The integrated dataset combines four monitoring projects in the kelp forests of the Santa Barbara Channel, USA. The surveying methods, raw data, and meta data files for each project can be found online.

Santa Barbara Coastal LTER (LTER): <http://sbc.lternet.edu>

Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO): <http://data.piscoweb.org/>

Kelp forest monitoring (KFM): <http://www.esapubs.org/archive/ecol/E094/245/>

San Nicolas Island monitoring (SNI): <http://www.esapubs.org/archive/ecol/E094/244/>

The data integration process includes data cleaning, column standardization, species validation, and data merging. The data cleaning was processed in seven steps: 1. Select the sites located within the Santa Barbara Channel; 2. Separate the biological survey data into three categories: fish survey, quad and swath count, and benthic cover; 3. Assign survey area to the corresponding record for each survey, and calculate the total point count for the benthic cover data; 4. Add life stage (adult and juvenile) information to the survey record; 5. Calculate plot average in cases where more than one diver surveyed each experimental plot (e.g. KFM roving diver fish count); 6. Remove redundant rows; 7. Convert date format into: YYYY-MM-DD.

After the cleaning process, the designated columns (for the integrated dataset) were selected and standardized indices were generated for site, subsite, transect, replicate plot, and taxon. Species’ scientific name were run through an online data base <http://www.marinespecies.org/> to obtain an authoritative name and taxon information. Finally, the standardized-format table from the four projects were merged into one csv file for each survey category.

For researchers who want to use the integrated datasets, there are some potential issues/cautions due to the difference among sampling methods of these four programs. Please refer to the protocol listed below for the details.