# Straightforward residual diagnostics for glm/glmer using the package DHARMa

Lisa Hülsmann

Standard residuals of generalized models are hard to interpret

`plot(yourGLM)`

- distributions are asymmetric

- residuals are never visually homogenous – even if the model is perfect

- more difficult for hierarchical models (random effects)

- but diagnostics are equally important
  - overdispersion (no variance parameter)
  - zero-inflation
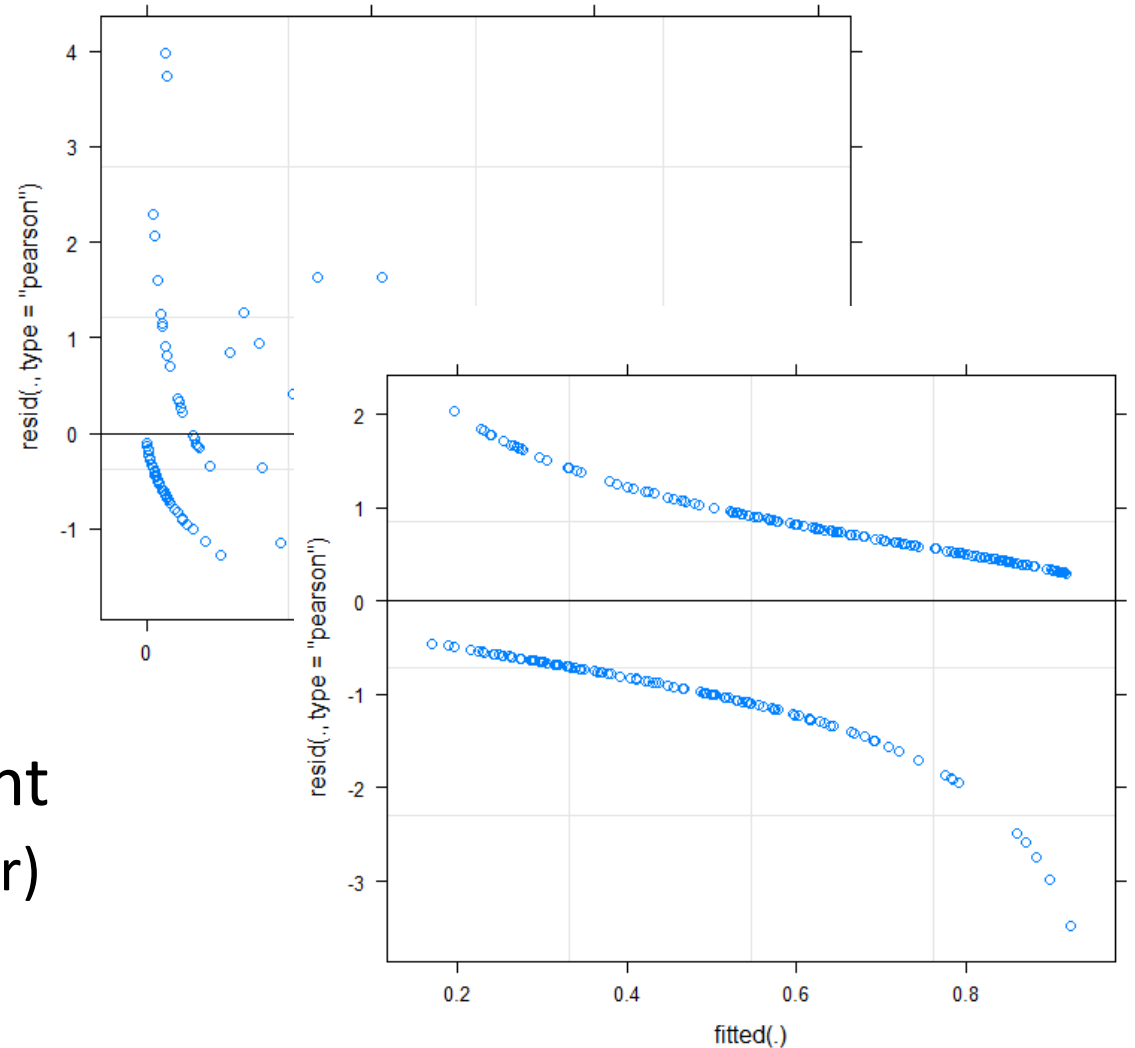
# Solution: Use the DHARMa package that is based on simulated residuals

- package developed by Florian Hartig
- user-friendly residual diagnostics for a large variety of models

- `simulateResiduals()`
  - generates residuals via simulation from the model
- `plotResiduals()`
  - plot residuals against predictors
- several test functions e.g.
  - `testDispersion()`
  - `testZeroInflation()`
  - …
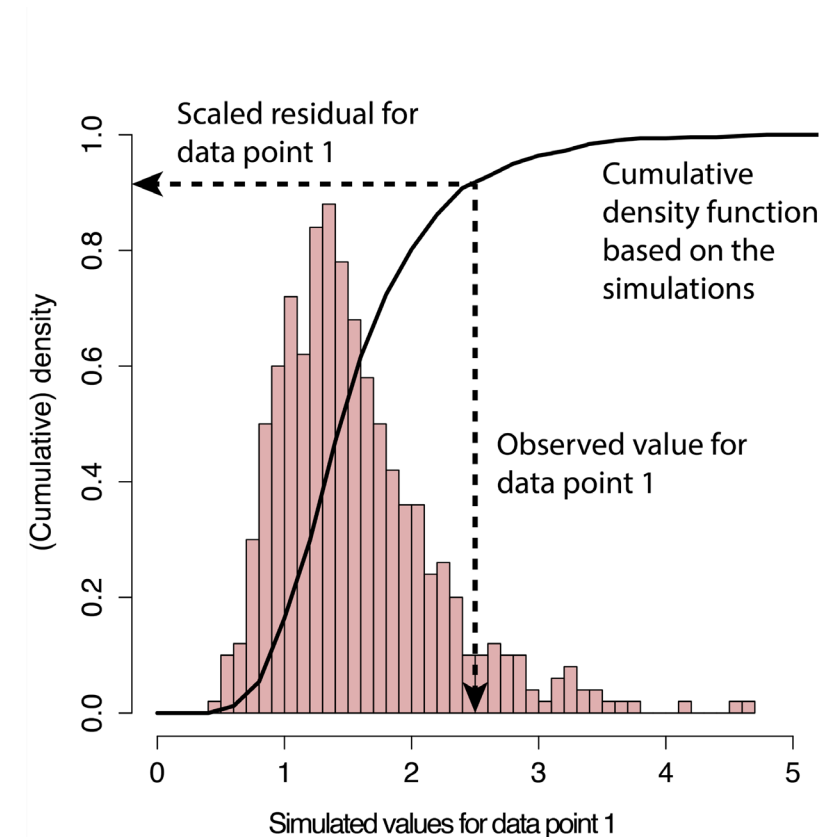
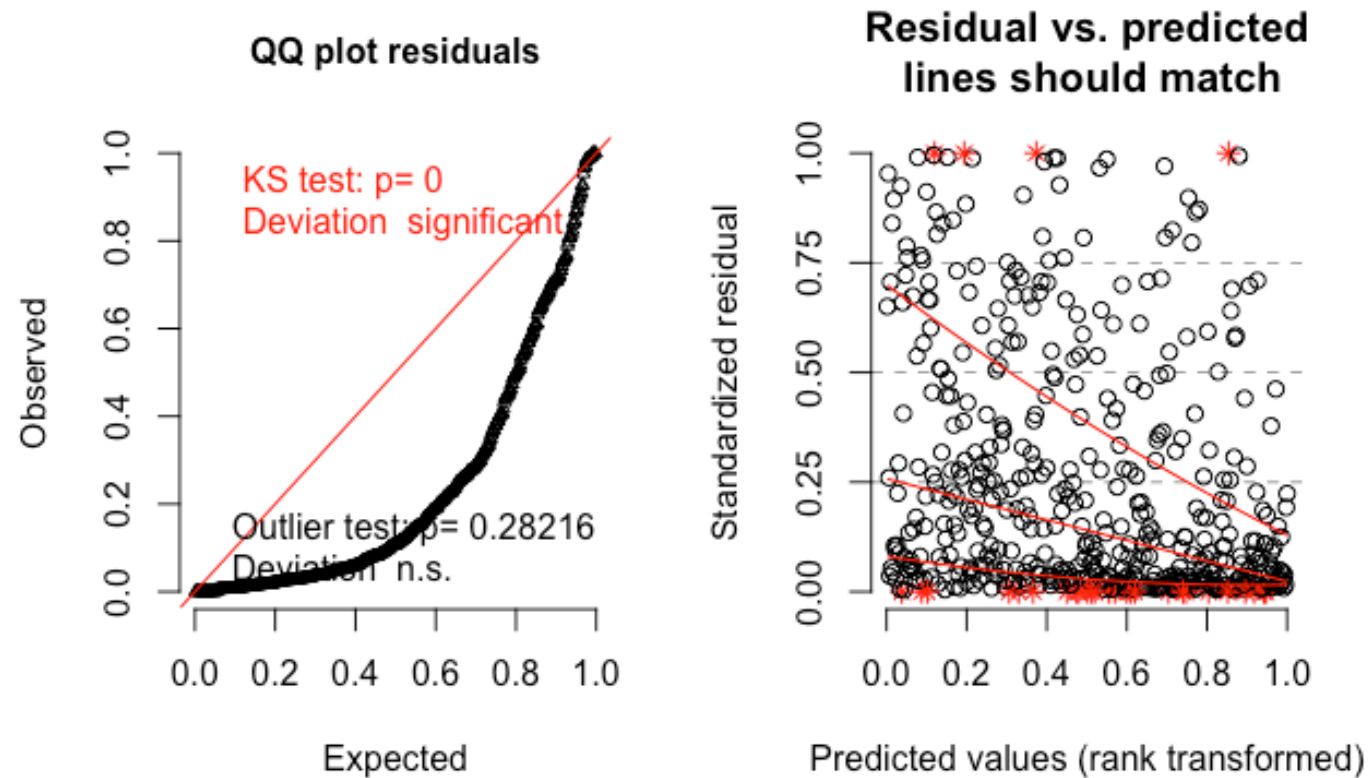# How to interpret the resulting plots

```
simulationOutput <- simulateResiduals(fittedModel = fittedModel)
plot(simulationOutput)
```
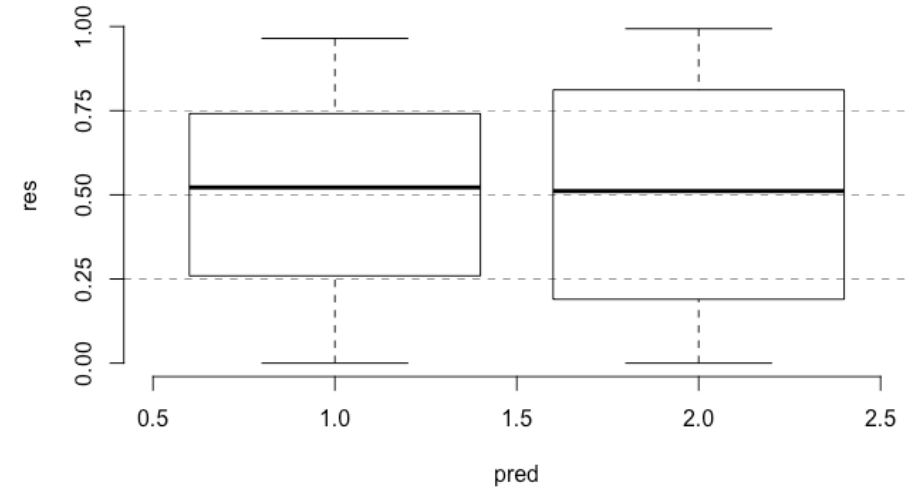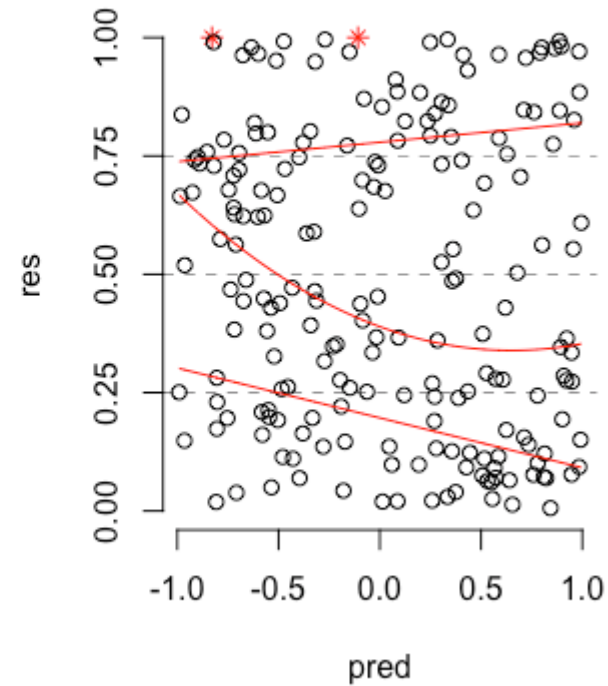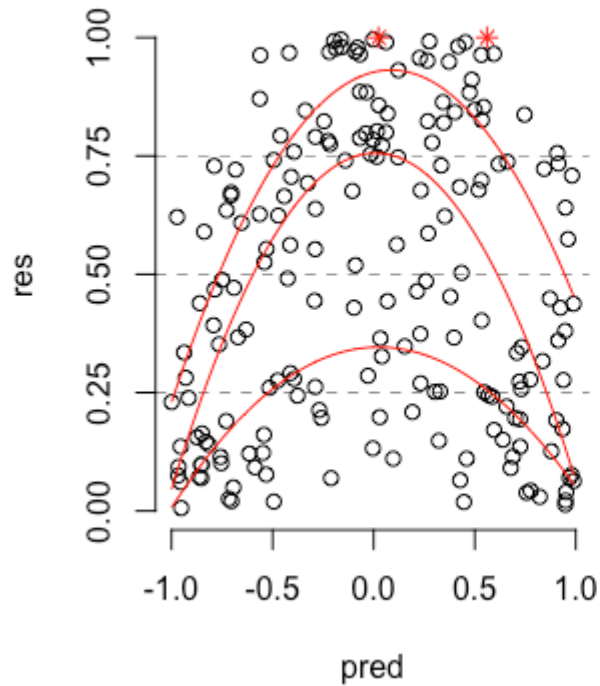

DHARMa scaled residual plots
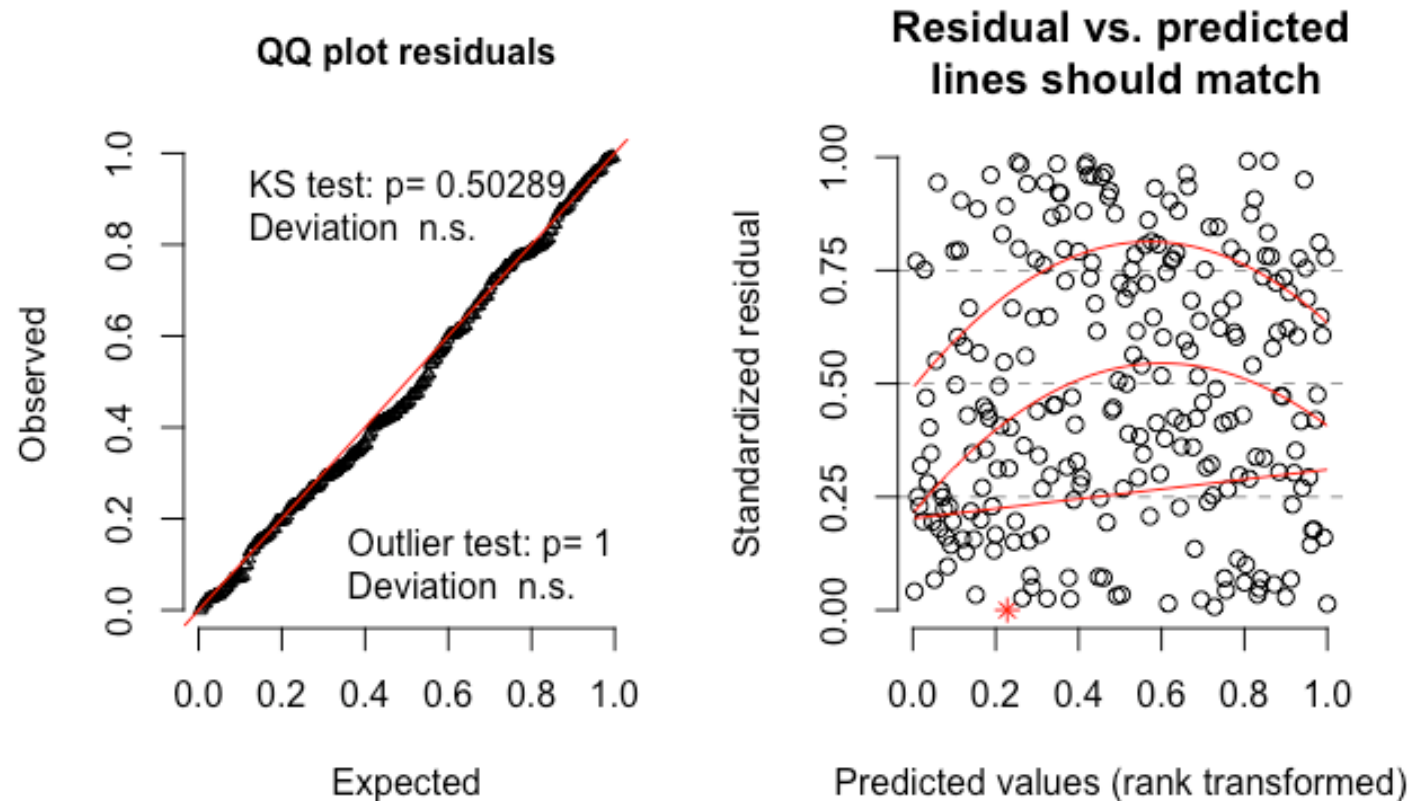
A misspecified model: here overdispersion


DHARMa scaled residual plots

Identify sources of mismatch: plot residuals against predictors

```
plotResiduals(YOURPREDICTOR, simulationOutput$scaledResiduals)
```

Several more advanced features

- calculating residuals per group

- temporal or spatial autocorrelation

- checking different levels of stochasticity in hierarchical models

- DHARMa can also be used for lm/lmer, glmmTMB, smaMM or posterior predictive simulations from Bayesian software
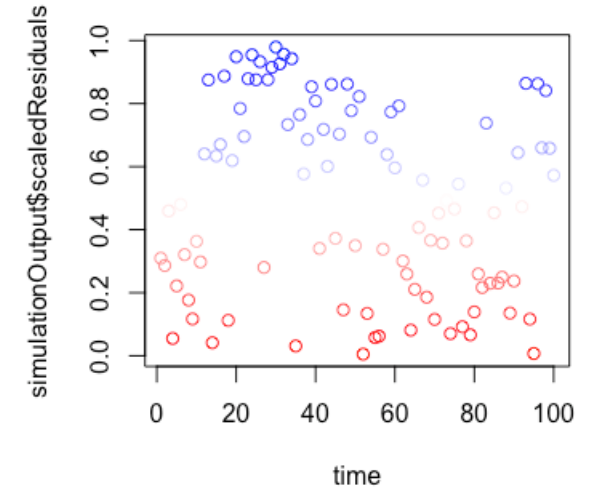
Check the vignette of DHARMa for more details.

You can also find nice examples for typical misspecifications there.

https://cran.r-project.org/web/packages/DHARMa/vignettes/DHARMa.html

Exercise

count	temperature	precipitation	plot
2	-0.13159197	-0.031600868	1
1	0.15144222	-0.958592258	1
1	-0.10013614	-0.401583887	1
0	-0.97284735	-0.977506335	1
0	0.63306416	-0.647717534	1

<u>Data</u>: abundance of myFavouriteSpecies
in response to temperature and precipitation,
with several observations made on a number of plots

<u>Task</u>:

- fit a glmm with both predictors and use plot as a random effect (family = poisson)
- use the following workflow to check the model

```
simulationOutput <- simulateResiduals(YOURMODEL)
plot(simulationOutput)
plotResiduals(YOURPREDICTOR, simulationOutput$scaledResiduals)
```

- also try out the functions `testDispersion()` and `testZeroInflation()`
- what is wrong in the model? correct it and run the tests again
- How are species abundances influenced by temperature and precipitation

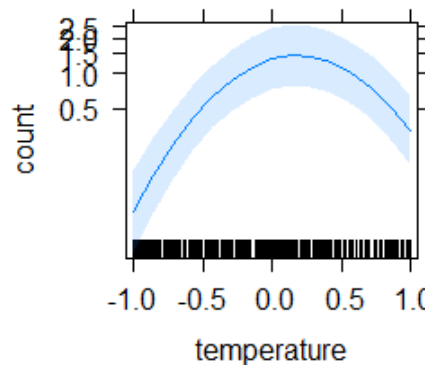# First model attempt

```
fittedModel <- glmer(count ~ temperature + precipitation + (1|plot),
                family = "poisson", data = myFavouriteSpecies)
```
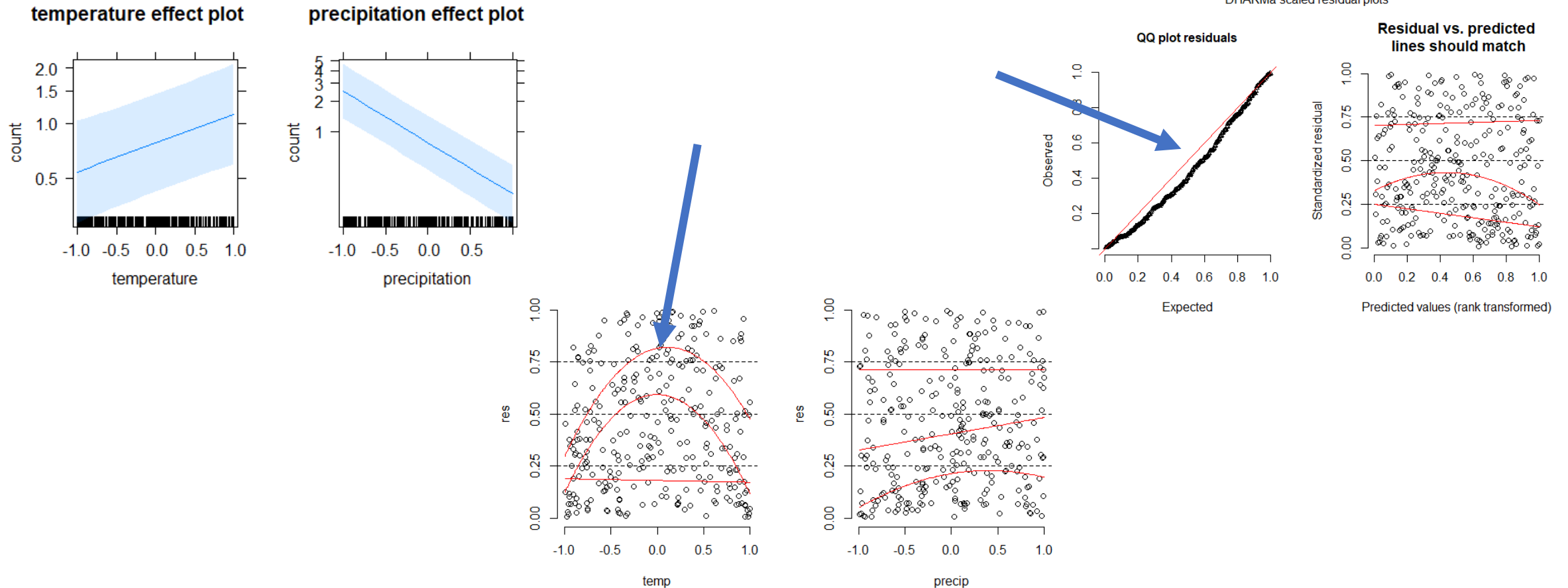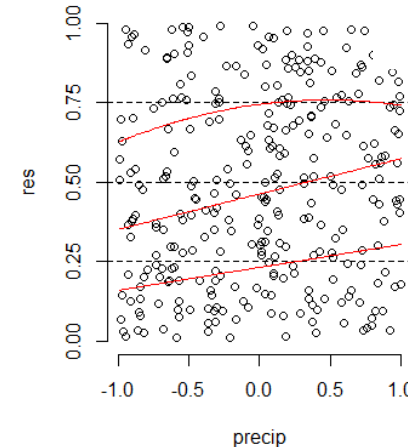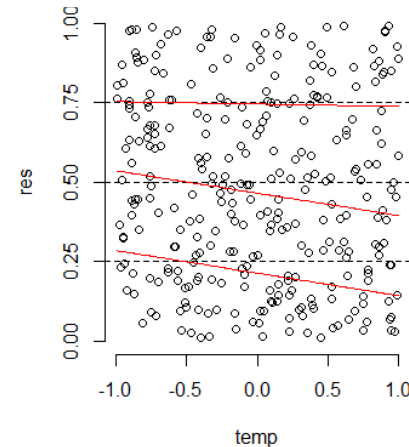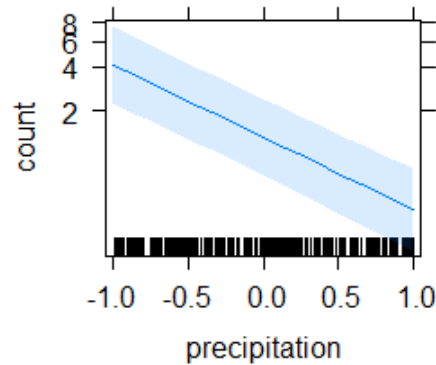
Clearly improved model

```
fittedModel <- glmer(count ~ temperature + I(temperature^2) + precipitation + (1|plot),
                family = "poisson", data = myFavouriteSpecies)
```
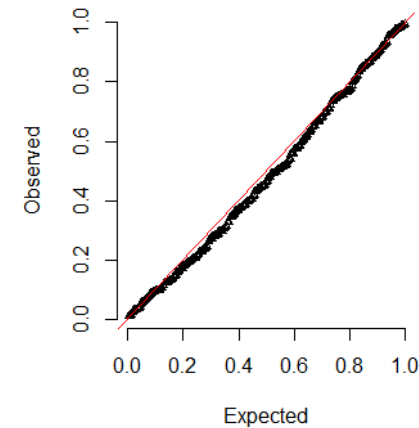

temperature effect plot

precipitation effect plot

DHARMa scaled residual plots

QQ plot residuals

Residual vs. predicted
lines should match