Introduction

Estimates of landed catch are a key component of many fishery management systems. Stock assessment models (referred to here as “assessments”) are often conditioned on time series of annual catch, usually under the assumption that catches are known without error. While some assessment models are able to incorporate uncertainty in catch (e.g. Stock Synthesis; Methot and Wetzel, 2013), reliable estimates of catch uncertainty are often unavailable and, as a result, ad-hoc approaches are often used to evaluate the effects of catch uncertainty (cite).

Over the past decade, the estimation of catch and associated uncertainty has become a focus for recreational fisheries in the United States (NAS 2017). Commercial fisheries, on the other hand, are often assumed to have precise estimates of catch by species. This is due, in part, to the use of landing receipts (aka “fish tickets”) that record the amount of fish landed into various sort groups, or “market categories.” It is quite common for multiple species to be landed within a single sort group. On the U.S. West Coast, for example, some market categories are named after a particular species but catches landed in those categories may consist of several species. Other sort groups that contain multiple species are clearly designated as mixed-species categories (e.g. “nearshore rockfish,” or species within a particular genus or family).

Importantly, the decision of how to sort species into market categories on a landing receipt is often left up to fishermen, dealers, or processors. In order to determine the composition of species in a given market category, port sampling programs intercept vessels offloading catch or during subsequent processing. These port sampling programs are critical when estimating catch at the species level, not only for mixed-species market categories but also for single-species categories since these often contain a mix of the intended (‘nominal’) species and other species that are similar in appearance and/or value. Depending on the species, the assumption that landings in a market category consist of a single species can lead to substantial errors in catch estimation, even when the category of interest is named after a single species and/or there is a requirement to sort that species into a specific market category.

Within market categories, the species composition of landed catch can vary spatially, temporally, by fishing gear, and catch disposition (e.g. fish sold alive or dead). These differences are attributable to many factors, including market preference, fishing behavior, regulatory constraints, and biological/ecological characteristics (e.g. spatial distribution) of the landed species. As a result, estimates of species composition for a given market category are often stratified over time (e.g. quarterly) and across other relevant strata (e.g. ports, gears, catch disposition). Sampling programs often have limited funds, and attempts to reduce bias in species composition estimates through the introduction of additional strata comes at a cost, namely reduced precision (Cochran 19xx; Tomlinson 1971).

On the U.S. West Coast, state sampling programs allocate effort both spatially and temporally, but many domains of interest (e.g. market category, gear type, catch disposition) remain unsampled or sparsely sampled due to a proliferation of categories over time, logistical constraints, and limited resources (Sen 1986; Crone 1995; [cite reliability paper]). Sampling rates are determined by either a fixed target (e.g. a percentage of recent landings; Tsou et al. 2015) or adjusted based on available funding. Ad-hoc ‘data borrowing’ algorithms based on expert opinion are used to calculate species compositions for unsampled strata and domains, but these algorithms do not produce estimates of uncertainty.

*[talk about how small samples, missing strata, and overdispersed data don’t play well with asymptotic results typically used in design-based surveys; since sampling rates are unlikely to increase without additional funding, need model-based approach; cite other model-based approaches in survey literature, including Faye-Herriott (sp?) model and Rao ‘small area’ estimation book*].

In this paper, we evaluate a model-based framework proposed by Shelton et al. (20xx) using commercial port sampling data collected in California, U.S.A. We describe data collected by the [official name of cooperative survey] over the period 1978-1990. We then extend the Shelton et al. model framework to address limitations of their approach. Specifically, we evaluate alternative likelihoods for overdispersed data, compare multiple hierarchical structures for the linear predictor, and generate posterior predictive distributions that integrate over uncertainty in spatial structure. Finally, we estimate landed catch by species for both sampled and unsampled strata, and propose a general framework for quantifying uncertainty including an efficient database design for dissemination of results at any level of aggregation. [*not sure if we want to include that last part about the DB*].

We use the phrases “sort group” and “market category” interchangeably in this paper.

Stuff for discussion:

* Sen (1986)
  + Recommended a minimum of 4 samples in each category (MC, gear, live) within a port-month stratum, roughly 52 samples per year. Redirect sampling to infrequently landed categories until 4 samples are obtained.
  + An increased number of categories increases the chance that a category will be missed by samplers.
  + Boats are first stage samples within a stratum, with clusters used to avoid sampling bias due to non-random sampling
* Cite Fay-Herriott and Datta and Ghosh 2012 papers in introduction to satisfy the survey stats crowd
* Why Shelton didn’t work
* Sort requirements do not eliminate the need for port sampling.
* The proliferation of market categories over time in the sampled catch has not been matched with an increase in sampling effort, effectively reducing the average number of samples per category over time (Figure X). This reduces precision of catch estimates, increases uncertainty in stock assessment outputs, and impedes efforts to monitor removals relative to catch targets.
* Fishermen and Dealers determine Market Categories for landed catch; issue with sampling – can’t get all categories; describe problem; “sort requirements” used to increase proportion of a particular species in a given market category, but other species are still landed in these categories (e.g. bocaccio in Figure X); can improve precision of important targets, but is not practical for large numbers of species; even for major targets, DOESN’T ELIMINATE THE NEED FOR SAMPLING; cite example of Dover sole – rex sole is small fraction, but of a HUGE landing; decline in sampling effort; need for model-based approach to impute missing strata; current approach is ad-hoc.
* Statistical framework; focus of estimation is the total landed catch, in weight, of a single species; extend Shelton et al. (20xx); model-based – allows for imputation, small-area estimation (Fey and Harriott); model selection based on predictive criteria; model averaging to account for model uncertainty; quantifies uncertainty.
* Model-based approach is best course of action given sparse data, but best solution is to increase sampling or reduce the number of strata.
* Recommend cost/benefit analysis to help identify optimal number of market categories given management goals.

References (incomplete)

NAS (National Academies of Sciences, Engineering, and Medicine). 2017. Review of the Marine Recreational Information Program. Washington, DC: The National Academies Press. <https://doi.org/10.17226/24640>

Sen, A.R. 1986. Methodological problems in sampling commercial rockfish landings. Fish. Bull. U.S. 84: 409-421.