Extinction risk from climate change: latitude

March 15, 2024

Load libraries and data

```r
rm(list = ls())
 root.dir = "C:/Users/mcu08001/Documents/1New Research/CC MetaRisk2/Analysis"

#load libraries
library(coda); library(ggplot2); library(rstan); library(bayesplot); library(
shinystan); library(loo); library(rstanarm); library(dplyr); library(ggpubr)
options(mc.cores = parallel::detectCores())
rstan_options(auto_write = FALSE)

#load data
dataP<-read.table("MetaRisk2 aggthres 5.txt",header=T); attach(dataP)

#other quantities
#betareg requires no 0s or 1s
koffset = 0.001 #the k that gives the best posterior predictive check
percent2 <- adj.percent
percent2[adj.percent == 0] = koffset;
percent2[adj.percent == 1] = 1 - koffset;
dataP$percent2 <- percent2;
data.use<-dataP

# Other quantities
N = length(data.use$percent2)
n.Study <- length(unique(data.use$Study)) #number of studies
Studyint<-as.integer(unclass(factor(data.use$Study)))
phi = data.use$Total.N

length(unique(dataP$Study))

## [1] 485
```

Test for effect of latitudinal band

Here I tested if extinction predictions change based on latitudinal band, represented by Antarctic, southern middle latitudes, tropics, northern middle latitudes, and Arctic regions. I predict that tropical regions will be characterized by the highest extinction risk.

```r
betamat <- (model.matrix(~Antarctic + S.Middle + Tropics + N.Middle + Arctic,
data=data.use, contrasts.arg=list(Antarctic = "contr.sum", S.Middle = "contr.
sum", Tropics="contr.sum", N.Middle = "contr.sum", Arctic = "contr.sum"))) #r
elative to global mean

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("2all_lat.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] -2.52181901 0.006836729 0.20764212 -2.92864266 -2.1129407 922.4315
## beta[2] -0.19524643 0.006944569 0.21558734 -0.60422639  0.2429012 963.7312
## beta[3]  0.00651311 0.003138724 0.08904241 -0.16866335  0.1765735 804.7994
## beta[4]  0.19330411 0.002553925 0.07695949  0.04367509  0.3421246 908.0467
## beta[5]  0.20565564 0.002981691 0.08842719  0.03100652  0.3777710 879.5214
## beta[6]  0.43739418 0.002851848 0.08568786  0.26841046  0.6033678 902.7885
##             Rhat
## beta[1] 1.004543
## beta[2] 1.003377
## beta[3] 1.003805
## beta[4] 1.003503
## beta[5] 1.007029
## beta[6] 1.000777

#checks
traceplot(mod,pars=params.to.monitor2,inc_warmup=FALSE)
```
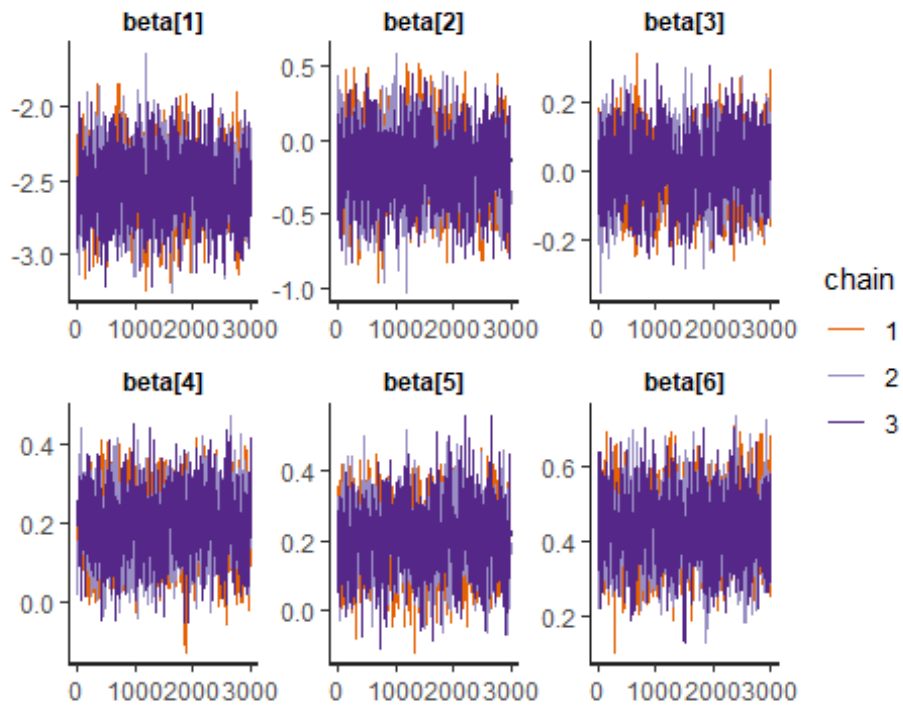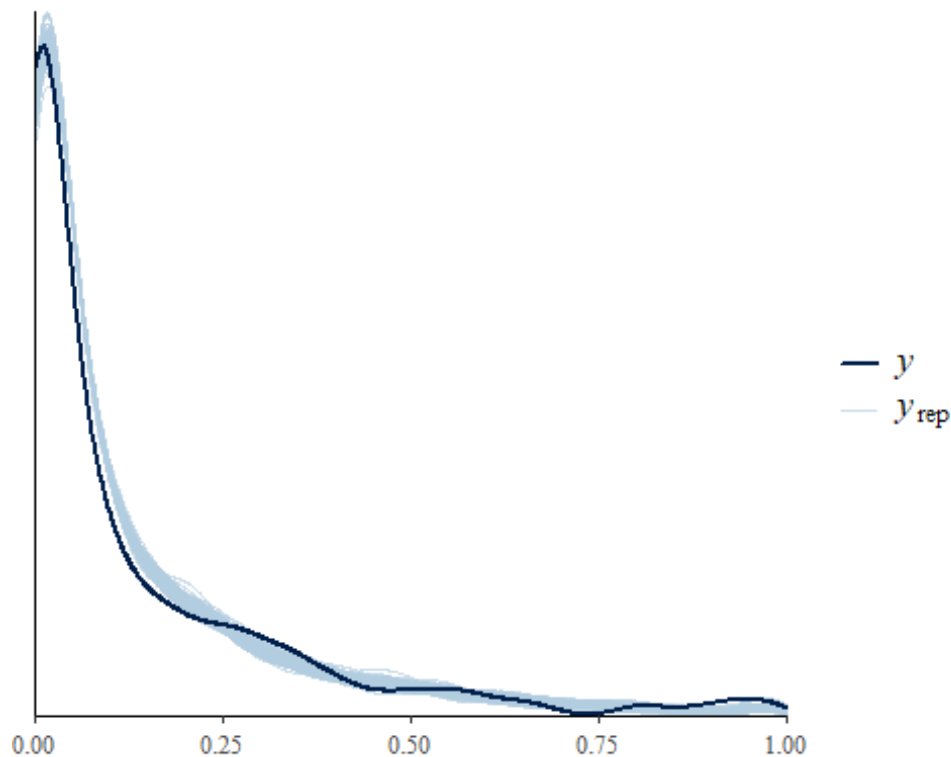
```r
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```

```
#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   7604.5  96.7
## p_loo      1994.6  25.3
## looic    -15209.0 193.4
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. ESS
## (-Inf, 0.7]   (good)      1381  42.7%   129
##    (0.7, 1]   (bad)       1554  48.0%   <NA>
##    (1, Inf)   (very bad)   300   9.3%   <NA>
## See help('pareto-k-diagnostic') for details.

#create data frame of looics from two models
load("2all_lat.rds")
loo.mod2=loo.mod # rename loo.mod so can load n
```

```
mod2 = mod
load("2all_interc.rds")

table.data<-data.frame(
  Model = c("Intercept-only model","Model including latitude"),
  LOOic = c(loo.mod$estimates[3],loo.mod2$estimates[3]),
  SE = c(loo.mod$estimates[6],loo.mod2$estimates[6])
)
knitr::kable(table.data, caption = "Table 1: Comparisons of LOOic between bas
eline and Model with latitudinal bands", format = "markdown")
```

Table 1: Comparisons of LOOic between baseline and Model with latitudinal bands

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including latitude	-15208.98	193.3504

```
Looic.diff = loo.mod2$estimates[3] - loo.mod$estimates[3]
cat("Difference in LOOic =", Looic.diff)

## Difference in LOOic = 4.787947
```

Results

Latitudinal variation in extinction risk is not supported by model comparison.

```
# dataP<-read.table("MetaRisk2 aggthres 5.txt",header=T);
# attach(dataP) # need to eliminate NA s for pre-industrial rise or stat prog
rams crash
#
# #betareg requires no 0s or 1s
# koffset = 0.001 #the k that gives the best posterior predictive check
# percent2 <- adj.percent
# percent2[adj.percent == 0] = koffset;
# percent2[adj.percent == 1] = 1 - koffset;
# dataP$percent2 <- percent2;
#
# data.use<-dataP
#
# #use one
# betamat <- model.matrix(~Antarctic, data = data.use, contrasts.arg = list(A
ntarctic = "contr.sum")) # did models with all and individually. However, bec
ause factors are not independent,
# betamat <- model.matrix(~S.Middle, data = data.use, contrasts.arg = list(S.
Middle = "contr.sum")) # did models with all and individually. However, becau
se factors are not independent,
# betamat <- model.matrix(~Tropics, data = data.use, contrasts.arg = list(Tro
pics = "contr.sum")) # did models with all and individually. However, because
factors are not independent,
# betamat <- model.matrix(~N.Middle, data = data.use, contrasts.arg = list(N.
Middle = "contr.sum")) # did models with all and individually. However, becau
```

```r
se factors are not independen
# betamat <- model.matrix(~Arctic, data = data.use, contrasts.arg = list(Arct
ic = "contr.sum")) # did models with all and individually. However, because f
actors are not independent,
#
# N = length(data.use$percent2)
# n.Study <- length(unique(data.use$Study)) #number of studies
# Studyint<-as.integer(unclass(factor(data.use$Study)))
# phi = data.use$Total.N
#
# #create model matrix for coefficients
# 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=5000, cores=3,iter=8000,
#          init = init.fn, save_warmup = FALSE, control=list(adapt_delta = 0.
9, max_treedepth = 15))#

n.total <- nrow(data.use)  #total N

#use common median threat
load("2all_interc.rds")
posterior2=as.data.frame(mod)
grand.mean = posterior2[["mu"]]
grand.mean.inv =invlogit(grand.mean)
r.mean.1 = grand.mean.inv
grand.mean.pred = quantile(grand.mean.inv,probs = c(0.025, 0.5, 0.975))
grand.mean.pred

##      2.5%        50%      97.5%
## 0.06568248 0.07596676 0.08662260

load("2antarct.rds")
mod.Ant <- mod
posterior <- as.data.frame(mod.Ant)
p.Ant <- sum(Antarctic == "Y")/n.total #prop of yes
p.other <- sum(Antarctic == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] - posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
#r.mean.1 = (p.other * beta.N + p.Ant * beta.Y)
beta.vs.r.mean.1 = (beta.Y) - (r.mean.1)
```

```r
param.vals.1 <- data.frame(
"r.mean.1" = r.mean.1,
"beta.N" = beta.N,
"beta.Y" = beta.Y,
"beta.vs.r.mean.1" = beta.vs.r.mean.1
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))

region.vals <- data.frame(
  region = "Antarctic",
  beta.Y.lo = pred.1[1,3],
  beta.Y = pred.1[2,3],
  beta.Y.hi = pred.1[3,3],
  beta.vs.mean.lo = pred.1[1,4],
  beta.vs.mean = pred.1[2,4],
  beta.vs.mean.hi = pred.1[3,4]
)
#*******************************************************************
************
load("2s_middle.rds")
#load("beta s.Middle.rds")
mod.SM <- mod
reg <- S.Middle
posterior <- as.data.frame(mod.SM)

p.Y <- sum(reg == "Y")/n.total #prop of yes
p.other <- sum(reg == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] - posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
#r.mean.1 = (p.other * beta.N + p.Y * beta.Y)
beta.vs.r.mean.1 = beta.Y - r.mean.1

param.vals.1 <- data.frame(
"r.mean.1" = r.mean.1,
"beta.N" = beta.N,
"beta.Y" = beta.Y,
"beta.vs.r.mean.1" = beta.vs.r.mean.1
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))
region.vals[2,] <- cbind("S.Middle", t(pred.1[,3]), t(pred.1[,4]))

#*******************************************************************
****************
load("2tropics.rds")
#load("beta tropics.rds")
mod.trop <- mod
```

```
reg <- Tropics
posterior <- as.data.frame(mod.trop)

p.Y <- sum(reg == "Y")/n.total #prop of yes
p.other <- sum(reg == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] - posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
#r.mean.1 = (p.other * beta.N + p.Y * beta.Y)
beta.vs.r.mean.1 = beta.Y - r.mean.1

param.vals.1 <- data.frame(
"r.mean.1" = r.mean.1,
"beta.N" = beta.N,
"beta.Y" = beta.Y,
"beta.vs.r.mean.1" = beta.vs.r.mean.1
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))
region.vals[3,] <- cbind("Tropics", t(pred.1[,3]), t(pred.1[,4]))
#*************************************************************************
*****************

load("2n_middle.rds")
mod.NM <- mod
reg <- N.Middle
posterior <- as.data.frame(mod.NM)

p.Y <- sum(reg == "Y")/n.total #prop of yes
p.other <- sum(reg == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] - posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
#r.mean.1 = (p.other * beta.N + p.Y * beta.Y)
beta.vs.r.mean.1 = beta.Y - r.mean.1

param.vals.1 <- data.frame(
"r.mean.1" = r.mean.1,
"beta.N" = beta.N,
"beta.Y" = beta.Y,
"beta.vs.r.mean.1" = beta.vs.r.mean.1
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))
region.vals[4,] <- cbind("N.Middle", t(pred.1[,3]), t(pred.1[,4]))
#*************************************************************************
*****************

load("2arctic.rds")
mod.Arc <- mod
reg <- Arctic
```

```
posterior <- as.data.frame(mod.Arc)

p.Y <- sum(reg == "Y")/n.total #prop of yes
p.other <- sum(reg == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] - posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
#r.mean.1 = (p.other * beta.N + p.Y * beta.Y)
beta.vs.r.mean.1 = beta.Y - r.mean.1

param.vals.1 <- data.frame(
"r.mean.1" = r.mean.1,
"beta.N" = beta.N,
"beta.Y" = beta.Y,
"beta.vs.r.mean.1" = beta.vs.r.mean.1
)
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))
region.vals[5,] <- cbind("Arctic", t(pred.1[,3]), t(pred.1[,4]))
#***************************************************************************
******************
save(region.vals,file ="Latitudinal bands analysis results 2.rds")
knitr::kable(region.vals, caption = "Table 2: Extinction risk by regiona and
difference from mean, with 95% credible intervals", format = "markdown", digi
ts = 3)
```

Table 2: Extinction risk by regiona and difference from mean, with 95% credible intervals

region	beta.Y.lo	beta.Y	beta.Y.hi	beta.vs.mean.lo	beta.vs.mean	beta.vs.mean.hi
Anta rctic	0.0249949401 60775	0.052633856 7046219	0.109086011 902341	- 0.052497246 410364	- 0.023023514 0343709	0.033578432 4863166
S.Mid dle	0.0687327691 11236	0.086038292 2771603	0.106918747 354409	- 0.010629685 0753243	0.010026463 9735213	0.033244853 2048658
Tropi cs	0.0617643807 765014	0.074058458 5066626	0.087860106 2055393	- 0.018152507 352489	- 0.001931240 60702644	0.015293308 5919268
N.Mi ddle	0.0542217675 28623	0.064120407 6880021	0.074605974 7758127	- 0.026127474 4151048	- 0.011766332 7521619	0.002664732 91247814
Arcti c	0.0285859124 004475	0.037645293 9766856	0.048975740 496411	- 0.052531251 0883054	- 0.038187580 8714051	- 0.023016344 2677905

Northern middle latitudes and Arctic appear to be characterized by lower predicted extinction risks than other regions.

```
#load("Latitudinal bands analysis results 2.rds")
#region.vals$region[1] = "Antarctic" #fix name
region.vals$beta.Y.lo <- as.numeric(region.vals$beta.Y.lo) # data is characte
```

```r
r and needs to be numeric
region.vals$beta.Y <- as.numeric(region.vals$beta.Y)
region.vals$beta.Y.hi <- as.numeric(region.vals$beta.Y.hi)
region.vals$beta.vs.mean.lo <- as.numeric(region.vals$beta.vs.mean.lo)
region.vals$beta.vs.mean <- as.numeric(region.vals$beta.vs.mean)
region.vals$beta.vs.mean.hi<- as.numeric(region.vals$beta.vs.mean.hi)
region.vals$region <- factor(region.vals$region, levels = c("Antarctic", "S.M
iddle", "Tropics", "N.Middle","Arctic"))

#Figures
Fig1a<-ggplot(data = region.vals)+
  geom_vline(xintercept=grand.mean.pred[2]) +
  geom_errorbar(aes(y = region, xmin = beta.Y.lo, xmax = beta.Y.hi), width =
0) +
  geom_point(stat = "identity", aes(y = region, x = beta.Y), color = "#416788
", size = 3, shape = 15) +
  xlab("Pre-industrial rise \n in temperature (C)") + xlim(c(0,.2)) +
  theme_classic()+ scale_y_discrete(labels = c("Antarctic", "S. Middle", "Tro
pics", "N. Middle", "Arctic")) +
  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
```

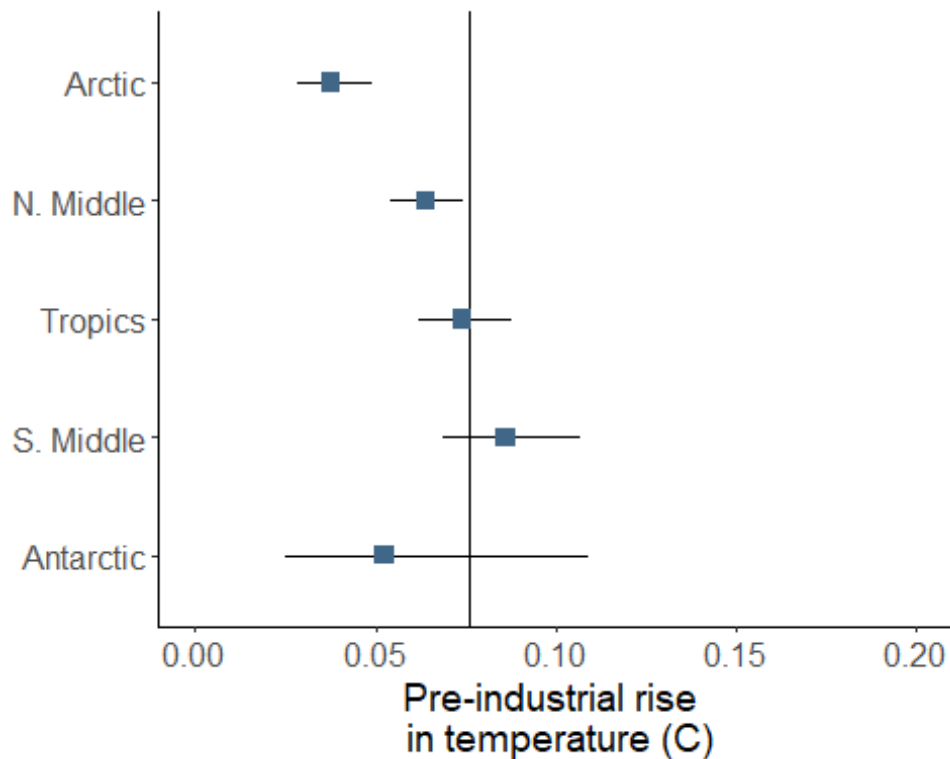*Fig. 1. Predicted extinction risk based on latitudinal bands.*

```r
#Differences from zero
glob.mean.over <- rep("n",nrow(region.vals))
glob.mean.over[region.vals$beta.vs.mean > 0 & region.vals$beta.vs.mean.lo > 0
] = "y"
glob.mean.over[region.vals$beta.vs.mean < 0 & region.vals$beta.vs.mean.hi < 0
] = "y"
region.vals$glob.mean.over <- glob.mean.over

Fig1b<-ggplot(data = region.vals)+
  geom_vline(xintercept=0) +
  geom_errorbar(aes(y = region, xmin = beta.vs.mean.lo, xmax = beta.vs.mean.h
i), width = 0) +
  geom_point(stat = "identity", aes(y = region, x = beta.vs.mean, color = glo
b.mean.over), size = 3, shape = 15) +
  scale_color_manual(values=c('grey','#5B507A'))+
  xlab("Percent difference \n from global mean") + xlim(c(-.1,.1)) +
  theme_classic()+
  theme(axis.title.y = element_blank(),axis.ticks.y = element_blank(),axis.li
ne.y = element_blank(),axis.text.y = element_blank(),
        axis.title=element_text(size=14),axis.text = element_text(size=12),le
gend.position = "none")+
  guides(size=F)
Fig1b
```
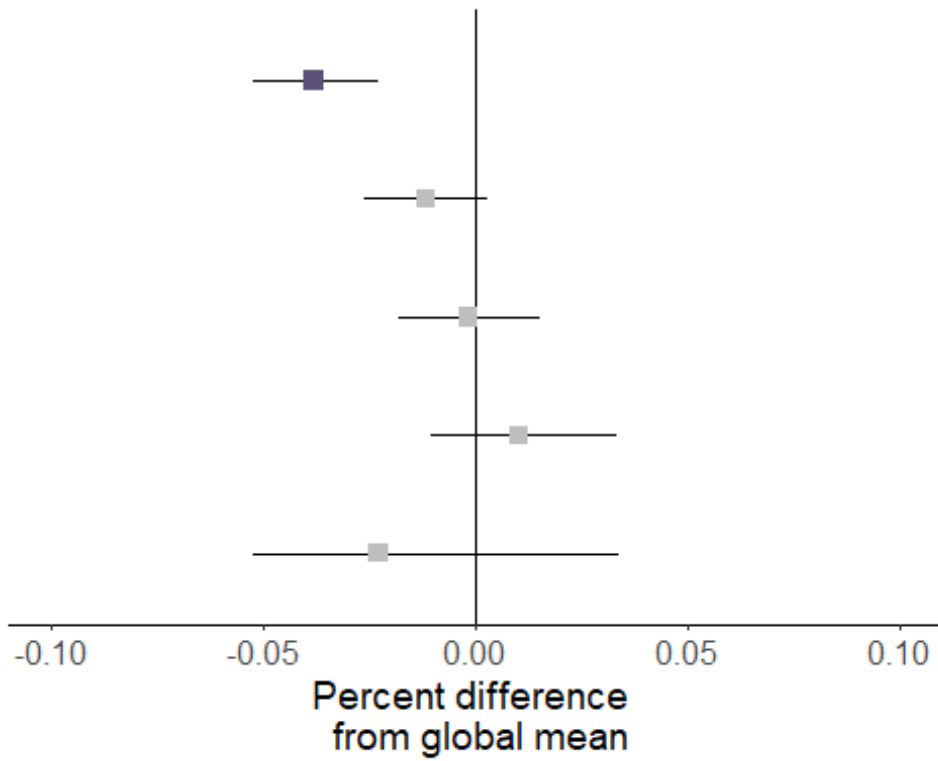
Fig. 2. Predicted extinction risk based on latitudinal bands.

```
ggarrange(Fig1a, NULL, Fig1b, ncol=3, widths = c(4,.4, 2))
```
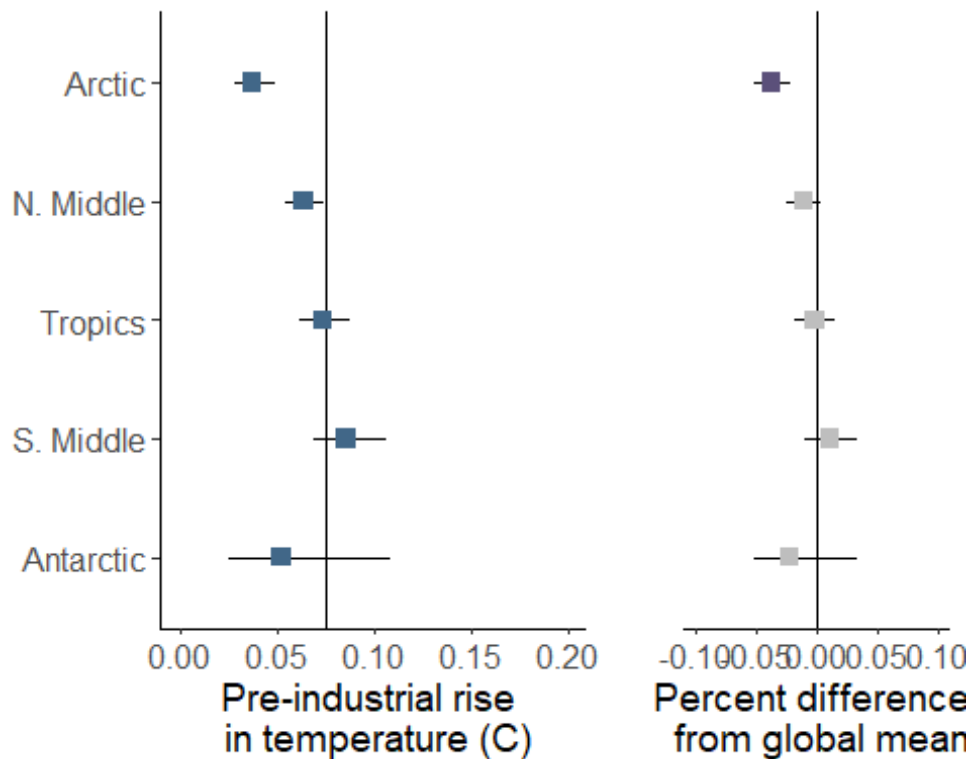
Fig. 3. Predicted extinction risk based on latitudinal bands.

```r
#ggsave("Metarisk2 lat bands.png",width=6,height=5,unit="in",dpi="print")

#catalog sample sizes
N.1.st<-length(unique(dataP$Study[dataP$Antarctic == "Y"]))
N.1.mod<-length((dataP$Study[dataP$Antarctic == "Y"]))

N.2.st<-length(unique(dataP$Study[dataP$S.Middle == "Y"]))
N.2.mod<-length((dataP$Study[dataP$S.Middle == "Y"]))

N.3.st<-length(unique(dataP$Study[dataP$Tropics == "Y"]))
N.3.mod<-length((dataP$Study[dataP$Tropics == "Y"]))

N.4.st<-length(unique(dataP$Study[dataP$N.Middle == "Y"]))
N.4.mod<-length((dataP$Study[dataP$N.Middle == "Y"]))

N.5.st<-length(unique(dataP$Study[dataP$Arctic == "Y"]))
N.5.mod<-length((dataP$Study[dataP$Arctic == "Y"]))

table.data<-data.frame(
  Factor = c("Antarctic", "S. Middle", "Tropics", "N. Middle", "Arctic"),
  Studies = c(N.1.st,N.2.st,N.3.st,N.4.st,N.5.st),
  Models = c(N.1.mod,N.2.mod,N.3.mod,N.4.mod,N.5.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
Antarctic	17	213
S. Middle	156	914
Tropics	240	1753
N. Middle	319	2131
Arctic	93	748

Conclusion

Latitudinal bands vary in risk, with lower risk in the Northern middle latitudes and Arctic.