

Behavioral Ecology and Sociobiology

Maternal, offspring and environmental influences on weaning age and supersuckling in Galápagos Sea Lions

Supplementary materials

Authors (*ORCID number*): Childs, Alexandra K. (0009-0003-4150-6745) *†¹; Feldmann, Carlina C.² (0009-0009-1167-7811); Dyck, Julia³ (0009-0008-3118-9404); Stoehr, Svenja¹ (0009-0008-0287-2140); Twiss, Sean D.⁴ (0000-0002-1923-8874); Krüger, Oliver^{1,5} (0000-0002-8066-4936).

Affiliations:

1 Department of Animal Behaviour, Faculty of Biology, Bielefeld University, Konsequenz 45, 33615, Bielefeld, Germany.

2 Department of Business Administration and Economics, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany.

3 Biostatistics and Medical Biometry, Medical School OWL, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

4 Department of Biosciences, Durham University, Durham DH1 3LE, UK.

5 Joint Institute for Individualisation in a Changing Environment (JICE), Bielefeld University and University of Münster, Konsequenz 45, Bielefeld 33615, Germany.

***Corresponding Author:** Alexandra K. Childs; achilds17@gmail.com

†Current Address: Department of Animal Behaviour, Faculty of Biology, Bielefeld University, Konsequenz 45, 33615 Bielefeld, Germany

Methods: Model formulation of the MSCRM

The multi-state capture-recapture model (MSCRM) belongs to the class of hidden Markov models (HMMs) comprising two stochastic processes: the state process and the observation process. HMMs assume that (partially) hidden states determine the distribution of the observations at each point in time.

S1.1. State Process

The state process $\{S_t\}_{t=1,\dots,T}$ is defined as a Markov chain with s_t the state at season t . It aims to describe the weaning process by distinguishing the unweaned state "U" and the weaned state "W". Additionally, the absent state "A" accounts for events of death and emigration. This absorbing state ensures that dead or absent animals which cannot be observed anymore are not considered weaned and the weaning process is therefore only informed by animals that can be observed. The upper half of the directed graph in Fig. S1 shows the possible state transitions. Notably, once the unweaned state is left, an animal cannot transition back to it. The same applies for the weaned state. Therefore, the following transition probability matrix

$$\Gamma = \begin{pmatrix} \gamma_{UU} & \gamma_{UW} & \gamma_{UA} \\ 0 & \gamma_{WW} & \gamma_{WA} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \phi \cdot (1 - \nu) & \phi \cdot \nu & 1 - \phi \\ 0 & \phi & 1 - \phi \\ 0 & 0 & 1 \end{pmatrix}$$

is assumed with γ_{kl} denoting the probability of switching from state k to l for all $k, l \in \{U, W, A\}$.

In defining a survival probability, that is the probability that an animal continues to be observable, and the assumption that survival probability is independent of the current state, we could then rewrite the entries in Γ in terms of this survival probability ϕ and a weaning probability ν . In practical terms, by estimating only γ_{UU} and γ_{UW} and considering that $\gamma_{UU} + \gamma_{UW} + \gamma_{UA} = 1$, we can derive $\phi = \gamma_{UU} + \gamma_{UW}$ and $\nu = \frac{\gamma_{UW}}{\gamma_{UU} + \gamma_{UW}}$.

Covariates that were assumed to affect the weaning and survival probability (ν, ϕ) could be included in the state process leading to an inhomogeneous Markov chain $\{S_t\}$ with Γ_t and therefore (ν_t, ϕ_t) varying over time. In order to do so, the transition probabilities to be estimated are linked to linear predictors

$$\tau_{j,t} = Y_t \beta_j \text{ for } j = 1, 2 \text{ and } t = 1, \dots, T$$

given the covariate matrix Y_t and the regression coefficients β_j such that

$$\gamma_{UU,t} = \frac{\exp(\tau_{1,t})}{1 + \exp(\tau_{1,t}) + \exp(\tau_{2,t})} \text{ and } \gamma_{UW,t} = \frac{\exp(\tau_{2,t})}{1 + \exp(\tau_{1,t}) + \exp(\tau_{2,t})}.$$

S1.2. Observation process

The observation process $\{X_t\}_{t=1,\dots,T}$ depends on the underlying state process $\{S_t\}$ as depicted in the bottom half of Fig. S1. The observations $x_{t,q}$ correspond to being observed suckling (S), being observed but not suckling (N), recovered dead (D), or not being observed (0) for each field season t and each animal q .

We construct three state-dependent probability distributions with observation- and state-specific sighting probabilities as follows:

$$X|k \sim \text{Categorical}(q_{\cdot|k}), \text{ for } k = U, W, A$$

$$q_{\cdot|U} = (\pi, \rho, 0, 1 - \pi - \rho)$$

$$q_{\cdot|W} = (0, \lambda, 0, 1 - \lambda)$$

$$q_{\cdot|A} = (0, 0, \epsilon, 1 - \epsilon).$$

Given the unweaned state U , we obtain a categorical distribution where individuals who were observed suckling have probability π , observed doing something else have probability ρ and unobserved have probability $1 - \pi - \rho$. Given the weaned state W and the absent state A , the distributions are essentially Bernoulli with λ the probability of being observed given the weaned state and ϵ the probability of being recovered dead.

We collect variables that are assumed to affect the resighting probability per field season t and thereby the distribution of observations in a covariate matrix Z_t and build a linear predictor

$$\eta_{i,t} = Z_t \alpha_i \text{ for } i = 1, 2, 3, 4 \text{ and } t = 1, \dots, T$$

with $\alpha_{i,t}$ the regression coefficients.

The four parameters $(\rho, \pi, \lambda, \epsilon)$ are then time varying $(\rho_t, \pi_t, \lambda_t, \epsilon_t)$ by being linked to the linear predictors as follows:

$$\rho_t = \frac{\exp(\eta_{1,t})}{1 + \exp(\eta_{1,t}) + \exp(\eta_{2,t})}$$

$$\pi_t = \frac{\exp(\eta_{2,t})}{1 + \exp(\eta_{1,t}) + \exp(\eta_{2,t})}$$

$$\lambda_t = \frac{1}{1 + \exp(\eta_{3,t})}$$

$$\epsilon_t = \frac{1}{1 + \exp(\eta_{4,t})}.$$

S1.3. Parameter estimation and state decoding

By framing the model as HMM, we leveraged inferential tools like the forward algorithm to estimate the latent weaning process and the Viterbi algorithm to decode the most likely state sequence. Note that we constructed a partially hidden Markov model in the sense that a suckling observation "S" and a death recovery "D" inform directly about the state having generated this observation. We used the package “LaMa” in R (R Core Team 2024) for conveniently building the HMM likelihoods and maximized them directly by using the numerical optimizer “nlm”. Confidence intervals are calculated as quantiles of Monte Carlo samples based on the variance-covariance matrix of the estimators as obtained via the Hesse matrix provided by “nlm”.

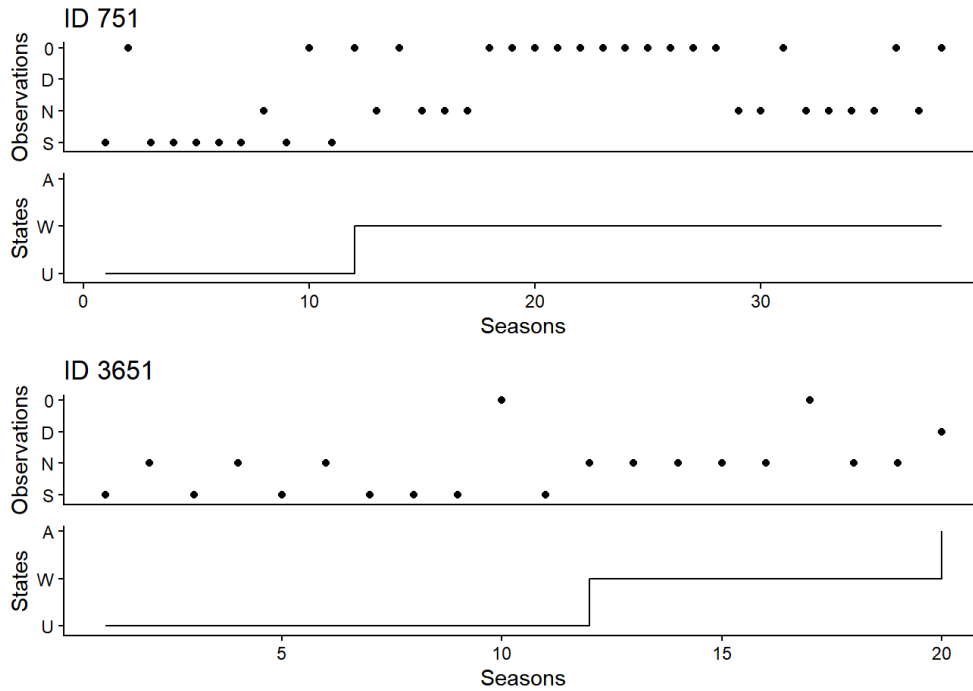


Fig. S1 Trajectories of two exemplary individuals' observations and corresponding states that are most likely (given model 1D) over time. Observations can be either observed suckling (S), observed, but not suckling (N), recovered dead (D), or not observed (0). States can be either unweaned (U), weaned (W), or absent (A).

Table S1 All packages used during data manipulation and analysis.

Package	Authors
gridExtra	Auguie 2017
lme4	Bates et al. 2015
Matrix	Bates, Maechler and Jagan 2025
lubridate	Grolemund and Wickham 2011
msm	Jackson 2011
LaMa	Koslik 2025
RTMB	Kristensen 2025
report	Makowski et al. 2023
patchwork	Pedersen 2025
nlme	Pinheiro, Bates and R Core Team 2025; Pinheiro and Bates 2000
R (base)	R Core Team 2025
ggplot2	Wickham 2016
readxl	Wickham and Bryan 2025
dplyr	Wickham et al. 2023
readr	Wickham, Hester and Bryan 2024
tidyr	Wickham, Vaughan and Girlich 2024
mgcv (Wood GAM)	Wood 2003, 2004, 2011, 2017; Wood, Pya and Säfken 2016
knitr	Xie 2014, 2015, 2025
zoo	Zeileis and Grothendieck 2005
kableExtra	Zhu 2024

Results

S2.1. Weaning age

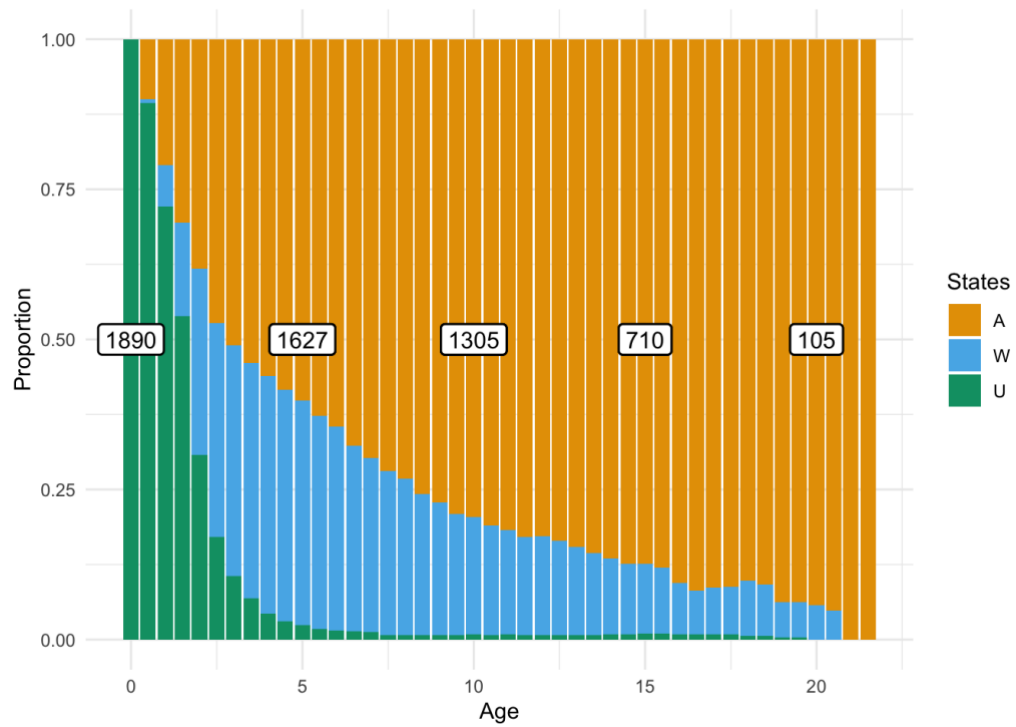


Fig. S2 Viterbi state proportions from MSCRM 1D (Table 1) showing the most likely state (unweaned [U], weaned [W], absent [A]) an individual is in at any age given its observational history. White-boxed numbers within the figure indicate the total number of possible observable animals at 5-year intervals.

Offspring effects

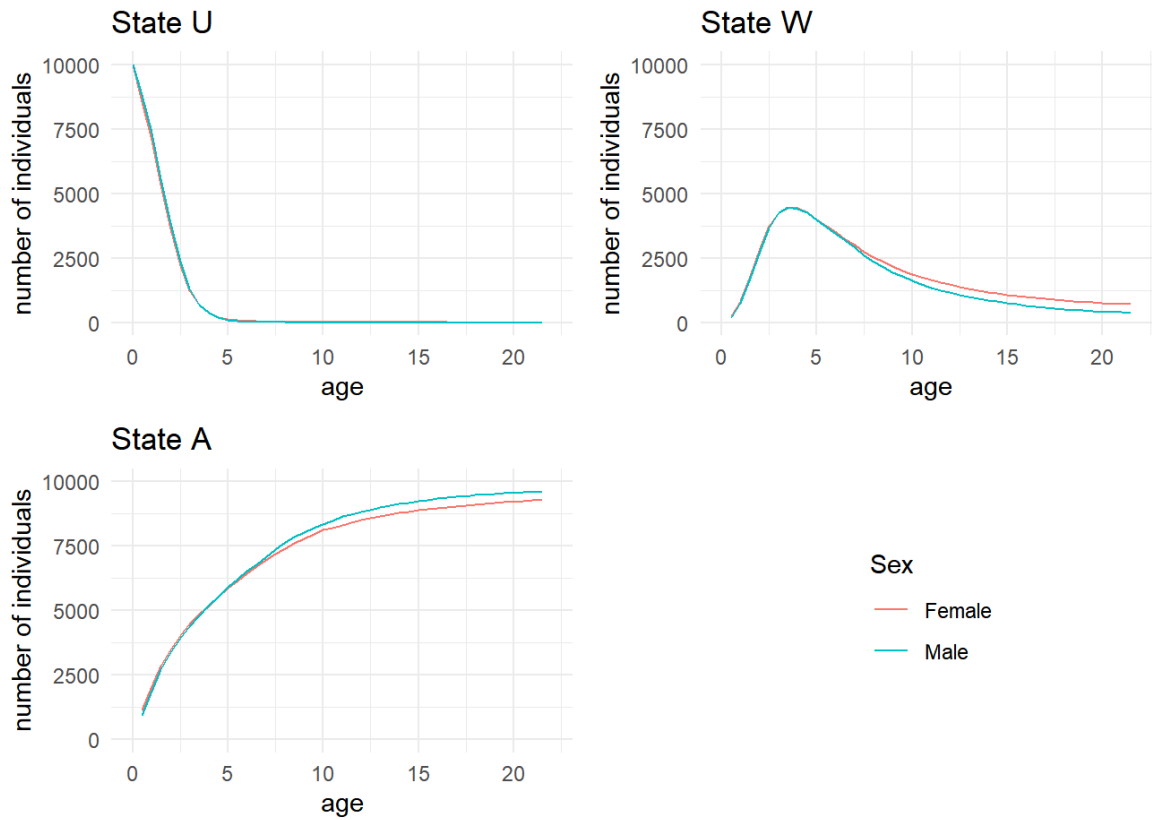


Fig. S3 A visual representation of the effect of offspring sex on the transition probability of each state, using simulated data from MSCRM 2. The y-axis represents the number of individuals in the simulated population and the x-axis the age of the offspring. Each state is shown as a line graph in a separate panel. The effect of the sex is illustrated by the difference between the blue (Male) and red (Female) lines.

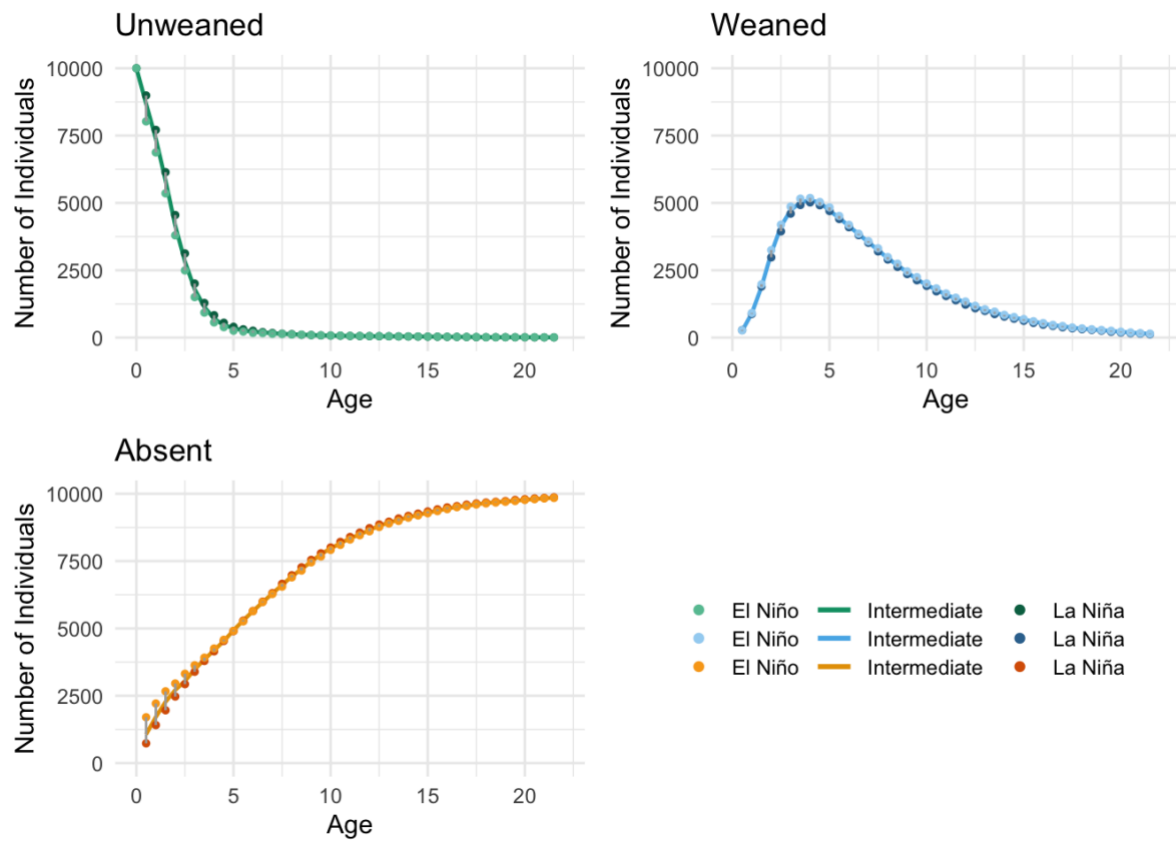


Fig. S4 ONI effect according to offspring age for all state behaviours: unweaned (U), weaned (W) and absent (A). Solid lines represent the number of individuals (n) in each state according to the simulated data based on MSCRM 3 during an Intermediate ONI category. Darker points represent the n during La Niña conditions and are overlain with paler points which represent the n during El Niño conditions.

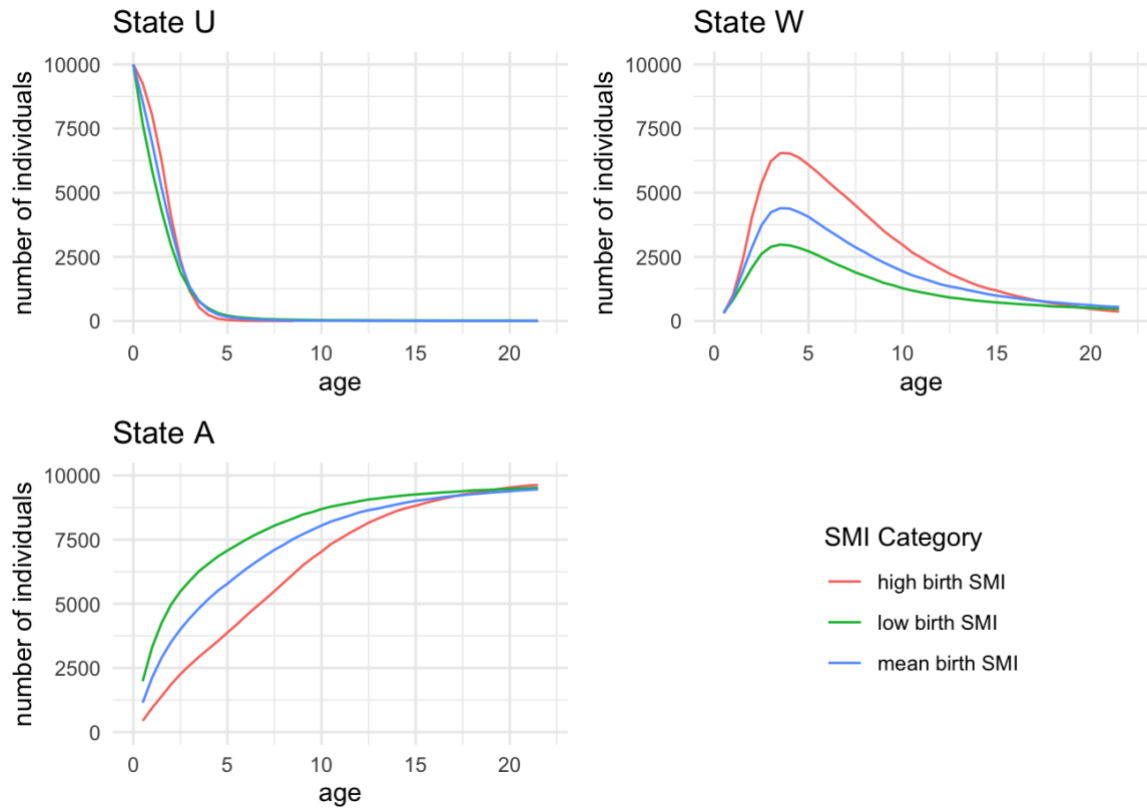


Fig. S5 Birth SMI effect - after Peig & Green (2009) - according to offspring age for all state behaviours: unweaned (U), weaned (W) and absent (A). Solid lines indicate the number of individuals in each state according to the simulated data based on MSCRM 4A, where red is low birth SMI (3.27), blue mean birth SMI (6.36) and green high birth SMI (11.20).

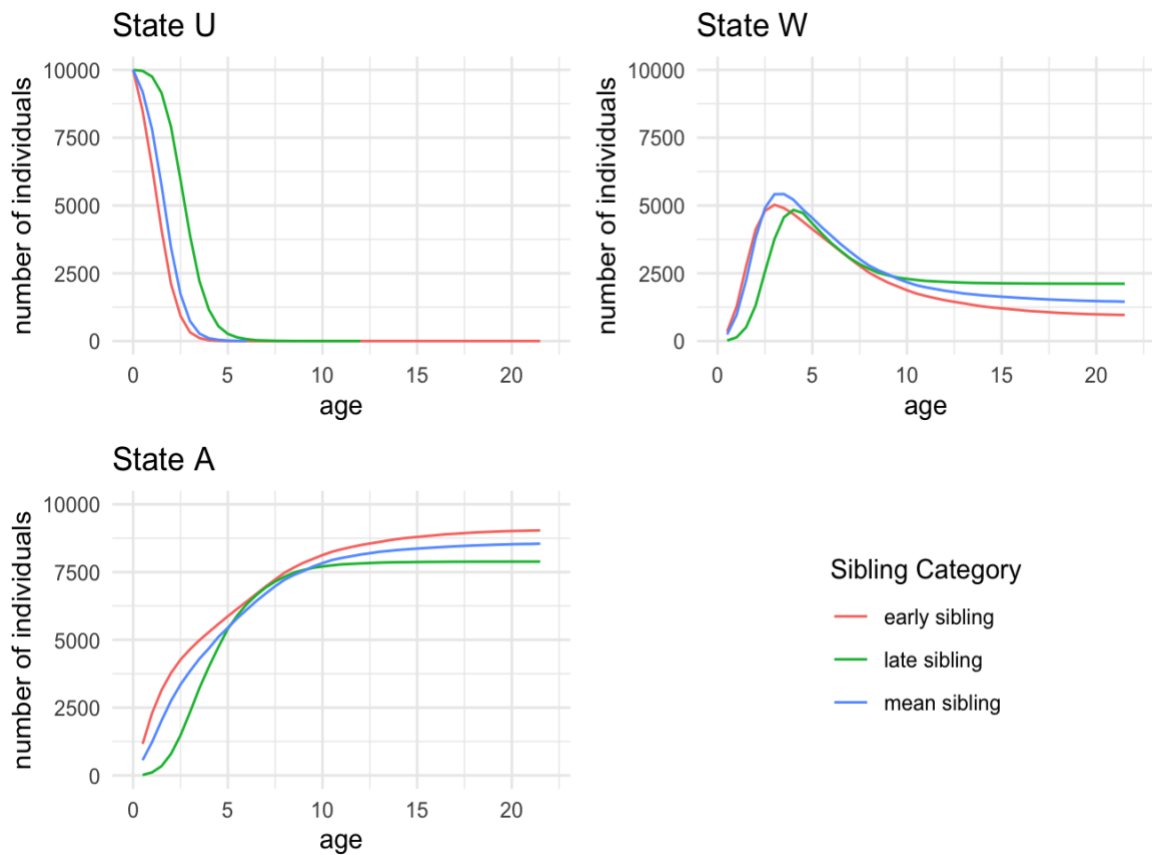


Fig. S6 Birth interval effect - calculated as time until the birth of the next sibling - according to offspring age for all state behaviours: unweaned (U), weaned (W) and absent (A). Solid lines indicate the number of individuals in each state according to the simulated data based on MSCRM 5A, where red is a short interval (early sibling [1 year]), blue is a mean interval (mean sibling [2 years]) and green a long interval (late sibling [7 years]).

Table S2 Percentage of unweaned offspring at 1 year intervals during the weaning window for the three sibling categories, extracted from the simulated state distribution of the ‘unweaned’ state based on MSCRM 5A.

Birth Interval	Age	Unweaned offspring (%)
Early Sibling (1 year)	1.5	41.0
	2.5	9.3
	3.5	1.1
	4.5	0.1
Mean Sibling (2 years)	1.5	57.4
	2.5	17.4
	3.5	2.8
	4.5	0.5
Late Sibling (7 years)	1.5	91.5
	2.5	59.6
	3.5	22.1
	4.5	5.5

S2.2. Supersucklers

Table S3 All fitted LRMs predicting the probability for an individual to become a supersuckler described by the predictors and random effect included. The number of parameters (np) and number of individuals (nind) for which the corresponding variables were available are used as a basis for comparison via the BIC. Variable notation $ONI_0 - ONI_4$ refers to the oscillation index at the given age of an individual, SMI_{10}, SMI_{90} to the SMI of an individual's first 10 or 30-90 days, *sibl* to time until sibling, m_{age} to the age of the mother at birth. Index refers to the steps in model selection. The index and BIC values in bold indicate the model that was preferred at each stage of model selection.

Index	random effect	predictors	BIC	nind	np
0A		-	488.66	766	1
0B		$sex, ONI_0, ONI_1, ONI_2, ONI_3, ONI_4$	509.75	766	7
1A	<i>cohort</i>	m_{age}	82.34	98	3
1B	<i>cohort</i>	-	77.93	98	2
2A	<i>cohort</i>	SMI_{10}	222.05	308	3
2B	<i>cohort</i>	SMI_{90}	221.86	308	3
2C	<i>cohort</i>	SMI_{10}, SMI_{90}	227.48	308	4
2D	<i>cohort</i>	-	216.39	308	2
3A	<i>cohort</i>	<i>sibl</i>	69.75	140	3
3B	<i>cohort</i>	-	65.47	140	2

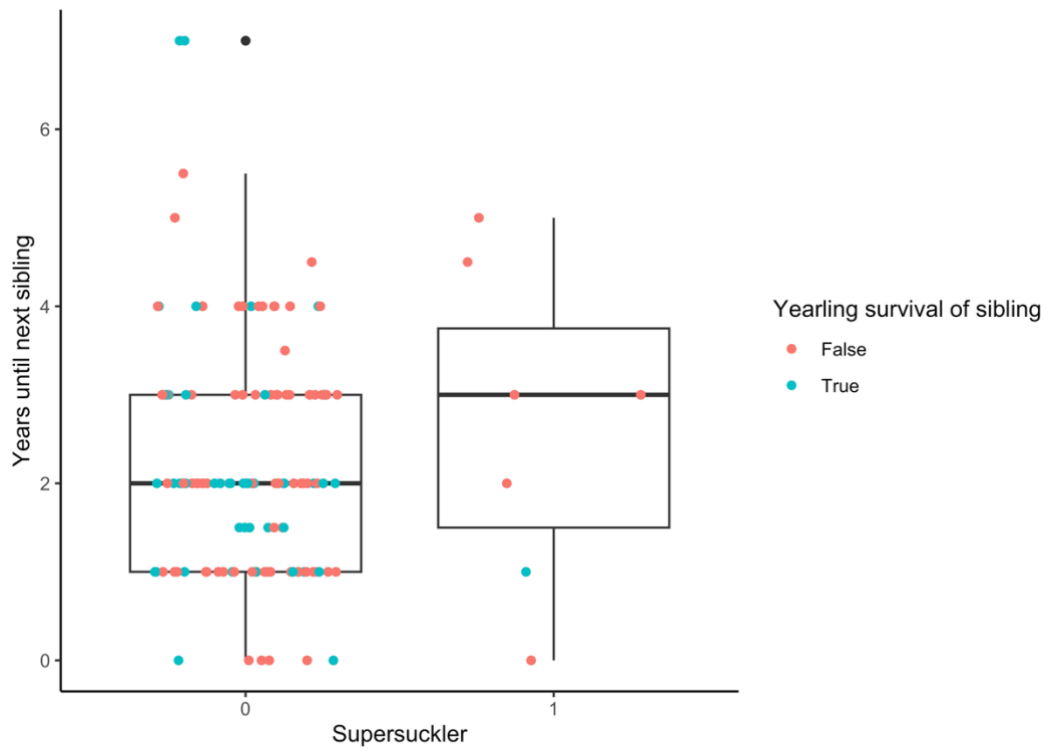


Fig. S7 A boxplot of the possible relationship between supersucklers and younger sibling survival to yearling age, where all younger siblings are plotted as jitter points according to the birth interval between them and the elder offspring and grouped according to the elder offsprings classification as a supersuckler (1, $N = 7$) or non-supersuckler (0, $N = 133$). Younger sibling survival to yearling age is denoted by the colour of their point where blue is 'survived' ($N = 51$) and red is 'did not survive' ($N = 89$).

Alternative Figure Descriptions

Fig. S1 Two sequences of observations and states each displaying the observations sequences for two exemplary individuals as well as the states according to Viterbi decoding based on MSCRM 1D.

Fig. S2 A stacked bar graph displaying the Viterbi state proportions from MSCRM Model 1D showing the proportion (y-axis) of individuals in each state (unweaned [green, U], weaned [blue, W], absent [yellow, A]) at a given age (x-axis). Age is binned by 6 months and the white-boxed numbers within the figure indicate the total number of possible observable animals at 5-year intervals.

Fig. S3 A visual representation of the effect of offspring sex on the transition probability of each state, using simulated data from MSCRM 2. The y-axis represents the number of individuals in the simulated population and the x-axis the age of the offspring. Each state is shown as a line graph in a separate panel. The effect of the sex is illustrated by the difference between the blue (Male) and red (Female) lines.

Fig. S4 A visual representation of the effect of ONI category on the transition probability of each state, using simulated data from MSCRM 3. The y-axis represents the number of individuals in the simulated population and the x-axis the age of the offspring. Each state is shown as a line graph in a separate panel. The effect each ONI category is separated into La Niña (pale points), Intermediate (solid line) and El Niño (dark points), illustrating the effect of this variable.

Fig. S5 A visual representation of the effect of offspring birth SMI on the transition probability of each state, using simulated data from MSCRM 4A. The y-axis represents the number of individuals in the simulated population and the x-axis the age of the offspring. Each state is shown as a line graph in a separate panel. The effect of birth SMI is separated into low (red), mean (blue) and high (green), clearly illustrating the effect of this variable.

Fig. S6 A visual representation of the effect of birth interval on the transition probability of each state, using simulated data from MSCRM 5A. The y-axis represents the number of individuals in the simulated population and the x-axis the age of the offspring. Each state is shown as a line graph in a separate panel. The effect of birth interval is separated into short interval (red), mean interval (blue) and long interval (green), clearly illustrating the effect of this variable.

Fig. S7 A box plot illustrating the possible relationship between offspring and their first (younger) sibling's survival to yearling age. Data points on the left-hand box plot represent offspring with known siblings who did not become supersucklers, those on the right-hand boxplot represent supersucklers with known siblings. Data points in blue indicate that the sibling survived to yearling age whereas red points indicate that the sibling died

before yearling age. The figure suggests a tentative relationship between supersuckling and younger siblings dying before yearling age but mostly highlights the need to increase the sample size for supersucklers.

References Cited Only in Supplementary Materials

- Auguie B (2017) gridExtra: Miscellaneous functions for “grid” graphics, version 2.3. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.gridExtra>
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bates D, Maechler M, Jagan M (2025) Matrix: Sparse and dense matrix classes and methods, version 1.7-3. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.Matrix>
- Grolemund G, Wickham H (2011) Dates and times made easy with lubridate. J Stat Softw 40:1–25. <https://doi.org/10.18637/jss.v040.i03>
- Jackson CH (2011) Multi-state models for panel data: The msm package for R. J Stat Softw 38:1–29. <https://doi.org/10.18637/jss.v038.i08>
- Kristensen K (2025) RTMB: R bindings for TMB, version 1.7. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.RTMB>
- Makowski D, Lüdtke D, Patil I, Thériault R, Ben-Shachar MS, Wiernik BM (2023) Automated results reporting as a practical tool to improve reproducibility and methodological best practices adoption. CRAN. <https://doi.org/10.32614/CRAN.package.report>
- Pedersen TL (2025) patchwork: The composer of plots, version 1.3.1. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.patchwork>
- Peig J, Green AJ (2009) New perspectives for estimating body condition from mass/length data: The scaled mass index as an alternative method. Oikos 118:1883–1891. <https://doi.org/10.1111/j.1600-0706.2009.17643.x>
- Pinheiro J, Bates D, R Core Team (2025) nlme: Linear and nonlinear mixed effects models, version 3.1-168. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.nlme>
- Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. Springer, New York. <https://doi.org/10.1007/b98882>
- R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer, New York. <https://ggplot2.tidyverse.org>
- Wickham H, Bryan J (2025) readxl: Read Excel files, version 1.4.5. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.readxl>
- Wickham H, François R, Henry L, Müller K, Vaughan D (2023) dplyr: A grammar of data manipulation, version 1.1.4. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.dplyr>
- Wickham H, Hester J, Bryan J (2024) readr: Read rectangular text data, version 2.1.5. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.readr>
- Wickham H, Vaughan D, Girlich M (2024) tidyr: Tidy messy data, version 1.3.1. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.tidyr>

Wood SN (2003) Thin-plate regression splines. *J R Stat Soc B* 65:95–114. <https://doi.org/10.1111/1467-9868.00374>

Wood SN (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J Am Stat Assoc* 99:673–686. <https://doi.org/10.1198/016214504000000980>

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc B* 73:3–36. <https://doi.org/10.1111/j.1467-9868.2010.00749.x>

Wood SN (2017) Generalized additive models: An introduction with R, 2nd edn. Chapman and Hall/CRC, Boca Raton

Wood SN, Pya N, Säfken B (2016) Smoothing parameter and model selection for general smooth models (with discussion). *J Am Stat Assoc* 111:1548–1575. <https://doi.org/10.1080/01621459.2016.1180986>

Xie Y (2014) knitr: A comprehensive tool for reproducible research. In: Stodden V, Leisch F, Peng RD (eds) *Implementing reproducible computational research*. Chapman and Hall/CRC, Boca Raton, pp 1–24

Xie Y (2015) *Dynamic documents with R and knitr*, 2nd edn. Chapman and Hall/CRC, Boca Raton

Xie Y (2025) knitr: A general-purpose package for dynamic report generation in R, version 1.50. R package. <https://yihui.org/knitr/>

Zeileis A, Grothendieck G (2005) zoo: S3 infrastructure for regular and irregular time series. *J Stat Softw* 14:1–27. <https://doi.org/10.18637/jss.v014.i06>

Zhu H (2024) kableExtra: Construct complex tables with “kable” and pipe syntax, version 1.4.0. R package. Comprehensive R Archive Network (CRAN). <https://doi.org/10.32614/CRAN.package.kableExtra>