Prediction and Evaluation

Summer and Gabe

```r
library(brms) # for statistics
```

```
Warning: package 'brms' was built under R version 4.5.2
```

```
Loading required package: Rcpp
```

```
Warning: package 'Rcpp' was built under R version 4.5.1
```

```
Loading 'brms' package (version 2.23.0). Useful instructions
can be found by typing help('brms'). A more detailed introduction
to the package is available through vignette('brms_overview').
```

```
Attaching package: 'brms'
```

```
The following object is masked from 'package:stats':

    ar
```

```r
library(tidyverse) # for data wrangling
```

```
Warning: package 'tidyverse' was built under R version 4.5.2
```

```
Warning: package 'ggplot2' was built under R version 4.5.1
```

```
Warning: package 'tibble' was built under R version 4.5.2
```

```
Warning: package 'purrr' was built under R version 4.5.1
```

```
Warning: package 'stringr' was built under R version 4.5.1


Warning: package 'forcats' was built under R version 4.5.1


Warning: package 'lubridate' was built under R version 4.5.2


-- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.1     v stringr   1.5.2
v ggplot2   4.0.0     v tibble    3.3.1
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0

-- Conflicts --------------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```r
library(lterdatasampler)
```

```
Warning: package 'lterdatasampler' was built under R version 4.5.2
```

```r
pie_crab <- lterdatasampler::pie_crab
```

```r
colnames(pie_crab)
```

```
[1] "date"        "latitude"      "site"          "size"
[5] "air_temp"    "air_temp_sd"   "water_temp"    "water_temp_sd"
[9] "name"
```

```r
?pie_crab
```

```
starting httpd help server ... done
```

Q1.1a How might *mean* annual *water* temperature affect crab size?

Higher temperatures could cause higher crab size due to increased metabolic rate, given that
they have sufficient food and it ain't too hot.

2

Q1.1b How might *mean* annual *air* temperature affect crab size?

Similar to the water temp, this would increase the crab size.

Q1.1c How might the *sd* (variability) of *water* temperature affect crab size?

Higher variability in temperatures would decrease the crab size due to it having to deal with temperature shifts

---

Q1.1d How might the *sd* (variability) of *air* temperature affect crab size?

Higher variability in temperatures would decrease the crab size due to it having to deal with temperature shifts

```
m.crab.lat.water <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + water_temp,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
      iter = 2000, warmup = 1000, chains = 4, cores = 4,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.water")
```

```
summary(m.crab.lat.water)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + water_temp
   Data: pie_crab (Number of observations: 392)
```

```
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     -23.00      6.96   -36.44    -9.73 1.00     1635     1928
latitude        0.81      0.12     0.58     1.04 1.00     1643     2028
water_temp      0.41      0.15     0.13     0.70 1.00     1662     1943

Further Distributional Parameters:
        Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma       2.81      0.10     2.62     3.02 1.00     2997     2424

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).
```
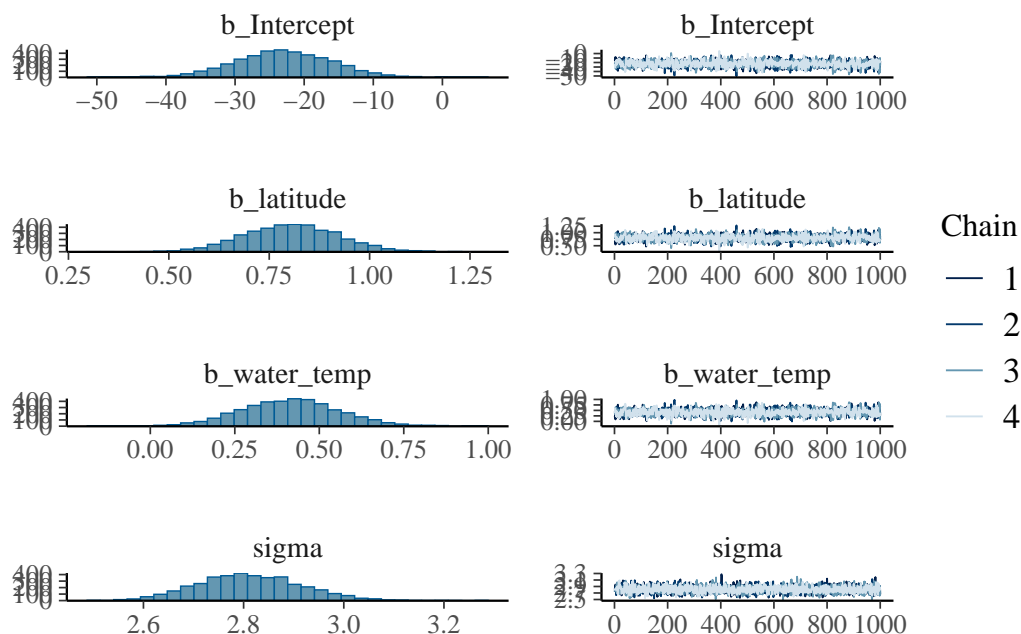
`plot(m.crab.lat.water)`

Q1.2a Assess the output

It seems that the model ran correctly as we have distributions that do not include zero, chains that are overlapping, and and R hat of 1. We looked through this and we looking for things out of place, as we know when one of these is wonky, it can be pretty evident in the graphs and seeing an R hat of anything other than 1.

Q1.2b Interpret the output

Interpret your model by answering:

1. What are the effects of your predictors? Remember to describe the effect using the units to make it biologically meaningful.

For every one degree C increase in water temp, there is a .41 mm increase in crab size. For every one degree increase in latitude, there is a .81 mm increase in crab size.

1. Are the effects reasonably different from zero? How do you know?

Yes, the posterior distributions do not include 0 and neither do the compatibility intervals.

```
m.crab.lat.air <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + air_temp,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
      iter = 2000, warmup = 1000, chains = 4, cores = 4,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.air")
```

```
summary(m.crab.lat.air)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + air_temp
```

```
   Data: pie_crab (Number of observations: 392)
  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    75.82     18.47    38.32   111.74 1.00     1795     1634
latitude     -0.96      0.34    -1.62    -0.28 1.00     1796     1641
air_temp     -1.64      0.38    -2.37    -0.88 1.00     1801     1635

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.77      0.10     2.58     2.98 1.00     2472     2168

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).
```
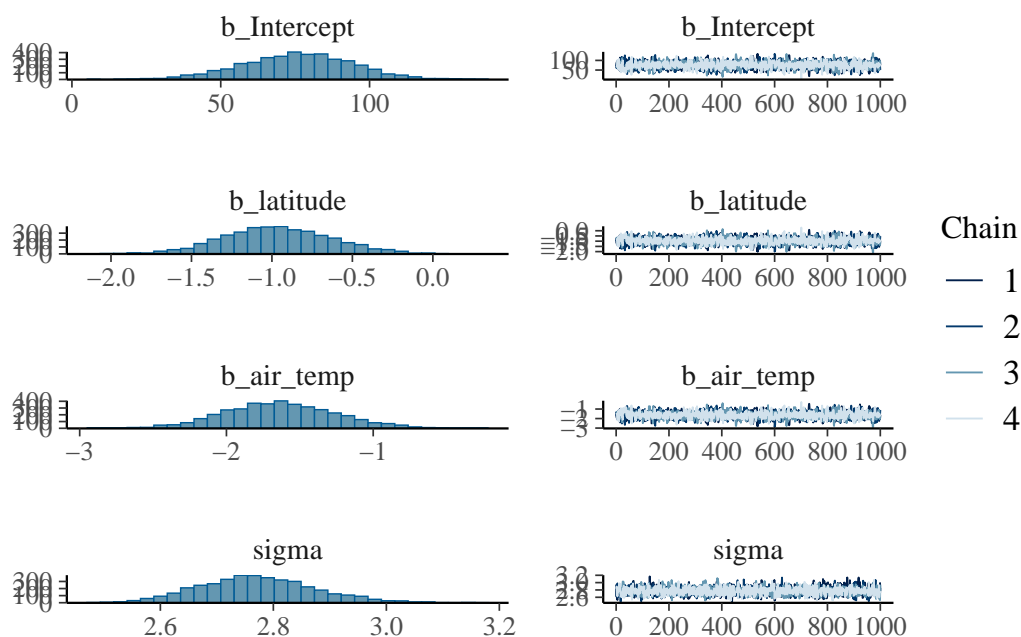
```
plot(m.crab.lat.air)
```

Q1.3a Assess the output

These all look good except for the posterior distribution for latitude which includes 0, but only on the tapering ends. The compatibility intervals do not contain zero.

Q1.3b Interpret the output

Interpret your model by answering:

1. What are the effects of your predictors? Remember to describe the effect using the units to make it biologically meaningful.

For every one degree of latitude increase, the crab size decreases by .96 mm. For every one degree C increase in air temp, the crab size decreases by 1.64 mm.

1. Are the effects reasonably different from zero? How do you know?

These are reasonably from zero because the compatibility interval doesn't include zero and the posterior distributions also do not include it for the most, with the exception of a small bin at the end of the latitude distribution.

```
# latitude and air model
m.crab.lat.air.water <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + air_temp + water_temp,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
      iter = 2000, warmup = 1000, chains = 4, cores = 6,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.air.water")
```

```
summary(m.crab.lat.air.water)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + air_temp + water_temp
   Data: pie_crab (Number of observations: 392)
  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     77.61     16.93    44.83   110.12 1.00     2404     2347
latitude      -1.06      0.31    -1.66    -0.45 1.00     2369     2368
air_temp      -2.40      0.37    -3.13    -1.68 1.00     2193     2340
water_temp     0.76      0.15     0.47     1.04 1.00     3431     2358

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.69      0.10     2.51     2.88 1.00     3460     2515

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).
```
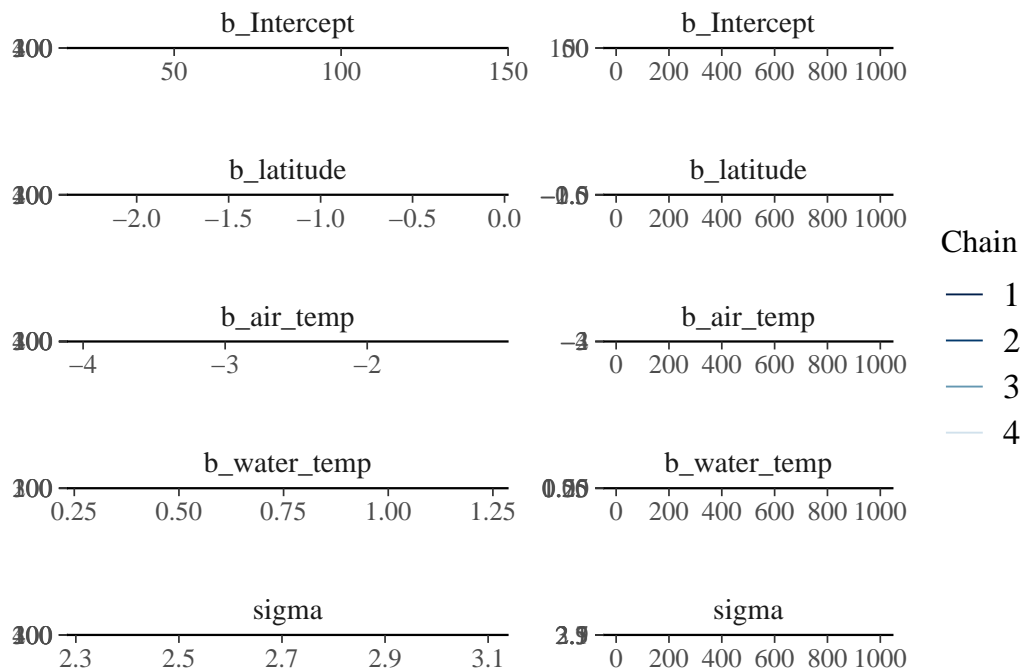
`plot(m.crab.lat.air.water)`

Q1.4a Assess the output

The model seems to have run correctly, and 0 is not included in any compatibility intervals or distributions. This seems to have worked better than the previous model.

Q1.4b Interpret the output

1. What are the effects of your predictors? Remember to describe the effect using the units to make it biologically meaningful.

For every one degree of latitude increase, the crab size decreases by 1.06 mm. For every one degree C increase in air temp, the crab size decreases by 2.4 mm. For every one degree C increase in water temp, the crab size increases by .76 mm.

1. Are the effects reasonably different from zero? How do you know?

Yes, the posterior and compatibility intervals did not overlap zero.

Q1.5 How do the models differ in their estimates?

In 2-4 sentences, compare the three models' estimates of the effect of latitude, water temp, and air temp; did estimates change across different models? Stay the same? Change in whether or not they are different from zero?

The first model we created showed that size increased with latitude, which did not hold up in either other model. When all included together, each predictor increased in their effect size on the crab size. All of them stayed different from zero however.

Q1.6 Why do you think a variable's sign changed?

You should have noticed the change in sign for a variable. In 1-2 sentences, and in the context of your knowledge about causal inference from DAGs from last week, describe why you think the variable may have changed signs (hint: remember pipes?).

We think this may have happened because the latitude is connected to both air and water temp via a pipe, and therefore its effect might be masked by temperature on the size, as size is more directly related to crab size in our DAG here. We'll have to condition on latitude in this situation to remove its effect.

```
# Look at "leave one out" results for all three models
# size ~ lat + mean water
loo(m.crab.lat.water)
```

```
Computed from 4000 by 392 log-likelihood matrix.

        Estimate   SE
elpd_loo   -963.0 13.6
p_loo         3.7  0.3
looic      1926.0 27.2
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.4, 0.9]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```

```
# size ~ lat + mean air
loo(m.crab.lat.air)
```

```
Computed from 4000 by 392 log-likelihood matrix.

        Estimate   SE
elpd_loo   -957.3 13.4
p_loo         4.1  0.3
looic      1914.7 26.9
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.5, 0.9]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```

```
# size ~ lat + mean water + mean air
loo(m.crab.lat.air.water)
```

```
Computed from 4000 by 392 log-likelihood matrix.

        Estimate   SE
elpd_loo   -945.1 14.2
p_loo         4.8  0.4
looic      1890.3 28.4
```

```
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.6, 1.0]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```

```
# Look at "leave one out" results for all three models
# size ~ lat + mean water
waic(m.crab.lat.water)
```

```
Computed from 4000 by 392 log-likelihood matrix.

          Estimate   SE
elpd_waic   -963.0 13.6
p_waic         3.7  0.3
waic        1926.0 27.2
```

```
# size ~ lat + mean air
waic(m.crab.lat.air)
```

```
Computed from 4000 by 392 log-likelihood matrix.

          Estimate   SE
elpd_waic   -957.3 13.4
p_waic         4.1  0.3
waic        1914.7 26.9
```
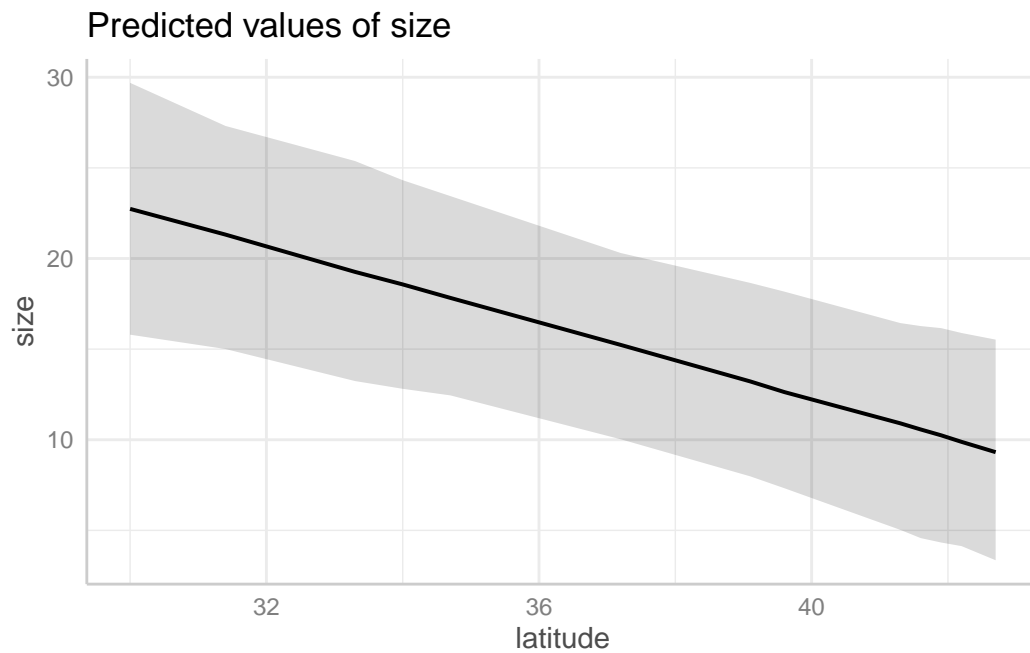
```
# size ~ lat + mean water + mean air
waic(m.crab.lat.air.water)
```

```
Computed from 4000 by 392 log-likelihood matrix.

          Estimate   SE
elpd_waic   -945.1 14.2
p_waic         4.8  0.4
waic        1890.3 28.3
```
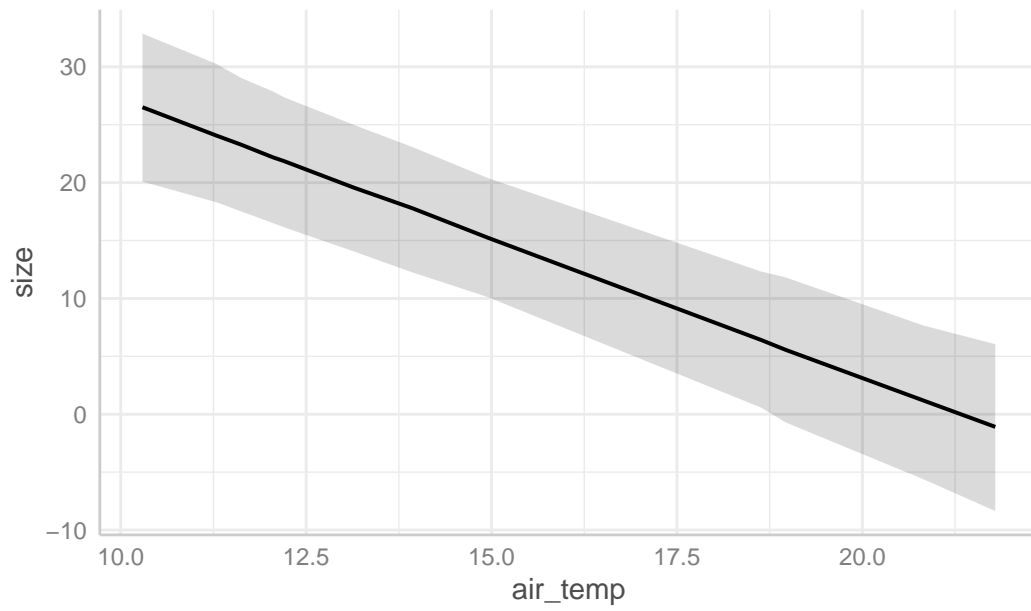
```
preds <- ggeffects::predict_response(m.crab.lat.air.water,
                            interval = "prediction")
plot(preds)
```
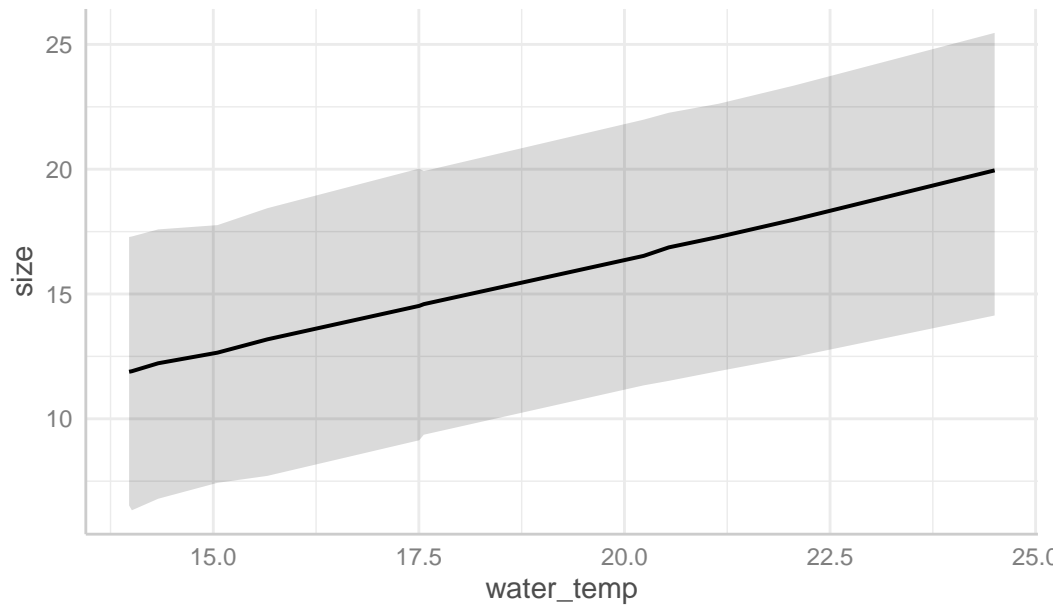
$latitude

Predicted values of size

$air_temp

Predicted values of size

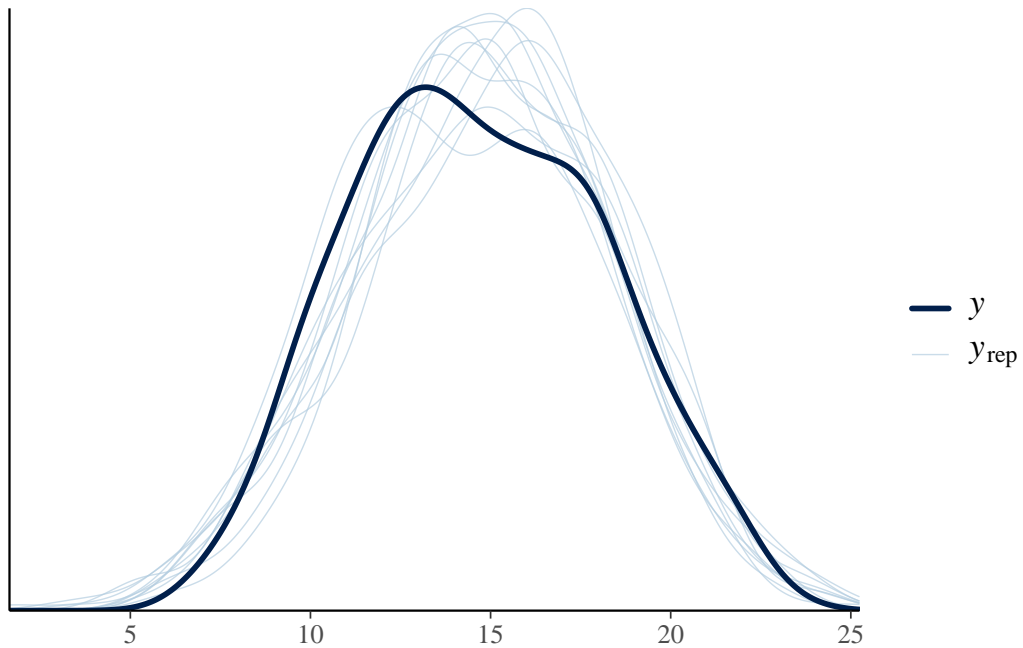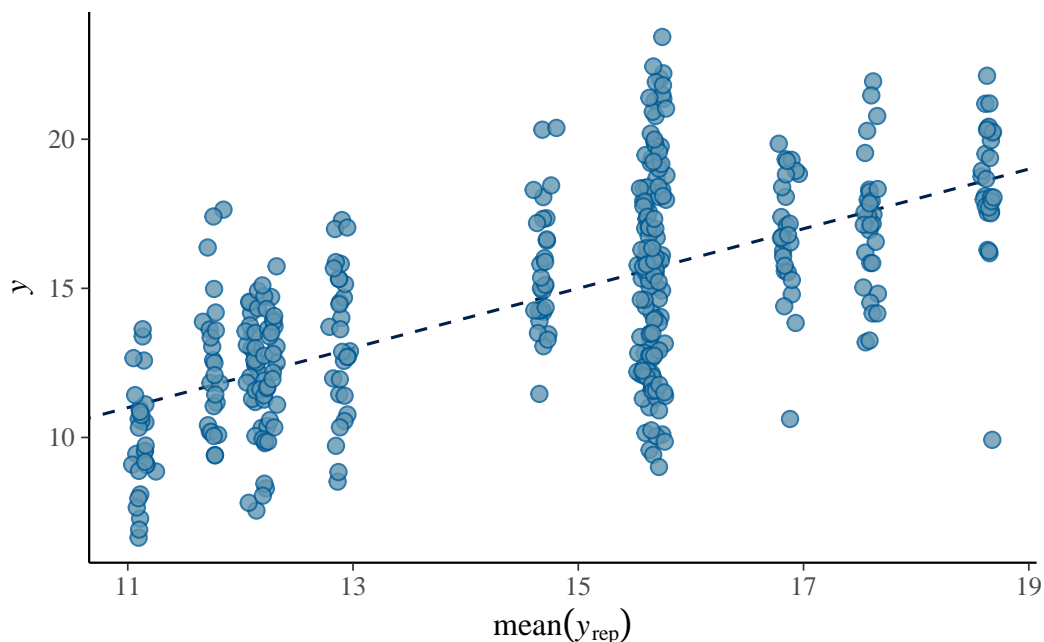$water_temp

Predicted values of size

```
pp_check(m.crab.lat.air.water, type = "dens_overlay")
```

Using 10 posterior draws for ppc type 'dens_overlay' by default.


```
pp_check(m.crab.lat.air.water, type = "scatter_avg")
```

Using all posterior draws for ppc type 'scatter_avg' by default.

```
m.crab.lat.water.sd <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + water_temp_sd,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
      iter = 2000, warmup = 1000, chains = 4, cores = 6,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.water.sd")
m.crab.lat.air.sd <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + air_temp_sd,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
```

```
      iter = 2000, warmup = 1000, chains = 4, cores = 6,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.air.sd")
m.crab.lat.air.water.sd <-
  brm(data = pie_crab, # Give the model the pie_crab data
      # Choose a gaussian (normal) distribution
      family = gaussian,
      # Specify the model here.
      size ~ latitude + air_temp_sd + water_temp_sd,
      # Here's where you specify parameters for executing the Markov chains
      # We're using similar to the defaults, except we set cores to 4 so the analysis runs fa
      iter = 2000, warmup = 1000, chains = 4, cores = 6,
      # Setting the "seed" determines which random numbers will get sampled.
      # In this case, it makes the randomness of the Markov chain runs reproducible
      # (so that both of us get the exact same results when running the model)
      seed = 4,
      # Save the fitted model object as output - helpful for reloading in the output later
      file = "output/m.crab.lat.air.water.sd")
```

```
summary(m.crab.lat.water.sd)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + water_temp_sd
   Data: pie_crab (Number of observations: 392)
  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        -3.78      1.53    -6.74    -0.80 1.00     3958     2740
latitude          0.48      0.03     0.42     0.55 1.00     4130     2988
water_temp_sd     0.02      0.13    -0.22     0.26 1.00     4308     3101


Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.84      0.10     2.65     3.06 1.00     5147     3246
```
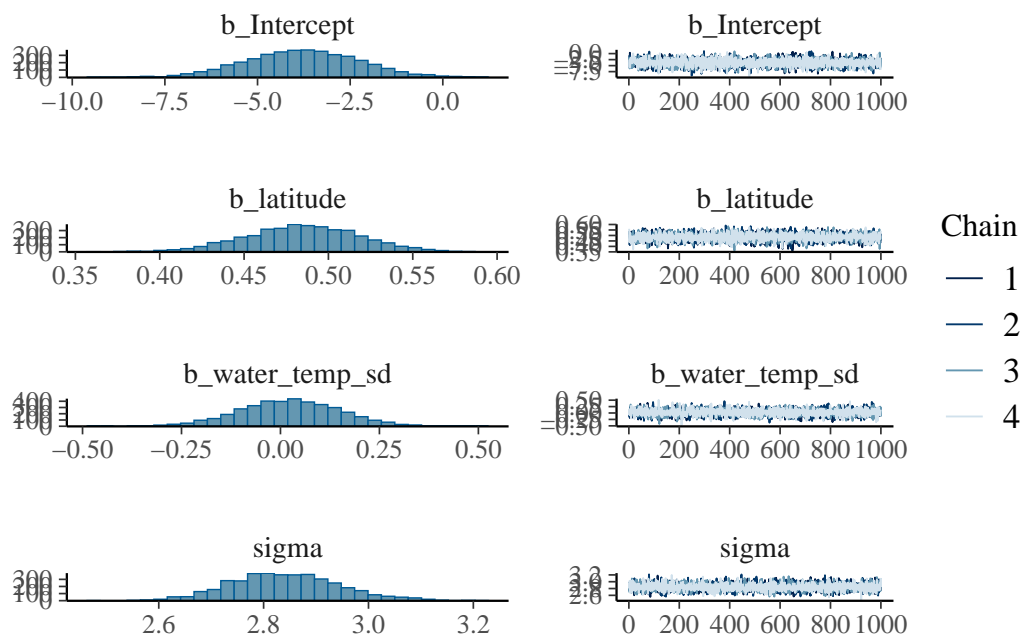
Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
plot(m.crab.lat.water.sd)
```



```
summary(m.crab.lat.air.sd)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + air_temp_sd
   Data: pie_crab (Number of observations: 392)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
```
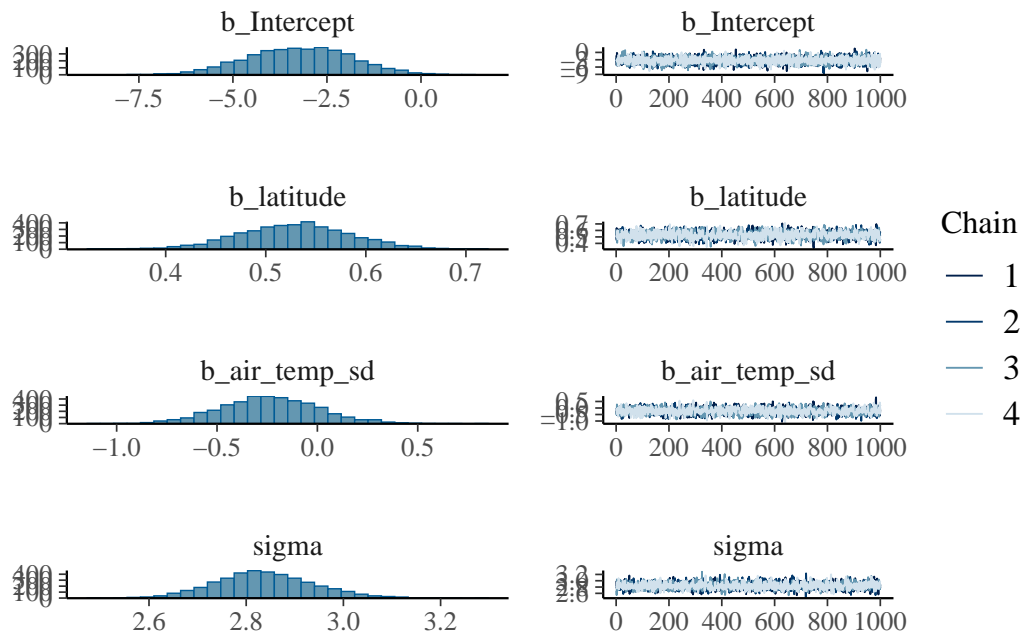
Regression Coefficients:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-3.24	1.38	-5.91	-0.58	1.00	4921	2967
latitude	0.53	0.06	0.42	0.64	1.00	2178	2586
air_temp_sd	-0.24	0.24	-0.71	0.26	1.00	2150	2404

```
Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.84      0.10     2.65     3.05 1.00     3557     2711


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(m.crab.lat.air.sd)
```



```
summary(m.crab.lat.air.water.sd)
```

```
 Family: gaussian
  Links: mu = identity
Formula: size ~ latitude + air_temp_sd + water_temp_sd
   Data: pie_crab (Number of observations: 392)
  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     -3.95      1.56    -7.02    -0.91 1.00     4084     3651
```

```
latitude            0.56      0.06      0.43     0.69 1.00       1963       2622
air_temp_sd        -0.42      0.30     -1.02     0.17 1.00       1827       2359
water_temp_sd       0.16      0.16     -0.17     0.48 1.00       2396       2508
```
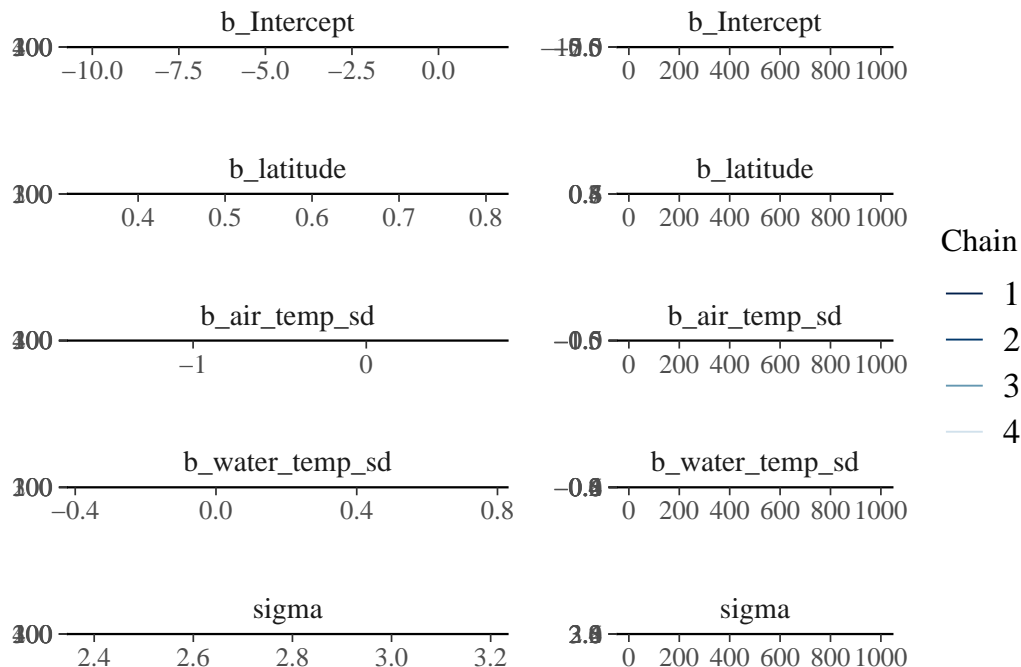
```
Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.84      0.10      2.65      3.05 1.00      3379       2820
```

```
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(m.crab.lat.air.water.sd)
```


Q1.4a Assess the output

Assess whether the model ran correctly by looking at R hat, the chains, and the posterior distributions using the plot() and summary() functions. Describe your thought process about whether the model ran correctly in 1-2 sentences.

1st model (water): The model seems to have run, and the posteriors and compatibilities of the values do not include zero. The chains are overlapping and the R hat is 1

19

2nd Model (air): The model seems to have run, and the posteriors and compatibilities of the values do include zero for air temp. The chains are overlapping and the R hat is 1

3rd Model (air and water): The model seems to have run, and the posteriors and compatibilities of the values do include zero for air temp. The chains are overlapping and the R hat is 1

Q1.4b Interpret the output

Interpret your model by answering:

1. What are the effects of your predictors? Remember to describe the effect using the units to make it biologically meaningful.

1st model: For every one degree of latitude increase, the crab size decreases by .48 mm. For every one degree C increase in water temp sd, the crab size increases by .02 mm.

2nd model: For every one degree of latitude increase, the crab size decreases by .06 mm. For every one degree C increase in air temp sd, the crab size decreases by .24 mm.

3rd model: For every one degree of latitude increase, the crab size decreases by .56 mm. For every one degree C increase in air temp, the crab size decreases by -.42 mm. For every one degree C increase in water temp, the crab size increases by .16 mm.

1. Are the effects reasonably different from zero? How do you know?

1st model: Yes, the compatibility intervals do not include zero.

2nd model: No, the compatibility intervals do include zero for air temp but not the other.

3rd model: No, the compatibility intervals do include zero for air temp but not the others.

Q2.4 How do the models differ in their parameter estimates?

In 2-4 sentences, compare the three models' estimates of the effect of latitude, water temp sd, and air temp sd; did estimates change across different models? Stay the same? Change in whether or not they are different from zero?

The only time there was a change in differences from zero were for air temp sd, however, it wasn't substantially different from zero. The water temp sd estimate increased when also conditioned on air temp sd however.

Q2.5 Calculate and compare PSIS and AIC values for each model

Calculate and compare the PSIS and AIC values for each model and answer:

1. Are the Pareto k estimates good?

All Pareto K estimates are good.

1. Which model has the lowest PSIS?

The latitude and air have the lowest PSIS

1. Which model has the lowest AIC?

The latitude and air have the lowest AIC

1. Do PSIS and AIC values agree on which model has the best out of sample prediction?

Yes they agree that the air temp and latitude model have best out of sample prediction.

```
waic(m.crab.lat.water.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

          Estimate   SE
elpd_waic   -967.1 13.3
p_waic         3.8  0.3
waic        1934.3 26.7
```

```
# size ~ lat + mean air
waic(m.crab.lat.air.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

          Estimate   SE
elpd_waic   -966.6 13.5
p_waic         3.5  0.3
waic        1933.1 26.9
```

```
# size ~ lat + mean water + mean air
waic(m.crab.lat.air.water.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

         Estimate   SE
elpd_waic   -967.1 13.2
p_waic         4.6  0.4
waic        1934.2 26.5
```

```
loo(m.crab.lat.water.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

         Estimate   SE
elpd_loo    -967.1 13.3
p_loo          3.8  0.3
looic       1934.3 26.7
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [1.0, 1.3]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```

```
# size ~ lat + mean air
loo(m.crab.lat.air.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

         Estimate   SE
elpd_loo    -966.6 13.5
p_loo          3.5  0.3
looic       1933.1 26.9
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.5, 1.1]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```

```
# size ~ lat + mean water + mean air
loo(m.crab.lat.air.water.sd)
```

```
Computed from 4000 by 392 log-likelihood matrix.

         Estimate   SE
elpd_loo   -967.1 13.2
p_loo         4.6  0.4
looic      1934.3 26.5
------
MCSE of elpd_loo is 0.0.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.5, 1.2]).

All Pareto k estimates are good (k < 0.7).
See help('pareto-k-diagnostic') for details.
```