insects

MDPI

Article

Phenological Trends and Projected Climatic Suitability of Five Wool Carder Bees of the Genus *Anthidium* Across the Mediterranean Basin

Okan Özgül [1,*], Max Kasparek [2,*] and Hasan Koç [3]

[1] Beekeeping Program, Department of Crop and Animal Production, Ula Ali Koçman Vocational School, Muğla Sıtkı Koçman University, Ula, 48640 Muğla, Türkiye
[2] Independent Researcher, Mönchhofstr. 16, 69120 Heidelberg, Germany
[3] Department of Biology, Faculty of Science, Muğla Sıtkı Koçman University, 48000 Muğla, Türkiye; khasan@mu.edu.tr
* Correspondence: okanozgul@mu.edu.tr (O.Ö.); kasparek@t-online.de (M.K.)

Simple Summary

Climate change may alter when bees become active and where they can find suitable conditions. We examined records of five wool carder bee species from the Mediterranean region collected between 1980 and 2024. We assessed changes in their early-season flight dates and estimated how climatically suitable areas might change by the end of this century. One species, *Anthidium florentinum*, showed a clear shift towards earlier flight dates, estimated at about 10 days per decade. Evidence for similar shifts in the other four species was less certain. Under a high-emissions climate scenario, climatically suitable conditions were projected to decline most in Spain, Morocco and Portugal, whereas France, Italy and Croatia showed small net gains. These projections require caution because the records are unevenly distributed and model performance declined when models were tested in geographic areas not used for training. The findings can guide further monitoring of Mediterranean bees and show where better records are needed to support conservation planning.

Abstract

Climate change may alter both the timing of bee activity and the climatic suitability of bee habitats, risking plant–pollinator mismatches, yet such effects are rarely assessed for Mediterranean bees. We analysed 4103 occurrence records (1980–2024) of five wool carder bee species (*Anthidium*) to evaluate first-flight trends and to project changes in climatically suitable areas for 2081–2100 using ensemble species distribution models driven by five general circulation models. Non-stationary phenological dynamics were detected in four species. Only *A. florentinum* showed a robust advance in first-flight timing of 10.0 days per decade that remained significant after False Discovery Rate (FDR) correction ($q < 0.001$), whereas the advance in *A. loti* was not significant ($q = 0.082$). The advance in *A. florentinum* persisted when only Iberian records were analysed. Under the SSP5-8.5 emissions scenario, six of 15 countries met the data-sufficiency criterion; among these, the largest net declines in climatically suitable sampled locations were projected for Spain (−15.3 percentage points), Morocco (−13.7) and Portugal (−12.5), whereas Italy, France and Croatia showed small net gains. Under spatial block cross-validation, model performance was lower than under random cross-validation but remained acceptable by AUC (0.75–0.81). Among countries meeting the data-sufficiency criterion, Morocco, Spain and Portugal exceeded the reference threshold of a Composite Vulnerability Index that combines projected habitat decline, phenological shift and projected warming. These results indicate uneven, country-specific

Check for updates

Academic Editor: Terd Disayathanoowat

Received: 25 September 2026
Revised: 15 October 2026
Accepted: 20 October 2026
Published: 23 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

vulnerability and highlight the need to account for non-stationarity, sampling bias and spatial validation when assessing climate risks to Mediterranean pollinators.

Keywords: phenology; species distribution modelling; climate change; Mediterranean Basin; *Anthidium*; phenological shifts; habitat loss; vulnerability index

1. Introduction

Phenology, the timing of recurring life-cycle events such as emergence, flight periods, or diapause, is highly sensitive to local environmental conditions in insects. The seasonal timing of insect life cycles is primarily driven by photoperiod and temperature [1,2]. Temperature directly controls insect metabolic rates; higher temperatures accelerate development and enable earlier seasonal emergence, a relationship formalised in mechanistic models linking individual-level thermal responses to population-level phenological patterns [3]. Consequently, altering climatic conditions can have detrimental effects on insect populations. This relationship is especially critical for bees (Hymenoptera: Apoidea), which serve as key pollinators of both wild and cultivated plants [4,5]. Climate warming can disrupt the phenological synchrony between plants and their pollinators [6–8]. A recent global analysis confirmed that plant and animal phenological responses to climate change are becoming increasingly asymmetric, with plants generally advancing more than animals [9].

Beyond phenological shifts, climate change is also expected to reshape the geographic ranges within which pollinators can persist, and these two dimensions—timing and distribution—are rarely assessed together for a single taxonomic group. Species distribution models (SDMs) that relate current occurrence records to climatic predictors have become a standard tool for projecting how suitable habitats may contract or expand under future climate scenarios [10]. Because different modelling algorithms can disagree substantially in their projections, particularly near the edges of a species' range, ensemble approaches that combine multiple algorithms are now widely used as a way of producing more robust distributional forecasts than any single model in isolation [11]. Bees offer an instructive case for combining these two lines of evidence. For example, long-term occurrence data across Europe and North America have shown that bumblebee species have generally failed to track warming at their northern range margins while contracting from the warmer edges of their distributions, a pattern that phenological data alone could not have revealed [12].

The Mediterranean Basin is warming and responding to climate change more rapidly than many other regions and has been repeatedly classified as a primary "climate change hotspot" [13,14]. Regional temperatures are projected to rise approximately 20% faster than the global average [15]. This rapid warming underlines the importance of studying the impacts of climate change on bee phenology in the Mediterranean. At the same time, the Mediterranean Basin is known as a biodiversity hotspot characterised by exceptionally high species diversity and a high rate of endemism [16,17]. Of the 2138 species of wild bees found in Europe [18], the highest diversity is noted in Mediterranean countries [19,20], including wool carder and resin bees of the megachilid tribe Anthidiini (Hymenoptera: Megachilidae) [21]. Overall, we know almost nothing about how bees in this region respond to phenological shifts in host-plant flowering [22].

To study possible climate-associated phenological shifts in the Mediterranean hotspot, we selected five medium-sized species of wool carder bees belonging to the genus *Anthidium*, that are relatively easy to identify and for which historical data from recent decades are available. These five species are solitary, polylectic bees [21,23]; thus, their generalist foraging habits reduce dependence on any single host plant, although their activity period

likely still tracks the broader availability of flowering resources. We combine multi-decadal occurrence records for these species with gridded climate data to examine phenology and geographic suitability within a single analytical framework. We tested whether first-flight timing has advanced and whether these trends remained stable through time, examined their association with spring growing degree days, and projected changes in climatically suitable habitats by the end of the century under two emissions scenarios. As additional context, we also compare each species' observed phenological shift against the shift expected from local warming trends alone. Because incidental occurrence records of the kind analysed here vary widely in sample size and reliability across species, years, and countries, we place particular emphasis on testing how sensitive our conclusions are to these sources of uncertainty, with the aim of distinguishing genuinely robust signals from those that remain provisional given current data.

2. Materials and Methods

2.1. Data Collection and Sampling Bias Correction

We selected for this study five *Anthidium* species: *Anthidium florentinum* (Fabricius, 1775), *A. loti* (Perris, 1852) s.l., *A. manicatum* (Linnaeus, 1758), *A. oblongatum* (Illiger, 1806), and *A. punctatum* Latreille, 1809. Occurrence data were obtained from the Global Biodiversity Information Facility (GBIF, https://doi.org/10.15468/dl.h3udkn (accessed on 15 April 2026)) using a standardised occurrence download, accessed on 15 April 2026 [24]. *Anthidium florentinum* is treated here as a single valid species, although DNA barcodes of European specimens fall into three clusters with non-overlapping Barcode Index Numbers (BINs), whose geographic limits and taxonomic status have not yet been resolved [25]. Similarly, we adopt the traditional species concept for *Anthidium loti* (*A. loti* s.l.), despite recent findings demonstrating that eastern populations constitute a distinct species, *Anthidium regulare* Eversmann, 1852; *A. loti* s.str. is confined to the Iberian Peninsula, France and western Liguria, whereas *A. regulare* occurs in the rest of Italy, the Balkans and the eastern Mediterranean [26]. To assess the effect of this taxonomic pooling, the phenological trend analysis was repeated for subsets defined by these published ranges. For *A. loti*, records from Spain, Portugal, France and western Liguria (43.75–44.50° N, ≤8.45° E) were assigned to *A. loti* s.str. and all other records to *A. regulare* [26]. For *A. florentinum*, whose cluster boundaries have not been published, the analysis was repeated on nested western subsets—the Iberian Peninsula alone, then adding France, Morocco and Algeria, and finally Italy—so that the result does not depend on the position of the boundary; the Iberian Peninsula is the origin of the reference specimen of one of the three clusters [25] (Table S11). The dataset was filtered to include only records with georeferenced coordinates (hasCoordinate = TRUE) and no geospatial issues. Only records classified as "Human observation" or "Preserved specimen" were retained. The temporal range was restricted to 1980–2024. All data were accessed under the GBIF data user agreement and are licensed under CC BY-NC 4.0 [24].

Additionally, we cross-checked our dataset against the recently published EBHD database [27] (Zenodo v1, December 2025 release). Within our study period (1980–2024), the two datasets provide identical coverage for the five target species—no additional records were found, confirming that our dataset is current and complete within the analysis window.

The study area was defined using a polygon covering the Mediterranean Basin, and records were limited to 23 countries within this region: Albania, Algeria, Bosnia and Herzegovina, Croatia, Cyprus, Egypt, France, Greece, Israel, Italy, Jordan, Lebanon, Libya, Malta, Montenegro, Morocco, Palestine, Portugal, Slovenia, Spain, Syria, Tunisia, and Türkiye. Of these candidate countries, eight (Egypt, Israel, Jordan, Libya, Lebanon, Malta, Palestine, and Tunisia) yielded no confirmed occurrence records for any of the five target species and

were therefore excluded a priori from all country-level habitat-loss and vulnerability analyses (Sections 2.4 and 2.8); this a priori exclusion is distinct from, and applied before, the data-sufficiency screening described in Section 2.4. The full per-country, per-species record breakdown is provided in Table S9. One record from North Macedonia, located within the study extent but outside the 23 candidate countries, was retained for phenological and distribution modelling but not for country-level analyses.

Only presence records (occurrenceStatus = present) were used in subsequent analyses. The final dataset comprised records from 69 contributing datasets across 28 publishers and 16 countries (15 of the 23 candidate countries plus North Macedonia; Table S9) and comprised a total of 4103 records, distributed across species as follows: *A. florentinum* (n = 2598), *A. manicatum* (n = 899), *A. oblongatum* (n = 310), *A. loti* s.l. (n = 202), and *A. punctatum* (n = 94).

To reduce sampling bias, the annual onset of flight activity was estimated using the 10th percentile day of year (P10 DOY), as this widely used metric has been shown to be a reliable indicator of phenological onset [28–30]. Prior to calculating P10, duplicate day-of-year (DOY) records within each year were removed, and only unique observation days were retained. Years with fewer than three unique observation days per species (minimum requirement) were excluded due to insufficient data. This sensitivity was tested against stricter ($n \geq 5$ and $n \geq 10$) thresholds (Section 3.1), and results should be interpreted alongside that sensitivity analysis [29]. The 10th percentile was preferred over the earliest annual record, which depends on a single observation, and over more complex onset estimators that require larger annual samples than are available here [28]; its sensitivity to the choice of estimator was assessed with a bias-corrected Weibull 10th percentile (Section 2.5), and the number of observation days behind each annual estimate is reported in Table S13. This approach limits the influence of extreme observations on percentile estimates and partially compensates for interannual variation in sampling intensity. The resulting P10 dates were analysed against observation year using the Mann–Kendall trend test and the Theil–Sen slope estimator (Section 2.5).

For species distribution modelling, occurrence records were restricted to the study extent (27–47° N, 18° W–37° E) and reduced to one record per species per 2.5′ grid cell, because records sharing a cell have identical predictor values; this retained 1156, 460, 169, 118 and 65 cells for *A. florentinum*, *A. manicatum*, *A. oblongatum*, *A. loti* and *A. punctatum*, respectively (Table S5). No further spatial thinning was applied; instead, residual spatial sampling bias was addressed by evaluating model performance under both random and spatial block cross-validation (Section 2.4). As absence data were not available, a target-group background was used [31]: georeferenced GBIF presence records of six bee families (Apidae, Megachilidae, Andrenidae, Halictidae, Colletidae and Melittidae; 658,109 records) were restricted to the same study extent and candidate countries and reduced to one record per 2.5′ cell (35,951 cells), so that background locations share the sampling bias of the target records. To limit the influence of the most intensively sampled countries, at most 3000 background cells were drawn per country, giving a common background pool of 18,519 cells. This cap reduced the combined share of Spain, France and Portugal from 62% of the available background cells to 45% of the pool (Table S12). For each species, background points were drawn from this pool at a 10:1 background-to-presence ratio [32].

2.2. Statistical Analysis and Software

All analyses were conducted in Python 3.13. The primary packages used included NumPy 2.1, pandas 3.0, SciPy 1.17 (Theil–Sen regression and statistical distributions; the Mann–Kendall test was implemented directly following Equation (5)), scikit-learn 1.9, xarray 2025.4 (for ERA5 processing), and Matplotlib 3.10 (for visualisation). Species dis-

tribution modelling was performed using the scikit-learn library, implementing Random Forest (RF), Gradient Boosting (GB), and Logistic Regression (LR) algorithms. Generalised additive model (GAM) analyses were approximated using scikit-learn's spline-based regression (SplineTransformer); no independent R/mgcv validation was performed. The threshold for statistical significance was set at $\alpha = 0.05$. Because the same statistical tests (Mann–Kendall trend test; GDD–DOY partial correlation) were repeated across five species, Benjamini–Hochberg false discovery rate (FDR) correction (Equation (1)) was applied within each test family; both uncorrected (raw) and FDR-corrected values are reported throughout, with FDR-corrected significance used as the basis for confirmatory claims.

$$q_i = \min_{j \geq i} \left[\frac{p_j \times n}{j} \right] \quad (1)$$

2.3. Climate Variables

Climatic variables used in species distribution modelling were obtained from the WorldClim v2.1 database at a spatial resolution of 2.5 arc-minutes (~4.6 km in the Mediterranean) [33]. All 19 bioclimatic (BIO) variables for the 1970–2000 reference period were initially considered as candidate predictors. This fixed baseline climatology period, defined by the WorldClim v2.1 release, was retained for consistency with standard SDM practice and comparability with the CMIP6 future projections built on the same baseline. To address multicollinearity, a stepwise Variance Inflation Factor (VIF) elimination was applied, whereby at each step the variable with the highest VIF value was removed until all remaining variables satisfied the threshold of VIF < 10 [34]. VIF values were calculated from WorldClim values extracted at the pooled presence and background locations. This procedure resulted in the retention of eight variables: bio1 (annual mean temperature), bio2 (mean diurnal temperature range), bio4 (temperature seasonality), bio8 (mean temperature of the wettest quarter), bio9 (mean temperature of the driest quarter), bio14 (precipitation of the driest month), bio15 (precipitation seasonality), and bio19 (precipitation of the coldest quarter), with a final maximum VIF of 7.49 (bio1). The highest pairwise Pearson correlation among the retained variables was observed between bio14 and bio15 ($r = -0.78$); both were retained because VIF—which accounts for shared variance with all other retained predictors jointly, not pairwise correlation alone—remained below the threshold of 10 for both variables (Section 3.7). All climatic layers (WorldClim baseline, CMIP6 projections) were aligned to the same 2.5 arc-minute grid prior to analysis; no downscaling or upscaling of individual variables was performed.

Future climate projections for 2081–2100 were derived from the WorldClim CMIP6 database (https://www.worldclim.org/data/cmip6/cmip6climate.html, accessed on 25 September 2026) using five General Circulation Models (GCMs): ACCESS-CM2, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0 and UKESM1-0-LL. Two emission scenarios were considered: SSP2-4.5 (intermediate pathway) and SSP5-8.5 (high-emission pathway). For each scenario, suitability was predicted separately with each GCM and then averaged across the five GCMs. Phenological DOY projections were treated as independent of emission scenarios.

For observational climate validation, ERA5—the fifth-generation atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), combining historical observations with model output via data assimilation to provide a physically consistent, gridded climate record—monthly averaged reanalysis data at a spatial resolution of 0.25°, covering 27–47° N and 18° W–37° E, were used [35]. Annual trends in near-surface air temperature (T2m) were calculated for the period 1980–2024. Spring growing degree days (GDD; base temperature 10 °C, January–April) were computed from ERA5 monthly mean 2 m air temperature (T2m), converted from Kelvin to

Celsius, following Equation (2) (where D_m denotes the number of days in month m), summed across January–April of each year. Regional GDD was calculated from the spatial mean of T2m across the ERA5 domain, whereas occurrence-level (local) GDD was calculated from the ERA5 grid cell nearest to each record (Section 2.5). ERA5 was preferred over WorldClim-derived climatologies for this calculation because it provides a continuous, physically consistent annual time series (1980–2024) rather than a single fixed-period climatological normal, allowing year-specific GDD to be matched to each year's phenological observations.

$$GDD = \sum_{m=1}^{4} \max(T_{2m,m} - 10\ ^{\circ}\mathrm{C},\ 0) \times D_m \quad (2)$$

2.4. Community-Level Ensemble Species Distribution Modelling

To model the actual and future geographic distribution of the five *Anthidium* species within a single joint framework, a community-level ensemble species distribution modelling (SDM) approach was adopted. In this approach, all five species were modelled using a shared set of predictor variables and a common target-group background pool (Section 2.1) rather than being fitted in isolation, allowing suitability patterns to be compared consistently across species and countries. For each species, three algorithms were implemented in parallel: Random Forest (RF) [36], Gradient Boosting (GB) [37], and Logistic Regression (LR). Model outputs were combined by calculating the arithmetic mean of suitability scores across all three algorithms per species, resulting in an ensemble suitability map (Equation (3)). Target-group background points were used at a 10:1 ratio, as described in Section 2.1. All models used the default settings of the scikit-learn library (Section 2.2), except that the Random Forest comprised 300 trees and Logistic Regression was fitted to standardised predictors with a maximum of 1000 iterations.

$$S_{ensemble} = \frac{1}{3}(P_{RF} + P_{GB} + P_{LR}) \quad (3)$$

Model performance was evaluated using two complementary cross-validation schemes: stratified random 5-fold CV, and spatial block CV using $1^{\circ} \times 1$ grid blocks (GroupKFold), the latter ensuring that spatially proximate records never appear in both training and test folds simultaneously—a conservative check against inflated performance due to spatial autocorrelation. For each fold, the area under the receiver operating characteristic curve (AUC-ROC) and the True Skill Statistic (TSS) were calculated for the ensemble prediction (arithmetic mean of RF, GB, and LR suitability scores); per-algorithm performance was not evaluated separately. Thresholds of AUC ≥ 0.70 and TSS ≥ 0.40 (TSS = sensitivity + specificity $- 1$) were adopted to indicate acceptable model performance [38,39]. Binary suitability was defined using a threshold equal to the training prevalence (presences/[presences + background] = $1/11 \approx 0.091$), which performs comparably to data-driven thresholds and is not tuned to the evaluation data [40]; the same threshold was used to calculate TSS. As a check, the maxSSS threshold (which maximises the sum of sensitivity and specificity) was computed for each species from five-fold out-of-fold predictions; it ranged from 0.09 to 0.11, and the resulting estimates of suitable area differed from those at the prevalence threshold by at most 4.9 percentage points under baseline climate and 12.6 percentage points under future scenarios (Table S6).

Habitat loss was defined, for each species and country, as the proportion of background cells classified as suitable under baseline climate (WorldClim v2.1, 1970–2000) that became unsuitable under future climate (2081–2100; five-GCM mean suitability), and habitat gain as the converse (Equation (4)). Net change was defined as the difference between the

numbers of suitable cells under future and baseline climate, expressed as a percentage of all evaluated cells (percentage points). Suitability was evaluated at the background cells of the 15 countries with confirmed target-species records (17,175 cells), that is, at locations where bees have been recorded rather than across a continuous grid.

$$\text{Loss}\backslash\% = \frac{n_{\text{suit,base}\cap\text{unsuit,fut}}}{n_{\text{suit,base}}} \times 100, \quad \text{Gain}\backslash\% = \frac{n_{\text{unsuit,base}\cap\text{suit,fut}}}{n_{\text{unsuit,base}}} \times 100 \tag{4}$$

To flag country-level estimates with limited support, we applied a pragmatic data-sufficiency criterion. A species–country estimate was considered adequately supported only if the country contained at least 20 baseline-suitable cells, at least 5% of its evaluated cells were classified as suitable, and at least 10 target-species occurrence records were available in that country. The first two conditions reduce instability caused by very small baseline denominators; the third prevents countries with only a handful of records from being treated as reliably characterised. A country met the criterion only if the estimates for all five species met it. Estimates not meeting this criterion were flagged with * and interpreted as directionally indicative only.

To ensure methodological consistency, the same eight VIF-selected predictor variables (bio1, bio2, bio4, bio8, bio9, bio14, bio15, bio19; Section 3.7), the same background pool and the same species-specific background samples were used in all species distribution models, so that the models evaluated by cross-validation (Section 3.3) are those used for the country-level habitat and CVI projections (Sections 3.5 and 3.6).

2.5. Phenological Trend Analysis

For each species, two complementary trend analyses were applied to the P10 time series. First, the statistical significance of temporal trends was assessed using the non-parametric Mann–Kendall trend test (Equation (5), where Φ denotes the standard normal cumulative distribution function; $\alpha = 0.05$), while the magnitude of the trend was estimated using the Theil–Sen median slope estimator (Equation (6)) [41,42]. Confidence intervals for the Theil–Sen slope estimates were calculated analytically at the 90% level using the theilslopes function from the scipy.stats module. As the Mann–Kendall *p*-values and Theil–Sen confidence intervals are derived from different statistical procedures, no direct numerical comparison was made between these two metrics. To assess sensitivity to the choice of phenological metric, trends were additionally estimated using a bias-corrected Weibull 10th percentile and the annual mean observation date.

To check whether each species' phenological trend remained consistent over time, Chow structural break tests (Equation (7), where RSS denotes the residual sum of squares, k is the number of estimated parameters (2, for intercept and slope), and n is the total number of observations; [43]) were used to look for candidate turning points—years at which the trend's slope changed abruptly—across the observation window for each species. When the relative deviation between the Theil–Sen slopes of the first and second halves (Equation (8)) exceeded 50%, the trend was operationally classified as non-stationary.

$$S = \sum_{i<j} \text{sign}(x_j - x_i); \quad Z = \frac{S \mp 1}{\sqrt{\text{Var}(S)}}; \quad p = 2(1 - \Phi(|Z|)) \tag{5}$$

$$\hat{\beta} = \text{median}\left\{ \frac{y_j - y_i}{x_j - x_i} : i < j \right\} \tag{6}$$

$$F = \frac{\left[RSS_{pooled} - (RSS_1 + RSS_2)\right]/k}{(RSS_1 + RSS_2)/(n - 2k)} \tag{7}$$

$$\Delta_{\%} = \frac{\left|\hat{\beta_2} - \hat{\beta_1}\right|}{\left|\hat{\beta_1}\right|} \times 100 \quad (8)$$

To investigate climate–phenology relationships, the association between P10 dates and spring growing degree days (GDD) was modelled using both linear regression and a spline-based generalised additive model approximation (GAM; DOY ~ spline (GDD_spring), implemented via scikit-learn's SplineTransformer).

Although the GAM exhibited high in-sample pseudo-R^2 values (0.15–0.88), five-fold cross-validation R^2 values (Equation (9)) were negative for all species (−0.29 to −59.0), indicating poor out-of-sample predictive performance consistent with overfitting in these small samples (n = 11–22) [44]. Cross-validated R^2 can be negative when prediction errors exceed the variation around the mean of the observations in a held-out fold. Because each fold contained few observations, values were particularly sensitive to low variation within a fold; the most extreme negative value should therefore not be interpreted as a proportional measure of overfitting. Given the discrepancy between in-sample fit and cross-validated performance, linear models were preferred for inferential analyses. To isolate the GDD–DOY relationship from shared temporal trends, a year-controlled partial correlation was additionally calculated (Equation (10)). The robustness of the local GDD association was examined across combinations of base temperature (5, 10 and 15 °C) and accumulation window (January–March, January–April, January–May and February–April), with FDR correction across all 60 species–specification combinations.

$$R^2_{cv} = 1 - \frac{\sum(y - \hat{y_{cv}})^2}{\sum(y - \bar{y})^2} \quad (9)$$

$$r_{partial} = \text{Pearson}(\text{resid}_{GDD}, \text{resid}_{DOY}), \qquad \text{resid}_x = x - (a + b \cdot \text{Year}) \quad (10)$$

2.6. Short-Term Extension of the Phenological Trend

To illustrate what continuation of the observed trend would imply over a policy-relevant horizon, the Theil–Sen slope (Section 2.5) was extended linearly from a 2012 reference point to 2050 (Equation (11)). The reference year is approximately the midpoint of the pooled species-specific qualifying observation years ($n \geq 3$ per year; Section 2.1), which are concentrated in the post-2000 period, and was chosen a priori to avoid endpoint bias. The extension was restricted to *A. florentinum*, the only species whose trend was significant after FDR correction (Section 3.1) and stationary (slope deviation between the two halves of the series was 11%). The shaded band in Figure 5 represents the analytical 90% confidence interval of the Theil–Sen slope, extended from the same reference point. The extension contains no thermal limits, photoperiodic constraints or climate forcing; it is therefore identical under SSP2-4.5 and SSP5-8.5 scenarios and is presented as a theoretical illustration of trend continuation, not as an ecological forecast. Longer extrapolations, and extrapolations for the four species with non-significant or non-stationary trends, were not made because linear continuation of their slopes leads to biologically implausible dates (Section 4).

$$\text{DOY}(t) = \text{DOY}_{2012} + \hat{\beta}\,(t - 2012),\ 2012 \leq t \leq 2050 \quad (11)$$

2.7. Phenological Debt Clock (PDC)

To contextualise whether each species' observed phenological shift keeps pace with the shift expected from local climate warming, we defined a heuristic Phenological Debt Clock (PDC). The term "debt" is used here as a conceptual hypothesis—a possible mismatch between the observed phenological change and the change expected from local warming—and not as an empirically measured biological deficit. For each species, a

climate-expected shift rate (days/decade) was derived by multiplying the occurrence-level (local) GDD–DOY sensitivity (s_local; Section 2.5) by the regional ERA5 growing-degree-day trend (t_GDD; –days/decade, 1980–2024; Equation (12)). The difference between the species' observed Theil–Sen rate and this expected rate (the "debt rate", days/decade; Equation (13)) was reported as a rate and was not extrapolated forward in time. Because this metric compounds two independently uncertain slope estimates, each species was classified into one of four confidence tiers based on Benjamini–Hochberg FDR-corrected significance (Section 2.2) of both the phenological trend and the local GDD-sensitivity test: CONFIRMED (both FDR-significant), PARTIAL (one of the two tests FDR-significant and the other nominally significant at $p < 0.05$ uncorrected but not after correction), SUGGESTIVE (both nominally significant at $p < 0.05$ uncorrected, but neither surviving FDR correction), or EXPLORATORY (all remaining combinations, including those in which neither test was nominally significant). The PDC is presented exclusively as a supplementary, hypothesis-generating indicator and was not used for any confirmatory inference.

$$\text{rate}_{\text{expected}} = s_{\text{local}} \times t_{\text{GDD}} \quad (12)$$

$$\text{rate}_{\text{debt}} = \text{rate}_{\text{observed}} - \text{rate}_{\text{expected}} \quad (13)$$

2.8. Composite Country Vulnerability Index

To provide an integrated assessment of climate-related vulnerability across countries in the Mediterranean Basin, a three-dimensional Composite Vulnerability Index (CVI) was developed. The CVI comprises the following components:

D1 (Habitat decline): The country-level net decline in climatically suitable cells under SSP5-8.5, calculated for each species as the decrease in suitable cells expressed as a percentage of all evaluated cells (set to zero where suitable area increased) and averaged across the five *Anthidium* species. Net decline was used rather than the loss percentage of Equation (4) because the latter is unstable when the baseline suitable area is very small.

D2 (DOY shift × habitat fraction): For each species, the absolute Theil–Sen P10 shift rate (days/decade; Section 2.5) multiplied by that species' D1 value expressed as a fraction; D2 is the mean of these products across the five species within a country.

D3 (Projected warming): The projected change in annual mean temperature (WorldClim bio1) between the 1970–2000 baseline and 2081–2100 under SSP5-8.5, calculated at each evaluated background cell (Section 2.4) as the mean of the five GCMs and averaged across the cells of each country. D3 therefore refers to the same future period and emission scenario as D1 and D2. As a sensitivity analysis, the CVI was recalculated with D3 defined as the observed Theil–Sen trend of annual mean near-surface temperature (ERA5, 1980–2024) at the ERA5 grid cells nearest to five randomly selected background records per country (Table S10).

Each component was normalised to a 0–1 range using min–max scaling (Equation (14)). The three components were then combined as a weighted average (Equation (15); D1:D2:D3 = 1/3:1/3:1/3 under the equal-weight scheme), yielding a CVI score ranging from 0 to 1. Following standard practice for vulnerability-index classification [45], an approximate high-risk reference threshold was set at the mean plus one standard deviation of the 15 countries' CVI scores (≈0.67), shown for visual reference in Figures 7 and 10. Because this threshold is defined relative to the 15 countries analysed, it identifies relatively high vulnerability within this set rather than an absolute level of risk. The value derived from the equal-weight scheme was also used as the reference line for the alternative weighting schemes (Section 3.9). No fixed categorical risk labels are defined beyond this

single reference line, as no external, cross-study calibration for such categories exists for this index.

$$D_{norm} = \frac{D - D_{min}}{D_{max} - D_{min}} \quad (14)$$

$$CVI = w_1 \cdot D_1 + w_2 \cdot D_2 + w_3 \cdot D_3, \;\; w_1 = w_2 = w_3 = \frac{1}{3} \quad (15)$$

As a sensitivity analysis for weighting schemes, two alternative configurations were also evaluated: a habitat-weighted scheme (D1:D2:D3 = 2:1:1) and a temperature-weighted scheme (D1:D2:D3 = 1:1:2).

3. Results

3.1. Phenological Shifts in Five Anthidium Species in the Mediterranean Basin (1980–2024)

The graphs presented in Figure 1 display, for each species-specific panel, the sampling bias-corrected 10th percentile first flight day of year (P10 date), with only species-years containing at least three unique observation days included in the analysis. The Theil–Sen median regression slopes fitted to each species' observation window are shown as dashed lines, while the 90% slope uncertainty intervals are represented by shaded bands (Theil–Sen; default implementation in scipy). The year of maximum structural instability, identified using Chow breakpoint tests, is indicated in each panel by a vertical dotted red line. These breakpoints mark the years of largest change in slope and are not attributed to any specific event. Panel annotations report the Theil–Sen slope (days decade^{-1}), the Mann–Kendall p-value (uncorrected), and flags for non-significance (n.s.) and non-stationarity where applicable. Titles explicitly note that significance markers are based on uncorrected p-values; FDR-corrected significance is reported separately in the text below; after FDR correction, only *A. florentinum* retained a significant trend ($q < 0.001$).

The results indicate pronounced non-stationarity in four of five species (*A. manicatum*, *A. oblongatum*, *A. loti*, *A. punctatum*; slope deviation between first and second halves exceeding 50%), constraining the validity of biological inference based on a single linear trend for these species. Only *A. florentinum* showed a statistically robust trend that survived Benjamini–Hochberg FDR correction ($p < 0.001$, $q < 0.001$; −10.0 days/decade), while *A. loti* showed a nominally significant trend ($p = 0.033$; −12.2 days/decade) that did not survive FDR correction ($q = 0.082$) and is interpreted as suggestive rather than confirmed. *Anthidium manicatum* ($p = 0.162$; −6.3 days/decade), *A. oblongatum* ($p = 0.065$; −19.0 days/decade), and *A. punctatum* ($p = 0.213$; −19.7 days/decade) showed no statistically significant trend, though all point estimates were in the same (advancing) direction. As a robustness check, the analysis was repeated with stricter minimum per-year sample thresholds ($n \geq 5$, $n \geq 10$ unique observation days per year per species); *Anthidium florentinum* remained significant at all thresholds, while results for the remaining species were unstable across thresholds (Table S1), reinforcing that only *A. florentinum* supports a statistically robust trend claim. The advance in *A. florentinum* was also detected with the bias-corrected Weibull 10th percentile (−13.8 days/decade, $p = 0.002$), whereas the annual mean observation date showed no significant trend (−2.1 days/decade, $p = 0.54$), suggesting that the signal is specific to the onset of activity rather than a shift in the entire flight period (Table S7). The *A. florentinum* trend did not depend on the inclusion of eastern Mediterranean records: it was −15.4 days/decade ($p = 0.008$) for Iberian records alone (a shorter series classified as non-stationary), −13.7 days/decade ($p \leq 0.004$) when French, Moroccan and Algerian records were added, and −14.3 days/decade ($p < 0.001$) when Italian records were also included, whereas the 219 eastern records (10 qualifying years) showed no significant trend on their own (Table S11). For *A. loti*, the 174 records of *A. loti* s.str. (86.1%) yielded a similar slope (−13.5 days/decade) that was not nominally significant ($p = 0.059$); the 28 *A. regulare*

records were too few for a separate analysis (three qualifying years). In six species-year combinations, removing a single observation year altered the statistical significance of the trend (leave-one-year-out sensitivity; Table S2), underscoring the fragility of trend estimates for species with fewer than ~15 years of qualifying data.

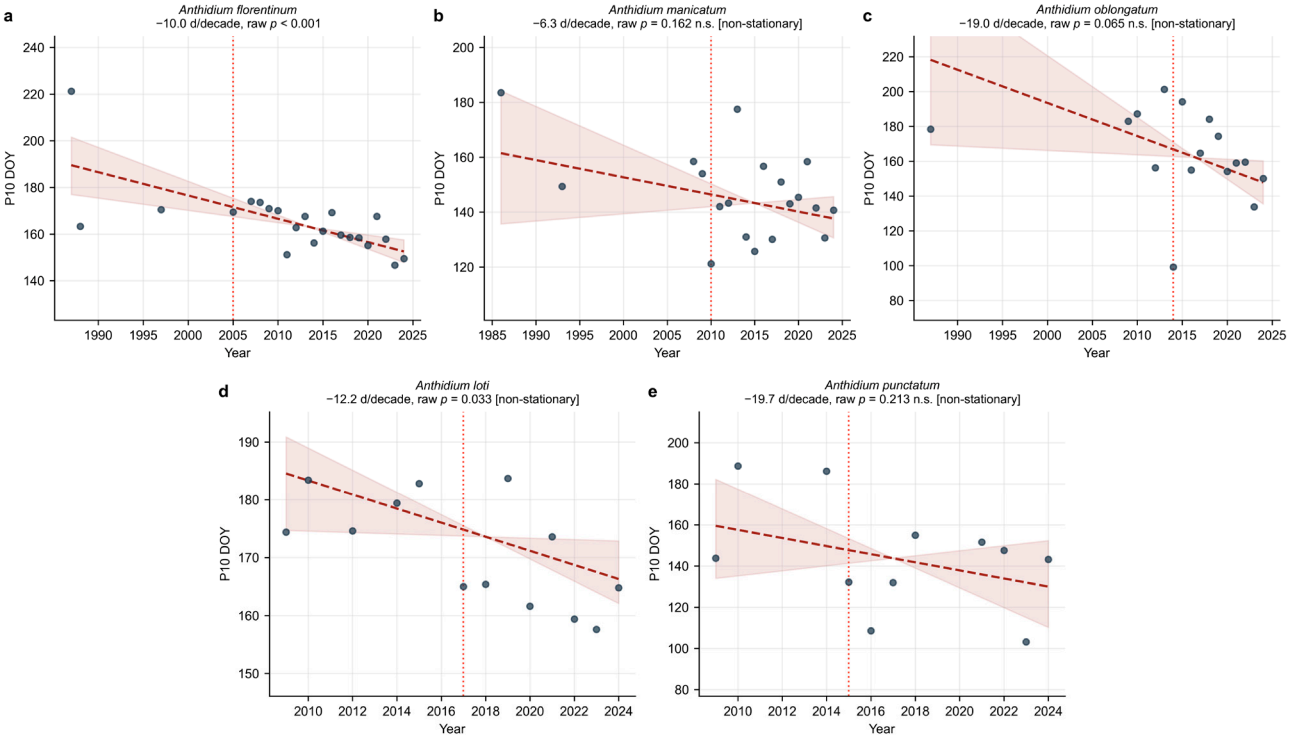


Figure 1. Phenological shifts in five wild bee species of the genus *Anthidium* (1980–2024). Panels: (**a**) *A. florentinum*, (**b**) *A. manicatum*, (**c**) *A. oblongatum*, (**d**) *A. loti*, (**e**) *A. punctatum*. Dark points show annual P10 values (day of year, DOY); the dashed line shows the Theil–Sen trend with its confidence band (shaded); the vertical dotted red line marks the Chow structural breakpoint; n.s., not significant; "non-stationary", slope differs between the two halves of the series. Panel annotations report the Theil–Sen slope (days decade^{-1}) and the Mann–Kendall *p*-value (uncorrected); FDR-corrected significance is reported separately (Section 3.1); after FDR correction, only *A. florentinum* remained significant ($q < 0.001$), whereas the nominally significant trend of *A. loti* ($p = 0.033$) did not ($q = 0.082$). A systematic leave-one-year-out sensitivity analysis, in which each observation year was removed in turn, is reported in Section 3.1: removing a single year altered trend significance in 6 of the tested species-year combinations, quantifying the extent to which individual years (including apparent outliers) influence the reported trends.

On the other hand, direct numerical comparison between the shaded uncertainty band and the Sen slope is not statistically appropriate, as these metrics are derived from different procedures—namely, the analytical Theil–Sen confidence interval and the analytical (normal-approximation) Mann–Kendall test, respectively.

3.2. Comparison of Linear and Generalised Additive Model (GAM) Climate-Phenology Responses in Five Anthidium Species

The graphs presented in Figure 2 illustrate, in the upper panels, the relationship between P10 date and spring growing degree days (GDD) using species-specific scatter plots; dashed lines represent ordinary least squares regression (in-sample R^2).

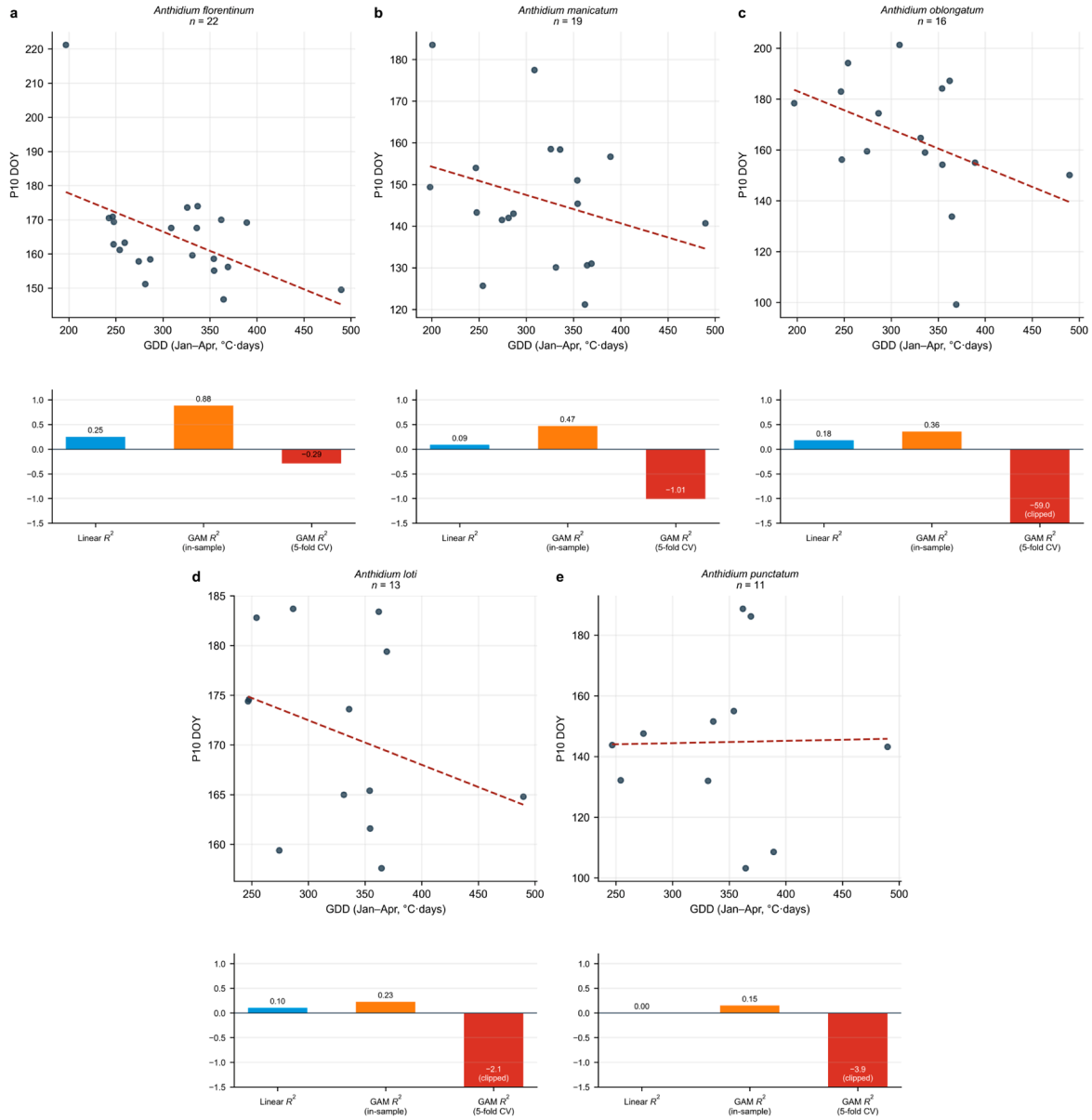


Figure 2. Comparison of linear and generalised additive model (GAM) climate-phenology responses in five *Anthidium* species. Panels: (**a**) *A. florentinum*, (**b**) *A. manicatum*, (**c**) *A. oblongatum*, (**d**) *A. loti*, (**e**) *A. punctatum*; *n*, number of years. Upper panels show species-specific scatter plots of P10 dates against spring growing degree days (GDD), with dashed lines representing ordinary least squares regression (in-sample R^2). Lower panels compare linear R^2 (blue), GAM in-sample pseudo-R^2 (orange), and GAM five-fold cross-validation (CV) R^2 (red) values; bars are clipped at −1.5, and the actual value is printed on the bar ("clipped").

The results indicate that the GAM exhibits substantial overfitting across all species under small sample sizes (n = 11–22). Although in-sample pseudo-R^2 values appeared high (0.15–0.88), five-fold CV R^2 values (Equation (9)) were strongly negative for all species (−0.29 to −59.0), meaning that out-of-sample predictions were less accurate than a horizontal line at the mean of the held-out observations, a degree of overfitting considerably more severe than typically reported, reflecting the very small sample sizes relative to spline flexibility. This finding is consistent with the literature highlighting the instability of flexible spline-based models in small ecological datasets [46–49]. In contrast, linear model in-sample R^2 values were modest (0.00–0.25), consistent with generally weak regional GDD–phenology relationships; notably, only *A. florentinum* retained a statistically

significant GDD–DOY partial correlation after controlling for year, and only when using occurrence-level (local) rather than regionally averaged GDD (Figure 3; partial $r = -0.47$, uncorrected $p = 0.028$; not significant after FDR correction, $q = 0.140$). This association was also sensitive to the GDD specification: across 60 combinations of species, base temperature and accumulation window, nominally significant partial correlations occurred only for *A. florentinum* (5 of 60 combinations), and none remained significant after FDR correction (Table S8). Accordingly, all inferential claims in this study are based exclusively on linear models, while GAM results are presented solely to illustrate the risk of overfitting rather than for inferential purposes. Because phenological trend extension (Section 2.6) is derived purely from the historical Theil–Sen trend and does not incorporate any GCM-derived climate forcing, it is, by construction, identical under SSP2-4.5 and SSP5-8.5; this scenario-independence reflects the projection method's design rather than an empirically tested calibration.

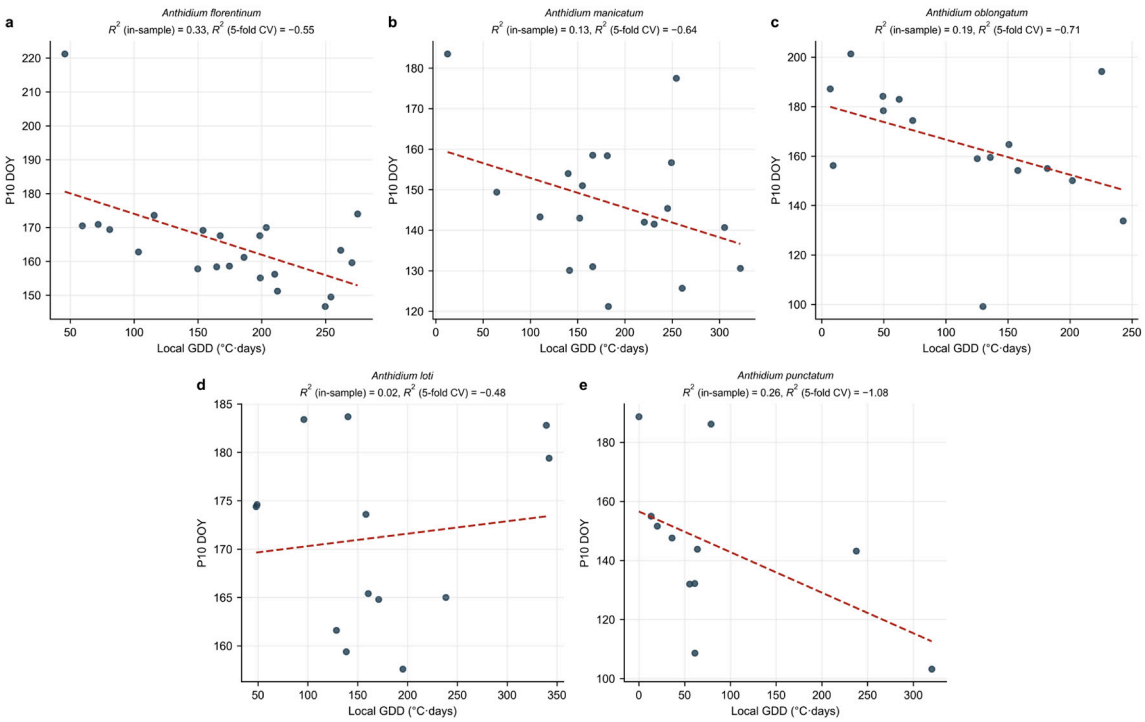


Figure 3. Comparison of occurrence-level (local) spring growing degree day (GDD) response for five *Anthidium* species, extracted at each record's own coordinate rather than as a regional average. Panels: (**a**) *A. florentinum*, (**b**) *A. manicatum*, (**c**) *A. oblongatum*, (**d**) *A. loti*, (**e**) *A. punctatum*. Points show annual P10 dates (day of year, DOY) against local GDD, and dashed lines show ordinary least squares regression; panel titles give the in-sample R^2 of the linear model and the five-fold cross-validated (CV) R^2. Linear R^2 values (in-sample) ranged from 0.02 to 0.33 across species, with only *A. florentinum* showing a statistically significant association (Section 3.2); five-fold cross-validated R^2 remained negative for all species, consistent with the regional-GDD model.

3.3. Community-Level Species Distribution Model (SDM) Validation Metrics in Five Anthidium Species

Figure 4 shows the performance of the ensemble model (arithmetic mean of RF, GB, and LR suitability scores) under both stratified random 5-fold and spatial block cross-validation schemes (Section 2.4); per-algorithm performance was not evaluated separately. The left panel shows the area under the receiver operating characteristic curve (AUC-ROC), while the right panel presents the True Skill Statistic (TSS).

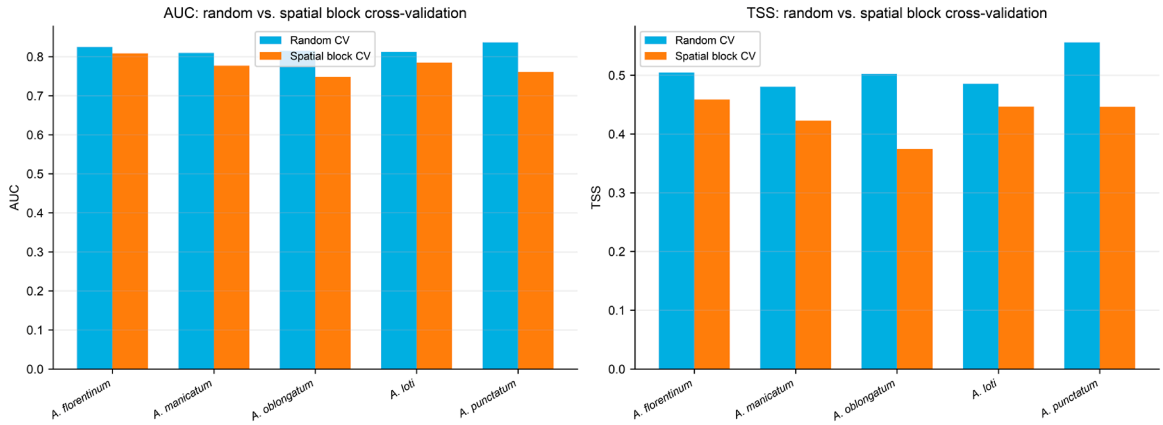


Figure 4. Community-level species distribution model (SDM) validation metrics in five *Anthidium* species. AUC = area under the receiver operating characteristic curve; TSS = True Skill Statistic. Both metrics are shown under random and spatial block cross-validation schemes (Section 2.4); TSS was calculated at the prevalence threshold (0.091).

Ensemble model performance (arithmetic mean of RF, GB, and LR suitability scores) was evaluated under both random and spatial block cross-validation. Under random CV, AUC ranged from 0.810 to 0.836 and TSS from 0.480 to 0.556. Under spatial block CV—a more conservative and ecologically realistic estimate of transferability to novel areas—AUC decreased to 0.748–0.809 and TSS to 0.374–0.458; all models exceeded the AUC threshold, and all but *A. oblongatum* (TSS = 0.374) exceeded the TSS threshold. The consistent decline from random to spatial CV (ΔAUC = 0.016–0.075) indicates that random cross-validation moderately overstates the models' ability to generalise to novel geographic areas, a well-documented consequence of spatial autocorrelation in occurrence data [50]. Habitat projections (Section 3.5) should therefore be interpreted as climatic associations within the sampled range. Multivariate Environmental Similarity Surface (MESS) analysis indicated that 7.7% (SSP2-4.5) and 26.0% (SSP5-8.5) of background cells fell outside the climatic range of the background calibration data under at least one of the five GCMs, mainly in North Africa, the eastern Mediterranean and the northern Adriatic–Balkan region (Figure S1; Table S12); projections for these areas, particularly under the SSP5-8.5 scenario, rely on climatic extrapolation and should be interpreted with strict caution.

3.4. Short-Term Extension of the First-Flight Trend of Anthidium florentinum to 2050

Figure 5 shows the observed P10 series of *A. florentinum* and its linear extension to 2050 (Section 2.6). Continuation of the observed trend would move the P10 date from DOY 164.6 (14 June) in 2012 to 146.6 (27 May) in 2030 and 126.6 (7 May) in 2050, an advance of 38 days (90% band of the slope: DOY 109.3–144.6 in 2050). Removing single observation years did not change the significance of the *A. florentinum* trend (Table S2). These values describe the continuation of the past trend and should not be read as predicted flight dates.

3.5. Projected Net Habitat Change by Country and Species Under the SSP5-8.5 Scenario (2081–2100)

In the analysis presented in Figure 6 (panel a: loss; panel b: gain), habitat loss was defined as the proportion of baseline-suitable background cells (ensemble suitability ≥ 0.091 under 1970–2000 climate) that fell below this threshold in 2081–2100, and habitat gain as the converse (Section 2.4). Future suitability was the mean of predictions under five CMIP6 GCMs (ACCESS-CM2, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL) for the SSP5-8.5 scenario. Values are shown per species and country, and countries are ordered by mean habitat loss across the five species (highest first).

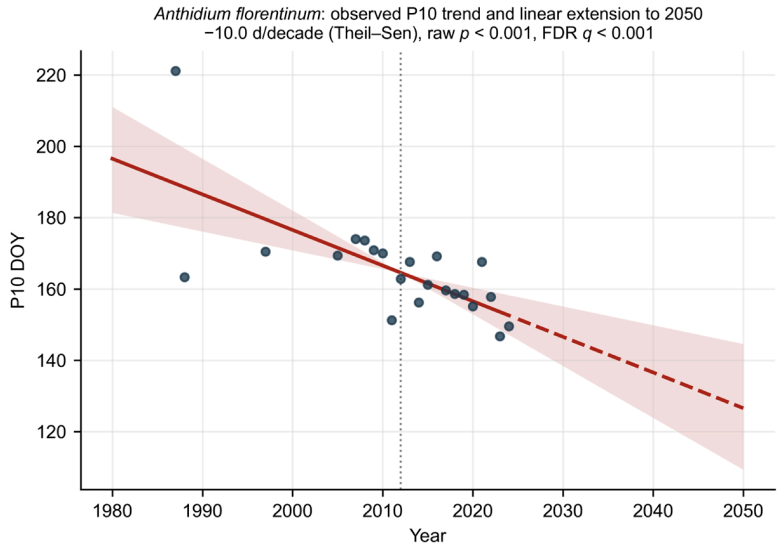


Figure 5. Observed first-flight dates (P10, 1980–2024) of *Anthidium florentinum* and linear extension of the Theil–Sen trend to 2050. Black points show annual P10 dates; the solid line shows the fitted trend over the observation period and the dashed line its extension; the shaded band is the 90% confidence interval of the slope, extended from the 2012 reference point (dotted vertical line). The extension is a theoretical illustration of trend continuation, not a forecast (Section 2.6).

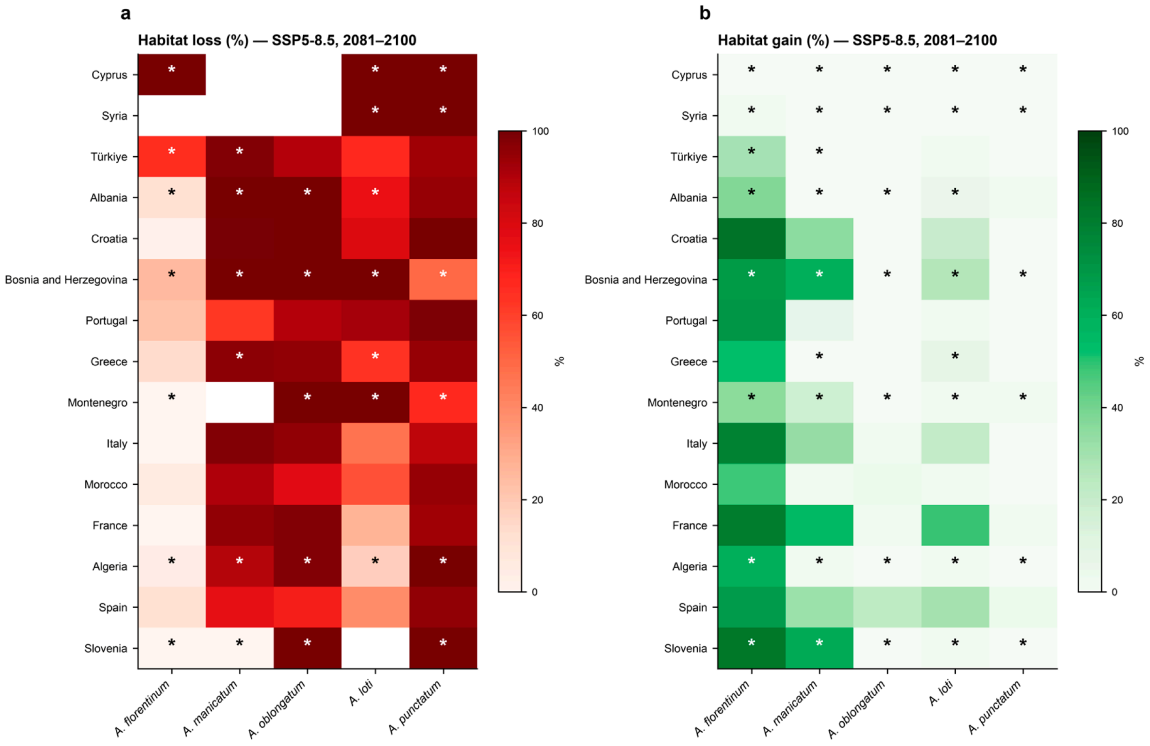


Figure 6. Projected habitat loss and gain by country and species under the SSP5-8.5 scenario (%, 2081–2100). (**a**) Habitat loss, as a percentage of baseline-suitable area; (**b**) habitat gain, as a percentage of baseline-unsuitable area. Countries with zero confirmed occurrence records for all five target species were excluded a priori (Section 2.1); * indicates species–country estimates that did not meet the data-sufficiency criterion defined in Section 2.4 (fewer than 20 baseline-suitable cells, less than 5% of evaluated cells suitable, or fewer than 10 target-species records in the country). Future suitability is the mean of five CMIP6 GCMs; countries are ordered by mean habitat loss across species. Because loss and gain percentages are calculated from different denominators (baseline-suitable versus baseline-unsuitable areas, respectively), a cell may be undefined (blank) in one panel but not the other when the relevant baseline area is zero; the * marker is shown wherever a value is present but unreliable.

Following the a priori exclusion of countries with no confirmed target-species occurrences (Section 2.1), 15 countries were retained for country-level analysis. Six of these met the data-sufficiency criterion defined in Section 2.4: Spain, Morocco, Portugal, Italy, France and Croatia. Averaged across species, the largest net declines in suitable cells were projected for Spain (−15.3 percentage points), Morocco (−13.7) and Portugal (−12.5), whereas Italy (+1.7), France (+8.0) and Croatia (+8.2) showed small net gains because gains of suitability at previously unsuitable locations exceeded losses (Table S3). Spain, Portugal and Morocco also showed net declines under SSP2-4.5 (−3.4, −3.9 and −9.9 percentage points, respectively), indicating that the direction of change was consistent across scenarios. The remaining nine countries did not meet the data-sufficiency criterion, either because fewer than 10 target-species records were available (Algeria, Bosnia and Herzegovina, Cyprus, Montenegro, Slovenia) or because baseline-suitable area was small for at least one species (Albania, Greece, Syria, Türkiye); these estimates are flagged with * in Figure 6 and should be read as directionally indicative only.

3.6. Country-Level Composite Vulnerability Index Under the SSP5-8.5 Scenario (2081–2100)

In Figure 7, the Composite Vulnerability Index (CVI) was calculated as a weighted average of three normalised components with equal weights (D1:D2:D3 = 1/3:1/3:1/3), resulting in values ranging from 0 to 1 (Section 2.8). The D1 component represents the mean net decline in climatically suitable cells across the five *Anthidium* species, D2 represents the species-averaged product of each species' absolute P10 shift rate and its habitat decline fraction, and D3 represents the normalised projected change in annual mean temperature between 1970 and 2000 and 2081–2100 under SSP5-8.5 (five-GCM mean), so that all three components refer to the same future period and scenario.

The results indicate that Morocco (CVI = 0.74), Spain (0.72) and Portugal (0.67) exceeded the approximate high-risk reference threshold (≈0.67), Portugal only marginally (0.668 vs. 0.666). All three met the data-sufficiency criterion (Section 2.4), and their high scores were driven mainly by habitat decline (D1) and its interaction with phenological shift (D2); Portugal combined the largest habitat decline with the second-lowest projected warming (4.8 °C). Türkiye (0.70) also exceeded the threshold but did not meet the data-sufficiency criterion, so its position should be interpreted with caution. The next-highest scores were observed for Italy (0.55), France (0.54), Algeria (0.52) and Croatia (0.49); for Italy and Croatia, these scores were driven largely by projected warming (D3) (Table S4). Projected warming ranged from 4.8 °C (Cyprus, Portugal) to 6.8 °C (Bosnia and Herzegovina). When D3 was instead based on the observed ERA5 temperature trend (1980–2024), the same four countries exceeded the corresponding threshold (≈0.62), and country scores were closely correlated with those based on projected warming (Spearman ρ = 0.96; Table S10). Several countries have very limited occurrence records (fewer than 10 in Algeria, Bosnia and Herzegovina, Cyprus, Montenegro and Slovenia), and their individual CVI rankings should accordingly be read as illustrative rather than robust country-specific assessments.

3.7. Stepwise VIF Elimination and Correlation Matrix of Retained Variables for the Mediterranean Basin Species Distribution Model

In Figure 8, a stepwise Variance Inflation Factor (VIF) elimination procedure was applied to the 19 WorldClim v2.1 BIO variables to address multicollinearity. At each step, the variable with the highest VIF value was removed until all remaining variables satisfied the condition VIF < 10, resulting in the retention of eight variables [34]: bio1 (annual mean temperature), bio2 (mean diurnal temperature range), bio4 (temperature seasonality), bio8 (mean temperature of the wettest quarter), bio9 (mean temperature of the driest quarter), bio14 (precipitation of the driest month), bio15 (precipitation seasonality), and bio19 (precipitation of the coldest quarter), with a final maximum VIF of 7.49 (bio1).

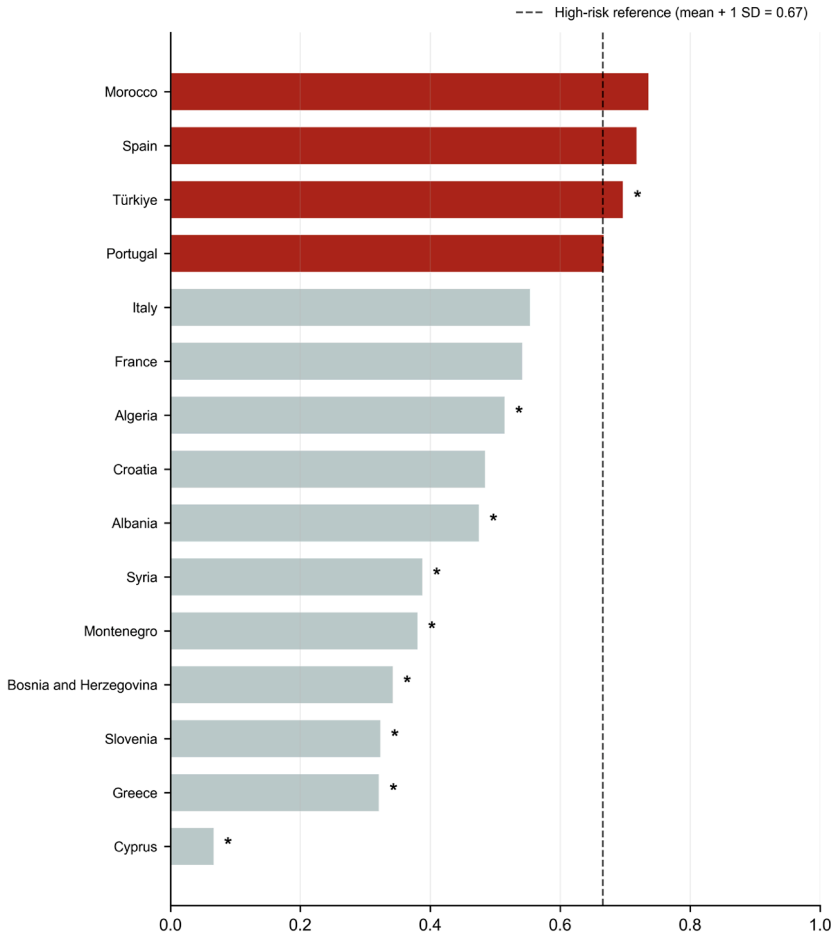


Figure 7. Country-level Composite Vulnerability Index under SSP5-8.5 (2081–2100), shown for the 15 countries with confirmed target-species occurrence records (Section 2.1). Dark red bars indicate countries at or above the high-risk reference threshold (dashed line; mean + 1 SD of the 15 countries' CVI scores, ≈0.67; Section 2.8); grey bars indicate countries below it. D3 is the projected change in annual mean temperature (five-GCM mean, SSP5-8.5). Am asterisk (*) marks countries whose D1/D2 components did not meet the data-sufficiency criterion (Section 2.4); only Croatia, France, Italy, Morocco, Portugal and Spain (unmarked) met this criterion.

The Pearson correlation matrix for the retained variables was calculated from WorldClim v2.1 (1970–2000) values extracted at the pooled presence and background locations. The highest correlation among variable pairs was observed between bio14 and bio15 ($r = -0.78$); both were retained because VIF—which accounts for shared variance with all other retained predictors jointly, not pairwise correlation alone—remained below the threshold of 10 for both variables. Nevertheless, this residual collinearity may increase coefficient variance in the Logistic Regression component of the ensemble model.

3.8. Phenological Debt Rate in Five Anthidium Species

In Figure 9, the Phenological Debt Clock (PDC) expresses, for each species, the difference (days per decade) between the observed rate of change in first-flight timing and the rate expected from local warming (Section 2.7). After Benjamini–Hochberg correction across the five-species test family, no species met the confirmatory threshold (CONFIRMED = 0 species). *Anthidium florentinum* falls in the PARTIAL tier: its phenological trend is FDR-significant ($q < 0.001$), while its GDD-sensitivity test is only nominally significant and does not survive FDR correction ($q = 0.140$). Its observed advance (−10.0 days/decade) exceeded the advance expected from its weak local GDD sensitivity

(−4.0 days/decade), giving a debt rate of −6.0 days/decade. The remaining four species fell in the EXPLORATORY tier (debt rates −3.8 to −15.1 days/decade); given the absence of a statistically robust underlying signal, these values are illustrative only and should not be interpreted as evidence of differential phenological debt across species. Because the PDC is a conceptual, hypothesis-generating indicator rather than a measured biological deficit, it was not used for any confirmatory inference.

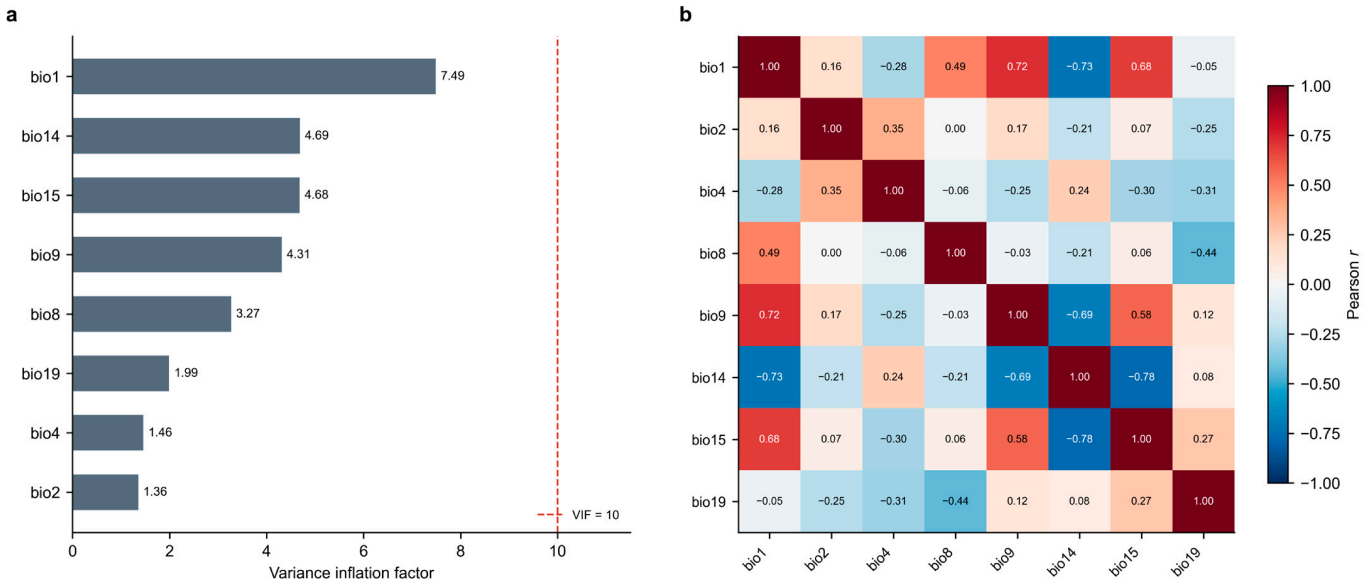


Figure 8. Stepwise VIF elimination and correlation matrix of retained variables for the Mediterranean Basin species distribution model. (**a**) Final Variance Inflation Factor of the eight retained variables (dashed line: VIF = 10); (**b**) Pearson correlation matrix, with colour scale from −1 (blue) to +1 (red).

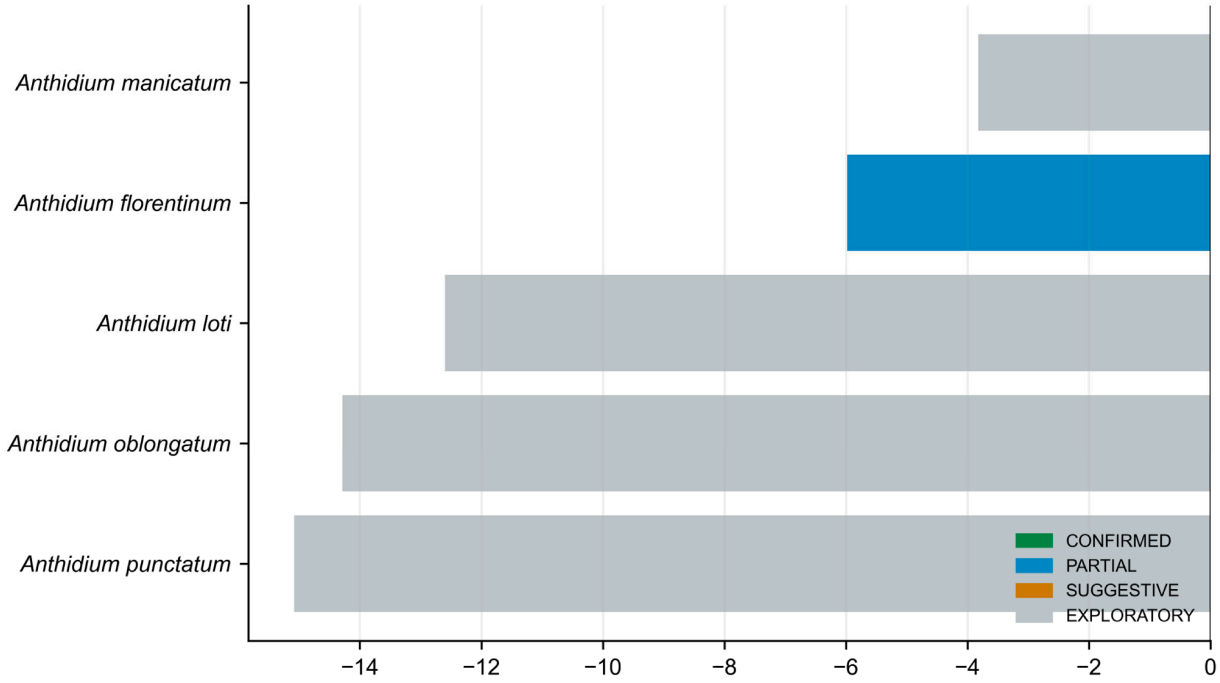


Figure 9. Phenological Debt Clock in five *Anthidium* species: phenological debt rate (days per decade; observed minus climate-expected shift in first-flight timing), coloured by confidence tier (CONFIRMED, PARTIAL, SUGGESTIVE, EXPLORATORY; Section 2.7). Negative values indicate an observed advance faster than expected from local warming. The PDC is a conceptual indicator, not a measured biological deficit.

3.9. Sensitivity of the Composite Vulnerability Index to Component Weighting Under the SSP5-8.5 Scenario (2081–2100)

Under the equal-weighting scheme, Morocco, Spain, Türkiye and Portugal exceeded the approximate high-risk threshold of 0.67 defined in Section 2.8 (Figure 10; derived from the equal-weight scheme and applied unchanged to the alternative schemes). Under the habitat-weighted scheme, the same four countries remained above the threshold, and the five highest-ranked countries were the same as under equal weighting (Morocco, Spain, Portugal, Türkiye, Italy). The temperature-weighted scheme altered the top tier: Türkiye ranked first, followed by Morocco, and only these two countries exceeded the threshold; Spain ranked third but fell below it, and Portugal, which had the second-lowest projected warming among the 15 countries, dropped to eleventh place.

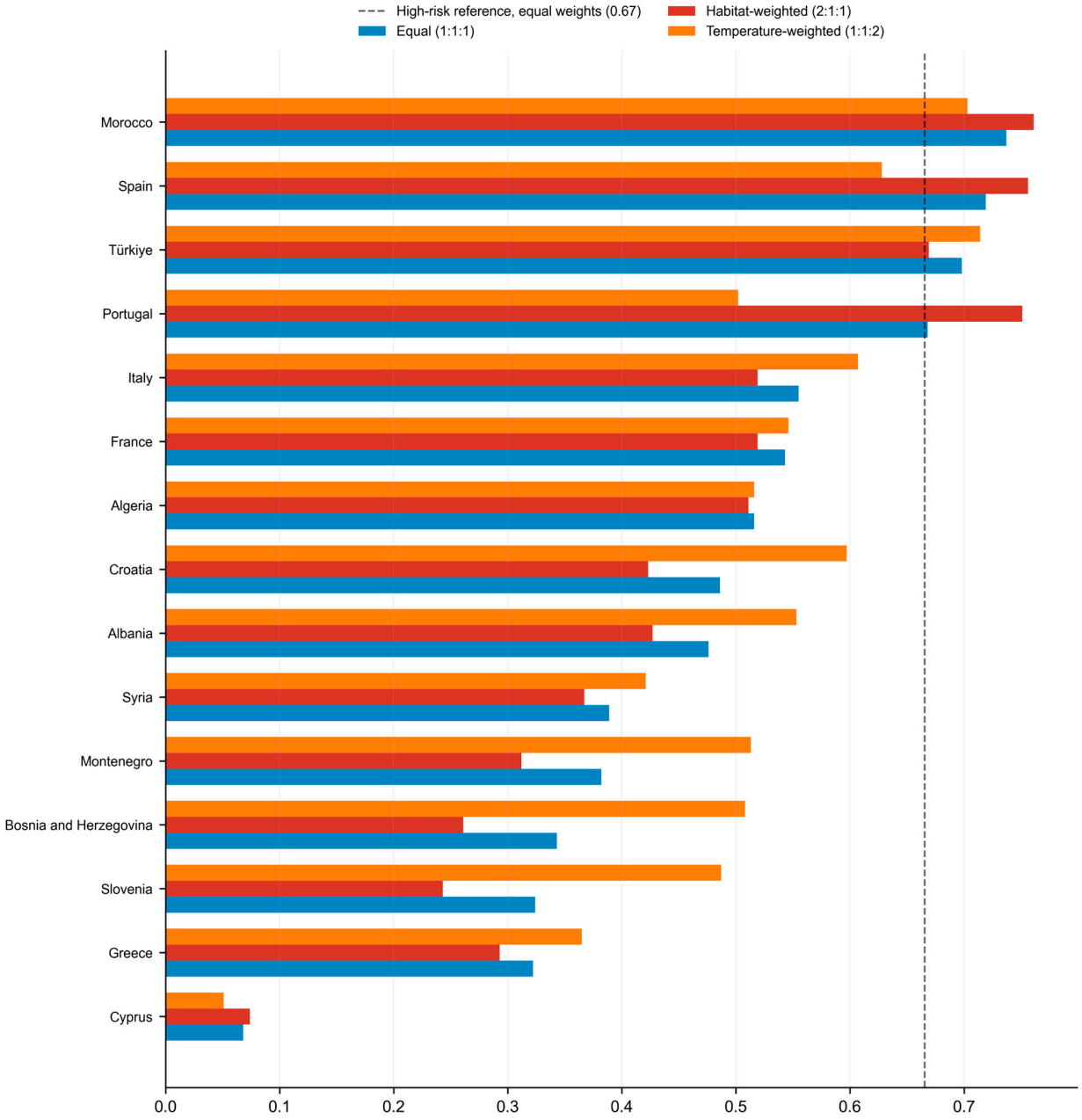


Figure 10. Sensitivity of the Composite Vulnerability Index to component weighting schemes under the SSP5-8.5 scenario (2081–2100). The dashed line marks the equal-weight high-risk reference threshold (≈0.67). D1 = habitat decline, D2 = phenological shift × habitat decline, D3 = projected change in annual mean temperature, five-GCM mean (Section 2.8).

Scores of countries with small habitat components but strong projected warming, such as Croatia, Montenegro, Bosnia and Herzegovina and Slovenia, increased markedly under the temperature-weighted scheme, confirming that the temperature component (D3) dominates the vulnerability scores of countries below the high-risk threshold (Section 3.6). Conversely, rankings under the habitat-weighted scheme were nearly identical to those under equal weighting.

Overall, Morocco, Spain, Türkiye and Italy ranked among the five highest-scoring countries under all three schemes, whereas the position of Portugal depended on the weight assigned to the temperature component. In the absence of empirical evidence to support alternative weighting schemes, the equal-weight assumption is retained as the primary, methodologically defensible choice [51].

4. Discussion

This study provides an integrated analytical framework examining phenological trends (1980–2024), climate–phenology relationships, projected habitat loss, and country-level composite vulnerability patterns for five *Anthidium* species in the Mediterranean Basin. Only *Anthidium florentinum*—the species with the largest sample size (n = 2598 records, 22 qualifying years)—showed a statistically robust, FDR-corrected significant advance in first flight timing (−10.0 days/decade). *Anthidium loti* s.l. showed a suggestive but not FDR-confirmed trend; the remaining three species showed no statistically significant trend, though all point estimates were directionally consistent (advancing). Slope deviations between the first and second halves of the observation window exceeded 50% for four of five species, underscoring that single-trend-line inference is not appropriate for most species in this dataset. This pattern also raises a question concerning the potential limitations of the underlying phenological metric itself: although Belitz et al. [29] found that 10th-percentile (P10) estimates generally performed well relative to more complex alternatives, they also found that estimator accuracy declines with smaller sample sizes—precisely the regime relevant to the sparsest years in this dataset. The median number of unique observation days per qualifying year was 33.5 for *A. florentinum* but only 8 for *A. loti* and 4 for *A. punctatum*, and fewer than five observation days were available in 18% of qualifying years for *A. florentinum*, 46% for *A. loti* s.l. and 64% for *A. punctatum* (Table S13); P10 estimates for the latter two species should therefore be regarded as imprecise. To assess whether this concern materially affected our conclusions, we repeated the trend analysis under progressively stricter minimum-sample thresholds ($n \geq 5$, $n \geq 10$; Section 3.1); *A. florentinum*'s trend remained significant across every threshold tested, whereas the marginal or non-significant trends in the other four species were themselves sensitive to threshold choice—indicating that the study's sole statistically confirmed result does not depend on inclusion of the sparsest, earliest observation years. This fragility is also evident when the slopes are extended forward: a linear continuation of the *A. punctatum* trend would fall below DOY 1 (before 1 January) by 2090, and that of *A. oblongatum* would reach the first days of the calendar year by 2100, outcomes that are biologically implausible for spring- and summer-active bees. For this reason, forward extension of phenological trends was restricted to *A. florentinum* and to 2050 (Section 2.6).

Within the climate–phenology context, GAMs showed severe overfitting under small sample sizes (cross-validated R^2, Equation (9), as low as −59.0), reinforcing the choice of linear models for inference. Ensemble species distribution models performed acceptably under random cross-validation (AUC 0.81–0.84, TSS 0.48–0.56) and declined moderately under spatial block cross-validation (AUC 0.75–0.81, TSS 0.37–0.46), indicating that random CV moderately overstates spatial transferability. Following the a priori exclusion of countries with no confirmed target-species occurrences (Section 2.1), six countries met the

data-sufficiency criterion defined in Section 2.4. Among these, net declines in climatically suitable cells were concentrated in the southwestern Mediterranean—Spain, Morocco and Portugal (12.5–15.3 percentage points under SSP5-8.5)—whereas France, Italy and Croatia showed small net gains, consistent with a poleward shift in climatic suitability. *Anthidium manicatum*'s projected regional range contraction should not be read as a global conservation concern: it is a cosmopolitan, highly adaptable species that continues to colonise new regions worldwide (including North America, South America, Australia, and New Zealand) ([21] and references therein), and its Mediterranean-basin sensitivity is better interpreted as a regional early-warning signal than as species-level extinction risk.

Forward extension of phenological trends was restricted to *A. florentinum* and to 2050 (Figure 5). These values are theoretical illustrations of what continuation of the observed trend would imply, not ecological forecasts: the extension (i) contains no thermal limits, photoperiodic constraints, or climate forcing and is therefore identical under SSP2-4.5 and SSP5-8.5; (ii) does not incorporate adaptive capacity, evolutionary responses, or interactions with host-plant phenology; and (iii) is subject to substantial parameter uncertainty given the limited number of qualifying years (n = 22). Habitat projections, in contrast, are based on the mean of five general circulation models, which captures part but not all of the inter-model uncertainty.

The difficulty of cleanly separating spatial and temporal climate signals from sparse, opportunistically collected occurrence data—a challenge also recently highlighted for dynamic occupancy modelling of wild bee losses [52]—underscores the need for the extensive sensitivity analyses reported throughout this study. Therefore, the Phenological Debt Clock (Figure 9) is presented as a complementary heuristic indicator—in which "debt" denotes a hypothesised mismatch between observed and temperature-expected change rather than an empirically measured biological deficit—and not as a validated mechanistic predictive tool; accordingly, all inferential claims made in this study rest exclusively on the linear trend analyses (Section 2.5) and the ensemble SDM cross-validation results (Section 2.4), not on the GAM comparisons or PDC rates, which are presented solely to illustrate methodological risk and contextual patterns, respectively.

Composite Vulnerability Index analyses identified Morocco, Spain and Portugal as the countries exceeding the approximate high-risk reference threshold among those meeting the data-sufficiency criterion (Portugal only marginally; Türkiye also exceeded it, but without sufficient data support), and their scores were driven mainly by projected habitat decline rather than by projected warming. For Italy and most lower-ranked countries, particularly in the Balkans, vulnerability scores were driven primarily by projected warming (D3). Because D3 refers to the same period and scenario as the habitat components, the index does not mix past and future conditions; replacing D3 with the observed 1980–2024 warming trend changed the ranking only slightly (Spearman ρ = 0.96) and identified the same countries above the threshold. Rankings were partly sensitive to the weighting scheme: Morocco, Spain, Türkiye and Italy remained among the five highest-ranked countries under all schemes, whereas Portugal dropped when the temperature component was up-weighted. Taken together, these results depict a geographically uneven Mediterranean response, with projected habitat decline concentrated in the Iberian Peninsula and north-western Africa and small net gains towards the northern margin of the study area.

Geographic sampling imbalance is a further constraint. Spain, France and Portugal together contribute 84% of the target-species records, whereas Morocco contributes 14 records and the other North African countries one or none, so the models are calibrated predominantly on Western European conditions. The target-group background partly mitigates this bias: because background cells come from records of six bee families collected by the same recording networks, they share the spatial sampling bias of the target records,

so the models contrast the climate of target-species records with that of sampled locations rather than with the full climatic range of the study area [31]. The cap of 3000 background cells per country further reduced the combined share of Spain, France and Portugal in the background from 62% to 45%, while North Africa contributed 6.4% of background cells compared with 0.3% of target records (Table S12). These measures cannot, however, supply information on the climatic niche of the species where they are hardly recorded. Under SSP5-8.5, novel climatic conditions (MESS < 0) affected 43% of evaluated cells in Morocco and 50% in Türkiye, compared with 2% in France and Portugal (Table S12), so projections for North Africa and the eastern Mediterranean rest on both sparse records and climatic extrapolation. Whether the projected habitat decline in the southwestern Mediterranean reflects a genuine ecological gradient or uneven data availability therefore cannot be fully resolved with the present dataset. Standardised monitoring of wild bees, including *Anthidium*, in North Africa and the eastern Mediterranean is urgently needed to test these projections and to calibrate future models across the full climatic range of the basin.

Taxonomic pooling is another potential source of bias. DNA barcodes of European *A. florentinum* fall into three clusters whose geographic limits have not yet been analysed [25], and eastern populations of *A. loti* s.l. belong to *A. regulare* [26]. Our records are dominated by the western Mediterranean: 59.4% of *A. florentinum* records come from the Iberian Peninsula and only 8.4% from the Balkans, Greece, Türkiye and Cyprus, while 86.1% of *A. loti* s.l. records belong to *A. loti* s.str. (Table S11). The *A. florentinum* advance was not driven by eastern records: it was at least as strong in every nested western subset, including the Iberian Peninsula alone (−15.4 days/decade), as in the full dataset (−10.0 days/decade), so this result does not depend on where the cluster boundaries lie. The *A. loti* trend, by contrast, lost nominal significance when restricted to *A. loti* s.str. ($p = 0.059$), which supports our interpretation of this trend as suggestive only. The species distribution models are likewise dominated by western records, but eastern cells were climatically partly distinct: 41% of the eastern *A. florentinum* cells (10.4% of all cells of the species) and 71% of the *A. regulare* cells (11.9% of *A. loti* s.l. cells) fell outside the climatic range of the corresponding western cells (MESS < 0), and eastern cells had higher precipitation seasonality and winter precipitation (bio15, bio19). Pooling may therefore slightly broaden the modelled climatic niches, and lineage-specific responses in the eastern Mediterranean cannot be assessed with the present data; this should be revisited once more georeferenced, genetically assigned records become available.

Phenological shifts in response to climatic change have been documented across a wide range of taxa. Earlier seasonal activity has been reported, e.g., in plants, birds, and various groups of insects, including advances in flowering dates and insect emergence [53–56]. Comparable evidence is also available specifically for bees. Wyver et al. [57], analysing presence-only records of 88 British wild bee species, reported an average advance of approximately 4 days per decade, whereas Bartomeus et al. [58] found an average advance of about 0.8 days per decade across ten North American bee species. Dorian et al. [59], analysing 70 solitary-bee species over 168 years in North America, found an overall advance of 0.43 days per decade, increasing to 1.62 days per decade after 1970. Their results further demonstrated substantial interspecific variation associated with ecological traits, including seasonality and nesting location. Elshoff et al. [60] found contrasting phenological responses to the same climate variation in various bumblebee (*Bombus* spp.) species. Similarly, Pawlikowski et al. [61] reported advances of 10–23 days in components of the main flight period of four common *Bombus* species in Central Europe over 35 years. These studies demonstrate that pronounced phenological responses are not restricted to a particular bee lineage, but may vary considerably among species. The study by Wyver et al. [62]

demonstrates that the ability to respond to changes in the climate niche (climatic envelope) greatly varies between species, with larger wild bee species with generalist diets and wide habitat breadths being significantly more successful at filling their potential climate envelopes in Great Britain than smaller, specialised species.

Our evidence for a substantial phenological advance in *A. florentinum*, with an FDR-confirmed trend of 10.0 days per decade, and directionally similar trends for the other four species examined, albeit with weaker or no FDR support, is thus relatively large compared with estimates reported for other bee taxa. The estimate for *A. florentinum* therefore falls towards the upper end of the range reported for bees, although direct comparisons should be made cautiously because the studies differ in geographic coverage, time span, data sources, phenological metrics, and statistical methodology. The magnitude of the trend observed in *A. florentinum* should nevertheless be interpreted with some caution. The phenological metric used here is based on an uncorrected percentile estimator applied to presence-only occurrence data. Such estimators can be sensitive to sample size and to the distribution of observations, particularly for percentiles close to the tails. Consequently, part of the observed magnitude may reflect characteristics of the available sampling record rather than a biological shift in exactly 10 days per decade.

An important question is whether such a seasonal shift necessarily results in ecological disruption. Evidence from North America suggests that bees and flowering plants can undergo broadly concurrent phenological advances, potentially maintaining temporal synchrony between pollinators and floral resources [58]. Conversely, large-scale analyses indicate that phenological shifts can alter the seasonal structure of pollinator assemblages. Duchenne et al. [63], for example, found that the mean flight date of more than 2000 European pollinator species shifted approximately six days earlier over 60 years, while the overall flight period shortened by about two days. They suggested that these changes may modify the seasonal distribution and overlap of pollination functions within pollinator assemblages.

The phenological advance documented here should therefore not be interpreted by itself as evidence of a negative population-level effect. Nor do our data establish warming as its cause: the advance coincides with regional warming, but the association between first-flight dates and spring growing degree days was weak and did not survive FDR correction (Section 3.2), and changes in recording effort, observer behaviour or the geographic composition of records could also shift percentile-based dates. We therefore describe the observed changes as climate-associated trends rather than climate-driven shifts. Indeed, phenological advancement and population change need not be directly coupled. Bonoan et al. [64], for example, found consistent advancement in the peak flight activity of the North American butterfly Fender's Blue (*Icaricia icarioides fenderi*) without a corresponding consistent decline in abundance. For *Anthidium*, determining the ecological consequences of the observed shift would require additional information on population trends, climatic drivers, floral-resource phenology, and potentially changes in the duration of the flight period.

For habitat loss, our projected net declines for Spain, Morocco and Portugal (12.5–15.3 percentage points of sampled locations) are broadly comparable to the range reported for European bumblebees, where climate change alone has already reduced suitable habitat by an average of 5% (locally up to 19%) since 1901 [65], and where 38–76% of species are projected to lose at least 30% of suitable territory by 2061–2080 under combined climate and land-use scenarios [66]. Several constraints of the distribution analyses should be noted. Habitat change was evaluated at grid cells where bees have been recorded rather than across a continuous land-surface grid, so country-level values describe sampled rather than total national area. The CVI reference threshold is relative to the 15 countries

analysed and does not represent an absolute level of risk. Finally, 26.0% of background cells experience climatic conditions outside the calibration range under SSP5-8.5 in at least one of the five GCMs, mainly in North Africa, the eastern Mediterranean and the northern Adriatic–Balkan region, where projections involve extrapolation.

5. Conclusions

Across the five Mediterranean *Anthidium* species, *A. florentinum* showed a robust, FDR-corrected advancement in first-flight timing that persisted when only Iberian records were analysed, whereas *A. loti* s.l. was suggestive and the remaining species yielded no statistically reliable temporal trend. At the country level, habitat estimates met the study's data-sufficiency criterion for six countries; net declines were concentrated in Spain, Morocco and Portugal, which also exceeded the study's approximate CVI reference threshold, whereas France, Italy and Croatia showed small net gains. The evidence thus points to uneven, species- and country-specific climate vulnerability rather than a uniform Mediterranean response.

Despite certain limitations, this study represents the first integrated phenological–distributional risk assessment for the genus *Anthidium*, combining long-term datasets, community-level ensemble modelling, and multidimensional vulnerability analytics within a single methodological framework at the scale of the Mediterranean Basin. Future research incorporating multiple phenological estimators, larger multi-GCM ensembles, floral resource data, host plant phenology interactions, landscape-level habitat connectivity, and local-scale adaptive capacity variables would help to determine the robustness of the observed patterns identified here and clarify their underlying mechanisms and ecological consequences. Given that the Mediterranean Basin is among the regions most sensitive to global climate change [13,14] and is considered a global biodiversity hotspot, the phenological trends and projected changes in climatic suitability of *Anthidium* bees documented here provide a reference framework for pollinator conservation policies in the region.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/insects17111097/s1, Table S1: Sensitivity of phenological trend estimates to the minimum per-year sample size threshold ($n \geq 3$, $n \geq 5$, $n \geq 10$); Table S2: Leave-one-year-out sensitivity analysis of phenological trends; Table S3: Projected habitat loss, gain and net change by species and country under SSP2-4.5 and SSP5-8.5 (2081–2100; five-GCM mean); Table S4: Composite Vulnerability Index (CVI) components (D3 = projected warming), total scores and sensitivity to weighting schemes (SSP5-8.5, 2081–2100); Table S5: Number of occurrence records and unique 2.5′ grid cells per species and of target-group background records; Table S6: Comparison of suitable area at the prevalence and maxSSS thresholds; Table S7: Phenological trends estimated with the 10th percentile, bias-corrected Weibull 10th percentile and mean observation date; Table S8: Sensitivity of the local GDD–DOY partial correlation to base temperature and accumulation window, with FDR correction; Table S9: Occurrence records by country and species and countries excluded from country-level analyses; Table S10: CVI calculated with the observed ERA5 temperature trend (1980–2024) as D3, compared with the version based on projected warming; Table S11: Occurrence records by lineage region, phenological trends of *A. florentinum* and *A. loti* s.l. recalculated by region, and climatic comparison of eastern and western cells; Table S12: Target-species records, target-group background cells before and after the per-country cap, and proportion of evaluated cells with novel climate (MESS < 0) by country; Table S13: Number of unique observation days behind each annual P10 estimate; Figure S1: Multivariate Environmental Similarity Surface (MESS) at background cells under SSP2-4.5 and SSP5-8.5 (2081–2100). For each cell, the minimum MESS across the five GCMs is shown; negative values indicate climatic conditions outside the range of the calibration data.

Author Contributions: Conceptualization, O.Ö. and M.K.; methodology, O.Ö.; software, O.Ö.; validation, O.Ö. and M.K.; formal analysis, O.Ö.; investigation, O.Ö.; resources, M.K.; data curation,

O.Ö.; writing—original draft preparation, O.Ö.; writing—review and editing, O.Ö., M.K. and H.K.; visualisation, O.Ö.; supervision, M.K.; project administration, O.Ö. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The occurrence data analysed in this study are publicly available from the Global Biodiversity Information Facility (GBIF) at https://doi.org/10.15468/dl.h3udkn (accessed on 15 April 2026). WorldClim v2.1 and CMIP6 climate layers are freely available at https://www.worldclim.org (accessed on 25 September 2026), and ERA5 reanalysis data are available from the Copernicus Climate Data Store at https://cds.climate.copernicus.eu (accessed on 24 August 2026). Derived datasets, Python analysis code (Python 3.13, including the analyses added during revision; package versions are given in Section 2.2) and Figure S1 supporting this study are openly available on Zenodo at https://doi.org/10.5281/zenodo.23351133. Additional supporting tables are provided in the Supplementary Materials.

Acknowledgments: This study used data from the GBIF network, comprising 69 datasets from 28 publishers, and we thank the institutions and individuals who collected, digitised, and mobilised these occurrence records. We acknowledge the WorldClim project for the bioclimatic and CMIP6 future-climate layers used in the species distribution models. This study contains modified Copernicus Climate Change Service information [2026]. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area Under the (ROC) Curve
CVI	Composite Vulnerability Index
CMIP6	Coupled Model Intercomparison Project Phase 6
FDR	False Discovery Rate
DOY	Day-of-year
MESS	Multivariate Environmental Similarity Surface
GB	Gradient Boosting
GAM	Generalised additive model
GCM	General Circulation Model
GDD	Spring growing degree days
LR	Logistic Regression
P10	10th-percentile day of year (first-flight estimator)
PDC	Phenological Debt Clock
RF	Random Forest
SSP	Shared Socioeconomic Pathway
SDM	Species distribution model
TSS	True Skill Statistic
VIF	Variance Inflation Factor

References

1. Saunders, D. Insect photoperiodism: Seasonal development on a revolving planet. *Eur. J. Entomol.* **2020**, *117*, 328–342. [CrossRef]
2. Hidalgo, S.; Chiu, J.C. Integration of photoperiodic and temperature cues by the circadian clock to regulate insect seasonal adaptations. *J. Comp. Physiol. A* **2024**, *210*, 585–599. [CrossRef] [PubMed]
3. Scranton, K.; Amarasekare, P. Predicting phenological shifts in a changing climate. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13212–13217. [CrossRef] [PubMed]
4. Klein, A.M.; Vaissière, B.E.; Cane, J.H.; Steffan-Dewenter, I.; Cunningham, S.A.; Kremen, C.; Tscharntke, T. Importance of pollinators in changing landscapes for world crops. *Proc. R. Soc. B* **2007**, *274*, 303–313. [CrossRef] [PubMed]

5. Potts, S.G.; Biesmeijer, J.C.; Kremen, C.; Neumann, P.; Schweiger, O.; Kunin, W.E. Global pollinator declines: Trends, impacts and drivers. *Trends Ecol. Evol.* **2010**, *25*, 345–353. [CrossRef] [PubMed]
6. Memmott, J.; Craze, P.G.; Waser, N.M.; Price, M.V. Global warming and the disruption of plant–pollinator interactions. *Ecol. Lett.* **2007**, *10*, 710–717. [CrossRef] [PubMed]
7. Parmesan, C. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Glob. Change Biol.* **2007**, *13*, 1860–1872. [CrossRef]
8. Freimuth, J.; Bossdorf, O.; Scheepens, J.F.; Willems, F.M. Climate warming changes synchrony of plants and pollinators. *Proc. R. Soc. B* **2022**, *289*, 20212142. [CrossRef] [PubMed]
9. Lang, W.; Zhang, Y.; Li, X.; Meng, F.; Liu, Q.; Wang, K.; Xu, H.; Chen, A.; Peñuelas, J.; Janssens, I.A.; et al. Phenological divergence between plants and animals under climate change. *Nat. Ecol. Evol.* **2025**, *9*, 261–272. [CrossRef] [PubMed]
10. Guisan, A.; Thuiller, W. Predicting species distribution: Offering more than simple habitat models. *Ecol. Lett.* **2005**, *8*, 993–1009. [CrossRef] [PubMed]
11. Araújo, M.B.; New, M. Ensemble forecasting of species distributions. *Trends Ecol. Evol.* **2007**, *22*, 42–47. [CrossRef] [PubMed]
12. Kerr, J.T.; Pindar, A.; Galpern, P.; Packer, L.; Potts, S.G.; Roberts, S.M.; Rasmont, P.; Schweiger, O.; Colla, S.R.; Richardson, L.L.; et al. Climate change impacts on bumblebees converge across continents. *Science* **2015**, *349*, 177–180. [CrossRef] [PubMed]
13. Giorgi, F. Climate change hot-spots. *Geophys. Res. Lett.* **2006**, *33*, L08707. [CrossRef]
14. Giorgi, F.; Lionello, P. Climate change projections for the Mediterranean region. *Glob. Planet. Change* **2008**, *63*, 90–104. [CrossRef]
15. Lionello, P.; Scarascia, L. The relation between climate change in the Mediterranean region and global warming. *Reg. Environ. Change* **2018**, *18*, 1481–1493. [CrossRef]
16. Myers, N.; Mittermeier, R.A.; Mittermeier, C.G.; Da Fonseca, G.A.; Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **2000**, *403*, 853–858. [CrossRef] [PubMed]
17. Thompson, J.D. *Plant Evolution in the Mediterranean: Insights for Conservation*; Oxford University Press: Oxford, UK, 2020.
18. Ghisbain, G.; Rosa, P.; Bogusch, P.; Flaminio, S.; Le Divelec, R.; Dorchin, A.; Kasparek, M.; Litman, J.; Mignota, M.; Müller, A.; et al. The new annotated checklist of the wild bees of Europe (Hymenoptera: Anthophila). *Zootaxa* **2023**, *5327*, 1–147. [CrossRef] [PubMed]
19. Orr, M.C.; Hughes, A.C.; Chesters, D.; Pickering, J.; Zhu, C.-D.; Ascher, J.S. Global patterns and drivers of bee distribution. *Curr. Biol.* **2021**, *31*, 451–458.e4. [CrossRef] [PubMed]
20. Reverté, S.; Miličić, M.; Ačanski, J.; Andrić, A.; Aracil, A.; Aubert, M.; Balzan, M.V.; Bartomeus, I.; Bogusch, P.; Bosch, J.; et al. National records of 3000 European bee and hoverfly species: A contribution to pollinator conservation. *Insect Conserv. Divers.* **2023**, *16*, 758–775. [CrossRef]
21. Kasparek, M. *The Resin and Wool Carder Bees (Anthidiini) of Europe and Western Turkey. Identification. Distribution. Biology.*; Chimaira: Frankfurt, Germany, 2022; 292p.
22. Danforth, B.N.; Minckley, R.L.; Neff, J.L. *The Solitary Bees. Biology, Evolution, Conservation*; Princeton University Press: Princeton, NJ, USA, 2019. [CrossRef]
23. Müller, A. Host-plant specialization in Western Palearctic Anthidine Bees (Hymenoptera: Apoidea: Megachilidae). *Ecol. Monogr.* **1996**, *66*, 235–257. [CrossRef]
24. GBIF Occurrence Download. Available online: https://doi.org/10.15468/dl.h3udkn (accessed on 15 April 2026).
25. Kasparek, M.; Boustani, M.; Fateryga, A.; Hölzler, G.; Ivgin Tunca, R.; Özgül, O.; Rasmont, P.; Schmid-Egger, C.; Soon, V.; Wood, T. The genetic barcode: A unique species identifier? Achievements and pitfalls in establishing a DNA barcoding reference library for European wool carder and resin bees (Megachilidae: Anthidiini). *Acta Entomol. Mus. Natl. Pragae* 2026, *in press*.
26. Kasparek, M.; Biella, P.; Ramazzotti, F.; Galimberti, A. Integrative taxonomy reveals cryptic speciation in the anthidiine wool carder bee *Anthidium loti* Perris, 1852 (Hymenoptera: Megachilidae) in the Western Palaearctic. *Eur. Zool. J.* **2025**, *92*, 1680–1694. [CrossRef]
27. Sentil, A.; Miličić, M.; Benrezkallah, J.; Ačanski, J.; Andrić, A.; Aubert, M.; Bartomeus, I.; Biella, P.; Boustani, M.; Carstensen, L.B.; et al. Synthesised database of wild bee and hoverfly records in Europe. *Sci. Data* **2026**, *13*, 227. [CrossRef] [PubMed]
28. Brooks, S.J.; Self, A.; Toloni, F.; Sparks, T. Natural history museum collections provide information on phenological change in British butterflies since the late-nineteenth century. *Int. J. Biometeorol.* **2014**, *58*, 1749–1758. [CrossRef] [PubMed]
29. Belitz, M.W.; Larsen, E.A.; Ries, L.; Guralnick, R.P. The accuracy of phenology estimators for use with sparsely sampled presence-only observations. *Methods Ecol. Evol.* **2020**, *11*, 1273–1285. [CrossRef]
30. Olsen, K.; Holm, T.E.; Pape, T.; Simonsen, T.J. Natural history museum collection and citizen science data show advancing phenology of Danish hoverflies (Insecta: Diptera, Syrphidae) with increasing annual temperature. *PLoS ONE* **2020**, *15*, e0232980. [CrossRef] [PubMed]
31. Phillips, S.J.; Dudík, M.; Elith, J.; Graham, C.H.; Lehmann, A.; Leathwick, J.; Ferrier, S. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol. Appl.* **2009**, *19*, 181–197. [CrossRef] [PubMed]

32. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
33. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]
34. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 7th ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2010.
35. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
36. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
37. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
38. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232. [CrossRef]
39. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [CrossRef] [PubMed]
40. Liu, C.; Berry, P.M.; Dawson, T.P.; Pearson, R.G. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **2005**, *28*, 385–393. [CrossRef]
41. Sen, P.K. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]
42. Theil, H. A rank-invariant method of linear and polynomial regression analysis. *Indag. Math.* **1950**, *12*, 173–185.
43. Chow, G.C. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **1960**, *28*, 591–605. [CrossRef]
44. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009.
45. Rogelis, M.C.; Werner, M.; Obregón, N.; Wright, N. Vulnerability indicators for natural hazards: An innovative selection and weighting approach. *Sci. Rep.* **2019**, *9*, 15254. [CrossRef] [PubMed]
46. Servén, D.; Brummitt, C. pyGAM: Generalized Additive Models in Python. *Zenodo* **2018**. [CrossRef]
47. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
48. Hastie, T.J.; Tibshirani, R.J. *Generalized Additive Models*; Chapman & Hall/CRC: London, UK, 1990.
49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. Available online: https://jmlr.org/papers/v12/pedregosa11a.html (accessed on 19 October 2026).
50. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
51. Füssel, H.-M.; Klein, R.J.T. Climate change vulnerability assessments: An evolution of conceptual thinking. *Clim. Change* **2006**, *75*, 301–329. [CrossRef]
52. Powney, G.D.; Bullock, J.M.; Boyd, R.J.; Carvell, C.; Edwards, M.; Edwards, R.; Kunin, W.E.; Isaac, N.J.B. Challenges in modelling spatio-temporal climatic correlates of local losses of wild bees using dynamic occupancy models. *Divers. Distrib.* **2025**, *31*, e70055. [CrossRef]
53. Bloomfield, J.P.; Bowler, D.E.; Gripenberg, S.; Harvey, M.C. Citizen science data reveals that earlier and shorter flight periods correlate with climate warming in the dark-edged bee-fly. *Insect Conserv. Divers.* **2026**, 1–13. [CrossRef]
54. Crick, H.Q.; Sparks, T.H. Climate change related to egg-laying trends. *Nature* **1999**, *399*, 423–424. [CrossRef]
55. Büntgen, U.; Piermattei, A.; Krusic, P.J.; Esper, J.; Sparks, T.; Crivellaro, A. Plants in the UK flower a month earlier under recent warming. *Proc. R. Soc. B* **2022**, *289*, 20212456. [CrossRef] [PubMed]
56. Hassall, C.; Owen, J.; Gilbert, F. Phenological shifts in hoverflies (Diptera: Syrphidae): Linking measurement and mechanism. *Ecography* **2017**, *40*, 853–863. [CrossRef]
57. Wyver, C.; Potts, S.G.; Edwards, M.; Edwards, R.; Roberts, S.; Senapathi, D. Climate-driven phenological shifts in emergence dates of British bees. *Ecol. Evol.* **2023**, *13*, e10284. [CrossRef] [PubMed]
58. Bartomeus, I.; Ascher, J.S.; Wagner, D.; Danforth, B.N.; Colla, S.; Kornbluth, S.; Winfree, R. Climate-associated phenological advances in bee pollinators and bee-pollinated plants. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20645–20649. [CrossRef] [PubMed]
59. Dorian, N.N.; McCarthy, M.W.; Crone, E.E. Ecological traits explain long-term phenological trends in solitary bees. *J. Anim. Ecol.* **2023**, *92*, 285–296. [CrossRef] [PubMed]
60. Elshoff, K.C.; Finn, S.R.; Treanore, E.D.; Richardson, L.L.; Crone, E.E. Bumble bee species display contrasting phenological responses to climate variation. *Ecology* **2026**, *107*, e70385. [CrossRef] [PubMed]
61. Pawlikowski, T.; Sparks, T.H.; Olszewski, P.; Pawlikowski, K.; Rutkowski, L.; Jakubowski, R. Rising temperatures advance the main flight period of *Bombus* bumblebees in agricultural landscapes of the Central European Plain. *Apidologie* **2020**, *51*, 652–663. [CrossRef]

62. Wyver, C.; Potts, S.G.; Roberts, S.P.M.; Senapathi, D. Life-history traits predict the ability of British wild bees to fill their climate envelopes. *Ecol. Entomol.* **2026**, *51*, 31–40. [CrossRef]
63. Duchenne, F.; Thébault, E.; Michez, D.; Elias, M.; Drake, M.; Persson, M.; Rousseau-Piot, J.S.; Pollet, M.; Vanormelingen, P.; Fontaine, C. Phenological shifts alter the seasonal structure of pollinator assemblages in Europe. *Nat. Ecol. Evol.* **2020**, *4*, 115–121. [CrossRef] [PubMed]
64. Bonoan, R.E.; Crone, E.E.; Edwards, C.B.; Schultz, C.B. Changes in phenology and abundance of an at-risk butterfly. *J. Insect Conserv.* **2021**, *25*, 499–510. [CrossRef]
65. De Tandt, B.; Serres, K.; Pietroiusti, R.; Erazo, D.; Massonnet, F.; Michez, D.; Thiery, W.; Ghisbain, G.; Dellicour, S. Declines in European bumblebee habitat suitability attributable to climate change. *Nat. Clim. Change* **2026**, *16*, 1222–1227. [CrossRef]
66. Ghisbain, G.; Thiery, W.; Massonnet, F.; Erazo, D.; Rasmont, P.; Michez, D.; Dellicour, S. Projected decline in European bumblebee populations in the twenty-first century. *Nature* **2024**, *628*, 337–341. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.