

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

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Litter size</b>	<b>2</b>
2.1	Data structure and models . . . . .	2
2.2	Analysis and results . . . . .	3
2.2.1	Litter size in wild populations . . . . .	3
2.2.2	Litter size in captivity . . . . .	7
<b>3</b>	<b>Probability of breeding in captivity</b>	<b>10</b>
3.1	Data description and model . . . . .	10
3.2	Analysis and results . . . . .	10
<b>4</b>	<b>Survival of pups in captivity</b>	<b>13</b>
4.1	Data structure and model . . . . .	13
4.2	Analysis and results . . . . .	14

## 1 Introduction

This document describes the analysis of variation in Arctic fox litter size, breeding probability and early survival of pups.

The main objective is to examine the relationship between litter size and mother's age. We compare data on litter size in wild populations with data from foxes kept in captivity as part of a conservation program. For foxes in captivity, data on pup survival until 8-10 weeks of age. Based on these data the relationship between pup survival and mother's age is examined.

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 are fed every day and thus experience no variation in food resources between years.

This document is part of the supplementary materials for the manuscript “**Reproductive senescence in cyclic environments: effects of maternal age on reproductive performance in the Arctic fox**”.

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

Further information about the study and the data can be found in the main manuscript.

## 2 Litter size

### 2.1 Data structure 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 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 include the random effect ‘mother ID’. In the datasets, each mother has 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 includes ‘year’ and ‘mother ID’ as random effects.

The model for litter size in wild populations is:

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

where  $Y_{i,k}$  and  $\lambda_{i,k}$  are observed litter size and expected litter size of mother  $i$  in area  $k$ .  $\text{ztCMP}$  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 a model with age of mother as a fixed effect, 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.

## 2.2 Analysis and results

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) # ggpredict function for predictions
library(ggbeeswarm) # geom_quasirandom for beeswarmplot
```

Reading in the three datasets:

```
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")
```

### 2.2.1 Litter size in wild populations

Fit polynomial model for litter size in wild populations:

```
w1 <- glmmTMB(litter_size ~ poly(mother_age, 2) +
              bs(rodent_index, 3) +
              (1|mother_id) + (1|mountain_area),
              data = wild,
              family=truncated_compois)
```

The assumptions for the model were checked using the DHARMA package (for details, see code in Quarto source file).

```
plotQQunif(w1, testDispersion = FALSE, main = " ")
```

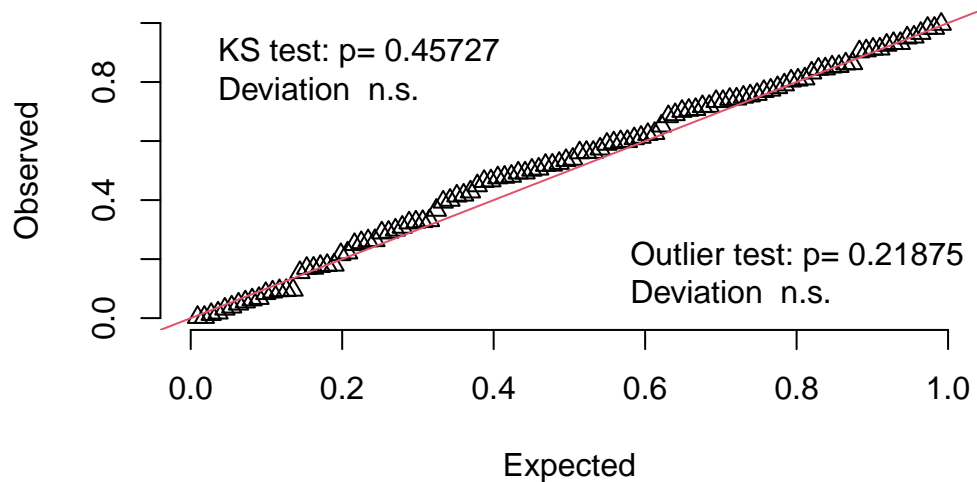


Figure S1. Scaled residual QQ plot for the model of litter size in wild populations. The figure also includes statistics used to examine model assumptions.

Predictions are calculated for each mother age from 1 to 8 keeping the rodent index equal to median value of the variable (`rodent_index = 4`). Random effects are ignored by using the argument `type = 'fixed'`.

```
(pw <- ggpredict(w1, terms="mother_age [all]",
  type = 'fixed',
  condition = c(rodent_index = 4)))
```

# Predicted values of litter\_size

mother_age	Predicted	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

Adjusted for:

```
* mother_id = NA (population-level)
* mountain_area = NA (population-level)
```

```

## Plot of predictions with observations
pw |> ggplot(aes(x = x, y = predicted)) +
  geom_errorbar(width = 0.2,
               aes(ymin = conf.low, ymax = conf.high)) +
  geom_point(color = 'deepskyblue4', size = 2.15) +
  scale_y_continuous(limits = c(0, 14.2),
                    expand = c(0, 0),
                    breaks = c(0, 2, 4, 6, 8, 10, 12, 14)) +
  scale_x_continuous(limits = c(0.9, 8.1),
                    breaks = 1:8) +
  labs(y="Litter size", x = "Age of mother (years)") +
  geom_quasirandom(data = wild,
                  mapping=aes(x=mother_age, y= litter_size),
                  alpha = 0.4, size = 0.85, width = 0.1) +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank())

```

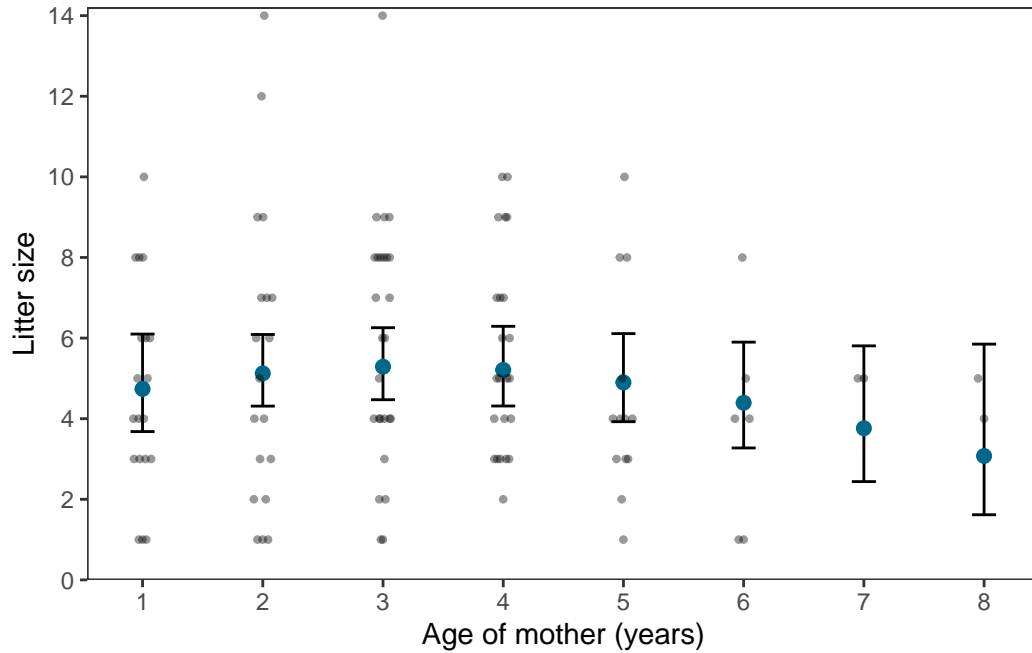


Figure S2. Predicted litter size as a function of female age in wild Arctic foxes. The predicted values are displayed as blue dots and observations as smaller, grey dots. Whiskers indicate 95% confidence intervals.

Predictions of litter size in wild populations in response to rodent index:

```
## Predictions of litter size in wild populations in response to rodent index
ggpredict(w1,
          terms = c("rodent_index [0:26.5, by=1]"),
          type = "fixed",
          condition = c(mother_age = 4)) |>
plot(colors = "#56B4E9") +
theme_bw() +
theme(plot.title = element_blank(),
      panel.grid.major.x = element_blank(),
      panel.grid.minor.x = element_blank(),
      panel.grid.major.y = element_blank(),
      panel.grid.minor.y = element_blank(),
      axis.text.x=element_text(size=9),
      axis.text.y=element_text(size=9),
      axis.title.x = element_text(color = "black", size = 10),
      axis.title.y = element_text(color = "black", size = 10)) +
labs(x = "Rodent index", y = "Litter size")
```

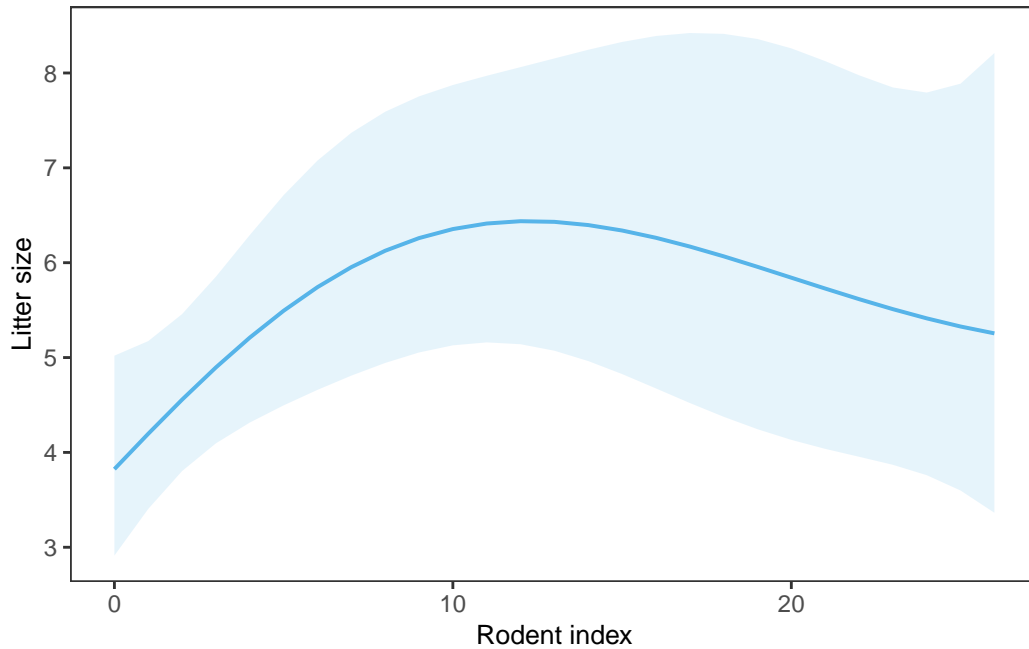


Figure S3. Relationship between rodent index and predicted litter size in wild Arctic foxes. The effect of rodent index on litter size is described by a B-spline function.

### 2.2.2 Litter size in captivity

To make the data on litter size in captivity and wild populations comparable, breeding attempts that did not produce pups are removed prior to analysis.

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

Fit polynomial model for litter size in captivity:

```
c1 <- glmmTMB(litter_size ~ poly(mother_age, 2) +
              (1|mother_id) + (1|year),
              data = captive_tr, family=truncated_compois)
```

AIC value for the model c1 is 382.58

We also fitted a model c2 including both mother\_age and rodent\_index as predictors:

```
c2 <- glmmTMB(litter_size ~ poly(mother_age, 2) +
              bs(rodent_index, 3) +
              (1|mother_id) + (1|year),
              data = captive_tr, family=truncated_compois)
```

AIC value for the model c2 including rodent\_index as a predictor is 382.10

The potential relationship between litter size and rodent index can be examined by plotting predicted values for litter size against rodent index:

```
## Predictions of litter size in captivity in response to rodent index
ggpredict(c2,
          terms = c("rodent_index [0:66, by=1]"),
          type = "fixed",
          condition = c(mother_age = 4)) |>
plot(colors = "#009E73") +
theme_bw() +
theme(plot.title = element_blank(),
      panel.grid.major.x = element_blank(),
      panel.grid.minor.x = element_blank(),
      panel.grid.major.y = element_blank(),
      panel.grid.minor.y = element_blank(),
      axis.text.x = element_text(size = 10),
      axis.text.y = element_text(size = 10),
      axis.title.x = element_text(color = "black", size = 12),
      axis.title.y = element_text(color = "black", size = 12),
      plot.margin = margin(t = 0, r = 10, b = 0, l = 10, unit = "pt")) +
labs(x = "Rodent index", y = "Litter size") +
scale_x_continuous(expand = c(0.02, 0)) +
scale_y_continuous(expand = c(0.02, 0))
```

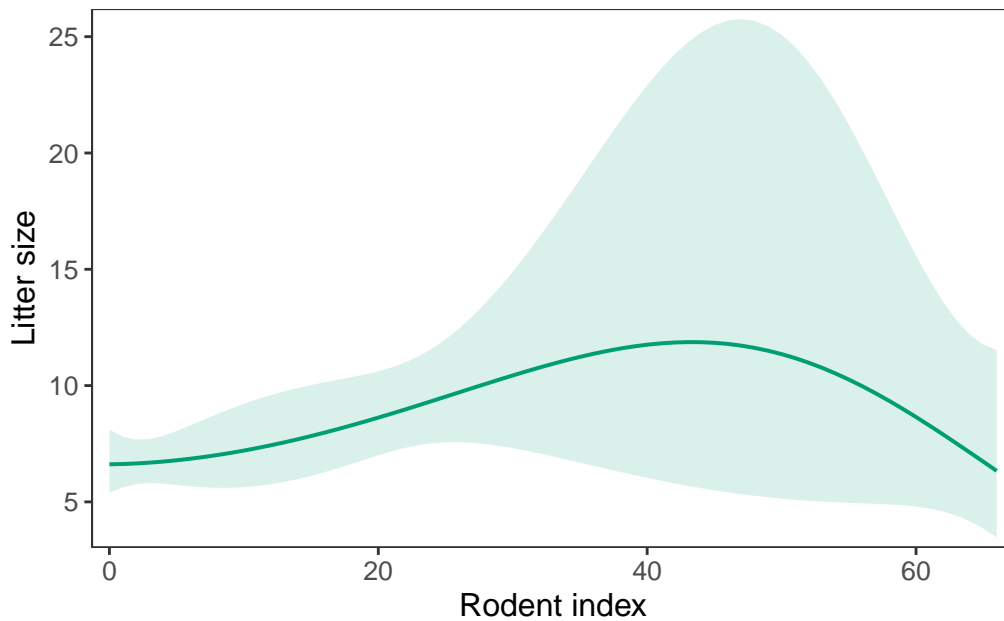


Figure S4. Relationship between rodent index and predicted litter size in captive Arctic foxes. The effect of rodent index on litter size is described by a B-spline function.

The plot does not display any clear relationship between variation in the rodent index and litter size in captivity. Furthermore, including the rodent index as a predictor in the model does not improve the likelihood of the model as assessed by the AIC values. We therefore exclude rodent index from the final model for litter size in captivity.

Predictions for foxes in captivity based on model c1:

```
pc <- ggpredict(c1, terms="mother_age [all]", type = 'fixed')
print(pc, n = Inf)
```

# Predicted values of litter\_size

mother_age	Predicted	95% CI
1	4.67	3.59, 6.07
2	5.80	4.91, 6.85
3	6.65	5.76, 7.69
4	7.05	6.08, 8.18
5	6.90	5.96, 8.00
6	6.24	5.29, 7.36
7	5.21	4.06, 6.70
8	4.02	2.67, 6.07
9	2.87	1.52, 5.42

Adjusted for:

```
* mother_id = NA (population-level)
* year = NA (population-level)
```



```
pc |> tibble() |> ggplot(aes(x = x, y = predicted)) +
  geom_errorbar(width = 0.2,
               aes(ymin = conf.low, ymax = conf.high)) +
  geom_point(color = 'indianred3', size = 2.15) +
  scale_y_continuous(limits = c(0, 14.2),
                    expand = c(0, 0),
                    breaks = c(0, 2, 4, 6, 8, 10, 12, 14)) +
  scale_x_continuous(limits = c(0.9, 9.1),
                    breaks = 1:9) +
  labs(y="Litter size", x = "Age of mother (years)") +
  geom_quasirandom(data = captive_tr,
                  mapping=aes(x=mother_age, y= litter_size),
                  alpha = 0.4, size = 0.85, width = 0.1) +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank())
```

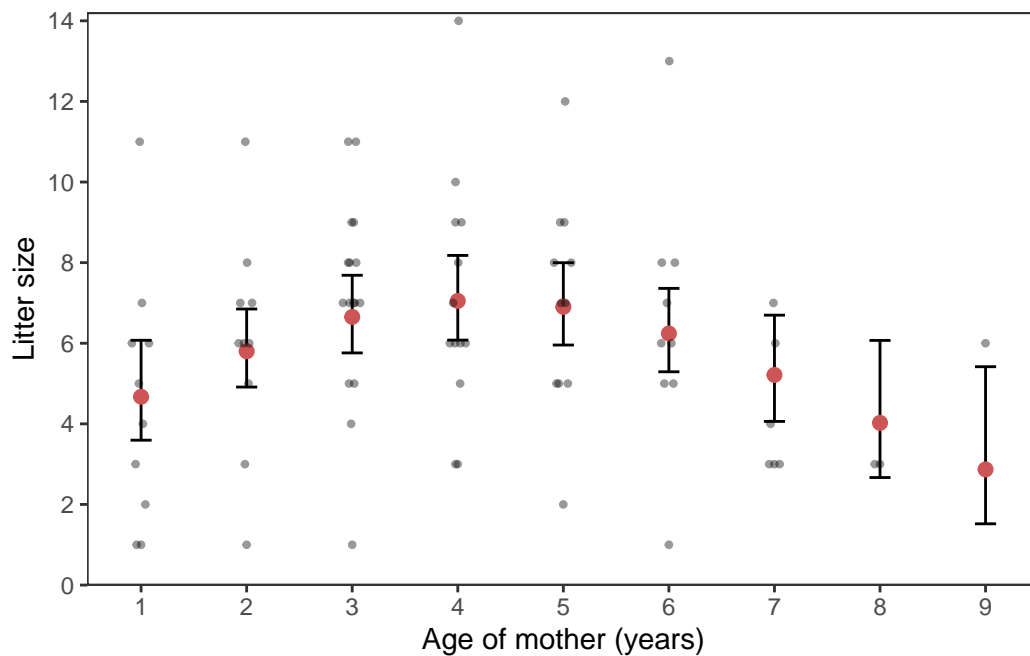


Figure S5. Predicted litter size as a function of mother age in captive Arctic foxes. The predicted values are displayed as red dots and observations as smaller, grey dots. Whiskers indicate 95% confidence intervals.

### 3 Probability of breeding in captivity

#### 3.1 Data description and model

The data from foxes in captivity include both females breeding and not breeding. This data is not available for wild foxes. For captive foxes we can therefore model the probability of producing a litter. The binary outcome ‘breeding’ or ‘no breeding’ of individual foxes in each year is used as the response variable and analysed using a GLMM with a binomial distribution. The model include the fixed effect ‘age of the mother’ using a second order polynomial 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:

$$\begin{aligned} b_{ij} &= \text{Prob}(Y_{ij} = 1|X_{ij}) \\ \text{logit}(b_{ij}) &= \beta_0 + \beta_1 X_{ij} + \beta_2 X_{ij}^2 + \delta_i + \gamma_j \\ Y_{ij} &\sim \text{Bern}(b_{ij}) \\ \delta_i &\sim N(0, \sigma_\delta) \\ \epsilon_j &\sim N(0, \sigma_\epsilon) \end{aligned}$$

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 fitted a second model including that also include ‘rodent index’ as a fixed effect using a B-spline function to account for non-linear responses:

$$\begin{aligned} b_{ij} &= \text{Prob}(Y_{ij} = 1|X_{ij}) \\ \text{logit}(b_{ij}) &= \beta_0 + \beta_1 X_{ij} + \beta_2 X_{ij}^2 + s(R_j) + \delta_i + \gamma_j \\ Y_{ij} &\sim \text{Bern}(b_{ij}) \\ \delta_i &\sim N(0, \sigma_\delta) \\ \epsilon_j &\sim N(0, \sigma_\epsilon) \end{aligned}$$

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

#### 3.2 Analysis and results

Initially, create the variable breeding:

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

Fit polynomial model for probability of breeding:

```
b1 <- glmmTMB(breeding ~ poly(mother_age, 2) +
              (1|mother_id) + (1|year),
              data = captive,
              family = binomial)
summary(b1)
```

```

Family: binomial ( logit )
Formula:          breeding ~ poly(mother_age, 2) + (1 | mother_id) + (1 | year)
Data: captive

```

AIC	BIC	logLik	deviance	df.resid
112.3	125.5	-51.2	102.3	98

Random effects:

Conditional model:

Groups	Name	Variance	Std.Dev.
mother_id	(Intercept)	0.03696	0.1922
year	(Intercept)	0.68027	0.8248

Number of obs: 103, groups: mother\_id, 21; year, 15

Conditional model:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	1.4238	0.3929	3.624	0.00029 ***
poly(mother_age, 2)1	-0.6927	2.6176	-0.265	0.79128
poly(mother_age, 2)2	-7.9121	2.9221	-2.708	0.00678 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

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

```

b2 <- glmmTMB(breeding ~ poly(mother_age, 2) + bs(rodent_index, 3) +
              (1|mother_id) + (1|year),
              data = captive,
              family = binomial)
summary(b2)

```

```

Family: binomial ( logit )
Formula:
breeding ~ poly(mother_age, 2) + bs(rodent_index, 3) + (1 | mother_id) +
(1 | year)
Data: captive

```

AIC	BIC	logLik	deviance	df.resid
111.0	132.1	-47.5	95.0	95

Random effects:

Conditional model:

Groups	Name	Variance	Std.Dev.
mother_id	(Intercept)	0.1689	0.4110
year	(Intercept)	0.2099	0.4582

Number of obs: 103, groups: mother\_id, 21; year, 15

Conditional model:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.395912	0.550651	0.719	0.4721
poly(mother_age, 2)1	-0.009526	2.676110	-0.004	0.9972
poly(mother_age, 2)2	-7.737750	2.979442	-2.597	0.0094 **
bs(rodent_index, 3)1	10.907175	8.433659	1.293	0.1959
bs(rodent_index, 3)2	-51.515345	75.812392	-0.679	0.4968
bs(rodent_index, 3)3	256.176605	470.985516	0.544	0.5865

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The model b1 for probability of breeding in captivity has an AIC value of 112.30. For comparison, the model b2 including 'rodent index' as a predictor has an AIC value of 111.00. 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.

```
## Predictions based on model b1
(pred_b1 <- ggpredict(b1, terms="mother_age [all]", type = 'fixed'))
```

# Predicted probabilities of breeding

mother_age	Predicted	95% CI
1	0.63	0.37, 0.83
2	0.79	0.61, 0.90
3	0.86	0.72, 0.94
4	0.89	0.75, 0.96
6	0.86	0.69, 0.94
7	0.77	0.55, 0.90
8	0.59	0.28, 0.85
10	0.11	0.01, 0.70

Adjusted for:

\* mother\_id = NA (population-level)

\* year = NA (population-level)

Not all rows are shown in the output. Use `print(..., n = Inf)` to show all rows.

```
pred_b1 |> tibble() |>
  ggplot(aes(x=x, y= predicted)) +
  geom_errorbar(data=tibble(pred_b1),
               mapping=aes(x=as.factor(x), ymin=conf.low, ymax=conf.high),
               width=0.2, linewidth=0.58, color="grey30") +
  geom_point(color = 'firebrick3', size = 3) +
  scale_y_continuous(limits = c(0, 1),
                    expand = c(0, 0),) +
  labs(y="Probability of breeding", x = "Age of mother (years)") +
  theme_bw() +
  theme(axis.line = element_line(colour = "black"),
        panel.grid.major.x = element_blank(),
```

```

panel.grid.minor.x = element_blank(),
panel.grid.major.y = element_blank(),
panel.grid.minor.y = element_blank()

```

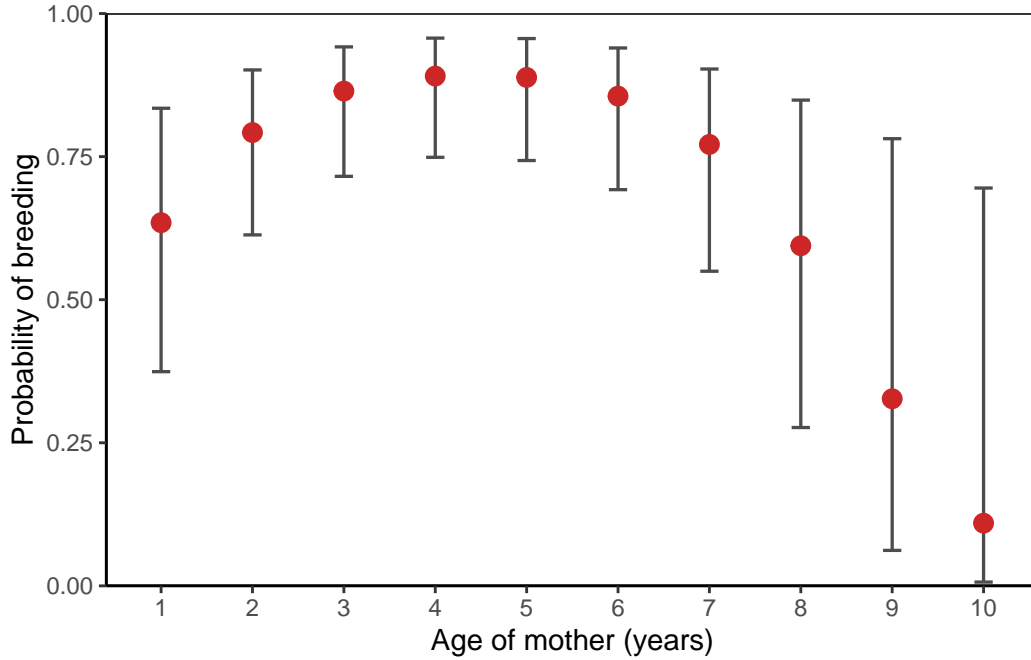


Figure S6. Probability of breeding in captive Arctic foxes plotted against age of mother.

## 4 Survival of pups in captivity

### 4.1 Data structure and model

The dataset includes information on survival of individual Arctic fox pups in the captive breeding station measured between emergence from the den until eight to ten weeks of age. Survival is analyzed using a generalized linear mixed model (GLMM) with a binomial distribution and logit link function. The age of the breeding female is included using a quadratic function to allow for a non-linear relationship between pup survival and mother age. The model also includes 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 is available to accurately estimate a random effect of ‘litter ID’. Additionally, the presence of litters with only one or two pups further complicates the estimation of a random effect associated with litter ID.

The structural form of the binomial model for pup survival:

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

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

## 4.2 Analysis and results

Fit polynomial model for pup survival:

```
s1 <- glmmTMB(survival ~ poly(mother_age, 2) +
              (1|mother_id) + (1|year),
              data = survival_data,
              family = binomial)
```

Fit an alternative model s2 for pup survival including the variable rodent\_index as a predictor:

```
s2 <- glmmTMB(survival ~ poly(mother_age, 2) +
              bs(rodent_index, 3) +
              (1|mother_id) + (1|year),
              data = survival_data,
              family = binomial)
```

The AIC values of models s1 and s2 are 281.44 and 285.12 respectively. Thus, including ‘rodent index’ as a predictor variable does not improve model performance measured by AIC. We therefore continue with the more parsimonious model s1.

Survival of pups in captivity is predicted using the ggpredict function from the ggeffects package:

```
## Predictions based on model s1
(pred_s1 <- ggpredict(s1, terms="mother_age [all]"))
```

# Predicted probabilities of survival

mother_age	Predicted	95% CI
1	0.92	0.68, 0.98
2	0.97	0.86, 0.99
3	0.98	0.92, 1.00
4	0.99	0.94, 1.00
5	0.99	0.95, 1.00
6	0.99	0.94, 1.00
7	0.98	0.90, 1.00
9	0.92	0.43, 0.99

Adjusted for:

```
* mother_id = NA (population-level)
*      year = NA (population-level)
```

Not all rows are shown in the output. Use ``print(..., n = Inf)`` to show all rows.

```

pred_s1 |> tibble() |>
  ggplot(aes(x=x, y= predicted)) +
  geom_errorbar(data=tibble(pred_s1),
               mapping=aes(x=as.factor(x), ymin=conf.low, ymax=conf.high),
               width=0.2, linewidth=0.58, color="grey30") +
  geom_point(color = 'aquamarine3', size = 3) +
  scale_y_continuous(limits = c(0, 1),
                    expand = c(0, 0),) +
  labs(y="Probability of survival", x = "Age of mother (years)") +
  theme_bw() +
  theme(axis.line = element_line(colour = "black"),
        panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank())

```

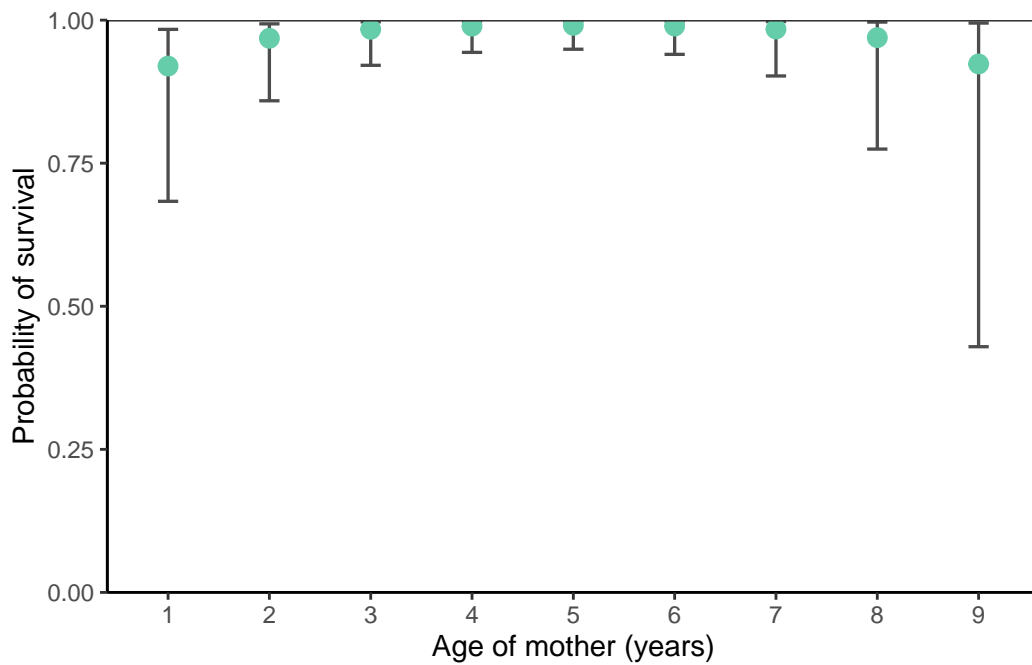


Figure S7. Probability of survival in captive Arctic fox pups from birth until 8-10 weeks of age plotted against age of mother.