Extinction risk from climate change: other threats

March 25th, 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; attach(data.use)

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
```

Test for overall threat effect

```r
#form matrix of factors
#create model matrix for coefficients
betamat <- data.frame(
  Intercept = rep(1,N),
  Threatened = ifelse(data.use$Threatened == "Y",1,0),
  Endemic = ifelse(data.use$Endemic == "Y",1,0),
  Non.clim.threat = ifelse(data.use$Non.clim.threat == "Y",1,0),
  Land.Use.Change = ifelse(data.use$Land.Use.Change == "Y",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=5000, cores=3,iter=8000,
#          init = init.fn, save_warmup = FALSE, control=list(adapt_delta = 0.
9, max_treedepth = 15))#

load("3allthreat.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] -3.0522288 0.003293059 0.1112868 -3.27376905 -2.8347429 1142.059
## beta[2]  0.4840492 0.004937104 0.1726892  0.15147406  0.8305115 1223.448
## beta[3]  0.7477895 0.004101094 0.1375451  0.48040237  1.0151293 1124.840
## beta[4]  0.2685812 0.004075115 0.1830882 -0.08701929  0.6264643 2018.557
## beta[5] -0.2510086 0.005293473 0.2023748 -0.64257554  0.1430060 1461.610
##             Rhat
## beta[1] 1.004185
## beta[2] 1.002559
## beta[3] 1.003848
## beta[4] 1.000069
## beta[5] 1.003224

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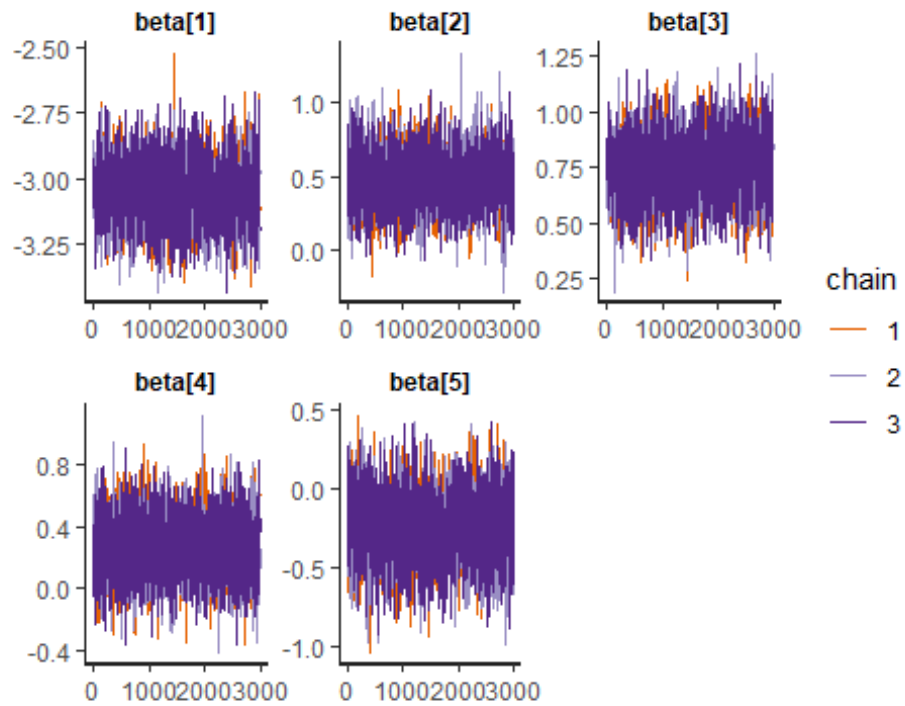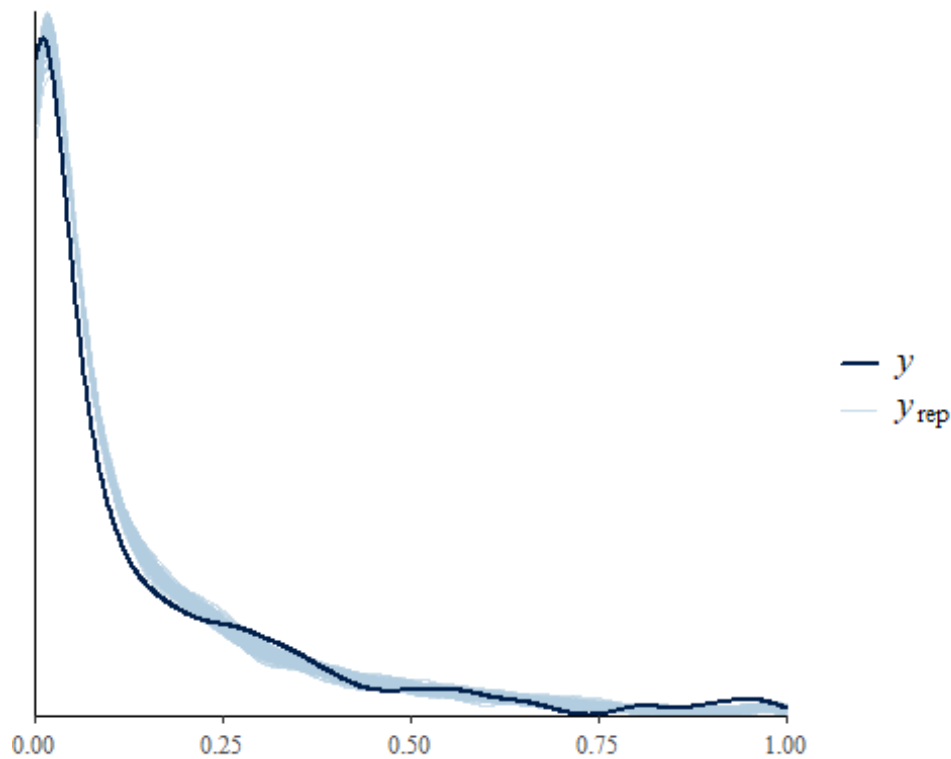
```
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```

```
#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   7619.0  96.6
## p_loo      1982.3  25.1
## looic    -15238.0 193.1
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. ESS
## (-Inf, 0.7]   (good)      1410  43.6%   103
##    (0.7, 1]   (bad)       1576  48.7%   <NA>
##    (1, Inf)   (very bad)   249   7.7%   <NA>
## See help('pareto-k-diagnostic') for details.

loo.mod2=loo.mod # rename loo.mod so can load n
mod2 = mod
load("2all_interc.rds")
```

```r
table.data<-data.frame(
  Model = c("Intercept-only model","Model including all threats"),
  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 all threats", format = "markdown")
```

Table 1: Comparisons of LOOic between baseline and all threats

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including all threats	-15237.99	193.1132

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

## Difference in LOOic = -24.21978
```

Test for effect of perceived threat

Here I tested if extinction predictions change based on whether species were considered threatened a priori, if non-climate threats were modeled, and if land use change was modeled through time. I predicted that extinction risks would increase for each of these categories relative to the overall mean.

Perceived threat

```r
#form matrix of factors
betamat <- model.matrix(~Non.clim.threat, data = data.use) #

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.t
o.monitor,
 #          chains = 3, warmup=5000, cores=7,iter=8000,
 #          init = init.fn, control=list(adapt_delta = 0.9, max_treedepth = 1
5))

load("2threat.rds")
params.to.monitor2=c("beta")#
```
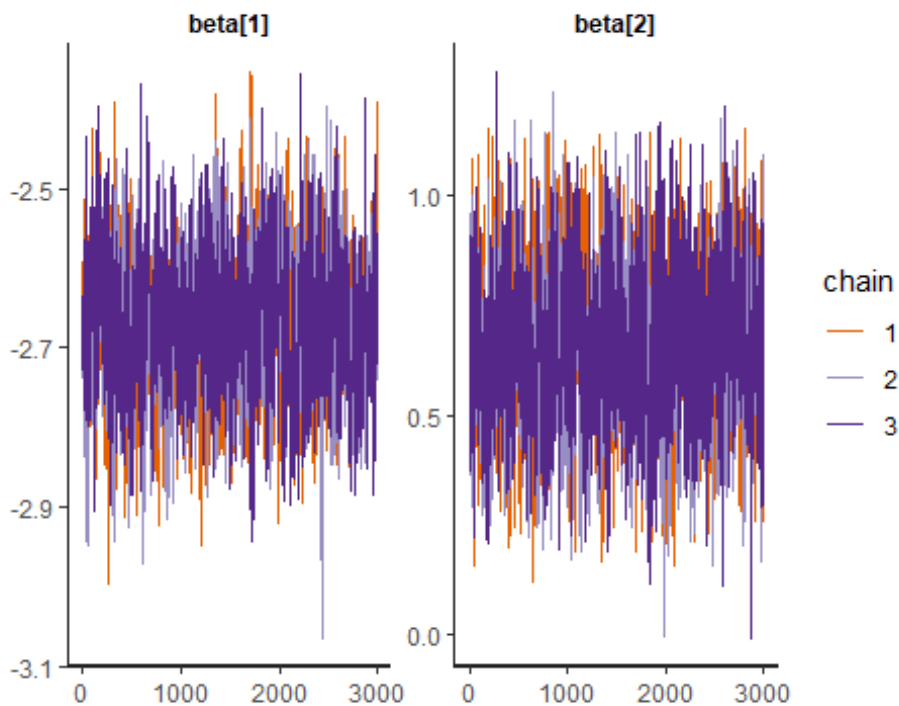
```r
sumx = summary(mod,probs=c(.025,0.975), digits=4, pars=params.to.monitor2)
sumx$summary

##                mean       se_mean         sd       2.5%        97.5%      n_eff
## beta[1] -2.6700120 0.002698749 0.08641382 -2.837096 -2.4965599 1025.277
## beta[2]  0.6561329 0.004945721 0.17275113  0.315909   0.9982394 1220.064
##              Rhat
## beta[1] 1.003345
## beta[2] 1.002287

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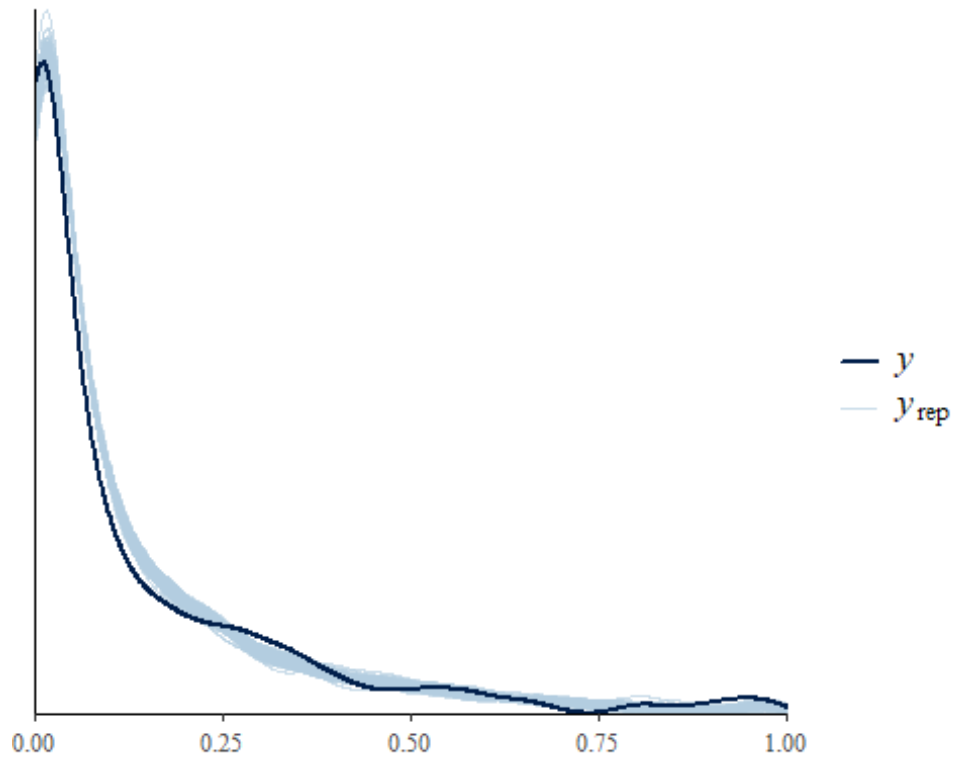


```r
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```

```
#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   7608.8  96.7
## p_loo      1990.7  25.0
## looic    -15217.6 193.4
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. ESS
## (-Inf, 0.7]   (good)      1405  43.4%   139
##    (0.7, 1]   (bad)       1574  48.7%   <NA>
##    (1, Inf)   (very bad)   256   7.9%   <NA>
## See help('pareto-k-diagnostic') for details.
```

Species thought to be threatened are more threatened.

```
loo.mod2=loo.mod # rename loo.mod so can load n
mod2 = mod
#load("beta interc only.rds")
load("2all_interc.rds")

table.data<-data.frame(
  Model = c("Intercept-only model","Model including threat"),
  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 2: Comparisons of LOOic between bas
eline and threat", format = "markdown")
```

Table 2: Comparisons of LOOic between baseline and threat

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including threat	-15217.62	193.3507

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

## Difference in LOOic = -3.854803
```

Some limited support for the model (LOOic is lower).

Endemicity

Here I tested if extinction predictions change based on whether species endemic or not to the geographic area analyzed in a study. I predicted that extinction risks would increase because endemic species tend to have smaller initial ranges and are more specialized.

```
#form matrix
betamat <- model.matrix(~Endemic, data = data.use) # 1 - No, 2 - Some, 3 - Ye
s

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.t
o.monitor,
 #             chains = 3, warmup=4000, cores=7,iter=7000,
 #           init = init.fn, control=list(adapt_delta = 0.9, max_treedepth = 1
5))

load("2endemic.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] -2.94626642 0.003626793 0.1059337 -3.1550903 -2.743742   853.1449
## beta[2] -0.06221211 0.017561385 1.4408485 -2.8636109  2.757582 6731.6124
## beta[3]  0.81524023 0.004917579 0.1365016  0.5452175  1.082759   770.4998
##              Rhat
## beta[1] 1.001334
## beta[2] 0.999872
## beta[3] 1.001257

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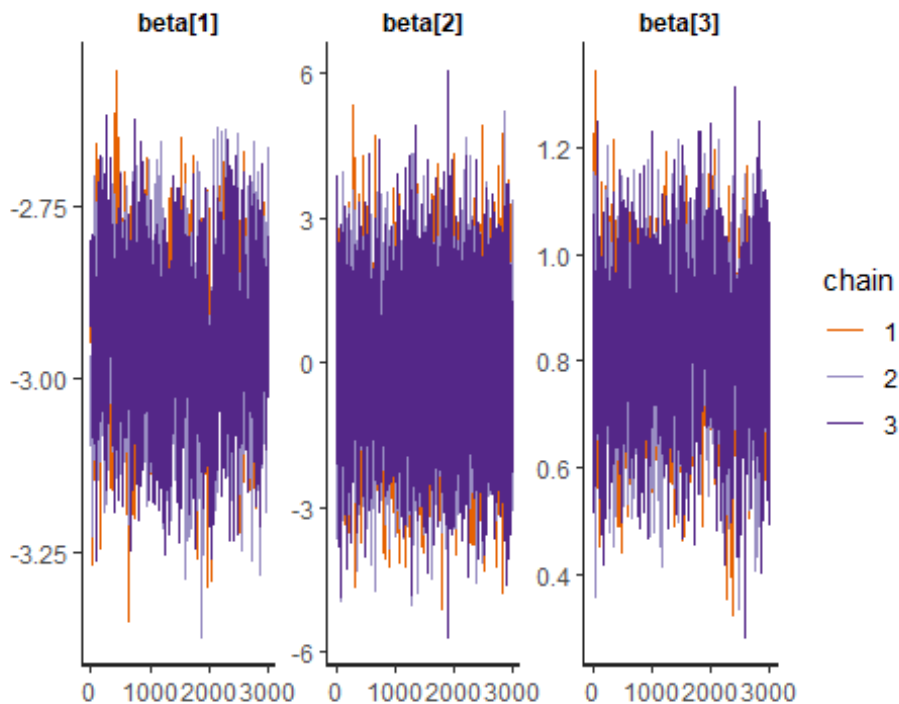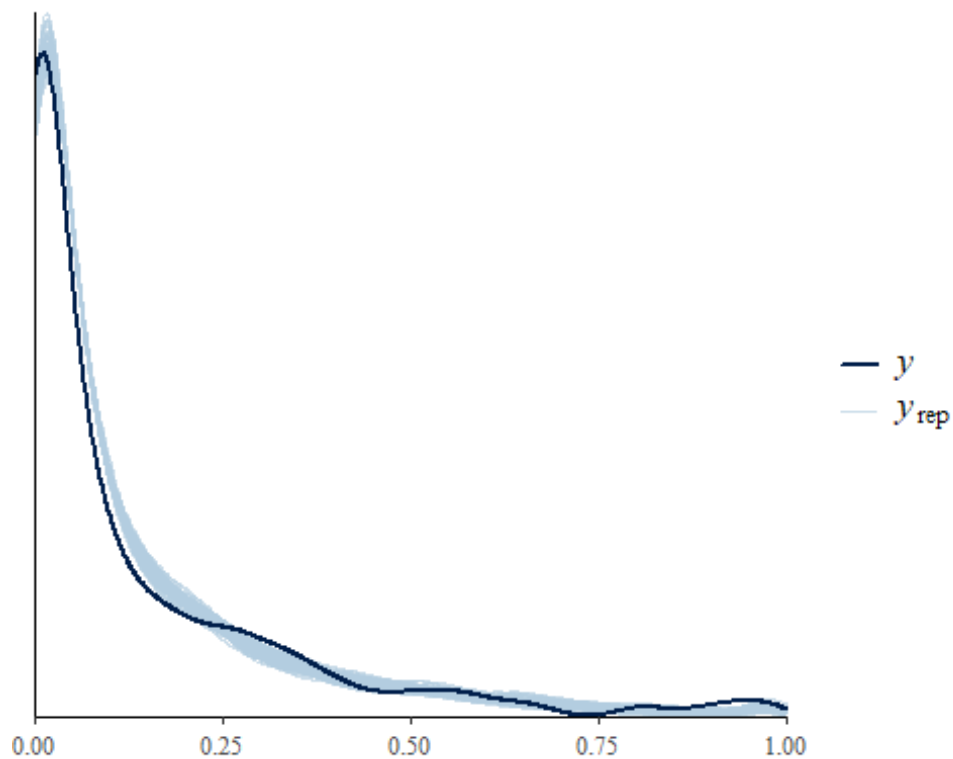
```r
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```



```r
#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   7608.6   96.4
## p_loo      1992.1   25.1
## looic    -15217.2  192.7
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. ESS
## (-Inf, 0.7]   (good)      1375  42.5%   125
##    (0.7, 1]   (bad)       1602  49.5%   <NA>
```

```
##     (1, Inf)   (very bad)  258   8.0%   <NA>
## See help('pareto-k-diagnostic') for details.

loo.mod2=loo.mod # rename loo.mod so can load n
load("2all_interc.rds")

table.data<-data.frame(
  Model = c("Intercept-only model","Model including threat"),
  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 3: Comparisons of LOOic between bas
eline and threat", format = "markdown")
```

Table 3: Comparisons of LOOic between baseline and threat

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including threat	-15217.23	192.7385

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

## Difference in LOOic = -3.463133
```

Some limited support for the model (LOOic is lower).

Non-climate threat

```
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 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 = 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

betamat <- model.matrix(~Non.clim.threat, data = data.use) #

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.t
o.monitor,
 #          chains = 3, warmup=4000, cores=7,iter=7000,
 #          init = init.fn, control=list(adapt_delta = 0.9, max_treedepth = 1
5))

load("2nc.threat.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] -2.5149171 0.002458583 0.0752971 -2.6625898 -2.3637582 937.9655
## beta[2]  0.1096649 0.005462529 0.1166680 -0.1141653  0.3398413 456.1590
##              Rhat
## beta[1] 1.000933
## beta[2] 1.006844

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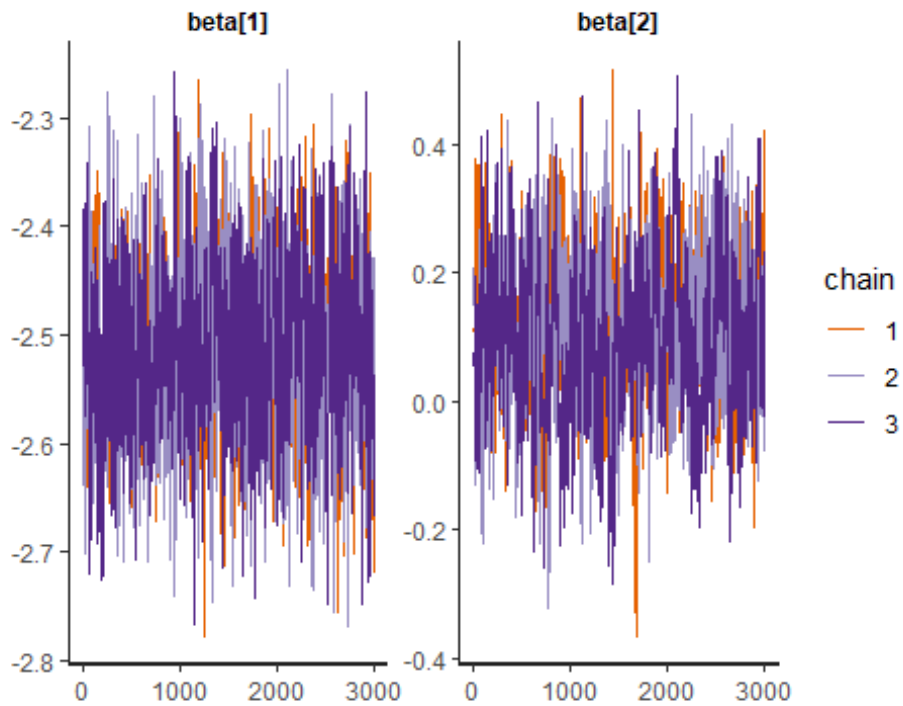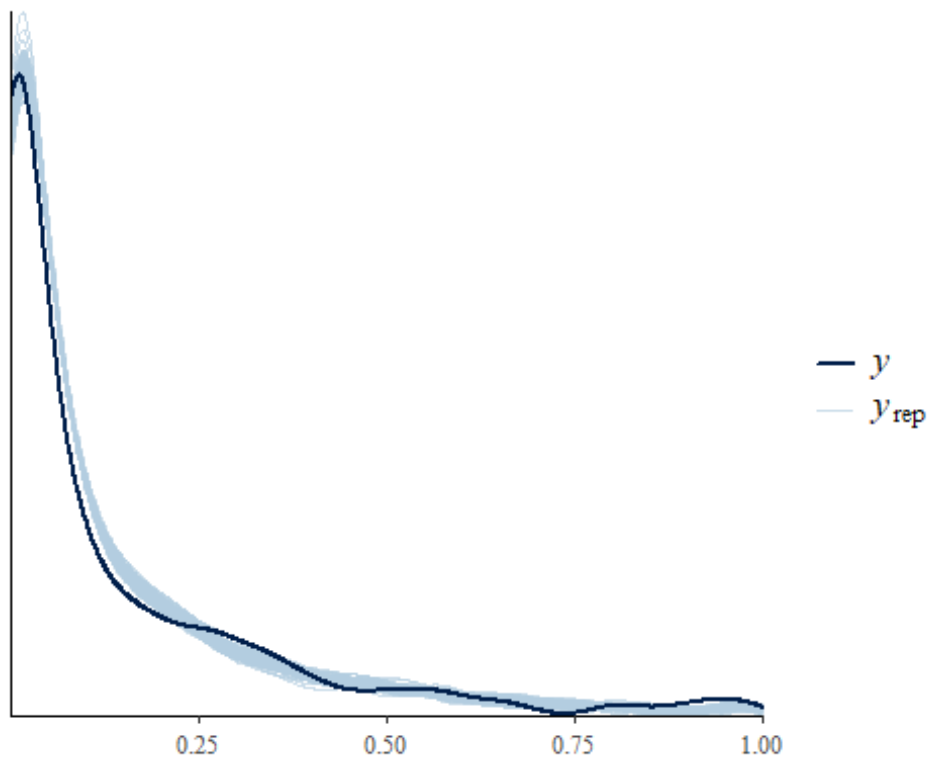
```
pp_check(
  stan.data$percent,
  rstan::extract(mod, par = 'y_rep')$y_rep[1:100, ],
  fun = 'dens_overlay'
)
```

```
#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   7597.3  96.8
## p_loo      2001.0  25.4
## looic    -15194.6 193.6
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. ESS
## (-Inf, 0.7]   (good)      1387  42.9%   125
##    (0.7, 1]   (bad)       1552  48.0%   <NA>
##    (1, Inf)   (very bad)   296   9.1%   <NA>
## See help('pareto-k-diagnostic') for details.

#save(mod, loo.mod, file="nc.threat.rds")
```

No support for non-climate threats as contributing substantially to increasing extinction threat.

```r
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 non-climate threat"),
  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 4: Comparisons of LOOic between bas
eline and non-climate threat", format = "markdown")
```

Table 4: Comparisons of LOOic between baseline and non-climate threat

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including non-climate threat	-15194.60	193.6465

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

## Difference in LOOic = 19.16506
```

No model support for non-climate threat (LOOic is higher).

Land-use over time

```r
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,
##      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 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 = 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

betamat <- model.matrix(~Land.Use.Change)

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.t
o.monitor,
 #          chains = 3, warmup=4000, cores=7,iter=7000,
 #          init = init.fn, control=list(adapt_delta = 0.9, max_treedepth = 1
5))
load("2LUchange.rds")
params.to.monitor2=c("beta")#
sumx = summary(mod,probs=c(.025,0.975), digits=4, pars=params.to.monitor2)
sumx$summary
```
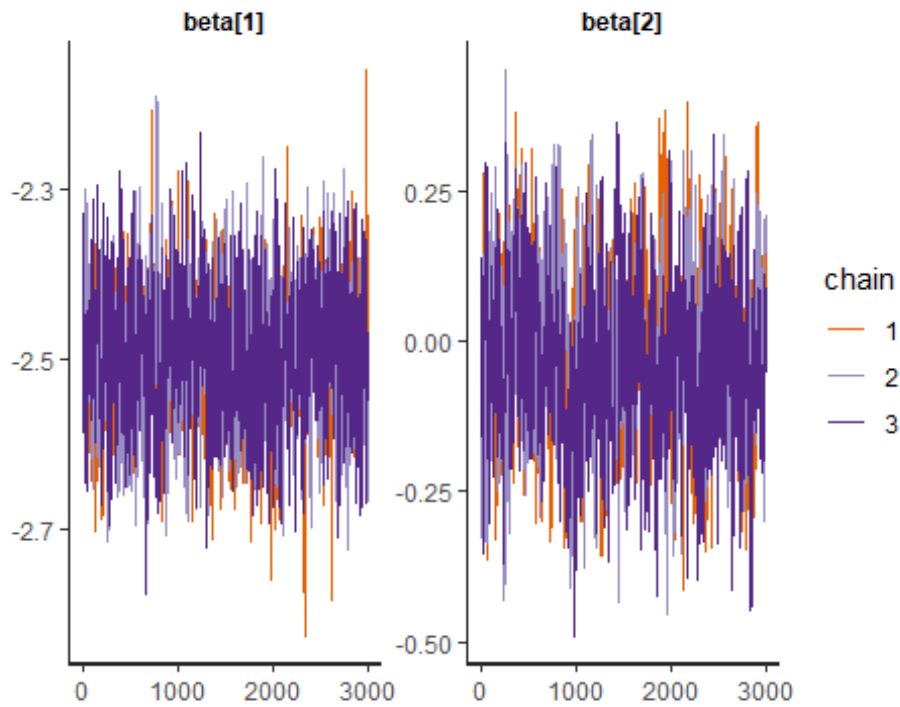
```
##                mean      se_mean         sd        2.5%      97.5%     n_eff
## beta[1] -2.49782910 0.002639123 0.07569736 -2.6478832 -2.35355 822.7015
## beta[2] -0.02149434 0.005538077 0.13194628 -0.2830907  0.24110 567.6443
##               Rhat
## beta[1] 1.000564
## beta[2] 1.008911
```

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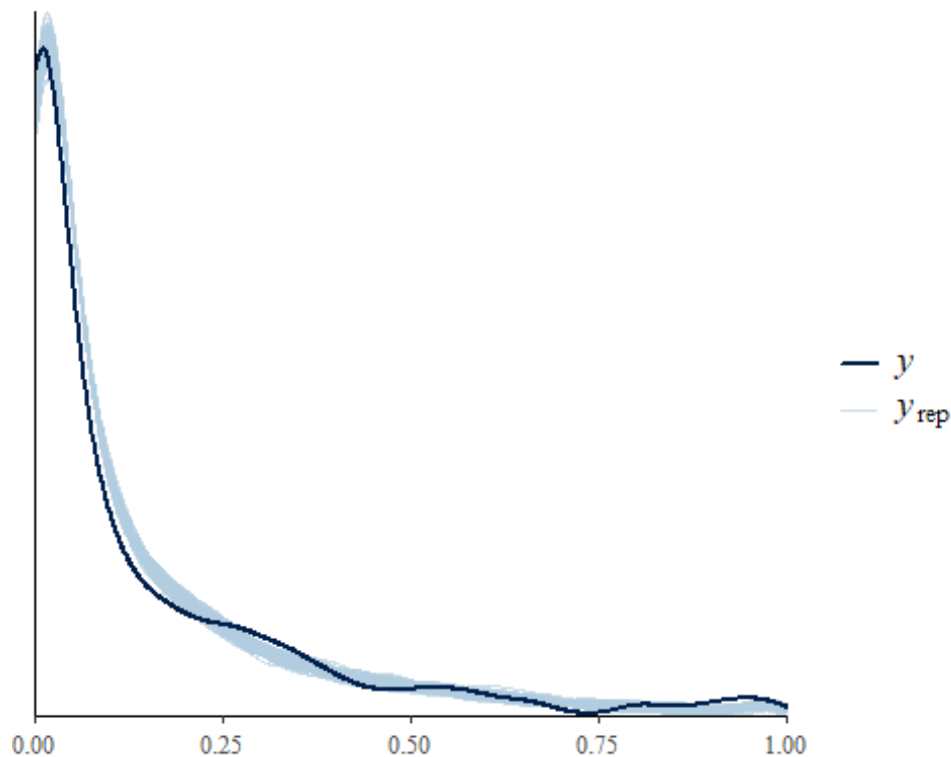
```
#calculate loo
# log_lik_1 <- extract_log_lik(mod, merge_chains = FALSE)
# r_eff <- relative_eff(exp(log_lik_1), cores = 6)
# loo.mod <- loo(log_lik_1, r_eff = r_eff, cores = 6)
loo.mod #

##
## Computed from 9000 by 3235 log-likelihood matrix.
##
##           Estimate    SE
## elpd_loo   7607.0   96.6
## p_loo      1991.8   25.1
## looic    -15214.0  193.3
## ------
## MCSE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                            Count Pct.    Min. ESS
## (-Inf, 0.7]  (good)        1432  44.3%   104
##    (0.7, 1]  (bad)         1548  47.9%   <NA>
##    (1, Inf)  (very bad)     255   7.9%   <NA>
## See help('pareto-k-diagnostic') for details.

#save(mod, loo.mod, file="LUchange.rds")

loo.mod2=loo.mod # rename loo.mod so can load n
mod2 = mod
```

```
#load("beta interc only.rds")
load("2all_interc.rds")

table.data<-data.frame(
  Model = c("Intercept-only model","Model including future land use change"),
  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 5: Comparisons of LOOic between bas
eline and land use change", format = "markdown")
```

Table 5: Comparisons of LOOic between baseline and land use change

Model	LOOic	SE
Intercept-only model	-15213.77	193.3243
Model including future land use change	-15214.01	193.2521

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

## Difference in LOOic = -0.2450977
```

Little model support for land use change affecting extinction risks.

```
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.pred <- invlogit(quantile(grand.mean, probs = c(0.025, 0.5, 0.975)
))

#**************************************************************************
*************
load("2threat.rds")
mod.a <- mod
posterior <- as.data.frame(mod.a)
p.y <- sum(data.use$Threatened == "Y")/n.total #prop of yes
p.n <- sum(data.use$Threatened == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.y * beta.Y
r.mean.1 = invlogit(grand.mean)
```

```r
beta.vs.r.mean.n = (beta.N) - (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.Y" = beta.Y,
"beta.vs.r.mean.n" = beta.vs.r.mean.n,
"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))

threat.vals <- data.frame(overall.mean = pred.1[2,1],
  beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.Y.lo = pred.1[1,3],
  beta.Y = pred.1[2,3],
  beta.Y.hi = pred.1[3,3],
  beta.N.mean.lo = pred.1[1,4],
  beta.N.mean = pred.1[2,4],
  beta.N.mean.hi = pred.1[3,4],
  beta.Y.mean.lo = pred.1[1,5],
  beta.Y.mean = pred.1[2,5],
  beta.Y.mean.hi = pred.1[3,5]
)

#*****************************************************************************
*************
load("2endemic.rds")
mod.a <- mod
posterior <- as.data.frame(mod.a)
p.y <- sum(data.use$Threatened == "Y")/n.total #prop of yes
p.n <- sum(data.use$Threatened == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[3]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.y * beta.Y
r.mean.1 = invlogit(grand.mean)
beta.vs.r.mean.n = (beta.N) - (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.Y" = beta.Y,
"beta.vs.r.mean.n" = beta.vs.r.mean.n,
"beta.vs.r.mean.y" = beta.vs.r.mean.y
)
```

```r
pred.1 = (apply(param.vals.1, 2, quantile, probs = c(0.025, 0.5, 0.975),na.rm
=TRUE))

threat.vals[2,] <- data.frame(overall.mean = pred.1[2,1],
  beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.Y.lo = pred.1[1,3],
  beta.Y = pred.1[2,3],
  beta.Y.hi = pred.1[3,3],
  beta.N.mean.lo = pred.1[1,4],
  beta.N.mean = pred.1[2,4],
  beta.N.mean.hi = pred.1[3,4],
  beta.Y.mean.lo = pred.1[1,5],
  beta.Y.mean = pred.1[2,5],
  beta.Y.mean.hi = pred.1[3,5]
)

#************************************************************************
*************

load("2nc.threat.rds")
mod.a <- mod
posterior <- as.data.frame(mod.a)
p.y <- sum(data.use$Non.clim.threat == "Y")/n.total #prop of yes
p.n <- sum(data.use$Non.clim.threat == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.y * beta.Y
r.mean.1 = invlogit(grand.mean)
beta.vs.r.mean.n = (beta.N) - (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.Y" = beta.Y,
"beta.vs.r.mean.n" = beta.vs.r.mean.n,
"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))

threat.vals[3,] <- cbind(overall.mean = pred.1[2,1],
  beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.Y.lo = pred.1[1,3],
  beta.Y = pred.1[2,3],
```

```r
  beta.Y.hi = pred.1[3,3],
  beta.N.mean.lo = pred.1[1,4],
  beta.N.mean = pred.1[2,4],
  beta.N.mean.hi = pred.1[3,4],
  beta.Y.mean.lo = pred.1[1,5],
  beta.Y.mean = pred.1[2,5],
  beta.Y.mean.hi = pred.1[3,5]
)
#*******************************************************************
*************

load("2LUchange.rds")
mod.a <- mod
posterior <- as.data.frame(mod.a)
p.y <- sum(data.use$Land.Use.Change == "Y")/n.total #prop of yes
p.n <- sum(data.use$Land.Use.Change == "N")/n.total #prop of nos
beta.Y = invlogit(posterior[["beta[1]"]] + posterior[["beta[2]"]])
beta.N = invlogit(posterior[["beta[1]"]])
#r.mean.1 = p.n * beta.N + p.y * beta.Y
r.mean.1 = invlogit(grand.mean)
beta.vs.r.mean.n = (beta.N) - (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.Y" = beta.Y,
"beta.vs.r.mean.n" = beta.vs.r.mean.n,
"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))

threat.vals[4,]<- cbind(overall.mean = pred.1[2,1],
  beta.N.lo = pred.1[1,2],
  beta.N = pred.1[2,2],
  beta.N.hi = pred.1[3,2],
  beta.Y.lo = pred.1[1,3],
  beta.Y = pred.1[2,3],
  beta.Y.hi = pred.1[3,3],
  beta.N.mean.lo = pred.1[1,4],
  beta.N.mean = pred.1[2,4],
  beta.N.mean.hi = pred.1[3,4],
  beta.Y.mean.lo = pred.1[1,5],
  beta.Y.mean = pred.1[2,5],
  beta.Y.mean.hi = pred.1[3,5]
)
#*******************************************************************
*************
threat.vals$threat = c("Preconceived Threat", "Endemic", "Other Threats","Lan
```

```
d use change")

save(threat.vals,file ="2Threat analysis results.rds")
threat.vals.r <- threat.vals
threat.vals.r[,1:13] <- round((threat.vals.r[,1:13]), digits = 4)
knitr::kable(threat.vals.r, caption = "Table 6: Extinction risk by threats an
d difference from mean, with 95% credible intervals", format = "markdown", fo
nt = 5)
```

Table 6: Extinction risk by threats and difference from mean, with 95% credible intervals

overall.mean	beta.N.lo	beta.N	beta.N.hi	beta.Y.lo	beta.Y	beta.Y.hi	beta.N.mean.lo	beta.N.mean	beta.N.mean.hi	beta.Y.mean.lo	beta.Y.mean	beta.Y.mean.hi	threat
0.076	0.0554	0.0648	0.0761	0.0910	0.1178	0.1518	-0.0253	-0.0113	0.0041	0.0132	0.0419	0.0769	Preconceived Threat
0.076	0.0409	0.0500	0.0604	0.0884	0.1063	0.1251	-0.0404	-0.0258	-0.0115	0.0095	0.0304	0.0517	Endemic
0.076	0.0652	0.0748	0.0860	0.0656	0.0882	0.1036	-0.0152	-0.0011	0.0136	-0.0135	0.0070	0.0292	Other Threats
0.076	0.0661	0.0676	0.0868	0.0574	0.0747	0.0973	-0.0139	0.0000	0.0152	-0.0217	-0.0013	0.0224	Land use change

```
#load("Threat analysis results.rds")
threat.vals$threat <- factor(threat.vals$threat, levels = c("Preconceived Thr
eat", "Endemic","Other Threats", "Land use change"))

#Figures
Fig2a<-ggplot(data = threat.vals)+
  geom_vline(xintercept=grand.mean.pred[2]) +
  geom_errorbar(aes(y = threat, xmin = beta.Y.lo, xmax = beta.Y.hi), width =
0) +
  geom_point(stat = "identity", aes(y = threat, x = beta.Y), color = "#416788
", size = 3, shape = 15) +
  xlab("Pre-industrial rise \n in temperature (C)") + xlim(c(0,.2)) +
  theme_classic()+
  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.

Fig2a
```


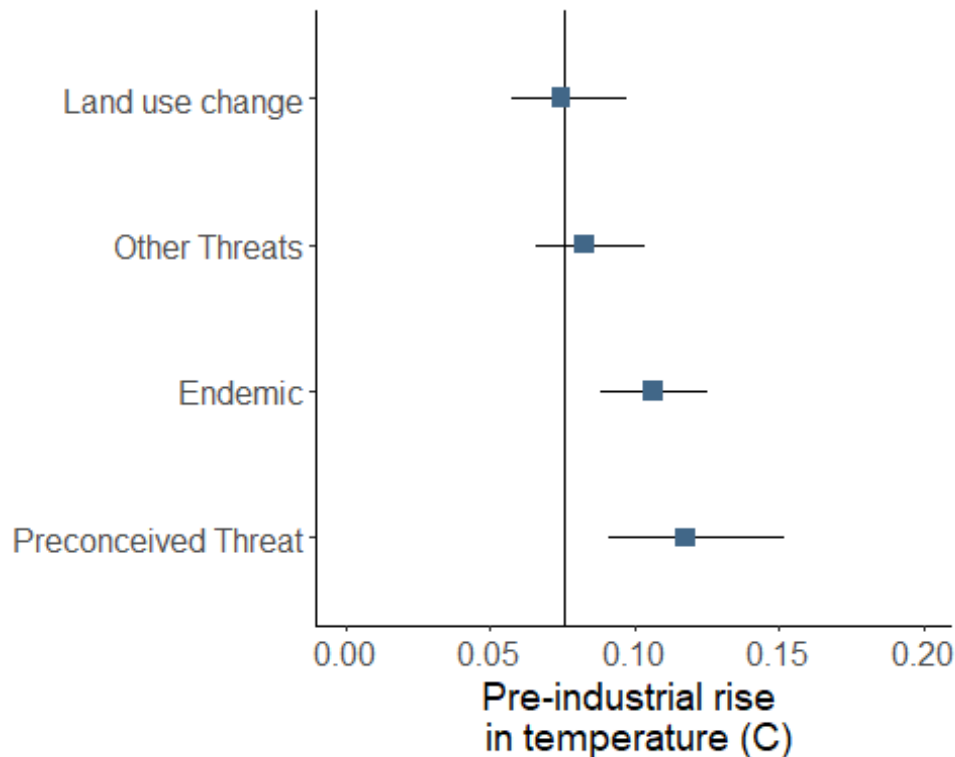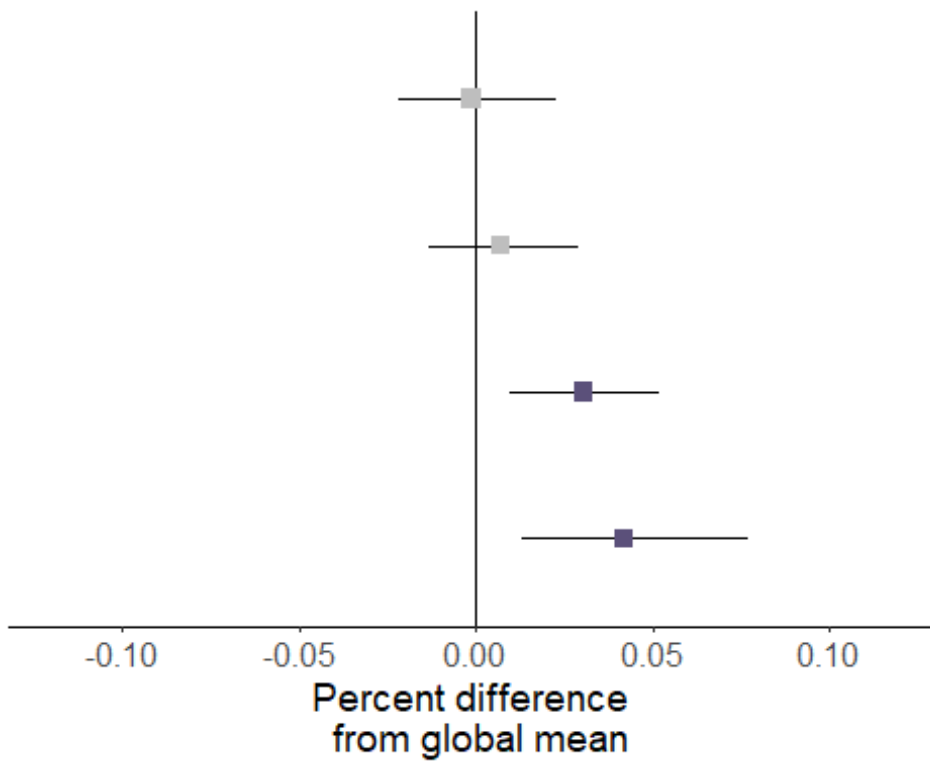*Fig. 1. Predicted extinction risk based on various threats.*

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

Fig2b <- ggplot(data = threat.vals)+
  geom_vline(xintercept=0) +
  geom_errorbar(aes(y = threat, xmin = beta.Y.mean.lo, xmax = beta.Y.mean.hi)
, width = 0) +
  geom_point(stat = "identity", aes(y = threat, 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.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)
Fig2b
```


Fig. 2. Predicted extinction risk based on various threats.

```
ggarrange(Fig2a, NULL, Fig2b, ncol=3, widths = c(4,.4, 2))
```
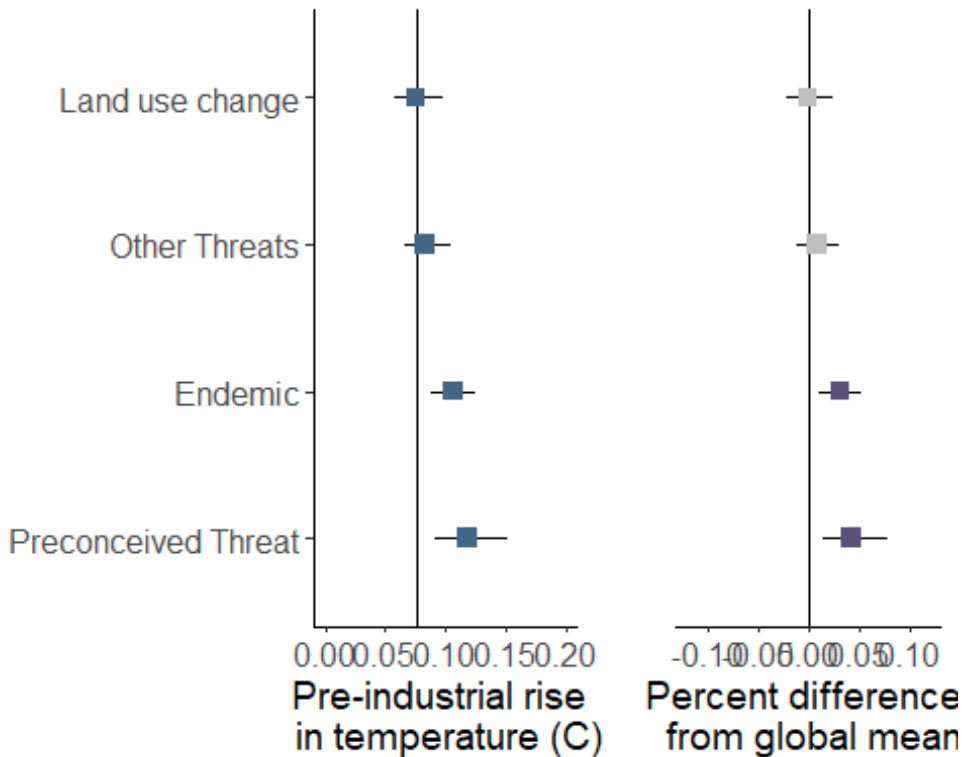
Fig. 3. Predicted extinction risk based on various threats.

```r
#ggsave("Metarisk2 threats.png",width=8,height=6,unit="in",dpi="print")

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

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

N.other.st<-length(unique(dataP$Study[dataP$Non.clim.threat == "Y"]))
N.other.mod<-length((dataP$Study[dataP$Non.clim.threat == "Y"]))

N.LU.st<-length(unique(dataP$Study[dataP$Land.Use.Change == "Y"]))
N.LU.mod<-length((dataP$Study[dataP$Land.Use.Change == "Y"]))

table.data<-data.frame(
  Factor = c("Preconceived","Endemic","Other","Land use change"),
  Studies = c(N.pre.st,N.end.st, N.other.st,N.LU.st),
  Models = c(N.pre.mod, N.end.mod, N.other.mod,N.LU.mod)
)
knitr::kable(table.data, caption = "Table 7: Number of studies and models for
each factor", format = "markdown")
```

Table 7: Number of studies and models for each factor

Factor	Studies	Models
Preconceived	127	802
Endemic	267	1676
Other	67	396
Land use change	39	216

Conclusion

Studies that focused on species that were considered to be threatened already or were endemic, indeed had higher extinction risks. However, contrary to expectations, models that included non-climate threats (usually land use change) and that modeled future land use changes did not indicate enhanced extinction risks.

Variation explained

```
#After Gelman 2019 R2 for Bayesian
#
#Load model and beta matrix - check if mu is modeled separately
load("3allthreat.rds")
posterior=as.data.frame(mod);
betamat <- data.frame(
  Intercept = rep(1,N),
  Threatened = ifelse(data.use$Threatened == "Y",1,0),
  Endemic = ifelse(data.use$Endemic == "Y",1,0),
  Non.clim.threat = ifelse(data.use$Non.clim.threat == "Y",1,0),
  Land.Use.Change = ifelse(data.use$Land.Use.Change == "Y",1,0))

#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 %*% t(betamat[i,]))#rows = samples, cols = i
     y.pred[,i] = (theta[,i] * data.use$Total.N[i])/(theta[,i] * data.use$To
tal.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 =  0.08371096 0.1121981 0.1526218

#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 =  0.7665317 0.7873023 0.8077524
```