kelpdecline: an R package to detect giant kelp decline in the Californias

```
library(kelpdecline)
```

The kelpdecline package aims to detect regions of kelp biomass in decline. The detection is based on a time series of Landsat estimated biomass (Cavanaugh et al. 2011, Bell et al. 2017). The figures and tables produced quantify the proportion of Landsat pixels (30 x 30 m) with biomass in decline within 0.25 x 0.25 degrees, Lat x Long, regions. The method proposed provides a framework to understand kelp decline in the context of a historical baseline and is valuable for all kelp forest conservation purposes. Our specific motivation for building this tool was to design an efficient sampling strategy for genetic monitoring studies (Klingbeil et al. 2022). Further considerations are found in the companion publication describing this package (Tennies & Alberto, in prep).

Input data

This package contains only two functions. The analysis starts by processing a NetCDF file containing a time series of kelp biomass curated by the Santa Barbara Coatal-Long Term Ecological Research (SBC-LTER) program. The data is available at https://sbclter.msi.ucsb.edu/data/catalog/package/?package=knb-lter-sbc.74. Make sure to read the license agreement before using this package. We quote below how SBC-LTER describes the file.

"This data file represents a time series of canopy area of giant kelp, *Macrocystis pyrifera*, and bull kelp, *Nereocystis luetkeana*, and canopy biomass of giant kelp derived from Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), Landsat 8 Operational Land Imager (OLI), and Landsat 9 Operational Land Imager 2 satellite imagery, along with relevant metadata. The kelp canopy is composed of the portions of fronds and stipes floating on the surface of the water. Canopy area (m) data are given for individual 30 x 30 meter pixels for all coastal areas of Baja California, Mexico, California, Oregon, and the outer coast of Washington (including offshore islands). Biomass data (wet weight, kg) are given for individual 30 x 30 meter pixels in the coastal areas extending from near Ano Nuevo, CA through the southern range limit in Baja California (including offshore islands), representing the range where giant kelp is the dominant canopy forming species."

The kelpdecline package uses only the biomass (giant kelp) layer in the NetCDF file. The package does not come with the NetCDF data file; users need to download it from the link above to run the examples below. The most recent file is about 2.3 GB, and the file is updated annually.

Example run

The first function nc_convert, converts the NetCDF file into a data.frame used internally by kelpdecline. Below, when providing the name of the NetCDF file in quotes, do not use the ../../ before the file name if the file is saved in the R working directory.

```
# The conversion function has a single argument that takes the name of the nc file
kelp.TS.DF<-nc_convert(nc_data_location="../../LandsatKelpBiomass_2022_Q4_withmetadata.nc")
```

The data.frame object *kelp.TS.DF* is then the main argument (*data*) of the main function *decline_finder*.

```
PropDeclineRaster <- decline_finder(data = kelp.TS.DF, baseline_threshold = 0.1,
    scarce_cutoff = 0.6, present_window = 16, hist_period = 100,
```

```
    window_lag = 0, lat_min = 27.01, lat_max = 37.5, lon_min = -123.5,
    lon_max = -114, table_name = "Output stats.txt")
```

The function saves a raster file to the global environment containing the proportions of pixels in decline within larger areas of 0.25 x 0.25, lat x long decimal degrees, hereafter called regions. A text tab-delimited text file is recorded to the working directory if a file name is supplied to the argument *table_name*. The default to *table_name* is NULL, so make sure to supply a file name to get an output table written to the working directory.

The output table produced by *decline_finder* is shown below for the full extent of the giant kelp biomass layer in the NetCDF file. We first read the file produced by *decline_finder* back into R to visualize it and run other analysis if need be. For better visualization of the output file, it is better to use spreadsheet software like Excel.

The most critical outputs in this table are in the first two columns: the proportion of pixels in decline and the total number of pixels for each region (row). Users should know the bounding coordinates for regions of interest (last columns). A high proportion of decline should be interpreted in the context of the number of pixels in that region, but also by the location of the region. For example, small kelp beds might be critical for stepping-stone connectivity. To better interpret the table below, we provide here row indexes for the following regions with substantial declines: Monterey Peninsula (4th row); San Miguel Island (23rd row); Carlsbad (46th row); Primo Tapia (53rd row); and Islas de San Benito (69th row). Note that changing the value for argument *scarce_cutoff* (read below) will likely result in a different number of rows in the output value, so always check for the boundary geographic coordinates for regions of interest.

```
Summary.Stats<-read.delim("Output stats.txt")
```

Table 1: The output table produced by running decline_finder with default values to its arguments. In this example, the table is stored in Summary.Stats object, in R memory, and in the Output stats.txt file, in the working directory

	decline_proportion	Pixels	AVG.biomass	kelp.area_km.sq	long_west	long_east	lat_south	lat_north
1	0.000	54	741.5	0.049	-122.50	-122.25	37.00	37.25
2	0.013	815	1107.5	0.733	-122.25	-122.00	36.75	37.00
3	0.008	626	940.1	0.563	-122.00	-121.75	36.75	37.00
4	0.432	2656	1150.1	2.390	-122.00	-121.75	36.50	36.75
5	0.104	5242	1467.2	4.718	-122.00	-121.75	36.25	36.50
6	0.009	846	1769.2	0.761	-122.00	-121.75	36.00	36.25
7	0.173	3040	1421.1	2.736	-121.75	-121.50	36.00	36.25
8	0.000	17	1150.4	0.015	-121.75	-121.50	35.75	36.00
9	0.094	2382	1355.9	2.144	-121.50	-121.25	35.75	36.00
10	0.007	458	1029.9	0.412	-121.50	-121.25	35.50	35.75
11	0.021	3333	1186.8	3.000	-121.25	-121.00	35.50	35.75
12	0.004	486	1023.5	0.437	-121.25	-121.00	35.25	35.50
13	0.081	1234	1261.0	1.111	-121.00	-120.75	35.25	35.50
14	0.185	211	1192.9	0.190	-121.00	-120.75	35.00	35.25
15	0.025	564	1106.9	0.508	-120.75	-120.50	35.00	35.25
16	0.006	783	1284.5	0.705	-120.75	-120.50	34.50	34.75
17	0.024	127	1364.3	0.114	-120.75	-120.50	34.25	34.50
18	0.108	2468	834.2	2.221	-120.50	-120.25	34.25	34.50
19	0.051	525	916.0	0.472	-120.25	-120.00	34.25	34.50
20	0.001	1820	932.1	1.638	-120.00	-119.75	34.25	34.50
21	0.000	98	855.7	0.088	-119.75	-119.50	34.25	34.50
22	0.020	50	846.3	0.045	-119.50	-119.25	34.25	34.50
23	0.800	2622	1227.9	2.360	-120.50	-120.25	34.00	34.25

	decline_proportion	Pixels	AVG.biomass	kelp.area_km.sq	long_west	long_east	lat_south	lat_north
24	0.756	41	747.2	0.037	-120.25	-120.00	34.00	34.25
25	0.317	104	799.6	0.094	-120.00	-119.75	34.00	34.25
26	0.000	14	341.2	0.013	-119.75	-119.50	34.00	34.25
27	0.000	70	450.2	0.063	-119.50	-119.25	34.00	34.25
28	0.193	984	872.9	0.886	-119.00	-118.75	34.00	34.25
29	0.424	33	813.4	0.030	-118.75	-118.50	34.00	34.25
30	1.000	5	1106.7	0.004	-120.50	-120.25	33.75	34.00
31	0.195	1765	904.3	1.588	-120.25	-120.00	33.75	34.00
32	0.013	792	722.8	0.713	-120.00	-119.75	33.75	34.00
33	0.008	244	466.3	0.220	-119.75	-119.50	33.75	34.00
34	0.000	524	537.7	0.472	-118.50	-118.25	33.75	34.00
35	0.000	291	605.1	0.262	-118.50	-118.25	33.50	33.75
36	0.000	482	853.7	0.434	-119.75	-119.50	33.25	33.50
37	0.000	6	719.3	0.005	-119.50	-119.25	33.25	33.50
38	0.143	7	365.6	0.006	-119.25	-119.00	33.25	33.50
39	0.000	149	563.9	0.134	-118.75	-118.50	33.25	33.50
40	0.020	251	710.0	0.226	-118.50	-118.25	33.25	33.50
41	0.739	23	690.5	0.021	-117.75	-117.50	33.25	33.50
42	0.000	18	741.8	0.016	-117.50	-117.25	33.25	33.50
43	0.000	686	990.1	0.617	-119.75	-119.50	33.00	33.25
44	0.020	450	849.9	0.405	-119.50	-119.25	33.00	33.25
45	0.012	1027	713.0	0.924	-118.75	-118.50	33.00	33.25
46	0.955	22	698.0	0.020	-117.50	-117.25	33.00	33.25
47	0.000	2799	763.8	2.519	-118.75	-118.50	32.75	33.00
48	0.006	1443	744.9	1.299	-118.50	-118.25	32.75	33.00
49	0.060	811	607.1	0.730	-117.50	-117.25	32.75	33.00
50	0.099	1511	660.9	1.360	-117.50	-117.25	32.50	32.75
51	0.234	64	803.9	0.058	-117.25	-117.00	32.25	32.50
52	1.000	1	64.5	0.001	-117.00	-116.75	32.25	32.50
53	0.577	758	893.1	0.682	-117.00	-116.75	32.00	32.25
54	0.056	54	771.1	0.049	-117.00	-116.75	31.75	32.00
55	0.000	4	827.8	0.004	-116.75	-116.50	31.75	32.00
56	0.248	4122	1292.8	3.710	-116.75	-116.50	31.50	31.75
57	0.000	573	1032.0	0.516	-116.75	-116.50	31.25	31.50
58	0.332	190	1137.8	0.171	-116.50	-116.25	31.25	31.50
59	0.917	12	105.9	0.011	-116.50	-116.25	31.00	31.25
60	0.000	30	639.7	0.027	-116.50	-116.25	30.75	31.00
61	0.831	703	1247.0	0.633	-116.25	-116.00	30.25	30.50
62	0.885	61	862.3	0.055	-116.00	-115.75	30.25	30.50
63	0.000	136	1004.4	0.122	-116.00	-115.75	30.00	30.25
64	0.696	1131	1037.5	1.018	-116.00	-115.75	29.75	30.00
65	0.069	58	167.8	0.052	-115.75	-115.50	29.75	30.00
66	0.877	440	996.3	0.396	-116.00	-115.75	29.50	29.75
67	0.326	279	1202.2	0.251	-115.75	-115.50	29.50	29.75
68	0.143	14	1235.9	0.013	-115.25	-115.00	29.25	29.50
69	0.966	910	1014.1	0.819	-115.75	-115.50	28.25	28.50
70	0.341	1205	1759.0	1.084	-115.50	-115.25	28.25	28.50
71	0.600	1451	1546.3	1.306	-115.25	-115.00	28.25	28.50
72	0.145	2939	1386.8	2.645	-115.50	-115.25	28.00	28.25
73	0.843	83	1353.4	0.075	-115.25	-115.00	28.00	28.25
74	0.013	6515	1406.6	5.864	-115.25	-115.00	27.75	28.00
75	0.000	380	1216.1	0.342	-115.25	-115.00	27.50	27.75

	decline_proportion	Pixels	AVG.biomass	kelp.area_km.sq	long_west	long_east	lat_south	lat_north
76	0.007	4844	1570.6	4.360	-115.00	-114.75	27.50	27.75
77	0.217	115	984.8	0.104	-114.75	-114.50	27.50	27.75

Other important arguments in *decline_finder* function

The remaining arguments in the function control how we compare a present period of kelp biomass to a historical baseline. Using the approach described below, we first classify kelp biomass decline at the Landsat pixel scale (30 x30 m) as a boolean (i.e., TRUE or FALSE) variable. Then, the proportion of Landsat pixels flagged as in decline is counted within regions of 0.25 x 0.25 degrees, lat x long.

A pixel is classified as in decline if the biomass recorded for that pixel in ALL the most recent quarters (argument: *present_window*) is below a proportion (argument: *baseline_threshold*) of historical average biomass. The period used to calculate the baseline average biomass per pixel is defined by the argument *hist_period*. The flexibility allows using different temporal baselines in specific comparisons. Likewise, it is possible to slide the focus "present" window to the past by a certain lag number of quarters (argument: *window_lag*), allowing us to estimate how kelp decline statistics varied in time.

Landsat pixels where biomass was rarely recorded throughout the time series affect the calculation of historical averages Given the effect these pixels have, and their poor indicator status for the overall trend of a kelp forest, we removed from the analysis pixels that throughout the *hist_period* had zero biomass in more than a given proportion (argument: *scarce_cutoff*) of all quarters in the series.

Geographic boundaries for the study extent are available with the arguments *lat_min*, *lat_max*, *long_min*, and *long_max*. These default to the full extent of the biomass layer in the NetCDF file.

Once the *decline_finder* function runs, users can quickly plot the produced raster file containing the proportions of decline by 0.25 x 0.25 degree region.

```
#simple raster plot
raster::plot(PropDeclineRaster)
```
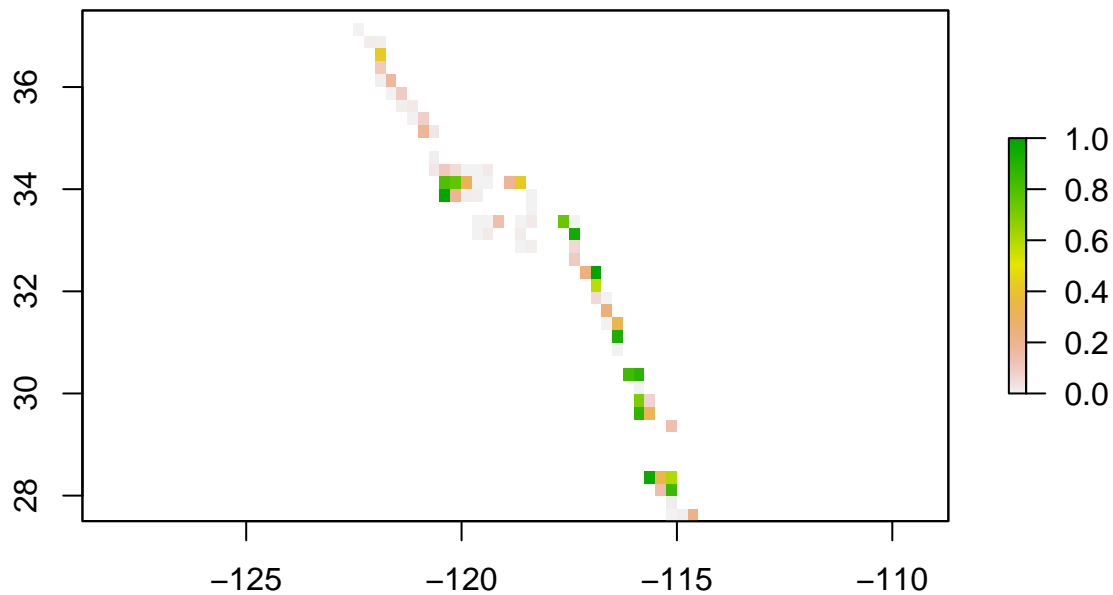
Below we improve the raster plot by adding shapefiles to visualize the coastline and using a different color scale. Note that the shapefiles used in the code below are not part of the package; users must obtain their own if they want to print the coastline.

```r
# making a custom color ramp
library("scales")
ramp <- scales::colour_ramp(c("darkgreen", "red1"))
ramp_discrete = ramp(seq(0, 1, length = 50))

# loading shapefiles.
cali_coast = raster::shapefile("../../california_coastline.shp")
raster::crs(cali_coast) = raster::crs(PropDeclineRaster)

baja_coast = raster::shapefile("../../bc_entidad.shp")
raster::crs(baja_coast) = raster::crs(PropDeclineRaster)

# plotting the raster with a better color scale
raster::plot(PropDeclineRaster, col = ramp(seq(0, 1, length = 10)),
    axes = T, main = "Proportion in decline")

# adding coastal lines
raster::lines(cali_coast)
raster::lines(baja_coast)
```

Proportion in decline
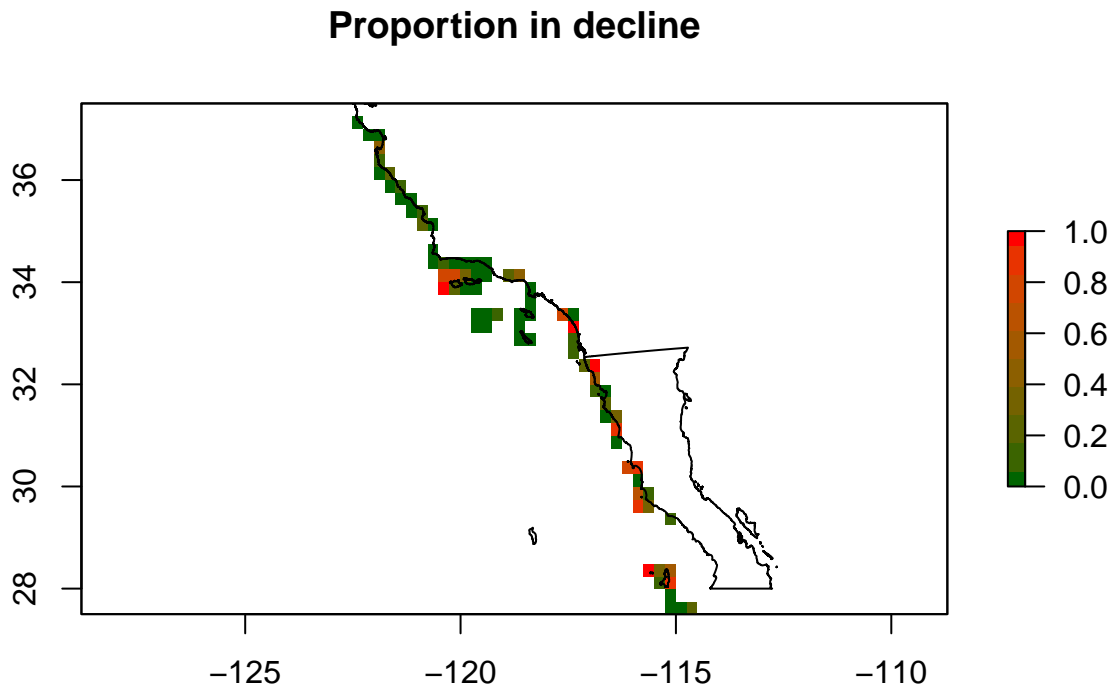
Sensitivity analysis

The function's default values should not be interpreted as the best practices to compare a present period with a historical baseline. These only reflect our opinion on reasonable initial values. The tool was built to provide a flexible method to explore the effects of parameter variation. Higher confidence about the proposed classification of kelp decline can be obtained for regions that consistently show high proportions of decline despite parameter variation. Below we explore this method for four regions of concern, initially identified with the default values. We chose to show here, as an example, four regions with different levels of decline: Monterey Bay (Central California), San Miguel Island (Southern California), Primo Tapia (Baja California), and Islas de San Benito (Baja California). Likewise, we show Point Loma (Southern California) as a region with stable biomass. As an example of a sensitivity analysis, we explored the effect of orthogonal combinations of *present_window* and *hist_threshold* in these five regions.

Be warned that the code below, exploring a large number of argument value combinations will take its time; about one hour on a normal laptop. Once the output files are produced the remaining code is fast.

```
# create vectors of parameter space over the two arguments

# baseline threshold from 0.1 to 1 in increments of 0.1
baseline_thresholdV <- seq(0.1, 1, 0.1)
# present_window from 8 to 24 quarters in increments of 2
# quarters
present_windowV <- seq(8, 24, 2)

# looping decline_finder over different combinations the
# arguments above
for (bt in baseline_thresholdV) {
```

```
    for (pw in present_windowV) {

        decline_finder(data = kelp.TS.DF, baseline_threshold = bt,
            scarce_cutoff = 0.6, present_window = pw, hist_period = 100,
            window_lag = 0, table_name = paste0("Stats.", bt,
                ".", pw, ".txt"))

        print(paste("pw: ", pw))  # helps to see where we are while waiting
    }
    print(paste("bt: ", bt))
}
```

The code above creates one output file for each combination of argument values in *baseline_threshold* (0.1 to 1, in increments of 0.1) and *present_window* (8 to 24 quarters in increments of 2). The output files will be saved in the working directory and named "Stats.*bt*.*pw*.txt", with *bt* and *pw* the values of *baseline_threshold* and *present_window* for a specific combination of values in the two arguments.

Next, we will read the output files back to R, using two nested *for* loops, and produce a perspective plot to visualize how the proportion of pixels in decline is affected by the combination of *baseline_threshold* and *present_window* values.

```
# baseline threshold from 0.1 to 1 in increments of 0.1
baseline_thresholdV <- seq(0.1, 1, 0.1)
# present_window from 8 to 24 quarters in increments of 2
# quarters
present_windowV <- seq(8, 24, 2)

# matrices to store the proportion in decline values for
# different argument combinations
Cell.Matrix.Monterey <- matrix(nrow = length(baseline_thresholdV),
    ncol = length(present_windowV))
Cell.Matrix.SanMiguel <- matrix(nrow = length(baseline_thresholdV),
    ncol = length(present_windowV))
Cell.Matrix.PrimoTapia <- matrix(nrow = length(baseline_thresholdV),
    ncol = length(present_windowV))
Cell.Matrix.SanBenito <- matrix(nrow = length(baseline_thresholdV),
    ncol = length(present_windowV))
Cell.Matrix.PointLoma <- matrix(nrow = length(baseline_thresholdV),
    ncol = length(present_windowV))

for (bt in 1:length(baseline_thresholdV)) {
    for (pw in 1:length(present_windowV)) {

        DFtemp <- utils::read.delim(paste0("Stats.", baseline_thresholdV[bt],
            ".", present_windowV[pw], ".txt"))

        # filter output for the specific coordinates for
        # different regions
        regionMonterey <- (DFtemp$long_west == -122 & DFtemp$long_east ==
            -121.75 & DFtemp$lat_south == 36.5 & DFtemp$lat_north ==
            36.75)
        regionSanMiguel <- (DFtemp$long_west == -120.5 & DFtemp$long_east ==
            -120.25 & DFtemp$lat_south == 34 & DFtemp$lat_north ==
            34.25)
```

```
        regionPrimoTapia <- (DFtemp$long_west == -117 & DFtemp$long_east ==
            -116.75 & DFtemp$lat_south == 32 & DFtemp$lat_north ==
            32.25)
        regionSanBenitos <- (DFtemp$long_west == -115.75 & DFtemp$long_east ==
            -115.5 & DFtemp$lat_south == 28.25 & DFtemp$lat_north ==
            28.5)
        regionPointLoma <- (DFtemp$long_west == -117.5 & DFtemp$long_east ==
            -117.25 & DFtemp$lat_south == 32.5 & DFtemp$lat_north ==
            32.75)

        Cell.Matrix.Monterey[bt, pw] <- DFtemp$decline_proportion[regionMonterey]
        Cell.Matrix.SanMiguel[bt, pw] <- DFtemp$decline_proportion[regionSanMiguel]
        Cell.Matrix.PrimoTapia[bt, pw] <- DFtemp$decline_proportion[regionPrimoTapia]
        Cell.Matrix.SanBenito[bt, pw] <- DFtemp$decline_proportion[regionSanBenitos]
        Cell.Matrix.PointLoma[bt, pw] <- DFtemp$decline_proportion[regionPointLoma]


    }
}
```

Finally, we can create perspective plots using the information in the Cell.Matrix objects for the different regions.

```
# For Monterey
par(cex = 0.7)
persp(baseline_thresholdV, present_windowV, Cell.Matrix.Monterey,
    theta = 140, phi = 10, zlim = c(0, 1), zlab = "proportion in decline",
    xlab = "baseline threshold", ylab = "present window", main = "Monterey peninsula",
    border = "grey65", col = "grey95", ticktype = "detailed")
```
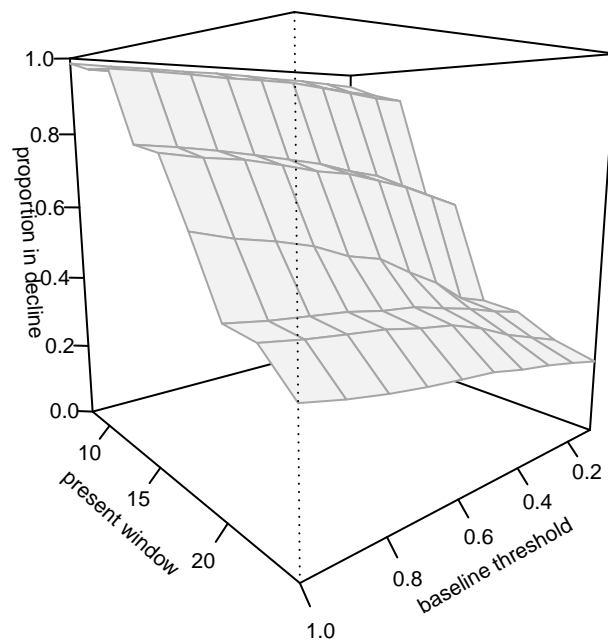
Monterey peninsula
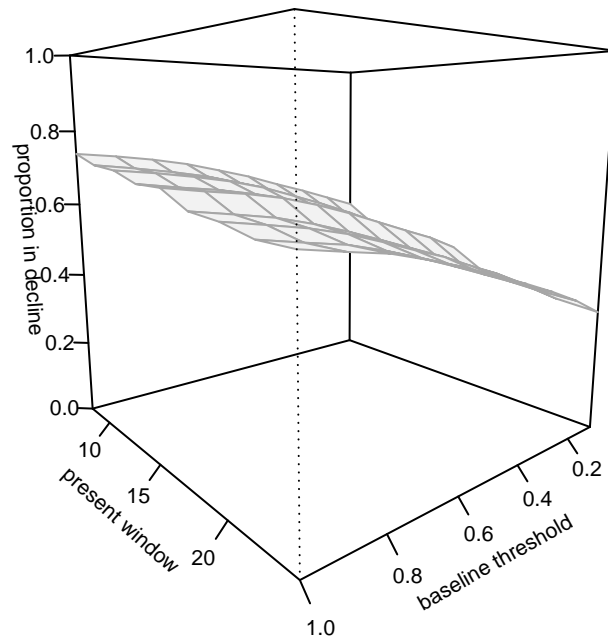

```
# For San Miguel Island
par(cex = 0.7)
persp(baseline_thresholdV, present_windowV, Cell.Matrix.SanMiguel,
    theta = 140, phi = 10, zlim = c(0, 1), zlab = "proportion in decline",
    xlab = "baseline threshold", ylab = "present window", main = "San Miguel Island",
    border = "grey65", col = "grey95", ticktype = "detailed")
```
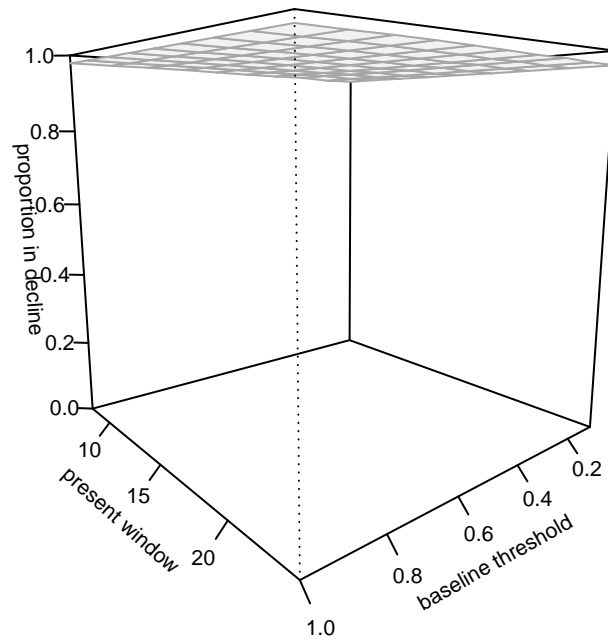
San Miguel Island
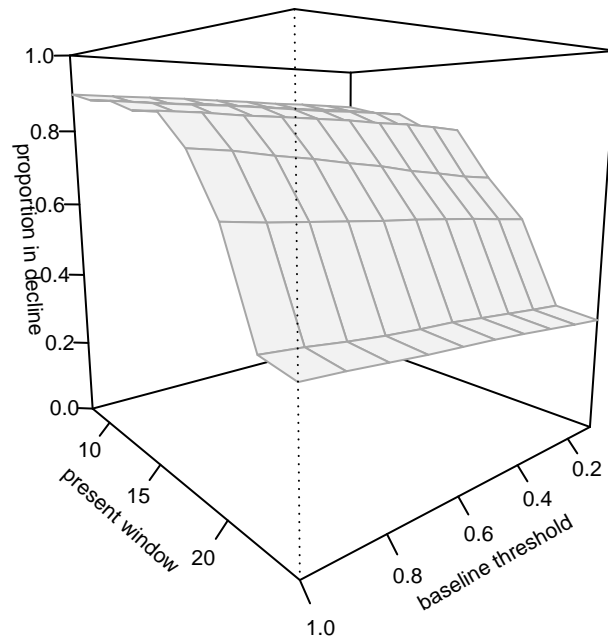

```r
# For Primo Tapia
par(cex = 0.7)
persp(baseline_thresholdV, present_windowV, Cell.Matrix.PrimoTapia,
    theta = 140, phi = 10, zlim = c(0, 1), zlab = "proportion in decline",
    xlab = "baseline threshold", ylab = "present window", main = "Primo Tapia",
    border = "grey65", col = "grey95", ticktype = "detailed")
```

Primo Tapia


```r
# For Islas de San Benito
par(cex = 0.7)
persp(baseline_thresholdV, present_windowV, Cell.Matrix.SanBenito,
    theta = 140, phi = 10, zlim = c(0, 1), zlab = "proportion in decline",
    xlab = "baseline threshold", ylab = "present window", main = "Islas de San Benito",
    border = "grey65", col = "grey95", ticktype = "detailed")
```

Islas de San Benito


```
# For Point Loma
par(cex = 0.7)
persp(baseline_thresholdV, present_windowV, Cell.Matrix.PointLoma,
    theta = 140, phi = 10, zlim = c(0, 1), zlab = "proportion in decline",
    xlab = "baseline threshold", ylab = "present window", main = "Point Loma",
    border = "grey65", col = "grey95", ticktype = "detailed")
```

Point Loma
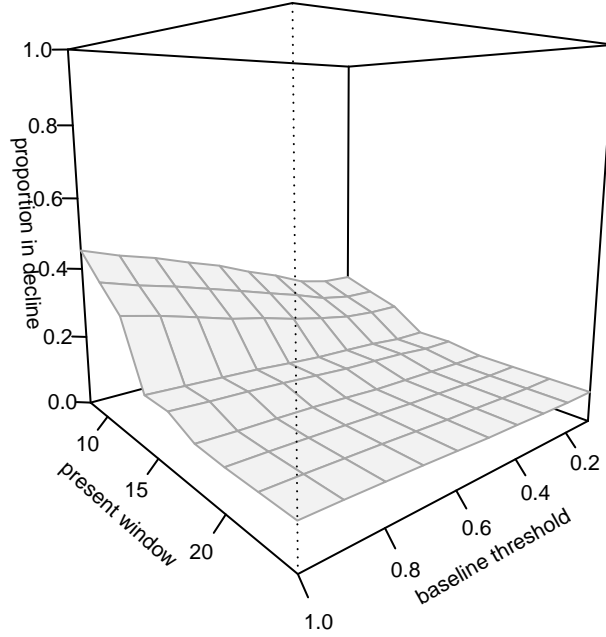
A few comments on the results from sensitivity analysis; The effect of *present_window* on proportion in decline is more prominent than *baseline_threshold*, for the range of parameters tested. Remember the argument *present_window* defines the length in quarters of the present window. For a pixel to be classified in decline, it must be below a threshold proportion of the historical biomass average for that pixel throughout all quarters in *present_window*. Therefore, the longer the present period, the higher the chances that some quarters have higher biomass than the threshold. This means that the effect of *present_window* in the proportion of pixels in decline is monotonically decreasing.

How fast the curve decreases over an axis of variable *present_window* periods is informative of how long a certain proportion of decline is present. For example, in the San Miguel and Primo Tapia regions, the proportion of pixels in decline falls fast from 8 to 16 quarters (2 to 4 years). In contrast, in Monterey and Islas de San Benito, we observe a flatter surface from 8 to 24 quarters (2-6 years). The conclusion is that the decrease in kelp biomass has been more permanent in these two regions, with Islas de San Benito looking very close to a case of local extinction (i.e., flat surface near 1.0 proportion of decline).

Note that another consequence of this algorithm is that a single (new) quarter with a large proportion of the pixels recovering above the threshold is sufficient to produce a flat surface at much lower levels of proportion in decline (z-axis). This makes sense because the method should capture full recovery whenever it occurs in a future quarter.

Acknowledgments and license

aimed at larger areas and used other methods (Bell et al. 2022 preprint).

References

Bell, T., Cavanaugh, K.C., Saccomanno, V.R., Cavanaugh, K.C., Houskeeper, H.F., Eddy, N., Schuetzenmeister, F., Rindlaub, N. and Gleason, M., (2022) Kelpwatch: A new visualization and analysis tool to explore kelp canopy dynamics reveals variable resistance and resilience to marine heat waves. bioRxiv, pp.2022-07.

Bell, T, K. Cavanaugh, D. Siegel. (2023) SBC LTER: Time series of quarterly NetCDF files of kelp biomass in the canopy from Landsat 5, 7 and 8, since 1984 (ongoing) ver 19. Environmental Data Initiative. https://doi.org/10.6073/pasta/630565d6a8bf54c7cbce6802284dd431. Accessed 2023-02-22.

Cavanaugh, K. C., Siegel, D. A., Reed, D. C., & Dennison, P. E. (2011). Environmental controls of giant-kelp biomass in the Santa Barbara Channel, California. *Marine Ecology Progress Series*, 429, 1-17.

Klingbeil, W., Montecinos, G., Alberto, F. (2022) Giant kelp genetic monitoring before and after disturbance reveals stable genetic diversity in Southern California. Frontiers in Marine Science. 9:947393.doi: 10.3389/fmars.2022.947393

Tennies, N., Alberto, F. (in prep) A tool for detecting giant kelp canopy biomass decline in the Californias. *Journal of Phycology*