Meta-Database of Marine Research in Mexico: Trends and Applications

Draft

Palacios-Abrantes, J; Arreguín-Sánchez, F; Cisneros, A; Cisneros-Mata, M.A; Rodríguez, L.

4/12/2017

Introduction

There is plenty evidence that climate change has been globally reshaping the marine ecosystem for the last century and is not expected to stop, rather than projected to continue, and intensify if anthropocentric emissions are not mitigated [Poloczanska, et al., 2013; Weatherdon et al 2016; Arreguín-Sánchez et al., 2016; Cheung et al., 2015; Edwards and Richardson, 2004; IPCC, william 1.5, Gatuso et al., 2015…etc!]. As oceans physicochemical characteristics evolve marine, fishes are expected to reduce in size [Pauly and Cheung, 2017] and shift in distribution [Cheung, 2009…10…15… etc!]. These effects will have direct impacts on landings [Cheung XXX], economy [Laham et all., 2017] and food security [Cisneros-Montemayor et al., 2017]. Such changes will be especially significant in developing nations as fisheries are most at risk from illegal fishing, weak governance, poor data knowledge, and population pressures and climate change [Golden et al., 2016]. It then becomes empirical for these nations to build adaptive capacity as defined by the Intergovernmental Panel on Climate Change (IPCC, 20XX) to mitigate the effects that climate change will have on coastal fishing communities.

With over 11 thousand kilometers of coastline, Mexico's coast houses XXX families [INEGI]. Many communities that live in the coast directly depend on landings for food and wealth [cita]. It is essential that Mexico incorporates climate change to the political agenda to foster adaptive capacity while incorporating sounding adaptation and mitigation plans [Mertz et al., 2009; Pershing et al., 2015; Rice and Garcia, 2011]. The production and availability of natural and socio-economic data is a key component of the actions needed to achieve these goals [Mertz et al., 2009; Moser et al., 2010]. As a result, data availability has been identified as a common barrier in many steps of the adaptive management process, from the planning stage to outcome evaluation [Moser et al., 2010].

In Mexico, numerous information covering the seas and coasts can be found in academic institutions (UNAM; CICESE; CINVESTAV) and NGOs (COBI; EDF; TNC) located (physically) both inside and outside the country, as well as government agencies (CONABIO; CONAPESCA). However, there is an overall assumption that in developing nations data are often limited and perceived as lacking [CITA]. In many cases, this is largely a result of the lack of knowledge about the availability of these data, rather than a complete (or partially) lack of it [CITA]. Data availability is key not only to better understand Mexico's marine and coastal environments, but to identify knowledge gaps so that research can be prioritized [Reichman et al., 2011]. Management of data and metadata (the information about every aspect of the data that is required to understand it) is a huge step to ensure proper preservation, stewardship, and access to information [Michener et al., 1997; Vincent et al., 2010]. And while diverse barriers often compromise the exchange of data among stakeholders [Reichman et al., 2011], having publicly accessible description on existing data (and metadata) is a huge step towards increasing collaboration for innovative research [Michener et al., 2006]. Having proper metadata increases data longevity, foster collaboration, and help examine analysis of the data itself; all important aspect of climate change studies [Michener et al., 2006]. The standardization of information through metadata specifications provides guidance for consistently describing data objects and data types (e.g., methods, units of measurement, and details of experimental design) [Reichman et al., 2011]. Moreover, it allows the creation of a catalog that holds data or information of different disciplines. Countries like Canada have recently created nation-wide repositories of marine-related research metadata [Cisneros-Montemayor, 2016].

Several initiatives in different levels of society in Mexico are currently working on grouping data-sets and information in single repositories. University level projects like The Gulf of Mexico Species Interactions database [GoMexSI] and DataMares [DataMares] are collecting information from specific areas (the gulfs of Mexico and California, respectively). Non-governmental organizations are also contributing with initiatives like MonitoreoNoroeste that focus on marine monitoring data [FMCN]. Finally, the National Commission for the Knowledge of Biodiversity (CONABIO) hold an important and robust list of biological and social data related to both marine and terrestrial environments [CONABIO]. Moreover, as part of a cutting-edge initiative, Mexico has recently created an information system solely devoted to climate change [SICC].

The current project aimed to create a meta-database for oceanographic, ecological, economic, fisheries and social data for marine ecosystems and marine-related sectors of Mexico. The information collected was then analyzed to determine trends and data gaps in Mexico. Finally, we developed a framework to evaluated Mexico's fisheries-related information in relation to climate change and identified fields of improvement. It is worth mention that the meta-database is an open source repository for public consultation and sharing. We believe that the meta-database of marine research will help facilitate efficient use of existing information and stimulate collaboration among multiple stakeholders.

Materials and Methods

Data Compilation

The first step towards building the meta-database was the selection of categories in witch data was going to be categorized. This step was crussial for the subsequent meta-analysis of the diffeent aspects of marine research in Mexico. The present effort concentrated in collecting information on data related to marine systems in Mexico from biology, ecology, fisheries, oceanography and social sciences. Because digital-available data in Mexico is scattered among different types of institutions, we started the search on global assessments (e.g. OBIS, FAO, NOAA, etc.) and subsequently went for more local data sources (e.g. Mexican universities catalogues, NGOs databases, etc.). While plenty information was found on these online sources, it was suspected that many of the metadata created was going to be found on personal off-line repositories (e.g. Long-track researchers, local ressources managers and NGOs). For this last step we organize **seven** different workshops/meetings in **five** marine-research hubs cities of Mexico. Hoping to collect information from top researchers we participated on **four** key academic events (EDF Mexico Meeting on Fisheries, 2017; State of the Gulf of Mexico, 2017; Sustainable Aquatic Ressources Workshop Mexico, 2017; International Meeting for Usage and Conservation of Fisheries Ressources, 2017). Because this is a collaborative initiative, we also partner up with other initiatives (e.g. DataMares, MonitoreoNoroeste, CONABIO, etc.) to have their data represented in the meta-dataset.

Metada Structure

Each record on the meta-dataset is part of a specific dataset included on a repository related to coastal and marine resources or regions. We first split each dataset found in the repository and then each topic represented in such dataset. Hence, each record in the meta-database represents a topic included in a dataset **(Figure 1)**. For biological data the topics were applied for each species (or group of species). Lastly, when the information had a spatial content, the information was disaggregated to the level of municipality. This dessagrageted method of including information creates a longer dataset but allows for a more specific type of meta-analysis.

The metadata gathered was categorized in 7 braod categories. Aquaculture that accounted for any type of information related to both in-land and off shore aquaculture. Conservation was used for any type of information gathered on marine protected areas, endangered species, or with a clear objective towards conservation. Ecology that accounted for any biological information related to the marine environment from genetics to exosystem data. The exception was biological information related to fisheries (e.g. catch, life history traits) that was grouped on its own category. Oceanography accounted for any chemical and/or phisical data found. Sociology was a broad category to descrive any type of information from the social science (e.g. landings revenue, human population dynamics, etc.). Finally, Tourism was used to categorize any type of data realted to the tourism activity.

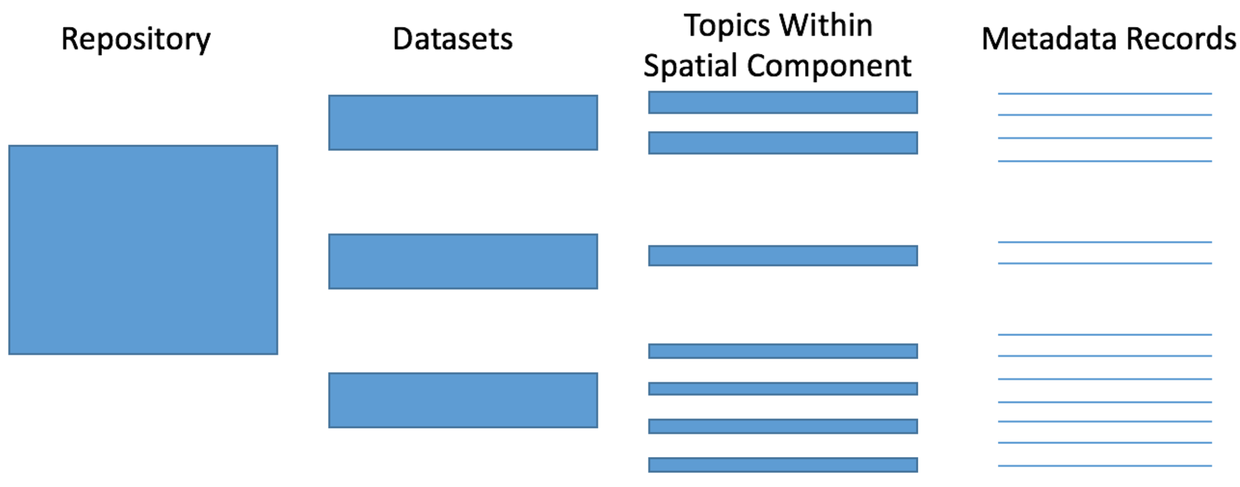

Figure 1:

Data Sources

The current effort to build the metadata collected data from 69 repositories gathered by 35 national and international institutions. The type of institutions included in this effort were academic (**n** = 43), gouvernamental agencies (**n** = GOV), **inter** (GOV) and **non** (**n** = Int) gouvernamental agencies, and international data sources (**n** = IGO).

Case study 1: Climate change data availlabillity

Case study 2: Marine Protected Areas and somehting else in Mexico

Results

General Numbers

The meta-database of marine research has 60.9 thousand records containing information for about 242.4 thousand data points. The research field with the most amount of records found was **Ecology** (**n** = 37111), followed by **Fisheries** (**n** = 20234) (Figure 1). However, when we analize the amount of data points by each register, we find that `Dis[1,1]` (**n** = 13.83) and **Conservation**(**n** = 11.38) have the higher rate of data-points per register.

There is a total of 20577 subjects in the metadata. The subject that has the most amount of data points is with 435 registers However, at the species level,* Ophiactis savignyi* is the most repeated one with 27 records. The data distribution is very diverse since the mean data point is 2.17 (sd 4.79). This is mainly because a large amount of records are one observation of species lists.

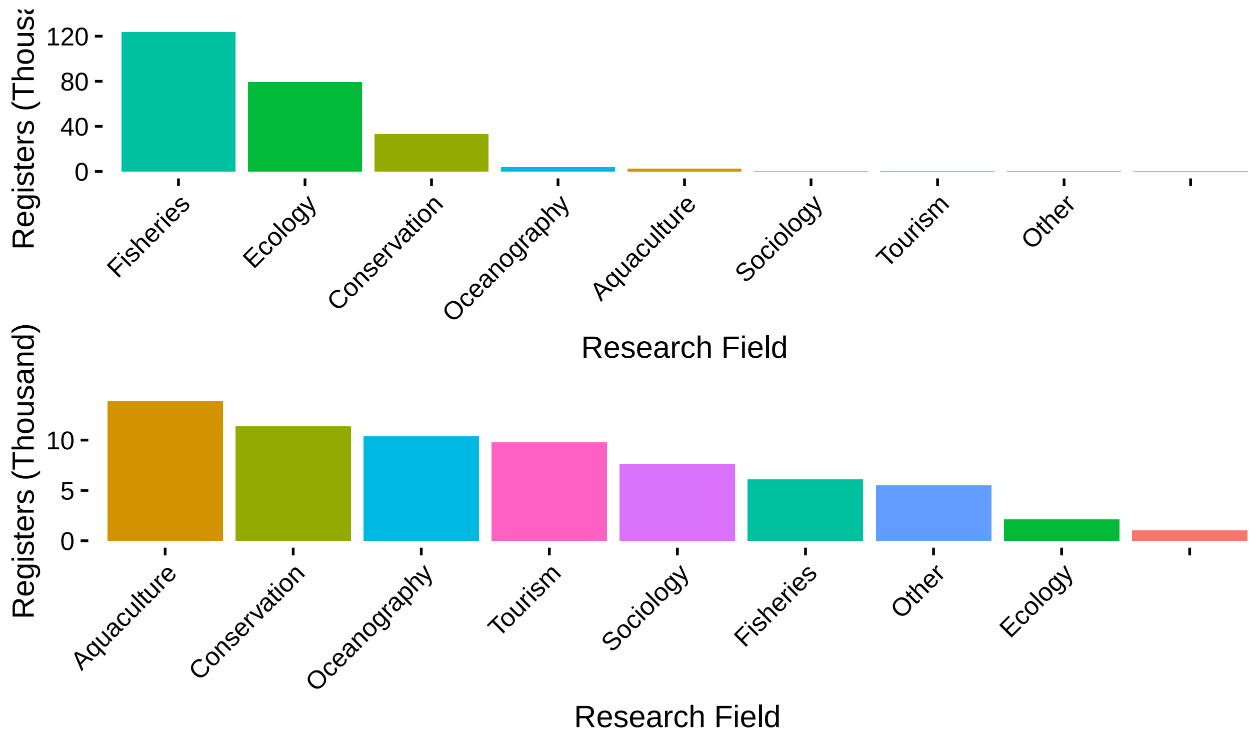

Figure 2: Number of data points per research field.

Geographic Analysis

In terms of regional data distribution, data points where evenly distributed between both the Pacific (84795) and the Atlantic oceans (61267). However, the great majority of data points are reported at a National level (61267). At a more detailed level we find that there is a clear dominance of data from both the Gulf of California (53456) and the Campeche Bank (36832), in respect to the other zones determined by the working group.

Discussion

Literature Review

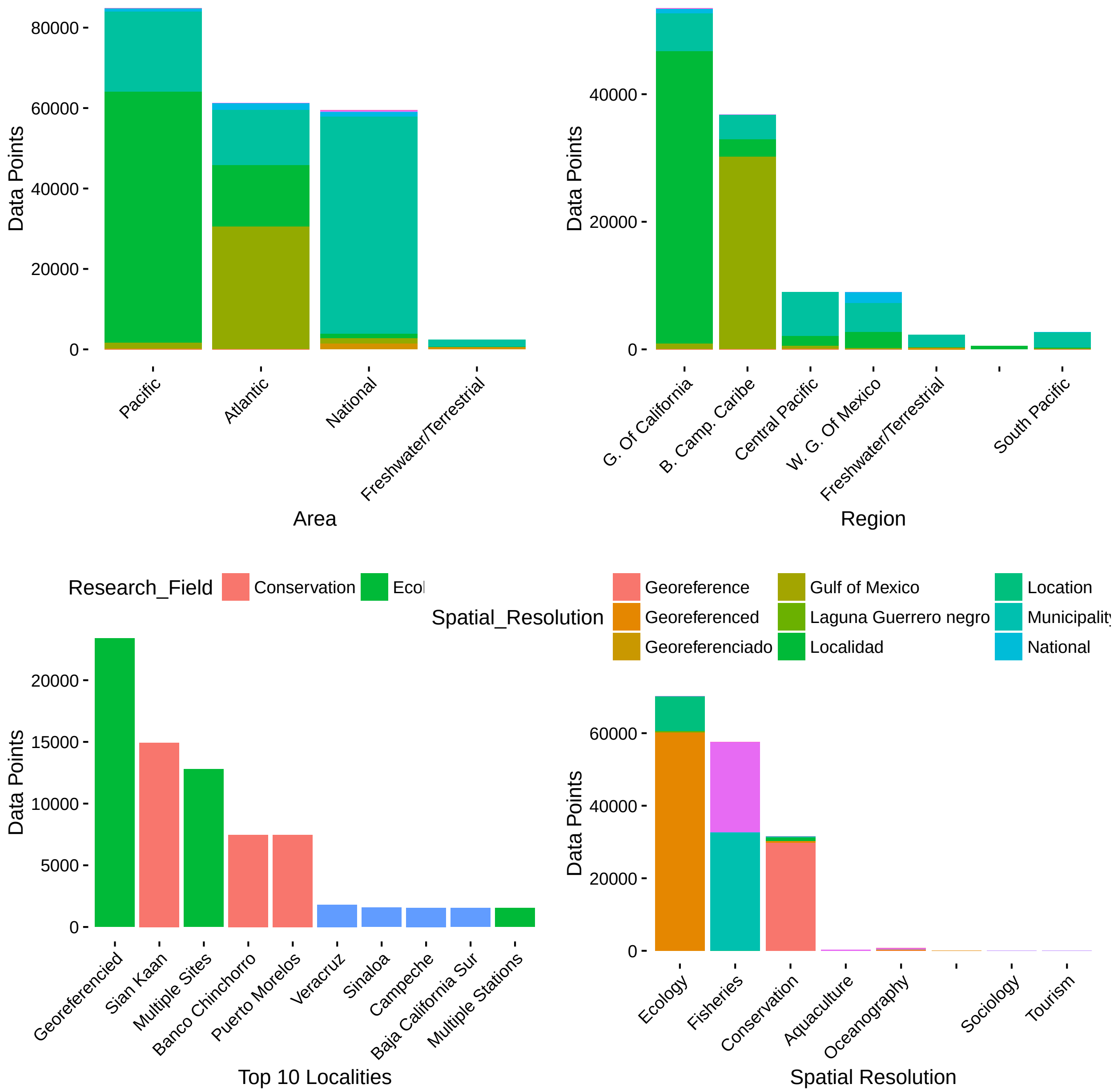

Figure 3: OPTION 1. Number of data points per geographic area for each research field.