Article

Poor air quality raises mortality in honey bees, a concern for all pollinators

Check for updates

Nico Coallier [1] ✉, Liliana Perez [2], Maxime Fraser Franco[1,3], Yenny Cuellar [2] & Julien Vadnais [2]

Human well-being relies on the presence and role of pollinators, as they contribute to the vitality of ecosystems, support the reproduction of wild plants, increase crop yields, and strengthen overall food security. While wild bee populations are dwindling due to climate and environmental change, there has been a notable 45% rise globally in the number of managed honey bee (*Apis mellifera*) colonies over the past five decades. Given their economic significance and their relative ease of tracking, honey bees have the potential to serve as bioindicators of global pollinator health. Consequently, honey bees have emerged as a keystone species requiring protection and conservation efforts. Here, we investigate the intricate relationship between air quality, environmental factors, and honey bee mortality across Canada and the United States. Using statistical and machine learning modeling, our findings underscore the honey bee's role as a bioindicator. We found that air quality is an important predictor of honey bee mortality. The risk of honey bee mortality increased with poor air quality (ozone and Air Quality Health Index) but was substantially reduced in regions with greater vegetation availability (Normalized Difference Vegetation Index). Therefore, our study offers a beacon of hope: improving management practices by increasing greenery can significantly mitigate the impact of deteriorating air quality on honey bees, providing a vital solution to safeguard our essential pollinators.

Over the past few years, there has been a rapid increase in extreme events like heat waves and droughts, primarily driven by human-induced climate change[1]. For instance, wildfires in high northern latitude regions are expected to increase in upcoming years[2]. Climate change is altering landscape fire patterns in numerous areas, leading to smoke emissions that carry various harmful air pollutants, posing potential risks to human health[3]. The large quantities of atmospheric pollutants, such as $PM_{2.5}$, contribute to the deterioration of air quality during the fire season[2]. Insects are also threatened by those chemicals. When present in sufficient concentrations, atmospheric pollutants from fire smoke can impair foraging and the development of some insect species such as honey bees, with potential consequences for population dynamics[3,4].

Pollinators are among the most threatened insects, facing notorious declines due to habitat loss, extreme weather events, and agro-chemical pollutants[5–7]. Importantly, their nutrition depends on access to flowering plants[8], and their capacity to locate them using scent has been found to be negatively affected by air pollution[9,10]. Research shows that elevated ozone concentrations (O_3) can affect pollinators' nutrition by degrading floral scent[9,11], by disrupting pheromone diffusion[12], and by reducing overall foraging activity[13,14]. In honey bees, air pollutants block chemical communication, impeding them from coordinate swarming by weakening the electroantennography response of their antennae to alarm pheromones[15]. Moreover, elevated levels of $PM_{2.5}$ in the atmosphere affect skylight polarization, which honey bees use for navigation during their foraging trips, resulting in longer trip durations[16,17]. Interactions between such stressors can synergistically affect honey bees' ability to withstand stress[18], suggesting that air pollution may compromise honey bee survival. Yet, few studies have assessed the potential impacts of air pollution on pollinator survival, often constrained by limited access to large-scale data on these animals (but see[18]). Empirical studies suggest that ensuring access to a high-quality diet can enhance honey bee resilience to climatic stressors such as heat[19]. Indeed, the availability and quality of vegetation resources are important regulators of honey bee populations and health[20,21]. In addition, vegetation cover has been proven to reduce various types of air pollutants such as PM_{10} and O_3[22–24]. However, other studies advance that air quality is probably the main driver of bee health and not vegetation availability, but these analyses were performed without considering their interactive effect[18]. Hence, the relationship between vegetation and bees' resilience to increased air contaminant levels remains unclear[3].

¹Nectar Technologies Inc., 6250 Rue Hutchison #302, Montréal, QC, H2V 4C5, Canada. ²Laboratoire de Géosimulation Environnementale (LEDGE), Département de Géographie, Université de Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal, QC, H2V 0B3, Canada. ³Département des Sciences Biologiques, Université du Québec à Montréal, Case postale 8888, Succursale Centre-ville, Montréal, QC, H3C 3P8, Canada. ✉e-mail: nico@nectar.buzz

Honey bees play a pivotal role as pollinators in (agro)ecosystems, enhancing wild plant variety and pollinating a large portion of the world's food resources[20]. Thus, using honey bees as bioindicators[25,26] of poor air quality would help determine whether and how vegetation availability mitigates its adverse effects, enabling the development of better biodiversity conservation, beekeeping, and land management practices. Here, we assess the impacts of air quality and vegetation availability on honey bee colony survival, combining statistical modeling and machine learning techniques tailored for survival analysis. Specifically, we address the following questions: (1) Will honey bee mortality increase with poor air quality? (2) Will high vegetation availability mitigate the negative impacts of poor air quality? Based on the model's predictions and air quality data for 2024, we developed a risk map of honey bee mortality for 2024.

To assess the impact of various air quality-related stressors on honey bee colony survival, our analysis incorporated a range of environmental factors proven to influence bee population health[16,21,27–30], including the daily average Air Quality Health Index (AQHI)[31], daily average ozone concentration (O_3 µg m^{-3}), daily average temperature (°C), daily total precipitation (mm), daily average wind speed (km h^{-1}), and the Normalized Difference Vegetation Index (NDVI) as a proxy of vegetation availability[18] (see Methods for further details). We employed anonymized data with a sample size of 103 477 hives tracked over three beekeeping seasons across Canada and the United States from 2020 to 2023 (Supplementary Fig. 1). We performed the analyses using two distinct methodologies: (1) one involved statistical modeling using generalized linear mixed models (GLMM) to account for differences in honey bee mortality among beekeeping operations, and (2) a machine learning algorithm, the Random Survival Forest (RSF)[32] to predict beehive mortality and evaluate which environmental factors were most important for predicting bee mortality. Our results show that air quality is an important predictor of beehive mortality, as mortality was greater in regions of poor air quality in Canada and the U.S. We also show that greater vegetation availability dampens the negative effects of poor air quality, suggesting that proper landscape management practices may attenuate the consequences of poor air quality on honey bee survival.

Results

For the statistical analysis, we report the best out of 33 GLMMs based on the Akaike Information Criterion (Supplementary Table 1), all predicting honey bee colony mortality (y). We found strong evidence that honey bee mortality increased with elevated ozone concentrations (Table 1 and Supplementary Fig. 2). Indeed, the baseline probability of mortality at mean ozone concentrations was 45.8%, increasing up to 73.5% (i.e. +27.8% increase) for a one unit increase in standard deviation of ozone concentration. We also found evidence of a strong effect of vegetation availability, where greater vegetation was associated with a reduced likelihood of bee mortality (Table 1 and Supplementary Fig. 2). As expected, both variables interacted with each other, such that greater vegetation availability attenuated the adverse effects of ozone on beehive mortality (Table 1 and Fig. 1). In addition, the risk of beehive mortality was reduced when wind speeds were slower, but the effect was weaker compared to ozone concentration and vegetation availability (Table 1 and Supplementary Fig. 2). Therefore, we did not find evidence that wind speed strongly attenuated the effect of ozone concentration. Other interactions were weak (Table 1). Overall, the fixed effects explained 15.5% of the total variance in beehive mortality ($R^2_{marginal} = 0.155$, Table 1).

The random effects show an important variance in intercepts and slopes among beekeeping operations ($R^2_{conditional} = 0.734$, Table 1), suggesting that the impact of ozone concentration on beehive mortality varies substantially among different locations (Supplementary Fig. 3). Moreover, operations that had greater average mortalities also displayed a steeper curve for ozone concentrations, revealing a more pronounced effect of ozone on beehive mortality for such operations (i.e. random effects correlation, Table 1 and Supplementary Fig. 3). These results highlight differences in how each beekeeping operation reacts to poor air quality, indicating local

and regional variation in the relationship between beehive mortality and air quality.

To assess the importance of environmental variables in predicting beehive mortality, we trained a RSF on 85% of the dataset and assessed its performance on the remaining 15%. The variables used for prediction were the average AQHI, the average NDVI, average wind speed, average precipitation, and average temperature. Optimal parameters for the RSF were selected using a random search. We evaluated the model's performance using Uno et al.'s (2011) modified concordance index[33] and the time-dependent AUC[34–36] (see Methods for further details). The model scored 76.5% for the concordance index, indicating that 76.5% of the compared pairs aligned with the survival event and its associated probability. The time-dependent AUC indicates that the model performed well, with an average AUC score of 93.6% (Supplementary Fig. 4).

The analysis of variable importance (see Methods for further details) revealed that the most important variables in predicting beehive mortality were precipitation, air quality (AQHI), and temperature, all having similar mean importance scores (Fig. 2). Contrary to our expectations, wind speed and vegetation availability (NDVI) were not as important as the aforementioned variables, having lower importance scores (Fig. 2).

Using the risk scores predicted by the RSF and air quality data from the OpenWeather API, we computed a risk map for honey beehives in 2024 for Canada and the U.S. (Fig. 3). Predictions indicate that the risks of honey bee mortality are greater in western U.S. as well as in the Northeast coast, regions characterized by lower air quality in 2024. Hence, conservation strategies could incorporate such risk maps to enable targeted approaches for protecting honey bee populations, directing resources towards areas with increased vulnerability due to poor air quality.

Discussion

Combining statistical and machine learning modeling with an extensive longitudinal dataset on honey bee health, our study shows that air quality is an important predictor of beehive mortality. Our predictive approach enabled us to build a risk map, based on air quality data, that outlines regions of critical importance for bee health monitoring. Importantly, we show that vegetation availability attenuates the negative impacts of poor air quality on beehive mortality. Therefore, our work underlines the value of using honey bees as bioindicators of poor air quality to encourage proper land management practices, and ultimately, protect multiple pollinator species.

Out of all the environmental variables we analyzed, mean air quality emerged as the second most influential in explaining honey bee mortality after precipitation. Empirical studies have outlined adverse effects of poor air quality on immune system functionality, navigation, olfaction, and foraging behavior[9,14,16,37,38]. Moreover, using field samples collected in Bangalore (India), Thimmegowda et al. (2020) showed that Giant Asian honey bees (*Apis dorsata*) exposed to higher air pollution levels had reduced chances of survival[18]. Our findings align with these observations, establishing a clear link between poor air quality and increased honey bee mortality. Importantly, Thimmegowda et al. (2020) identified reduced foraging as one of the most probable causes linking poor air quality to reduced survival. Consequently, air pollutants may significantly contribute to the decline of pollinator populations by directly impairing their foraging ability, thereby limiting their capacity to provide essential ecosystem services such as pollination. Crop pollination is a major source of revenue for beekeepers in North America, contributing up to $3.18 billion in additional harvest revenue in Canada alone[39,40]. Such findings highlight a pressing need for policies aimed at reducing air pollution in order to maintain global food security.

We also observed substantial variation in how colonies from each beekeeping operation responded to elevated ozone concentrations, suggesting that the effects of air quality may be of more critical importance in certain areas. At the regional level, these differences could be due to seasonal changes in adverse air quality events and in vegetation availability that our analyses do not take into account. For instance, in Quebec, forest fires typically influence air quality during spring and early summer while in

Table 1 | Results of the generalized linear mixed model with the lowest AIC estimating the effect of environmental variables on the beehives' mortality (model 33)

| Random Effects			Variance	Standard deviation	Corr.	
Intercept			3.561	1.887	0.69	
Slope (Mean O^3 | beehive location)			3.610	1.900		
Fixed Effects	Estimate	S. E.	Odds Ratio	Probability	z-value	*p* value
Intercept	−0.170	0.453	0.843	0.457	−0.377	>0.05
O^3	1.192	0.454	3.295	0.767	2.627	<0.01*
NDVI	−0.420	0.010	0.657	0.396	−43.178	<0.001*
Wspd	−0.118	0.009	0.889	0.471	−13.473	<0.001*
NDVI:O^3	−0.126	0.010	0.882	0.469	−12.417	<0.001*
NDVI:Wspd	−0.035	0.008	0.966	0.491	−4.602	<0.01*
O^3:Wspd	−0.005	0.009	0.996	0.499	−0.478	>0.05
NDVI:O^3:Wspd	−0.040	0.007	0.961	0.490	−3.577	<0.001*
Observations	112 616					
Beehive locations	17					
R$^2_{marginal}$	0.155					
R$^2_{conditional}$	0.734					
AIC	128 766					
log-Likelihood	−63 372					

Honey bee mortality ~ mean NDVI + mean O^3 + mean Wspd + mean NDVI:mean O^3 + mean NDVI:mean Wspd + mean O^3:mean Wspd + mean NDVI:meanO^3:mean Wspd + (mean O^3 \| beehive location)
O^3, Average ozone concentration; NDVI, Normalized Difference Vegetation Index; Wspd, Average wind speed.
* indicates a significant effect (*p* value < 0.05).
The model was fitted using seasonal data from 2021 to 2023 for beekeeping operations located in Canada and the U.S. Canadian provinces include Quebec, Alberta, and Nova Scotia. American states include California, Florida, Maine, North Dakota, New-York, and Wyoming. We standardized all variables to mean and unit variance prior to fitting the model.

Fig. 1 | Predicted interactive effect of atmospheric ozone concentration and vegetation availability (NDVI) on honey bee hive mortality estimated by the GLMM with the lowest AIC. The probability of death is represented by a color gradient, where lighter colors (yellow) indicate higher chances of mortality. The chances of mortality are at their highest when the NDVI is low and ozone concentrations are elevated. Predictions cover 99% of the data, as very high NDVI values were infrequent in our observations.
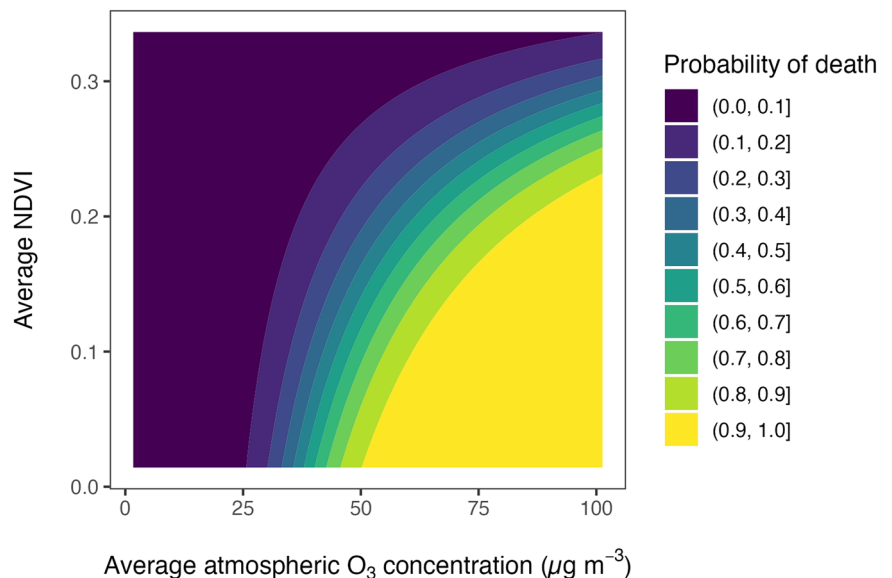
western regions, such events are more prevalent in late summer[41]. Yet, median NDVI values were similar among studied regions (Supplementary Fig. 5). At the local level, differences in beekeeping practices associated with the use of agrochemicals, pest management, proper monitoring, and operation size, may also explain why the effect of ozone concentration on honey bee mortality was stronger in some beekeeping operations[42–44]. Local attributes of vegetation density and availability may also drive the observed differences among operations in the relationship between mortality and ozone concentration, which we did not analyze here.

Although not as critical for predicting honey bee mortality, we observed increased chances of survival in regions with greater vegetation availability. One explanation could be that vegetation availability provided access to a larger pool of resources. Landscape changes alter the access to resources for honey bees, which are typically exposed to expansive farmlands and cropping systems[45–48]. These habitats are characterized by a limited variety of dominant crops such as almonds, cranberries, blueberries, apples, and raspberries[49]. Studies have shown that the abundance and diversity of surrounding floral species influence how much pollen and nectar pollinators collect, and the range of pollen types they gather[46,50]. Pollen is rich in both macro- and micronutrients and plays a crucial role in the physiological development and the immunity of honey bees, while floral nectar is rich in micronutrients and phytochemicals, acting as the primary
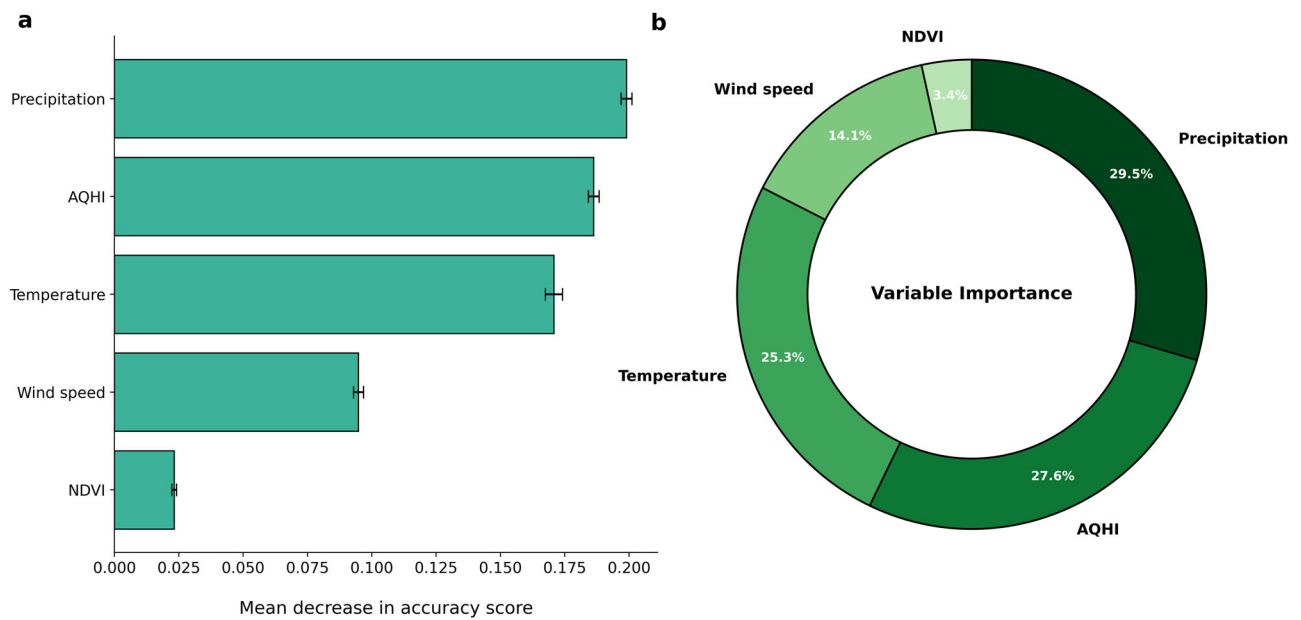
Fig. 2 | Variable importance measuring the contribution of each variable used in the Random Survival Forest to predict honey bee hive mortality. Each variable importance was computed using the permutation importance algorithm. Panel (**a**) shows, for each variable, the average decrease in model performance from 15 permutations with its standard deviation. Panel (**b**) shows the relative contribution of each variable.
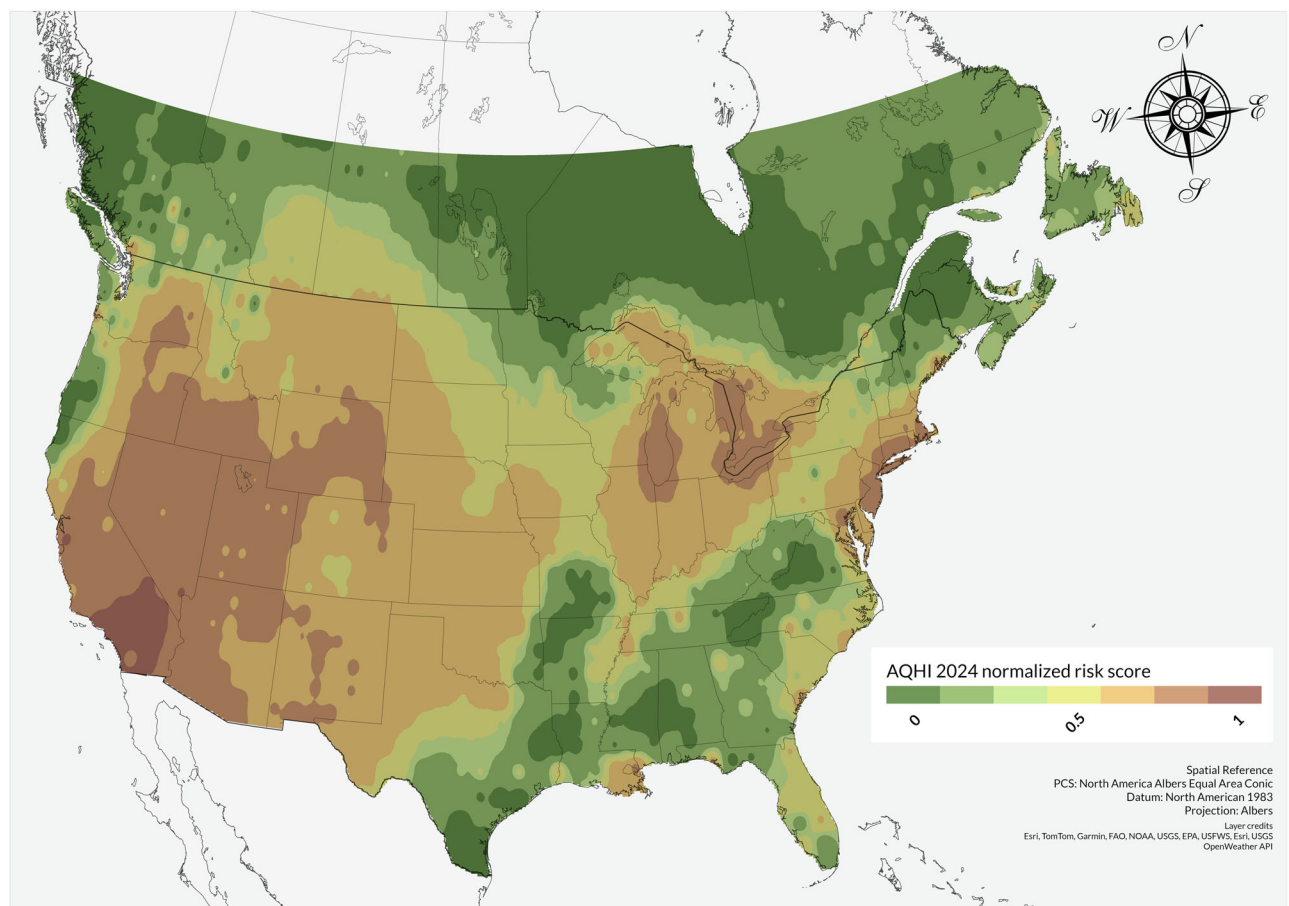

Fig. 3 | Risk map of honey beehive mortality for 2024 in Canada and the U.S. The map displays extrapolated risk scores for 2024. Those were generated using predicted risk scores from the RSF model and air quality data (AQHI) obtained from the OpenWeather API. The AQHI values represent averages from April 1 to September 1, 2024.

source of carbohydrates to sustain daily activities[46,51–53]. Pollen and nectar also increase honey bees' resilience against external stressors[38–40]. As such, shortages in these elements have been found to alter brood production and reduce survival in mason bees (*Osmia bicornis*) and honey bees[54,55].

Our findings also suggest that vegetation availability has a protective effect on honey bees exposed to poor air quality. Indeed, hives exposed to poor air quality in areas with greater vegetation showed better survival rates compared to those in areas with lower vegetation. On the one hand, access to more resources may help mitigate the effects of air pollutants on foraging efficiency by making it easier for bees to locate food. For example, yellow-faced bumblebees (*Bombus vosnesenskii*) will tend to forage further from their colony to find patches with a richer diversity of flowers[48]. Thus, bee-keepers may benefit from sending their colonies for pollination in yards surrounded by rich vegetation. On the other hand, it is also possible that vegetation acts as a protective barrier by absorbing parts of the atmospheric air pollutants. This can be observed by a negative correlation between air quality and NDVI[24,56] (Supplementary Fig.6). Research shows that the primary mechanism by which trees and shrubs remove air pollutants is via uptake through leaf stomata and through the plant surface[22]. Additional mechanisms include microclimate regulation and wind speed reductions which contribute in reducing ground-level ozone concentrations[57]. Vegetation also provides a physical barrier by intercepting airborne particles, and reduces the risk of pesticide exposure in farmlands[58]. However, the timing of poor air quality events may influence the protective capacity of vegetation. For example, late summer decreases in NDVI could exacerbate honey bee mortality if air quality is unfavorable at this time of the year such as during forest fires. During the spring of 2023, Canada experienced exceptional weather conditions marked by warmth and dryness, setting the stage for a record-breaking fire season with up to 17 million hectares (ha) of land burned across the country[41]. Quebec and Alberta had exceptionally early starts according to the Canadian Interagency Forest Fire Center. In Alberta, wildfire activity started at the end of April and continued past the legislated end of wildfire season into November. In Quebec, 182 fires were ignited by lightning on June 1st alone and the total area burned (4.5 million ha) was greater than the sum of the area burned over the last 20 years. In the U.S., western regions have experienced an increase in the frequency and the spread of wildfires over the past 20 years[59–61], with evidence pointing to wildfires being the main cause of above-average $PM_{2.5}$ concentrations[62]. The relationship between honey bee mortality and air quality combined with the increasing trend of fires in both countries is alarming. Projections suggest that wildfire-induced $PM_{2.5}$ concentrations are expected to nearly double by 2050[63], raising concerns about the future well-being of honey bee populations. Based on air quality data for 2024, our risk map reflects this trend in the Western U.S., but also in the Great Lakes and Northeast coast as the AQHI was much higher (i.e. poor quality) in these regions (Fig. 3). Air quality in the U.S. is also influenced by the smoke from Canadian wildfires[64]. While changes in air quality due to forest fires can be highly dynamic, our model suggests that bee health in these regions should be closely monitored to mitigate long-term consequences. Therefore, increasing vegetation cover should not be viewed as a standalone measure but rather as part of an holistic approach that includes proper monitoring alongside other strategies (e.g. biomarkers, integrated systems, pollinator ecology) to protect bee health[65–68].

One limitation of our study is that our predictive model may underestimate seasonal mortality in regions of Northern latitudes such as Quebec, Alberta, or Eastern U.S. due to mortality rates increasing during winter. For instance, hives are typically overwintering in Canada between mid November and April and mortality is usually not monitored by beekeepers during this period. Winter mortality in Canada is often calculated at the beginning of the new season in spring, leading to right-censoring of the survival data for the beekeeping season (Supplementary Fig. 7). Therefore, regions with warmer climates (e.g. California, Florida) may obtain more accurate mortality data due to keeping hives active during winter by providing them with supplemental food. Another caveat of our study is that the NDVI can only be seen as a proxy of resource availability and does not

consider species richness nor relative abundances (i.e. resource diversity). Hence, future implementations of our approach could integrate additional metrics to assess how different vegetation characteristics influence the relationship between air quality and bee mortality.

This study unveils a critical link between honey bee mortality and air quality in Canada and the U.S. Our research solidifies evidence linking poor air quality to compromised honey bee health[18,37], and establishes a clear link between honey bee mortality and limited vegetation availability. Interestingly, the survival response of honey bee colonies to poor air quality varied among beekeeping operations, prompting questions about the interplay between ecological phenomena and beekeeping practices–a concern that translates into considerable workload and economic implications for beekeepers. Yet, our observations make a strong case that vegetation availability is crucial to mitigate the consequences of poor air quality on honey bee health, emphasizing the need for research and mitigation strategies to address the upcoming challenges associated with escalating wildfires in Canada and the U.S. Hence, maintaining robust and diverse vegetation cover emerges as a simple solution to support ecosystem health and protect honey bee populations.

Methods

Honey bees dataset
We extracted anonymized data from beekeepers who use Nectar 1.19, an apiary tracking and management platform tailored for commercial beekeepers. The platform empowers beekeepers to maintain a comprehensive record of all actions, locations, movements, mortality, and other information concerning their beekeeping operations. Therefore, we had access to historical data from the 2020–2021, 2021–2022, and 2022–2023 beekeeping seasons for beekeepers located in Canada and the U.S. Our dataset is structured as an open cohort, allowing new colonies to enter at any given point in time during the beekeeping seasons. We only analyzed operations with at least one full year of follow-up data to effectively track mortality rates throughout a whole beekeeping season. We defined a colony as dead if its recorded death date fell within the current beekeeping season up until the start of the next season. The beekeepers are located in the U.S. states of California, Wyoming, Florida, Maine, North Dakota, and New York, and the Canadian provinces of Alberta, Nova Scotia, and Quebec. The final dataset contained a total size of over 103 477 hives with 112 626 observations.

Environmental dataset
Using publicly available data, we examined environmental factors likely to impact honey bees survival based on the literature[3,16,21,27,29]. The variables we extracted include daily measurements of the Air Quality Health Index (AQHI), daily average ozone concentration (O_3 µg m^{-3}), daily average temperature (°C), daily total precipitation (mm), daily average wind speed (km h^{-1}), and a measure of vegetation availability derived from satellite imagery, the Normalized Difference Vegetation Index (NDVI).

The AQHI is used by the government of Canada to assess and communicate health risks associated with air pollution. This metric represents a general assessment of air quality on an unbounded scale from 1 to 10+ by combining air pollutant concentrations, with higher values indicating worse air quality and increased health risks. The pollutants include Ozone (O_3), particulate matter 2.5 ($PM_{2.5}$), and Nitrogen Dioxide (NO_2). We extracted the air quality data using Airpyllution[69], a Python package offering access to current and historical (accessible from November 27th, 2020) air quality information via the OpenWeather API. Following the formula developed by Stieb et al., (2008), we calculated the seasonal average AQHI over the period each hive spent in its yard[31].

We used the NDVI to assess whether vegetation availability reduced the negative impacts of poor air quality on honey bee survival. This index can be used as an indicator of vegetative water, energy, and carbon balance[70]. Moreover, it holds significant importance as a parameter in ecosystem models[71], models estimating vegetation net primary production[72], and models predicting crop yields[73], and can be used as indicators of resource availability and habitat quality[18]. We calculated the NDVI values using the

near infrared and the red bands of the Harmonized Sentinel-2 MSI L2 products[74]. Both bands are offered at a 10 meter resolution. For each hive, we defined a 3 km buffer around the yard location, to take into account the foraging distance of honey bees[47,75]. Prior to performing the NDVI calculations for each hive, we used cloud and water bodies masking to exclude extreme NDVI values. We then computed the seasonal average of the NDVI by extracting values in 10-day increments over the period each hive spent in the yard for a given season. Lastly, for the amount of precipitation and the wind speed, we averaged the values over the period spent in the yard using the daily local values. We performed the calculations in Python, using the Earth Engine client library[76] for the NDVI and the Meteostat library[77] for precipitation and wind speed.

Statistical model

We fitted 33 generalized linear mixed models (GLMM) with a binomial error distribution, all estimating the relationship between honey bee colony mortality (y) as a function of combinations of environmental factors (X). We defined models that included either main effects exclusively or combinations of main and interaction effects between environmental predictors. We also tested different configurations of random effects, with models defined with random intercepts exclusively or with random slopes for either the region of study (i.e. U.S. state or Canadian province) or the operation managing the beehive. Considering the computational complexities associated with linear mixed models, we performed a convergence check before performing model comparisons. We evaluated and compared the predictive performance among models using the Akaike Information Criterion[78] (Supplementary Table 1). We standardized predictor variables to mean and unit variance prior to fitting the models. The fixed effects of the most parsimonious model (Table 1 and Supplementary Table 1) include the average atmospheric ozone concentration ($\mu g\ m^{-3}$), vegetation availability (NDVI), the average wind speed ($km\ h^{-1}$), and all combinations of two-way interactions as well as a three-way interaction between the three variables. Random effects include random intercepts and slopes relating mortality to ozone concentration for each operation managing the beehive ($n = 17$). We fitted the models in Python using the Lmer function of the Pymer4[79] package. We computed the predictions and the figures as well as Nakagawa and Schielzeth's R^2 for GLMM[80] in R version 4.4.1 with the packages lme4 (version 1-35.4), ggplot2 (version 3.5.1), and performance (version 0.12.3).

Machine learning model

To predict honey bee mortality and assess the importance of environmental variables, we fitted a Random Survival Forest (RSF) using Python's RandomSurvivalForest model[32]. RSF analysis extends Breiman's[81] random forest machine learning technique, specifically designed for investigating time-to-event observations. This method offers a flexible alternative to traditional survival analysis by eliminating the need for strict parametric or proportional hazards assumptions[82], making it well suited for datasets with right-censored observations such as honey beehives that survived the season. We applied a min-max standardization to ensure that all variables had similar scales before fitting the model. We constructed the RSF model by first performing a grid search across common tuning parameters, followed by a randomized search to refine the hyperparameters. We tested different numbers of trees, ranging from 100 to 1000 in increments of 100, and set the number of variables considered at each split to either the square root or base-2 logarithm of the total number of predictors. We set the candidate values for the maximum depth of the trees to 10, 30, and 60, while the minimum number of samples required at each split and at each leaf node were each specified as 10, 20, or 40. We configured the randomized search algorithm to perform three-fold cross-validation over five iterations to identify the optimal combination of hyperparameters.

We trained the RSF model using 85% of the dataset and then evaluated its performance on the remaining 15%. We used the modified concordance index[33] and the time-dependent ROC-AUC[34–36] to evaluate the model's performance. The modified C-index assesses the model's ability to accurately rank two random individuals (i.e. hives) according to their observed survival times, and assumes random censoring and censoring that is independent of the variables. The modified concordance index is based on inverse probability of censoring weights and does not depend on the distribution of censoring times in the test data, making it more consistent and robust than Harell's concordance index. Its values range between 0 and 1, with 0.5 being a random model and 1 a perfect model. In contrast, the time-dependent ROC-AUC assesses how well the model distinguishes between hives that died and those that survived at specific time intervals. It allows us to assess at a given time point how well a model can distinguish hives who will experience an event by time (sensitivity) from those that will not (specificity). This approach provides insight into the model's predictive performance over time.

Lastly, we assessed the importance of variables in predicting the survival probability of honey bee colonies using scikit-learn's permutation importance algorithm[83]. This method evaluates the decrease in the model's performance when a particular variable's values are randomly shuffled. A large drop in performance indicates that the variable is crucial for accurate predictions. In our study, we computed variable importance based on the decrease in the C-index when each feature was permuted. We used 15 permutations for each variable.

Map generation

Using the Airpyllution package and the OpenWeather API, we extracted hourly air pollutants data for a grid of points equally spaced at 1-degree intervals, from April 1st to September 1st 2024. From the extracted Ozone (O_3), particulate matter 2.5 ($PM_{2.5}$), and Nitrogen Dioxide (NO_2) data, we applied the formula by Stieb et al., (2008) to compute the average AQHI over the period for each point on the grid. We then used the predicted scores from the RSF model to extrapolate new predicted scores for 2024. To do so, we fitted an elastic net regression model of the RSF risk scores (y) as a function of the 2024 AQHI values (x) using the ElasticNet function from scikit-learn. We then applied a min-max transformation of the extrapolated risk scores, and performed an inverse-distance weighted interpolation to create a continuous spatial map of newly predicted risk scores. The final map was generated in ArcGIS Pro version 3.3.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Commercial and privacy restrictions impede us from making the dataset used for this study freely available. Requests for the full dataset should be addressed to Nico Coallier. The provided sample dataset contains data from Nectar research hives and can be found here: https://github.com/Nectar-Technologies/ledge-nectar/tree/main/aqi/example_data.

Code availability

To maximize transparency given restrictions on data availability, the code and a sample dataset used to produce the results are available in the following GitHub repository: https://github.com/Nectar-Technologies/ledge-nectar.

References

1. Zacharakis, I. & Tsihrintzis, V. A. Integrated wildfire danger models and factors: A review. *Sci. Total Environ.* **899**, 165704–165704 (2023).
2. Chen, D. et al. Simulating spatio-temporal dynamics of surface PM2.5 emitted from Alaskan wildfires. *Sci. Total Environ.* **898**, 165594–165594 (2023).
3. Liu, Y. et al. Systematic mapping and review of landscape fire smoke (LFS) exposure impacts on insects. *Environ. Entomol.* **51**, 871–884 (2022).

4. McFrederick, Q. S., Fuentes, J. D., Roulston, T., Kathilankal, J. C. & Lerdau, M. Effects of air pollution on biogenic volatiles and ecological interactions. *Oecologia* **160**, 411–420 (2009).

5. Potts, S. G. et al. Safeguarding pollinators and their values to human well-being. *Nature* **540**, 220–229 (2016).

6. Kazenel, M. R., Wright, K. W., Griswold, T., Whitney, K. D. & Rudgers, J. A. Heat and desiccation tolerances predict bee abundance under climate change. *Nature* **628**, 342–348 (2024).

7. Sánchez-Bayo, F. & Wyckhuys, K. A. G. Worldwide decline of the entomofauna: A review of its drivers. *Biol. Conserv.* **232**, 8–27 (2019).

8. Dobson, H. E. M. Floral Volatiles in Insect Biology. in *Insect-Plant Interactions* vol. V (CRC Press, Boca Raton, 1994).

9. McFrederick, Q. S., Kathilankal, J. C. & Fuentes, J. D. Air pollution modifies floral scent trails. *Atmos. Environ.* **42**, 2336–2348 (2008).

10. Papazian, S. & Blande, J. D. Dynamics of plant responses to combinations of air pollutants. *Plant Biol.* **22**, 68–83 (2020).

11. Farré-Armengol, G. et al. Ozone degrades floral scent and reduces pollinator attraction to flowers. *N. Phytol.* **209**, 152–160 (2016).

12. Arndt, U. Air pollutants and pheromones — A problem? *Chemosphere* **30**, 1023–1031 (1995).

13. Phillips, B. B. et al. Impacts of multiple pollutants on pollinator activity in road verges. *J. Appl. Ecol.* **58**, 1017–1029 (2021).

14. Ryalls, J. M. W. et al. Anthropogenic air pollutants reduce insect-mediated pollination services. *Environ. Pollut.* **297**, 118847 (2022).

15. Visscher, P. K., Vetter, R. S. & Robinson, G. E. Alarm pheromone perception in honey bees is decreased by smoke (Hymenoptera: Apidae). *J. Insect Behav.* **8**, 11–18 (1995).

16. Cho, Y. et al. Foraging trip duration of honeybee increases during a poor air quality episode and the increase persists thereafter. *Ecol. Evol.* **11**, 1492–1500 (2021).

17. Dolezal, A. G. et al. Interacting stressors matter: diet quality and virus infection in honeybee health. *R. Soc. Open Sci.* **6**, 181803 (2019).

18. Thimmegowda, G. G. et al. A field-based quantitative analysis of sublethal effects of air pollution on pollinators. *Proc. Natl. Acad. Sci.* **117**, 20653–20661 (2020).

19. Vanderplanck, M. et al. Ensuring access to high-quality resources reduces the impacts of heat stress on bees. *Sci. Rep.* **9**, 12596–12596 (2019).

20. Roulston, T. H. & Goodell, K. The Role of Resources and Risks in Regulating Wild Bee Populations. *Annu. Rev. Entomol.* **56**, 293–312 (2011).

21. Zhang, G. et al. Native vegetation embedded in landscapes dominated by corn and soybean improves honey bee health and productivity. *J. Appl. Ecol.* **60**, 1032–1043 (2023).

22. Nowak, D. J., Crane, D. E. & Stevens, J. C. Air pollution removal by urban trees and shrubs in the United States. *Urban Urban Green.* **4**, 115–123 (2006).

23. Escobedo, F. J., Kroeger, T. & Wagner, J. E. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* **159**, 2078–2087 (2011).

24. Sheng, Q., Ji, Y., Zhou, C., Zhang, H. & Zhu, Z. Spatiotemporal Variation and Pattern Analysis of Air Pollution and Its Correlation with NDVI in Nanjing City. *China.: A Landsat-Based Study For.* **14**, 2106 (2023).

25. Cunningham, M. M. et al. Honey bees as biomonitors of environmental contaminants, pathogens, and climate change. *Ecol. Indic.* **134**, 108457 (2022).

26. Di Fiore, C. et al. Honeybees as Bioindicators of Heavy Metal Pollution in Urban and Rural Areas in the South of Italy. *Atmosphere* **13**, 624–624 (2022).

27. Leonard, R. J., Pettit, T. J., Irga, P., McArthur, C. & Hochuli, D. F. Acute exposure to urban air pollution impairs olfactory learning and memory in honeybees. *Ecotoxicology* **28**, 1056–1062 (2019).

28. Le Conte, Y. & Navajas, M. Climate change: Impact on honey bee populations and diseases. *Rev. Sci. Tech.* **27**, 499–510 (2008).

29. vanEngelsdorp, D. & Meixner, M. D. A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *J. Invertebr. Pathol.* **103**, S80–S95 (2010).

30. Calovi, M., Grozinger, C. M., Miller, D. A. & Goslee, S. C. Summer weather conditions influence winter survival of honey bees (Apis mellifera) in the northeastern United States. *Sci. Rep.* **11**, 1553 (2021).

31. Stieb, D. M. et al. A new multipollutant, no-threshold air quality health index based on short-term associations observed in daily time-series analyses. *J. Air Waste Manag. Assoc.* **58**, 435–450 (2008).

32. Pölsterl, S. scikit-survival: a library for time-to-event analysis built on top of scikit-learn. *J. Mach. Learn Res* **21**, 212:8747–212:8752 (2020).

33. Uno, H., Cai, T., Pencina, M. J., D'Agostino, R. B. & Wei, L. J. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat. Med.* **30**, 1105–1117 (2011).

34. Uno, H., Cai, T., Tian, L. & Wei, L. J. Evaluating prediction rules for t-year survivors with censored regression models. *J. Am. Stat. Assoc.* **102**, 527–537 (2007).

35. Hung, H. & Chiang, C.-T. Estimation methods for time-dependent AUC models with survival data. *Can. J. Stat.* **38**, 8–26 (2010).

36. Lambert, J. & Chevret, S. Summary measure of discrimination in survival models based on cumulative/dynamic time-dependent ROC curves. *Stat. Methods Med. Res.* **25**, 2088–2102 (2016).

37. Mayack, C. et al. Poor Air Quality Is Linked to Stress in Honeybees and Can Be Compounded by the Presence of Disease. *Insects* **14**, 689–689 (2023).

38. Girling, R. D., Lusebrink, I., Farthing, E., Newman, T. A. & Poppy, G. M. Diesel exhaust rapidly degrades floral odours used by honeybees. *Sci. Rep.* **3**, 2779 (2013).

39. Topitzhofer, E., Breece, C., Wyns, D. & Sagili, R. Revenue Sources for a Commercial Beekeeping Operation in the Pacific Northwest.

40. Agriculture and Agri-Food Canada. Statistical Overview of the Canadian Honey and Bee Industry and the Economic Contribution of Honey Bee Pollination, 2021. https://agriculture.canada.ca/en/sector/horticulture/reports/statistical-overview-canadian-honey-and-bee-industry-2021 (2021).

41. Canadian Interagency Forest Fire Centre (CIFFC). *Canada Report*. (2023).

42. Underwood, R. M., Traver, B. E. & López-Uribe, M. M. Beekeeping Management Practices Are Associated with Operation Size and Beekeepers' Philosophy towards in-Hive Chemicals. *Insects* **10**, 10 (2019).

43. Andrews, E. To save the bees or not to save the bees: honey bee health in the Anthropocene. *Agric. Hum. Values* **36**, 891–902 (2019).

44. El Agrebi, N. et al. Risk and protective indicators of beekeeping management practices. *Sci. Total Environ.* **799**, 149381 (2021).

45. Brodschneider, R. & Crailsheim, K. Nutrition and health in honey bees. *Apidologie* **41**, 278–294 (2010).

46. Lau, P. W., Esquivel, I. L., Parys, K. A., Hung, K.-L. J. & Chakrabarti, P. The nutritional landscape in agroecosystems: a review on how resources and management practices can shape pollinator health in agricultural environments. *Ann. Entomol. Soc. Am.* **116**, 261–275 (2023).

47. Steffan-Dewenter, I. & Kuhn, A. Honeybee foraging in differentially structured landscapes. *Proc. R. Soc. Lond. B Biol. Sci.* **270**, 569–575 (2003).

48. Jha, S. & Kremen, C. Resource diversity and landscape-level homogeneity drive native bee foraging. *Proc. Natl. Acad. Sci.* **110**, 555–558 (2013).

49. Agriculture and Agri-Food Canada. Native Pollinators and Agriculture in Canada. https://www.publications.gc.ca/collections/collection_2014/aac-aafc/A59-12-2014-eng.pdf (2014).

50. Dimou, M. & Thrasyvoulou, A. Seasonal variation in vegetation and pollen collected by honeybees in Thessaloniki, Greece. *Grana* **46**, 292–299 (2007).

51. Pasquale, G. D. et al. Influence of Pollen Nutrition on Honey Bee Health: Do Pollen Quality and Diversity Matter? *PLOS ONE* **8**, e72016 (2013).
52. Mao, W., Schuler, M. A. & Berenbaum, M. R. Honey constituents up-regulate detoxification and immunity genes in the western honey bee Apis mellifera. *Proc. Natl. Acad. Sci.* **110**, 8842–8846 (2013).
53. Pasquale, G. D. et al. Variations in the Availability of Pollen Resources Affect Honey Bee Health. *PLOS ONE* **11**, e0162818 (2016).
54. Helm, B. R. et al. The geometric framework for nutrition reveals interactions between protein and carbohydrate during larval growth in honey bees. *Biol. Open* **6**, 872–880 (2017).
55. Austin, A. J. & Gilbert, J. D. J. Solitary bee larvae prioritize carbohydrate over protein in parentally provided pollen. *Funct. Ecol.* **35**, 1069–1080 (2021).
56. Zhou, M., Huang, Y. & Li, G. Changes in the concentration of air pollutants before and after the COVID-19 blockade period and their correlation with vegetation coverage. *Environ. Sci. Pollut. Res.* **28**, 23405–23419 (2021).
57. Akbari, H., Pomerantz, M. & Taha, H. Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. *Sol. Energy* **70**, 295–310 (2001).
58. Hisamoto, S., Ikegami, M., Goka, K. & Sakamoto, Y. The impact of landscape structure on pesticide exposure to honey bees. *Nat. Commun.* **15**, 8999 (2024).
59. Burke, M. et al. The changing risk and burden of wildfire in the United States. *Proc. Natl. Acad. Sci.* **118**, e2011048118 (2021).
60. Iglesias, V., Balch, J. K. & Travis, W. R. U. S. fires became larger, more frequent, and more widespread in the 2000s. *Sci. Adv.* **8**, eabc0020 (2022).
61. Rickly, P. S. et al. Influence of Wildfire on Urban Ozone: An Observationally Constrained Box Modeling Study at a Site in the Colorado Front Range. *Environ. Sci. Technol.* **57**, 1257–1267 (2023).
62. Li, Y. et al. Dominance of Wildfires Impact on Air Quality Exceedances During the 2020 Record-Breaking Wildfire Season in the United States. *Geophys. Res. Lett.* **48**, e2021GL094908 (2021).
63. Sarangi, C. et al. Projected increases in wildfires may challenge regulatory curtailment of PM2.5 over the eastern US by 2050. *Atmos. Chem. Phys.* **23**, 1769–1783 (2023).
64. Dong, M. et al. Maps: Tracking Air Quality and Smoke From Wildfires. *The New York Times* (2023).
65. Parreño, M. A. et al. Critical links between biodiversity and health in wild bee conservation. *Trends Ecol. Evol.* **37**, 309–321 (2022).
66. López-Uribe, M. M., Ricigliano, V. A. & Simone-Finstrom, M. Defining Pollinator Health: A Holistic Approach Based on Ecological, Genetic, and Physiological Factors. *Annu. Rev. Anim. Biosci.* **8**, 269–294 (2020).
67. Committee, E. S. et al. A systems-based approach to the environmental risk assessment of multiple stressors in honey bees. *EFSA J.* **19**, e06607 (2021).
68. Zaman, A. & Dorin, A. A framework for better sensor-based beehive health monitoring. *Comput. Electron. Agric.* **210**, 107906 (2023).
69. Alexander, C., King, D. & Liow, M. Airpyllution: A Python package for air pollution analysis. https://github.com/UBC-MDS/airpyllution/tree/main (2024).
70. Rafique, R., Zhao, F., De Jong, R., Zeng, N. & Asrar, G. R. Global and Regional Variability and Change in Terrestrial Ecosystems Net Primary Production and NDVI: A Model-Data Comparison. *Remote Sens.* **8**, 177 (2016).
71. Hunt, E. R. et al. Carbon Sequestration from Remotely-Sensed NDVI and Net Ecosystem Exchange. in *From Laboratory Spectroscopy to Remotely Sensed Spectra of Terrestrial Ecosystems* (ed. Muttiah, R. S.) 161–174 (Springer Netherlands, Dordrecht, 2002). https://doi.org/10.1007/978-94-017-1620-8_8.
72. Running, S. W. et al. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. *BioScience* **54**, 547–560 (2004).
73. Kouadio, L., Newlands, N. K., Davidson, A., Zhang, Y. & Chipanshi, A. Assessing the Performance of MODIS NDVI and EVI for Seasonal Crop Yield Forecasting at the Ecodistrict Scale. *Remote Sens.* **6**, 10193–10214 (2014).
74. European Space Agency. MSI Level-2A BOA Reflectance Product. Collection 1. https://doi.org/10.5270/S2_-znk9xsj (2021).
75. Couvillon, M. J. et al. Honey bee foraging distance depends on month and forage type. *Apidologie* **46**, 61–70 (2015).
76. Gorelick, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
77. Lamprecht, C. S. Meteostat Python. (2024).
78. *Model Selection and Multimodel Inference*: A Practical Information-Theoretic Approach. (Springer, New York, NY, 2002). https://doi.org/10.1007/b97636.
79. Jolly, E. Pymer4 Connecting R and Python for Linear Mixed Modeling. *J. Open Source Softw.* **3**, 862–862 (2018).
80. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).
81. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
82. Verschut, T. A. & Hambäck, P. A. A random survival forest illustrates the importance of natural enemies compared to host plant quality on leaf beetle survival rates. *BMC Ecol.* **18**, 33–33 (2018).
83. Pedregosa, F. et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Acknowledgements

Author contributions

N.C. and L.P. conceived and designed the research. N.C. and M.F.F. conducted the analyses with the assistance of Y.C. and J.V. N.C. managed the database. M.F.F., Y.C., N.C., and L.P. contributed to writing, reviewing, and editing the final manuscript. All of the authors approved the manuscript.

Competing interests

The authors declare the following competing interest: N.C is the CTO at Nectar Technologies. M.F.F. is a senior data scientist at Nectar Technologies. All other authors have no competing interests to declare.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02082-x.

Correspondence and requests for materials should be addressed to Nico Coallier.

Peer review information *Communications Earth & Environment* thanks Jonathan Jiang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis and Alice Drinkwater. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints