Clean Waters

People value marine waters that are free of pollution and debris for aesthetic and health reasons. Contamination of waters comes from oil spills, chemicals, eutrophication, algal blooms, disease pathogens (e.g., fecal coliform, viruses, and parasites from sewage outflow), floating trash, and mass kills of organisms due to pollution. People are sensitive to these phenomena occurring in areas they access for recreation or other purposes as well as for simply knowing that clean waters exist. This goal scores highest when contamination from all potential sources is zero.

We include four components of pollution in the clean waters goal: **pathogens** (p), **trash** (t), **water quality** (w), and **sediment quality** (s). This decision was meant to represent a comprehensive list of the contamination categories that are commonly considered in assessments of coastal clean waters (Borja et al. 2008) and for which we could obtain data. The status of this goal is calculated as the geometric mean of the four data layers included.

$$CW_{i,t} = 100 * \sqrt[4]{p_{i,t} + t_{i,t} + s_{i,t} + w_{i,t}}$$

We used a geometric mean, as is commonly done for water quality indices (Liou et al. 2004), because a very bad score for any one subcomponent would pollute the waters sufficiently to make people feel the waters were 'too dirty' to enjoy for recreational or aesthetic purposes (e.g., a large oil spill trumps any other measure of pollution). However, in cases where a subcomponent was zero, we added a value of 0.01 (on a scale of 0 to 1) to prevent the overall score from going to zero. Given that there is uncertainty around our pollution estimates, a zero score resulting from one subcomponent seemed too extreme.

Although clean waters are relevant and important anywhere in the ocean, coastal waters drive this goal both because the problems of pollution are concentrated there and because people predominantly access and care about clean waters in coastal areas. The nearshore area is what people can see and where beach-going, shoreline fishing, and other activities occur. Furthermore, the data for high seas areas is limited and there is little meaningful regulation or governance over the input of pollution into these areas. We therefore calculate this goal only for the first 3 nm of ocean for each country's EEZ. We chose 3 nautical miles for several reasons, but found the status results to be relatively insensitive to different distances.

Goal Layers

Beach closures (pathogens)

Data description
This layer is derived from the U.S. Environmental Protection Agency (EPA) BEach Advisory and Closing Online Notification (BEACON) dataset. The most recent version (2.0) was used. The beach action (advisory and closures) dataset from the BEACON database was downloaded for all states in the assessment (ME, NY, NH, MA, RI, CT). This data contains beach-level information for closures including the length of the closure and the reson for closure. The Beach Days dataset was also downloaded and used to get an estimate of length of swim seasons.

Methods summary
We used the proportion of the swim season with beaches closed due to pathogens in the water as a proxy for the impact of pathogens in coastal waters (Clean Waters) and also as a measure of limited recreation access (Tourism & Recreation). First we looked at the average length of swim season by state using the Beach Days dataset. The season length varied from 97 to 104 so we set a single reference point as 100 days free of closures, representing the average length of the swimming season for the region.

State	Average # of beach days
CT	98
MA	101
ME	99.3
NH	97.3
RI	97.7
NY	104

Next we looked at beach closures by region. The beach action dataset identified the reason why a beach was closed which for the Northeast included elevated bacteria, sewage or runoff from rainfall. For each beach the total number of days closed due to these reasons was calculated and then divided by the reference point of 100 days to calculate the proportion of the season each beach was closed. This was then averaged across each region and done for every year 2005 to 2017.

Clean Waters goal layer
The pathogens data layer used in the Clean Waters goal had values for each region and year set equal to the average proportion of a regions swim season with beaches open, or the inverse of the pressure layer, where higher values indicate higher score.

Gapfilling:
None needed

Coastal trash

Data description: The data comes from the Ocean Conservancy's International Coastal Cleanup reports. The data was manually copied and pasted from the Ocean Conservancy website as it was not provided in an easy to use format. Each report table was added to an excel spreadsheet with one sheet for each year. The years included were 2005 - 2017.

Methods summary: The data was filtered to only those states in our regions (NY to Maine) as well as the Mid-Atlantic states for setting a reference point. The total pounds of trash collected per volunteer was calculated and used as the metric for coastal trash presence. The target reference point for this goal was zero pounds of trash collected per person and the limit reference point (maximum possible which would warrant a zero score) was set to reflect the highest amount collected over the same time period across all states on the US Mid-Atlantic and Northeast coasts (Maryland with 94.5 pounds of trash per person) plus a 10% buffer, as we did not think that the highest recorded historical level of trash collected necessarily reflects the highest potential trash that could be collected across the region in future years.

Clean Waters goal layer
Each region was scored by dividing the calculated trash per person by the reference point of 104 pounds for each year.

Gapfilling: Data was available for all years 2005 - 2017, and for each state in the region, so no gapfilling was required. The data was reported by state so both regions in Massachusetts were given the same score.

Sediment Quality

Data description
The sediment quality layer is derived from the US Environmental Protection Agency National Coastal Condition Assessment (NCCA) which has data available for 2010, 2005/2006 and 2001. The report includes information across four indices for randomly sampled sites to represent the condition of all coastal waters:

Biological Benthic Index, Water Quality Index (Phosphorus, Nitrogen, Dissolved Oxygen, Water Clarity, Chlorophyll a), Sediment Quality Index (contaminants & toxicity), Fish Quality Index. For the sediment quality layer we use the Sediment Quality Index (SQI) component of the report which measures contaminants & toxicity. The Sediment Quality Index provides data per site in categories of "Good", "Fair" and "Poor" according to thresholds established by the EPA. Some sites were assigned the category "Missing". These were removed before analysis.

SQI data was downloaded from the NCCA data portal for the 2010 assessment year. Data for 2005/2006 was not available in a usable format on the portal. Personal communication with people at the EPA led to retrieval of the 2005/2006 information which was emailed to us on August 27, 2018.

Methods summary
The data was filtered for sites in the Northeast region, identified by postal code. Each sampling site had an associated weight assigned by the EPA intended to account for how representative the site is for the region. Code was provided by contacts at the EPA to accurately account for these weights. Using the `spsurvey` R package we applied a categorical data analysis across all sampled sites. This returned an estimate (`p`) for the percent of the region that site represents, and their assigned category. We then assigned scores to each category (Good = 100, Fair = 50 and Poor = 0), multiplied them by the estimate (`p`) per site, and then summed up the score per region to get a full picture of sediment quality in each region.

Clean Waters goal layer. The sediment quality layer for the Clean Waters goal had values for each region and year set equal to the sum of site-weighted scores per region.

Gapfilling. This dataset was spatially comprehensive covering our entire region, but only contains two time points, 2005/2006 and 2010. The most recent field season was conducted in 2015 but unfortunately the data from that season has not yet been made available. This necessitated temporal gapfilling. Scores for each layer were the same for 2005 and 2006, and then gapfilled linearly by region for the years between 2006 and 2010. Scores for 2011-2017 were set equal to the 2010 scores.

Water Quality

Data description
The water quality layer is derived from the US Environmental Protection Agency National Coastal Condition Assessment (NCCA) which has data available for 2010, 2005/2006 and 2001. The report includes information across four indices for randomly sampled sites to represent the condition of all coastal waters: Biological Benthic Index, Water Quality Index (Phosphorus, Nitrogen, Dissolved Oxygen, Water Clarity, Chlorophyll a), Sediment Quality Index (contaminants & toxicity), Fish Quality Index. For our water quality layer we use only the Water Quality Index component of the report which measures Phosphorus, Nitrogen, Dissolved Oxygen, Water Clarity, and Chlorophyll a for each monitoring site. The Water Quality Index (WQI) provides data per site in categories of "Good", "Fair" and "Poor" according to thresholds established by the EPA. Some sites were assigned the category "Missing". These were removed before analysis.

WQI data was downloaded from the NCCA data portal for the 2010 assessment year. Data for 2005/2006 was not available in a usable format on the portal. Personal communication with people at the EPA led to retrieval of the 2005/2006 information which was emailed to us on August 27, 2018.

Methods summary
The data was filtered for sites in the Northeast region, identified by postal code. Each sampling site had an associated weight assigned by the EPA intended to account for how representative the site is for the region. Code was provided by contacts at the EPA to accurately account for these weights. Using the `spsurvey` R package we applied a categorical data analysis across all sampled sites. This returned an estimate (`p`) for the percent of the region that site represents, and their assigned category. We then assigned scores to each category (Good = 100, Fair = 50 and Poor = 0), multiplied them by the estimate (`p`) per site, and then summed up the score per region to get a full picture of water quality in each region.

Clean Waters goal layer

The water quality layer for the Clean Waters goal had values for each region and year set equal to the sum of site-weighted scores per region.

Gapfilling

This dataset was spatially comprehensive covering our entire region, but only contains two time points, 2005/2006 and 2010. The most recent field season was conducted in 2015 but unfortunately the data from that season has not yet been made available. This necessitated temporal gapfilling. Scores for each layer were the same for 2005 and 2006, and then gapfilled linearly by region for the years between 2006 and 2010. Scores for 2011-2017 were set equal to the 2010 scores.

References   Borja, A., Bricker, S. B., Dauer, D. M., Demetriades, N. T., Ferreira, J. G., Forbes, A. T., et al. (2008). Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. Marine Pollution Bulletin 56, 1519–1537. doi:10.1016/j.marpolbul.2008.07.005.

Liou, S.-M., Lo, S.-L., and Wang, S.-H. (2004). A Generalized Water Quality Index for Taiwan. Environ Monit Assess 96, 35–52. doi:10.1023/B:EMAS.0000031715.83752.a1.

U.S. Environmental Protection Agency, BEACON 2.0 (Beach Advisory and Closing Online Notification). (2019). Beach closures. Retrieved from https://watersgeo.epa.gov/beacon2/

U.S. Environmental Protection Agency. 2016. National Aquatic Resource Surveys. National Coastal Condition Assessment 2010 (data and metadata files). Available from U.S. EPA web page:https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys. Data from the National Aquatic Resource Surveys. Date accessed: 2018-07-31.

Ocean Conservancy. 2019. International Coastal Cleanup Reports. Accessed May 15 2020. https://oceanconservancy.org/trash-free-seas/international-coastal-cleanup/annual-data-release/