

Analysis of Reef Fish and Invertebrate Assemblage Structure on High and Low Human Impact Hawaiian Coral Reefs

Authors: James Sturges & Tari Lopez

Last Modified: 2 December 2022

README

This script analyzes data of Hawaiian coral reef metrics from 5 different sites on the big island of Hawaii. These data were compiled by Drs. Jason Smith & Jeremy Claisse and were the result of ongoing student-lead sampling efforts. James Sturges was a member of the 2019 class and 2022 sampling team.

```
#### Packages ----
```

```
library(tidyverse)
library(readxl)
library(vegan)
```

```
#### Site Metadata ----
```

```
# 5 sites (3 high impact 2 low impact)
# All sites are located on the dry side of the Big Island
# Kahaluu is the only site with a high and low impact site adjacent to each other
dat_metasite <- read_excel("Coral_Reef_Data_2014_2018_2022_Final.xlsx", "Site Metadata")
```

```
##### Fish Metadata ----
```

```
# metadata includes 3 name types (Common, Latin, and Hawaiian)
# also includes diet categories (Herbivore, Omnivore, and Predator)
# Note that not all fish have a Hawaiian name so there are NAs
fish_meta <- read_excel("Coral_Reef_Data_2014_2018.xlsx", "Fish Metadata")
```

```
#### Fish Community Data ----
```

```
#Creates fish df in wide format from raw data
dat_fish <- read_excel("Coral_Reef_Data_2014_2018_2022_Final.xlsx", "Fish")
```

```
# Removes spp. Manybar Goatfish
# Was not analyzed prior to 2022 sampling year
# changes col name to just Group for student lead sampling groups
```

```
dat_fish <- dat_fish %>%
  select(-"Manybar Goatfish") %>%
  rename(student_group = Group)
```

```
# change fish df to long format
```

```

dat_fish <- dat_fish %>%
  pivot_longer(cols = !Year:student_group, names_to = "com.name", values_to = "count")

# joins metadata for spp name and dietary strategy
dat_fish <- dat_fish %>%
  left_join(fish_meta)

```

Invertebrate Community Data ----

```

# creates a df for all invertebrates observed on swath transects
dat_inv <- read_excel("Coral_Reef_Data_2014_2018_2022_Final.xlsx", "Invertebrates")

# converted to long format
dat_inv <- dat_inv %>%
  pivot_longer(cols = !Year:Group, names_to = "Taxa", values_to = "Abundance")

```

Substrate and Coral Cover Data ----

```

# creates a dataframe for substrate percent cover estimates from visual surveys
# note that coral coverage values are reported as total sum of live and dead coral spp
# We estimated individual cover for live and dead coral by multiplying percent dead estimates with coral cover
dat_sub <- read_excel("Coral_Reef_Data_2014_2018_2022_Final.xlsx", "Substrate",
  range = cell_cols("A:X"))

```

We estimated percent cover of live and dead coral for each spp (3)

```

dat_sub <- dat_sub %>%
  mutate(Pocillopora_dead = round(Pocillopora_Live_Death*Perc_Pocillopora_dead/100,1),
    Pocillopora_live = round(Pocillopora_Live_Death*(100-Perc_Pocillopora_dead)/100,1),
    Porites_dead = round(Porites_Live_Death*Perc_Porites_dead/100,1),
    Porites_live = round(Porites_Live_Death*(100-Perc_Porites_dead)/100,1),
    Pavona_dead = round(Pavona_Live_Death*Perc_Pavona_dead/100,1),
    Pavona_live = round(Pavona_Live_Death*(100-Perc_Pavona_dead)/100,1),
  )

```

We want to look at how much of the coral reef was dead at the time of each transect
create table of just percentages of existing coral cover that is dead

```

perc_dead <- dat_sub %>%
  select(Year, Site, Replicate, Group, Perc_Pocillopora_dead, Perc_Porites_dead, Perc_Pavona_dead)

```

convert table to long format, remove characters so Taxa is Genus

```

perc_dead <- perc_dead %>%
  pivot_longer(cols = !Year:Group, names_to = "Taxa", values_to = "Perc_dead") %>%
  mutate(Taxa = str_remove(Taxa, "Perc_"),
    Taxa = str_remove(Taxa, "_dead"))

```

remove dat_sub columns that are not a percent cover, & combined live & dead % covers

```

dat_sub <- dat_sub %>%
  select(!c(Perc_Pocillopora_dead, Perc_Porites_dead, Perc_Pavona_dead, Pocillopora_Live_Death, Porites_Live_Death))

```

to tidy (long) format

```

dat_sub <- dat_sub %>%
  pivot_longer(cols = !Year:Group, names_to = "Substrate", values_to = "Perc_Cover")

```

```

#### Fish Assemblage Structure ----
# This code chunk creates the community dataframe in wide format with site and year as the unique ID.
# converts fish counts to density values based on the length of the transect
# Note that transects were 10 meters long but 2 meters wide
# We did not limit the vertical height of the transect
# thus we divided by 20 to estimate fish per m-squared
dat_fish <- dat_fish %>%
  mutate(density_m2 = count/20)

dat_fish_transect_avg <- dat_fish %>%
  group_by(Year, Site, com.name) %>%
  summarize(avg_density_m2 = mean(density_m2))

# creates wide format for multivariate analysis
wide_dat_fish_transect_avg <- dat_fish_transect_avg %>%
  pivot_wider(names_from = com.name, values_from = avg_density_m2)

# join with site meta data
wide_dat_fish_transect_avg <- wide_dat_fish_transect_avg %>%
  left_join(dat_metasite)

# create a unique ID column that combines site and year (each point on NMDS)
wide_dat_fish_transect_avg <- wide_dat_fish_transect_avg %>%
  mutate(site_year = paste(Site, Year))

# Replaces underscore with a space for label
names(wide_dat_fish_transect_avg) <- str_replace_all(names(wide_dat_fish_transect_avg), c(" " = "_"))

# Creates the community data frame
comm_dat_fish_transect_avg <- wide_dat_fish_transect_avg %>%
  column_to_rownames(var = "site_year") %>%
  select(Achilles_Tang:Zebra_Moray)

#### 1st NMDS object created ----
# Used for site and year specific patterns
# Would expect sites to be most similar to themselves across years
# point color will show the same site
# This is a good way to look at site and year specific trends without considering human impact
NMDS_comm_dat_fish_transect_avg <- metaMDS(comm_dat_fish_transect_avg,
  trymax = 999,
  distance = "bray",
  autotransform = F)
scores(NMDS_comm_dat_fish_transect_avg)

tibble_comm_dat_fish_transect_avg <- as_tibble((NMDS_comm_dat_fish_transect_avg$points),
  rownames = ("site_year"))

wide_dat_fish_transect_avg <- wide_dat_fish_transect_avg %>%
  left_join(tibble_comm_dat_fish_transect_avg)

# Using envfit we can see which species are influence the ordination of our plots
fish_spp_fit <- envfit(NMDS_comm_dat_fish_transect_avg, wide_dat_fish_transect_avg, permutations = 999)

```

```

site_scores <- wide_dat_fish_transect_avg %>%
  select(Year, Site, site_year, Human_Impact, MDS1, MDS2) %>%
  column_to_rownames(var = "site_year")

#determine which spp are significantly influencing the ordination
spp_scores <- as.data.frame(scores(fish_spp_fit, display = "vectors"))
spp_scores <- cbind(spp_scores, Species = rownames(spp_scores))
spp_scores <- cbind(spp_scores, pval = fish_spp_fit$vectors$pvals)
sig_spp_scores <- subset(spp_scores, pval <= 0.01)
sig_spp_scores <- sig_spp_scores %>%
  filter(!Species %in% "MDS1") %>%
  filter(!Species %in% "MDS2") %>%
  filter(!Species %in% "Latitude") %>%
  filter(!Species %in% "Longitude") %>%
  filter(!Species %in% "Year")

# only 3 species were significant in our ordination
# There is no evidence that year, site, or level of human impact altered fish assemblage structure
# could be the result of sampling bias from student lead fish counts

# Plot 1 of 1st NMDS object
# Fish assemblage structure with site as a color metric
# easily view temporal trends for the same site
plot_wide_dat_fish_transect_avg_site <- ggplot(site_scores,
  aes(MDS1, MDS2,
    fill = Site)) +
  geom_text((aes(label = Year)), vjust = -1.1, hjust = .5, size = 3) +
  geom_point(aes(shape = Site), size = 4) +
  ggtitle("Fish Assemblage by Site") +
  theme_classic() +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  geom_text(x = 1.1, y = 0.7, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_fish_transect_avg$stress, 2))) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  theme(axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.ticks = element_blank(),
    axis.text = element_blank())

```

```

## Warning: The 'size' argument of 'element_rect()' is deprecated as of ggplot2 3.4.0.
## i Please use the 'linewidth' argument instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

```

```

plot_wide_dat_fish_transect_avg_site

```

Fish Assemblage by Site

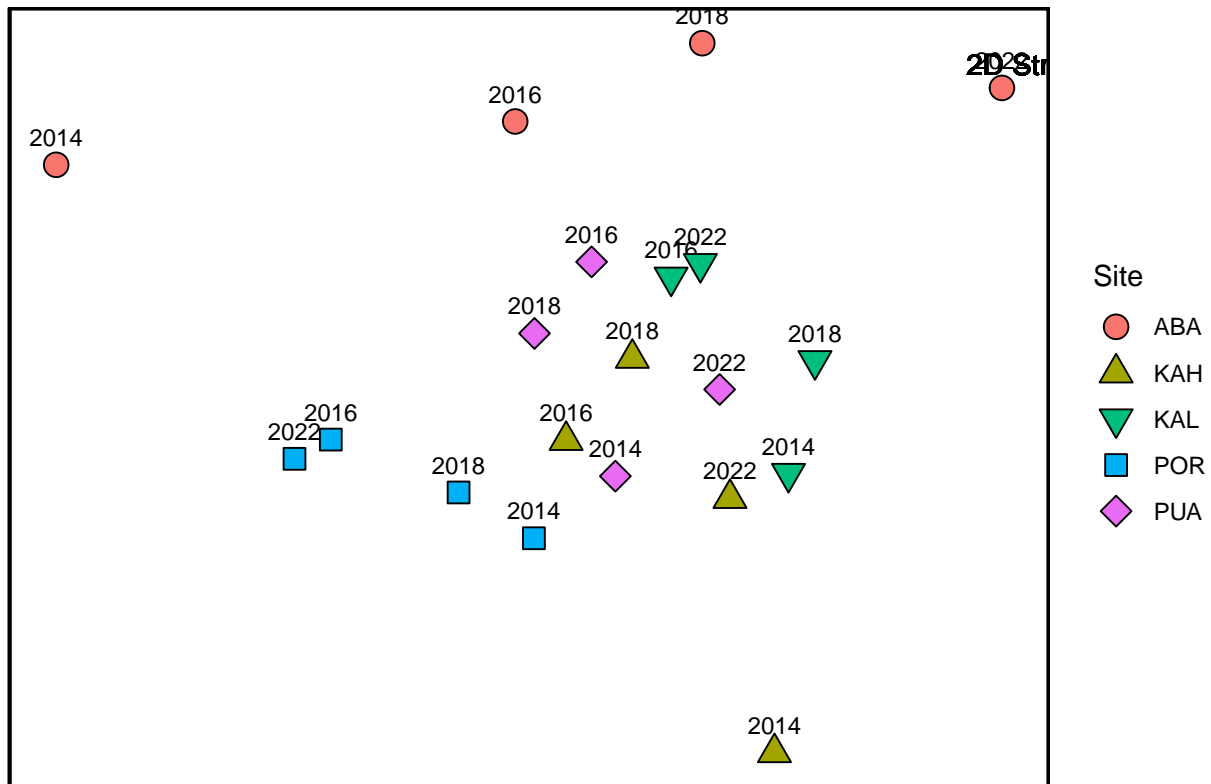


Figure 1 caption ----

#NMDS ordination of fish assemblage structure on five different Hawaiian reefs (color). Sampling at each

Plot 2 of 1st NMDS object ----

```
plot_wide_dat_fish_transect_avg_impact <- ggplot(site_scores,
  aes(MDS1, MDS2,
    shape = Site)) +
  geom_text(aes(label = Year), vjust = -1.0, hjust = 0.5, size = 2.5) +
  geom_point(aes(fill = Human_Impact, shape = Site), size = 3) +
  scale_fill_manual(values = c('High' = "firebrick4", 'Low' = "dodgerblue2")) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  ggtitle("Fish Assemblage with Human Impact") +
  theme_classic() +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  geom_text(x = 1.05, y = 0.7, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_fish_transect_avg$stress, 2)))) +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
  theme(axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.ticks = element_blank(),
    axis.text = element_blank())

plot_wide_dat_fish_transect_avg_impact
```

Fish Assemblage with Human Impact

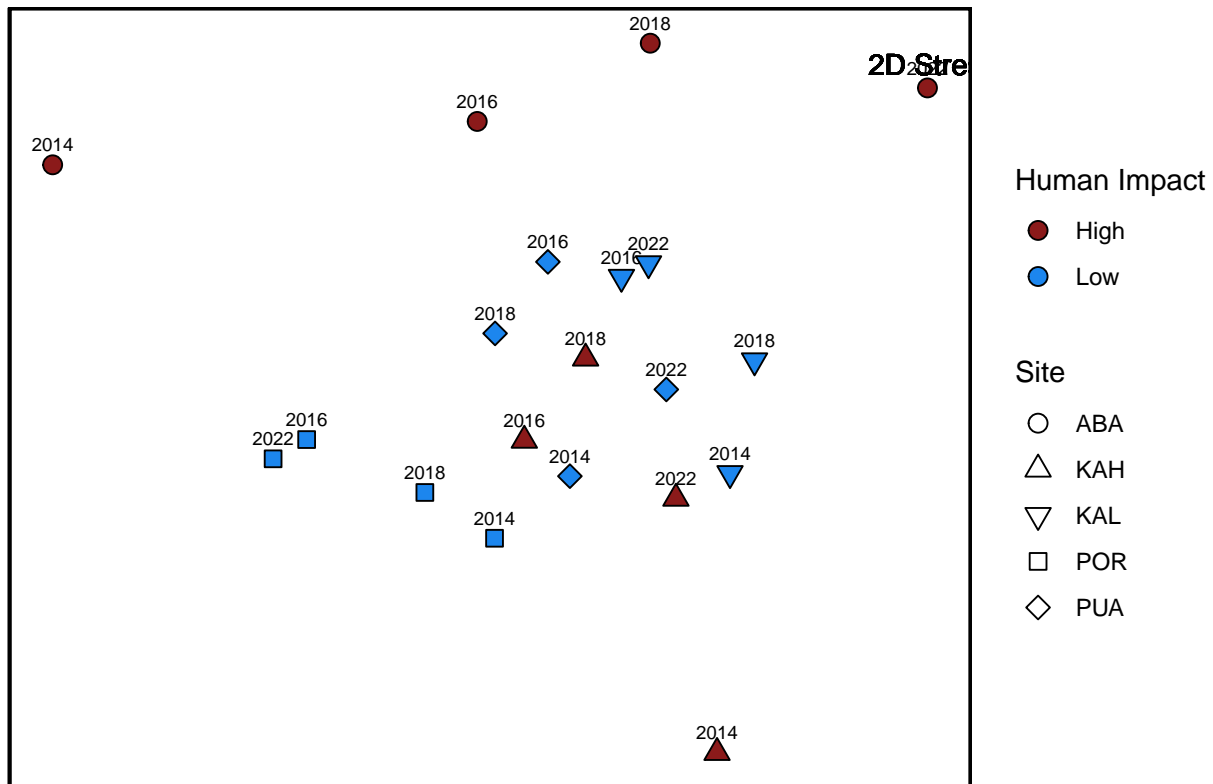


Figure 2 caption ----

#NMDS ordination of the same fish assemblages estimates used in Fig 1. This ordination has the points c

```
plot_NMDS_comm_fish_wide_transect_avg <- ggplot(site_scores, aes(x = MDS1, y = MDS2)) +
  geom_point(aes(fill = Site, shape = Site), size = 3) +
  # scale_fill_manual(values = c('High' = "firebrick4", 'Low' = "dodgerblue2")) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  ggtitle("Fish Assemblage by Site") +
  geom_text(aes(label = Year), vjust = -1.1, hjust = .5, size = 3) +
  geom_text(x = 1.1, y = 0.7, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_fish_transect_avg$stress1, 2))))
labs(shape = "Site") +
theme_classic()+
# guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid"),
  axis.title.x = element_blank(),
  axis.title.y = element_blank(),
  axis.ticks = element_blank(),
  axis.text = element_blank(),
  legend.position = "right",
  legend.text = element_text(size = 12),
  legend.title = element_text(size = 12))

#plot_NMDS_comm_fish_wide_transect_avg
```

```

plot_fish_spp_vectors <- plot_NMDS_comm_fish_wide_transect_avg +
  geom_segment(data = sig_spp_scores, aes(x = 0, xend = NMDS1, y = 0, yend = NMDS2),
    arrow = arrow(length = unit(.5, "cm")),
    colour = "grey10", lwd = 0.5) +
  ggrepel::geom_text_repel(data = sig_spp_scores,
    aes(x = NMDS1, y = NMDS2, label = Species), direction = "both", segment.size
plot_fish_spp_vectors

```

Fish Assemblage by Site

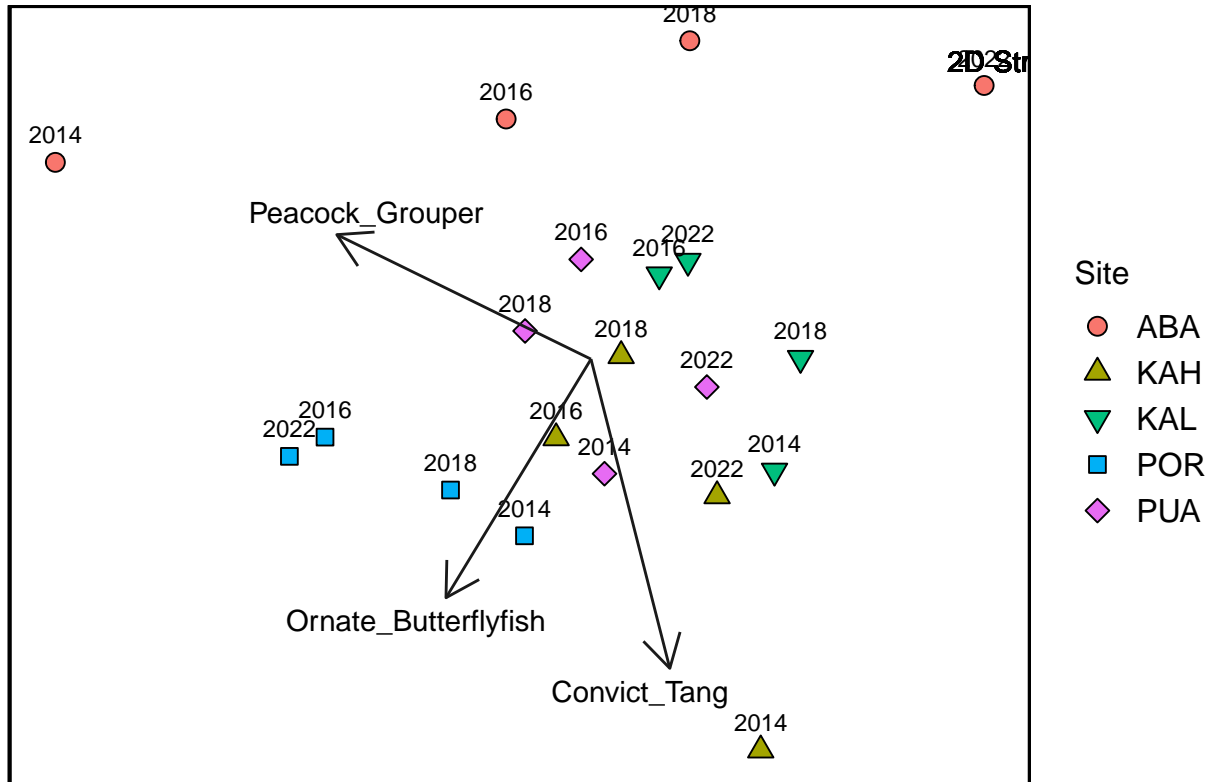


Figure 3 caption ----

#NMDS ordination of fish assemblage structure on five different Hawaiian reefs (color). Sampling at each

```

plot_NMDS_comm_fish_wide_transect_avg_impact <- ggplot(site_scores, aes(x = MDS1, y = MDS2)) +
  geom_point(aes(fill = Human_Impact, shape = Site), size = 3) +
  scale_fill_manual(values = c('High' = "firebrick4", 'Low' = "dodgerblue2")) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  ggtitle("Fish Assemblage with Human Impact") +
  geom_text(aes(label = Year), vjust = 2, hjust = .2, size = 2.5) +
  geom_text(x = 1.1, y = 0.7, label = c(paste("2D Stress:", round(NMDS_comm_dat_fish_transect_avg$stress
  labs(colour = "Human Impact", shape = "Site") +
  theme_classic() +
  guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid"),
    axis.title.x = element_blank(),

```

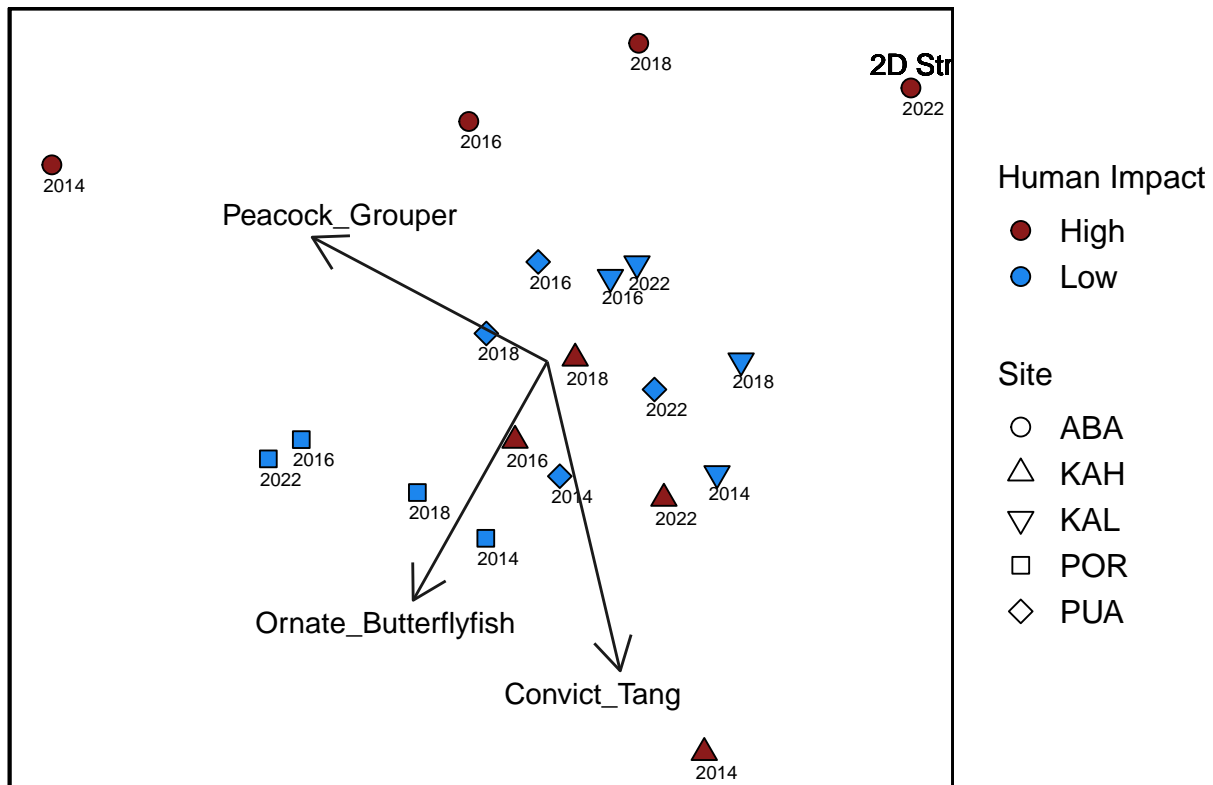
```
axis.title.y = element_blank(),
axis.ticks = element_blank(),
axis.text = element_blank(),
legend.position = "right",
legend.text = element_text(size = 12),
legend.title = element_text(size = 12))
```

```
#plot_NMDS_comm_fish_wide_transect_avg_impact
```

```
plot_fish_spp_vectors_impacts <- plot_NMDS_comm_fish_wide_transect_avg_impact +
  geom_segment(data = sig_spp_scores, aes(x = 0, xend = NMDS1, y = 0, yend = NMDS2),
    arrow = arrow(length = unit(.5, "cm")),
    colour = "grey10", lwd = 0.5) +
  ggrepel::geom_text_repel(data = sig_spp_scores,
    aes(x = NMDS1, y = NMDS2, label = Species), direction = "both", segment.size = 1)

plot_fish_spp_vectors_impacts
```

Fish Assemblage with Human Impact



```
#### Figure 4 caption ----
```

```
##NMDS ordination of fish assemblage structure on five different Hawaiian reefs (shape). Sampling at ea
```

```
#### Invert Community Structure ----
```



```

dat_inv <- dat_inv %>%
  mutate(density_m2 = Abundance/20)

dat_inv_transect_avg <- dat_inv %>%
  group_by(Year, Site, Taxa) %>%
  summarize(avg_density_m2 = mean(density_m2))

wide_dat_inv_transect_avg <- dat_inv_transect_avg %>%
  pivot_wider(names_from = Taxa, values_from = avg_density_m2)

wide_dat_inv_transect_avg <- wide_dat_inv_transect_avg %>%
  left_join(dat_metasite)

wide_dat_inv_transect_avg <- wide_dat_inv_transect_avg %>%
  mutate(site_year = paste(Site, Year))

names(wide_dat_inv_transect_avg) <- str_replace_all(names(wide_dat_inv_transect_avg), c(" " = "_"))

comm_dat_inv_transect_avg <- wide_dat_inv_transect_avg %>%
  column_to_rownames(var = "site_year") %>%
  select(Banded_urchin:Top_Snail)

NMDS_comm_dat_inv_transect_avg <- metaMDS(comm_dat_inv_transect_avg,
  trymax = 200,
  distance = "bray",
  autotransform = F)

tibble_comm_dat_inv_transect_avg <- as_tibble((NMDS_comm_dat_inv_transect_avg$points),
  rownames = ("site_year"))

wide_dat_inv_transect_avg <- wide_dat_inv_transect_avg %>%
  left_join(tibble_comm_dat_inv_transect_avg)

inv_spp_fit <- envfit(NMDS_comm_dat_inv_transect_avg, wide_dat_inv_transect_avg, permutations = 999)
site_scores_inv <- wide_dat_inv_transect_avg %>%
  select(Year, Site, site_year, Human_Impact, MDS1, MDS2) %>%
  column_to_rownames(var = "site_year")
spp_scores_inv <- as.data.frame(scores(inv_spp_fit, display = "vectors"))
spp_scores_inv <- cbind(spp_scores_inv, Species = rownames(spp_scores_inv))
spp_scores_inv <- cbind(spp_scores_inv, pval = inv_spp_fit$vectors$pvals)

sig_spp_scores_inv <- subset(spp_scores_inv, pval <= 0.05)
sig_spp_scores_inv <- sig_spp_scores_inv %>%
  filter(!Species %in% "MDS1") %>%
  filter(!Species %in% "MDS2") %>%
  filter(!Species %in% "Year") %>%
  filter(!Species %in% "Latitude") %>%
  filter(!Species %in% "Longitude")

hulls_inv <- site_scores_inv %>%
  group_by(Human_Impact) %>%
  slice(chull(MDS1, MDS2))

```

```

plot_wide_dat_inv_transect_avg_site <- ggplot(wide_dat_inv_transect_avg,
      aes(MDS1, MDS2,
            fill = Site)) +
  geom_text((aes(label = Year)), vjust = -1.1, hjust = .5, size = 3) +
  geom_point(aes(shape = Site), size = 4) +
  ggtitle("Invertebrate Assemblage by Site") +
  theme_classic() +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  geom_text(x = 1.1, y = 0.35, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_inv_transect_avg$stress, 2))) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.ticks = element_blank(),
        axis.text = element_blank())

plot_wide_dat_inv_transect_avg_site

```

Invertebrate Assemblage by Site

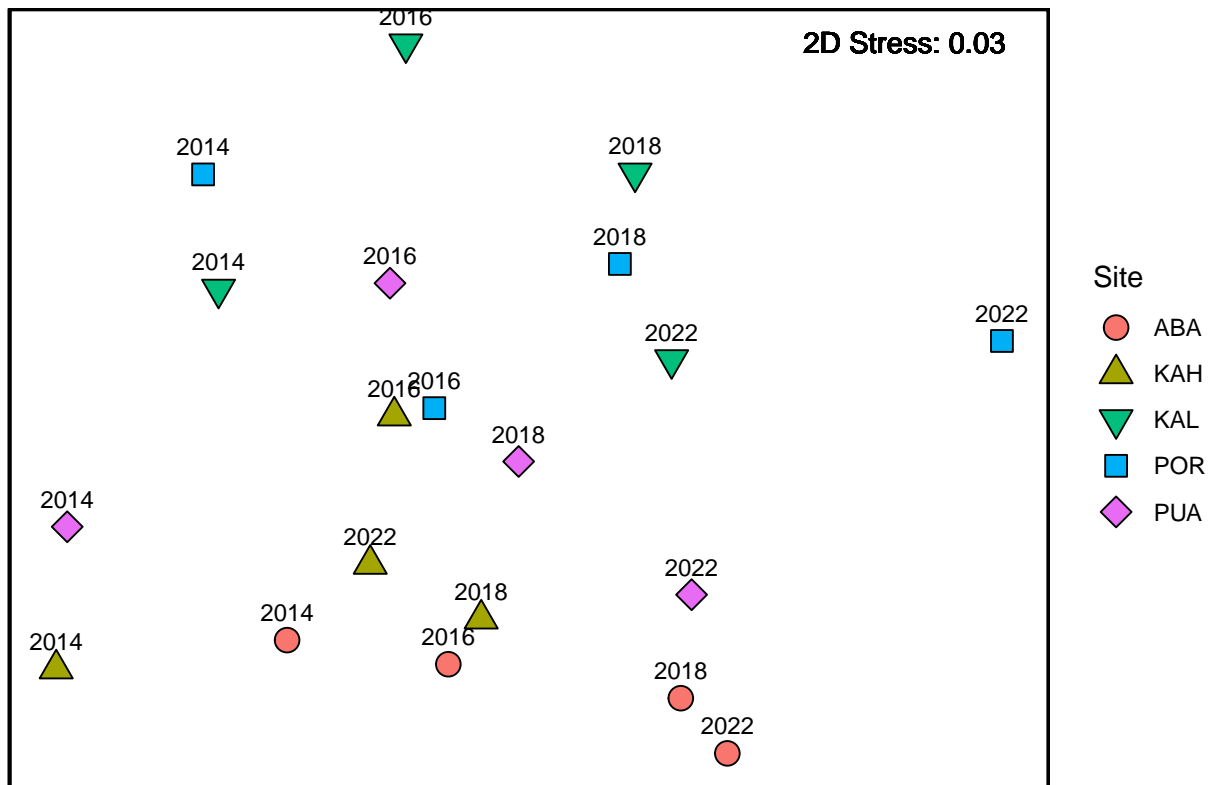


Figure 5 caption ----

#

```

plot_wide_dat_inv_transect_avg_impact <- ggplot(wide_dat_inv_transect_avg,
      aes(MDS1, MDS2,
            shape = Site)) +
  geom_text((aes(label = Year)), vjust = -1.0, hjust = 0.5, size = 2.5) +

```

```

geom_point(aes(fill = Human_Impact, shape = Site), size = 3) +
scale_fill_manual(values = c('High' = "firebrick4", 'Low' ="dodgerblue2")) +
scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
ggtitle("Invertebrate Assemblage with Human Impact") +
theme_classic() +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  geom_text(x = 1.1, y = 0.35, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_inv_transect_avg$stress, 2)))) +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid")) +
  geom_polygon(data = hulls_inv,
    aes(x = MDS1, y = MDS2,
      fill = as_factor(Human_Impact),
      group = as_factor(Human_Impact)), alpha = 0.2) +
guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
theme(axis.title.x = element_blank(),
  axis.title.y = element_blank(),
  axis.ticks = element_blank(),
  axis.text = element_blank())

```

plot_wide_dat_inv_transect_avg_impact

Invertebrate Assemblage with Human Impact

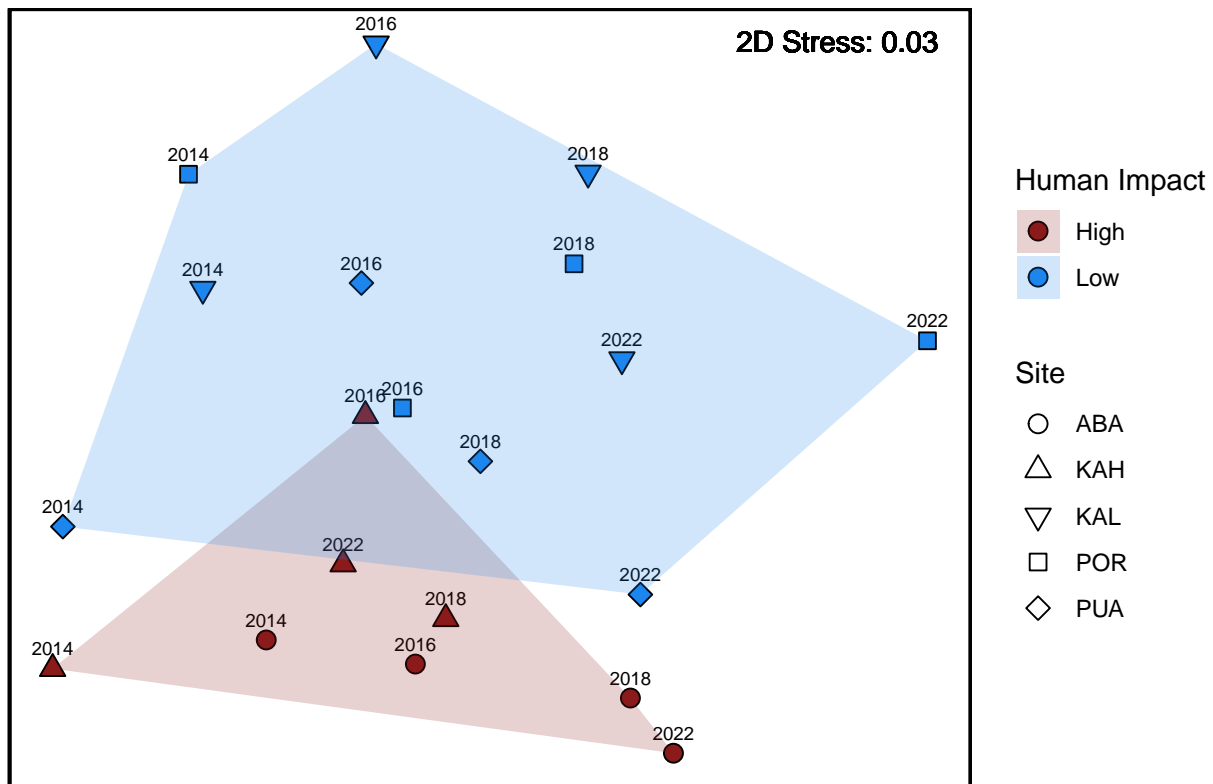


Figure 6 caption ----

#

```

plot_NMDS_comm_inv_wide_transect_avg <- ggplot(site_scores_inv, aes(x = MDS1, y = MDS2)) +
  geom_point(aes(fill = Site, shape = Site), size = 3) +
  # scale_fill_manual(values = c('High' = "firebrick4", 'Low' = "dodgerblue2")) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  ggtitle("Invertebrate Assemblage by Site") +
  geom_text(aes(label = Year), vjust = 2, hjust = .2, size = 2.5) +
  geom_text(x = 1.1, y = 0.35, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_inv_transect_avg$stress, 2))),
  labs(shape = "Site") +
  theme_classic()+
  # guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.position = "right",
        legend.text = element_text(size = 12),
        legend.title = element_text(size = 12))

#plot_NMDS_comm_inv_wide_transect_avg

plot_inv_spp_vectors <- plot_NMDS_comm_inv_wide_transect_avg +
  geom_segment(data = sig_spp_scores_inv, aes(x = 0, xend = NMDS1, y = 0, yend = NMDS2),
    arrow = arrow(length = unit(.5, "cm")),
    colour = "grey10", lwd = 0.5) +
  ggrepel::geom_text_repel(data = sig_spp_scores_inv,
    aes(x = NMDS1, y = NMDS2, label = Species), direction = "both", segment.size = 1)

plot_inv_spp_vectors

```

Invertebrate Assemblage by Site

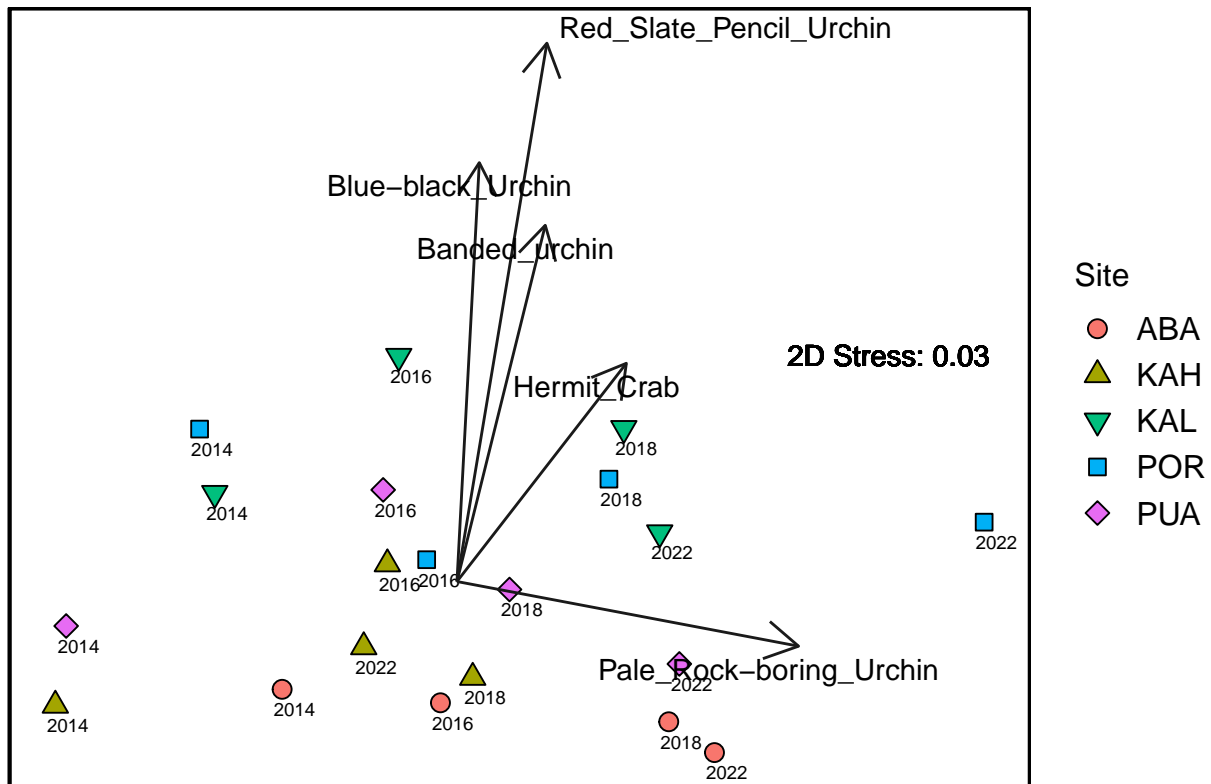


Figure 7 caption ----

#

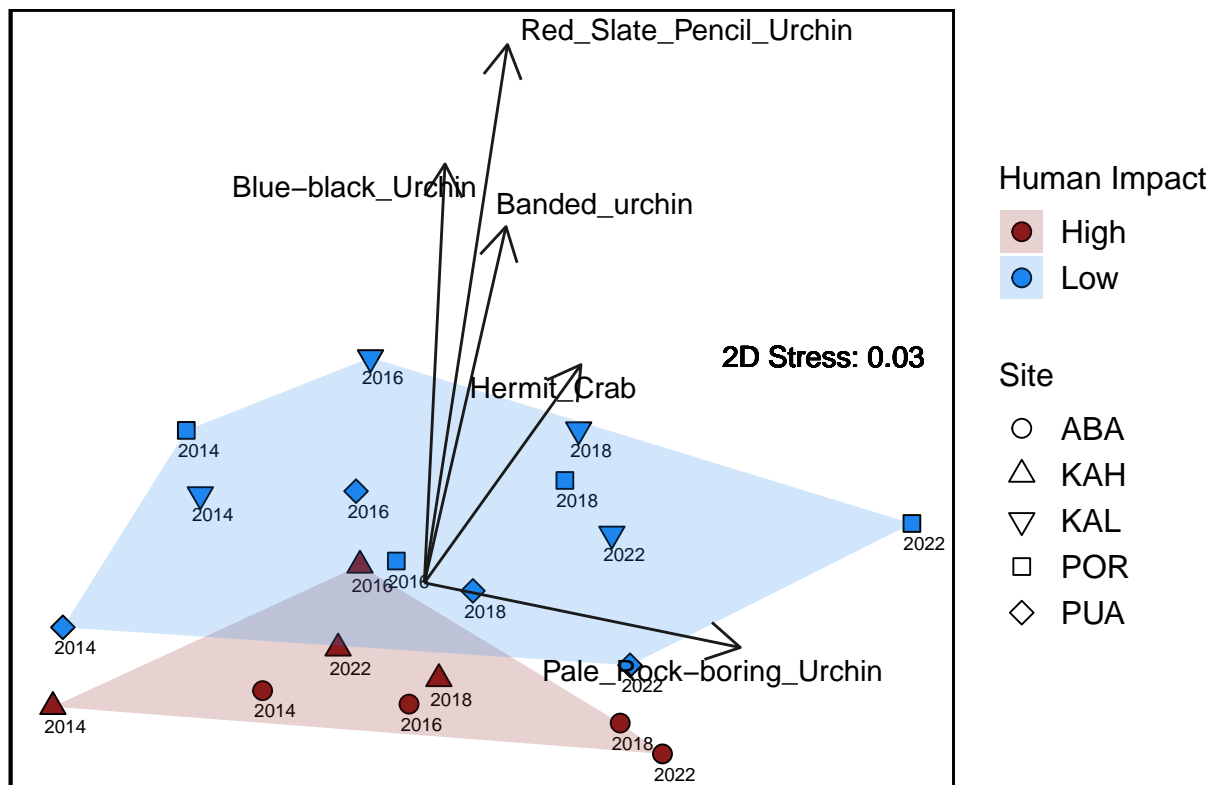
```
plot_NMDS_comm_inv_wide_transect_avg <- ggplot(site_scores_inv, aes(x = MDS1, y = MDS2)) +
  geom_point(aes(fill = Human_Impact, shape = Site), size = 3) +
  scale_fill_manual(values = c('High' = "firebrick4", 'Low' = "dodgerblue2")) +
  scale_shape_manual(values = c(21, 24, 25, 22, 23)) +
  ggtitle("Invertebrate Assemblage with Human Impact") +
  geom_text(aes(label = Year), vjust = 2, hjust = .2, size = 2.5) +
  geom_text(x = 1.1, y = 0.35, label = c(paste("2D Stress:" ,round(NMDS_comm_dat_inv_transect_avg$stress, 2)))) +
  labs(colour = "Human Impact", shape = "Site") +
  theme_classic() +
  geom_polygon(data = hulls_inv,
    aes(x = MDS1, y = MDS2,
      fill = as_factor(Human_Impact),
      group = as_factor(Human_Impact)), alpha = 0.2) +
  guides(fill= guide_legend("Human Impact", override.aes = list(shape = 21))) +
  theme(panel.background = element_rect(fill = NA, colour = "black", size = 1, linetype = "solid"),
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.ticks = element_blank(),
    axis.text = element_blank(),
    legend.position = "right",
    legend.text = element_text(size = 12),
    legend.title = element_text(size = 12))
```

```
#plot_NMDS_comm_inv_wide_transect_avg
```

```
plot_inv_spp_vectors <- plot_NMDS_comm_inv_wide_transect_avg +
  geom_segment(data = sig_spp_scores_inv, aes(x = 0, xend = NMDS1, y = 0, yend = NMDS2),
    arrow = arrow(length = unit(.5, "cm")),
    colour = "grey10", lwd = 0.5) +
  ggrepel::geom_text_repel(data = sig_spp_scores_inv,
    aes(x = NMDS1, y = NMDS2, label = Species), direction = "both", segment.size
```

plot_inv_spp_vectors

Invertebrate Assemblage with Human Impact



```
#### Figure 8 caption ----
```

```
#
```