GAM_spatial_SD_Prior_Simulation

Adam C. Smith

08/03/2022

Priors for the spatial variance in GAM parameters

The stratum-level population trajectories are fit using a spatially explicit, hierarchical GAM smooth. This spatial hierarchical structure allows the model to share information on the shape of the population trajectory among neighbouring strata and across the full range of the sites monitored for a given species. The variation among strata in the shape of the population trajectory is estimated using an intrinsic conditional autoregressive (iCAR) structure. This iCAR component of the model estimates the parameters of the GAM smooth as drawn from a normal distribution centered on the means of the parameter values in the neighbouring strata and with an estimated standard deviation (SD). The value of the SD controls the amount of variation in the shape of the trajectories among strata.

An appropriate prior for the SD parameter is not intuitive, because the GAM parameters do not directly reflect a biological process for which prior knowledge might provide an informative prior. Priors on variance parameters can be unintentionally informative if they put substantial prior mass at improbable levels of variation ((Gelman 2006; Lemoine 2019; Wesner and Pomeranz 2021)). Therefore, we conducted a prior simulation to translate the SD priors into a variation in long-term population trends, for which we do have some biological intuition. The simulation was designed to ensure that the prior on the spatial SD was largely uninformative, allowing for a reasonably broad range of long-term (1966-2019) population trend estimates across a species range, and that it did not contain substantial prior mass at highly improbable levels of variation.

We compared the variation in stratum-level trends under the alternative priors to observed variation in trends from a collection of realised long-term trend estimates from a different statistical model applied to the North American Breeding Bird Survey ((Link, Sauer, and Niven 2020)).

We compared three possible prior distributions for the Standard Deviation of the among-stratum variance on the GAM parameters.

1. half-normal

2. half-t-distribution with 3 degrees of freedom

3. gamma with shape parameter = 2

And, for each of the prior distributions, we compared 5 different values to set the prior-scale: (0.5, 1, 2, 3, and 4). Given the log-link in the trend model and the scaling of the low-rank thin-plate regression spline with identifiability constraints ((Wood 2020)), these 5-values of prior-scales should cover the range of plausible parameter values.

Spoiler

We suggest a half t-distribution, with a scale parameter = 1, fits the realised distributions of the SD in long-term trends among regions for most bird species surveyed by the BBS (i.e., most bird species with the best information on population trends at continental scales and for ~50 years). In addition, a half-normal prior with scale parameter = 2, or a gamma parameter with shape = 2 and scale = 2, seem to fit similarly well. The half-t prior results in prior distribution of SD in long-term trends that fit the realized distributions and includes long tails that cover the range of plausible trend estimates without including large amounts of prior probability mass at implausibly extreme values.
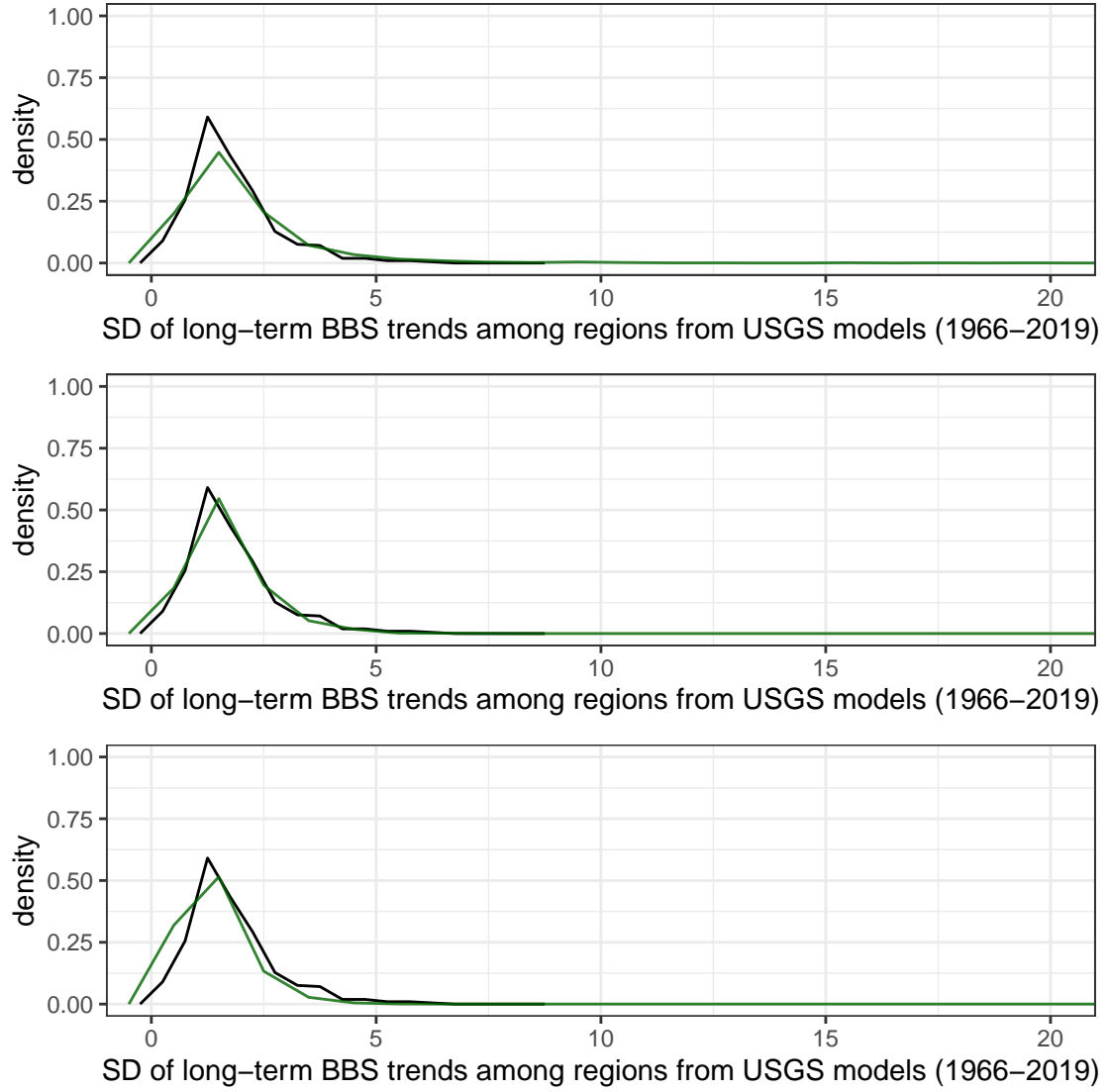
```
print(overp_t1/overp_gamma1/overp_norm1)
```

Figure 1: Observed distributions of the SD of long-term population trends from the BBS data using non-GAM models (in black), and the simulated prior distribution (in green) of long- and short-term trends from the spline smooth basis used in this paper, with a half-t (df = 3 and scale parameter = 2) prior distribution on the standard deviation of the spline parameters

Details on the prior simulations

Here's the code that runs the simulations in Stan.

```r
# this is all set-up
library(tidyverse)
library(cmdstanr)
library(mgcv)
library(patchwork)


for(pp in c("norm","t")){
  for(prior_scale in c(0.5,1,2,3,4)){

  tp = paste0("GAM_",pp,prior_scale,"_rate")

  #STRATA_True <- log(2)
  output_dir <- "output/"
  out_base <- tp
  csv_files <- paste0(out_base,"-",1:3,".csv")


  if(pp == "norm"){
    pnorm <- 1
  }
  if(pp == "t"){
    pnorm <- 2
  }

  if(!file.exists(paste0(output_dir,csv_files[1]))){

    nyears = 54 #to match the time-scales of BBS and CBC analyses
dat = data.frame(year = 1:nyears)
nknots = 13
nknots_realised = nknots-1

M = mgcv::smoothCon(s(year,k = nknots, bs = "tp"),data = dat,
                    absorb.cons=TRUE,#this drops the constant
                    diagonal.penalty=TRUE) ## If TRUE then the smooth is reparameterized to turn the


    year_basis = M[[1]]$X



    stan_data = list(#scalar indicators
      nyears = nyears,
```

```r
    #GAM structure
    nknots_year = nknots_realised,
    year_basis = year_basis,

    prior_scale = prior_scale,
    pnorm = pnorm
    )




  # Fit model ----------------------------------------------------------

  print(paste("beginning",tp,Sys.time()))

  mod.file = "models/GAM_prior_sim.stan"

  ## compile model
  model <- cmdstan_model(mod.file)


  # Initial Values -----------------------------------------------------


  init_def <- function(){
    list(sdbeta = runif(1,0.01,0.1),
         BETA_raw = rnorm(nknots_realised,0,0.01))}

  stanfit <- model$sample(
    data=stan_data,
    refresh=100,
    chains=2, iter_sampling=1000,
    iter_warmup=500,
    parallel_chains = 2,
    #pars = parms,
    adapt_delta = 0.8,
    max_treedepth = 14,
    seed = 123,
    init = init_def,
    output_dir = output_dir,
    output_basename = out_base)


 save(list = c("stanfit","stan_data","csv_files",
               "out_base"),
      file = paste0(output_dir,"/",out_base,"_gamye_iCAR.RData"))



  }

}#end prior_scale loop
}#end pp loop
```

5

The simulation model

The Stan simulation model is very simple. We implemented it in Stan to match the implementation in the full model (although this simulation doesn't require MCMC sampling).

```
// simple GAM prior simulation

data {
  int<lower=1> nyears;
  //scale of the prior distribution
  real<lower=0>  prior_scale;
  // indicator for the prior distribution 0 = t, 1 = normal
  int<lower=0,upper=1> pnorm;
// data for spline s(year)
  // number of knots in the basis function for year
  int<lower=1> nknots_year;
  // basis function matrix
  matrix[nyears, nknots_year] year_basis;
}

parameters {
  real<lower=0> sdbeta;     // sd of spline coefficients
  // unscaled spline coefficients
  vector[nknots_year] BETA_raw;
}

transformed parameters {
  vector[nknots_year] BETA;  //scaled spline parameters
  vector[nyears] smooth_pred;

    BETA = sdbeta * BETA_raw; //scaling the spline parameters
     smooth_pred = year_basis * BETA; //log-scale smooth trajectory
  }

model {

//Conditional statements to select the prior distribution
if(pnorm == 1){
 sdbeta ~ normal(0,prior_scale); //prior on sd of GAM parameter
}
if(pnorm == 0){
  sdbeta ~ student_t(3,0,prior_scale); //prior on sd of GAM parameter
}

   BETA_raw ~ normal(0,1); //non-centered parameterisation

}

 generated quantities {
  //estimated smooth on a count-scale
   vector[nyears] nsmooth = exp(smooth_pred);

  }
```

We then summarized the estimated trajectories as well as all possible 1-year, 10-year, and long-term trends from the alternative priors.

```
source("Functions/posterior_summary_functions.R") ## loads the saved simulation results

nsmooth_out <- NULL
trends_out <- NULL
summ_out <- NULL

for(pp in c("norm","t")){
  for(prior_scale in c(0.5,1,2,3,4)){

    tp = paste0("GAM_",pp,prior_scale,"_rate")

    #STRATA_True <- log(2)
    output_dir <- "output/"
    out_base <- tp


load(paste0(output_dir,"/",out_base,"_gamye_iCAR.RData"))

summ = stanfit$summary()

summ <- summ %>%
  mutate(prior_scale = prior_scale,
         distribution = pp)

nsmooth_samples <- posterior_samples(stanfit,
                                     parm = "nsmooth",
                                     dims = c("Year_Index"))



BETA_samples <- posterior_samples(stanfit,
                                  parm = "BETA",
                                  dims = c("k"))
BETA_wide <- BETA_samples %>%
  pivot_wider(.,id_cols = .draw,
              names_from = k,
              names_prefix = "BETA",
              values_from = .value)


nsmooth_samples <- nsmooth_samples %>%
  left_join(., BETA_wide,by = ".draw") %>%
  mutate(prior_scale = prior_scale,
         distribution = pp)



nyears = max(nsmooth_samples$Year_Index)
# function to calculate a %/year trend from a count-scale trajectory
trs <- function(y1,y2,ny){
  tt <- (((y2/y1)^(1/ny))-1)*100
```

```
}

for(tl in c(2,11,nyears)){ #estimating all possible 1-year, 10-year, and full trends
  ny = tl-1
  yrs1 <- seq(1,(nyears-ny),by = ny)
  yrs2 <- yrs1+ny
  for(j in 1:length(yrs1)){
    y2 <- yrs2[j]
    y1 <- yrs1[j]

nyh2 <- paste0("Y",y2)
nyh1 <- paste0("Y",y1)
trends <- nsmooth_samples %>%
  filter(Year_Index %in% c(y1,y2)) %>%
  select(.draw,.value,Year_Index) %>%
  pivot_wider(.,names_from = Year_Index,
              values_from = .value,
              names_prefix = "Y") %>%
  rename_with(.,~gsub(pattern = nyh2,replacement = "YE", .x)) %>%
  rename_with(.,~gsub(pattern = nyh1,replacement = "YS", .x)) %>%
  group_by(.draw) %>%
  summarise(trend = trs(YS,YE,ny))%>%
  mutate(prior_scale = prior_scale,
         distribution = pp,
         first_year = y1,
         last_year = y2,
         nyears = ny)
trends_out <- bind_rows(trends_out,trends)
}
}
nsmooth_out <- bind_rows(nsmooth_out,nsmooth_samples)
summ_out <- bind_rows(summ_out,summ)
print(paste(pp,prior_scale))

  }#prior_scale
}# pp

save(file = "output/GAM_prior_sim_summary.RData",
     list = c("nsmooth_out",
              "trends_out",
              "summ_out"))
```

Comparing simulation priors to realised data

Realised trend estimates

First, here is the distribution of long-term trends from a different model for the BBS data from 1966-2019, for 426 species.

```
bbs_trends_usgs <- read.csv("data/BBS_1966-2019_core_best_trend.csv")
```

```
## selecting survey-wide trend estimates
bbs_trends_usgs_long <-bbs_trends_usgs %>%
  filter(Region == "SU1")%>%
  select(Trend,Species.Name) %>%
  mutate(abs_trend = abs(Trend)) %>% #calculating absolute values of the trends
  arrange(-abs_trend)

G_long_usgs <- max(bbs_trends_usgs_long$abs_trend)

realised_long_bbs_hist <- ggplot(data = bbs_trends_usgs_long,
                          aes(abs_trend,after_stat(density)))+
  geom_freqpoly(breaks = seq(0,13,0.5),center = 0)+
  xlab("Absolute value of long-term BBS trends USGS models (1966-2019)")+
  theme_bw()+
  scale_y_continuous(limits = c(0,1))
print(realised_long_bbs_hist)
```
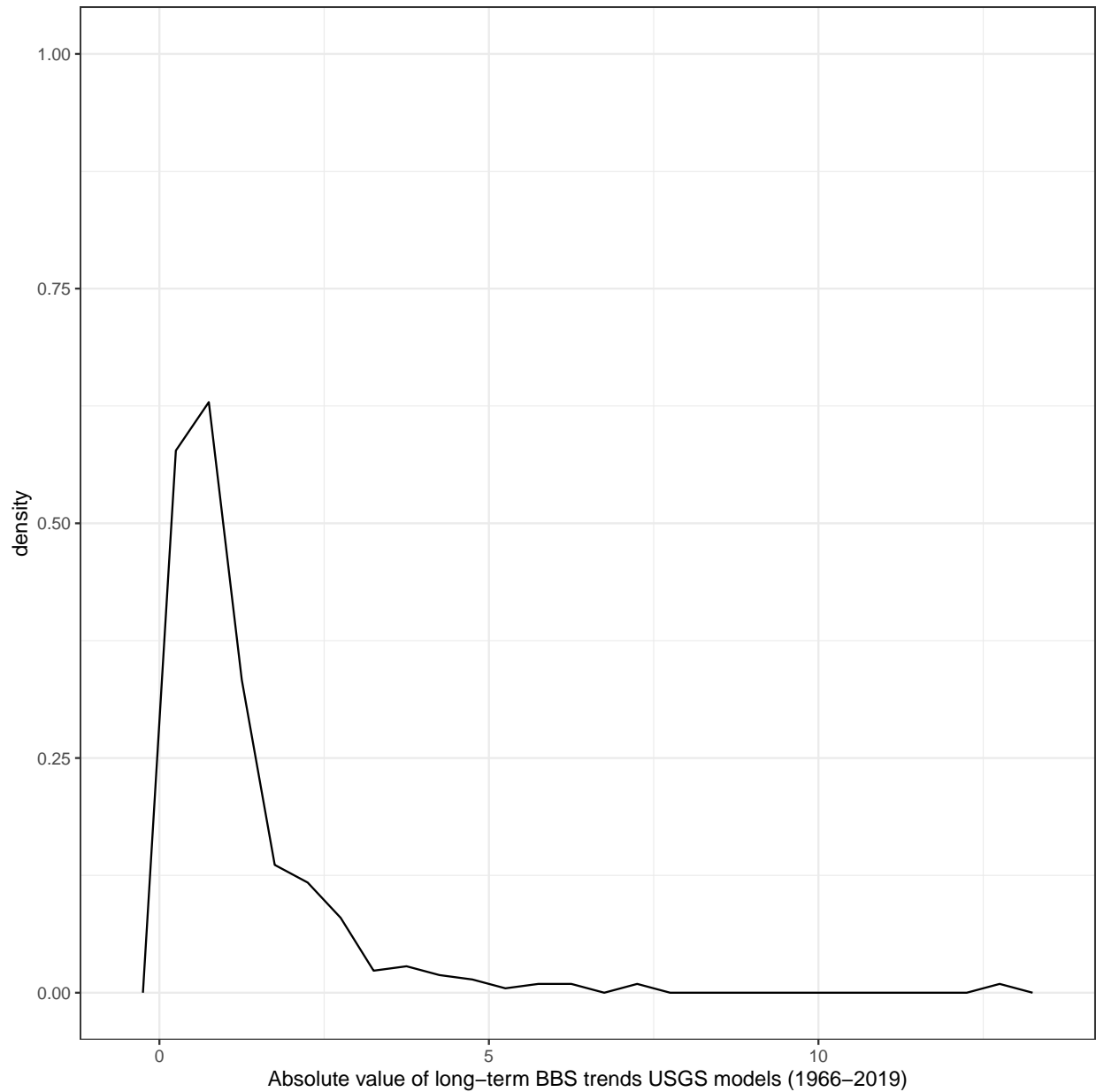
The maximum absolute value of an observed long-term trends are for Cave Swallow and Eurasian Collared Dove, which have increased at an annual rate of 12.9 %/year. This annual rate of change implies an approximate 63000 % overall increase in the populations since 1966. As such, we feel this represents example of an "extreme" long-term trend that is unlikely to be observed in most of the BBS dataset. For example, the next largest values of trend in the data is a 7.3 %/year increase in Canada Goose populations.

```
kable(bbs_trends_usgs_long[1:3,], booktabs = TRUE,
      digits = 3) %>%
  kable_styling(font_size = 8)
```

And the largest absolute value of trends for a declining species is < 4%/year for King Rail, Bank Swallow, and Lark Bunting.

Trend	Species.Name	abs_trend
12.92	Cave Swallow	12.92
12.91	Eurasian Collared-Dove	12.91
7.27	Canada Goose	7.27

Trend	Species.Name	abs_trend
-3.82	King Rail	3.82
-3.72	Bank Swallow	3.72
-3.66	Lark Bunting	3.66

```r
bbs_declines_usgs_long <-bbs_trends_usgs_long %>%
  filter(Trend < 0)%>%
  arrange(-abs_trend)

kable(bbs_declines_usgs_long[1:3,], booktabs = TRUE,
      digits = 3) %>%
  kable_styling(font_size = 8)
```

Summarize the long-term trends from the prior simulations

```r
load("output/GAM_prior_sim_summary.RData")

trends_out <- trends_out %>%
  mutate(abs_trend = abs(trend),# absolute values of trends
         scale_factor = factor(prior_scale,ordered = TRUE)) #just for plotting

trends_long <- trends_out %>%  #select long-term trends only
  filter(nyears == 53 )

trends_normal <- trends_long %>%
  filter(distribution == "norm")
```

The distributions for the normal priors do a reasonable job of covering the range of possible long-term trend values, but the heavier-tailed t-distributions seem to better fit the shape of the distributions.

```r
  overp_normal <- realised_long_bbs_hist +
    geom_freqpoly(data = trends_normal,
                  aes(abs_trend,after_stat(density),
                      colour = scale_factor),
                  breaks = seq(0,100,1),center = 0,
                  alpha = 0.8)+
    scale_colour_viridis_d(begin = 0.5,alpha = 0.8,
                           "Prior Scale\nSD half-normal")+
  coord_cartesian(xlim = c(0,20))


  trends_t <- trends_long %>%
  filter(distribution == "t")
```

```
overp_t <- realised_long_bbs_hist +
  geom_freqpoly(data = trends_t,
                aes(abs_trend,after_stat(density),
                    colour = scale_factor),
                breaks = seq(0,100,1),center = 0,
                alpha = 0.8)+
  scale_colour_viridis_d(begin = 0.5,alpha = 0.8,
                         "Prior Scale\nSD half-t df=3")+
  coord_cartesian(xlim = c(0,20))
```
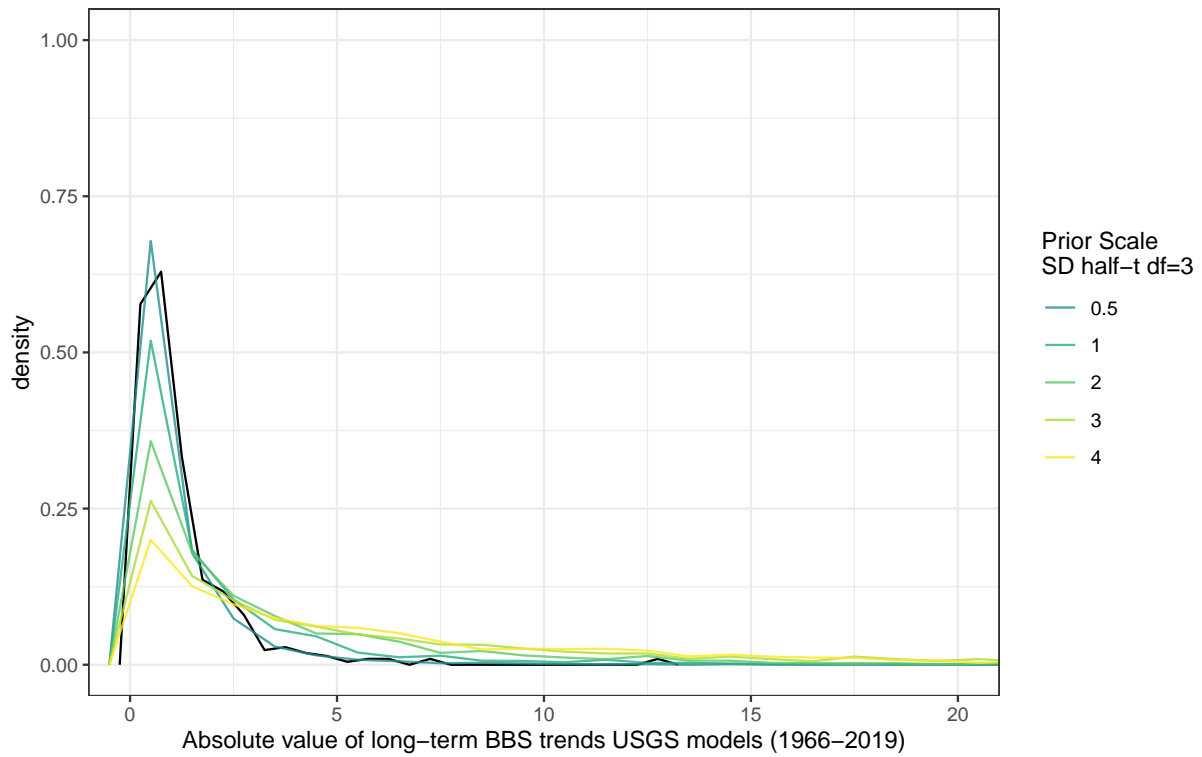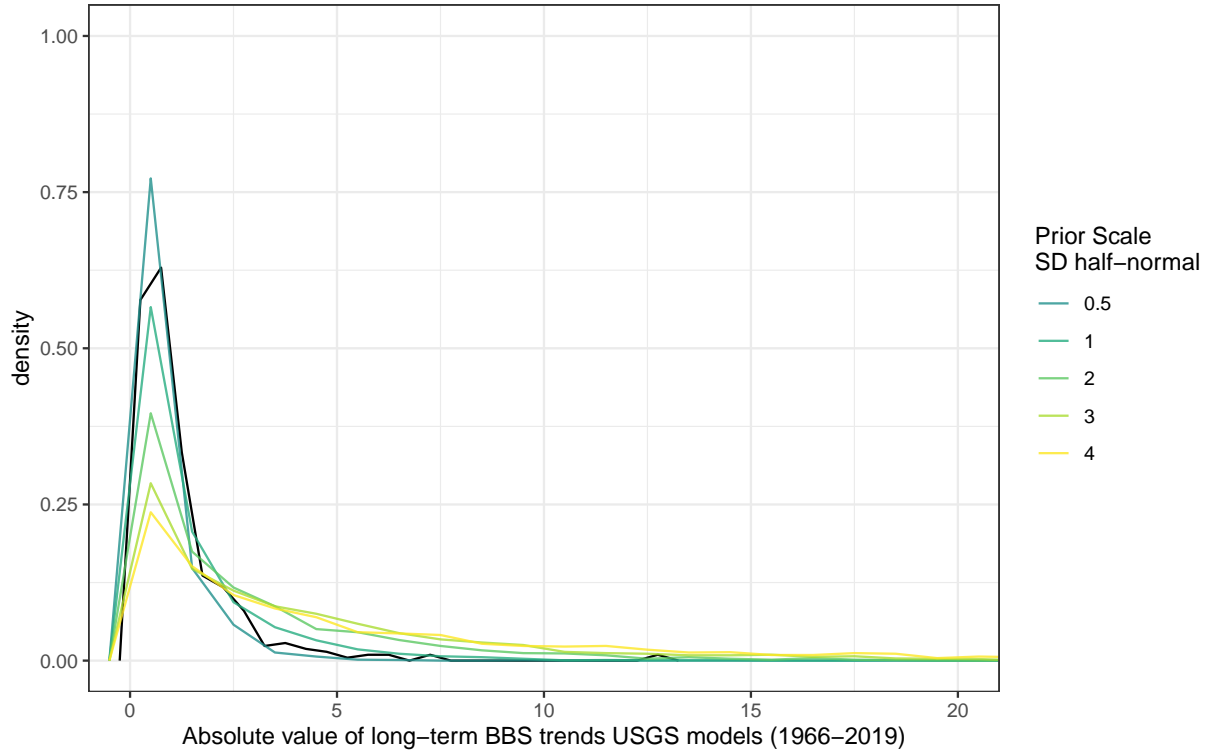
The half-t-distributions with a scale value of 2 or 3 fit the shape of the realised trend distribution reasonably well, and the long-tail includes significant prior mass at and beyond the observed maximum absolute values of trends.

```
print(overp_normal / overp_t)
```

In addition, each of the prior distributions include some prior mass at trend values beyond the realised maximum values. And the broadest priors (scales of 3 or 4), include prior mass at trend values that are extremely unlikely for a wild population to sustain over >50 years. For example, more than 10% of the two widest distributions are at values larger than the largest observed value of long-term trends, and the 99th

distribution	prior_scale	median_abs_t	U80	U90	U99	pGTmax
norm	0.5	0.403	1.123	1.777	3.686	0.000
norm	1.0	0.790	2.201	3.584	8.115	0.002
norm	2.0	1.465	4.383	6.800	16.521	0.021
norm	3.0	2.570	6.712	10.186	27.547	0.067
norm	4.0	3.056	8.876	13.785	32.492	0.112
t	0.5	0.524	1.590	2.439	7.597	0.002
t	1.0	0.927	2.863	4.746	15.990	0.016
t	2.0	1.743	5.549	8.890	30.854	0.049
t	3.0	2.947	9.169	14.859	39.754	0.122
t	4.0	4.087	11.833	18.989	52.542	0.174

percentiles extend to values that are truly extreme (40 - 50%/year).

```
quant_long_tends <- trends_long %>%
  group_by(distribution,prior_scale) %>%
  summarise(median_abs_t = median(abs_trend),
            U80 = quantile(abs_trend,0.80),
            U90 = quantile(abs_trend,0.90),
            U99 = quantile(abs_trend,0.99),
            pGTmax = length(which(abs_trend > G_long_usgs))/length(abs_trend))
```

```
## `summarise()` has grouped output by 'distribution'. You can override using the `.groups` argument.
```

```
kable(quant_long_tends, booktabs = TRUE,
      digits = 3) %>%
  kable_styling(font_size = 8)
```

Short-term trends

The maximum observed short-term trend from the same USGS analysis of BBS trends is for Eurasian Collared-dove, which have a ten-year trend rate of 23.5%/year. Although the full set of short-term trends is not available for download, this extreme rate of short-term increase provides a useful benchmark against which we can compare the prior distributions of short-term trends. This ten-year trend rate implies a 730 % overall increase in the population over this time. This is an extreme increase over a short-period of time.

```
G_short_usgs <- 23.5 #Eurasian Collared Dove trend for short-term 2009-2019 analysis USGS
    # short-term trends not included in Science Base, but visible here:
    # https://www.mbr-pwrc.usgs.gov/bbs/reglist19v3.shtml
```

We can also use archived estimates from the Canadian Wildlife Service for 1970 - 2017, which were derived from a different model than the GAM used here.

```
bbs2017_Canada <- read.csv("data_basic/All BBS trends 2017 w reliab.csv")

bbs_short_survey_wide <- bbs2017_Canada %>%
  filter(region.type == "Continental",
         trendtype == "short-term",
         startyear == 2007,
         trendtime == "full") %>%
```

```
  select(species,trend) %>%
  mutate(abs_trend = abs(trend)) %>%
  arrange(-abs_trend)
```

The realised distribution of short-term trends from this archived analysis shows that short-term trends are generally more extreme than long-term trends.

```
realised_short_bbs_hist <- ggplot(data = bbs_short_survey_wide,
                           aes(abs_trend,after_stat(density)))+
  geom_freqpoly(breaks = seq(0,20,0.5),center = 0)+
  xlab("Absolute value of long-term BBS trends CWS model (2007-2017)")+
  theme_bw()+
  scale_y_continuous(limits = c(0,1))
print(realised_short_bbs_hist)
```
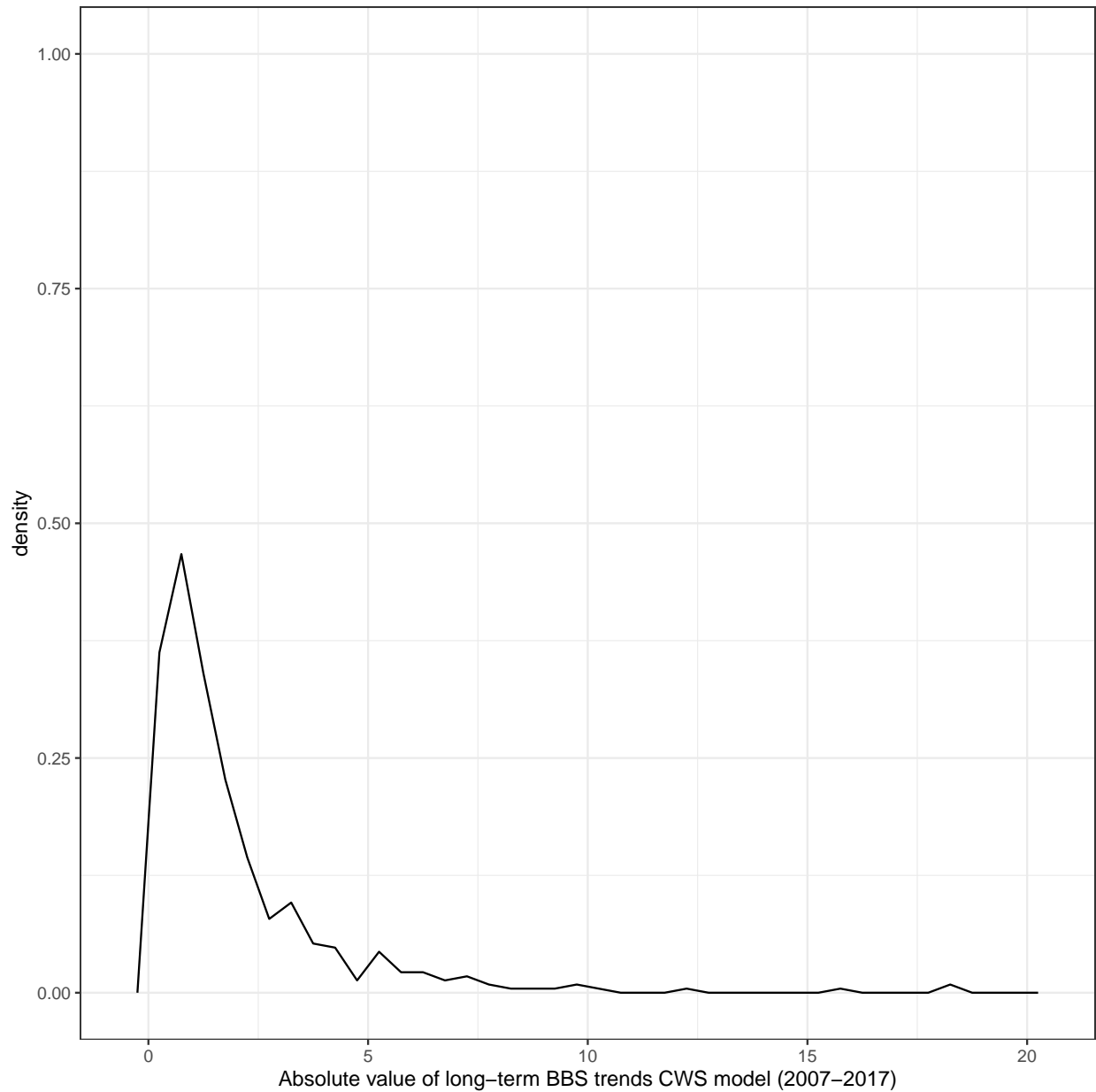
The short-term trends from the prior distributions represent all possible ten-year trends, not just the final ten-years of the trajectory.

```
trends_short <- trends_out %>%  #select long-term trends only
  filter(nyears == 10 )

trends_short_normal <- trends_short %>%
  filter(distribution == "norm")

trends_short_t <- trends_short %>%
  filter(distribution == "t")
```

Overplotting the simulation prior distributions of short-term trends with the realised collection of short-term trends.

```
overp_normal_short <- realised_short_bbs_hist +
  geom_freqpoly(data = trends_short_normal,
                aes(abs_trend,after_stat(density),
                    colour = scale_factor),
                breaks = seq(0,100,1),center = 0,
                alpha = 0.8)+
  scale_colour_viridis_d(begin = 0.5,alpha = 0.8,
                         "Prior Scale\nSD half-normal")+
coord_cartesian(xlim = c(0,20))




overp_t_short <- realised_short_bbs_hist +
  geom_freqpoly(data = trends_short_t,
                aes(abs_trend,after_stat(density),
                    colour = scale_factor),
                breaks = seq(0,100,1),center = 0,
                alpha = 0.8)+
  scale_colour_viridis_d(begin = 0.5,alpha = 0.8,
                         "Prior Scale\nSD half-t df=3")+
coord_cartesian(xlim = c(0,20))
```
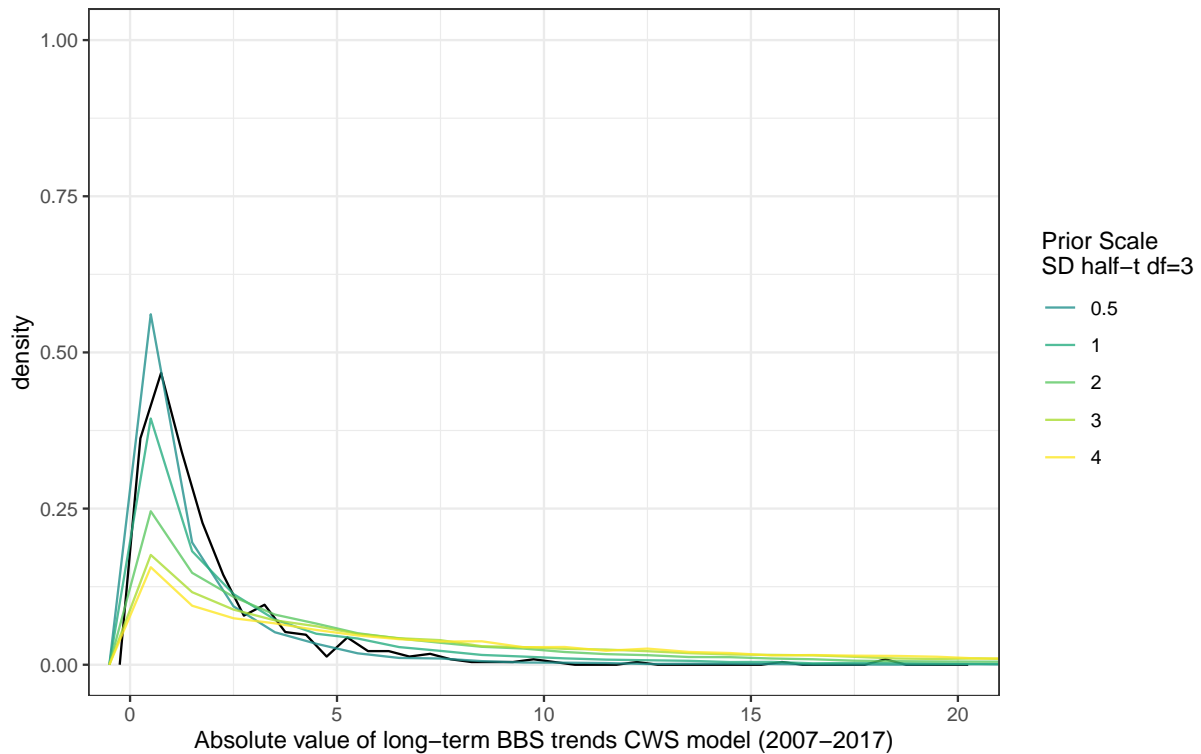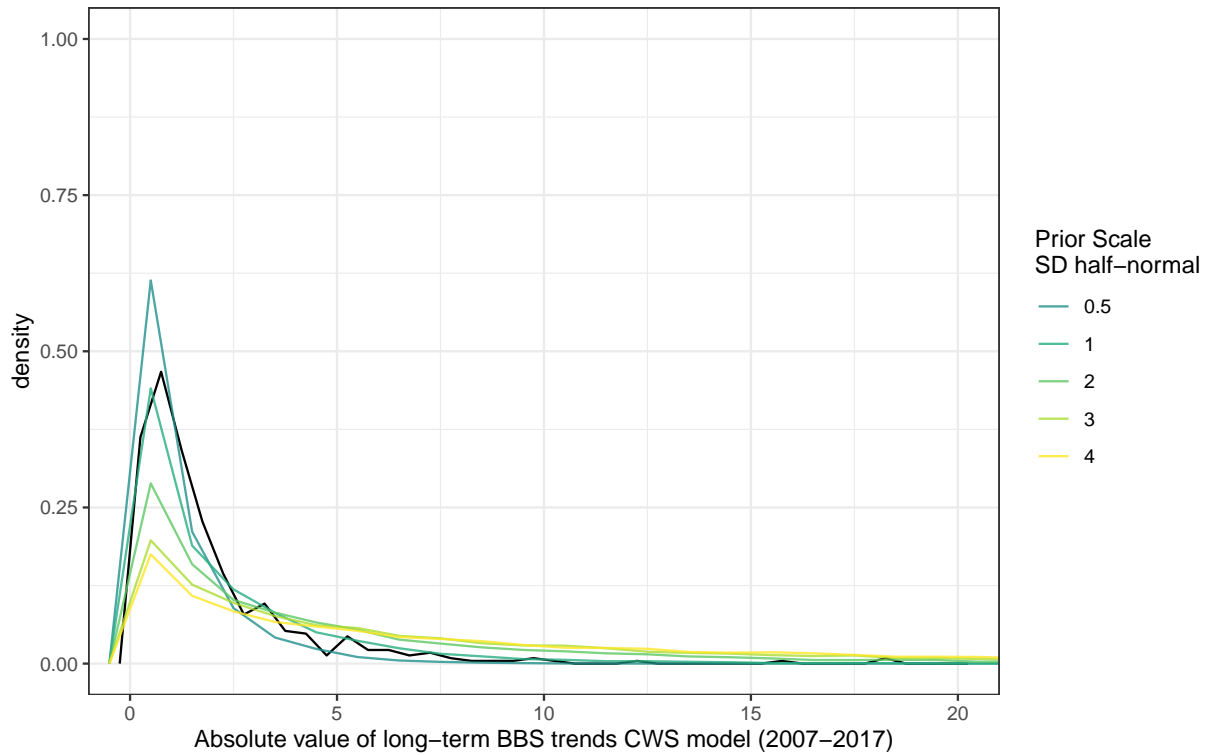
The t-distribution scale parameter $= 2$, does a reasonably good job of covering the range of realistic short-term trends in the BBS data, and includes a long tail that includes much more extreme values when supported by the data. The distributions with larger scale parameters place significant prior density at rates of short-term change that are highly improbable at continental scales. For example, approximately 10% of the two widest distributions are at values larger than the largest observed value, and the 99th percentiles extend to values that are truly extreme (64 - 92%/year).

```
print(overp_normal_short / overp_t_short)
```

Absolute value of long−term BBS trends CWS model (2007−2017)

Absolute value of long−term BBS trends CWS model (2007−2017)

Below, you can see that the half-t-distribution with scale parameter = 2 has a long-tail that will allow for extreme trends when supported by the data (e.g., it has approximately 5% of the prior distribution covering values larger than the maximum observed value and a 99th percentile far beyond the most extreme observed value).

distribution	prior_scale	median_abs_t	U80	U90	U99	pGTmax
norm	0.5	0.687	1.823	2.818	6.123	0.000
norm	1.0	1.257	3.589	5.465	13.168	0.000
norm	2.0	2.480	7.281	11.498	26.876	0.015
norm	3.0	4.052	11.359	17.401	41.275	0.052
norm	4.0	5.178	14.872	22.987	57.107	0.097
t	0.5	0.796	2.395	3.953	11.813	0.001
t	1.0	1.526	4.738	7.834	26.698	0.012
t	2.0	2.995	8.872	14.497	48.119	0.041
t	3.0	4.832	14.696	23.364	64.604	0.099
t	4.0	6.237	18.549	29.864	92.292	0.147

```
quant_short_tends <- trends_short %>%
  group_by(distribution,prior_scale) %>%
  summarise(median_abs_t = median(abs_trend),
            U80 = quantile(abs_trend,0.80),
            U90 = quantile(abs_trend,0.90),
            U99 = quantile(abs_trend,0.99),
            pGTmax = length(which(abs_trend > G_short_usgs))/length(abs_trend))
```

```
## `summarise()` has grouped output by 'distribution'. You can override using the `.groups` argument.
```

```
kable(quant_short_tends, booktabs = TRUE,
      digits = 3) %>%
  kable_styling(font_size = 8)
```

References

Gelman, Andrew. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)." *Bayesian Analysis* 1 (3): 515–34. https://doi.org/10.1214/06-BA117A.

Lemoine, Nathan P. 2019. "Moving Beyond Noninformative Priors: Why and How to Choose Weakly Informative Priors in Bayesian Analyses." *Oikos* 128 (7): 912–28. https://doi.org/10.1111/oik.05985.

Link, William A., John R. Sauer, and Daniel K. Niven. 2020. "Model Selection for the North American Breeding Bird Survey." *Ecological Applications* 30 (6): e02137. https://doi.org/10.1002/eap.2137.

Wesner, Jeff S., and Justin P. F. Pomeranz. 2021. "Choosing Priors in Bayesian Ecological Models by Simulating from the Prior Predictive Distribution." *Ecosphere* 12 (9): e03739. https://doi.org/10.1002/ecs2.3739.

Wood, Simon N. 2020. "Inference and Computation with Generalized Additive Models and Their Extensions." *TEST* 29 (2): 307–39. https://doi.org/10.1007/s11749-020-00711-5.