

# Bayesian inference of an individual-based mutualistic network

03\_01

## Net 03\_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("networks/nets_raw", paste0(params$net, "_int.csv"))) |>
  arrange(ind)
mat <- as.matrix(web[, -1])
mat <- apply(mat, c(1,2), as.integer)
rownames(mat) <- web$ind

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

## If there is only one column with sampling effort, use it:
if (!net %in% c("01_01", "01_02", "02_01", "02_02", "02_03", "10_01", "11_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 == "11_01") {
  effort <- effort |>
```

```

    mutate(se_cam_months = se_cam_days/30) |>
    pull(se_cam_months, name = "ind")
}

if (net == "15_01") {
  effort <- effort |>
    mutate(se_cam_months = se_cam_days/30) |>
    pull(se_cam_months, 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")
}

# For Pistacia and Juniperus, use constant sampling effort
if (net %in% c("01_01", "01_02", "02_01", "02_02", "02_03")) {
  effort <- rep(10, nrow(mat))
  names(effort) <- web$ind
}

## Some nets may require adjusting of the count data or effort values
if (net %in% c("02_01")) {
  mat <- mat/10 # divide counts by 10 to make modelling feasible
  mat[mat > 0 & mat < 1] <- 1 # don't miss rare counts
  mat <- round(mat)
  mat <- apply(mat, c(1,2), as.integer)
}

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

# summary(mat)
summary(as.numeric(mat))

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000  0.000   0.000   2.423   2.000  41.000

# if (max(mat) > 1000) {
#   stop("More than 1000 counts in some cell(s)")
# }

summary(effort)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         4         4         4         4         4         4

```

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

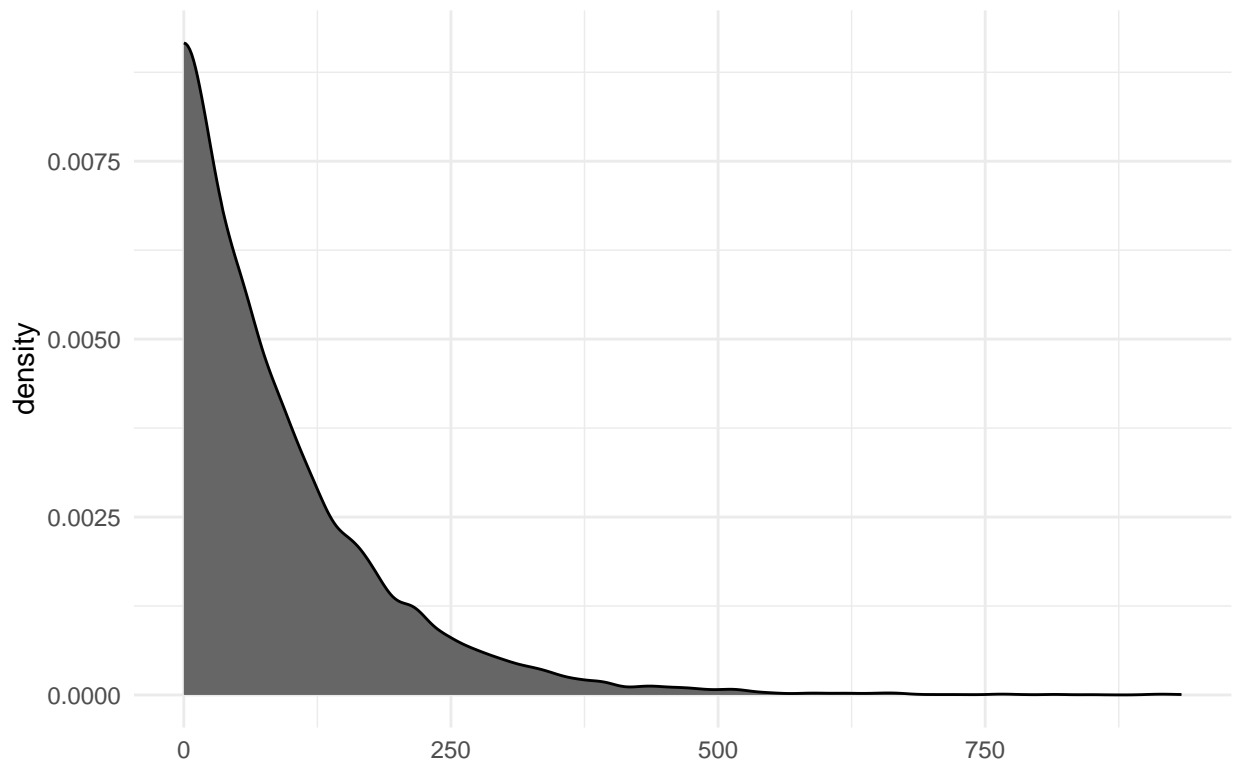
## Bayesian inference of network structure

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

beta <- params$beta

plot_prior(beta)
```

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



```
fit <- fit_model(dt,
  refresh = 0,
  beta = 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 4 finished in 11.7 seconds.  
## Chain 1 finished in 11.9 seconds.  
## Chain 3 finished in 11.8 seconds.  
## Chain 2 finished in 11.9 seconds.  
##  
## All 4 chains finished successfully.  
## Mean chain execution time: 11.9 seconds.  
## Total execution time: 12.1 seconds.
```

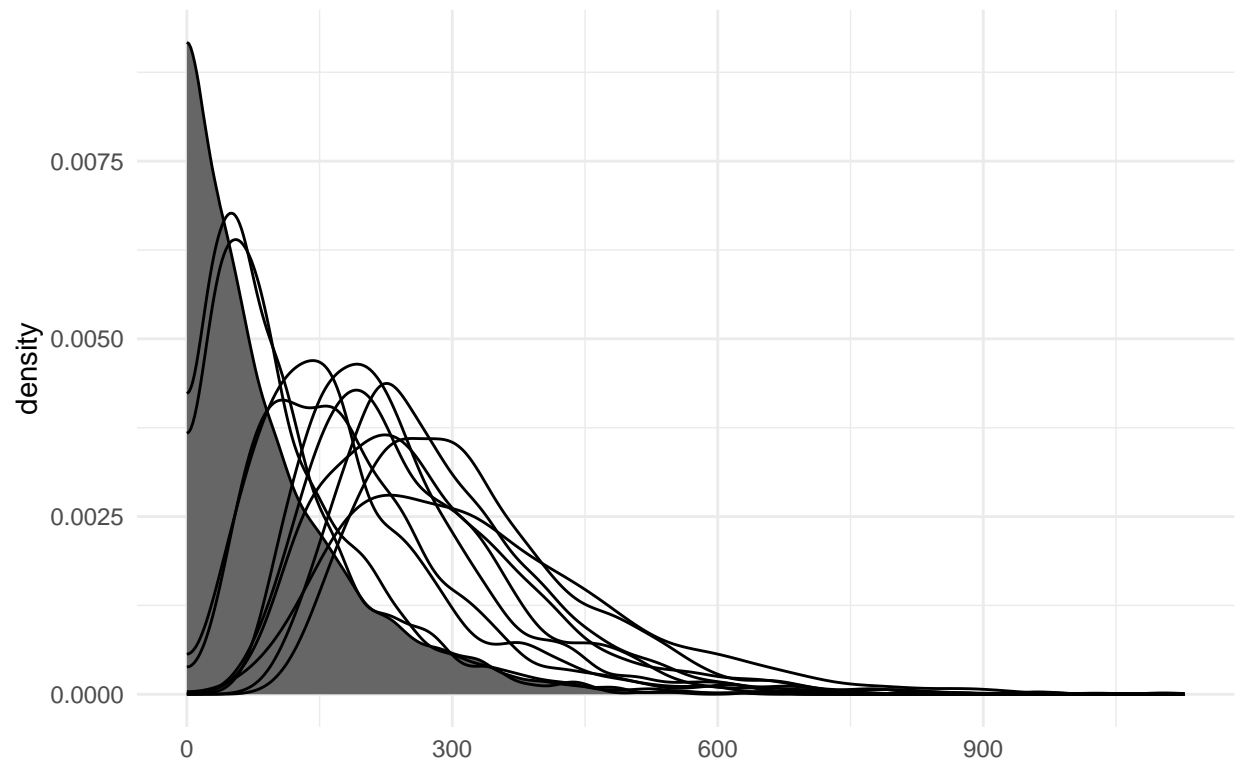
```
get_seed(fit)
```

```
## [1] 2047405762
```

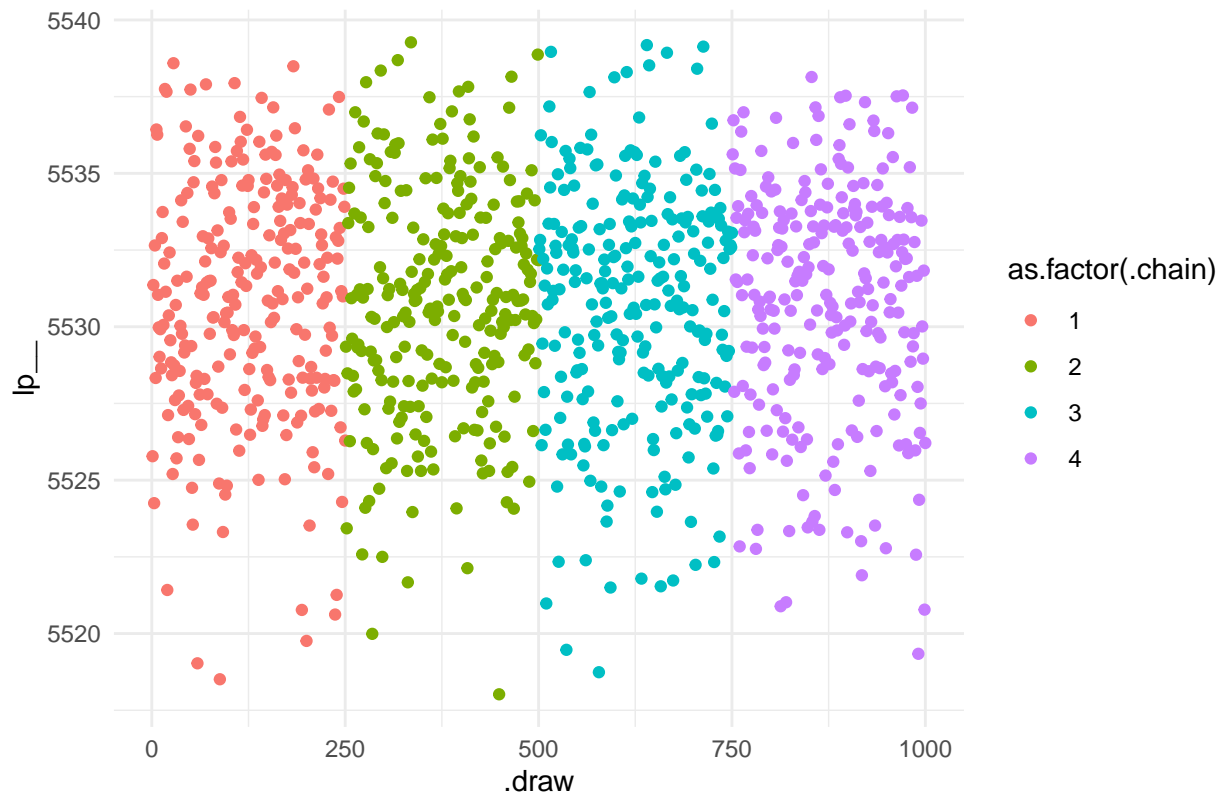
```
check_model(fit, data = dt)
```

```
## Processing csv files: /var/folders/zr/_6trblk14l373mf22yjy42740000gn/T/RtmpRtXnKT/varying_preferences  
##  
## 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.
```

Preference (r) parameter: prior (dark grey) vs posterior (light grey) distrib



Log posterior across chains



## Posteriors

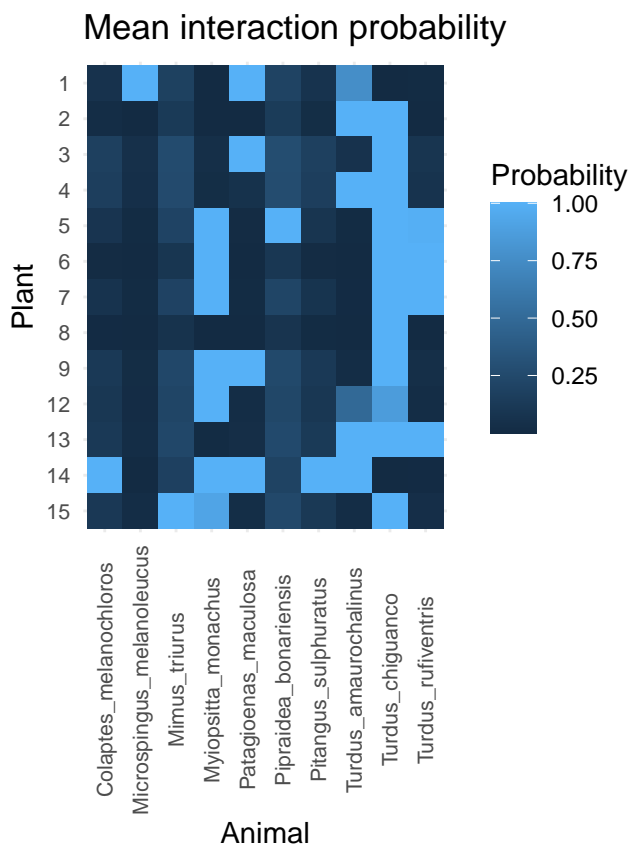
Get posterior distributions:

```
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 in
##   <chr> <chr>      <int>      <int> <int>      <dbl>      <dbl>      <dbl>      <dbl>
## 1 1    Turdus_chiguanco 1          1      1      0.270      330.      0.0750      0.0966 0.0
## 2 2    Turdus_chiguanco 1          1      1      0.270      330.      0.159      0.0966 1.0
## 3 3    Turdus_chiguanco 1          1      1      0.270      330.      0.0320      0.0966 1.0
## 4 4    Turdus_chiguanco 1          1      1      0.270      330.      0.0277      0.0966 1.0
## 5 5    Turdus_chiguanco 1          1      1      0.270      330.      0.0580      0.0966 1.0
## 6 6    Turdus_chiguanco 1          1      1      0.270      330.      0.180      0.0966 1.0
```

Mean edge probability:

```
plot_interaction_prob(post)
```

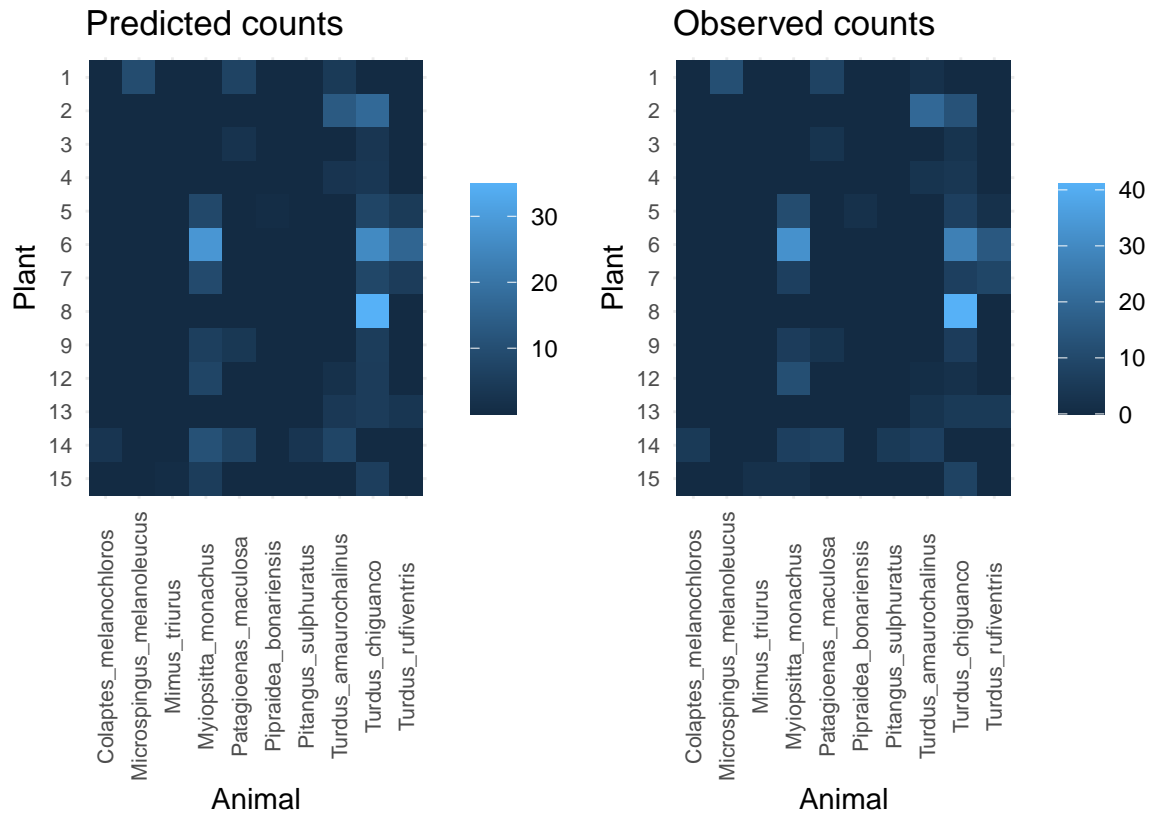


## Generate predicted visits for each pairwise interaction

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

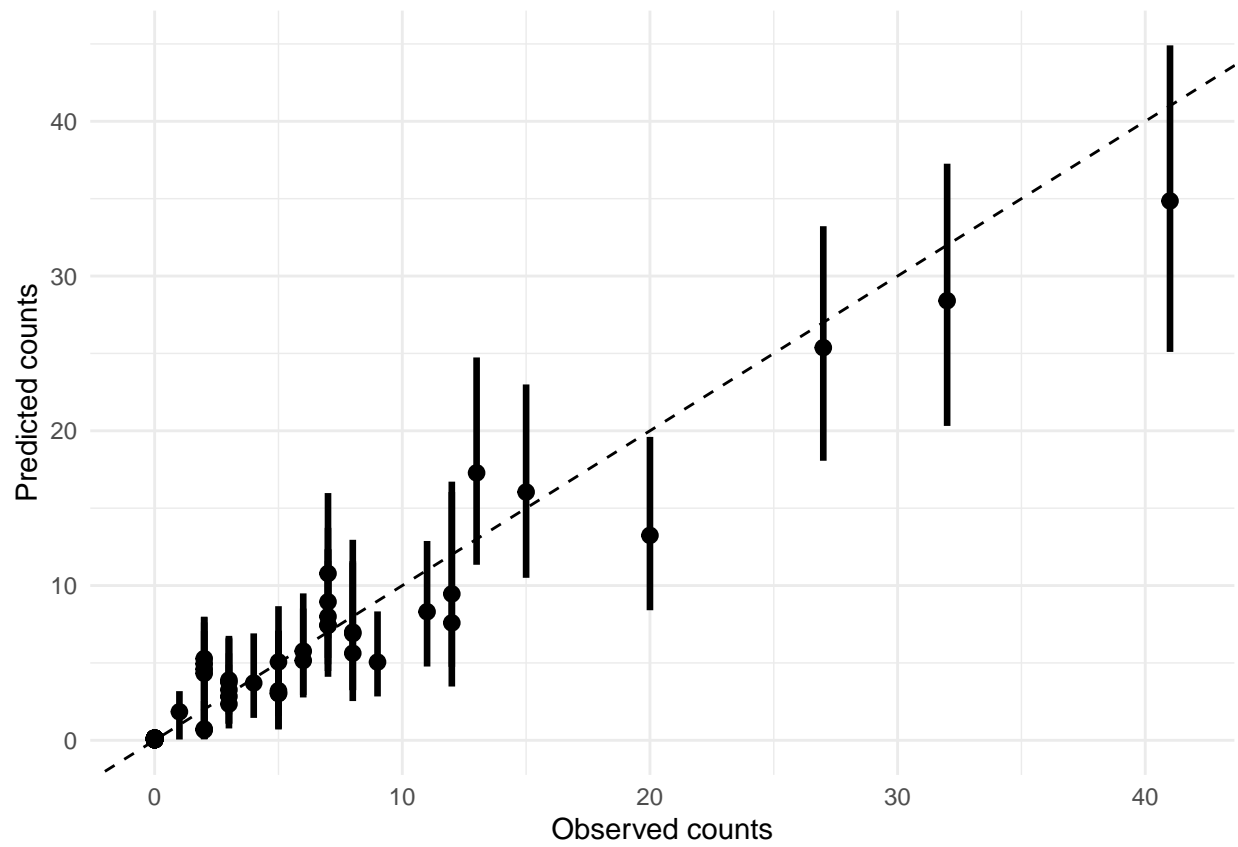
Compare observed and predicted visits by the model:

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

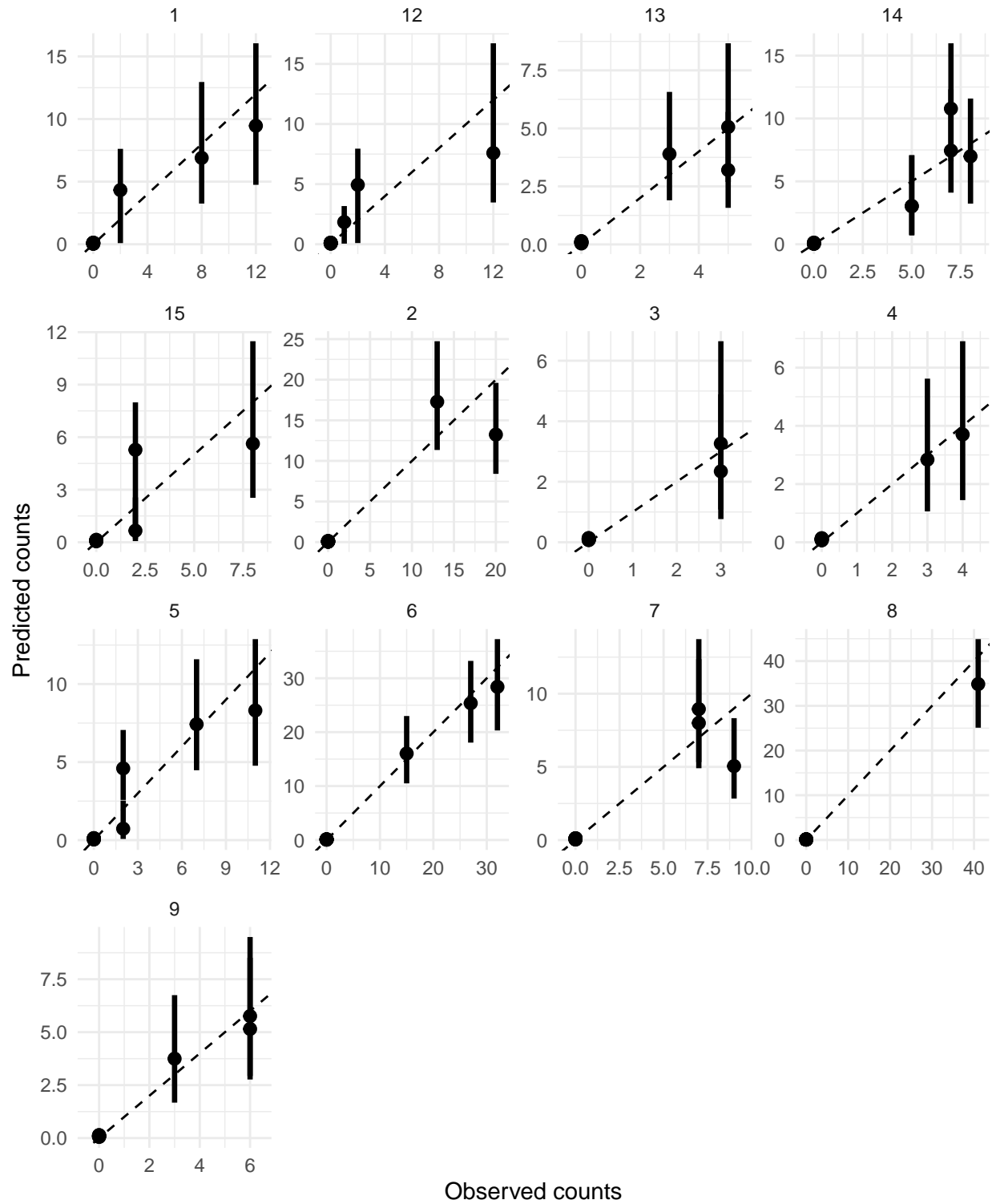


```
plot_counts_pred_obs(post.counts, dt)
```





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



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