# Modelling probability of breeding, litter size and pup survival in Arctic foxes

Kristine Ulvund, Knut Anders Hovstad, Craig R. Jackson, Lars Rød-Eriksen, Veronika Areskoug, Torbjørn Ergon, Nina E. Eide, Øystein Flagstad, Arild Landa

## Table of contents

Introduction	1
Probability of breeding in captivity	2
Data description and models	2
Analysis and results	3
Litter size	5
Data description and models	5
Analysis and results	6
Litter size in captivity	6
Litter size in wild populations	
Survival of pups in captivity 1	.0
Data description and model	0
Analysis and results	1
Life-history components 1	.3
Analysis and results	.3
References 1	6

### Introduction

This document is part of the supplementary materials for the manuscript "Reproductive senescence and effects of maternal age on reproductive performance in the Arctic fox", and describes the analysis of variation in Arctic fox breeding probability, litter size and early pup survival.

The main objective was to examine the relationship between mother's age and probability of breeding, litter size and early pup survival (from den emergence to 8-10 weeks of age) in captive foxes and mother's age and litter size in wild foxes. The data on captive foxes are from a conservation breeding programme. Data on litter size in wild populations were collected from den sites in five different populations.

The most important prey for Arctic foxes in the wild is lemmings and other species of small rodents. The cyclic variation in abundance of small rodents create large variations in availability of prey between years. In contrast, foxes in the captive breeding station were fed every day and thus experienced no variation in food resources between years.

To run the R code in this document, it is necessary to first download these three datasets:

- (1) litter\_size\_captive\_breeding\_station.csv
- (2) litter\_size\_wild\_populations.csv
- (3) survival\_captive\_breeding\_station.csv

All analyses were conducted using R version 4.2.2 and the models were fitted using glmmTMB version 1.1.10 (Brooks et al. 2017). The assumptions for all models were checked using the DHARMa package version 0.4.7 (Hartig 2024, see code in Quarto source file for details). Further information about the study and the data can be found in the main manuscript.

Loading the required packages:

```
library(tidyverse)  # data management and plotting (ggplot2)
library(glmmTMB)  # glmmTMB function used to fit the models
library(splines)  # bs function for B-splines
library(DHARMa)  # functions to check model assumptions
library(ggeffects)  # ggemmeans function used to estimate marginal means
library(ggbeeswarm)  # geom_quasirandom for beeswarmplot
library(MuMIn)  # AICc function
library(knitr)  # functions to create tables
library(kableExtra)  # kbl function to create tables
```

Reading in the three data sets:

```
captive <- read_csv('data/litter_size_captive_breeding_station.csv')
wild <- read_csv('data/litter_size_wild_populations.csv')
survival_data <- read_csv("data/survival_captive_breeding_station.csv")</pre>
```

## Probability of breeding in captivity

## Data description and models

The data from foxes in captivity included both females breeding and not breeding. These data are not available for wild foxes. For captive foxes we could therefore model the probability of producing a litter. The binary outcome 'breeding' or 'no breeding' of individual foxes in each year was used as the response variable and analysed using a GLMM with a binomial distribution. The models included the fixed effect 'mother age' using either a quadratic or cubic function. The 'mother ID' and 'year' were included as random effects to account for correlation in the data.

The structural form of the model for probability of breeding with a quadratic function for 'mother age':

$$\begin{split} Y_{ij} \sim \text{Bernoulli}(b_{ij}) \\ \text{logit}(b_{ij}) &= \beta_0 + X_{ij}\beta_1 + X_{ij}^2\beta_2 + \delta_i + \gamma_j \\ \delta_i &\sim N(0, \sigma_\delta) \\ \epsilon_i &\sim N(0, \sigma_\epsilon) \end{split}$$

where  $b_{ij}$  is the probability of breeding for mother i in year j.

 $\delta_i$  is the random effect of mother i and  $\gamma_j$  the random effect of year j.

To examine the potential effect of rodents on probability of breeding in captivity, we also fitted a second model including rodent index as a predictor:

$$\begin{split} Y_{ij} \sim \text{Bernoulli}(b_{ij}) \\ \text{logit}(b_{ij}) &= \beta_0 + X_{ij}\beta_1 + X_{ij}^2\beta_2 + s(R_j) + \delta_i + \gamma_j \\ \delta_i \sim N(0, \sigma_\delta) \\ \epsilon_j \sim N(0, \sigma_\epsilon) \end{split}$$

s() is a B-spline function that allows for a smooth, non-linear relationship between  $logit(b_{ij})$  and the rodent index R in year j.

## Analysis and results

Initially, we created the variable breeding:

```
captive <- mutate(captive, breeding = ifelse(litter_size == 0, 0, 1))</pre>
```

Fit a model with only intercept and random effects:

```
b0 <- glmmTMB(breeding ~ 1 + (1|mother_id) + (1|year),
data = captive, family = binomial)
```

Fit quadratic model for probability of breeding:

Fit an alternative quadratic model including the variable rodent\_index as a predictor:

Fit cubic model for probability of breeding:

Fit an alternative cubic model including the variable rodent index as a predictor:

Table S1 shows that models which include mothers age using a quadratic function has lower AICc values than models with a cubic function. Adding 'rodent index' as a predictor to the models resulted in only a small change in AICc values. Considering the small difference in AIC between model b1 and b2, and no clear pattern of covariation between breeding and rodent abundance in the data, we decided to use the most parsimonious model b1.

**Table S1** Model comparison for probability of breeding. The table presents the fixed effects and corrected Akaike Information Criterion (AICc) values for each model.

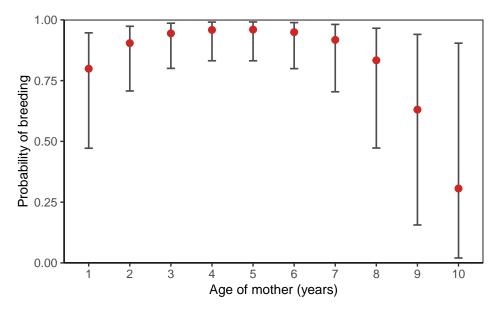
Model	Fixed effects	AICc
b0	intercept only	92.30
b1	poly(mother_age, 2)	88.69
b2	$poly(mother\_age, 2) + bs(rodent index, 3)$	88.90
b3	poly(mother_age, 3)	90.60
b4	$poly(mother\_age, 3) + bs(rodent index, 3)$	91.12

Estimated marginal means for probability of breeding are calculated using the ggemmeans function from the R package ggeffects (Lüdecke 2024). The R code used to generate Table S2 and Figure S1 is available in the Quarto source file included within the Supplementary Materials.

```
## Estimated marginal means for probability of breeding based on model b1
pb <- ggemmeans(b1, terms="mother_age [all]", type = 'fixed')</pre>
```

Table S2 Estimated marginal means for probability of breeding with 95% confidence intervals.

Age of mother	Estimated mean	95% CI
1	0.80	0.47,  0.95
2	0.90	0.71, 0.97
3	0.94	0.80,  0.99
4	0.96	0.83,  0.99
5	0.96	0.83,  0.99
6	0.95	0.80,  0.99
7	0.92	0.70, 0.98
8	0.83	0.47,  0.97
9	0.63	0.16, 0.94
10	0.31	0.02,0.90



**Figure S1** Probability of breeding in captive Arctic foxes plotted against age of mother. Whiskers indicate 95% confidence intervals.

## Litter size

## Data description and models

For the *wild populations*, data on litter size were only recorded for breeding events that actually produced pups and where the pups were observed after emerging from the den. The data does not include information about unsuccessful breeding events or foxes that did not breed at all.

For *captive foxes*, litter size data include zeros. For comparison with *wild foxes*, we have removed the zeros for captive foxes and fitted a zero-truncated model also for these data. This implies that we model the variation in litter size conditional on producing at least one pup that survived until marking.

In both datasets, there are repeated observations for some of the females (litter size for the same female in different years). To account for this dependence, we included the random effect 'mother ID'. In the datasets, each mother had a single litter per year.

For wild populations the effect of year can be expected to vary among mountain areas, but the data does not permit the estimation of a random effect of year within mountain area. The model for litter size in wild populations therefore included 'area' in addition to 'mother ID' as random effects. The model for litter size in captivity included 'year' and 'mother ID' as random effects.

The structural form of the model for litter size in wild populations:

$$\begin{split} Y_{i,k} &\sim \text{ztCMP}(\lambda_{i,k}, \phi) \\ \log(\lambda_{i,k}) &= \beta_0 + X_i \beta_1 + X_i^2 \beta_2 + s(R_k) + \delta_i + \epsilon_k \\ \delta_i &\sim N(0, \sigma_\delta) \\ \epsilon_k &\sim N(0, \sigma_\epsilon) \end{split}$$

where  $Y_{i,k}$  and  $\lambda_{i,k}$  are observed litter size and expected litter size of mother i in area k. ztCMP is the zero-truncated Conway-Maxwell Poisson distribution with dispersion scale  $\phi$ , as parameterized in the glmmTMB package (Brooks et al. 2017).  $X_i$  is the age of mother i included as a second order polynomial with intercept  $\beta_0$  the regression coefficients  $\beta_1$  and  $\beta_2$ .  $R_k$  is rodent index from area k and s() is a B-spline function that allows for a smooth, non-linear relationship between the response litter size and rodent index. The B-spline is implemented using the bs function from the splines package.  $\delta_i$  is the random effect of mother i and  $\epsilon_k$  is the random effect of mountain area k (both random effects assumed to be normal with zero mean).

Litter size in captivity was analysed using models that included the fixed effect 'mother age' as either a quadratic or cubic function, and 'mother ID' and 'year' as random effects. Initially, we also included 'rodent index' as a fixed effect but the data did not provide any support for a relationship between litter size and 'rodent index'. The 'rodent index' was therefore omitted from the final model for litter size in captivity.

## Analysis and results

#### Litter size in captivity

To ensure comparability between the litter size data from captive and wild populations, we excluded breeding attempts that did not result in pups before analysis. We then applied zero-truncated models to both datasets.

```
## Removing zeros for truncated models
captive_tr <- captive[captive$litter_size != 0,]</pre>
```

Fit model with only intercept and random effects:

Fit model with quadratic function for 'mother age':

Fit model with quadratic function for 'mother age' and 'B-spline' function for 'rodent index':

Fit model with cubic function for 'mother age':

In addition, we tried to fit a model with only 'rodent index' as fixed effect but we could not get this model to converge.

The model with a quadratic function for 'mother age' had the lowest AICc value (Table S3). The difference in AICc to the model with only intercept was 4.58 and thus 'mother age' significantly improved the model. The models with 'rodent index' as predictor did not improve the AICc value.

**Table S3** Model comparison for litter size in captive foxes. The table presents the fixed effects and corrected Akaike Information Criterion (AICc) values for each model.

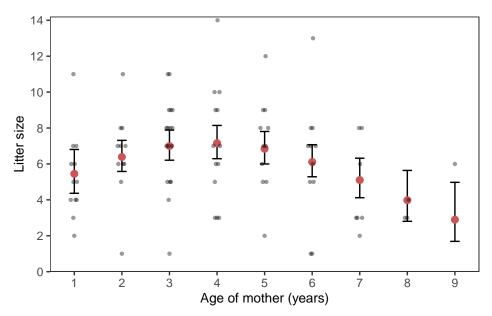
Model	Fixed effects	AICc
c0	intercept only	425.88
c1	poly(mother_age, 2)	421.31
c2	$poly(mother\_age, 2) + bs(rodent index, 3)$	422.28
c3	poly(mother_age, 3)	423.00

Estimated marginal means for litter size in captivity were calculated based on model c1, and the results displayed in Table S4 and Figure S2.

```
pc <- ggemmeans(c1, terms = "mother_age [all]", type = "fixed")</pre>
```

Table S4 Estimated marginal means for litter size in captivity with 95% confidence intervals.

Age of mother	Estimated mean	95% CI
1	5.45	4.36, 6.81
2	6.39	5.58, 7.31
3	7.00	6.21, 7.88
4	7.16	6.29, 8.14
5	6.84	6.00, 7.81
6	6.11	5.28, 7.07
7	5.10	4.12, 6.32
8	3.98	2.81, 5.64
9	2.90	1.69, 4.97



**Figure S2** Litter size as a function of female age in captive Arctic foxes. Estimated marginal means are displayed as red dots with whiskers indicating 95% confidence intervals. Observations are displayed as smaller, grey dots.

#### Litter size in wild populations

Fit model with only intercept and random effects:

Fit model with quadratic function for 'mother age':

Fit model with B-spline function for 'rodent index':

Fit a model with quadratic function for 'mother age' and B-spline function for 'rodent index':

Fit model with cubic function for 'mother age':

Fit model with a cubic function for 'mothers age' and a B-spline function for 'rodent index'

The best model, as indicated by AICc, included only 'rodent index' as a fixed effect in addition to the intercept (Table S5). Using a cubic function for 'mother age' did not improve the model. The model including both a quadratic function for 'mother age' and a B-spline function for 'rodent index' had a somewhat higher AICc value compared to the model with only 'rodent index'. Consequently, we cannot conclude that litter size in wild foxes depends on maternal age. Given that the primary objective of this study was to investigate the effects of maternal age on Arctic fox reproductive output, we opted to present the full model that includes a quadratic function for 'mother age'.

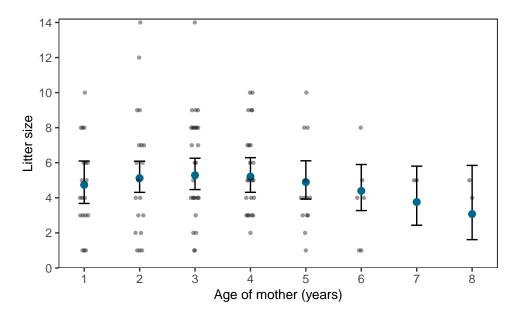
**Table S5** Model comparison for litter size in wild populations. The table presents the fixed effects and corrected Akaike Information Criterion (AICc) values for each model.

Model	Fixed effects	AICc
w0	intercept only	530.32
w1	poly(mother_age, 2)	531.31
w2	bs(rodent index, 3)	527.31
w3	$poly(mother\_age, 2) + bs(rodent index, 3)$	529.11
w4	$poly(mother\_age, 3)$	531.27
w5	$poly(mother\_age, 3) + bs(rodent index, 3)$	530.68

Estimated marginal means for litter size in wild populations were calculated based on model w3, and results are provided in Table S6 and also illustrated in Figure S3.

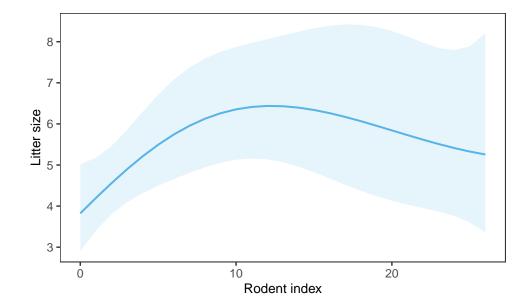
Table S6 Estimated marginal means for litter size in wild populations with 95% confidence intervals.

Age of mother	Estimated mean	95% CI
1	4.74	3.68, 6.10
2	5.12	4.31, 6.09
3	5.29	4.47, 6.26
4	5.21	4.31, 6.29
5	4.90	3.93, 6.11
6	4.39	3.27, 5.90
7	3.76	2.44, 5.81
8	3.07	1.62, 5.85



**Figure S3** Litter size as a function of female age in wild Arctic foxes. Estimated marginal means are displayed as blue dots and observations as smaller, grey dots. Whiskers indicate 95% confidence intervals.

Estimated litter size in wild populations in response to rodent index is given in Figure S4.



**Figure S4** Relationship between rodent index and estimated litter size in wild Arctic foxes. The effect of rodent index on litter size is described by a B-spline function.

## Survival of pups in captivity

### Data description and model

The dataset includes information on survival of individual Arctic fox pups in the captive breading station measured between emergence from the den until eight to ten weeks of age. Survival was analyzed using a generalized linear mixed model (GLMM) with a binomial distribution and logit link function. The age of the breeding female was included using a quadratic function to allow for a non-linear relationship between pup survival and mother age. The model also included random effects for year and 'mother ID' to account for correlation in the data. Due to the dataset's limited size and the high probability of survival, insufficient information was available to accurately estimate a random effect of 'litter ID'. Additionally, the presence of litters with only one or two pups further complicated the estimation of a random effect associated with litter ID.

The structural form of the binomial model for pup survival:

$$\begin{split} Y_{ij} \sim \text{Bernoulli}(b_{ij}) \\ \text{logit}(b_{ij}) &= \beta_0 + X_{ij}\beta_1 + X_{ij}^2\beta_2 + s(R_j) + \delta_i + \gamma_j \\ \delta_i \sim N(0, \sigma_\delta) \\ \epsilon_j \sim N(0, \sigma_\epsilon) \end{split}$$

s() is a B-spline function that allows for a smooth, non-linear relationship between  $\log \operatorname{id}(b_{ij})$  and the rodent index R in year j.

## Analysis and results

Fit model with only intercept and random effects:

```
::: {.cell}
```

:::

Fit quadratic model for pup survival:

Fit model for pup survival with 'rodent index' as predictor:

Fit model including 'mother age' and 'rodent index':

Fit model with cubic function for 'mother age':

Fit model with cubic function for 'mother age' and B-spline for 'rodent index':

**Table S7** Model comparison for probability of early pup survival in captivity. The table presents the fixed effects and corrected Akaike Information Criterion (AICc) values for each model.

Model	Fixed effects	AICc
s0	intercept only	291.09
s1	poly(mother_age, 2)	281.55
s2	bs(rodent index, 3)	293.57
s3	$poly(mother\_age, 2) + bs(rodent index, 3)$	285.39
s4	$poly(mother\_age, 3)$	279.09
s5	$poly(mother\_age, 3) + bs(rodent index, 3)$	282.32

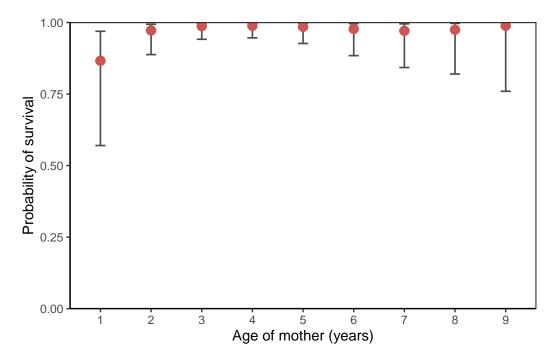
The best model with respect to AICc was model s4 including a cubic function for 'mother age' as the only fixed effect in addition to the intercept (Table S7).

Survival of pups in captivity is estimated using the ggemmeans function from the ggeffects package and the results displayed in Table S8 and Figure S5.

```
## Estimated marginal means based on model s4
ps <- ggemmeans(s4, terms = "mother_age [all]", type = "fixed")</pre>
```

Table S8 Estimated marginal means for early pup survival in captivity with 95% confidence intervals.

Age of mother	Estimated mean	95% CI
1	0.87	0.57,  0.97
2	0.97	0.89,  0.99
3	0.99	0.94, 1.00
4	0.99	0.95, 1.00
5	0.98	0.93, 1.00
6	0.98	0.88, 1.00
7	0.97	0.84, 1.00
8	0.97	0.82, 1.00
9	0.99	0.76, 1.00



**Figure S5** Probability of survival in captive Arctic fox pups from birth until 8-10 weeks of age plotted against age of mother.

## Life-history components

In this section, we estimate mean expected surviving pups and corresponding 95% confidence intervals based on estimated probability of breeding, litter size and pup survival. Variation in the estimated means between age groups, was used to calculate the contribution of variance in each life-history component to variance in overall reproductive output.

## Analysis and results

The expected number of surviving offspring in captivity  $(\widehat{R}_a)$ , depending on mother age (a), can be computed by multiplying the predicted breeding probability  $(\widehat{B}_a)$ , litter size  $(\widehat{L}_a)$ , and offspring survival  $(S_a)$  for each mother age.

$$\widehat{R}_a = \widehat{B}_a \widehat{L}_a \widehat{S}_a$$

To compute an approximate 95% confidence interval, we first computed the approximate sampling variance and confidence interval on the log-scale, and then back transform the confidence interval to the arithmetic scale.

Since  $\log(\widehat{R}_a) = \log(\widehat{B}_a) + \log(\widehat{L}_a) + \log(\widehat{S}_a)$  and the three components were estimated independently, we got

$$\operatorname{Var}\left(\log(\widehat{R}_a)\right) = \operatorname{Var}\left(\log(\widehat{B}_a)\right) + \operatorname{Var}\left(\log(\widehat{L}_a)\right) + \operatorname{Var}\left(\log(\widehat{S}_a)\right)$$

Since  $L_a$  was estimated using a log-link,  $\operatorname{Var}\left(\log(\widehat{L}_a)\right)$  is simply the sampling variance for the linear predictor in this model. The probabilities  $B_a$  and  $S_a$  were estimated using a logit-link,  $p=\frac{1}{1+e^{-\eta}}$  where p is either  $B_a$  or  $S_a$  and  $\eta$  is the corresponding linear predictor. To compute approximations for  $\operatorname{Var}\left(\log(p)\right)$ , we used the delta-method approximation (Pawitan 2001).

$$\mathrm{Var}\left(\log(p)\right) \approx \left(\frac{\partial \log(p)}{\partial \eta}\right)^2 \mathrm{Var}(\eta) = \left(\frac{\partial - \log(1 + e^{-\eta})}{\partial \eta}\right)^2 \mathrm{Var}(\eta)$$

The chain rule gives

$$\frac{\partial -\log(1+e^{-\eta})}{\partial \eta} = -\frac{1}{1+e^{-\eta}}(-e^{-\eta}) = 1-p$$

Hence,

$$\mathrm{Var}\left(\log(p)\right) \approx \left(1-p\right)^2 \mathrm{Var}(\eta)$$

Compute expected numbers of surviving pups (with 95% confidence intervals) for each age group:

```
# summary(b1) # Breeding probability model
# summary(c1) # litter size model
# summary(s1) # pup survival model
# Computing linear predictors (etas)
# NB! polynomial contrasts are orthogonal and hence depend on the distribution
# in the data. Hence, computing predictors from original data and extracting
# the values corresponding to each mother age.
X_B <- model.matrix(~ poly(mother_age, 2), captive)</pre>
X_L <- model.matrix(~ poly(mother_age, 2), captive_tr)</pre>
X_S <- model.matrix(~ poly(mother_age, 2), survival_data)</pre>
X_B <- X_B[match(1:9,captive$mother_age), ]</pre>
X_L <- X_L[match(1:9, captive_tr$mother_age), ]</pre>
X_S <- X_S[match(1:9,survival_data$mother_age), ]</pre>
eta_B <- X_B %*% fixef(b1)$cond # breeding P
eta L <- X L %*% fixef(c1)$cond # litter size given breeding
eta_S <- X_S %*% fixef(s1)$cond # pup survival P</pre>
# Variances of linear predictors
V_eta_B <- X_B %*% vcov(b1)$cond %*% t(X_B) |> diag()
V_eta_L <- X_L %*% vcov(c1)$cond %*% t(X_L) |> diag()
V_eta_S <- X_S %*% vcov(s1)$cond %*% t(X_S) |> diag()
# The three components and expected number of survivors
(L <- exp(eta_L))
       [,1]
2 5.450454
4 6.388306
1 6.995824
6 7.158008
11 6.842985
14 6.112223
26 5.100973
33 3.977470
47 2.897750
(B \leftarrow 1/(1+exp(-eta_B)))
         [,1]
2 0.7991215
4 0.9042762
1 0.9444447
6 0.9586504
9 0.9599365
16 0.9494076
28 0.9176143
35 0.8336041
50 0.6306734
```

```
(S \leftarrow 1/(1+exp(-eta_S)))
          [,1]
17 0.9197091
36 0.9682568
   0.9840032
21 0.9894670
52 0.9908795
89 0.9895999
159 0.9844022
206 0.9694191
291 0.9234192
R <- B*L*S
# Variances of log(B), log(L), log(S), and log(Lambda)
V_log_L <- V_eta_L</pre>
V_{\log_B} \leftarrow (1-B)^2 * V_{\max_B}  # delta method approx.
V \log S \leftarrow (1-S)^2 * V \text{ eta } S \# \text{ delta method approx},
V_log_R <- V_log_B + V_log_L + V_log_S</pre>
# Computing approx. 95% CI (central limit theorem) and making a data frame
SE_log_R <- sqrt(V_log_R)</pre>
expected_survivors <- data.frame(</pre>
  mother_age = 1:9,
  R = R
  lwr_R = exp(log(R) - 2*SE_log_R),
  upr_R = exp(log(R) + 2*SE_log_R)
expected_survivors
   mother_age
                      R
                             lwr_R
                                      upr_R
2
            1 4.005862 2.6726193 6.004197
4
             2 5.593419 4.5858751 6.822327
            3 6.501475 5.5991053 7.549273
1
6
            4 6.789749 5.8549403 7.873811
9
            5 6.508921 5.6020334 7.562620
16
            6 5.742640 4.8453979 6.806028
28
            7 4.607716 3.5671143 5.951884
35
            8 3.214240 2.0173854 5.121153
```

9 1.687580 0.6048753 4.708287

50

The variance components are calculated like this:

variance in reproductive success respectively.

```
V1B <- var(log(B))
V1L <- var(log(L))
V1S <- var(log(S))

V1T <- V1B + V1L + V1S

V1B/V1T

[,1]
[1,] 0.1695419

V1L/V1T

[,1]
[1,] 0.8226681

V1S/V1T

[,1]
```

Variance in the point estimates of breeding probability, litter size given breeding, and pup survival at the logarithmic scale across the mother age groups accounted for about 54.1%, 45.5% and 0.4% of the

## References

[1,] 0.007790049

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal, 9(2), 378-400 https://doi.org/10.32614/RJ-2017-066

Hartig, F. (2024). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.7, https://CRAN.R-project.org/package=DHARMa.

Lüdecke, D. (2024). ggeffects: Marginal Means And Adjusted Predictions Of Regression Models. Retrieved October 20, 2024, from https://strengejacke.github.io/ggeffects/articles/ggeffects.html

Pawitan Y. (2001). In all likelihood: statistical modelling and inference using likelihood. Oxford: Oxford University Press.