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. Without this information, assessment authors often rely on ad-hoc sensitivity analyses which may or may not be incorporated into management advice and/or fail to propagate catch uncertainty into quantities of interest to managers.

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 availability of landing receipts (aka “fish tickets”) which serve as a record of the weight of fish landed into various market categories (sort groups). As noted by Pearson et al. (2008), it is important to recognize that species and market categories are not synonymous. On the U.S. West Coast, for example, it is common for multiple species to be landed within a single market category (CALCOM 2017, PacFIN 2017). This is expected for categories that are clearly designated as mixed-species categories (e.g. “nearshore rockfish,” or species within a particular genus or family). However, some categories that are named after a single species still contain several species, to varying degrees, even after regulations require sorting into a particular category (Pearson et al. 2008).

The decision of how to sort species into market categories on a landing receipt is typically made by the fishermen, dealers, or processors. As a result, trained port samplers intercept vessels offloading catch or during subsequent processing in order to determine the species composition of catch landed in a given market category (Sen 1984, Crone 1995, Tsou et al. 2015). These species composition data are used to partition the weight of landed catch in a market category across species, a process commonly referred to as “catch expansion” (Pearson and Almany 1995). To calculate total landings for a single species, the expanded catch is summed across all market categories in which the species was landed.

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, port 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; Pearson et al. 2008; Tsou et al. 2015). Ad-hoc ‘data borrowing’ algorithms based on expert opinion are used to calculate species compositions for unsampled strata and domains, but these algorithms have unknown bias and do not produce estimates of uncertainty. In contrast, model-based estimators are increasingly used to estimate quantities of interest for domains with small sample sizes and/or unsampled strata (sometimes referred to as ‘small area’ estimation; Rao 2003). Shelton et al. (2012) developed a Bayesian hierarchical statistical framework for species composition data that pools information among sparsely sampled strata, predicts species compositions for unsampled strata, and can be combined with landing receipts to estimate total landings by species, across market categories and other strata, with associated estimates of uncertainty. Shelton et al. considered hierarchical pooling only among quarters within a single year, comparing generalized linear and hierarchical linear models to trawl data from a single port in California. The authors also underscored the need to better understand performance of alternative models, and to overcome issues with computation time, particularly since commercial port sampling data sets often include hundreds of landed strata spanning decades, multiple ports, gear types, and other domains of interest.

In this paper, we evaluate the model-based framework proposed by Shelton et al. (2012) using commercial port sampling data collected in California, U.S.A. We describe species composition data collected by the California Cooperative Groundfish Survey (CCGS 2017) over the period 1978-1990. We then extend the Shelton et al. framework to address limitations of their approach. Specifically, we evaluate alternative likelihoods to address overdispersion, compare multiple hierarchical structures for pooling information through time, and integrate model predictions across uncertainties in the spatial model structure. Finally, we estimate landed catch by species for both sampled and unsampled strata, and summarize a general framework for quantifying uncertainty including an efficient database design for dissemination of results at any level of aggregation.

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
* 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.
* 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)

CCGS (California Cooperative Groundfish Survey). 2017. Pacific States Marine Fisheries Commission, 350 Harbor Boulevard, Belmont, California, 94002. URL: 128.114.3.187.

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.