Fisheries Research 269 (2024) 106860

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Catch per unit effort modelling for stock assessment: A summary of good practices

Simon D. Hoyle[a,b,c,*], Robert A. Campbell[d], Nicholas D. Ducharme-Barth[e], Arnaud Grüss[f], Bradley R. Moore[a], James T. Thorson[g], Laura Tremblay-Boyer[h], Henning Winker[i], Shijie Zhou[j], Mark N. Maunder[k,l]

[a] *National Institute of Water and Atmospheric Research Ltd (NIWA), 217 Akersten St, Port Nelson, Nelson 7010, New Zealand*
[b] *Indian Ocean Tuna Commission, PO Box 1011, Victoria, Seychelles*
[c] *Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand*
[d] *CSIRO Environment, Private bag No. 1, Aspendale, Victoria 3195, Australia*
[e] *NOAA National Marine Fisheries Service, Pacific Islands Fisheries Science Center, 1845 Wasp Boulevard, Building 176, Honolulu 96818, HI, USA*
[f] *National Institute of Water and Atmospheric Research Ltd (NIWA), 301 Evans Bay Parade, Greta Point, Wellington 6021, New Zealand*
[g] *Habitat and Ecological Processes Research Program, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way N.E., Seattle, WA 98115, USA*
[h] *CSIRO Environment, GPO Box 1538, Hobart, TAS 7001, Australia*
[i] *Department of Aquatic Resources, Institute of Marine Research, Swedish University of Agricultural Sciences, 45330, Sweden*
[j] *CSIRO Environment, 306 Carmody Road, St Lucia, Brisbane, QLD 4067, Australia*
[k] *Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA*
[l] *Center for the Advancement of Population Assessment Methodology, Scripps Institution of, Oceanography, La Jolla, CA, USA*

A R T I C L E I N F O

Handled by A.E. Punt

Keywords:
Catch-per-unit-effort (CPUE) standardization
Indices of relative abundance
Good practices
Data preparation
Modelling methods
Stock assessments

A B S T R A C T

Indices of abundance based on fishery catch-per-unit-effort (CPUE) are important components of many stock assessments, particularly when fishery-independent surveys are unavailable. Standardizing CPUE to develop indices that better reflect the relative abundance requires the analyst to make numerous decisions, which are influenced by factors that include the biology of the study species, the structure of the fishery of interest, the nature of the available data, and the objectives of the analysis such as how standardized data will be used in a subsequent assessment model. Alternative choices can substantially change the index, and hence stock assessment outcomes and management decisions. To guide decisions, we provide advice on good practices in 16 areas, focusing on decision points: fishery definitions, exploring and preparing data, misreporting, data aggregation, density and catchability covariates, environmental variables, combining CPUE and survey data, analysis tools, spatial considerations, setting up and predicting from the model, uncertainty estimation, error distributions, model diagnostics, model selection, multispecies targeting, and using CPUE in stock assessments. Often the most influential outcome of exploring and analysing catch and effort data is that analysts better understand the population and the fishery, thereby improving the stock assessment.

1. Introduction

Indices of abundance based on fishery catch-per-unit-effort (CPUE) are important components of many stock assessments for fish and other marine species, particularly when fishery-independent sources of information about population trends, such as research surveys, are unavailable (Maunder and Punt, 2004). This is the case, for example, in many fisheries for pelagic or lower value species where surveys are too costly or impractical (Bishop, 2006). In such situations, the main approach is to develop models that standardize fishery-dependent catch and effort data to produce an index of abundance (often simply referred to as the 'index') that better reflect the relative abundance by accounting for the other factors that influence catch rates (Maunder and Punt, 2004; Ye and Dennis, 2009). This practice dates at least to Beverton and Holt, Section 12) (1957).

CPUE standardization requires the analyst to make numerous

* Corresponding author at: National Institute of Water and Atmospheric Research Ltd (NIWA), 217 Akersten St, Port Nelson, Nelson 7010, New Zealand.
E-mail address: simon.hoyle@gmail.com (S.D. Hoyle).

https://doi.org/10.1016/j.fishres.2023.106860
Received 30 April 2023; Received in revised form 15 September 2023; Accepted 16 September 2023
Available online 30 September 2023

decisions. These decisions are influenced by factors that include the biology of the study species, the structure of the fishery of interest, the nature of the available data, and the objectives of the analysis (including how standardized data will be used in a subsequent assessment model). Alternative choices can change the indices and, consequently, the stock assessment outcomes and resulting advice to fishery managers. To guide decisions and choices, in this paper we provide advice on good practices in 16 areas for analysts who develop indices of abundance for use in stock assessments (Table 1, Fig. 1). For each of the 16 areas, we focus on decision points – occasions where analysts need to select an approach – particularly on instances where those choices can affect the resulting index. Analysts are encouraged to consider the proposed good practices and apply them where practical. It is important for the analyst to understand the implications of applying or not applying these practices and to be able to justify their decisions.

We begin by considering the fishery definition in the stock assessment – i.e., for what stock component are we estimating CPUE? Next, we consider the available data, including issues associated with data preparation, quality, misreporting, resolution, and aggregation. We then discuss processes that can affect catch rates, such as variables associated with catchability versus density. We then move to aspects of the analysis itself, including the tools used, methods to identify fishing fleet targeting strategies, issues associated with space, setting up and predicting from the model, error distributions and uncertainty estimation, model diagnostics, and model selection. Finally, we consider approaches for employing CPUE in stock assessments, including the use of index fisheries, splitting the time series, selectivity changes, and catchability change.

The advice provided in this paper can be implemented using various software packages. However, given their free availability and widespread use, we primarily focus on examples using the R software environment and its associated libraries (R Core Team, 2022).

2. Basic equations and definitions

The relation between catch rates (CPUE) and stock abundance is based on the catch equation which, as a first order approximation, relates the number of fish in the catch, C, fishing effort, E, and the average fish population density, D, on the fishing grounds:

$$C = qED$$

where q is a fixed constant of proportionality known as the catchability coefficient and is related to the efficiency of the fishing gear (i.e., the proportion of the stock removed by one unit of effort). From this equation:

Table 1
The 16 areas of good practices in catch-per-unit-effort (CPUE) standardization discussed in this paper.

Section/subsection	Area of good practices
Section 3	Fishery definitions
Subsection 4.1	Exploring and preparing data
Subsection 4.2	Misreporting and biases
Subsection 4.3	Data aggregation
Subsection 4.4	Density and catchability covariates
Subsection 4.5	Environmental variables
Subsection 4.6	Combining fishery and survey data
Subsection 5.1	Model fitting methods
Subsection 5.2	Spatial considerations
Subsection 5.3	Multispecies targeting
Subsection 5.4	Error distributions
Subsection 5.5	Uncertainty estimation
Subsection 5.6	Model diagnostics
Subsection 5.7	Model selection
Subsection 5.8	Assembling an index from a fitted model
Section 6	Using CPUE indices in stock assessments

$$CPUE = \frac{C}{E} = qD = \frac{qN}{A}$$

where N is the number of fish on the fishing grounds (a three-dimensional volume, usually defined by the two-dimensional surface) and A is the spatial area of the fishing grounds. It follows that changes in CPUE are due either to changes in the stock density or to changes in the catchability coefficient. If the changes in q can be accounted for ('filtered out'), then the remaining changes in CPUE can be related to those in stock density. This is the basic idea underlying what is known as the standardization of catch rates.

The concept of abundance needs some elaboration, particularly in relation to the concept of availability. Following the definitions proposed by Marr (1951), true abundance is the absolute number of individuals in a population, availability is the degree (a percentage) to which a population is accessible to the efforts of a fishery, and apparent abundance is the abundance as affected by availability, or the absolute number of fish accessible to the fishery. The concept of availability can be further decomposed into two components: one that is directly under the influence of the fishing gear (i.e., the selectivity of the gear defined as the probability of capture of any fish dependent on individual traits such as size) and one which is influenced by factors other than the fishing gear (see below).

From these definitions, if M represents the true abundance, N measures the apparent abundance, s measures the selectivity of the fishing gear, and a represents the other component of availability, then $N = saM$ and substituting into the above equation gives:

$$CPUE = \frac{saqM}{A}$$

From this equation, it is seen that the relationship between CPUE and the true abundance of fish within a given spatial region is influenced by both the selectivity (s) and other components of availability (a) of the fish to the fishing gear, and the efficiency of the fishing gear (q); thus, to adequately standardize CPUE, the analyst needs to understand what influences each of these factors. Note, in most of the fishery literature, the selectivity parameter combines the two parameters selectivity and availability as defined here (i.e., the distinction between the s and a terms is rarely made). Selectivity and availability often differ by factors such as age, size, sex, and stage and can be informed by the composition data based on these characteristics. To simplify the following descriptions and discussions, we simply refer to selectivity, availability, and composition data in the following without specifically referring to these characteristics.

Availability will be influenced by the environmental conditions prevailing at the time of the fishing operation and/or behavioural attributes of the species being targeted. For example, oceanographic conditions may influence the vertical distribution of the habitats preferred by both pelagic and semi-demersal target species and, consequently, the overlap of these habitats with the fishing gear (Hinton and Nakano, 1996; Maunder et al., 2006; Monnahan et al., 2021). Behavioural characteristics of a species, such as diurnal vertical migrations, may also influence the vertical availability of fish to the gear (Abascal et al., 2010; He et al., 1997). Covariates describing environmental conditions that likely impact local availability to fishing gear are one type of 'catchability covariate'. Alternatively, oceanography can also affect the local density, i.e., by causing fish to move outside of the fished area and these covariates are often referred to as habitat or 'density' covariates (O'Leary et al., 2020). The distinction between 'catchability' and 'density' covariates is discussed more thoroughly in Sections 4.4 and 5.8.

In addition to localized environmental conditions, catchability will be influenced by the types of fishing gear used and how they are deployed. The effectiveness of a given fishing gear (i.e., its ability to catch the available fish being targeted) will also depend upon a range of factors that are under the influence of the fisher. For example, the catch

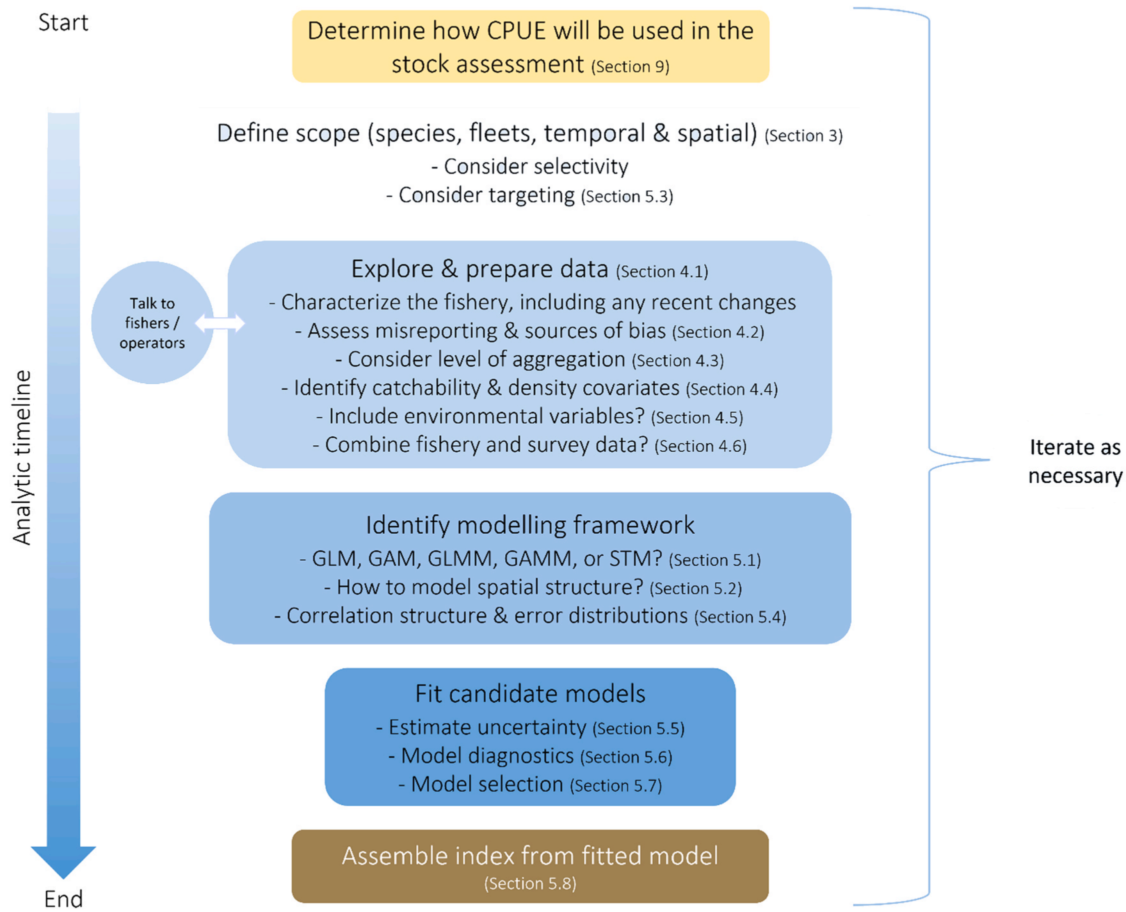

Fig. 1. Flow chart summarizing the stages of an analysis of catch per unit effort data.

of broadbill swordfish (*Xiphias gladius*) increases with the use of lightsticks and squid bait on shallow sets deployed in the afternoon, while albacore tuna (*Thunnus alalunga*) prefer pilchard baits on deep sets deployed in the morning (Campbell, 2019; Campbell et al., 2017). Individual fishers will also strive to improve the effectiveness of the fishing operations to increase the catch rates, and the combined effects of learning and improving technology over time lead to a phenomenon commonly known as 'effort creep.' Towards this end, fishers will often 'experiment' with the way the fishing gear is deployed to maximize its effectiveness. To understand this variability and allow analysts to standardize the resulting CPUE for these differences, it is important that the details of how the fishing gear is deployed are fully recorded in the vessel logbook.

3. Fishery definitions

A CPUE index in any stock assessment is associated with the part of a stock sampled by a specific fishery (termed a 'fleet' or sometimes 'métier,' defined here as a specific fleet deploying a specific fishing method). The assessment model defines the relationship between the index and the overall stock via the catch equation (catchability coefficient and the selectivity function in age/length/sex/stage structured assessment models), and for multi-region models, the spatial structure.

Defining the fishery (and its representation as a fleet in the assessment model) is, therefore, a fundamental aspect of CPUE standardization, requiring a detour into the discussion of selectivity for structured assessment models. Most structured stock assessments make the separability assumption, i.e., that fishing mortality is the product of catchability, annual fishing effort, and selectivity (Quinn and Deriso, 1999). As shown above, selectivity is a combination of availability and contact selectivity, and many assessments assume that selectivity is constant through time, even though in practice it tends to vary over time (Sampson and Scott, 2012). Wrongly assuming constant selectivity can bias the assessment results (Hilborn and Walters, 1992; Martell and Stewart, 2014), although so can time-varying selectivity in the presence of unbalanced composition sampling, substantial data gaps, and/or model misspecification. These biases are driven in three main ways: (1) by affecting the proportion of the population abundance vulnerable to the fishery, which will affect model results via the fit to the CPUE; (2) by affecting removals from the population and, therefore, the population structure; and (3) by affecting the composition data, resulting in a mismatch between the observed and expected size/age/sex composition which can substantially affect results (particularly population scaling) via the composition likelihood. Joint modelling can be used to ensure that CPUE and composition data are set up consistently in the model (see below).

To reduce these biases, fisheries should be defined in ways that minimize selectivity change through time. An important source of selectivity variation is changing spatial and seasonal distribution of fishing effort through time and its interaction with spatiotemporal variation in population structure and availability.

The analyst should explore spatial and seasonal patterns in availability by size, age, and/or maturity to identify fishery definitions that will be robust to changing effort distribution. Tree-based analytical methods can be used to identify optimal spatial arrangements for fisheries by exploring spatial variation in composition and CPUE trends (Lennert-Cody et al., 2013). Similarly, generalized additive models (GAMs) can be employed to identify spatial, seasonal, and environmental patterns in length, maturity, and sex ratio which are then used to define fisheries for CPUE standardization (Devine et al., 2022; Hoyle et al., 2017). Care needs to be taken when the spatial definition does not include the whole population (see below).

Such exploratory analyses are also an effective way to improve understanding of the biological structure of the stock and the fisheries. This understanding is often more important for assessment outcomes than changes in trends of the indices themselves.

Gear selectivity can also change through time, the effect of which can be minimized by separating gear types with different selectivity into different fisheries. For example, a change from J-hooks to circle-hooks in the Australian longline fishery for broadbill swordfish resulted in fewer larger fish being caught (Campbell et al., 2019; Pilling and Brouwer, 2017). Managing selectivity variation with fishery definitions is a useful way to account for the major sources of selectivity variation and to predict catch and CPUE with plausible selectivity for strata with gaps in the composition data time series. However, the number of fishery definitions can be limited by the difficulty of managing them with the stock assessment package being used, and by the information available to distinguish them and to estimate selectivity.

An alternative or complementary approach to account for selectivity variation is to jointly model the CPUE and composition data (Maunder et al., 2020) using a spatiotemporal model (STM) to create a joint index of relative abundance and stock composition. The approach augments the standardization with the observed composition data to generate two sets of predicted composition data: one to estimate the index fleet selectivity and the other to estimate the extraction fleet selectivity. The index composition data are spatially weighted by the predicted standardized abundance (the index), while the extraction fleet composition data are spatially weighted by the catch. The index fleet can be designed to have stable selectivity through time by accounting for spatial and seasonal covariate effects on availability, as well as catchability variation associated with vessels, equipment, and fishing technique. In contrast, extraction fleet selectivity will vary through time. Some constraints with the joint modelling approach are that STMs can be data hungry and computationally slow and that their implementation requires more expertise than traditional approaches to set up and run. These factors have tended to limit their application, but simplifications are available to reduce computational demands (e.g., don't model the composition data but use the raw compositions). In addition, complexities such as spatial variation in covariate effects may be difficult to parameterize. This approach will be further discussed in Sections 5 and 6.

It is important to consider the spatial domain of the stock assessment: how this domain relates to the population domain of the species, how this domain can be subdivided into a hypothetical 'stock-wide sampling frame' composed of spatial sampling units, and what portion of the sampling frame is informed by CPUE data for a given fleet. If the fleet samples only a small proportion of this stock-wide sampling frame, then there is likely to be substantial non-random variation through time in the proportion of the assessed stock available to that fleet. This variation in availability through time is represented as variation in catchability for a standardized index from the fleet if it is used to represent changes in stock-wide abundance or size/age/sex composition (Wilberg et al., 2009).

The best approach to define fisheries and their associated CPUE will depend on the specifics of each stock assessment, the available data, and the analyst's skills and time. Indices that represent a higher proportion of the stock are more likely to be representative (i.e., fewer processes could drive changes in availability that are confounded with density). However, there may be constraints that restrict the components that can be included in a representative index such as spatiotemporal patterns or covariate effects that are too complex to include in a single STM. It may be useful to combine the use of fishery definitions to account for large and stable effects on selectivity with STMs to account for remaining variation within each fishery. On the other hand, it may be better to combine multiple fisheries by calibrating them (with catchability and selectivity estimated) in an STM to try and span the spatial range of the stock. Such choices are a topic for ongoing research.

4. Data issues

4.1. Exploring and preparing data

Before standardizing CPUE it is essential to thoroughly explore the catch and effort dataset, to develop understanding of the data, generate hypotheses, and inform the standardization strategies to consider. Understanding can be developed by graphically summarising the data (both the response variable and covariates), exploratory modelling (e.g., using random forest (RF) models), and by consulting with key stakeholders. There should be special focus on identifying changes in the distributions of covariates through space or time, as covariates that change are the most likely to affect the index (Bentley et al., 2012). This preliminary exploration will also allow the analyst to anticipate issues with model variables such as data entry errors, missing data, and outliers.

Exploratory analyses should consider the sources of the data and the constraints imposed by data collection and storage methods through time. Logbooks and observer forms often change through time in ways that affect data quality, such as in the recording of spatial and temporal resolution, which covariates are reported, the precision and detail reported, species resolution, whether catches are recorded in number or weight, whether a species is recorded, and reporting of sex and size. Data storage methods may also be influential.

Graphical summaries can be undertaken in consecutive steps from low-to-high resolution. Aggregated summaries (e.g., histograms or density plots) of each variable across the full dataset provide an overall description of the response variable and potential covariates. These explorations should consider the proportions of zeroes (where relevant) and the range of the data prior to cleaning. Response variable summaries will inform the set of error distributions to consider (Campbell, 2015; Hoyle et al., 2014b); see also Section 5.4, including delta (hurdle) and mixture models (Langley, 2019). Aggregated summaries also help identify outliers and data entry errors (see below) and the need for possible data transformation. Examining individual values taken by continuous variables is often necessary to identify and remove outliers.

Spatial and temporal trends in the catch data and candidate covariates should be described through, for instance, time-series of boxplots and maps of summary metrics (mean, median, upper and lower quantiles). Where possible, maps of the variable of interest should be disaggregated by temporal strata to help detect any spatiotemporal trends. Spatial GAMs can be useful for exploring patterns and identifying statistically meaningful trends (see Section 5.2).

To help identify factors that may influence catch rates, relationships between nominal CPUE and candidate covariates should be explored. These data analyses will inform the type of modelling framework to use for continuous covariates, as not all model types allow non-linear relationships (see Section 5.1). Exploratory modelling using a flexible modelling framework such as RFs or boosted regression trees (BRTs) could also be undertaken to identify covariates explaining the highest proportion of the variability in the nominal CPUE and the potential shape of the relationship (e.g., linear or quadratic).

Analysts should maximize the number of records available for the standardization to improve model performance and estimates of uncertainty. As such, the treatment of missing data in covariates needs careful consideration as most standardization methods can only include records that have values for all the model covariates (Forrestal et al., 2019). Analysts should tabulate missing values for each covariate and try to identify causes and correlates of missingness. Options for records with missing values include dropping the full record, dropping the covariate, inferring the value from other information (e.g., a missing fishing gear value may be inferred from the gear typically used by the fisher), or imputation (e.g., median values per vessel). Approaches for imputing values range from simple rule-based procedures (Walters,

2003) to more sophisticated geostatistical (Munoz et al., 2010; Thorson et al., 2015) and Bayesian algorithms (Shemla and McAllister, 2006). When including a covariate with a high proportion of missing entries (>5–10%) that changes the standardized index, the analyst should confirm that the change is due to the covariate itself (e.g., by using influence plots) and not due to the removal of records where the covariate was missing.

The identification of outliers and their treatment is a key step in preparing data for CPUE standardization. Model predictive performance is likely to be poorer in the lower and upper tails of covariates given insufficient data, and outlying values are more likely to be reporting errors. Common approaches include omitting values outside ranges that are either fixed or pre-determined based on quantiles. The analyst should also check whether outliers are true records of infrequent fishing behaviour (for instance) or data entry mistakes.

Lastly, stakeholder engagement is an important, yet often overlooked, component of developing a CPUE standardization. It supports all the steps outlined above. Fishers, fishery observers, and even fish sellers can provide valuable insights to clarify potential data entry issues, confirm or invalidate data trends identified in exploratory analyses, advise on key covariates to include and their bounding values (e.g., net length, number of hooks per line), and explain motivations for changes in fishing behaviour (Tesfamichael et al., 2014). When detailed logbook data are available, fishers can also highlight the targeting and gear setting practices they consider will influence the catch rates of the species being targeted. Importantly, engagement can improve trust in the analysis as well as facilitate constructive dialogue when the final indices are considered for management.

4.2. Misreporting and biases

Misreporting catch of the species of interest is a consistent and sometimes major problem in many fisheries (Pitcher et al., 2002; Rudd and Branch, 2017), which can affect the data used in CPUE standardization. Indices are more prone to bias when rates of misreporting are higher or change through time. To understand the potential for bias, the analyst needs to become familiar with fishing industry dynamics and the data collection processes and to explore the data. It is useful to construct a timeline of changes (e.g., regulations, gear specifications, logbook forms, observer training, market conditions etc.) that may affect reporting for the species and fishery being evaluated. This can help to generate hypotheses to consider while modelling, particularly when a sharp change is observed between time periods.

One cause of misreporting is errors in species identification, which is particularly important for bycatch species. Such errors can also affect target species in both commercial (e.g., Beerkircher et al., 2009; Peatman et al., 2019; Webber and Starr, 2022) and recreational (Jones, 2004) fisheries. The ways species are reported can also vary. For example, sharks can be recorded either at the species level or at a more generic grouping level. CPUE indices may need to be restricted to periods with reliable species identification (e.g., Noriega et al., 2011).

Sampling bias can affect catch estimates based on estimates of species composition made by both fishers and observers. Examples include biases associated with grab sampling (Peatman et al., 2019), inconsistent reporting and possible layering in the catch (Webber and Starr, 2022), and the stratification used for statistical analysis (Duparc et al., 2020).

Changes in misreporting can also be linked to administrative factors such as changes in logbooks, regulations, the introduction of e-monitoring, or observer training. Both under-reporting and over-reporting can result from attempts to avoid quota limits, such as when the harvest location is misreported (e.g., Hoyle et al., 2015b). Deliberate over-reporting can be linked to attempts to establish a catch history in anticipation of future quota allocation (Buck, 1995). Under-reporting can be a feature of a management system, such as when only the top N captured species are reported (e.g, New Zealand inshore trawl fisheries, Langley, 2019).

Bycatch species are more likely to be under-reported when they are of little interest to the crew (due to their low commercial value), and rates may vary between vessels and change through time. For example, rates of shark bycatch reporting in Japanese southern bluefin tuna (*Thunnus maccoyii*) longline fisheries varied substantially between vessels and increased substantially in 2008 for reasons that remain unclear (Hoyle et al., 2017). When rates of misreporting vary between vessels, analysts can focus on vessels that report more reliably (Grüss et al., 2022). It should be noted that not all data may be needed to estimate a precise index. Limiting the analysis to reliable data that has less variability in catchability and selectivity will be adequate if they have sufficient spatial, temporal, and covariate coverage.

Under-reporting can also occur due to discarding or high-grading, which can be based on size or relative commercial value (e.g., tuna discards in the Indian Ocean, Huang and Liu, 2010), or on prescribed conservation measures. The stage in the fishing process when discarding occurs can affect whether it is recorded as part of the catch. Analyses may need to be adjusted to account for changes in discard rates (e.g., Hoyle et al., 2019b). Changes in reporting procedures, such as the introduction of electronic monitoring, may also change the rates of reporting of discard species (Emery et al., 2019). This is important if discards are included in the total catch used in CPUE analyses. Depredation of fishing sets by large marine predators (e.g., Peterson et al., 2014; Roche et al., 2007) can also result in under-reporting or zero-inflation of catch and may require additional modelling or adjustment of the data.

Misreporting is difficult to address but there are options in some circumstances. If available, analysts can compare catch rates between data types to identify groups of unreliable records, such as comparisons between observer data and logbook data (Hoyle et al., 2017), or between logbooks before and after the introduction of electronic monitoring (Emery et al., 2019). Since catch rates vary with covariates and targeting, predicting expected logbook catch rates based on models fitted to observer data is more reliable and powerful than simply comparing raw catch rates between datasets (Hoyle et al., 2017; Kai, 2019).

4.3. Data aggregation

Analysis of aggregated CPUE data may be required or convenient given the availability and quality of data. Due to privacy concerns, public domain fisheries data are often available only at aggregated spatial and temporal scales (Hinz et al., 2013). Aggregation may begin with the fishing logbook, e.g., with effort reported at the day or even the trip level. Data may also be aggregated across species due to identification problems or for convenience. Although analyses of aggregated data may be the only option when finer-scale data (e.g., set-by-set or species level data) are unavailable, their results should be treated with caution and critically evaluated given the issues raised in the following paragraphs.

Indices derived from data aggregated spatially, temporally, or across species may not vary in proportion to the target stock (or complex). Covariates that affect catch rates at the set or vessel level (e.g., gear settings) or summarize the effects on catchability of multiple factors (e.g., vessel IDs) are often unavailable when CPUE data are aggregated to coarser spatial or temporal strata. This loss of information may limit the ability of CPUE standardization to correct for changes in catchability over time (e.g., effort creep: Kleiven et al., 2022; Palomares and Pauly, 2019). Aggregating data temporally and using a coarser temporal definition of effort (e.g., day, trip, quarter) than the higher-resolution effective unit of effort (e.g., fishing sets, pot throws, or hours searched) can change the relationship between CPUE and abundance (i.e., the covariate changes over a smaller temporal or spatial scale, often intentionally to improve catch rates, but these changes are not apparent in the aggregated data). For example, defining CPUE as catch per day could mask a population decline if search time within each day or the number of schools fished changes to maintain stability in catch magnitude (Ducharme-Barth et al., 2022; Hsu et al., 2022). Spatial and temporal aggregation of CPUE data may create a mismatch between the actual oceanographic conditions and the oceanographic covariate associated with the aggregated CPUE data. If oceanographic covariates are modelled as density covariates (see Section 4.4), this mis-match could lead to spurious estimated relationships between predicted density and the covariate and, if using STMs, could result in inappropriate spatial imputations of density (Ducharme-Barth et al., 2022). For example, fishers may move to find a specific habitat type within a larger spatial stratum to increase catch rates of a certain species. If fishers target fish aggregations or areas of preferred fish habitat such as eddies or seamounts at a finer resolution than the spatial scale of the aggregated CPUE data, then localized, sequential depletion of aggregations could introduce hyperstability into the index (Cardinale et al., 2011; Sadovy and Domeier, 2005). Lastly, aggregating CPUE data across species is likely to exaggerate fishery-induced changes to the combined abundance, given that catchability varies across species (Kleiber and Maunder, 2008). Similarly, if productivity varies across the aggregated species, fishery-induced decline in abundance of the less productive species might be masked by the abundance trend of the more productive species (Dulvy et al., 2000).

In addition to the risk of bias, analysing aggregated data poses statistical modelling challenges, such as applying an appropriate variance structure. Commonly used error structures (e.g., Lognormal and Poisson) for CPUE standardization models are generally robust to heteroscedastic data (e.g., non-constant variance assumption where, typically, the variance increases in proportion to the mean). However, this mean-variance relationship is inverted in aggregated data if areas of high catch rates also attract higher concentrations of effort, thus reducing the variance and violating distributional assumptions (e.g., Hoyle, 2021). Furthermore, spatially aggregating data makes it difficult for models to accurately represent the spatiotemporal correlation structure. For instance, in GAMs, where the spatial correlation is modelled implicitly by varying the degree of the spatial smoothing applied, aggregating data spreads the spatial and temporal extent of abundance hotspots, and generates less flexible spatial smooths (e.g., fewer degrees of freedom) which may further inappropriately 'smear' the hotspots when making spatial predictions.

Perhaps most importantly, aggregating data can remove fine-scale information that allows analysts to understand fishing behaviour. The key process of data exploration is limited by loss of information about, for example, individual vessel behaviour, fine-scale fish and vessel distributions, and movement patterns. Analysis methods to identify targeting strategy based on species composition (see Section 5.3) may remain possible (e.g., Fu et al., 2016), but with less resolution and sensitivity than with operational data.

Although they do not balance the concerns above, there can also be advantages associated with aggregating data before standardization, particularly for fisheries where pseudo-replication and serial autocorrelation are issues. For example, a vessel that obtains high CPUE along an oceanic front may continue to fish the front for as long as high catches persist, so that sets cannot be considered independent. Similar effects are known from sequential trawl hauls. It is often computationally challenging to apply the appropriate autocorrelation or random effect structure to address the consequent violation of independence, especially for large datasets. Given the risks of bias due to aggregation on the one hand, and pseudo-replication on the other, aggregation by vessel, month, statistical cell and/or other factors may in some cases be the lesser of two evils. A further option is to subsample the data in a random or structured way to reduce pseudo-replication. Trade-offs can be explored through simulation, or by working with more computationally tractable subsets of the operational data.

4.4. Density and catchability covariates

Since Beverton and Holt, Section 12) (1957), stock assessment scientists have standardized fishery CPUE to remove the confounding effects of vessel size and power. For example, Maunder and Punt (2004) state "*Explanatory variables should, however, be considered in an analysis only if there is an* a priori *reason that they may influence catchability.*" However, spatial ecologists do the exact opposite by including covariates and then conditioning upon their effect when predicting densities across space. To alleviate this confusion, Thorson (2019a) distinguishes between 'catchability' and 'density' covariates: both are included in a linear predictor to explain catch-and-effort data, but only density (and not catchability) covariates are conditioned upon when predicting densities across space. The index is constructed by aggregating predictions across a specified spatial and temporal domain such that the effect of catchability is removed from the index.

Distinguishing between catchability and density covariates leads to the question: how does the analyst know whether a covariate affects density, catchability, or both? For example, species might migrate to maintain a desired daytime foraging temperature (Lehodey et al., 1997) such that local temperature predicts population densities. Simultaneously, temperature might affect digestion and metabolic rates, leading to a greater attraction towards a baited fishing hook (and hence a higher CPUE for a given density and soak duration). In this case, an argument could be made that temperature affects both catchability and density. To resolve these hypothesized mechanisms, an analyst could combine fishery CPUE with auxiliary information, e.g., by measuring behaviour directly using satellite and/or conventional tags. Similarly, analysts can make inferences about density, catchability, and availability via:

1. *Local depletion*: In freshwater sampling, analysts can sample and remove individuals in multiple survey passes, using the decline and quantity of removals to identify catchability. Similarly, in sessile marine organisms, a decline in fishery CPUE over a short season can be compared with preceding fishery removals to estimate catchability.
2. *Paired sampling*: Similarly, bottom trawl data can be paired with vertically disaggregated acoustical backscatter to identify vertical distribution and infer availability to gear operating at a given vertical layer (Monnahan et al., 2021).
3. *Process studies*: Finally, an analyst might use a small-scale experiment (e.g., temperature-dependent feeding experiments) to identify how temperature might affect bait attraction and then assume this relationship to 'subtract out' this effect.

Regardless of the process, we recommend that analysts explicitly outline their rationale when specifying that a covariate affects density or catchability. In cases where the correct process is difficult to determine, analysts should explore the implications of excluding the covariate from the index and including it as either density or catchability. These alternative hypotheses can form the basis of alternative CPUE index scenarios considered in the stock assessment.

4.5. Environmental variables

As noted in Section 4.4, aspects of the aquatic environment, such as water temperature, surface chlorophyll-*a* concentration, oxygenation, and light at depth, affect both fish density and catchability. The task of CPUE standardization is to eliminate the effect of catchability variables so the analyst can employ CPUE data to infer differences in fish density across time and space. Given this estimate of density across time and space, the analyst can then sum density in each area to generate an abundance index for use in stock assessment, or average over years to get an estimate of habitat utilization for use in spatial management.

Regardless of whether analysts specify a covariate as affecting either density or catchability, all covariates are used to inform model fit to observations. Treatment between the two types of covariates differs at the prediction stage. Catchability covariates are fixed at a base value when predicting density across space and time, thereby 'filtering out' variation associated with different values in the fitted data. By contrast, a density covariate has an assigned value at every location across a modelled spatial and temporal domain, and the analyst conditions upon these values when predicting densities. It should be noted that catchability covariates are only needed for catch events, while density covariates are needed for all temporal and spatial strata included in the index irrespective of whether there is catch. Including a variable that affects density as a catchability covariate can result in biased inference. For example, if recruitment decreases as the average temperature increases (e.g., due to climate change), including annual water temperature as a catchability covariate in an index-standardization model and subsequently using a constant value in the prediction will mask the true abundance decline over time. However, including average temperature as a density covariate (and using its value when predicting densities) can result in improved estimates and forecasts (e.g., O'Leary et al., 2020). The converse is also true — wrongly attributing to density an effect that is linked to catchability will introduce bias.

Several environmental variables are commonly used in CPUE models. Variables such as sea surface temperature, sea surface salinity, sea level height, chlorophyll-*a* concentration, and dissolved oxygen concentration have been found to be significant predictors of spatial distribution for some species (Campbell, 2015; Han et al., 2022; Liu et al., 2022; Tian et al., 2009), particularly for pelagics. Some environmental variables, such as moon phase, eddies, water turbidity, and cloud cover, may affect the availability and/or catchability of certain species and gear types.

Environmental variables are often not recorded by fishers at the time of fishing, but rather obtained from other sources such as meteorological organisations. This is particularly true for density covariates, which must be available for all locations and time periods. Variables with a long time-lag (e.g., temperature in the months preceding the fishing event) should be used with caution in CPUE models. They may affect productivity during the period leading to fishing and therefore local density (a density covariate), but are often autocorrelated, where current values may affect catchability (a catchability covariate). Similarly, analysts sometimes use covariates that are derived from a large spatial area or obtained from satellite or regional ocean modelling systems (e.g., Phillips et al., 2014). In all cases, it is important to provide some ecological justification for (1) whether those variables are affecting density or catchability and (2) why those are better than using vessel-based measurements of environmental conditions, if available. The error in measuring the covariates may also need to be considered in the model, particularly if the error varies among spatiotemporal strata.

Despite being important factors affecting fish distribution, many environmental variables are confounded with spatial and seasonal coordinates that are often employed as predictors in CPUE models. Hence, spatial-temporal effects can act as proxies for comprehensive dynamic environmental variables, especially if interactions between these effects are included to help account for temporal changes in environmental conditions at any spatial location. If the environmental conditions generally exhibit consistent spatial and temporal patterns, which is often the case, explicitly including these environmental variables may explain little additional variation. In the cases where using environmental variables does increase accuracy, their inclusion may also increase the annual coefficients of variation (CVs) compared to the models without the environmental variables, likely due to the added requirement of estimating a relatively imprecise relationship between catch rates and environmental variables (Forrestal et al., 2017; Forrestal et al., 2019).

Goodyear (2016) explored relationships between environmental factors, three-dimensional variation in habitat, and longline CPUE in a case study for blue marlin (*Makaira nigricans*). Results of analysing data generated from this model generally favoured the inclusion of environmental and habitat variables but were affected by the approaches taken by each analyst, particularly for variable categorization and model selection (Forrestal et al., 2017; Forrestal et al., 2019).

Two research areas closely related to CPUE standardization are species distribution modelling (SDM) and habitat suitability modelling (HSM). Both SDM and HSM heavily use environmental variables to model fish distributions and their habitat preferences (Bosch et al., 2018; Lee and Terrell, 1988; Maunder et al., 2006; Pickens et al., 2021; Rowden et al., 2017; Zhang and Li, 2017). These studies involve similar data and modelling techniques to CPUE analyses. They also often consider a third spatial dimension, depth, which can be particularly important for some gears such as pelagic longline or for bottom-fish species. (see also Hinton, 1996); Hinton and Nakano (1996) developed a method to calculate indices of abundance for pelagic longline that match the spatiotemporal-vertical habitat with the longline gear and species habitat preference. This method was extended into a statistical framework by Maunder et al. (2006).

4.6. Combining fishery and survey data

In some jurisdictions, analysts produce a separate CPUE index for each fishery fleet or métier. This then results in a multitude of abundance indices, and equally weighting these indices implicitly results in the assessment model averaging them. To resolve and simplify this situation, Conn (2010) used a state-space model to combine multiple indices into a single 'consensus' index, an approach that has been refined using Dynamic Factor Analysis (DFA) (Peterson et al., 2017; Peterson et al., 2021). However, DFA has several drawbacks, including that: (1) it must assign some implicit weight to each constituent index when combining them, and these constituent indices representing small or large areas are often given equal weight; and (2) residual variation in constituent indices is ignored, and the analyst must make some decision about which DFA index represents changes in the stock or is attributed to correlated variation in catchability.

Given the difficulty of reconciling differences in multiple indices within an assessment, or pre-processing using DFA, analysts increasingly seek to include data from multiple fisheries and/or surveys during index standardization. Joint analysis expands data coverage spatially and temporally. It also ensures consistency of analytical methods, thereby removing a key source of differences between indices. This occurs, e.g., in combining nearshore and offshore survey data (Perretti and Thorson, 2019), combining multiple surveys to achieve a basin-scale index (Maureaud et al., 2021; Ono et al., 2018), or fleets from multiple nations to increase the spatial and temporal coverage of an index (Ducharme-Barth et al., 2020; Hoyle et al., 2019a; Hoyle et al., 2015b). Prior to joint analyses of data from multiple fleets, individual national datasets should be thoroughly explored (e.g., Hoyle et al., 2015a; Hoyle and Okamoto, 2015; Hoyle et al., 2015c) to identify and eliminate sources of data conflict. Joint indices for Atlantic tropical tunas (Hoyle et al., 2019a; Hoyle et al., 2019c) were judged to have improved the resulting stock assessments (Anonymous, 2019; Walter et al., 2020) by reducing data conflicts, improving model diagnostics, and ensuring broad and consistent spatial and temporal coverage.

However, there is less research regarding how to combine data from multiple fisheries and/or surveys. As one exception, Grüss and Thorson (2019) combined data from different surveys to generate an abundance index for Gulf of Mexico red snapper (*Lutjanus campechanus*). Notably, the authors first analysed each data source individually to confirm that any apparent conflict in the indices could be explained by differences in the spatial extent of each fleet. Rufener et al. (2021) combined fishery and survey data, after confirming minor apparent data conflict.

Given these successes, we suspect that there will be ongoing efforts to combine CPUE from multiple fisheries and/or surveys. We recommend the following good practices:

1. The analyst should standardize each dataset individually using broadly consistent methods and compare resulting density maps and

abundance indices both visually and using goodness-of-fit criteria (e.g., Alglave et al., 2022; Rufener et al., 2021);
2. Differences between dataset should be explained mechanistically where possible, such as by including covariates that can explain the differences in catchability.

These recommendations arise from the expectation that data conflicts imply model misspecification (the "Law of Conflicting Data", Maunder and Piner, 2017). The analyst should also consider that selectivity differences between surveys and fisheries may require simultaneously modelling the catch rate and length composition data to estimate differences in selectivity, in which case calibrating catchability and selectivity may require spatial overlap.

5. Analysis

5.1. Modeling framework

As outlined in previous sections, the objective of CPUE standardization is to remove the effects of changes in fishing practice through time (i.e., changes to catchability) from observed catch rates such that the trend in standardized CPUE reflects the true relative abundance of the stock component selected by the fishery. In practice, this is accomplished by building a statistical model that predicts CPUE as a function of a set of variables thought to impact local abundance (density) or catchability. Recent standardizations mostly use generalized linear models (GLMs) or an expansion thereof, implemented in the R Statistical Computing system (R Core Team, 2022); this section focuses on tools available within this framework.

There are five key factors to consider when selecting a modelling framework which will limit the set of approaches available to the analysts: (1) the form of the relationship between CPUE and candidate explanatory variables; (2) the probability (error) distribution of the CPUE; (3) the type of correlation that might be present within the CPUE dataset; (4) whether CPUE will be jointly standardized with composition data; and (5) how results are processed to yield an overall index of abundance. Often, the analyst will have to prioritize one of these five factors, as different software packages do not always allow the implementation of all modelling structures jointly. There are also logistical considerations (e.g., computational power) for large datasets that may prevent the implementation of some features.

A CPUE standardization will likely consider both categorical and continuous covariates. Vessel identity, bait type, and area are examples of categorical covariates; hooks-between-floats, net length, and sea surface temperature are examples of continuous covariates. Continuous covariates can sometimes be implemented as categorical, depending on the desired statistical treatment (e.g., Grüss et al., 2019). In the GLM framework and its relatives, categorical covariates are typically modelled as deviates from a 'baseline' intercept level. In a conventional GLM, categorical covariates are estimated as fixed effects, that is, there are no constraints on the value of the coefficient assigned to each level. This is the default treatment for a categorical covariate and will be available in most modelling frameworks.

An alternative is to fit categorical covariates as random effects. In a mixed-effect (or hierarchical) framework (e.g., generalized linear mixed model (GLMM); generalized additive mixed model (GAMM)), categorical covariates can be fitted as fixed or random effects. With random effects, level coefficients are constrained to belong to a distribution (usually Gaussian) of the expected values of the coefficients (Thorson and Minto, 2015). This results in the 'shrinkage' of covariate levels towards an overall mean, with more shrinkage for values less informed by data, which are otherwise more likely to be extreme. Mixed effects are now available in an increasing number of R packages, including *nlme*, *lme4*, *glmer*, *glmmTMB,* and *mgcv* (see also Bolker, 2022), and are straightforward to implement in bespoke Template Model Builder (TMB) code (see below).

Random effects have limited application for categorical covariates with too few levels to estimate the shape of the distribution, e.g., bait type. However, they can be useful for covariates with multiple (>12) levels that may be expected to belong to a common, normal distribution, such as a vessel effect. In these instances, some levels might have too few records to reliably estimate their model coefficients as fixed effects. Poorly sampled levels are not an issue in a mixed-effect framework as they are shrunk towards the coefficients of better-informed levels. Nevertheless, fixed effect levels with very low record numbers often have minimal influence on the indices if retained and may be omitted if they cause model convergence problems. Note also that vessel effects will not be normally distributed if, for example, there are groups of vessels with similar catchability or if there are temporal trends in the catchability of vessels joining and leaving the fishery. In such cases, the analyst may add structure to the random effects model or may prefer to use fixed effects. Simulation testing is useful for selecting a strategy that provides a reliable index.

A common strategy for vessels is to define a 'core fleet' comprised of vessels meeting an arbitrary set of activity thresholds (e.g., as a function of catch and/or time present in the fleet; (Kendrick and Bentley, 2011; McKenzie and Parsons, 2012). The core vessels are more likely to have characteristics, such as consistent targeting strategies and reporting behaviour, that indicate stable catchability. Treating vessel ID as a random effect can be an alternative to defining a core fleet since it avoids making arbitrary decisions about threshold rules for the core fleet and expands the number of records available to the analysis (Grüss et al., 2023a). Care should be taken when defining a 'core fleet' as more experienced operators may be able to maintain high catches even if abundance declines.

The relationship between CPUE and continuous covariates can take a variety of shapes, including linear, saturating, and dome shaped. In most instances, the relationship can be assumed to be non-linear as a starting point, as linearity is a strong constraint. Non-linear relationships are best implemented via splines, which are a useful improvement over polynomials as they are not constrained in the shapes they can take. Polynomials have other undesirable properties such as the fit in one data range being affected by data in other parts of the range (Harrell, 2001; Magee, 1998). The *splines* library (e.g., via the *ns* function) allows GLMs to include splines with a user-specified number of knots, including in base-R packages such as *stats::glm*.

A preferred alternative for fitting non-linear relationships is the R package *mgcv* which implements GAMs using penalized splines (Wood, 2017). With penalized splines, the analyst specifies a maximum of knots, and the algorithm attempts to maximize fit to the data while minimising spline 'wiggliness.' The inclusion of splines has improved the performance of CPUE standardization models by predicting more realistic relationships between CPUE and continuous covariates. However, care must still be taken to prevent overfitting or extrapolating far beyond the range of available data. If the splines are allowed to be too flexible, the CPUE standardization model could be capturing noise. In *mgcv*, the default optimization method GCV tends to overfit, and REML is recommended instead (Wood, 2017). The relative penalty on 'wiggliness' can be controlled via the gamma parameter which scales the effective sample size. Even with data that are independent and identically distributed (i.i.d.), it is recommended to increase gamma to 1.4 from the default *mgcv* value of 1 when using method=GCV (Wood, 2017). The analyst can also elect to allow for fewer knots, e.g., for one-dimensional splines; three to four knots will be enough to realistically represent most relationships between CPUE and a continuous covariate (Grüss et al., 2019; Roberts et al., 2016). However, restricting the number of knots will worsen the fit to the data and may increase the number of covariates retained in the model. The *mgcv* package implements many other useful features, such as cyclic (a.k.a. periodic) regression splines (see `*?mgcv::smooth.terms*') which constrain the effects fitted to the minimum and maximum values of a continuous covariate to be equal. This is useful when (for example) time of day or month of year are included as continuous covariates.

It is up to the analyst to select a probability distribution for the response variable (see Section 5.4). The choice of distributions depends firstly on whether the response variable is discrete (e.g., catch occurrence, catch in number of individuals) or continuous (e.g., catch in weight, the ratio of catch to a measure of effort). Common modelling frameworks (GLMs, GAMs) can now handle most standard error distributions, including the normal, log-normal, Gamma, binomial, Poisson, and negative binomial distribution models. For the Tweedie distribution (see Section 5.4), an additional R package (Tweedie: Dunn, 2017) may be needed, or it can be approximated using a Poisson-linked delta model (Thorson, 2018). The Weibull distribution, a positive continuous distribution that can handle overdispersion in the data, can be implemented in the R package *survival* (Therneau et al., 2022). The R package *gamlss* allows the analyst to fit a diverse set of error distributions (including zero-inflated applications of more common distributions and some that are less commonly available such as the beta distribution) and specify model structure explicitly for the different distribution parameters (Rigby and Stasinopoulos, 2005). Alternatively, the R package *brms* (see below) allows the analyst to specify custom distributions (e.g., Tremblay-Boyer and Neubauer, 2019). The choice of error distribution should be validated by inspecting residual diagnostics (See Section 5.6).

Most statistical frameworks assume by default that residuals are i.i.d. However, correlation between records in fisheries datasets arises from multiple sources. Spatial correlation can be handled implicitly by specifying a two-dimensional spline for longitude and latitude in a GAM, or explicitly in a geostatistical or spatiotemporal modelling framework by estimating the Matérn covariance function, for example, using the R packages *VAST* (Thorson, 2019a). The *R-INLA* package (Lindgren and Rue, 2015) can also be used for other geostatistical or spatiotemporal modelling applications (e.g., Cosandey-Godin et al., 2015; Pinto et al., 2019; Zhou et al., 2019) and provides additional options to account for assumptions about spatial and/or temporal correlation. Correlation structure between records belonging to the same strata (e.g., fishing trip, observer) can also be specified with generalized estimating equations (e.g., Coelho et al., 2020; Peatman and Nicol, 2021) using the R package *geepack* (Højsgaard et al., 2006).

Model fitting can be implemented in a frequentist or Bayesian statistical framework using maximum likelihood estimation (MLE) or Markov Chain Monte Carlo (MCMC), respectively. Most GLM/GAM tools are implemented via MLE or some derivation thereof. TMB (library TMB: Kristensen et al., 2016) can be employed for large datasets as model fitting will be considerably faster due to automatic differentiation and Laplace approximation; but the analyst will have to specify the CPUE model manually (e.g., model matrix, likelihood function) in the C+ + code embedded in TMB files, unless appropriate code is already available (e.g., VAST; Thorson, 2019a). An alternative is to fit CPUE in a Bayesian framework using MCMC with the R package *brms* (Bürkner, 2017). Key advantages of this approach include (1) the option to specify priors to inform or constrain effects, and (2) a more intuitive and better integrated estimate of index uncertainty. One downside is that model fitting will be slower, especially for large datasets with complicated covariate structures and may be less stable if some relationships are poorly informed by the available data.

Machine learning methods such as artificial neural networks (Maunder and Hinton, 2006), support vector machines (Li et al., 2015), regression trees (Watters and Deriso, 2000), and RFs (Chambers and Hoyle, 2015; Li et al., 2015) have been used to model CPUE but relatively infrequently to date. They can achieve high predictive performance due to their flexibility, but results may be difficult to interpret, and they are prone to overfitting.

5.2. Spatial considerations

All fish populations and fisheries exhibit spatial structure to some extent. Nevertheless, the introduction of spatial considerations into CPUE standardization raises several issues. First comes the question of how exactly space should be considered in a model and whether a complex model (e.g., an STM) will necessarily perform better than a simpler model (e.g., a GAM). Second comes the question of how to factor in the existence of sub-stocks (i.e., stock components that are distinguished for management purposes and whose productivities may differ).

Classically, CPUE standardization is performed using a GLM, which is very often a two-step, delta GLM to account for the presence of many zeros in the data (Lo et al., 1992; Stefansson, 1996). The simplest way to consider space in a GLM consists of dividing the study region into area strata and including the fixed effect of area stratum – a categorical covariate – in the GLM (e.g., Forrestal et al., 2017). The inclusion of area strata in a model aims to account for spatial heterogeneity in stock density, and stock density is assumed to be homogeneous within each area stratum (Bishop, 2006); areas should be sufficiently small and numerous to accommodate strong spatial patterns in CPUE. A time + area model assumes that the temporal variability is the same in each area and only the means differ. We refer to this approach as the 'GLM' approach (Table 2). Often, the GLM employed for CPUE standardization includes a time-area interaction to account for potentially different temporal trends among area strata (e.g., Campbell, 2004; Carruthers et al., 2011; Nakano, 1998). The resulting index should be calculated by weighting the effect for each area by the size of that area. However, different trends among areas (i.e., the year-area interaction is significant) may imply that a spatial stock assessment model is required. Alternatively, the time-area interaction can be integrated into a GLMM as a random effect (Chang, 2003; Maunder and Punt, 2004; Miyabe and Takeuchi, 2003), which treats it as a nuisance parameter; we refer to this approach as the 'GLMMint' approach.

Even though the 'GLM' and 'GLMMint' approaches represent (simple) ways to consider space in CPUE standardization, they come with the issue of defining the area strata, which can substantially affect model performance. Ideally, data are rich enough to define fine-scale spatial

Table 2
Overview of the catch-per-unit-effort (CPUE) standardization approaches with spatial considerations, along with a selection of R packages mentioned in the text.

Approach	Overview	R packages
GLM	Method using generalized linear models (GLMs) that integrate fixed year and area effects.	*stats::glm, mgcv, brms, TMB*
GLMMint	Method using generalized linear mixed models (GLMMs) that integrate fixed year and area effects and a random year-area interaction term.	*TMB, brms, nlme, lme4, glmer, glmmTMB, mgcv*
Spatial GAM	Method using generalized additive models (GAMs) that integrate an interaction term between longitude and latitude representing spatial variation (long-term latent variation) at a broad spatial scale.	*mgcv, brms, TMB*
Spatiotemporal GAM	Method using GAMs that integrate an interaction term between longitude, latitude and year representing spatiotemporal variation (latent variation that changes over time) at a broad spatial scale, in addition or in lieu of an interaction term between longitude and latitude.	*mgcv, brms, TMB, R-INLA*
STM	Method using models that explicitly specify a correlation or semi-variance function for latent variables representing unmeasured processes that vary over space and time. These can be specified with similar structure to a spatio-temporal GAM but also allow more flexibility for explicit movement, size/age structure, or other ecological mechanisms.	*TMB, R-INLA, VAST, sdmTMB*

strata that do not constrain estimates. For example, tuna longline fishery CPUE analyses usually define spatial strata as 5° cells (Anon, 2013). However, some degree of aggregation is often required and there are several approaches that can be used. In the 'ad hoc' approach, area strata are the management areas employed in the study region or are based on environmental layers such as bottom depth contours (Forrestal et al., 2019; Huang et al., 2020; Huang et al., 2007). In the 'binary recursive partitioning approach,' an algorithm is used to divide the study region into several area strata in a sequential and recursive manner (Ichinokawa and Brodziak, 2010). Finally, in the 'spatial clustering' approach, a *k*-medoids algorithm is employed to partition a spatial grid covering the study region into several area strata based on the proximity and mean value of CPUE in each spatial grid cell (Ono et al., 2015). Using a simulation experiment based on Pacific saury (*Cololabis saira*) data, Hsu et al. (2022) found that the GLMM relying on the ad hoc approach had the poorest performance and that the GLMMs relying on the spatial clustering approach had the best performance. Moreover, regarding the spatial clustering approach, Hsu et al. (2022) found that, under preferential sampling (a classical situation with CPUE data), assigning equal weights to spatial proximity and mean CPUE values in the *k*-medoids analysis was preferable to giving more weight to mean CPUE values.

A more flexible approach than the 'GLM' and 'GLMMint' approaches to considering space in CPUE standardization is the 'spatial GAM' approach, which consists of fitting a GAM including an interaction term between longitude and latitude (Braccini et al., 2021; Grüss et al., 2019; McKechnie et al., 2013). The spatial interaction term (e.g., a tensor product smooth) depicts spatial variation (long-term latent variation) at a broad scale (Denis et al., 2002; Grüss et al., 2021). The spatial GAM approach can be extended to a 'spatiotemporal GAM' approach via the inclusion of an interaction term between longitude, latitude, and time, which represents spatiotemporal variation (latent variation that changes over time) at a broad scale (Hoyle, 2020; Zhou et al., 2019). By accounting for spatial and/or spatiotemporal variation at a broad scale, such GAMs can generate more stable predictions for spatial areas where data are sparse (Hoyle, 2020; McKechnie et al., 2013). Using a simulation experiment based on blue marlin data from the Atlantic Ocean, Grüss et al. (2019) found that spatial GAMs tended to outperform simpler CPUE standardization models.

STMs constitute an even more flexible approach for CPUE standardization by depicting spatial and spatiotemporal variation at a very fine scale, thereby yielding very precise estimates (Anderson et al., 2022; Shelton et al., 2014; Thorson et al., 2015). Over the recent years, STMs implemented with R package *VAST* (Thorson, 2019a) have been increasingly employed to standardize CPUE data (e.g., Cao et al., 2017; Ducharme-Barth et al., 2022; Grüss et al., 2019; Kanamori et al., 2021; Xu et al., 2019). Because STMs represent latent spatial and spatiotemporal variation at a very fine scale, the inclusion of habitat variables that are inherently spatial or spatiotemporal (e.g., sea surface temperature) in these models generally does not improve their predictive capabilities (e.g., Han et al., 2021; Hsu et al., 2022; Thorson, 2015). Using a simulation experiment based on skipjack tuna (*Katsuwonus pelamis*) data for the western and central Pacific, Ducharme-Barth et al. (2022) found that including local environmental covariates or regional oceanographic indices in STMs did not meaningfully improve model performance beyond what was achieved using spatiotemporal random effects, and even degraded model performance in some cases. However, we note that other STM case studies have suggested small but important improvements resulting from including density covariates (Thorson, 2019b). They may have the most influence when predicting into areas with poor sampling coverage.

The simulation experiment conducted by Grüss et al. (2019) indicated that, overall, the STM approach implemented using VAST performed better than simpler approaches, including the spatial GAM approach. However, in some instances, the VAST and spatial GAM approaches performed similarly or the spatial GAM approach performed slightly better (Grüss et al., 2019). Spatiotemporal GAMs are often easier to apply for exploration and may have the flexibility to fit model structures that are unavailable in the VAST framework. Thus, we suggest that the analyst should not assume a priori that STMs will outperform the spatial or the spatiotemporal GAM approach, or even the GLMMint approach, in their case study. Instead, the analyst should ideally carry out a simulation experiment to better understand the capabilities of STMs in their case study under different scenarios. Such simulation experiments permit evaluation of the accuracy, error, and confidence interval coverage of STMs compared to simpler approaches. When operating models are not available to generate simulated data, the fish abundances needed for the simulation experiment can be obtained by fitting an STM without any vessel effect or other catchability effects, as in Thorson et al. (2015) and Hsu et al. (2022).

When the stock of interest is made of several sub-stocks, analysts may develop a separate CPUE standardization model for each one (McKenzie and Parsons, 2012). This approach is not necessarily warranted with STMs. Specifically, the fine-scale spatial areas occupied by each sub-stock can be identified in the STM input, and the STM can then compute an abundance index for each sub-stock (Grüss et al., 2023a; Thorson, 2022). Even in the presence of stock-structure, a single STM that appropriately weights composition data spatially by the estimated CPUE may be able to implicitly adjust for any spatial structure in the stock and the fishery (Maunder et al., 2020). However, further research is needed to explore this issue.

If covariate effects or data availability vary spatially or through the time series, it may be difficult to develop a single STM that includes all influential covariate effects. It is important to include relevant covariate effects since models that fail to include them will often produce biased indices of abundance. In general, analysts should explore multiple modelling approaches to ensure that the resulting indices are consistent with the best available information about abundance trends.

Interestingly, instead of being employed for standardizing CPUE, STMs can instead be used to identify stock structuring (Lindegren et al., 2022), investigate existing stock structure hypotheses (Grüss et al., 2023a), or evaluate the impacts of modifying sub-stock spatial boundaries. They can also be applied to estimate relative population scale across regions for use in multi-region stock assessments (Ducharme-Barth et al., 2020; Hoyle and Langley, 2020). Inclusion of density covariates when using an STM to develop relative population scale across regions (i.e., 'regional scaling') should be carefully considered given that they can influence predicted values in poorly sampled areas of the model domain (Ducharme-Barth et al., 2020). Given the influence of 'regional scaling' in multi-region, spatially explicit stock assessments, we recommend developing and exploring alternative CPUE index scenarios with different implied 'regional scaling'.

When choosing the sampling frame for the CPUE analysis, analysts must consider whether to include regions at the edge of the fishery that are fished infrequently. It may be uncertain whether these regions are unfished due to low CPUE or for unrelated reasons (e.g., they are too far from port). They may also be subject to spatial changes in stock distribution due to environmental events. Including such regions in the modelling frame will increase uncertainty, as the abundance in such regions will be inferred from limited information. One option is to limit the spatial domain to a 'core' fishery defined as those regions where most of the catch is taken, or that are fished for a majority of the modelling period (Campbell, 2015; Grüss et al., 2023b; Xu and Lennert-Cody, 2022). For example, Campbell (2015) limited the spatial domain to those one-degree squares where the total catch of swordfish was greater than 500 fish over the period modelled, while Xu and Lennert-Cody (2022) defined a core fishing ground for skipjack for the floating object and unassociated fisheries in the eastern Pacific Ocean as all one-degree squares with at least 11 and 6 years of CPUE data between 2000 and 2021, respectively. Sensitivity of the resulting abundance index to different spatial domain definitions should be considered, as limiting the analysis only to a core region could mask hyperstability, or changes in the species' range.

Non-stationarity of the spatiotemporal correlation also needs to be considered. Rates of spatiotemporal correlation are unlikely to be constant either within habitat types or between them. However, approaches that allow nonstationary spatiotemporal correlation are rare, and it may be more practical to do separate analyses for different areas (i.e., habitat types) and subsequently compare the estimates of correlation structure to determine whether combining them with a stationary correlation structure is appropriate. If not, then the results of the separate analyses may need to be appropriately combined.

With fishery-dependent data, there is a growing need to account for preferential sampling (the likely correlation between sampling location and abundance) which can bias indices that fail to account for it (Ducharme-Barth et al., 2022). Preferential sampling can be accommodated using joint models for sampling intensity and density (Alglave et al., 2022; Conn et al., 2017; Pennino et al., 2019; Rufener et al., 2021) which make the strong assumption that abundance should approach zero in un-sampled areas. Research is needed to develop approaches for modelling preferential sampling in a CPUE standardization framework either as an inhomogeneous Poisson process (Diggle et al., 2010) or using random-utility models to consider fisher sampling location due to economic or regulatory factors (Girardin et al., 2017).

Finally, the spatial dynamics of a fishery may change significantly over time (for example, effort may contract spatially in response to economic or management factors while the stock may spatially contract in response to overfishing) and relative abundance indices based on catch and effort data can become biased unless consideration is given to such changes. Campbell (2016) developed a general framework for indices of stock abundance which uses alternative hypotheses to consider uncertainty about how to structure the analysis, particularly concerning the spatiotemporal dynamics of the fishery.

5.3. Multispecies targeting

CPUE standardization aims to account for the non-random catchability change caused by targeted fishing operations, which typically seek to optimize their profits. This is further complicated in multispecies fisheries in which fishers can make discrete operational choices to increase the catchability of specific species or species complexes. Allocating targeted effort to one species over another can be related to the choice of fishing-ground, habitat-type, fishing-technique, or the way the gear is deployed (Palmer et al., 2009; Pelletier and Ferraris, 2000; Winker et al., 2013). In the best case, unaccounted variation in targeting only inflates the uncertainty in the index of abundance, but large short-term shifts or long-term trends in targeting can systematically change catchability and, therefore, severely bias abundance indices. The effects on CPUE of variability in targeting must be accounted for to stabilise catchability and estimate reliable abundance indices for multispecies fisheries. Targeting may include the choice to fish at the season and locations in which high catches of a sought-after target species are expected. Such target changes can be effectively accounted for with adequate STMs (Thorson et al., 2016b). However, other operational adjustments are non-spatial (e.g., bait type, fishing depth, gear deployment) or occur at much finer spatiotemporal levels (e.g., dynamic temperature gradients, time of day, or habitat features) than are reported. For example, many small-scale and recreational hook and line fisheries report spatial information at the coarse scale of a single location, whereas multiple fishing locations and fishing techniques may have been employed during the same trip. Information about targeting strategies may be unreported, or reported infrequently and inconsistently among skippers, locations, and time periods. In other words, the factors that might indicate the target/fishing strategy are latent variables. Consequently, finer scale variations in targeting are often unobservable and are hereafter referred to 'fishing tactics' (Okamura et al., 2017; Thorson et al., 2016b; Winker et al., 2013). The principal idea behind the various approaches employed to account for these latent changes in fishing tactics is to make use of the multispecies information contained in the catch and effort data.

Subsetting approaches aim to select records where effort was likely allocated towards the species of interest (Hoyle et al., 2022; Stephens and MacCall, 2004). These include using catch proportions by species to determine a threshold for subsetting the data (Biseau, 1998; Helle et al., 2015; Klaer and Smith, 2012) to only include records from a core area or specific fishing season (Hoyle et al., 2022). However, such subsetting can be sensitive to the subjective choice of the threshold and associated with risks introducing hyperstability into the standardized index. For example, the effect of subsetting to non-zero catches can be illustrated via a delta-lognormal model in which the binomial distribution model is used to estimate the encounter probability, and the lognormal model is employed to estimate the scale of positive catch rates. For any situation in which the estimated indices from both model components show a unidirectional trend, an index based on a subset of positive CPUE alone would be hyperstable relative to standardized CPUE from the delta-lognormal model. Unsurprisingly, simulation experiments yield biased abundance trends when zeros are excluded, with the exception of those species that are ubiquitous and occur in almost all (> 95%) of the CPUE records (Langley, 2019). This bias is likely increased if thresholds are set higher (e.g., 30–50% of the catch weight) and most severe for less abundant species given that encounter probabilities below 30% become approximately proportional to abundance (Kerwath et al., 2019). Similarly, a species with declining abundance may contract towards its core area where CPUE may then remain hyperstable compared to the overall decline (Harley et al., 2001; Thorson et al., 2016c). Optimally, the target species would be known for a random subset of the data, and this could be used to determine the best model to predict targeting for the trips with unknown target.

Stephens and MacCall (2004) applied a logistic regression of multispecies presence–absence information to subset trip records to locations (habitats) where the species under assessment was likely to be present but may not have been caught ('true' zeros) and locations where the species was unlikely to occur ('false' zeros). This approach therefore aims to retain 'true' zeros in the abundance trend. By design, this approach is most suitable for species that co-occur within multispecies assemblages, where the mere presence-absence of co-occurring species can provide a strong predictor of habitat suitability. Another approach is to identify 'indicative' vessels based on vessel characteristics and catch history of the species of interest (Helle et al., 2015; Punt et al., 2000). This approach requires a detailed understanding of the fleet and assumes that the indicative vessels employ consistent fishing tactics. The selection criteria for indicative vessels rely on a degree of subjectivity which may affect the CPUE indices (Helle et al., 2015). Using indicative vessels alone may be insufficient for CPUE standardization if vessels change targeting over time. For example, skipper behaviour may be more influential than the vessel effect (Palmer et al., 2009), but economic (e.g., fish price, market demands, fuel costs) or regulatory factors (e.g., quota limitations, area closures, or by-catch restrictions) may also lead to changes in targeting (Abbott et al., 2015).

As an alternative to subsetting, several approaches have been proposed that use covariates to account for variations in fishing tactics. One approach is to use the catch rates of alternative target or bycatch species as covariates to correct for the effort directed away from the target species or species under consideration (Glazer and Butterworth, 2002; Su et al., 2008). Although the catch rates of alternative species do not hold direct information about the catch of the species of interest, the information in the predictor variables derived from these covariates is not entirely independent of the response CPUE and may have unpredictable impacts on the standardized CPUE trends. Another approach is to derive categorical covariates from ranked catch ratios of two species as an indicator of targeting strategy. These catch ratios, which often include the species of interest, are then grouped by percentile frequency in descending order, either over the entire time series (e.g., Mejuto et al., 2009) or by year (e.g., Hiraoka et al., 2012). Each percentile group represents a factor level of the categorical variable. However, if the ratios are grouped over the entire time series, then there will inevitably be confounding effects between changes in abundance and the catch ratio grouping. If the abundance of the species of interest changes over time, the true trend abundance tends to be removed from the CPUE index, but if the other species changes, this can introduce a biased trend into the abundance index for the species of interest (Chang et al., 2011). Further, such bias is likely to be aggravated if individual CPUE records are aggregated. However, if the ratios are grouped by year, it can be demonstrated that the catch ratio covariate has no influence on the year-effect of interest and is, thus, completely ineffective for removing targeting-induced variation from the index of abundance (Hoyle et al., 2014b).

If discrete fishing tactics result in distinctive species composition in the catch, one can employ clustering techniques to categorize multi-species CPUE records into groups with similar catch compositions (He et al., 1997). The identified clusters are assumed to represent fishing tactics which may be treated as categorical variables in the standardization model to adjust for differences in catchability associated with each cluster (He et al., 1997). This two-stage approach has been widely applied for tuna CPUE standardization of large pelagic longline fisheries (Carvalho et al., 2010; Hoyle et al., 2022; Hoyle et al., 2015b). A related two-stage ordination approach is the 'Direct Principal Component' (DPC) procedure (Winker et al., 2013; Winker et al., 2014), which uses continuous principal component scores (PCS), derived from a Principal Component Analysis (PCA) of the catch composition data, as either linear or nonlinear predictor variables to adjust for the effect of variations in fishing tactics. The DPC method relaxes the assumption of fishing tactics being discrete and simulations indicated that the DPC can also perform adequately if CPUE records originated from mixtures of fishing tactics (Winker et al., 2014). Both ordination approaches can provide valuable insights into the targeting dynamics. For example, clusters or PCA species loadings can be visually explored at different spatiotemporal scales (e.g., seasons, area) or by individual vessels. However, several statistical caveats require careful consideration. If the species of interest represents a dominant component of a fishing tactic, it typically needs be included in the catch composition used to identify targeting, and the information contained in the resultant predictor variables composition is not entirely independent from the response in the standardization model and is, thus, not strictly orthogonal. Therefore, additional transformations (e.g., arcsine-square-root or fourth root) of the species composition data have been recommended (Campbell et al., 2017; He et al., 1997; Winker et al., 2014). Simulation testing has indicated that this confounding effect was not a major concern for multispecies scenarios of moderate complexity (He et al., 1997; Winker et al., 2014). A drawback of this non-independence is that model selection criteria (e.g., AIC, Bayesian Information Criterion (BIC)) are unlikely to be appropriate since they select too many non-meaningful targeting signatures, leading to confounding and over-precision in the abundance trend (Winker et al., 2014). This makes it generally challenging to identify meaningful interactions between targeting effects and other covariates. Considering that several subjective decisions must be made (Campbell et al., 2017), the identified targeting signatures should be consistent with prior knowledge that distinct targeting practices exist. Otherwise, there is a risk of mistaking spatiotemporal abundance patterns for targeting,

Recently, two approaches have been proposed to overcome the confounding effects of the approaches discussed above (Okamura et al., 2017; Thorson et al., 2016b). Both infer the latent fishing tactic from the residual structure. Okamura et al. (2017) proposed the two-stage 'directed residual mixture' (DRM) approach which entails first fitting a standardization model to each species with all covariates other than fishing tactics, then applying a Gaussian mixture model to estimate discrete latent factors for fishing tactics from the resulting mixing proportions of residuals. In the third stage, the estimated components are included as additional factorial covariates in a final standardization mode for the individual species. DRM provided an unbiased estimator for a deterministic proof-of-concept scenario under somewhat idealized conditions with two species and two fishing tactics for which previous approaches failed (Okamura et al., 2017). However, there is a lack of follow-up case studies on the efficiency of the DRM in the presence of zeros and over-dispersion, which are common features of real-world datasets. Moreover, the 'spatial dynamic factor analysis' (SDFA) (Thorson et al., 2016b), which is now embedded in the R package *VAST*, presents a statistically coherent method for simultaneously estimating spatial-temporal variation, random vessel effect, fishing tactics, and relative fish abundance for multispecies species within a single model where the latent fishing tactic effect is estimated from remaining unexplained residual correlations among the multispecies catch rates. Simulations demonstrated that SDFA performs well when targeting is linked to the spatial allocation of fishing effort. However, the fishing tactics model showed limited ability to correct for targeting effects in cases where spatial reporting was aggregated to a coarser resolution than the 'true' simulated dynamics or if spatial information was absent. Further research is needed to better understand the data requirements for SDFA and to evaluate the impacts of missing covariates on DRM's ability to account for fishing tactics.

5.4. Error distributions

As described in the Maunder and Punt Sections 2.1, 3.1, 3.3 and 4.3) (2004) review, it remains important to select an error distribution and a modelling approach that match the structure of the data. The analyst should first determine an appropriate class of error distribution for the data. An appropriate error structure is needed to accurately represent variance around the standardized index and to determine the trend (Dick, 2004) and scale (Thorson et al., 2021) of the standardized index.

For discrete observations (e.g., catches recorded as counts), the Poisson distribution may be suitable if the variance is approximately equal to the mean although, in practice, this is rarely the case. If variance is larger than the mean, then the negative binomial distribution may be appropriate (Walsh and Brodziak, 2015), while the flexible Conway-Maxwell-Poisson distribution can account for either over- or under-dispersed data (Lynch et al., 2014). Additionally, if data are collected discretely, then effort can be included as an offset (Maunder and Punt, 2004), or the numerical catch can be converted to catch rate (e.g., number of fish caught per 1000 hooks), which can be modelled continuously. If the numbers of individuals caught per observation are large, then they can be approximated as continuous variables (Maunder and Punt, 2004).

Continuous observations (e.g., catch per unit effort or catch recorded as biomass) are typically right-skewed; an appropriate distribution choice can be informed using Taylor's power law (Taylor, 1961), $\mathrm{var}(Y) = \alpha\mu_Y^p$. If the variance is proportional to the square of the mean ($p = 2$), then the lognormal or Gamma distribution may be appropriate and, if proportional to the cube of the mean ($p = 3$), the data might be fitted using the inverse Gaussian distribution model (Dick, 2004). However, a recent study tested the different distribution models in STMs and found that the Gamma and Tweedie distributions provided the best performance for estimating index scale, followed by the lognormal, and the inverse Gaussian performed worst (Thorson et al., 2021). Alternative distributions may be distinguished using residual analysis, mean-variance plots, and Akaike's Information Criterion (AIC) (Dick, 2004).

When selecting the error distribution, the proportions of zero observations in the data should be scrutinized. If there are more zeros than predicted by the distribution, they may be dealt with using a mixture modelling approach. A discrete variable with excess zeros is typically fitted with a zero-inflated model which uses a separate likelihood component to account for the extra probability of a zero observation (Lambert, 1992; Minami et al., 2007; Walsh and Brodziak, 2015; Zuur et al., 2012). Given a continuous response variable, excess zeros can be accommodated using a two-stage hurdle or delta modelling approach where catch-rate is modelled conditional on a positive encounter (Lo et al., 1992). The Tweedie distribution can take the shape of several distributions in the exponential family depending on the value of p. When p is in the range ($1 < p < 2$), then the Tweedie is equivalent to a compound Poisson-Gamma distribution which can simultaneously account for zero observations and positive catch (Shono, 2008a). Applying either a Poisson-link delta-gamma or Tweedie model structure may be a sensible default option (Thorson et al., 2021).

In recent years, research has focused on more appropriately modelling and accounting for correlation in the data generation process. These approaches should be incorporated into routine CPUE standardization analyses where appropriate. 'Poisson-link delta models' represent a computationally efficient alternative to the compound Poisson-gamma distribution and can explicitly account for potential positive correlation between encounter rate and positive catch (Thorson, 2018).

Newly developed flexible modelling approaches (e.g., VAST; Thorson 2019) more readily model multivariate data within a CPUE standardization framework. Multivariate approaches can account for correlation between modelled categories such as species (Thorson and Barnett, 2017; Thorson et al., 2016a); or age/length/sex/stage categories (Kai et al., 2017; Maunder et al., 2020). Accounting for correlation between species can improve predictions of spatial density for rare species (Thorson and Barnett, 2017), and accounting for correlation across size classes can allow for size-class specific indices (Kai et al., 2017) or the standardization of composition data (Maunder et al., 2020; Thorson and Haltuch, 2019).

5.5. Uncertainty estimation

The uncertainty associated with the CPUE indices used in a stock assessment includes observation error associated with the estimated time effects in the index and process error associated with variation in catchability (the relationship between the time effect and abundance) (Francis et al., 2003).

Observation error in the temporal effects can be estimated as part of the model fitting process and constructing the index, as discussed in Section 5.8. Many GLM analyses estimate categorical time effects relative to the base time (e.g., the first year in the time series). These can be recalculated as canonical confidence intervals to associate uncertainty with all temporal effects (Francis 1999). Observation error for two-step delta or hurdle models can be estimated using parametric (Shono, 2008b) or bootstrap (Ichinokawa and Takeuchi, 2012) methods, and subsequently applied within the predict-then-aggregate process. Alternatively, uncertainty can be characterised by draws from the posterior predictive distribution if the model is fitted in a Bayesian framework (e.g., using the R package brms; (Bürkner, 2017)).

However, standardization usually underestimates observation error in the temporal effects because it does not fully account for dependencies in the input data (Hoyle et al., 2014b). Dependencies occur among consecutive sets by a vessel and among sets by vessels from the same company which may communicate with one another. There is also overdispersion due to both aggregation of fish and the unavailability to the analyst of important factors affecting catch rates, such as local environmental factors that affect fish distribution. Some of this overdispersion can be accounted for by including random effects in the CPUE standardization model (Rufener et al., 2021; Thorson and Minto, 2015; Xu et al., 2019). Many sources of catchability variation are not amenable to standardization (Wilberg et al., 2009). As a result, uncertainty estimates from CPUE standardization models (observation error) are often seen as, at best, useful for suggesting relative uncertainty among year effects. They can also be useful as an estimate of the minimum uncertainty, and to determine relative uncertainty between indices.

Process errors occur when the average catchability across the fleet varies between time intervals. Effectively, the relationship changes between the abundance index and the true abundance leading to more uncertainty in the abundance index than predicted by the observation error. This variability cannot be measured directly but can be estimated indirectly with stock assessment models (Francis et al., 2003). This variation is often estimated or assumed to be larger than the observation error in the time effects (Maunder and Punt, 2004), particularly for large industrial fisheries with low observation error due to high effort data sample sizes.

5.6. Diagnostics

Traditional model validation diagnostics such as residual analysis are important for determining if the analyst has specified the model correctly, treated zero observations appropriately, and selected the proper error distribution (i.e., the model assumptions are not violated). However, non-Gaussian error structures, mixture (e.g., delta or zero inflated) distributions, and/or mixed-effects frameworks commonly applied in contemporary CPUE standardization analyses complicate using traditional diagnostic approaches. One solution is to define residuals as probability-integral-transform (PIT) residuals (Warton et al., 2017). PIT-residuals are generated in a 'model-free' bootstrap by comparing observations to a distribution of predicted values for the given observation generated from the fitted model. Prior to calculating the PIT-residual, a simple check of the distribution of the observations against the distribution of the bootstrapped/predicted values (e.g., Parker et al., 2017) can be used to identify if the distributional assumptions are being met (e.g., are coverage levels similar?). R package *DHARMa* (Hartig, 2020) is useful for calculating PIT-residuals and has several tests for assessing if observations match the distributional assumptions of the model, outlier detection, identification of over- or under-dispersion, and zero-inflation. In a mixed effects framework, if models are solved using the Laplace approximation to integrate across the random effects (e.g., TMB; Kristensen et al., 2016), the reduced Laplace approximation for strongly non-linear models may limit the utility of PIT-style residuals (Thygesen et al., 2017). Accuracy of the Laplace approximation can be determined by refitting the model to simulated data (Rufener et al., 2021; Thygesen et al., 2017). Alternatively, residuals can be calculated using the posterior estimate from MCMC samples (Rufener et al., 2021) or by using one-step predictions (Thygesen et al., 2017). One-step prediction residuals are a useful tool for validating model's fit to sequential data such as CPUE analyses, where an observation Y_t is compared to a prediction in time t, $\widehat{Y_{t|t-1}}$, conditioned on data up to $t-1$.

In addition to model validation, diagnostics can be used to develop intuition on how the CPUE standardization model is transforming the nominal index. Step plots (sequentially plotting how the index changes with each additional covariate or model component) show incremental changes to the index (Bentley et al., 2012). However, if there are interactions between covariates, then their impact could be masked within a step plot. Influence and coefficient-distribution-influence (CDI; Bentley et al., 2012) plots are useful in combination with step plots. CDI plots quantify the relative influence of each covariate on the final fitted model by incorporating both the estimated coefficients for individual covariates and the sampling across covariate levels each year. Influence and CDI plots can be readily created for fixed effect GLMs using the *influ* package (https://github.com/trophia/influ) in R. Hsu et al. (2022) extended the *influ* package to STMs implemented with the R package *VAST*. Applied in a spatial or spatiotemporal modelling framework, a large annual influence value for the year-area interaction term could indicate a misspecification of the spatial structure that failed to account for a large shift in spatial sampling (Hsu et al., 2022). Counterfactual analyses (Hansell et al., 2022; Pearl, 2009) can also be employed to evaluate the impact of key structural assumptions (e.g., fitting a model with spatial random effects held at zero and comparing the resultant index with a model freely estimating spatial effects in order to identify their impact on the standardized index).

Additional diagnostics can be applied to specifically interrogate STMs. Calculated residuals should be evaluated for temporal, spatial (i.e., Moran's I test), and spatiotemporal patterns. In addition to being useful for evaluating model performance, cross-validation can be used to identify spatial outliers (Conn et al., 2018; Marshall and Spiegelhalter, 2003). Spatial and spatiotemporal random effects are often assumed to be normally distributed with mean-zero. The estimated random effects should be checked to ensure that they match the assumed distributions. Lastly, STMs can predict abundance (or density) into un-sampled areas. Predictions in un-sampled areas will be influenced by adjacent observations given the modelled spatial/spatiotemporal covariance structure, and by any modelled relationships with density covariates. These predictions should be scrutinized for consistency with 'common sense' expectations of what abundance should be given understanding of the species biology and fishery dynamics. Departures from expectations may indicate the need to change how the model makes predictions into unfished areas, such as by applying a preferential sampling model or changing other aspects of the model structure to account for unfished areas caused by economic or regulatory factors (e.g., with the use of a random utility model).

Hinton and Maunder (2004) propose an 'omnibus test' that simply compares the total likelihood from the stock assessment model for all data types combined and uses one index of abundance versus using an alternative index of abundance. This test measures which index of abundance is most consistent with the stock assessment model and the other data. We recommend further research regarding omnibus-tests when specifying CPUE standardization models.

5.7. Model selection

Model selection is an important part of the analysis process. Analysts often employ automated procedures such as stepwise forward or backward selection algorithms, although these can be problematic (Sribney, 1996; Wiegand, 2010). Throwing all covariates into an automated selection procedure is unlikely to identify the most appropriate model and neglects an opportunity to generate useful understanding. Henderson and Velleman (1981) noted that automated techniques often hide important features of the data from the analyst and provided the axiom that "the data analyst knows more than the computer." Modelling data interactively can change understanding and lead to different inferences.

The analyst should first consider the objective of their CPUE standardization. When generating an index rather than making inferences about the covariates themselves, the goal is to correct for covariate influence and avoid potential bias in the index (i.e., covariates that explain annual variation in abundance should not be used to explain catchability). The focus on prediction rather than estimation has implications for model selection, making AIC an appropriate tool if statistical assumptions are met (Aho et al., 2014; Akaike, 1973). CPUE analyses for understanding are also important for stock assessment but have different objectives and may use different analytical approaches.

Model selection and inference based on information criteria are well-established with clear guidelines available (e.g., Burnham and Anderson, 2004). These include identifying plausible predictor variables and combinations thereof, evaluating model performance, and averaging across plausible models. Evaluating model performance should include considering the goodness of fit of the selected models (Mac Nally et al., 2018).

The analyst should determine whether a parameter is likely to affect CPUE, a priori, and whether it affects density or catchability. Starting with a set of plausible models saves analysis time and reduces the risk of overfitting (a.k.a. data dredging or *p*-hacking).

Predictive cross-validation methods can be useful for selecting between models with different structures (e.g., Charsley et al., 2022; Maunder and Hinton, 2006; Shono, 2008a). They can be time-consuming but are very flexible and broadly applicable.

Model selection based on likelihood is affected by lack of independence. Operational data usually represent a time series of fishing events by the same vessels, and there may be information sharing between vessels. The 'intra-vessel', 'intra-trip', or other correlation can be represented by treating that factor (e.g., vessel or trip) as a random effect. Such a treatment is appropriate to account for the reduction in effective sample size resulting from simple correlation in the observations. However, random effect distributional assumptions may be inaccurate given, for example, trends in fishing power or targeting through time.

Operational datasets can have very high sample sizes, and that combined with lack of independence can make almost any variable statistically significant based on likelihood, even if it has negligible influence on the year effect. Since the goal is to obtain an index of abundance, including a covariate that neither explains much deviance nor influences the index of abundance is not usually a problem. However, alternative data selection criteria are often used to help develop models that are simple, manageable, and fast to compute.

When the objective is accurate prediction for the year effect, a primary consideration for variable selection is influence. Influence is affected by a combination of the extent to which the covariate affects the response and the extent to which its value changes through time. For example, a variable that affects catch rates but is completely balanced (such as moon phase in many cases) may have no influence on the resulting abundance index. Such variables can be useful nonetheless if they improve precision and help estimate other effects. They are also useful for understanding the biology of the species and the nature of the fishery, an important benefit of modelling CPUE.

One commonly employed criterion is the proportion of deviance explained by a variable (R^2). This criterion is useful because it is related to influence and robust to lack of independence, but the appropriate threshold level will be case-dependent. For example, covariates will explain a higher proportion of residual variance when data are more aggregated, allowing the threshold to be set at a higher level. However, conditional R^2 (and consistent AIC) showed inconsistent model selection in the simulation experiment conducted by Hsu et al. (2022).

Sometimes it is important to retain covariates that do not meet selection thresholds because it is clear a priori that they will affect CPUE. For example, catch rates almost always vary consistently between vessels, and the vessels in the fleet usually change through time. Analysts may, therefore, assume that catchability varies among vessels even if it does not reach the model selection threshold and should confirm that vessel ID is not influential before dropping it from a model.

Some tools have internal automatic variable selection, such as the GAMs implemented with R package *mgcv*, which can automatically choose the degrees of freedom for a spline and can be configured to shrink effects to zero simultaneously without the need for a stepwise process (Marra and Wood, 2011). Model selection in R package *mgcv* is based on prediction error criteria or likelihood-based methods, with the likelihood-based methods less prone to local minima (Wood, 2017).

5.8. Assembling an index from a fitted model

Previous reviews of CPUE standardization typically assumed that the index is constructed by first fitting a statistical model to available data (see Section 5.1) and then extracting a coefficient that represents the partial effect of year to use as the index; we call this the 'year-effect as 'index' method in the following. For example, Maunder and Punt (2004) state: "Most methods used to standardize catch and effort data estimate a year effect on which an index of abundance can be based." However, the 'year-effect as index' method has many limitations. Most generally, treating the partial effect of year as the index does not include the effects of other covariates that vary among years. One interpretation of this practice is that all the other covariates are implicitly treated as 'catchability covariates', and their effect is filtered out when treating the year effect as the abundance index.

One special case of concern arises when including a year x area interaction as a predictor variable in the CPUE-standardization model. Maunder and Punt (2004) dealt with this topic in detail and outlined a few alternative treatments, including averaging across year x area estimates or fitting separate models to different areas. However, research since then has converged upon a generic approach which we here call the 'predict-then-aggregate' method for index construction (Walters, 2003 provides a good foundation for the method). This approach is based on sampling theory and involves the following steps.

1. Explicitly define a sampling frame that ideally corresponds to the stock being assessed. Sampling units in the frame might be defined spatially (i.e., subdividing the stock range into a set of non-overlapping strata) or via other partitions (i.e., port).
2. Predict CPUE that would have occurred for each sampling unit, conditioning upon values of density covariates that occur at each sampling unit and dropping the partial effect of catchability covariates (e.g., by setting them at a defined reference level). Note that predictions of CPUE into un-sampled sampling units should be carefully scrutinized and considered (see McKechnie et al., 2013; Walters, 2003). For example, predictions may need to use one approach for areas that are un-sampled because low catch rates make fishing uneconomic and a different approach for areas where management measures exclude fishing effort.
3. Aggregate across sampling units which can often be done by taking the area-weighted (i.e., the area of available habitat) sum of CPUE across these units.
4. Calculate the variance of the aggregate, either using the delta method, quantiles from MCMC samples in a Bayesian model, nonparametric bootstrapping of the data, or other methods. We expect that these methods will generally give similar results (Magnusson et al., 2013), but emphasize that it is necessary to propagate information about covariance among predictions to correctly calculate the variance of the 'predict-then-aggregate' approach.

For examples of this approach see Campbell (2004) and Campbell (2015).

Area weighting should, where possible, use the area of available habitat in each spatial cell (Maunder and Punt, 2004), assuming that density is uniform within each cell. For pelagic fishery CPUE, approximate ocean areas are often determined from the area of grid cells at a latitude, after subtracting any land area. However, the habitat available to individual species in the pelagic ocean can vary seasonally, with environmental variables, and long-term with climate change (Goodyear, 2016). Inclusion of oceanographic 'density' covariates in the standardization model and or applying a post-hoc environmental filter can be used to restrict index calculation to viable cells/spatial areas. Methods for determining these areas are an area of active research. Similarly, cells that are too large can cause bias when they include areas of both high and low density, with fishing concentrated in the high-density areas. Uncertainty about habitat areas is not often considered but can be included via alternative scenarios or Monte Carlo simulation.

Habitat areas in benthic and reef fisheries are often difficult to determine, and uncertainty can be large. As a result, strata are sometimes weighted by catch (aggregated across the time series), as a proxy for relative area (e.g., Ralston, 1999). Catch weighting tends to introduce bias because fishing intensity always varies spatially. Higher effort usually occurring in areas with higher catch rates, but some areas may be closed to fishing and others less fished because they are harder to access. Catch weighting is likely to give more weight to areas that are more heavily fished, so will tend to exaggerate depletion.

When the index-standardization model does not include any interactions of year with other covariates, we expect that the 'predict-then-aggregate' method will provide the same index as the 'year-effect as index' method (Campbell, 2015). However, when a density covariate interacts with the year effect (e.g., year x area), or the value of a density covariate varies through time (e.g., temperature), the predict-then-aggregate method is required.

For a response variable that was transformed before fitting the model, predictions must be back transformed before aggregating. For two-stage approaches such as hurdle models, sampling unit predictions are obtained by combining predictions from both stages before aggregating.

For a hurdle model with the binomial component fitted using a logit transformation, the reference levels of the catchability covariates affect the predicted annual probabilities of non-zero catch. This in turn affects the final index because the probability of non-zero catch is constrained by a maximum of 1. When binomial predictions are back-transformed, index variability can be reduced depending on the reference level selected for the catchability covariates. To avoid bias in index trends and variability, analysts can adjust the mean of the index either via the choice of reference levels or preferably by adding a constant to the logit-scale predictions before back-transforming to the probability scale (Hoyle et al., 2022). In most situations, we recommend adjusting the scale so that the annual mean of the binomial component of the index equals the mean of the annual proportions of nonzero catch. The predicted probabilities of non-zero catch may themselves be used in some cases to indicate abundance after transforming (Hoyle et al., 2011).

6. Using CPUE indices in stock assessments

As discussed earlier, each CPUE index in a stock assessment is associated with a fleet. Section 3 mentioned an approach described by Maunder et al. (2020) to jointly standardize CPUE and composition data and use the results to define two separate fleets in the model with different purposes. One fleet is associated with the CPUE index and is assigned composition data weighted by the spatial distribution of the CPUE (representing abundance) and (usually) time-invariant selectivity so that changes in these composition data represent changes in the composition of the underlying population. The other fleet (not associated with an index) is assigned composition data weighted by the time-varying spatial distribution of the catch so that time-varying selectivity can be employed to extract fish of the appropriate size/age composition.

Often two or more indices or other datasets provide information that is to some extent in conflict. The 'law of conflicting data' (Maunder and Piner, 2017) states that since data are facts, conflict between datasets implies model misspecification. If information from different sources is in conflict, the analyst should try to identify the cause and resolve it. If resolution is not possible, the analyst should create a set of alternative models that omit conflicting datasets; in each model the remaining datasets are well fitted (Francis, 2011; Schnute and Hilborn, 1993). When conflicts occur, there may be a need to weight the alternative hypotheses, which introduces some subjectivity. In general, information about population trends in CPUE should have priority over information in the composition data (Francis, 2011) because it is usually more reliable.

It is a relatively common practice in some assessment cultures to split CPUE series by time into separate non-overlapping sections and to include each index in a separate fishery. Splitting can occur in response to changes in availability of species-specific catch data (e.g., North Pacific blue shark, ISC Shark Working Group, 2022), evidence of changing operational patterns that potentially change catchability or selectivity that can't be included in the CPUE standardisation, or changes to logbooks such as the addition of new data fields (e.g., North Pacific striped marlin, ISC, 2019). However, splitting is often counterproductive and should be avoided where possible unless there is another reliable index of abundance to cover the gap (e.g., Hoyle et al., 2012; Hoyle, 2011). Splitting the time series can waste much of the abundance information in the CPUE data and often substantially changes model outcomes. At best, it increases uncertainty, but it can introduce considerable bias. This should only be the case if something in the model is misspecified (e.g., misspecification of growth and the associated fits to the length composition data which affect model scaling). However, in practice, assessment models always include some degree of misspecification, such as conflict across the time series between different periods of composition data. Without the stabilizing constraint through time provided by the continuous index, the model estimates abundance scale independently before and after the split by adjusting catchability. Splitting the time series therefore often results in abundance changes that reduce the internal data conflict but may be inconsistent with expectations about long-term catchability. At minimum, analysts should consider whether the implied catchability change at the split is consistent with expectations.

There are usually better alternatives to splitting a CPUE series. If the split is due to data availability, analysts should strive to retain the same or similar catchability and selectivity for both indices since there is no evidence that either has changed. If the split is due to catchability changes linked to new technology, analysts should consider whether the changes estimated in the assessment are plausible. If a change that likely increased catchability has the opposite effect in the model, it is usually better not to split the series. If selectivity has changed with effort moving to a spatially or seasonally different part of the stock, a better approach is to analyse the dataset as separate spatial or seasonal CPUE series that each take a consistent part of the stock so that the change in effort distribution does not affect either index. Alternatively, a joint model of composition and CPUE data (Maunder et al., 2020) can be used as described earlier in this section.

Transitions such as these often involve alternative hypotheses about how the changes have affected catchability. Given this uncertainty, a useful approach for the analyst is, rather than splitting the index, to identify a set of plausible hypotheses about how catchability has changed and construct alternative indices based on these hypotheses (Campbell, 2016). This approach has more widespread value because analysts often end up with two or more alternative indices, each of which has support. Given the importance of CPUE indices for determining stock assessment outcomes and the many ways that estimated CPUE indices can diverge from the true abundance trend, it is important to consider the range of plausible options as alternative stock assessment scenarios, preferably within a model ensemble.

Trends in catchability over time can bias CPUE indices and can be of great importance for stock assessments. Increases in catch efficiency, or fishing power, have played a critical role in the history of fisheries (Scherrer and Galbraith, 2020; Squires and Vestergaard, 2013). Wilberg et al. (2009) describe causes of increased catchability such as changes in fishing practices and technology (Garrod, 1964). Technological creep can mask declines in abundance by increasing catchability (Kleiven et al., 2022); and it is observed in almost all analyses involving time series of fishing effort, particularly if the analyses exceed one decade in temporal coverage (Palomares and Pauly, 2019). For example, Squires (1992) found that the most important sources of technical progress in the Pacific coast trawl fishery were electronics, the application of scientific rather than craft principles to vessel and equipment design, and to harvesting methods.

Estimates of mean rates of fishing power increases across fisheries are variable. Eigaard et al. (2014) estimated a mean rate of increase of 3.2% per year, while Palomares and Pauly (2019) reviewed 51 estimates of about 2–4% per year and estimated an expected rate of 1.3% per year for studies that cover a 100-year period, although this is likely to be an underestimate (Scherrer and Galbraith, 2020).

One difficulty in estimating rates of fishing power change is obtaining the requisite data (Scherrer and Galbraith, 2020). A major component of fishing power is the ability to locate fish; though technology supports this in many ways, few are recorded for analysis. Examples include the installation of GPS systems, provision of increasingly informative environmental data from satellites and models, upgraded communication technology to share information between vessels and with fishing companies, and scientific progress in understanding the factors that affect fish distribution, which may derive from public research or from fishers' interrogation of their own stored catch and effort data. Catchability change associated with vessel turnover can be accounted for by including the vessel ID in the model (e.g., Hoyle and Okamoto, 2011), but datasets rarely indicate when a particular vessel installed a particular piece of technology or when the behaviour of a vessel, fishing company, or fleet adapted to use the available information more effectively.

Where catchability increases are considered likely but estimates are unavailable, ignoring them will positively bias stock status estimates (e.g., Han et al., 2023; Ye and Dennis, 2009). Wilberg et al. (2009) recommend a default assumption that catchability varies over time and multiple methods of including time-varying catchability should be applied. To allow for uncertainty about fishing power, stock assessments (particularly for target species) should consider a range of reasonable scenarios regarding long-term catchability trends, from low to high but noting that 0% is rarely plausible.

7. Concluding remarks

CPUE standardization is an influential component of the stock assessment process, as well as a powerful tool for generating understanding of the stock and the fishery. Rather than a statistical exercise, it should be seen as a core part of the fisheries research program. As such, the initial analytical steps of exploring and characterising the dataset, including talking to fishers, are key. In parallel, the understanding developed during CPUE standardization can often motivate changes in the assessment model structure itself, a contribution often more influential for assessment outcomes than small changes in the CPUE indices themselves.

An important development in the field of CPUE standardisation since Maunder and Punt (2004) has been the increase of STM applications to fisheries datasets. However, while STMs are attractive models with powerful features, they do not constitute the best option in all situations. They are also complicated to implement for new analysts because of the different statistical concepts involved. We recommend that simpler approaches be tried first to develop understanding before, if warranted, transitioning to more sophisticated ones (like STMs), if only to understand the influence of spatio-temporal structure on one's catch-and-effort dataset. Of note, a course based on the VAST STM (Thorson, 2022; Thorson, 2019a) is available (https://github.com/James-Thorson/2018_FSH556), as are examples in the VAST Wiki (https://github.com/James-Thorson-NOAA/VAST). In addition, as STMs may be sensitive to model settings and choices, we encourage analysts to consult the growing literature (e.g., Commander et al., 2022; Dambly et al., 2023) describing potential analytic trade-offs. We also emphasize that simulation experiments represent valuable tools for testing and comparing CPUE standardization modelling approaches, so that the analyst may select approaches to produce indices likely to be unbiased. Finally, the development of methods to explicitly model fisher location choice to improve predictions in un-sampled areas would constitute a promising extension to STMs.

In Table 3, we reiterate the major recommendations introduced in the preceding sections. Analysts are encouraged to consider the good practices identified in the 16 areas covered in this review and to apply them where practical. Not all will be feasible to apply to every analysis, but it is important to understand the implications of not applying them.

In Table 4 (based on ISC Albacore Working Group, 2016) we summarize information requirements for presenting the results of CPUE analyses. See also Hoyle et al. (2014a) and IOTC (2015). In Table 5 (based on ISC Albacore Working Group, 2016), we list criteria that can be used to assess the strengths and weaknesses of candidate abundance indices.

Finally, we raise a note of caution about the potential reliability of indices of abundance based on CPUE data. The relationship between CPUE indices and relative biomass or abundance can be proportional, but there are many scenarios in which this relationship changes or

Table 3
Summary of good practices in catch-per-unit-effort (CPUE) standardization.

1. Indices of abundance are often the most influential part of an assessment—invest accordingly.
Data
2. Defining fleets is key. Definitions depend on stock and fishery structure, the type of CPUE analysis (time + space versus spatiotemporal model (STM)), and the assessment approach (biomass dynamic versus age/size-structured, conventional versus index fishery).
a. Explore and understand data (catch, effort, CPUE, sizes, ages, maturity, gear types, logbook types, vessel turnover, misreporting, etc.). Plot everything.
b. Talk to fishers and other stakeholders.
3. Structural changes based on understanding the system can be very important for the assessment, whereas many other issues just cause small changes in the index trends.
4. Revisit data exploration when updating indices. Do not simply 'turn the handle'.
5. Identify likely covariates before you start modelling. Avoid data dredging.
6. Differentiate between catchability and density variables.
7. Always include the variables that affect catchability – this is usually more important than the type of model you use. Think about potential for bias due to missing variables.
8. Consider targeting and target change through time and how to address it. Understand the fleet well enough to know what the targeting strategies might be.
Analysis
9. Generalized additive models (GAMs) and STMs are better than generalized linear models (GLMs). GAMs are best for exploration. STMs can be better for the final model(s). Each has unique capabilities.
10. Model the whole stock if you can do so without dropping important covariates or relationships due to data gaps or difficult spatial interactions.
11. Test your model by simulation.
12. Build multiple models using different approaches to develop your understanding of how the models are working and to consider alternative hypotheses. Start simple.
13. Use influence plots to understand how the variables and their values affect the indices.
14. Construct the index via the "predict-then-aggregate" approach rather than the previous practice of treating the partial effect of year as the index.
Assessment
15. Use the index fishery approach if you can.
16. Assume effort creep. There are catchability changes that your model has not captured. Catchability increases are almost inevitable in the long term.
17. Do not blindly split indices and assume the model will scale them correctly – this is unlikely and dangerous.
18. Do not include several conflicting indices in an assessment at the same time. Do include alternative indices as alternative assessment scenarios.

Table 4
Information requirements to support the acceptance of abundance indices.

Fishery description	Describe fishery including catch, effort, size composition of catch, nominal CPUE by area, season, history of fishery development and changes.
Analysis description	Describe data selection, CPUE standardization model, and CPUE estimates. Include any data filtering, outlier removal.
Statistical Results	Provide model diagnostics and goodness-of-fit criteria relative to alternative model configurations; tables, etc.
Nominal/ Standardized	Comparison plot of nominal and standardized indices.
Diagnostic plots	QQ, residuals, etc.
Point estimate & variability	Characterize uncertainty in estimates of standardized CPUE; SE or CV of standardized CPUE (generated or assumed).

Source: Adapted from ALBWG (2013.

breaks down (Cooke and Beddington, 1984; Dunn et al., 2000; Harley et al., 2001; National Research Council, 2000; Ye and Dennis, 2009). Although there are no easy answers, the recommendations presented here will help analysts to obtain the best information available.

Funding

This work was supported in part by NIWA Strategic Science Investment Funding and by the Indian Ocean Tuna Commission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Table 5
Criteria for evaluating the strengths and weaknesses of candidate abundance indices.

Criterion	Description
Spatial distribution	Proportion of stock covered by fishery; latitude and longitude
Size/age range	Distribution of size or ages in catch
Fishing ground map	Show area of operations for each fishery by season/ decade
Relative contribution	Proportion of total catch in the fishery
Temporal coverage	Time period of data collection
Temporal consistency	Change in spatial location of fishing grounds over temporal period, e.g., decadal changes/seasonal changes
Temporal consistency in size composition	Decadal and seasonal changes in size of fish captured
Statistical soundness	Standardization method, diagnostic plots, and CPUE variability provided
Targeting	Primary target, by-catch species
Drivers of catchability change	Time series of external factors affecting catchability (e.g., management practices, fishing technology, targeting changes)
Socio-economic factors	Time series of price, demand, technological changes (e.g., freezers), etc.

Source: Adapted from ALBWG (2013.

CRediT authorship contribution statement

Simon D. Hoyle: Conceptualization, Writing – original draft, Writing – review & editing, Visualization, Supervision, Funding acquisition. **Robert A. Campbell:** Writing – original draft, Writing – review & editing. **Nicholas D. Ducharme-Barth:** Writing – original draft, Writing – review & editing, Visualization. **Arnaud Grüss:** Writing – original draft, Writing – review & editing. **Bradley R. Moore:** Writing – original draft, Writing – review & editing, Visualization. **James T. Thorson:** Writing – original draft, Writing – review & editing. **Laura Tremblay-Boyer:** Writing – original draft, Writing – review & editing. **Henning Winker:** Writing – original draft, Writing – review & editing. **Shijie Zhou:** Writing – original draft, Writing – review & editing. **Mark N. Maunder:** Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

Acknowledgments

We thank Steven Holmes, Jill Coyle, Carolina Minte-Vera, and three anonymous reviewers for their advice. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of their employers.

References

Abascal, F.J., Mejuto, J., Quintans, M., Ramos-Cartelle, A., 2010. Horizontal and vertical movements of swordfish in the Southeast Pacific. ICES J. Mar. Sci. 67, 466–474.
Abbott, J.K., Haynie, A.C., Reimer, M.N., 2015. Hidden flexibility: institutions, incentives, and the margins of selectivity in fishing. Land Econ. 91, 169–195.
Aho, K., Derryberry, D., Peterson, T., 2014. Model selection for ecologists: the worldviews of AIC and BIC. Ecology 95, 631–636.
Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Springer Verlag.
Alglave, B., Rivot, E., Etienne, M.-P., Woillez, M., Thorson, J.T., Vermard, Y., 2022. Combining scientific survey and commercial catch data to map fish distribution. ICES J. Mar. Sci. 79, 1133–1149. https://doi.org/10.1093/icesjms/fsac032.

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A., 2022. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv.

Anon. Report of the IOTC CPUE Workshop, San Sebastian, Spain, 21–22 October, 2013. Indian Ocean Tuna Commission; 2013.

Anonymous. Report of the 2018 ICCAT bigeye tuna stock assessment meeting. SCRS/2018/010. Collective Volume of Scientific Papers ICCAT. 75:1721–1855; 2019.

Beerkircher, L., Arocha, F., Barse, A., Prince, E., Restrepo, V., Serafy, J., Shivji, M., 2009. Effects of species misidentification on population assessment of overfished white marlin *Tetrapturus albidus* and roundscale spearfish *T. georgii*. Endanger. Species Res. 9, 81–90.

Bentley, N., Kendrick, T.H., Starr, P.J., Breen, P.A., 2012. Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. ICES J. Mar. Sci. 69, 84–88. https://doi.org/10.1093/icesjms/fsr174.

Beverton, R.J.H., Holt, S.J., 1957. On the dynamics of exploited fish populations. Ministry of Agriculture and Fisheries, London: UK.

Biseau, A., 1998. Definition of a directed fishing effort in a mixed-species trawl fishery, and its impact on stock assessments. Aquat. Living Resour. 11, 119–136.

Bishop, J., 2006. Standardizing fishery-dependent catch and effort data in complex fisheries with technology change. Rev. Fish. Biol. Fish. 16, 21–38. https://doi.org/10.1007/s11160-006-0004-9.

Bolker, B. GLMM FAQ. 〈https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html〉; 2022.

Bosch, S., Tyberghein, L., Deneudt, K., Hernandez, F., De Clerck, O., 2018. In search of relevant predictors for marine species distribution modelling using the MarineSPEED benchmark dataset. Diversity and Distributions 24 (2), 144–157. https://doi.org/10.1111/ddi.12668. Wiley Online Library.

Braccini, M., Denham, A., O'Neill, M.F., Lai, E., 2021. Spatial and temporal patterns in catch rates from multispecies shark fisheries in Western Australia. Ocean Coast. Manag. 213, 105883 https://doi.org/10.1016/j.ocecoaman.2021.105883.

Buck, E.H., 1995. Individual transferable quotas in fishery management. Congr. Res. Serv., Libr. Congr. 95, 849.

Bürkner, P.-C., 2017. brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw. 80, 1–28.

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference understanding AIC and BIC in model selection. Sociol. Methods Res. 33, 261–304. https://doi.org/10.1177/0049124104268644.

Campbell, R. Aggregate and size-based standardised CPUE indices for longline target species caught within the ETBF - 2019 Update. 24th meeting of the Tropical Tuna Resource Assessment Group, 17–18 July 2019. Mooloolaba, Australia; 2019.

Campbell, R.; Williams, K.; Williams, D. Summary of size data collected from the Eastern Tuna and Billfish Fishery and associated indicators. 24th meeting of the Tropical Tuna Resource Assessment Group, 17–18 July 2019. Mooloolaba. Australia; 2019.

Campbell, R.A., 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fish. Res. 70, 209–227. https://doi.org/10.1016/j.fishres.2004.08.026.

Campbell, R.A., 2015. Constructing stock abundance indices from catch and effort data: some nuts and bolts. Fish. Res. 161, 109–130. https://doi.org/10.1016/j.fishres.2014.07.004.

Campbell, R.A., 2016. A new spatial framework incorporating uncertain stock and fleet dynamics for estimating fish abundance. Fish Fish 17, 56–77. https://doi.org/10.1111/faf.12091.

Campbell, R.A., Zhou, S., Hoyle, S.D., Hillary, R., Haddon, M., Auld, S., 2017. Developing innovative approaches to improve CPUE standardisation for Australia's multispecies pelagic longline fisheries. Fisheries Research and Development Corporation, Canberra, Australia.

Cao, J., Thorson, J.T., Richards, R.A., Chen, Y., 2017. Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. Can. J. Fish. Aquat. Sci. 74, 1781–1793.

Cardinale, M., Nugroho, D., Jonson, P., 2011. Serial depletion of fishing grounds in an unregulated, open access fishery. Fish. Res. 108, 106–111. https://doi.org/10.1016/j.fishres.2010.12.007.

Carruthers, T.R., Ahrens, R.N.M., McAllister, M.K., Walters, C.J., 2011. Integrating imputation and standardization of catch rate data in the calculation of relative abundance indices. Fish. Res. 109, 157–167. https://doi.org/10.1016/j.fishres.2011.01.033.

Carvalho, F.C., Murie, D.J., Hazin, F.H., Hazin, H.G., Leite-Mourato, B., Travassos, P., Burgess, G.H., 2010. Catch rates and size composition of blue sharks (*Prionace glauca*) caught by the Brazilian pelagic longline fleet in the southwestern Atlantic Ocean. Aquat. Living Resour. 23, 373–385.

Chambers, M.; Hoyle, S.D. Estimates of non-member catch of SBT in the Indian and Pacific Oceans, CCSBT-ESC/1509/10. 20th Extended Scientific Committee of the CCSBT. Incheon, Republic of Korea; 2015.

Chang, S.-K., 2003. Analysis of Taiwanese white marlin catch data and standardization of catch rates. ICCAT Col. Vol. Sci. Pap. 55, 453–466.

Chang, S.-K., Hoyle, S., Liu, H.-I., 2011. Catch rate standardization for yellowfin tuna (*Thunnus albacares*) in Taiwan's distant-water longline fishery in the Western and Central Pacific Ocean, with consideration of target change. Fish. Res. 107, 210–220. https://doi.org/10.1016/j.fishres.2010.11.004.

Charsley, A.R., Sibanda, N., Hoyle, S., Crow, S., 2022. Comparing the performance of three common species distribution modelling frameworks for freshwater environments through application to eel species in New Zealand. Can. J. Fish. Aquat. Sci.

Coelho, R., Infante, P., Santos, M.N., 2020. Comparing GLM, GLMM, and GEE modeling approaches for catch rates of bycatch species: A case study of blue shark fisheries in the South Atlantic. Fish. Oceanogr. 29, 169–184. https://doi.org/10.1111/fog.12462.

Commander, C.J., Barnett, L.A., Ward, E.J., Anderson, S.C., Essington, T.E., 2022. The shadow model: how and why small choices in spatially explicit species distribution models affect predictions. PeerJ 10, e12783.

Conn, P.B., 2010. Hierarchical analysis of multiple noisy abundance indices. Can. J. Fish. Aquat. Sci. 67, 108–120. https://doi.org/10.1139/F09-175.

Conn, P.B., Thorson, J.T., Johnson, D.S., 2017. Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. Methods Ecol. Evol. 8, 1535–1546. https://doi.org/10.1111/2041-210x.12803.

Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R., Hooten, M.B., 2018. A guide to Bayesian model checking for ecologists. Ecol. Monogr. 88, 526–542. https://doi.org/10.1002/ecm.1314.

Cooke, J., Beddington, J., 1984. The relationship between catch rates and abundance in fisheries. Math. Med. Biol.: A J. IMA 1, 391–405.

Cosandey-Godin, A., Krainski, E.T., Worm, B., Flemming, J.M., 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Can. J. Fish. Aquat. Sci. 72, 186–197. https://doi.org/10.1139/cjfas-2014-0159.

Dambly, L.I., Isaac, N.J., Jones, K.E., Boughey, K.L., O'Hara, R.B., 2023. Integrated species distribution models fitted in INLA are sensitive to mesh parameterisation. Ecography e06391.

Denis, V., Lejeune, J., Robin, J., 2002. Spatio-temporal analysis of commercial trawler data using General Additive models: patterns of Loliginid squid abundance in the north-east Atlantic. ICES J. Mar. Sci. 59, 633–648. https://doi.org/10.1006/jmsc.2001.1178.

Devine, J.; Ballara, S.; Hoyle, S. Fisheries characterisations for barracouta (Thyrsites atun) and preliminary standardised CPUE analyses in BAR 4 and BAR 5, 1990 to 2020. New Zealand Fisheries Assessment Report:13; 2022.

Dick, E.J., 2004. Beyond 'lognormal versus gamma': discrimination among error distributions for generalized linear models. Fish. Res. 70, 351–366. https://doi.org/10.1016/j.fishres.2004.08.013.

Diggle, P.J., Menezes, R., Su, Tl, 2010. Geostatistical inference under preferential sampling. J. R. Stat. Soc.: Ser. C. (Appl. Stat.) 59, 191–232. https://doi.org/10.1111/j.1467-9876.2009.00701.x.

Ducharme-Barth, N.; Vincent, M.; Vidal, T.; Hamer, P. Analysis of Pacific-wide operational longline dataset for bigeye and yellowfin tuna catch-per-unit-effort (CPUE). Technical Report SC16-SA-WP-03; 2020.

Ducharme-Barth, N.D., Grüss, A., Vincent, M.T., Kiyofuji, H., Aoki, Y., Pilling, G., Hampton, J., Thorson, J.T., 2022. Impacts of fisheries-dependent spatial sampling patterns on catch-per-unit-effort standardization: a simulation study and fishery application. Fish. Res. 246, 106169 https://doi.org/10.1016/j.fishres.2021.106169.

Dulvy, N.K., Metcalfe, J.D., Glanville, J., Pawson, M.G., Reynolds, J.D., 2000. Fishery stability, local extinctions, and shifts in community structure in skates. Conserv. Biol. 14, 283–293. https://doi.org/10.1046/j.1523-1739.2000.98540.x.

Dunn, A., Harley, S., Doonan, I., Bull, B., 2000. Calculation and interpretation of catch-per-uniteffort (CPUE) indices. N. Z. Fish. Assess. Rep. 1, 44.

Dunn, P.Tweedie, 2017. Evaluation of Tweedie exponential family models. R. Package Version 2.

Duparc, A., Aragno, V., Depetris, M., Floch, L., Cauquil, P., Lebranchu, J., Gaertner, D., Marsac, F., Bach, P., 2020. Assessment of the species composition of major tropical tunas in purse seine catches: a new modelling approach for the tropical tuna treatment processing (case of the French fleet in Atlantic Ocean). Collect Vol. Sci. Pap. ICCAT 76, 951–982.

Eigaard, O.R., Marchal, P., Gislason, H., Rijnsdorp, A.D., 2014. Technological development and fisheries management. Rev. Fish. Sci. Aquac. 22, 156–174. https://doi.org/10.1080/23308249.2014.899557.

Emery, T.J., Noriega, R., Williams, A.J., Larcombe, J., 2019. Changes in logbook reporting by commercial fishers following the implementation of electronic monitoring in Australian Commonwealth fisheries. Mar. Policy 104, 135–145. https://doi.org/10.1016/j.marpol.2019.01.018.

Forrestal, F., Goodyear, C.P., Schirripa, M., Babcock, E., Lauretta, M., Sharma, R., 2017. Testing robustness of CPUE standardization using simulated data: findings of initial blind trials. Collect Vol. Sci. Pap. ICCAT 74, 391–403.

Forrestal, F.C., Schirripa, M., Goodyear, C.P., Arrizabalaga, H., Babcock, E.A., Coelho, R., Ingram, W., Lauretta, M., Ortiz, M., Sharma, R., 2019. Testing robustness of CPUE standardization and inclusion of environmental variables with simulated longline catch datasets. Fish. Res. 210, 1–13. https://doi.org/10.1016/j.fishres.2018.09.025.

Francis, C.R.; Hurst, R.J.; Renwick, J.A. Quantifying annual variation in catchability for commercial and research fishing. 2003.

Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138.

Fu, D.; Roux, M.-J.; Clarke, S.C.; Francis, M.P.; Dunn, A.; Hoyle, S.D. Pacific-wide sustainability risk assessment of bigeye thresher shark (*Alopias superciliosus*). NIWA Client Report No 2016089WN for project WCP16301. Wellington; 2016.

Garrod, D., 1964. Effective fishing effort and the catchability coefficient q. Rapp. Et. Process verbaux Des. réunions du Cons. Int. pour l'Explor. De. la Mer. 155, 66–70.

Girardin, R., Hamon, K.G., Pinnegar, J., Poos, J.J., Thébaud, O., Tidd, A., Vermard, Y., Marchal, P., 2017. Thirty years of fleet dynamics modelling using discrete-choice models: What have we learned? Fish Fish. 18, 638–655.

Glazer, J., Butterworth, D., 2002. GLM-based standarization of the catch per unit effort series for South African west coast hake, focusing on adjustments for targeting other species. Afr. J. Mar. Sci. 24, 323–339.

Goodyear, C.P., 2016. Modeling the time-varying density distribution of highly migratory species: atlantic blue marlin as an example. Fish. Res. 183, 469–481.

Grüss, A., Thorson, J.T., 2019. Developing spatio-temporal models using multiple data types for evaluating population trends and habitat usage. ICES J. Mar. Sci. 76, 1748–1761. https://doi.org/10.1093/icesjms/fsz075.
Grüss, A., Walter III, J.F., Babcock, E.A., Forrestal, F.C., Thorson, J.T., Lauretta, M.V., Schirripa, M.J., 2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fish. Res. 213, 75–93. https://doi.org/10.1016/j.fishres.2019.01.008.
Grüss, A., Pirtle, J.L., Thorson, J.T., Lindeberg, M.R., Neff, A.D., Lewis, S.G., Essington, T. E., 2021. Modeling nearshore fish habitats using Alaska as a regional case study. Fish. Res. 238, 105905 https://doi.org/10.1016/j.fishres.2021.105905.
Grüss, A., Moore, B.R., Pinkerton, M.H., Devine, J.A., 2022. VAST (vector autoregressive spatio-temporal) modelling of macrourid relative abundance in the Ross Sea region to support bycatch management. WG-FSA-2022/48. CCAMLR, Hobart, Australia.
Grüss, A., McKenzie, J.R., Lindegren, M., Bian, R., Hoyle, S.D., Devine, J.A., 2023a. Supporting a stock assessment with spatio-temporal models fitted to fisheries-dependent data. Fish. Res. 262, 106649.
Grüss, A., Moore, B.R., Pinkerton, M.H., Devine, J.A., 2023b. Understanding the spatio-temporal abundance patterns of the major bycatch species groups in the Ross Sea region Antarctic toothfish (Dissostichus mawsoni) fishery. Fish. Res. 262, 106647.
Han, H., Yang, C., Zhang, H., Fang, Z., Jiang, B., Su, B., Sui, J., Yan, Y., Xiang, D., 2022. Environment variables affect CPUE and spatial distribution of fishing grounds on the light falling gear fishery in the Northwest Indian Ocean at different time scales. Front. Mar. Sci. 9, 939334 https://doi.org/10.3389/fmars.2022.939334.
Han, Q., Grüss, A., Shan, X., Jin, X., Thorson, J.T., 2021. Understanding patterns of distribution shifts and range expansion/contraction for small yellow croaker (*Larimichthys polyactis*) in the Yellow Sea. Fish. Oceanogr. 30, 69–84.
Han, Q., Shan, X., Jin, X., Gorfine, H., 2023. Contrasting stock status trends obtained from survey and fishery CPUE, taking Larimichthys polyactis in Yellow Sea Large Marine Ecosystem as an example. Ecol. Indic. 147, 110032.
Hansell, A.C., Becker, S.L., Cadrin, S.X., Lauretta, M., Walter III, J.F., Kerr, L.A., 2022. Spatio-temporal dynamics of bluefin tuna (*Thunnus thynnus*) in US waters of the northwest Atlantic. Fish. Res. 255, 106460 https://doi.org/10.1016/j.fishres.2022.106460.
Harley, S.J., Myers, R.A., Dunn, A., 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. 58, 1760–1772. https://doi.org/10.1139/cjfas-58-9-1760.
Harrell, F.E., 2001. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer.
Hartig, F. DHARMa: residual diagnostics for hierarchical (Multi-Level/Mixed) regression models. 2020.
He, X., Bigelow, K.A., Boggs, C.H., 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fish. Res. 31, 147–158.
Helle, K., Pennington, M., Hareide, N.-R., Fossen, I., 2015. Selecting a subset of the commercial catch data for estimating catch per unit effort series for ling (*Molva molva* L.). Fish. Res. 165, 115–120.
Henderson, H.V., Velleman, P.F., 1981. Building multiple regression models interactively. Biometrics 391–411.
Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice Dynamics and Uncertainty. Chapman and Hall, New York.
Hinton, M.G., 1996. Standardizing catch and effort statistics using physiological, ecological, or behavioral constraints, and environmental data, with applications to blue marlin (*Makaira nigricans*) and swordfish (*Xiphias gladius*) of the Pacific Ocean. Scripps Institution of Oceanography. University of California, San Diego.
Hinton, M.G.; Maunder, M.N. Methods for standardizing CPUE and how to select among them. Collective volume of scientific papersInternational Commission for the Conservation of Atlantic Tunas/Recueil de documents scientifiquesCommission internationale pour la Conservation des Thonides de l'Atlantique/Coleccion de documentos cientificosCo:no; 2004.
Hinton, M.G., Nakano, H., 1996. Standardizing catch and effort statistics using physiological, ecological, or behavioral constraints and environmental data, with an application to blue marlin (*Makaira nigricans*) catch and effort data from the Japanese longline fisheries in the Pacific. Inter Am. Trop. Tuna Comm. Bull. 21, 171–200.
Hinz, H., Murray, L.G., Lambert, G.I., Hiddink, J.G., Kaiser, M., 2013. Confidentiality over fishing effort data threatens science and management progress. Fish Fish. 14, 110–117. https://doi.org/10.1111/j.1467-2979.2012.00475.x.
Hiraoka, Y.; Kanaiwa, M.; Yokawa, K. Estimation of abundance indices for blue shark in the North Pacific. ISC/12/SHARKWG-1/09. Working document submitted for publication to the ISC Shark Working Group Workshop, 28 May – 4 June, 2012; 2012.
Højsgaard, S., Halekoh, U., Yan, J., 2006. The R package geepack for generalized estimating equations. J. Stat. Softw. 15, 1–11.
Hoyle, S. Potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance. CCSBT Extended Scientific Committee for the 26th Meeting of the Scientific Committee. Online; 2021.
Hoyle, S.; Huang, H.; Kim, D.N.; Lee, M.K.; Matsumoto, T.; Walter, J. Collaborative study of bigeye tuna CPUE from multiple Atlantic Ocean longline fleets in 2018. Collect Vol Sci Pap ICCAT. 75:2033–2080; 2019a.
Hoyle, S.; Kleiber, P.; Davies, N.; Langley, A.D.; Hampton, J.. Stock assessment of skipjack tuna in the western and central Pacific Ocean, WCPFC-SC-7-SA/WP-04. WCPFC Scientific Committee Regular Session, Pohnpei, FSM, 9–17 August 2011, 7th; 2011.
Hoyle, S.; Hampton, J.; Davies, N. Stock assessment of albacore tuna in the South Pacific ocean. Scientific Committee, Eighth Regular Session, 7–15 August 2012, Busan, Republic of Korea; 2012.
Hoyle, S.D. Stock assessment of albacore tuna in the South Pacific ocean. Scientific Committee, Seventh Regular Session, 9–17 August 2011, Pohnpei, Federated States of Micronesia; 2011.
Hoyle, S.D. Investigation of potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance, CCSBT-ESC/2008/29. CCSBT Extended Scientific Committee 25. online; 2020.
Hoyle, S.D., Langley, A.D., 2020. Scaling factors for multi-region stock assessments, with an application to Indian Ocean tropical tunas. Fish. Res. 228, 105586 https://doi.org/10.1016/j.fishres.2020.105586.
Hoyle, S.D.; Okamoto, H. Analyses of Japanese longline operational catch and effort for bigeye and yellowfin tuna in the WCPO, WCPFC-SC7-SA-IP-01. Western and Central Pacific Fisheries Commission, 7th Scientific Committee. Pohnpei, Federated States of Micronesia; 2011.
Hoyle, S.D., Okamoto, H., 2015. Descriptive analyses of the Japanese Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Comm. Work. Part. Trop. Tunas.
Hoyle, S.D.; Chang, S.-T.; Fu, D.; Kim, D.N.; Lee, S.I.; Matsumoto, T.; Chassot, E.; Yeh, Y.-M. Collaborative study of bigeye and yellowfin tuna CPUE from multiple Indian Ocean longline fleets in 2019, with consideration of discarding. IOTC-2019-WPM10–16. 10th Working Party on Methods, 17–19 October, 2019. Donostia-San Sebastian, Spain: Indian Ocean Tuna Commission; 2019b.
Hoyle, S.D.; Lauretta, M.; Lee, M.K.; Matsumoto, T.; Sant'Ana, R.; Yokoi, H.; Su, N.-J. Collaborative study of yellowfin tuna CPUE from multiple Atlantic Ocean longline fleets in 2019. 2019c.
Hoyle, S.D.; Langley, A.D.; Campbell, R.A. Guidelines for presenting CPUE indices of abundance for WCPFC stock assessments. WCPFC-SC10–2014/ SA-IP-11. WCPFC Scientific Committee 10. Majuro, Republic of the Marshall Islands.; 2014a.
Hoyle, S.D.; Langley, A.D.; Campbell, R.A. Recommended approaches for standardizing CPUE data from pelagic fisheries, WCPFC-SC10–2014/ SA-IP-10. WCPFC Scientific Committee 10. Majuro, Republic of the Marshall Islands.; 2014b.
Hoyle, S.D., Lee, S.I., Kim, Z.G., 2015a. Descriptive analyses of the Korean Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Comm. Work. Part. Trop. Tunas.
Hoyle, S.D., Okamoto, H., Yeh, Y.-m, Kim, Z.G., Lee, S.I., Sharma, R., 2015b. IOTC–CPUEWS02 2015: report of the 2nd CPUE Workshop on Longline Fisheries, 30 April – 2 May 2015. Indian Ocean Tuna Comm.
Hoyle, S.D., Yeh, Y.-M., Chang, S.-T., Wu, R.-F., 2015c. Descriptive analyses of the Taiwanese Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Comm. Work. Part. Trop. Tunas.
Hoyle, S.D.; Semba, Y.; Kai, M.; Okamoto, H. Development of porbeagle shark stock abundance indicators using Japanese data. New Zealand Fisheries Assessment Report 2017/07; 2017.
Hoyle, S.D., Lee, S.I., Kim, D.N., 2022. CPUE standardization for southern bluefin tuna (Thunnus maccoyii) in the Korean tuna longline fishery, accounting for spatiotemporal variation in targeting through data exploration and clustering. PeerJ 10, e13951.
Hsu, J., Chang, Y.-J., Ducharme-Barth, N.D., 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. Fish. Res. 255, 106440 https://doi.org/10.1016/j.fishres.2022.106440.
Huang, H.-W., Liu, K.-M., 2010. Bycatch and discards by Taiwanese large-scale tuna longline fleets in the Indian Ocean. Fish. Res. 106, 261–270. https://doi.org/10.1016/j.fishres.2010.08.005.
Huang, W.-B., Lo, N.C., Chiu, T.-S., Chen, C.-S., 2007. Geographical distribution and abundance of Pacific saury, *Cololabis saira* (Brevoort)(Scomberesocidae), fishing stocks in the Northwestern Pacific in relation to sea temperatures. Zool. Stud. 46, 705–716.
Huang, W.-B.; Chang, Y.-J.; Hsieh, C. Standardized CPUE of Pacific saury (*Cololabis saira*) caught by the Chinese Taipei stick-held dip net fishery up to 2019. NPFC-2020-SSC PS06-WP05; 2020.
Ichinokawa, M., Brodziak, J., 2010. Using adaptive area stratification to standardize catch rates with application to North Pacific swordfish (*Xiphias gladius*). Fish. Res. 106, 249–260. https://doi.org/10.1016/j.fishres.2010.08.001.
Ichinokawa, M.; Takeuchi, Y. Estimation of coefficient of variances in standardized CPUE of Pacific bluefin tuna caught by Japanese coastal longline with a nonparametric method. Working document submitted for publication to the ISC Pacific bluefin tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC); 2012.
IOTC. Guidelines for the Presentation of CPUE Standardisations and Stock Assessment Models. Extract from the 17th Session of the Scientific Committee. Appendix IV. Indian Ocean Tuna Commission; 2015.
ISC. Stock Assessment Report for Striped Marlin (*Kajikia audax*) in the Western and Central North Pacific Ocean through 2017. WCPFC-SC15–2019/SA-WP-09. WCPFC Scientific Committee, Fifteenth Regular Session, 12–20 August 2019, Pohnpei, Federated States of Micronesia 2019.
ISC Albacore Working Group. Annex 4: Report of the Albacore Working Group Workshop. 8–14 November 2016. Nanaimo, British Columbia, Canada. International Scientific Committee for Tuna and Tuna-like Species In the North Pacific Ocean; 2016.
ISC Shark Working Group. Stock assessment and future projections of blue sharks in the North Pacific Ocean through 2020, WCPFC-SC9–2013/SA-WP-11. 18th WCPFC Scientific Committee, 10–18 August 2022. Online; 2022.
Jones, C.M., 2004. The use of recreational CPUE statistics in groundfish stock assessments. Rep. CIE Recreat. CPUE Indices.

Kai, M., 2019. Spatio-temporal changes in catch rates of pelagic sharks caught by Japanese research and training vessels in the western and central North Pacific. Fish. Res. 216, 177–195. https://doi.org/10.1016/j.fishres.2019.02.015.
Kai, M., Thorson, J.T., Piner, K.R., Maunder, M.N., 2017. Spatiotemporal variation in size-structured populations using fishery data: an application to shortfin mako (*Isurus oxyrinchus*) in the Pacific Ocean. Can. J. Fish. Aquat. Sci. 74, 1765–1780. https://doi.org/10.1139/cjfas-2016-0327.
Kanamori, Y., Nishijima, S., Okamura, H., Yukami, R., Watai, M., Takasuka, A., 2021. Spatio-temporal model reduces species misidentification bias of spawning eggs in stock assessment of spotted mackerel in the western North Pacific. Fish. Res. 236, 105825.
Kendrick, T.H., Bentley, N., 2011. Fishery characterisations and catch-per-unit-effort indices for three sub-stocks of red gurnard in GUR 1, 1989–90 to 2008–09. Ministry of Fisheries, Wellington, New Zealand.
Kerwath, S.E., Parker, D., Winker, H., Potts, W., Mann, B., Wilke, C., Attwood, C., 2019. Tracking the decline of the world's largest seabream against policy adjustments. Mar. Ecol. Prog. Ser. 610, 163–173.
Klaer, N.L., Smith, D.C., 2012. Determining primary and companion species in a multi-species fishery: implications for TAC setting. Mar. Policy 36, 606–612.
Kleiber, P., Maunder, M.N., 2008. Inherent bias in using aggregate CPUE to characterize abundance of fish species assemblages. Fish. Res. 93, 140–145. https://doi.org/10.1016/j.fishres.2008.03.013.
Kleiven, A.R., Espeland, S.H., Stiansen, S., Ono, K., Zimmermann, F., Olsen, E.M., 2022. Technological creep masks continued decline in a lobster (*Homarus gammarus*) fishery over a century. Sci. Rep. 12, 1–10. https://doi.org/10.1038/s41598-022-07293-2.
Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H.J., Bell, B., 2016. TMB: automatic differentiation and laplace approximation. J. Stat. Softw. 70, 1–21.
Lambert, D.Zero-inflated, 1992. Poisson regression, with an application to defects in manufacturing. Technometrics 34, 1–14. https://doi.org/10.2307/1269547.
Langley, A., 2019. An investigation of the performance of CPUE modelling approaches–a simulation study. N. Z. Fish. Assess. Rep. 57, 50.
Lee, L.A.; Terrell, J.W. Habitat suitability index models: Flathead catfish. Biological Report 82: US Department of the Interior, Fish and Wildlife Service, Research and Development; 1988.
Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., Picaut, J., 1997. El Nino Southern Oscillation and tuna in the western Pacific. Nature 389, 715–718. https://doi.org/10.1038/39575.
Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A., Minami, M., 2013. Defining population spatial units: Simultaneous analysis of frequency distributions and time series. Fish. Res. 139, 85–92. https://doi.org/10.1016/j.fishres.2012.10.001.
Li, Z., Ye, Z., Wan, R., Zhang, C., 2015. Model selection between traditional and popular methods for standardizing catch rates of target species: a case study of Japanese Spanish mackerel in the gillnet fishery. Fish. Res. 161, 312–319.
Lindegren, M., van Deurs, M., Maureaud, A., Thorson, J.T., Bekkevold, D., 2022. A spatial statistical approach for identifying population structuring of marine fish species: European sprat as a case study. ICES J. Mar. Sci. 79, 423–434. https://doi.org/10.1093/icesjms/fsac007.
Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. 63, 1–25. https://doi.org/10.18637/jss.v063.i19.
Liu, S., Liu, Y., Li, J., Cao, C., Tian, H., Li, W., Tian, Y., Watanabe, Y., Lin, L., Li, Y., 2022. Effects of oceanographic environment on the distribution and migration of Pacific saury (*Cololabis saira*) during main fishing season. Sci. Rep. 12, 1–13. https://doi.org/10.1038/s41598-022-17786-9.
Lo, N.C.-h, Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognornial models. Can. J. Fish. Aquat. Sci. 49, 2515–2526.
Lynch, H.J., Thorson, J.T., Shelton, A.O., 2014. Dealing with under-and over-dispersed count data in life history, spatial, and community ecology. Ecology 95, 3173–3180. https://doi.org/10.1890/13-1912.1.
Mac Nally, R., Duncan, R.P., Thomson, J.R., Yen, J.D., 2018. Model selection using information criteria, but is the "best" model any good? J. Appl. Ecol. 55, 1441–1444. https://doi.org/10.1111/1365-2664.13060.
Magee, L., 1998. Nonlocal behavior in polynomial regressions. Am. Stat. 52, 20–22.
Marr, J.C., 1951. On the use of the terms abundance, availability and apparent abundance in fishery biology. Copeia 1951, 163–169.
Marra, G., Wood, S.N., 2011. Practical variable selection for generalized additive models. Comput. Stat. Data Anal. 55, 2372–2387.
Marshall, E., Spiegelhalter, D., 2003. Approximate cross-validatory predictive checks in disease mapping models. Stat. Med. 22, 1649–1660. https://doi.org/10.1002/sim.1403.
Martell, S., Stewart, I., 2014. Towards defining good practices for modeling time-varying selectivity. Fish. Res. 158, 84–95. https://doi.org/10.1016/j.fishres.2013.11.001.
Maunder, M.N.; Hinton, M.G. Estimating relative abundance from catch and effort data, using neural networks. 2006.
Maunder, M.N., Piner, K.R., 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fish. Res. 192, 16–27. https://doi.org/10.1016/j.fishres.2016.04.022.
Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70, 141–159. https://doi.org/10.1016/j.fishres.2004.08.002.
Maunder, M.N., Hinton, M.G., Bigelow, K.A., Langley, A.D., 2006. Developing indices of abundance using habitat data in a statistical framework. Bull. Mar. Sci. 79, 545–559.
Maunder, M.N., Thorson, J.T., Xu, H., Oliveros-Ramos, R., Hoyle, S.D., Tremblay-Boyer, L., Lee, H.H., Kai, M., Chang, S.-K., Kitakado, T., 2020. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fish. Res. 229, 105594 https://doi.org/10.1016/j.fishres.2020.105594.
Maureaud, A.A., Frelat, R., Pécuchet, L., Shackell, N., Mérigot, B., Pinsky, M.L., Amador, K., Anderson, S.C., Arkhipkin, A., Auber, A., 2021. Are we ready to track climate-driven shifts in marine species across international boundaries?-A global survey of scientific bottom trawl data. Glob. Change Biol. 27, 220–236. https://doi.org/10.1111/gcb.15404.
McKechnie, S.; Hoyle, S.; Harley, S. Longline CPUE series that account for changes in the spatial extent of fisheries. WCPFC-SC9–2013/SA-IP-05. WCPFC 9th Scientific Committee. Pohnpei, Federated States of Micronesia: Western and Central Pacific Fisheries Commission; 2013.
McKenzie, J.R.; Parsons, D.M. Fishery Characterisations and Catch-per-unit-effort Indices for Three Sub-stocks of Snapper SNA 1, 1989–90 to 2009–10: Ministry for Primary Industries; 2012.
Mejuto, J., García-Cortés, B., Ramos-Cartelle, A., De la Serna, J., 2009. Standardized catch rates for the blue shark (*Prionace glauca*) and shortfin mako (*Isurus oxyrinchus*) caught by the Spanish surface longline fleet in the Atlantic Ocean during the period 1990-2007. Collect Vol. Sci. Pap. ICCAT 64, 1509–1521.
Minami, M., Lennert-Cody, C.E., Gao, W., Román-Verdesoto, M., 2007. Modeling shark bycatch: the zero-inflated negative binomial regression model with smoothing. Fish. Res. 84, 210–221. https://doi.org/10.1016/j.fishres.2006.10.019.
Miyabe, N., Takeuchi, Y., 2003. Standardized bluefin CPUE from the Japanese longline fishery in the Atlantic including those for mixing studies. ICCAT Col. Vol. Sci. Pap. 55, 1190–1207.
Monnahan, C.C., Thorson, J.T., Kotwicki, S., Lauffenburger, N., Ianelli, J.N., Punt, A.E., 2021. Incorporating vertical distribution in index standardization accounts for spatiotemporal availability to acoustic and bottom trawl gear for semi-pelagic species. ICES J. Mar. Sci. 78, 1826–1839. https://doi.org/10.1093/icesjms/fsab085.
Munoz, B., Lesser, V.M., Smith, R.A., 2010. Applying multiple imputation with geostatistical models to account for item nonresponse in environmental data. J. Mod. Appl. Stat. Methods 9, 27.
Nakano, H., 1998. Stock status of Pacific swordfish, *Xiphias gladius*, inferred from CPUE of the Japanese longline fleet standardized using general linear models. US Nat. Mar. Fish. Serv., NOAA Tech. Rep. NMFS 142, 195–209.
National Research Council, 2000. Improving the collection, management, and use of marine fisheries data. National Academies Press.
Noriega, R., Werry, J.M., Sumpton, W., Mayer, D., Lee, S.Y., 2011. Trends in annual CPUE and evidence of sex and size segregation of *Sphyrna lewini*: management implications in coastal waters of northeastern Australia. Fish. Res. 110, 472–477. https://doi.org/10.1016/j.fishres.2011.06.005.
O'Leary, C.A., Thorson, J.T., Ianelli, J.N., Kotwicki, S., 2020. Adapting to climate-driven distribution shifts using model-based indices and age composition from multiple surveys in the walleye pollock (Gadus chalcogrammus) stock assessment. Fish. Oceanogr. 29, 541–557.
Okamura, H., Morita, S.H., Funamoto, T., Ichinokawa, M., Eguchi, S., 2017. Target-based catch-per-unit-effort standardization in multispecies fisheries. Can. J. Fish. Aquat. Sci. 75, 452–463.
Ono, K., Punt, A.E., Hilborn, R., 2015. Think outside the grids: an objective approach to define spatial strata for catch and effort analysis. Fish. Res. 170, 89–101. https://doi.org/10.1016/j.fishres.2015.05.021.
Ono, K., Ianelli, J.N., McGilliard, C.R., Punt, A.E., 2018. Integrating data from multiple surveys and accounting for spatio-temporal correlation to index the abundance of juvenile Pacific halibut in Alaska. ICES J. Mar. Sci. 75, 572–584. https://doi.org/10.1093/icesjms/fsx174.
Palmer, M., Quetglas, A., Guijarro, B., Moranta, J., Ordines, F., Massutí, E., 2009. Performance of artificial neural networks and discriminant analysis in predicting fishing tactics from multispecific fisheries. Can. J. Fish. Aquat. Sci. 66, 224–237.
Palomares, M.L., Pauly, D., 2019. On the creeping increase of vessels' fishing power. Ecol. Soc. 24 https://doi.org/10.5751/Es-11136-240331.
Parker, D., Winker, H., West, W., Kerwath, S.E., 2017. Standardization of the catch per unit effort for swordfish (*Xiphias gladius*) for the South African longline fishery. Collect Vol. Sci. Pap. ICCAT 74, 1295–1305.
Pearl, J., 2009. Causal inference in statistics: an overview. Stat. Surv. 3, 96–146.
Peatman, T.; Nicol, S. Updated purse seine bycatch estimates in the WCPO. WCPFC-SC17–2021/ST-IP-06. WCPFC Scientific Committee, Seventeenth Regular Session. Online Meeting: Western and Central Pacific Fisheries Commission; 2021.
Peatman, T., Fukofuka, S., Park, T., Williams, P., Hampton, J., Smith, N., 2019. Better purse seine catch composition estimates: progress on the Project 60 work plan. WCPFC-SC15-2019/ST-WP 02.
Pelletier, D., Ferraris, J., 2000. A multivariate approach for defining fishing tactics from commercial catch and effort data. Can. J. Fish. Aquat. Sci. 57, 51–65.
Pennino, M.G., Paradinas, I., Illian, J.B., Muñoz, F., Bellido, J.M., López-Quílez, A., Conesa, D., 2019. Accounting for preferential sampling in species distribution models. Ecol. Evol. 9, 653–663. https://doi.org/10.1002/ece3.4789.
Perretti, C.T., Thorson, J.T., 2019. Spatio-temporal dynamics of summer flounder (*Paralichthys dentatus*) on the Northeast US shelf. Fish. Res. 215, 62–68. https://doi.org/10.1016/j.fishres.2019.03.006.
Peterson, C.D., Belcher, C.N., Bethea, D.M., Driggers III, W.B., Frazier, B.S., Latour, R.J., 2017. Preliminary recovery of coastal sharks in the south-east United States. Fish Fish. 18, 845–859.
Peterson, C.D., Wilberg, M.J., Cortés, E., Latour, R.J., 2021. Dynamic factor analysis to reconcile conflicting survey indices of abundance. ICES J. Mar. Sci. 78, 1711–1729. https://doi.org/10.1093/icesjms/fsab051.
Peterson, M.J., Mueter, F., Criddle, K., Haynie, A.C., 2014. Killer whale depredation and associated costs to Alaskan sablefish, Pacific halibut and Greenland turbot longliners. PLoS One 9, e88906.

Phillips, A.J., Ciannelli, L., Brodeur, R.D., Pearcy, W.G., Childers, J., 2014. Spatio-temporal associations of albacore CPUEs in the Northeastern Pacific with regional SST and climate environmental variables. ICES J. Mar. Sci. 71, 1717–1727. https://doi.org/10.1093/icesjms/fst238.
Pickens, B.A., Carroll, R., Schirripa, M.J., Forrestal, F., Friedland, K.D., Taylor, J.C., 2021. A systematic review of spatial habitat associations and modeling of marine fish distribution: a guide to predictors, methods, and knowledge gaps. PloS One 16, e0251818. https://doi.org/10.1371/journal.pone.0251818.
Pilling, G.; Brouwer, S. Report from the SPC pre-assessment workshop, Noumea, April 2017. WCPFC-SC13–2017/IP-02. 13th Regular Session of the Scientific Committee, 9–17 August 2017. Rarotonga, Cook Islands; 2017.
Pinto, C., Travers-Trolet, M., Macdonald, J.I., Rivot, E., Vermard, Y., 2019. Combining multiple data sets to unravel the spatiotemporal dynamics of a data-limited fish stock. Can. J. Fish. Aquat. Sci. 76, 1338–1349. https://doi.org/10.1139/cjfas-2018-0149.
Pitcher, T.J., Watson, R., Forrest, R., Valtýsson, H.Þ, Guénette, S., 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish Fish 3, 317–339. https://doi.org/10.1046/j.1467-2979.2002.00093.x.
Punt, A.E., Walker, T.I., Taylor, B.L., Pribac, F., 2000. Standardization of catch and effort data in a spatially-structured shark fishery. Fish. Res. 45, 129–145.
Quinn, T.J., Deriso, R.B., 1999. Quantitative Fish Dynamics. Oxford University Press.
R Core Team, 2022. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Ralston, S. Trends in standardized catch rate of some rockfishes (Sebastes spp.) from the California trawl logbook database. 1999.
Rigby, R.A., Stasinopoulos, D.M., 2005. Generalized additive models for location, scale and shape. J. R. Stat. Soc.: Ser. C. (Appl. Stat.) 54, 507–554. https://doi.org/10.1111/j.1467-9876.2005.00510.x.
Roberts, J.J., Best, B.D., Mannocci, L., Fujioka, E., Halpin, P.N., Palka, D.L., Garrison, L. P., Mullin, K.D., Cole, T.V., Khan, C.B., 2016. Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. Sci. Rep. 6, 1–12.
Roche, C., Guinet, C., Gasco, N., Duhamel, G., 2007. Marine mammals and demersal longline fishery interactions in Crozet and Kerguelen Exclusive Economic Zones: an assessment of depredation levels. CCAMLR Sci. 14, 67–82.
Rowden, A.A., Anderson, O.F., Georgian, S.E., Bowden, D.A., Clark, M.R., Pallentin, A., Miller, A., 2017. High-resolution habitat suitability models for the conservation and management of vulnerable marine ecosystems on the Louisville Seamount Chain, South Pacific Ocean. Front. Mar. Sci. 4, 335. https://doi.org/10.3389/fmars.2017.00335.
Rudd, M.B., Branch, T.A., 2017. Does unreported catch lead to overfishing? Fish Fish. 18, 313–323.
Rufener, M.C., Kristensen, K., Nielsen, J.R., Bastardie, F., 2021. Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. Ecol. Appl. 31, e02453 https://doi.org/10.1002/eap.2453.
Sadovy, Y., Domeier, M., 2005. Are aggregation-fisheries sustainable? Reef fish fisheries as a case study. Coral Reefs 24, 254–262.
Sampson, D.B., Scott, R.D., 2012. An exploration of the shapes and stability of population–selection curves. Fish Fish. 13, 89–104. https://doi.org/10.1111/j.1467-2979.2011.00417.x.
Scherrer, K., Galbraith, E., 2020. The risk of underestimating long-term fisheries creep. Ecol. Soc. 25 https://doi.org/10.5751/Es-11389-250118.
Schnute, J.T., Hilborn, R., 1993. Analysis of contradictory data sources in fish stock assessment. Can. J. Fish. Aquat. Sci. 50, 1916–1923.
Shelton, A.O., Thorson, J.T., Ward, E.J., Feist, B.E., 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Can. J. Fish. Aquat. Sci. 71, 1655–1666.
Shemla, A., McAllister, M.K., 2006. Bayesian generalized linear models to standardize and impute missing data in the Atlantic bluefin tuna (*Thunnus thynnus thynnus*) TaskII catch and effort database. ICCAT Collect. Vol. Sci. Pap. 59, 750–768.
Shono, H., 2008a. Application of the Tweedie distribution to zero-catch data in CPUE analysis. Fish. Res. 93, 154–162. https://doi.org/10.1016/j.fishres.2008.03.006.
Shono, H., 2008b. Confidence interval estimation of CPUE year trend in delta-type two-step model. Fish. Sci. 74, 712–717.
Squires, D., 1992. Productivity measurement in common property resource industries: an application to the Pacific coast trawl fishery. The. Rand J. Econ. 23, 221–236. https://doi.org/10.2307/2555985.
Squires, D., Vestergaard, N., 2013. Technical change in fisheries. Mar. Policy 42, 286–292. https://doi.org/10.1016/j.marpol.2013.03.019.
Sribney, B. What are some of the problems with stepwise regression? ⟨https://www.stata.com/support/faqs/statistics/stepwise-regression-problems/⟩; 1996.
Stephens, A., MacCall, A., 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fish. Res. 70, 299–310. https://doi.org/10.1016/j.fishres.2004.08.009.
Su, N.J., Yeh, S.Z., Sun, C.L., Punt, A.E., Chen, Y., Wang, S.P., 2008. Standardizing catch and effort data of the Taiwanese distant-water longline fishery in the western and central Pacific Ocean for bigeye tuna, *Thunnus obesus*. Fish. Res. 90, 235–246.
Taylor, L.R., 1961. Aggregation, variance and the mean. Nature 189, 732–735.
Tesfamichael, D., Pitcher, T.J., Pauly, D., 2014. Assessing changes in fisheries using fishers' knowledge to generate long time series of catch rates: a case study from the Red Sea. Ecol. Soc. 19 https://doi.org/10.5751/Es-06151-190118.
Therneau, T.M.; Lumley, T.; Atkinson, E.; Crowson, C. A Package for Survival Analysis in R. ⟨https://cran.r-project.org/web/packages/survival/index.html⟩; 2022.
Thorson, J. VAST model structure and user interface. ⟨https://github.com/James-Thorson-NOAA/VAST⟩; 2022.
Thorson, J.T., 2015. Spatio-temporal variation in fish condition is not consistently explained by density, temperature, or season for California Current groundfishes. Mar. Ecol. Prog. Ser. 526, 101–112.
Thorson, J.T., 2018. Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. Can. J. Fish. Aquat. Sci. 75, 1369–1382. https://doi.org/10.1139/cjfas-2017-0266.
Thorson, J.T., 2019a. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish. Res. 210, 143–161. https://doi.org/10.1016/j.fishres.2018.10.013.
Thorson, J.T., 2019b. Measuring the impact of oceanographic indices on species distribution shifts: The spatially varying effect of cold-pool extent in the eastern Bering Sea. Limnol. Oceanogr. 64, 2632–2645.
Thorson, J.T., Barnett, L.A., 2017. Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. ICES J. Mar. Sci. 74, 1311–1321. https://doi.org/10.1093/icesjms/fsw193.
Thorson, J.T., Haltuch, M.A., 2019. Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. Can. J. Fish. Aquat. Sci. 76, 401–414. https://doi.org/10.1093/icesjms/fsw193.
Thorson, J.T., Minto, C., 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J. Mar. Sci. 72, 1245–1256. https://doi.org/10.1093/icesjms/fsu213.
Thorson, J.T., Shelton, A.O., Ward, E.J., Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. 72, 1297–1310. https://doi.org/10.1093/icesjms/fsw193.
Thorson, J.T., Fonner, R., Haltuch, M.A., Ono, K., 2016a. Winker, H. Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. Can. J. Fish. Aquat. Sci. 73, 1–14.
Thorson, J.T., Ianelli, J.N., Larsen, E.A., Ries, L., Scheuerell, M.D., Szuwalski, C., Zipkin, E.F., 2016b. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Glob. Ecol. Biogeogr. 25, 1144–1158. https://doi.org/10.1093/icesjms/fsw193.
Thorson, J.T., Pinsky, M.L., Ward, E.J., 2016c. Model-based inference for estimating shifts in species distribution, area occupied and centre of gravity. Methods Ecol. Evol. 7, 990–1002. https://doi.org/10.1093/icesjms/fsw193.
Thorson, J.T., Cunningham, C.J., Jorgensen, E., Havron, A., Hulson, P.-J.F., Monnahan, C.C., von Szalay, P., 2021. The surprising sensitivity of index scale to delta-model assumptions: recommendations for model-based index standardization. Fish. Res. 233, 105745 https://doi.org/10.1016/j.fishres.2020.105745.
Thygesen, U.H., Albertsen, C.M., Berg, C.W., Kristensen, K., Nielsen, A., 2017. Validation of ecological state space models using the Laplace approximation. Environ. Ecol. Stat. 24, 317–339. https://doi.org/10.1007/s10651-017-0372-4.
Tian, S., Chen, X., Chen, Y., Xu, L., Dai, X., 2009. Standardizing CPUE of Ommastrephes bartramii for Chinese squid-jigging fishery in Northwest Pacific Ocean. Chin. J. Oceanol. Limnol. 27, 729–739. https://doi.org/10.1007/s00343-009-9199-7.
Tremblay-Boyer, L.; Neubauer, P. Historical catch reconstruction and CPUE standardization for the stock assessment of oceanic whitetip shark in the Western and Central Pacific Ocean. WCPFC-SC15/SA-IP-17. Scientific Committee, 15th Regular Session. Pohnpei, Federated States of Micronesia: Western and Central Pacific Fisheries Commission; 2019.
Walsh, W.A., Brodziak, J., 2015. Billfish CPUE standardization in the Hawaii longline fishery: model selection and multimodel inference. Fish. Res. 166, 151–162. https://doi.org/10.1016/j.fishres.2014.07.015.
Walter, J.; Urtizberea Ijurco, A.; Hiroki, Y.; Satoh, K.; Ortiz, M.; Kimoto, A.; Ailloud; Matsumoto, T. Stock Synthesis model for Atlantic yellowfin tuna, SCRS/2019/121. ICCAT's Collective Volume of Scientific Papers; 2020.
Walters, C., 2003. Folly and fantasy in the analysis of spatial catch rate data. Can. J. Fish. Aquat. Sci. 60, 1433–1436.
Warton, D.I., Thibaut, L., Wang, Y.A., 2017. The PIT-trap—A "model-free" bootstrap procedure for inference about regression models with discrete, multivariate responses. PloS One 12, e0181790. https://doi.org/10.1371/journal.pone.0181790.
Watters, G.; Deriso, R. Catches per unit of effort of bigeye tuna: A new analysis with regression trees and simulated annealing. Bulletin Inter-American Tropical Tuna Commission/Boletin Comision Interamericana del Atun Tropical. La Jolla CA. [BullIATTC/BolCIAT] Vol 21; 2000.
Webber, D.; Starr, P. Characterisation and CPUE from 1989–90 to 2018–19 of jack mackerel off the west coast of New Zealand (JMA 7). New Zealand Fisheries Assessment Report:06; 2022.
Wiegand, R.E., 2010. Performance of using multiple stepwise algorithms for variable selection. Stat. Med. 29, 1647–1659.
Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2009. Incorporating time-varying catchability into population dynamic stock assessment models. Rev. Fish. Sci. 18, 7–24.
Winker, H., Kerwath, S.E., Attwood, C.G., 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fish. Res. 139, 118–131.
Winker, H., Kerwath, S.E., Attwood, C.G., 2014. Proof of concept for a novel procedure to standardize multispecies catch and effort data. Fish. Res. 155, 149–159.
Wood, S.N., 2017. Generalized additive models: an introduction with R. Chapman and Hall/CRC.
Xu, H.; Lennert-Cody, C.E. Standardizing the purse-seine indices of abundance and associated length compositions for skipjack tuna in the eastern Pacific Ocean. Document SAC-13 INF-K. Scientific Advisory Committee, 13th meeting, 16–20 May 2022. Online: Inter-American Tropical Tuna Commission; 2022.

Xu, H., Lennert-Cody, C.E., Maunder, M.N., Minte-Vera, C.V., 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Fish. Res. 213, 121–131.

Ye, Y., Dennis, D., 2009. How reliable are the abundance indices derived from commercial catch–effort standardization? Can. J. Fish. Aquat. Sci. 66, 1169–1178.

Zhang, J.; Li, S. A Review of Machine Learning Based Species' Distribution Modelling. 2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII): IEEE; 2017.

Zhou, S., Campbell, R.A., Hoyle, S.D., 2019. Catch per unit effort standardization using spatio-temporal models for Australia's Eastern Tuna and Billfish Fishery. ICES J. Mar. Sci. 76, 1489–1504. https://doi.org/10.1093/icesjms/fsz034.

Zuur, A.F., Saveliev, A.A., Ieno, E.N., 2012. Zero inflated models and generalized linear mixed models with R. Highland Statistics Limited.