to determine the extent of the *noise* captured by the model, but also in terms of spatial predictions to visualise usage of reef systems over times.

As abalone stocks decline, and eventually rebuild, fleet movement and individual diver choices on where to fish will signal important information in the depletion and recovery process. With hard copy logbook reporting, we are unable to determine whether a stock declines or stock recovery is homogeneous across an SAU, or whether there is significant spatial structure in changes that occur. This is particularly important where once important fishing grounds fail to recover, or recover more slowly. Traditionally information on this question is provided by fishers, although fishers rarely fish everywhere and may not always be aware of all activity in a particular area, and is this prone to gaps in knowledge (although that may no always be acknowledged). The GPS data logger program provides critical detail on where and to what extent components of a fishery change, and the rate at which they change. The investigation of fleet dynamics is an important new direction to better inform managers on the spatial structure of recovery (O'Farrell et al. 2017).

8.3 Feasibility of spatial data for stock status

Tasmania now has a 12 year time-series of GPS and Depth logger data, covering on average around ~90% of fishing activity in the Tasmanian abalone fishery. For states with a shorter or fractured time-series, area-based metrics can still provide valuable insights on the trend of fishery performance by examining relative change from the time-series start-year, as deviations from time-based trends are evident after three to four years. data investigations (not included in this report) also suggests that the dive polygon based spatia performance measures can be informative at lower levels of fishery coverage. While an adequate time-series is being acquired from which to directly determine reference points for area based catch rates, an interim measure could be to fit a model between time and area, and predicted area based TRPs and LRPs from the more extensive time based catch rate data set.

Scales at which abalone fisheries are managed, are necessarily larger than the scale at which the fishery might be assessed. Total Allowable Commercial Catch (TACC) is often set at larger geographical scales such as fishery Zones, with effort management undertaken at within-zone scale in Spatial Assessment Units (Spatial

Management Units, Fishery Areas). At still smaller scales, catch (but not necessarily catch rates) can vary widely. Of concern during a declining phase of a fishery is that smaller sections of coastline are abandoned, and there may be delays before those areas are re-integrated to the fishery (e.g. Figure 3.13). Thus catch rates at an SAU level may only represent the more productive and resilient sections of the coastline, and without access to the spatial detail of the fishery provided by GPS logger programs, the assumption will (incorrectly) be that the entire SAU is displaying the same trend. Most abalone fishery assessment processes as well as the Status of Australian Fish Stocks (SAFS) program are victim to the complications of Dynamic Pool assumptions Mundy, Haddon, and McAllister (2023). Many of the Dynamic Pool assumptions required by integrated assessment models are equally applicable to the calculation of mean annual catch rate, and, their interpretation. Regardless of whether catch rates utilise area or time as a measure of effort, catch rates are an index of local abundance at the time of fishing. Further, they are an index of what was taken and not what remains, and not representative of the entire stock. Access to spatial fisheries data partially solves this dilemma by providing the ability to examine fishing effort at relevant scales, either by inclusion of within SAU terms in catch rate standardisations (Chapter 3), or by examining the dynamic of fishing activity in the 1 Hectare hexagon grid layers (Chapter 5, Chapter 6, Chapter 7).

8.4 Utility of logbook catch rate data for stock status

Docket or Logbook based catch and effort data remain valuable to the determination of stock status of abalone fisheries. This is with the caveat that fishing time reported in mandatory Docket book systems is reported accurately, as found for the Tasmanian abalone fishery (Chapter 1). Docket book data is free, and there is a long time-series of data available for all abalone fisheries. Providing that harvest strategies include appropriate precautionary settings when relying on Docket book catch rate data, there is no reason to stop relying on this valuable data source. The Tasmanian abalone Empirical Harvest Strategy achieves this by using an asymmetric control rule, which is more conservative as a fishery declines below the catch rate Target Reference Point compared to a rebuilding fishery which is above the TRP.

8.5 Further work

The investigation of spatial mixed effect models in this study was limited and focused on the impact of including location on the annual mean catch rate. Further work is required to better understand the mechanics of these random spatial terms, and how best to utilise these tools to examine the spatio-temporal dynamic in our dive fishery. More work is required in this space to better understand to what extent mixed effect spatial linear models are capturing the variability in man daily catch rates (Chapter 3), and, whether they assist by providing greater confidence in the mean estimates provided.

Documenting and quantifying fleet dynamics remains a relatively untouched challenge. Tasmania Divers are paid a dive rate per Kilogram of abalone caught, and not by time. Unless other economic or weather factors come into play, there is a strong incentive for them to move on from areas where catch achieved by hour of effort declines below the individuals threshold. This may manifest in divers moving multiple times during a day, or, not returning to that fishing ground in the near future. Its this latter behaviour that will provide much greater insight to the status of stocks at appropriate spatial scales, and at scales smaller than SAUs.

8.6 Recommendations

1. Catch per Unit Area (CPUA) to be included as part of the catch rate analysis in the Tasmanian Fishery Assessment.

Extension and Adoption

It is expected that there will be general interest in the approaches used in this study, and the findings in each of the sections. The capacity for other groups or fisheries to utilise the work presented here will be limited to those fisheries a time-series of geo-referenced fishery-dependent data. Some of the methods for exploring spatio-temporal dynamics may be of use to fisheries with access to VMS data, where fishing activity is summarised with a similar grid over the known fishing grounds.

Extension

Outcomes of this project have been presented in multiple research, management and industry forums. The work on hyperstable catch rates and the effect of environmental variables (temperature and wave power) was presented at the 11th International Abalone Symposium in Auckland March 2023. The concept of switching from a time-based effort (CPUE) to area-based effort (CPUA) for calculating catch rates has been presented to Tasmanian port meetings, and feedback sought on whether CPUA better reflects their experiences in parts of the fishery they know well. The importance of scale work was presented at the recent Australasian Abalone mini-Research Symposium held in Hobart September 2024.

The new approach to standardisation and the inclusion of Catch per Unit Area (CPUA) in the Harvest Strategy has been added to the drat Revised Tasmanian Abalone harvest Strategy. The revised Harvest Strategy document is currently undergoing a mandatory public consultation process. Code developed for the hyperstability component has been provided to SARDI and used to examine hyperstability in some of there jurisdictions that have access to geo-referenced fishery dependent data.

Adoption

Inclusion of spatial information in fishery assessments

The use of CPUA as a index of abundance has been proposed for inclusion as part of the standard metrics for assessment in the draft Revised Abalone Harvest Strategy. It is expected that this change will be approved following the mandatory consultation process currently under way.

Strategy for catch and effort standardisation

Within this project we explored generic approaches for statistical standardisation of catch and effort data, regardless of whether the data origins where electronic from data logger equipment or fisher reported docket-book returns. This generic model has also been proposed as a replacement for the current standardisation strategy in the draft Revised Harvest Strategy document. As for the inclusion of spatial metrics, it is expected this will also be approved following the consultation process.

Part V

Appendix

This appendix contains a number of sections that are typically included in the FRDC Final Report Standards. There are also several additional pieces of exploratory work that have been included here, that contain work that was done as part of this project but are only indirect (Chapter 9 provides an examination of the relationship between FIS and overlapping dive polygon catch rates), or were part of an extensive investigation into potential standardisation models and potential terms for inclusion (sec-chapter14, sec-chapter15, and sec-chapter16). In particular whether environmental variables contribute meaningfully to explaining variance in the catch rate data. While these later chapters utilised Docket return catch rate data, the exercise was hugely informative for the progression of the project and have considerable value to assist readers understanding of how we arrived where we did.

Researchers

The following researchers contributed to the work documented in this report

- Dr Craig Mundy
- Dr Bill Venables
- Adj Professor Cathy Dichmont
- Adj Professor Malcolm Haddon

Intellectual Property

No commercial valuable Intellectual Property was developed during this project.

References

Abraham, Edward R, and Philipp Neubauer. 2015. "Relationship Between Small-Scale Catch-Per-Unit-Effort and Abundance in New Zealand Abalone (pāua, *Haliotis Iris*) Fisheries." *PeerJ PrePrints* 3 (September): e1388v2. https://doi.org/10.7287/peerj.preprints.1388v2.

Allaire, JJ. 2023. "Quarto: R Interface to 'Quarto' Markdown Publishing System." https://CRAN.R-project.org/package=quarto.

Anderson, C. Sean, Ward, J. Eric, English, A. Philina, Barnett, and A. K. Lewis. 2022. "sdmTMB: An r Package for Fast, Flexible, and User-Friendly Generalized Linear Mixed Effects Models with Spatial and Spatiotemporal Random Fields." *bioRxiv* 2022.03.24.485545. https://doi.org/10.1101/2022.03.24.485545.

Anuchiracheeva, S., H. Demaine, G. P. Shivakoti, and K. Ruddle. 2003. "Systematizing Local Knowledge Using GIS: Fisheries Management in Bang Saphan Bay, Thailand." *Ocean & Coastal Management* 46: 10491068.

Barnett, Lewis A. K., Eric J. Ward, and Sean C. Anderson. 2021. "Improving Estimates of Species Distribution Change by Incorporating Local Trends." *Ecography* 44 (3): 427–39. https://doi.org/10.1111/ecog.05176.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. https://doi.org/10.18637/jss.v067.i01.

Bolker, Benjamin M., Mollie E. Brooks, Connie J. Clark, Shane W. Geange, John R. Poulsen, M. Henry H. Stevens, and Jada-Simone S. White. 2009. "Generalized Linear Mixed Models: A Practical Guide for Ecology and Evolution." *Trends in Ecology & Evolution* 24 (3): 127–35. https://doi.org/10.1016/j.tree.2008.10.008.

Brooks, Mollie E., Kasper Kristensen, Koen J. van, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler, and Benjamin M. Bolker. 2017. "glmmTMB Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling" 9. https://doi.org/10.32614/RJ-2017-066.

Brown, Violet A. 2021. "An Introduction to Linear Mixed-Effects Modeling in R." *Advances in Methods and Practices in Psychological Science* 4 (1): 251524592096035. https://doi.org/10.1177/2515245920960351.

Cardinale, Massimiliano, Duto Nugroho, and Patrik Jonson. 2011. "Serial Depletion of Fishing Grounds in an Unregulated, Open Access Fishery." *Fisheries Research* 108 (1): 106–11. https://doi.org/https://doi.org/10.1016/j.fishres.2010.12.007.

Clark, Michael. 2023. "Mixedup: Miscellaneous Functions for Mixed Models." https://m-clark.github.io/mixedup.

Clark, Tom S., and Drew A. Linzer. 2015. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods* 3 (2): 399–408. https://doi.org/10.1017/psrm.2014.32.

Curran-Everett, Douglas. 2013. "Explorations in Statistics: The Analysis of Ratios and Normalized Data." *Advances in Physiology Education* 37 (3): 213–19. https://doi.org/10.1152/advan.00053.2013.

Evans, Rhian, Philina A. English, Sean C. Anderson, Stéphane Gauthier, and Clifford L. K. Robinson. 2021. "Factors Affecting the Seasonal Distribution and Biomass of E. Pacifica and T. Spinifera Along the Pacific Coast of Canada: A Spatiotemporal Modelling Approach." *PLOS ONE* 16 (5): e0249818. https://doi.org/10.1371/journal.pone.0249818.

Gaertner, Daniel, and Michel Dreyfus-Leon. 2004. "Analysis of Non-Linear Relationships Between Catch Per Unit Effort and Abundance in a Tuna Purse-Seine Fishery Simulated with Artificial Neural Networks." *ICES Journal of Marine Science* 61 (5): 812–20. https://doi.org/10.1016/j.icesjms.2004.05.002.

Gohel, David, and Panagiotis Skintzos. 2023. "Flextable: Functions for Tabular Reporting." https://CRAN.R-project.org/package=flextable.

Hahsler, Michael, and Matthew Piekenbrock. 2022. "Dbscan: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Related Algorithms." https://CRAN.R-project.org/package=dbscan.

Hall, G. B., and C. H. Close. 2007. "Local Knowledge Assessment for a Small-Scale Fishery Using Geographic Information Systems." *Fisheries Research* 83: 11–22.

Hall, G. B., A. M. Moore, P. Knight, and N. Hankey. 2009. "The Extraction and Utilization of Local and Scientific Geospatial Knowledge Within the Bluff Oyster Fishery, New Zealand." *Journal of Environmental Management* 90: 20552070.

Hamilton, Richard J., Glenn R. Almany, Don Stevens, Michael Bode, John Pita, Nate A. Peterson, and J. Howard Choat. 2016. "Hyperstability Masks Declines in Bumphead Parrotfish (Bolbometopon Muricatum) Populations." *Coral Reefs* 35 (3): 751763. https://doi.org/10.1007/s00338-016-1441-0.

Harley, Shelton J, Ransom A Myers, and Alistair Dunn. 2001. "Is Catch-Per-Unit-Effort Proportional to Abundance?" *Canadian Journal of Fisheries and Aquatic Sciences* 58 (9): 1760–72. https://doi.org/10.1139/f01-112.

Hodges, James S., and Brian J. Reich. 2010. "Adding Spatially-Correlated Errors Can Mess up the Fixed Effect You Love." *The American Statistician* 64 (4): 325–34. https://doi.org/10.1198/tast.2010.10052.

Hsu, Jhen, Yi-Jay Chang, and Nicholas D. Ducharme-Barth. 2022. "Evaluation of the Influence of Spatial Treatments on Catch-Per-Unit-Effort Standardization: A Fishery Application and Simulation Study of Pacific Saury in the Northwestern Pacific Ocean." *Fisheries Research* 255 (November): 106440. https://doi.org/10.1016/j.fishres.2022.106440.

Karpov, K. A., P. L. Haaker, I. K. Taniguchi, and L. Rogers-Bennett. 2000. "Serial Depletion and the Collapse of the California Abalone (Haliotis Spp.) Fishery." *Can. Spec. Publ. Fish. Aquat. Sci.* 130: 1124.

Kotwicki, Stan, James N. Ianelli, and André E. Punt. 2014. "Correcting Density-Dependent Effects in Abundance Estimates from Bottom-Trawl Surveys." *ICES Journal of Marine Science* 71 (5): 1107–16. https://doi.org/10.1093/icesjms/fst208.

Krainski, Elias Teixeira, Finn Lindgren, and Haavard Rue. 2023. "INLAspacetime: Spatial and Spatio-Temporal Models Using 'INLA'." https://CRAN.R-project.org/package=INLAspacetime.

Kristensen, Kasper, Anders Nielsen, Casper W. Berg, Hans Skaug, and Bradley M. Bell. 2016. "TMB: Automatic Differentiation and Laplace Approximation" 70. https://doi.org/10.18637/jss.v070.i05.

Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner, and Dominique Makowski. 2021. "Performance: An r Package for Assessment, Comparison and Testing of Statistical Models" 6: 3139. https://doi.org/10.21105/joss.03139.

Maunder, Mark N., and AndrÃÂ© E. Punt. 2004. "Standardizing Catch and Effort Data: A Review of Recent Approaches." *Fisheries Research* 70 (2Ã¢â‚¬â€œ3): 141 159. https://doi.org/http://dx.doi.org/10.1016/j.fishres.2004.08.002.

Miller, K. J., B. T. Maynard, and C. N. Mundy. 2009. "Genetic Diversity and Gene Flow in Collapsed and Healthy Abalone Fisheries." *Molecular Ecology* 18 (2): 200211. https://doi.org/10.1111/j.1365-294X.2008.04019.x.

Mundy, C. 2012. "Using GPS Technology to Improve Fishery Dependent Data Collection in Abalone Fisheries."

Mundy, C., M. Haddon, and J. McAllister. 2023. "Chapter 9 - Assessment of Commercial Abalone Fisheries." In, edited by Peter A. Cook and Sandra E. Shumway, 42:291–330. Abalone: Biology, Ecology, Aquaculture and Fisheries. Elsevier. https://doi.org/10.1016/B978-0-12-814938-6.00009-9.

Mundy, C., H. Jones, and D. Worthington. 2018. "Implementing a Spatial Assessment and Decision Process to Improve Fishery Management Outcomes Using Geo-Referenced Diver Data."

Mundy, C., K. J. Miller, and L. Lyall. 2010. "Factors Limiting the Resilience and Recovery of Fishing Abalone Populations."

Mundy, C., S. J. Pyke, J. McAllister, and H. Jones. 2018. "Development of a Juvenile Abalone Monitoring Method."

Nguyen, Vivian M., Nathan Young, Marianne Corriveau, Scott G. Hinch, and Steven J. Cooke. 2019. "What Is "Usable" Knowledge? Perceived Barriers for Integrating New Knowledge into Management of an Iconic Canadian Fishery." *Canadian Journal of Fisheries and Aquatic Sciences* 76 (3): 463–74. https://doi.org/10.1139/cjfas-2017-0305.

O'Farrell, Shay, James N. Sanchirico, Iliana Chollett, Marcy Cockrell, Steven A. Murawski, Jordan T. Watson, Alan Haynie, Andrew Strelcheck, and Larry Perruso. 2017. "Improving Detection of Short-Duration Fishing Behaviour in Vessel Tracks by Feature Engineering of Training Data." *ICES Journal of Marine Science* 74 (5): 1428–36. https://doi.org/10.1093/icesjms/fsw244.

Oliver, Eric C. J., Jessica A. Benthuysen, Nathaniel L. Bindoff, Alistair J. Hobday, Neil J. Holbrook, Craig N. Mundy, and Sarah E. Perkins-Kirkpatrick. 2017. "The Unprecedented 2015/16 Tasman Sea Marine Heatwave." *Nature Communications* 8 (July): 112. http://dx.doi.org/10.1038/ncomms16101.

Orensanz, J. M., A. M. Parma, G. Jerez, N. Barahona, M. Montecinos, and I. Elias. 2005. "What Are the Key Elements for the Sustainability of "s-Fisheries"? Insights from South America." *Bulletin of Marine Science* 76 (2): 527556. <Go to ISI>://000228346100017.

Parma, Ana M., J. M. (Lobo) Orensanz, Inés Elías, and Gabriel Jerez. 2003. "Diving for Shellfish and Data: Incentives for the Participation of Fishers in the Monitoring and Management of Artisanal Fisheries Around Southern South America." In, edited by Stephen J. Newman, Daniel J. Gaughan, Gary Jackson, Michael C. Mackie, Brett Molony, Jill St John, and Patricia Kailola.

Pebesma, Edzer, and Roger Bivand. 2023. "{Spatial Data Science: With Applications in r}." https://doi.org/10.1201/9780429459016.

Pedersen, Thomas Lin. 2023. "Patchwork: The Composer of Plots." https://CRAN.R-project.org/package=patchwork.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rose, G. A., and D. W. Kulka. 1999. "Hyperaggregation of Fish and Fisheries: How Catch-Per-Unit-Effort Increased as the Northern Cod (Gadus Morhua) Declined." *Canadian Journal of Fisheries and Aquatic Sciences* 56 (S1): 118–27. https://doi.org/10.1139/f99-207.

Rue, Haavard, Sara Martino, and Nicholas Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with Discussion)." 71.

Seijo, J. C., E. P. Pérez, and J. F. Caddy. 2004. "A Simple Approach for Dealing with Dynamics and Uncertainty in Fisheries with Heterogeneous Resource and Effort Distribution." *Mar. Freshwater Res.* 55 (3): 249256. https://doi.org/10.1071/MF04040.

Silk, Matthew J., Xavier A. Harrison, and David J. Hodgson. 2020. "Perils and Pitfalls of Mixed-Effects Regression Models in Biology." *PeerJ* 8 (August): e9522. https://doi.org/10.7717/peerj.9522.

Tanaka, K. R., A. L. Schmidt, T. L. Kindinger, J. L. Whitney, and J. C. Samson. 2022a. "Spatiotemporal

Assessment of Aprion Virescens Density in Shallow Main Hawaiian Islands Waters, 2010–2019 | NOAA Fisheries." https://www.fisheries.noaa.gov/resource/document/spatiotemporal-assessment-aprion-virescens-density-shallow-main-hawaiian-islands.

———. 2022b. "Spatiotemporal Assessment of Aprion Virescens Density in Shallow Main Hawaiian Islands Waters, 2010–2019 | NOAA Fisheries." https://www.fisheries.noaa.gov/resource/document/spatiotemporal-assessment-aprion-virescens-density-shallow-main-hawaiian-islands.

Thompson, Patrick L., Sean C. Anderson, Jessica Nephin, Carolyn K. Robb, Beatrice Proudfoot, Ashley E. Park, Dana R. Haggarty, and Emily M. Rubidge. 2023. "Integrating Trawl and Longline Surveys Across British Columbia Improves Groundfish Distribution Predictions." *Canadian Journal of Fisheries and Aquatic Sciences* 80 (1): 195–210. https://doi.org/10.1139/cjfas-2022-0108.

Thorson, James T., Cheryl L. Barnes, Sarah T. Friedman, Janelle L. Morano, and Margaret C. Siple. 2023. "Spatially Varying Coefficients Can Improve Parsimony and Descriptive Power for Species Distribution Models." *Ecography* 2023 (5): e06510. https://doi.org/10.1111/ecog.06510.

Ward, Eric J, Lewis A K Barnett, Sean C Anderson, Christian J C Commander, and Timothy E Essington. 2022. "Incorporating Non-Stationary Spatial Variability into Dynamic Species Distribution Models." *ICES Journal of Marine Science* 79 (9): 2422–29. https://doi.org/10.1093/icesjms/fsac179.

Ward, Hillary G. M., Paul J. Askey, and John R. Post. 2013. "A Mechanistic Understanding of Hyperstability in Catch Per Unit Effort and Density-Dependent Catchability in a Multistock Recreational Fishery." *Canadian Journal of Fisheries and Aquatic Sciences* 70 (10): 1542–50. https://doi.org/10.1139/cjfas-2013-0264.

Wickham, Hadley. 2016. "Ggplot2: Elegant Graphics for Data Analysis." https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse" 4: 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Lionel Henry. 2023. "Purrr: Functional Programming Tools." https://CRAN.R-project.org/package=purrr.

Wood, S. N. 2011. "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models" 73: 3–36.

Zhou, Shijie, Robert A. Campbell, and Simon D. Hoyle. 2019. "Catch Per Unit Effort Standardization Using Spatio-Temporal Models for Australia's Eastern Tuna and Billfish Fishery." *ICES Journal of Marine Science* 76 (6): 1489–1504. https://doi.org/10.1093/icesjms/fsz034.

9 FIS vs CPUE

9.1 Introduction

Fishery Independent Surveys (FIS) of abalone abundance conducted by researchers are assumed to be the gold standard for assessing biomass. Rarely though is there a direct comparison of FIS abundance and CPUE, where fishing activity overlaps with survey sites. Tasmania has run a small pilot study over several years, undertaking a transect based FIS program at a select number of sites across the East coast. While the small number of sites are not adequate for estimation of stock biomass, or as general indicator of abundance, Without going into details around what FIS results might mean (indicator of biomass, Dynamic Pool assumptions, representativeness etc.), data from the pilot study and our spatial data collection program can form the basis of an investigation on the relationship between FIS and CPUE.

9.2 Methods

The Tasmanian FIS data set is comprised of randomised belt transects within fixed sites (LEG), and, randomly placed fixed Abalone Recruitment Modules (ARM) at the same sites.

- ARM = data on cryptic juvenile abalone abundance.

- LEG = data on emergent abalone abundance visible to divers.

The belt transects followed the design of Mundy, Miller, and Lyall (2010), with a fixed site of 60m x 30m and 10 replicated 15m x 1m belt transects, randomised along and either side of a central baseline. Counts of abalone are converted to a mean density (abalone/m2) for each survey at each site. A description of the ARM design and data are detailed in Mundy et al. (2018). Counts of juvenile abalone recorded beneath ARMs are converted to a mean density (abalone/m2) for each survey at each site. Spatial coordinates (Eastings and Northings, GDA94) obtained using handheld GPS, are included in the dataset.

9.2.1 Spatial overlay of FIS and fishing events: Dive polygons

Individual Fishing events that overlapped any FIS site are identified using a simple spatial join. KUD polygons are extracted where the polygon contains or intersects the centroid of the FIS site. Once all overlapping KUDs are found, the time difference between the KUD event and the survey events is determined, and then a matching pair with the smallest time difference interval is identified for use in the analyses, and run for both the FIS-Transect data and the FIS-ARM data.

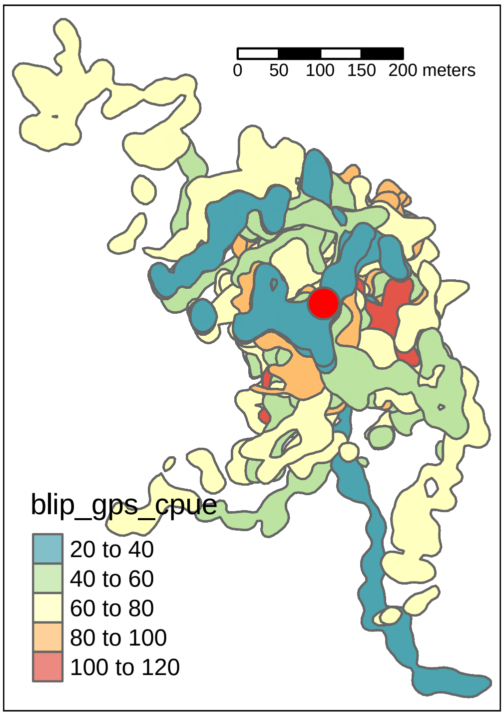


Figure 9.1: Dive polygons that overlaps the Black Reef FIS site. The red dot indicates the center of the FIS site.

The majority of matched fishing events and transect surveys are within 100 days (Figure 9.2), although there are some matching pairs where the time difference exceeds 12 months. When the time difference between spatially overlapping events is this large, we should be cautious about assuming the relationship is current. Here we restrict the matching pairs by setting a threshold time difference of ± 60 days of the transect survey event. The above process is repeated for the ARM dataset for abalone juvenile abundance.

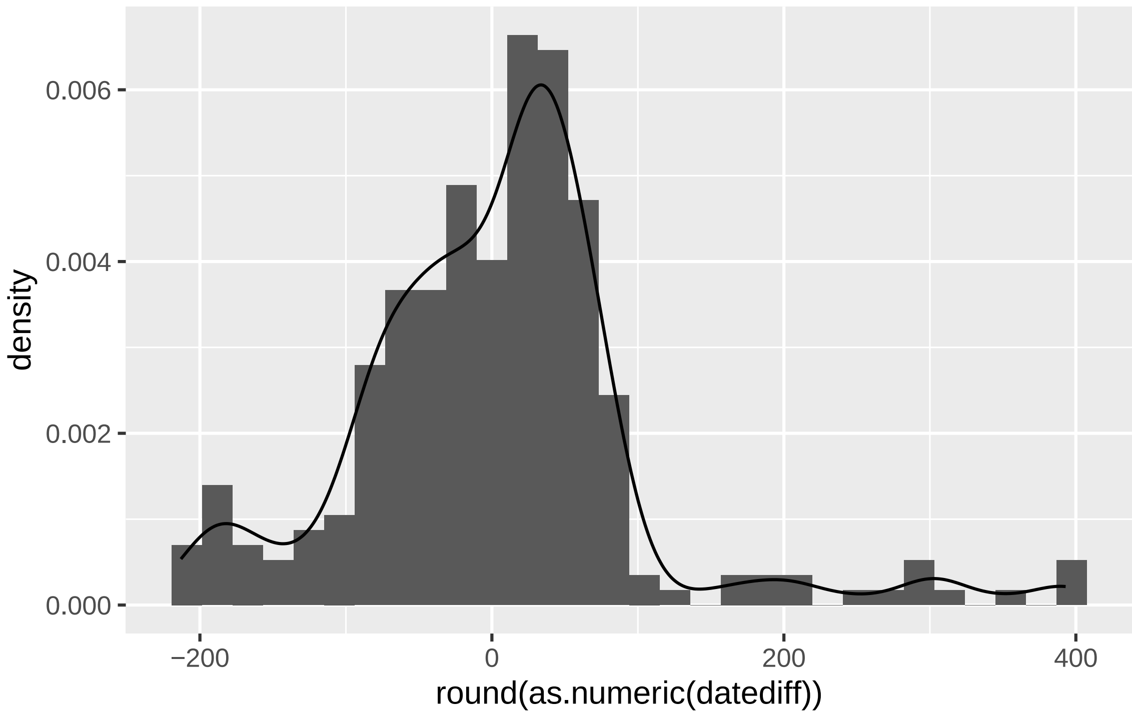


Figure 9.2: Frequency distribution of time-difference between fishing events and survey events.

9.2.2 Spatial overlay of FIS and fishing: Hexagonal grid polygons

The alternative to examining the relationship between individual fishing events and research surveys is to explore the relationship between survey density and annual harvest per one hectare grid cells (Figure 9.3). We apply the same approach used for the KUD polygons, but with a subtle twist. As part of the point-in-polygon analysis that generates the grid data sets, we determine the date of first and last fishing events in each cell, for each year. If we calculate the difference between the research survey date and the date of first fishing in each grid cell that overlays the survey site, we can make sure that we assign the grid cell from the current year of fishing to the corresponding research density.

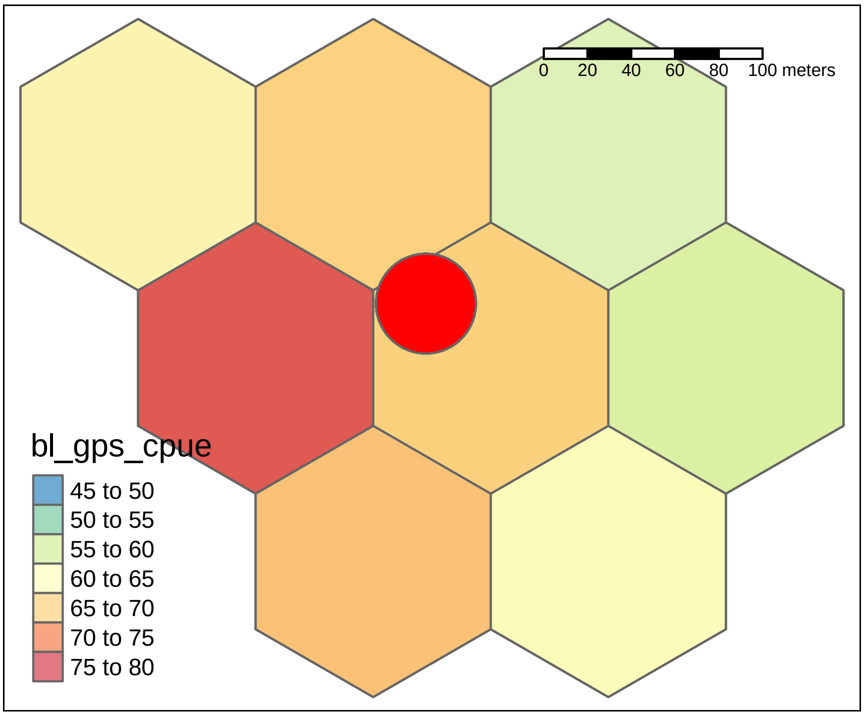


Figure 9.3: Grid cells within 50m of the Black Reef FIS site. The red dot indicates the center of the FIS site. Grid cells are colour coded by cpue. Note: only the directly overlapping grid cell is retained for analyses.

9.3 Results

9.3.1 FIS-Transect

9.3.1.1 Relationship between FIS and KUD fishing events

The relationship between CPUE and FIS abundance within a window of ± 90 days appears to be linear, and positive (Figure 9.4), although relatively weak with an R^2 of 0.2. This roughly positive relationship is encouraging in that the two indices of biomass are not contradictory. Shrinking the window of time difference to ± 30 days, has a negligible effect on the model fit, with R^2 remaining at approximately 0.2 (Figure 9.5 A). Log-transforming the data achieved a minor improvement in R^2 of 0.26 (Figure 9.5 B).

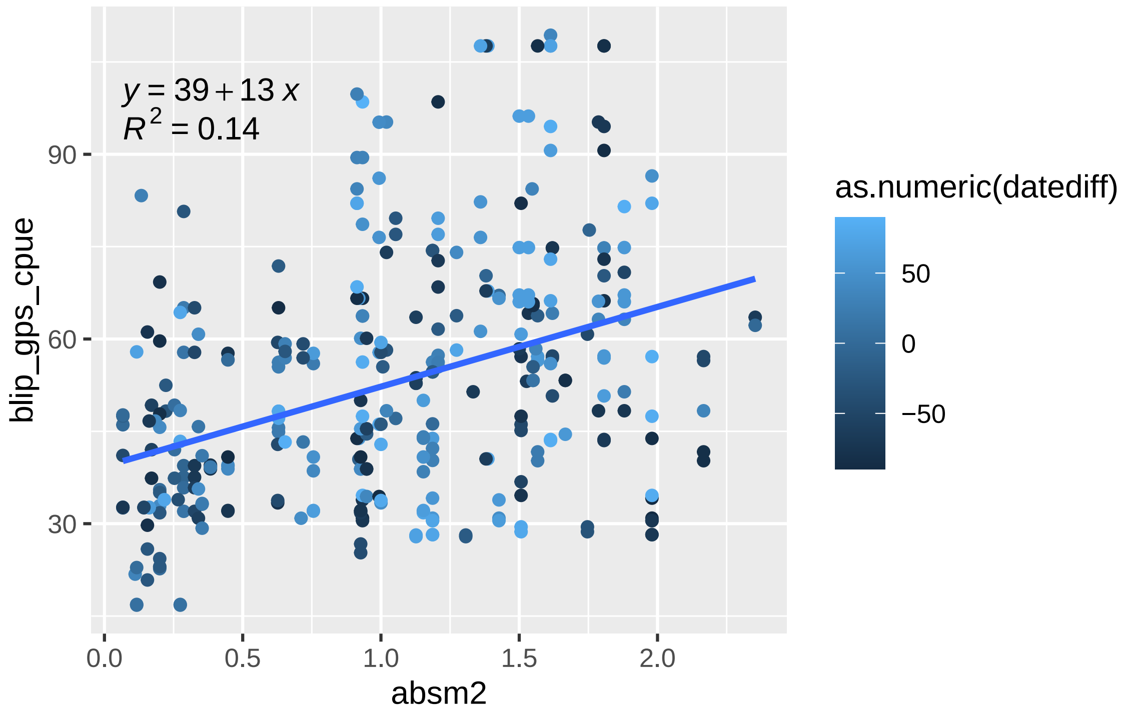


Figure 9.4: Relationship between FIS abalone density and CPUE, where fishing and survey events fall within a 90 day window.

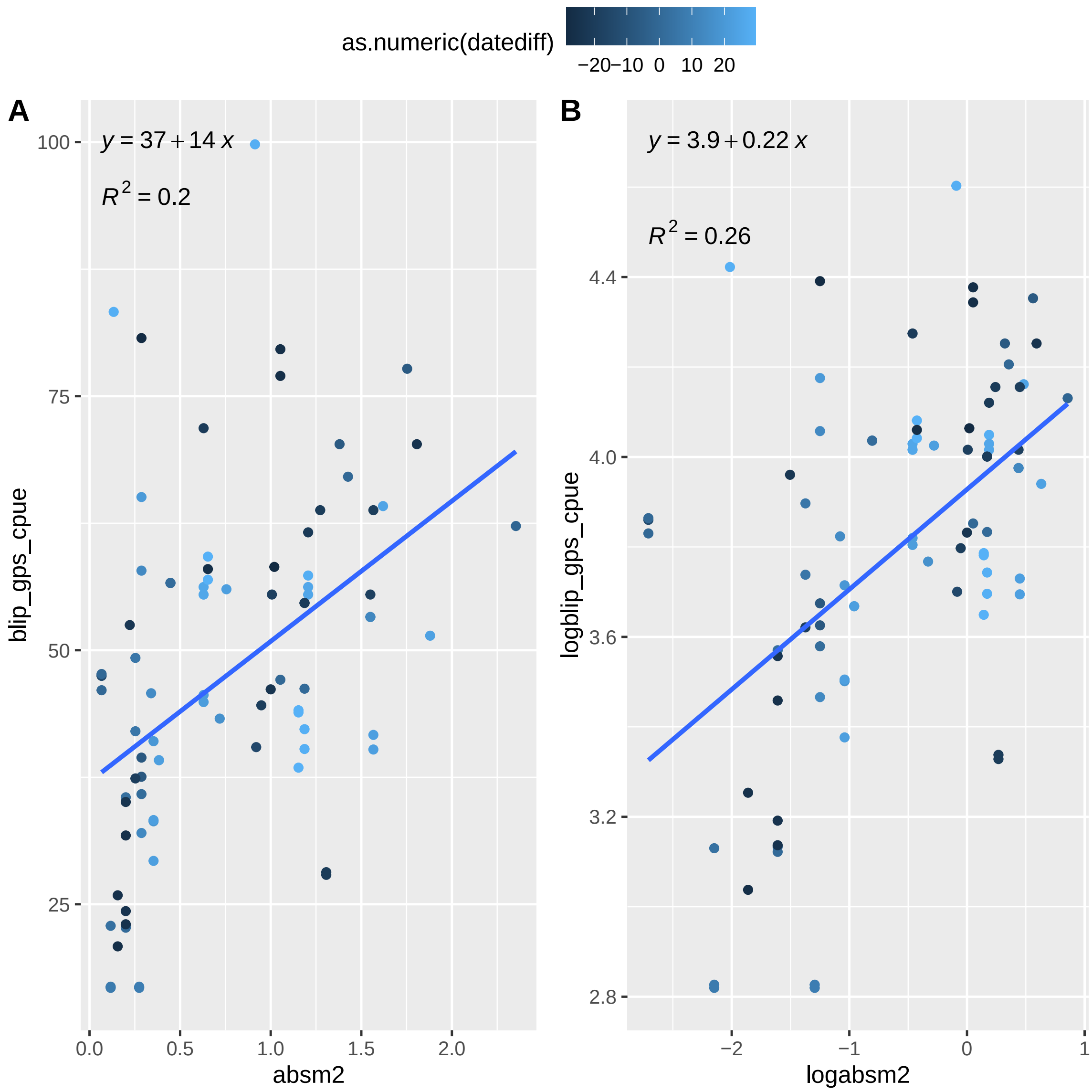


Figure 9.5: Relationship between FIS abalone density and CPUE, where fishing and survey events fall within a 30 day window.

9.3.1.2 Relationship between FIS and GRID

The relationship between annual catch per 1 hectare grid cell and FIS density, is marginally better, however R^2 remains low at 0.36, suggesting FIS density is a poor predictor of annual catch (Figure 9.6).

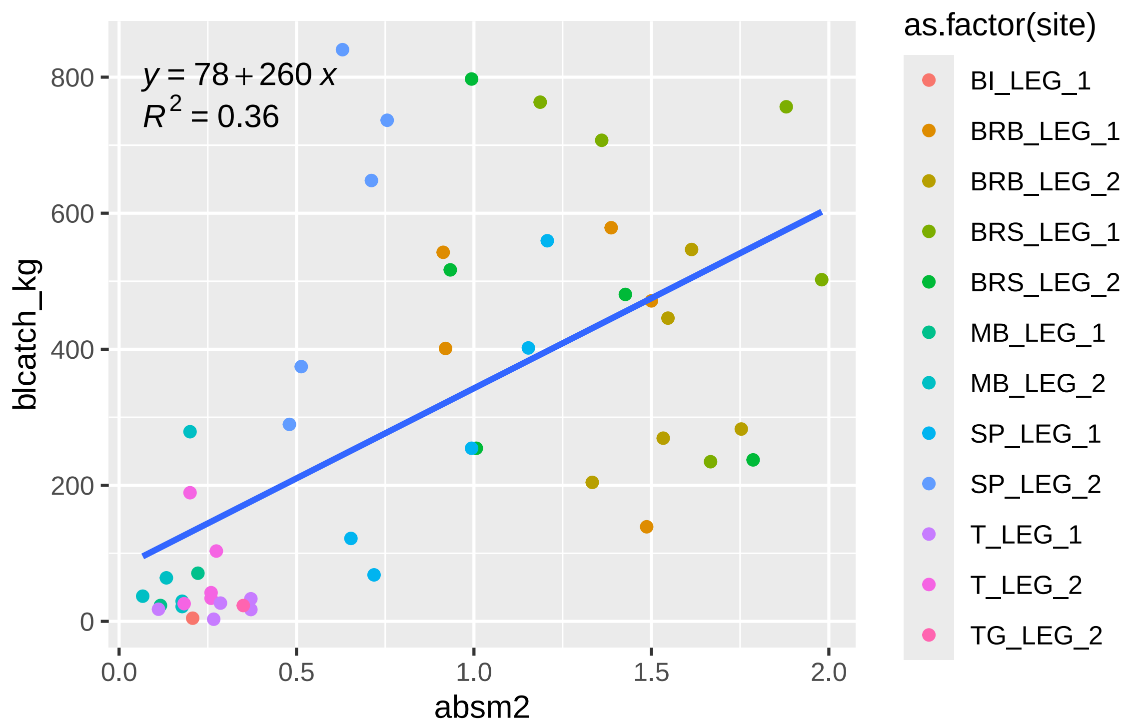


Figure 9.6: Relationship between FIS abalone density and annual catch in the overlaying grid cell.

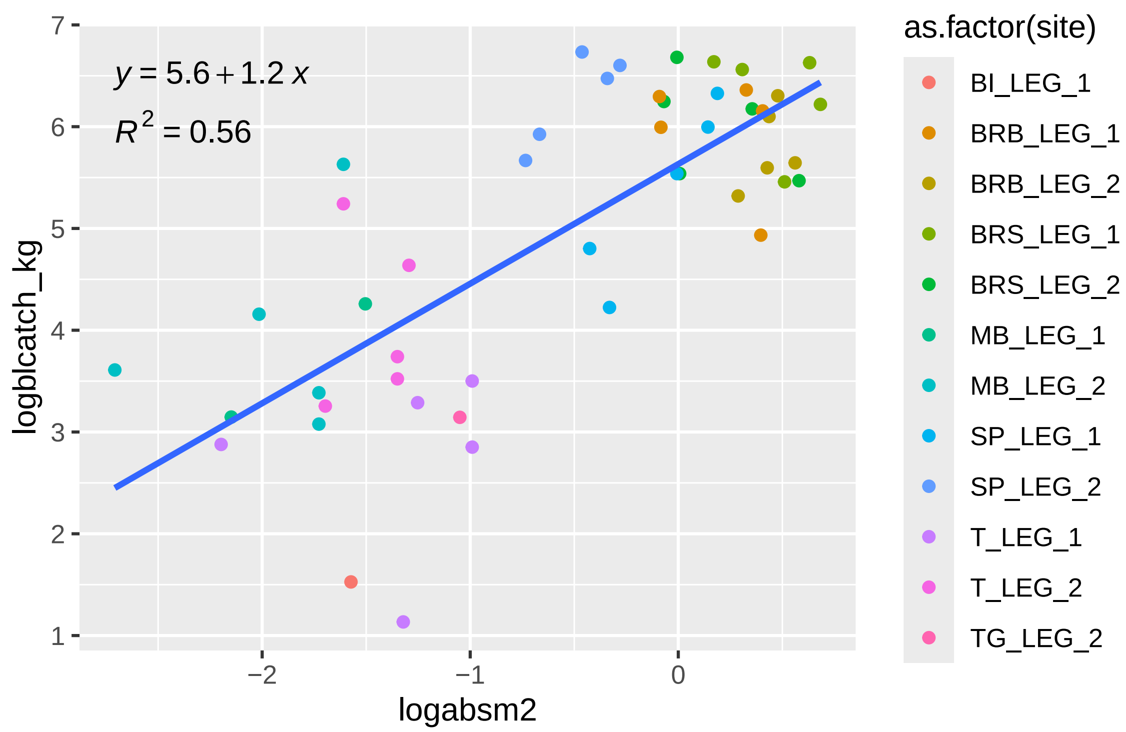


Figure 9.7: Effect of log transformation on the fit between FIS abalone density and annual catch in the overlaying grid cell.

9.3.2 Overlay grid on FIS-ARM sites

The juvenile abundance data from research surveys provides an opportunity to explore the relationship between spawning biomass and juvenile abundance, where annual catch in a grid cell is a proxy for biomass. There appears to be no effective relationship between FIS-ARM density and our proxy for biomass (Kg/Ha harvested per grid cell), with R^2 effectively zero (@fig-grid-fisARM). At best we might conclude that there are no observations where juvenile abundance is high and annual catch is also high. Anything else is wishful thinking or a guess.

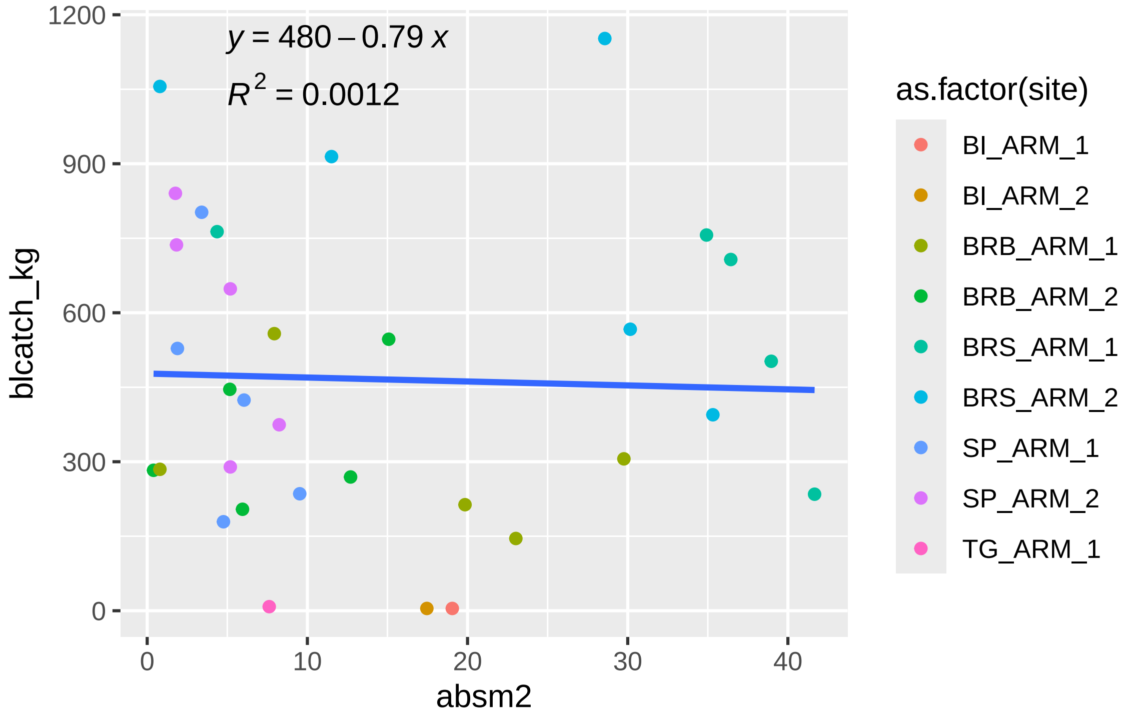


Figure 9.8: Relationship between FIS-ARM juvenile abalone density and annual catch in the overlaying grid cell.

9.4 Discussion

This chapter examined simple relationships between catch rates (individual dive events) and annual catch (grid cell) using spatial methods and, two forms of research derived abalone density - visual surveys of emergent abalone, and visual surveys of cryptic juvenile abalone on recruitment modules. In all cases, the relationship between abundance of emergent abalone and density was linear and positive, rather than contradictory. However, this study found the fit to be generally poor, and offers very low confidence that a strong predictive relationship between FIS density and catch rates or annual catch/Hectare might be found. Restricting the time elapsed between completion of the FIS and overlapping fishing events to 30 days, gave almost no improvement to the model fit. Log transformation allowed some improvement in the fit, but not for the relationship between FIS and individual fishing events that overlapped with a study site.

Within a grid cell or within an individual dive polygon, there is likely to be further spatial complexity in abundance of abalone. It may be that for grid based analyses, the 1Ha grid scale is too large, and in affect smoothing what is likely to be a rather heterogeneous field. Reducing the grid cell size to 0.5 or 0.25 hectares may be provide a much better relationship between geo-referenced fishery data and FIS data. Further, rather than an annual summation of GPS points within a grid cell, this could be reduced to shorter time-steps for example quarterly. Similarly, rather than using the 90% isopleth to represent dive polygons, a smaller isopleth such as the 50% or 25% isopleth may capture some of the underlying within dive polygon heterogeneity.

10 Climate Data Part 1: Climate plots

10.1 Data loading and preparation

The data set used in this Part was as at the 2023-03-22. It only looks at Blacklip (BL) data. The aim of this Part is to simply investigate and display the data with a particular focus on the environmental data and how they relate to catch and each other.

10.1.1 Data names and structure

Some of the variable names in the loaded data set are as follows:

```
n = 166913, p = 42
```

Variable	class		Variable	class		Variable	class
diver_id	integer	|	rawblhours	numeric	|	totalcpue	numeric
fishing_date	POSIXct	|	adjblhours	numeric	|	blprop.factor	factor
fishyear	numeric	|	cpue	numeric	|	LnCE	numeric
yearq	character	|	propblip	numeric	|	plaindate	Date
fishmonth	numeric	|	take_vessel	character	|	number	numeric
blockno	numeric	|	tran_vessel	character	|	expcat	factor
subblockno	character	|	ports	character	|	wave_power	numeric
region	character	|	numdivers	numeric	|	wave_hght	numeric
newzone	character	|	lastname	character	|	wave_dir	numeric
catch	numeric	|	firstname	character	|	wave_sd	numeric
totalcatch	numeric	|	entitle-ment.num	character	|	wind_mode	character
hours	numeric	|	numdocks	integer	|	C	numeric

From this we have a few climate variables as continuous variables such as:

- `wave_hght`: wave height per day,
- `wave_power`: wave power per day,
- `wave_dir`: wave direction,
- `C`: sea surface temperature (which we will rename to `sst` later).

As factors:

- `wind_mode`: a character vector of direction so we may need something more numeric value later. There is an equivalent for wave mode.

Also to relate the logbook data with each other, such as:

- `totalcatch`: the catch of BL and GL abalone
- `catch`: catch of BL
- `hours`: the hours taken to catch this
- `rawblhours`: raw BL hours
- `asjblhours`: adjusted BL hours (many NAs!!)
- `cpue`: catch/hours
- `LnCe`: log(cpue)
- `propblip`: proportion of Blacklip in the total catch

10.1.2 Missing value composition

A quick check on the variables which have missing values (NAs) in them, and how many NAs each has, is as follows:

```
miss <- abCEbl %>% is.na() %>% colSums() %>% .[. > 0]
cbind(NAs = miss, `%` = round(100*miss/nrow(abCEbl), 2)) %>% kable()
```

	NAs	%
subblockno	40154	24.06
rawblhours	30033	17.99
adjblhours	158866	95.18
lastname	59	0.04
totalcpue	136880	82.01
blprop.factor	32	0.02
expcat	5793	3.47
wave_power	423	0.25
wave_hght	423	0.25
wave_dir	423	0.25
wave_sd	423	0.25
wave_mode	423	0.25
wind_speed	423	0.25
wind_dir	423	0.25
wind_sd	423	0.25
wind_mode	423	0.25
C	344	0.21

10.1.3 Data processing and convenience extensions

It is convenient to work with a working data set, leaving the initial loaded data intact.

The first extension is to add a `diver_ex` variable that gives a rough measure of the experience each diver has, in years, at that point of the data record.

```
vars <- abCEbl %>%
  ungroup() %>%
  select(plaindate, diver_id) %>%
  unique() %>%
  arrange(diver_id, plaindate) %>%
  within({
    diver_year <- as.POSIXlt(plaindate)$year + 1900
    # diver_ex <- ave(diver_year, diver_id, FUN = function(x) x - min(x, na.rm = TRUE))
    # # this code has a bug if someone takes a gap year or more
    diver_ex <- ave(diver_year, diver_id,
                    FUN = function(x) (cumsum(!duplicated(x)) - 1))
  })
BLdat <- left_join(abCEbl, vars, by = c("plaindate", "diver_id")) %>% data.frame()
```

The next step is to do various tidying up, filter from 1991 and placing some guards on key variables:

```
BLdat <- BLdat             %>%                        # start with this
  subset(fishyear > 1991) %>%                        # remove any old data
  droplevels()            %>%                        # remove unused fishers
  rename(zone = newzone,
         logCPUE = LnCE)  %>%                        # new name for an old variable
  within({
    diver_id <- factor(diver_id)
    year <- factor(fishyear)
    month <- factor(month.abb[fishmonth], levels = month.abb)
    block <- factor(substring(100 + blockno, 2))
    two_divers <- 0 + (numdivers == 2)  # binary to allow for 2-diver events
    hours <- pmax(1/12, pmin(hours,24)) # always good to check no weird hours
    logCPUE <- pmax(-2.5, pmin(logCPUE, 8.5)) # put a guard on logCPUE to stop later crashing
    zone <- factor(zone, levels = c("BS", "N", "W", "E"))
    sst <- C
  })                      %>%
  select(diver_id, diver_ex, year, month, plaindate, zone, block,two_divers, propblip,
         catch, totalcatch, hours, cpue, logCPUE, wave_hght, wave_power, wave_dir, wave_mode,
         wind_speed, wind_dir, wind_sd, wind_mode, sst) %>%
  data.frame(stringsAsFactors = FALSE)
```

Now we add in the numerical compass direction (compass rose) variable, but first note that one of the compass directions recorded in the data does not have the conventional acronym:

```
with(BLdat, table(wave_mode, useNA = "a"))
```

```
wave_mode
    E   ENE   ESE     N   NNE   NWN     S   SSE   SSW     W   WNW   WSW  <NA>
 5430  7206  6392  1125  2840  1161 15101 10633 49475  9198  2491 55438   423
```

The direction listed as `NWN` should presumably be `NNW`. Make this change speculatively and add the `compass` numerical variable.

```
#cpoints <- setNames(seq(0, by = 22.5, length.out = 16),
#                     c("N", "NNE", "NE", "ENE", "E", "ESE", "SE", "SSE",
#                       "S", "SSW", "SW", "WSW", "W", "WNW", "NW", "NNW"))
BLdat <- BLdat %>%
  within({
    wave_mode <- sub("NWN", "NNW", as.character(wave_mode))
    #compass <- cpoints[wave_dir]
    #wave_mode <- factor(wave_mode, levels = names(cpoints))
  })
```

```
with(BLdat, table(wave_mode, useNA = "a"))
```

```
wave_mode
    E   ENE   ESE     N   NNE   NNW     S   SSE   SSW     W   WNW   WSW  <NA>
 5430  7206  6392  1125  2840  1161 15101 10633 49475  9198  2491 55438   423
```

How the favoured wave directions vary over `zone` can be seen in the following circular bar graphs. *Noting that these are the wave directions on the day of fishing, and should not be considered as indicative of the wave climate in each zone.*

```
# library(data.table)
# windbreaks <- c(0,5,10,15,20,25,30,35,40,45,50,180)
# windlabels <- c("0-5","5-9","10-14","15-19","20-24","25-29","30-34", "35-39","40-44","45-49"
# setDT(BLdat)[ , windgroup := cut(wind_speed,
#                                  breaks = windbreaks,
#                                  right = FALSE,
#                                  labels = windlabels)]
#
# wavebreaks <- c(0,5,10,15,20,25,30,35,40,45,50,750)
# wavelabels <- c("0-5","5-9","10-14","15-19","20-24","25-29","30-34", "35-39","40-44","45-49"
# setDT(BLdat)[ , wavegroup := cut(wave_power,
#                                  breaks = wavebreaks,
#                                  right = FALSE,
#                                  labels = wavelabels)]
```

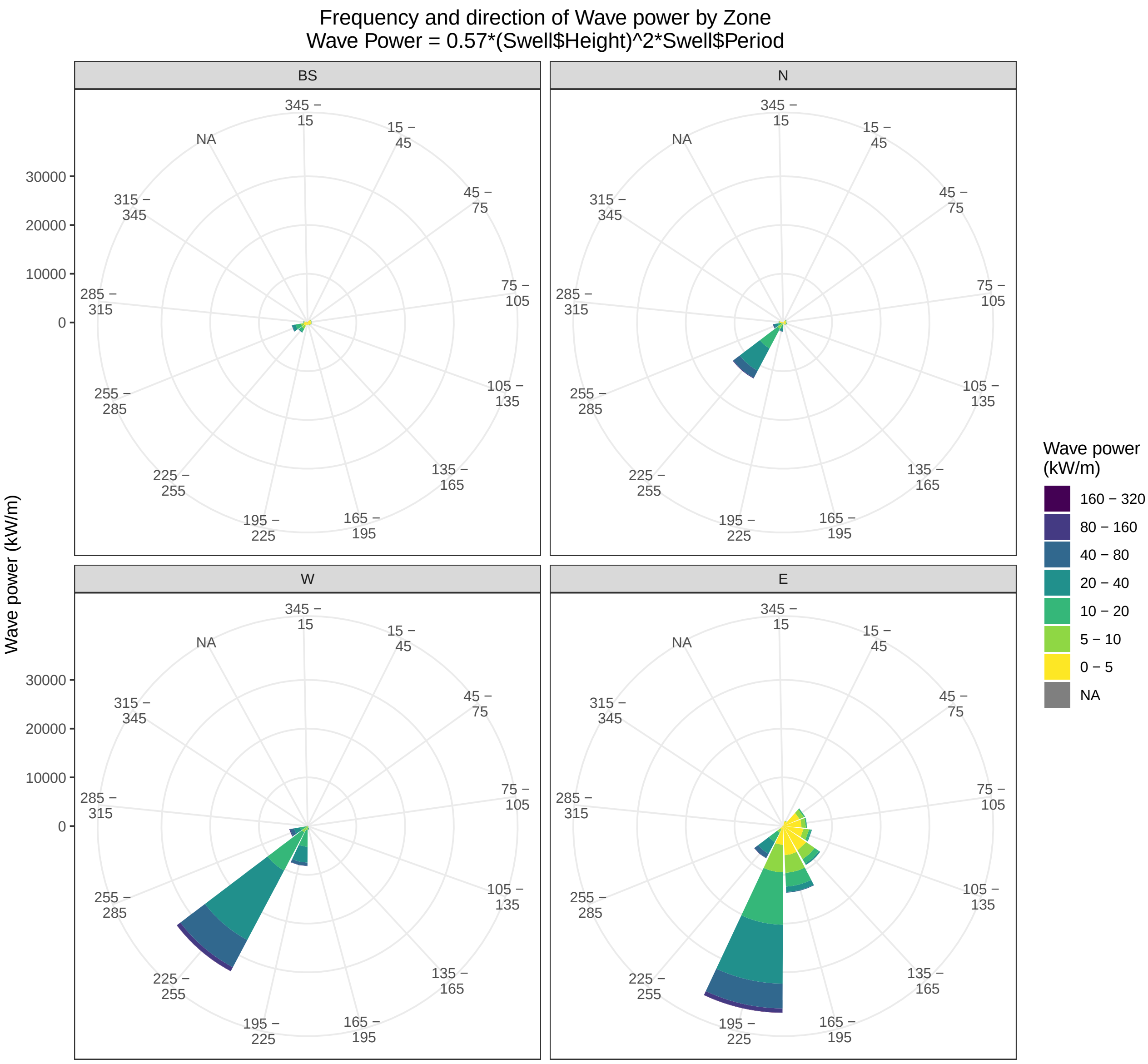
```r
source("Windrose.R")
library(viridis)
```

```
Loading required package: viridisLite
```

```r
p.sr1 <- plot.windrose(data = BLdat,
                       KeyTitle="Wave power\n(kW/m)",
                       spd = "wave_power",
                       dir = "wave_dir",
                       spdmax = 750,
                       spdseq = c(0,5,10,20,40,80,160,320))
                       #spdseq = c(seq(0, 750, 150))

chart_title <- paste("Frequency and direction of Wave power by Zone","\nWave Power = 0.57*(Swe

p.sr4 <- p.sr1 + facet_wrap(~zone, ncol = 2,drop=FALSE) + labs(title=chart_title,y="Wave power
p.sr4
```



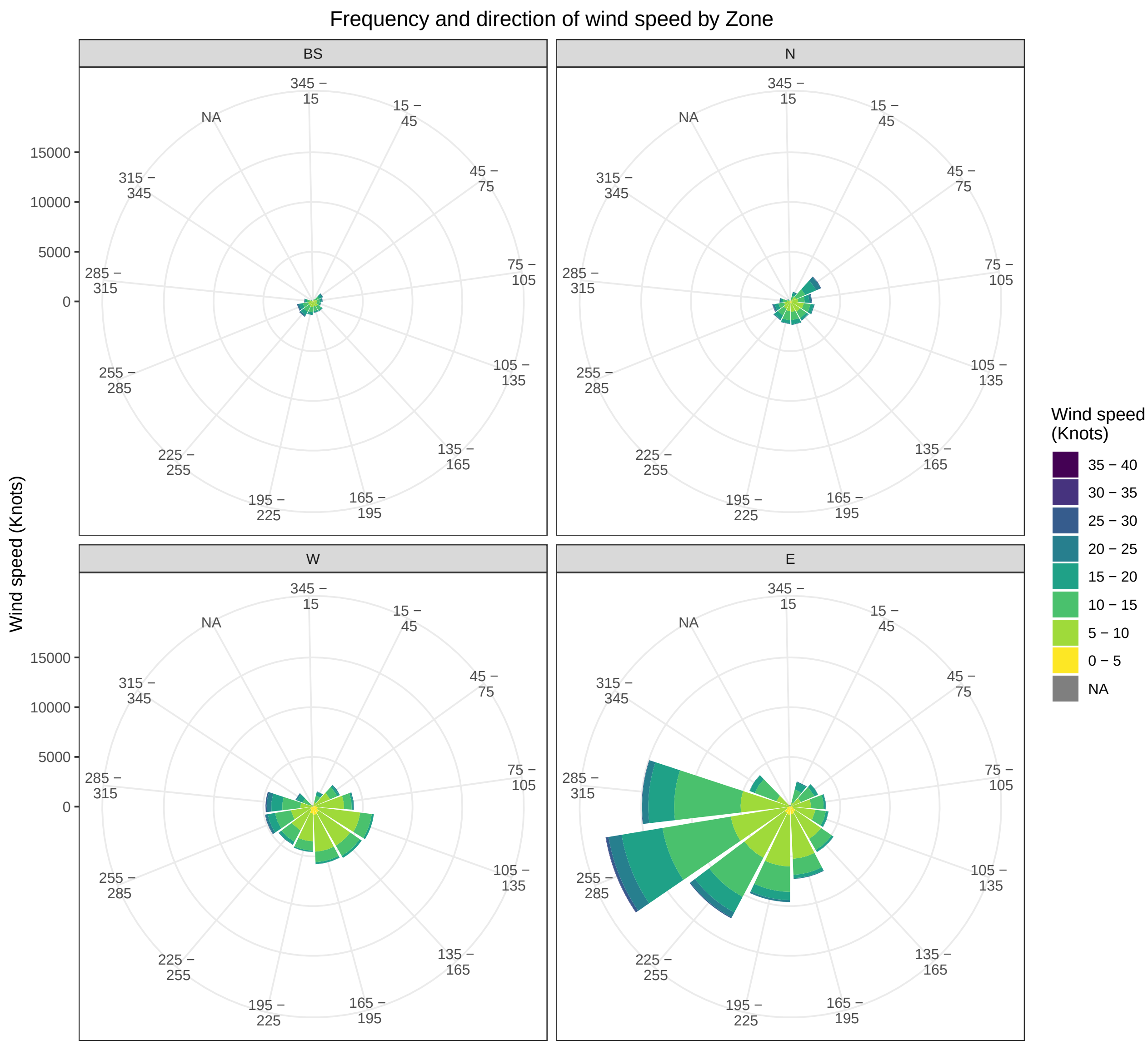
How the favoured wind speeds vary over `zone` can be seen in the following circular bar graphs. *Noting that these are the wind speeds on the day of fishing, and should not be considered as indicative of the wind climate in each zone.* On the East coast of Tasmania, Winds with a westerly component are essentially offshore (with a few exceptions), thus we see fishing activity on days when winds appear relatively strong compared to fishing days the Western zone.

```
source("Windrose.R")
library(viridis)
wind_sr1 <- plot.windrose(data = BLdat,
                          KeyTitle="Wind speed \n(Knots)",
                          spd = "wind_speed",
                          dir = "wind_dir",
                          spdmax = 40,
                          spdseq = c(0,5,10,15,20,25,30,35,40))
```

```
                        #spdseq = c(seq(0, 750, 150))

chart_title <- "Frequency and direction of wind speed by Zone"

wind_sr4 <- wind_sr1 + facet_wrap(~zone, ncol = 2,drop=FALSE) + labs(title=chart_title,y="Wind
wind_sr4
```



10.1.4 Sea surface temperature

A crude summary of the sea surface temperatures included in the data is as follows:

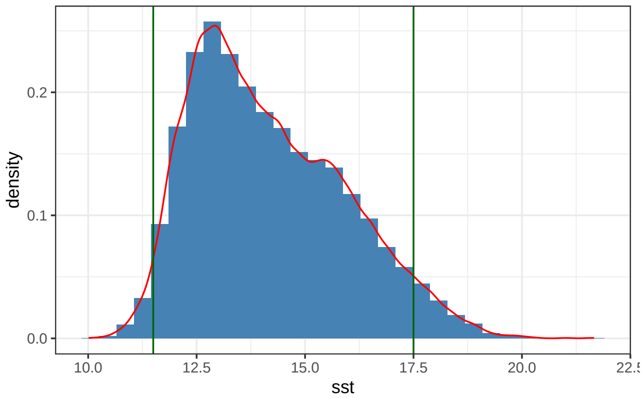
```
ggplot(BLdat) + aes(x = sst) +
  geom_histogram(aes(y = ..density..), bins = 30, fill = "steelblue") +
```

```
  geom_density(aes(y = ..density..), colour = "red") +
  theme_bw() + geom_vline(xintercept = c(11.5, 17.5), colour = "darkgreen")
```

```
Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.
i Please use `after_stat(density)` instead.
```

```
Warning: Removed 344 rows containing non-finite outside the scale range
(`stat_bin()`).
```

```
Warning: Removed 344 rows containing non-finite outside the scale range
(`stat_density()`).
```

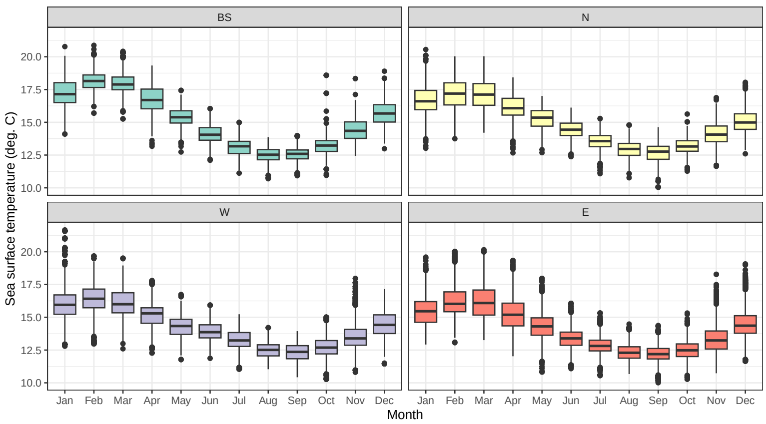


From this the bulk of the temperatures are about 11.5-17.5 degrees; and over 20 is an extreme (max = 21.663847).

SST, naturally, has a strong seasonal pattern and hence is strongly confounded with `month`. As `month` is a useful conditioning variable, the two will be in competition when considered as important predictors of catch.

There is also some variation across zones, with the Bass Strait and Northern zones being slightly warmer in summer than Western and Eastern zones, as shown in the following diagram.

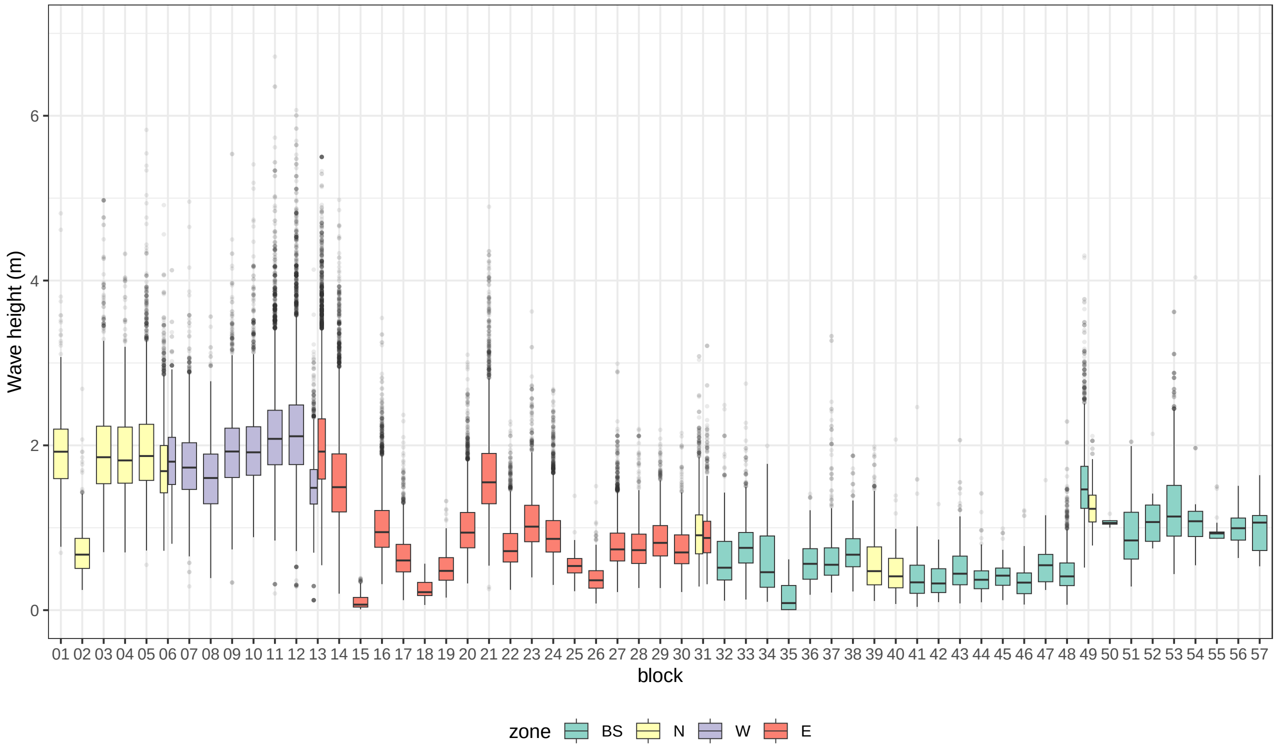
```
Warning: Removed 344 rows containing non-finite outside the scale range
(`stat_boxplot()`).
```



We know that SST effects are much more complex than simply a plot by month. If it is used in any model it would at least have an offset of about 6-8 years as papers relate the impact of SST more on recruitment events than adults. However, an extreme event may impact adults too. These concepts will be taken up at another stage as we are simply describing the data here.

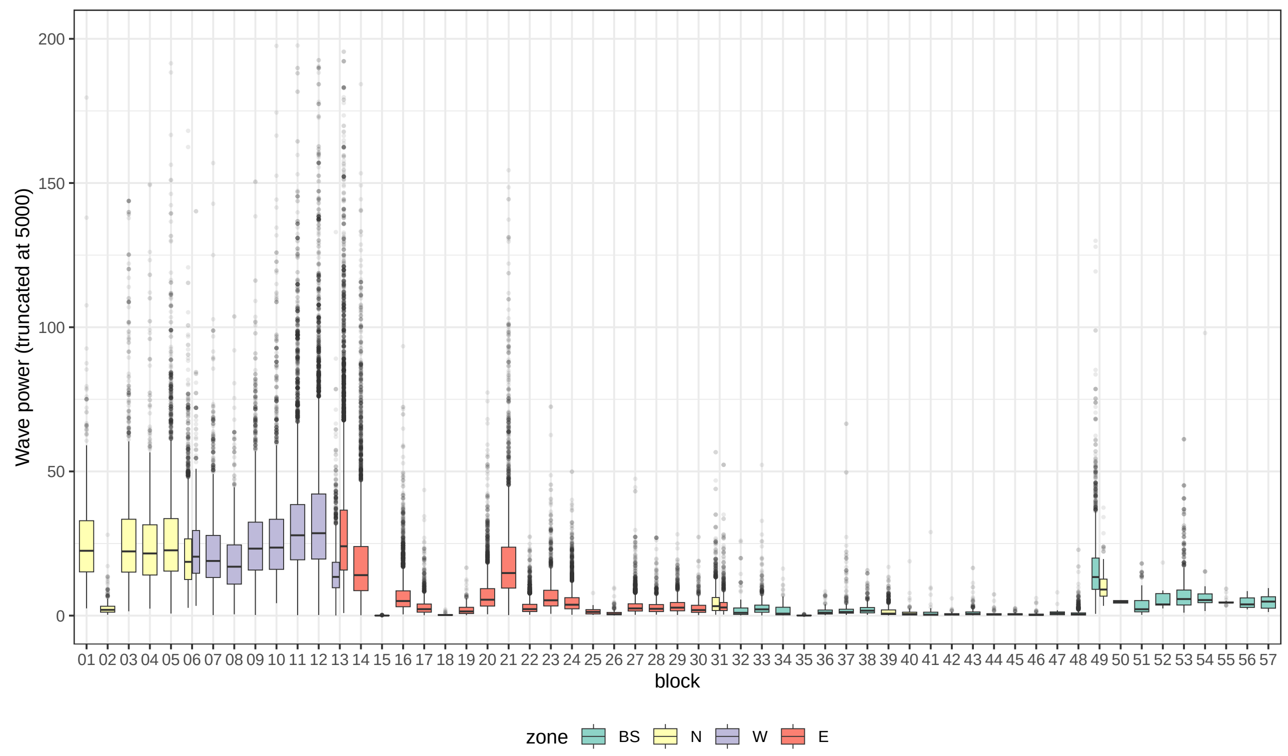
10.1.5 Wave height and wave power

First a boxplot of wave height over blocks (and zones):



The distribution is most variable in the North and West zones.

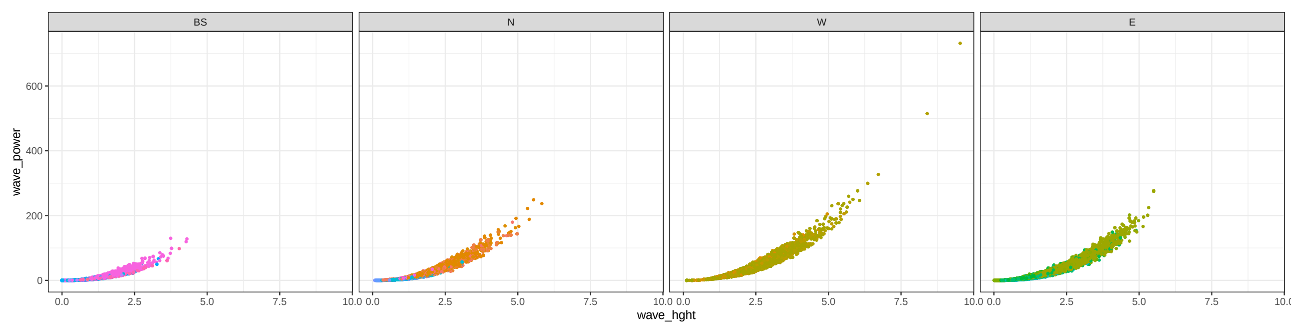
Now for the wave power values. Note that these values have been *truncated* to below 5000, to allow the pattern in the bulk of the data to be visually appreciated. These few (88) very large values will exert very high leverage in any analysis involving wave power, unless steps are taken to mitigate the effect.



–>

10.1.6 Collinearity between wave properties

The following figure shows the relationship between wave height and power, separately for each of the four zones:



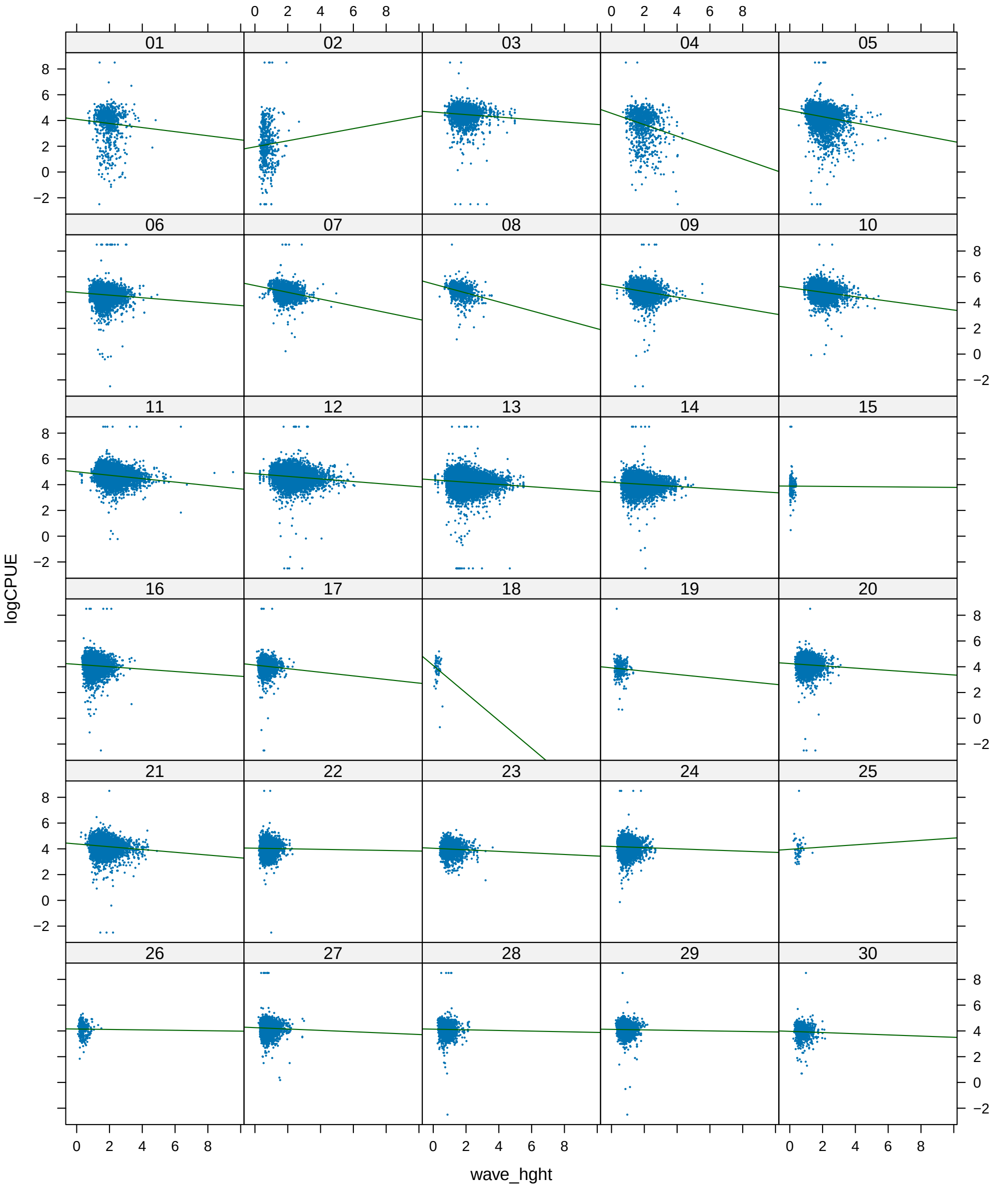
From this, one can see there is a strong curved relationship between wave height and wave power. This is true by block. There are 423 NAs in the `wave_hght` compared to 423 however, it does seem to be often the case when `wave_hght` is NA then `wave_power` is 0. Assuming that a zero is in fact a missing value for either variable, we arrive at the table:

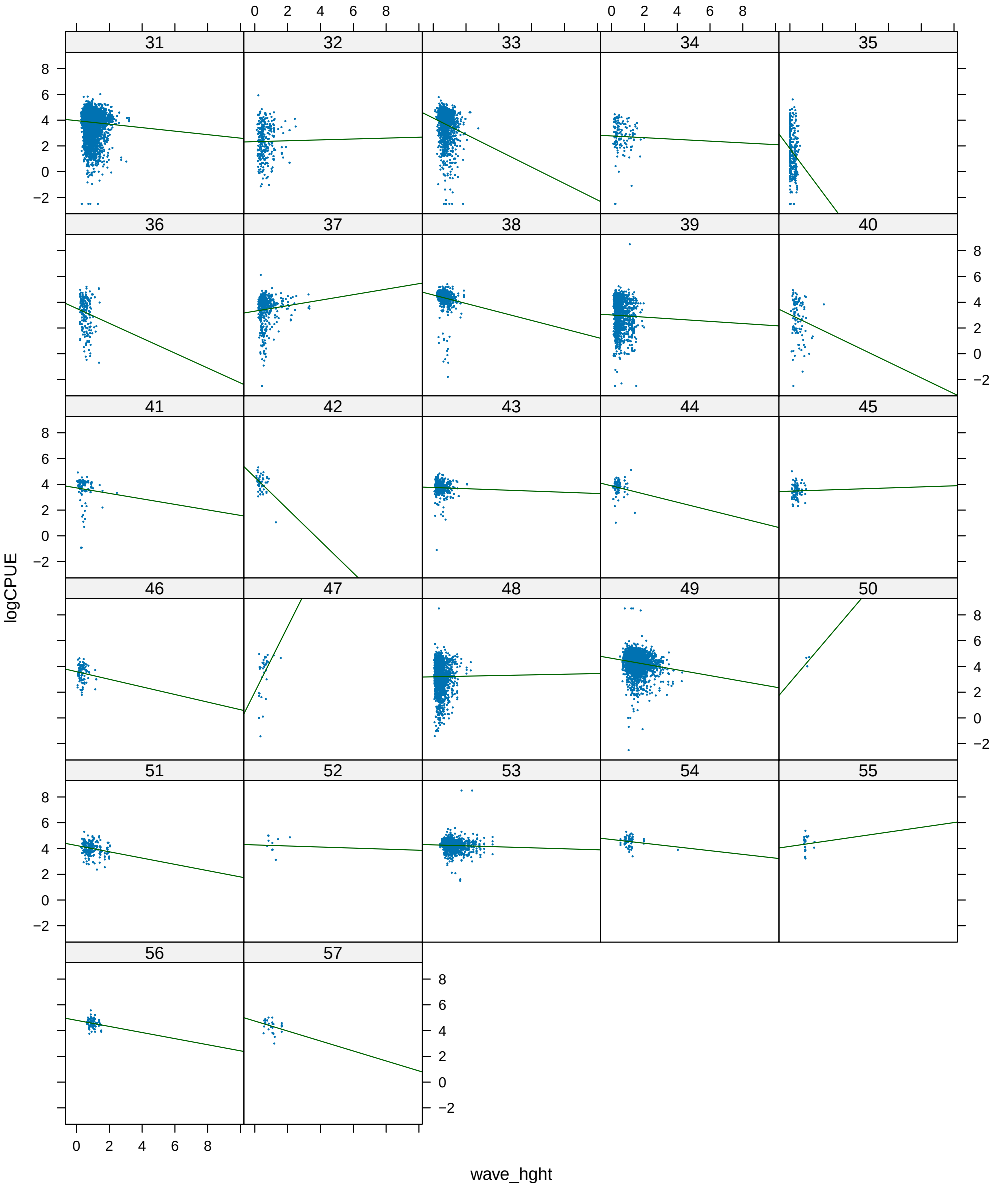
	FALSE	TRUE
FALSE	166477	0
TRUE	0	436

10.2 Marginal relationships to CPUE

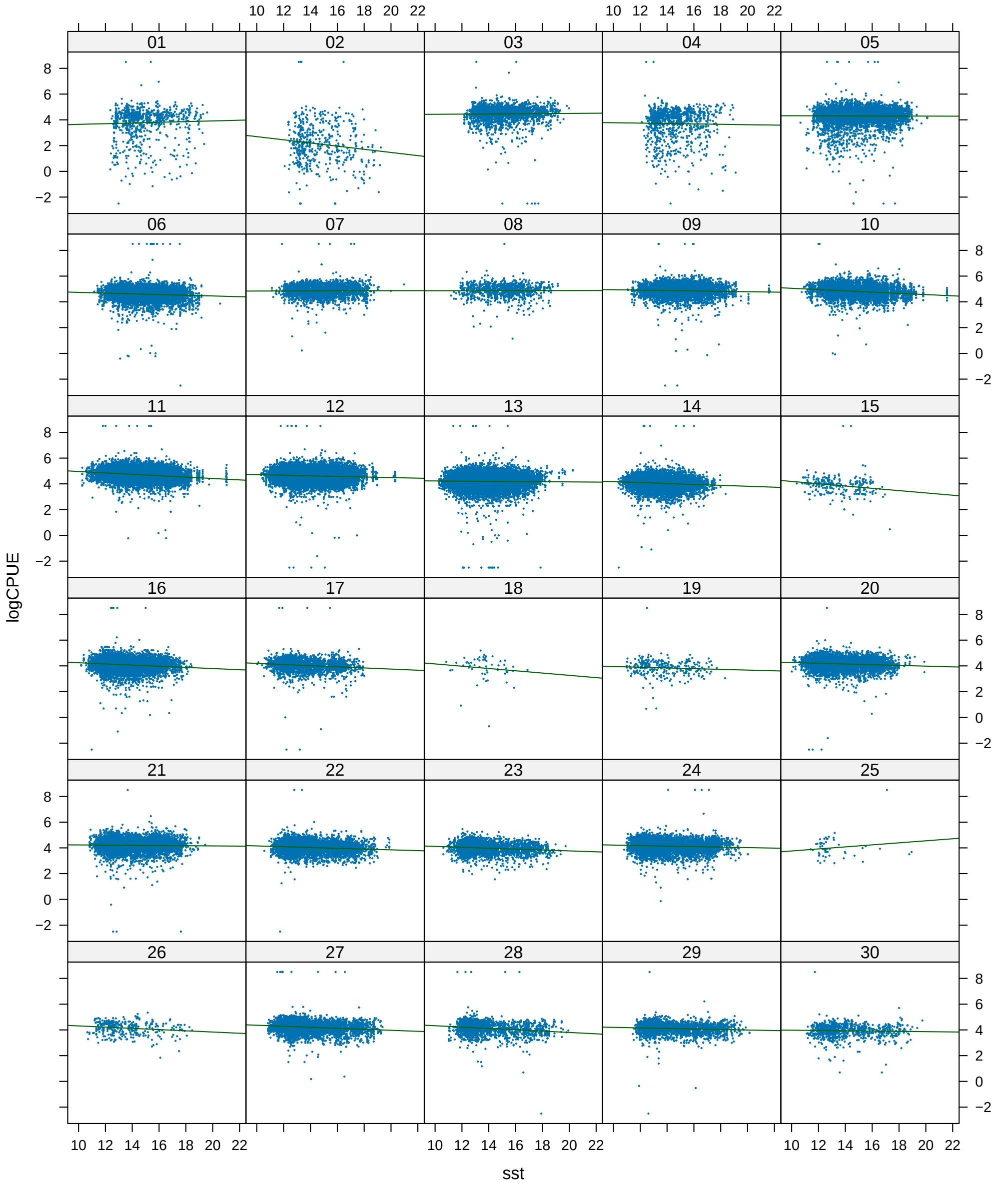
Putting it all together.

First with wave height over blocks:

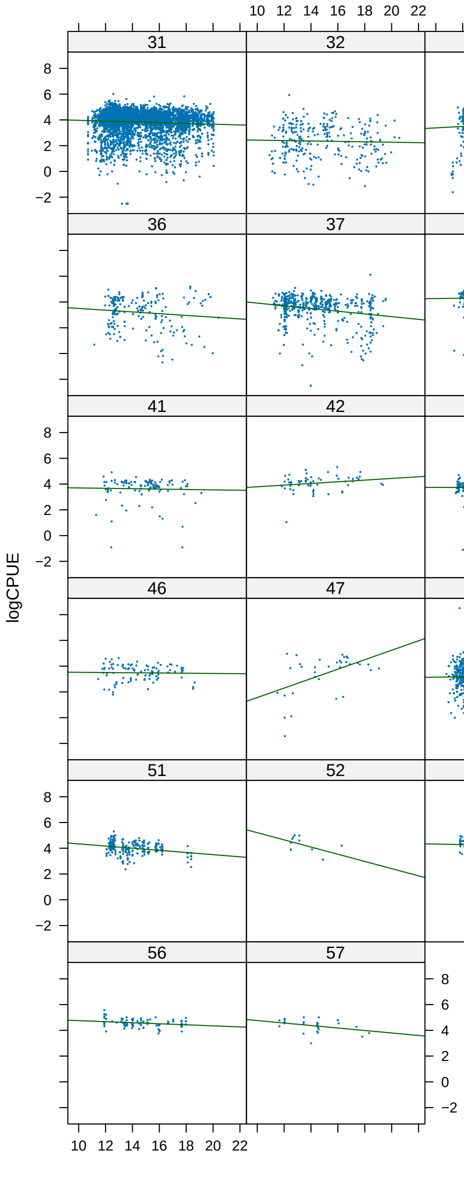




Then, with SST over blocks:



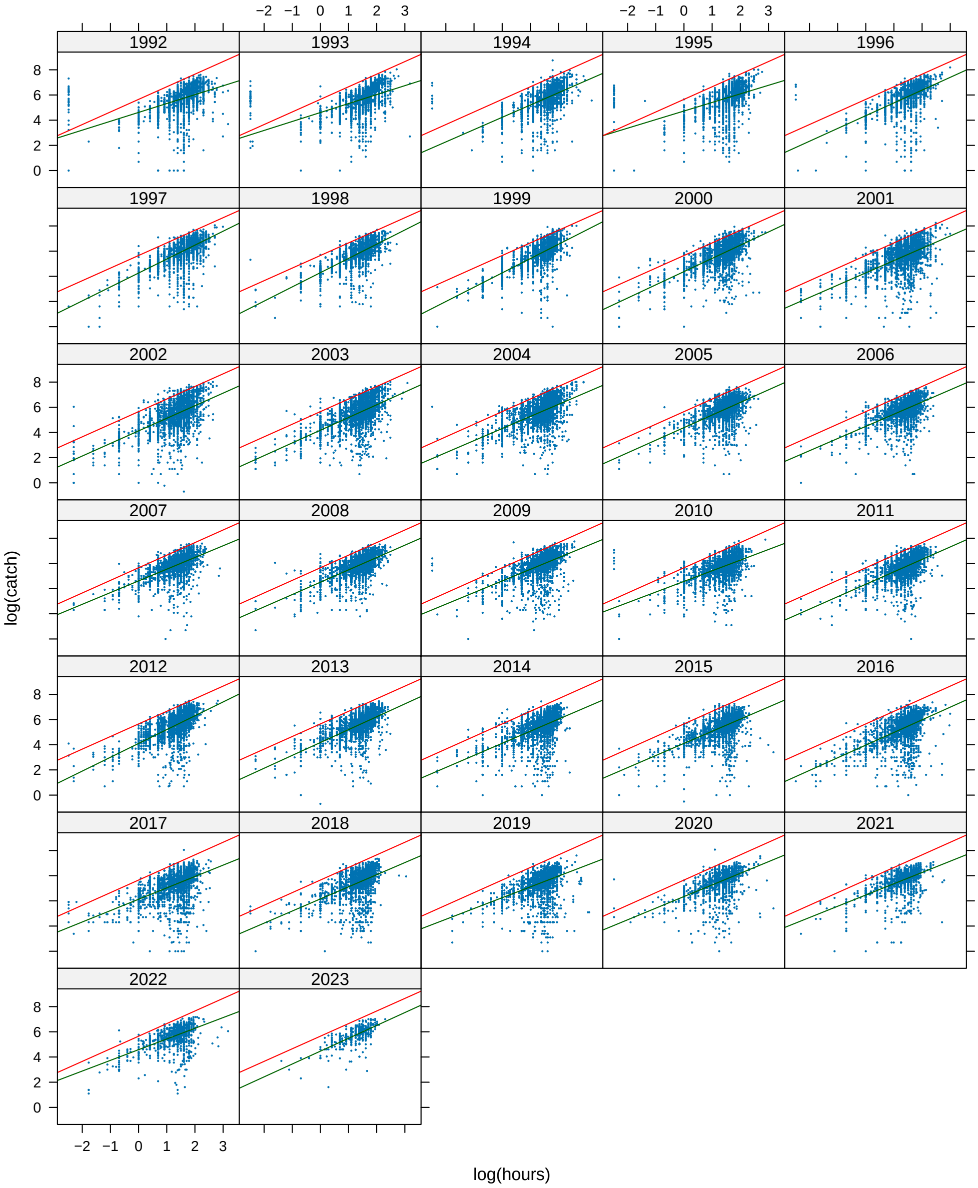
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
logCPUE
sst



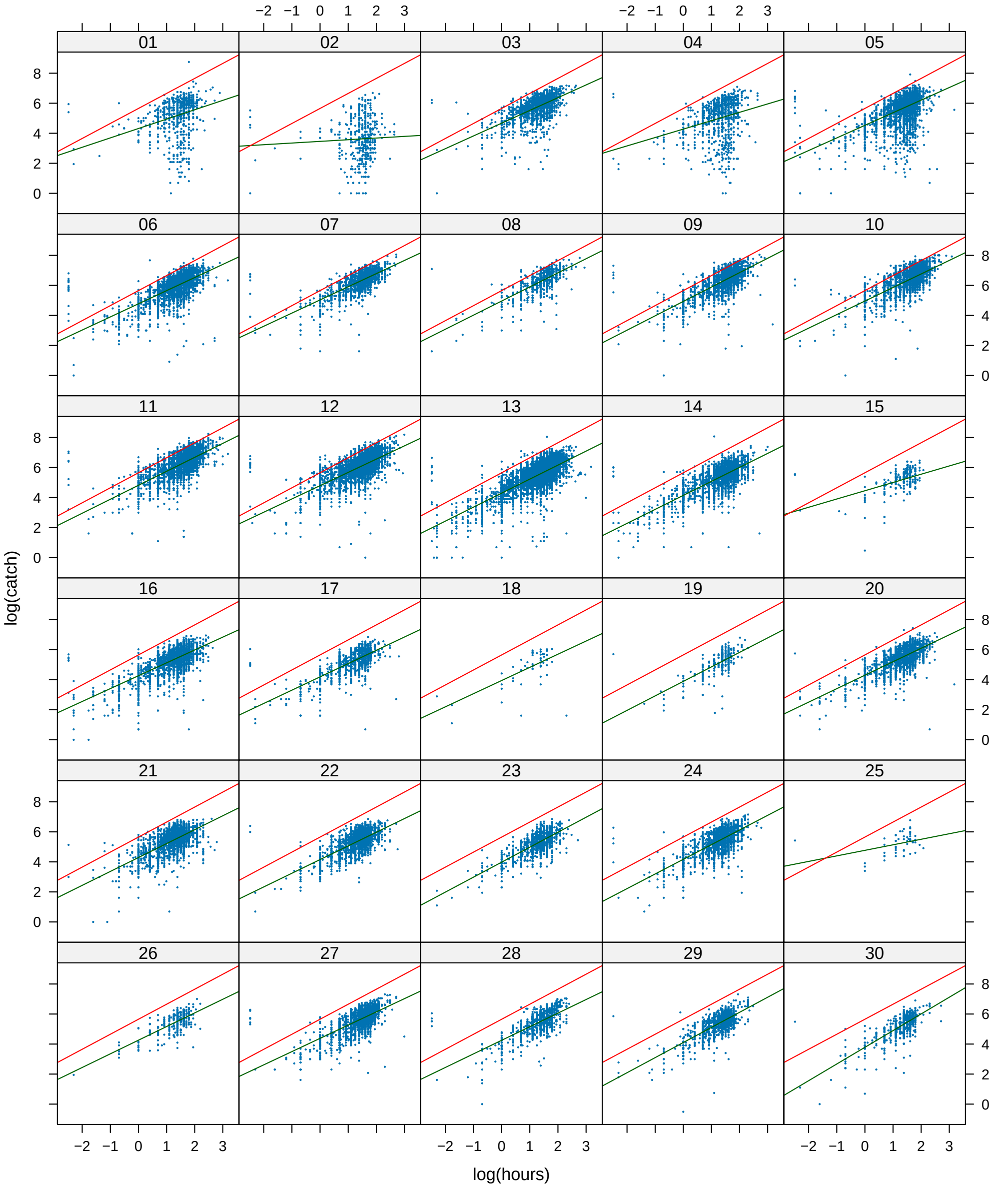
10.3 Relationship between catch and hours

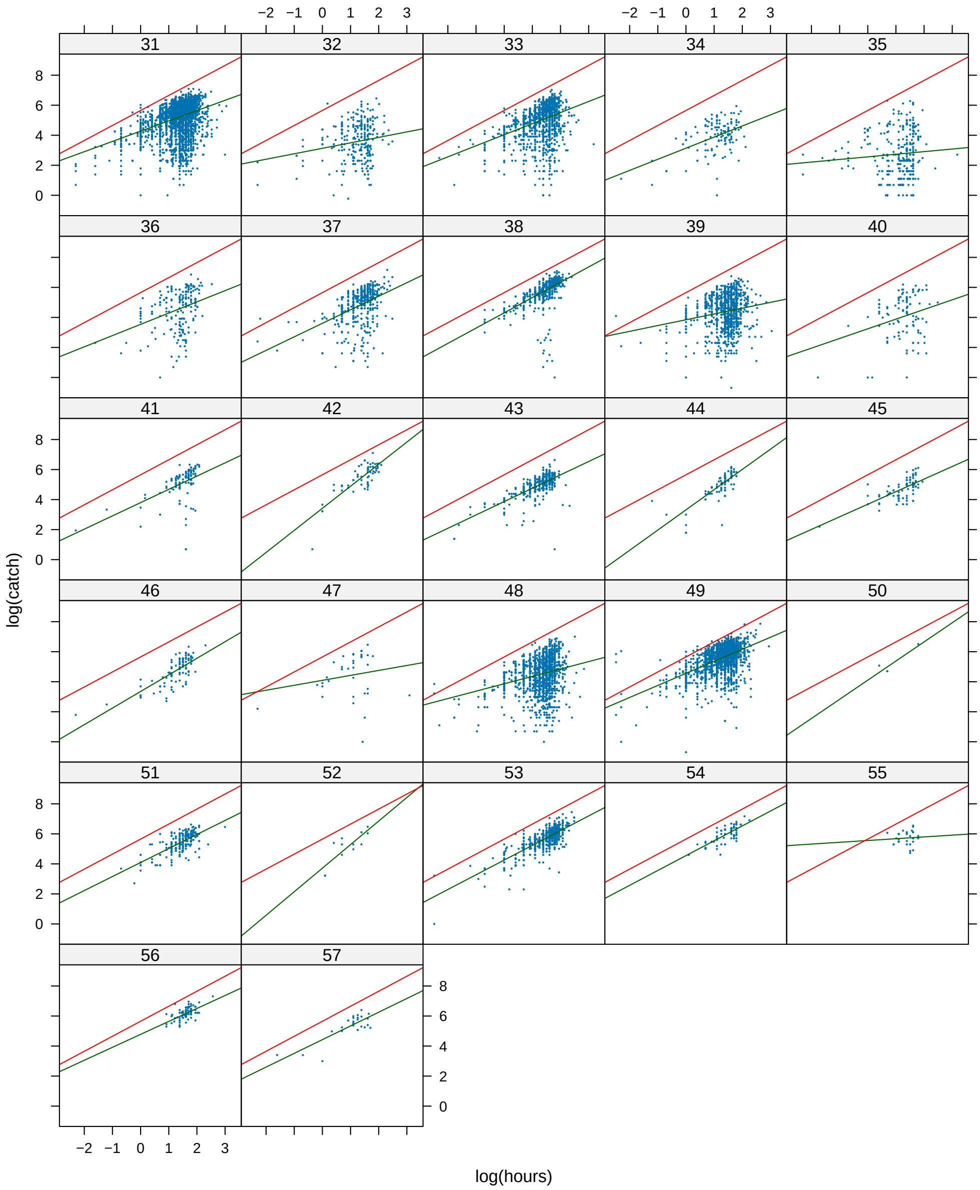
This is a first look at the relationship between `log(catch)` and `log(hours)` ignoring predictors other than `year` or `block`. This is just a quick, first step, visualisation.

First, a check with `year` as the conditioning variable. In the following layout the *green* line is the local regression line for the panel, whereas the *red* line, the same in all panels, is a reference line with slope 1. Parallelism between the two lines offers some support for catch being proportional to effort, (where the latter is nominal effort in hours).



A second view looks at the relationship conditional on `block` and marginalised over `year`. This is shown below.


01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
log(catch)
log(hours)
−2 −1 0 1 2 3
0 2 4 6 8



Note that in these displays, zero catches have been excluded rather than re-jigged in some way to keep the log from failing. However `hours` *has* been re-jigged to ensure that no record with a positive Blacklip catch has an `hours` value of less than 5 minutes (arbitrarily chosen!).

Although it is always important to have a first simple look at the data, there is nothing in these analyses that show that simple relationships are all that is needed to characterise the data.

As a result, Part 2 will explore things further in the context of a standardisation.

11 Climate Data Part 2: Standardisation

11.1 Preamble

This is Part 2 of our exploration of how the climate data may be useful in enhancing the standardisation of Blacklip Abalone. Part 1 was mainly concerned with visualising and exploring the climate data itself.

The data we use is the same as the extended version produced and used in Part 1.

Variable	class	missing	Variable	class	missing
diver_id	factor	0	totalcatch	numeric	0
diver_ex	numeric	0	hours	numeric	0
year	factor	0	cpue	numeric	0
month	factor	0	logCPUE	numeric	0
plaindate	Date	0	wave_hght	numeric	436
zone	factor	0	wave_power	numeric	436
block	factor	0	wave_dir	numeric	423
two_divers	numeric	0	wave_mode	character	423
propblip	numeric	0	wind_speed	numeric	423
catch	numeric	0	wind_dir	numeric	423

11.2 Standardisation models

11.2.1 Fixed or random?

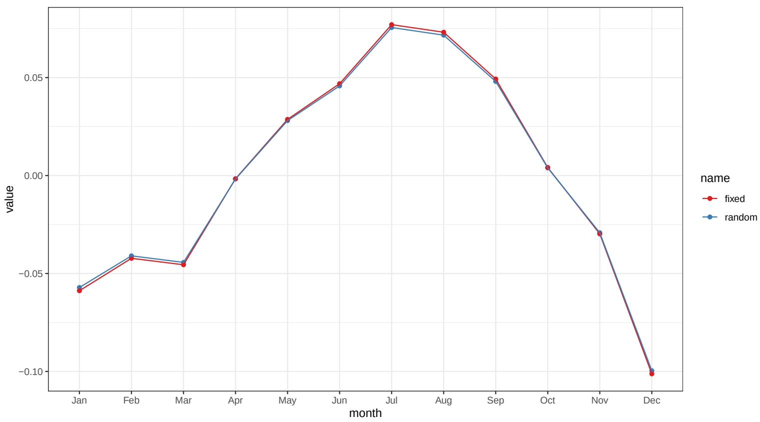
In previous notes we used the following model to build an integrated standardisation over the fishery.

```
mod_1 <- lmer(logCPUE ~ I(propblip-1) + two_divers + (1|year) + (1|month) +
                 (1|diver_id) + (1|block) + (1|block:year) +  (1|block:year:month),
               control = lmerControl(optimizer = "bobyqa"),
               data = BLdat)
```

This has `year` and `month` as random main effects, with random interactions involving them as well. It is interesting to see how different the estimates of these effects are when the main effects for `block`, `year` and `month` are changed to fixed effects, but leaving the interactions involving them as random. This is done in the following model fit:

```
mod_1_fe <- lmer(logCPUE ~ I(propblip-1) + two_divers + year + month +
                 (1|diver_id) + block + (1|block:year) +  (1|block:year:month),
               control = lmerControl(optimizer = "bobyqa"),
               data = BLdat)
```

Now we compare the corresponding fixed and random estimates, first for `month`:[1]

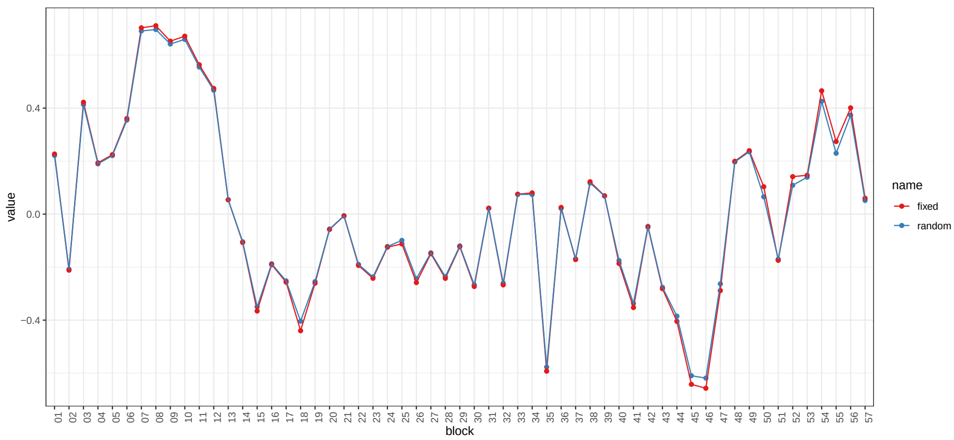


The difference is barely perceptible. Now for the `year` effect:

[1]In these comparisons the fixed effects are standardised so that, like the random effects, they have mean zero.

Here the shrinkage is more obvious, but the only year in which it is appreciable is 2023 and around the 2017 years.

Finally we look at the comparison with respect to the `block` main effect. Here there is little shrinkage except for a few blocks.



These comparisons do not argue in favour of using fixed effect terms *per se*, because the `diver_id` term and the *interaction* terms are still modelled as random. There is no logical or inferential problem with this, but it is still the case that the terms with the really large numbers of parameters are modelled as random, and these terms play a big part in ensuring the robustness, in a technical sense, of the models in standardisation.

If it were possible to fit a fully fixed effect model, (which seems doubtful, one of the reasons a block by block analysis may have been used elsewhere), there is no doubt that a model that includes interactions and diver terms could display substantial examples of useful shrinkage occurring.

11.2.2 Is catch proportional to dive hours?

Before leaving this model we look at the question of whether the primary underlying assumption, namely that, other things being equal,

$$C \propto H$$

where C is catch and H the dive hours spent by the diver(s) to make that catch.

In the context of the present model, this assumption can be checked by including `log(hours)` as a fixed effect predictor in the model. If the above proportional relationship applies, the coefficient of this term should be ≈ 0; any substantial deviation would indicate that the proportional relationship assumption is not credible, (or the model itself is inadequate).

```
mod_1a <- lmer(logCPUE ~ log(hours) + I(propblip-1) + two_divers + (1|year) + (1|month) +
                  (1|diver_id) + (1|block) + (1|block:year) +  (1|block:year:month),
               control = lmerControl(optimizer = "bobyqa"),
               data = BLdat)
```

The fixed effect terms have estimates as shown in the following:

	Estimate	Std. Error	t value
(Intercept)	4.2341	0.0558	75.8137
log(hours)	-0.0943	0.0024	-39.7968
I(propblip - 1)	2.5332	0.0114	222.1672
two_divers	-0.0908	0.0051	-17.7608

The coefficient of `log(hours)`, $\beta = -0.0943$, is small,[2] indicating that nominal effort might be better measured in an attenuated version of hours, namely

$$C \propto H^{1+\beta} \quad \text{where} \quad 1 + \beta = 1 + (-0.0943) = 0.9057$$

This initially slightly shocking suggestion could be explained by the possibility that as dive times in the same area increase they become progressively less productive of catch.

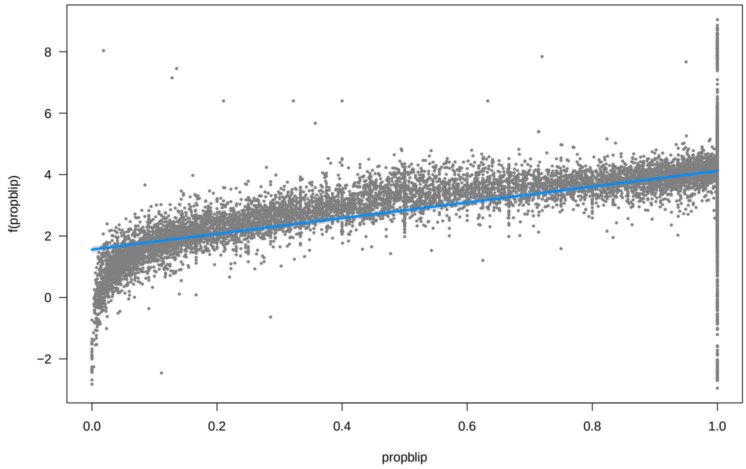
Note that in Part 1 of these notes, in the figure showing the marginal relationship of `log(catch)` with `log(hours)` per `block`, the green local regression lines had slopes *generally slightly less* than the red guide line with slope 1. This is in line with the current finding.

Note also that the presence of two divers rather than one also slightly decreases the expected catch, but this time proportionally by the multiplier $\exp(-0.0908) = 0.9132$. This may have a similar explanation: doubling the number of divers in one location may not be as productive as having each in a different location.

11.2.2.1 Minor model enhancements

Looking at the partial residuals shows an interesting pattern with respect to `propblip`:

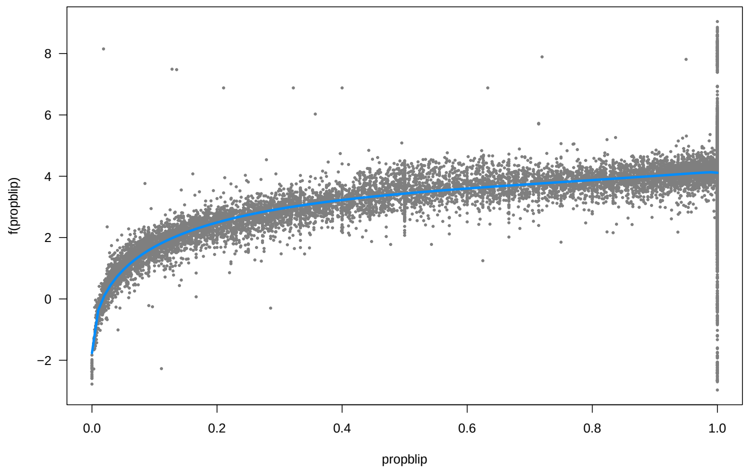
[2] The high significance is simply due to the very large sample size.



Observations with propblip < 1 are few, so the need to accommodate this feature is not strong. We can go some way to doing this, however, by using a polynomial term in a transformed propblip variable, as follows:

```
mod_2 <- lmer(logCPUE ~ poly(asin(sqrt(1-propblip)), 4) + two_divers + (1|year) + (1|month) +
                (1|diver_id) + (1|block) + (1|block:year) +  (1|block:year:month),
              control = lmerControl(optimizer = "bobyqa"),
              data = BLdat)
```

The partial plot of the term now looks like this:

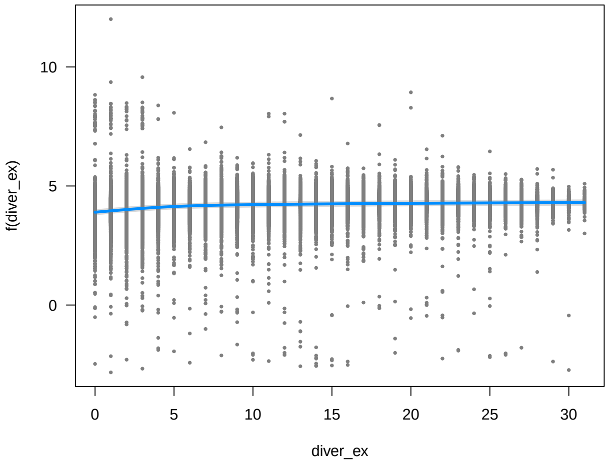


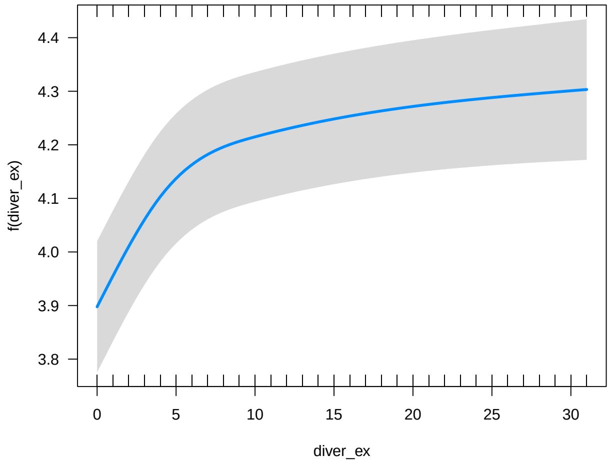
where the peculiar pattern now appears to be captured.

Previous investigations suggested that diver_ex (diver experience, in years) was not a very influential predictor. It is nevertheless interesting to see how this variable influences logCPUE in a conditional, rather than a marginal setting. To do this we add a flexible term in diver_ex to the model and look at its contribution to the model.

```
mod_2 <- update(mod_2, . ~ . + ns(diver_ex, 3))
vdex <- visreg(mod_2, "diver_ex", data = BLdat, plot = FALSE) ## for model inspection
```

We now look at the contribution to the model, with and without the partial residuals.





The figure on the left suggests, correctly, that the contribution is minuscule. Expanding the diagram as in the figure on the right suggests that the contribution is continuously increasing but that increase in experience is most dramatic in the first few years.

We will use `mod_2` as the starting point for extensions involving the environmental variables, to which we next turn.

11.3 Environmental predictors of catch

We now turn to assessing the extent to which adding weather and environmental predictors can be useful in standardisation models such as the one above.

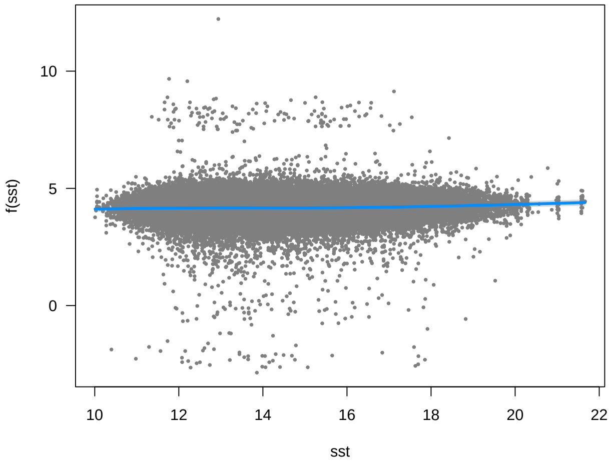
11.3.1 Sea surface temperature

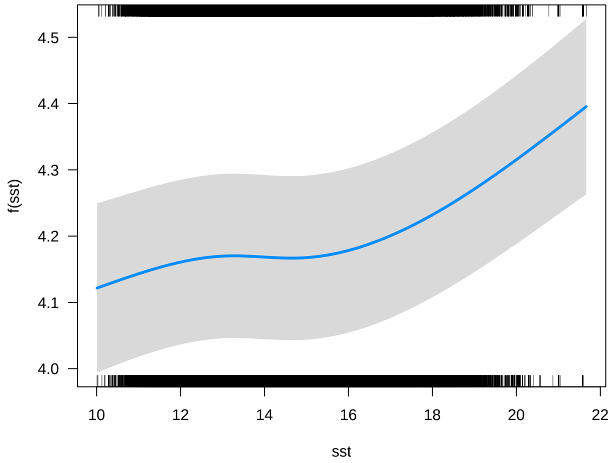
Sea surface temperature, `sst`, is the only environmental variable with no missing values. As there is a large number of missing values in the other environmental predictors, (wave height, power and direction) we consider the case of `sst` separately first, using the complete data set.

Previous marginal views in Part 1 suggested that `sst` was unlikely to be very influential as a predictor. We now look formally at this using the modelling approach, and hence conditionally rather than marginally. We again consider a flexible term, for exploratory purposes

```
mod_2_sst <- update(mod_2, . ~ . + ns(sst, 3))
vsst <- visreg(mod_2_sst, "sst", data = BLdat, plot = FALSE) ## for model inspection
```

Looking at the contribution to the model, again with and without partial residuals, gives the following:





The left hand diagram reinforces the view that the contribution to the model is again essentially negligible. The expanded diagram on the right suggests a general increase in catch with increase in `sst` with a maximum possible gain in catch rate of $\sim 25\%$ over the observed range, but hardly, if ever, achieved.

We conclude that `sst` does not make a useful contribution and omit it from the exploratory study to follow.

11.3.2 Wave height, power and wind direction

At this point we need to use a reduced data set with the cases where these three environmental variables are missing excluded.

```
BLdat_clean <- BLdat %>%
  filter(complete.cases(wave_hght, wave_power, wave_dir, wave_mode)) %>%
  droplevels()
```

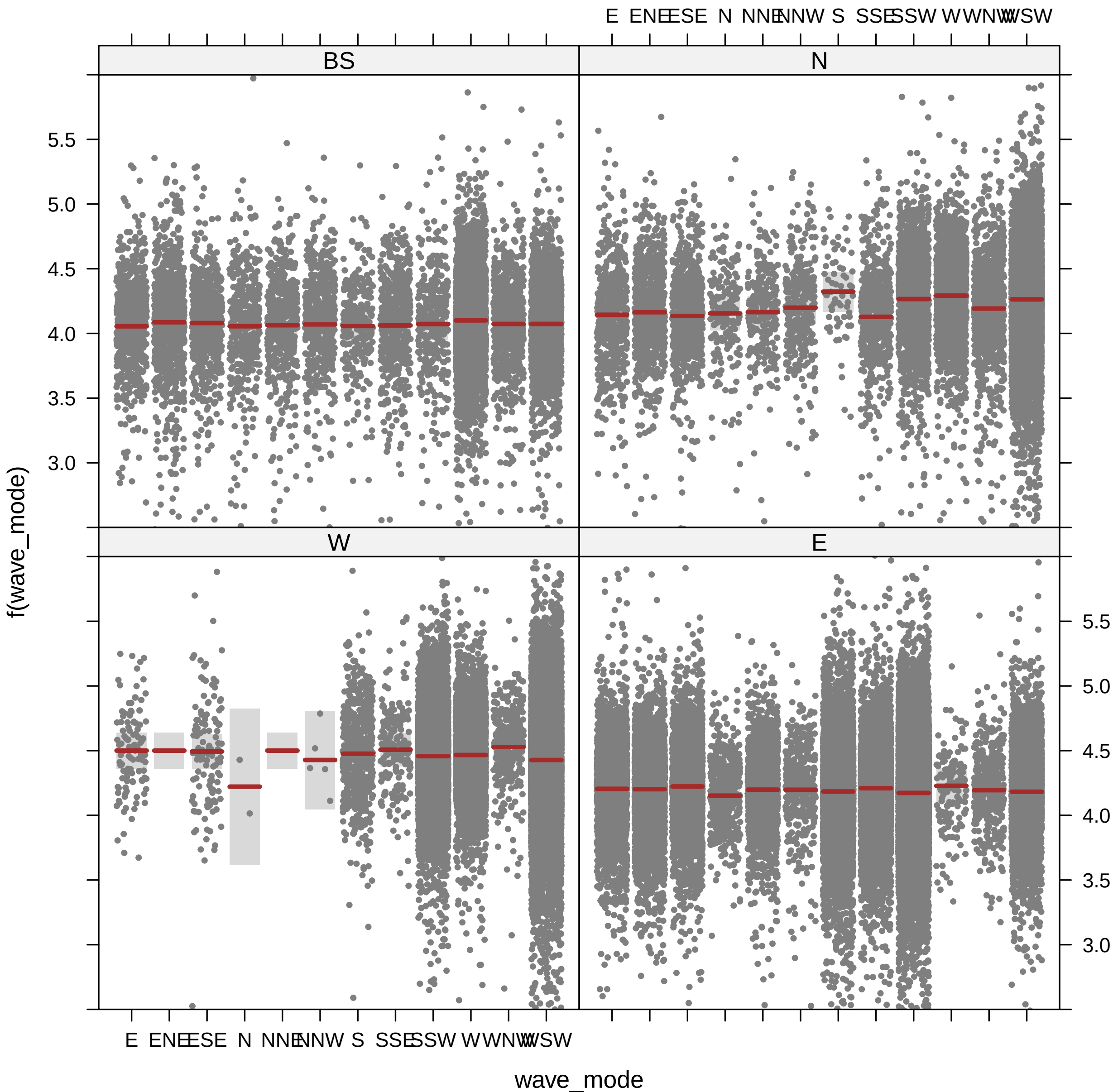
Turning first to wave direction, (a factor), it seems reasonable to consider this variable in relation to `zone`, though note that not all directions were observed in every zone. Here is a table of observed frequencies:

	E	ENE	ESE	N	NNE	NNW	S	SSE	SSW	W	WNW	WSW
BS	773	996	647	399	504	508	248	515	380	3427	736	2370
N	578	826	743	186	241	296	65	642	1910	1953	1078	12974
W	112	0	128	2	0	5	611	183	9009	3673	296	33086
E	3967	5384	4874	538	2095	352	14175	9293	38169	145	381	7004

The exploratory model is fitted as follows:

```
mod_2_wave_mode <- update(mod_2, . ~ . + zone/wave_mode, data = BLdat_clean)
vdir <- visreg(mod_2_wave_mode, "wave_mode", by = "zone", data = BLdat_clean, plot = FALSE)
```

Visualisation of the contributions now has to be conditional on `zone`. The plot below shows the partial residuals as well as the size of the effect. Note that the partial residuals have been clipped to remove the more extreme outliers:



With the partial residuals removed entirely, some variability both within and between zones is evident, though again the size of the effect is essentially negligible:

The major differences between the four panels are due to the `zone` main effect, which in the working assessment model is a component of the `block` main random effect. Within zones the difference in `logCPUE` between compass directions, that is, the `zone:wave_mode` interaction, is tiny.

The case for including a `wave_mode` term of any kind in the assessment model seems weak at best. We will not include it further here.

Turning to wave height and power, there are two issues with these predictors that need to be addressed before their predictive usefulness can be assessed.

- They are, as we have noted, highly correlated and
- There are two observations in the record which have very high leverage for both variables.

The two observations in question for high leverage are the following:

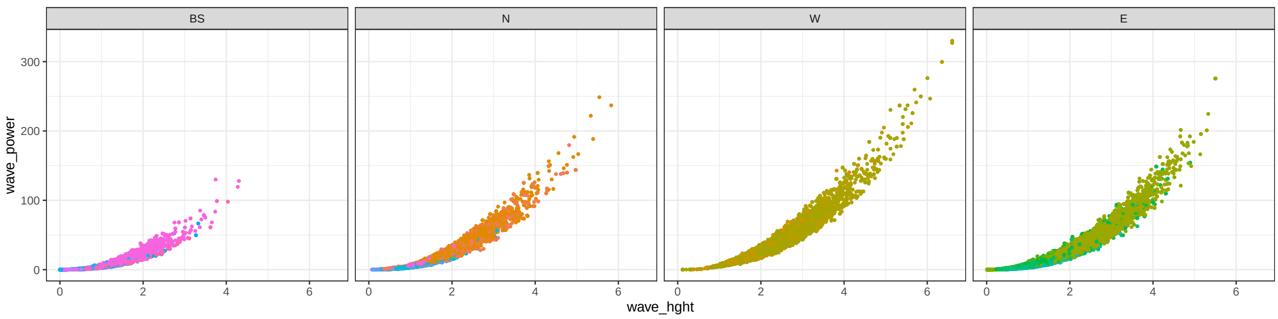
diver_id	plain-date	zone	block	prop-blip	catch	hours	wave_hght	wave_power	wave_mode	sst	logCPUE
5403	2011-07-09	W	11	1	1300	9	9.54	731.61	WSW	13.32	4.97
5403	2011-07-10	W	11	1	950	7	8.40	514.64	WSW	13.16	4.91

These are for the same diver, in the same area, on successive days. For comparison, all other values for `wave_hght` are less than 6.7 and all other values of `wave_power` are less than 330.

The `log(CPUE)` values for these two observations are ≈ 4.9, which is high for the zone but not excessively so. Nevertheless with either of these two predictors in the model, these observations will be given very high leverage and my skew the importance of the variable(s).

For exploratory purposes here we will temporarily cut back these values to 6.6 for `wave_hght` and 330 for `wave_power`. This is simply to gain a better view of the potential usefulness of the variables in any future analysis, *even if these values prove to be genuine*. (Such a move is called 'Winsorisation' in statistics. Any future analysis could compare results with and without Winsorisation.)

With the range contracted, it now becomes clear from the diagram that the highest wave height and power records are in the West and East zones. Colours are different blocks.

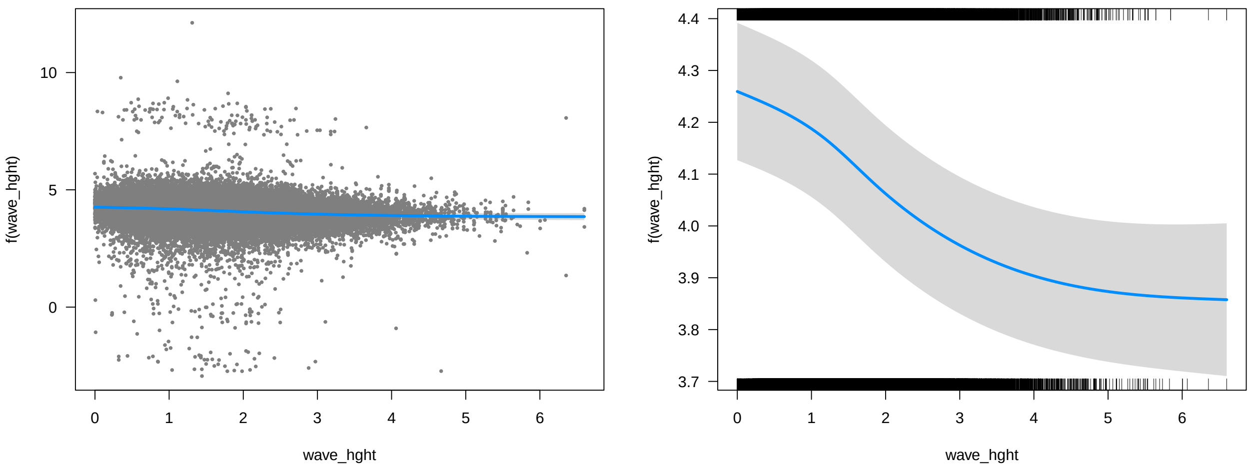


Given the tight curvilinear relationship between `wave_power` and `wave_hght` it is unlikely there will be much to choose between including them in the standardisation model, and no benefit to gain in choosing both together. Hence we look at their possible contributions separately.

First look at wave height:

```
mod_2_wvh <- update(mod_2, . ~ . + ns(wave_hght, 3), data = BLdat_clean)
vish <- visreg(mod_2_wvh, "wave_hght", data = BLdat_clean, plot = FALSE) ## for model inspecti
```

The contribution to the model due to this additional cubic polynomial term, first with partial residuals and then without, leads to the following:

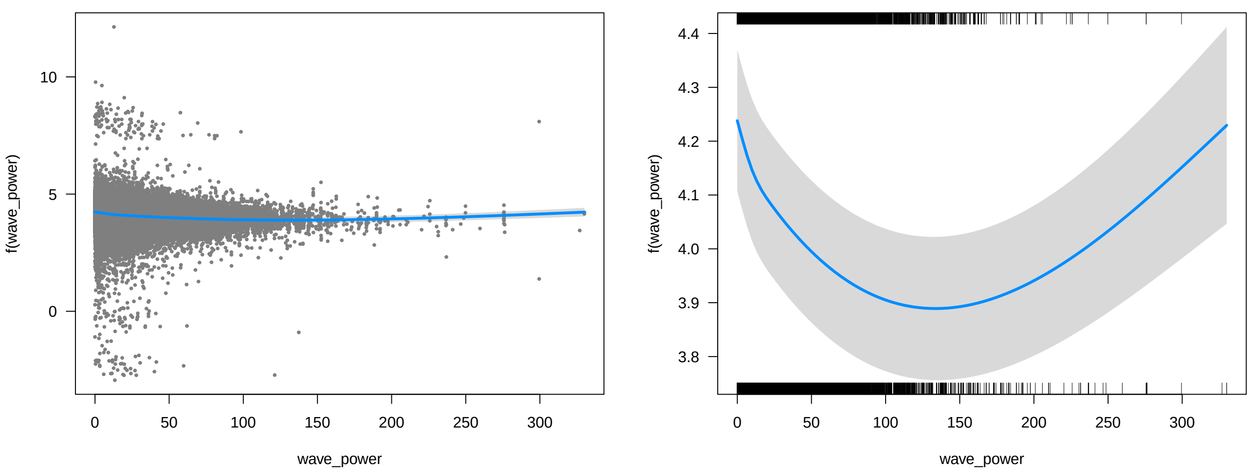


Significance is a poor guide again, here. Instead the range and shape of the partial contribution is more informative. The shape appears somewhat curious with an initial decrease, followed by a leveling out. In any case the range of the component is again very small, though marginally larger than, for example, that of `sst` as shown previously.

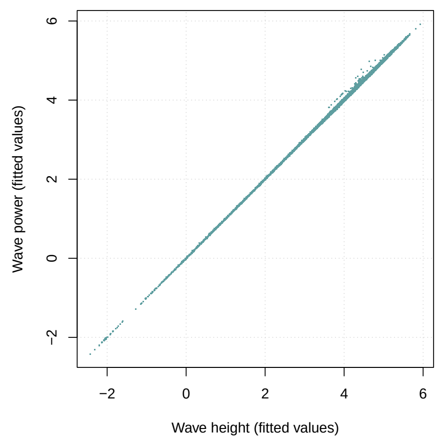
Turning to wave power, we fit the same kind of model:

```
mod_2_wvp <- update(mod_2, . ~ . + ns(wave_power, 3), data = BLdat_clean)
visp <- visreg(mod_2_wvp, "wave_power", data = BLdat_clean, plot = FALSE) ## for model inspect
```

The contribution to the model, in the same form, is shown below:



In this case there is a pronounced final upswing, but the range of the component is again tiny, and much the same as it was for wave height. The difference in shapes between the two components may be explained, in part, by the curved nature of the tight connection between the predictors themselves, as noted several times. In any case, both model lead to very similar fitted values, which we can show in a simple plot:



This further emphasises that either wave height or wave power can function as an effective surrogate for the other.

11.4 Final comments

11.4.1 Environmental variables

In this modelling approach we have tried to gain some insight on the question of whether or not environmental variables are likely to be useful in catch rate standardisation. The clear answer is that there is little if anything to gain by including them. Although this conclusion has been reached using a particular modelling strategy for standardisation, we strongly suspect the same outcome will apply whatever method is used.

The most likely reason why environmental variables cannot be shown to be important in this context is the observational nature of the data itself: rather than data from a designed, balanced experiment with randomization and controls, we only have realised data from the fishery in action, where the goal is to maximise outcomes, not to test the effect of the environmental variables. Decisions when and where to fish are made taking account of environmental conditions and presumably carefully optimising with respect to them. This leads to inevitable confounding when trying to tease out the subtle effects of variable environmental conditions: the variation is strongly curtailed.

11.4.2 Side issues

We also covered two side issues that may be of some interest.

The first of these was the intriguing possibility that some attenuated version of `hours` fished might offer a better basis for (nominal) effort. Specifically a power such as $H^{0.9}$ might be a better measure of effective effort, particularly for the very long episodes of daily fishing time per diver. We leave this without further comment.

The second side issue was a simple demonstration that the choice between fixed and random *main* effects may not be crucial in practice. We showed this for the case of `month`, `year` and `block` main effects only. Where it is likely to be important, however, is where you have large numbers of effects, as in the case of `diver_id`, and where the information basis is patchy at best, as in the case of interactions between `block`, 'year`and`month' in our model. In these latter cases the integrative power of random effect models becomes crucial, avoiding the parameter explosion and resulting instability of their fixed effect counterparts. This issue could be further explored, but it would not be a brief or simple exercise to do well. There are strong theoretical reasons behind the choice of random effects in these circumstances, however.

12 Climate Data Part 3: Standardisation Results

12.1 Preamble

This is Part 3 of the triptych. Part 1 was mainly concerned with visualising and exploring the climate data itself. Part 2 looked at the possibilities for including climate data, such as we have, in a standardisation model. The overall conclusion was that what we had at that stage did not appear to offer any useful enhancements, at least not within a notional global standardisation model. The present Part 3 is a follow-up on that notional standardisation model itself, with a few tweaks and comparisons with potential alternatives.

The data we use is the same as the extended version produced and used in Part 1, but *the short record form 2023 has been excluded* to avoid unnecessary complications.

The records by zone and year are as follows:

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Bass Strait	–	–	162	135	142	109	44	60	71	107
North	–	–	909	643	359	490	386	364	328	437
East	–	–	2920	3489	3585	3513	3140	3626	4152	4166
West	–	–	1179	1012	806	711	633	908	685	883

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Bass Strait	207	377	336	507	462	481	627	212	543	552
North	517	869	980	754	795	786	814	771	761	874
East	3843	4426	4054	3863	3383	2758	2393	2315	2507	2454
West	1501	1537	1633	1841	1981	1933	2025	2112	2110	1966

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Bass Strait	459	587	702	657	734	585	456	449	424	431
North	947	1066	1024	924	998	938	920	870	527	552
East	3327	2817	2642	2426	2566	2462	2706	2161	1359	1084
West	1938	1790	2028	2038	2007	2060	1671	1613	1525	1776

	2020	2021	2022
Bass Strait	348	292	255
North	356	302	230
East	792	732	716
West	1359	951	802

12.2 Standardisation strategies

In this section we present our current thinking on alternative approaches to the standardisation issue, initially for the Blacklip Abalone only.

As we understand it, the current standardisation method is a 'block by block' approach, where a model *selection* process, as well as model fitting takes place in every block separately. The block-by-block approach:

- Is parameter rich, implicitly fitting a separate set of parameters, including variance parameters, for all 57 blocks. The effect of the separate, though structured, model *selection* aspect of this process is difficult to assess, but it adds a level of complication when simplicity is a virtue.
- In working separately on each block, it potentially ignores any advantages that taking into account the *context* might have, particularly for blocks with scant data.

A 'block by block' approach represents one extreme case for taking context information into account when assessing individual blocks: i.e. use none at all: 'Every block is special'. And, of course, to some extent every block is. Nevertheless we suggest that it is worthwhile to investigate alternative approaches which do take into account context information by modelling at different spatial scales. We do so here using two modelling strategies:

1. A global model using a single model fitted to data from the entire Blacklip fishery, and
2. A zone by zone strategy where separate models of the same form are fitted to the data from each of the zones and catch standardisation for any block uses the model pertinent to the zone in which it resides.

With regard to 2., we note that according to the data, we have some blocks appear to be split between two zones. These blocks, and the zones to which they belong in parentheses, are: 06 (N, W), 13 (W, E), 31 (N, E), 49 (BS, N). In these cases the blocks in question will be included in *both* zone models, but only *once* in the global model, of course.

Note also that this is primarily a feasibility study. We only use zones as a convenient (semi-)partition of the blocks into a workable number of spatially compact groups; alternative partitions may well be more appropriate if this strategy is ultimately adopted.

12.3 Fixed and random effects

Before proceeding to the model details, we make some points about the choice of fixed and random effects, as previously argued in the *Exegesis* document.

- From a practical point of view, in standardisation as here the distinction between fixed and random terms is slight. The main difference in this context is the way in which they are estimated. Random terms are estimated with *model-based penalisation*, which makes them very similar in essence to smooth terms in generalised additive models, which are estimated with *data-based penalisation* using generalised cross-validation.
- Fixed effect terms, after estimation, have the logical status of parameter estimates. Random effect terms have a logical status more akin to *residuals* and have *variance component* estimates associated with them, as well as the BLUP estimates, which can be likened to parameters in some qualified sense.
- Prediction with mixed effect models can take several forms. In any prediction
  - All fixed effect terms must be included in any prediction.
  - Random effect terms may be included or excluded, depending on purpose. If all random effect terms are excluded, the prediction is said to be *marginal* to the random terms.
  - If any random effect is included, the prediction is said to be *conditional* on it.
    In this case any excluded random terms will contribute their variance components to the overall error when considering uncertainty.
  - In predictions, excluding a random effect amounts to assuming it has a zero contribution to the estimate of the mean.

12.3.1 A global model

For the entire fishery model, (1. above), after some exploration of alternatives the form to which we were finally led is as follows:

```
mod_1 <- lmer(logCPUE ~ poly(asin(sqrt(1-propblip)), 4) + ## fixed
                two_divers +  ns(diver_ex, 3) +           ## fixed
                year +                                    ## fixed
                (1|diver_id) + (1|month/year/block) +     ## _within_ years
                (1|block/year),                           ## _betweein_ years
              control = lmerControl(optimizer = "bobyqa"),
              data = BLdat)
```

We work on the assumption that for each block standardisation is required on an *annual* basis; the catch and effort data is supplied on a daily basis, so some averaging is required to arrive at an annual standardised CPUE. This averaging is achieved in predictions, (or ‘backcasts’ as they are known), in the model above by

- retaining terms in prediction that are *constant* within years, (which include all the fixed effects, of course) and
- using the random terms that *vary* within years effectively as components of the overall error, but estimated in particular strata.

The random effect terms that are constant within years are all included in `(1|block/year)`. The other random terms account for within-year variability at several levels, including the lowest, the per observation level, implicitly.

Note that the diver random effect, (1|diver_id), has both between and within year components. This is an unavoidable weakness of the design: some information on the stock status is inevitably lost through confounding with diver. Hopefully this loss is small when there are several divers.

Some comments on the role of the other terms are as follows.

- ns(diver_ex, 3) has been included as a fixed effect to allow some account to be taken of "diver experience", as noted in Part 2. This has the effect of reducing the variance component associated with the diver_id random effect, which reduces the uncertainty, but also of requiring some specific value to be given for the fixed effect term, as it is necessarily included in the predictions.
- The main effect for year has been modelled as fixed, though it could have been a random term without appreciable difference to the result.
- The random terms (1|month/year/block) account for a month random main effect, and a month:year random interaction and a month:year:block three-way interaction. There is little choice but to make these main effects and interactions random, but the month main effect could have been modelled as fixed. Here the advantage of modelling it as random is that it may be excluded from the predictions, thus facilitating averaging within years as noted above. Had it been modelled as fixed, however, predictions would have had to be made at some specific month of the year. Choosing such a month seems arbitrary, though possibly not an insurmountable obstacle.
- The three-way term (1|block:(year:month)) contained in the above, is present to 'iron out' differences within years across months at the block level. As it varies within year, it is excluded from the predictions so it contributes to the overall error.
  Note that for the data set we have this term alone leads to 9726 BLUPs. For this reason alone it needs to be a random term, as estimating so many fixed effect parameters is both practically difficult and statistically risky. This is one of the big advantages of modern random effect model fitting technology, that such large terms can be handled reasonably comfortably, but only as random effects.
- The term (1|block/year) is the work horse for the predictions. In the block main effect it estimates a static difference between the blocks themselves across the fishery and in the block:year interaction it measures, deviations from this block average across the years.
  The combination of the year main effect, the block main effect and the block:year interaction will be used to reflect how this block CPUE (in the log scale) has fluctuated, on average, across years.

12.3.2 Per zone models

To give some insight on the role of spatial scale in the standardisation, as indicated above, we fitted separate models to the blocks in the four administrative zones, and used the appropriate zone model to provide the standardisation for each block. This admittedly fairly crude device is meant to fall somewhere between an all-fishery global scale and the single block local scale.

To be specific, the same model was fitted to each zone, as in the code here:

```
mod_1_BS <- update(mod_1, subset = zone == "BS")
mod_1_N  <- update(mod_1, subset = zone == "N")
mod_1_W  <- update(mod_1, subset = zone == "W")
mod_1_E  <- update(mod_1, subset = zone == "E")
```

We also noted that four blocks apparently straddle two adjacent zones. These were: 06 (N, W), 13 (W, E), 31 (N, E), 49 (BS, N). In these cases their results will also be split across the two zonal models. Predictions (standardisations) will be given for those blocks *twice* using each of the zonal models to which they contribute.

12.4 Crude CPUE

To provide the models with some reference point grounded more directly with the data itself, we simply use a crude CPUE measure for each block, for each year it has been fished. That is, for any given block and year y, the crude CPUE is, using an obvious notation:

$$\mathsf{CPUE}_y = \frac{\sum \mathsf{catch}_y}{\sum \mathsf{hours}_y}$$

and, of course, the crude log CPUE is just its natural logarithm.

This crude measure ignores all of the identified extraneous influences of which the model approach takes account, but it does at least give an easily understood reference point in the data.

12.5 Results

12.5.1 Setting fixed effect values

Standardisation will be effected by prediction methods using a set of specified values for the fixed effects, (apart from `year` itself), and omitting the random effects on which the predictions are not conditional, namely the `diver_id` and `month` main effect and interactions.

The values used for the fixed effects are

- `propblip = 1`, i.e. the block is fully a Blacklip fishing area,
- `diver_ex = 6`, i.e. we are dealing with a 6 year experience diver, (which is the median value), and
- `two_divers = 0`, i.e. only one diver is involved, again the most common value.

The `year` fixed effect will cover 1992 to 2019 inclusive, the 28 years for which we have full season results.

12.5.2 The log and natural scale: bias corrections

We also give parallel results in the log and in the natural scale. Modelling is done in the log scale, so standardisation in the natural scale will require a bias correction if the predictions are to be of the mean, rather than the median in the natural scale (as is appropriate). The simple bias corrections are all based on the mean of the lognormal distribution, which is given by

$$\log Y \sim \mathsf{N}(\mu, \sigma^2) \implies \mathsf{E}[Y] = \exp\left(\mu + \tfrac{1}{2}\sigma^2\right)$$

So the back-transformed values need to be increased by the factor exp$(\frac{1}{2}\sigma^2)$. The appropriate variance, σ^2 to use for random effect models such as ours is not immediately clear, but we will use the sum of the variance components for random effects omitted from the prediction, plus the residual variance itself. In symbols

$$\sigma^2 = \sigma^2_{\texttt{diver\_id}} + \sigma^2_{\texttt{month}} + \sigma^2_{\texttt{year:month}} + \sigma^2_{\texttt{block:(year:month)}} + \sigma^2_{\texttt{residual}}$$

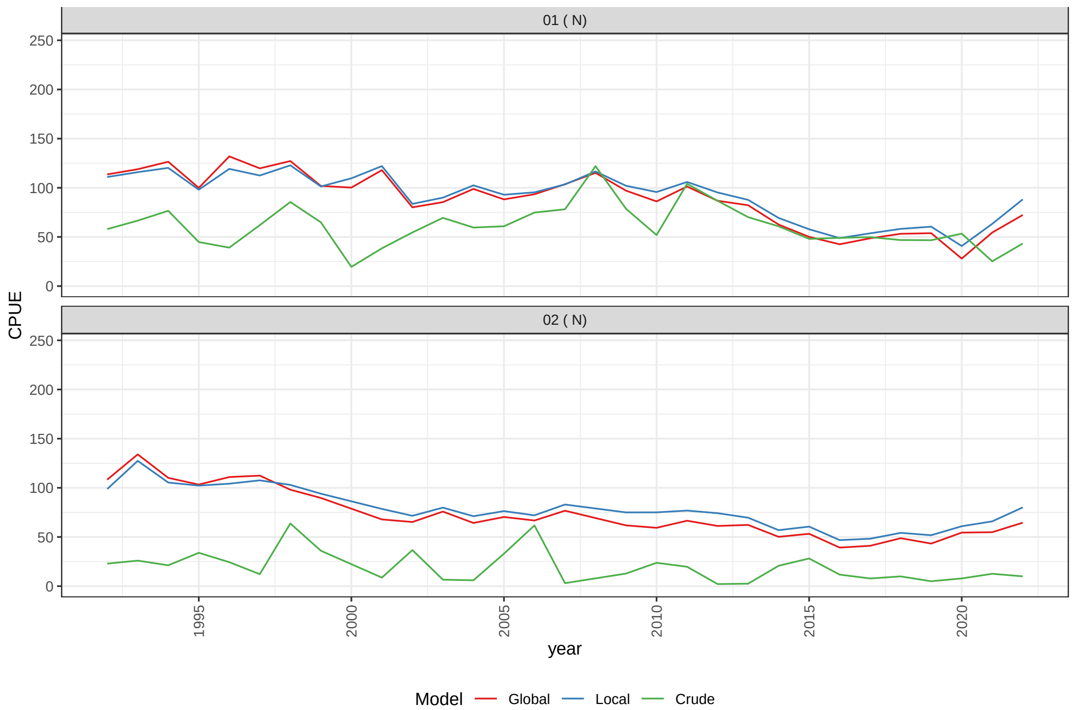
since each component is part of the overall error, independently, when predictions are conditional *only* on the `block/year` random effects.

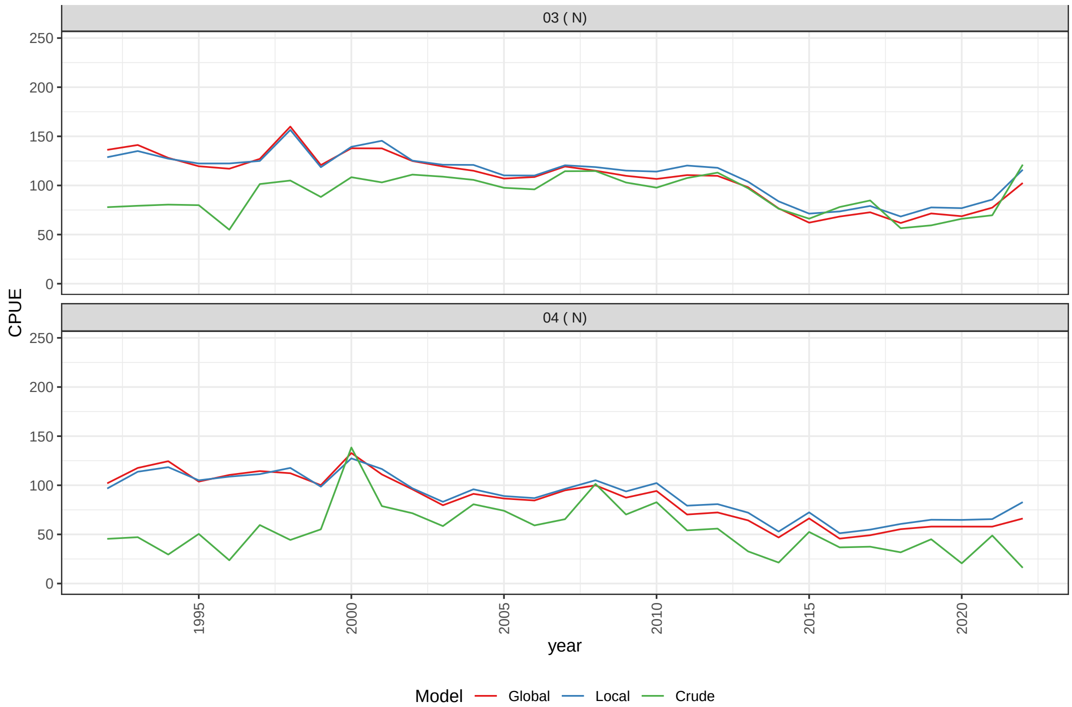
Notice that this also raises another source of difference between the three spatial scales, as the variance (component) estimates are based on different data sets in each zone. (We contend that this also poses a problem for any block by block analysis, but will not pursue that issue here.)

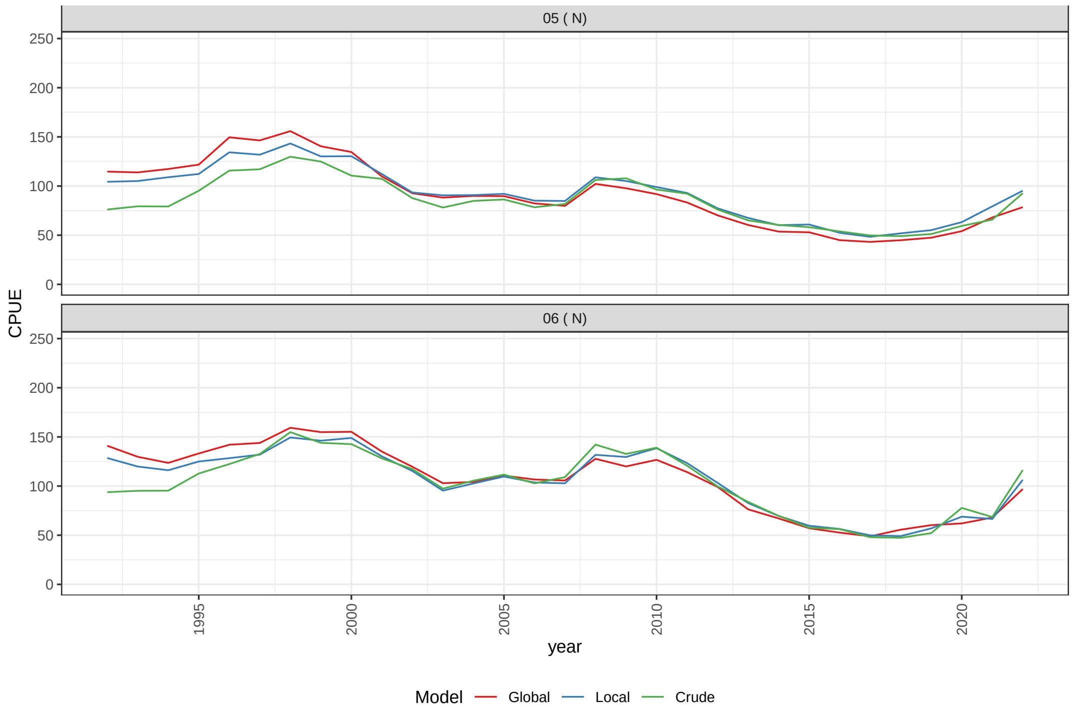
12.5.3 Block by block comparisons

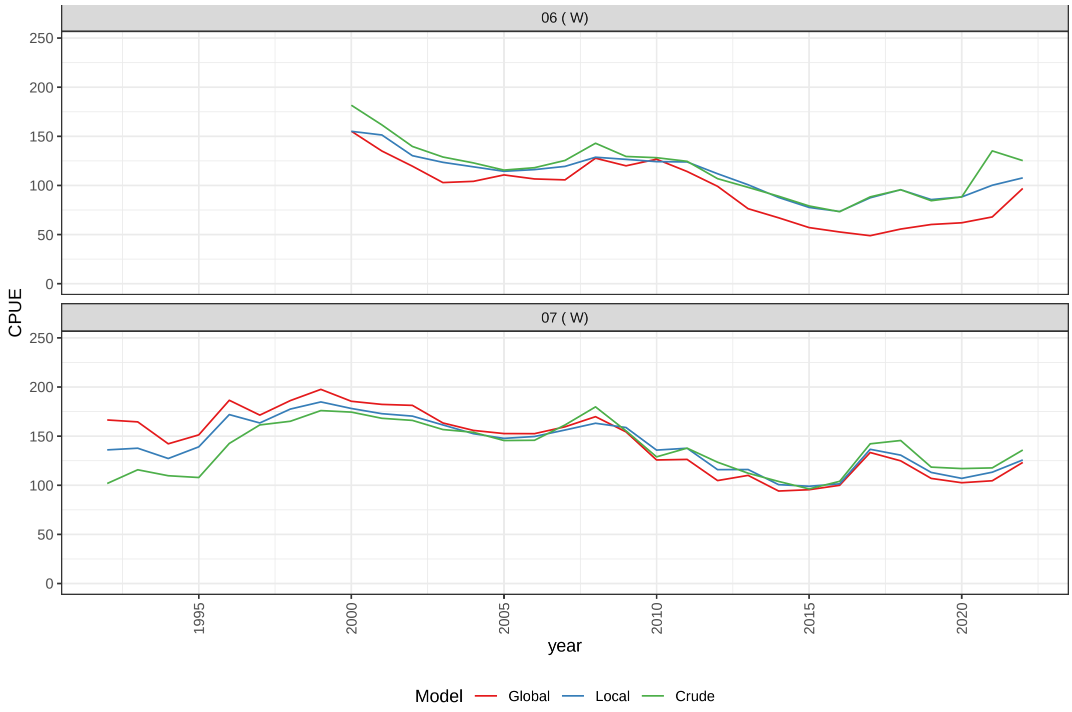
The diagrams following show, block by block, the results of the two standardisation methods described above, along with the crude CPUE. The diagram on the left is the log(CPUE) and that on the right is the corresponding CPUE, bias corrected.

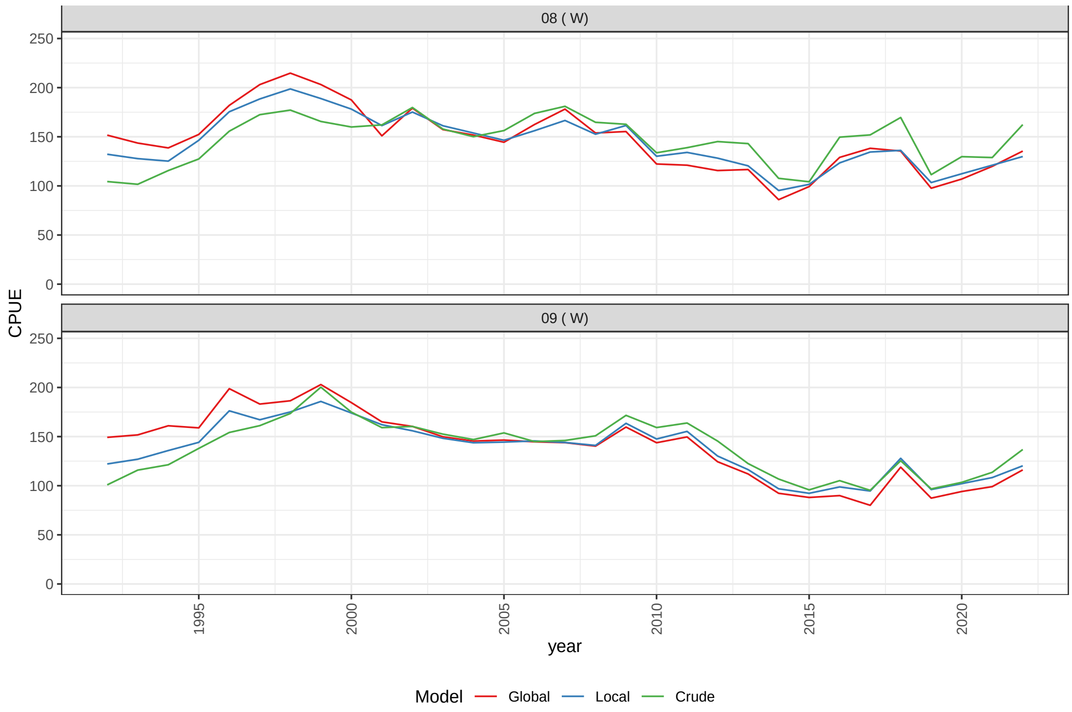
The blocks are listed in order of their numerical labelling, and the header of each panel shows the block and the zone to which it belongs. Recall that some blocks are split between zones, namely 06 (N, W), 13 (W, E), 31 (N, E), 49 (BS, N), and are thus listed twice.

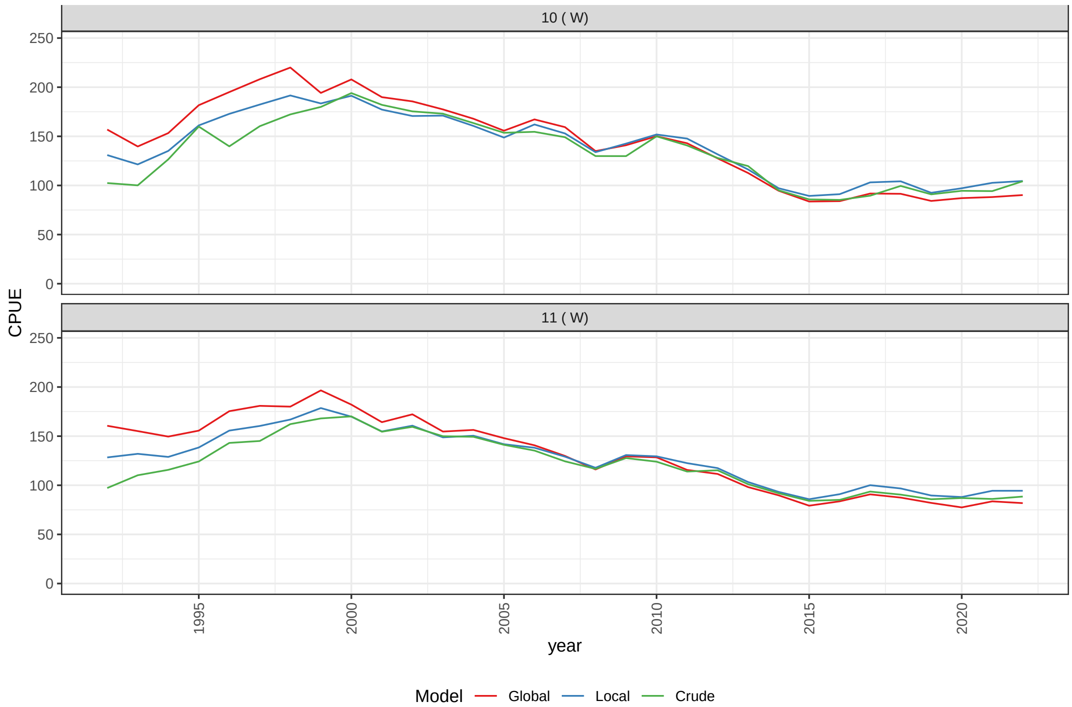


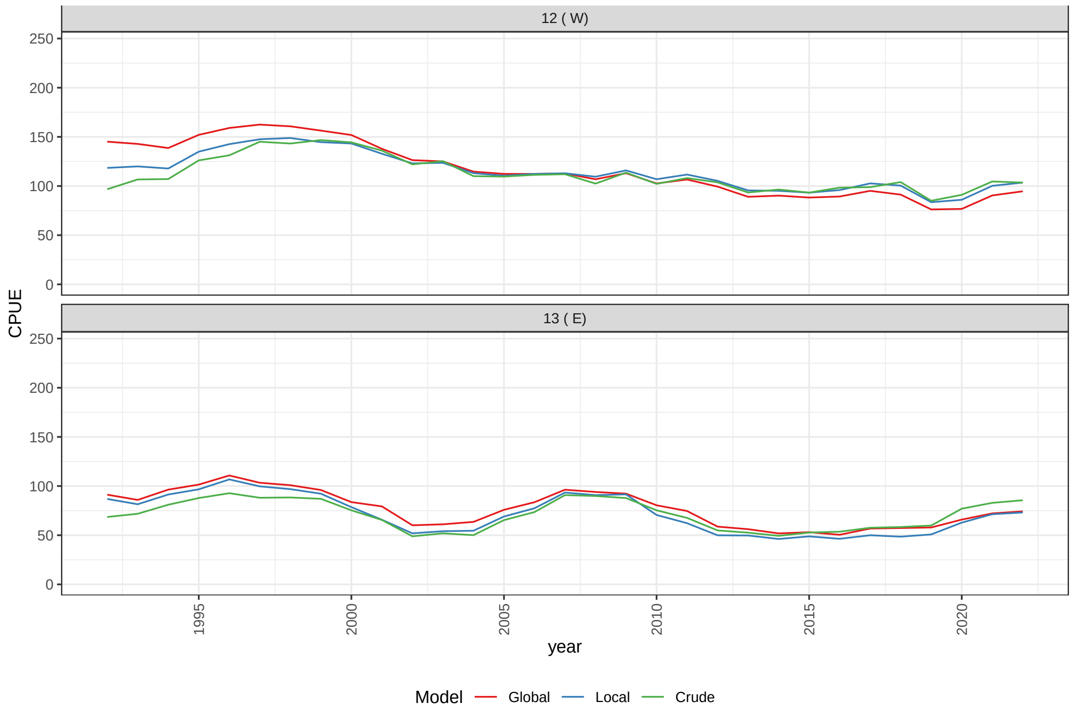

03 (N)
04 (N)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

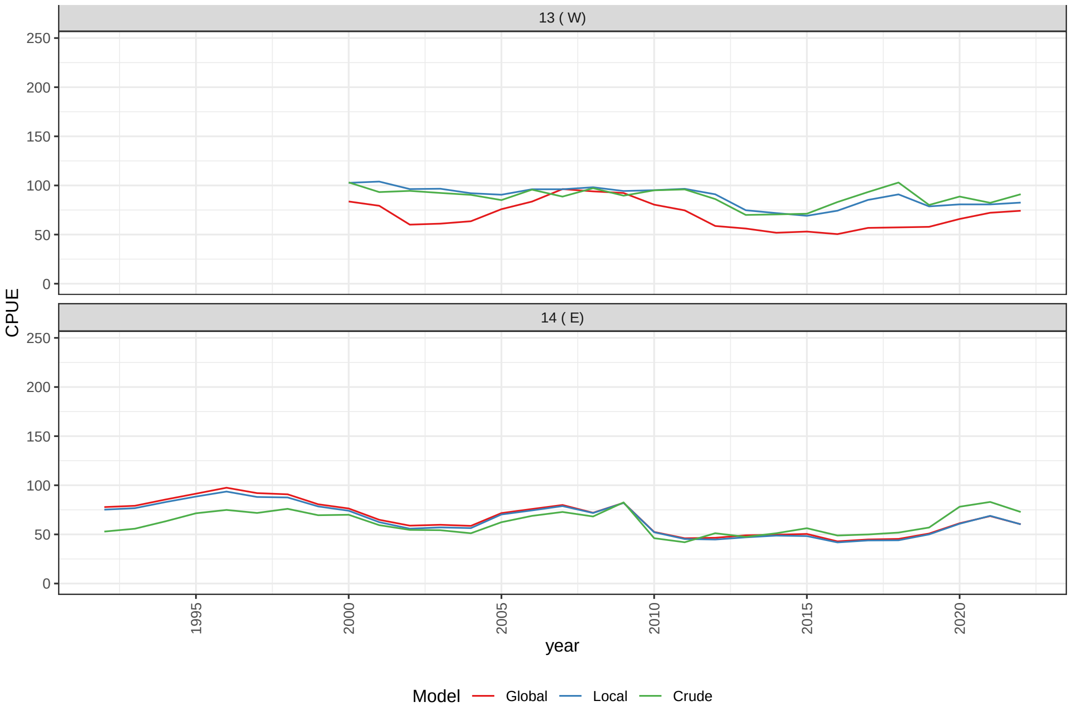

05 (N)
06 (N)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

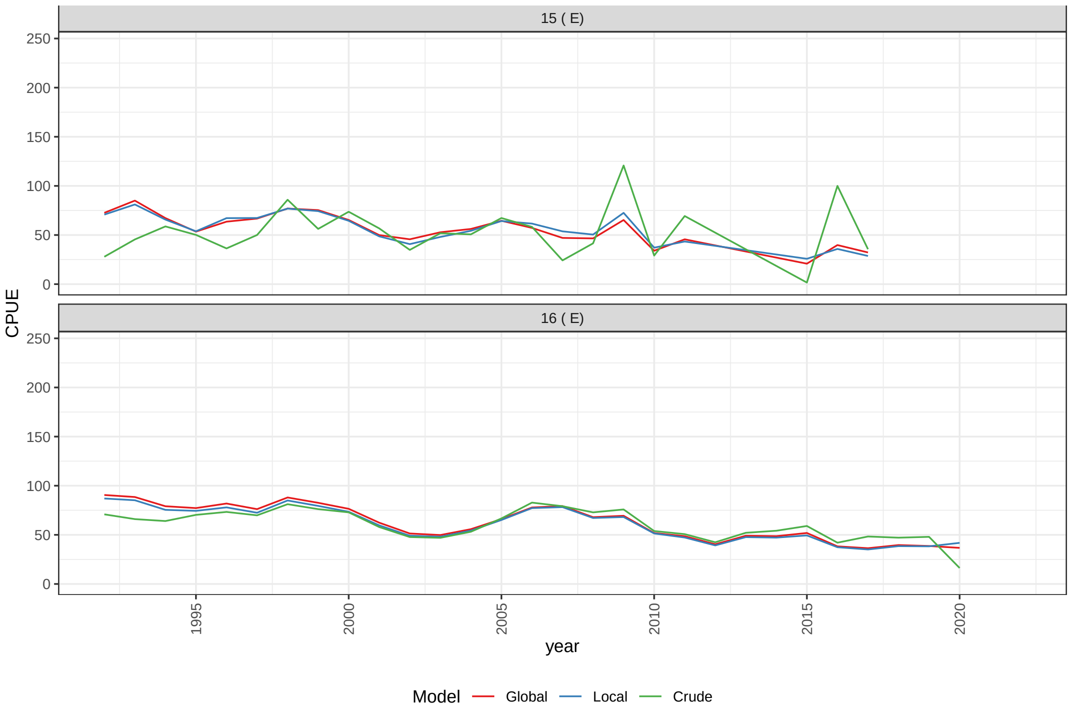

06 (W)
07 (W)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

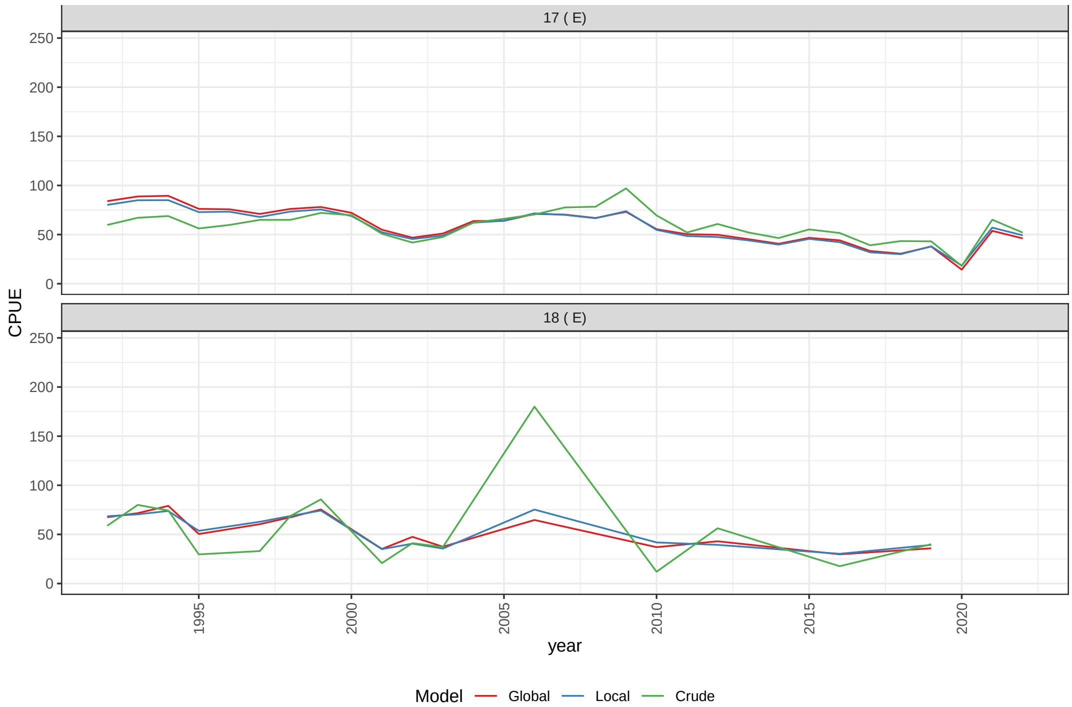

08 (W)
09 (W)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

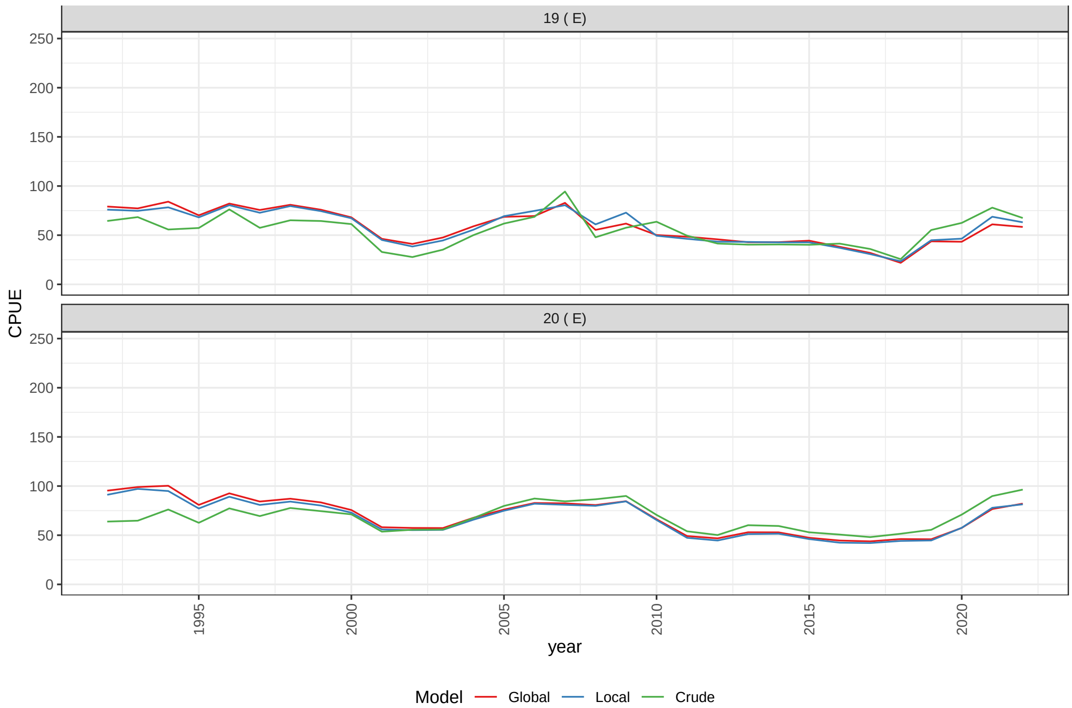

10 (W)
11 (W)
CPUE
year
Model
Global
Local
Crude

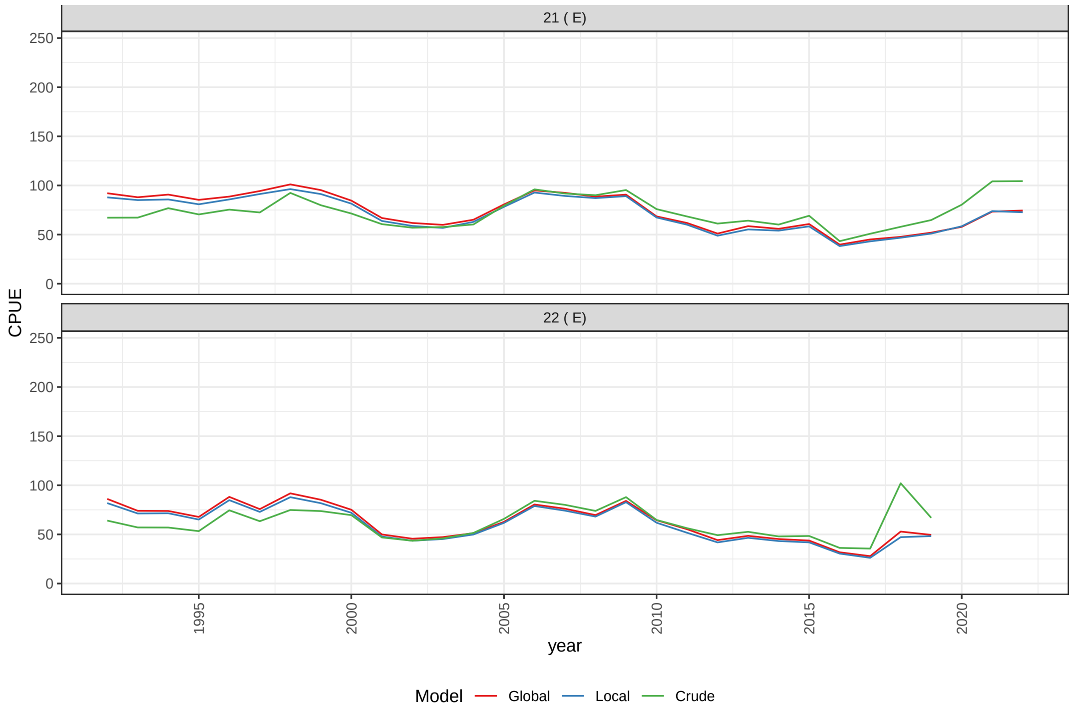

12 (W)
13 (E)
CPUE
year
Model
Global
Local
Crude

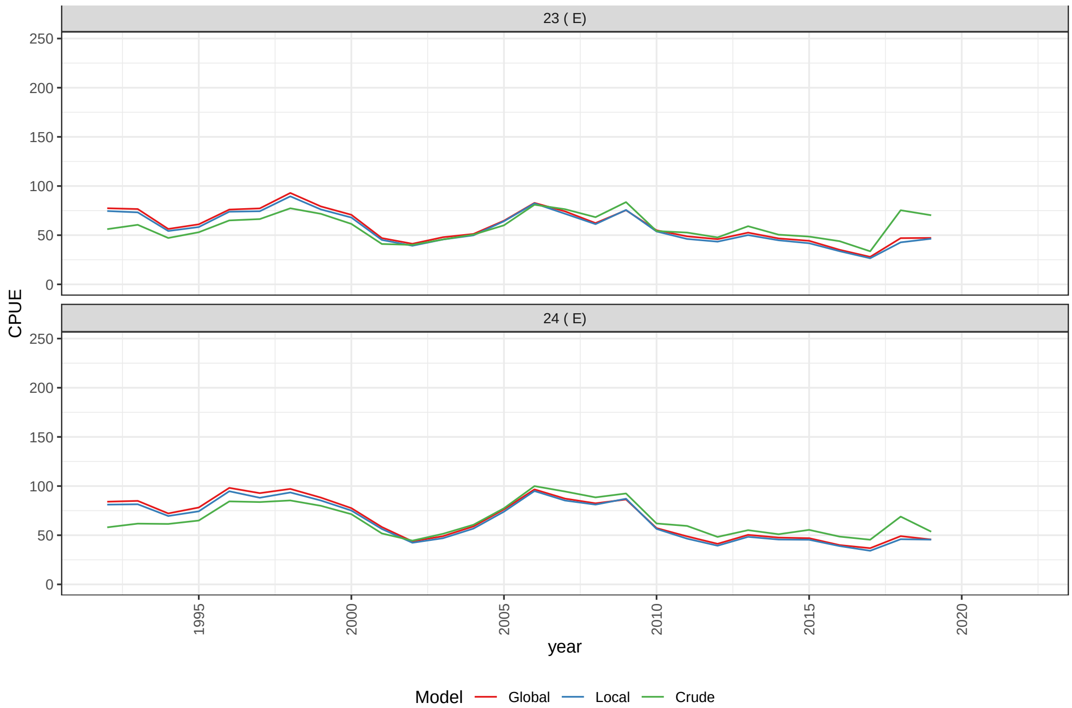

13 (W)
14 (E)
CPUE
0
50
100
150
200
250
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

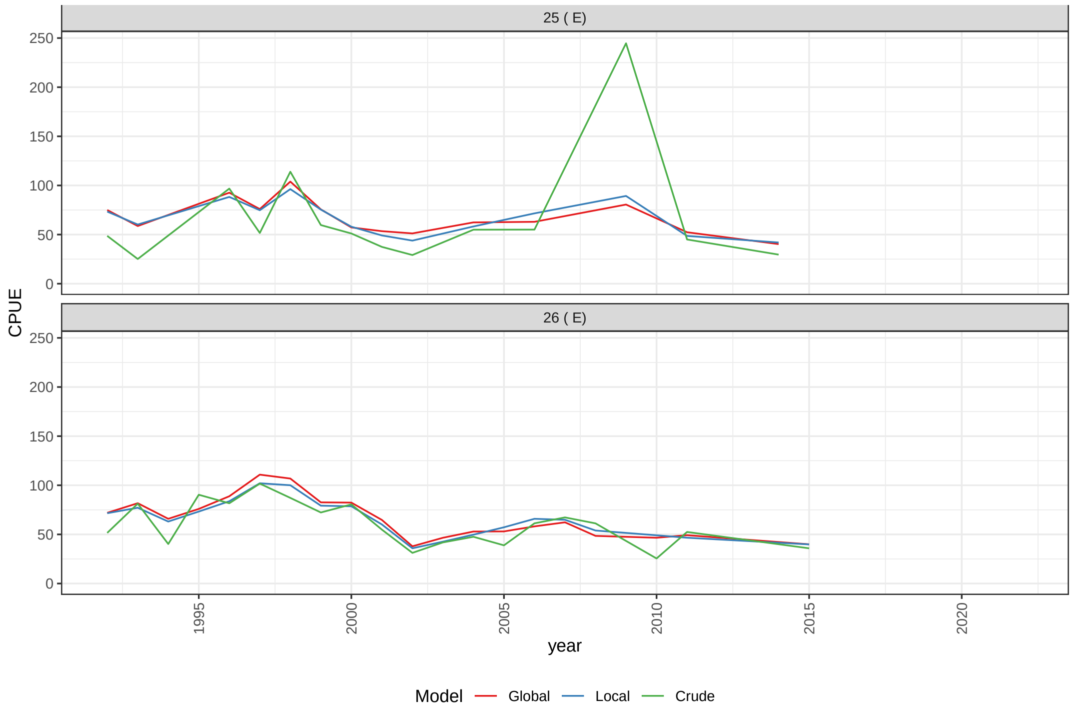

15 (E)
16 (E)
CPUE
0
50
100
150
200
250
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

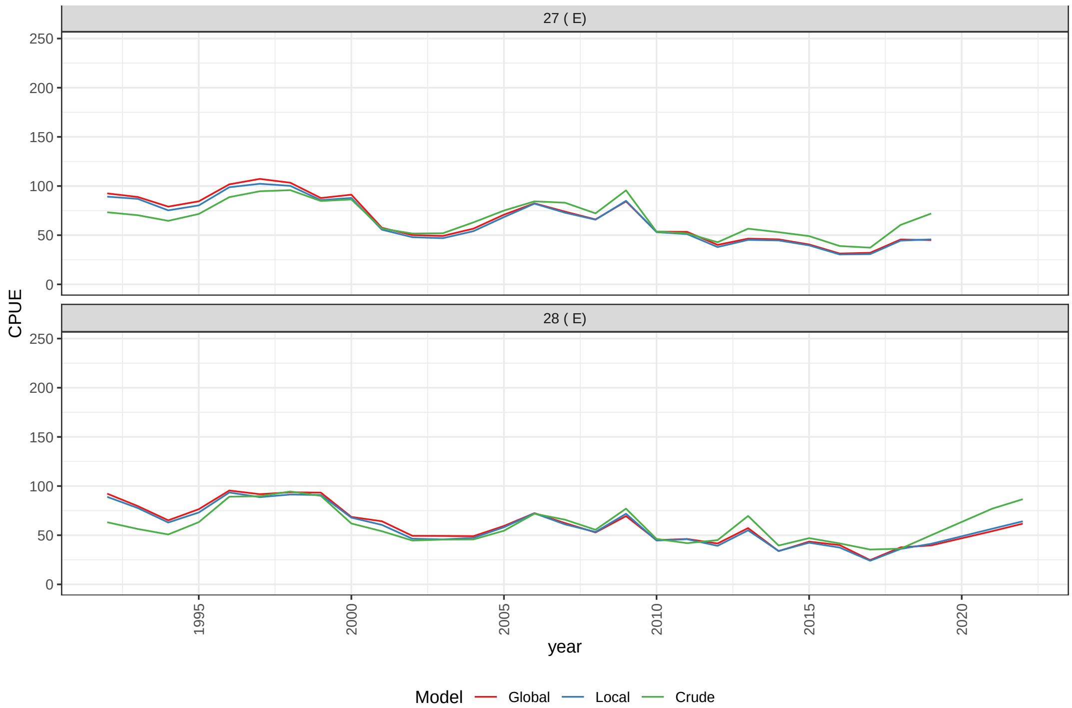

17 (E)
18 (E)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model Global Local Crude

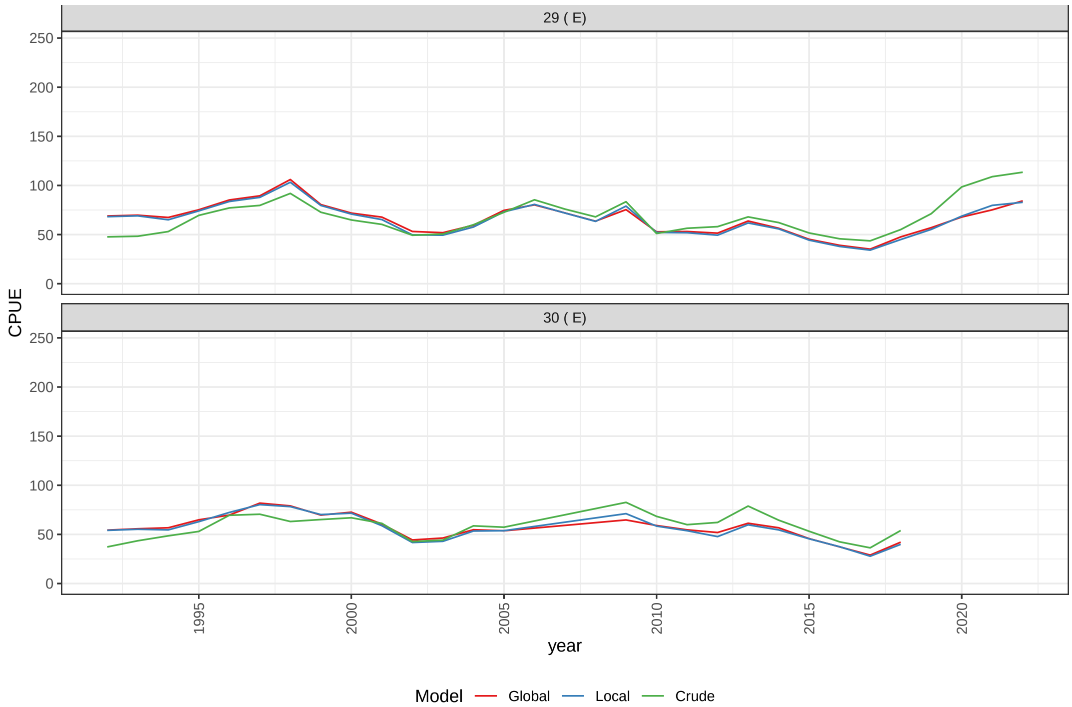

19 (E)
20 (E)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model Global Local Crude

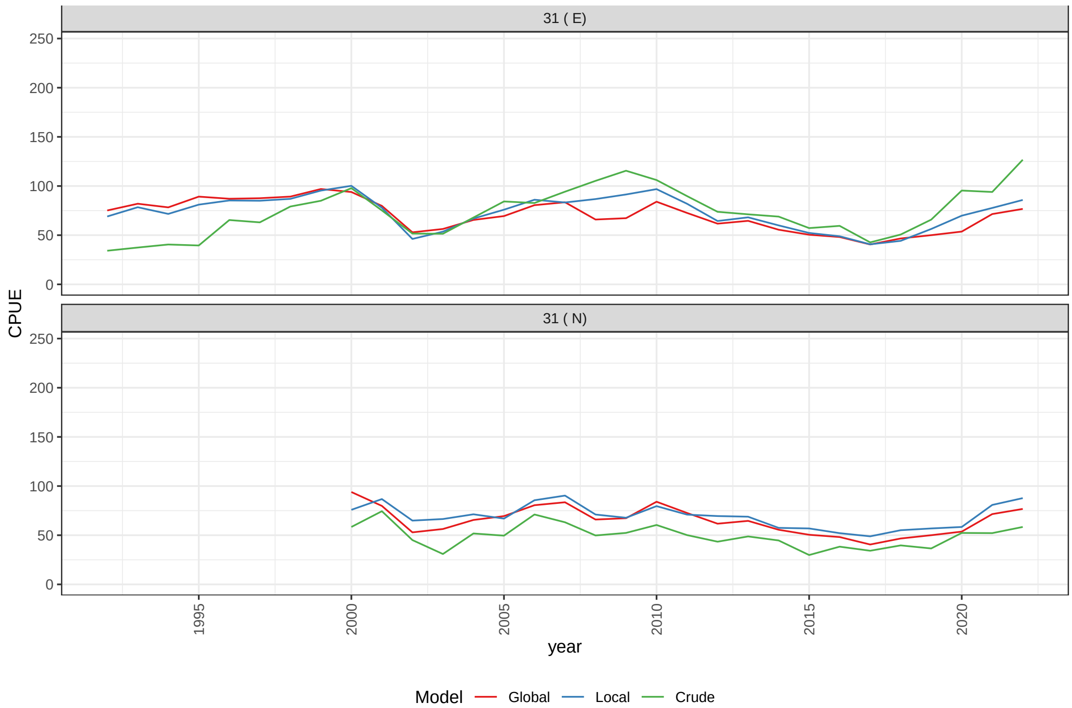

21 (E)
22 (E)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

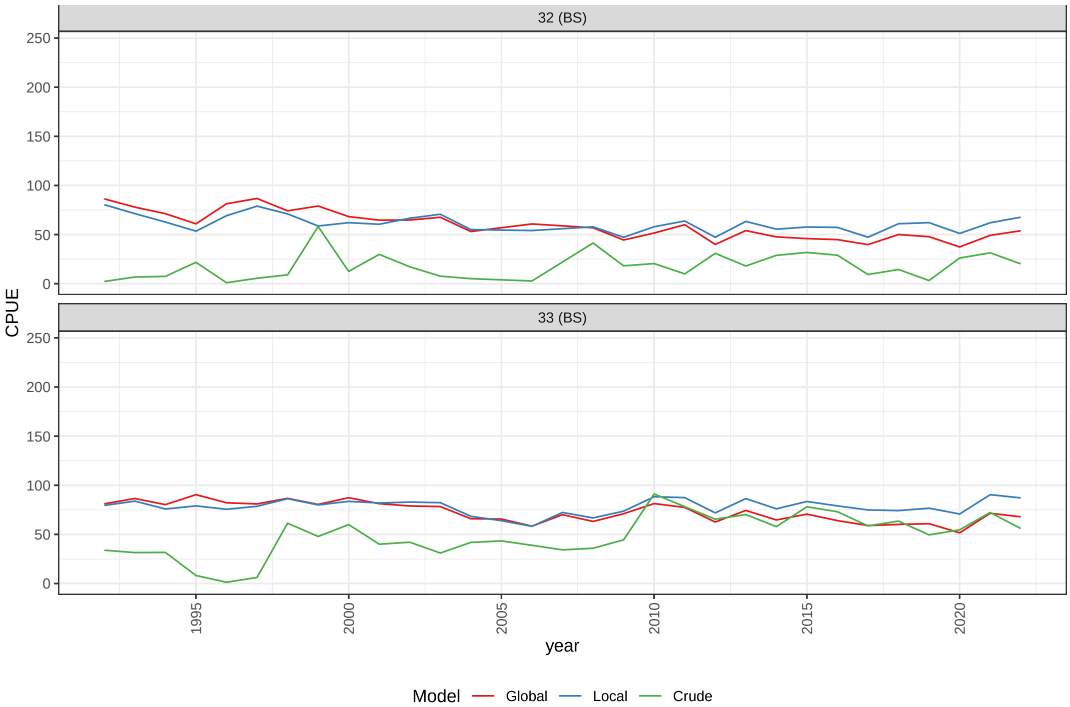

23 (E)
24 (E)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude


25 (E)
26 (E)
CPUE
0
50
100
150
200
250
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

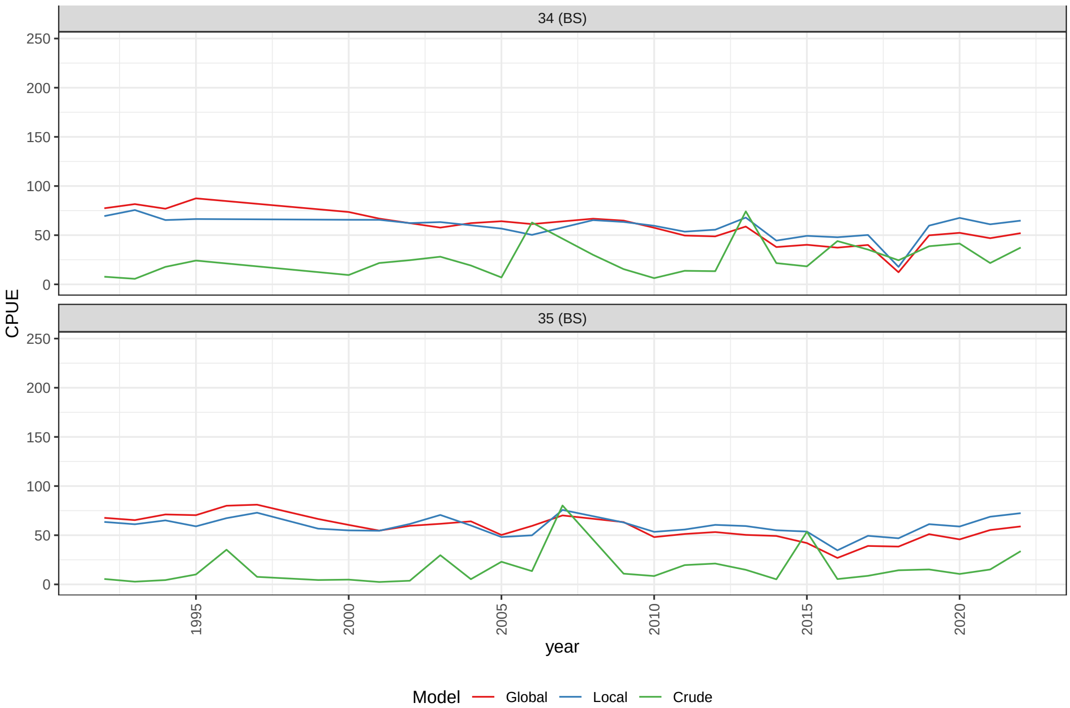

27 (E)
28 (E)
CPUE
0
50
100
150
200
250
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude


29 (E)
30 (E)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

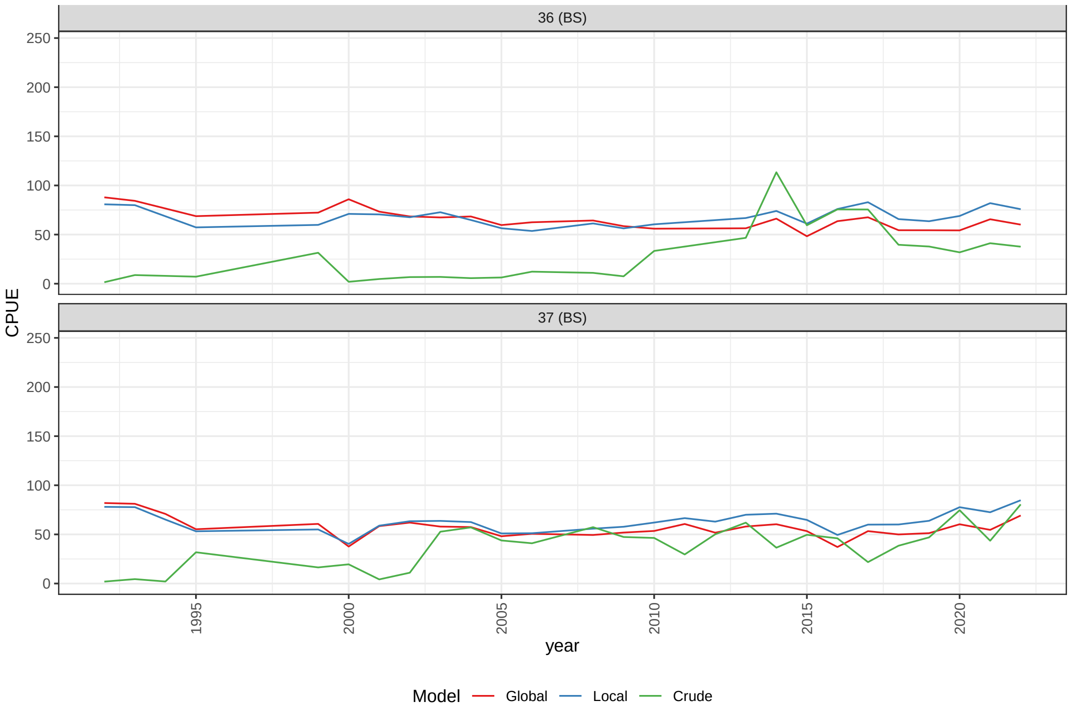

31 (E)
31 (N)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

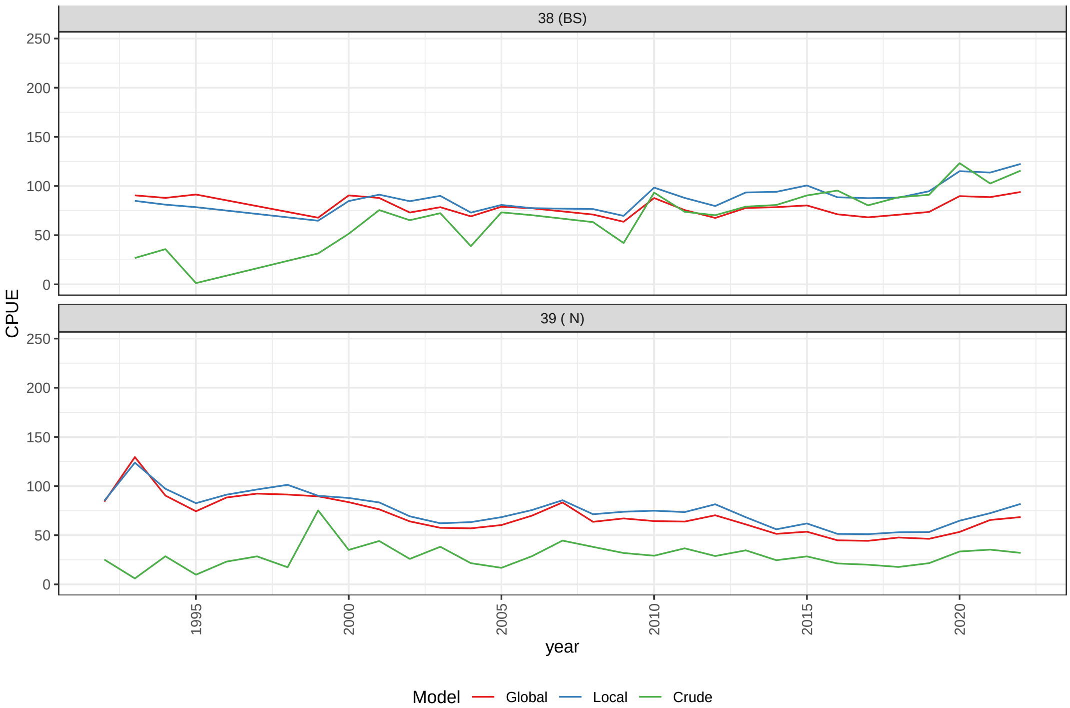


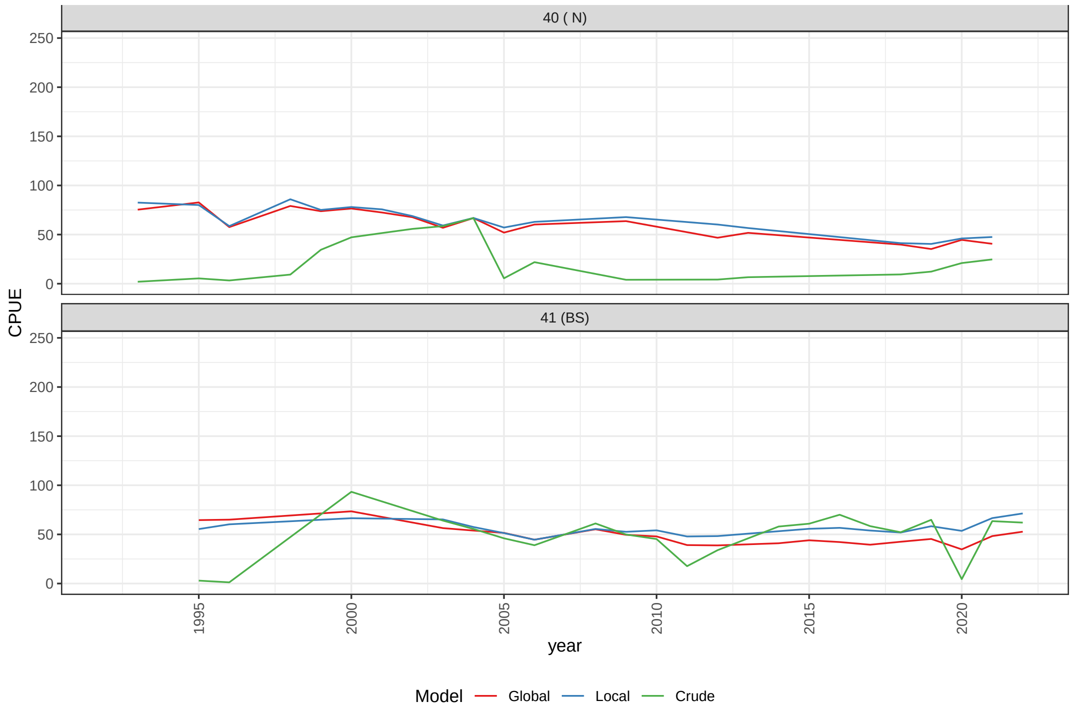
32 (BS)
33 (BS)
CPUE
year
Model
Global
Local
Crude

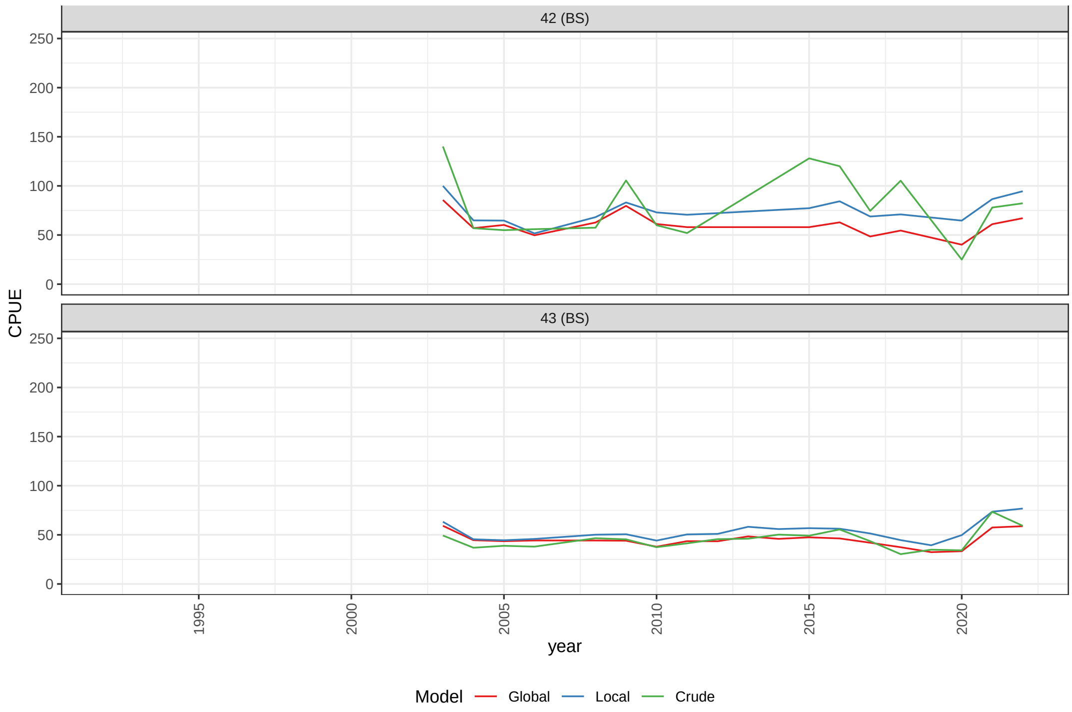


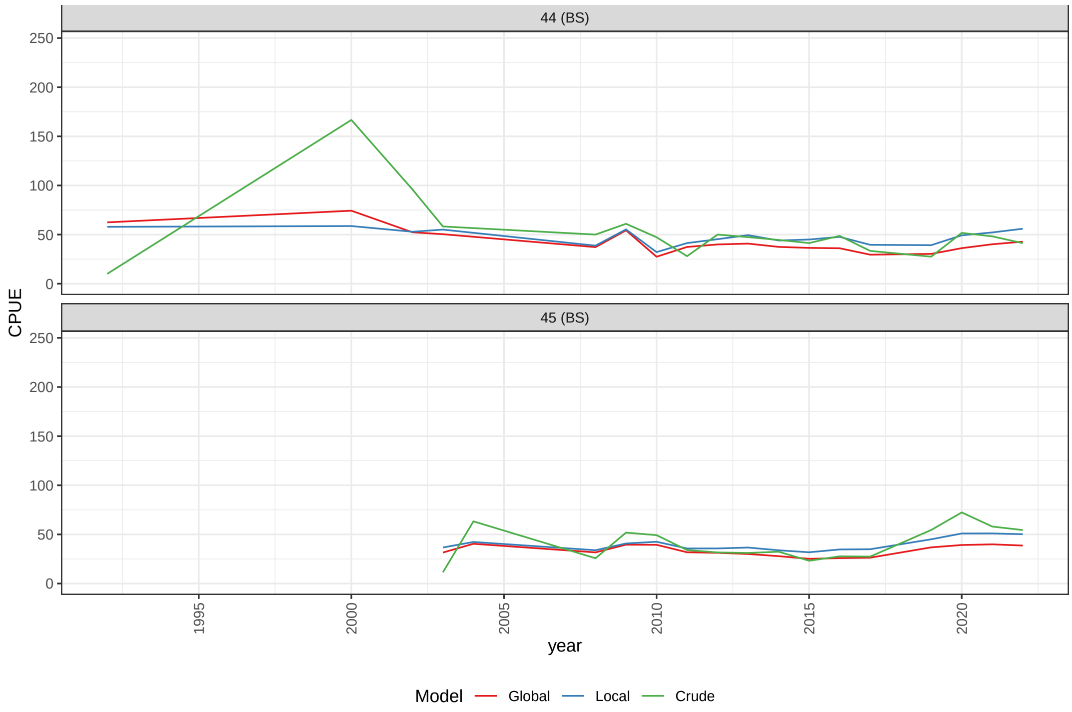
34 (BS)
35 (BS)
CPUE
year
Model
Global
Local
Crude

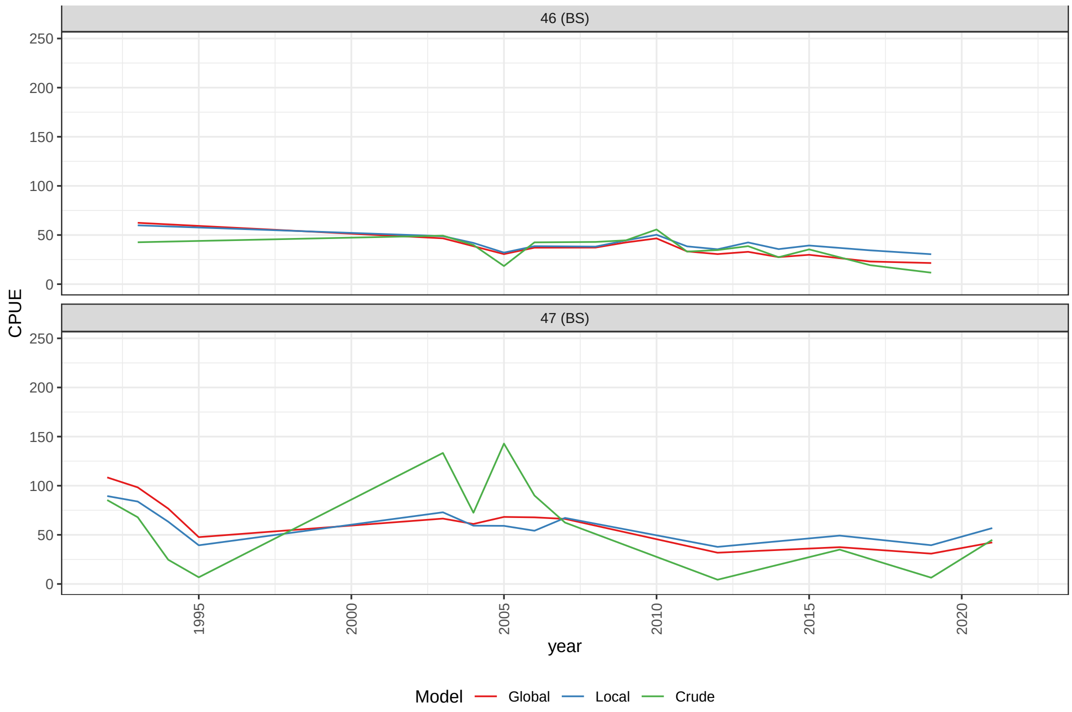

36 (BS)
37 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

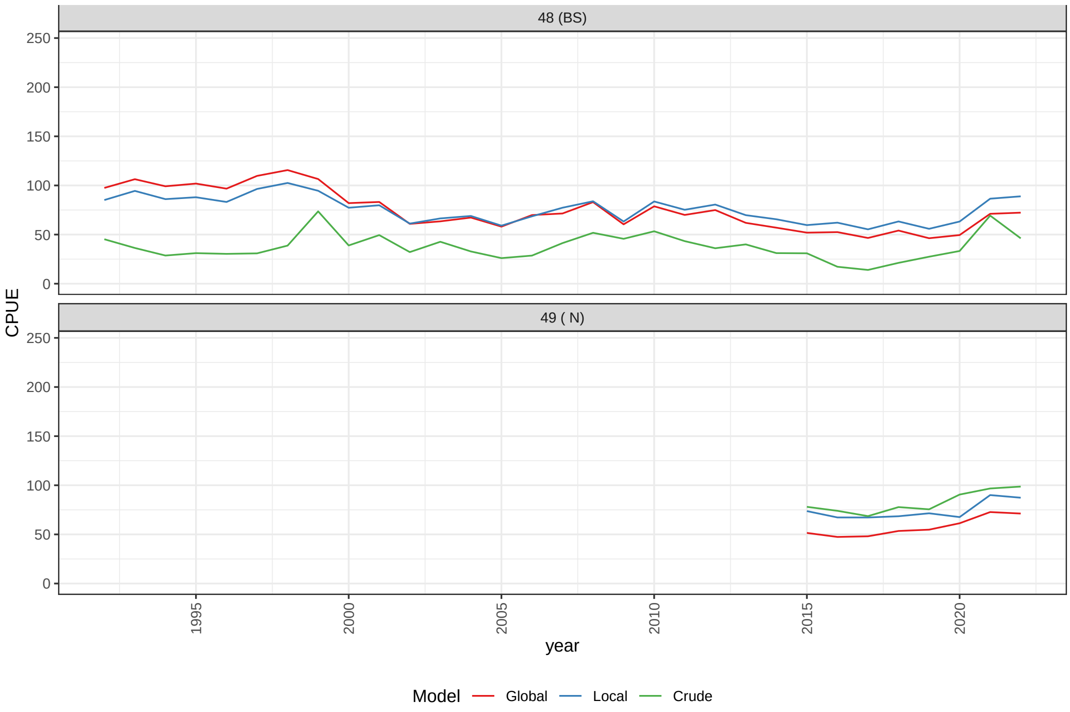

38 (BS)
39 (N)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

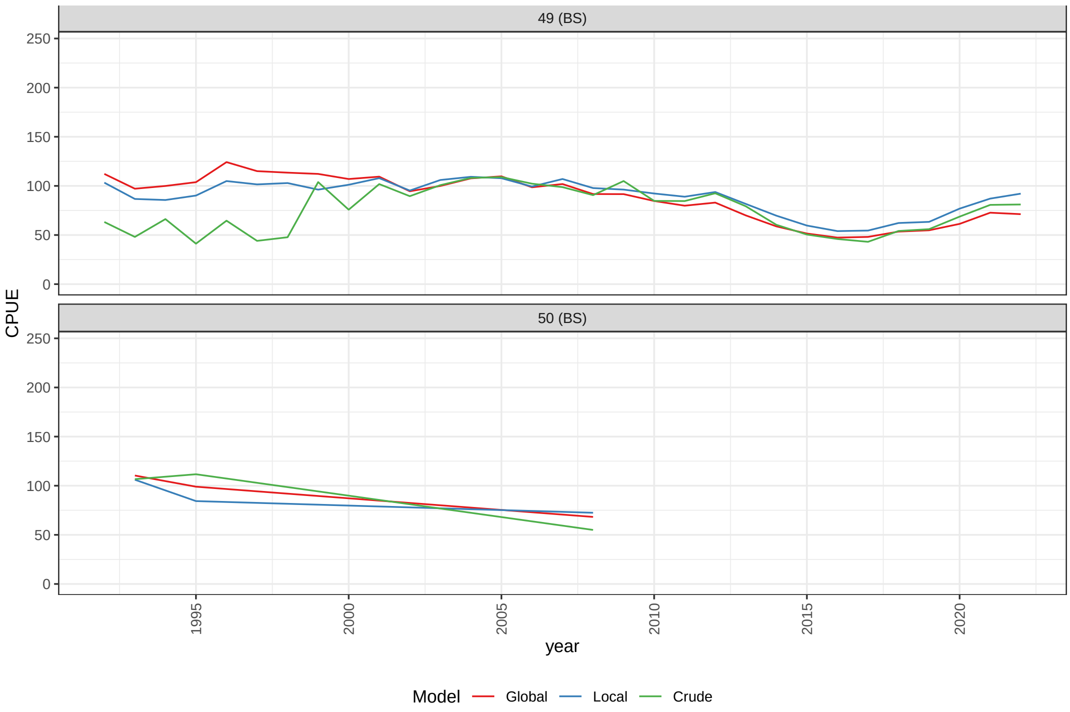

40 (N)
41 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

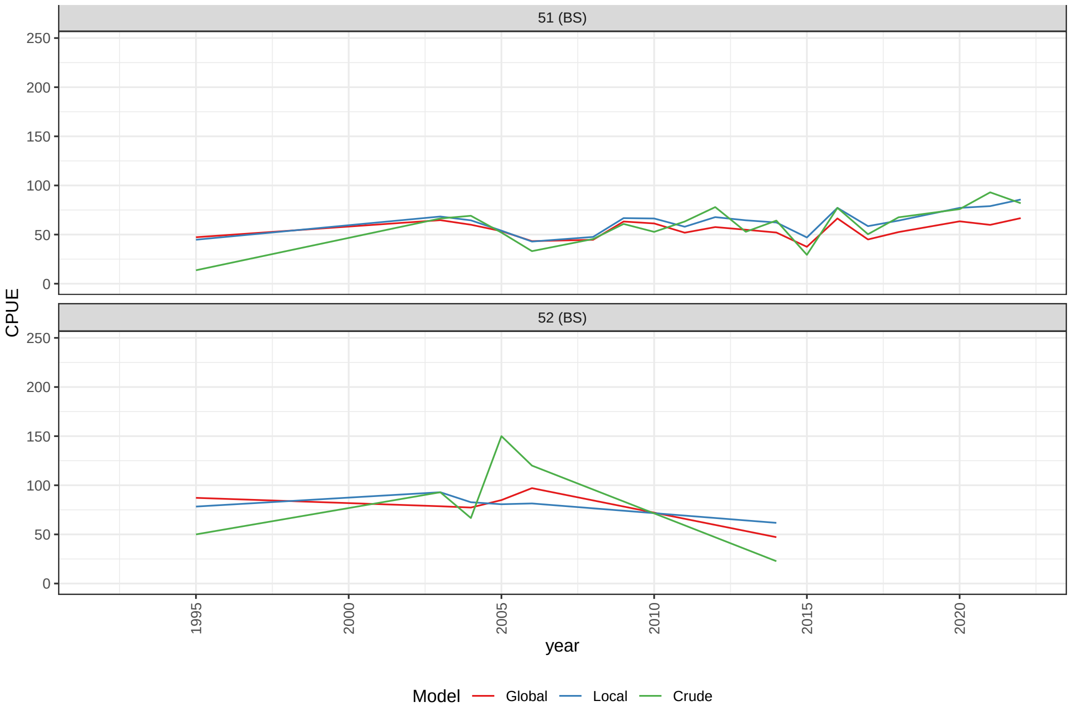

42 (BS)
43 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

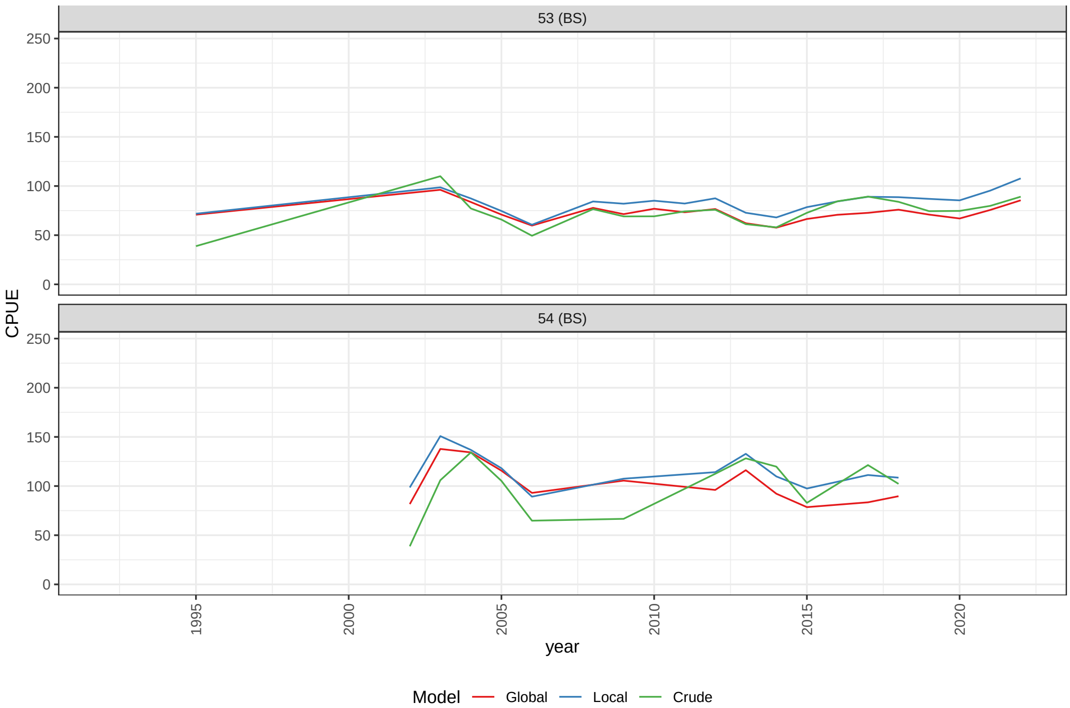

44 (BS)
45 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

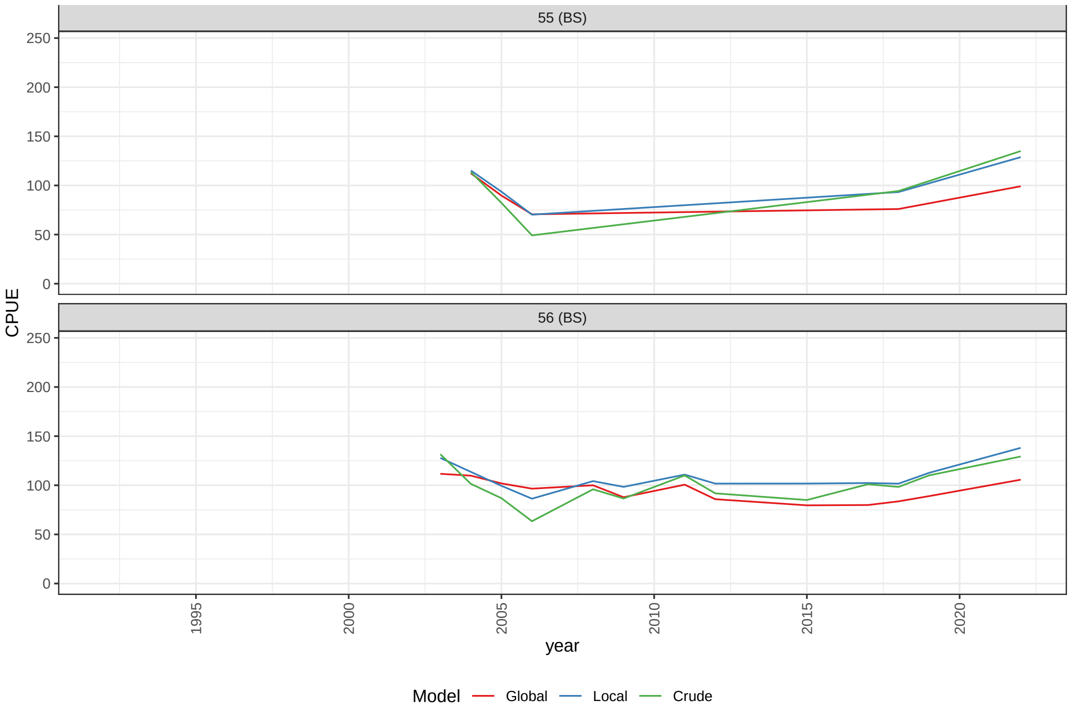

46 (BS)
47 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude

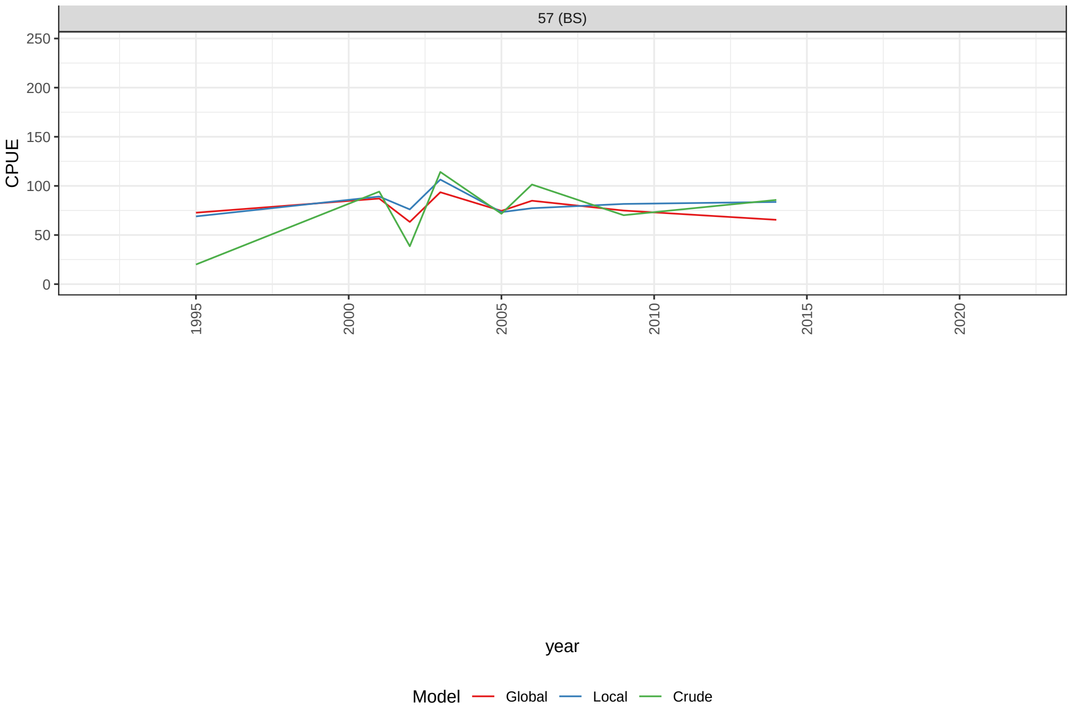

48 (BS)
49 (N)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude


49 (BS)
50 (BS)
CPUE
250
200
150
100
50
0
1995
2000
2005
2010
2015
2020
year
Model
Global
Local
Crude


51 (BS)
52 (BS)
CPUE
year
Model Global Local Crude


53 (BS)
54 (BS)
CPUE
year
Model Global Local Crude





12.6 Discussion

12.6.1 Spatial issues

One message that seems fairly clear to us is that the spatial scale on which the model is fitted to the data is not such a crucial issue, as long as it takes in enough blocks to allow the penalisation aspect scope to do a useful job.

The per-zone models lead to very similar standardisation results for their blocks as does a single, all-of-fishery model.

What does seem to be more of an issue is the size of the blocks themselves. (Here 'size' refers to both the spatial extent and the amount of data the block generates.) Where blocks are data sparse the model appears almost to ignore what little data is there and penalise almost right back to the grand mean level, i.e. what would have resulted using just a fixed `year` term to map biomass change and ignoring the blocks. We suggest it might be useful to investigate some adjustment of the target regions, perhaps combining adjacent blocks where data is sparse.

One objective of the project is to investigate how, and to what extent, spatial information may be useful in the assessment. We suggest that a useful first step would be simply to link lats and longs to the current logbook data. This would allow the longer logbook series to be used to check if it is at least likely that spatial information will likely prove useful in this regard. Even if lats and longs were only available for the centroids of the blocks themselves, this could be a start. Currently all we have in the logbook data is the block label, making even adjacency unknown.

12.6.2 Data issues

These notes have so far been confined to the Blacklip data, only. It is clear from the data that there has been some initial splitting of data into Blacklip and Greenlip catch, and presumably of fishing hours as well. It would be very satisfying to know the details. There are some anomalies in the record that we have had to work around to make the models work, such as zero or near zero fishing hours but with large catches.

One outcome that leaps out is the importance of the `propblip` variable, presumably already used to split the catch and effort, but remaining an important predictor in the standardisation model. This suggests that perhaps the split has not been quite as good as it might be. It would be useful to know how this variable was derived, what is its accuracy and how has it already been used.

The data record contains no zero catches, which surprises me, (but it does contain zero or very near zero fishing hours, as noted). This also seems surprising. Were zero catch records excluded for some reason?

12.6.3 Standardisation issues

We suggest that this approach to standardisation, using a modelling approach that uses data from some large number of blocks simultaneously has a number of advantages and should be considered, possibly not as a replacement for the present methods but as a complementary method. At present the modelling approach offers some insight that a block-by-block modelling approach cannot; using random effects it can use the context far more effectively, providing answers for sparse data blocks that the previous method cannot, or cannot do safely. On the other hand it does rely on a strong assumption that such context information is indeed relevant and useful.

We do not suggest the modelling approach we have outlined here is necessarily best, particularly not at the early stage of development sketched here, but we do suggest that the general strategy is worth considering.