Appendix S1

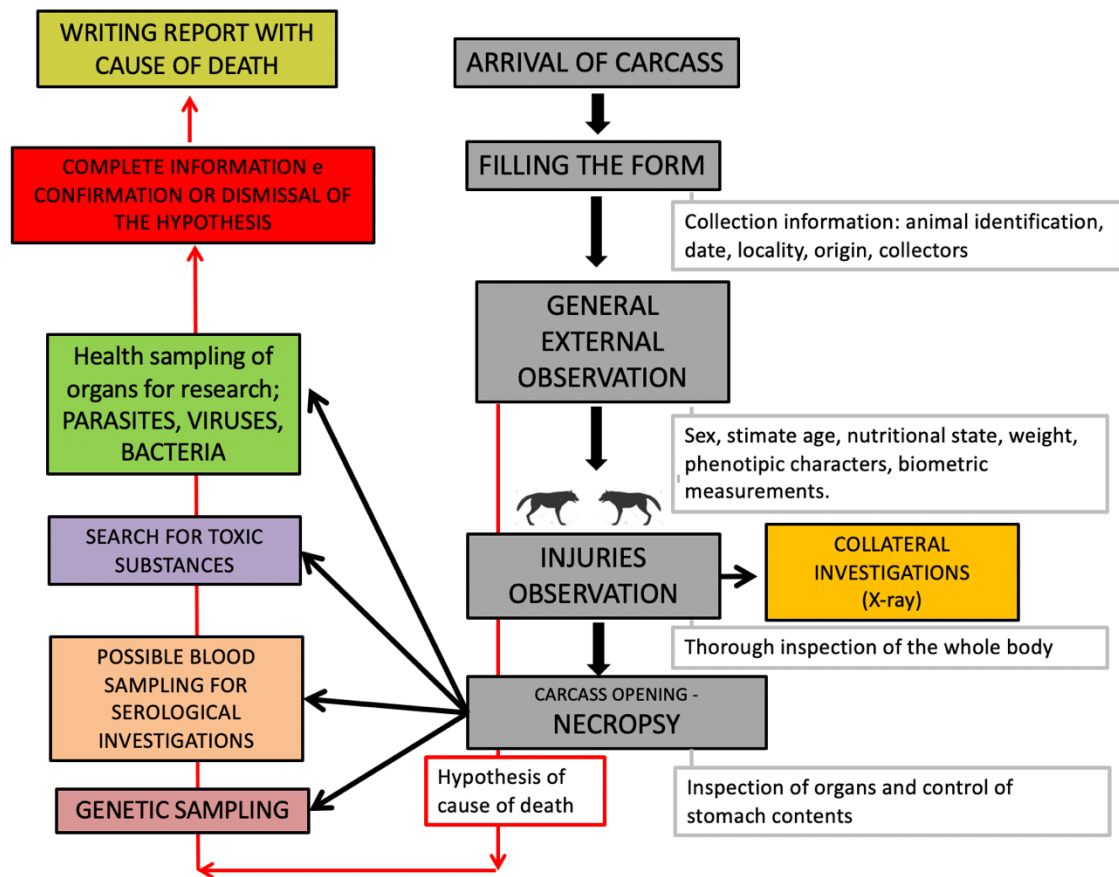**Protocol for necropsy and genetic analysis**

Fig. S1: Overview of the data collection process.

Necropsy examinations on wolf carcasses were conducted by the IZSLER, the Italian acronym of Experimental Zooprophylactic Institute of Lombardy and Emilia Romagna, situated in Parma, Piacenza, Reggio Emilia, Modena, Bologna, Forlì-Cesena, and by the Wildlife and Exotic Service of the University of Bologna.

At the arrival of each carcass, a first form containing the following information was filled (Figure S1): subject's identification data with the attribution of a unique ID code, the discovery location (reported as GPS coordinates), the sex, the weight (in kg) and the nutritional status (obtained through the direct observation and palpation of the locations where the fat accumulates, thus: flank fossa, ribs, lumbar vertebrae and pelvic bones; see below). The age of the animal was estimated on the basis of dental development, body size and weight (Morner et al., 2005). Here, all individuals were aged using 3 categories as follows: class 1: 12 months; class 2: 1–2 years; class 3: > 2 years. The aging (based on

months of life) was executed in relation to the reproductive cycle of the Italian wolf, which defines May as pups' birth month (Boitani, 1981).

The biometric information and phenotypic characteristics were also recorded: in particular, were noted: total length (from the nose to the junction of the tail), length of the tarsus, length of the tail, height of the ear, chest circumference, neck circumference and finally the presence of spurs, stripes, white nails and the coloring of the coat (Figure S1). Then, an external examination of the carcass was made by disposing the carcass in lateral decubitus: such position allows for straight evaluation of the nutritional and health status of the subject (Figure S1). During this first inspection, we established whether radiographic investigations were needed to identify foreign bodies (such as bullets, blunt objects) or fractures in the skeletal system (i.e. due to traumatic-contusive lesions) (Figure S1). At this stage, it was also possible to assess the nutrition, skin and mucosal status as well as the explorable lymph nodes and proceed with tongue sampling to be preserved in 95% ethyl alcohol for the genetic determination of the species (Figure S1). After finalizing the first physical inspection, the carcass was placed in dorsal decubitus to proceed with its flaying. The complete necropsy examination started with the opening of the abdominal cavity and was followed by the opening of the thoracic one. At each of these two steps, all organs were inspected and evaluated individually before being sampled for additional laboratory analysis (Figure S1).

All individuals were also examined for the presence of zoonotic parasitic infections such as Trichinella spp. and Leishmania infantum. Specifically, diaphragm and the tibialis muscles were analyzed to detect *Trichinella* spp. (Kapel et al., 2005), whereas spleen and popliteal lymph nodes were examined to detect *Leishmania infantum* (Gomes et al., 2007). In all individuals, stomach and liver were also analyzed to detect the presence of toxic substances (such as zinc phosphide, strychnine, organophosphate pesticides, metaldehyde and anticoagulants) (Berny, 2007), Figure S1. Moreover, further investigations on the possible presence of *Sarcoptes scabiei* were only performed in case of suspicious skin lesions, which could have been caused by sarcoptic mange. In the case of mangy subjects, staging

was established according to the methodology described in Pence and Ueckermann (2002). Each step of the necropsy process was photographed.

We collected and stored in 95% ethanol a fragment of lingual tissue in 141 carcasses to carry out genetic analyses to confirm the belonging of samples to the Italian subspecies and detect possible traces of dog ancestry.

DNA extraction was performed using the genomic DNeasy Blood Tissue Kit (Qiagen Inc., Hilden, Germany), following the manufacturer's instructions, in a QIAcube (Qiagen Inc., Hilden, Germany) extractor. Subsequently, each DNA sample was amplified by Polymerase Chain Reaction (PCR), sequenced at 500 bp of the mtDNA control-region (diagnostic for the Italian wolf population) (Montana et al., 2017), and genotyped through a multiple-tube approach at 39 unlinked autosomal microsatellites (STR, which differentiate between wolves, dogs and their first two generation hybrids) as listed in Randi et al. (2014). A marker on the Amelogenin gene was used to identify the sex of the samples. Additionally, four Y-chromosome microsatellites were used to identify paternal haplotypes (Randi et al., 2014), whereas a dominant 3-bp deletion (named KB or CBD103DG23) at the b-defensin CBD103 gene, correlated to black coat and likely of dog origin (the K-locus) (Anderson et al., 2009; Galaverni et al., 2017) was also genotyped. Any sign of past hybridization with domestic dog (*Canis lupus familiaris*) was investigated using a Bayesian genetic clustering procedures implemented in STRUCTURE 2.3.4 (Falush et al., 2003) and described in Caniglia et al. (2020). In particular, we categorized an individual as wolf, hybrid or introgressed based on their membership proportions to the wolf cluster (qwolf). We further considered as introgressed individuals presenting a mtDNA or Y-haplotype of canine origin, or the deletion at the K-locus (Randi et al., 2014) regardless of their qwolf values (Caniglia et al., 2020). Extraction, amplification and post-amplification procedures were carried out in three separate rooms at the Unit for Conservation Genetics (BIO-CGE Ozzano dell'Emilia, Bologna, Italy), which is part of the Italian Institute for Environmental Protection and Research (ISPRA).

Not all wolves covered by this study were genetically tested, in fact 33.5% were not sent for species attribution. This information gap is given by the lack of national coordination

and a solid network that guarantees the genetic determination of species for each canid examined. So, a total of 137 out of the 141 analysed samples was successfully genotyped at 39 STR, Amelogenin and K loci, four STR Y-Linked and mtDNA control-region. Among them, 105 (74.5%) resulted pure wolves, showing a qwolf >= 0.995 and no other genetic signals of dog origin; 13 (9.2%) resulted wolf x dog hybrids, showing a qwolf <= 0.954, and 23 (16.3%) were classified as introgressed individuals, showing a lower genetic dog derivate component represented by qwolf < 0.995 but > 0.954 or by the presence of a dog Y-haplotype (N=4) and/or the Kb allele (N=3). No dog mtDNA haplotype was identified, since all the analysed animals shared the private Italian wolf mtDNA haplotype (named W14 by Randi et al. 2000 and by Montana et al. 2017).

The molecular gender, genetically identified by analysing the Amelogenine marker, always con- firmed the autopsy visual examination.

The causes of mortality were grouped into three main categories: anthropic, natural and unknown. A schematic representation of each category (numbers included and contributing causes) is shown in Figure 2, details on each group are shortly defined as: *i*) Natural: it included health-related (i.e., presence of disease and/or starvation to death) and natural competition (interspecific and intraspecific) as causes of mortality; *ii*) Anthropic: it included vehicle collisions caused by cars and trains, hereafter called ACC, and illegal killing (linked to poisoning, fire weapons and other minor) – hereafter called IK. In case no precise information relative to the cause of death was identified during the necropsies, the subjects examined were assigned to the category "unknown" (Brownlie and Munro, 2016). Moreover, additional processes that were not directly causes of death, but could still trigger the physiological imbalance, were noted as "contributing causes of death".

Fig. S2: Causes of death of recovered wolves, in three periods between 2005 and 2021. Data collection in 2005 started in October and 2021 data refer to January and February only.
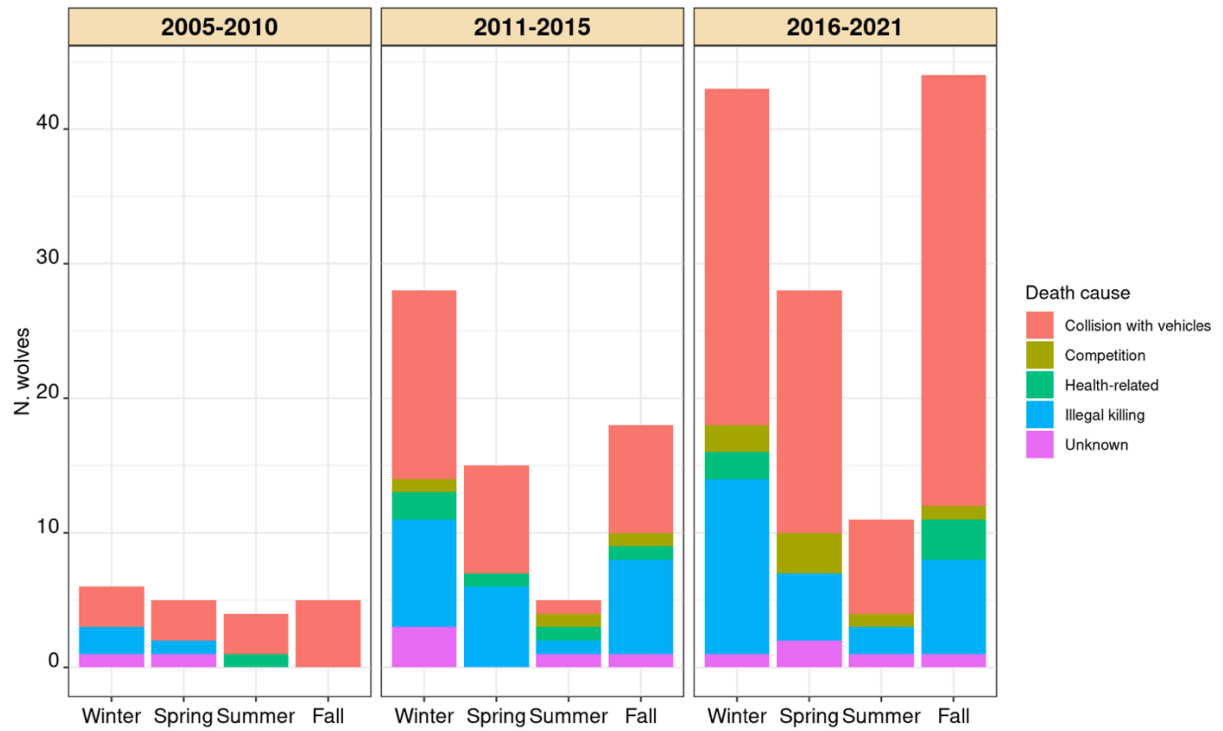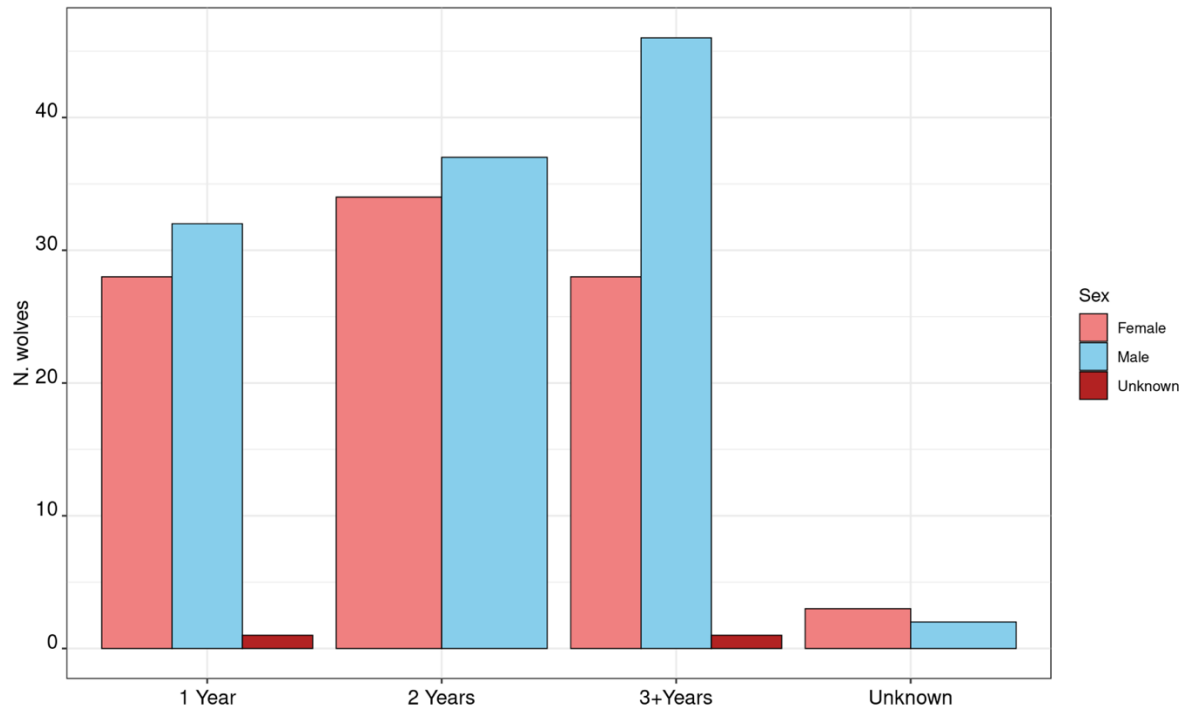
Fig. S3: Distribution of recovered wolves between the two sexed and three age classes (1 year, 2 years, 3 or more years).

References

Anderson, T.M., vonHoldt, B.M., Candille, S.I., Musiani, M., Greco, C., Stahler, D.R., Smith, D.W., Padhukasahasram, B., Randi, E., Leonard, J.A., Bustamante, C.D., Ostrander, E.A., Tang, H., Wayne, R.K., Barsh, G.S., 2009. Molecular and evolutionary history of melanism in North American gray wolves. Science 323, 1339–1343. https://doi.org/10.1126/science.1165448

Berny, P., 2007. Pesticides and the intoxication of wild animals. J Vet Pharmacol Ther 30, 93–100. https://doi.org/10.1111/j.1365-2885.2007.00836.x

Boitani, L., 1981. *Lupo Canis lupus*. In: Pavan, M. (Eds.), Distribuzione e biologia di 22 specie di Mammiferi in Italia (pp. 61–67). Rome. https://www.ibs.it/distribuzione-biologia-di-22-specie-libri-vintage-vari/e/2560038844628

Brooks Brownlie, H.W., Munro, R., 2016. The Veterinary Forensic Necropsy: A Review of Procedures and Protocols. Vet. Pathol. 53, 919–928. https://doi.org/10.1177/0300985816655851

Caniglia, R., Galaverni, M., Velli, E., Mattucci, F., Canu, A., Apollonio, M., Mucci, N., Scandura, M. Fabbri, E., 2020. A standardized approach to empirically define reliable assignment thresholds and appropriate management categories in deeply introgressed populations. Sci. Rep. 10, 2862. https://doi.org/10.1038/s41598-020-59521-2

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567-87. doi: 10.1111/j.1471-8286.2007.01758.x

Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., Randi, E., 2017. Disentangling timing of admixture, pattern of introgression, and phenotypic indicators in a hybridizing wolf population. Mol. Biol. Evol. 34, 2324-2339. https://doi.org/10.1093/molbev/msx169

Gomes, A.H.S., Ferreira, I.M.R., Lima, M.L.S.R., Cunha, E.A., Garcia, A.S., Araújo, M.F.L., Pereira-Chioccola, V.L., 2007. PCR identification of Leishmania in diagnosis and control of canine leishmaniasis. Vet. Parasitol. 144, 234-241. https://doi.org/10.1016/j.vetpar.2006.10.008

Kapel, C. M. O., Webster, P., Gamble, H. R., 2005. Muscle distribution of sylvatic and domestic Trichinella larvae in production animals and wildlife. Vet. Parasitol. 132, 101-105. https://doi.org/10.1016/j.vetpar.2005.05.036

Montana, L., Caniglia, R., Galaverni, M., Fabbri, E., Ahmed, A., Bolfı´kova´, B.Č., Czarnomska, S.D., Galov, A., Godinho, R., Hindrikson, M., Hulva, P., Jędrzejewska, B., Jelenčič, M., Kutal, M., Saarma, U., Skrbinsˇek, T., Randi, E., 2017. Combining phylogenetic and demographic inferences to assess the origin of the genetic diversity in an isolated wolf population. PloS one 12, e0176560. https://doi.org/10.1371/journal.pone.0176560

Mörner, T., Eriksson, H., Brojer, C., Nilsson, K., Uhlhorn, H., Agren, E., Segerstad, C.H., Jansson, D.S., Gavier-Widen, D., 2005. Diseases and mortality in free-ranging brown

bear (*Ursus arctos*), gray wolf (*Canis lupus*), and wolverine (*Gulo gulo*) in Sweden. J. Wildl. DisJ. Wildl. Dis. 41, 298–303. doi:10.7589/0090- 3558-41.2.298.

Pence, D.B., Ueckermann, E., 2002. Sarcoptic manage in wildlife. Rev Sci Tech Title 21, 385–398. https://europepmc.org/article/med/11974622

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., Bigi, D., Bolfíková, B. C., Smetanová, M., Caniglia, R., 2014. Multilocus detection of wolf x dog hybridization and guidelines for marker selection. PLoS one 9:e86409. https://doi.org/10.1371/journal.pone.0086409

Randi, E., Lucchini, V., Christensen, M. F., Mucci, N., Funk, S. M., Dolf, G., Loeschcke, V., 2000. Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization. Conserv. Biol. 14, 464-473. https://doi.org/10.1046/j.1523-1739.2000.98280.x

Appendix 2 - Model selection and diagnostics

Introduction

The following documents aim to provide a quick overview about model selection and diagnostics. A more comprehensive overview about statistical analyses can be found in the reproducible script.

Emilia-Romagna region

To start with, let's start with the Emilia-Romagna data, whose datasets contains the following variables of interest:

- "illegal.ever" - a dichotomous variable indicating if any illegally killed wolf was found at a certain municipality, between 2005 and 2021. Municipalities with more illegally killed wolves are nevertheless coded in this binary way, because those with 2 or 3 wolves were few.
- "predations" - the total number of domestic animals (cattle, sheep, goats) that had been killed by wolves in each municipality.
- "log.pred.events" - a natural logarithm of the total number of predations by wolves on domestic animals (number of events). The variable was log-converted to smooth out its extremes.
- "farm.density" - the total number of farms in each municipality, divided per its total surface in hectares.
- "human density" - the total number of residents in each municipality, divided per its total surface in hectares.
- "wolf.presence" - a dummy variable indicating whether wolves had been recorded on a certain municipality on some consecutive monitoring initiatives, between 2006 and 2016 (see Apollonio et al., 2016, reference n. 45).
- "marginal.area" - a dummy variable indicating if a certain municipality was considered to be marginal for agriculture. Based on the National Agricultural Network (https://www.reterurale.it/flex/cm/pages /ServeBLOB.php/L/IT/IDPagina/1)
- "wolf.found offset" - a variable indicating the number of wolves that were found, between 2005 and 2021 on a certain municipality. Used as an offset.

All continuous variables were standardized an centered, before their inclusion as predictors.

We compared the following models, with a backwise selection approach based on leave-one-out cross validation, and supplemented by some other measures of fitness. Such as the WAIC, area under the curve and classification accuracy:

Let's calculate the VIF on an analogous frequentist logistic regression, before fitting the full model (without BYM structure)

```
vif(glm(illegal.ever ~ log.pred.events + farm.density.std + human.density.std + wolf.presence +
        marginal.area + wolf.found.std,
        family=binomial(link="logit"), data=d.er))
```

```
##   log.pred.events  farm.density.std human.density.std     wolf.presence
##          1.449720          1.123598          1.424633          1.714927
##     marginal.area     wolf.found.std
##          1.916923          1.508884
```

VIF looks good and there is no sign of strong collinearity between predictors:

```r
#mod.a.er <- brm(illegal.ever ~ predations.std + farm.density.std + human.density.std +
#              wolf.presence + marginal.area,
#              family=bernoulli(link = "logit"), data=d.er,
#              chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#              control = list(adapt_delta = 0.99, max_treedepth = 10),
#              prior=c(prior(normal(0, 1), class=b, coef = predations.std),
#                      prior(normal(0, 1), class=b, coef = farm.density.std),
#                      prior(normal(0, 1), class=b, coef = human.density.std),
#                      prior(normal(0, 1), class=b, coef = wolf.presence1),
#                      prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.b.er <- brm(illegal.ever ~ predations.std + farm.density.std + human.density.std +
#              wolf.presence + marginal.area +
#              offset(log(wolf.found.offset)), family=bernoulli(link = "logit"),
#              data=d.er,
#              chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#              control = list(adapt_delta = 0.99, max_treedepth = 10),
#              prior=c(prior(normal(0, 1), class=b, coef = predations.std),
#                      prior(normal(0, 1), class=b, coef = farm.density.std),
#                      prior(normal(0, 1), class=b, coef = human.density.std),
#                      prior(normal(0, 1), class=b, coef = wolf.presence1),
#                      prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.c.er <- brm(illegal.ever ~ log.pred.events.std + farm.density.std +
#              human.density.std + wolf.presence + marginal.area +
#              offset(log(wolf.found.offset)), family=bernoulli(link = "logit"),
#              data=d.er,
#              chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#              control = list(adapt_delta = 0.99, max_treedepth = 10),
#              prior=c(prior(normal(0, 1), class=b, coef = log.pred.events.std),
#                      prior(normal(0, 1), class=b, coef = farm.density.std),
#                      prior(normal(0, 1), class=b, coef = human.density.std),
#                      prior(normal(0, 1), class=b, coef = wolf.presence1),
#                      prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.d.er <- brm(illegal.ever ~ log.pred.events.std + farm.density.std +
#              human.density.std + wolf.presence + marginal.area +
#              offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#              family=bernoulli(link = "logit"),
#              data=d.er, data2 = list(W = m.er),
#              chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#              control = list(adapt_delta = 0.99, max_treedepth = 10),
#            prior=c(prior(normal(0, 1), class=b, coef = log.pred.events.std),
#                      prior(normal(0, 1), class=b, coef = farm.density.std),
#                      prior(normal(0, 1), class=b, coef = human.density.std),
#                      prior(normal(0, 1), class=b, coef = wolf.presence1),
#                      prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.e.er <- brm(illegal.ever ~ log.pred.events.std + human.density.std + wolf.presence +
#              offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#              family=bernoulli(link = "logit"),
#              data=d.er, data2 = list(W = m.er),
#              chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
```

```
#             control = list(adapt_delta = 0.99, max_treedepth = 12),
#             prior=c(prior(normal(0, 1), class=b, coef = log.pred.events.std),
#                   prior(normal(0, 1), class=b, coef = human.density.std),
#                   prior(normal(0, 1), class=b, coef = wolf.presence1)))
#
#mod.f.er <- brm(illegal.ever ~ log.pred.events.std + wolf.presence +
#             offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#             family=bernoulli(link = "logit"),
#             data=d.er, data2 = list(W = m.er),
#             chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#             control = list(adapt_delta = 0.99, max_treedepth = 12),
#             prior=c(prior(normal(0, 1), class=b, coef = log.pred.events.std),
#                   prior(normal(0, 1), class=b, coef = wolf.presence1)))

#mod.g.er <- brm(illegal.ever ~ log.pred.events.std +
#             offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#             family=bernoulli(link = "logit"),
#             data=d.er, data2 = list(W = m.er),
#             chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#             control = list(adapt_delta = 0.99, max_treedepth = 12),
#             prior=c(prior(normal(0, 1), class=b, coef = log.pred.events.std)))

#mod.h.er <- brm(illegal.ever ~ offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#             family=bernoulli(link = "logit"),
#             data=d.er, data2 = list(W = m.er),
#             chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#             control = list(adapt_delta = 0.99, max_treedepth = 12))

#mod.i.er <- brm(illegal.ever ~ 1 + car(W, type = "bym2"),
#             family=bernoulli(link = "logit"),
#             data=d.er, data2 = list(W = m.er),
#             chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#             control = list(adapt_delta = 0.99, max_treedepth = 12))
```

From Table 1 you can see that a model with the logarithm of predations as a predictor, the total number of wolves found as an offset and a Besag-York-Mollié structure seems to be the best choice. In the BYM structure, we used "Queen neighbours": two municipalities were deemed to be neigbours if they shared at least one point in common, on their perimeter. Although differences are not so pronounces between models. Let's see if this model converged:

```
mod <- mod.g.er
summary(mod)
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: illegal.ever ~ log.pred.events.std + offset(log(wolf.found.offset)) + car(W, type = "bym2")
##     Data: d.er (Number of observations: 328)
## Samples: 4 chains, each with iter = 5000; warmup = 1000; thin = 10;
##          total post-warmup samples = 1600
##
## Correlation Structures:
##        Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sdcar      0.92      0.59     0.05     2.28 1.00      699     1143
##
```

```
## Population-Level Effects:
##                    Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept             -2.88      0.41    -3.86    -2.28 1.00      991     1273
## log.pred.events.std    0.59      0.19     0.24     0.99 1.00     1411     1414
##
## Samples were drawn 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(mod)
```
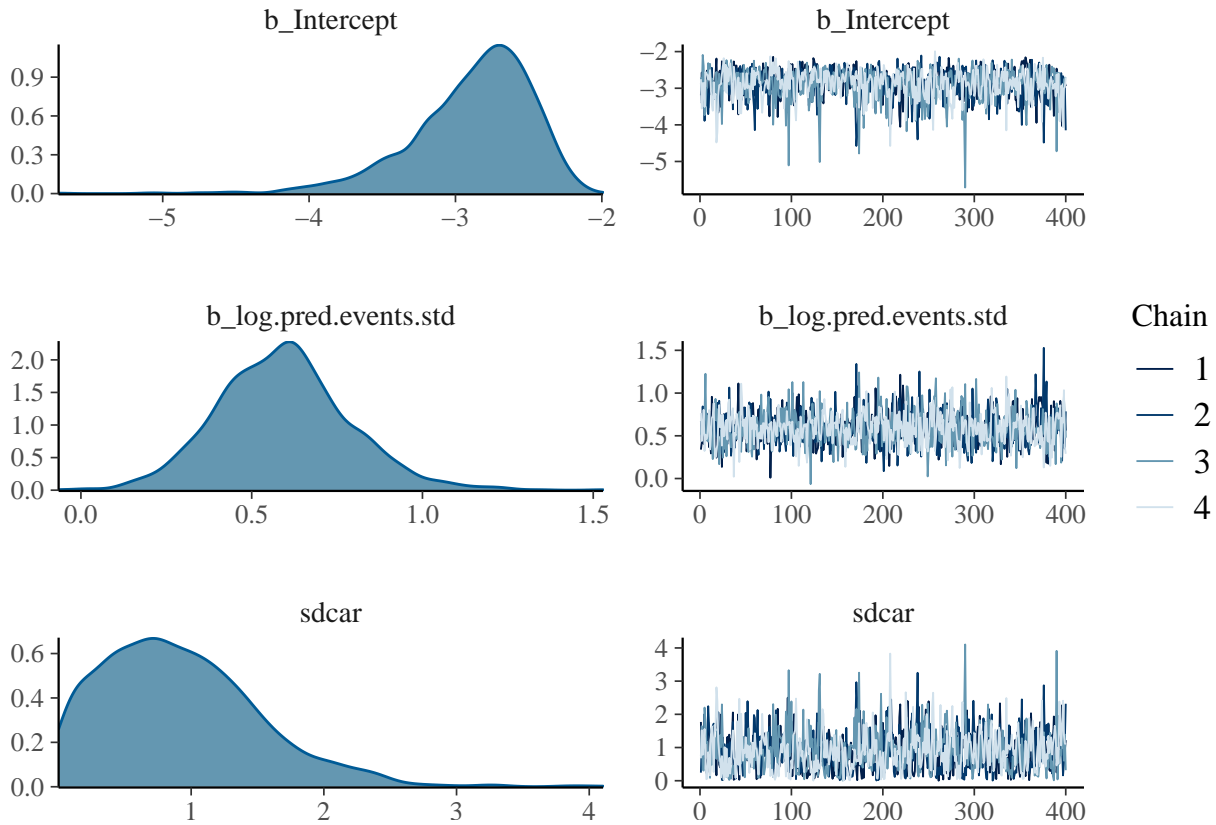

Yest, MCMC have mixed well and model parameters seem to have a nice posterior distribution. But will the model fit the data? Let's compare the observed response variable with simulated datasets from the posterior predictive distribution:

```
brms::pp_check(mod)
```

```
## Using 10 posterior samples for ppc type 'dens_overlay' by default.
```

```
## Warning: Using CAR terms without a grouping factor is deprecated. Please use
## argument 'gr' even if each observation represents its own location.
```
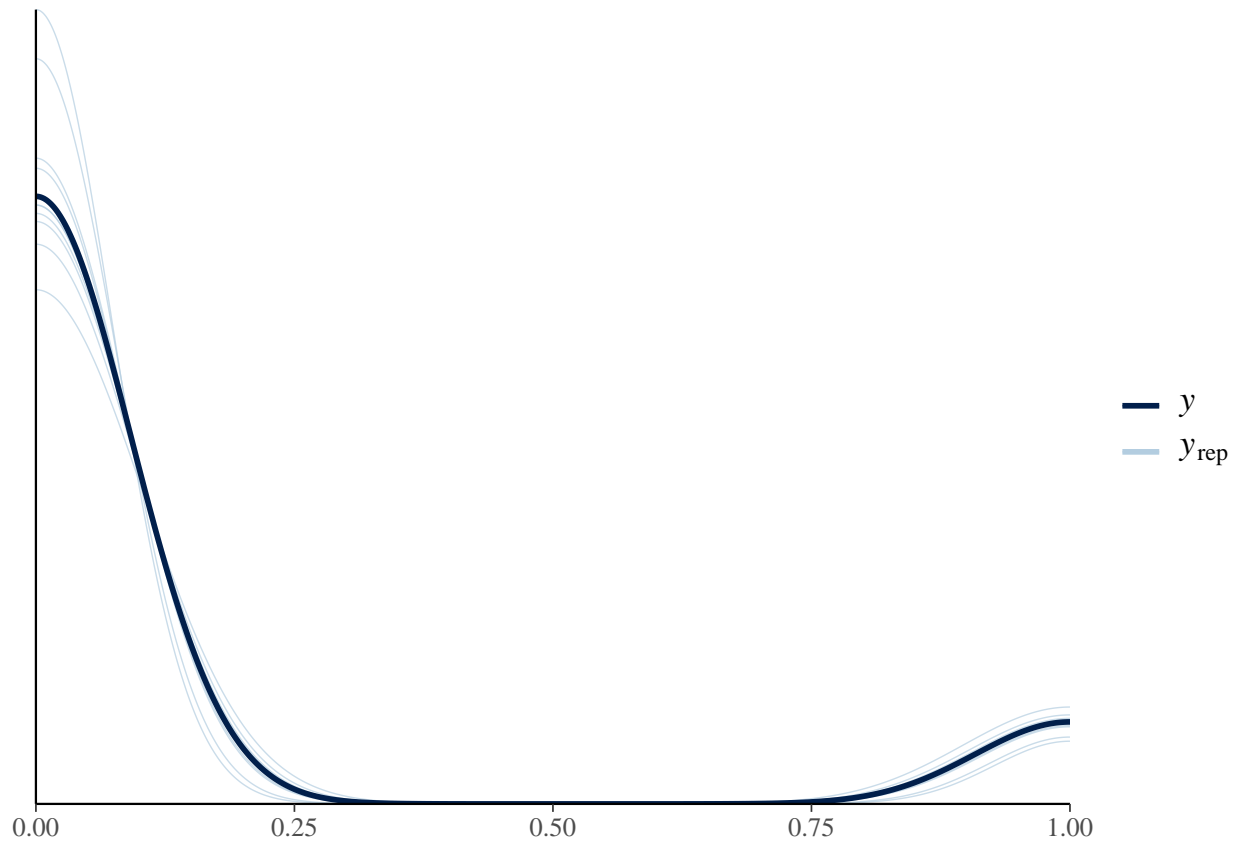
Now, let try to make model binned residuals plot. In this plot, data are partitioned into categories based on their fitted values and then average residuals are plotted against the average fitted value. In case of a good fit, one would expect 95% of residuals to fall within the error bounds:

```
binned_residuals(mod)
```

```
## Warning: Using CAR terms without a grouping factor is deprecated. Please use
## argument 'gr' even if each observation represents its own location.
```

```
## Warning: Following potential variables could not be found in the data: W
```

```
## Warning: Probably bad model fit. Only about 64% of the residuals are inside the error bounds.
```
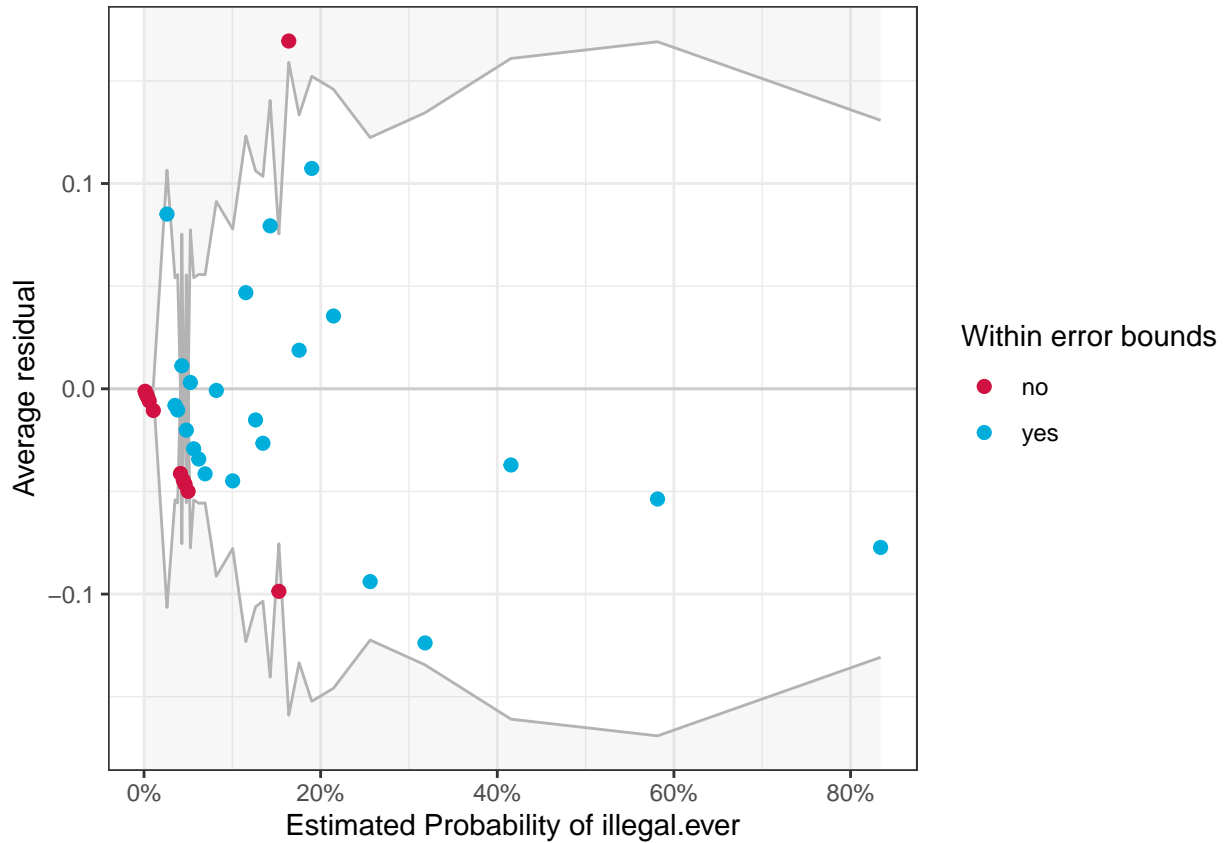
About 45% of model residuals are outside of error bounds: the model does not fit really well the data, as confirmed by our progressive elimination of the various covariates. Let' check also the role of the logarithm of predations:

```
binned_residuals(mod, term=c("log.pred.events.std"))
```

```
## Warning: Using CAR terms without a grouping factor is deprecated. Please use
## argument 'gr' even if each observation represents its own location.
```

```
## Warning: Following potential variables could not be found in the data: W
```

```
## Ok: About 100% of the residuals are inside the error bounds.
```

```
## `geom_smooth()` using formula 'y ~ x'
```
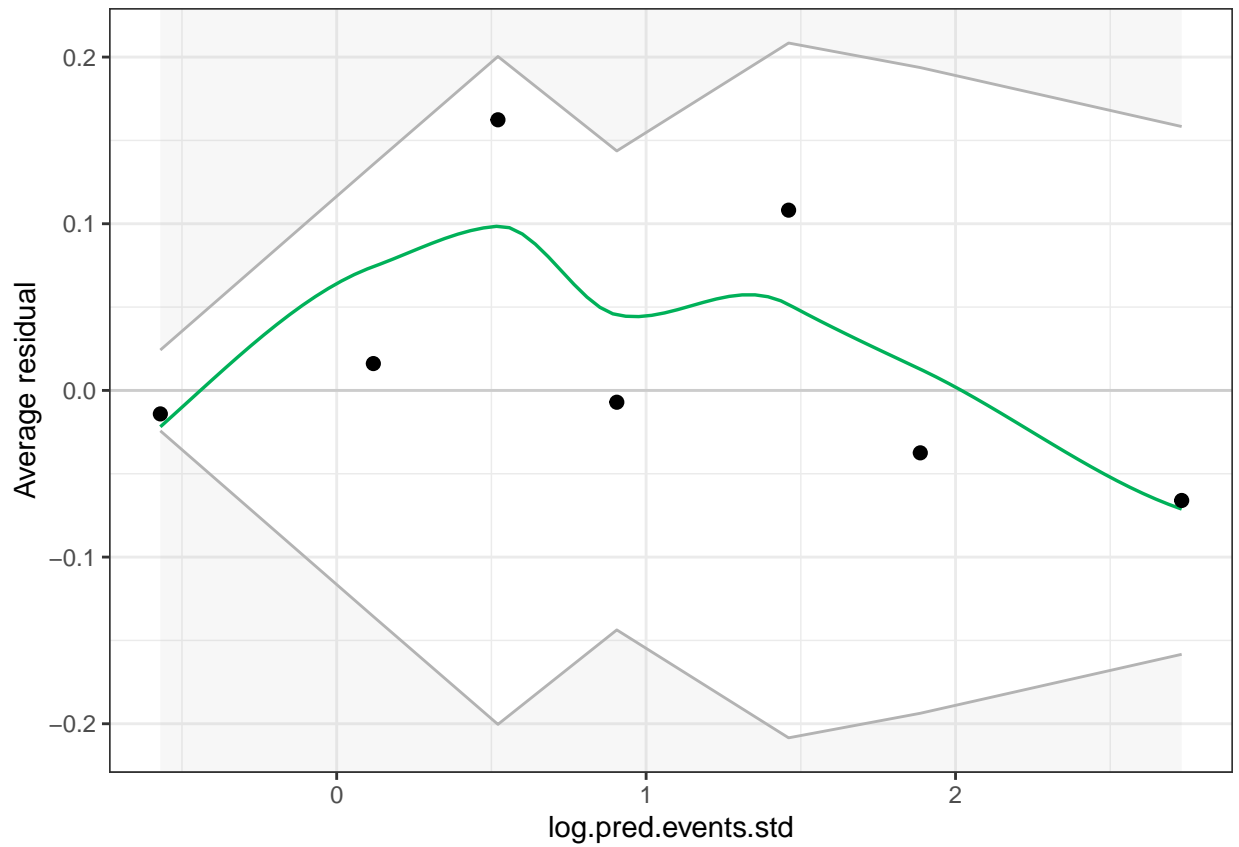
Indeed, all residuals fall into error bounds. This indicate that the association with the covariate was modelled relatively well. Now let's plot the marginal effect of the covariate:

```
conditional_effects(mod)
```
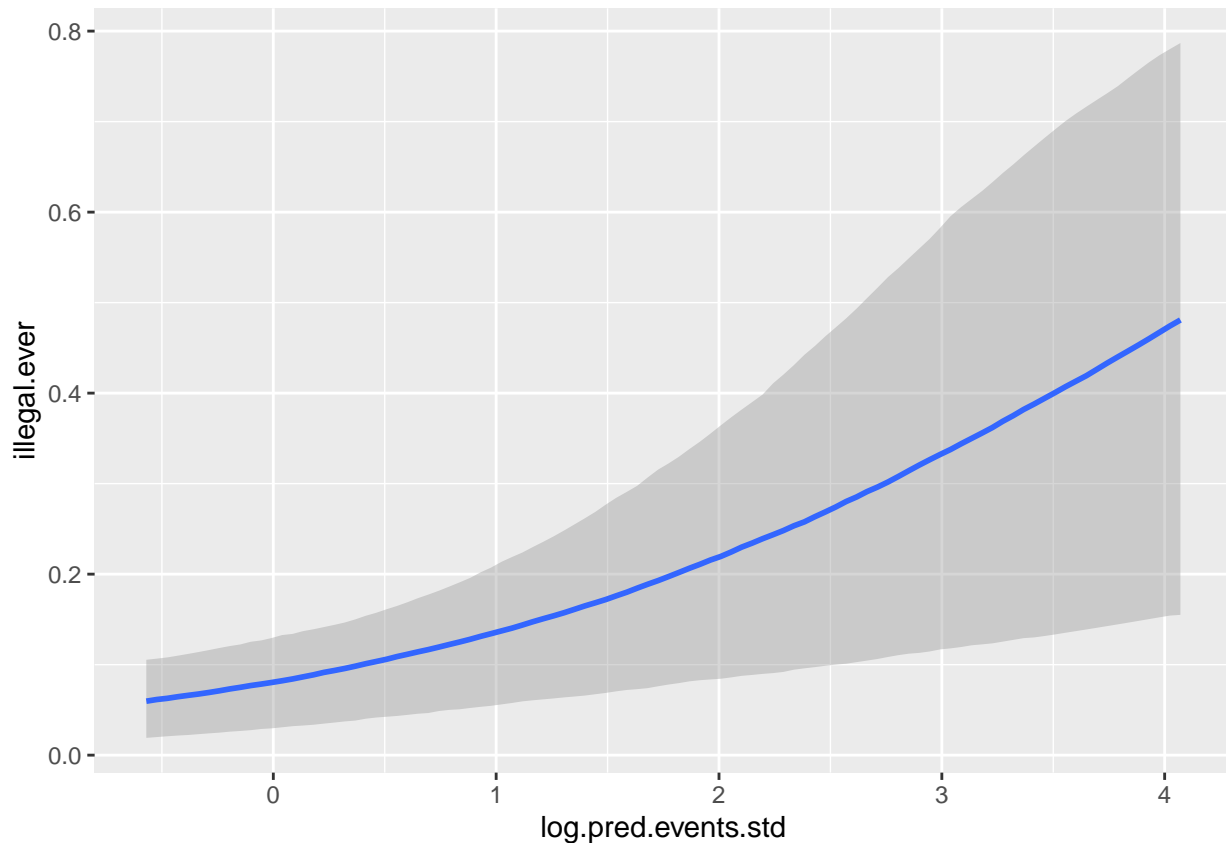
From what we can see, the marginal effect is positive but relatively low. The probability of having recorded at least a dead wolf in a certain municipality, between 2015 and 2021 increased by about 10% for a 1-unit increase in the standardized and centered logarithm of the number of predation events in that municipality.

Entire study area

Now consider the whole dataset about the Tuscany and Emilia-Romagna regions, which contains the following variables of interest:

- "illegal.ever" - a dichotomous variable indicating if any illegally killed wolf was found at a certain municipality, between 2005 and 2021. Municipalities with more illegally killed wolves are nevertheless coded in this binary way, because those with 2 or 3 wolves were few.
- "farm.density" - the total number of farms in each municipality, divided per its total surface in hectares.
- "human density" - the total number of residents in each municipality, divided per its total surface in hectares.
- "wolf.presence" - a dummy variable indicating whether wolves had been recorded on a certain municipality on some consecutive monitoring initiatives, between 2006 and 2016 (see Apollonio et al., 2016, reference n. 45).
- "marginal.area" - a dummy variable indicating if a certain municipality was considered to be marginal for agriculture. Based on the National Agricultural Network (https://www.reterurale.it/flex/cm/pages /ServeBLOB.php/L/IT/IDPagina/1)
- "wolf.found offset" - a variable indicating the number of wolves that were found, between 2005 and 2021 on a certain municipality. Used as an offset.

Let's calculate the VIF on an analogous frequentist logistic regression, before fitting the full model (without BYM structure):

```
vif(glm(illegal.ever ~ farm.density + human.density + wolf.presence +
        marginal.area + wolf.found,
        family=binomial(link="logit"), data=d))
```

```
##  farm.density human.density wolf.presence marginal.area    wolf.found
##      1.100959      1.187683      1.336548      1.342534      1.299716
```

VIF looks good and there is no sign of strong collinearity between predictors:

```
#mod.a.tot <- brm(illegal.ever ~ farm.density.std + human.density.std + wolf.presence +
#                 marginal.area,
#                 family=bernoulli(link = "logit"), data=d,
#                 chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                 control = list(adapt_delta = 0.99, max_treedepth = 10),
#                 prior=c(prior(normal(0, 1), class=b, coef = farm.density.std),
#                         prior(normal(0, 1), class=b, coef = human.density.std),
#                         prior(normal(0, 1), class=b, coef = wolf.presence1),
#                         prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.b.tot <- brm(illegal.ever ~ farm.density.std + human.density.std + wolf.presence +
#                 marginal.area +
#                 offset(log(wolf.found.offset)), family=bernoulli(link = "logit"),
#                 data=d,
#                 chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                 control = list(adapt_delta = 0.99, max_treedepth = 10),
#                 prior=c(prior(normal(0, 1), class=b, coef = farm.density.std),
#                         prior(normal(0, 1), class=b, coef = human.density.std),
#                         prior(normal(0, 1), class=b, coef = wolf.presence1),
#                         prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.c.tot <- brm(illegal.ever ~ farm.density.std + human.density.std + wolf.presence +
#                 marginal.area +
#                 offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#                 family=bernoulli(link = "logit"),
#                 data=d, data2 = list(W = m),
#                 chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                 control = list(adapt_delta = 0.99, max_treedepth = 10),
#                 prior=c(prior(normal(0, 1), class=b, coef = farm.density.std),
#                         prior(normal(0, 1), class=b, coef = human.density.std),
#                         prior(normal(0, 1), class=b, coef = wolf.presence1),
#                         prior(normal(0, 1), class=b, coef = marginal.area1)))
#
#mod.d.tot <- brm(illegal.ever ~ human.density.std + wolf.presence +
#                 offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#                 family=bernoulli(link = "logit"),
#                 data=d, data2 = list(W = m),
#                 chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                 control = list(adapt_delta = 0.99, max_treedepth = 10),
#                 prior=c(prior(normal(0, 1), class=b, coef = human.density.std),
#                         prior(normal(0, 1), class=b, coef = wolf.presence1)))
#
#mod.e.tot <- brm(illegal.ever ~ human.density.std +
#                 offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#                 family=bernoulli(link = "logit"),
#                 data=d, data2 = list(W = m),
```

```
#                chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                control = list(adapt_delta = 0.99, max_treedepth = 10),
#                prior=c(prior(normal(0, 1), class=b, coef = human.density.std)))
#
#mod.f.tot <- brm(illegal.ever ~ 1 +
#                offset(log(wolf.found.offset)) + car(W, type = "bym2"),
#                family=bernoulli(link = "logit"),
#                data=d, data2 = list(W = m),
#                chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                control = list(adapt_delta = 0.99, max_treedepth = 10))
#
#mod.g.tot <- brm(illegal.ever ~ 1 + car(W, type = "bym2"),
#                family=bernoulli(link = "logit"),
#                data=d, data2 = list(W = m),
#                chains=4, cores=4, iter=5000, warmup=1000, thin = 10,
#                control = list(adapt_delta = 0.99, max_treedepth = 10))
```

From Table 1 we can see that a model with only the offset variable and the BYM correlations structure is probably the best one. Let's keep it as the best candidate model. But did it converge?

```
mod.tot <- mod.g.tot
summary(mod.tot)
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: illegal.ever ~ 1 + car(W, type = "bym2")
##     Data: d (Number of observations: 494)
## Samples: 4 chains, each with iter = 5000; warmup = 1000; thin = 10;
##          total post-warmup samples = 1600
##
## Correlation Structures:
##       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sdcar     4.05      1.64     1.88     8.27 1.00     1197     1461
##
## Population-Level Effects:
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -5.06      1.65    -9.17    -2.98 1.00     1246     1411
##
## Samples were drawn 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(mod.tot)
```
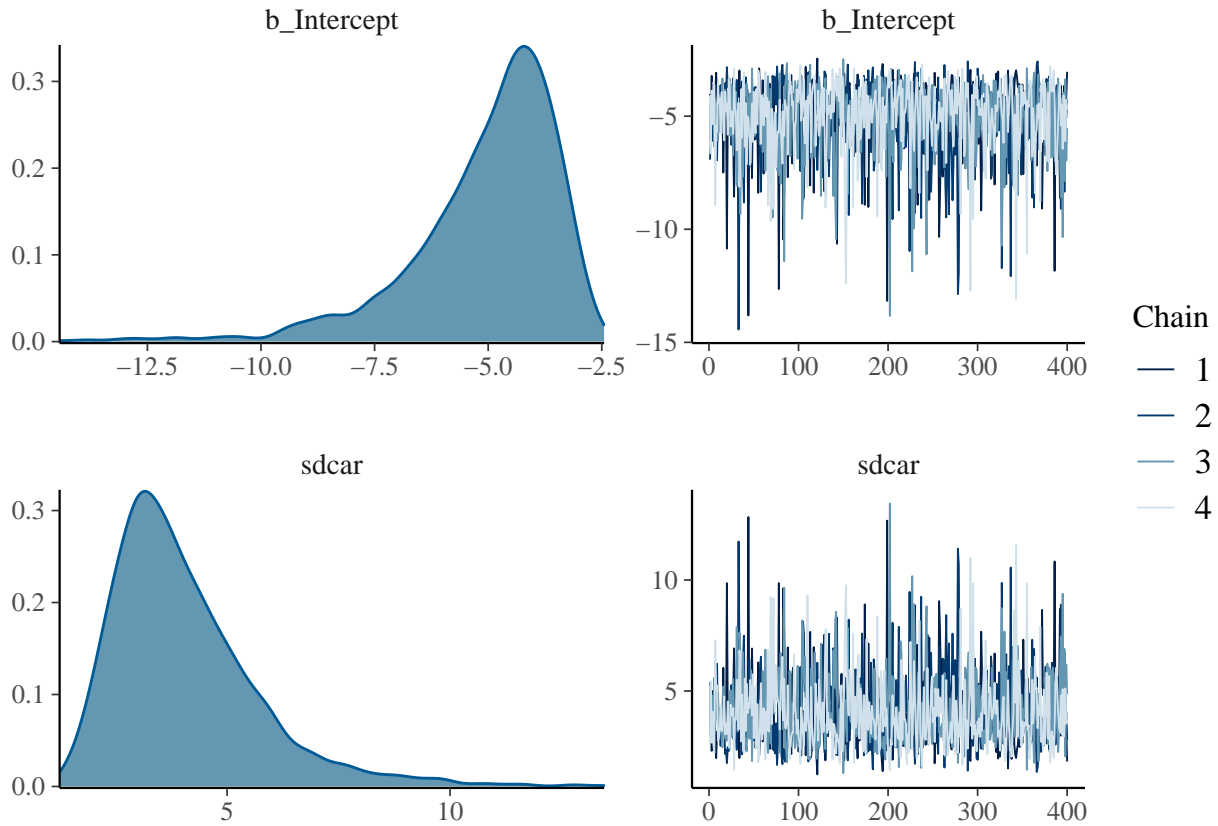
b_Intercept

b_Intercept

sdcar

sdcar

Chain

— 1

— 2

— 3

— 4

Yest, MCMC have mixed well and model parameters seem to have a nice posterior distribution. But will the model fit the data? Let's compare the observed response variable with simulated datasets from the posterior predictive distribution:

```
brms::pp_check(mod.tot)
```

```
## Using 10 posterior samples for ppc type 'dens_overlay' by default.
```

```
## Warning: Using CAR terms without a grouping factor is deprecated. Please use
## argument 'gr' even if each observation represents its own location.
```
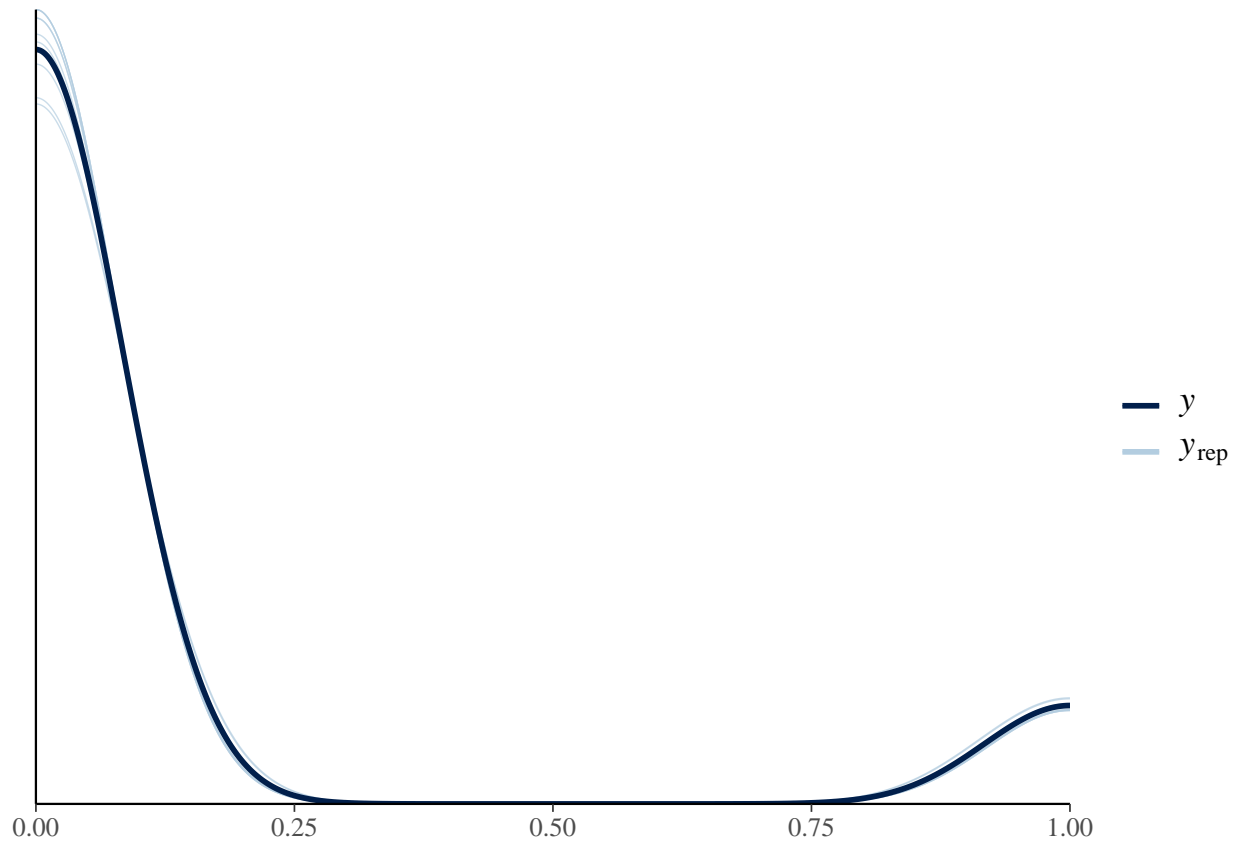
Now, let try to make model binned residuals plot. In this plot, data are partitioned into categories based on their fitted values and then average residuals are plotted against the average fitted value. In case of a good fit, one would expect 95% of residuals to fall within the error bounds:

```
binned_residuals(mod.tot)
```

```
## Warning: Using CAR terms without a grouping factor is deprecated. Please use
## argument 'gr' even if each observation represents its own location.
```

```
## Warning: Following potential variables could not be found in the data: W
```

```
## Warning: Probably bad model fit. Only about 34% of the residuals are inside the error bounds.
```
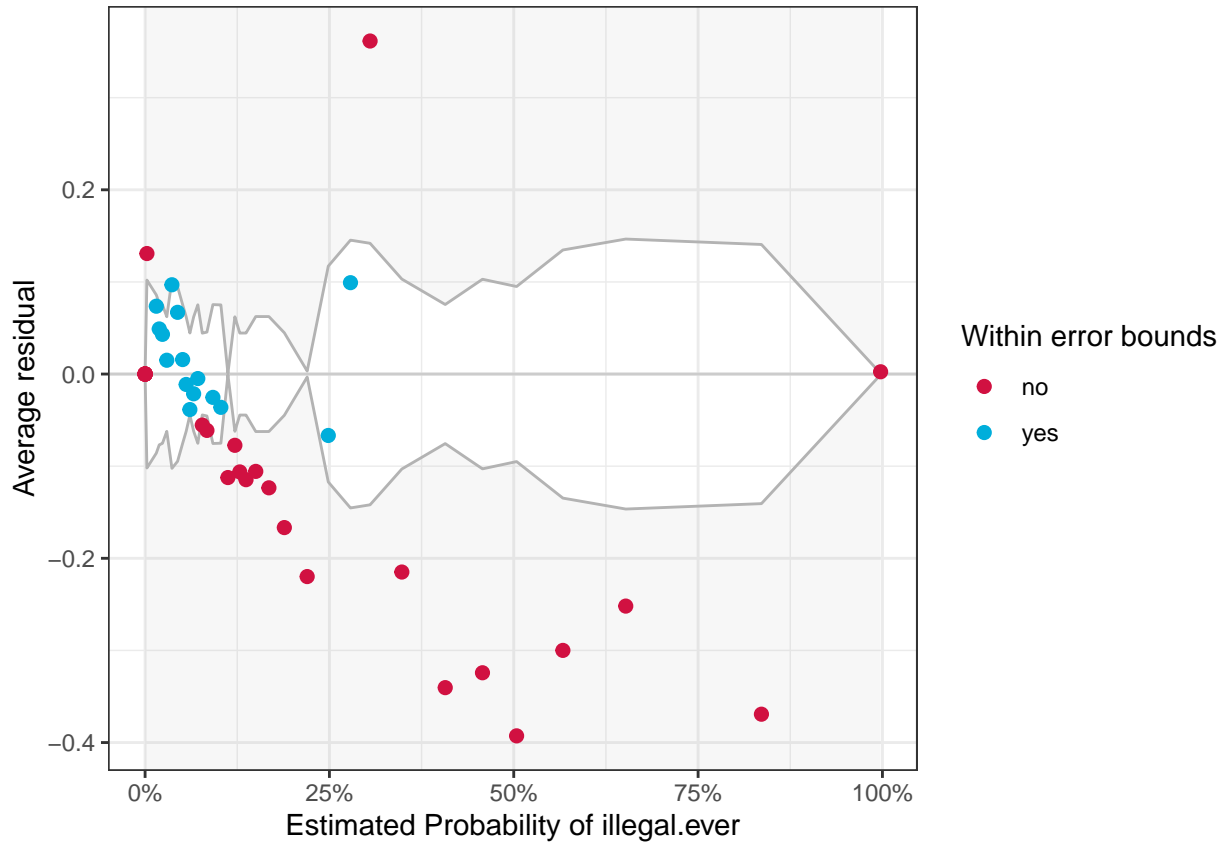
This result confirm the low predictive validity of the model, which not surprisingly did not identify any particular covariate for predicting the presence of one or more killed wolves in a municipality.