

Supplementary materials of the article: **Complex drivers of phenology in the pine processionary moth: lessons from the past**

Jean-Claude Martin¹, Xavier Mesmin², Maurane Buradino¹, Jean-Pierre Rossi² & Carole Kerdelhué²

Agricultural and Forest Entomology

Appendix 1: Assessment of Model 1 validity hypotheses – diagnostic plots and discussion

Model 1 and 2 are mathematically related and residual distributions were consequently almost identical. Only diagnostics for Model 1 are shown here. Each GLMM validity hypothesis is discussed below with associated graphics. ‘DHARMA’ simulations were performed with default arguments of function ‘simulateResiduals’ except for the number of simulations that was raised to 1000.

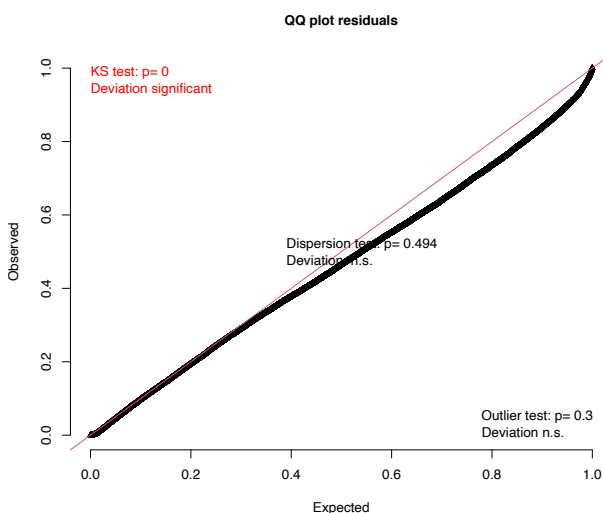


Figure S1.1 Assessment of model validity hypotheses for correct error distribution (KS test), dispersion (Dispersion test) and frequency of outliers (Outlier test) as obtained from package DHARMA.

Residuals distribution

Figure S1.1 shows that the fitted model significantly deviates from the expectations regarding the expected Gaussian distribution with identity link (significant KS test). Concerning residuals distribution, it is expected that slight deviations from normality become significant when sample size increases. Concurrently, tests and estimates based on a Gaussian distribution are increasingly reliable when sample size increases (Faraway, 2014, p. 81; Knief and Forstmeier, 2021). As we have

here a large dataset (41847 data points) we are very likely in this case. It is therefore necessary to assess the magnitude of the deviation from the expected distribution. Figure S1.2 shows that this deviation is minor: the obtained residuals clearly follow a Gaussian distribution, but slightly leptokurtic (Sokal and Rohlf, 1995, p. 114) as compared to the expectation. To our knowledge, no argument is available to correct the kurtosis in packages currently implementing generalized linear mixed models. So the deviation seems really minor, and no correction can be made.

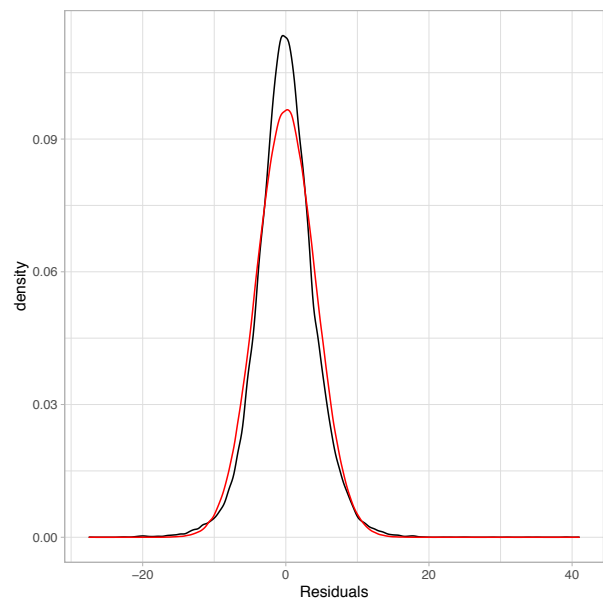


Figure S1.2 Comparison between the expected (in red) and obtained (in black) residuals distribution. The discrepancy between the two curves leads to the KS test = 0 in Fig. S1.1.

Dispersion

The dispersion of data points does not deviate significantly from the expectations (Fig. S1.1, “Dispersion test”).

Outlier frequency

The outlier frequency does not differ significantly from the expected frequency (p-value obtained with the argument *type* = “bootstrap” in the function ‘testOutliers’ of package ‘DHARMa’, as recommended for integer-valued responses).

Homoscedasticity – model formulation

Figure S1.3 shows that the expectation regarding the homogeneity of variances along the gradient of predicted values is not met.

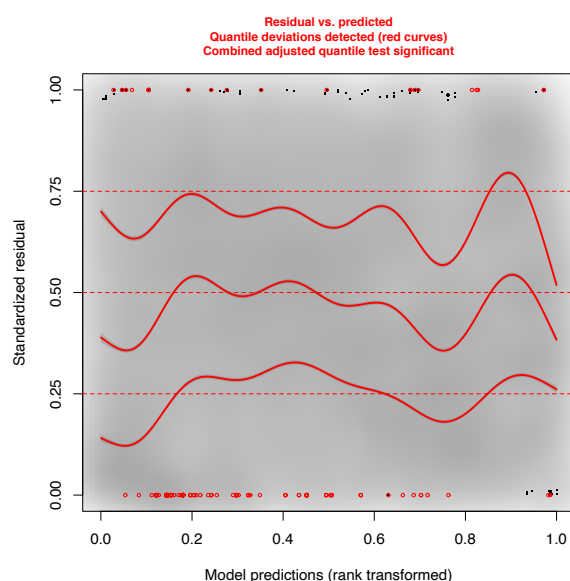


Figure S1.3 Assessment of the model validity hypothesis for homogeneity of residual variances along the fitted values as obtained from package ‘DHARMa’ (Hartig, 2020). The grey pattern on the background depicts the standardized residuals according to model predictions for each data point (dark = many data points in this area of the plot, light = few data points in this area of the plot). Solid lines depict the evolution of quantiles 25, 50 and 75 of the standardized residuals and dashed lines show the expected evolution of these quantiles with homogeneity of variances. The deviation of the solid lines from the dashed lines is significant.

The pattern observed suggests changes in the model formulation to adjust the fitted regressions to slight deviations from normality (*i.e.* points not exactly aligned in Fig. 2C and 2D for instance). As the sample size is large, these slight deviations from linearity become significant here.

The deviations from linearity might be captured by using polynomial terms, but Fig. S1.3 suggests that polynomials of degree 2 or 3 would not be sufficient. Moreover, adding polynomial terms to meet the homoscedasticity assumption would lead to overfitting, and would considerably decrease the interpretability of the results.

Finally, Fig. S1.4 indicates that even though the model formulation may miss deviations from linearity, Model 1 overall fits the observed data very well.

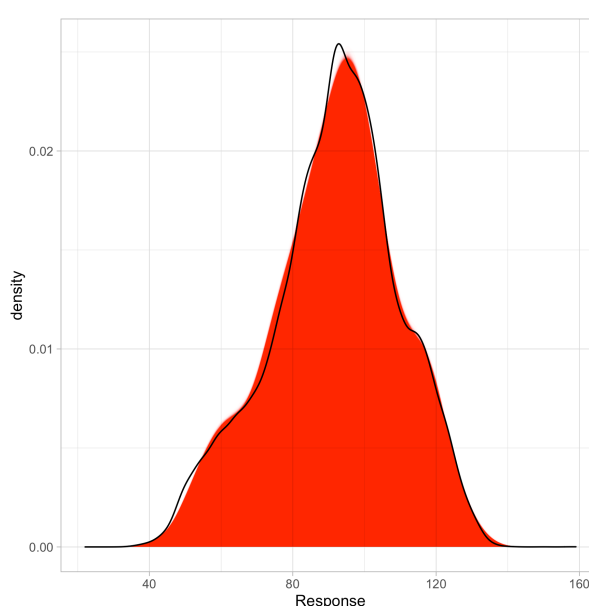


Figure S1.4 Comparison between the response distribution predicted from the coefficients estimated in Model 1 (in red, 1000 curves overlaid with transparency, based on the 1000 simulations of ‘simulateResiduals’ of ‘DHARMa’ conditioned on the fitted random effects) and the response distribution actually obtained (in black).

On the grounds of these diagnostics, we believe that these deviations from the expected residual distribution in Model 1-2 do not come from a misspecification of our model, and therefore that the coefficients and p-values retrieved from these models are reliable.

Appendix 2: Assessment of Model 3 validity hypotheses – diagnostic plots and discussion

‘DHARMa’ simulations were performed with default arguments of function ‘simulateResiduals’ except for the number of simulations that was raised to 1000.

Hypotheses for correct distribution, dispersion and outlier frequencies were all met (Fig. S2.1)

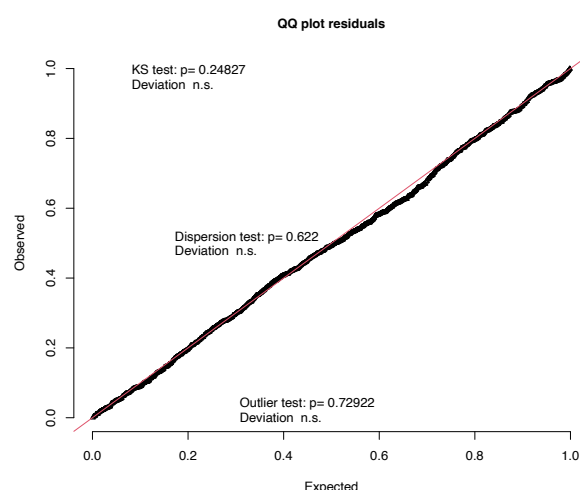


Figure S2.1 Assessment of model validity hypotheses for correct error distribution (KS test), dispersion (Dispersion test) and frequency of outliers (Outlier test) as obtained from package DHARMa.

Homoscedasticity – model formulation

Conversely, the expectation regarding the homogeneity of variances along the gradient of predicted values was not met because of a slight deviation from the expected distribution for high positive residuals (Fig. S2.2).

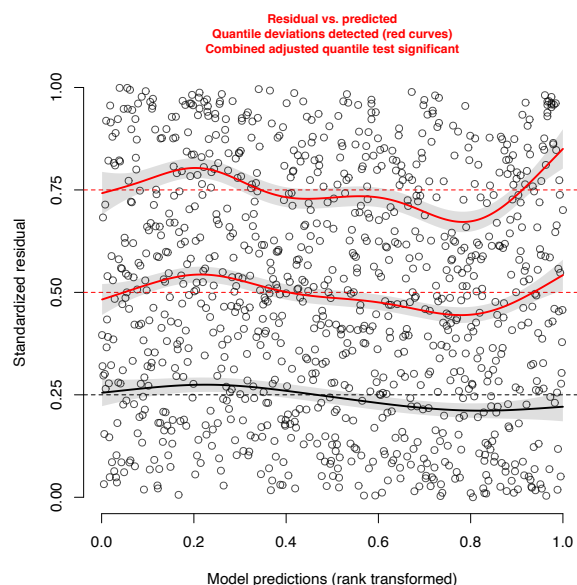


Figure S2.2 Assessment of the model validity hypothesis for homogeneity of residual variances along the fitted values as obtained from package ‘DHARMa’ (Hartig, 2020). The dots depict the standardized residuals according to model predictions for each data point. Solid lines depict the evolution of quantiles 25, 50 and 75 of the standardized residuals and dashed lines show the expected evolution of these quantiles with homogeneity of variances. The deviation of the solid lines from the dashed lines is significant ($p = 0.041$).

In our opinion, the slight deviation observed in Fig. S2.2 does not call into question the results obtained from Model 3.

When the outlier from site F781 in 1978 was removed, all hypotheses were validated.

Appendix 3: Assessment of Model 5 validity hypotheses – diagnostic plots and discussion

‘DHARMa’ simulations were performed with default arguments of function ‘simulateResiduals’ except for the number of simulations that was raised to 1000.

Hypotheses for correct distribution, dispersion and outlier frequencies were all met (Fig. S3.1)

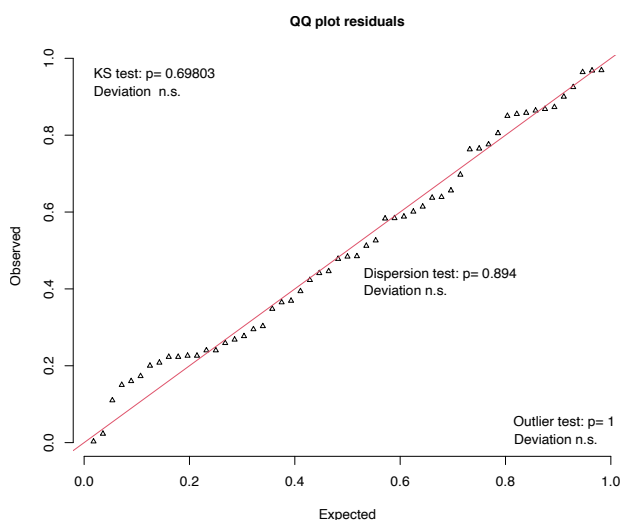


Figure S3.1 Assessment of model validity hypotheses for correct error distribution (KS test), dispersion (Dispersion test) and frequency of outliers (Outlier test) as obtained from package DHARMa.

Homoscedasticity – model formulation

Conversely, the expectation regarding the homogeneity of variances along the gradient of predicted values was not met because of a clear U-shaped pattern in the distribution (Fig. S3.2).

This suggested the addition of a quadratic effect in the model formula. However, we did not do so because:

- The hypotheses we wanted to test in this specific model were about the slope and intercept of the linear correlation between y and x. Adding a quadratic effect would not have allowed to test this hypothesis;
- The quadratic pattern between y and x was only supported by the two data points with the lowest median emergence date on the

abscissae (see Fig. 4B, emergence before the 1st of July). Removing these points corrected the pattern in the residuals ($p=0.263$) and the estimates associated with this model indicated that the slope did not differ significantly from 1 ($p=0.681$) and that the intercept did not differ significantly from 0 ($p=0.583$). So the conclusions of this model strictly respecting all GLMM hypotheses were the same than in those presented in the Main Text.

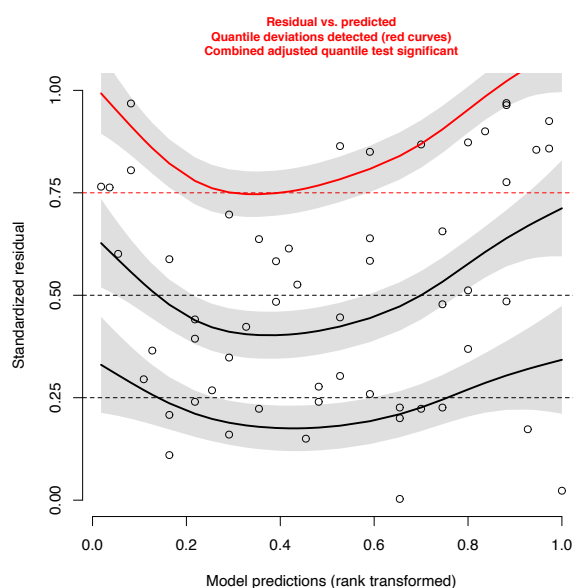


Figure S3.2 Assessment of the model validity hypothesis for homogeneity of residual variances along the fitted values as obtained from package ‘DHARMa’ (Hartig, 2020). The dots depict the standardized residuals according to model predictions for each data point. Solid lines depict the evolution of quantiles 25, 50 and 75 of the standardized residuals and dashed lines show the expected evolution of these quantiles with homogeneity of variances. The deviation of the solid lines from the dashed lines is significant ($p = 0.002$).

On these grounds, we chose to keep Model 5 with all the available data and without any quadratic effect.

Appendix 4: Description of raw data on the number of caterpillars caught and on adult phenology of univoltine individuals

Figure S4.1 Number of caterpillars monitored during the study for each site and year

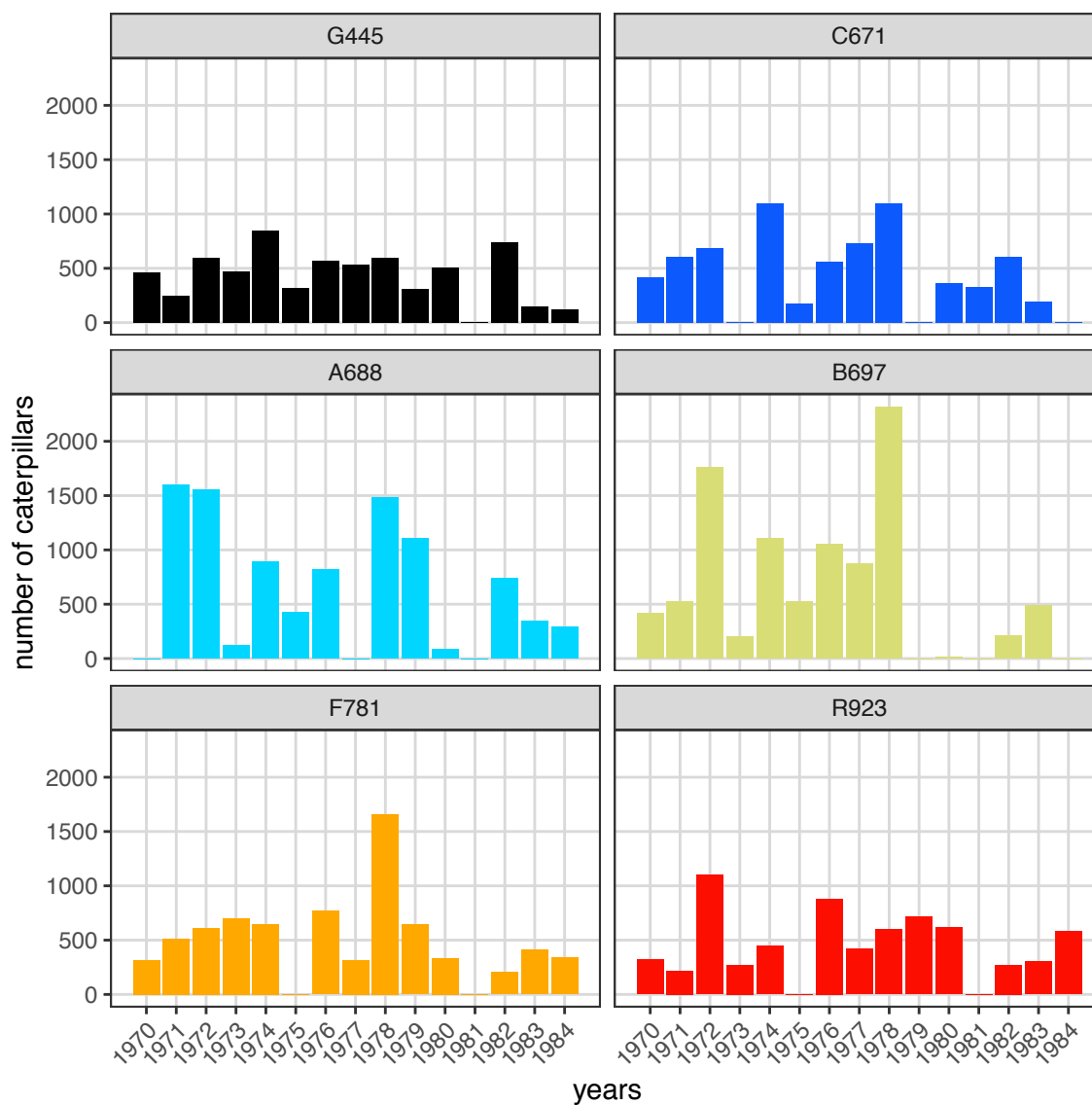
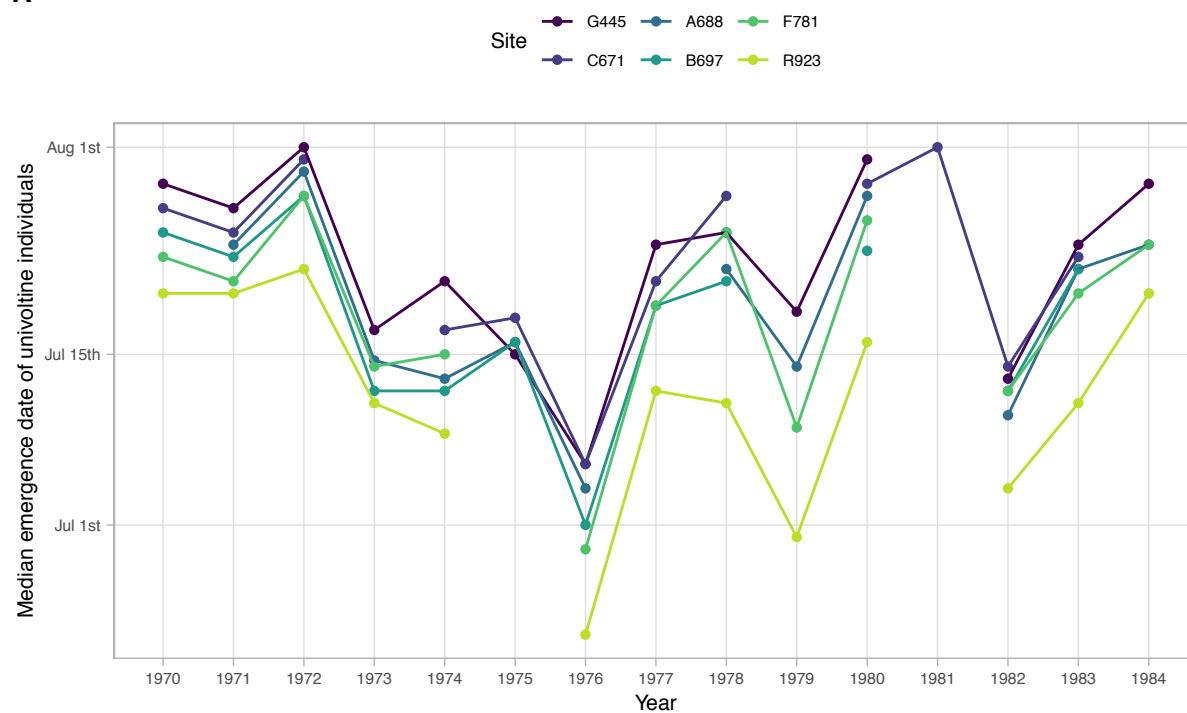


Figure S4.2 Variability across site and year of adult phenology expressed as the median (A) or mean (B) emergence date of univoltine individuals (DirEm_N median).

A



B

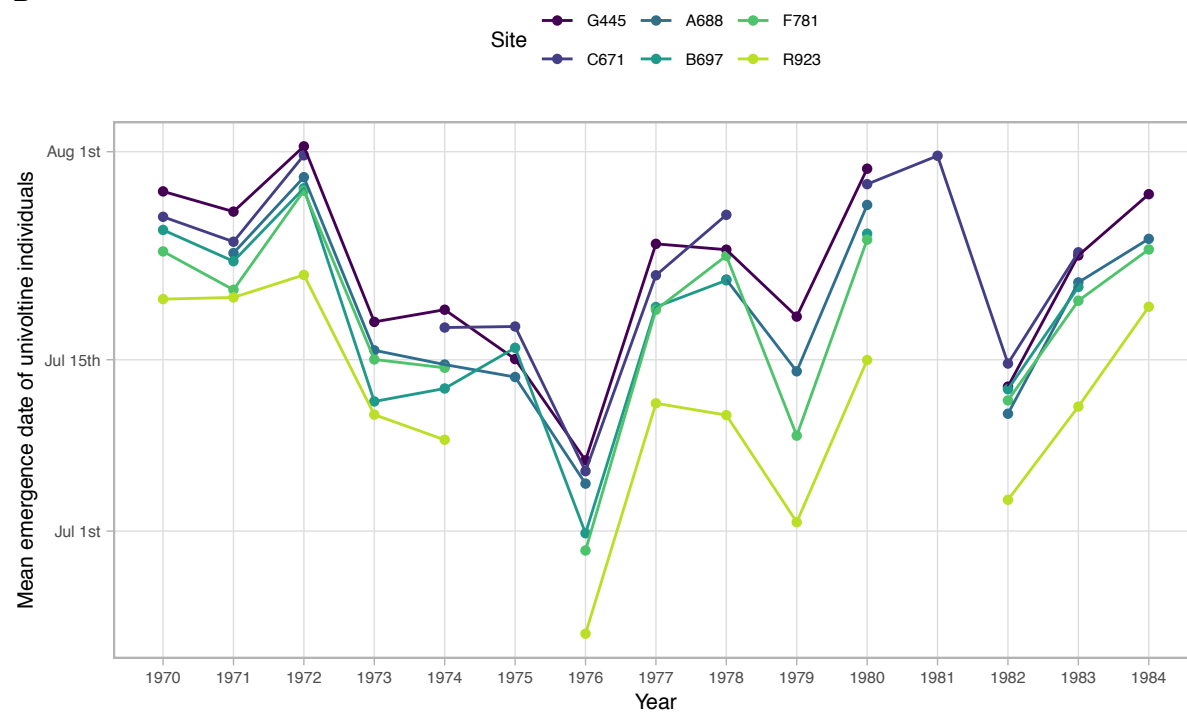
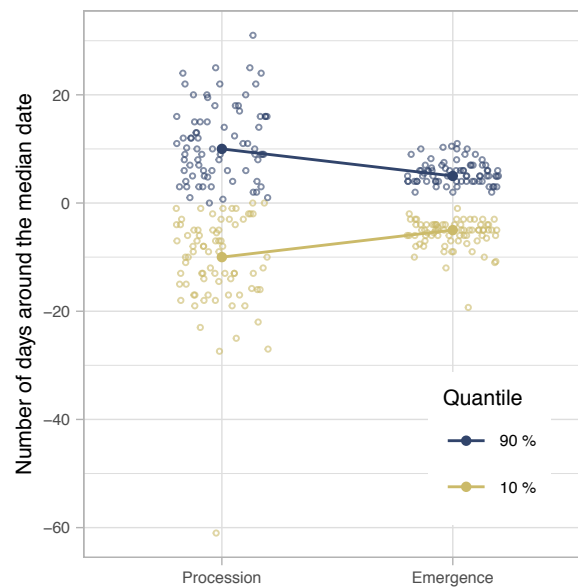


Figure S4.3 Reduction of the temporal dispersion of emergence dates of univoltine individuals as compared to that of their procession dates. Each circle represents either the 10 % (yellow) or the 90 % (darkblue) quantile date of the procession (left) or of the emergence (right) on a given “site X year” combination. For instance, for any yellow (resp., darkblue) circle of the left side of the figure (i.e. showing processions), 90 % of the processions occurred later (resp., earlier) on that particular “site X year” combination. The points show the median value of their related circles. The inter-quantile range is significantly lower for emergence than for procession ($\chi^2 = 100.1$, $df = 1$, $P < 0.001$; model checking all DHARMA hypotheses, performed with ‘glmmTMB’, linearly parameterized negative binomial distribution “nbinom1”, log link, and random effects on site and year to consider inter-dependence among data due to the sampling design).



Appendix 5: Complementary information for Model 3 fitted with or without data collected in site F781 in 1978

Figure S5.1 Distribution of the proportion of moths entering prolonged diapause per site and per year (all batches of each site-year are pooled). The data obtained in site F781 in 1978, considered as outliers, are highlighted in red.

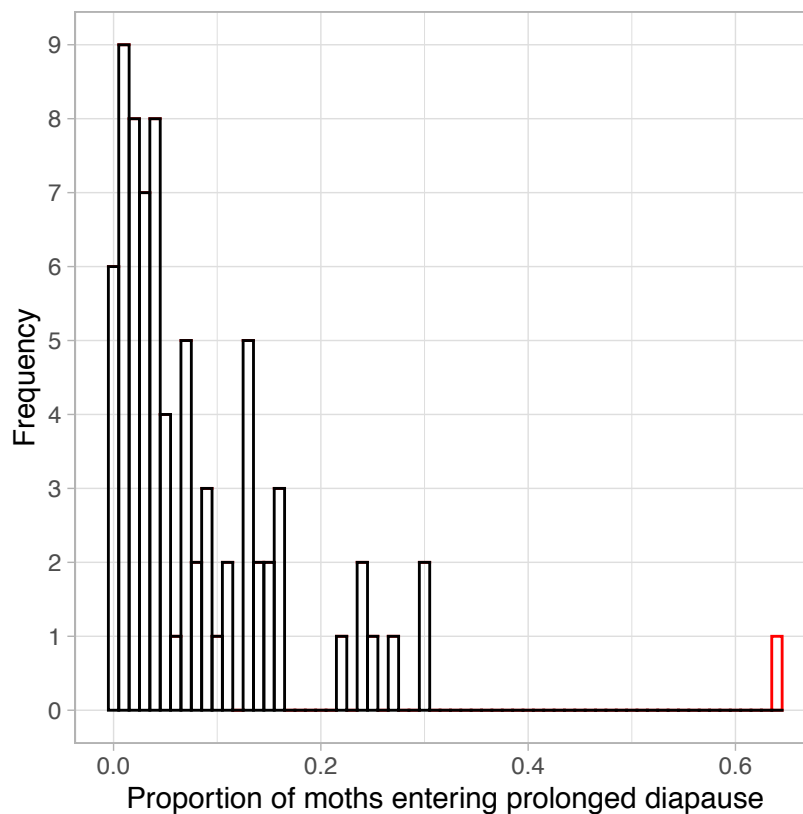


Figure S5.2 Correlation between the proportion of moths entering prolonged diapause and population density (*i.e.* the total number of caterpillars entering procession at a given site in a given year). Raw data are displayed as squares which color shows the number of moths for each (x, y) coordinate. The regression line from Model 3 fitted with all data is displayed as a black line.

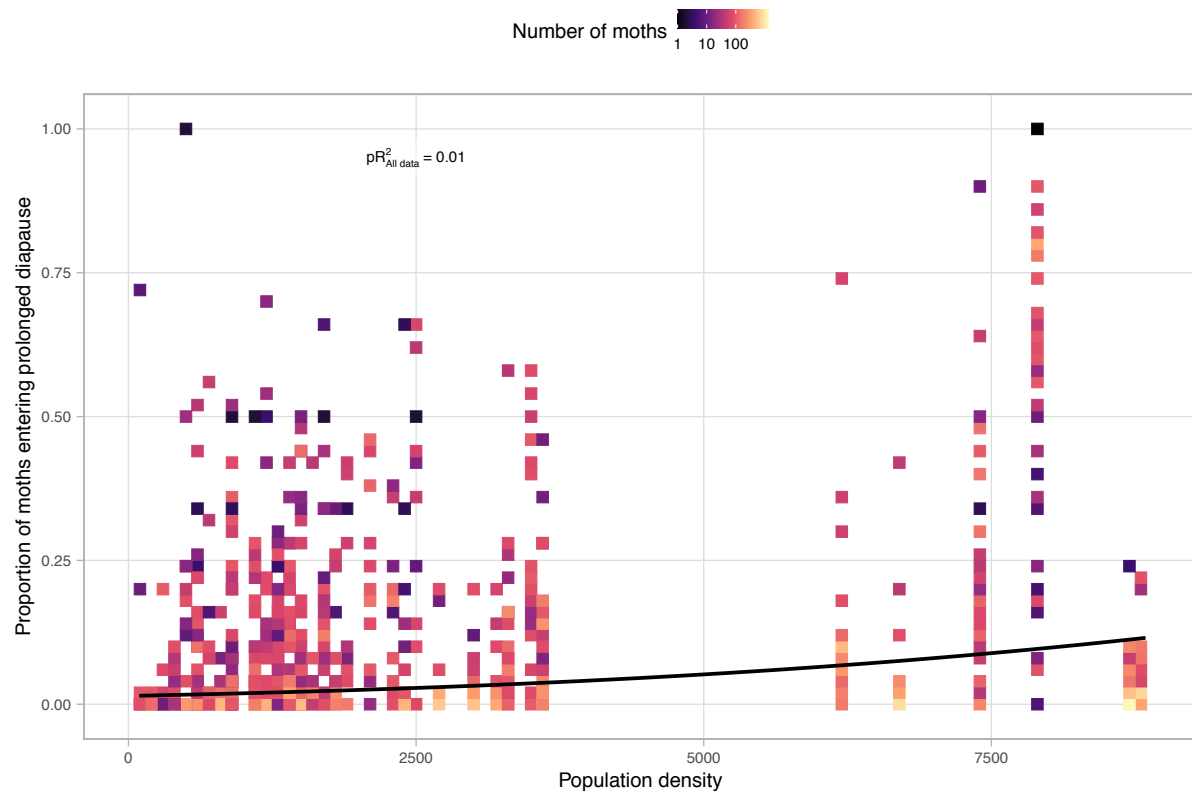
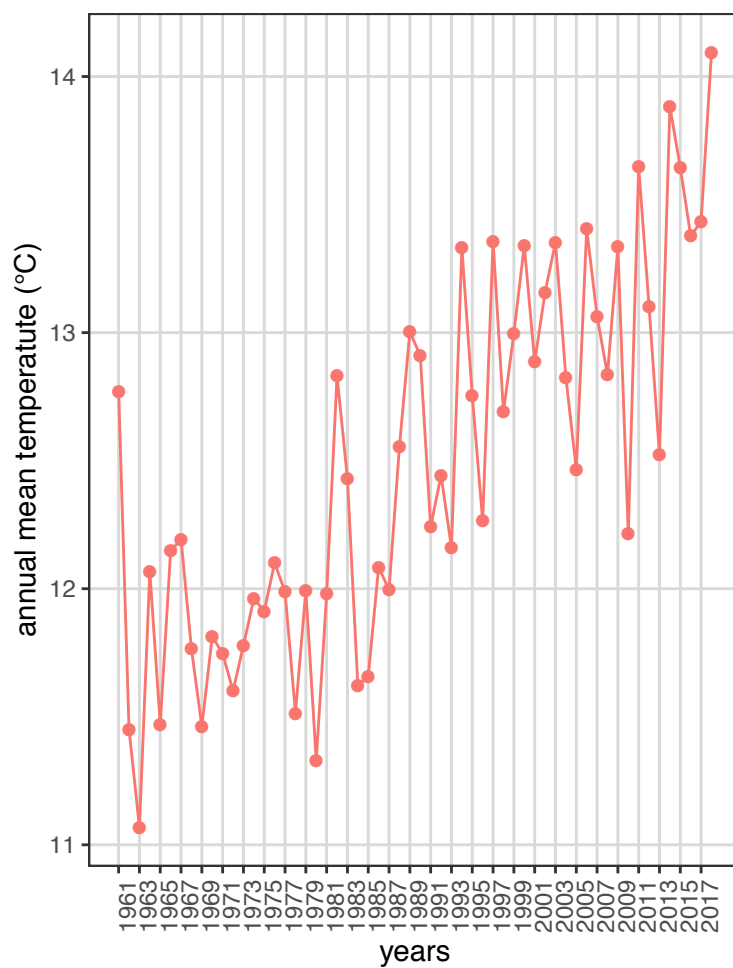


Table S5.1 Values and significance of deviance analyses performed on the fixed effects of Model 3 fitted without data collected in site F781 in 1978. “-” stands for “unsignificant fixed effect removed during model selection”.

Model	Fixed effects	χ^2	df	P
Model 3: Proportion of moths entering prolonged diapause	Procession date	35.6	1	< 0.001
	Winter temperature	9.4	1	0.002
	Elevation	8.5	1	0.003
	Population density	-	-	-

Appendix 6: Evolution of mean annual temperatures between 1969 and 2018 at the study site. Climatic data were retrieved from WorldClim v2.1



References

- Faraway, J.J., 2014. Linear Models with R, Second Edition. ed, Texts in Statistical Science. Chapman & Hall/CRC Taylor & Francis group.
- Hartig, F., 2020. Package “DHARMA” Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.
- Knief, U., Forstmeier, W., 2021. Violating the normality assumption may be the lesser of two evils. Behav Res. <https://doi.org/10.3758/s13428-021-01587-5>
- Sokal, R.R., Rohlf, F.J., 1995. Biometry: the principles and practice of statistics in biological research, 3rd ed. ed.