

Bayesian inference of an individual-based mutualistic network

20_01

Net 20_01

```
library(BayesianNetworks)
library(network.tools)
library(tidyverse)
theme_set(theme_minimal())
options(mc.cores = 4)
```

Data

Load dataset and sampling effort per individual plant:

```
web <- readr::read_csv(here::here("data/nets_raw", paste0(params$net, "_int.csv"))) |>
  arrange(ind)
```

```
## Rows: 19 Columns: 15
## -- Column specification -----
## Delimiter: ","
## chr (1): ind
## dbl (14): Erithacus_rubecula, Sylvia_atricapilla, Turdus_philomelos, Turdus_merula, Psammmodromus_alg
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
mat <- as.matrix(web[, -1])
mat <- apply(mat, c(1,2), as.integer)
rownames(mat) <- web$ind
```

```
# create numeric vector of sampling effort for each plant with names = plant id
effort <- readr::read_csv(here::here("data/nets_attr", paste0(params$net, "_attr.csv"))) |>
  select(ind, starts_with("se_")) |>
  filter(ind %in% web$ind) |>
  arrange(ind)
```

```
## Rows: 20 Columns: 8
## -- Column specification -----
## Delimiter: ","
## chr (3): ind, fruit_type, fruit_color
## dbl (5): x, y, se_bc_days, se_bc_area, seeds_per_fruit
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
## If there is only one column with sampling effort, use it:
if (!net %in% c("10_01", "15_01", "18_01", "18_02", "20_01", "21_01", "21_02")) {
  effort <- effort |>
    pull(starts_with("se_"), name = "ind")
}

# Otherwise, select sampling effort column in some specific nets:
if (net == "10_01") {
  effort <- effort |>
    mutate(se_cam_days = se_cam_h/24) |>
    pull(se_cam_days, name = "ind")
}

if (net == "15_01") {
  effort <- effort |>
    pull(se_cam_days, name = "ind")
}

if (net %in% c("18_01", "18_02", "20_01")) {
  effort <- effort |>
    mutate(se_bc_months = se_bc_days/30) |>
    pull(se_bc_months, name = "ind")
}

if (net %in% c("21_01", "21_02")) {
  effort <- effort |>
    pull(se_obs_h, name = "ind")
}

## Some nets may require adjusting of the count data or effort values
## Insert that here eg.
# if (params$net == "01_01") {
#   mat <- round(mat/10)
#   mat <- apply(mat, c(1,2), as.integer)
# }

stopifnot(identical(length(effort), nrow(mat)))
stopifnot(identical(names(effort), rownames(mat)))

summary(as.numeric(mat))
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
##         0         0         0         1         0        32
```

```
if (max(mat) > 500) {
  stop("More than 500 counts in some cell(s)")
}

summary(effort)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      7.967   7.967   7.967   8.149   8.400   8.400
```

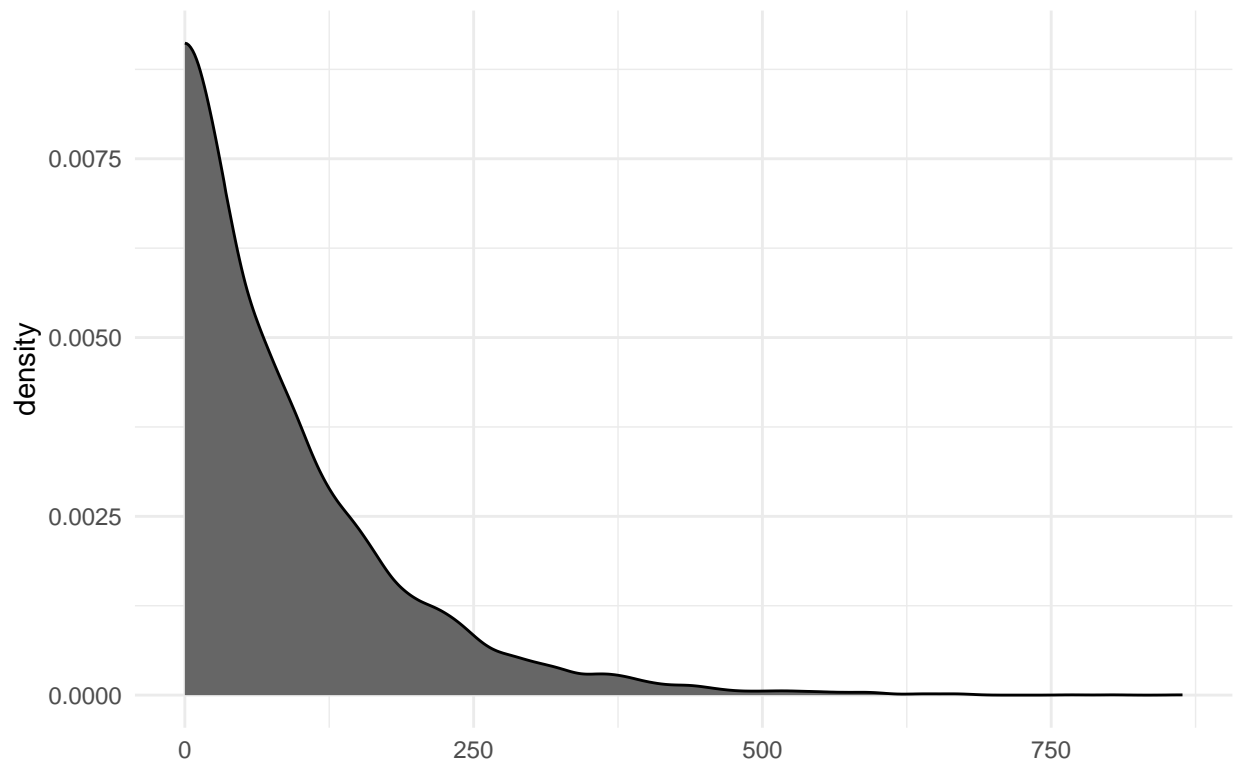
```
if (max(effort) > 500) {
  stop("Sampling effort > 500 for some plants")
}
```

Bayesian inference of network structure

```
dt <- prepare_data(mat, sampl.eff = effort)

plot_prior(params$beta)
```

Prior probability for r (preference) parameter with $\beta = 0.01$



```
fit <- fit_model(dt,
  refresh = 0,
  beta = params$beta,
  model = params$model,
  # max_treedepth = 15,
  # init = function() list(r = runif(1, 0, 20000)),
  iter_warmup = params$iter,
  iter_sampling = params$iter,
  thin = 4 * params$iter / 1000)
```

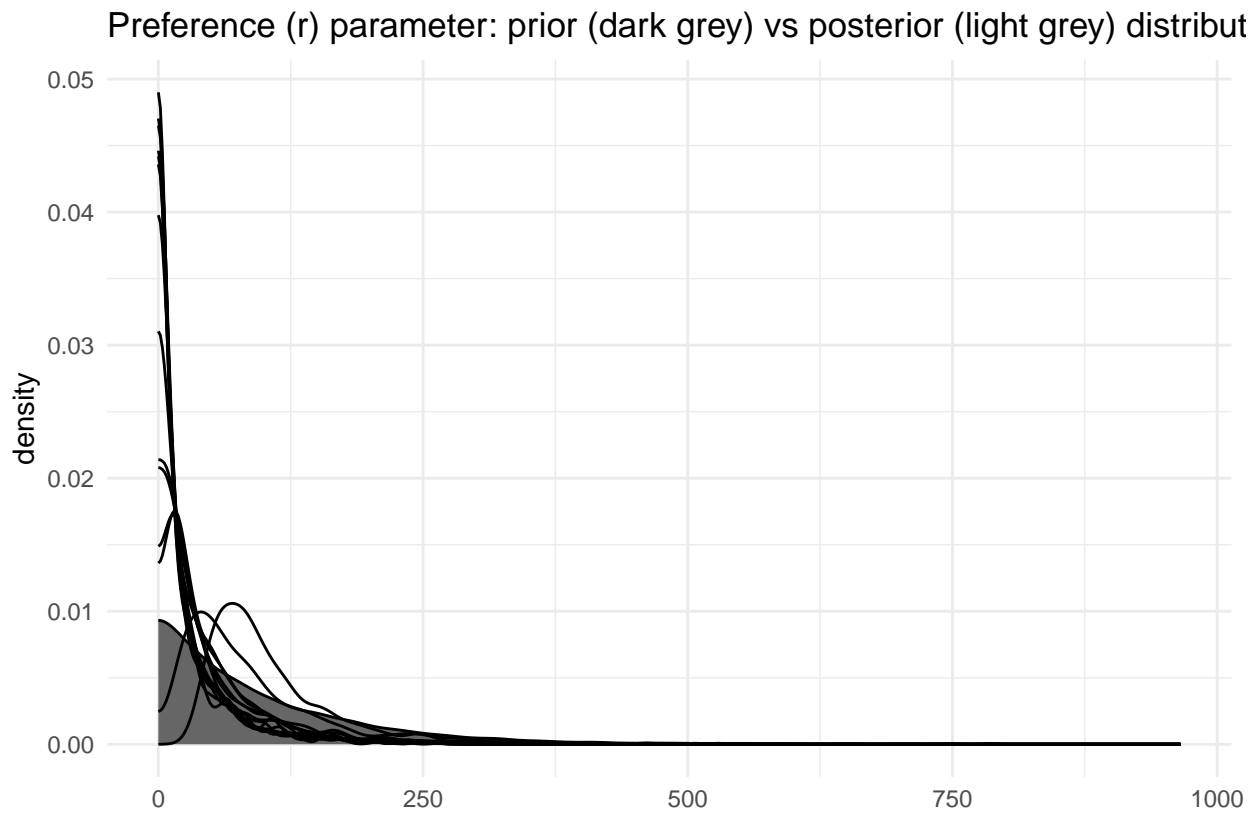
```
## Running MCMC with 4 parallel chains...
##
## Chain 2 finished in 28.9 seconds.
## Chain 4 finished in 29.9 seconds.
## Chain 1 finished in 30.8 seconds.
## Chain 3 finished in 31.2 seconds.
##
## All 4 chains finished successfully.
## Mean chain execution time: 30.2 seconds.
## Total execution time: 31.2 seconds.
```

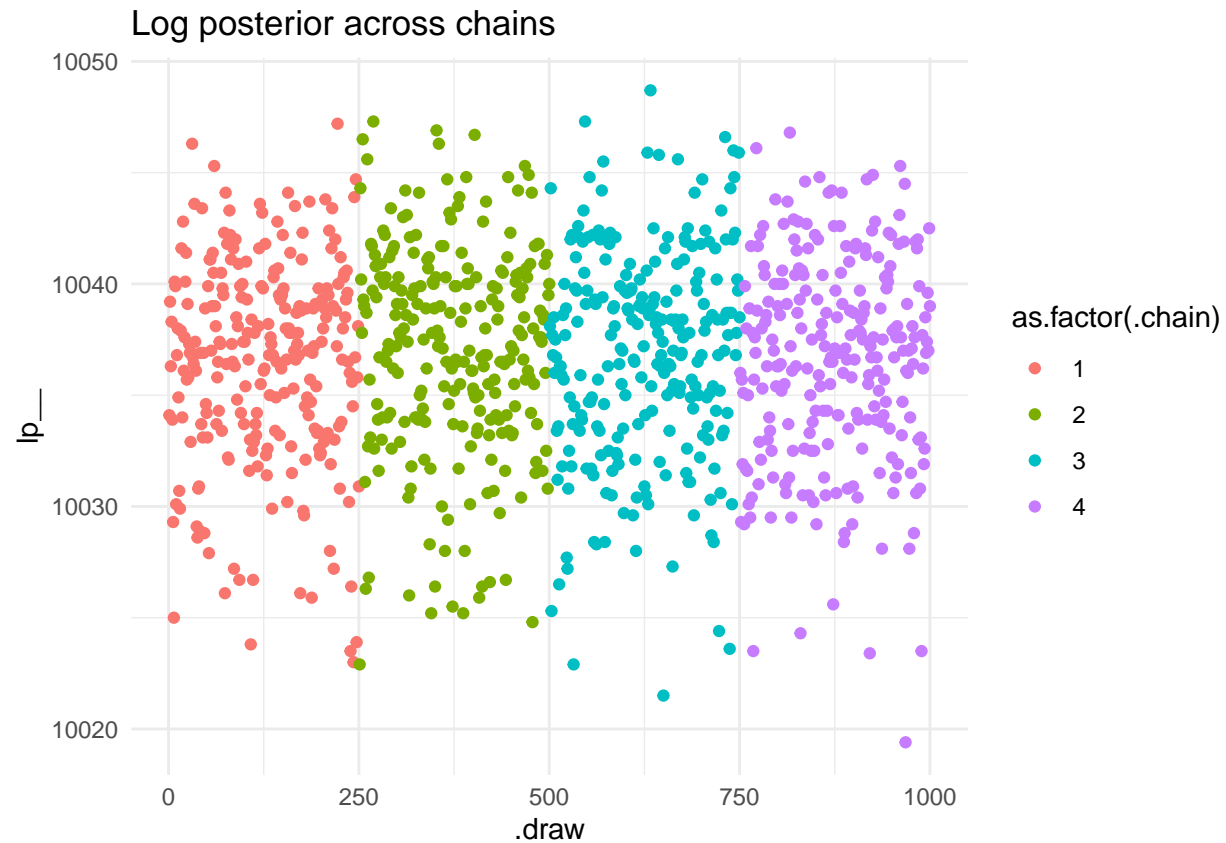
```
get_seed(fit)
```

```
## [1] 535778454
```

```
check_model(fit, data = dt)
```

```
## Processing csv files: C:/Users/frodr/AppData/Local/Temp/RtmpkDNfRE/varying_preferences-202406241437-
##
## Checking sampler transitions treedepth.
## Treedepth satisfactory for all transitions.
##
## Checking sampler transitions for divergences.
## No divergent transitions found.
##
## Checking E-BFMI - sampler transitions HMC potential energy.
## E-BFMI satisfactory.
##
## Effective sample size satisfactory.
##
## Split R-hat values satisfactory all parameters.
##
## Processing complete, no problems detected.
```





Posteriors

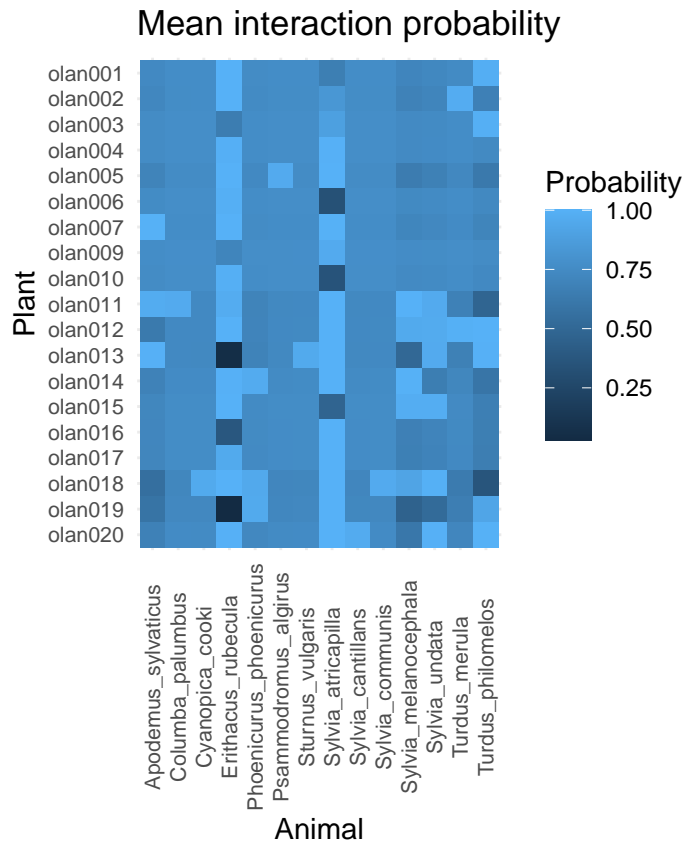
Get posterior distributions:

```
post <- get_posterior(fit, dt)
head(post)
```

```
## # A tibble: 6 x 11
## # Groups:   Animal, Plant [6]
##   Plant   Animal   .chain .iteration .draw connectance preference plant.abund animal.abund int.pr
##   <chr>   <chr>     <int>    <int>  <int>      <dbl>      <dbl>      <dbl>      <dbl>    <dbl>
## 1 olan001 Erithacus_~ 1         1      1        0.875       96.2       0.0224      0.0689    1.0
## 2 olan002 Erithacus_~ 1         1      1        0.875       96.2       0.0350      0.0689    1
## 3 olan003 Erithacus_~ 1         1      1        0.875       96.2       0.00869     0.0689    0.8
## 4 olan004 Erithacus_~ 1         1      1        0.875       96.2       0.00774     0.0689    0.9
## 5 olan005 Erithacus_~ 1         1      1        0.875       96.2       0.0468      0.0689    1
## 6 olan006 Erithacus_~ 1         1      1        0.875       96.2       0.00972     0.0689    0.9
```

Mean edge probability:

```
plot_interaction_prob(post)
```

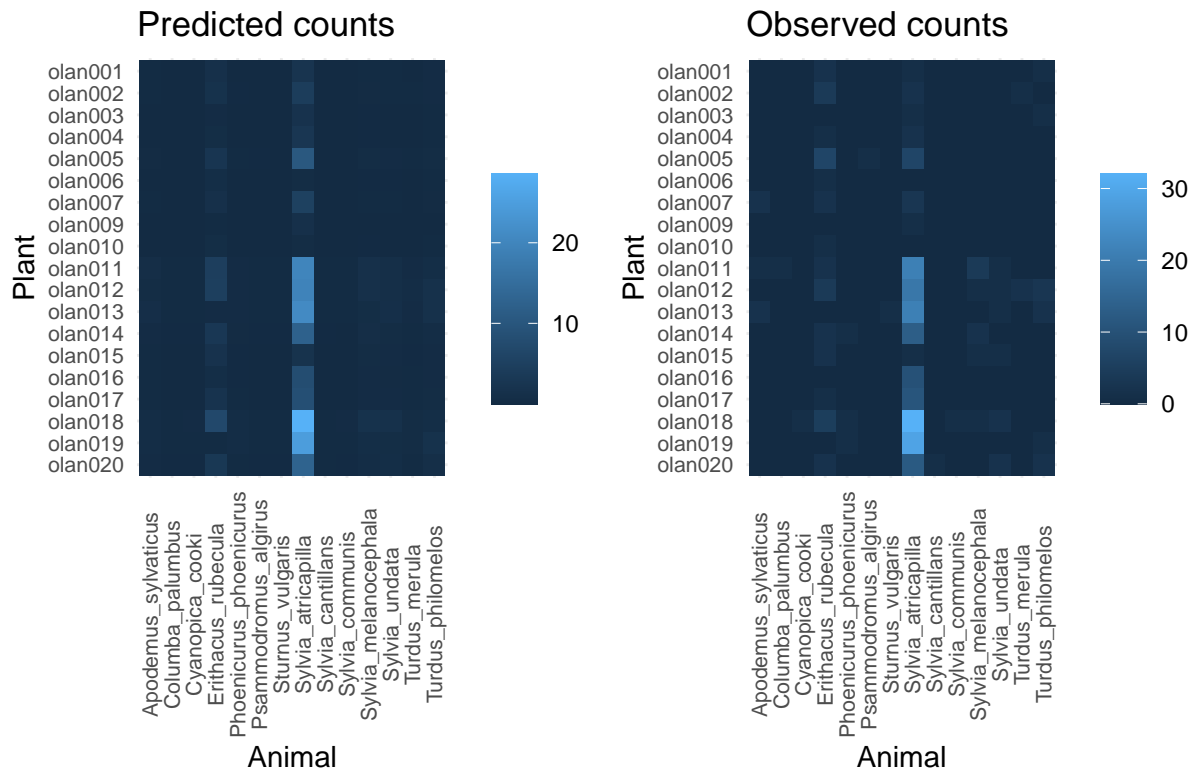


Generate predicted visits for each pairwise interaction

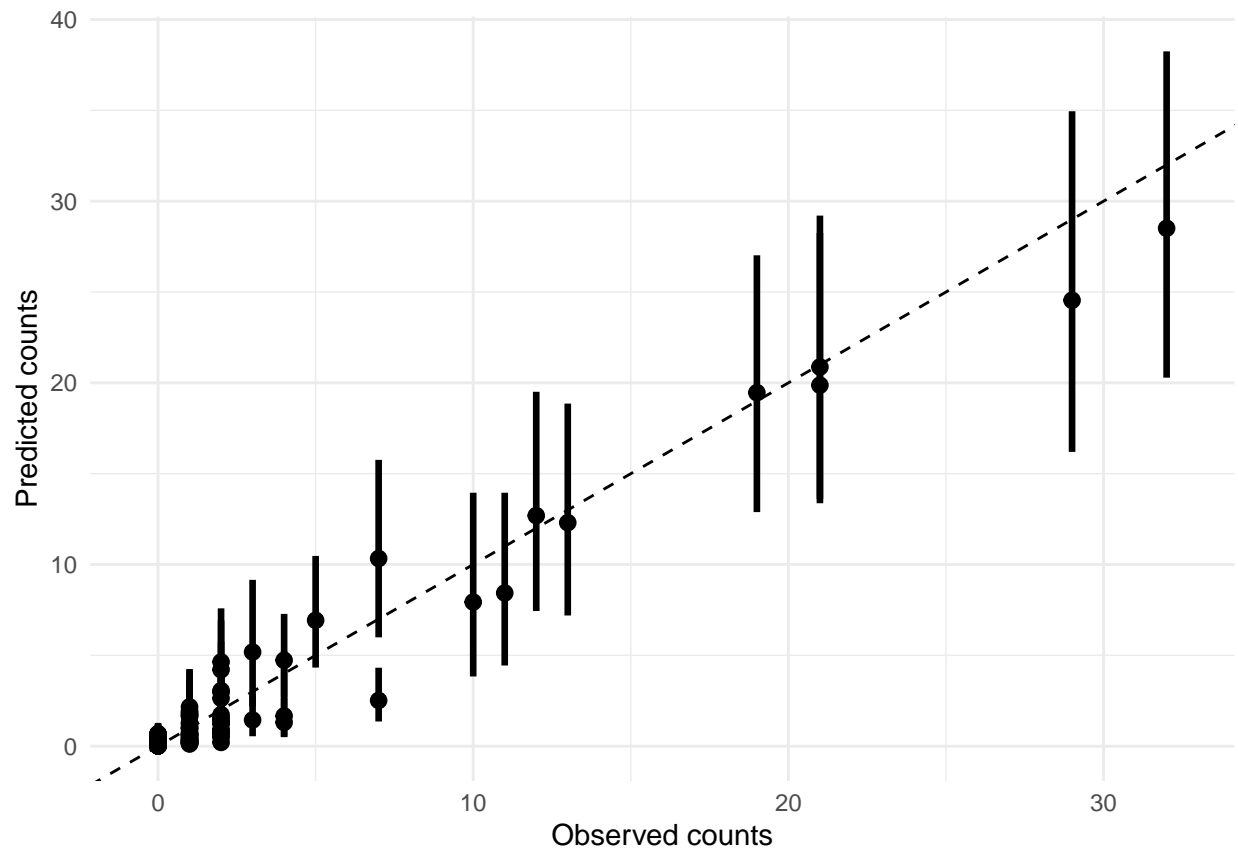
```
post.counts <- predict_counts(fit, dt)
```

Compare observed and predicted visits by the model:

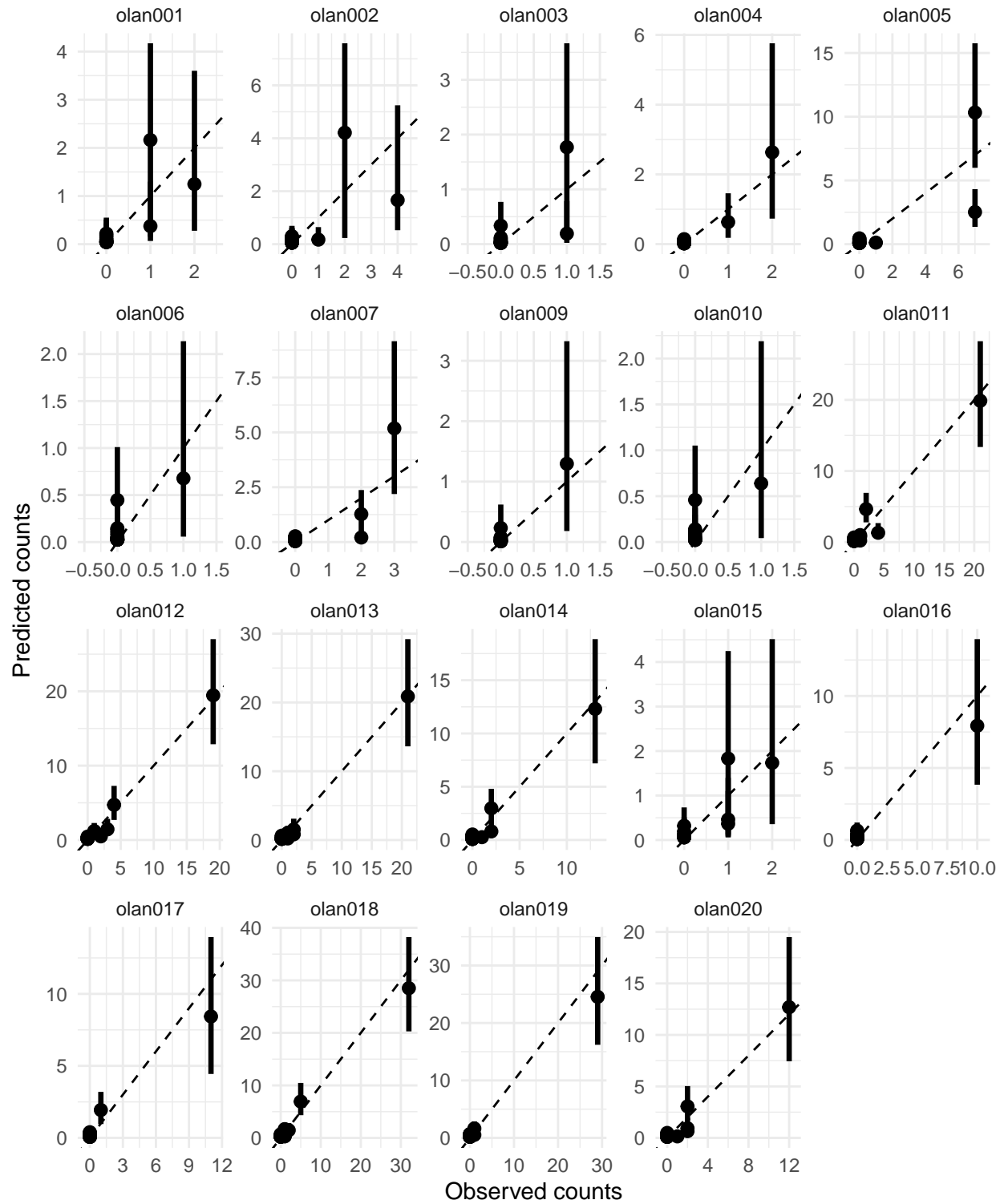
```
p <- plot_counts_pred(post.counts, sort = FALSE)
o <- plot_counts_obs(mat, sort = FALSE, zero.na = FALSE)
library(patchwork)
p + o
```



```
plot_counts_pred_obs(post.counts, dt)
```

```
plot_counts_pred_obs(post.counts, dt, byplant = TRUE, scales = "free")
```



```
saveRDS(post.counts, here::here(paste0("data/nets_post/", params$net, "_post_counts.rds")))
```