ABSTRACT

The European honey bee (*Apis mellifera*) is both a crucial pollinator for agricultural and natural ecosystems, and an agricultural commodity in its own right. However, honey bees are experiencing heavy mortality in North America and Europe due to a complex suite of factors. Weather affects both the bees themselves and the plants that support them. Surrounding land use, particularly proportion of agricultural and urban areas, determines forage resource abundance and pesticide exposure risk. Finally, management decisions, including treatment to control parasitic *Varroa destructor* mites, contribute to colony success and failure. We used three years of data from a survey of Pennsylvania beekeepers to assess the importance of weather, topography, land use, and management factors on overwintering mortality of managed honey bee colonies at both apiary and colony levels. A Random Forest model for mite-treated apiaries predicted overwintering survival with 73.3% accuracy for colonies and 65.7% for apiaries, as determined by cross-validation. Growing degree days was the most important predictor at both levels. Neither topographic nor management variables were important predictors. A weather-only model was used to predict colony survival probability across Pennsylvania for the three years of the study, and to create a composite map of survival probability for 1981-2019 (long-term probability mean value of 59.5%). Although three years of data were not enough to adequately capture the range of possible climatic conditions, the model nonetheless performed well within its constraints. The Random Forest approach is suited to understanding complex nonlinear drivers of survival, and to predicting outcomes given current conditions or projected climate changes.

Keywords: European honey bee, pollinator health, Random Forest, overwintering mortality

Abbreviations

- USDA-NASS: United States Department of Agriculture - National Agricultural Statistics Service
- CDL: Cropland Data Layer
- FRI: Forage Resource Index
- ITL: Insect Toxic Load
- RF: Random Forest
- OOB: out-of-bag

INTRODUCTION

Honey bees (*Apis mellifera*) contribute more than $20 billion in pollination services to US agriculture (Calderone, 2012), with additional economic value from downstream industrial sectors (Chopra et al., 2015). Honey production generates more than $300 million in income annually for US beekeepers (USDA, 2019). Winter colony mortality has a strong negative effect on economic and ecosystem potentials, with 30-40% of US colonies dying each winter (Bee Informed Team, 2019). Winter mortality is known to vary regionally in both the US and Europe (e.g. Seitz et al., 2015, Brodschneider et al, 2018). Overwintering survival in Pennsylvania is on average about 53.5% (2016-2019) (Bee Informed Team, 2019).

Unlike other insect species, honey bee colonies are not dormant during the winter: they remain active and maintain the hive temperature by forming a thermoregulating cluster (Döke et al., 2015). This enables them to survive long periods of cold temperatures without access to forage (reviewed in Döke et al., 2015, Seeley et al., 1985, Currie et al., 2015). The colony ceases foraging for nectar and pollen and relies on its existing stores. Furthermore, brood rearing ceases, and the colony is dependent on the survival of a long-lived cohort of bees that is produced in the fall: these bees will live for several months, while worker bees produced in the summer will only live for a few weeks. Thus, factors which undermine the ability of the bees to collect and store adequate amounts of food during the summer and autumn, or to thermoregulate effectively during the winter, can contribute to colony mortality. These factors include pathogens and parasites (some of which can be managed by the beekeeper), land use in the surrounding areas which influences forage quality and exposure to pesticides, and weather factors which influence the availability of forage, the thermoregulatory ability of the bees in the winter, and the amount of time before bees are able to initiate brood rearing the spring. Modeling and predicting honey bee winter survival requires consideration of all of these factors.

Management practices, notably control of the parasitic *Varroa destructor* mite, directly affect bee health and impacts winter survival (Genersch et al., 2010, van Dooremalen et al., 2012). Winter mortality of honey bee colonies is strongly correlated with uncontrolled mite populations (Genersch et al., 2010, van Dooremalen et al., 2012). Varroa mites are ectoparasites which feed on pupae and adult bees (Nazzi et al., 2016). Varroa mites transmit viruses and immunocompromise bees, which result in increased viral levels and symptoms (Grozinger et al., 2019, Annoscia et al., 2019). Parasitized, virus-infected bees have reduced nutritional stores and a reduced lifespan (Amdam et al., 2004). Thus, high levels of Varroa reduce the probability of winter survival (Dainat et al., 2012). Beekeepers have several management options that they can use to control Varroa populations and improve winter survival (Haber et al., 2019), though some beekeepers prefer to avoid treatments (Underwood et al., 2019).

Indeed, under some conditions, it may be more economically viable to simply replace a colony in the spring (Degrandi-Hoffman et al., 2019).

Many studies have evaluated how survival correlates with particular land use practices, such as the percentage of agricultural land or the percentage of certain crops in the area surrounding the hive. However, while several studies have indicated that honey bees show reduced growth or higher mortality with increasing urban or agricultural land use (Ricigliano et al., 2019, Clermont et al., 2015), others have found that agricultural land use is positively correlated with colony survival (Kuchling et al., 2018, Sponsler et al., 2015). These measures of land use do not necessarily correlate directly with forage quality, as bees can collect substantial resources from weeds in agricultural areas, and crops can vary greatly in the resources they provide to bees or their pesticide regimes (Requier et al., 2015, Colwell et al., 2017, Sponsler et al., 2019). Indices of forage quality and of pesticide loading based on surrounding land cover have been developed that are intended to incorporate specific effects of crop and habitat types on a broad scale (Kennedy et al., 2013, Koh et al., 2016, Douglas et al., 2019), but thus far these have not been applied to studies of honey bee winter survival or health.

Seasonal weather conditions affect both forage availability and thermoregulatory success, and directly and indirectly influence honey bee health (Schweiger et al., 2010). Weather conditions in the early spring or fall can change the timing of availability of foraging resources, lengthen or shorten the time in which these are available for bees, and alter the time in which bees can actively forage (Bartomeus et al., 2011, Scaven et al., 2013). Indeed, even small variations in growing season temperature can dramatically change the numbers of available flowers and the amount of nectar they produce (Mu et al., 2013). When winter conditions drop below 10°C, the bees form a thermoregulating cluster (Döke et al., 2015). Outside temperature conditions influence the efficiency of maintaining these temperatures, with optimal external temperatures of -5° to 10°C (Dainat et al., 2012). Some temperature fluctuation in the winter allows the cluster to relocate to areas in the hive with available honey stores (Currie et al., 2015). Honey bees must maintain the colony at optimal temperatures to rear healthy brood, and raise the temperature of the brood nest to ~33°C in late winter/early spring to initiate brood rearing (Currie et al., 2015). Previous research in Austria found that warmer and drier climates were associated with higher winter losses (Switanek et al., 2017), but that study did not include measurements of land use practices.

Our objective was to evaluate the factors that influence the overwintering survival of European honey bee colonies in Pennsylvania, including beekeeper management practices, weather and topographic variables that affect temperature and moisture, and the composition of the surrounding landscape in terms of the availability of foraging resources and potential pesticide load contributed by agricultural land use (Kennedy et al., 2013, Koh et al., 2016, Douglas et al., 2019). As well as identifying the most important drivers of mortality, we develop a model that can be used to predict the probability of overwintering success, both for the current year and as a function of projected future climate changes. To the best of our knowledge this is the first study on honey bee overwintering survival that combines weather, topography, and derived land use factors.

METHODS

The complex nature of the factors influencing overwintering survival in the European honey bee necessitated the integration of multiple datasets that represent temperature and moisture conditions, the availability of foraging resources, and the use of pesticides. Such diverse datasets can best be addressed with randomization-based machine learning techniques that do not require the data to meet standard statistical assumptions or expect relationships between overwintering success and environment to follow any predetermined form.

Data sources

Honey bee overwintering survival. Our main dataset originates from the Pennsylvania State Beekeepers Association Winter Loss Survey, from 2016 to 2019. The survey collected information about the management habits of the beekeepers. These include the colony numbers reported for November and April, which were used to calculate survival. Each beekeeper received a randomly assigned ID to protect personally identifiable information. This ID was regenerated each year, and thus individual apiaries could not be followed through time. The survey included an option to provide geographic coordinates for the beekeeper's apiary, and only responses that included these data were used. Only non-migratory operations were included; these colonies remained in the same place year-round. Beekeepers could only provide information for one location; we assume that the data represents data from a single apiary because most of the beekeepers reported fewer than 10 hives (larger numbers of hives would have suggested that they were spread across locations). The survey also asked respondents if colonies were moved during the year, and those that indicated that their colonies were moved were deleted from the data set. After these filtering steps, 342 apiaries, with 1,726 colonies, remained (Figure 1).

Weather and topographic variables. For each reported point, we generated annual and seasonal weather variables from 4 km gridded daily temperature and precipitation data (PRISM Climate Group, 2019). These included standard bioclimatic and agronomic indices (Table 1). Novel bee-specific weather indices have been developed based on honey bee biology and behavior, and relevant for overwintering survival, such as winter days within bee optimal thermoregulating temperature range, and number of winter days suitable for foraging, with maximum temperature above 16C and total precipitation below 3mm (Table 1; Seeley et al., 1985, Busby, 1991). Topographic variables such as slope and aspect were included in the analysis because they modify the local climate at finer scales than can be represented by the gridded climate data (Wang et al., 2017). These were calculated from 30m resolution gridded elevation data (USGS, 2014, GRASS Development Team, 2018).

Preliminary analysis of this dataset clearly demonstrated that treating for Varroa mites was a key factor in determining overwintering survival across all three years (Figure 2). Only 17% of beekeepers did not treat in some way. Because of the clear effect of the treatment and the small number of untreated colonies, we chose to analyze only the treated apiaries. The survey data was extremely unbalanced, with 1,429 colonies within 257 apiaries. The apiaries contained from 1 to 34 colonies, with a median value of 3. Our objective was to predict survival at the colony level, but we analyzed the data at both apiary and colony scales to ensure that results were consistent at both levels. We used a binary classification to model survival at the colony scale. To provide comparable results, and because of the highly unbalanced dataset, we modeled survival at the apiary scale as a binary variable as well, denoting whether an apiary experienced any colony mortality.

Forage resource index and insecticide toxic load. Two bee-specific distance-weighted landscape descriptors were generated for each apiary. These descriptors were based on the 30m resolution USDA-NASS Cropland Data Layer (CDL) for 2017 (Boryan et al., 2011), which provided information on the land use categories (including specific crop types). The seasonal Forage Resource Index (FRI) was developed based on expert opinion, and described the quality and abundance of floral resources for each land use category (Kennedy et al., 2013, Koh et al., 2016, Lonsdorf et al., 2009). As in Koh et al. 2016 (Koh et al., 2016), we generated the seasonal FRI at each apiary location using a distance decay function that accounts for the usual foraging distance of honey bees (5 km) (Beekman et al., 2000). The Insect Toxic Load (ITL) characterized the amount of active ingredient used for each insecticide based on statewide records of per-hectare use by crop type, and converted this to an aggregated insect toxic load using honey bee LD50s (Douglas et al., 2019). The same CDL data and distance-weighting function were used for the FRI and ITL to maintain consistency.

Statistical analyses

A probability Random Forest (RF), a flexible tree-based machine learning approach, was used to analyze overwintering mortality in relation to environmental and landscape factors within colonies that had been treated for Varroa mites. Random Forests develop a large number of decision trees using a random sampling of variables, then average across all trees to produce an ensemble (forest) fit (Wang et al., 2017, Breiman et al., 2001). The RF technique has been demonstrated to be very efficient when working with dataset comprising a large number of predictors (Berk, 2008), and when the relationship between variables is nonlinear or complex, because it is a flexible distribution-free method (Shoemaker et al., 2018). Given the complexity and nonlinearity of the dataset used in this project, RF was preferred to a linear regression method, and allowed the development of a reliable empirical model without prior knowledge of the relationship between the phenomena and the predictors (Auret et al., 2012).

All analyses were conducted in R 3.6.2 (R Core Team, 2019), using ranger 0.11.2 for RF models of survival probability and permutation-based variable importance (Wright et al., 2017), with caret 6.0-84 for model evaluation (Kuhn, 2019). Permutation-based variable importance measures the improvement in model accuracy due to inclusion of each variable (Breiman, 2001). The form of the relationship between survival probability and the major independent variables was assessed using partial dependence plots (pdp 0.7.0 package; Greenwell, 2017). Maps were produced with the raster (3.0-12; Hijmans et al., 2020) and sp (1.4-0; Pebesma et al., 2020) packages.

Our initial intent was to train the model for the winters of 2016-2017 and 2017-2018 and test with 2018-2019, but 2016 was warm and dry, 2017 was very wet, and 2018 was warm and wet. Instead, we used cross-validation stratified by year to assess model accuracy. Ten repetitions of a 10-fold cross-validation were used to tune the model on a gridded parameter search with the number of trees between 2,000 and 5,000 on an increment of 500, and number of variables per tree from 3 to 8.

An independent set of ten repetitions of a 10-fold cross-validation using the tuned parameters was used to obtain the error estimates. The final model was fitted on the full dataset, in order to obtain the most reliable estimates of variable importance and the best model for prediction. Such a model overestimates accuracy, so cross-validation error estimates are given. These estimates show how the model is likely to perform when presented with new data. The same cross-validation and analysis methods were used at both the apiary and colony scales.

RESULTS AND DISCUSSION

The analysis was conducted on the 257 apiaries that used Varroa mite treatments, comprising 1,429 colonies across the three years of the survey (2016-2019). Overwintering survival was calculated as the number of colonies in November versus the number of colonies alive in April. These apiaries were distributed throughout Pennsylvania (Figure 1).

The apiary model had cross-validated out-of-bag (OOB) error of 0.222 and an accuracy of 65.7% (95% confidence interval 59.6% - 761.5%); the colony model had an OOB error of 0.19 and an accuracy of 73.3% (95% confidence interval 70.9% - 75.5%) (Figure 3). The best apiary model used 3 variables and 4,000 trees, and the colony model 3 variables and 4,500 trees. Variable number was much more influential than number of trees.

The full model was nearly twice as likely to predict that colonies survived when they died than that they died when they actually survived, suggesting that there is an additional source of mortality we have not considered (Table 2). Management factors for which data were not available, including supplemental feeding, are potential explanations. Furthermore, other studies have found that the timing of mite treatment and weather conditions during mite treatment can significantly influence treatment efficacy and subsequent winter survival (Beyer et al., 2018). Finally, the survey did not include information about levels of or evidence for parasites or pathogens, and thus we could not evaluate whether these parameters correlated with survival. Interestingly, the beekeeper's years of experience had no relationship to colony survival, though it was identified as important elsewhere (Jacques et al., 2017).

Growing degree days in the prior summer was the strongest predictor of overwintering survival for both the models (Table 1 for colony model; apiary model not shown). Precipitation of the wettest quarter and maximum temperature of the warmest month were nearly as important, but showed the same patterns and are not presented individually. Random Forest models are robust to correlated predictors, but will indicate both are important. A study in Austria found that hot, dry summers reduced overwintering survival (Switanek et al., 2017). This is generally consistent with our findings, although our more nuanced analysis of climatic variables found adverse effects of both too-cool and too-hot summers. Precipitation also showed a unimodal relationship with survival; neither too dry nor too wet. Topographic factors were not important, possibly because colonies thermoregulate, mitigating the effects of microclimate (Döke et al., 2015, Currie et al., 2015).

The landscape variables, FRI and ITL, did not contribute substantially to the colony survival model. The foraging index is based on expert opinion, while the insecticide toxic load index does not take into account variation in local crop management practices or exposure rates of bees; thus, there is clearly room for improvement in methods for assessing the suitability of surrounding land use for pollinator use. Moreover, supplementary feeding from the beekeeper may have mitigated impacts of floral resource availability, while insecticide exposure can have complex effects on bees which may not be captured by winter survival rates (Sponsler et al., 2019).

Because the climatic variables were the most important overall, we also modeled colony survival using only weather variables. This facilitated prediction, since the 30m topographic variables require substantially more computing resources to process than the 4km climate data, it is difficult to estimate the FRI and ITL across a landscape, and because a weather-only model can be applied to both past and projected future climate data without requiring corresponding land use information. The weather-only model performed about as well as the full model, with an OOB error of 0.19 and 73% accuracy (95% confidence interval 70.6% - 75.2%). Variable importances were very similar to the full model.

The maps of predicted honey bee survival for the three winters studied showed considerable variability, both between years and across the state (Figure 5). Winter 2016-2017 showed a predicted mean survival of 49.2% (range 5% - 97.6%); 2017-2018 had a predicted mean survival of 59.2%, (range 9.8% - 100%); and 2018-2019 winter had a mean predicted survival of 59.7% (range 17% - 100%). Mean predicted long-term survival probability across Pennsylvania for 1981-2019 was 59.5% (range 5.3% - 100%). The mean is consistent with values reported elsewhere (53.5%) (Bee Informed Team, 2019). No part of Pennsylvania was always good or always bad for honey bee survival; there was substantial spatial and temporal variability.

CONCLUSIONS

The use of Varroa mite treatments was the major factor determining overwintering survival of European honey bee colonies in Pennsylvania, which is consistent with the results of other studies (e.g., Genersch et al., 2010, van Dooremalen et al., 2012). Climatic factors, particularly summer temperatures and winter precipitation, were the strongest predictors for treated colonies. Topographic factors and landscape quality factors were not important, contrary to expectations. The landscape indices may require further modification, or other factors may need to be included to better capture the effects of floral quality and insecticide use on colony survival. The effects of specific management practices, such as providing supplemental feed, could not be assessed with the available data. Further beekeeper surveys will include additional questions on management to facilitate reanalysis.

Because we only had three years of georeferenced survival data, the Random Forest model was not trained on the complete range of weather conditions possible in Pennsylvania, and will be less reliable when used to predict on novel conditions. Nonetheless, this model worked well in the three years for which it was developed, with a high accuracy given the extreme variability within the dataset, and the multitude of factors that affect overwintering survival. We anticipate updating the model regularly as new survey data become available.

The weather-only Random Forest model can be used for broad predictive purposes at state or regional scales, and will become more reliable as further survey data are incorporated. Because it does not rely on landscape or management factors, this model can be used to characterize changes in overwintering survival with the changing climate. With slight modifications to use current data, this modeling approach can be used to estimate survival probabilities for the upcoming winter. The results of this project have been used to develop a real-time tool to predict honey bee survival probability as a function of GDD. The tool has been incorporated into the Beescape decision support tool (https://beescape.org/), used by beekeepers and technical advisors. Our results clearly demonstrate both the predictive power of weather variables on analyses of honey bee overwintering survival, and the efficacy of using machine learning methods such as Random Forest that are capable of identifying complex nonlinear relationships with correlated predictors.

REFERENCES

1. Amdam G. V., Hartfelder K., Norberg K., Hagen A., Omholt S.W., (2004), Altered physiology in worker honey bees (Hymenoptera: Apidae) infested with the mite Varroa destructor (Acari: Varroidae): a factor in colony loss during overwintering?, Journal of Economic Entomology, Volume 97, Issue 3, 1 June 2004, Pages 741–747, https://doi.org/10.1093/jee/97.3.741
2. Annoscia D., Brown S.P., Di Prisco G., De Paoli E., Del Fabbro S., Frizzera D., Zanni V., Galbraith D. A., Caprio E., Grozinger C. M., Pennacchio F., Nazzi F., (2019) Haemolymph removal by Varroa mite destabilizes the dynamical interaction between immune effectors and virus in bees, as predicted by Volterra's model, Proceedings of the Royal Society B Biological Sciences, DOI: https://doi.org/10.1098/rspb.2019.0331
3. Auret L., Aldrich C., (2012), Interpretation of nonlinear relationships between process variables by use of random forests, Minerals Engineering, Volume 35, Pages 27-42, DOI: https://doi.org/10.1016/j.mineng.2012.05.008
4. Bartomeus I., et al., (2011), Climate-associated phenological advances in bee pollinators and bee-pollinated plants, PNAS 2011 108 (51) 20645-20649, DOI: https://doi.org/10.1073/pnas.1115559108
5. Bee Informed Team. Loss & Management Survey, accessed October 2019, https://beeinformed.org/citizen-science/loss-and-management-survey/

6. Beekman M., Ratnieks F.L.W., (2000), Long-range foraging by the honey-bee, Apis mellifera L., Functional Ecology 2000 14, 490-496. DOI: https://doi.org/10.1046/j.1365-2435.2000.00443.x
7. Berk R.A., (2008), Statistical Learning from a Regression Perspective, New York: Springer
8. Beyer M., et al., (2018), Winter honey bee colony losses, Varroa destructor control strategies, and the role of weather conditions: Results from a survey among beekeepers, Research in Veterinary Science, Volume 118, 2018, Pages 52-60. DOI: https://doi.org/10.1016/j.rvsc.2018.01.012
9. Boryan C., Yang Z., Mueller R., Craig M., (2011), Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program, Geocarto International, 26:5, 341-358. DOI: 10.1080/10106049.2011.562309
10. Breiman L., (2001), Random Forests. Machine Learning 45, 5–32. DOI: 10.1023/A:1010933404324
11. Brodschneider R., et al., (2018), Multi-country loss rates of honey bee colonies during winter 2016/2017 from the COLOSS survey, Journal of Apicultural Research, 57:3, 452-457. DOI: 10.1080/00218839.2018.1460911
12. Busby J.R., (1991), BIOCLIM: a bioclimate analysis and prediction system. Plant Prot. Q. 6(1) 8-9. AGR:IND91051921
13. Calderone N. W., (2012), Insect Pollinated Crops, Insect Pollinators and US Agriculture: Trend Analysis of Aggregate Data for the Period 1992–2009. PLOS ONE 7(5): e37235 . DOI: https://doi.org/10.1371/journal.pone.0037235
14. Chopra S. S., Bakshi B. R., Khanna V., (2015), Economic Dependence of U.S. Industrial Sectors on Animal-Mediated Pollination Service, Environmental Science & Technology 2015 49(24), 14441-14451, DOI: https://doi.org/10.1021/acs.est.5b03788
15. Clermont A., Eickermann M., Kraus F., Hoffmann L., Beyer M., (2015), Correlations between land covers and honey bee colony losses in a country with industrialized and rural regions, Science of The Total Environment, 532(1-13). DOI: https://doi.org/10.1016/j.scitotenv.2015.05.128
16. Colwell M. J., WilliamsG. R., Evans R.C., Shutler D., (2017), Honey bee-collected pollen in agro-ecosystems reveals diet diversity, diet quality, and pesticide exposure, Ecol Evol. 2017; 7: 7243– 7253. DOI: https://doi.org/10.1002/ece3.3178
17. Currie R. W., Spivak M., Reuter G. S., (2015), 'Wintering Management of Honey Bee Colonies' in *The Hive and the Honey Bee*, by Graham J. M., 2015, ed. Dadant & Sons, Hamilton, USA, pp: 629-662. ISBN : 0915698099

18. Dainat B., Evans J.D., Chen Y.P., Gauthier L., Neumann P., (2012), Dead or Alive: Deformed Wing Virus and Varroa destructor Reduce the Life Span of Winter Honeybees, Applied and Environmental Microbiology, DOI: 10.1128/AEM.06537-11
19. Degrandi-Hoffman G., Graham H., Ahumada F., Smart M., Ziolkowski N., (2019), The Economics of Honey Bee (Hymenoptera: Apidae) Management and Overwintering Strategies for Colonies Used to Pollinate Almonds, Journal of Economic Entomology, Volume 112, Issue 6, December 2019, Pages 2524–2533. DOI: https://doi.org/10.1093/jee/toz213
20. Döke M. A., Frazier M., Grozinger C. M., (2015), Overwintering honey bees: biology and management, Current Opinion in Insect Science 2015, 10:185–193. DOI: https://doi.org/10.1016/j.cois.2015.05.014
21. Douglas M.R., Sponsler D.B., Lonsdorf E.V., Grozinger C.M., (2019), Rising insecticide potency outweighs falling application rate to make US farmland increasingly hazardous to insects, bioRxiv. DOI: https://doi.org/10.1101/71576
22. Genersch E., von der Ohe W., Kaatz H., Schroeder A., Otten C., et al., (2010), The German bee monitoring project: a long term study to understand periodically high winter losses of honey bee colonies. *Apidologie*, Springer Verlag, 41 (3) ⟨10.1051/apido/2010014⟩. ⟨hal-00892101⟩
23. GRASS Development Team, 2018. Geographic Resources Analysis Support System (GRASS) Software, Version 7.4. Open Source Geospatial Foundation. https://grass.osgeo.org
24. Greenwell B.M., (2017), pdp: An R package for constructing partial dependence plots. The R Journal 9, 421–436. ISSN 2073-485
25. Grozinger C. M., Flenniken M. L., (2019), Bee Viruses: Ecology, Pathogenicity, and Impacts, Annual Review of Entomology 64:1, 205-226. DOI: 10.1146/annurev-ento-011118-111942
26. Haber A. I., Steinhauer N. A., vanEngelsdorp D., (2019), Use of Chemical and Nonchemical Methods for the Control of Varroa destructor (Acari: Varroidae) and Associated Winter Colony Losses in U.S. Beekeeping Operations, Journal of Economic Entomology, Volume 112, Issue 4, 1509–1525, https://doi.org/10.1093/jee/toz088
27. Hijmans R.J., et al., (2020), raster: Geographic Data Analysis and Modeling, https://cran.r-project.org/web/packages/raster/raster.pdf
28. Jacques A., Laurent M., (2017), EPILOBEE Consortium, M. Ribière-Chabert, M. Saussac, et al., A pan-European epidemiological study reveals honey bee colony survival depends on beekeeper education and disease control. PLOS ONE 12(3): e0172591. DOI: https://doi.org/10.1371/journal.pone.0172591

29. Kennedy C.M., et al., (2013), A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems, Ecology Letters 16: 584–599. DOI: https://doi.org/10.1111/ele.12082
30. Koh I., et al., (2016), Modeling the status, trends, and impacts of wild bee abundance in the United States, Proceedings of the National Academy of Sciences, 113 (1) 140-145. DOI: 10.1073/pnas.1517685113
31. Kuchling S., Kopacka I., Kalcher-Sommersguter E., et al., (2018), Investigating the role of landscape composition on honey bee colony winter mortality: A long-term analysis. *Sci Rep* 8, 12263. DOI: 10.1038/s41598-018-30891-y
32. Kuhn M., (2019), *contributions from Wing J., et al.*, caret: Classification and Regression Training R package version 6.0-84. https://CRAN.R-project.org/package=caret
33. Lonsdorf E., Kremen C., Ricketts T., Winfree R., Williams N., Greenleaf S., (2009), Modelling pollination services across agricultural landscapes, *Annals of Botany*, Volume 103, Issue 9, Pages 1589–1600. DOI: https://doi.org/10.1093/aob/mcp069
34. Mu J., Peng Y., Xi X., Wu X., Li G., Niklas K.J., Sun S., (2015), Artificial asymmetric warming reduces nectar yield in a Tibetan alpine species of Asteraceae, Annals of Botany, Volume 116, Issue 6, November 2015, Pages 899–906, https://doi.org/10.1093/aob/mcv042
35. Nazzi F., Le Conte Y., (2016), Ecology of Varroa destructor, the major ectoparasite of the Western honey bee, Apis mellifera, Annual Review of Entomology. 61:417–32. DOI: 10.1146/annurev-ento-010715-023731
36. Pebesma E., et al., (2020), sp: Classes and Methods for Spatial Data, https://cran.r-project.org/web/packages/sp/sp.pdf
37. PRISM Climate Group, Oregon State University. (2019). http://www.prism.oregonstate.edu/
38. R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
39. Requier F., Odoux J., Tamic T., Moreau N., Henry M., Decourtye A., Bretagnolle V., (2015), Honey bee diet in intensive farmland habitats reveals an unexpectedly high flower richness and a major role of weeds, Ecological Applications, 25: 881-890. DOI: https://doi.org/10.1890/14-1011.1
40. Ricigliano V. A., Mott B. M., Maes P. W., Floyd A. S., Fitz W., Copeland D.C., Meikle W.G., Anderson K.E., (2019), Honey bee colony performance and health are enhanced by apiary proximity to US Conservation Reserve Program (CRP) lands, Scientific Reports 4894-9-1. DOI: https://doi.org/10.1038/s41598-019-41281-3
41. Scaven V.L., Rafferty N.E., (2013), Physiological effects of climate warming on flowering plants and insect pollinators and potential consequences for their

interactions, *Current Zoology*, Volume 59, Issue 3, 1 June 2013, Pages 418–426. DOI: https://doi.org/10.1093/czoolo/59.3.418

42. Schweiger O., et al., (2010), Multiple stressors on biotic interactions: how climate change and alien species interact to affect pollination, 2010, Biol. Rev., 85, pp. 777–795. DOI: https://doi.org/10.1111/j.1469-185X.2010.00125.x
43. Seeley T.D., Visscher P.K., (1985), Survival of honeybees in cold climates: the critical timing of colony growth and reproduction. Ecological Entomology, 10: 81-88. DOI: 10.1111/j.1365-2311.1985.tb00537
44. Seitz N., et al., (2015), A national survey of managed honey bee 2014–2015 annual colony losses in the USA, Journal of Apicultural Research, 54:4, 292-304). DOI: 10.1080/00218839.2016.1153294
45. Shoemaker K.T., Heffelfinger L.J., Jackson N.J., Blum M.E., Wasley T., Stewart K.M., (2018), A machine-learning approach for extending classical wildlife resource selection analyses, Ecology and evolution, 8(6), 3556–3569. DOI: https://doi.org/10.1002/ece3.3936
46. Sponsler D. B., Johnson R.M., (2015), Honey bee success predicted by landscape composition in Ohio, USA. PeerJ 3:e838. DOI: https://doi.org/10.7717/peerj.838
47. Sponsler D.B., Grozinger C.M., Hitaj C., Rundlöf M., Botías C., Code A., Lonsdorf E.V., Melathopoulos A.P., Smith D.J., Suryanarayanan S., Thogmartin W.E., Williams N.M., Zhang M., Douglas M.R., (2019), Pesticides and pollinators: A socioecological synthesis, Science of The Total Environment, 662 (1012-1027). DOI: https://doi.org/10.1016/j.scitotenv.2019.01.016
48. Switanek M., et al., (2017), Modelling seasonal effects of temperature and precipitation on honey bee winter mortality in a temperate climate, Science of the Total Environment, 579 1581–1587. DOI: http://dx.doi.org/10.1016/j.scitotenv.2016.11.178
49. Underwood R.M., Traver B.E., López-Uribe M.M.,(2019), Beekeeping Management Practices Are Associated with Operation Size and Beekeepers' Philosophy towards in-Hive Chemicals. Insects 10(1), 10. DOI: https://doi.org/10.3390/insects10010010
50. USDA National Agricultural Statistics Service, (2019), Honey. ISSN: 1949-1492 https://usda.library.cornell.edu/concern/publications/hd76s004z?locale=en. Accessed 15 September, 2019
51. USGS, 2014, National Elevation Dataset. https://gdg.sc.egov.usda.gov/. Accessed Apr 2014
52. van Dooremalen C., Gerritsen L., Cornelissen B., van der Steen J. J. M., van Langevelde F., et al., (2012), Winter Survival of Individual Honey Bees and Honey Bee Colonies Depends on Level of Varroa destructor Infestation. PLOS ONE 7(4): e36285. DOI: https://doi.org/10.1371/journal.pone.0036285

53. Wang A., Goslee S.C., Miller D.A., Sanderson M.A., Gonet J.M., (2017), Topographic variables improve climatic models of forage species abundance in the northeastern United States. Appl Veg Sci, 20: 84-93. DOI: 10.1111/avsc.12284
54. Wright M.N., Ziegler A., (2017), Ranger: A fast implementation of Random Forests for high dimensional data in C++ and R. Journal of Statistical Software 77, 1–17. DOI: 10.18637/jss.v077.i01

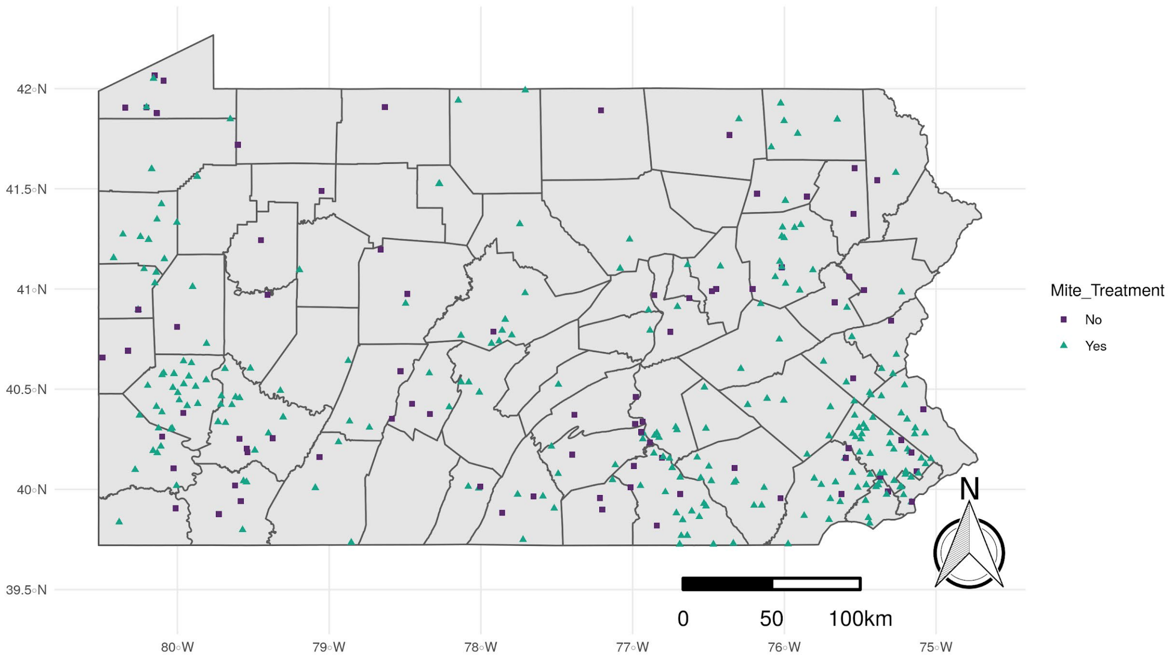

Figure 1. Locations of Pennsylvania beekeeper survey respondents from 2016-2019, stratified by use of treatment for Varroa mites (257 treated and 85 untreated apiaries).

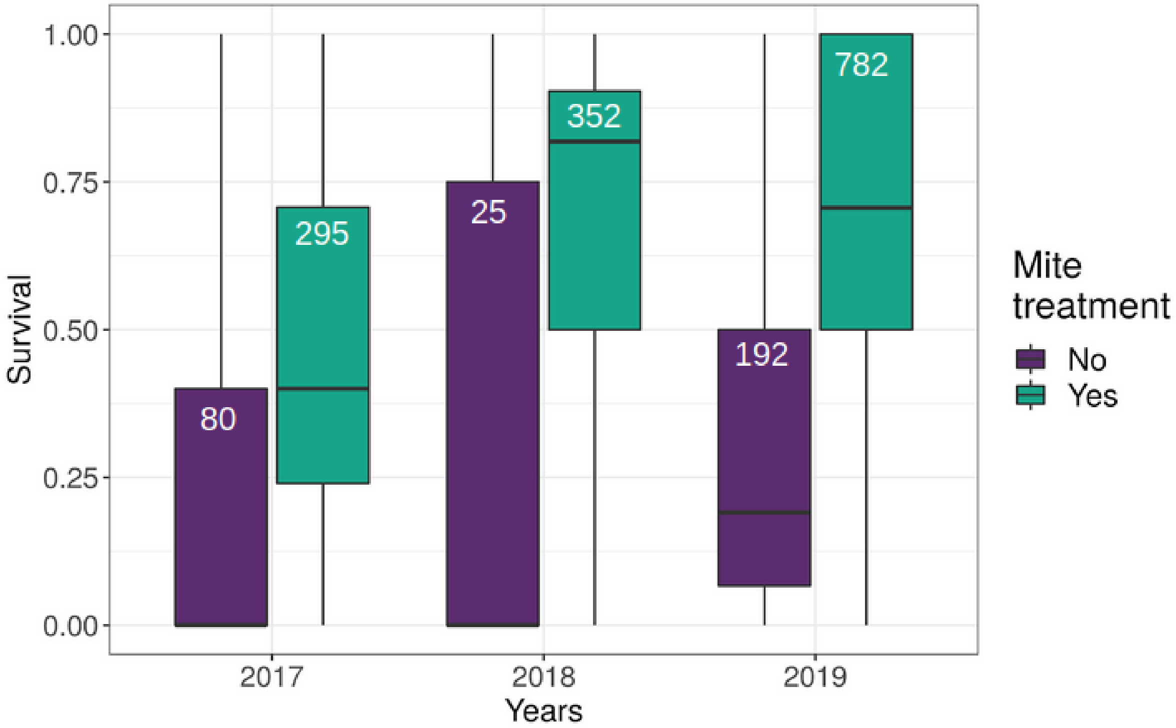

Figure 2. Survival of mite-treated and untreated honey bee colonies by year. In each of the three years, 80 out of 375 (21%), 25 out of 377 (7%) and 192 out of 974 (20%) colonies were untreated (297 out of 1,726, or 17% overall).

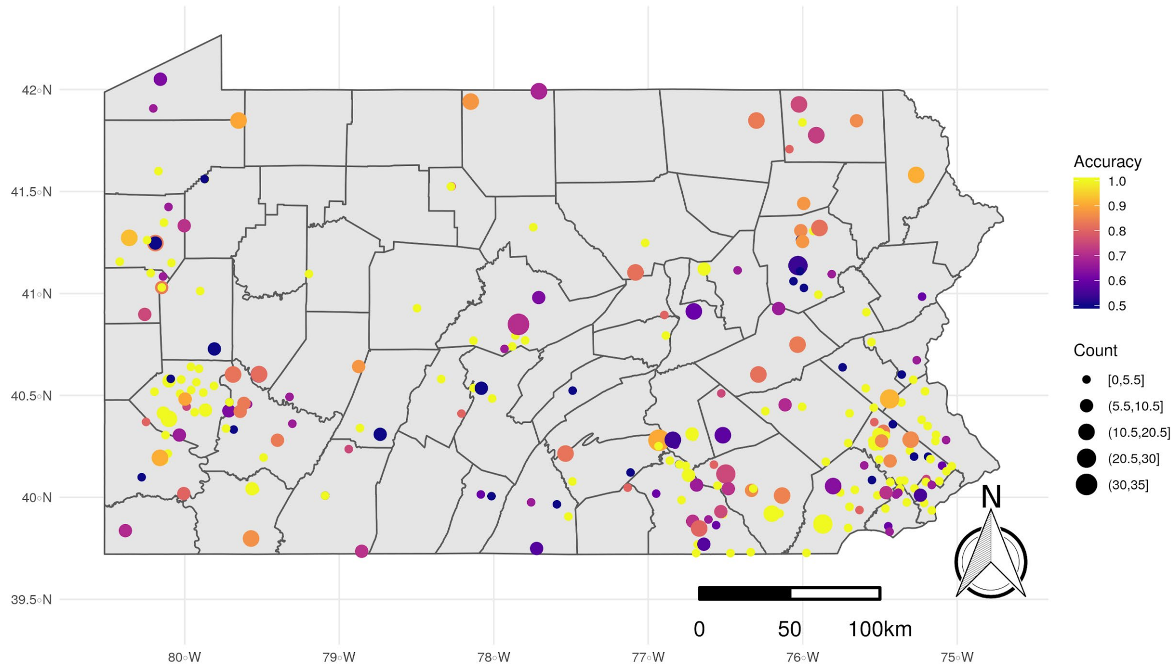

Figure 3. Prediction accuracy of the Random Forest model of overwintering survival probability aggregated by apiary for simplicity. The test set contained 1,429 colonies within 257 apiaries. The color indicates the mean overall model accuracy at that apiary, and the circle size is proportional to the number of colonies in November.

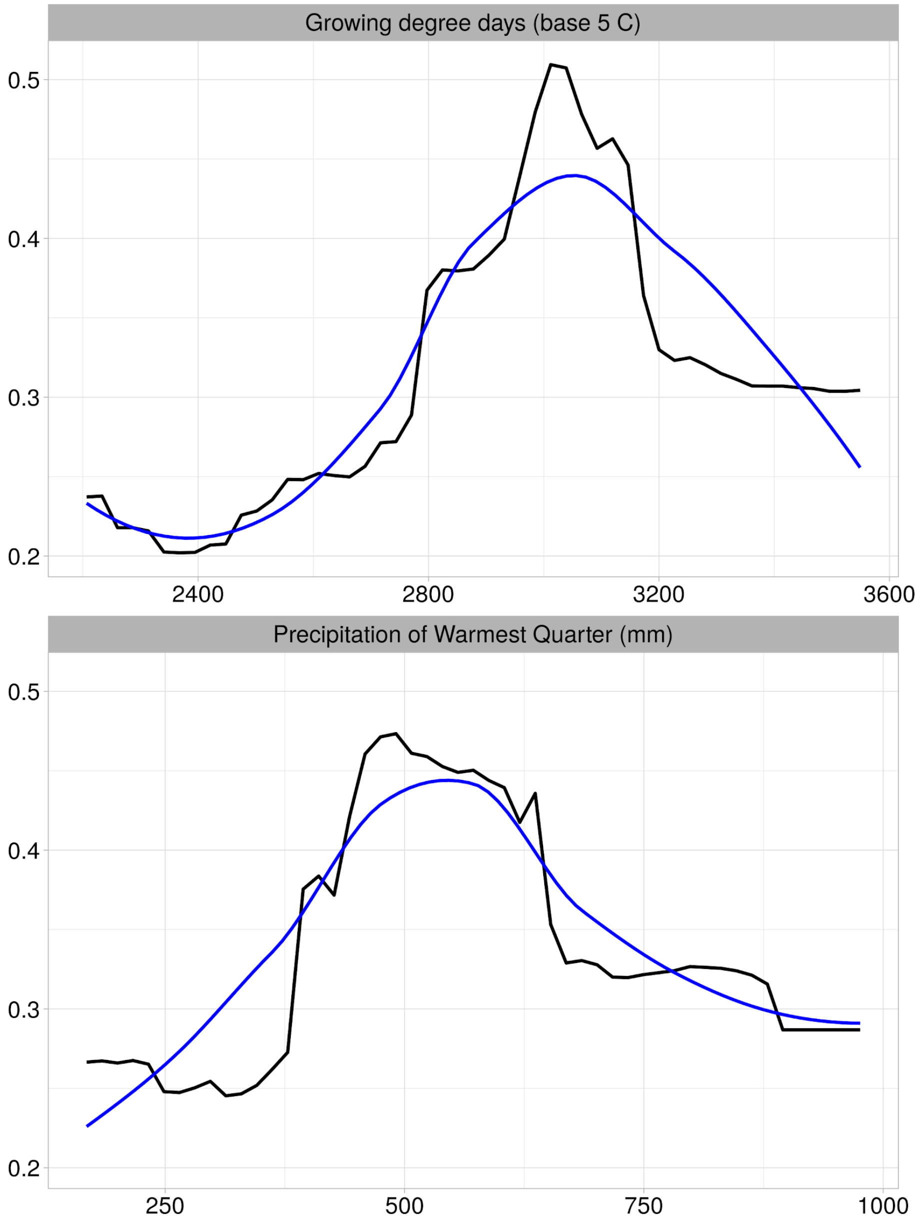

Figure 4. Partial dependence plots for the two most important variables. These plots describe the relationship between the explanatory variable named in the header (along the x axis) and the probability of overwintering survival (y axis), given all the other variables in the model. The black line represents the modeled relationship between survival and the variables, while the blue line shows a spline-smoothed fit.

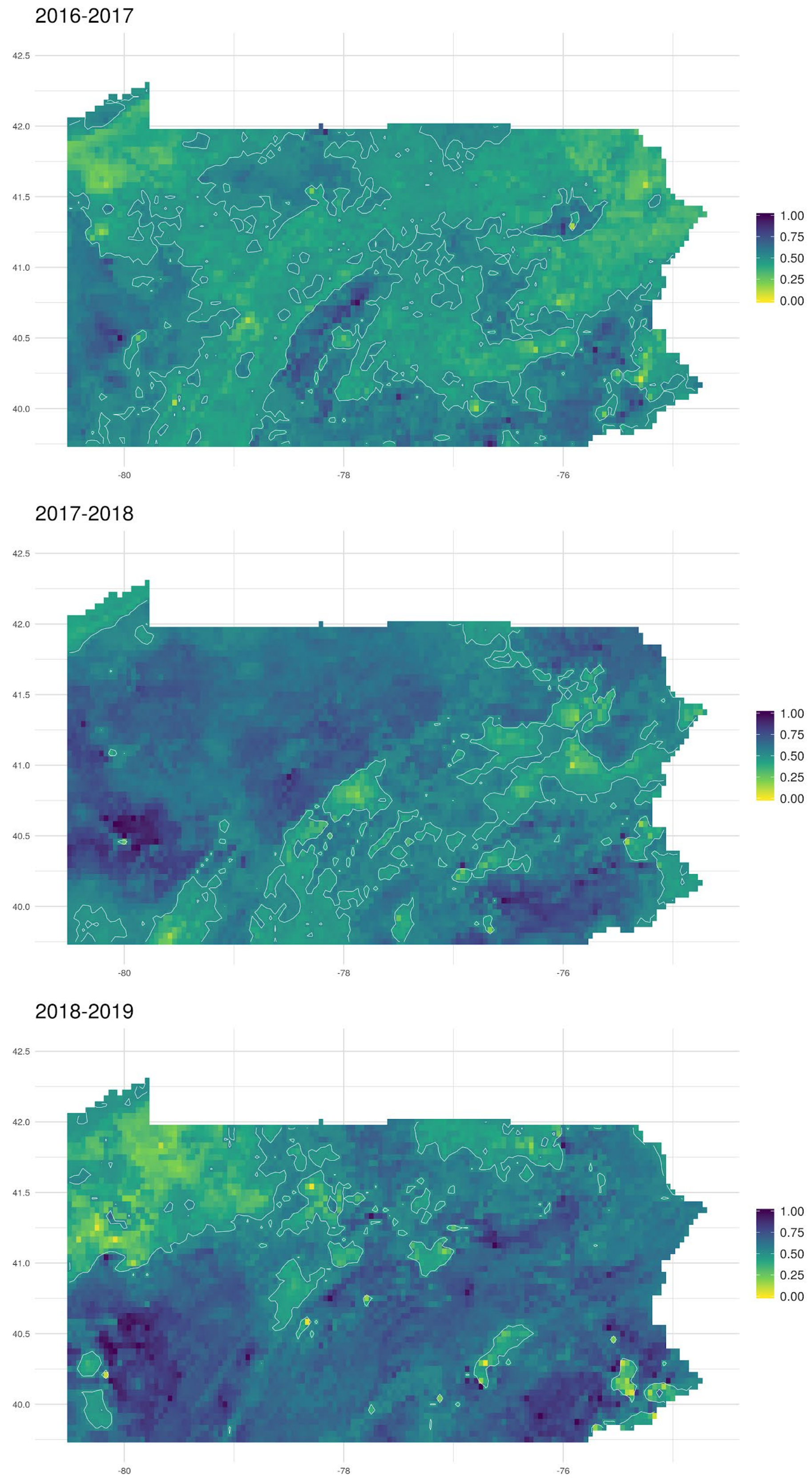

Figure 5. Weather-based prediction maps of the probability of honey bee colony survival for the most recent three years of PRISM data. Contour lines show the 0.5 probability level.

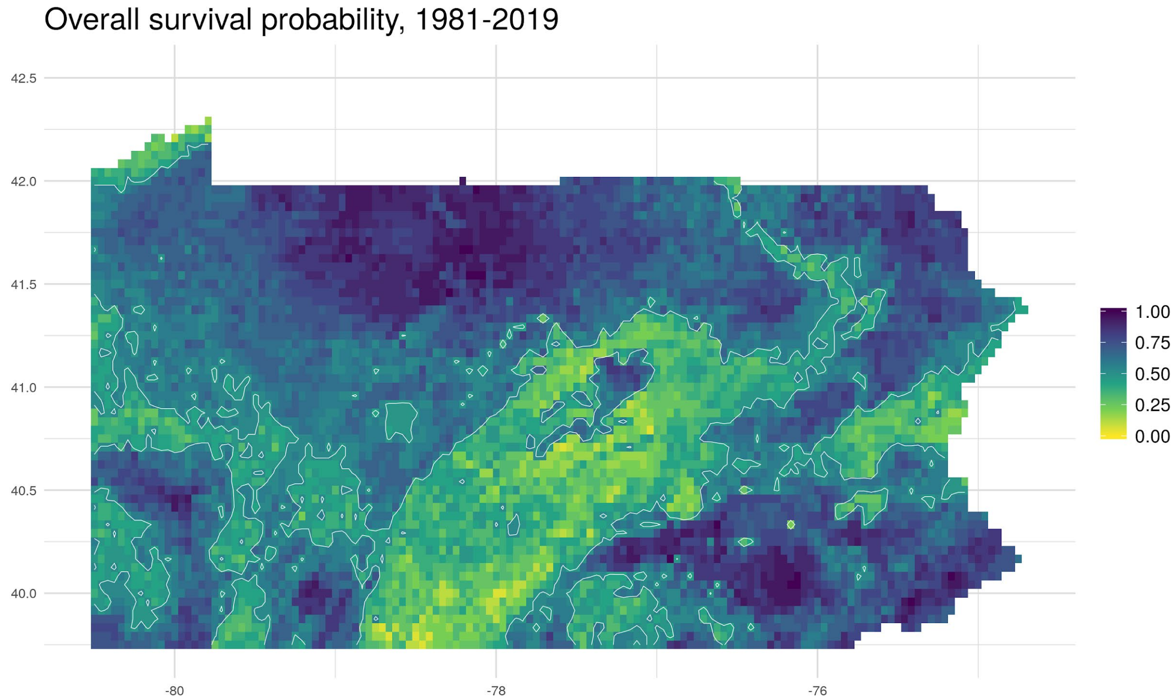

Figure 6. Mean probability of colony survival for 1981-2019. Contour lines show the 0.5 probability level.

Table 1. Weather and topographic variables hypothesized to affect honey bee overwintering survival. Weather variables include both BIOCLIM (Busby, 1991) and agronomic indices, as well as bee-specific variables developed for this study. Larger permutation-based variable importance values are more influential. The top two variables are bold. Autumn: September, October, November. Winter: December, January, February.

Variable description	Unit	Variable importance
Weather		
Winter minimum temperature	c	0.0151
Winter total precipitation	mm	0.0164
Winter days within the bee-optimal temperature range -5C to +10C	d	0.0111
Winter days with maximum temperature above 16C and precipitation below 3mm	d	0.0119
Winter minimum temperature variation	C	0.0118
Autumn total precipitation	mm	0.0128
Growing degree days (base 5 C)	**C**	**0.0252**
Days between rain events > 0.25 mm	mm	0.0127
BIOCLIM 2: Mean diurnal temperature range	C	0.0151
BIOCLIM 3: Temperature isothermality		0.0196
BIOCLIM 4: Temperature seasonality	C	0.0132
BIOCLIM 5: Maximum temperature of warmest month	C	0.0201
BIOCLIM 6: Minimum Temperature of Coldest Month	C	0.0131
BIOCLIM 7: Temperature Annual Range	C	0.0130
BIOCLIM 8: Mean Temperature of Wettest Quarter	C	0.0140
BIOCLIM 9: Mean Temperature of Driest Quarter	C	0.0147

BIOCLIM 12: Annual Precipitation	mm	0.0148
BIOCLIM 16: Precipitation of Wettest Quarter	mm	0.0202
BIOCLIM 17: Precipitation of Driest Quarter	mm	0.0122
BIOCLIM 18: Precipitation of Warmest Quarter	**mm**	**0.0213**
BIOCLIM 19: Precipitation of Coldest Quarter	mm	0.0157
Topography		
Elevation	m	0.0154
Slope		0.0126
Potential incident solar radiation, 21 Dec	$Wh*m^{-2}*d^{-1}$	0.0156
Profile curvature	m^{-1}	0.0111
Terrain curvature	m^{-1}	0.0110
Topographic wetness index		0.0116
East/West orientation of slope		0.0102
North/South orientation of slope		0.0103
Landscape		
Distance-weighted Insect Toxic Load		0.0112
Distance-weighted Forage Quality autumn		0.0104
Management		
Beekeeper years of experience		0.0040
Number of colonies in November		0.0057

Table 2. Confusion matrix for model predictions of colony-level honey bee overwintering survival for the full Random Forest model.

a. Full model		**Actual**	
		Mortality	**Survival**
Predicted	**Mortality**	337 (24%)	87 (6%)
	Survival	192 (13%)	813 (57%)