APPENDIX S7. Sensitivity analysis of classification scheme model-fit to sample size. To test the effect of sample size on model-fit, we reran our full model (DOY ~ Elevation + Year + Temperature*Phenophase) for each classification scheme for a range of subset fractions of 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 of the original data set (corresponding number of specimens: 53, 105, 210, 315, 420, 526, and 788, respectively). For each fraction, we randomly subsampled 100 unique data sets for that given fractional amount. We then tested the fit based on Akaike information criterion (AIC) and R^2 values of each classification scheme on each of the 100 subsampled data sets per fraction (total of 700 separate analyses). (A, B) The relative AIC scores and R^2 for each classification scheme by subsample fraction. The values have been standardized (mean-centered) within each fraction for easy comparison between fractions. (C, D) Frequency of best-fit classification schemes by fraction based on AIC and R^2 values. The frequency of classification schemes is out of the 100 randomly subsampled data sets for fraction.