**Supplementary Material for**

# "Extreme hot weather has stronger impacts on avian reproduction in forests than in cities" by

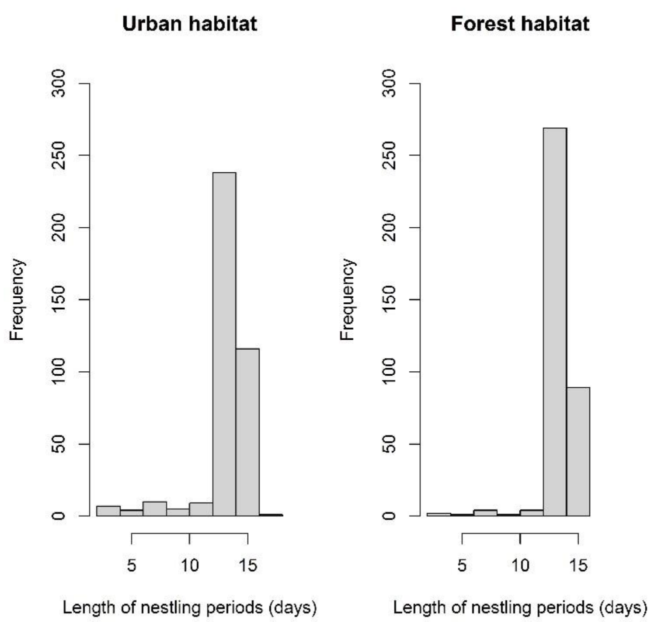
Ivett Pipoly, Bálint Preiszner, Krisztina Sándor, Csenge Sinkovics, Gábor Seress, Ernő Vincze, Veronika Bókony, András Liker

**correnspponding author**: Ivett Pipoly, pipoly.ivett@mk.uni-pannon.hu, +3688 624773, and András Liker, aliker@almos.uni-pannon.hu, University of Pannonia, H-8200, Veszprém, Egyetem street 10. Building N 14, Hungary

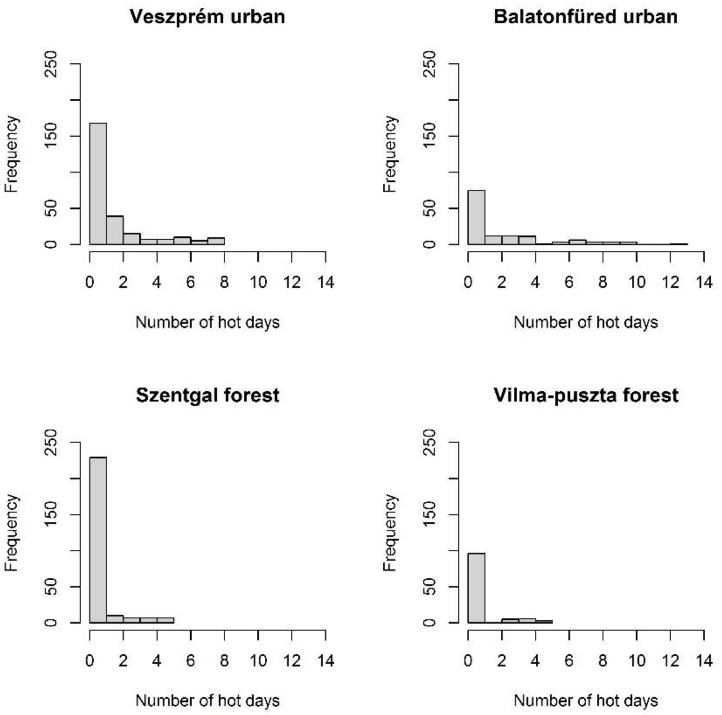
## 1. Data distributions

The distribution of length of nestling periods was similar in urban and forest populations (Fig. S1; range was 3-18 days in urban, and 3-16 days in forest habitat). The distribution of number of hot days was skewed in all study populations, and varied with habitat type (Fig. S2), as there were more number of hot days detected in urban than forest sites. However, the number of hot days covered a relatively large part of some nestling periods (Fig. S3), as the proportion of hot days during nestling development (i.e. number of hot days / length of nestling period) ranged between 0.00 and 0.8667 for urban, and between 0.00 and 0.3571 for forest nestlings.

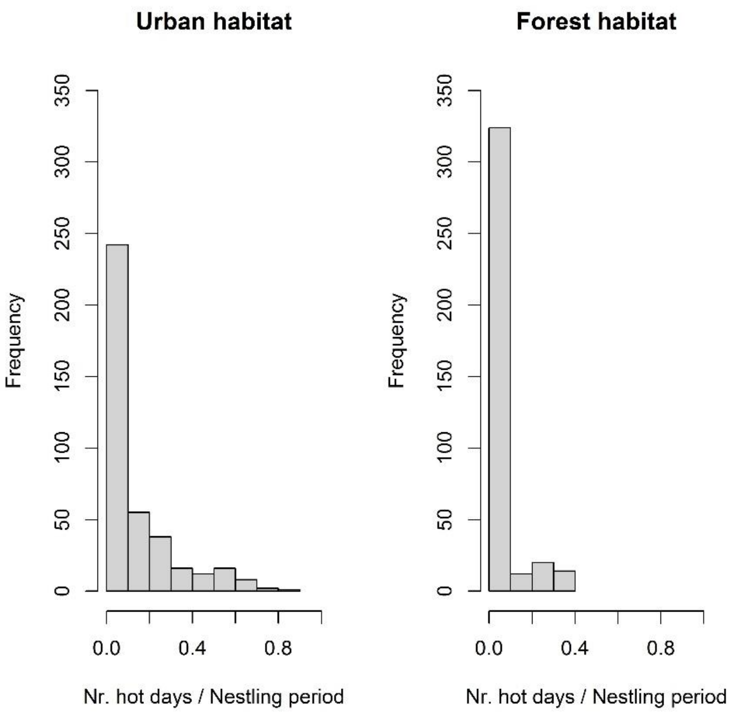
**Supplementary Figure 1.** Distribution of nestling period length (in days), including those broods where no nestling reached the age of 14-16 days. N=390 nestling periods for urban, and N=370 nestling periods for forest populations.



**Supplementary Figure 2.** Distribution of number of hot days in each of our study sites.




**Supplementary Figure 3**. Distribution of proportion of hot days during nestling development.



**Supplementary Table 1. Testing the effect of pair ID as random term.** A) in the simple models, and B) in the supported multi-predictor models, with likelihood ratio tests. △AIC values shows the difference between the model without pairID and the model containing pair ID as random term (i.e. positive △AIC value mean lower AIC value for the model with pair ID). All other model parameter is the same in the models as described in the main text. For nestling mass, results of two supported multi-predictor models are separated with "/".

| | **A) Simple model** | | **B) Multi-predictor model** | |
|---|---|---|---|---|
| | **ΔAIC** | **p** | **ΔAIC** | **p** |
| **Nestling mass** | 2.18 | 0.041 | 2.30 / 1.67 | 0.038 / 0.057 |
| **Nestling tarsus length** | -1.99 | 0.911 | -1.02 | 0.322 |
| **Nestling mortality** | -2.00 | 0.998 | -2.00 | 0.993 |

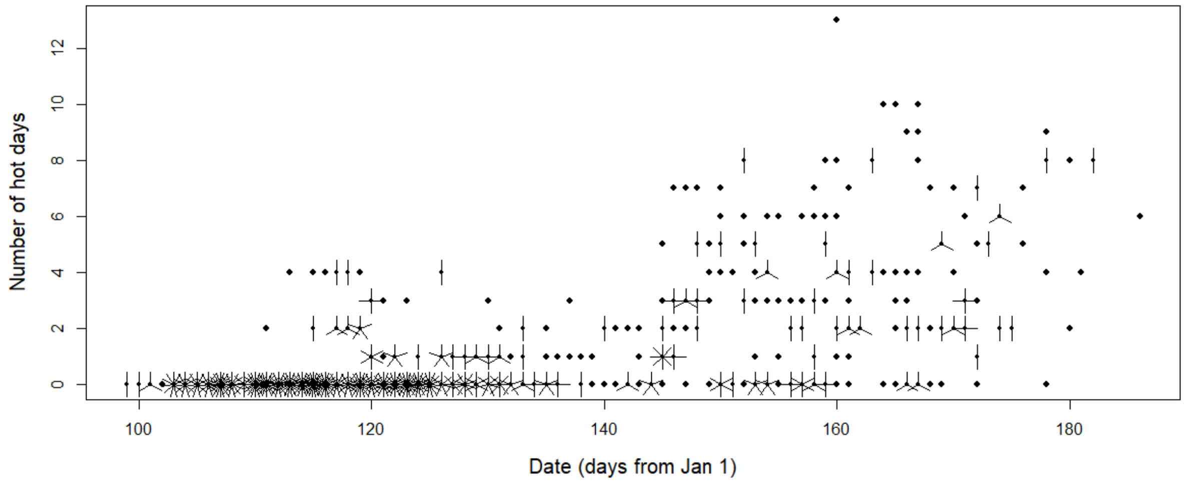
## 2. Seasonality

The number of hot days increases over the breeding season (Fig. S4). We included broods from the whole breeding season from six years, 70% of which were the first annual broods of the pairs and 30% were second annual broods. Naturally, both temperature and date can differ between first and second annual broods because they are in different parts of the breeding season, and also there are several possible resources or reproductive traits that may change seasonally, and these seasonal changes may differ between different types of habitats. Here we explain how we treated seasonality in our analyses.

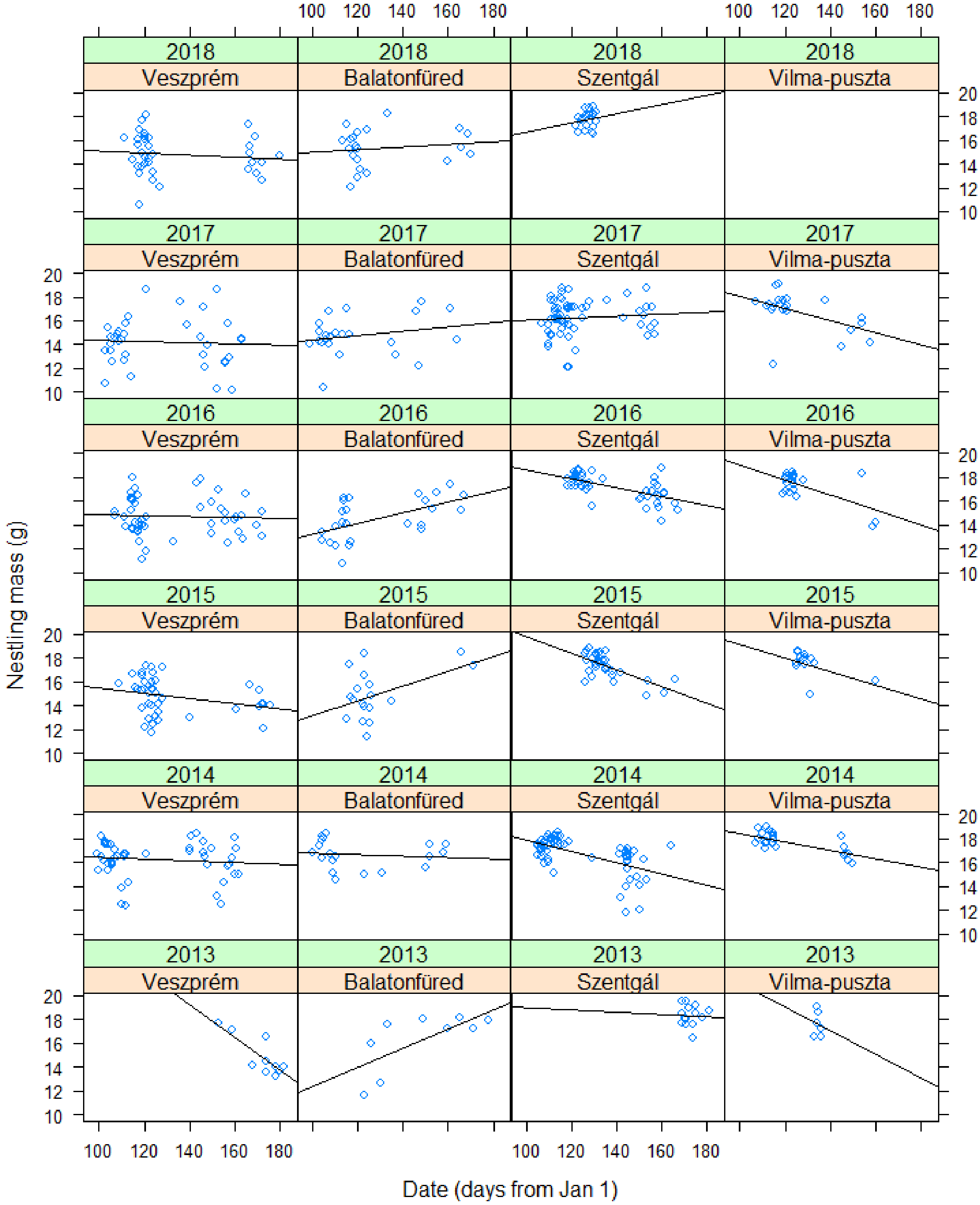
First, graphical inspection of the data showed that the seasonal difference in nestling size and mortality between first and second annual broods is well described with the linear effect of hatching date (Fig. S5-S7), and we did not detect any consistent pattern in the relationships between nestling size and date within first broods or within the second broods (Fig. S5-S7). For some sites in some years, the graphs suggested non-linear relationships (Fig. S5-S7), thus, we added the quadratic term of hatching date in the multi-predictor models. Based on these graphs, we decided to use hatching date only and not to include first *versus* second brood as an additional predictor in our analyses, as this would have increased multi-collinearity severely without considerable extra explanatory power.

Furthermore, we did not include the interaction between date and study site in any of our models, for the following reasons. First, because of the strong multi-collinearity between the number of hot days and date (Fig. S4), the effects of heat-related and non-heat-related seasonal effects cannot be separated in our analyses. Second, this very high multi-collinearity caused model-fitting algorithms to fail in some of the models containing the interactions of study site with both date and the number of hot days.

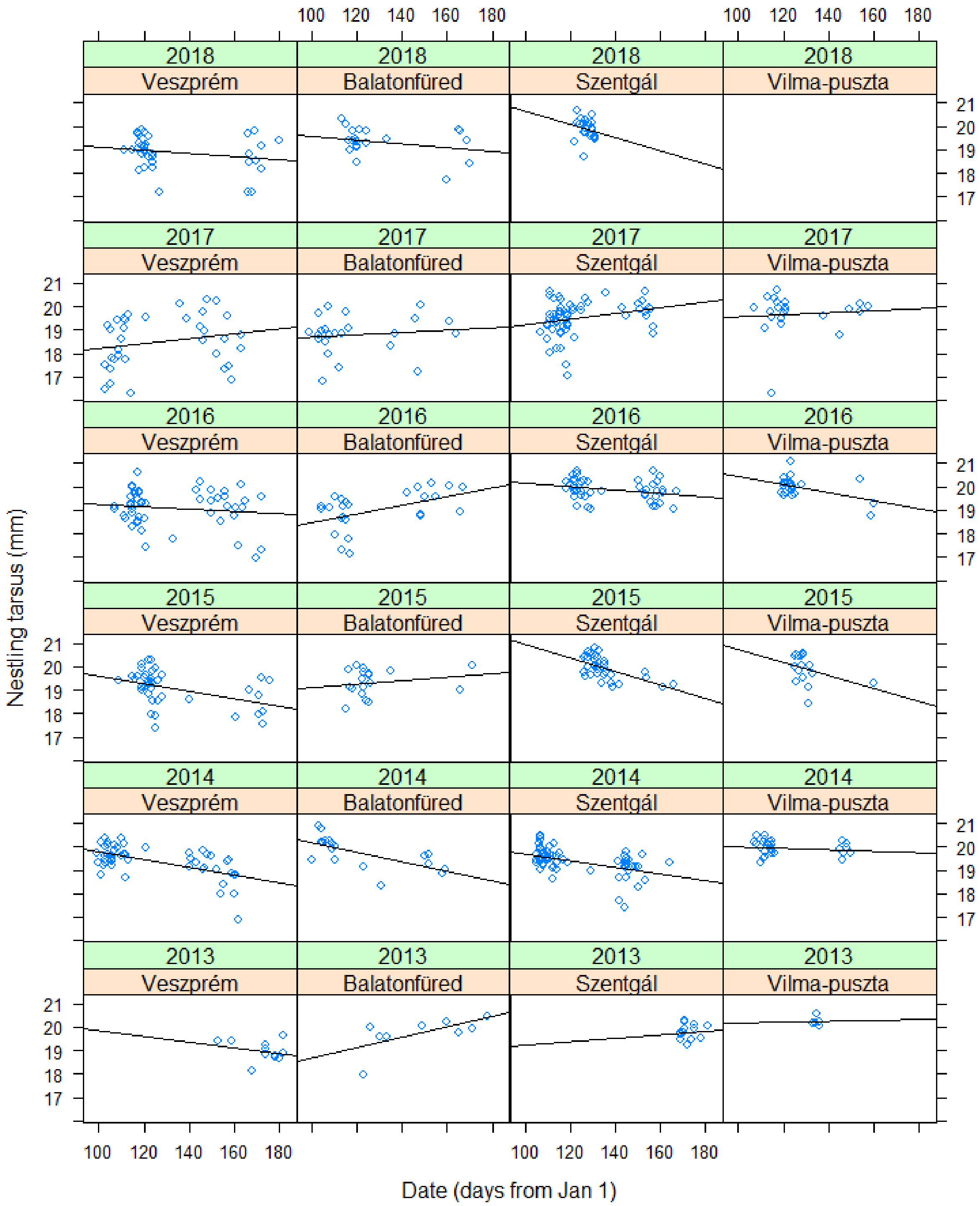
**Supplementary Figure 4.** Seasonal change in the number of hot days during great tit nestling periods.



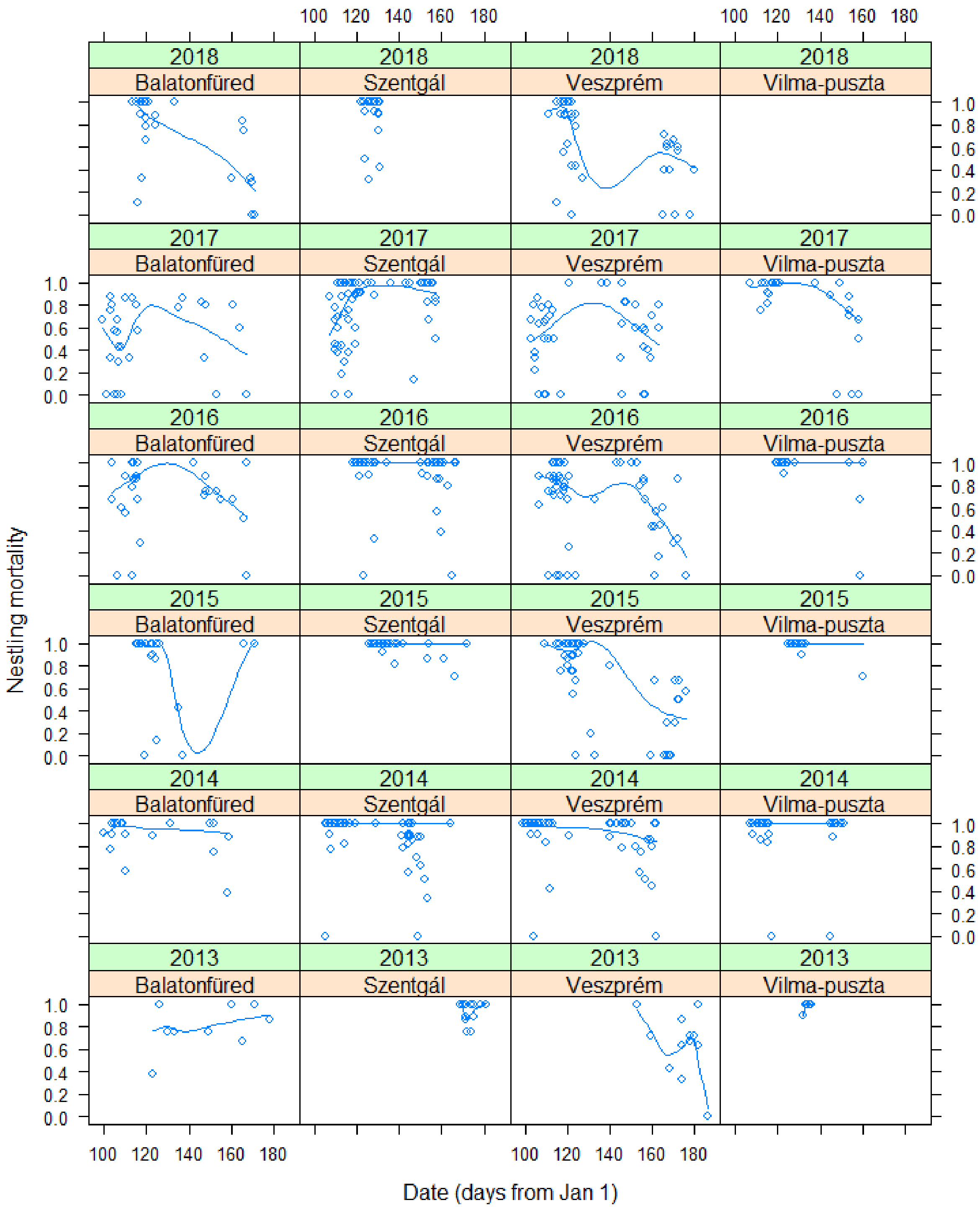
**Supplementary Figure 5.** Relationship between average nestling mass and hatching date of broods, separately in each study site (rows) and in each study year (columns). The 2018 Vilma-puszta data are missing due to a permanent technical failure by the weather station in that year.



**Supplementary Figure 6.** Relationship between average nestling tarsus length and hatching date of broods, separately in each study site (rows) and in each study year (columns). The 2018 Vilma-puszta data are missing due to a permanent technical failure by the weather station in that year.



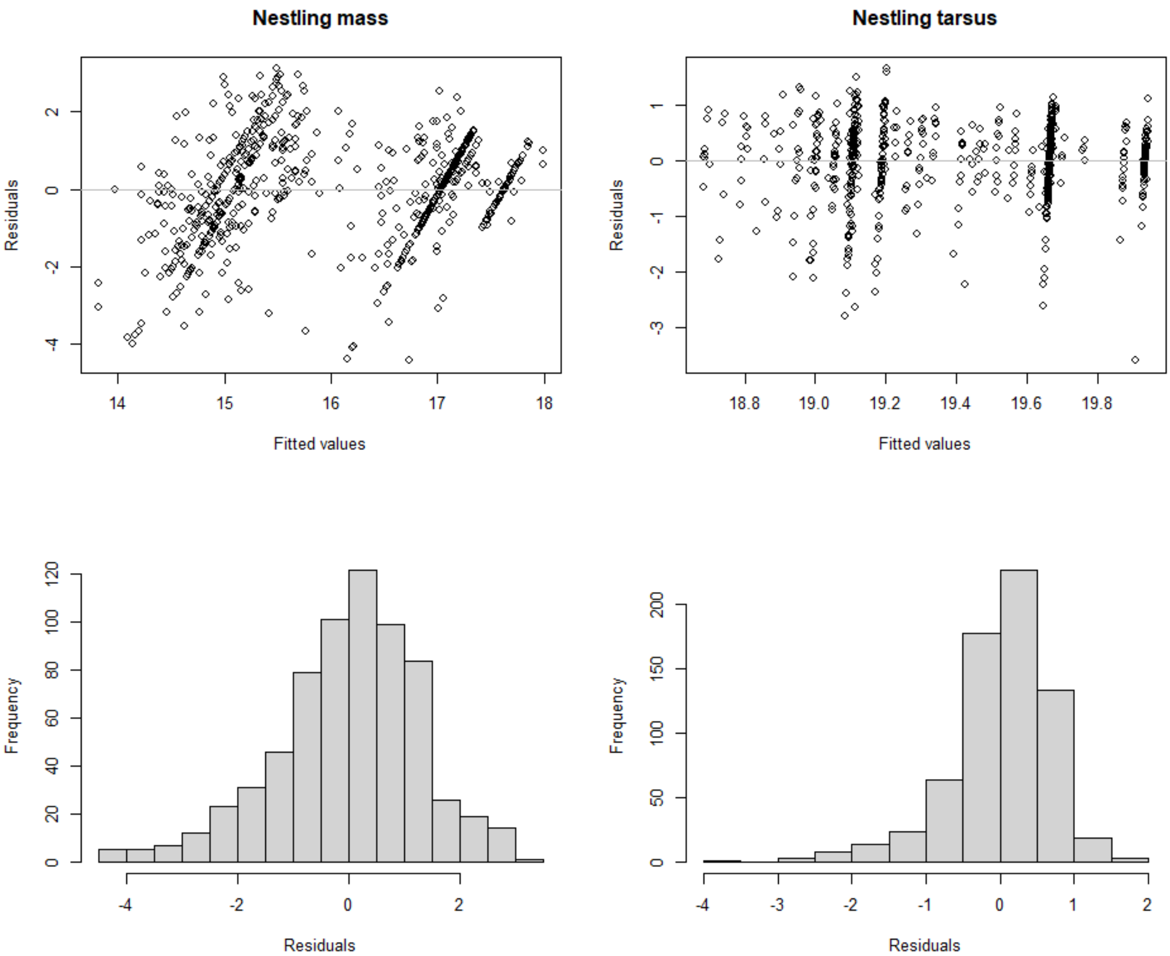
**Supplementary Figure 7.** Relationship between nestling mortality and hatching date of broods, separately in each study site (rows) and in each study year (columns). The 2018 Vilma-puszta data are missing due to a permanent technical failure by the weather station in that year.



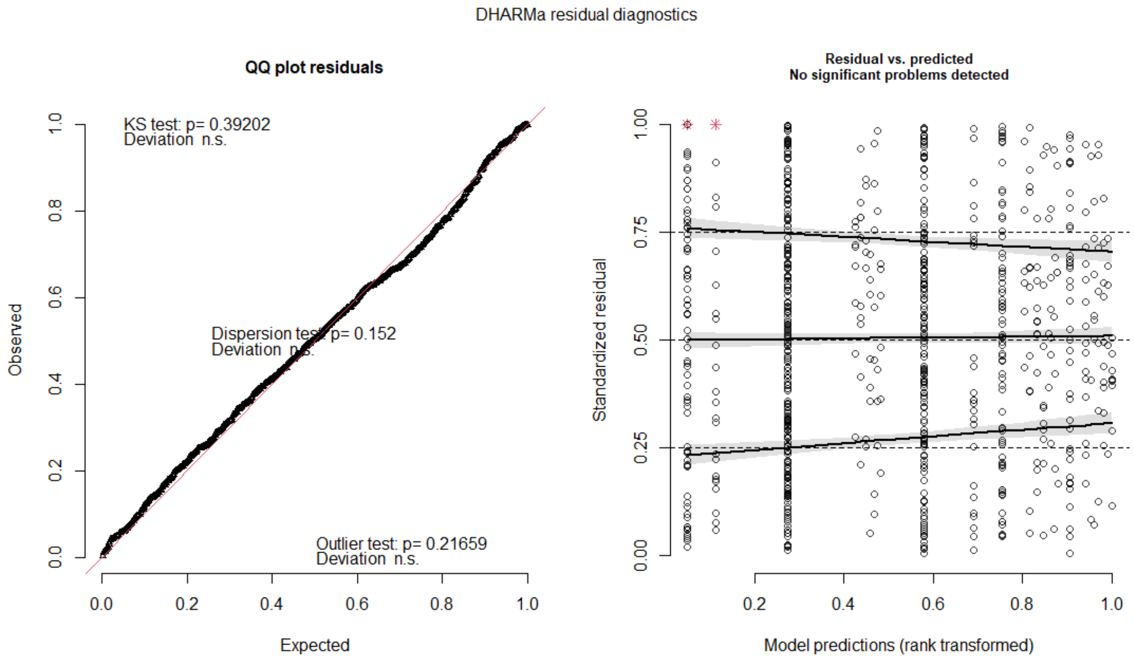
### 3. Model diagnostics

Diagnostic plots for LMM models with nestling size as response variable (average nestling mass and tarsus length of broods) containing number of hot days, study sites and their two-way interaction as predictors showed that no serious deviations from homoscedasticity, normality, and linearity are present in these models (Fig. S8). There were no influential outlier points, as Cook's distance values were low (< 0.08). Similarly, the GLMM model for nestling mortality containing the same predictors had a good model fit according to model diagnostics plots (Fig. S9).

**Supplementary Figure 8**. Diagnostic plots for LMM models with average body mass of nestlings (left) and average tarsus length of nestlings (right) as response variables.



**Supplementary Figure 9.** Diagnostic plot for GLMM model with nestling mortality as response variable. DHARMa R package was used for checking model diagnostics. The dispersion parameter in this model is not significantly different from 1 (estimated as 0.896; p = 0.152).



## 4. Multi-collinearity in multi-predictor models

**Supplementary Table 2. Variance inflation factors (VIF) for predictors in the full and the supported multi-predictor models in AICc-based model selection, for each response variable**. VIF values were calculated with *vif* function of "car" R package as $(GVIF^{(1/(2*Df))})^2$, to accurately present multi-collinearity for categorical variables as well (based on Fox & Monette, 1992, *Journal of the American Statistical Association*, **87**:417, 178-183). VIF values > 2.00 suggest multi-collinearity for a given predictor. Predictors were standardised in these models. Date2 represents the quadratic term of hatching date. Values for first and second supported models for nestling mass are separated with "/".

| Predictors | Nestling mass | | Nestling tarsus length | | Nestling mortality | |
|---|---|---|---|---|---|---|
| | VIF | Df | VIF | Df | VIF | Df |
| *Full model* | | | | | | |
| Nr. hot days | 3.21 | 1 | 3.21 | 1 | 3.11 | 1 |
| Study site | 1.38 | 3 | 1.38 | 3 | 1.24 | 3 |
| Average temperature | 7.41 | 1 | 7.40 | 1 | 7.61 | 1 |
| Year | 1.31 | 5 | 1.32 | 5 | 1.23 | 5 |
| Date | 370.53 | 1 | 369.64 | 1 | 407.83 | 1 |
| $Date^2$ | 347.42 | 1 | 346.76 | 1 | 376.80 | 1 |
| Brood size | 2.30 | 1 | 2.30 | 1 | --- | |
| Brood age | 1.05 | 1 | 1.05 | 1 | --- | |
| *Supported model(s)* | | | | | | |
| Nr. hot days | 1.54 / 1.37 | 1 | 2.24 | 1 | 2.98 | 1 |
| Study site | 1.15 / 1.06 | 3 | 1.12 | 3 | 1.16 | 3 |
| Average temperature | --- | | 2.19 | 1 | 6.00 | 1 |
| Year | 1.11 / 1.06 | 5 | --- | | 1.19 | 5 |
| Date | --- | | --- | | 4.41 | 1 |
| $Date^2$ | --- | | --- | | --- | |
| Brood size | 1.80 / --- | 1 | 1.44 | 1 | --- | |
| Brood age | 1.03 / 1.03 | 1 | 1.04 | 1 | --- | |

## 5. AICc-based model selection results for multi-predictor models

**Supplementary Table 3. Result of AICc-based model selection for average nestling mass.** Supported models are highlighted in bold. Note that all models except for the "null model" included the fixed effects of study site, number of hot days, and their interaction (these are not shown among the predictors in the first column); the "Null-1" model contains only this interaction, and the "Null-2" model contains only the random effect of pair ID (which is also included in all other models). "MeanT" refers to the average temperature of the nestling period; "Date2" is the quadratic term of hatching date.

| Predictors | df | logLik | AICc | ΔAICc | Akaike weight |
|---|---|---|---|---|---|
| **Age + Brood size + Year** | **17** | **-1220.19** | **2475.31** | **0** | **0.55** |
| **Age + Year** | **16** | **-1221.97** | **2476.78** | **1.47** | **0.26** |
| Date2 + Date + Age + Brood size + Year | 19 | -1220.17 | 2479.49 | 4.19 | 0.07 |
| MeanT + Age + Brood size + Year | 18 | -1221.48 | 2480.00 | 4.69 | 0.05 |
| MeanT + Age + Year | 17 | -1223.37 | 2481.67 | 6.36 | 0.02 |
| Date2 + Date + Age + Year | 18 | -1222.4 | 2481.84 | 6.53 | 0.02 |
| MeanT + Date2 + Date + Age + Brood size + Year | 20 | -1221.12 | 2483.52 | 8.22 | 0.01 |
| Brood size + Year | 16 | -1226.23 | 2485.29 | 9.99 | 0 |
| MeanT + Date2 + Date + Age + Year | 19 | -1223.17 | 2485.51 | 10.2 | 0 |
| Year | 15 | -1227.84 | 2486.41 | 11.11 | 0 |
| Date2 + Date + Brood size + Year | 18 | -1226.43 | 2489.91 | 14.61 | 0 |
| MeanT + Brood size + Year | 17 | -1227.61 | 2490.15 | 14.84 | 0 |
| Date2 + Date + Year | 17 | -1227.96 | 2490.85 | 15.54 | 0 |
| MeanT + Year | 16 | -1229.11 | 2491.05 | 15.75 | 0 |
| Date2 + Date + Age + Brood size | 14 | -1232.31 | 2493.25 | 17.95 | 0 |
| MeanT + Date2 + Date + Brood size + Year | 19 | -1227.39 | 2493.94 | 18.63 | 0 |
| Age + Brood size | 12 | -1234.77 | 2494.02 | 18.71 | 0 |
| MeanT + Date2 + Date + Year | 18 | -1228.84 | 2494.72 | 19.42 | 0 |
| MeanT + Date2 + Date + Age + Brood size | 15 | -1232.37 | 2495.48 | 20.17 | 0 |
| MeanT + Age + Brood size | 13 | -1235.07 | 2496.68 | 21.38 | 0 |
| Brood size | 11 | -1241.41 | 2505.22 | 29.91 | 0 |
| Date2 + Date + Brood size | 13 | -1239.51 | 2505.57 | 30.26 | 0 |
| MeanT + Date2 + Date + Brood size | 14 | -1239.99 | 2508.62 | 33.31 | 0 |
| MeanT + Brood size | 12 | -1242.58 | 2509.63 | 34.32 | 0 |
| MeanT + Date2 + Date + Age | 14 | -1245.06 | 2518.77 | 43.46 | 0 |
| Age | 11 | -1249.02 | 2520.44 | 45.13 | 0 |
| MeanT + Age | 12 | -1249.51 | 2523.49 | 48.18 | 0 |
| Date2 + Date + Age | 13 | -1248.72 | 2524.00 | 48.69 | 0 |
| MeanT + Date2 + Date | 13 | -1251.77 | 2530.09 | 54.78 | 0 |
| Null-1 | 10 | -1255.36 | 2531.04 | 55.74 | 0 |
| Date2 + Date | 12 | -1254.36 | 2533.20 | 57.89 | 0 |
| MeanT | 11 | -1256.61 | 2535.61 | 60.31 | 0 |
| Null-2 | 3 | -1369.13 | 2744.30 | 269 | 0 |

**Supplementary Table 4. Result of AICc-based model selection for average nestling tarsus length.** Supported model is highlighted in bold. Note that all models except for the "null model" included the fixed effects of study site, number of hot days, and their interaction (these are not shown among the predictors in the first column); the "Null-1" model contains only this interaction, and the "Null-2" model contains only the random effect of pair ID (which is also included in all other models). "MeanT" refers to the average temperature of the nestling period; "Date2" is the quadratic term of hatching date.

| Predictors | df | logLik | AICc | ΔAICc | Akaike weight |
|---|---|---|---|---|---|
| **MeanT + Age + Brood size** | **13** | **-686.49** | **1399.52** | **0** | **0.94** |
| MeanT + Date2 + Date + Age + Brood size | 15 | -687.74 | 1406.21 | 6.68 | 0.03 |
| MeanT + Age + Brood size + Year | 18 | -686.07 | 1409.18 | 9.65 | 0.01 |
| MeanT + Brood size | 12 | -692.46 | 1409.39 | 9.86 | 0.01 |
| Date2 + Date + Age + Brood size | 14 | -691.27 | 1411.17 | 11.65 | 0 |
| MeanT + Date2 + Date + Age + Brood size + Year | 20 | -685.2 | 1411.68 | 12.16 | 0 |
| Age + Brood size | 12 | -693.68 | 1411.83 | 12.31 | 0 |
| MeanT + Date2 + Date + Brood size | 14 | -693.67 | 1415.99 | 16.46 | 0 |
| Brood size | 11 | -697.1 | 1416.59 | 17.07 | 0 |
| Age + Brood size + Year | 17 | -690.98 | 1416.89 | 17.36 | 0 |
| MeanT + Brood size + Year | 17 | -691.88 | 1418.69 | 19.17 | 0 |
| Date2 + Date + Brood size | 13 | -696.21 | 1418.98 | 19.45 | 0 |
| MeanT + Date2 + Date + Brood size + Year | 19 | -690.66 | 1420.49 | 20.97 | 0 |
| Brood size + Year | 16 | -695.18 | 1423.18 | 23.66 | 0 |
| Date2 + Date + Age + Brood size + Year | 19 | -692.23 | 1423.63 | 24.1 | 0 |
| Date2 + Date + Brood size + Year | 18 | -696.89 | 1430.82 | 31.3 | 0 |
| MeanT + Date2 + Date + Age + Year | 19 | -698.68 | 1436.53 | 37.01 | 0 |
| MeanT + Date2 + Date + Year | 18 | -702.76 | 1442.56 | 43.03 | 0 |
| Age + Year | 16 | -708.91 | 1450.64 | 51.12 | 0 |
| MeanT + Age + Year | 17 | -707.92 | 1450.76 | 51.24 | 0 |
| Year | 15 | -712.47 | 1455.66 | 56.14 | 0 |
| MeanT + Date2 + Date + Age | 14 | -713.53 | 1455.70 | 56.17 | 0 |
| Date2 + Date + Age + Year | 18 | -709.65 | 1456.34 | 56.82 | 0 |
| MeanT + Year | 16 | -712.36 | 1457.55 | 58.03 | 0 |
| Date2 + Date + Year | 17 | -712.68 | 1460.30 | 60.78 | 0 |
| MeanT + Age | 12 | -718.76 | 1461.98 | 62.46 | 0 |
| MeanT + Date2 + Date | 13 | -718.12 | 1462.79 | 63.27 | 0 |
| MeanT | 11 | -723.78 | 1469.97 | 70.44 | 0 |
| Age | 11 | -724.26 | 1470.91 | 71.39 | 0 |
| Null-1 | 10 | -727.16 | 1474.64 | 75.12 | 0 |
| Date2 + Date + Age | 13 | -726.24 | 1479.04 | 79.52 | 0 |
| Date2 + Date | 12 | -729.11 | 1482.70 | 83.17 | 0 |
| Null-2 | 3 | -783.63 | 1573.30 | 173.78 | 0 |

**Supplementary Table 5. Result of AICc-based model selection for nestling mortality.** Supported model is highlighted in bold. Note that all models except for the "null model" included the fixed effects of study site, number of hot days, and their interaction (these are not shown among the predictors in the first column); the "Null-1" model contains only this interaction, and the "Null-2" model contains only the random effects of pair ID and observation ID (which are also included in all other models). "MeanT" refers to the average temperature of the nestling period; "Date2" is the quadratic term of hatching date.

| Predictors | df | logLik | AICc | ΔAICc | Akaike weight |
|---|---|---|---|---|---|
| **MeanT+Date2+Date+Year** | **18** | **-1032.7** | **2102.33** | **0** | **1** |
| Date2+Date+Year | 17 | -1055.19 | 2145.20 | 42.87 | 0 |
| MeanT+Date2+Date | 13 | -1074.94 | 2176.37 | 74.04 | 0 |
| Year | 15 | -1088.84 | 2208.32 | 105.99 | 0 |
| MeanT+Year | 16 | -1088.53 | 2209.79 | 107.46 | 0 |
| Date2+Date | 12 | -1105.89 | 2236.20 | 133.87 | 0 |
| MeanT | 11 | -1125.98 | 2274.31 | 171.98 | 0 |
| Null-1 | 10 | -1130.95 | 2282.19 | 179.86 | 0 |
| Null-2 | 3 | -1203.8 | 2413.62 | 311.29 | 0 |

## 6. The effect of the number of hot days during the nestling period on reproductive success in great tits

**Supplementary Table 6. Estimated marginal means for effect of number of hot days.** The slope of the relationship with 95% confidence interval (CI) is shown for each site, from A) the simple models (containing number of hot days, study site, and their two-way interaction) and B) the supported multi-predictor models (Table 1). For nestling mass, results of the two supported multi-predictor models are separated by "/". Slopes significantly different from zero (i.e. zero not included between the lower and upper limit of CI) are highlighted in bold.

| Study sites | Average nestling mass | | | | Average nestling tarsus length | | | | Nestling mortality* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Slope | SE | lower CI | upper CI | Slope | SE | lower CI | upper CI | Slope | SE | lower CI | upper CI |
| **A) simple model** | | | | | | | | | | | | |
| Veszprém city | -0.081 | 0.046 | -0.172 | 0.010 | **-0.052** | **0.021** | **-0.094** | **-0.011** | **0.198** | **0.082** | **0.037** | **0.360** |
| Balatonfüred city | **0.163** | **0.049** | **0.065** | **0.260** | 0.036 | 0.023 | -0.009 | 0.081 | 0.011 | 0.090 | -0.165 | 0.187 |
| Szentgál forest | -0.106 | 0.085 | -0.273 | 0.061 | -0.050 | 0.038 | -0.126 | 0.026 | 0.161 | 0.176 | -0.183 | 0.505 |
| Vilma-puszta forest | **-0.489** | **0.131** | **-0.749** | **-0.230** | -0.059 | 0.060 | -0.178 | 0.060 | **0.864** | **0.236** | **0.401** | **1.327** |
| **B) multi-predictor model** | | | | | | | | | | | | |
| Veszprém city | **-0.201 / -0.269** | **0.098 / 0.095** | **-0.394 / -0.457** | **-0.007 / -0.082** | **-0.146** | **0.049** | **-0.243** | **-0.050** | 0.246 | 0.201 | -0.148 | 0.639 |
| Balatonfüred city | **0.271 / 0.236** | **0.098 / 0.098** | **0.076 / 0.042** | **0.466 / 0.430** | 0.002 | 0.049 | -0.094 | 0.098 | 0.113 | 0.193 | -0.266 | 0.492 |
| Szentgál forest | **-0.473 / -0.609** | **0.185 / 0.179** | **-0.840 / -0.962** | **-0.107 / -0.255** | -0.120 | 0.080 | -0.279 | 0.038 | 0.527 | 0.362 | -0.182 | 1.236 |
| Vilma-puszta forest | **-0.768 / -0.866** | **0.258 / 0.256** | **-1.278 / -1.373** | **-0.258 / -0.359** | -0.112 | 0.120 | -0.349 | 0.125 | **1.118** | **0.442** | **0.251** | **1.985** |

For average nestling mass and tarsus length, number of pairs was 535 and number of broods was 674. Sample sizes: Veszprém n= 222, Balatonfüred n= 115, Szentgál forest n= 242, Vilma-puszta forest n= 95 broods.
For nestling mortality, number of pairs was 600 and number of broods was 760. Sample sizes: Veszprém n= 260, Balatonfüred n= 130, Szentgál forest n= 260, Vilma-puszta forest n= 110 broods.
*Estimates for nestling mortality are on the logit scale

## 7. Sensitivity analysis of simple interaction models

As the range of number of hot days differed between urban and forest sites (urban: 0 - 13, forest: 0 - 5), we repeated the simple models with a subset of data where both habitat types had data points, i.e. only the broods with less than 6 hot days were included from each study site. These analyses resulted in similar estimates for the urban-forest habitat differences as the analyses of the whole data set, with differences for nestling mass and mortality remaining significant (Supplementary Table 7, compare with Table 2: Simple models in the main text).

**Supplementary Table 7. Differences between urban and forest habitats in the effect of hot days on nestlings' body mass, tarsus length, and mortality.** The table shows linear contrasts comparing the average slope of two urban versus two forest sites, calculated from the estimates of the simple models, excluding nestling periods with more than 5 hot days. Positive contrasts mean more positive slopes (i.e. less negative effects of hot days) in the urban habitat. Significant habitat differences ($p < 0.05$) are highlighted in bold.

| Response | contrast ± SE | df | t | p |
|---|---|---|---|---|
| **Nestling mass** | **0.343 ± 0.104** | **111** | **3.311** | **0.001** |
| **Nestling tarsus length** | 0.062 ± 0.047 | 111 | 1.325 | 0.188 |
| **Nestling mortality*** | **-0.443 ± 0.204** | **Inf** | **-2.174** | **0.030** |

* Estimates for nestling mortality are on the logit scale