

Extinction risk from climate change: ecosystem

Mar. 25th, 2024

Load libraries and data

Test for effect of habitat

Here I tested if extinction predictions change based on habitat, including terrestrial, marine, freshwater, islands, and mountains. I predicted that extinction risks would increase on islands and mountains, and possibly freshwater also.

```
#create model matrix for coefficients
data.use$habs <- ifelse(data.use$Region == "Marine", "Marine", ifelse(data.use$
Fresh == "Y", "Fresh", "Terrestrial"))

betamat <- data.frame(
  Intercept = rep(1, N),
  Mtn = ifelse(Mtn == "Y", 1, 0),
  Island = ifelse(Island == "Y", 1, 0),
  Terrestrial = ifelse(data.use$habs == "Terrestrial", 1, 0),
  Marine = ifelse(data.use$habs == "Marine", 1, 0),
  Freshwater = ifelse(data.use$habs == "Fresh", 1, 0))
#
stan.data<-list(N = N, percent = data.use$percent2, betamat = betamat, phi =
phi, S = n.Study, P = ncol(betamat), Study = Studyint)
params.to.monitor=c("beta", "y_rep", "stu", "sigma_stu", "eta", "log_lik")
#
init.beta=rep(0, ncol(betamat)-1)
init.fn<- function (chain_id) {
  list(beta = c(-2.5, init.beta))
}
# mod=stan(file="MetaRisk2 RSTAN beta mat.stan", data=stan.data, pars=params.to
.monitor,
#          chains = 3, warmup=7000, cores=3, iter=10000,
#          init = init.fn, save_warmup = FALSE, control=list(adapt_delta = 0.
9, max_treedepth = 15))#
#
load("2allhabs.rds")
params.to.monitor2=c("beta")#
sumx = summary(mod, probs=c(.025, 0.975), digits=4, pars=params.to.monitor2)
sumx$summary

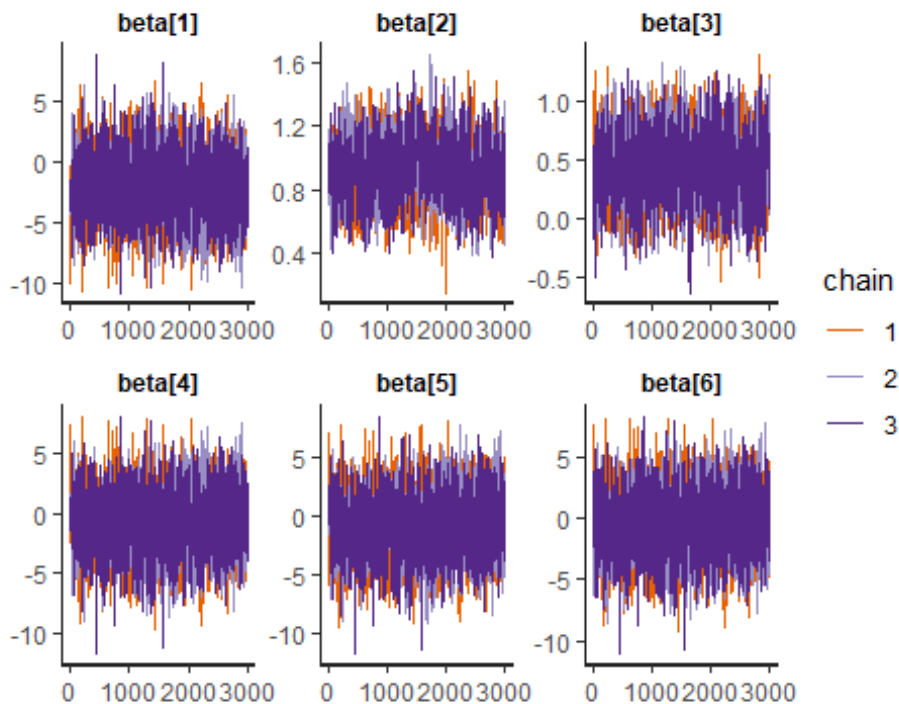
##              mean      se_mean      sd      2.5%      97.5%      n_eff
Rhat
## beta[1] -1.9250455 0.049327814 2.5323442 -6.8509049 3.0802876 2635.492 1.0
```

```

00263
## beta[2]  0.9286076 0.005515180 0.1832946  0.5711938 1.2977893 1104.537 1.0
02544
## beta[3]  0.4399901 0.006983944 0.2744104 -0.1024399 0.9813104 1543.831 1.0
00573
## beta[4] -0.8283834 0.049446233 2.5344846 -5.8520490 4.1154775 2627.319 1.0
00210
## beta[5] -0.8220796 0.049688843 2.5460994 -5.8161981 4.1158819 2625.627 1.0
00406
## beta[6] -0.4117927 0.049507821 2.5345511 -5.4171638 4.5472902 2620.924 1.0
00126

#
#checks
traceplot(mod,pars=params.to.monitor2,inc_warmup=FALSE)

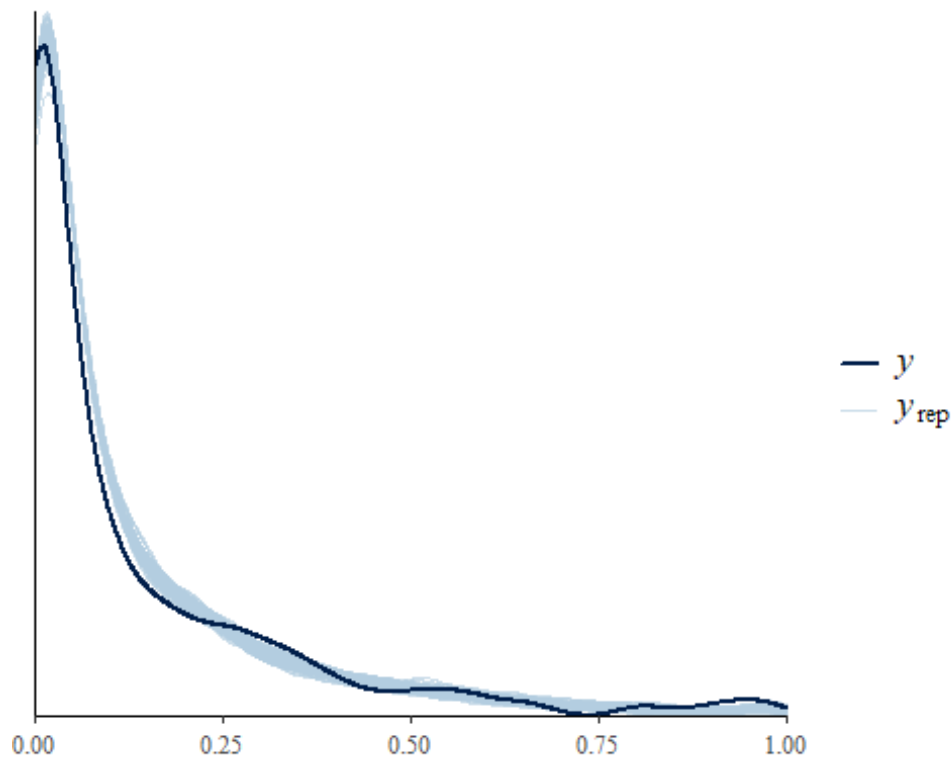
```



```

#
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)

```



```
#calculate loo
# log_lik_1 <- extract_log_lik(mod, merge_chains = FALSE)
# r_eff <- relative_eff(exp(log_lik_1), cores = 6)
# loo.mod <- loo(log_lik_1, r_eff = r_eff, cores = 6)
loo.mod #
```

```
##
## Computed from 9000 by 3235 log-likelihood matrix.
##
##           Estimate      SE
## elpd_loo    7613.1   96.6
## p_loo       1985.2   24.9
## looic      -15226.3 193.2
## -----
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                Count Pct.    Min. ESS
## (-Inf, 0.7]  (good)   1412  43.6%   117
##  (0.7, 1]   (bad)    1584  49.0%   <NA>
##  (1, Inf)   (very bad)  239   7.4%   <NA>
## See help('pareto-k-diagnostic') for details.
```

```
#create data frame of looics from two models
load("2allhabs.rds")
```

```
table.data<-data.frame(
```

```

Model = c("Intercept-only model", "Model including habitat"),
LOOic = c(loo.int$estimates[3], loo.mod$estimates[3]),
SE = c(loo.int$estimates[6], loo.mod$estimates[6])
)
knitr::kable(table.data, caption = "Table 1: Comparisons of LOOic between baseline and all-habitat model", format = "markdown")

```

Table 1: Comparisons of LOOic between baseline and all-habitat model

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including habitat	-15226.29	193.2120

```

Looic.diff = loo.mod$estimates[3] - loo.int$estimates[3]
cat("Difference in LOOic =", Looic.diff)

## Difference in LOOic = -12.52453

```

Three habitats: Marine, terrestrial and freshwater

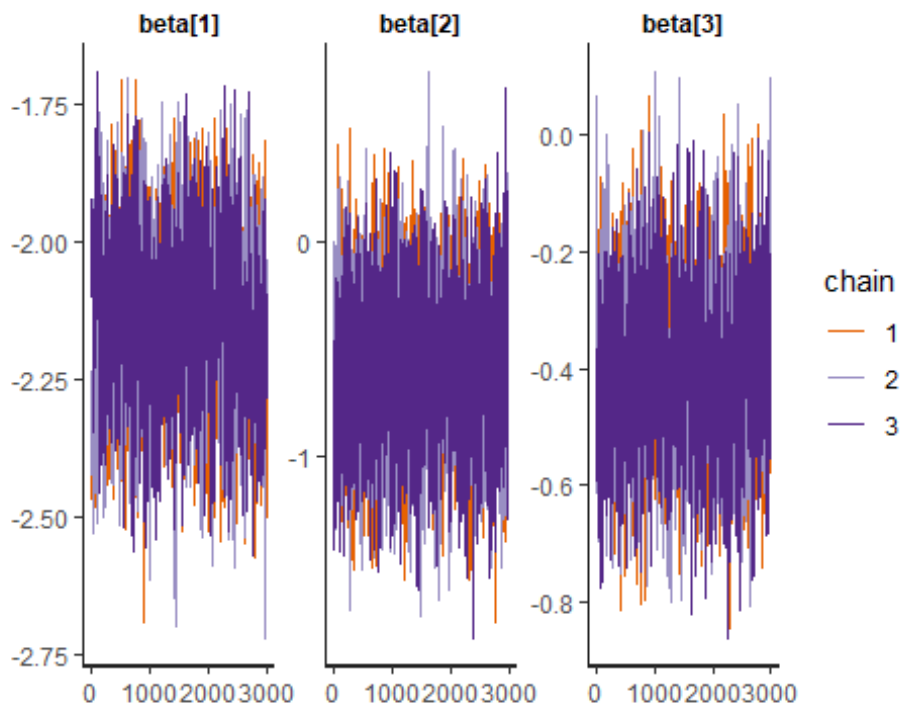
```

#Marine is also a part of region analysis, but want to explore it as a habitat and keep region based on continent.
data.use$habs <- ifelse(data.use$Region == "Marine", "Marine", ifelse(data.use$Fresh == "Y", "Fresh", "Terrestrial"))
betamat <- model.matrix(~data.use$habs)
#
#Need to do each separately, because habitats overlap, e.g., freshwater on islands
#betamat <- model.matrix(~Mtn) #modeled as yes, no, some
#betamat <- model.matrix(~Island) #modeled as yes, no, some
#
stan.data<-list(N = N, percent = data.use$percent2, betamat = betamat, phi = phi, S = n.Study, P = ncol(betamat), Study = Studyint)
params.to.monitor=c("beta", "y_rep", "stu", "sigma_stu", "eta", "log_lik")
#
init.beta=rep(0, ncol(betamat)-1)
init.fn<- function(chain_id) {
  list(beta = c(-2.5, init.beta))
}
# mod=stan(file="MetaRisk2 RSTAN betamat.stan", data=stan.data, pars=params.to.monitor,
#          chains = 3, warmup=7000, cores=7, iter=10000,
#          init = init.fn, control=list(adapt_delta = 0.9, max_treedepth = 15))
#
load("2habsL2.rds")
params.to.monitor2=c("beta")#
sumx = summary(mod, probs=c(.025, 0.975), digits=4, pars=params.to.monitor2)
sumx$summary

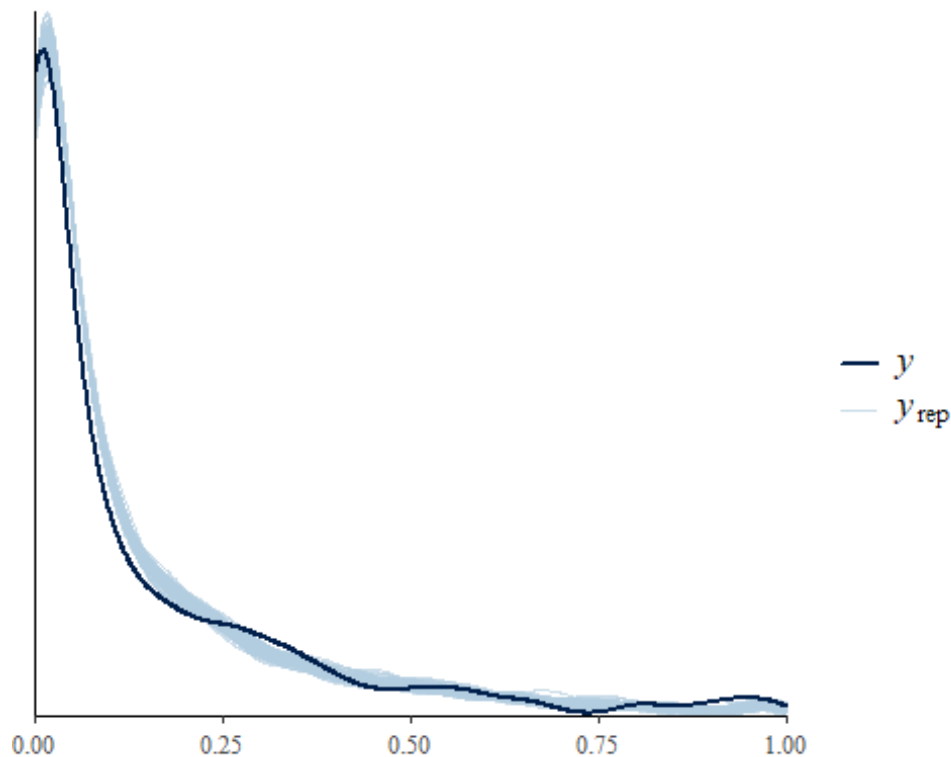
```

```
##               mean      se_mean      sd      2.5%      97.5%      n_eff
## beta[1] -2.1466812 0.004442924 0.1406338 -2.4290611 -1.87723932 1001.940
## beta[2] -0.5921805 0.008801843 0.3401566 -1.2504375  0.06831581 1493.519
## beta[3] -0.3933419 0.004232193 0.1374896 -0.6640107 -0.12989104 1055.380
##               Rhat
## beta[1] 1.003050
## beta[2] 1.002641
## beta[3] 1.004113

#
#checks
traceplot(mod, pars=params.to.monitor2, inc_warmup=FALSE)
```



```
#
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```



```
#calculate loo
# log_lik_1 <- extract_log_lik(mod, merge_chains = FALSE)
# r_eff <- relative_eff(exp(log_lik_1), cores = 6)
# loo.mod <- loo(log_lik_1, r_eff = r_eff, cores = 6)
loo.mod #
```

```
##
## Computed from 9000 by 3235 log-likelihood matrix.
##
##           Estimate      SE
## elpd_loo    7603.5   96.7
## p_loo       1994.4   25.0
## looic      -15206.9 193.5
## -----
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##               Count Pct.    Min. ESS
## (-Inf, 0.7] (good)   1381 42.7%   124
## (0.7, 1]   (bad)    1588 49.1%   <NA>
## (1, Inf)   (very bad) 266  8.2%   <NA>
## See help('pareto-k-diagnostic') for details.
```

```
#create data frame of looics from two models
```

```
load("2habsL2.rds")
```

```
#
```

```
table.data<-data.frame(
```

```

Model = c("Intercept-only model", "Model including habitat"),
LOOic = c(loo.int$estimates[3], loo.mod$estimates[3]),
SE = c(loo.int$estimates[6], loo.mod$estimates[6])
)
knitr::kable(table.data, caption = "Table 2: Comparisons of LOOic between baseline and model with three habitats", format = "markdown")

```

Table 2: Comparisons of LOOic between baseline and model with three habitats

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including habitat	-15206.94	193.4666

```

Looic.diff = loo.mod$estimates[3] - loo.int$estimates[3]
cat("Difference in LOOic =", Looic.diff)

## Difference in LOOic = 6.824013

#create data frame of looics from two models
data.use<-dataP; attach(data.use)

## The following object is masked _by_ .GlobalEnv:
##
##   percent2

## The following objects are masked from data.use (pos = 3):
##
##   Adaptation, adj.percent, Antarctic, Arctic, Author, ave.percent,
##   Climate.Mod, concat, Demography.LH, Disp.Mod, Dispersal, Earth.Sys,
##   Endemic, Fresh, Island, Land.Use.Change, max.percent, min.percent,
##   Model.Type, Mtn, N.Middle, Non.clim.threat, Other, Other.Habitat,
##   percent2, Physiology, Pop.diff, Pre.Ind.Rise, Region, S.Middle,
##   Scenario, Sp.int, Study, Taxa, Threatened, Time, Total.N, Tropics,
##   version, WtSp, Year, Year.Pred

## The following objects are masked from data.use (pos = 4):
##
##   Adaptation, adj.percent, Antarctic, Arctic, Author, ave.percent,
##   Climate.Mod, concat, Demography.LH, Disp.Mod, Dispersal, Earth.Sys,
##   Endemic, Fresh, Island, Land.Use.Change, max.percent, min.percent,
##   Model.Type, Mtn, N.Middle, Non.clim.threat, Other, Other.Habitat,
##   Physiology, Pop.diff, Pre.Ind.Rise, Region, S.Middle, Scenario,
##   Sp.int, Study, Taxa, Threatened, Time, Total.N, Tropics, version,
##   WtSp, Year, Year.Pred

## The following objects are masked from dataP:
##
##   Adaptation, adj.percent, Antarctic, Arctic, Author, ave.percent,
##   Climate.Mod, concat, Demography.LH, Disp.Mod, Dispersal, Earth.Sys,

```

```

##      Endemic, Fresh, Island, Land.Use.Change, max.percent, min.percent,
##      Model.Type, Mtn, N.Middle, Non.clim.threat, Other, Other.Habitat,
##      Physiology, Pop.diff, Pre.Ind.Rise, Region, S.Middle, Scenario,
##      Sp.int, Study, Taxa, Threatened, Time, Total.N, Tropics, version,
##      WtSp, Year, Year.Pred

## The following object is masked from package:base:
##
##      version

n.total <- nrow(data.use) #total N
#
#Load("beta mtn.rds")
load("2mtn.rds")
mod.mtn <- mod
posterior <- as.data.frame(mod.mtn)
p.y <- sum(data.use$Mtn == "Y")/n.total #prop of yes
p.s <- sum(data.use$Mtn == "S")/n.total #prop of some
p.n <- sum(data.use$Mtn == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[3]"]])
beta.S = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.s * beta.S + p.y * beta.Y
r.mean.1 = invlogit(grand.mean)
beta.vs.r.mean.n = (beta.N) - (r.mean.1)
beta.vs.r.mean.s = (beta.S) - (r.mean.1)
beta.vs.r.mean.y = (beta.Y) - (r.mean.1)
#
param.vals.1 <- data.frame(
  "r.mean.1" = r.mean.1,
  "beta.N" = beta.N,
  "beta.S" = beta.S,
  "beta.Y" = beta.Y,
  "beta.vs.r.mean.n" = beta.vs.r.mean.n,
  "beta.vs.r.mean.s" = beta.vs.r.mean.s,
  "beta.vs.r.mean.y" = beta.vs.r.mean.y
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975), na.rm
=TRUE))
#
hab.vals <- data.frame(beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.S.lo = pred.1[1,3],
  beta.S = pred.1[2,3],
  beta.S.hi = pred.1[3,3],
  beta.Y.lo = pred.1[1,4],
  beta.Y = pred.1[2,4],
  beta.Y.hi = pred.1[3,4],
  beta.N.mean.lo = pred.1[1,5],

```



```

beta.N.mean = pred.1[2,5],
beta.N.mean.hi = pred.1[3,5],
beta.S.mean.lo = pred.1[1,6],
beta.S.mean = pred.1[2,6],
beta.S.mean.hi = pred.1[3,6],
beta.Y.mean.lo = pred.1[1,7],
beta.Y.mean = pred.1[2,7],
beta.Y.mean.hi = pred.1[3,7]
)
#*****
*****
load("3islandAUS.rds")
mod.isl <- mod
params.to.monitor2=c("beta")#
sumx = summary(mod.isl,probs=c(.025,0.975), digits=4, pars=params.to.monitor2
)
sumx$summary

##              mean      se_mean      sd      2.5%      97.5%      n_eff
## beta[1] -2.4183376 0.003588374 0.09794597 -2.61207572 -2.2281731 745.0368
## beta[2] -0.5568152 0.005931125 0.15261105 -0.86132078 -0.2623271 662.0607
## beta[3]  0.4281474 0.007098608 0.16309201  0.09861945  0.7540473 527.8604
##              Rhat
## beta[1] 1.002298
## beta[2] 1.001871
## beta[3] 1.003018

invlogit(c(-2.34,-2.34-.6,-2.34+.17))

## [1] 0.08786391 0.05021127 0.10247703

posterior <- as.data.frame(mod.isl)
p.y <- sum(Island == "Y")/n.total #prop of yes
p.s <- sum(Island == "S")/n.total #prop of some
p.n <- sum(Island == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[3]"]])
beta.S = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.s * beta.S + p.y * beta.Y
beta.vs.r.mean.n = (beta.N) - (r.mean.1)
beta.vs.r.mean.s = (beta.S) - (r.mean.1)
beta.vs.r.mean.y = (beta.Y) - (r.mean.1)
#
param.vals.1 <- data.frame(
  "r.mean.1" = r.mean.1,
  "beta.N" = beta.N,
  "beta.S" = beta.S,
  "beta.Y" = beta.Y,
  "beta.vs.r.mean.n" = beta.vs.r.mean.n,
  "beta.vs.r.mean.s" = beta.vs.r.mean.s,
  "beta.vs.r.mean.y" = beta.vs.r.mean.y

```

```

)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975), na.rm
=TRUE))
#
hab.vals[2,] <- cbind(beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.S.lo = pred.1[1,3],
  beta.S = pred.1[2,3],
  beta.S.hi = pred.1[3,3],
  beta.Y.lo = pred.1[1,4],
  beta.Y = pred.1[2,4],
  beta.Y.hi = pred.1[3,4],
  beta.N.mean.lo = pred.1[1,5],
  beta.N.mean = pred.1[2,5],
  beta.N.mean.hi = pred.1[3,5],
  beta.S.mean.lo = pred.1[1,6],
  beta.S.mean = pred.1[2,6],
  beta.S.mean.hi = pred.1[3,6],
  beta.Y.mean.lo = pred.1[1,7],
  beta.Y.mean = pred.1[2,7],
  beta.Y.mean.hi = pred.1[3,7]
)
#*****
*****
load("2habsL2.rds")
mod.habs <- mod
posterior <- as.data.frame(mod.habs)

#marine
p.m <- sum(Region == "Marine")/n.total #prop of yes
p.f <- sum(Fresh == "Y")/n.total #prop of yes
p.t <- (1 - p.m - p.f) #prop of yes

beta.m = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.f = invlogit(posterior[["beta[1]"]])
beta.t = invlogit(posterior[["beta[1]"]] + posterior[["beta[3]"]])

#r.mean.1 = p.m * beta.m + p.f * beta.f + p.t * beta.t
beta.vs.r.mean.m = (beta.m) - (r.mean.1)
beta.vs.r.mean.f = (beta.f) - (r.mean.1)
beta.vs.r.mean.t = (beta.t) - (r.mean.1)

param.vals.1 <- data.frame(
  "r.mean.1" = r.mean.1,
  "beta.m" = beta.m,
  "beta.f" = beta.f,
  "beta.t" = beta.t,
  "beta.vs.r.mean.m" = beta.vs.r.mean.m,

```

```

"beta.vs.r.mean.f" = beta.vs.r.mean.f,
"beta.vs.r.mean.t" = beta.vs.r.mean.t
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975), na.rm
=TRUE))
nas<-c(NA,NA,NA)

hab.vals[3:5,] <- cbind(
  beta.N.lo = nas,
  beta.N = nas,
  beta.N.hi = nas,
  beta.S.lo = nas,
  beta.S = nas,
  beta.S.hi = nas,
  beta.Y.lo = pred.1[1,2:4],
  beta.Y = pred.1[2,2:4],
  beta.Y.hi = pred.1[3,2:4],
  beta.N.mean.lo = nas,
  beta.N.mean = nas,
  beta.N.mean.hi = nas,
  beta.S.mean.lo = nas,
  beta.S.mean = nas,
  beta.S.mean.hi = nas,
  beta.Y.mean.lo = pred.1[1,5:7],
  beta.Y.mean = pred.1[2,5:7],
  beta.Y.mean.hi = pred.1[3,5:7]
)
#*****
*****
hab.vals$habitat = c("Mountains","Islands","Marine","Freshwater","Terrestrial
")

#show values
hab.vals.r <- hab.vals
hab.vals.r[,1:18] <- round((hab.vals.r[,1:18]), digits = 4)
opts <- options(knitr.kable.NA = "")
knitr::kable(hab.vals.r, caption = "Table 1: Extinction risk by taxa, with 95
% credible intervals", format = "markdown", font = 4, longtable = T)

```

Table 3: Extinction risk by taxa, with 95% credible intervals

bet a.N. lo	be ta. N	bet a.N. hi	bet a.S. lo	be ta. S	bet a.S. hi	bet a.Y. lo	be ta. Y	bet a.Y. hi	beta. N.me an.lo	beta. N.me an	beta. N.me an.hi	beta.S .mean .lo	beta. S.me an	beta.S .mean .hi	beta. Y.me an.lo	beta. Y.me an	beta. Y.me an.hi	habi tat
0.0 63 5	0. 08 59	0.1 14 1	0.0 49 6	0. 05 95	0.0 70 6	0.1 10 4	0. 14 79	0.1 93 4	- 0.014 9	0.01 01	0.039 2	- 0.031 3	- 0.01 64	- 0.001 7	0.033 0	0.07 20	0.119 2	Mou ntai ns
0.0 68 4	0. 08 19	0.0 97 2	0.0 38 6	0. 04 86	0.0 60 5	0.0 95 7	0. 12 03	0.1 50 0	- 0.011 6	0.00 59	0.024 5	- 0.041 7	- 0.02 73	- 0.011 8	0.017 5	0.04 41	0.075 0	Islan ds
						0.0 34 3	0. 06 07	0.1 06 8							- 0.044 2	- 0.01 52	0.031 9	Mari ne

bet a.N. lo	be ta. N	bet a.N. hi	bet a.S. lo	be ta. S	bet a.S. hi	bet a.Y. lo	be ta. Y	bet a.Y. hi	beta. N.me an.lo	beta. N.me an	beta. N.me an.hi	beta.S .mean .lo	beta. S.me an	beta.S .mean .hi	beta. Y.me an.lo	beta. Y.me an	beta. Y.me an.hi	habi tat
						0.0 81 0	0. 10 47	0.1 32 7							0.003 0	0.02 87	0.058 5	Fres hwa ter
						0.0 63 2	0. 07 31	0.0 84 3							- 0.017 1	- 0.00 27	0.011 7	Terr estri al

Terrestrial habitats have lower predicted extinction risks and mountain habitats have higher predicted extinction risks than other habitats.

```
#Load("Habitat analysis results.rds")
hab.vals$beta.Y.lo <- as.numeric(hab.vals$beta.Y.lo) # data is character and
needs to be numeric
hab.vals$beta.Y <- as.numeric(hab.vals$beta.Y)
hab.vals$beta.Y.hi <- as.numeric(hab.vals$beta.Y.hi)
hab.vals$beta.Y.mean.lo <- as.numeric(hab.vals$beta.Y.mean.lo)
hab.vals$beta.Y.mean <- as.numeric(hab.vals$beta.Y.mean)
hab.vals$beta.Y.mean.hi <- as.numeric(hab.vals$beta.Y.mean.hi)
hab.vals$habitat <- factor(hab.vals$habitat, levels = c("Terrestrial", "Fresh
water", "Marine", "Islands", "Mountains"))

#Figures
xinter <- pred.1[2,1]
Fig1a<-ggplot(data = hab.vals)+
  geom_vline(xintercept=xinter) +
  geom_errorbar(aes(y = habitat, xmin = beta.Y.lo, xmax = beta.Y.hi), width =
0) +
  geom_point(stat = "identity", aes(y = habitat, x = beta.Y), color = "#41678
8", size = 3, shape = 15) +
  xlab("Pre-industrial rise \n in temperature (C)") + xlim(c(0,.2)) +
  theme_classic()+
  #scale_y_discrete(labels = c("Mountains", "Islands", "Marine", "Freshwater", "T
errestrial")) +
  theme(axis.title.y = element_blank(),axis.title=element_text(size=14),axis.
text = element_text(size=12))+
  guides(size=F)

## Warning: The `scale` argument of `guides()` cannot be `FALSE`. Use "none
" instead as
## of ggplot2 3.3.4.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

Fig1a

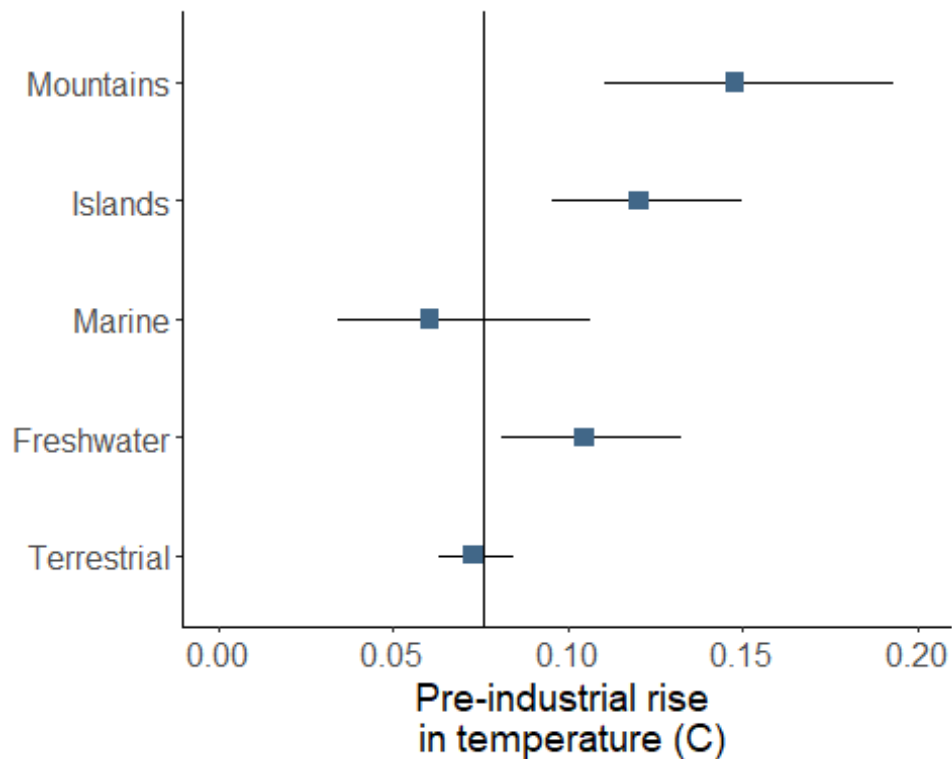


Fig. 1. Predicted extinction risk based on habitat.

#Differences from zero

```
glob.mean.over <- rep("n", nrow(hab.vals))
glob.mean.over[hab.vals$beta.Y.mean > 0 & hab.vals$beta.Y.mean.lo > 0] = "y"
glob.mean.over[hab.vals$beta.Y.mean < 0 & hab.vals$beta.Y.mean.hi < 0] = "y"
hab.vals$glob.mean.over <- glob.mean.over
```

```
Fig1b <- ggplot(data = hab.vals)+
  geom_vline(xintercept=0) +
  geom_errorbar(aes(y = habitat, xmin = beta.Y.mean.lo, xmax = beta.Y.mean.hi),
    width = 0) +
  geom_point(stat = "identity", aes(y = habitat, x = beta.Y.mean, color = glob.mean.over),
    size = 3, shape = 15) +
  scale_color_manual(values=c('grey', '#5B507A'))+
  xlab("Percent difference \n from global mean") + xlim(c(-.12, .12)) +
  theme_classic()+
  theme(axis.title.y = element_blank(), axis.ticks.y = element_blank(), axis.line.y = element_blank(),
    axis.text.y = element_blank(),
    axis.title = element_text(size=14), axis.text = element_text(size=12), legend.position = "none")+
  guides(size=F)
Fig1b
```

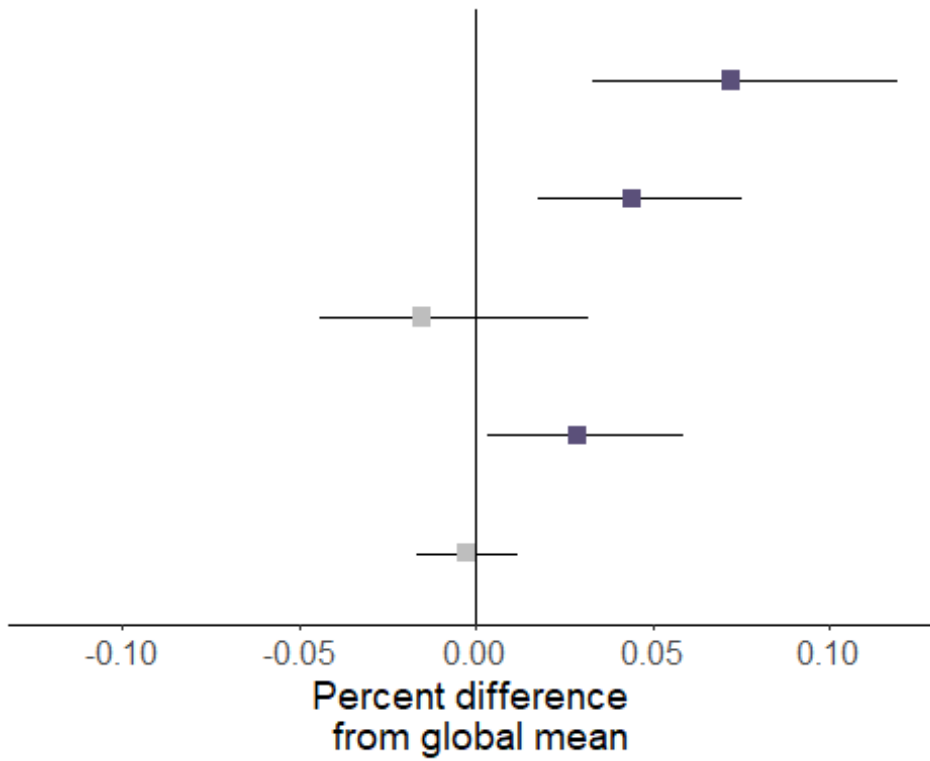


Fig. 2. Predicted extinction risk based on habitat.

```
ggarrange(Fig1a, NULL, Fig1b, ncol=3, widths = c(4,.4, 2))
```

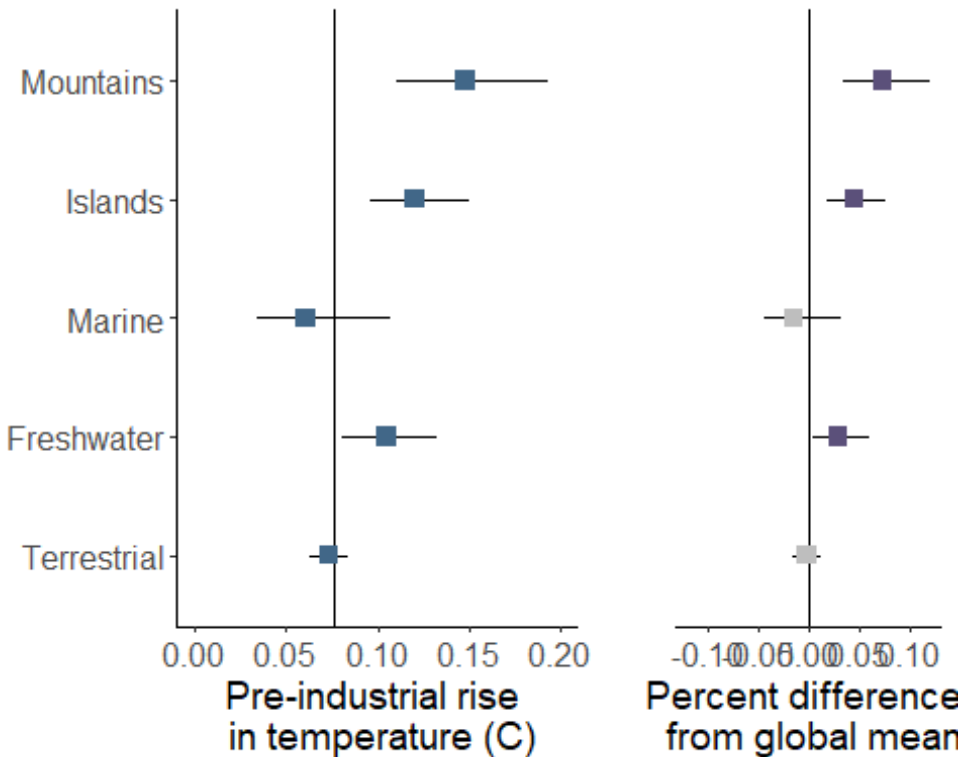


Fig. 3. Predicted extinction risk based on habitat.

```
ggsave("Metarisk2_habs.png",width=4,height=3,unit="in",dpi="print")

threat.YN<-data.frame(
  threat = c("Mountains - N","Mountains - Y","Islands - N","Islands - Y"),
  YN = rep(c("N","Y"),2),
  beta = c(hab.vals$beta.N[1],hab.vals$beta.Y[1],hab.vals$beta.N[2],hab.vals$beta.Y[2]),
  beta.lo = c(hab.vals$beta.N.lo[1],hab.vals$beta.Y.lo[1],hab.vals$beta.N.lo[2],hab.vals$beta.Y.lo[2]),
  beta.hi = c(hab.vals$beta.N.hi[1],hab.vals$beta.Y.hi[1],hab.vals$beta.N.hi[2],hab.vals$beta.Y.hi[2])
)

#set colors
boolColors <- as.character(c("N"="#416788", "Y"="#8cb369"))
boolScale <- scale_colour_manual(name="YN", values=boolColors)

#Figures
Fig2<-ggplot(data = threat.YN)+
  geom_vline(xintercept=grand.mean.pred[2]) +
  geom_errorbar(aes(y = threat, xmin = beta.lo, xmax = beta.hi), width = 0) +
  geom_point(stat = "identity", aes(y = threat, x = beta, color = YN), size = 3, shape = 15) +
  xlab("Pre-industrial rise \n in temperature (C)") + xlim(c(0.05,.20)) +
  theme_classic() + boolScale +
```

```

  theme(axis.title.y = element_blank(),axis.title=element_text(size=14),axis.
text = element_text(size=12))+
  guides(color = guide_legend(""))
Fig2

```

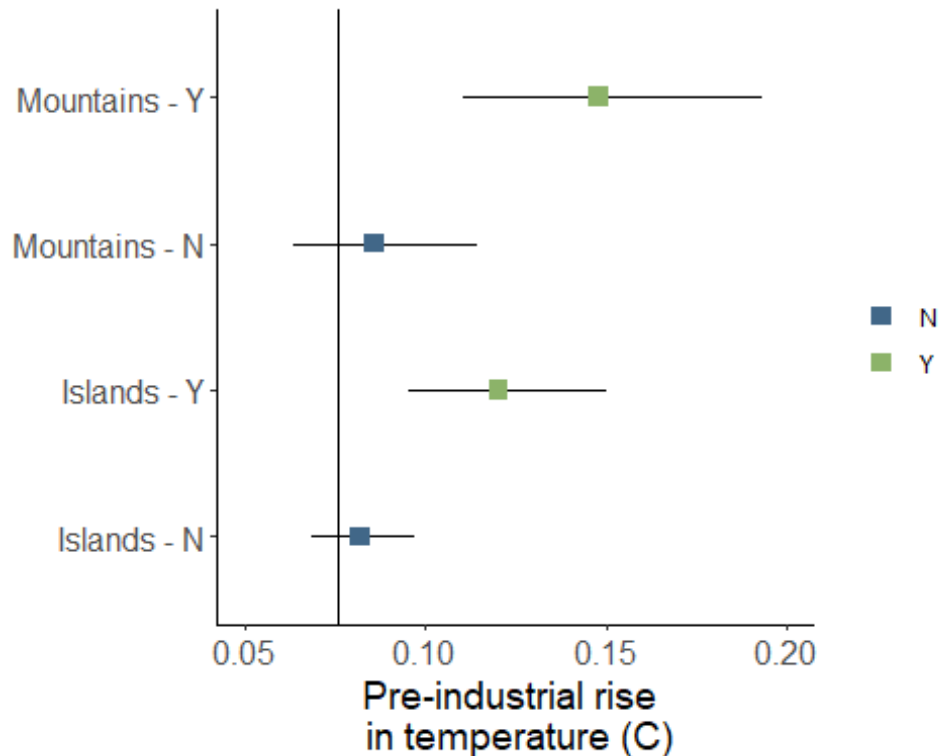


Fig. 4. Predicted extinction risk based on habitats.

```

ggsave("Metarisk2 hab YN.png",width=6,height=5,unit="in",dpi="print")

#catalog sample sizes
habs <- ifelse(data.use$Region == "Marine","Marine",ifelse(data.use$Fresh ==
"Y", "Fresh","Terrestrial"))
#
N.mtn.st<-length(unique(dataP$Study[dataP$Mtn == "Y"]))
N.mtn.mod<-length((dataP$Study[dataP$Mtn == "Y"]))

N.isl.st<-length(unique(dataP$Study[dataP$Island == "Y" | dataP$Region == "Au
straliaNewZealand"]))
N.isl.mod<-length((dataP$Study[dataP$Island == "Y" | dataP$Region == "Austral
iaNewZealand"])))

N.mar.st<-length(unique(dataP$Study[dataP$Region == "Marine"]))
N.mar.mod<-length((dataP$Study[dataP$Region == "Marine"]))

N.fresh.st<-length(unique(dataP$Study[dataP$Fresh == "Y"]))
N.fresh.mod<-length((dataP$Study[dataP$Fresh == "Y"]))

```



```

N.terr.st<-length(unique(dataP$Study[habs == "Terrestrial"]))
N.terr.mod<-length((dataP$Study[habs == "Terrestrial"]))

table.data<-data.frame(
  Factor = c("Mountain","Island","Marine","Freshwater","Terrestrial"),
  Studies = c(N.mtn.st,N.isl.st,N.mar.st,N.fresh.st,N.terr.st),
  Models = c(N.mtn.mod,N.isl.mod,N.mar.mod,N.fresh.mod,N.terr.mod)
)
knitr::kable(table.data, caption = "Table 3: Number of studies and models for
each factor", format = "markdown")

```

Table 3: Number of studies and models for each factor

Factor	Studies	Models
Mountain	89	628
Island	79	498
Marine	30	116
Freshwater	72	292
Terrestrial	402	2827

Conclusion

Habitats vary in extinction risk, with lower risks for species in terrestrial habitats and higher risks for species in mountainous habitats.

Variation explained

```

#After Gelman 2019 R2 for Bayesian
#
load("3allhabsAUS.rds") #with Australia as island
posterior=as.data.frame(mod);

#create model matrix for coefficients
data.use$habs <- ifelse(data.use$Region == "Marine","Marine",ifelse(data.use$
Fresh == "Y", "Fresh","Terrestrial"))
betamat <- data.frame(
  Intercept = rep(1,N),
  Mtn = ifelse(data.use$Mtn == "Y",1,0),
  Island = ifelse(data.use$Island == "Y" | data.use$Region == "AustraliaNewZe
aland",1,0),
  Terrestrial = ifelse(data.use$habs == "Terrestrial",1,0),
  Marine = ifelse(data.use$habs == "Marine",1,0),
  Freshwater = ifelse(data.use$habs == "Fresh",1,0))

betamat <- as.matrix(betamat)
#Variables and matrices
S = 9000; #samples
K = ncol(betamat); #factors
p.mat <- as.matrix(posterior[,1:K])

```

```

y = dataP$percent2
y.mat = t(matrix(rep(y,S), nrow = N, ncol = S))
y.mean <- mean(y)

#Calculate y.pred for fixed effects only
y.pred <- matrix(rep(NA, N*S), nrow = S, ncol = N)
theta <- y.pred
for (i in 1:N) {
  theta[,i] = invlogit(p.mat %*% betamat[i,])#rows = samples, cols = i
  y.pred[,i] = (theta[,i] * data.use$Total.N[i])/(theta[,i] * data.use$Total.N[i] + (1-theta[,i]) * data.use$Total.N[i])
}

#Calcluate residual variance
res.f = y.mat - y.pred
RSS.f = rowSums((res.f)^2)
res.v.f = 1/(N -1) * RSS.f

#Calculate fit variance
pred.v.f = 1/(N-1) * rowSums((y.pred)^2)

#Calculate R2
R2.v.f = pred.v.f/(pred.v.f + res.v.f)
cat("fixed effects R2 = ", quantile(R2.v.f,probs = c(0.025, 0.5, 0.975),na.rm = T))

## fixed effects R2 = 0.09545623 0.1286327 0.1715944

#Total model With random effects
y.pred.c <- (as.matrix(posterior[, (K+1):(N+K)])) #calculated in STAN, with all RE and weightings

#Calculate residual variance
res.c = y.mat - y.pred.c
RSS.c = rowSums((res.c)^2)
res.v.c = 1/(N-1) * RSS.c

#Calculate fit variance

pred.v.c = 1/(N-1) * rowSums(y.pred.c^2)

#Calculate full model R2
R2.v.c = pred.v.c/(pred.v.c + res.v.c)
print(" ")
## [1] " "

cat("Overall model R2 = ", quantile(R2.v.c,probs = c(0.025, 0.5, 0.975),na.rm = T))

```

```
## Overall model R2 = 0.7646649 0.7859718 0.8058091
```