

Practical 4: Linear models

Benjamin Rosenbaum

Table of contents

Regression, additive	2
Regression, interaction	13
One categorical predictor	19
Two categorical predictors	25
Additive model	26
Interaction model	30
Categorical and continuous predictors	33
Additive model	34
Interaction model	38
Exercise	42

We learn about linear models with continuous or categorical predictors, namely linear regression, ANOVA, ANCOVA

Research questions are answered via model selection (LOO), but also with comparison of posterior predictions (counterfactuals, “what-if” scenarios). With categorical predictors, the emmeans package is helpful here.

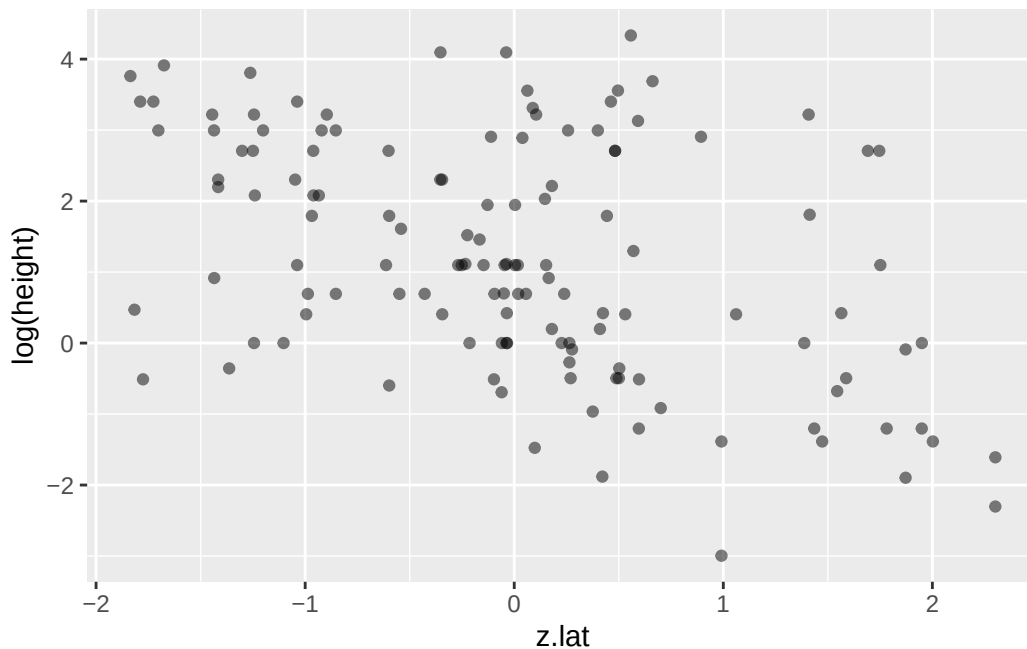
```
rm(list=ls())
library("brms")
library("bayesplot")
library("performance")
library("ggplot2")
library("emmeans")
library("ecostats")
library("Data4Ecologists")
```

```
library("cowplot")
try(dev.off())
```

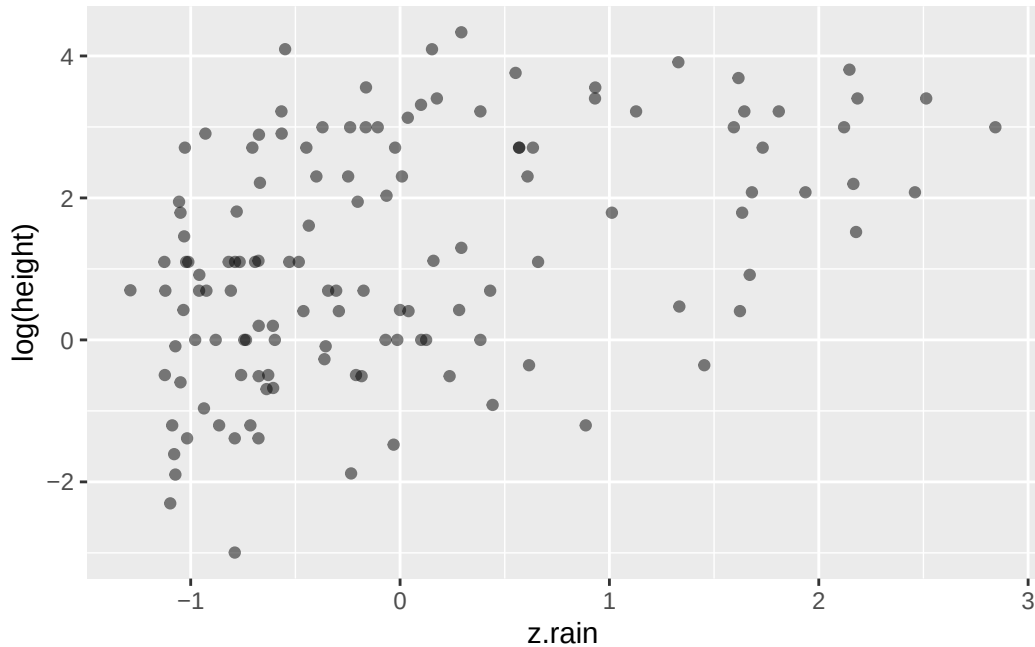
Regression, additive

We use the same global plants dataset as before (from the `ecostats` package). We use an additional predictor rainfall and we scale both predictors (`scale()`, `mean=0`, `sd=1`). This makes things easier especially when involving interactions.

```
data(globalPlants)
globalPlants$z.lat = scale(globalPlants$lat)
globalPlants$z.rain = scale(globalPlants$rain)
ggplot(globalPlants, aes(z.lat, log(height))) + geom_point(alpha=0.5)
```



```
ggplot(globalPlants, aes(z.rain, log(height))) + geom_point(alpha=0.5)
```



Here, we want to examine the latitudinal gradient in plant height, while controlling for rainfall.

Deterministic part: $\mu = b_0 + b_1 \cdot lat + b_2 \cdot rain$
 Stochastic part: $\log(height) \sim \text{Normal}(\mu, \sigma)$

We use vaguely informative priors, we expect a negative relation with latitude, a positive one with rainfall.

```
fit.lm.add = brm(log(height) ~ z.lat + z.rain,
  prior =
    prior(normal(-1,1), class=b, coef=z.lat) +
    prior(normal(+1,1), class=b, coef=z.rain),
  data = globalPlants)
```

We check for convergence as usual, everything OK here.

```
summary(fit.lm.add, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: log(height) ~ z.lat + z.rain
Data: globalPlants (Number of observations: 131)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
```

```
total post-warmup draws = 4000
```

Priors:

```
b_z.lat ~ normal(-1, 1)
```

```
b_z.rain ~ normal(+1, 1)
```

```
Intercept ~ student_t(3, 1.1, 2.5)
```

```
<lower=0> sigma ~ student_t(3, 0, 2.5)
```

Regression Coefficients:

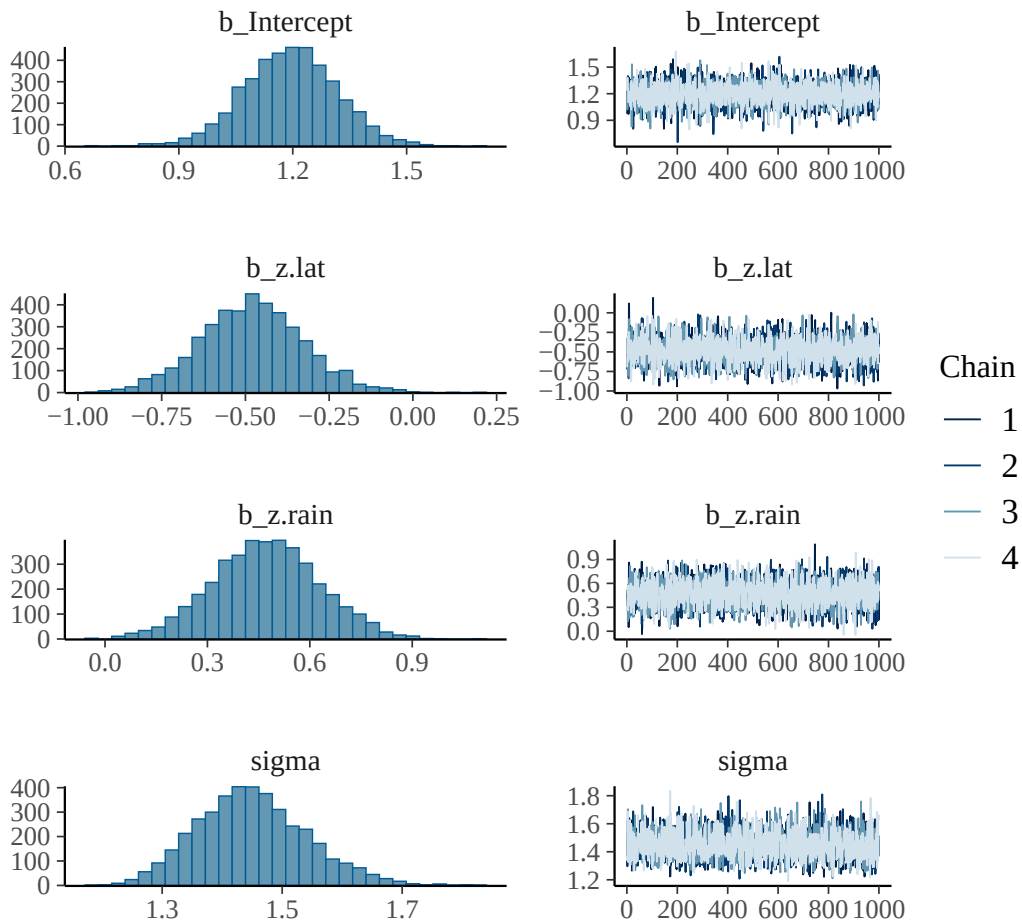
	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.19	0.12	0.95	1.43	1.00	3766	2606
z.lat	-0.48	0.15	-0.79	-0.18	1.00	3259	2543
z.rain	0.47	0.16	0.16	0.77	1.00	3431	2928

Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.45	0.09	1.29	1.64	1.00	3632	2694

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

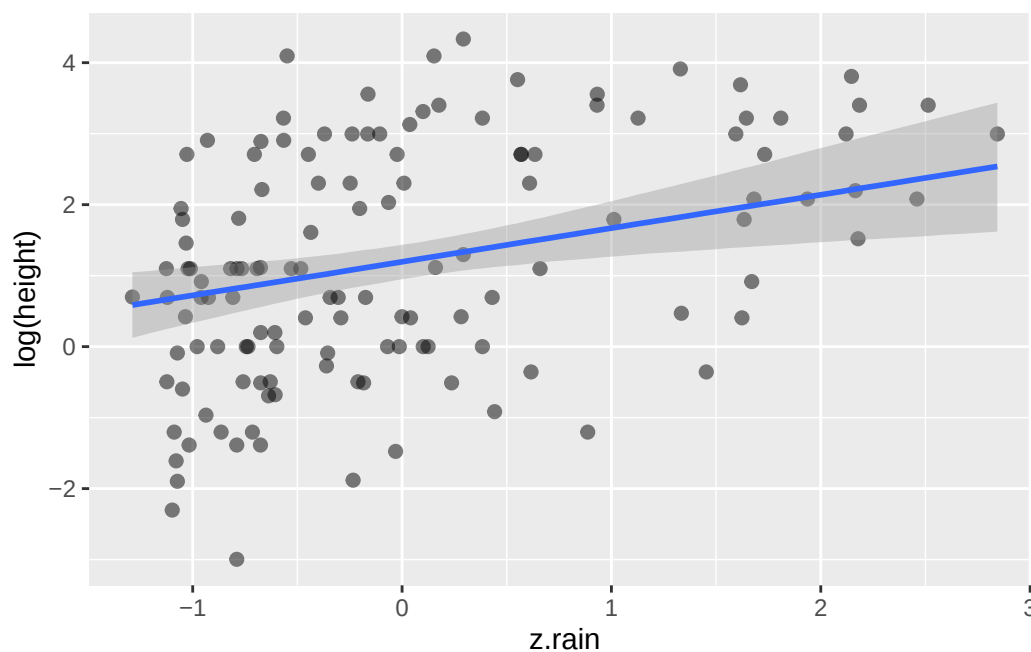
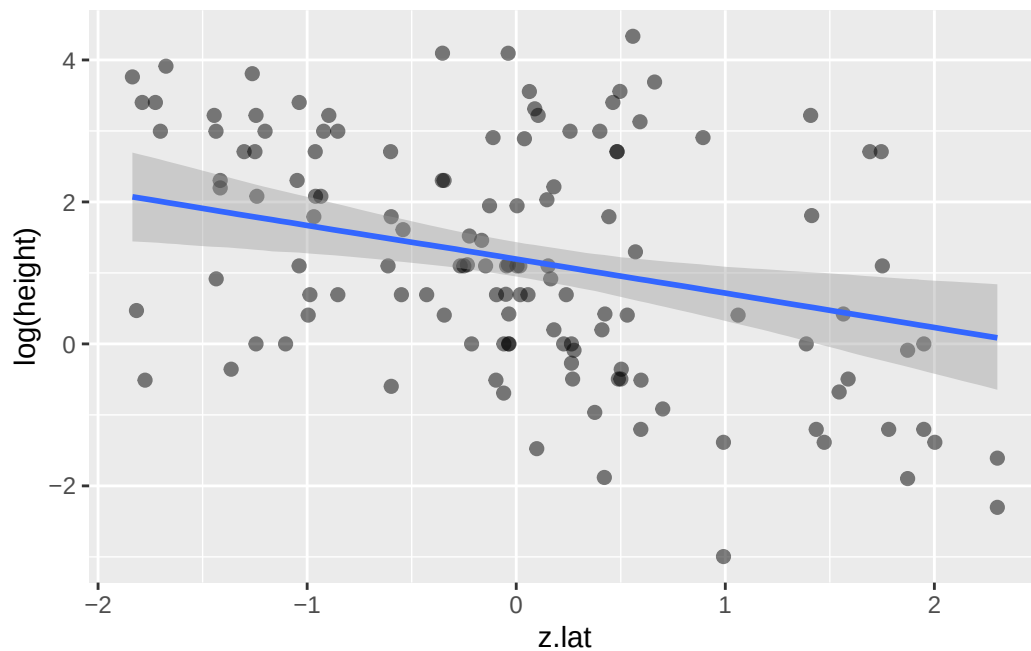
```
plot(fit.lm.add)
```



Now we check predictions, to evaluate if the model fits the data well.

`conditional_effects()` will plot predictions against each predictor, while the other one is held constant at its mean (here =0 because we scaled it).

```
plot(conditional_effects(fit.lm.add),
     points=TRUE,
     point_args=c(alpha=0.5),
     ask=FALSE)
```



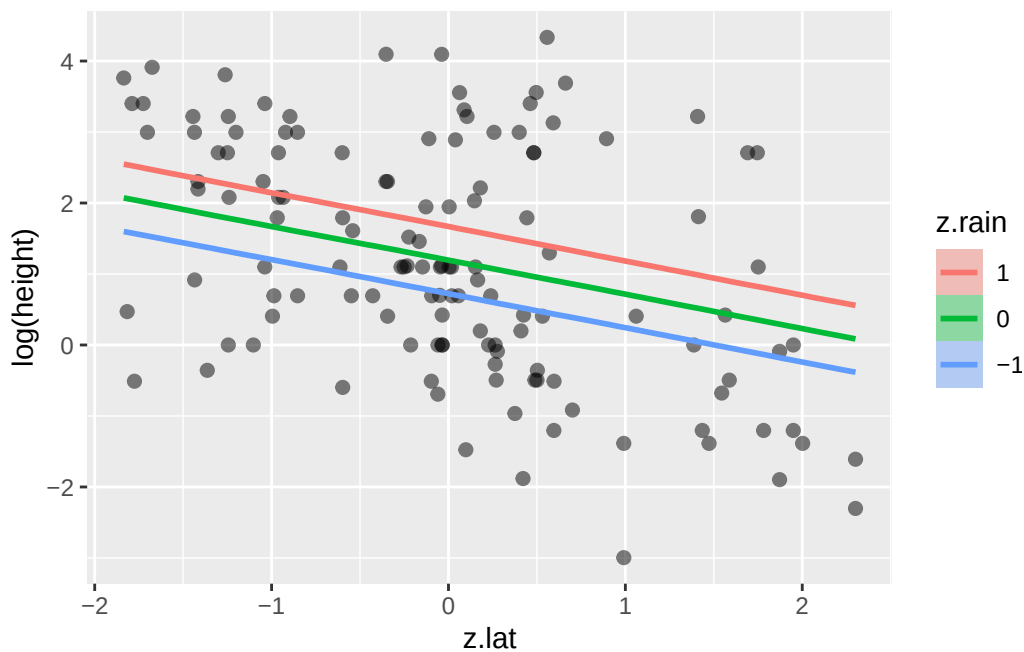
You can also specify the predictor with `effects=""`, if you want to plot them separately.

```
plot(conditional_effects(fit.lm.add, effects="z.lat"),
     points=TRUE,
     point_args=c(alpha=0.5))
plot(conditional_effects(fit.lm.add, effects="z.rain"),
```

```
points=TRUE,
point_args=c(alpha=0.5))
```

Although the model does not contain an interaction, "z.lat:z.rain" will plot fitted effects of z.lat for 3 levels of z.rain (mean-1sd, mean, mean+1sd). prob=... chooses the quantiles of model uncertainty. With prob=0 we do not plot any uncertainty and just the mean regression line for better visibility.

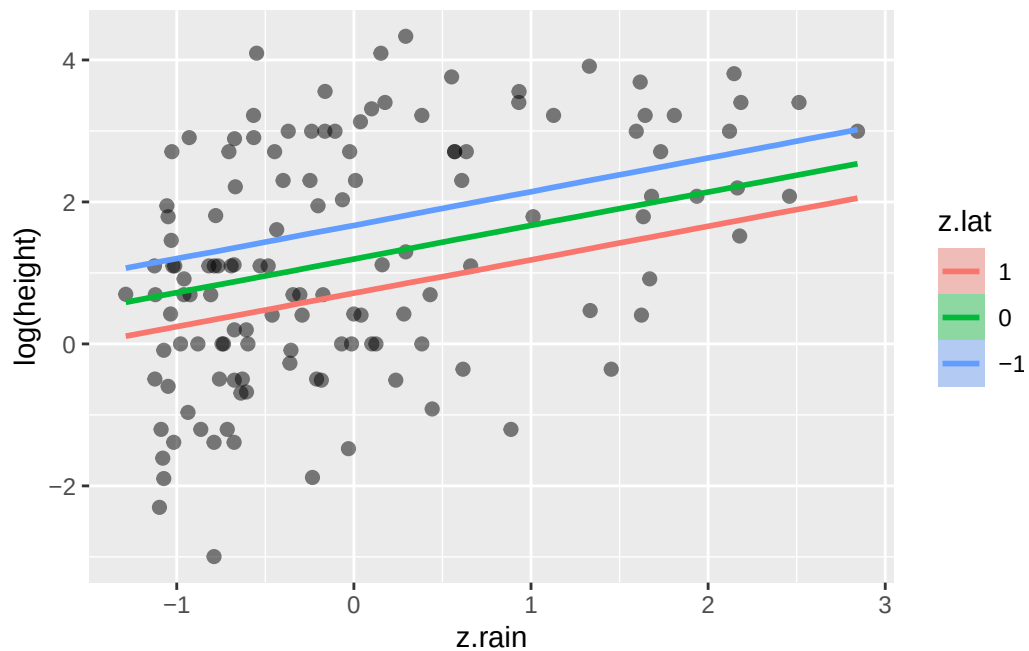
```
plot(conditional_effects(fit.lm.add, effects="z.lat:z.rain", prob=0),
     points=TRUE,
     point_args=c(alpha=0.5))
```



Note that lines are always parallel in an additive model.

We can also plot the effects of the 2nd predictor z.rain for 3 levels of z.lat by switching the order "z.rain:z.lat".

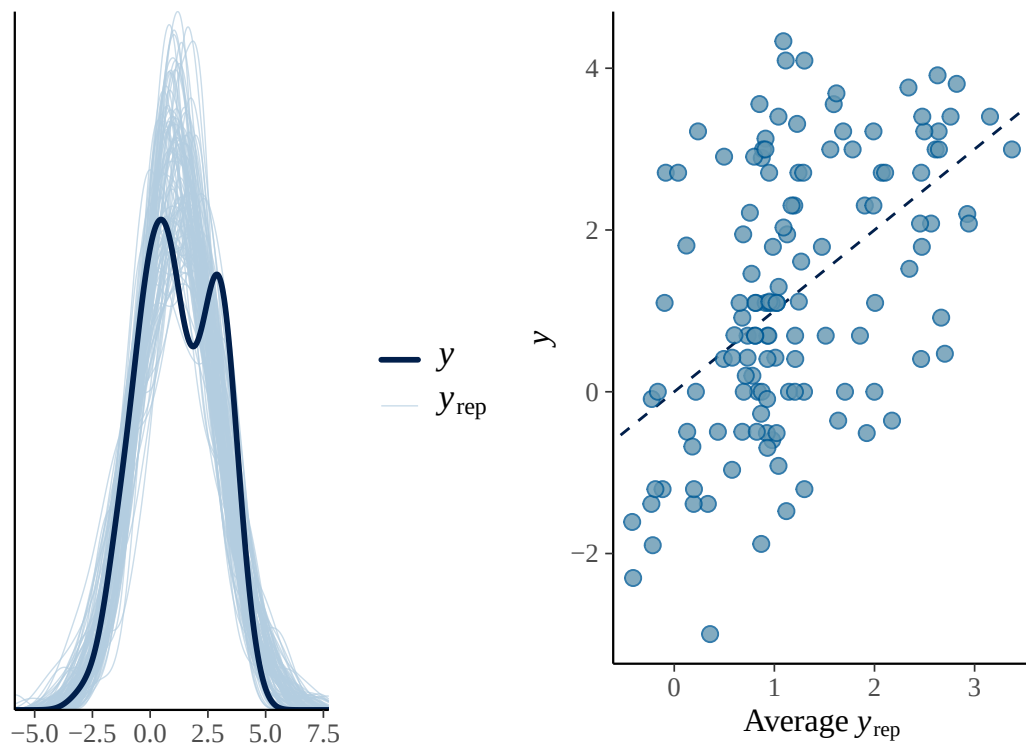
```
plot(conditional_effects(fit.lm.add, effects="z.rain:z.lat", prob=0),
     points=TRUE,
     point_args=c(alpha=0.5))
```



With more than one predictor, it's getting more difficult to assess the quality of model fit from these plots.

Posterior predictive plots, on the other hand, are independent from the number of predictor variables.

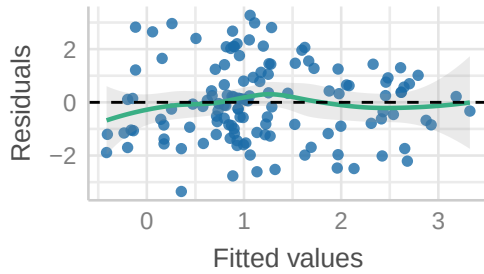
```
p1 = pp_check(fit.lm.add, ndraws=100)
p2 = pp_check(fit.lm.add, type="scatter_avg")
plot_grid(p1,p2)
```

```
check_model(fit.lm.add, check=c("linearity","homogeneity","qq","normality"))
```

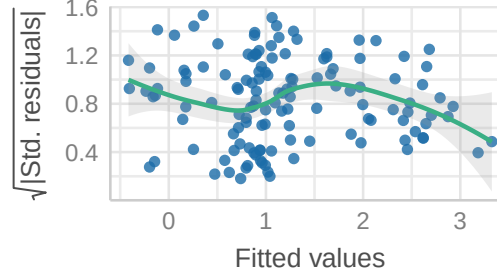
Linearity

Reference line should be flat and horizontal



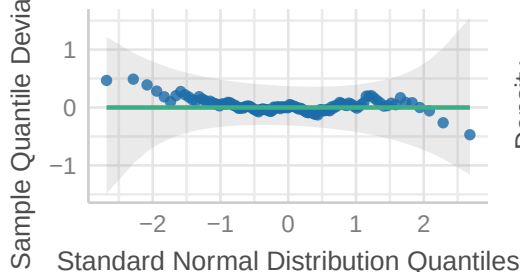
Homogeneity of Variance

Reference line should be flat and horizontal



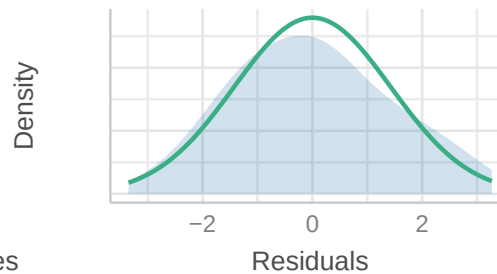
Normality of Residuals

Points should fall along the line



Normality of Residuals

Distribution should be close to the normal cu



There is still a lot of unexplained variation, but at least linear model assumptions seem to be satisfied.

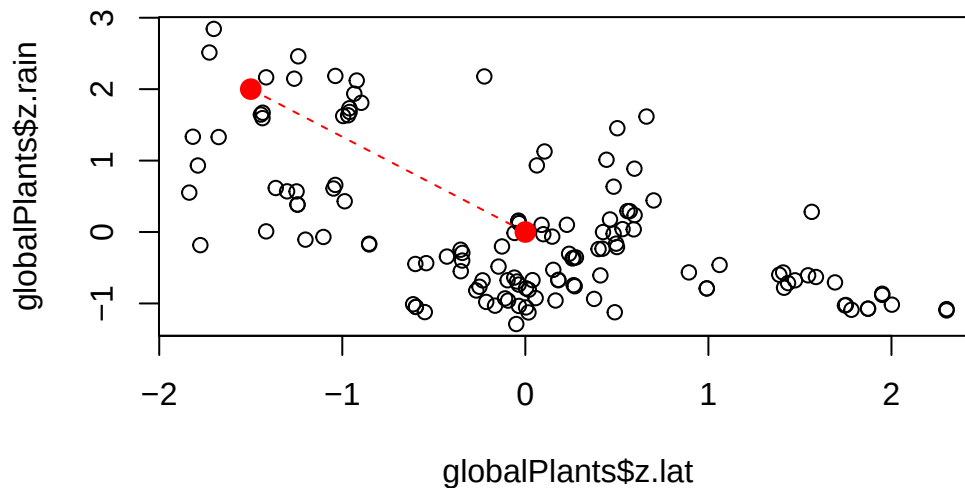
```
bayes_R2(fit.lm.add)
```

	Estimate	Est.Error	Q2.5	Q97.5
R2	0.2603617	0.05539821	0.1479813	0.3644788

Question: Is the average plant height at mean latitude and mean rainfall different from a tropical scenario (close to equator, high rainfall).

We can make **counterfactual predictions** for these 2 scenarios to answer this question.

```
plot(globalPlants$z.lat, globalPlants$z.rain)
points(0,0,col="red", pch=16, cex=1.5)
points(-1.5,2,col="red", pch=16, cex=1.5)
lines(c(-1.5,0),c(2,0), col="red", lty=2)
```



Computing predictions is not enough to make a quantitative statement on the question.

```
fitted(fit.lm.add, newdata=data.frame(z.lat=-1.5, z.rain=2))
```

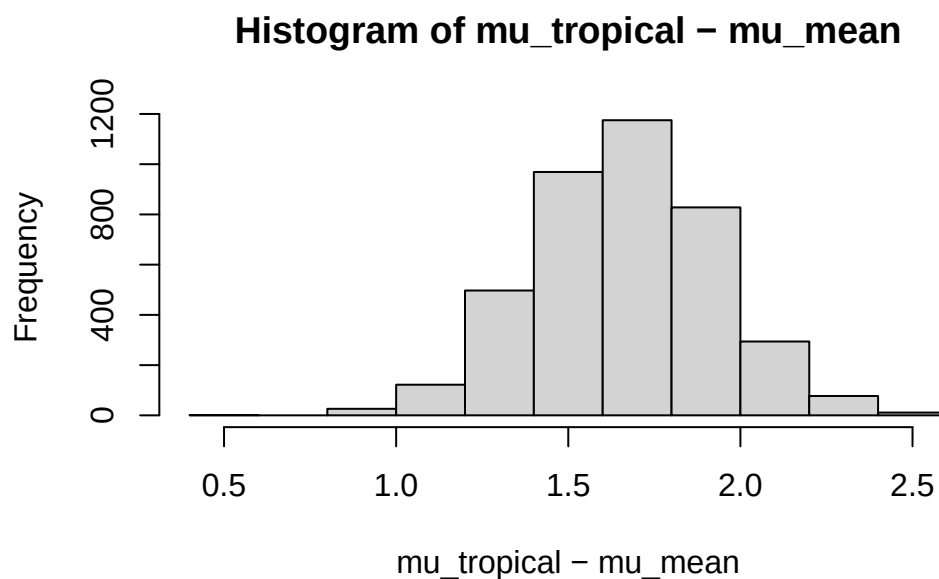
```
      Estimate Est.Error      Q2.5      Q97.5
[1,]  2.856993  0.2908321  2.286048  3.413919
```

```
fitted(fit.lm.add, newdata=data.frame(z.lat= 0,   z.rain=0))
```

```
      Estimate Est.Error      Q2.5      Q97.5
[1,]   1.19341  0.1235972  0.9500366  1.43438
```

We need to extract full posterior predictive distributions and compute the distribution of predicted difference. Then we can look at mean difference and credible intervals

```
mu_tropical = posterior_epred(fit.lm.add, newdata=data.frame(z.lat=-1.5, z.rain=2))
mu_mean     = posterior_epred(fit.lm.add, newdata=data.frame(z.lat= 0,   z.rain=0))
hist(mu_tropical-mu_mean)
```



```
mean(mu_tropical-mu_mean)
```

```
[1] 1.663584
```

```
quantile(mu_tropical-mu_mean, probs=c(0.05, 0.95))
```

```
      5%      95%
1.235225 2.088131
```

Alternatively, we can use the `hypothesis()` function and get the same results

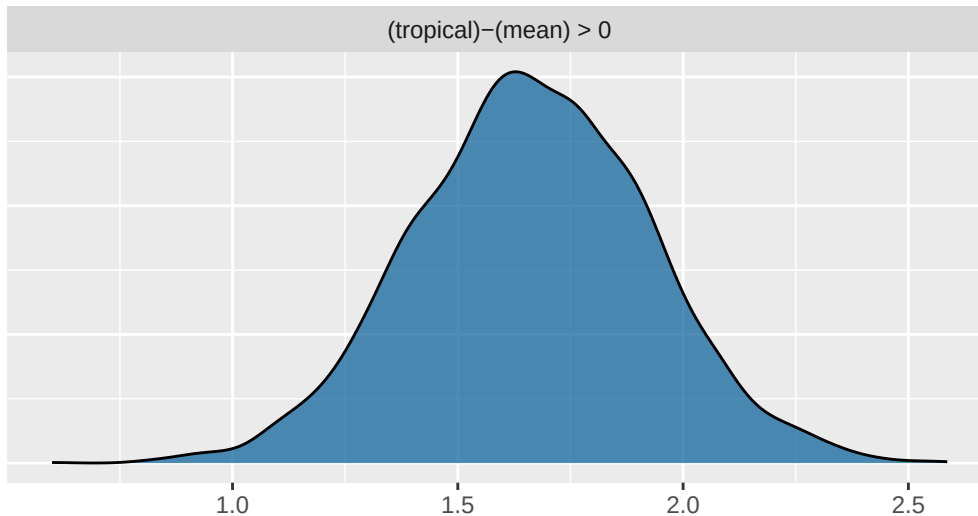
```
mus = fitted(fit.lm.add,
             newdata=data.frame(z.lat = c(-1.5,0),
                                z.rain=c( 2 ,0)),
             summary=F)
mus = as.data.frame(mus)
names(mus) = c("tropical","mean")
hypothesis(mus, "tropical>mean")
```

Hypothesis Tests for class :

	Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
1	(tropical)-(mean) > 0	1.66	0.26	1.24	2.09	Inf	1	*

'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
 '*': For one-sided hypotheses, the posterior probability exceeds 95%;
 for two-sided hypotheses, the value tested against lies outside the 95%-CI.
 Posterior probabilities of point hypotheses assume equal prior probabilities.

```
hypothesis(mus, "tropical>mean") |> plot()
```



Regression, interaction

Question: Does effect of rain change with latitude?

-> need an interaction model

Deterministic part: $\mu = b_0 + b_1 \cdot lat + b_2 \cdot rain + b_3 \cdot lat \cdot rain$

Stochastic part: $\log(height) \sim \text{Normal}(\mu, \sigma)$

We use the same priors as before for 2 main effects (slopes when other predictor is =0, here =mean). Vague prior is put on interaction, with zero mean.

Mean-centering makes main effects b_1, b_2 meaningful and prior choice simpler!

```
fit.lm.int = brm(log(height) ~ z.lat * z.rain,
  prior =
    prior(normal(-1,1), class=b, coef=z.lat) +
    prior(normal(+1,1), class=b, coef=z.rain) +
    prior(normal( 0,1), class=b, coef=z.lat:z.rain),
  data = globalPlants)
```

Convergence checks looking good!

```
summary(fit.lm.int, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: log(height) ~ z.lat * z.rain
Data: globalPlants (Number of observations: 131)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

Priors:

```
b_z.lat ~ normal(-1, 1)
b_z.lat:z.rain ~ normal(0, 1)
b_z.rain ~ normal(+1, 1)
Intercept ~ student_t(3, 1.1, 2.5)
<lower=0> sigma ~ student_t(3, 0, 2.5)
```

Regression Coefficients:

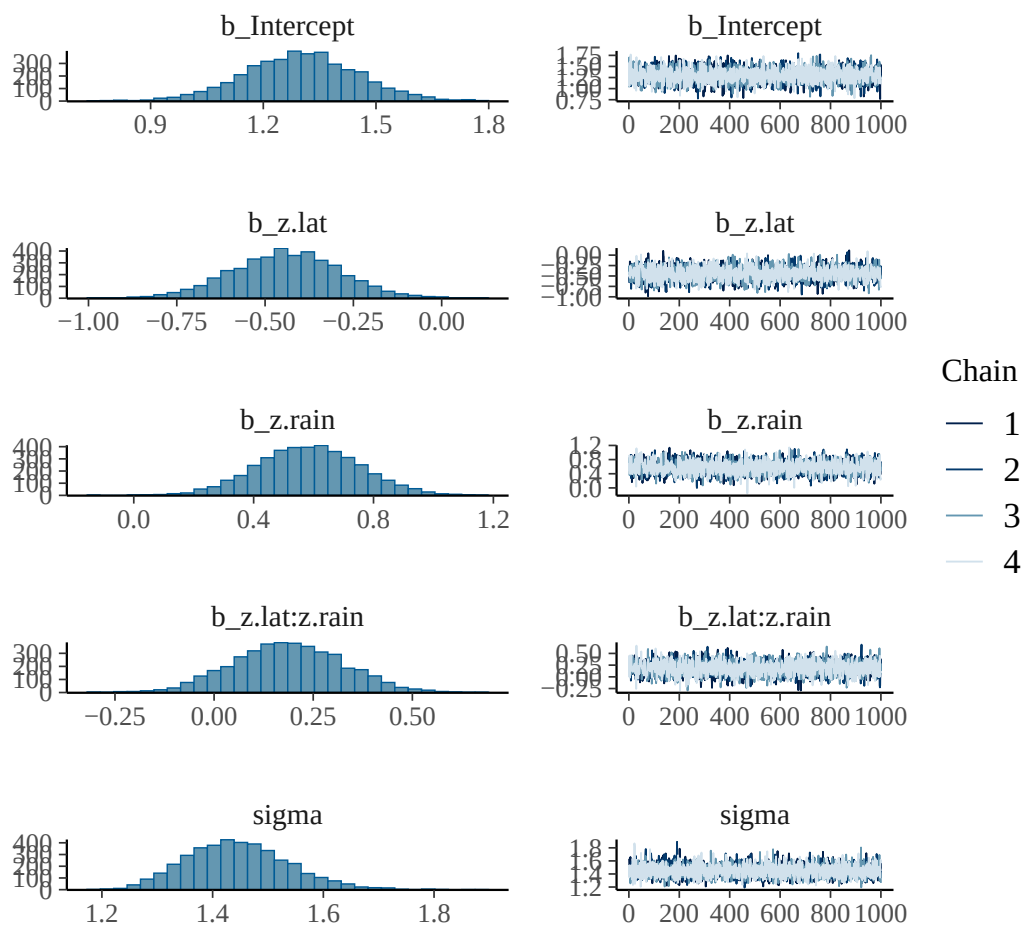
	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.30	0.15	1.00	1.60	1.00	4041	2882
z.lat	-0.44	0.15	-0.74	-0.14	1.00	3475	3299
z.rain	0.58	0.17	0.25	0.92	1.00	3220	3186
z.lat:z.rain	0.19	0.14	-0.08	0.46	1.00	3317	3186

Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.45	0.09	1.28	1.64	1.00	4272	2845

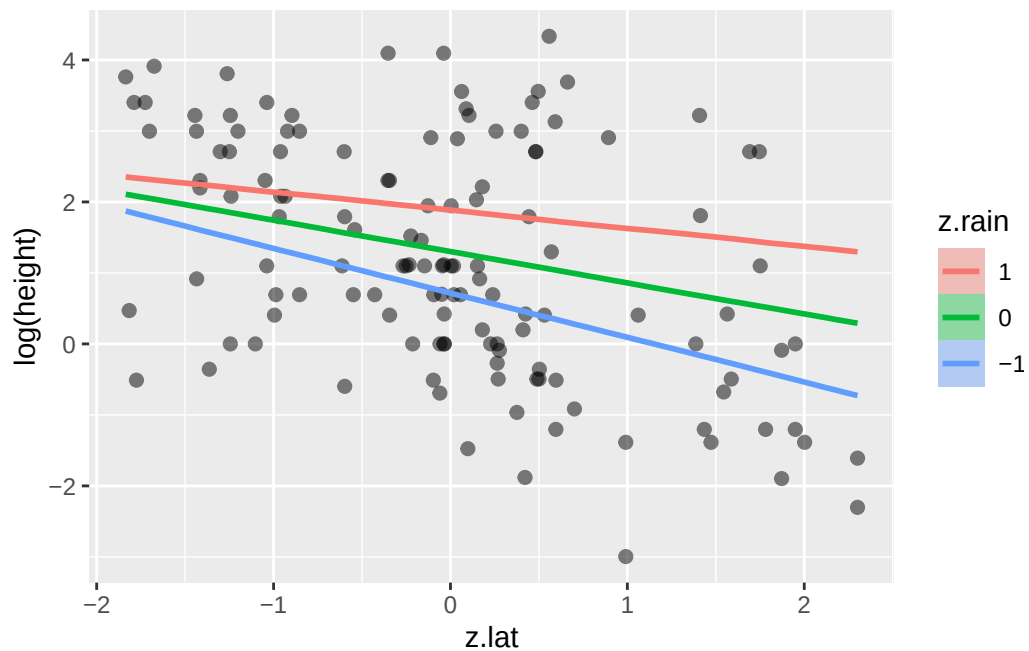
Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
plot(fit.lm.int)
```

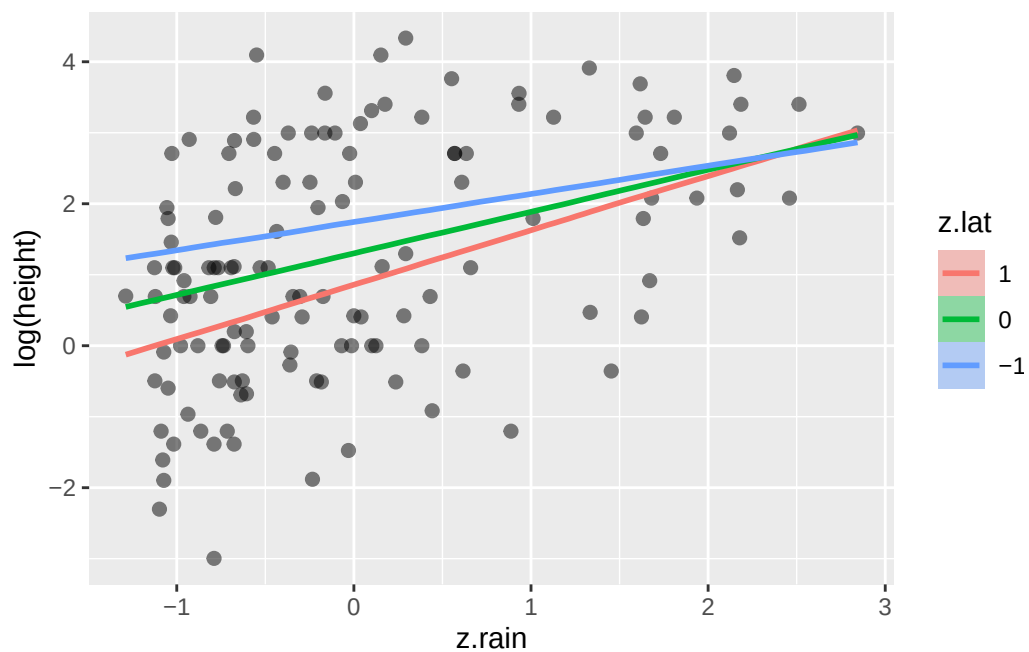


Now `conditional_effects` shows interaction for 3 levels of 2nd predictor

```
plot(conditional_effects(fit.lm.int, effects="z.lat:z.rain", prob=0),
     points=TRUE,
     point_args=c(alpha=0.5))
```



```
plot(conditional_effects(fit.lm.int, effects="z.rain:z.lat", prob=0),
     points=TRUE,
     point_args=c(alpha=0.5))
```

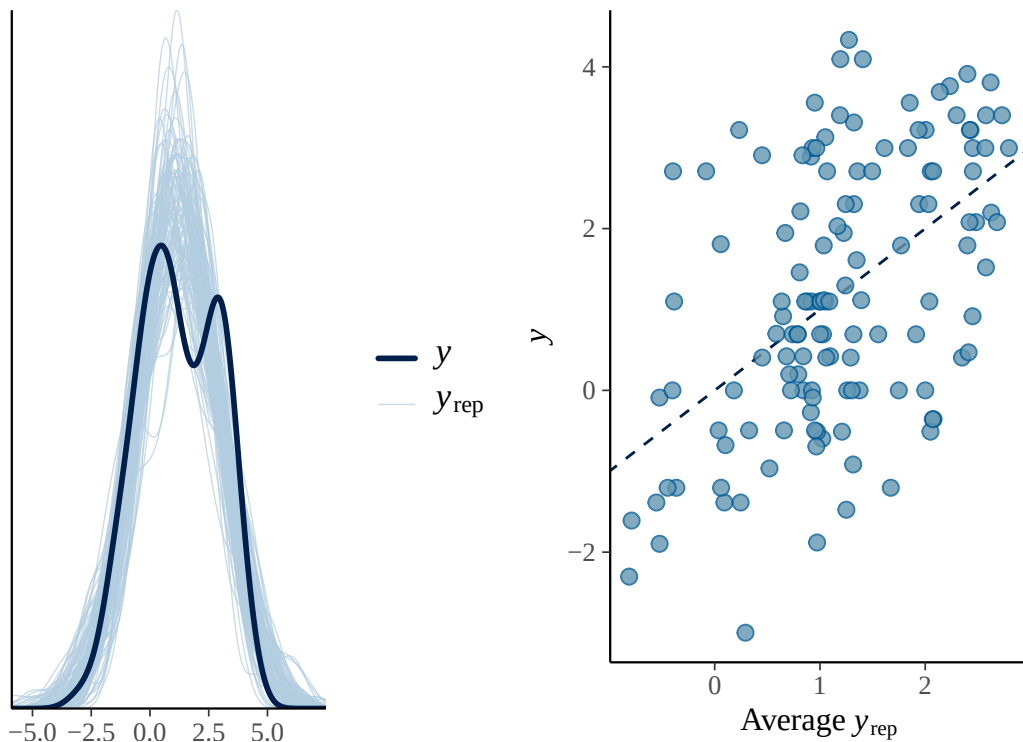


We see some interaction, the effect of rain (slope) becomes stronger (more important) with

latitude.

We quickly do some posterior predictive checks before we move on to the quantification of the research question. PPC look okay-ish, but there is some feature in the data (bimodal histogram) that predictions don't reproduce. This could indicate missing predictors, but we'll leave it here.

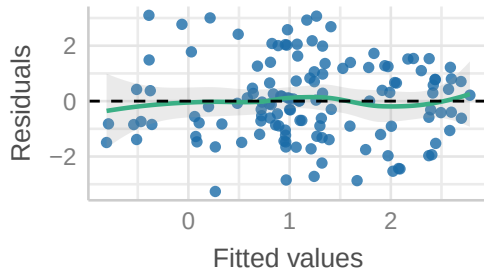
```
p1 = pp_check(fit.lm.int, ndraws=100)
p2 = pp_check(fit.lm.int, type="scatter_avg")
plot_grid(p1,p2)
```



```
check_model(fit.lm.int, check=c("linearity","homogeneity","qq","normality"))
```

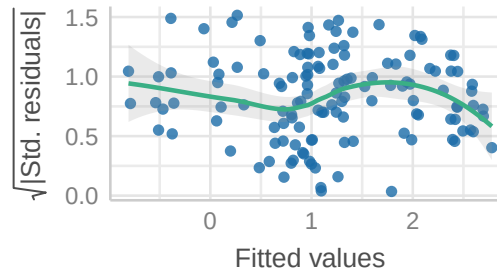
Linearity

Reference line should be flat and horizontal



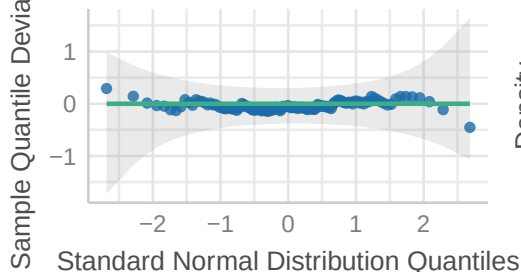
Homogeneity of Variance

Reference line should be flat and horizontal



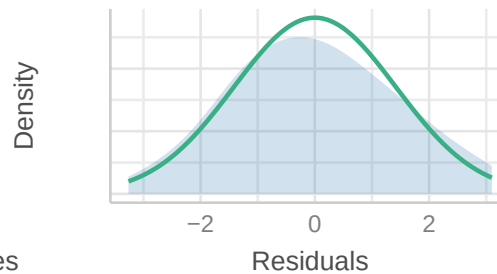
Normality of Residuals

Points should fall along the line



Normality of Residuals

Distribution should be close to the normal cu



Is the interaction meaningful?

```
hypothesis(fit.lm.int, "z.lat:z.rain>0")
```

Hypothesis Tests for class b:

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
1 (z.lat:z.rain) > 0	0.19	0.14	-0.04	0.42	9.78	0.91	

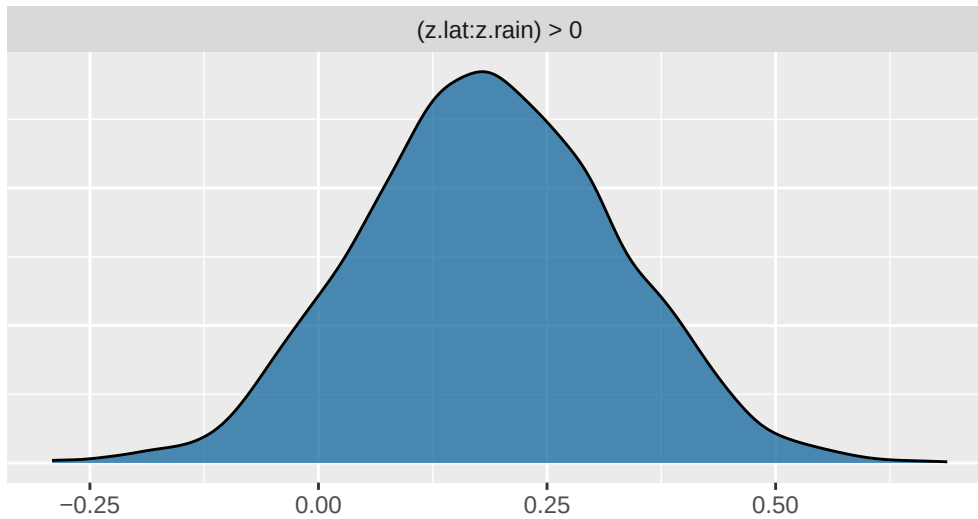
'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.

'*': For one-sided hypotheses, the posterior probability exceeds 95%;

for two-sided hypotheses, the value tested against lies outside the 95%-CI.

Posterior probabilities of point hypotheses assume equal prior probabilities.

```
hypothesis(fit.lm.int, "z.lat:z.rain>0") |> plot()
```



We see some (weak) evidence for a positive interaction ($b = 0.19$, $P(b > 0) = 0.91$)

```
L00(fit.lm.int, fit.lm.add)
```

	elpd_diff	se_diff
fit.lm.add	0.0	0.0
fit.lm.int	-0.2	1.4

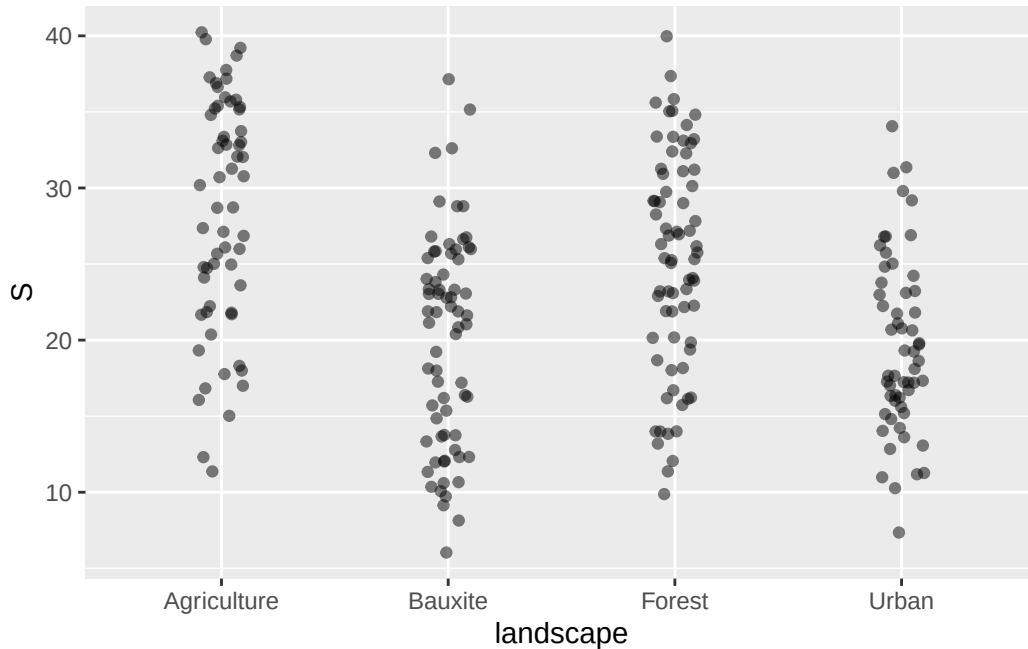
But no, we can't say that the interaction model produces better predictions than the additive model. Data does not sufficiently support hypothesis on an interaction.

One categorical predictor

New dataset, bird species richness in different landscapes from Data4Ecologists package

Question: Does species richness change with landscape type? (1-way ANOVA)

```
data(birds)
ggplot(birds, aes(x=landscape, y=S)) +
  geom_jitter(width=0.1, alpha=0.5)
```



We can either dummy-code (default) or effect-code the model (`~0+...` removes “intercept”). Which priors does brms use for a categorical predictor?

```
default_prior(S ~ landscape, data = birds)
```

	prior	class	coef	group	resp	dpar	nlpar	lb	ub	source
	(flat)	b								default
	(flat)	b	landscapeBauxite							(vectorized)
	(flat)	b	landscapeForest							(vectorized)
	(flat)	b	landscapeUrban							(vectorized)
	student_t(3, 23, 8.9)	Intercept								default
	student_t(3, 0, 8.9)	sigma						0		default

```
default_prior(S ~ 0+landscape, data = birds)
```

	prior	class	coef	group	resp	dpar	nlpar	lb	ub	source
	(flat)	b								default
	(flat)	b	landscapeAgriculture							(vectorized)
	(flat)	b	landscapeBauxite							(vectorized)
	(flat)	b	landscapeForest							(vectorized)
	(flat)	b	landscapeUrban							(vectorized)
	student_t(3, 0, 8.9)	sigma						0		default

For dummy-coding, a prior for intercept (reference level) is given, but not on effects (differences to reference level).

For effect-coding, no priors are given at all on the group-level means.

→ Either way, we should assign meaningful priors.

```
fit.anova1.dummy = brm(S ~ landscape,
                        prior = prior(normal(0,10), class=b), # on differences
                        data = birds)
fit.anova1.effect = brm(S ~ 0+landscape,
                        prior = prior(normal(20,10), class=b), # on actual means
                        data = birds)

summary(fit.anova1.dummy, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ landscape
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

```
Priors:
b ~ normal(0, 10)
Intercept ~ student_t(3, 23, 8.9)
<lower=0> sigma ~ student_t(3, 0, 8.9)
```

Regression Coefficients:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	28.24	0.91	26.43	30.04	1.00	3108	2357
landscapeBauxite	-8.14	1.20	-10.45	-5.80	1.00	3572	3457
landscapeForest	-3.16	1.22	-5.53	-0.73	1.00	3530	2949
landscapeUrban	-8.58	1.27	-11.14	-6.14	1.00	3468	3127

Further Distributional Parameters:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	7.00	0.31	6.42	7.61	1.00	3836	2703

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
summary(fit.anova1.effect, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ 0 + landscape
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

```
Priors:
b ~ normal(20, 10)
<lower=0> sigma ~ student_t(3, 0, 8.9)
```

Regression Coefficients:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
landscapeAgriculture	28.31	0.89	26.56	30.08	1.00	5534	3048
landscapeBauxite	20.07	0.83	18.46	21.71	1.00	4863	3028
landscapeForest	24.99	0.82	23.40	26.58	1.00	5084	3194
landscapeUrban	19.62	0.93	17.83	21.43	1.00	4974	2923

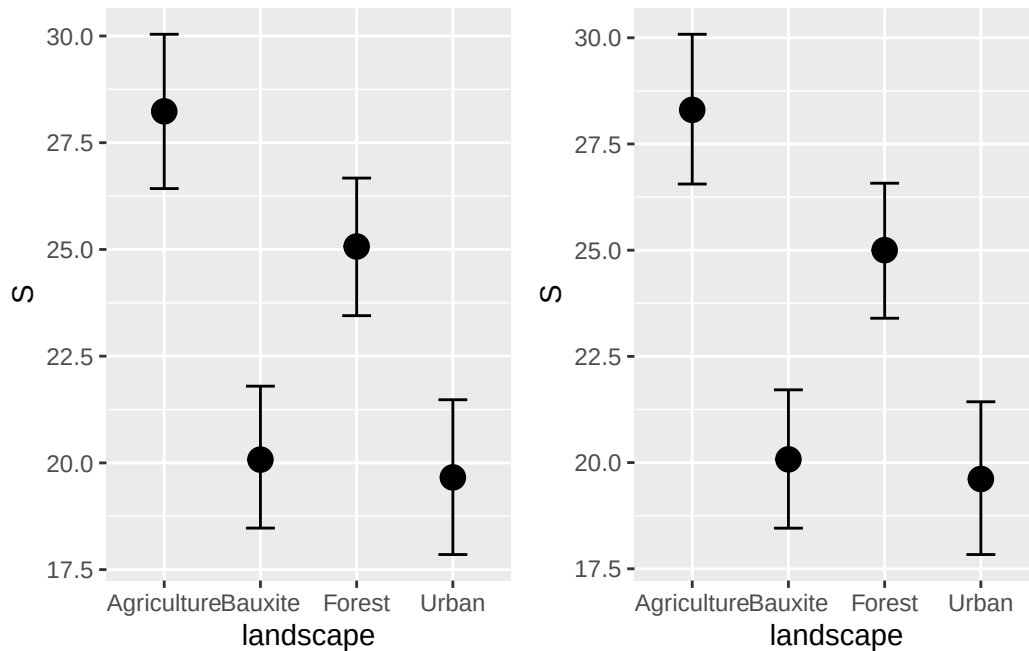
Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	7.00	0.32	6.42	7.67	1.00	4414	2967

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

Both models produce the same predictions

```
p1 = plot(conditional_effects(fit.anova1.dummy), plot=FALSE)
p2 = plot(conditional_effects(fit.anova1.effect), plot=FALSE)
plot_grid(p1[[1]], p2[[1]])
```



While effect-coding (`~0+landscape`) gives group-level means in the summary (they are the parameters here), for dummy-coding (`~landscape`) we can use `emmeans` to get these predictions

```
emmeans(fit.anova1.dummy, ~landscape)
```

landscape	emmean	lower.HPD	upper.HPD
Agriculture	28.2	26.4	30.0
Bauxite	20.1	18.5	21.8
Forest	25.1	23.5	26.7
Urban	19.7	17.8	21.5

Point estimate displayed: median
HPD interval probability: 0.95

Also all their pairwise difference / contrasts

```
pairs(emmeans(fit.anova1.dummy, ~landscape))
```

contrast	estimate	lower.HPD	upper.HPD
Agriculture - Bauxite	8.143	5.840	10.48
Agriculture - Forest	3.168	0.791	5.59

Agriculture - Urban	8.582	6.124	11.12
Bauxite - Forest	-4.988	-7.225	-2.72
Bauxite - Urban	0.432	-1.929	3.07
Forest - Urban	5.411	3.004	7.89

Point estimate displayed: median
HPD interval probability: 0.95

For most pairwise comparisons, the mean difference is far away from 0 (CI does not cover 0).
Is there an overall test to check if species richness changes with landscape (ANOVA)?

We test the model against an intercept-only model (LOO model comparison)

```
fit.anova1.null = brm(S ~ 1, # brms already chooses intercept prior
  data = birds)
```

```
summary(fit.anova1.null)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ 1
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

Regression Coefficients:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	23.28	0.48	22.32	24.19	1.00	3889	2783

Further Distributional Parameters:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	7.83	0.34	7.19	8.53	1.00	3638	2496

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

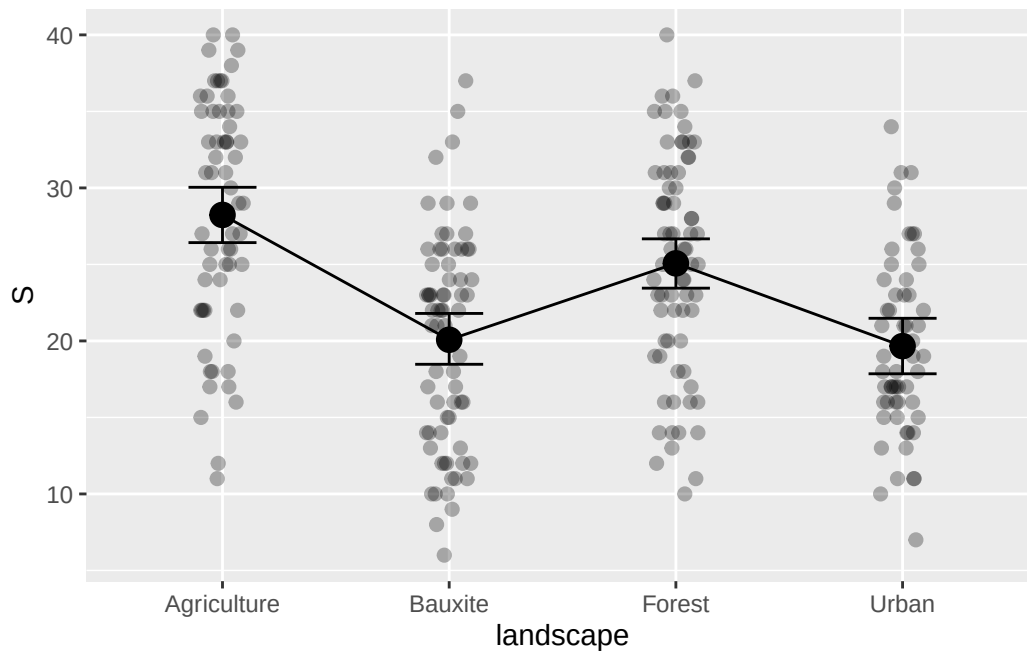
```
LOO(fit.anova1.null, fit.anova1.dummy)
```

	elpd_diff	se_diff
fit.anova1.dummy	0.0	0.0
fit.anova1.null	-27.1	7.3

Yes, we see a strong support for the ~landscape model.

Finally, here's a nicer plot for the fitted values vs data

```
p1 = plot(conditional_effects(fit.anova1.dummy,"landscape"),
          points=TRUE,
          point_args=list(width=0.1, alpha=0.3),
          plot = FALSE
        )
p1[[1]] + geom_line(aes(group="landscape"))
```



Two categorical predictors

We include a 2nd predictor “area”, here as a categorical with 2 levels (small/large).

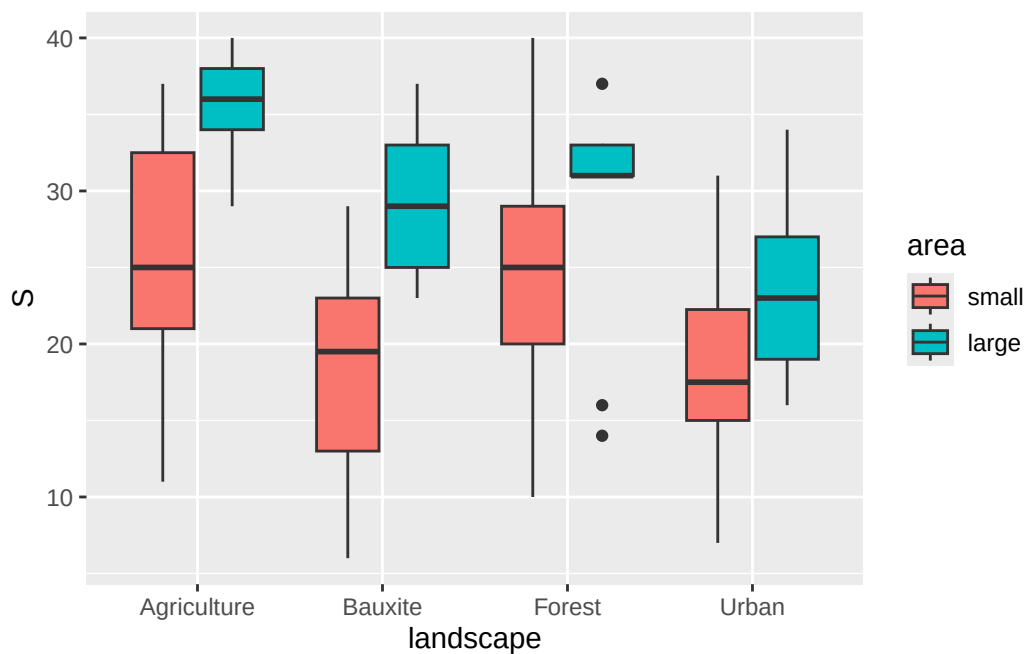
Question: Surely, richness is higher in larger areas, but does the difference depend on landscape type?

→ We fit an additive and an interaction model (additive works with dummy-coding only)

```
birds$area = cut(birds$log.area., 2, labels=c("small", "large"))
head(birds)
```

	patch	S	landscape	area	log.area.	year
1	ag1a	24	Agriculture	small	0.5453297	2005
2	ag1b	15	Agriculture	small	-0.2107610	2005
3	ag1c	25	Agriculture	small	0.3492867	2005
4	ag1d	35	Agriculture	large	0.9180241	2005
5	ag2a	32	Agriculture	small	0.1378772	2005
6	ag2b	40	Agriculture	large	1.7729067	2005

```
ggplot(birds, aes(landscape, S)) +  
  geom_boxplot(aes(fill=area))
```



Additive model

```
fit.anova2.add = brm(S ~ landscape+area,  
  prior = prior(normal(0,10), class=b),  
  data = birds)
```

With multiple categorical predictors (& their interaction), the estimated parameters in the summary table become less interpretable. Nevertheless, always look at the summary table to check Rhat values for convergence.

```
summary(fit.anova2.add, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ landscape + area
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

```
Priors:
b ~ normal(0, 10)
Intercept ~ student_t(3, 23, 8.9)
<lower=0> sigma ~ student_t(3, 0, 8.9)
```

Regression Coefficients:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	26.08	0.86	24.42	27.81	1.00	3331	2941
landscapeBauxite	-7.00	1.11	-9.17	-4.81	1.00	3682	3168
landscapeForest	-2.00	1.11	-4.10	0.21	1.00	3324	3110
landscapeUrban	-7.63	1.15	-9.89	-5.44	1.00	3398	3040
arealarge	7.73	1.05	5.65	9.74	1.00	4909	2991

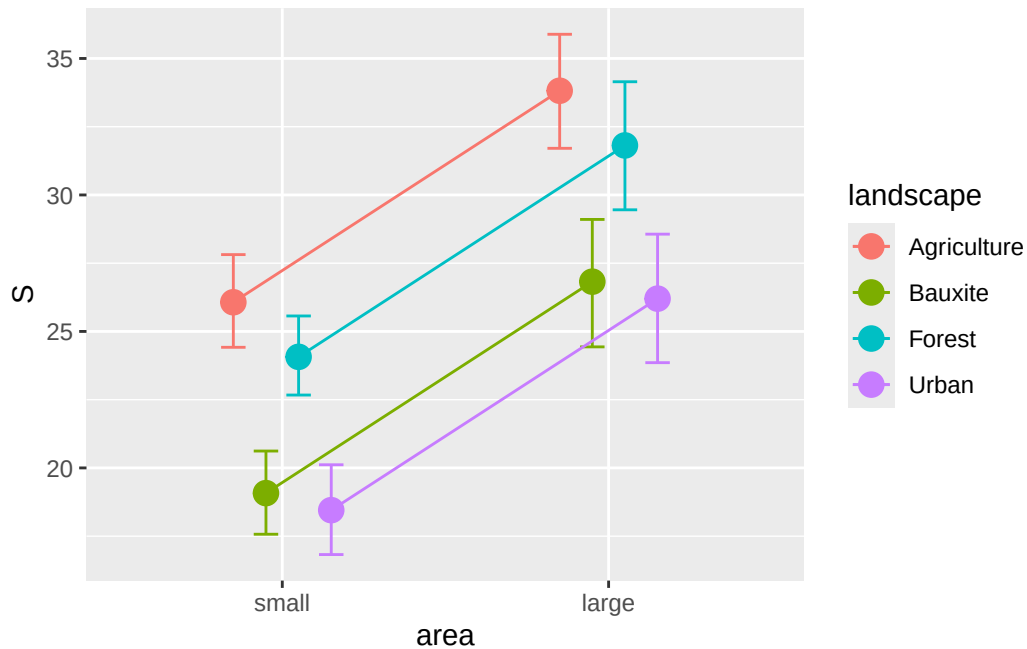
Further Distributional Parameters:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	6.37	0.28	5.84	6.97	1.00	4727	2275

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

Additive effects means parallel lines

```
p1 = plot(conditional_effects(fit.anova2.add, effects="area:landscape"),
          plot=FALSE)
p1[[1]] + geom_line(aes(group=landscape),
                    position=position_dodge(0.4))
```



Means and contrasts averaged over landscapes (i.e. landscape is averaged out, and means are displayed for `~area`)

```
emmeans(fit.anova2.add, ~area)
```

area	emmean	lower.HPD	upper.HPD
small	21.9	21.1	22.7
large	29.7	27.9	31.6

Results are averaged over the levels of: landscape
 Point estimate displayed: median
 HPD interval probability: 0.95

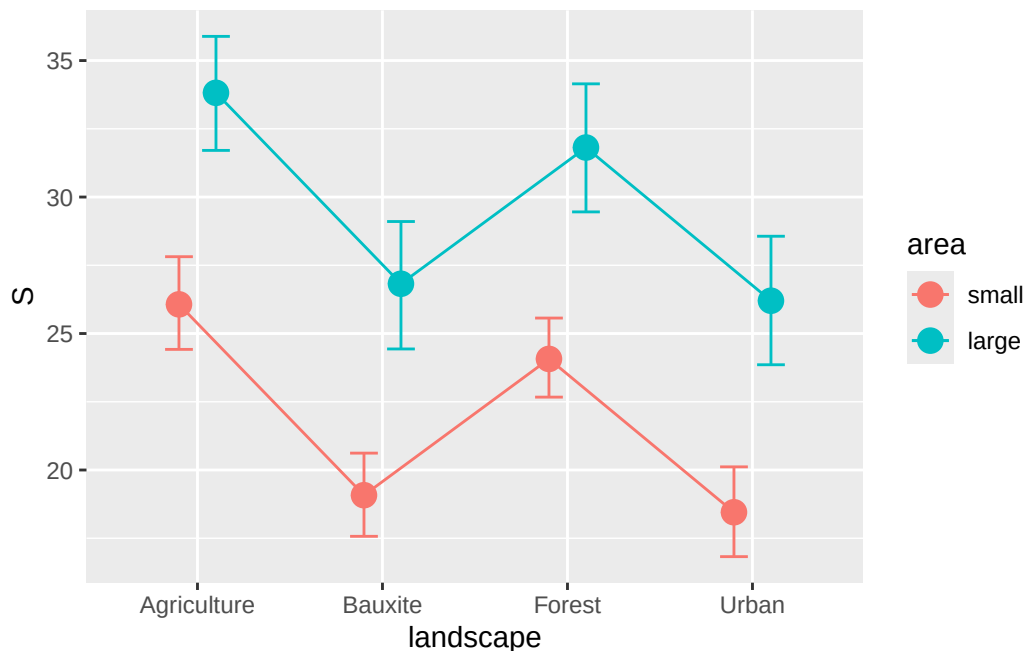
```
emmeans(fit.anova2.add, ~area) |> pairs()
```

contrast	estimate	lower.HPD	upper.HPD
small - large	-7.75	-9.81	-5.74

Results are averaged over the levels of: landscape
 Point estimate displayed: median
 HPD interval probability: 0.95

Now the other way round! Again parallel lines because of additive effects.

```
p1 = plot(conditional_effects(fit.anova2.add, effects="landscape:area"),
          plot=FALSE)
p1[[1]] + geom_line(aes(group=area),
                    position=position_dodge(0.4))
```



Means and contrasts averaged over area (the same kind of predictions we did in the previous 1-way ANOVA $S \sim \text{landscape}$)

```
emmeans(fit.anova2.add, ~landscape)
```

landscape	emmean	lower.HPD	upper.HPD
Agriculture	29.9	28.4	31.6
Bauxite	22.9	21.3	24.6
Forest	27.9	26.2	29.6
Urban	22.3	20.5	24.0

Results are averaged over the levels of: area

Point estimate displayed: median

HPD interval probability: 0.95

```
emmeans(fit.anova2.add, ~landscape) |> pairs()
```

contrast	estimate	lower.HPD	upper.HPD
Agriculture - Bauxite	7.006	4.772	9.10
Agriculture - Forest	2.022	-0.225	4.08
Agriculture - Urban	7.595	5.560	10.01
Bauxite - Forest	-5.010	-7.236	-3.04
Bauxite - Urban	0.607	-1.628	2.76
Forest - Urban	5.615	3.425	7.76

Results are averaged over the levels of: area

Point estimate displayed: median

HPD interval probability: 0.95

Interaction model

Full interaction model means individual means for all level-combinations, although it's not immediately obvious from the summary table.

```
fit.anova2.int = brm(S ~ landscape*area,
                     prior = prior(normal(0,10), class=b),
                     data = birds)
```

```
summary(fit.anova2.int, prior=TRUE)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ landscape * area
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

Priors:

```
b ~ normal(0, 10)
Intercept ~ student_t(3, 23, 8.9)
<lower=0> sigma ~ student_t(3, 0, 8.9)
```

Regression Coefficients:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	25.60	0.95	23.74	27.47	1.00	1883	2426
landscapeBauxite	-6.91	1.26	-9.35	-4.44	1.00	2277	2585
landscapeForest	-1.09	1.24	-3.54	1.29	1.00	2369	2873
landscapeUrban	-6.76	1.28	-9.27	-4.25	1.00	2458	2894

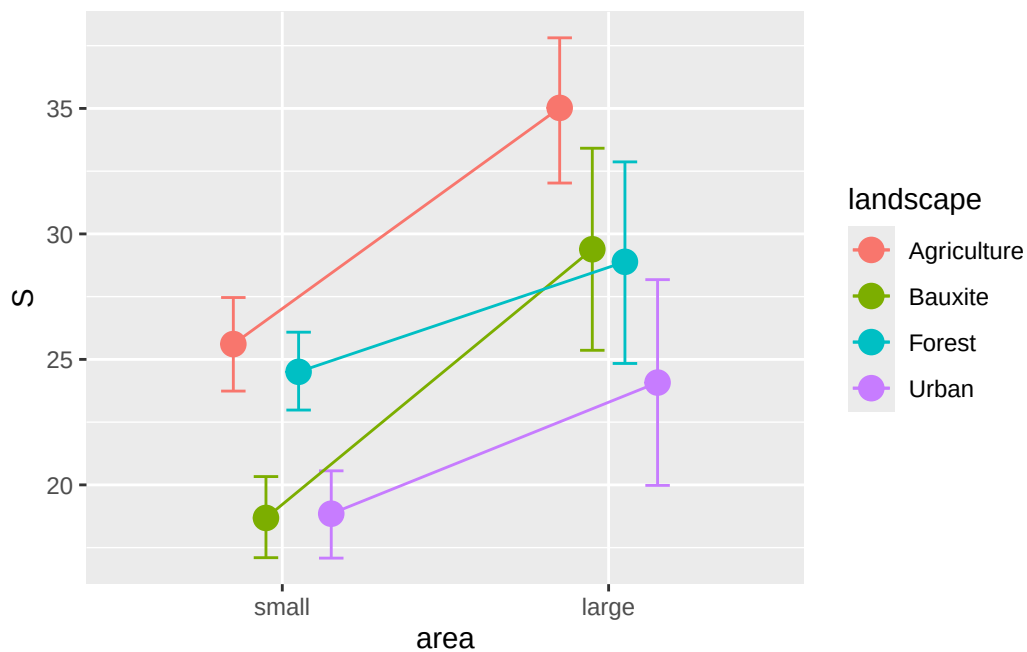
arealarge	9.43	1.73	5.93	12.75	1.00	1863	2210
landscapeBauxite:arealarge	1.26	2.75	-4.28	6.74	1.00	2765	3276
landscapeForest:arealarge	-5.05	2.76	-10.48	0.43	1.00	2801	3115
landscapeUrban:arealarge	-4.18	2.79	-9.58	1.46	1.00	2739	2958

Further Distributional Parameters:

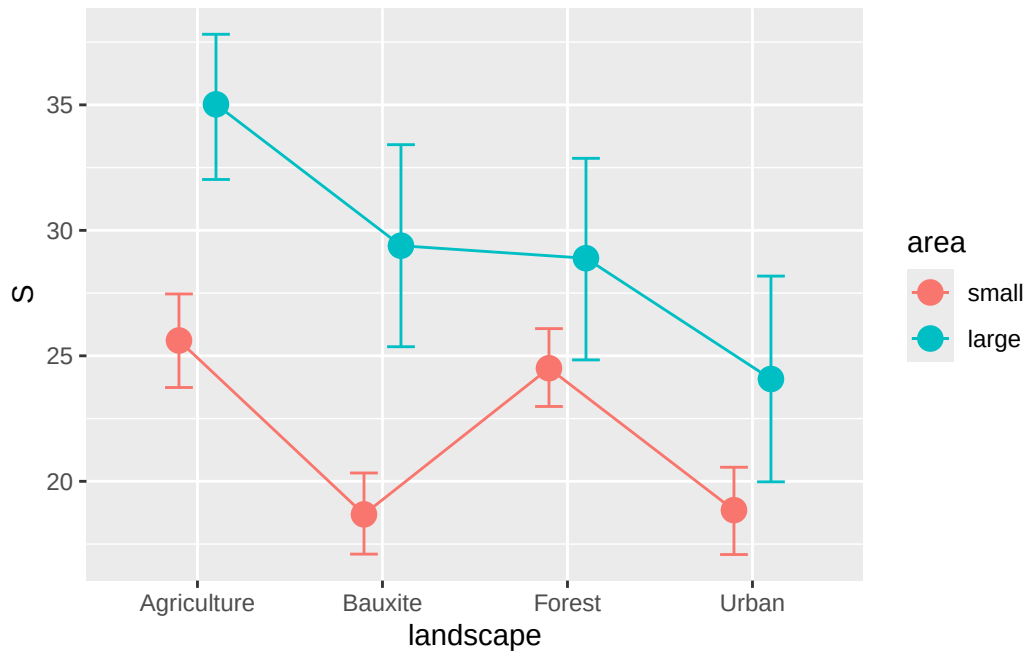
	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	6.31	0.28	5.80	6.89	1.00	4680	2607

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
p1 = plot(conditional_effects(fit.anova2.int, effects="area:landscape"),
          plot=FALSE)
p1[[1]] + geom_line(aes(group=landscape),
                    position=position_dodge(0.4))
```



```
p1 = plot(conditional_effects(fit.anova2.int, effects="landscape:area"),
          plot=FALSE)
p1[[1]] + geom_line(aes(group=area),
                    position=position_dodge(0.4))
```



Careful when averaging over a prediction in interaction models

```
emmeans(fit.anova2.int, ~area)
```

NOTE: Results may be misleading due to involvement in interactions

area	emmean	lower.HPD	upper.HPD
small	21.9	21.1	22.8
large	29.3	27.4	31.3

Results are averaged over the levels of: landscape

Point estimate displayed: median

HPD interval probability: 0.95

```
emmeans(fit.anova2.int, ~area) |> pairs()
```

NOTE: Results may be misleading due to involvement in interactions

contrast	estimate	lower.HPD	upper.HPD
small - large	-7.45	-9.53	-5.28

Results are averaged over the levels of: landscape

Point estimate displayed: median

HPD interval probability: 0.95

Since the summary table doesn't give us the predicted means right away, we can calculate them

```
emmeans(fit.anova2.int, ~area:landscape)
```

area	landscape	emmean	lower.HPD	upper.HPD
small	Agriculture	25.6	23.8	27.5
large	Agriculture	35.0	32.3	38.1
small	Bauxite	18.7	17.1	20.3
large	Bauxite	29.4	25.5	33.5
small	Forest	24.5	23.0	26.1
large	Forest	28.9	24.8	32.7
small	Urban	18.9	17.1	20.6
large	Urban	24.1	20.1	28.3

Point estimate displayed: median
HPD interval probability: 0.95

Now back to our research question, if the area effect on species richness changes with landscape type.

```
L00(fit.anova2.int, fit.anova2.add)
```

	elpd_diff	se_diff
fit.anova2.int	0.0	0.0
fit.anova2.add	-0.7	2.3

No, the data does not support the hypothesis and we conclude that the area effect is independent of landscape type here. There is just a small difference in elpd (in favor of the interaction), but compared to the associated standard error we have to treat both models as equally performing. The principle of parsimony dictates to prefer the less complex model, here the additive one.

Categorical and continuous predictors

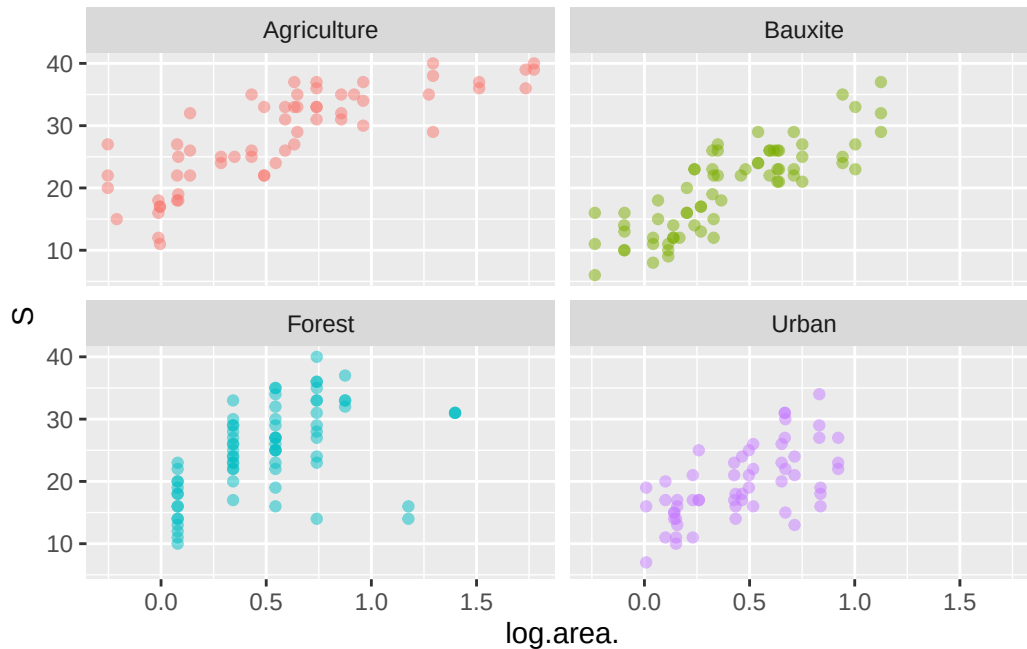
Now we use the full resolution of area as continuous predictor.

Same question: Does the area-effect change between landscape types?

→ We test additive vs interaction ANCOVA

If chosen as an exercise, you can leave out the priors for now and focus on correct model definition.

```
ggplot(birds, aes(x=log.area., y=S, col=landscape)) +
  geom_point(alpha=0.5) +
  facet_wrap(~landscape) +
  theme(legend.position="none")
```



Additive model

First, we check which priors to assign.

In the additive model, we need priors for discrete landscape effects (intercept effects) and for area (slope).

→ We use overall prior for all effects (`class=b`) and override this only for area slope (`class=b`, `coef=log.area.`)

```
default_prior(S ~ landscape+log.area., data=birds)
```

Warning: Rows containing NAs were excluded from the model.

prior	class	coef	group	resp	dpar	nlpar	lb	ub	source
(flat)	b								default
(flat)	b	landscapeBauxite							(vectorized)
(flat)	b	landscapeForest							(vectorized)

	(flat)	b	landscapeUrban	(vectorized)
	(flat)	b	log.area.	(vectorized)
student_t(3, 23, 8.9)	Intercept			default
student_t(3, 0, 8.9)	sigma		0	default

```
fit.ancova.add = brm(S ~ landscape+log.area.,
  prior =
    prior(normal(0,10), class=b) +
    prior(normal(10,10), class=b, coef=log.area.),
  data = birds)
```

```
summary(fit.ancova.add)
```

```
Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ landscape + log.area.
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000
```

Regression Coefficients:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	20.90	0.80	19.31	22.45	1.00	3208	2893
landscapeBauxite	-5.97	0.87	-7.66	-4.20	1.00	3451	3091
landscapeForest	-2.27	0.88	-3.98	-0.53	1.00	3679	2941
landscapeUrban	-7.05	0.92	-8.82	-5.25	1.00	3864	3108
log.area.	12.69	0.83	11.09	14.35	1.00	4651	2589

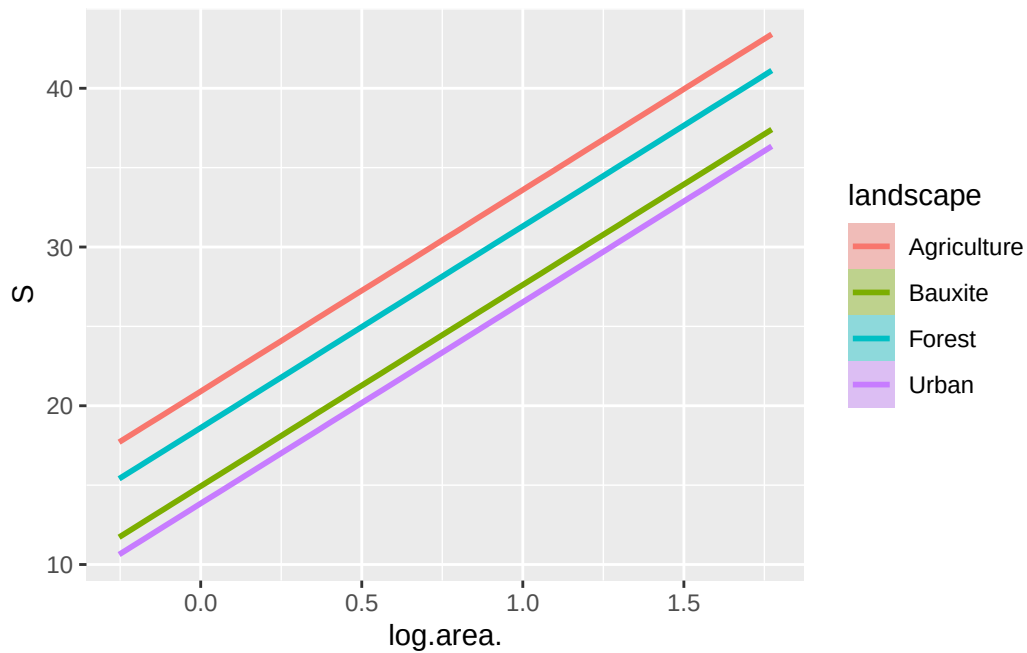
Further Distributional Parameters:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	4.98	0.23	4.55	5.44	1.00	4695	2665

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

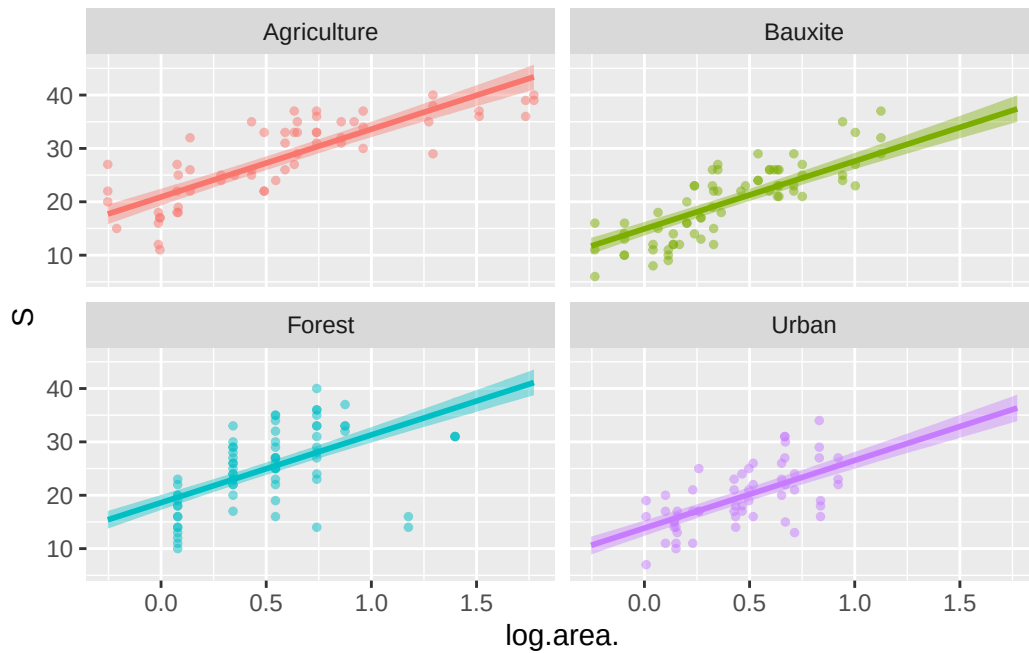
In the additive model, each landscape type has its own intercept, but a joint slope with area.

```
plot(conditional_effects(fit.ancova.add, effects="log.area.:landscape", prob=0))
```



We can do some ggplot2 magic to plot separately for each landscape type

```
p1 = plot(conditional_effects(fit.ancova.add, effects="log.area.:landscape"),
          points=T,
          point_args=c(size=1.0, alpha=0.5),
          plot=F)
p1[[1]] +
  facet_wrap(~landscape) +
  theme(legend.position="none")
```



We can compute intercepts (at mean area) & contrasts with emmeans

```
emmeans(fit.ancova.add, ~landscape)
```

landscape	emmean	lower.HPD	upper.HPD
Agriculture	27.1	25.8	28.4
Bauxite	21.1	20.0	22.3
Forest	24.8	23.6	26.0
Urban	20.0	18.8	21.3

Point estimate displayed: median

HPD interval probability: 0.95

```
emmeans(fit.ancova.add, ~landscape) |> pairs()
```

contrast	estimate	lower.HPD	upper.HPD
Agriculture - Bauxite	5.97	4.170	7.61
Agriculture - Forest	2.29	0.559	4.00
Agriculture - Urban	7.06	5.262	8.82
Bauxite - Forest	-3.69	-5.335	-2.06
Bauxite - Urban	1.07	-0.574	2.86
Forest - Urban	4.77	3.067	6.49

Point estimate displayed: median
 HPD interval probability: 0.95

Interaction model

Again, we check which priors to assign.

Here, we need priors for all landscape effects (intercept effects), slope in reference level, and changes in slopes

-> I could not find a good way to shorten this, assign priors manually

```
default_prior(S ~ landscape*log.area., data=birds)
```

Warning: Rows containing NAs were excluded from the model.

	prior	class	coef	group	resp	dpar	nlpar	lb	ub
	(flat)	b							
	(flat)	b	landscapeBauxite						(vec
	(flat)	b	landscapeBauxite:log.area.						(vec
	(flat)	b	landscapeForest						(vec
	(flat)	b	landscapeForest:log.area.						(vec
	(flat)	b	landscapeUrban						(vec
	(flat)	b	landscapeUrban:log.area.						(vec
	(flat)	b	log.area.						(vec
student_t(3, 23, 8.9)		Intercept							
student_t(3, 0, 8.9)		sigma						0	

```
fit.ancova.int = brm(S ~ landscape*log.area.,
  prior =
    prior(normal(0,10), class=b, coef=landscapeBauxite) +
    prior(normal(0,10), class=b, coef=landscapeForest) +
    prior(normal(0,10), class=b, coef=landscapeUrban) +
    prior(normal(10,10), class=b, coef=log.area.) +
    prior(normal(0,5), class=b, coef=landscapeBauxite:log.area.) +
    prior(normal(0,5), class=b, coef=landscapeForest:log.area.) +
    prior(normal(0,5), class=b, coef=landscapeUrban:log.area.),
  data = birds)
```

```
summary(fit.ancova.int)
```

```

Family: gaussian
Links: mu = identity; sigma = identity
Formula: S ~ landscape * log.area.
Data: birds (Number of observations: 257)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
       total post-warmup draws = 4000

```

Regression Coefficients:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	21.41	0.92	19.66	23.27	1.00	2482	2642
landscapeBauxite	-7.79	1.25	-10.21	-5.41	1.00	2655	2936
landscapeForest	-2.00	1.35	-4.64	0.66	1.00	2511	2626
landscapeUrban	-7.59	1.50	-10.51	-4.63	1.00	2243	2435
log.area.	11.78	1.11	9.66	13.98	1.00	2325	2639
landscapeBauxite:log.area.	4.15	1.88	0.54	7.83	1.00	2615	2678
landscapeForest:log.area.	-0.63	1.94	-4.37	3.26	1.00	2791	2916
landscapeUrban:log.area.	1.03	2.40	-3.70	5.67	1.00	2777	2472

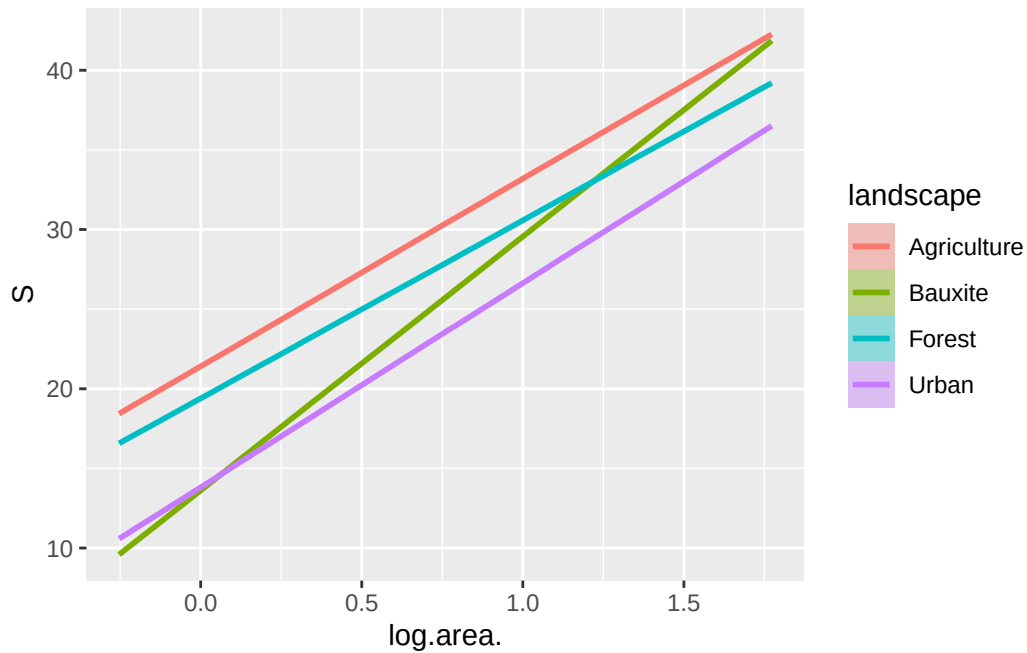
Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	4.93	0.23	4.50	5.39	1.00	4487	2715

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

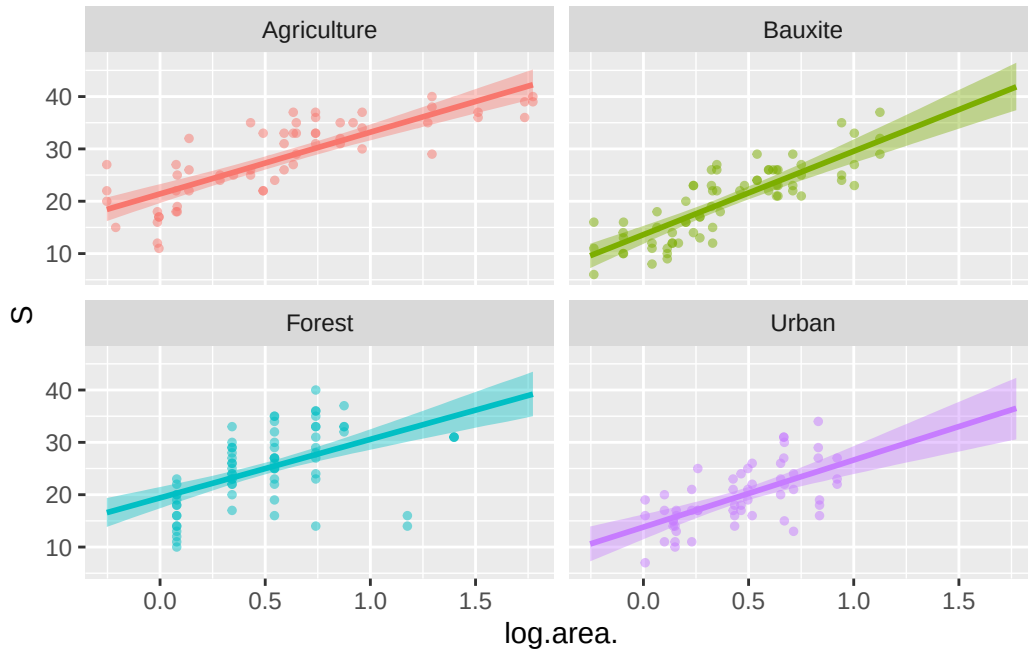
In the interaction model, each landscape type has its own intercept & own slope.

```
plot(conditional_effects(fit.ancova.int, effects="log.area.:landscape", prob=0))
```



We can do some ggplot2 magic to plot separately for each landscape type

```
p1 = plot(conditional_effects(fit.ancova.int, effects="log.area.:landscape"),
          points=T,
          point_args=c(size=1.0, alpha=0.5),
          plot=F)
p1[[1]] +
  facet_wrap(~landscape) +
  theme(legend.position="none")
```

Here we can compute slopes and their contrasts with a new function `emtrends()`. `var=` specifies the continuous predictor, for which slopes are presented.

```
emtrends(fit.ancova.int, ~landscape, var="log.area.")
```

landscape	log.area..trend	lower.HPD	upper.HPD
Agriculture	11.8	9.65	14.0
Bauxite	15.9	12.65	18.9
Forest	11.2	7.92	14.4
Urban	12.8	8.52	17.2

Point estimate displayed: median
HPD interval probability: 0.95

```
emtrends(fit.ancova.int, ~landscape, var="log.area.") |> pairs()
```

contrast	estimate	lower.HPD	upper.HPD
Agriculture - Bauxite	-4.171	-7.911	-0.622
Agriculture - Forest	0.617	-3.094	4.449
Agriculture - Urban	-1.072	-5.721	3.658
Bauxite - Forest	4.743	0.359	9.244
Bauxite - Urban	3.124	-1.853	8.500
Forest - Urban	-1.635	-6.907	4.108

Point estimate displayed: median
HPD interval probability: 0.95

Now back to our research question, if the area effect (slopes) on species richness changes with landscape type.

```
L00(fit.ancova.add, fit.ancova.int)
```

	elpd_diff	se_diff
fit.ancova.int	0.0	0.0
fit.ancova.add	-0.7	2.0

Again, no strong support for the interaction model. Area-effect does not vary strongly between landscape types (same result as the 2-way ANOVA).

For the sake of brevity I left out convergence checks and posterior predictive checks, but you should always include them in any analysis

```
plot(fit)
pp_check(fit, ndraws=100)
pp_check(fit, type="scatter_avg")
check_model(fit, check=c("linearity", "homogeneity", "qq", "normality"))
```

Exercise

Now that we have covered the most important aspects of linear models, apply this to a dataset of your choice. Maybe you have a dataset and some hypotheses to test in your own research project. Otherwise, the `palmerpenguins` package has a nice dataset to fit some linear models:

allisonhorst.github.io/palmerpenguins/index.html

```
library("palmerpenguins")
str(penguins)
```

```
tibble [344 x 8] (S3: tbl_df/tbl/data.frame)
 $ species      : Factor w/ 3 levels "Adelie","Chinstrap",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ island       : Factor w/ 3 levels "Biscoe","Dream",...: 3 3 3 3 3 3 3 3 3 3 ...
 $ bill_length_mm : num [1:344] 39.1 39.5 40.3 NA 36.7 39.3 38.9 39.2 34.1 42 ...
```

```
$ bill_depth_mm      : num [1:344] 18.7 17.4 18 NA 19.3 20.6 17.8 19.6 18.1 20.2 ...
$ flipper_length_mm: int [1:344] 181 186 195 NA 193 190 181 195 193 190 ...
$ body_mass_g        : int [1:344] 3750 3800 3250 NA 3450 3650 3625 4675 3475 4250 ...
$ sex                : Factor w/ 2 levels "female","male": 2 1 1 NA 1 2 1 2 NA NA ...
$ year               : int [1:344] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 ...
```