The Ocean Health Index Toolbox Manual

Contents

1 Introduction

1.1 The Ocean Health Index

The Ocean Health Index (OHI) project was created with a strong foundation in communication and collaboration. Following the initial publication by Halpern et al. 2012 in *Nature*, we have conducted annual global assessments to track changes in global ocean health. At the same time, we have been developing the **Ocean Health Index Toolbox** software that will enable others to conduct assessments on their own using the OHI framework. Here, the **Ocean Health Index Manual** describes how to use the Toolbox software.

OHI scores by goal are calculated at the scale of the reporting unit, which is called a **region** and then combined using a weighted average to produce the score for the overall area assessed, called a **study area**.

1.2 The Ocean Health Index Toolbox

The Ocean Helath Index Toolbox Application (App) is open-source software developed to facilitate Index calculations at any scale. With the Toolbox App, the Ocean Health Index framework can be customized to incorporate the data, indicators, and priorities regarding ocean-derived benefits throughout the area of interest. It can also compare how different management scenarios could affect overall ocean health, which can inform effective strategies for ocean resource management.

The Toolbox App runs using the statistical programming language R and displays input data and calculated scores graphically with maps, histograms, and tables. Goal scores are calculated for each region separately and then combined using a weighted average to produce the score for the overall study area assessed. The Toolbox can be used to calculate scores in smaller-scale **regional assessments** after all data has been gathered and decisions have been made regarding goal models, pressures and resilience, and reference points. The Toolbox runs on both a Mac and PC.

Modifications can be made to the global studies and to new regional studies, allowing calculation of scores with updated data and visualized with the Toolbox App. We recommend first exploring the results from the global assessment and becoming familiar with the Toolbox App before beginning a regional assessment.

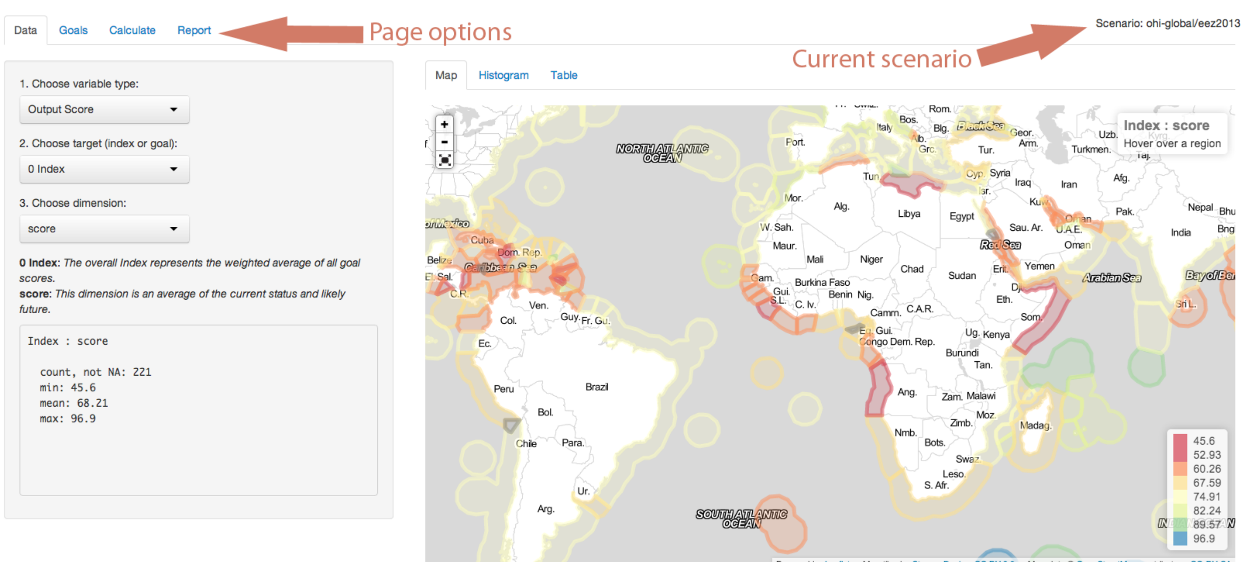
Requirements: an Internet connection for the initial Toolbox App installation.

1. Install the latest version of R.
2. Launch the Toolbox Application.
3. Explore global results using the Toolbox App

2 Overview of the OHI Toolbox App

The Toolbox Application (App) runs through a web browser and has three page options: *Data*, *Calculate*, and *Report*. The *Data* page provides several ways of viewing the data (*Map*, *Histogram*, *Table*), and is the default home-screen when the Toolbox is opened. The *Calculate* page is used to calculate goal scores based on the information in the file system (data can be changed and tested). The *Report* page provides different methods of reporting the data.

The scenario is displayed in the upper-right corner of the Data page. This identifies the scale (example: global) and year (example: 2013) of the information displayed.

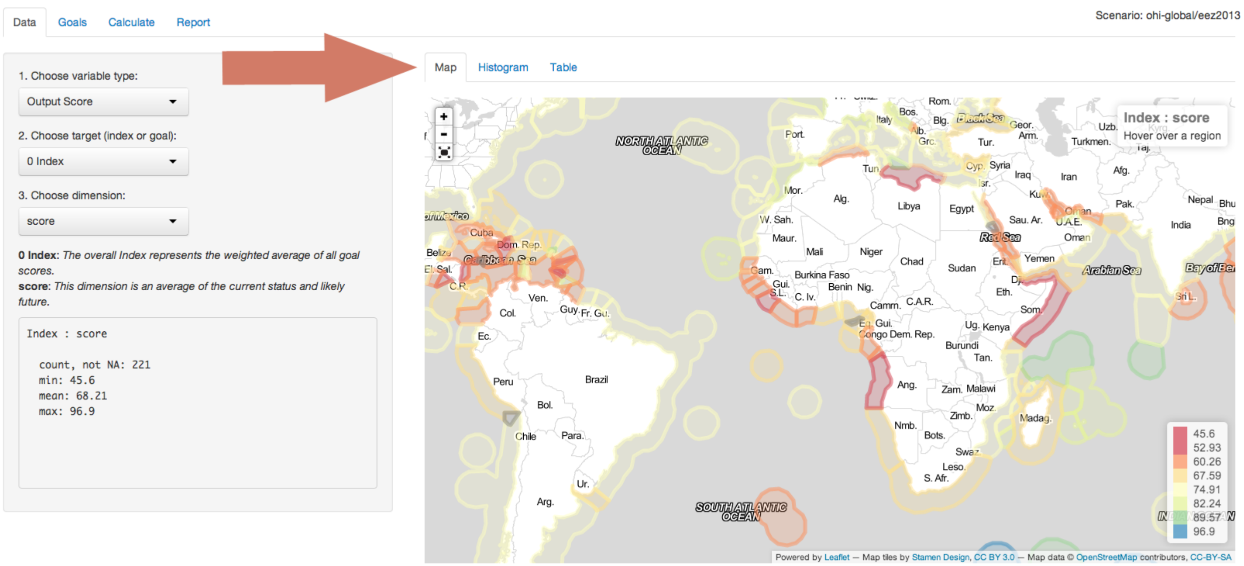


2.1 Data Page

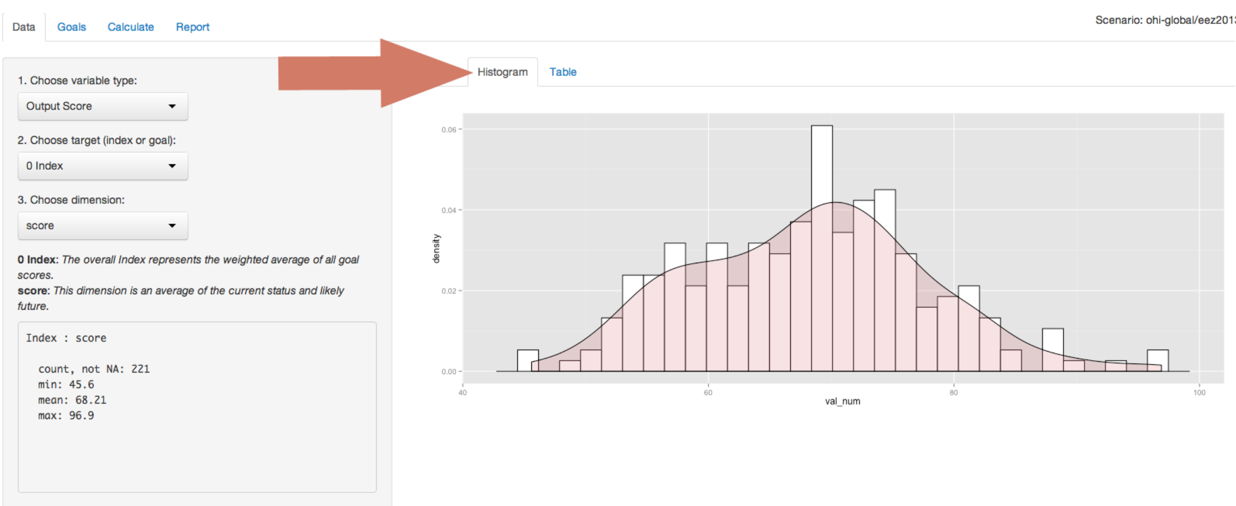
2.1.1 Overview of display options

The Data Page displays input data or calculated scores for each goal parameter, and presents the information as a Map, Histogram, or Table. These options are presented as tabs located above the map (Map view is the default display option for the Data page).

- Data displayed in Map form:



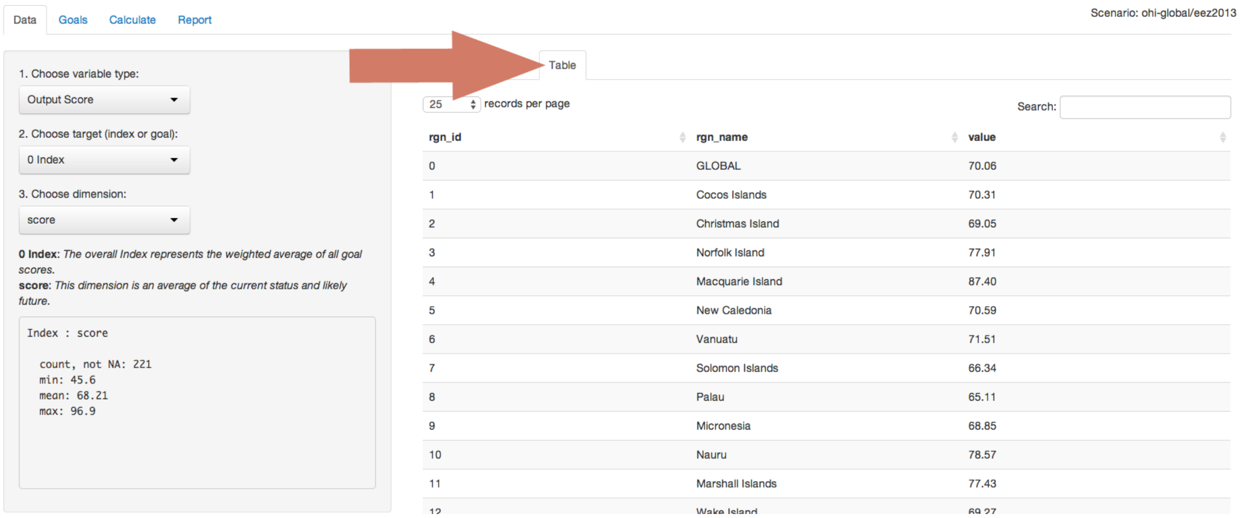
- Data displayed in Histogram form:



- Data displayed in Table form:

For all visualization options, statistics and metadata are displayed below the drop-down menus on the left side of the page.

Insert figure here



2.1.2 Overview of variable options

The Data page has drop-down menus from which the user chooses the data to be displayed by the Toolbox:

1. **variable type**
   - *Output Score: the calculated score*
   - *Input Layer: the data layer used to calculate the score*
2. **target**
   - *Overall Index*
   - *Ten goals*
   - *Eight sub-goals*
3. **dimension** (for score variables) or layer (for data layer variables)
4. **category** (this will appear for data layer variables with categories)
5. **year** (this will appear for data layer variables with multiple years)

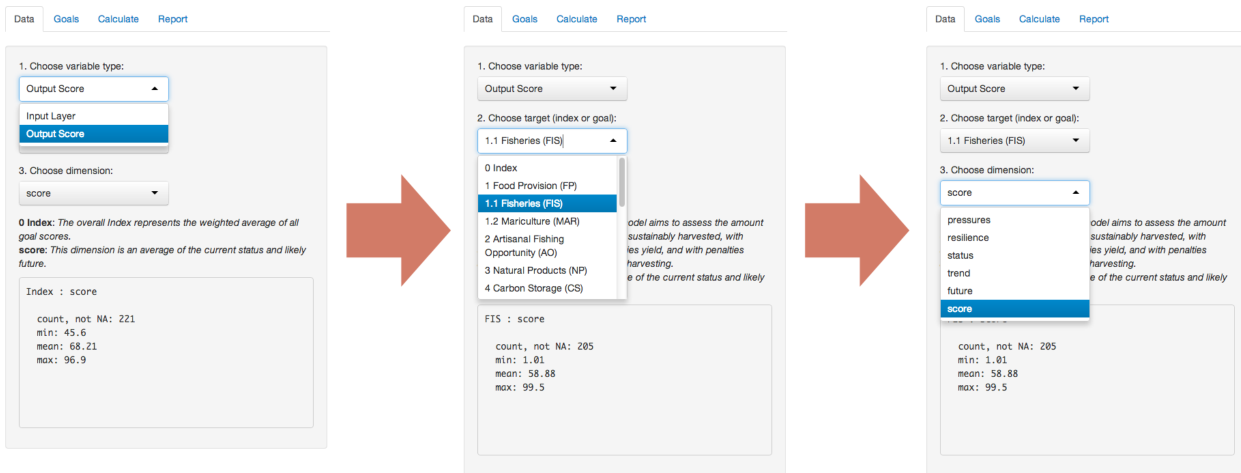
For example:

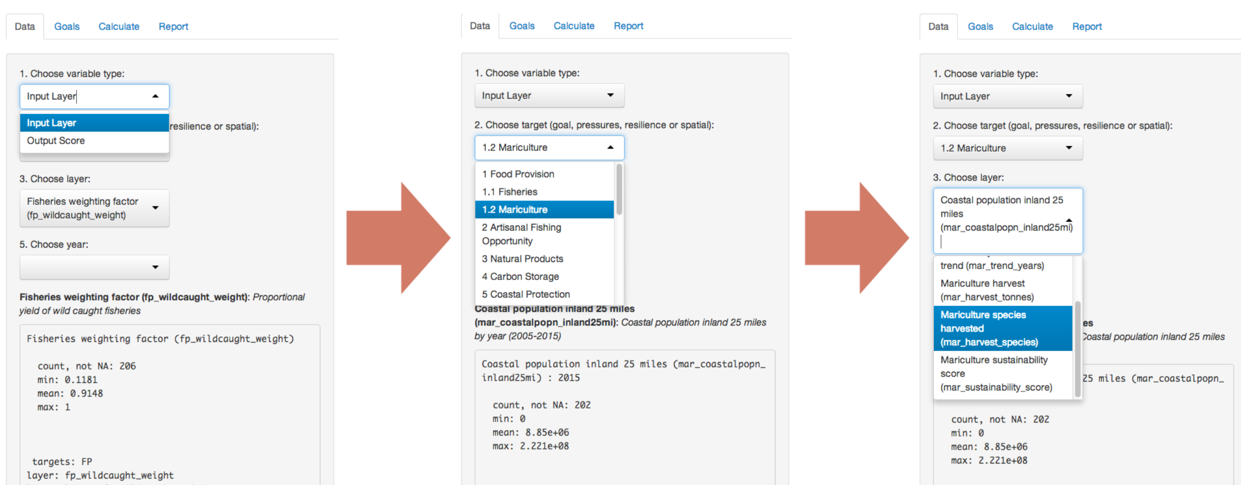
> **Output Score > Fisheries (FIS) > score** shows the final calculated score for the fisheries sub-goal in the current scenario (2013).

> **Input Layer > Mariculture > Mariculture species harvested** shows the different harvested species that were used to calculate the mariculture sub-goal in the current scenario (2013).

2.1.2.1 Output Score If the user selects 'Output Score' as the variable type (which is the default), they will then be able to choose a target and a dimension. The target is a goal or sub-goal, and the dimension indicates the calculation to be reported.

- Target options (with numeric identifiers):
  - Overall Index
  - Ten goals and eight sub-goals
- Dimension options:
  - Status, trend, future state, pressures, resilience, and score


Data
Goals
Calculate
Report
1. Choose variable type:
Output Score
Input Layer
Output Score
3. Choose dimension:
score
0 Index: *The overall Index represents the weighted average of all goal scores.*
score: *This dimension is an average of the current status and likely future.*
Index : score
count, not NA: 221
min: 45.6
mean: 68.21
max: 96.9
2. Choose target (index or goal):
1.1 Fisheries (FIS)
0 Index
1 Food Provision (FP)
1.1 Fisheries (FIS)
1.2 Mariculture (MAR)
2 Artisanal Fishing Opportunity (AO)
3 Natural Products (NP)
4 Carbon Storage (CS)
FIS : score
count, not NA: 205
min: 1.01
mean: 58.88
max: 99.5
pressures
resilience
status
trend
future
score


Data
Goals
Calculate
Report
1. Choose variable type:
Input Layer
Input Layer
Output Score
resilience or spatial):
3. Choose layer:
Fisheries weighting factor (fp_wildcaught_weight)
5. Choose year:
Fisheries weighting factor (fp_wildcaught_weight): *Proportional yield of wild caught fisheries*
Fisheries weighting factor (fp_wildcaught_weight)
count, not NA: 206
min: 0.1181
mean: 0.9148
max: 1
targets: FP
layer: fp_wildcaught_weight
2. Choose target (goal, pressures, resilience or spatial):
1.2 Mariculture
1 Food Provision
1.1 Fisheries
1.2 Mariculture
2 Artisanal Fishing Opportunity
3 Natural Products
4 Carbon Storage
5 Coastal Protection
Coastal population inland 25 miles (mar_coastalpopn_inland25mi): *Coastal population inland 25 miles by year (2005-2015)*
Coastal population inland 25 miles (mar_coastalpopn_inland25mi) : 2015
count, not NA: 202
min: 0
mean: 8.85e+06
max: 2.221e+08
Coastal population inland 25 miles (mar_coastalpopn_inland25mi)
trend (mar_trend_years)
Mariculture harvest (mar_harvest_tonnes)
Mariculture species harvested (mar_harvest_species)
Mariculture sustainability score (mar_sustainability_score)

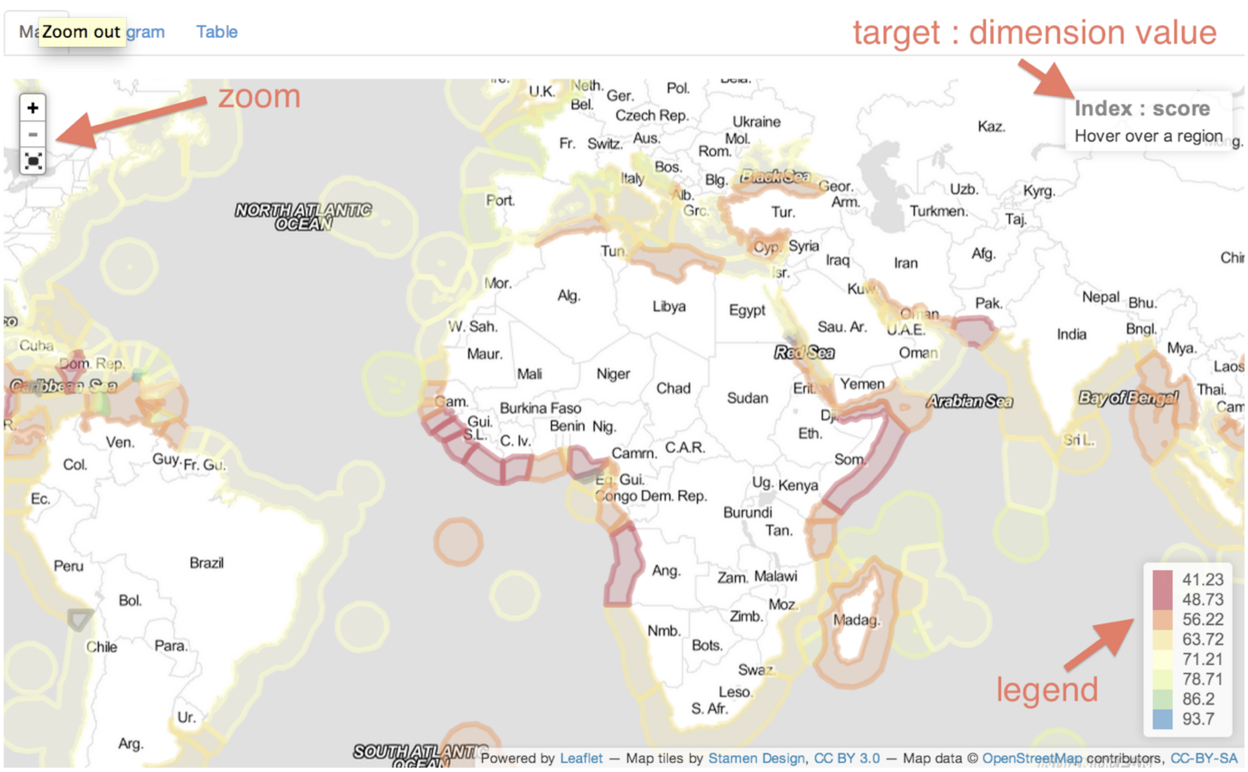
2.1.2.2 Input Layer

If the user selects "Input Layer" as the variable type, they will be able to choose a target and a specific layer that is associated with that target. If that layer has multiple categories or years available, the user will be able to select a preference, with a default being the first category alphabetically and the most recent year.

- Target options (with numeric identifiers):
  - Ten goals and eight sub-goals
  - Pressures
  - Resilience
  - Spatial
- Layers options:
  - Specific data layer associated with the target
- Category options:
  - This option will appear if the layer identified has categories
- Year options:
  - This option will appear if the layer identified has multiple years

2.1.3 Map

The map displays data for every region as reported in the scenario.



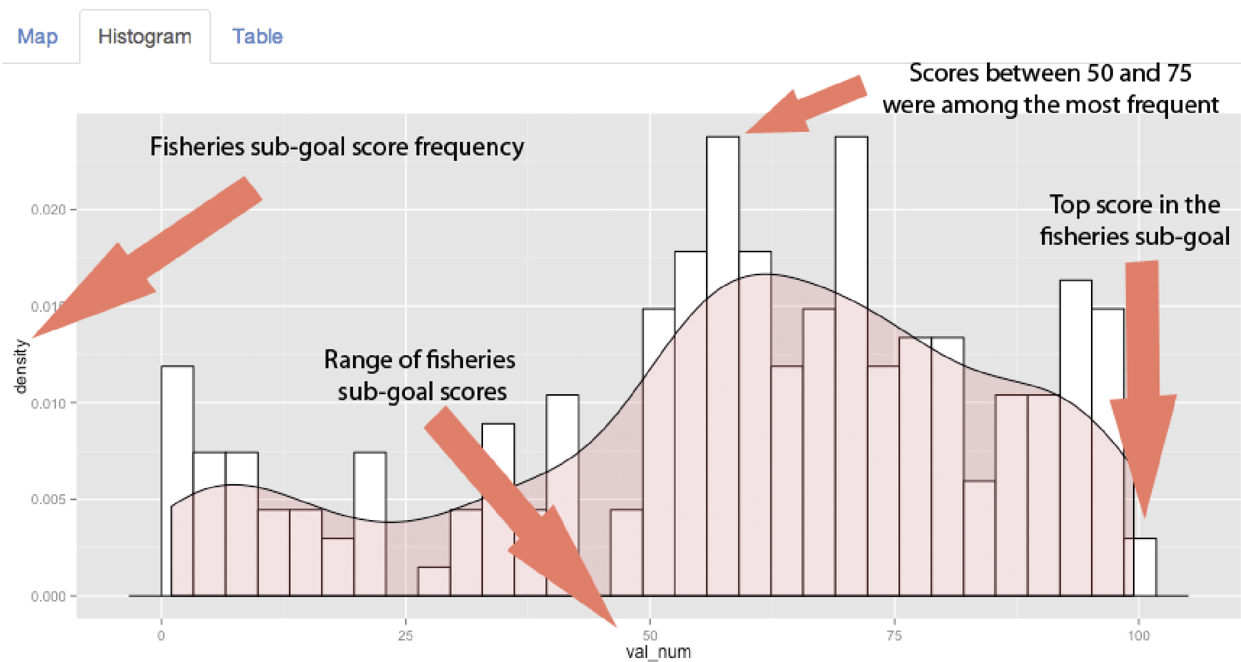
Zooming in and out is possible with the plus/minus in the upper left of the map, and a color legend is displayed in the lower right corner of the map. The legend provides a linear colormap of the data and the range of values will change as different variables are selected.

The target and dimension chosen from the drop-down menu is also displayed as text on the top right corner of the map. Below this information, the region name, (region number), and value appear if the user's mouse hovers over a specific region. When the value of that region is displayed, the region will darken on the map. If the mouse is not indicating a specific region, this area will instruct to 'hover over a region'.

2.1.4 Histogram

The histogram shows the distribution of the selected variable as the number of observations per value bin (white bars) and a smoothed density function (pink shading).

> For example, if the **Output Score > Fisheries > score** sub-goal is selected, the information displayed by the histogram is as follows:



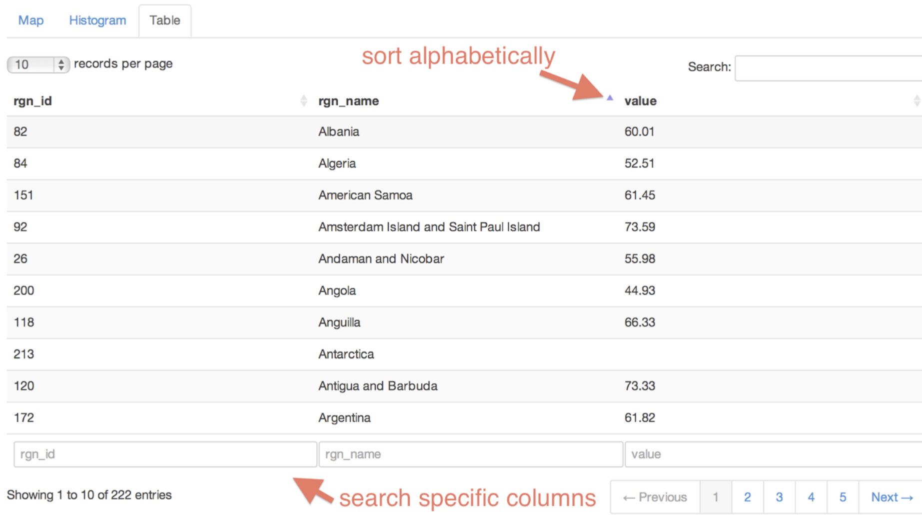
2.1.5 Table

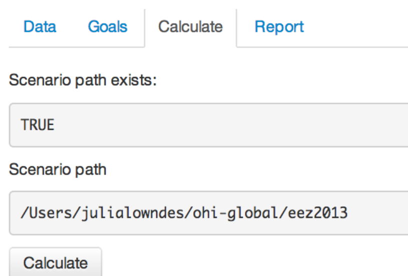
The table displays the target value for each region and the overall study area. It provides an identifying code (rgn_id), name (rgn_name), and value (value) for each.

It is possible to search the data using the search bar at the upper right of the table (caps-sensitive). It is also possible to search within individual columns by entering the search term at the bottom of each column; this option is more easily displayed by changing the number of records per page to 10 using the drop-down menu. Data may also be sorted the arrows at the left of each column name. Sorting by rgn_id is the default.

2.2 Calculate Page

The Calculate page is used to calculate the Index. The Calculate page displays the "Scenario path" in the user's home directory where the folders are located.





2.3 Report Page

The Report page allows the user to create a report of results as .html and .pdfs files.

It is currently possible to create flower plots and tables with several options, including saving global figures or figures per country. In future versions of the Toolbox App, more options for visualizations, including maps, histograms, and equations, will be possible. It will also be possible to compare different scenarios.

3 Explore Global Results with the Toolbox App

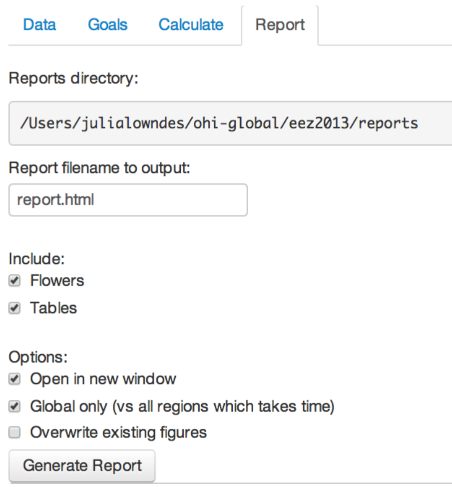
develop

3.1 develop

4 Introduction to OHI Regional Assessments

Regional assessments use the Ocean Health Index framework to study smaller spatial scales, often where policy and management decisions are made. Regional assessments incorporate local priorities, higher-resolution data and indicators, and use tailored goal models and reference points, which produce scores better reflecting local realities. The OHI is designed to combine existing work into a comprehensive assessment of ocean health.

The process for developing a regional assessment can be as valuable as the final calculated scores, since it creates an ocean alliance that combines knowledge and cultural values from many different perspectives and disciplines. Calculating Ocean Health Index scores at regional scales



requires engagement from as many different groups as possible, including research institutions, government agencies, policy groups, non-governmental organizations, and the civil and private sectors.

The Ocean Health Index framework is flexible to accommodate regional priorities: goal models can be modified to incorporate the best available local data and indicators, or excluded entirely to best represent regional attributes. Similarly, pressures and resilience measures can be refined using local data and indicators. Calculated Index scores are only as 'good' as the data on which they are based, so finding the best data or indicators available is fundamental.

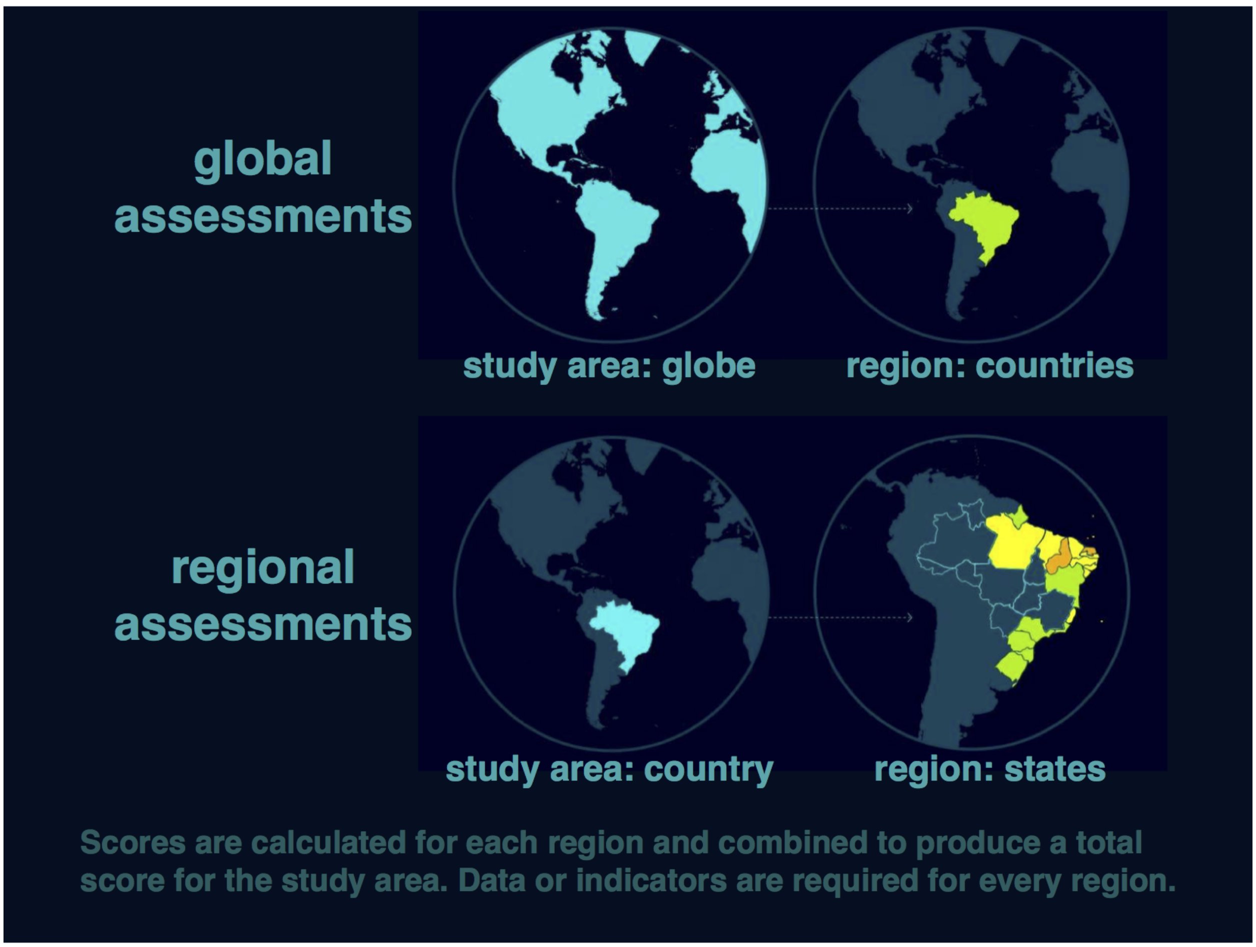
Scores by goal are calculated weighted average to produce the score for the overall area assessed, called a **study area**. In regional assessments, coastal states or provinces are often the **regions** within a country, which is the **study area**.

Regional assessments have been published in peer-review journals for the U.S. West Coast (Halpern et al. 2014) and for Brazil (Elfes et al. 2014), and more are underway.

4.1 Before getting started

Before beginning a regional assessment, it is important to have planned several things:

- Develop strategic plan and timeline
  - determine resources available
  - outline timeline with any meetings and workshops necessary
  - engage management early for results to be most useful (optional)
- Assemble a qualified team with diverse skills
  - have a broad scientific understanding and experience with environmental policy
  - are comfortable with handling large data sets, making decisions, and thinking creatively
  - can work with the software program R and user-created packages
  - can work with ArcGIS or other spatial analysis software
  - collaborate well in a multidisciplinary team, remotely and in person
- Funding
  - greatly depends on the local context
  - potentially needed for a management and scientific team, workshops and meetings (including travel), communications, policy engagement, and operating costs
- Policy and management interest (optional)


global
assessments
study area: globe
region: countries
regional
assessments
study area: country
region: states
Scores are calculated for each region and combined to produce a total
score for the study area. Data or indicators are required for every region.

– inform government policies to improve ocean health is most effective if there is interest and engagement from policy makers
– requires ongoing communication during Index development to best inform management actions that could have measurable impacts.
– repeated assessments as new data become available enable tracking ocean health through time and evaluating management priorities

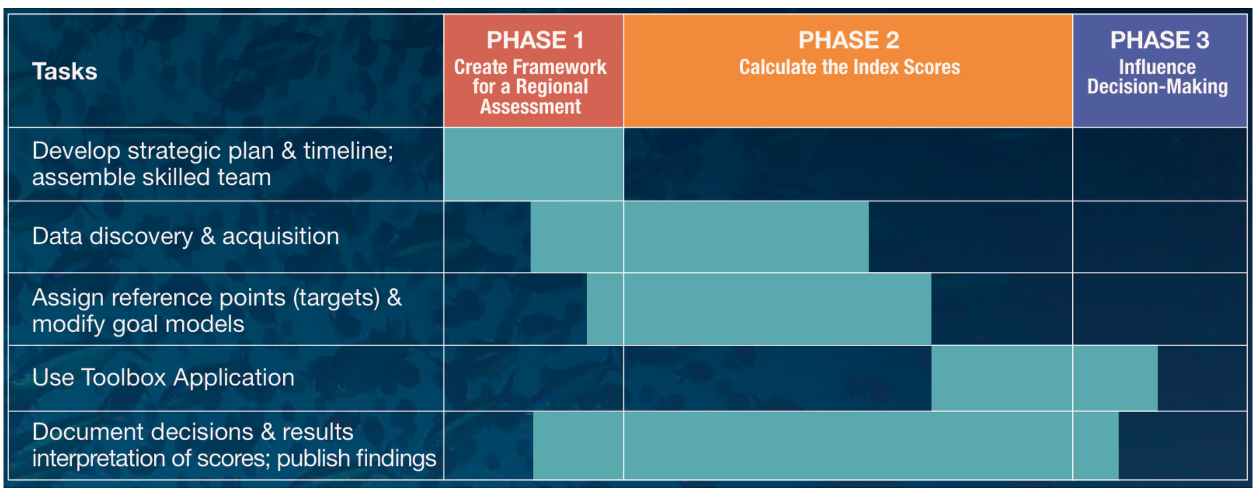
5 Conducting a Regional Assessment

5.1 What to expect when conducting a regional assessment

There are key considerations and processes that will be a part of every regional assessment, however, how each regional assessment is conducted will be unique depending on the local context. For example: what data and indicators are available will determine how goal models are able to be tailored to the region, and what skillsets and resources are involved will affect the time it takes to complete the assessment.

The time required to complete a regional assessment depends on the local context and available resources, but the proportion of time required for different phases of the assessment is more predictable.

Expect for strategic planning, discovering and gathering data, and developing reference points and models, to comprise > 80% of the time alloted for the regional assessment.



5.2 What is provided

The Ocean Health Index Toolbox can be used to calculate scores for regional assessments. When you begin a regional assessment, you will be provided with a **repository**: a set of Toolbox files and software scripts specific to your study area. The data in the files provided serve as a template: they are the data used in the global assessments, extracted for your study area. These data files will need to be updated with locally available data and indicators so as to give a more fine-scale assessment of ocean health in your study area. However, if better data are not available, it is possible to use the files provided.

Our team of scientists and managers is prepared to provide guidance for regional assessments, from initial meetings to discussions about model development to disseminating results.

5.3 Recommendations

We recommend to remember the following as you develop your approach:

- People are part of ocean health
  - economic and social data are included, and certain terrestrial data near the coast
- Document all decisions made
  - document data sources, urls, date of access, reference points set, why models were modified the way they were (and not other potential ways)...
  - important for transparency, describing methods and explaining results in reports and publications, and for reproducibility (for any future comparable assessments in your study area)
- Assessments can use a mixture of regional-, national- and global-scale data
- Scores are calculated by region
  - comparing scores between sub-national regions is a main reason of conducting a regional assessment
  - regional scores are combined by weighted average to calculate scores for the study area
  - goal models are the same across all regions; data are specific to the region. For example, habitat-based goal models are the same in each region, but which habitats contribute to the final scores will depend on which habitats are present in each region
- Data do not need to be at the same spatial or temporal scale
  - each assessment should represent the best understanding of ocean health at the point in time. For example if fisheries data are available from 1980-2011 and tourism from 2008-2012; these can be used together to calculate Index scores, which will be the representation of current ocean health

5.4 Checklist: To Do Before Using the Toolbox

Most of the time spent while conducting a regional assessment occurs before using the Toolbox, although being familiar with the Toolbox's composition and structure can help guide you. The following will not be checked-off in sequence: there is a lot of back and forth as you discover data, develop reference points and models, revisit other data possibilities.

Here is a checklist of decisions and steps that must be done before the Toolbox can calculate regional goal scores:

- Understand the philosophy of the Ocean Health Index
  - understand what goals represent and how they are modeled
  - understand what pressures and resilience are included
  - understand how reference points are set
- Identify local characteristics and priorities
  - should any goals be removed?
  - should any goals be redefined?
  - should other goals be added?
  - should goals be weighted unequally?
  - what are pressures to the local system?
  - what resilience measures (laws, regulations) are in place locally?
  - what local cultural preferences or priorities should be captured in the assessment?
- Decide the spatial scale for regions within the study area

  - we recommend multiple regions
  - regions within the study area should be at the smallest scale the data allow
  - in what spatial scale are data most frequently reported?
  - in what spatial scale are policy decisions made (optional)?

- Be familiar with the global inputs to your Toolbox repository
  - understand the data and models used in global assessments
  - prioritize which data and models should be updated
  - use the Toolbox repository as a registry to organize data layers

- Discover and gather appropriate data and indicators
  - what local data and indicators are available for goal models?
  - what local data and indicators are available pressures and resilience?
  - how would reference points be set for local data?
  - do local data and indicators capture the philosophy of the Ocean Health Index?
  - are local data and indicators at the appropriate spatial and temporal scales?
  - process and format data and indicators to create data layers* for the Toolbox

- Modify goal models and set reference points:
  - can goal goal models be refined using locally available data and indicators?
  - can reference points be refined using locally available data and indicators?

* Each data component that is included in the OHI is called a **data layer** because it will be combined with others to create the most complete picture of ocean health. Many data layers are rescaled from 0-1 to be combined with others on the same unitless scale.

5.5 Discovering and Gathering Appropriate Data and Indicators

The OHI spans disciplines and integrates diverse data and sources to give a comprehensive assessment of ocean health. A hallmark of the OHI is that it uses freely-available data to create models that capture the philosophy of individual goals, and finding appropriate data requires research and creativity. There are many decisions to make when gathering from disparate sources, deciding reference points, and developing goal models.

Index scores are a reflection of data quality, and thus accessing the best data and indicators available is of highest importance.

Existing data and indicators can be gathered from many sources across environmental, social, and economic disciplines, including:

- government reports
- government websites
- academic literature
- masters and PhD theses
- university websites
- non-profit organizations

All data will be rescaled to specific reference points (targets) before being combined; therefore setting these reference points at the appropriate scale is a fundamental component of any OHI assessment. This requires the regional assessment team to interpret the philosophy of each Index goal and sub-goal using the best available data and indicators.

5.5.1 Data gathering responsibilities

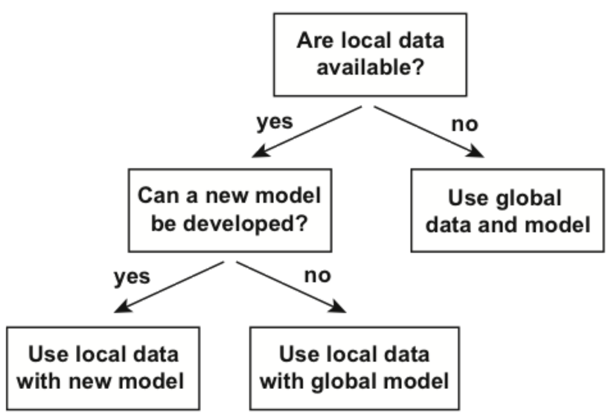
Gathering appropriate data requires searching for and accessing freely-available data. It is important that team members responsible for data discovery make thoughtful decisions about whether data are appropriate for the regional assessment, and that they also get feedback from the full team to discuss the merits of different data sources. Data discovery and acquisition can be quite an iterative process, as there are both practical and philosophical reasons for including or excluding data, in addition to requiring access to the data.

Index scores can be recalculated annually as new data become available. This can establish a baseline of ocean health and serve as a monitoring mechanism to evaluate the effectiveness of actions and policies in improving the status of overall ocean health. This is good to keep in mind while looking for data: will it be available again in the future? It is also important to document the sources of all data so that it is both transparent where it came from and so that you will be able to find it again in the future.

5.5.2 The process of data discovery

In terms of philosophical considerations, the most important thing to remember when gathering data is that the data must contribute to measuring ocean health. Many data sources that enhance our knowledge of marine processes may not directly convey information about ocean health and may not be appropriate within the OHI framework. Because of this, compiled indicators can sometimes be more suitable than would raw data measuring single marine attributes.

Begin by understanding and comparing the best approaches used in assessments that have been completed, including global assessments, Brazil, and the US West Coast. If finer-resolution local data were available in the study area, they could be used either in a newly developed regional goal model using locally appropriate and informed approaches, or in the existing global goal model. When local data were not available, the same global-scale data were used with the original global goal model, which is least desirable because it does not provide more information than the global study. When looking for data, the following decision tree may be useful. This should be a goal-by-goal process:



Searching for data requires exploring data sources beyond any single discipline, and a good place to start is with an internet search. Internet searches can lead to published data in government and non-governmental organization reports, peer-reviewed articles, and masters and doctoral dissertations. Not everything will be freely available online but it is sometimes possible to request access.

It is good practice to keep detailed notes of attributes of each potential data layer, since there may be different options to work with. Searching for data by goal is a good approach, although some data layers will be used for multiple goals.

5.5.3 Requirements for data layers

Four requirements to remember when investigating (or 'scoping') potential data layers are:

1. relevance to ocean health
2. how to the reference point
3. spatial scale
4. temporal scale.

5.5.3.1 Relevance to ocean health There must be a clear connection between the data and ocean health, and determining this will be closely linked to each goal model.

5.5.3.2 Reference point As each data layer must be scaled to a reference point, as you consider different data sources it is important to think about or identify what a reasonable reference point may be. Ask the following types of questions as you explore data possibilities:

- Is there a known relationship associated with these data?
- Have policy targets been set regarding these data?
- Would a historic target be appropriate?
- Could a region within the study area be set as a spatial target?

5.5.3.3 Appropriate spatial scale Data must be available for every region within the study area.*

5.5.3.4 Appropriate temporal scale Data must be available for at least five years to calculate the trend. Longer time series are preferable because this can be used to set temporal reference points.*

* It is not always possible to meet the spatial and temporal requirements with each data layer. In these cases it can still be possible to use these data if appropriate gap-filling techniques are used. It is important that data satisfy as many of these requirements as possible, and in cases where creative ways of working with such data are not possible, it might be better to exclude these data from the analyses and try a different approach.

5.5.4 Example: US West Coast data discovery

Below are examples of some decisions made when exploring available data for the US West Coast regional assessment. Determining whether certain data could be included started with a good understanding of the data layers and models included in the global assessment, and because the US West Coast is a data-rich region, finer-resolution local data could be used in place of many of the global data layers.

5.5.4.1 Reasons data were excluded There are a lot of existing data that contribute to our scientific understanding of ocean processes and interactions but that are not ideal for the OHI. Reasons to exclude data occur both on a practical level (do data adhere to the requirements above?) and on a philosophical level, which requires reflecting on the relationship with ocean health. Some common reasons for excluding data are listed below:

- **The data do not cover the entire area of the reporting region**. The state of California had excellent, long-term data on public attendance at state parks that would have been quite useful in the calculation of the tourism and recreation goal. However, Oregon and Washington did not have these same data so they were not used.

- **There is not a clear and scientifically proven connection between the metric described by the data and ocean health**. Along the US West Coast, kelp beds are a very important habitat because of their contribution to biodiversity and coastal protection. However, kelp coverage is quite variable and is driven primarily by abiotic natural forcing (wave/storm disturbance and temperature) and thus kelp coverage is not a good metric of ecosystem health. For these reasons kelp coverage was not included in the assessment.

- **The feature being measured may provide benefits to people, but this feature is not derived from the ocean**. Sea walls and riprap provide coastal protection to many people along the US West Coast. However, these structures are not a benefit that is derived from the ocean itself, so only biogenic habitats were included in the calculation of this goal. These data can be included as a pressure due to habitat loss.

- **Data collection is biased and might misrepresent ocean health**. The US Endangered Species Act identifies a species list focused on species of concern within the US. As such, these data are biased in the context of ocaean health since they assess only species whose populations may be in danger. For the calculation of the biodiversity goal, using these data would be inappropriate because this goal represents the status of all species in the region, not just those that are currently of conservation concern. Using these data may have shown the status of biodiversity to be lower than it really is because the selection of species to assess was already biased towards species of concern.

5.5.4.2 Creative approaches to using data

- **Time series data are not long enough to calculate a trend or a reference point (when a historical reference point is most appropriate)**. For the US West Coast, available data measure the current extent of seagrass habitats, however, these only exist for one time point in most areas so could not be used to calculate the trend or set a historical reference point. As these were the best data available for habitat coverage, we built a model to calculate the status and trend of seagrass habitats using other data that were available over time. A reasonable approach was to model the pressures exerted on seagrasses over time as a proxy for seagrass health.

5.6 Formatting Data for the Toolbox

5.6.1 develop

Transfer information from this .xlsx sheet into this doc

Each data layer must then be formatted in a specific way to be used by the OHI Toolbox App; (See: formatting_data_for_toolbox.

5.6.2 info

6 Installing the Toolbox for a Regional Assessment

[[develop

6.1 Getting Started

Ocean Health Index assessments have a specific organization of data files, registry files, and R scripts that work together to do calculations and present the final scores graphically. These files and scripts are packaged together in a folder called a **repository**. Each Ocean Health Index assessment has its own repository.

Repositories are stored online through GitHub, which enables easy collaboration with versioning capabilities. Learn more about GitHub repositories and OHI.

It is possible to conduct a regional assessment by downloading a template repository from GitHub. We recommend having a GitHub user account and syncing with the online repository: this faciliates collaboration within your team but also allows us to provide support when you need it. However, it is possible to work independently without GitHub (after the initial download), although we cannot offer much support in this case.

Requirements: an Internet connection for the initial Toolbox App installation and continued access if syncing with GitHub (highly recommended); proficiency with data management and the coding language **R**; proficiency in spatial software such as ArcGIS (potentially, depending on the regional assessment strategy).

1. Begin a regional assessment
2. Become familiar with the Toolbox file system
3. Modify the Toolbox to calculate regional scores.

6.2 OHI and GitHub

The Ocean Health Index (OHI) Toolbox is written in R and uses GitHub, an open-source development platform, to develop and share software and data. GitHub has powerful versioning capabilities, which allows changes to be archived and tracked by each user. This is incredibly useful to not only document what work has been done, but how it differs from work done in the past, and who is responsible for the changes: it provides a record of all changes made. GitHub enables easy collaboration with others, including the OHI team.

OHI has several **repositories** ('repos') where data and code are stored. From the GitHub glossary:

> A repository is the most basic element of GitHub. They're easiest to imagine as a project's folder. A repository contains all of the project files (including documentation), and stores each file's revision history. Repositories can have multiple collaborators and can be either public or private.

6.2.1 OHI regional assessments and GitHub

OHI-Science is the 'owner' of the OHI repositories stored on GitHub, and individual users contribute to these repositories when they have permission. A user works locally on their own computer, but syncs their work back to the repo on the server. Because there can be multiple users working on the same repo, there are specific steps involved, and GitHub has specific words for each of these steps, including **cloning** a repo from the server to a local repository, **committing** any changes made with a description, **pulling** any updates from the server's repo, and finally **pushing** committed changes back to the server.

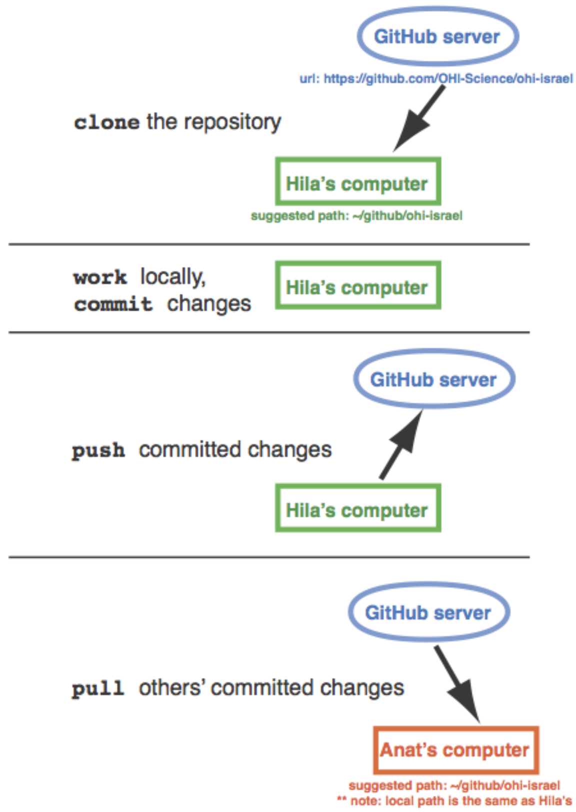
	`github.com/ohi-science/[repo]` (web)	`github.com/[user]/[repo]` (web)	`~/github/[repo]` (local)
->		-> fork	-> clone
<-	merge pull request {admin} <-	<- pull request	<- push, <-> commit

6.2.1.1 GitHub flow with *Fork & Pull Model*

where: * `[repo]` is one of OHI-Science repositories: ohicore, ohiprep, etc. * `[user]` is your GitHub username

see also: * GitHub flow in the browser

This example illustrates GitHub's collaborative workflow with the `ohi-israel` repo owned by `OHI-Science`:



6.3 Accessing repositories with GitHub

We recommend that groups interested in conducting OHI assessments do so using GitHub, which is an online interface for version-control software called *git*. This will enable collaboration and transparency, and will provide access to the latest developments in the Toolbox software. To get started, follow the steps below. The OHI team will create a repository for your regional assessment.

It is also possible to conduct a regional assessment without GitHub. We do not recommend this because it is more difficult for you to track progress and decisions, and for us to help or advise you. However, if this is preferable, see the instructions here. We can always get your filesystem onto GitHub later.

6.3.1 Install R and RStudio

Make sure you have the most current version of R and RStudio. (Hint: type `sessionInfo()` into the R console to see which version of R you have installed, and update it if this is not the most recent version available on CRAN).

Figure 1: software required for regional assessments with version control

6.3.2 Create a GitHub account

GitHub is an online interface for *git*, and to use it you must create a GitHub account. You will use this username and password when you install *git* and access your GitHub repository (next two steps below).

6.3.3 Install *git* software

git is version-control software that you will need to install on your computer.

Download *git* and follow the install instructions. There are specific settings to select for **Windows** and **Mac**.

6.3.4 Set up your Git Identity

After downloading and installing, you will then need to set up your Git Identity, which identifies you with any changes made. You will use the command line:

- **Mac**: launch the Terminal application (Applications > Utilities > Terminal)
- **Windows**: go to command line in Windows (Start > Run > cmd)

Substitute your GitHub username instead of jdoe:

```
git config --global user.name jdoe
```

and then: substitute the email address you used to create your GitHub account:

```
git config --global user.email johndoe@example.com
```

You can check settings with the following:

```
git config --list
```

Quit the Terminal after typing:

```
exit
```

6.3.5 Clone your ohi-[assessment] repo to your computer

The OHI team will create a repository for your assessment once you provide your GitHub username: email your username to bbest@nceas.ucsb.edu or lowndes@nceas.ucsb.edu.

The repository will be backed up and stored on github.com and you will also have a local version on your computer.

** **Important:** Please create a folder called `github` in your root directory and save the repository there. This will make collaborating possible, particularly between Macs and Windows machines. The directory for the assessment `~/github/ohi-[assessment]` will therefore be:

- on a Mac: `/Users/[User]/github/ohi-[assessment]`
- on Windows: `C:\Users\[User]\Documents\github\ohi-[assessment]`

Downloading a repository from GitHub is called **cloning** and there are multiple ways to do it. Cloning allows you to make changes on your local computer and sync them to the online repository, described elsewhere (GitHub-OHI repositories).

6.3.5.1 Cloning options

There are several options to clone the ohi-[assessment] repository to your local machine:

1) The GitHub App for Mac or Windows. After installing the GitHub App, log in with your username and passord. Then follow instructions for cloning (see the *Getting started guide* on a Mac and Windows)
2) The command line (Windows) or Terminal (Mac). For example, clone `ohi-israel` into the `~/github` folder:

```
cd ~/github
git clone https://github.com/OHI-Science/ohi-israel.git
cd ohi-israel
```

3) RStudio. This is best after the initial clone, since RStudio occassionally has trouble with setting the username / password.

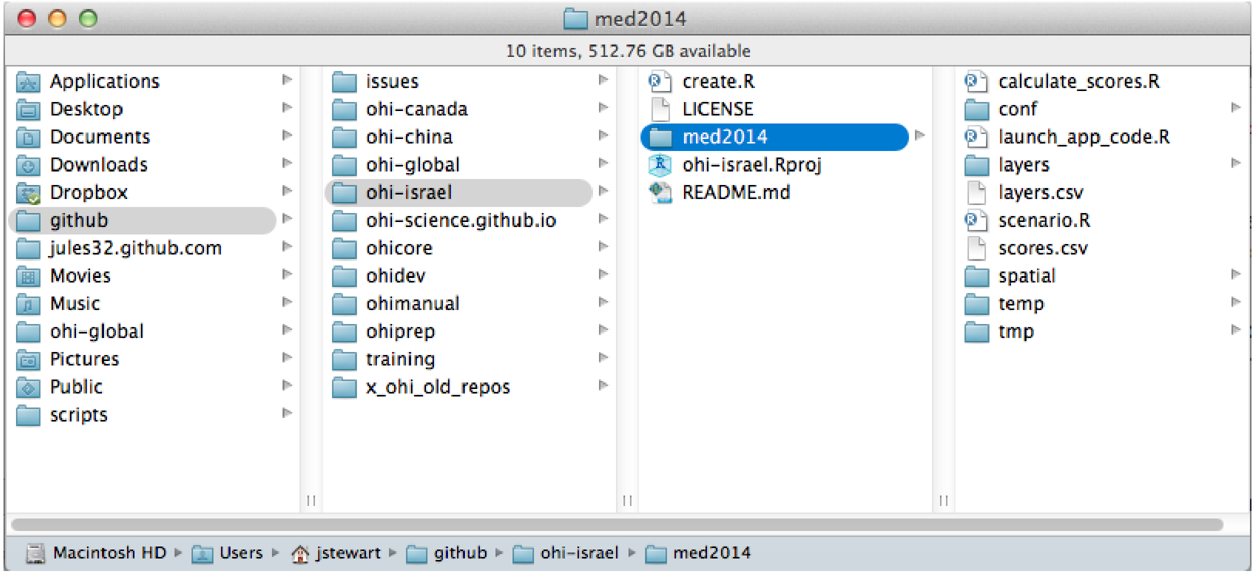
6.3.6 Working locally

All of the files available from the online server repository `https://github.com/OHI-Science/ohi-[assessment]` are now cloned on your local computer, in the location you specified. We recommend cloning the repository to this file path: `~/github/ohi-[assessment]`.

You will work on your computer to change the files in `~/github/ohi-[assessment]` to reflect the desired modifications your team has identified for your regional assessment. All changes within this folder will be tracked by GitHub when you commit and sync these changes, *even if you make these changes outside of R or RStudio*. This means that you can delete or paste files in the Mac Finder or Windows Explorer, edit `layers.csv` and `goals.csv` in Excel or a text editor, and modify `functions.r` in RStudio, and commit all these changes to GitHub to track the changes you make.

6.3.7 Committing and pushing changes locally

Committing and pushing changes can be done using RStudio or the Github App.



6.3.8 Using RStudio to commit and push changes to GitHub ohi-[assessment] repository

Launch your `ohi-[assessment]` project in RStudio by double-clicking the `ohi-[assessment].Rproj` file. From RStudio, Commit locally, associating a message with each set of changes.

When you modify or add a file, the file will appear in the 'Git' window. In the example below, the file `test.R` was created.

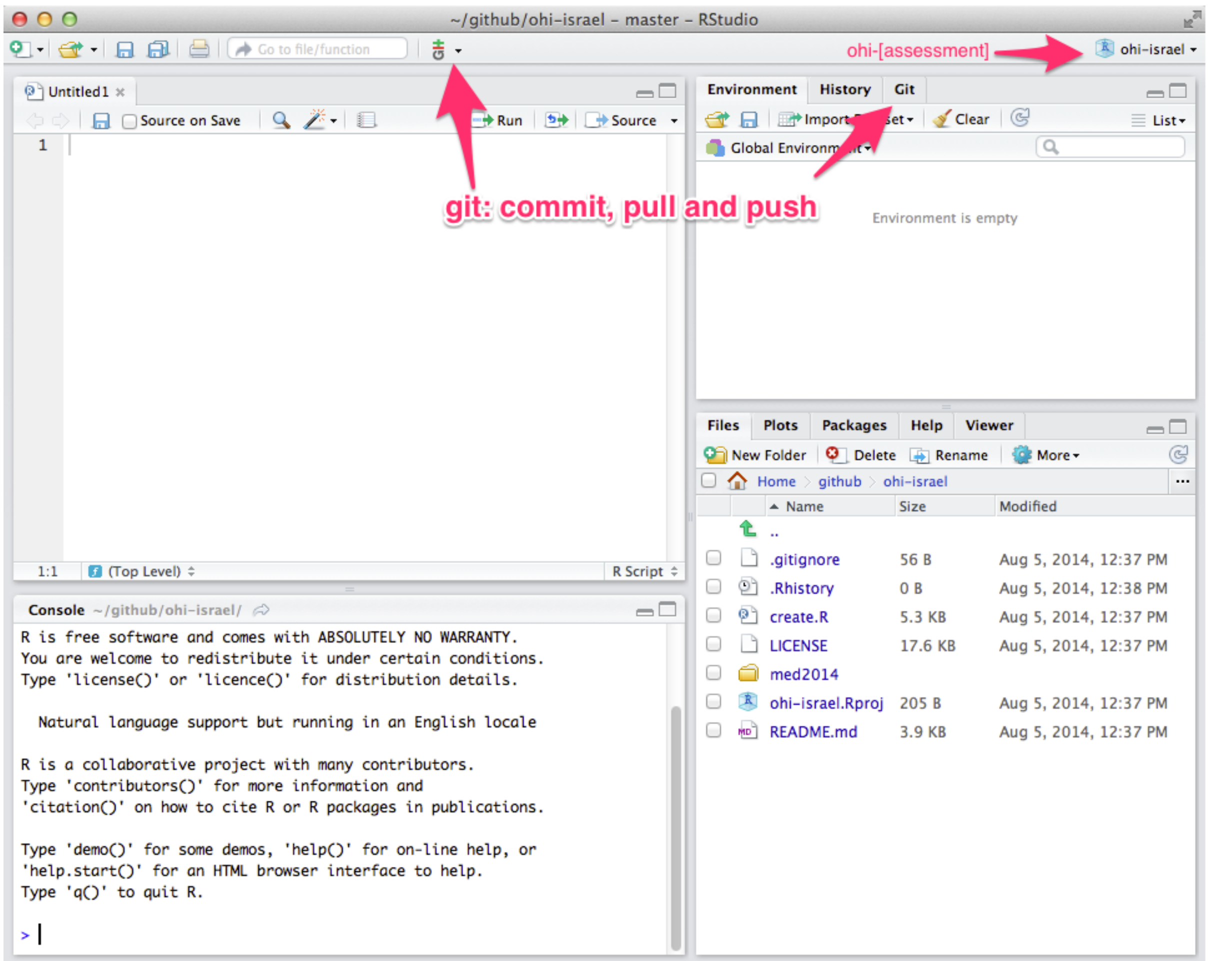
1. Clicking the 'Staged' box and the 'Commit' button opens a new window where you can review changes.
2. Type a commit message that is informative to the changes you've made.

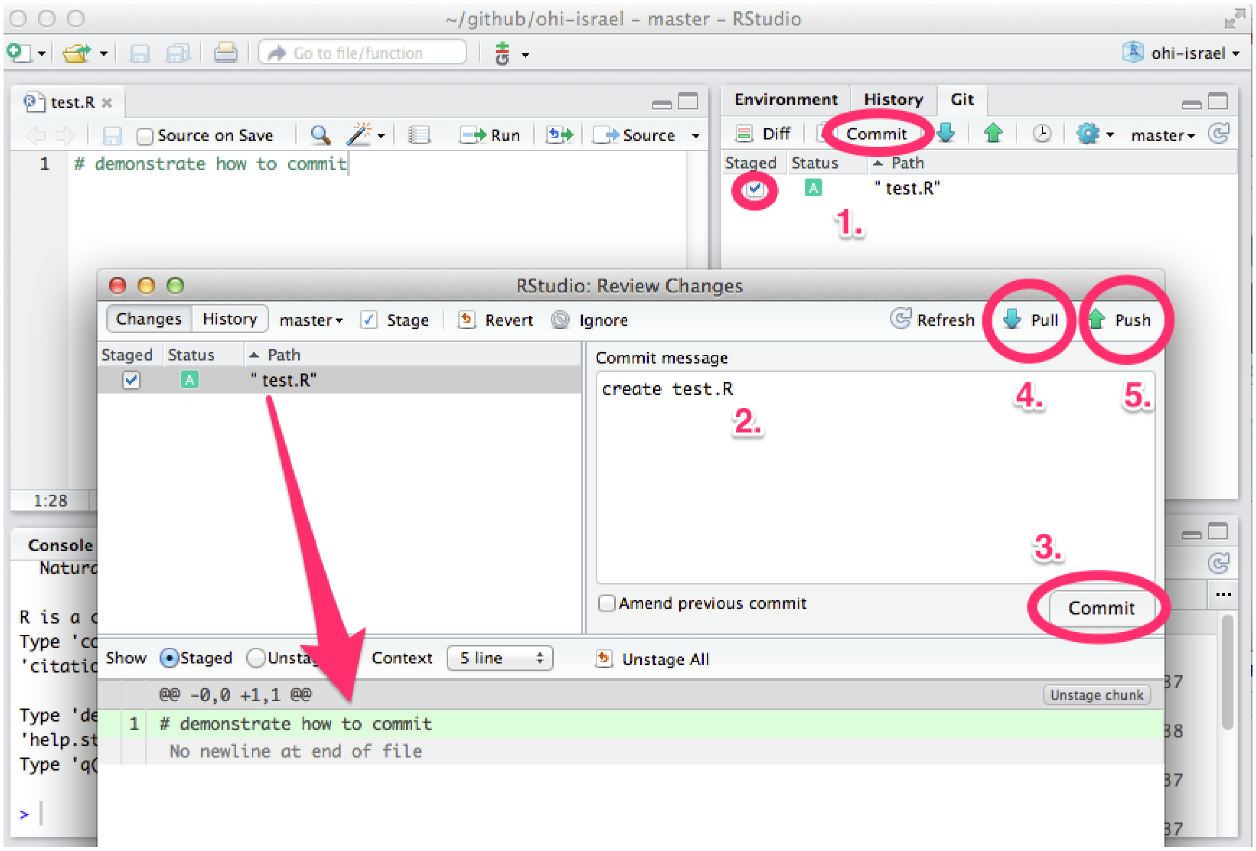
- Note 1: there will often be multiple files 'staged' at the same time, and so the same commit message will be associated with all of the updated files. It is best to commit changes often with informative commit messages.
- Note 2: clicking on a staged file will identify additions and deletions within that file for your review

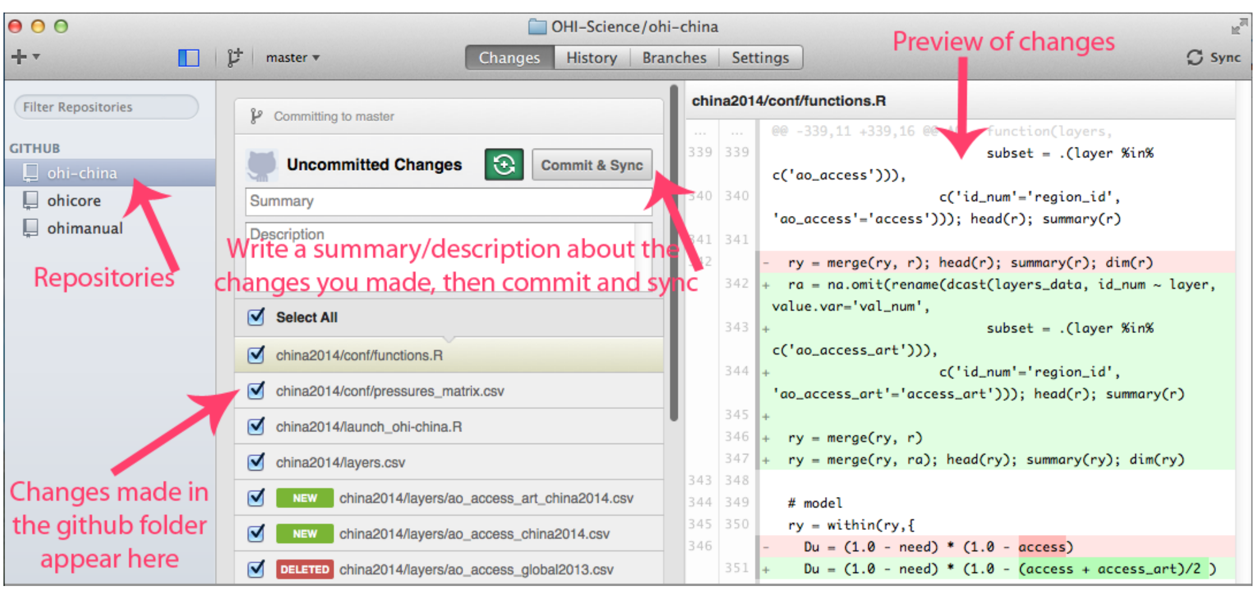
3. Click 'Commit' to commit the changes and the commit message
4. Pull any changes that have been made to the online repo. This is important to ensure there are no conflicts with updating the online repo.
5. Push your committed changes to the online server at `github.com/OHI-Science/ohi-[assessment]`. Your changes are now visable online.

6.3.9 Using Github App to push commits to GitHub ohi-[assessment] repository

The Github App can also be used to commit and sync any changes made locally in the github folder. To do so, write a summary and description in the respective message bars in the Github App window, then click on commit and sync.


~/github/ohi-israel – master – RStudio
ohi-[assessment]
ohi-israel
git: commit, pull and push
Untitled1
Source on Save
Run
Source
Environment
History
Git
Import Dataset
Clear
List
Global Environment
Environment is empty
1:1
(Top Level)
R Script
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
github
ohi-israel
Name
Size
Modified
.gitignore 56 B Aug 5, 2014, 12:37 PM
.Rhistory 0 B Aug 5, 2014, 12:38 PM
create.R 5.3 KB Aug 5, 2014, 12:37 PM
LICENSE 17.6 KB Aug 5, 2014, 12:37 PM
med2014
ohi-israel.Rproj 205 B Aug 5, 2014, 12:37 PM
README.md 3.9 KB Aug 5, 2014, 12:37 PM
Console ~/github/ohi-israel/
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.


~/github/ohi-israel – master – RStudio
test.R
Source on Save
Run
Source
demonstrate how to commit
Environment
History
Git
Diff
Commit
master
Staged
Status
Path
"test.R"
1.
RStudio: Review Changes
Changes
History
Stage
Revert
Ignore
Refresh
Pull
Push
Commit message
create test.R
2.
4.
5.
3.
Amend previous commit
Commit
Show
Staged
Context
5 line
Unstage All
@@ -0,0 +1,1 @@
Unstage chunk
demonstrate how to commit
No newline at end of file
Console


OHI-Science/ohi-china
master
Changes
History
Branches
Settings
Preview of changes
Sync
Filter Repositories
GITHUB
ohi-china
ohicore
ohimanual
Repositories
Committing to master
Uncommitted Changes
Commit & Sync
Summary
Description
Write a summary/description about the changes you made, then commit and sync
Select All
china2014/conf/functions.R
china2014/conf/pressures_matrix.csv
china2014/launch_ohi-china.R
china2014/layers.csv
NEW
china2014/layers/ao_access_art_china2014.csv
NEW
china2014/layers/ao_access_china2014.csv
DELETED
china2014/layers/ao_access_global2013.csv
Changes made in the github folder appear here
china2014/conf/functions.R

6.3.10 Keeping syncronized

Always pull before pushing commits. This is the best way to avoid merging errors due to other users in your group modifying files.

6.3.10.1 More Information

- presentation: Reproducible science with the Ocean Health Index
- wiki: Using GitHub

6.4 Accessing repositories without GitHub

6.4.1 OHI regional assessments without GitHub

It is possible to conduct a regional assessment without having a GitHub account. You can do a one-time download from the GitHub repository and then work locally on your computer without syncing to the GitHub server. We do not recommend this because it is more difficult for you to track progress and decisions, and for us to help or advise you. We recommend accessing a repository with GitHub but provide instructions here of how to access a repository without Github. We can always get your filesystem onto GitHub later.

6.4.2 Install the latest version of R

Download **R** and install this on your computer. If you already have R installed, check the website for updates. There are frequent updates to the R software, and the current version is identified on the websited. Compare what is available from their wesite with what you already have on your computer by typing `sessionInfo()` into your R console. (This will also identify packages you have installed).

While not required, we highly recommend working with **RStudio**, which is an interface that makes working with R much easier. RStudio also has great compatibility with GitHub so you can collaborate with your team directly from your R code. RStudio does not get updated as often as R does, but it is good to check for updates regularly.

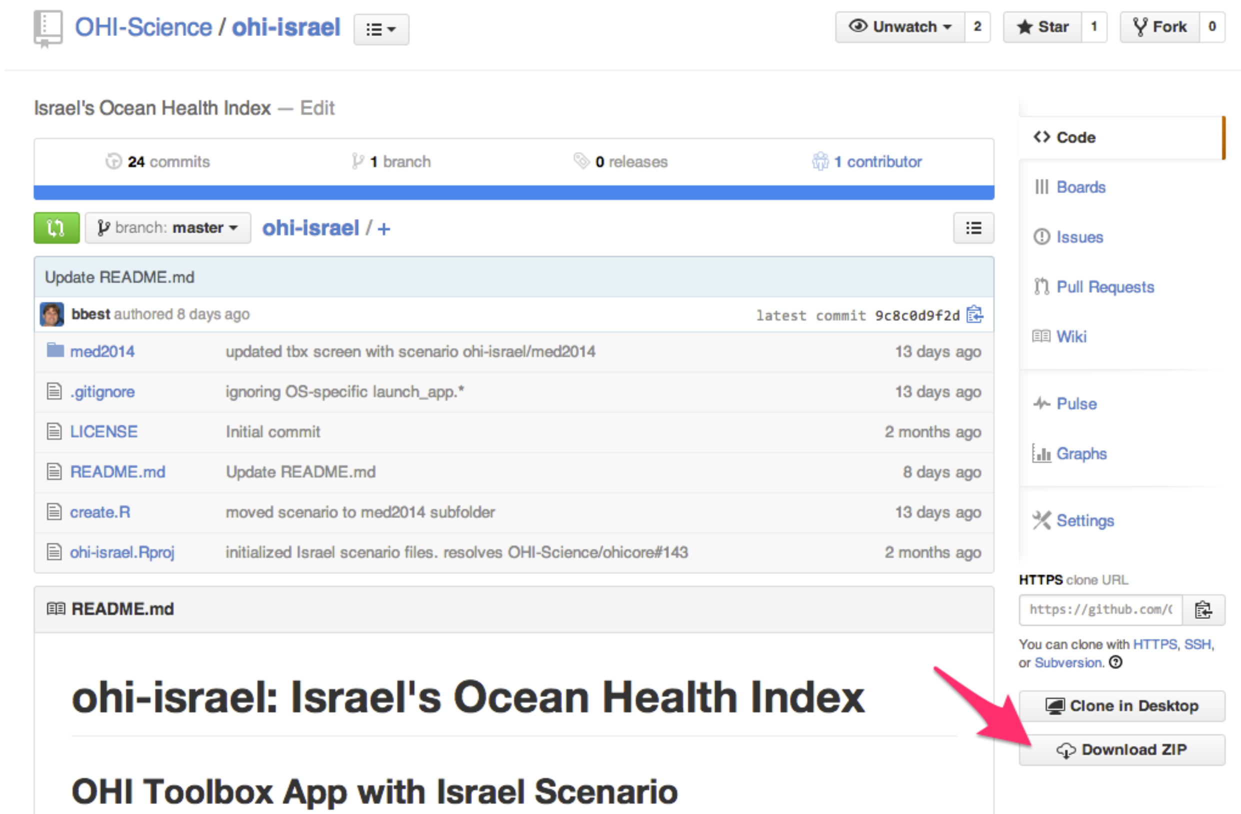
6.4.3 Download the repo to your computer

Contact the OHI team (bbest@nceas.ucsb.edu or lowndes@nceas.ucsb.edu) to create a repository for your group. The repository will be stored on github.com and called `OHI-Science/ohi-[assessment]`: for example, `github.com/OHI-Science/ohi-israel`. When your repository has been created, enter the url in to an internet browser (for example, `github.com/OHI-Science/ohi-israel`). Click `Download ZIP`.

6.4.4 Working locally

Unzip the downloaded folder `ohi-[assessment]-master` and save the folder on your computer. We recommend creating a folder called `github` in your root directory and saving your repository to this file path: `~/github/ohi-[assessment]`. This will make collaborating much simpler since everyone will work with the same file path.

You will work on your computer to change the files in `~/github/ohi-[assessment]` to reflect the desired modifications your team has identified for your regional assessment.



6.5 OHI Toolbox File System

This section describes the files within the ohi-[assessment] folder that you have accessed by either cloning through GitHub and RStudio or downloading to your computer from GitHub.

6.5.1 Assessments and scenarios

Within the **ohi-**[**assessment**] folder is the **scenario** folder. The scenario folder contains all the data, functions and other files required to calculate the Ocean Health Index. To calculate the Index for a different region or with new data or models, you will modify the files within this folder (default data is from the global assessment).

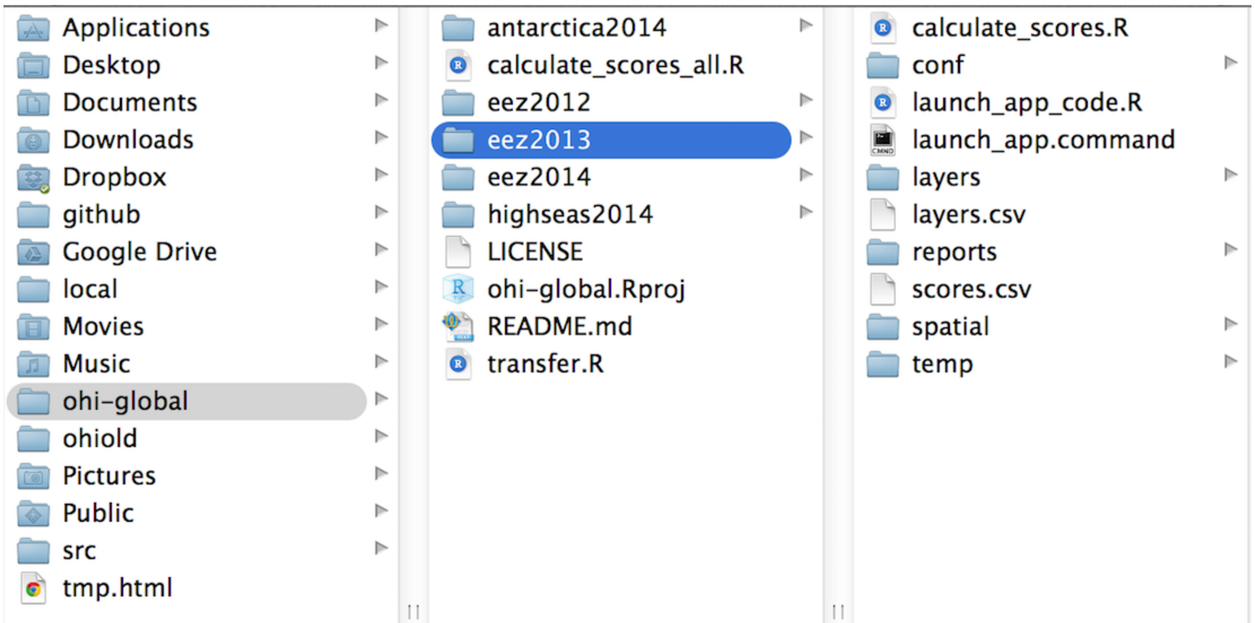
In this example, **ohi-global** is the assessment folder and **eez2013** is the scenario.

6.5.2 *layers.csv*

`layers.csv` is the registry that manages all data to be used in the Toolbox.

Each row of information represents a specific data layer that has been prepared and formatted properly for the Toolbox. The first columns contain information inputted by the user; other columns are generated later by the Toolbox App as it confirms data formatting and content. The first columns have the following information:

- **targets** indicates how the data layer related goals or dimensions. Goals are indicated with two-letter codes and sub-goals are indicated with three-letter codes, with pressures, resilience, and spatial layers indicated separately.

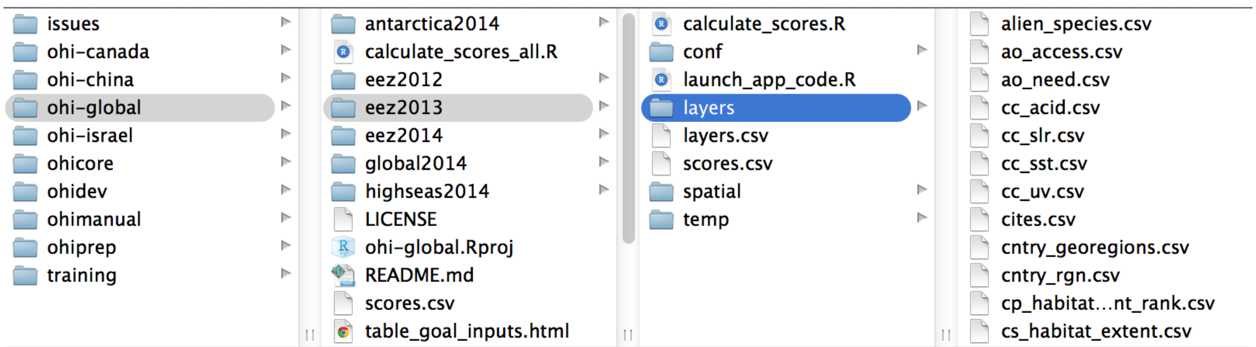


1	targets	layer	name	description	fld_value	units	filename
2	AO	ao_access	Fisheries manage	The opportunity fo	value	value	ao_access.csv
3	AO	ao_need	Purchasing powe	The per capita pu	value	value	ao_need.csv
4	CW	cw_coastalpopn_tren	Coastal human p	Coastal populatio	trend	trend score	cw_coastalpopn_tr
5	CW	cw_fertilizer_trend	Fertilizer consum	Statistics on fertili	trend.score	trend score	cw_fertilizer_trend.
6	CW	cw_pathogen_trend	Trends in access	Trends in percent	trend	trend score	cw_pathogen_trend
7	CW	cw_pesticide_trend	Pesticide consum	Statistics on pesti	trend.score	trend score	cw_pesticide_trend
8	FIS	fis_b_bmsy	B/Bmsy estimates obtained using th		b_bmsy	B / B_msy	fis_b_bmsy.csv
9	FIS	fis_meancatch	Catch data for ea	Reported data inc	mean_catch	metric tons	fis_meancatch.csv

- **layer** is the identifying name of the data layer, which will be used in R scripts like `functions.R` and .csv files like `pressures_matrix.csv` and `resilience_matrix.csv`. This is also displayed on the Toolbox App under the drop-down menu when the variable type is 'input layer'.
- **name** is a longer title of the data layer; this is displayed on the Toolbox App under the drop-down menu when the variable type is 'input layer'.
- **description** is further description of the data layer; this is also displayed on the Toolbox App under the drop-down menu when the variable type is 'input layer'.
- **fld_value** indicates the units along with the units column.
- **units** some clarification about the unit of measure in which the data are reported
- **filename** is the .csv filename that holds the data layer information, and is located in the folder 'layers'.

6.5.3 *layers* folder

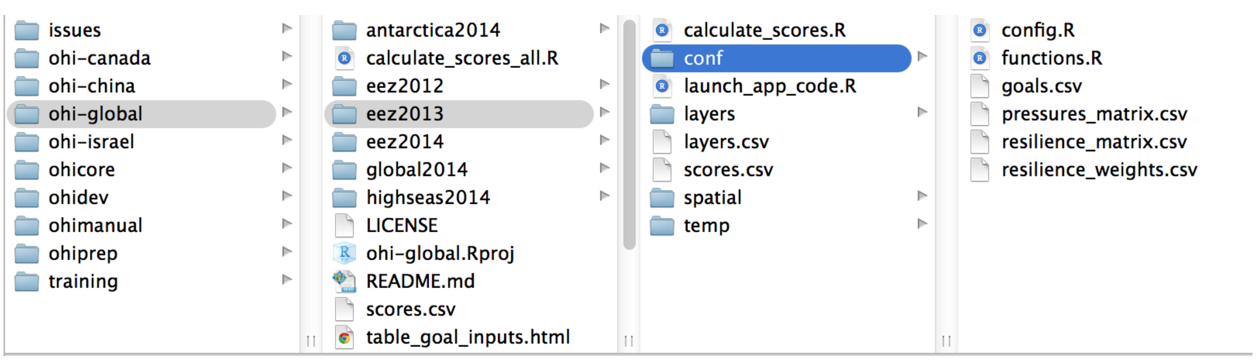
The `layers` folder contains every data layer as an individual .csv file. The names of the .csv files within the layers folder correspond to those listed in the *filename* column of the `layers.csv` file described above. All *.csv* files can be read with text editors or with Microsoft Excel or similar software.



Note that each .csv file within the `layers` folder has a specific format that the Toolbox expects and requires. Comma separated value files (*.csv* files) can be opened with text editor software, or will open by default by Microsoft Excel or similar software. Open a `layers/*.csv` file: note the unique region identifier (*rgn_id*) with a single associated *score* or *value*, and that the data are presented in 'long format' with minimal columns. Please see this tutorial for further details and instructions on data formatting requirements.

6.5.4 *conf* folder

The `conf` folder includes includes R functions (*config.R* and *functions.R*) and .csv files containing information that will be accessed by the R functions (*goals.csv*, *pressures_matrix.R*, *resilience_matrix.csv*, and *resilience_weights.csv*).



6.5.4.1 *config.r*

`config.r` is an R script that configures labeling and constants appropriately.

6.5.4.2 *functions.r*

`functions.r` contains functions for each goal and sub-goal model, which calculate the status and trend using data layers identified as 'layers' in `layers.csv`.

6.5.4.3 *goals.csv*

`goals.csv` is a list of goals and sub-goals and their weights used to calculate the final score for each goal. Other information includes the goal description that is also presented in the Toolbox App. `goals.csv` also indicates the arguments passed to functions.R. These are indicated by two columns: *preindex_function* (functions for all goals that do not have sub-goals, and functions for all sub-goals) and *postindex_function* (functions for goals with sub-goals).

6.5.4.4 *pressures_matrix.csv*

`pressures_matrix.csv` describes the layers ('layers' column in layers.csv) needed to calculate pressure categories. The matrix has weights assigned that were determined by Halpern et al. 2012 (Nature) based on scientific literature and expert opinion.

6.5.4.5 *resilience_matrix.csv*

`resilience_matrix.csv` describes the layers ('layers' column in layers.csv) needed to calculate resilience categories.

6.5.4.6 *resilience_weights.csv*

`resilience_weights.csv` describes the weight of various resilience layers, were determined by Halpern et al. 2012 (Nature) based on scientific literature and expert opinion.

6.5.5 spatial folder

The spatial folder contains a single file, *regions_gcs.js*. This is a spatial file in the GeoJSON format; it has the appropriate study area and regions for the assessment. This file will be created by the OHI team for all regional assessments.

6.5.6 *calculate_scores.r*

`calculate_scores.r` will run the Toolbox calculations using the .csv files in the *layers* folder that are registered in *layers.csv* and the configurations identified in *config.r*. Scores will be saved in *scores.csv*.

6.5.7 scores.csv

`scores.csv` is a record of the calculated scores for the assessment (Global 2013 scores). Scores are reported for each dimension (future, pressures, resilience, score, status, trend) for each reporting region, and are presented in 'long' format.

6.5.8 Relaunching the Toolbox

After the initial Toolbox setup, further launches of the Toolbox Application can be done without the software program R. Instead, PC users can double-click the `launchApp.bat` file and Mac users can double-click the `launchApp.command` file.

7 Using the Toolbox for a Regional Assessment

** **Note: this page is under development**

This page explains how to incorporate all of the pre-Toolbox decisions your team has made for your regional assessment into the OHI framework and your repository. Having a good understanding of how the Toolbox is structured can also help identify what must be modified for a regional assessment, particularly with data and models. This page assumes a good understanding of the Toolbox file system.

The most common modifications you will make to your repository are changes with:

- updating or adding new data layers
- modifying goal models
- removing goals

7.1 Modifying and creating data layers

To modify existing or create new data layers, data must be appropriately formatted.

There are several steps to follow when working with data layers:

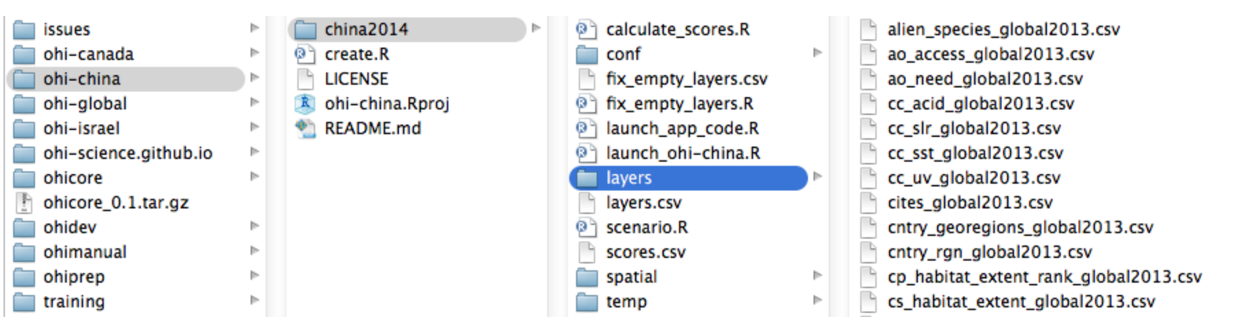
1. Create data layer with proper formatting
2. Save the layer in the `layers` folder
3. Register the layer in `layers.csv`
4. Check (and update when appropriate) `pressures_matrix.csv` and `resilience_matrix.csv` (located in the `ohi-[assessment]/[scenario]/conf` folder)

7.1.1 Create data layers with proper formatting

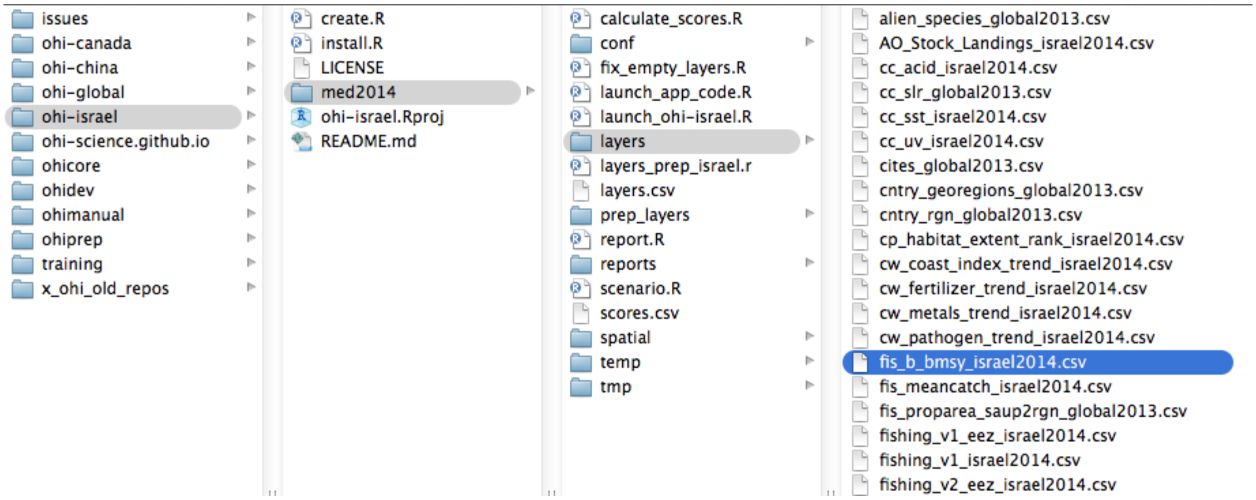
develop

7.1.2 Save data layers in the *layers* folder

Data layers are *.csv* files and are located in the `ohi-[assessment]/[scenario]/layers` folder. The layers provided in your regional assessment repo are the global values from the 2013 assessment: these layers all have a suffix of `_global2013.csv`. These data are at coarse-resolution and should be exhanged for local, high-resolution data when possible.

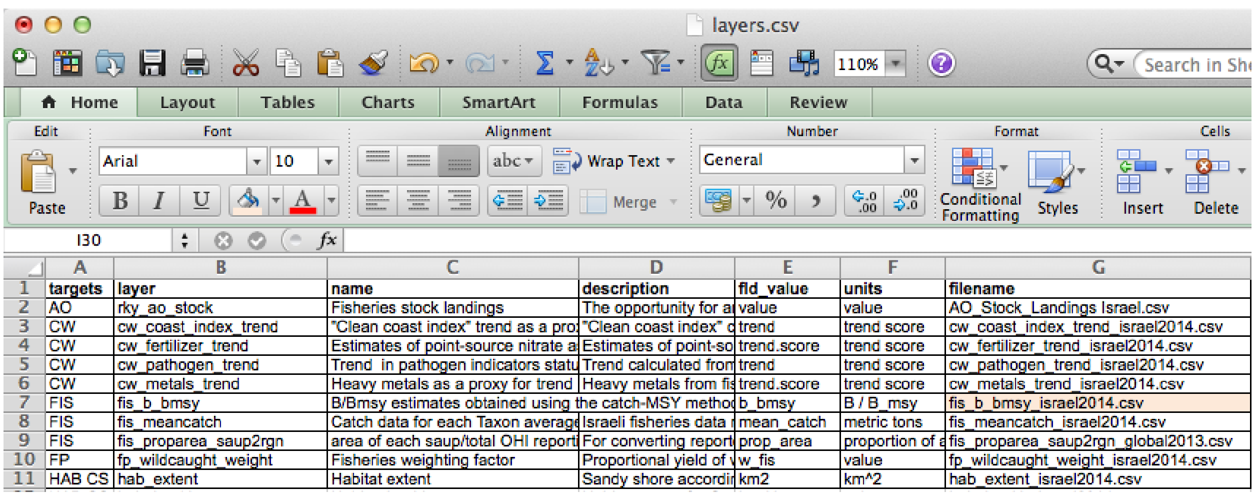


When you modify existing or create new data layers, we recommend saving this as a new *.csv* file with a suffix identifying your regional assessment (example: `_israel2014.csv`). Modifying the layer name provides an easy way to track which data layers have been updated regionally, and which rely on global data.



7.1.3 Register data layers in *layers.csv*

When there are new filenames associated with each layer, they will need to be registered in `ohi-[assessment]/[scenario]/layers.csv`. If a layer simply has a new filename, only the *filename* column needs to be updated:



However, if a new layer has been added (for example when a new goal model is developed), you will need to add a new row in the registry for the new data layer and fill in the first eight columns (columns A-H); other columns are generated later by the Toolbox App as it confirms data formatting and content:

- **targets:** Add the the goal/dimension that the new data layer relates to. Goals are indicated with two-letter codes and sub-goals are indicated with three-letter codes, with pressures, resilience, and spatial layers indicated separately.
- **layer:** Add an identifying name for the new data layer, which will be used in R scripts like functions.R and .csv files like pressures_matrix.csv and resilience_matrix.csv.
- **name:** Add a longer title for the data layer: this will be displayed in the Toolbox interface.
- **description:** Add a longer description of the new data layer this will be displayed in the Toolbox interface.
- **fld_value:** Add the appropriate units for the new data layer (which will be referenced in subsequent calculations).

- **units:** Add a description about the 'units' chosen in the 'fld_value' column above.
- **filename:** Add a filename for the new data layer that matches the name of the csv file that was created previously in the 'layers' folder.
- **fld_id_num:** Area designation that applies to the newly created data layer, such as: *rgn_id* and *fao_id*.

7.1.4 Check pressures and resilience matrices

under development

7.2 Modifying goal models

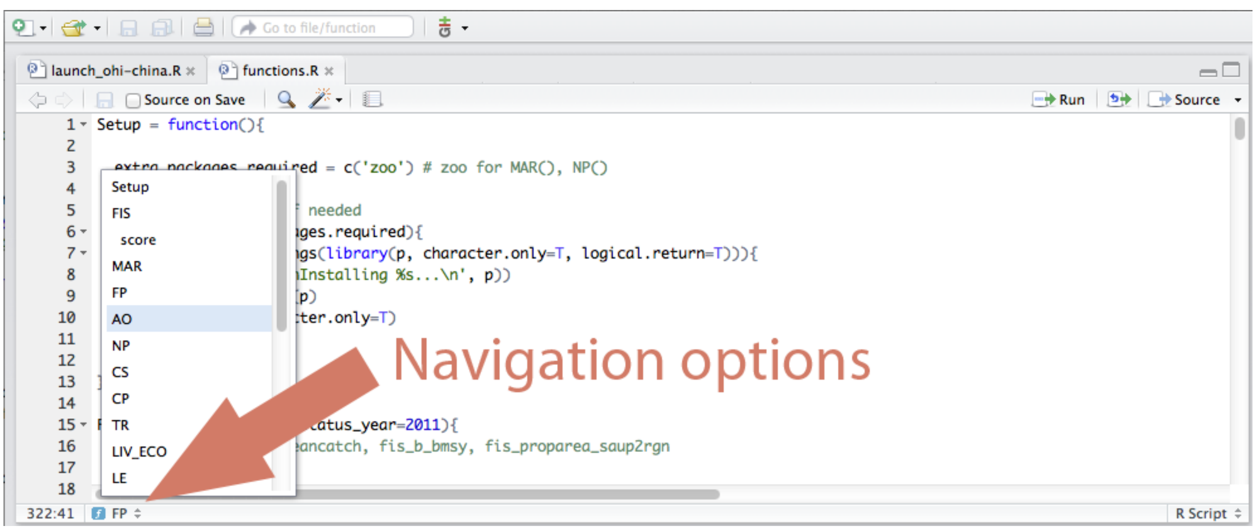
In the discussion on data layers above, when an existing layer is still used as before but has a new *filename*, nothing further needs to be done for the Toolbox to incorporate this updated layer. However, if a new layer has been added to the `layers` folder and registered in `layers.csv` (and potentially added to the pressures or resilience matrices), the Toolbox will still not use it unless it is incorporated into a goal model.

There are several steps to follow when working with goal models:

1. Update `functions.r`
2. Check and possibly update `goals.csv`

7.2.1 Update *functions.r*

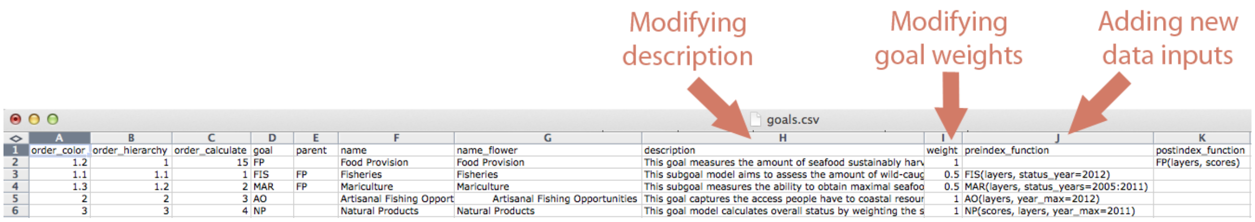
To incorporate a new data layer into a goal model, open `functions.R`: this script contains all the models for each goal and sub-goal. In RStudio, there is a navigation pane that can be used to navigate between them:



7.2.2 Check and possibly update *goals.csv*

goals.csv provides input information for *functions.r*, particularly about goal weighting and function calls. It also includes descriptions about goals and sub-goals, which is presented in the Toolbox Application.

Changing goal weights will be done here by editing the value in the *weight* column. Weights do not need to be 0-1 or add up to 10; weights will be scaled as a percentage of the goal totals. `goals.csv` also indicates the arguments passed to `functions.r`. These are indicated by two columns: *preindex_function* (functions for all goals that do not have sub-goals, and functions for all sub-goals) and *postindex_function* (functions for goals with sub-goals).



When updating layers or goal models, it is important to ensure that information called from *goals.csv* is correct:

- check the years
- etc...

7.3 Removing goals

If a goal is not relevant in your region, it is possible to remove the goal completely from the calculation. There are four places where you will need to remove the reference to this goal:

1. `functions.r`
2. `goals.csv`
3. `pressures_matrix.csv`
4. `resilience_matrix.csv`

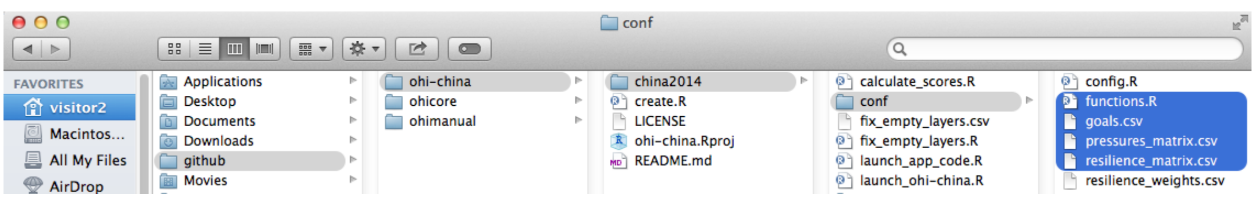


Figure 2: *Failing to delete all referenced layers after the goal is deleted will prompt a number of error messages.*

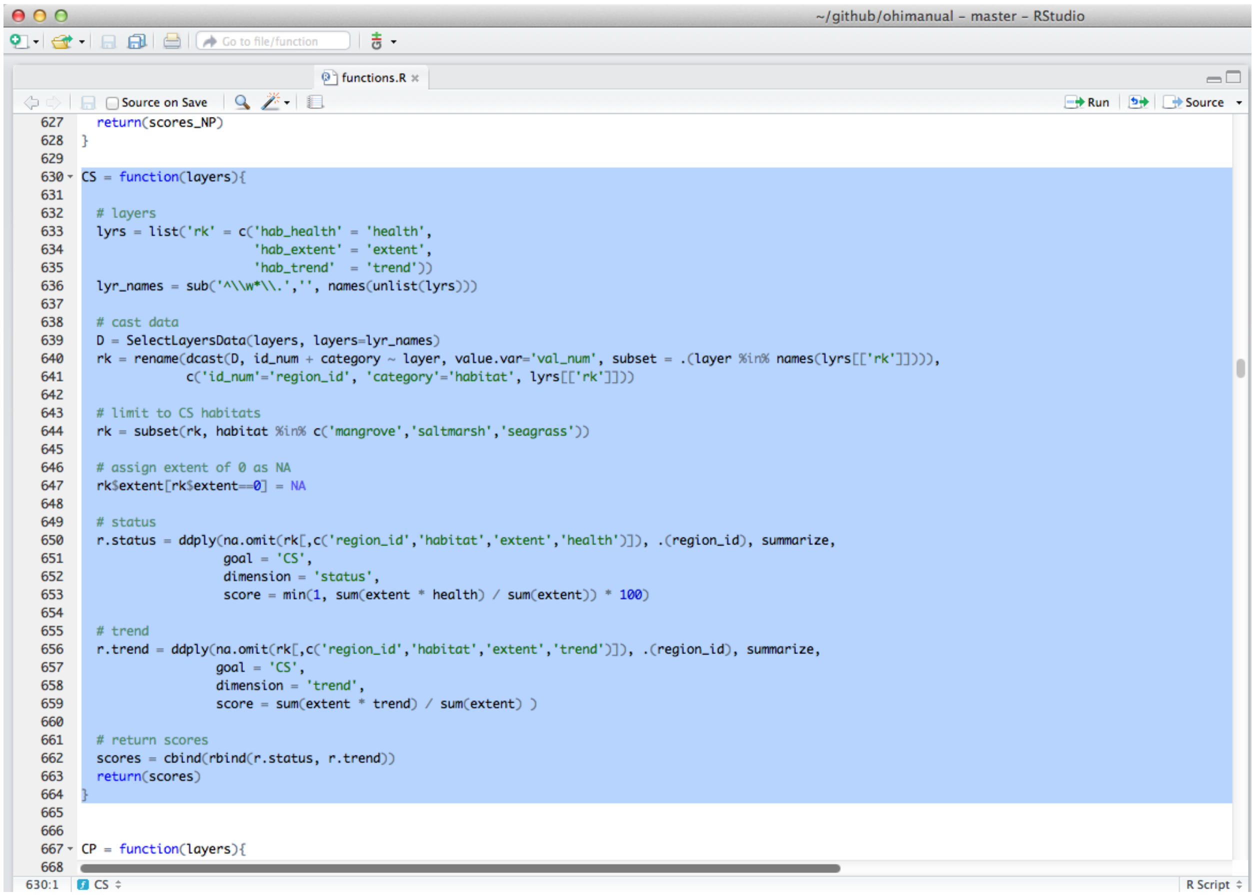
7.3.1 Example: Removing 'Carbon Storage' goal

1) Remove the CS goal model from `functions.r`:

2) Remove the CS row from `goals.csv`:

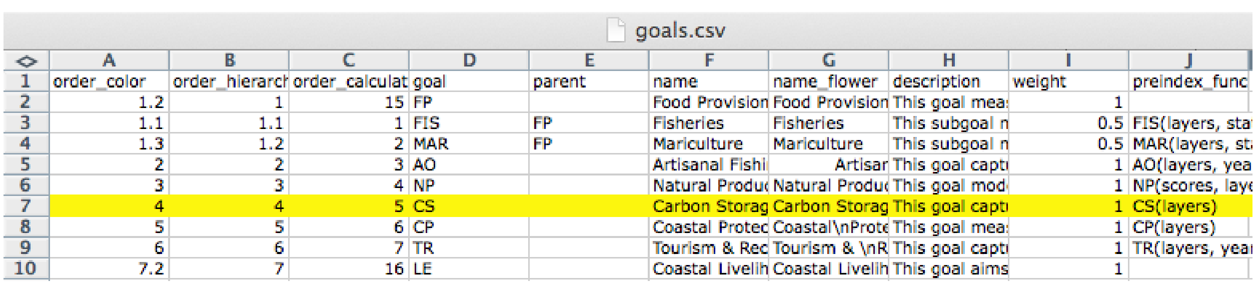
3) Remove all CS rows from `pressures_matrix.csv`:

4) Remove all CS rows from `resilience_matrix.csv`:



```r
  return(scores_NP)
}

CS = function(layers){

  # layers
  lyrs = list('rk' = c('hab_health' = 'health',
                       'hab_extent' = 'extent',
                       'hab_trend'  = 'trend'))
  lyr_names = sub('^\\w*\\.','', names(unlist(lyrs)))

  # cast data
  D = SelectLayersData(layers, layers=lyr_names)
  rk = rename(dcast(D, id_num + category ~ layer, value.var='val_num', subset = .(layer %in% names(lyrs[['rk']]))),
              c('id_num'='region_id', 'category'='habitat', lyrs[['rk']]))

  # limit to CS habitats
  rk = subset(rk, habitat %in% c('mangrove','saltmarsh','seagrass'))

  # assign extent of 0 as NA
  rk$extent[rk$extent==0] = NA

  # status
  r.status = ddply(na.omit(rk[,c('region_id','habitat','extent','health')]), .(region_id), summarize,
                   goal = 'CS',
                   dimension = 'status',
                   score = min(1, sum(extent * health) / sum(extent)) * 100)

  # trend
  r.trend = ddply(na.omit(rk[,c('region_id','habitat','extent','trend')]), .(region_id), summarize,
                  goal = 'CS',
                  dimension = 'trend',
                  score = sum(extent * trend) / sum(extent) )

  # return scores
  scores = cbind(rbind(r.status, r.trend))
  return(scores)
}


CP = function(layers){
```

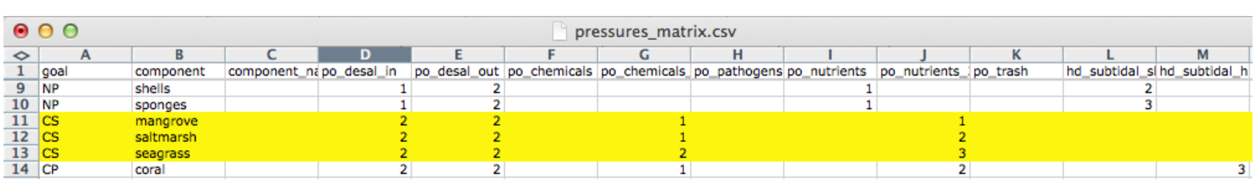
Figure 3: Delete the highlighted text that references the CS layers and calculates CS goal status, trend, and scores



goals.csv

	A	B	C	D	E	F	G	H	I	J
1	order_color	order_hierarch	order_calculat	goal	parent	name	name_flower	description	weight	preindex_func
2	1.2	1	15	FP		Food Provision	Food Provision	This goal mea:	1	
3	1.1	1.1	1	FIS	FP	Fisheries	Fisheries	This subgoal n	0.5	FIS(layers, sta
4	1.3	1.2	2	MAR	FP	Mariculture	Mariculture	This subgoal n	0.5	MAR(layers, st
5	2	2	3	AO		Artisanal Fishi	Artisar	This goal captu	1	AO(layers, yea
6	3	3	4	NP		Natural Produc	Natural Produc	This goal mod	1	NP(scores, laye
7	4	4	5	CS		Carbon Storag	Carbon Storag	This goal captu	1	CS(layers)
8	5	5	6	CP		Coastal Protec	Coastal\nProte	This goal mea:	1	CP(layers)
9	6	6	7	TR		Tourism & Rec	Tourism & \nR	This goal captu	1	TR(layers, yea
10	7.2	7	16	LE		Coastal Livelih	Coastal Livelih	This goal aims	1	

Figure 4: Delete the highlighted row that contains the CS goal



pressures_matrix.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	goal	component	component_na	po_desal_in	po_desal_out	po_chemicals	po_chemicals_	po_pathogens	po_nutrients	po_nutrients_	po_trash	hd_subtidal_sl	hd_subtidal_h
9	NP	shells		1	2				1			2	
10	NP	sponges		1	2				1			3	
11	CS	mangrove		2	2		1			1			
12	CS	saltmarsh		2	2		1			2			
13	CS	seagrass		2	2		2			3			
14	CP	coral		2	2		1			2			3

Figure 5: Delete the highlighted rows that contain CS pressures

resilience_matrix.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	goal	component	alien_species	cites	fishing_v1	fishing_v1_ee:	fishing_v2_ee:	fishing_v3	fishing_v3_ee:	habitat	habitat_combo	habitat_combo	li_gci
7	SPP		alien_species	cites			fishing_v2_eez					habitat_combo_eez	
8	CS										habitat_combo		
9	CS	mangrove only									habitat_combo		
10	CW												
11	FIS						fishing_v2_eez					habitat_combo_eez	

Figure 6: Delete the highlighted rows that contain CS resilience

7.4 Example modifications

7.4.1 Adding a new layer to a goal model

In this example we will walk through the following steps:

1. decide to add artisanal access component to the model because of locally available data
2. prepare the data file; save layer ao_access_art
3. register in `layers.csv`
4. update goal model in `functions.r`
5. update goal call in `goals.csv`

1) and 2) is done outside of the Toolbox

3. register in `layers.csv`

Sheets | Charts | SmartArt Graphics | WordArt

	A	B	C	D	E	F	G	H
1	targets	layer	name	description	fld_value	units	filename	fld_id_num
2	AO	ao_access	Fisheries manageme	The opportunity f	value	value	ao_access_china2014.csv	rgn_id
3	AO	ao_access_art	Example data	Made-up data	value	value	ao_access_art_china2014.csv	rgn_id
4	AO	ao_need	Purchasing power pa	The per capita pu	value	value	ao_need_global2013.csv	rgn_id
5	CW	cw_coastalpopn_tre	Coastal human popu	Coastal populatio	trend	trend score	cw_coastalpopn_trend_global2013.csv	rgn_id
6	CW	cw_fertilizer_trend	Fertilizer consumpti	Statistics on fertil	trend score	trend score	cw_fertilizer_trend_global2013.csv	rgn_id

4. update goal model

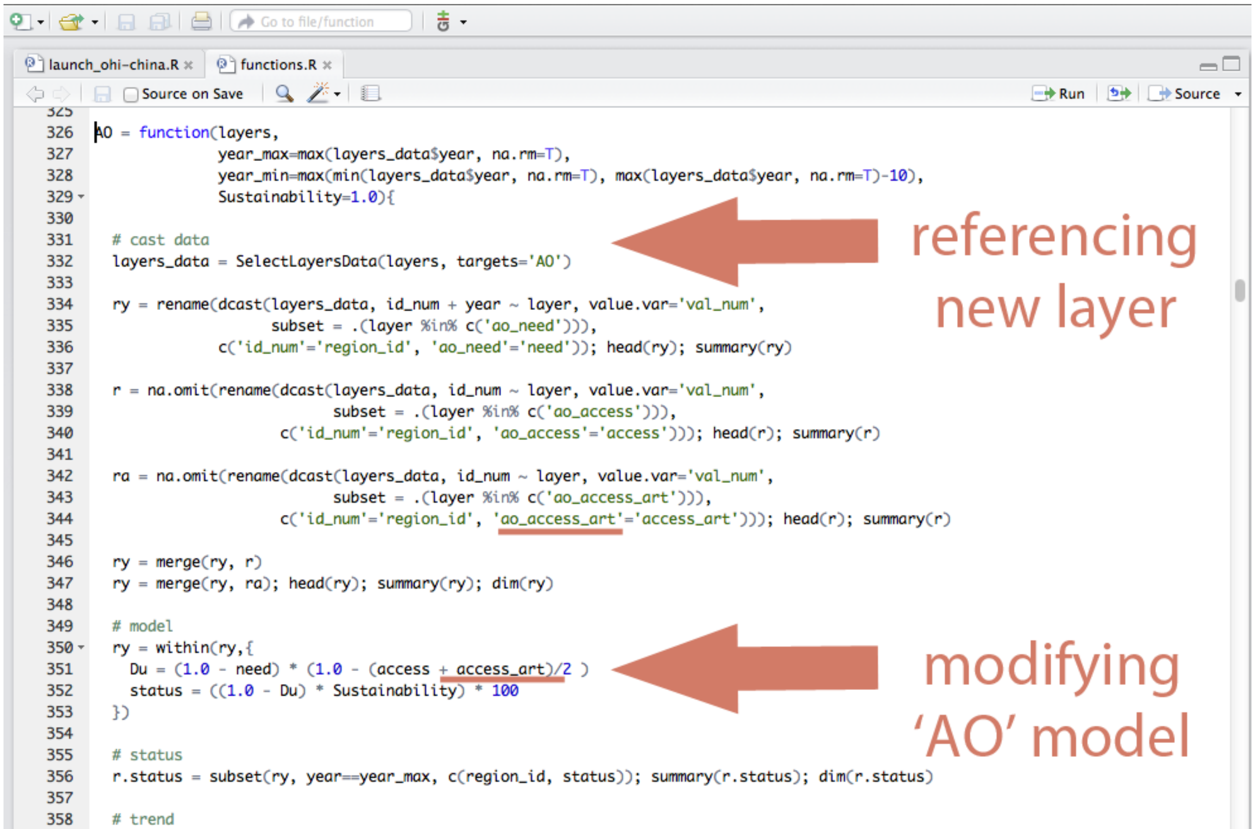
5. develop

7.5 Notes about R

The Toolbox is written in R, and relies heavily on a few R packages created to faciliate data handling and manipulation. The primary R package used is called `dplyr` by Hadley Wickham. The `dplyr` package allows for 'chaining' between functions, which is represented with a `%>%`. See for documentation.

8 Frequently Asked Questions (FAQs)

This document provides answers to some frequently asked questions about conducting regional assessments using the Ocean Health Index. A few questions are related to general concepts in the Ocean Health Index, but mostly those topics are covered at http://www.oceanhealthindex.org/About/FAQ/. Here, the FAQ are primarily technical questions regarding regional assessments and using the OHI Toolbox. This document will be updated continually as we have more questions. Questions are arranged by theme, and have the format Q: (question) and A: (answer).



8.1 Overall

8.2 Conceptual

Q: Are regional assessment scores comparable with global assessment scores?

A: Regional Index scores cannot be directly compared to global Index scores, or to other regional Index scores calculated through separate efforts. This is because data and indicators (both what they measure and their quality), reference points (set using local knowledge and priorities), and specific goal models are often different for the areas being compared.

However, because scores for each goal are scaled to a reference point, qualitative comparisons can be made. For example, a score of 71 in the US West Coast compared to 66 in Brazil says that the US West coast is closer to fully meeting its sustainable goals (i.e., meeting regional reference points). Furthermore, use of the same Ocean Health Index framework across regional assessments permits fruitful discussion and general comparisons even if data inputs differ. Ocean Health Index assessments at any scale always work within a standardized definition of ocean health, using information to capture the philosophy of the ten goals that have been identified (and undergone scientific peer-review) prior to compiling relevant data. Use of the ten-goal framework is important both to ensure that all aspects of ocean health are captured and to allow better comparison across regional assessments than would be possible if the different regions used different methods.

Q: Where is climate change measured in the Index?

A: Four different aspects of climate change – increases in sea surface temperature (SST), sea level rise (SLR), ultraviolet radiation (UV), and ocean acidification (OA) – are included as pressures to many goals in the Index, including Natural Products, Carbon Storage, Coastal Protection, Sense of Place, Livelihoods & Economies and Biodiversity. Mitigation of climate change through carbon storage is one of the ten goals.

Q: Why are food provision and artisanal fishing opportunities goals separated?

A: These goals measure different aspects of how people relate to fishing. The catch of fish made by artisanal (=small-scale, subsistence type) fisheries is captured in the food provision goal. Jobs, wages and income from both the food provision and artisanal fishing goals are captured in the livelihoods & economies goal. The purpose of the artisanal fishing opportunity goal is to evaluate the opportunity for people to pursue this fishing in relation to their need to do so.

8.3 Timing and Resources

Q: How much does it cost to produce a regional assessment?

A: Regional assessments can be completed at(varying costs depending on the local context.(Funds are needed for a management and scientific team, workshops and meetings (including travel), communications, policy engagement, and operating costs. Therefore, securing funding is an important component to satisfactorily complete the assessment. We encourage the development of a local proposal or strategic action plan that details a timeline of activities and the resources needed to accomplish them.

Q: How many people are required in a team?

A: rather than a specific number of individuals, what is required are specific skillsets. For example, if the scientific analysts were capable of effectively conducting the R analysis, then a dedicated R analyst would not be required. In current assessments, teams range between 2 and 8 people.

Q: How long does it take to calculate OHI at a regional scale?

A: This depends on many things: budget and number of people involved, the scale of the study area and whether new regions will need to be created, how easily data can be acquired, how much local data can be incorporated, how many goal models need to be changed. Additionally, decisions about setting reference points require input from experts.

Q: How much time will modifications by an R analyst take?

A: This will depend on if you are changing any models, and potentially data layers–but a lot of changing data layers just requires registering them properly in layers.csv (and maybe pressures_matrix.csv and resilience_matrix.csv if they are pressures or resilience files) and having the functions.r file call those layers. That is more ‘bookkeeping’ than actual R programming.

Q: How much time will modifications by a GIS analyst take?

A: this will depend on how many layers you are processing: you are clipping spatial data? That will take some time because there are quite a few files, but maybe not too long since it is pretty small scale and once there is a clipping mask created I think you apply it to other files.

Q: Which goals require a GIS analyst?

A: All goals using spatial data could potentially require a GIS analyst. These goals are commonly: habitat-based goals and sub-goals: (Coastal Protection, Carbon Storage, Habitats—a sub-goal of Biodiversity), Food Provision, Sense of Place, Species—a sub-goal of Biodiversity, Clean Waters

8.4 Structure

Q: Can we remove or add goals to the OHI?

A: A lot of deliberation went into defining the ten goals, and they seem to do a pretty good job of covering many if not most ocean uses, so additional goals may not be necessary. But it could be that they eclipse or replace an existing goal.

8.5 Reference points

Q: Can planning targets can be used as the reference points?

A: Yes, planning targets can be used as reference points. This won't be appropriate for every goal, but there are cases where this seemed best (example: iconic species sub-goal in the global assessment, mariculture sub-goal in the US West Coast assessment).

Q: What is sector evenness?

A: Sector evenness (also called a diversity index) is an economic concept that is included in OHI to enable comparison across many different sectors included in the Livelihoods & Economies goal. This goal evaluates jobs, wages and revenues for nine marine employment sectors. The distribution of employment across these nine sectors is an effective indicator of resilience. If total employment within a community is primarily based in one or two sectors, the overall economic system will be excessively vulnerable to downturns in those sectors. Conversely, if employment is spread relatively evenly throughout all nine sectors, the overall system will be more robust and resistant to such disturbances. Overall revenue within the community will remain more stable during such downturns, and workers displaced by a downturn in their sector may be able to find employment in another sector without leaving the community.

8.6 Appropriate data layers

Q: Shipping and port activity are hardly affected by the health of the ecosystem. Why are these included in the Index?

A: Shipping and port activity are included as pressures only

Q: Can oil spills be included in OHI?

A: Yes, oil spills could be included as a pressure and in the Clean Waters goal.

Q: Is seasonal (non-permanent) sea ice included in OHI habitats?

A: No, sea ice only includes permanent sea ice.

Q: Can seaweeds be included in the Carbon Storage goal?

A: Because they store carbon for less than 100 years, seaweeds and corals are not included in the carbon storage goal. While the pelagic oceanic carbon sink (phytoplankton) plays a large role in the sequestration of anthropogenic carbon, the pelagic ocean mechanisms are not amenable to local or regional management intervention. Phytoplankton and contribute to carbon fixation when they die and sink to the sea bottom at sufficient depth, because it is effectively out of circulation. However, if those phytoplankton are eaten, the carbon is cycled back into the system and not sequestered. Something that could potentially be included in the carbon storage goal is mollusc shells, if they are added to a landfill and not recycled in the sea. So if information on mariculture production and waste disposal are available, this could be an interesting addition to carbon storage at a regional scale.

Q: Is coastal engineering included in Coastal Protection? What if it reduces erosion?

A: We did not include an assessment of the protection afforded by man-made structures, such as jetties and seawalls, because these structures cannot be preserved without maintenance, may have other negative side effects (e.g. alter sedimentation rates causing erosion in new locations), thus they do not constitute long-term sustainable services. Coastal engineering (jetties, harbours, marina and breakwater) is not natural, and is mostly seen as a pressure. It will also be evident in the status of due to decreased natural habitat. It gets tricky when structures are built to help reduce coastal erosion–they are still manmade and therefore not a natural benefit that the ocean provides. But if available data allow, it might be possible to include tradeoff effects: maybe in areas where natural habitats are degraded and man-made structures have been built to reduce erosion, we could reduce the pressure that would otherwise be applied.

Q: How is seawater used for cooling on-shore power plants incorporated into OHI?

A: The use of cooling water for on-shore power plants would be a pressure on the ocean, since it causes entrapment of fishes, larvae, etc, and usually is circulated back into the ocean at higher temperatures (and maybe other chemicals, minerals, etc). Since the energy is coming from land-based activities, there isn't a service that the ocean is providing that 'benefits' people, it is only a pressure from the OHI perspective.

Q: How is freshwater production through desalination incorporated into OHI?

A: Desal would be incorporated into OHI in several places. The benefit is that there is freshwater produced, which could be incorporated into the Natural Products goal (or potentially into its own goal). Data required would be the volume of freshwater created based on the volume of seawater involved and spatial extent. Setting the reference point would not be based on how much can be produced, but some other targets perhaps set by government (percentage of the population served). Similar to the mariculture sub-goal and tourism goals, any negative effects caused by desal that affect other goals (example: species) do not influence the ability to obtain desalination targets now and in the future. Therefore, the sustainability coefficient only measures the ability to sustain that goal, but not the impacts on other goals: instead, they are taken into account as pressures when calculating the other goals. Desal should be included as a pressure similar to cooling on-shore power plants since the discharge brine is dense, doesn't plume very well and there are chemicals involved.

8.7 Food Provision

Q: Could the culture of marine fish in closed pools on-shore be included in the Mariculture sub-goal?

A: This should not be included because onshore aquaculture does not require a marine environment.

Q: Can aquaculture farms that receive seawater supply and return seawater back to the sea be included in the food provision goal?

A: This would be more appropriately included in the Mariculture sub-goal, and with finer-scale data additional pressures due to the intake pipes and the processed brine back into the marine system could be incorporated as well. Natural Products

Q: If natural products are all produced through on-land aquaculture, should this goal be removed?

A: In this case you would probably have good reason to exclude the natural product goal due if this was defendable through discussions with experts and any reports/papers on the topic. This would also depend on the origin of these natural products–are they from the region's waters? Habitat-based goals

Q: How is coral health calculated?

A: Coral health was estimated by compiling point data from multiple studies of percent live coral cover. In other words, estimates of coral cover within transects of certain sites were repeated in time and we used that rate of change in time as an indication of health of the reefs in the whole region. The difficulty lies in 1) having enough different locations sampled that you can say something about the whole region and 2) finding studies that did repeated measures in time, in the same location, over at least 20 years. In the Global 2013 assessment, there were so few datasets that satisfied this condition that we had to pool observations from different locations.

Q: Is it possible to calculate habitat goals when there is only one year of habitat data?

A: With only one year of habitat data, it is not possible to calculate the trend (which requires 5 years of data). Instead, it might be best to use the available habitat data to calculate the current status and then to overlay pressures for the last 5 years to calculate trend.

8.8 Livelihoods & Economies

Q: Benefits gained from Wild-caught fisheries, Mariculture, Tourism & Recreation are included in specific goals. Why are these counted again in Livelihoods & Economies?

A: The quantity of fish, mariculature, and participation in T&R are considered separately in goals whereas the monetary component is captured in L&E.

Q: Why are revenue data from shipping, boat building, ports and harbors included as revenue? Do these activities rely on a healthy ocean?

A: These sectors are included in the Ocean Health Index because the demand for some of those boats (fishing boats, sailboats, yachts) is dependent on a healthy ocean.

Q: Why isn't oil and gas industries included in revenue?

A: The Natural Products goal does not include non-living items such as oil, gas, and mining products, because these practices are not considered to be sustainable. They are also done at such large scales that including them would essentially make OHI an index for oil and mining–and they are not truly an ocean product. Because these products are not included in terms of quantity extracted, it did not seem appropriate to include information regarding jobs, wages or revenue.

8.9 Tourism & Recreation

Q: How do I calculate the sustainability term for TR?

A: The best way is to use a local indicator or measure of tourism sustainability or competitiveness, otherwise use the TTCI value from the Global 2013 assessment for the study area (applied evenly across all regions.

8.10 Natural Products

Q: Where do Natural Products come from?

A: In the global assessments, Natural Products data come from the UN's Food and Agriculture Administration (www.fao.org/fishery/statistics/software/fishstatj/en). These data are compiled and reported by product for each country, and available by downloading the FishStatJ software.

8.11 Sense of Place

Q: Data are only available for marine protected areas, not terrestrial protected areas. Can we still calculate the Lasting Special Places sub-goal?

A: Yes, it is possible to calculate only the marine component of this sub-goal: this is not ideal but OHI is flexible to work with the data available.

8.12 Pressures

Q: How are single ecological pressures (si in Equation S8) calculated?

A: Data included in pressures calculations are accessed in the same manner as any other data layer, and rescaled from 0-1 with an appropriate reference point. For further information, see HowTo_GatherAppropriateData and HowTo_CalculatePressures from ohi-science.org.

Q: Does the pressures matrix need to be changed?

A: It is likely that the pressures matrix will not need to be changed. The weights assigned in the matrix were set using information from the literature and by experts; the matrix was created by Halpern et al. 2012.

Q: How is commercial high and low bycatch calculated?

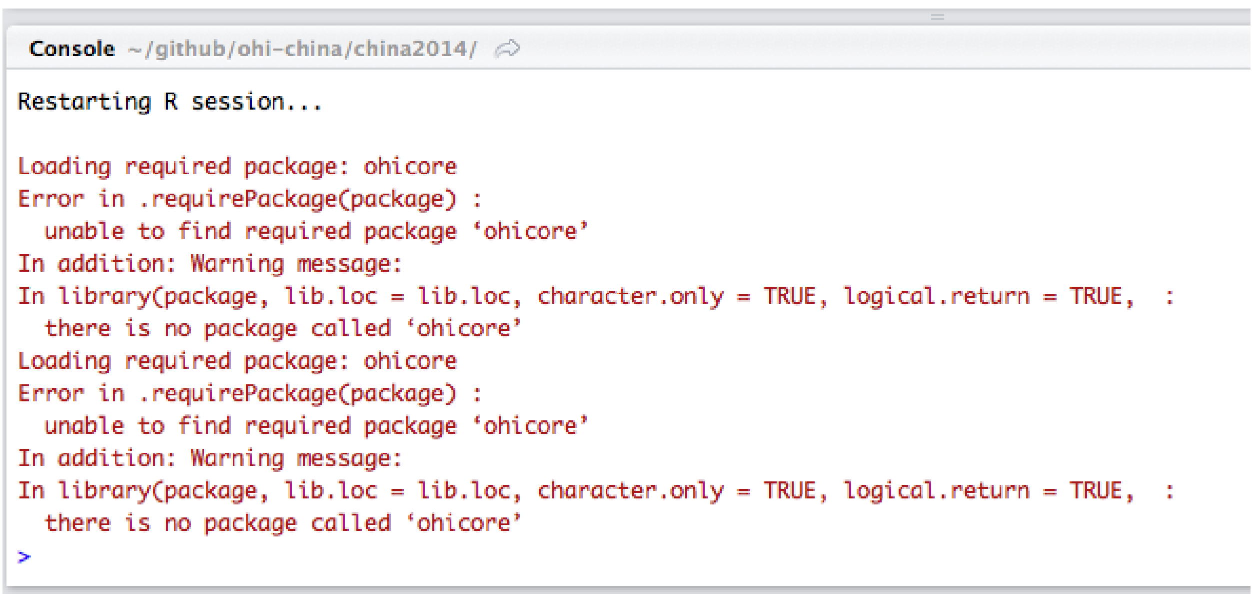
A: Commercial high and low bycatch are categorical values that were set based on fishing gear type. This began as a list of gear types used, producing a range of potential bycatch frequencies (from local reports when possible), which can be rescaled.

9 Toolbox Troubleshooting

The Toolbox prints messages during its processing to help guide error checking and debugging. Here are a few troubleshooting tips.
(** note: this page is under development).

9.1 Loading RWorkspace on Restart

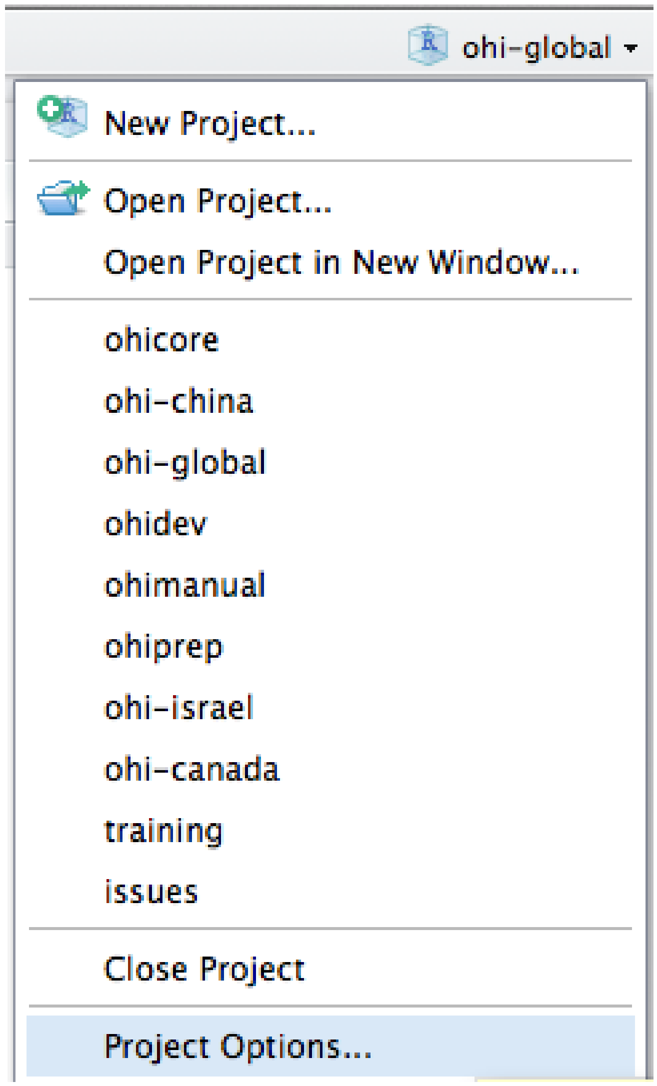
When you restart your R Session (Session > Restart R on a Mac), if you see that it is trying to load `ohicore`, it may give you an error:

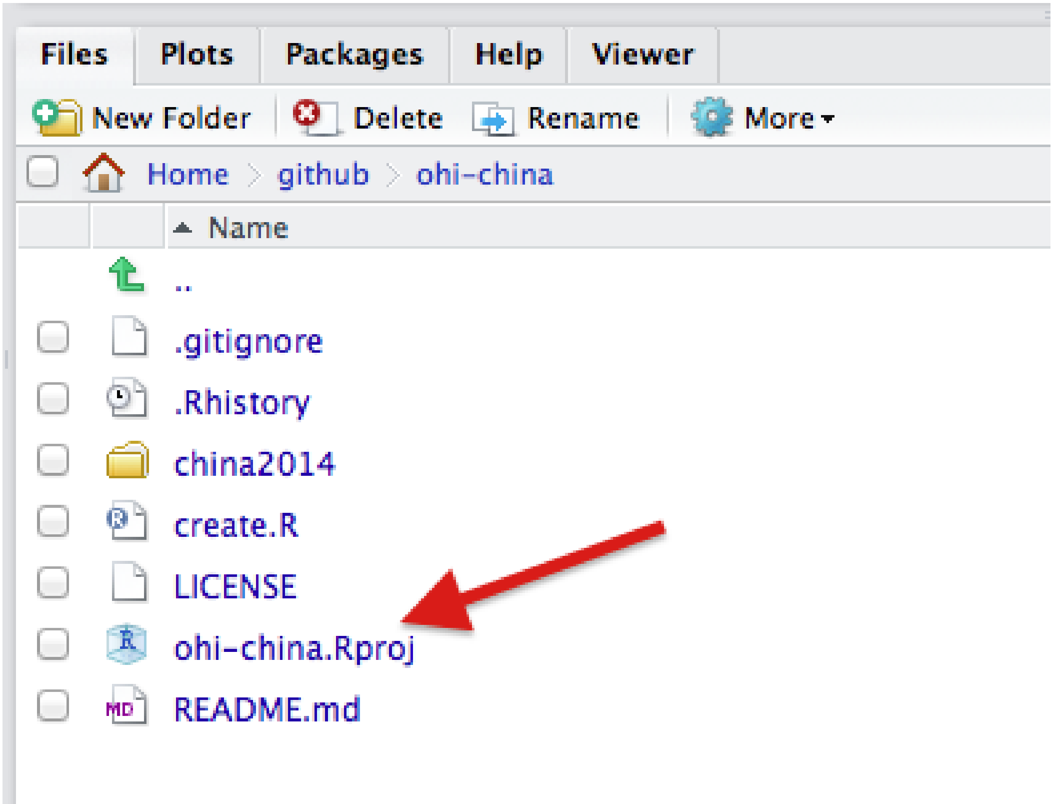


You do not want it to load `ohicore` or to save anything in your workspace. You will need to change the default setting from you **.Rproj** file. Steps to do this:

1. Go to Project Options, either in the pull-down menu or by double-clicking the .Rproj file:

2. Change all options to **No:**


ohi-global
New Project...
Open Project...
Open Project in New Window...
ohicore
ohi-china
ohi-global
ohidev
ohimanual
ohiprep
ohi-israel
ohi-canada
training
issues
Close Project
Project Options...



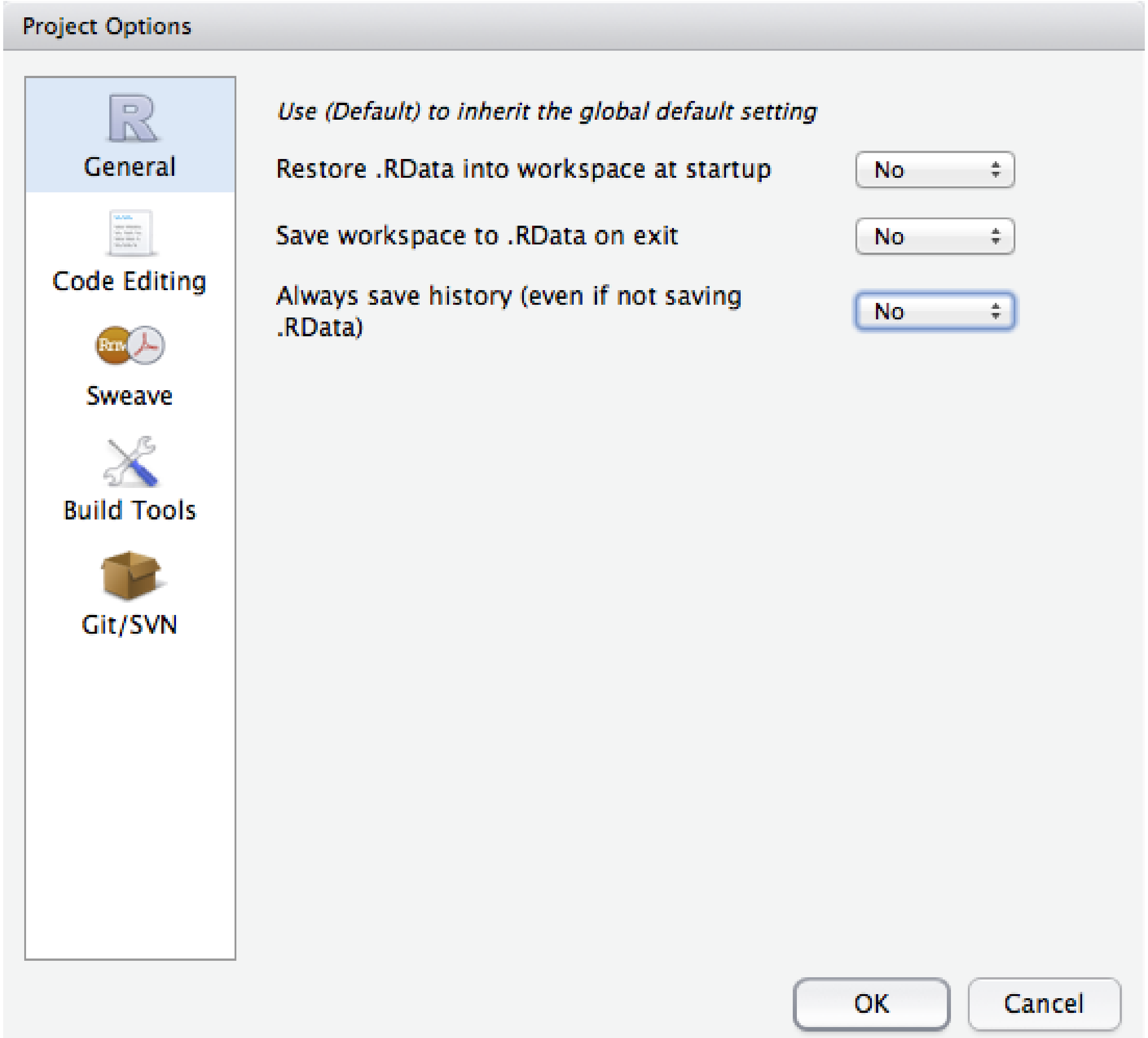
9.2 Calculating Pressures. . .

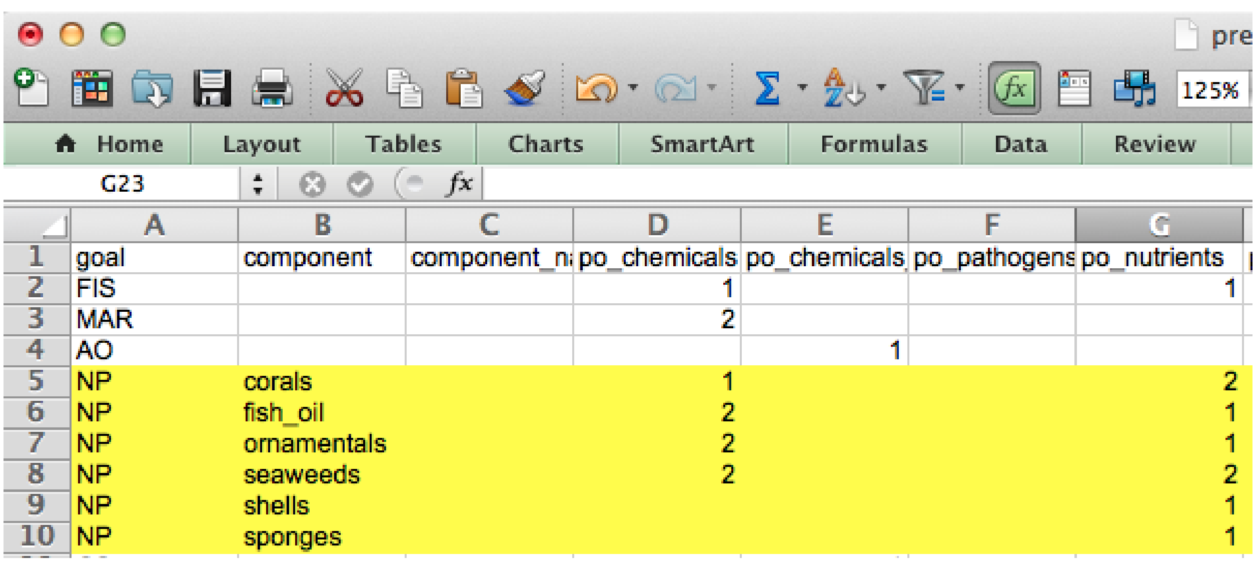
9.2.1 'The following components for [goal] are not in the aggregation layer [layer]. . . '

Example:

```
Running Setup()...
Calculating Pressures...
The following components for NP are not in the aggregation layer np_harvest_product_weight categories (
Error in data.frame(names(P), P) :
  arguments imply differing number of rows: 0, 1
```

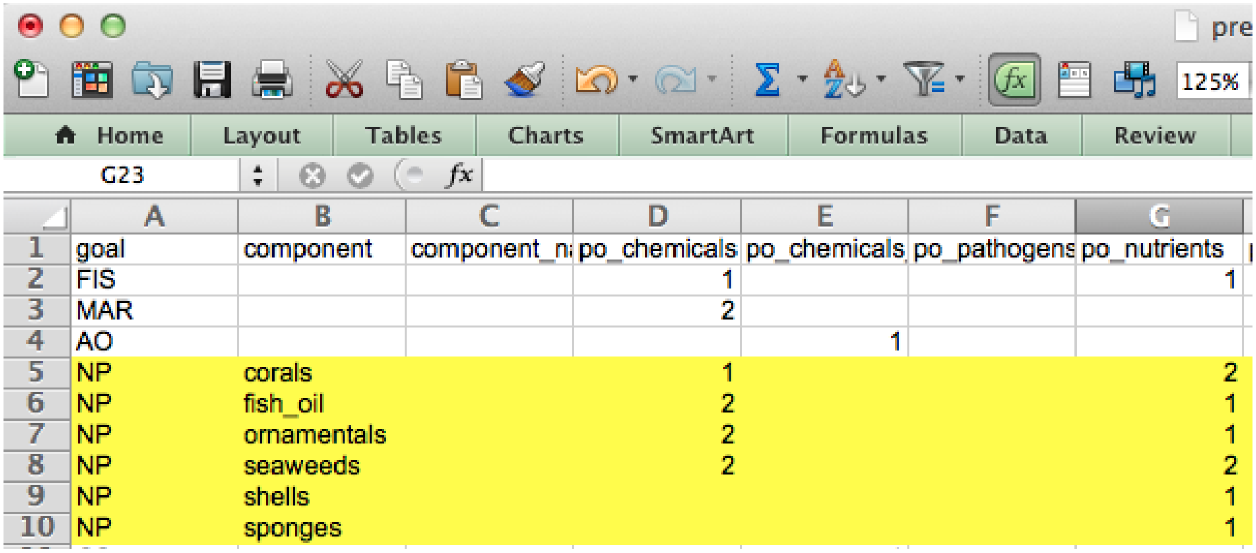
This error means you should update your pressures matrix because it expects there to be components that your region does not have.


Project Options
General
Code Editing
Sweave
Build Tools
Git/SVN
Use (Default) to inherit the global default setting
Restore .RData into workspace at startup
No
Save workspace to .RData on exit
No
Always save history (even if not saving .RData)
No
OK
Cancel


Home
Layout
Tables
Charts
SmartArt
Formulas
Data
Review
G23
125%
goal component component_n po_chemicals po_chemicals po_pathogens po_nutrients
FIS 1 1
MAR 2
AO 1
NP corals 1 2
NP fish_oil 2 1
NP ornamentals 2 1
NP seaweeds 2 2
NP shells 1
NP sponges 1

9.2.2 'Error in matrix. . . '

Example: >

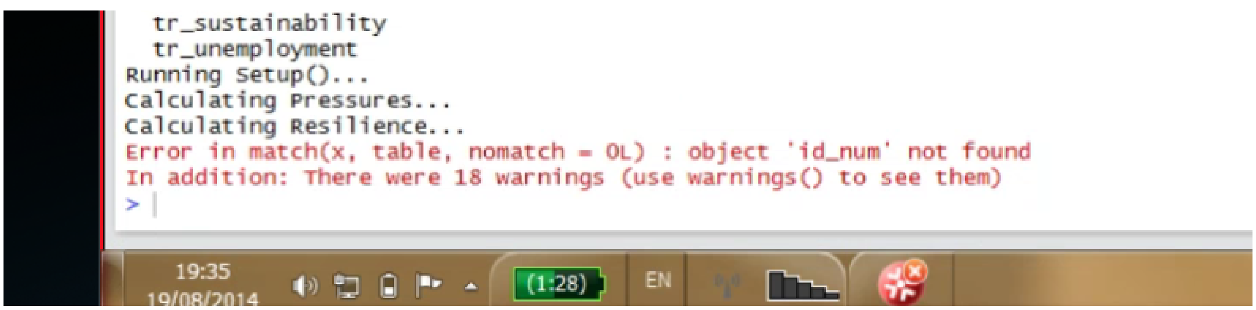


	A	B	C	D	E	F	G
1	goal	component	component_n	po_chemicals	po_chemicals	po_pathogens	po_nutrients
2	FIS			1			1
3	MAR			2			
4	AO				1		
5	NP	corals		1			2
6	NP	fish_oil		2			1
7	NP	ornamentals		2			1
8	NP	seaweeds		2			2
9	NP	shells					1
10	NP	sponges					1

This error means there is an empty column in `pressures_matrix.csv`, and the Toolbox cannot handle empty columns.

9.3 Calculating Resilience . . .

9.3.1 'Error in match(x, table, nomatch = OL) : object id_num not found'



This error means you should check that there is at least one entry for each goal (for each row) in `resilience_matrix.csv`.