**CHAPTER TWO**

**Building a complete picture of commercial fisheries on the US West coast**

*Abstract*

Commercial fisheries systems have previously been recognized as an opportunity to

*Introduction*

Commercial fisheries are important ecologically and economically. This is underscored by the ecological and social disruption a collapse in a commercial fishery can cause, i.e. the cod collapse which resulted in both a state of emergency and reports of PTSD among fishermen and persistent change in the state of the ecosystem. As such, many wealthy countries have devoted resources to collect data on how, when and where fish are caught to better monitor ecological health of these populations. The result of this monitoring is a rich store of data on the ecological and economic patterns of fishing. Indeed, previous authors have recognized that fisheries are essentially ecological perturbations and might, provided the appropriate caveats, be good sources of data to test ecological theories at scales impossible to achieve in the lab {Jensen:2011bu}.

Fisheries have for a long time been recognized as a canonical example of the tragedy of the commons, and much work in fisheries has been devoted to understanding and aligning human incentives to ecological outcomes. However fisheries data, especially with the advent of VMS and the high coverage of observer data which records movement (i.e. fine scale search behavior) and haul-level harvest (i.e. fitness), also provides data relevant for behavioral ecology in terms of foraging and predator-prey interactions but has been largely underutilized.

For nations with industrial fisheries, the data typically collected consists of records of licenses, harvest (fish tickets), records of fishing locations (logbook and observer data), and now increasingly continuous tracks of movement (VMS). Unfortunately, these datasets span multiple temporal and spatial scales and were not collected with the goal of marrying them to one another. Thus VMS data, for example, exists on the US west coast starting in 2009, which results in almost 7 years of continuous vessel locations with (XXX meter accuracy). But because vessels on the US west coast switch between fisheries on every trip, it’s impossible to use this data unless you can assign the movement trajectories to the appropriate fishery. Similarly landing tickets on the US west coast aren’t linked with licenses on a trip-basis. Thus you can see the licenses for a vessel, but not which ones were used for which fishery. This makes even identifying the fishery as a unit of participation impossible.

Existing work examining harvesting behavior, hereafter fleet dynamics, has historically only used one of these datasets at a single time in a single fishery due to the silo-ing of data collection. Studies that have examined catch and search are often coarse, and typically focus on how fishermen choose what area in which to fish as a function of expected profits. This has proven to be powerful as a way to predict how effort is reallocated spatially, but the spatial resolution is coarse where fishing areas often correspond to areas on the order XX nautical miles squared despite fishing often occurring on spatial scales orders of magnitude finer in scale. Newer management regulations, and the advent of cheap and accurate GPS technology has resulted in vessel-monitoring systems (VMS) in many fisheries where the positions of a fishing vessel can be recorded regularly (i.e. hourly, etc.). This provides a wealth of fine scale movement data, akin to the explosion that has occurred in the creation of the sub-discipline of movement ecology. Work that has used these fisheries has focused on identifying fishing behavior from movement paths, and focused on fisheries where vessels are extremely specialized. Studies resulting from this work have focused on single fisheries often in one behavioral mode, i.e. using VMS data alone to look at patterns in movement (levy walk – Joo?). Yet major commercial fisheries are often characterized by hundreds of different species caught with diverse gears in a wide range of habitats.

In this chapter I demonstrate how these datasets can be combined in order to examine large scale dynamics such as how vessels participate across fisheries, to fine scale foraging behaviors.

build a data pipeline which sews together data describing foraging and fitness for the US west coast commercial fisheries between 2009 and 2013. I begin with landings data and use catch data to define fisheries, of which there are 10 that account for 90% of the landings and revenue on the US west coast. This results in a dataset of trips identified to fisheries. Using these trips I create a set of algorithms that match trips to the movement (VMS) data. Finally using VMS data that have been filtered into fisheries, I use the pink shrimp trawl fishery as an example and identify bouts of fishing behavior and define fishing grounds.

In order to examine use the US west coast as an example of a commercial fisheries social ecological system and demonstrate how the rich possibilities these datasets offer are by sewing disparate datasets together to capture fisheries dynamics at multiple scales. Using landing tickets, observer data and VMS I am able to identify fishing grounds and fishing bouts from movement of vessels, identify fisheries participation from the contents of catch, and calculate indices of effort (searching time, and path distance) often missing from any one of these datasets used alone.

In this chapter I suggest that fisheries data also represent an opportunity to examine the dynamics of social-ecological systems, and specifically how people change their resource consumption patterns in response to changes in ecological and economic conditions, an interaction crucial to the further development of sustainability science and the study of social ecological systems.

*Methods*

**Description of Data Sources**

**Defining Realized Fisheries**

Fisheries are defined as harvest assemblages caught with a specific gear (van Putten et al. 2012; Boonstra and Hentati Sundberg 2014). The Pacific Fisheries Management Council (PFMC) has developed a set of sector-based definitions similar to this approach for the federally managed groundfish landings (www.pcouncil.org), but no equivalent exists for non-groundfish fisheries (Northwest Fisheries Science Center 2015). In order to treat the landings dataset uniformly, I applied a métier analysis to this landing data (Deporte et al. 2012) to build a set of realized fisheries. A métier analysis identifies realized fisheries by clustering the species composition of landings. This methodology requires choices in the way similarity among trips are measured, a clustering algorithm for grouping similar trips together, and a constraint that the methods can scale across hundreds of thousands of landings. In the following I specify our rational for these choices.

For our distance metric I used the Hellinger distance *D* (P. Legendre and Legendre 2012) to calculate the similarity in revenue profiles between trips and generated a pairwise distance matrix. This distance metric has the benefit that it is asymmetric, where the presence of a species in both trips is considered more informative than the absence of a species. The Hellinger distance between the species composition of two fishing trips *A* and *B* is defined as

|  |  |
| --- | --- |
|  | (1) |

where *ai* is the fraction of revenue derived from species *i* on trip *A*, *bi* is the fraction of revenue derived from species *i* on trip *B*, and *S* indicates the total number of species collected in both trips. With this metric, trips *A* and *B* become increasingly similar (and the Hellinger distance declines) as the proportion of revenue attributable to each of the *S* species becomes increasingly matched.

We identified realized fisheries as groups of trips with similar target assemblages using the infoMap community detection algorithm (Rosvall and Bergstrom 2008). This algorithm examines networks for subgraphs more interconnected to one another than the network in which it is embedded. To generate the required network I transformed the distance matrix into a similarity matrix by subtracting the distance metric’s upper limit (i.e. ) from each pairwise distance. The result is a weighted, undirected network where trips are connected by edges proportional to their similarity. However, because our dataset contained 340,466 unique trips, I were not able to perform clustering using a single matrix containing all pairwise similarities. To obtain manageable matrix sizes I used one year of landings (2010) which I split by gear. Pairwise distances among trips and community detection were calculated within each gear partition, which grouped trips into target assemblage categories. To classify the 2009, 2012 and 2013 trips to fisheries, I assigned each unclassified trip to the same realized fishery as the 2010 trip to which it was closest in multi-dimensional space using a k-nearest neighbors algorithm.

A challenge in testing the effectiveness of this classification method, and part of the reason for its need, is that there is not an independent classification of US west coast fisheries that I could use to compare the results. To address this issue, I tested the reliability of our classification approach by evaluating the extent to which it identified known spatial and temporal structure of well-described US west coast fisheries and fishery sectors. Specifically, because I did not bound our clusters spatially, temporally, or by vessel characteristics, I were able to compare our emergent realized fisheries to existing sector definitions of groundfish, and groundfish impacting fisheries provided by the Northwest Fisheries Science Center Observer Program (Northwest Fisheries Science Center 2015).