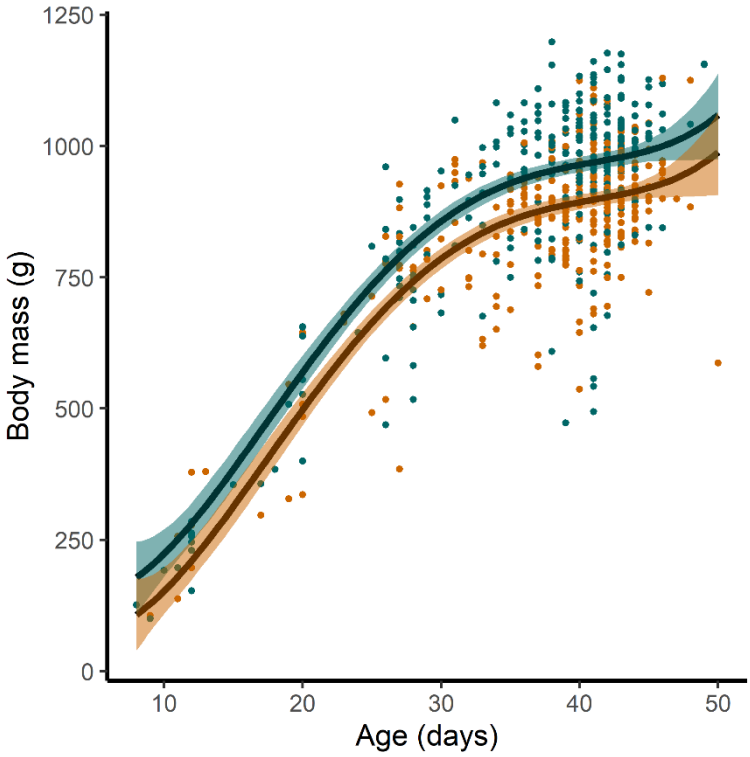

Fig. S1. Body mass in relation to age for female (blue) and male (orange) red kite nestlings. Shown are model predictions (lines), 95% Credible Intervals (shaded), and raw data points (dots).

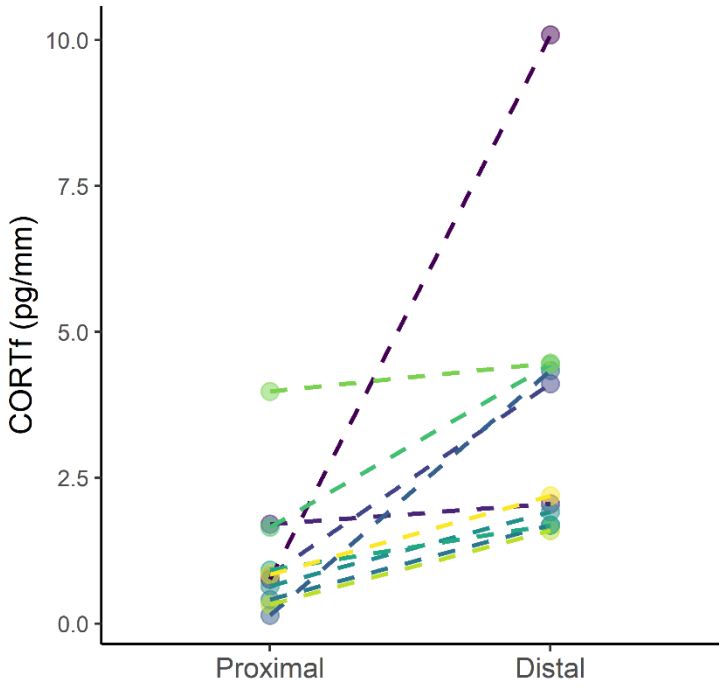

Fig. S2. CORTf values of proximal and distal feather segments (n = 11 individuals).

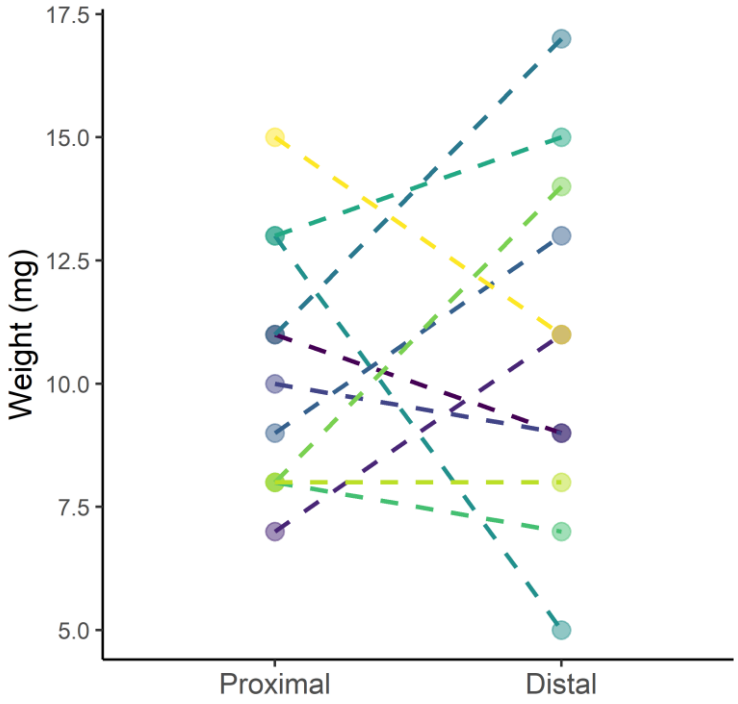

Fig. S3. Weight in milligrams of proximal and distal feather segments (n = 11 individuals).

Table S1. Summary of the Linear Mixed Model investigating age and sex differences in nestling body mass. Effects with 95 % Credible Interval (CrI) excluding zero are shown in bold. Residuals from this model were used as a proxy for body condition at each sampling event.

Fixed effects	*Estimate*	*95 % CrI*
(Intercept)	**916.41**	**902.90 – 929.92**
Age	**3296.74**	**3114.25 – 3479.23**
Age^2	**-1255.34**	**-1423.18 – -1087.50**
Age^3	76.19	-93.01 – 245.39
Age^4	**248.05**	**88.51 – 407.60**
Sex [m]	**-71.14**	**-89.91 – -52.36**
Random Effect	*s.d.*	
Individual ID	87.08	

Supplementary Information

Supplementary Materials and Methods

CORTf enzyme immunoassay

CORTf concentration was measured using an enzyme immunoassay (EIA) (Munro & Stabenfeldt, 1984; Müller et al., 2006). Samples were re-suspended in phosphate buffer and incubated in presence of a CORT-antibody (1:8000) (Chemicon; cross-reactivity: 11-dehydrocorticosterone 0.35%, Progesterone 0.004%, 18- hydroxydeoxycorticosterone 0.01%, Cortisol 0.12%, 18-hydroxycorticosterone 0.02% and Aldosterone 0.06%). As enzyme label we used a horse-radish peroxidase complex linked to CORT (1:400000) and 2,2'Azino-bis(3-ethylbenzo-thiazoline-6-sulfonicacid) diammonium salt (ABTS) as substrate. The concentration of CORT was calculated using a standard curve run in duplicate on each plate and expressed as pg CORT/mm feather (following Bortolotti et al., 2008; Jenni-Eiermann et al., 2015). Plates were read with a BioTek ELX808IU spectrophotometer at 405 nm (reference wavelength 655 nm) after 1.5 hours after substrate addition.

Validation of stress measures

CORTf

To assess the validity of using one feather as a representation of the individual state, we tested the inter-individual repeatability by comparing two separate feathers of 30 individuals. Repeatability was assessed using intra-class correlation coefficients (ICC), calculated with the package ICC in R (Wolak et al., 2012) using the variance components from a one-way ANOVA. The ICC for feathers from the same individual was 0.84 (confidence interval CI=0.69-0.92), suggesting high repeatability of single feathers and, thus, supporting the adequacy of the method. In a recent study on Alpine Swifts *Tachymarptis melba* we validated the extraction efficiency of the EIA by spiking 21 feathers of known concentrations. The mean recovery was 93.37% ± 2.33 (mean ± s.e.m., range 74.06–111.88) (Jenni-Eiermann et al., 2022).

H/L

We verified two potential methodological issues: First, we assessed whether increased

counting error occurred at high leucocyte densities. We assigned 80 samples to a leucocyte density category (low, medium or high). We then randomly chose 3 samples per category and assessed H/L from 100, 200, 300 and 400 leucocytes. There was no pattern in the coefficient of variation along density categories (*Kruskal-Wallis rank sum test, p-value: 0.332*) nor high variation between H/L calculated in different numbers of leucocytes (coefficient of variation: CV = 0.05 – 0.126). Second, we tested whether a sample of 100 leucocytes yielded an adequate repeatability. For this, we randomly selected 3 smears, measured H/L in 100, 200, 300 and 400 leucocytes and repeated the count 5 times per slide. Both repetition of the same slide and counts in different leucocyte amounts yielded a CV lower than 0.11 indicating that H/L calculated using 100 leucocytes was an adequate, repeatable measure (Pendl 2018, *pers. comm.; Lentfer et al., 2015)*.

Influence of disturbance on H/L ratio

We checked whether the time between beginning of climbing and blood sample influenced the H/L ratio by fitting a generalised mixed model with binomial data distribution. The response variable H/L was taken into the model as a two-part vector of number of heterophils (H) and lymphocytes (L). We added the time difference between blood sampling and beginning of the climbing event as explanatory variable and scaled to standard deviation (s.d.) = 1 and centred prior to modelling. Brood ID was included as a random intercept. There was no significant effect of the time difference on the H/L ratio (effect size = -0.016, 95% CrI = -0.017 – 0.039).

Measuring body condition

We defined body condition as the residual of a linear mixed model (LMM; package lme4; Bates et al., 2015) with body mass as response and age as explanatory variable, while controlling for sex-specific growth differences. Bird ID was included as random effect to account for repeated measurements of the same individual (see Table S1, Fig. S1). Age was obtained either by recording the precise hatching dates through cameras or by extrapolation from a growth curve that considered not only feather length (P8), but also hatching rank and was based on a separate data set of nestling measurements (Nägeli et al., 2021). The relationship between body mass and age was not linear during development as mass gain or feather growth, respectively, might be prioritized at times. AIC favoured a model that included 4 polynomial terms of age, and this was, furthermore, supported by fitting the same data in an additive

model using the package "BAMLSS"(Umlauf et al., 2018), using a smoother for the age covariate. The posterior distribution of the smoothed age effect showed a polynomial relationship of degree four, as suggested by the AIC of the LMM (Table S1; Fig. S1).

Analysis of the feather segments

Lattin et al. (2011) reported for the first time a non-linear negative relationship between feather sample weight and CORTf, which was originally thought to derive from changes in methanol-based extraction efficiency with samples of different weights ("the small sample artefact"). However, Berks et al. (2016) showed that this negative relationship persisted even when correcting for the methanol:weight ratio, suggesting that instead it may be due to significant cross-reactivity to other metabolites in low weight samples. We separately analysed red kite feather segments and investigated the existence of this small sample artefact among lighter-weighted feather samples. Distal segments had higher CORTf concentration (Fig. S2), but their weight was not lower than the proximal ones (two-sided, paired *Wilcoxon signed-rank test, p = 0.64;* Fig. S3). Further, no relationship between feather mass and CORTf was detected among all samples (Pearson's cor = 0.06, p = 0.11). This indicates that the higher CORTf concentrations in distal segments are not the result of the small sample artefact.

References

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Berk, S. A., McGettrick, J. R., Hansen, W. K., & Breuner, C. W. (2016). Methodological considerations for measuring glucocorticoid metabolites in feathers. *Conservation Physiology*, *4*(1). https://doi.org/10.1093/conphys/cow020

Bortolotti, G. R., Marchant, T. A., Blas, J., & German, T. (2008). Corticosterone in feathers is a long-term, integrated measure of avian stress physiology. *Functional Ecology*, *22*(3), 494–500. https://doi.org/10.1111/j.1365-2435.2008.01387.x

Jenni-Eiermann, S., Helfenstein, F., Vallat, A., Glauser, G., & Jenni, L. (2015). Corticosterone: Effects on feather quality and deposition into feathers. *Methods in Ecology and Evolution*, *6*(2), 237–246. https://doi.org/10.1111/2041-210X.12314

Jenni-Eiermann, S., Marin, J. O., & Bize, P. (2022). Variation in feather corticosterone levels in Alpine swift nestlings provides support for the hypo-responsive hypothesis. *General and Comparative Endocrinology*, *316*, 113946.

Lattin, C. R., Reed, J. M., DesRochers, D. W., & Romero, L. M. (2011). Elevated corticosterone in feathers correlates with corticosterone-induced decreased feather quality: a validation study. *Journal of Avian Biology*, *42*(3), 247–252. https://doi.org/10.1111/j.1600-048X.2010.05310.x

Lentfer, T. L., Pendl, H., Gebhardt-Henrich, S. G., Fröhlich, E. K. F., & Von Borell, E. (2015). H/L ratio as a measurement of stress in laying hens – methodology and reliability. *British Poultry Science*, *56*(2), 157–163. https://doi.org/10.1080/00071668.2015.1008993

Müller, C., Jenni-Eiermann, S., Blondel, J., Perret, P., Caro, S. P., Lambrechts, M., & Jenni, L. (2006). Effect of human presence and handling on circulating corticosterone levels in breeding blue tits (Parus caeruleus). *General and Comparative Endocrinology*, *148*(2), 163–171. https://doi.org/10.1016/j.ygcen.2006.02.012

Munro, C., & Stabenfeldt, G. (1984). Development of a microtitre plate enzyme immunoassay for the determination of progesterone. *Journal of Endocrinology*, *101*(1), 41-49.

Nägeli, M., Scherler, P., Witczak, S., Catitti, B., Aebischer, A., van Bergen, V., Kormann, U., & Grüebler, M. U. (2021). Weather and food availability additively affect reproductive output in an expanding raptor population. *Oecologia*, *1*, 3. https://doi.org/10.1007/s00442-021-05076-6

Umlauf, N., Klein, N., & Zeileis, A. (2018). BAMLSS: Bayesian Additive Models for Location, Scale, and Shape (and Beyond). *Journal of Computational and Graphical Statistics*, *27*(3), 612–627. https://doi.org/10.1080/10618600.2017.1407325

Wolak, M. E., Fairbairn, D. J., & Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, *3*(1), 129–137. https://doi.org/10.1111/j.2041-210X.2011.00125.x