A Spatial Approach to Analyzing the Effects of Weather Patterns on Honey Bee (*Apis mellifera*) Colony Loss Across the United States

by

Kaj Overturf

A thesis submitted to the Graduate Faculty of Auburn University in partial fulfilment of the requirements for the Degree of Master of Science in Geography

Auburn, Alabama

May 7, 2022

Approved by

Stephanie Rogers, Chair, Assistant Professor of Geosciences

Geoffrey Williams, Assistant Professor of Entomology and Plant Pathology

Roberto Molinari, Assistant Professor of Mathematics and Statistics

Abstract

Western honey bees (*Apis mellifera*) have experienced high rates of colony loss over the past decade. Honey bees are critical to agricultural production, providing 15 billion to the United States (U.S.) economy each year via their pollination services. Due to their importance, studies have attempted to determine what factors are behind these observed losses. However, relatively few studies have considered the effects of weather in their analyses despite weather being a key driver in ecological systems. Those studies that have considered the effects of weather have not utilized spatial analysis despite studies in other fields finding that issues such as non-stationarity, where relationships between variables differ across space, and the modifiable areal unit problem (MAUP) may affect analysis results.

The objectives of Chapter 2 were to determine which weather variables best predict winter colony loss rates at a national scale in the U.S. and then compare the results from a traditional non-linear approach, using a generalized linear regression (GLR), to results from a spatial approach, using a geographically weighted regression (GWR) which takes into consideration data non-stationarity, thus helping to elucidate how changes may occur across space. The best supported variables at the national scale were mean maximum temperature during the month of November, mean precipitation during the month of February, mean windspeed, and mean elevation. The GWR had an AIC score that was 328.80 points lower than the GLR and had an R-squared value of 20% versus 13.5% for the GLR. Thus, these results show that a spatial approach is more statistically robust than the traditional GLR and indicate that the effects of weather on colony loss are non-stationary.

The objective of Chapter 3 was to determine how MAUP would change model results when a linear regression was run at six different scales - zip code, county, state, and level one, two, and three ecoregions. MAUP is the phenomenon of different data aggregation methods resulting in different results. Results indicated that the effects of the two variables analyzed – mean temperature and mean precipitation – differed substantially between the various scales of analysis. Additionally, the R-squared values changed drastically, with of low of 8% at the zip code level and a high of 73% for level two ecoregions. These results are consistent with findings from other fields and indicate that MAUP should be considered when analyzing an aggregated honey bee colony loss dataset.

The results of this thesis show that spatial phenomena such as non-stationarity and MAUP can alter the results of honey bee colony loss analyses, sometimes to a great degree. This highlights the need for more localized management strategies, as the effects of weather, and potentially other variables, on colony loss vary by location. Additionally, this thesis shows the need for more research involving spatial analysis in this field. Future studies may seek to analyze other variables, such as varroa mite levels or pesticide presence, alongside weather variables to create a fuller model that may further aid local management decisions.

Acknowledgments

I would like to thank my advisor Dr. Stephanie Rogers who has guided me through graduate school and without whom this thesis would not be possible. I would like to additionally thank my committee members, Dr. Geoffrey Williams and Dr. Roberto Molinari, for their input and guidance throughout this project, and the Bee Informed Partnership for providing the data for this thesis. Finally, I would like to thank the members of my lab, as well as all the other colleagues and friends I have made while at Auburn for their help and support over the past two years.

Table of Contents

List of Tables

List of Figures

Chapter 1

Literature Review

1.1 Introduction

Weather can have a significant effect on the population dynamics of insects, with temperature and rainfall affecting abundance and survival (Williams 1951; Debach 1958; Williams 1961; Drake 1994). Despite knowledge that weather affects population dynamics, it is often unclear how and to what degree (Knape and Valpine 2011). This information may be particularly important for understanding colony loss of honey bees. Western honey bees (*Apis melifera*), hereafter referred to as honey bees, have seen high colony losses over the past decade, with an average annual loss rate of 39% in the United States and 43% of colonies having been lost between April of 2019 and 2020 (Bruckner et al. 2020). Pollinator species are responsible for pollinating around 100 crops that account for 90 percent of food consumed in the majority of the countries of the world (Kluser and Peduzzi 2007) with honey bees responsible for around 15.5 percent on their own (Nabhan and Buchmann 1997). The pollination services provided by honey bees have been estimated to be worth anywhere between 4.5 and 40 billion dollars to the U.S. economy each year (Gill 1991), with Calderone (2012) finding their contribution to be worth 15.12 billion dollars as of 2009. Due to the importance of honey bees to our food production and economy, many studies have researched the reasons for the high losses we have observed. These studies have looked at a range of factors, from *Varroa destructor*, a parasitic mite (Zee et al. 2015; Steinhauer et al. 2018), to diseases such as American foulbrood (*Paenibacillis larvae)* (vanEngelsdorp and Meixner 2010; Steinhauer et al. 2018) to beekeeping practices (Steinhauer et al. 2018; El Agrebi et al. 2021, to nutrition (Perry et al. 2015; Smart et al. 2016; Steinhauer et al. 2018). Past studies have found that numerous factors, likely working in conjunction, can lead to honey bee colony loss (Dainat et al. 2012; Havard et al. 2019; Bird et al. 2020; Bruckner et al. 2021). However, despite the large number of studies on honey bee loss, the effects of weather patterns on their survival remains somewhat understudied (Havard et al. 2019). Additionally, most studies have not utilized any form of spatial analysis to this point.

1.2 Spatial Background

The first law of geography states that all things are related, but nearer things are more related than distant things (Tobler 1965). An important concept in the natural world, relating to this law, is spatial heterogeneity, or spatial variance. This is where attribute values for various features, be they temperature, elevation, etc., tend to differ across space. This is considered a central causal factor in ecological systems; variation across space is critical for explaining the phenomena we observe (Pickett and Cadenasso 1995). Despite this, spatial analysis in honey bee colony loss research is lacking, with very few studies including any aspect of it. Two exceptions include Zee et al. (2014) and Zee et al. (2015) which both found that the effects of predictor variables such as varroa mite loads and pesticides varied by region in Europe. There have been no studies to date in this field that address two issues relating to space: non-stationarity and the modifiable areal unit problem (MAUP).

1.2.1 Geographically Weighted Regressions

Geographically weighted regressions (GWR) build upon Tobler's first law by considering spatially local effects of predictor variables. This method was first described by Brunsdon et al. (1996). One issue with traditional linear modelling is the potential for non-stationarity, where the relationships between model variables are not consistent across space (Brunsdon et al. 1996). GWRs address this issue by creating a local regression for each feature in the dataset using a defined neighborhood size, either set as a distance band or a number of neighbors. Once a neighborhood size is defined, regressions are run using one of two local weighting functions: Bisquare or Gaussian. These methods are largely similar, with features in the neighborhood further away from the target feature being weighted less. The Bisquare method gives no weight to features outside of the defined neighborhood while the Gaussian method does at a decreasing rate (ESRI 2022).

GWRs allow for the analysis of spatial data which may not be ideal for a traditional non-spatial regression (Brunsdon et al. 1996). There have been many studies to date that have utilized GWR's including Xu et al. (2019), which used them to analyze the effects of various factors on pollution levels, Zarei et al. (2016), which used them to determine factors behind brown bear den selection, and Pirdavani et al. (2014), which sought to determine the effects of a teleworking policy on traffic safety. Despite work from past studies indicating that the effects of predictor variables vary across space, there have been no honey bee studies to date that have utilized GWRs. This is potentially problematic, as management suggestions may be made without the realization that the effects of the variables analyzed may not be constant across the study area.

1.2.2 The Modifiable Areal Unit Problem (MAUP)

Many studies deal with aggregated datasets, where point data, representing measurements such as temperature, rainfall, etc. are averaged over an area of interest (e.g., state or census tract), typically represented as a polygon. Aggregating data introduces the issue of MAUP, first described by Openshaw and Taylor (1979). MAUP is the phenomenon of different data aggregation methods producing different results, which can lead to different interpretations depending on the method used (Openshaw and Taylor 1979). There are two aspects of MAUP: the scale of analysis, also known as the scale effect, and the unit definition, also known as the zone effect (Openshaw and Taylor 1979; Horner and Murray 2002). The scale effect is based on the number of polygons used in the analysis, with fewer, larger polygons resulting in reduced variation (Openshaw and Taylor 1979; Dark and Bram 2007). The zone effect is based on the changes in results that can be caused by using the same number of polygons in different arrangements within a study area (Openshaw and Taylor 1979; Dark and Bram 2007). One common example of the zone effect is gerrymandering, where political districts are drawn in different shapes to provide support for one political party over another (Stehle 2022). Both of these aspects have been found to alter the results of analyses in studies in fields such as human geography, physical geography, and landscape ecology (Openshaw and Taylor 1979; Fotheringham and Wong 1991; Jelinski and Wu 1996; Horner and Murray 2002; Dark and Bram 2007). Some examples include Horner and Murray (2002), which found that aggregation methods have a large impact on excess commuting calculations, Dark and Bram (2007), which presented examples of MAUP in physical geography, such as hydrologic modelling, and Jelinski and Wu (1996), which found that MAUP may affect results from landscape analysis. Despite previous work in the aforementioned fields, MAUP has not been considered in research involving honey bee colony loss. Should MAUP have similar effects as those found in other fields, ignoring this issue when analyzing aggregated datasets may lead to a misunderstanding of the effects of predictor variables on colony loss.

1.3 Previous Work on the Effects of Weather on Colony Loss

Climate and weather are arguably the greatest drivers of natural systems (Daly and Bryant 2013). Climate is the long-term average of weather conditions, while weather is experienced on a shorter-term, daily, basis. Despite the importance of climate and weather in natural systems, there have been few studies that have investigated how weather affects honey bee colony loss to date (Havard et al. 2019). Switanek et al. (2017) conducted a study in Austria that sought to determine how temperature and precipitation affected colony loss rates. This study calculated conditions for every month of the year and found that higher temperatures each month were correlated with higher losses, with the exceptions of February and November, where higher temperatures were correlated with lower losses. Greater amounts of precipitation were correlated with fewer colony losses for every month except for October (Switanek et al. 2017). Calovi et al. (2021) found that four factors: growing degree days, maximum temperature of the warmest month, precipitation of the warmest quarter, and precipitation of the wettest quarter, best predicted winter mortality in Pennsylvania. Beyer et al. (2018) found that warm and wet conditions during winter and cool and wet conditions in July were associated with higher colony losses in Luxembourg (Table 1.1). Despite some of this previous work, the effects of weather on honey bee colony loss are still somewhat unclear and remain a gap in knowledge in the honey bee research field (Havard et al. 2019).

Table 1.1. Summary of weather effects linked to honey bee (*Apis mellifera*) colony loss found by three previous studies.

Positive effect (increased loss)	**Negative effect (decreased loss)**	**Study**
Not applicable: Non-linear effects	Not applicable: Non-linear effects	Calovi et al. 2021
Warm and wet conditions in winter Cool and wet conditions in July	None found	Beyer et al. 2018
Higher temperatures for most months	Higher temperatures for February and November Greater amounts of precipitation for most months	Switanek et al. 2017

1.4 Seasonal Effects on Honey Bee Colony Loss

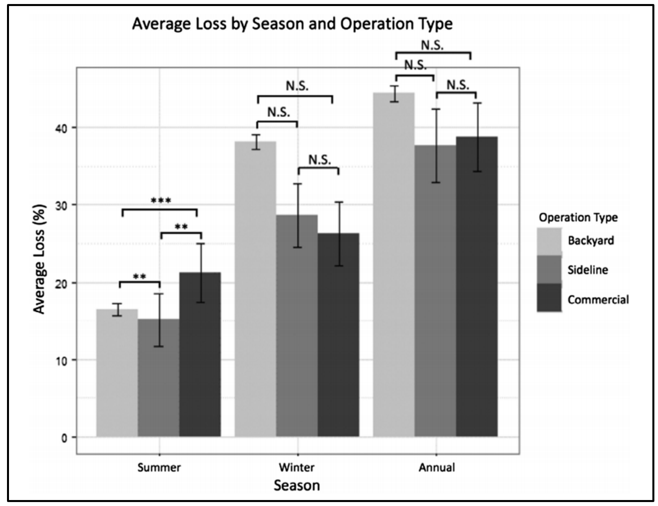

Figure 1.1. Average loss of honey bee (*Apis mellifera*) colonies by season for 2015-2016 in the United States, split by operation type: Backyard (small), Sideline (medium), and Commercial (large). Stars indicate varying levels of significance between losses for different operation types, while N.S. indicates no significance. From Kulhanek et al. (2017).

Past studies have analyzed honey bee colony loss during both summer and winter seasons, with beekeepers generally experiencing greater colony loss during winter (Figure 1.1) (Steinhauer et al. 2014; Kulhanek et al. 2017). Several factors help explain why this period of the year sees greater colony loss. Honey bees must live off food they have previously collected during the winter months as flowers will no longer be in bloom. Additionally, they will not produce additional bees to replace those that die until spring, with the exception of some southern areas of the U.S. such as parts of Florida, Alabama and Georgia. Cold temperatures can be deadly to honey bees. Honey bees will form a cluster – a congregation of bees intended to keep the colony warm during cold weather – at around 15 degrees Celsius (Phillips et al. 1914; Free and Spencer-Booth 1958). The exterior of clusters have been measured as low as 4.5 degrees Celsius, but at temperatures below 8 degrees Celsius honey bees will become inactive and eventually die (Simpson 1961). Thus, long periods of cold temperatures may be detrimental to honey bee colony survival. While some areas of the country may experience higher losses during summer than winter, likely due to the mild nature of the winters, the overall average clearly shows a higher loss rate during winter months (Bruckner et al. 2020). Thus, the focus of this study is on winter mortality.

1.5 Conclusion

Honey bees are critical to agricultural production and the economy, which makes the high levels of annual colony loss observed in the U.S. concerning. Despite the many studies that have sought to determine the causes of colony loss, the effects of weather on colony loss rates remain understudied despite knowledge that weather is a key driver in ecological systems. Even fewer studies have incorporated spatial aspects into their analyses, with no honey bee studies considering the impacts that non-stationarity or MAUP may have on their results. Previous studies in other fields have shown that these issues can greatly affect model results, making their investigation in the field of honey bee research critical. Investigations into the effects of weather on colony loss using spatial approaches may yield new insights into the causes of the high colony loss rates we have observed and may result in more region-specific management suggestions than were previously possible.

1.6 References

Beyer, M., J. Junk, M. Eickermann, A. Clermont, F. Kraus, C. Georges, A. Reichart, and L. Hoffmann. 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* 118 (June): 52–60. https://doi.org/10.1016/j.rvsc.2018.01.012.

Bird, G., A. E. Wilson, G. R. Williams, and N. B. Hardy. 2021. "Parasites and Pesticides Act Antagonistically on Honey Bee Health." *Journal of Applied Ecology* 58 (5): 997–1005. https://doi.org/10.1111/1365-2664.13811.

Bruckner, S., L. Straub, P. Neumann, and G. R. Williams. 2021. "Synergistic and Antagonistic Interactions Between Varroa Destructor Mites and Neonicotinoid Insecticides in Male Apis Mellifera Honey Bees." *Frontiers in Ecology and Evolution* 9. https://doi.org/10.3389/fevo.2021.756027

Bruckner, S., N. Steinhauer, J. Engelsma, A. M. Fauvel, K. Kulhanek, E. Malcolm, A. Meredith, M. Milbrath, E. Niño, J. Rangel, K. Rennich, D. Reynolds, R. Sagili, J. Tsuruda, D. vanEngelsdorp, S. D. Aurell, M. Wilson, and G. Williams. 2020. "2019-2020 Honey Bee Colony Losses in the United States: Preliminary Results." Unpublished work.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical Analysis* 28 (4): 281–98. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

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. https://doi.org/10.1371/journal.pone.0037235.

Calovi, M., C.M. Grozinger, D.A. Miller, and S. C. Goslee. 2021. “Summer weather conditions influence winter survival of honey bees (*Apis mellifera*) in the northeastern United States.” *Scientific Reports* 11: 1553. https://doi.org/10.1038/s41598-021-81051-8

Dainat, B., J. D. Evans, Y. P. Chen, L. Gauthier, and P. Neumann. 2012. “Predictive Markers of Honey Bee Colony Collapse.” *PLOS ONE* 7 (2): e32151. https://doi.org/10.1371/journal.pone.0032151.

Daly, C., and K. Bryant. 2013. “The PRISM climate and weather system—an introduction.” Online. Northwest Alliance for Computational Science and Engineering, Oregon State University, Corvallis, USA. http://prism. oregonstate.edu/.

Dark, S. J., and D. Bram. 2007. “The Modifiable Areal Unit Problem (MAUP) in Physical Geography.” *Progress in Physical Geography: Earth and Environment* 31 (5): 471–79. https://doi.org/10.1177/0309133307083294.

Debach, P. 1958. “The Role of Weather and Entomophagous Species in the Natural Control of Insect Populations12.” *Journal of Economic Entomology* 51 (4): 474–84. https://doi.org/10.1093/jee/51.4.474.

Di Prisco, G., V. Cavaliere, D. Annoscia, P. Varricchio, E. Caprio, F. Nazzi, G. Gargiulo, and F. Pennacchio. 2013. “Neonicotinoid Clothianidin Adversely Affects Insect Immunity and Promotes Replication of a Viral Pathogen in Honey Bees.” *Proceedings of the National Academy of Sciences* 110 (46): 18466. https://doi.org/10.1073/pnas.1314923110.

Drake, V. 1994. "The Influence of Weather and Climate on Agriculturally Important Insects: An Australian Perspective." *Australian Journal of Agricultural Research* 45 (3): 487. https://doi.org/10.1071/AR9940487.

El Agrebi, N., N. Steinhauer, S. Tosi, L. Leinartz, D. C. de Graaf, and C. Saegerman. 2021. "Risk and Protective Indicators of Beekeeping Management Practices." *Science of The Total Environment* 799 (December): 149381. https://doi.org/10.1016/j.scitotenv. 2021.149381.

ESRI. "How Geographically Weighted Regression (GWR) Works—ArcGIS Pro | Documentation." 2022. Accessed January 24, 2022. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-geographicallyweightedregression-works.htm.

Fotheringham, A.S. and Wong, D.W.S. 1991. "The Modifiable Areal Unit Problem in Multivariate Statistical Analysis." *Environment and Planning* 23(7): 1025-1044.

Free, J. B., and Y. Spencer-Booth. 1958. "Observations on the Temperature Regulation And Food Consumption of Honeybees."

Gill, R. 1991. "The value of Honeybee Pollination to Society." *Acta Horticulturae* 288, 6th pollination symposium.

Havard, T., M. Laurent, and M. Chauzat. 2019. "Impact of Stressors on Honey Bees (Apis Mellifera; Hymenoptera: Apidae): Some Guidance for Research Emerge from a Meta-Analysis." *Diversity* 12 (1): 7. https://doi.org/10.3390/d12010007.

Henry, M., M. Beguin, F. Requier, O. Rollin, J.-F. Odoux, P. Aupinel, J. Aptel, S. Tchamitchian, and A. Decourtye. 2012. "A Common Pesticide Decreases Foraging Success and Survival in Honey Bees." *Science* 336 (6079): 348–50. https://doi.org/10.1126/science.1215039.

Hennessy, G., C. Harris, C. Eaton, P. Wright, E. Jackson, D. Goulson, and F. F. L. W. Ratnieks. 2020. “Gone with the Wind: Effects of Wind on Honey Bee Visit Rate and Foraging Behaviour.” *Animal Behaviour* 161 (March): 23–31. https://doi.org/10.1016/j.anbehav. 2019.12.018.

Horner, M. W., and A. T. Murray. 2002. “Excess Commuting and the Modifiable Areal Unit Problem.” *Urban Studies* 39 (1): 131–39. https://doi.org/10.1080/00420980220099113.

Jelinski, E. D. and Wu, J. 1996. “The modifiable area unit problem and implications for landscape ecology.” *Landscape Ecology* 11(3):129-140.

Kluser, S., and P. Peduzzi. 2007. “Global Pollinator Decline: A Literature Review.” UNEP/GRID- Europe.

Knape, J., and P. D. Valpine. 2011. “Effects of Weather and Climate on the Dynamics of Animal Population Time Series.” *Proceedings of the Royal Society B: Biological Sciences* 278 (1708): 985–92. https://doi.org/10.1098/rspb.2010.1333.

Kulhanek, K., N. Steinhauer, K. Rennich, D. M. Caron, R. R. Sagili, J. S. Pettis, J. D. Ellis, M. E. Wilson, J. T. Wilkes, D. R. Tarpy, R. Rose, K. Lee, J. Rangel, and D. vanEngelsdorp. 2017. “A National Survey of Managed Honey Bee 2015–2016 Annual Colony Losses in the USA.” *Journal of Apicultural Research* 56 (4): 328–40. https://doi.org/ 10.1080/00218839. 2017.1344496.

Morton, H. L., and J. O. Moffett. 1972. “Ovicidal and Larvicidal Effects of Certain Herbicides on Honey Bees.” *Environmental Entomology* 1 (5): 611–14. https://doi.org/10.1093/ee/ 1.5.611.

Morton, H. L., J. O. Moffett, and R. H. Macdonald. 1972. "Toxicity of Herbicides to Newly Emerged Honey Bees." *Environmental Entomology* 1 (1): 102–4. https://doi.org/10.1093/ee/1.1.102

Nabhan, G., and S. Buchmann. "Services Provided by Pollinators" in Nature's Services: Societal Dependence on Natural Ecosystems. Edited by Gretchen Daily. Page 138. Washington, D. C. Island Press. 1997.

Openshaw, S. and P. J. Taylor, 1979. "A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem." *Statistical Applications in the Spatial Sciences*, 127–144. London: Pion.

Perry, C. J., E. Søvik, M. R. Myerscough, and A. B. Barron. 2015. "Rapid Behavioral Maturation Accelerates Failure of Stressed Honey Bee Colonies." *Proceedings of the National Academy of Sciences* 112 (11): 3427–32. https://doi.org/10.1073/pnas.1422089112.

Pettis, J. S., D. vanEngelsdorp, J. Johnson, and G. Dively. 2012. "Pesticide Exposure in Honey Bees Results in Increased Levels of the Gut Pathogen Nosema." *Naturwissenschaften* 99 (2): 153–58. https://doi.org/10.1007/s00114-011-0881-1.

Pettis, J. S., E. M. Lichtenberg, M. Andree, J. Stitzinger, R. Rose, and D. vanEngelsdorp. 2013. "Crop Pollination Exposes Honey Bees to Pesticides Which Alters Their Susceptibility to the Gut Pathogen Nosema Ceranae." *PLOS ONE* 8 (7): e70182. https://doi.org/10.1371/journal.pone.0070182.

Phillips, E. F., and G. S. Demuth. 1914. "The Temperature of the Honeybee Cluster in Winter." U.S. Department of Agriculture.

Pickett, S. and M. Cadenasso. 1995. "Landscape Ecology: Spatial Heterogeneity in Ecological Systems." *Science* 269: 331-334.

Pirdavani, A., T. Bellemans, T. Brijs, B. Kochan, and G. Wets. 2014. "Assessing the Road Safety Impacts of a Teleworking Policy by Means of Geographically Weighted Regression Method." *Journal of Transport Geography* 39 (July): 96–110. https://doi.org/10.1016/j.jtrangeo.2014.06.021.

Rangel, J., and D. Tarpy. 2016. "The combined effects of miticides on the mating health of honey bee (*Apis mellifera* L.) queens." *Journal of Apicultural Research* 54 (3)

Simpson, J. 1961. "Nest Climate Regulations in Honey Bee Colonies." *Science.*133 (3416): 1327-1333.

Smart, M., J. Pettis, N. Rice, Z. Browning, and M. Spivak. 2016. "Linking Measures of Colony and Individual Honey Bee Health to Survival among Apiaries Exposed to Varying Agricultural Land Use." *PLOS ONE* 11 (3): e0152685. https://doi.org/10.1371/journal.pone.0152685.

Stehle, S. 2022. "Temporal Aggregation Bias and Gerrymandering Urban Time Series." *GeoInformatica* 26 (1): 233–52. https://doi.org/10.1007/s10707-021-00452-z.

Steinhauer, N., K. Kulhanek, K. Antúnez, H. Human, P. Chantawannakul, M. Chauzat, and D. vanEngelsdorp. 2018. "Drivers of Colony Losses." *Current Opinion in Insect Science*, 26: 142–48. https://doi.org/10.1016/j.cois.2018.02.004.

Steinhauer, N. A, K. Rennich, M. E. Wilson, D. M. Caron, E. J. Lengerich, J. S. Pettis, R. Rose, J. Skinner, D. Tarpy, J. Wilkes, and D. vanEngelsdorp. 2014. "A National Survey of Managed Honey Bee 2012–2013 Annual Colony Losses in the USA: Results from the Bee Informed Partnership." *Journal of Apicultural Research* 53 (1): 1–18. https://doi.org/10.3896/IBRA.1.53.1.01.

Switanek, M., K. Crailsheim, H. Truhetz, and R. Brodschneider. 2017. “Modelling Seasonal Effects of Temperature and Precipitation on Honey Bee Winter Mortality in a Temperate Climate.” *Science of The Total Environment* 579 (February): 1581–87. https://doi.org/10.1016/j.scitotenv.2016.11.178.

Tobler, W. R. 1965. “Computation of the Correspondence of Geographical Patterns.” *Papers of the Regional Science Association* 15 (1): 131–39. https://doi.org/10.1007/BF01947869.

vanEngelsdorp, D., and M. D. Meixner. 2010. “A Historical Review of Managed Honey Bee Populations in Europe and the United States and the Factors That May Affect Them.” *Journal of Invertebrate Pathology* 103 (January): S80–95. https://doi.org/10.1016/j.jip.2009.06.011.

Williams, C.B. 1951. “Changes in Insect Populations in the Field in Relation to Preceding Weather Conditions.” *Proceedings of the Royal Society of London. Series B - Biological Sciences* 138 (890): 130–56. https://doi.org/10.1098/rspb.1951.0011.

Williams, C. B. 1961. “Studies in the Effect of Weather Conditions on the Activity and Abundance of Insect Populations.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 244 (713): 331–78. https://doi.org/10.1098/rstb.1961.0011.

Zee, R., A. Gray, L. Pisa, and T. de Rijk. 2015. “An Observational Study of Honey Bee Colony Winter Losses and Their Association with Varroa Destructor, Neonicotinoids and Other Risk Factors.” *PLOS ONE* 10 (7): e0131611. https://doi.org/10.1371/journal.pone.0131611.

Zee, R., R. Brodschneider, V. Brusbardis, J.D. Charrière, R. Chlebo, M. Coffey, B. Dahle, et al. 2014. “Results of International Standardised Beekeeper Surveys of Colony Losses for

Winter 2012–2013: Analysis of Winter Loss Rates and Mixed Effects Modelling of Risk Factors for Winter Loss." *Journal of Apicultural Research* 53 (1): 19–34. https://doi.org/10.3896/IBRA.1.53.1.02.

Zarei, A. A., S. Abedi, M. Mahmoudi, and S. Peyravi Lati. 2016. "Assessment of Brown Bear's (Ursus arctos syriacus) Winter Habitat Using Geographically Weighted Regression and Generalized Linear Model in South of Iran." *Iranian Journal of Applied Ecology* 4 (14): 75–88. https://doi.org/10.18869/acadpub.ijae.4.14.75.

Chapter 2

Investigating Spatial and Non-Spatial Approaches for Understanding Honey Bee (*Apis mellifera*) Winter Colony Loss Across the United States

2.1 Introduction

Western honey bees (*Apis mellifera*; hereafter honey bees) have experienced high colony losses in the northern hemisphere over the last decade, with a mortality rate of over 43% in the United States (U.S.) between April of 2019 and 2020 alone (Bruckner et al. 2020). Honey bees are responsible for pollinating many important crops and, as a result, contribute significantly to the economy and food security (Southwick and Southwick 1992). Their monetary contribution to the agricultural industry in the U.S. has been estimated to be approximately 15 billion dollars per year (Calderone 2012). Due to the importance of honey bees, many studies have investigated the reasons for these high losses, which range from pests and diseases such as *Varroa destructor* (Zee et al. 2015; Steinhauer et al. 2018) and American foulbrood (*Paenibacillis larvae*) among many others (vanEngelsdorp and Meixner 2010; Steinhuaer et al. 2018), to beekeeping practices (Steinhauer et al. 2018; El Agrebi et al. 2021), to nutrition (Perry et al. 2015; Smart et al. 2016; Steinhauer et al. 2018). Past studies have found that numerous factors, likely working in conjunction, have detrimental effects on honey bee colonies (Dainat et al. 2012; Havard et al. 2019; Bird et al. 2020; Bruckner et al. 2021). Despite the numerous studies on honey bee colony loss, the effects of weather patterns on their survival remains understudied (Havard et al. 2019).

Weather is arguably one of the greatest drivers of natural systems (Daly and Bryant 2013). Despite the importance of weather in natural systems, relatively few studies have considered how weather variables may affect the loss rates of honey bee colonies. A few examples of studies that have considered the effects of weather on colony loss include Switanek et al. (2017), Beyer et al. (2018), and Calovi et al. (2021), although none of these studies have encompassed an area the size and scale of the continental U.S. Switanek et al. (2017) conducted a study in Austria that sought to determine how temperature and precipitation affected colony loss rates. This study calculated weather conditions for each month of the year and found that higher monthly temperatures were correlated with higher losses, with the exceptions of February and November, where higher temperatures were correlated with lower losses. Additionally, they found that greater amounts of precipitation were correlated with fewer colony losses for every month except for October. These results differ from Beyer et al. (2018), that found warm and wet conditions during winter months and cool and wet conditions in July were associated with higher colony losses in Luxembourg. Calovi et al. (2021) found that four factors: growing degree days, maximum temperature of the warmest month, precipitation of the warmest quarter, and precipitation of the wettest quarter, best predicted winter mortality in the State of Pennsylvania. Despite some of this previous work, the effects of weather on honey bee colony loss are still unknown at the national scale in the U.S. and remain somewhat of a gap in knowledge in the honey bee research field (Havard et al. 2019).

Another gap in knowledge in the field of honey bee colony loss research is how the effects of environmental variables, such as temperature and precipitation, change across space. Studies that analyze effects across large, heterogenous areas often deal with non-stationarity, where relationships between predictor and response variables differ from location to location (Brunsdon et al. 1996). This phenomenon has been found to affect results in studies ranging from wildlife distribution modeling (Osborne et al. 2007) to hydrological modelling (Deb et al. 2019). Failing to consider non-stationarity may lead to a misunderstanding of the effects variables have on colony loss. Despite this issue, studies investigating the effects of weather on honey bee colony loss have failed to integrate spatial analysis, such as the geographically weighted regression (GWR), which may aid in dealing with non-stationarity (Brunsdon et al. 1996), largely due to the lack of high resolution, spatially explicit datasets.

Given the lack of research involving weather and spatial analysis in this field, the research questions of this study were: 1) which combination of weather variables best explain the observed winter colony loss rates over a nine-year span, and during which month of the year do these variables best explain the observed losses, and 2) does a spatial model (i.e., a GWR) outperform the non-spatial model (i.e., a generalized linear regression (GLR))? It was predicted that weather conditions (i.e., temperature, precipitation, dewpoint, and windspeed) would affect colony loss rates differently across the U.S. due to the presence of non-stationarity, resulting in the GWR being more statistically robust than the GLR.

2.2 Methods

2.2.1 Data Management

2.2.1.1 Colony Loss Data

Colony loss data for the winters of 2011-2012 through 2019-2020 were provided by the Bee Informed Partnership's (BIP) annual national colony loss survey, a citizen science initiative (vanEngelsdorp et al. 2012; Spleen et al. 2013; Steinhauer et al. 2014; Lee et al. 2015; Kulhanek et al. 2017; Bruckner et al. 2020). This dataset details the number of colonies each responding apiary owner managed and the number that were lost during the winter months (defined as between 1 October - 1 April). These values were used to calculate the percent of colonies each apiary owner lost each winter, with responses linked to the zip code the colonies were kept in. While migratory colonies represent the majority of all colonies in the U.S., this study subset the dataset to only stationary beekeepers in order to link weather conditions to loss in each zip code. A total of 9718 zip codes had colony loss data from stationary beekeepers reported over this nine-year period.

Colony loss data were aggregated at the zip code level over the entire nine-year period to reduce the amount of variability within each zip code. This aggregation was calculated by dividing the number of colonies lost within each zip code by the number of colonies managed and multiplying by 100 to obtain the loss rate per 100 colonies over the entire nine-year period. The loss dataset was then joined to the zip code polygon shapefile using zip code as the unique identifier. Zip codes with fewer than ten colonies were removed from the dataset to avoid issues resulting from low sample sizes (Jenkins and Quintana-Ascencio, 2020). This reduced the zip code file from 9718 to 5806 polygons (Figure 2.1).

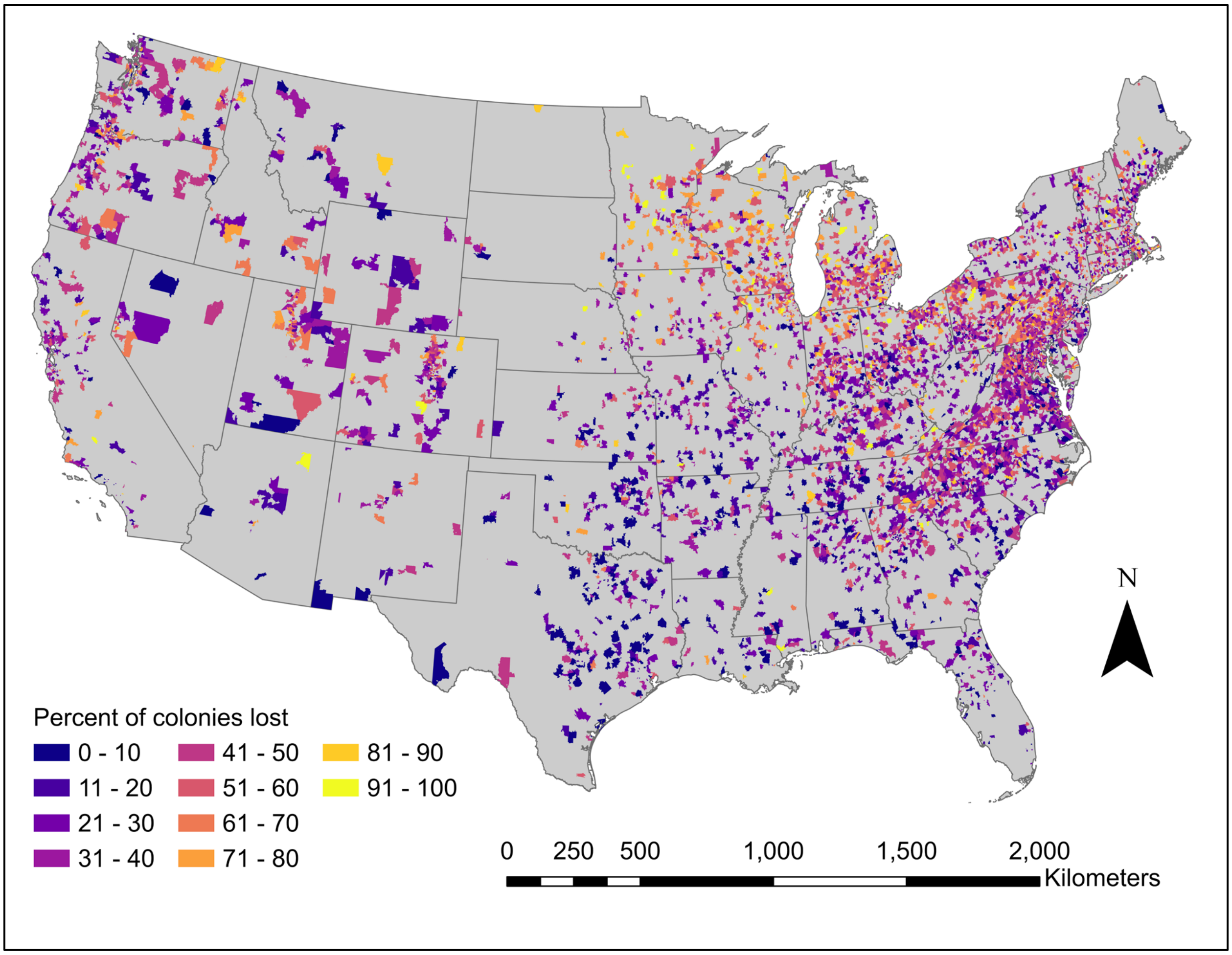

Figure 2.1. Map of winter honey bee (*Apis mellifera*) colony loss rates for stationary beekeepers from the Bee Informed Partnership (BIP) survey data from the winter of 2011-2012 to the winter of 2019-2020. Data were recorded at the zip code level with zip codes having fewer than ten colonies removed. Colors indicate the percent of colonies lost, with purples representing lower losses and yellows representing higher losses.

2.2.1.2 Weather Variables

Seven broad predictor variables were selected for analysis: six weather variables - mean, minimum, and maximum temperature, total precipitation, mean dewpoint, mean windspeed - and elevation, a confounding variable. Five of these predictor variables – all the temperature variables, as well as dewpoint and precipitation – were further divided by month so that each had a total of 13 variables, one for each month and one annual variable. This resulted in a total of 67 candidate predictor variables. All weather data, except wind speed, were acquired from PRISM (Parameter-elevation Regressions on Independent Slopes Model) (Daly and Bryant 2013). These raster datasets detail average weather conditions for each month and year of the study period at 4 km spatial resolution. A dataset for average wind speed at 10 m above the surface from 2008 to 2017 was acquired from the Global Wind Atlas (Badger and Jørgensen 2011). This dataset has a 250 m spatial resolution. For the purposes of this study, the average values from 2008-2017 were assumed to reflect the conditions during the study period. Elevation data at a 1 km resolution were acquired from the USGS (2007). U.S. zip code area polygons were acquired from ESRI (2021).

Next, weather data were aggregated before also being joined to the zip code polygon file. To do this, monthly weather conditions were firstly averaged across the nine-year span to find the mean weather conditions for each set of months using the *Raster Calculator* tool in ArcGIS Pro. For example, the mean temperature for all Januarys during the nine-year period was calculated. This resulted in twelve rasters for each weather variable, with the exception of windspeed, which had one raster for total average values instead of monthly values. Additionally, an average annual raster was created for each variable by simply averaging the nine yearly rasters for each. The *Zonal Statistics* tool was then used to find mean values of each predictor variable for each zip code. An XY point layer was then created which represented the centroid of each zip code and the *Extract Multi-Values to Points* tool was used to link all weather and elevation data to this file. These points were then spatially joined to the corresponding zip code polygons.

2.2.2 Model Selection

To determine the best set of predictor variables for further analysis, a series of GLRs were run in R (R Core Team 2022) using RStudio (RStudio Team 2022) for each of the 67 candidate variables. The Akaike information criterion scores (AIC) (Akaike 1987) for each regression were compared (Table 2.1). Lower AIC scores indicate a better model fit (Akaike 1987). To determine which variables would be candidates for the final model, the variables were organized into groups by theme. For example, all precipitation variables were considered part of the precipitation group. The best supported variable, based on AIC score, for each group was selected as a candidate for further analysis. These variables were November mean maximum temperature, March mean temperature, march mean minimum temperature, February mean precipitation, February mean dewpoint, mean windspeed, and mean elevation (Table 2.1).

Table 2.1. AIC, delta AIC, model likelihood (ModelLik), and AIC weight (AICWt) scores for all candidate variables for explaining variation in honey bee (*Apis mellifera*) winter colony loss rates, ranked by AIC score. The number of variables per model is represented by K. The best model for each covariate (temperature, precipitation, dewpoint, windspeed, elevation) was included as a candidate for the final model. The best supported model for each category is shown in bold.

Model names	K	AIC	Delta_AIC	ModelLik	AICWt
Avg Max Temp November	**3**	**51639.80**	**0**	**1**	**0.37**
Avg Max Temp January	3	51639.99	0.18	0.913215	0.34
Avg Max Temp February	3	51640.38	0.58	0.748442	0.28
Avg Max Temp Annual	3	51645.84	6.04	0.048917	0.02
Avg Max Temp December	3	51657.05	17.24	0.00018	6.65E-05
Avg Max Temp March	3	51661.76	21.95	1.71E-05	6.31E-06
Avg Max Temp April	3	51670.11	30.30	2.63E-07	9.71E-08
Avg Max Temp October	3	51673.14	33.33	5.78E-08	2.13E-08
Avg Mean Temp March	**3**	**51683.72**	**43.91**	**2.91E-10**	**1.07E-10**
Avg Mean Temp February	3	51683.84	44.04	2.74E-10	1.01E-10
Avg Mean Temp April	3	51686.13	46.33	8.72E-11	3.22E-11
Avg Mean Temp Annual	3	51688.37	48.56	2.85E-11	1.05E-11
Avg Mean Temp January	3	51708.83	69.03	1.03E-15	3.78E-16
Avg Mean Temp November	3	51718.37	78.56	8.72E-18	3.22E-18
Avg Mean Temp December	3	51735.65	95.84	1.54E-21	5.69E-22
Avg Mean Temp October	3	51738.89	99.09	3.05E-22	1.12E-22
Avg Min Temp March	**3**	**51745.73**	**105.92**	**9.98E-24**	**3.68E-24**
Avg Min Temp April	3	51750.07	110.27	1.14E-24	4.20E-25
Avg Min Temp February	3	51756.59	116.78	4.38E-26	1.61E-26
Avg Max Temp September	3	51765.72	125.91	4.55E-28	1.68E-28
Avg Mean Dewpoint February	**3**	**51768.76**	**128.96**	**9.93E-29**	**3.66E-29**
Avg Mean Dewpoint April	3	51783.16	143.35	7.44E-32	2.74E-32
Avg Min Temp Annual	3	51785.73	145.92	2.06E-32	7.58E-33
Avg Mean Dewpoint March	3	51802.82	163.02	4.00E-36	1.47E-36
Avg Mean Temp September	3	51808.59	168.79	2.23E-37	8.23E-38
Avg Min Temp January	3	51817.68	177.87	2.38E-39	8.76E-40
Avg Mean Dewpoint January	3	51846.90	207.09	1.07E-45	3.95E-46
Avg Max Temp May	3	51859.49	219.68	1.98E-48	7.30E-49
Avg Min Temp December	3	51865.94	226.13	7.87E-50	2.90E-50
Avg Mean Dewpoint Annual	3	51868.15	228.35	2.60E-50	9.58E-51
Avg Mean Temp August	3	51874.90	235.09	8.91E-52	3.29E-52
Avg Mean Temp May	3	51878.37	238.56	1.57E-52	5.80E-53
Avg Mean Dewpoint December	3	51878.76	238.95	1.29E-52	4.77E-53

Avg Max Temp August	3	51881.09	241.29	4.03E-53	1.49E-53
Avg Min Temp October	3	51885.00	245.20	5.70E-54	2.10E-54
Avg Min Temp November	3	51886.29	246.49	2.99E-54	1.10E-54
Avg Mean Dewpoint November	3	51902.06	262.26	1.13E-57	4.15E-58
Avg Min Temp September	3	51927.13	287.33	4.05E-63	1.49E-63
Avg Min Temp May	3	51936.32	296.51	4.11E-65	1.51E-65
Avg Max Temp June	3	51967.88	328.08	5.73E-72	2.12E-72
Avg Mean Temp June	3	51968.06	328.26	5.25E-72	1.94E-72
Avg Min Temp August	3	51982.08	342.28	4.74E-75	1.75E-75
Avg Mean Dewpoint October	3	51998.8	358.99	1.11E-78	4.10E-79
Avg Mean Dewpoint May	3	52004.52	364.71	6.35E-80	2.34E-80
Avg Max Temp July	3	52018.75	378.94	5.17E-83	1.91E-83
Avg Mean Temp July	3	52021.77	381.96	1.14E-83	4.22E-84
Avg Min Temp June	3	52028.52	388.71	3.91E-85	1.44E-85
Avg Min Temp July	3	52099.71	459.90	1.36E-100	5.02E-101
Avg Mean Dewpoint September	3	52124.3	484.49	6.22E-106	2.30E-106
Avg Mean Dewpoint June	3	52136.71	496.91	1.25E-108	4.63E-109
Avg Mean Dewpoint August	3	52179.7	539.89	5.80E-118	2.14E-118
Avg Mean Dewpoint July	3	52217.4	577.60	3.77E-126	1.39E-126
Avg Precipitation February	**3**	**52236.54**	**596.73**	**2.64E-130**	**9.74E-131**
Avg Precipitation March	3	52251.68	611.88	1.36E-133	5.01E-134
Avg Precipitation December	3	52273.24	633.43	2.83E-138	1.04E-138
Avg Precipitation Annual	3	52306.36	666.55	1.82E-145	6.72E-146
Avg Precipitation January	3	52321.67	681.86	8.61E-149	3.18E-149
Avg Precipitation April	3	52335.8	695.99	7.36E-152	2.72E-152
Avg Precipitation August	3	52369.51	729.71	3.51E-159	1.30E-159
Avg Precipitation November	3	52371.3	731.49	1.44E-159	5.31E-160
Avg windspeed	**3**	**52401.37**	**761.57**	**4.24E-166**	**1.56E-166**
Avg Elevation	**3**	**52403.19**	**763.39**	**1.71E-166**	**6.30E-167**
Avg Precipitation July	3	52404.81	765.01	7.61E-167	2.81E-167
Avg Precipitation May	3	52409.73	769.92	6.51E-168	2.40E-168
Avg Precipitation September	3	52413.1	773.30	1.20E-168	4.44E-169
Avg Precipitation June	3	52415.05	775.24	4.55E-169	1.68E-169
Avg Precipitation October	3	52424.45	784.64	1.52E-171	4.13E-171

Following this, a correlation matrix was generated to determine the degree of correlation between each candidate predictor variable to eliminate highly correlated variables to reduce noise in the modelling process (Table 2.2). This study considered any value over 0.6 to indicate high correlation (Peters 2020). November mean maximum temperature was the best performing variable overall, so variables such as February dewpoint temperature and March mean/minimum temperature, which were highly correlated to November mean maximum temperature, were not included as candidates for further work to avoid the effects of multicollinearity, which can result in uncertain coefficient values (Table 2.2). A series of multi-variable models with combinations of the candidate variables were then created, with the model with the best AIC score being selected as the final model for analysis (Table 2.3).

Table 2.2. Correlation matrix for candidate predictor variables for explaining variation in honey bee (*Apis mellifera*) winter colony loss rates -mean windspeed (Wind), March mean temperature (Tmean), November mean maximum temperature (Tmax), March mean minimum temperature (Tmin), February mean precipitation (Precipitation), mean elevation (Elevation), and February mean dew point (Dewpoint) – within zip codes covered by the Bee Informed partnership (BIP) dataset for the U.S. Values range from -1 to 1 with values closer to 1 indicating correlation and values closer to -1 indicating inverse correlation.

	Wind	Tmean	Tmax	Tmin	Precipitation	Elevation	Dewpoint
Wind		-0.25	-0.29	-0.24	-0.32	-0.06	-0.29
Tmean			0.97	0.99	0.36	-0.2	0.94
Tmax				0.95	0.35	-0.25	0.91
Tmin					0.4	-0.28	0.96
Precipitation						-0.28	0.51
Elevation							-0.33
Dewpoint							

Table 2.3. AIC, delta AIC, model likelihood (ModelLik), and AIC weight (AICWt) scores for possible final models for explaining variation in honey bee (*Apis mellifera*) winter colony loss rates. The number of variables per model is represented by K. Variables include November mean maximum temperature (temperature), February mean precipitation (precipitation), mean elevation (elevation), and mean windspeed (wind). Dewpoint was not included due to its high correlation with temperature and precipitation.

Model names	K	AIC	Delta_AIC	ModelLik	AICWt
Temperature + Precipitation + Elevation + Wind	6	51582.69	0	1	0.999
Temperature + Precipitation + Wind	5	51599.48	16.80	0.000225	0.0002
Temperature + Precipitation + Elevation	5	51609.85	27.16	1.26E-06	1.26E-06
Temperature + Precipitation	4	51618.88	36.20	1.38E-08	1.38E-08

2.2.3 Regression Analysis

With the predictor variables determined, the *Generalized Linear Regression (GLR)* tool from ArcGIS Pro was used to analyze the effect of weather variables on the winter colony loss rate. This was a non-spatial approach that did not consider the location of the zip codes during analysis. Following this, the *Geographically Weighted Regression (GWR)* tool was used to determine if considering spatial variation would improve the model. A GWR is a local form of a linear regression that can model spatially varying relationships and allows coefficient values to differ across space by calculating a regression equation for every feature by incorporating its neighborhood of features, which can be defined in several ways (Brunsdon et al. 1996). The GWR was set to define the neighborhoods needed for analysis by number of neighbors using the golden search method, which searches for the number of neighbors that give the lowest AIC score (ESRI 2022). GWR's weight neighbors nearer to the target feature heavier than those further away. The Bisquare weighting method was selected, which slowly reduces the weight of features further away and assigns features outside of a neighborhood a weight of zero (ESRI 2022). All work was done with the North America equidistant conic projection, which allows for distance calculations to be conducted accurately.

2.3 Results

2.3.1 Best Supported Months from Model Selection

The ability of weather patterns to explain colony loss, based on AIC score, differed between months. Two variables in the final model had monthly data associated with them: mean maximum temperature and mean total precipitation. Winter colony loss rates were best explained by November mean maximum temperature (Figure 2.2a) and February mean precipitation (Figure 2.2b). AIC scores for winter (1 October to 1 April) were consistently lower, indicating a better model fit, than the rest of the year (Figure 2.2).

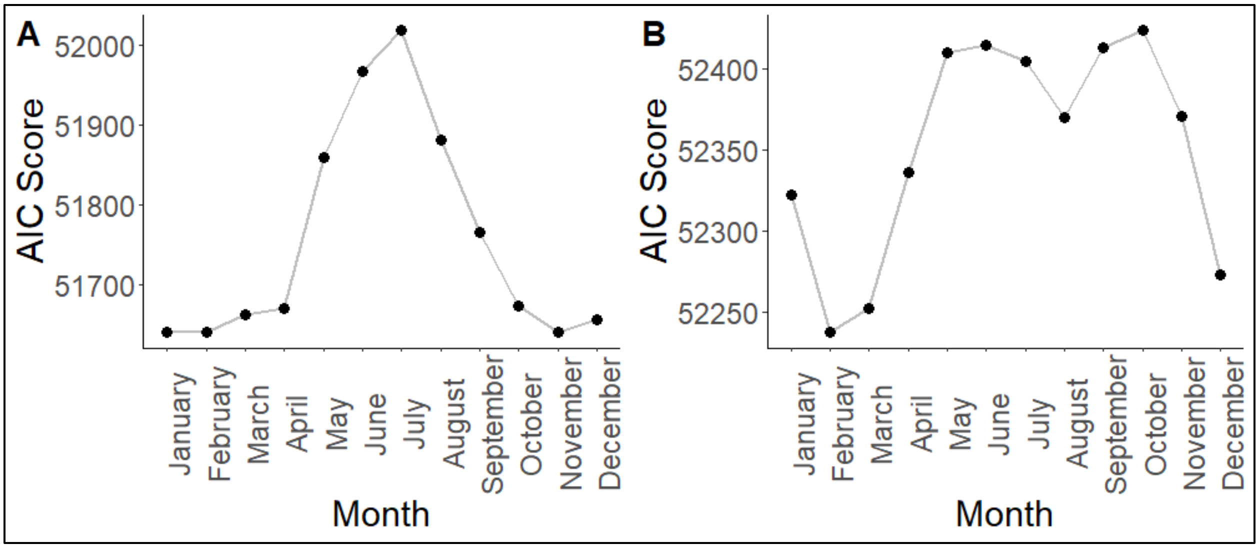

Figure 2.2. A comparison of Akaike's Information Criterion (AIC) scores between months for mean maximum temperature (A) and mean total precipitation (B). Lower scores indicate a better model fit and thus better support for explaining the variation in honey bee *(Apis mellifera*) winter colony loss rates.

2.3.2 Generalized Linear Regression (GLR)

The GLR model for winter colony loss rates between 2011-2020 poorly explained variation in loss rates, having an adjusted R-squared value of 0.135. The effect sizes for the four predictor variables were as follows: for every 1 m/s increase in mean wind speed, 2.27 ± 0.84 fewer colonies were lost; for every 1 mm increase in mean February precipitation, 0.054 ± 0.0041 fewer colonies were lost; for every degree Celsius warmer the mean maximum temperature was during November, 1.8 ± 0.13 fewer colonies were lost; and for every 1 m increase in mean elevation, 0.0031 ± .00145 fewer colonies were lost. All variables were highly significant (Table 2.4). In the case of all variables, increased values were associated with reduced colony losses. This model was non-stationary (Koenker statistic of <0.00000), meaning that the relationships between the predictors and response were not consistent across the study area.

Table 2.4. Coefficient values, standard error, z-statistic, and probability, for the generalized linear regression model used to explain variation in honey bee (*Apis mellifera*) winter colony loss rates. Asterisks indicate significant models at P <0.05.

Variable	Coefficient	Standard Error	Z-Statistic	Probability
Intercept	74.60	2.09	35.75	<0.000000*
Windspeed (m/s)	-2.27	0.43	-5.35	<0.000000*
February Precipitation (mm)	-0.054	0.0081	-6.69	<0.000000*
November Temperature (°C)	-1.80	0.068	-26.39	<0.000000*
Elevation (m)	-0.0031	0.00074	-4.28	<0.000000*

3.3 Geographically Weighted Regression (GWR)

The golden search identified 367 as the ideal number of neighbors for creating local regressions. The AIC score for the GWR model was 51,336.78 compared to 51,665.58 for the GLR model, an improvement of 328.80 points. The adjusted R-squared for the GWR model was 0.20, an improvement from the global GLR model of 0.065, or 6.5 percentage points Unlike the GLR model, the R-squared value was allowed to differ across zip codes, ranging from effectively zero to 0.30, with lowest values scattered throughout the U.S. and highest values clustered in Illinois (Figure 2.3). The effects of the four predictor variables (November mean maximum temperature, February mean precipitation, mean windspeed, and mean elevation) varied by zip code (Figure 2.4), compared to the GLR model where the effects were constant across the entire study area (Table 2.4).

Coefficient values, which show regional variation among model variables, were mapped for each predictor variable (Figure 2.4). Coefficient values for mean windspeed ranged from -9.96 to 8.53 with standard errors ranging from 1.39 to 5.14 (Figure 2.4a and 2.4b, respectively). Lower negative coefficient values, indicating that increased predictor values were associated with lower colony loss rates, were scattered throughout the U.S., although concentrations were found in northern Pennsylvania (PA), Indiana (IN), southern Tennessee (TN), northern Alabama (AL), and much of the west (for a labelled map of the U.S. see Figure 2.5). Higher positive values, indicating that increased predictor values were associated with higher colony loss rates, were concentrated in eastern PA/western New Jersey (NJ) and various parts of Appalachia such as West Virginia (WV) and western North Carolina (NC). Standard errors, which are a measure of coefficient uncertainty, were consistently highest in parts of the northeast, Virginia (VA), and Iowa (IA).

Coefficient values for February mean precipitation ranged from -0.82 to 1.06 with standard errors ranging from 0.018 to 0.49 (Figure 2.4c and 2.4d, respectively). Lower negative coefficient values were concentrated in parts of PA, NJ, VA, and Ohio (OH) while higher positive values were concentrated in IA. Standard errors were consistently high in the upper Midwest, especially northern OH, and along the coast of VA.

Coefficient values for November mean maximum temperature ranged from -10.23 to 9.79 with standard errors ranging from 0.39 to 5.98 (Figure 2.4e and 2.4f, respectively). Lower negative coefficient values were concentrated in parts of the Midwest, especially Indiana (IN) and western Kentucky (KY). Higher positive values were concentrated in northeastern PA, western New York (NY), Southern NJ, and parts of OH and eastern KY. Standard errors were consistently high in parts of IN, OH, and KY.

Coefficient values for mean elevation ranged from -0.12 to 0.15 with standard errors ranging from 0.0025 to 0.0672 (Figure 2.4g and 2.4h, respectively). Lower negative coefficient values were concentrated in OH and southern New England. Higher positive values were concentrated in eastern PA, southern NJ, and parts of the Midwest including northern Illinois (IL) and southern Wisconsin (WI). Standard errors were consistently high in parts of the Midwest such as IL.

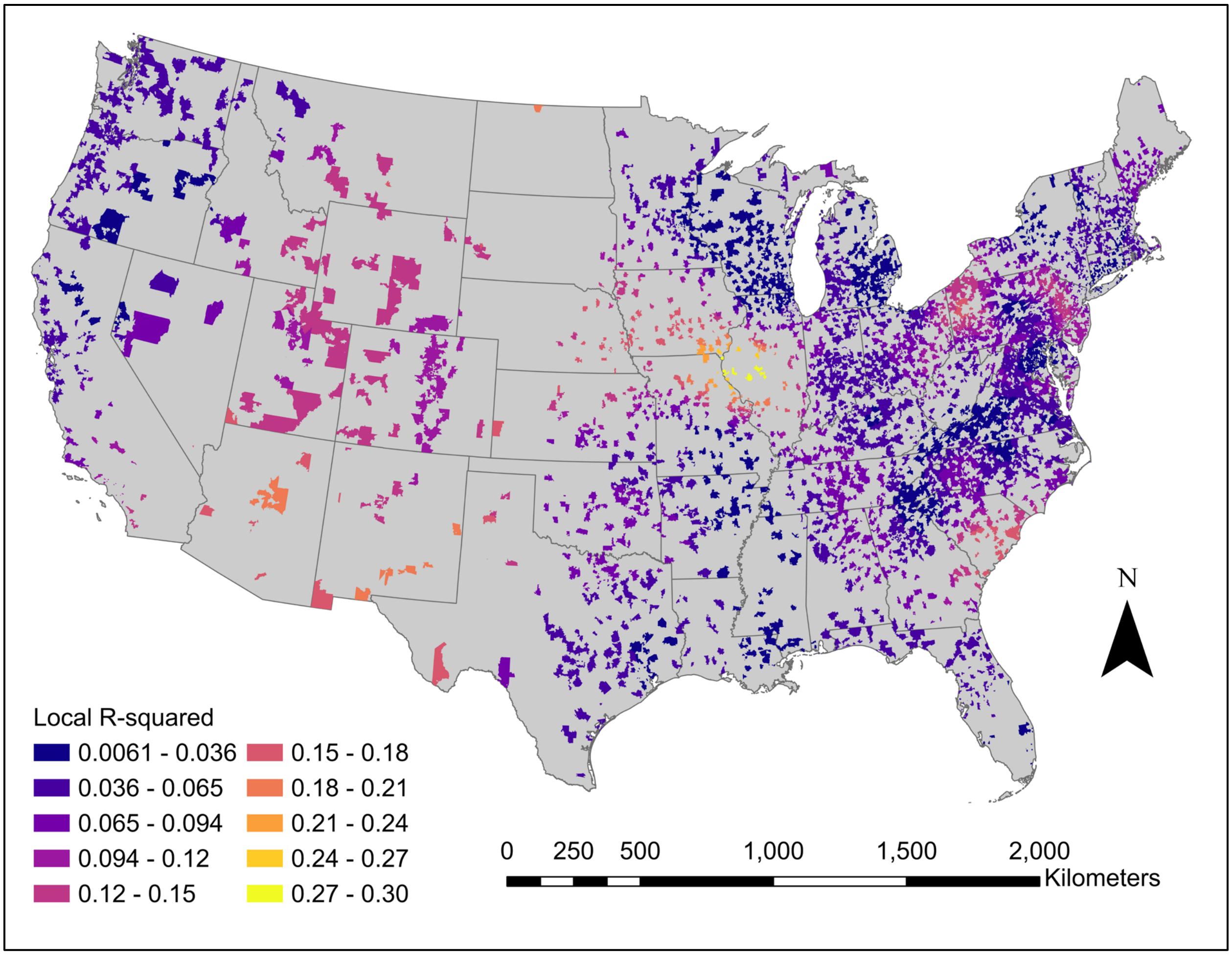

Figure 2.3. Local R-squared values for each zip code, calculated with a geographically weighted regression (GWR). Colors indicate the percent of variation in managed honey bee (*Apis mellifera*) colony loss explained by the four predictor variables included in the analysis: November mean maximum temperature, February mean precipitation, mean windspeed, and mean elevation.

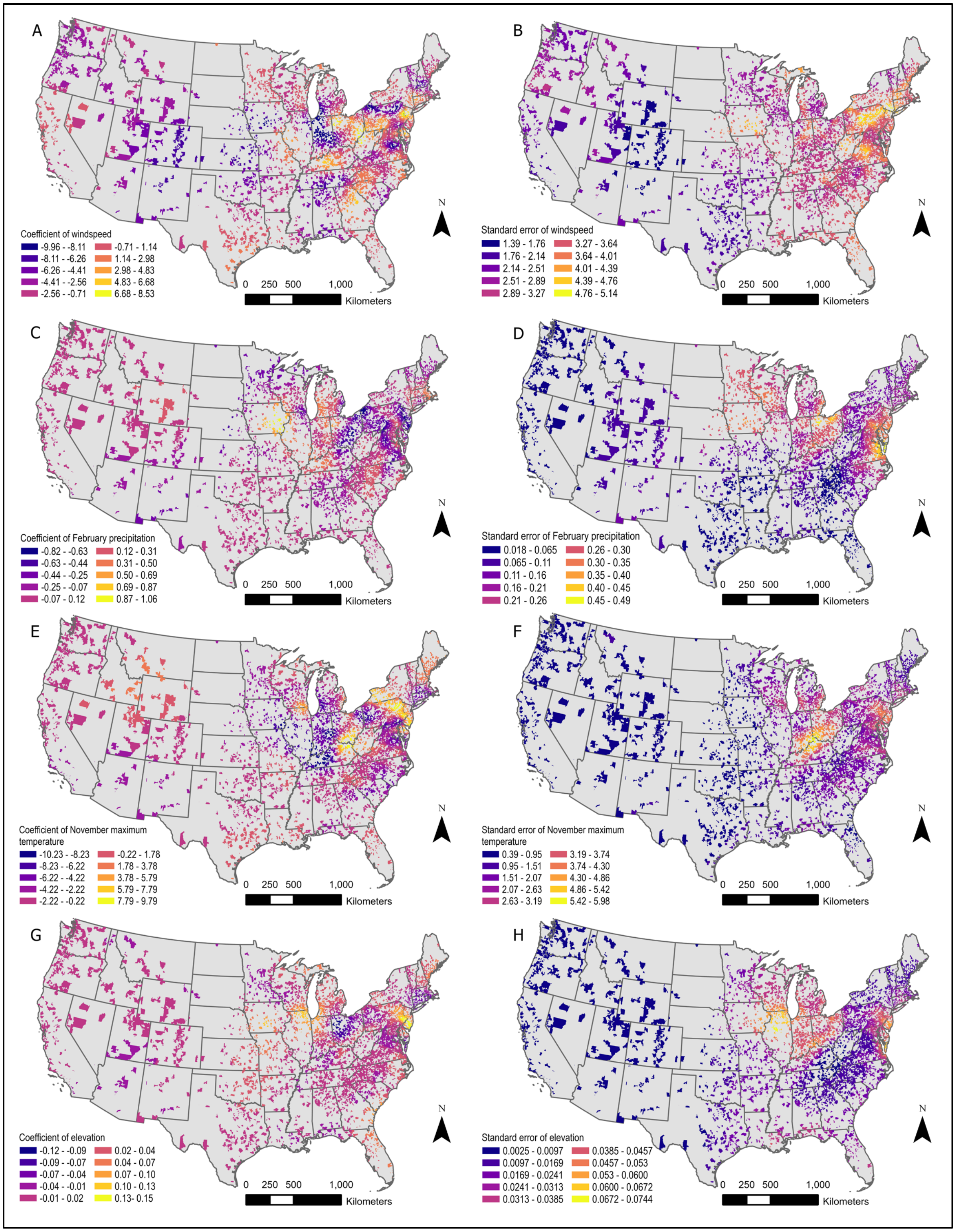

Figure 2.4. Coefficient and standard error values for the four predictor variables of honey bee (*Apis mellifera*) winter colony loss rates (November mean maximum temperature, February mean precipitation, mean windspeed, mean elevation) calculated at the zip code level with a geographically weighted regression (GWR).

2.4 Discussion

While honey bee colony loss can be linked to numerous factors working in conjunction, this study sought to determine the effects of weather on winter colony loss in the continental U.S. and to determine if a spatial approach (GWR) would yield better results, in terms of AIC score and R-squared value, than a traditional non-spatial (GLR) approach. This was the first honey bee study to utilize a GWR for analysis of colony loss and the first to analyze loss rates at a nationwide scale in the U.S.

2.4.1 Best Supported Months from Model Selection

The model with the most support for explaining the variation in colony loss rates, based on the lowest AIC score, included November mean maximum temperature and February mean precipitation. This indicates that temperature and precipitation during these months produce a better model fit than during others, when analyzed at the nationwide scale, meaning that variation in loss rates were most related to mean maximum temperature and mean precipitation during November and February, respectively. There is a trend seen in the AIC graphs (Figure 2.2) of winter months having consistently lower AIC scores than other months of the year. This suggests that temperature and precipitation during winter can better explain the observed winter loss rates than other months of the year. Although it differs by region of the country, this period tends to be associated with colder temperatures, lack of forage resources, and no brood development (Dadant and Sons 2015). These factors, in addition to other stressors that may be present, may cause a colony to be especially susceptible to loss at those times of the year. However, while these months were the best supported in the nationwide analysis, they would likely differ from place to place were this model to be built at a more local scale.

2.4.2 Regression Analysis

The GWR model outperformed the GLR equivalent in terms of both R-squared and AIC values, indicating that a spatial approach better explains the variation in observed loss rates across the U.S. These findings are consistent with GWR analysis in other studies such as Xu et al. (2019), which found that GWRs outperformed a non-spatial model when analyzing the effects of various factors on pollution levels across space. The findings of this study are not surprising when considering how geographically large and heterogeneous the continental U.S. is. A model that assumes equal effects of predictor variables across an entire country is unlikely to be as accurate given what we know from past studies (Brunsdon et al. 1996).

As expected, the GLR was non-stationary, meaning that the relationships were not constant across the study area, making the GWR necessary to expose more localized relationships. For example, without the GWR we would not be able to see that increased November mean maximum temperature has a strong association with higher loss rates in northeastern PA while having a strong association with lower loss rates in IN (Figure 2.4e). GWR models consider spatially local neighborhoods which allows for the calculation of local R-squared values and effect sizes at the zip code scale (Brunsdon et al. 1996). This resulted in a better fitting model than was obtained from the GLR, which only analyzed loss at the national scale.

R-squared values for each zip code did not exceed 0.30, with many having values close to zero. This indicates that for certain local neighborhoods, such as those in Michigan (MI), VA, and WI the predictor variables were not important for explaining the variation in the observed loss rates (Figure 2.3). These results can partially be attributed to the scale of analysis. Low R-squared values may be due to the fact that predictor values may simply not change much across certain local scales or that other drivers of colony loss at this scale were not included. These additional variables likely include parasites, pesticides, and other causes of colony loss that have been previously researched (Havard et al. 2019).

The November mean maximum temperature was by far the most important predictor variable of those analyzed at the national scale (GLR), having a much higher AIC weight than other predictors (Table 2.2). While windspeed, precipitation, and elevation were significant predictor variables (Table 2.4), they were not as well supported by the AIC table, having a weight of -9.74E-131 or lower compared to the value of 0.37 for November mean maximum temperature (Table 2.2). This indicates that the proportion of predictive power provided by these variables is very low. The results found for temperature, indicating that higher temperatures were associated with lower winter loss rates, corroborate somewhat with past studies which have shown that southern states, an area of higher mean temperatures, generally experience lower winter colony loss rates (Kulhanek et al. 2017).

The effects of all variables differed substantially across space in the GWR model, meaning that temperature was not necessarily the most important variable for predicting loss in all locations (Figure 2.4). These results suggest that while the R-squared values for most zip codes were low, the effects of the predictor variables on colony loss rates were substantial in some locations. For example, the effects of mean windspeed and November mean maximum temperature are quite high in parts of eastern PA (Figure 2.4a and 2.4e). Although the reasons behind the large and widely varying effects are not necessarily clear, it is evident that changes in predictor variable values within each zip code neighborhood have the potential to result in vastly higher or lower colony loss rates in some areas of the country. Possible reasons for the vastly different coefficient values observed between zip codes once again include the presence of local, undetermined factors.

The results of this study provide support to previous work on the effects of weather on colony loss. Temperature and precipitation, the most common variables analyzed previously, were found to have significant effects on colony loss. Both Switanek et al. (2017) and Beyer et al. (2018) found that warmer temperatures were linked to higher losses year-round and during winter, respectively, with Switanek et al. (2017) noting exceptions for two months: February and November, where warmer temperatures were linked to lower losses. The GLR model from this study supports Switanek et al. (2017), also finding that higher temperatures in November were linked to decreased losses, potentially as a result of less cold stress, although the GWR model produced a different result. These two studies reported conflicting results for precipitation, with Switanek et al. (2017) finding that increased precipitation year-round was linked to decreased loss in Austria, whereas, Beyer et al. (2018) found that increased summer precipitation was linked to increased loss in Luxembourg. The GWR model showed that the effect of all variables varied substantially across space, with some areas experiencing increased loss and some experiencing decreased loss as predictor values increased, possibly due to the regionality of landscape and weather patterns and the presence of other, undetermined, variables.

The results of this study show how the effects of weather vary substantially from region to region, possibly due to the heterogeneous nature of weather and geographic characteristics. Additionally, these results may explain the inconsistent findings regarding the effects of weather on loss in previous studies. This information is critical for honey bee colony loss research, as it suggests that effects found in some studies may not be translatable to other regions or potentially more local scales within their own study area.

2.4.3 Limitations and Future Work

While this study produced several key takeaways, there are areas of the analysis that could be improved upon in future studies. The BIP Colony Loss and Management Survey is one of the most robust, long running, and complete beekeeper surveys in the U.S. Even so, there are geographic issues associated with collecting and analyzing data at the zip code scale although zip code is one of the easiest ways to collect locational information from the general public (Grubesic 2008). For this study, the most pertinent issue with zip codes was the discrepancy in their sizes across the U.S. For example, the average zip code in Wyoming (WY) covers approximately 1,430 km^2, while the average zip code in NJ covers just 33 km^2 (Grubesic 2008). This adds a level of uncertainty as colonies within the loss dataset could be located anywhere within a zip code. This means that, for larger zip codes, the predictor variable means are less likely to represent the conditions colonies within those zip codes truly experienced when compared to smaller zip codes. However, the usage of mean conditions across the study period somewhat reduces this issue. Although these aggregations introduce slight uncertainty, zip codes were the smallest possible unit of analysis for this study and thus more accurately reflect the conditions experienced by colonies than other aggregation methods.

The spatial distribution of stationary survey respondents was uneven as there was greater zip code coverage in the northeastern quadrant of the continental U.S. and large gaps in some other parts of the country, notably in the west. This undoubtably has some effect on the model selection for the GLR as zip codes in the northeast are more numerous and thus had an outsized impact on the model. Additionally, certain neighborhoods in the GWR were likely affected by the presence of isolated zip codes. Ideally, colony loss data would have been available for every zip code in the

U.S., which would help with these issues, but this was unfortunately not realistic as some zip codes either lack stationary beekeepers, or those beekeepers have not participated in the BIP survey.

Another potential issue when working with GWR models is local multicollinearity, where variables that are not correlated in the global model become correlated at more local scales. The initial GLR model was built to avoid multicollinearity, but when the same model is run in different locations at a more localized scale, local multicollinearity may become an issue. This can result in estimates with higher degrees of uncertainty, which should be considered when interpreting results. This study did not seek to create a perfect GWR model, but rather sought to compare a GWR model using the same variables as the best supported GLR model. Future studies may refine the methods used to create a better local model.

Despite the limitations present, this analysis was able to confirm that a spatially oriented approach is better supported when compared to a non-spatial equivalent and confirmed that the effects of weather on honey bee colony loss differ across space. While this study established the potential usefulness of the GWR for this work, future work may seek to conduct an analysis that can better control for missing zip code data and local multicollinearity and include additional variables related to colony loss. Further work may also look at year-to-year variation in the effects of weather and the effect extreme weather events have on loss rates.

2.5 References

Akaike, H. 1987 “Factor Analysis And AIC.” Psychometrika 52 (3): 317-332.

Badger, J., and E. Jørgensen, H. 2011. “A high resolution global wind atlas - improving estimation of world wind resources.” In Energy Systems and Technologies for the coming Century: Proceedings (pp. 215-225). Danmarks Tekniske Universitet, Risø Nationallaboratoriet for Bæredygtig Energi. Denmark. Forskningscenter Risoe. Risoe-R No. 1776(EN)

Beyer, M., J. Junk, M. Eickermann, A. Clermont, F. Kraus, C. Georges, A. Reichart, and L. Hoffmann. 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* 118 (June): 52–60. https://doi.org/10.1016/j.rvsc.2018.01.012.

Bird, G., A. E. Wilson, G. R. Williams, and N. B. Hardy. 2021. “Parasites and Pesticides Act Antagonistically on Honey Bee Health.” *Journal of Applied Ecology* 58 (5): 997–1005. https://doi.org/10.1111/1365-2664.13811.

Bruckner, S., L. Straub, P. Neumann, and G. R. Williams. 2021. “Synergistic and Antagonistic Interactions Between Varroa Destructor Mites and Neonicotinoid Insecticides in Male Apis Mellifera Honey Bees.” *Frontiers in Ecology and Evolution* 9. https://doi.org// 10.3389/fevo.2021.756027.

Bruckner, S., N. Steinhauer, J. Engelsma, A. M. Fauvel, K. Kulhanek, E. Malcolm, A. Meredith, M. Milbrath, E. Niño, J. Rangel, K. Rennich, D. Reynolds, R. Sagili, J. Tsuruda, D. vanEngelsdorp, S. D. Aurell, M. Wilson, and G. Williams. 2020. “2019-2020 Honey Bee Colony Losses in the United States: Preliminary Results.” Unpublished work.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical Analysis* 28 (4): 281–98. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

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. https://doi.org/10.1371/journal.pone.0037235.

Calovi, M., C.M. Grozinger, D.A. Miller, and S. C. Goslee. 2021. "Summer weather conditions influence winter survival of honey bees (*Apis mellifera*) in the northeastern United States". *Scientific Reports* 11**:** 1553. https://doi.org/10.1038/s41598-021-81051-8

Dadant and Sons. "The Hive and the Honey Bee." (U.S.A., Dadant and Sons, 1949), 89.

Dainat, B., J. D. Evans, Y. P. Chen, L. Gauthier, and P. Neumann. 2012. "Predictive Markers of Honey Bee Colony Collapse." *PLOS ONE* 7 (2): e32151. https://doi.org/10.1371/ journal.pone.0032151.

Daly, C., and K. Bryant. 2013. "The PRISM climate and weather system—an introduction." Online. Northwest Alliance for Computational Science and Engineering, Oregon State University, Corvallis, USA. http://prism. oregonstate.edu/.

Deb, P., A. S. Kiem, and G. Willgoose. 2019. "Mechanisms Influencing Non-Stationarity in Rainfall-Runoff Relationships in Southeast Australia." *Journal of Hydrology* 571 (April): 749–64. https://doi.org/10.1016/j.jhydrol.2019.02.025.

El Agrebi, N., N. Steinhauer, S. Tosi, L. Leinartz, D. C. de Graaf, and C. Saegerman. 2021. "Risk and Protective Indicators of Beekeeping Management Practices." *Science of The Total Environment* 799 (December): 149381. https://doi.org/10.1016/j.scitotenv. 2021. 149381.

ESRI. 2022. “How Geographically Weighted Regression (GWR) Works—ArcGIS Pro | Documentation.” Accessed January 24, 2022. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/howgeographicallyweightedregression-works.htm.

ESRI. 2021. “United States ZIP Code Boundaries 2020 - Overview.” Accessed January 21, 2022. https://www.arcgis.com/home/item.html?id=46b350fa939149debfd9cc71566b43b4.

Grubesic, T. H. 2008. “Zip Codes and Spatial Analysis: Problems and Prospects.” *Socio-Economic Planning Sciences* 42 (2): 129–49. https://doi.org/10.1016/j.seps.2006.09.001.

Havard, T., M. Laurent, and M. Chauzat. 2019. “Impact of Stressors on Honey Bees (Apis Mellifera; Hymenoptera: Apidae): Some Guidance for Research Emerge from a Meta-Analysis.” *Diversity* 12 (1): 7. https://doi.org/10.3390/d12010007.

Jenkins, D. G., and P. F. Quintana-Ascencio. 2020. “A Solution to Minimum Sample Size for Regressions.” *PLOS ONE* 15 (2): e0229345. https://doi.org/10.1371/journal.pone. 0229345.

Kulhanek, K., N. Steinhauer, K. Rennich, D. M. Caron, R. R. Sagili, J. S. Pettis, J. D. Ellis, M. E. Wilson, J. T. Wilkes, D. R. Tarpy, R. Rose, K. Lee, J. Rangel, and D. vanEngelsdorp. 2017. “A National Survey of Managed Honey Bee 2015–2016 Annual Colony Losses in the USA.” *Journal of Apicultural Research* 56 (4): 328–40. https://doi.org/ 10.1080/00218839. 2017.1344496.

Lee, K., N. Steinhauer, K. Rennich, M. E. Wilson, D. R. Tarpy, D. M. Caron, R. Rose, K. Delaplane, K. Baylis, E. Lengerich, J. Pettis, J. Skinner, J. Wilkes, R. Sagili, and D, vanEngelsdorp. 2015. “A National Survey of Managed Honey Bee 2013–2014 Annual

Colony Losses in the USA." *Apidologie* 46 (3): 292–305. https://doi.org/10.1007/s13592-015-0356-z.

Osborne, P., G. Foody, and S. Suarez-Seoane. 2007. "Non-stationarity and local approaches to modelling the distributions of wildlife." *Diversity and Distributions* 13: 313-323. https://doi.org/10.1111/j.1472-4642.2007.00344.x

Perry, C. J., E. Søvik, M. R. Myerscough, and A. B. Barron. 2015. "Rapid Behavioral Maturation Accelerates Failure of Stressed Honey Bee Colonies." *Proceedings of the National Academy of Sciences* 112 (11): 3427–32. https://doi.org/10.1073/pnas.1422089112.

Peters, J. 2020. "Modeling Species Distribution and Habitat Suitability of American Ginseng (Panax Quinquefolius) in Virginia." Master's Thesis. James Madison University, Harrisonburg, Virginia.

Pickett, S. and M. Cadenasso. 1995. "Landscape Ecology: Spatial heterogeneity in Ecological Systems." *Science* 269: 331-334.

R Core Team. 2022. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team. 2022. "RStudio: Integrated Development for R." RStudio, Inc., Boston, MA. http://www.rstudio.com/

Smart, M., J. Pettis, N. Rice, Z. Browning, and M. Spivak. 2016. "Linking Measures of Colony and Individual Honey Bee Health to Survival among Apiaries Exposed to Varying Agricultural Land Use." *PLOS ONE* 11 (3): e0152685. https://doi.org/10.1371/journal.pone.0152685.

Southwick, E. E., and L. Southwick. 1992. “Estimating the Economic Value of Honey Bees (Hymenoptera: Apidae) as Agricultural Pollinators in the United States.” *Journal of Economic Entomology* 85 (3): 621–33. https://doi.org/10.1093/jee/85.3.621.

Spleen, A. M., E. J. Lengerich, K. Rennich, D. Caron, R. Rose, J. S. Pettis, M. Henson, et al. 2013. “A National Survey of Managed Honey Bee 2011–12 Winter Colony Losses in the United States: Results from the Bee Informed Partnership.” *Journal of Apicultural Research* 52 (2): 44–53. https://doi.org/10.3896/IBRA.1.52.2.07

Steinhauer, N., K. Kulhanek, K. Antúnez, H. Human, P. Chantawannakul, M. Chauzat, and D. vanEngelsdorp. 2018. “Drivers of Colony Losses.” *Current Opinion in Insect Science*, 26: 142–48. https://doi.org/10.1016/j.cois.2018.02.004.

Steinhauer, N., K. Rennich, M. E. Wilson, D. M. Caron, E. J. Lengerich, J. S. Pettis, R. Rose, J. Skinner, D. Tarpy, J. Wilkes, and D. vanEngelsdorp. 2014. “A National Survey of Managed Honey Bee 2012–2013 Annual Colony Losses in the USA: Results from the Bee Informed Partnership.” *Journal of Apicultural Research* 53 (1): 1–18. https://doi.org/10.3896/IBRA.1.53.1.01.

Switanek, M., K. Crailsheim, H. Truhetz, and R. Brodschneider. 2017. “Modelling Seasonal Effects of Temperature and Precipitation on Honey Bee Winter Mortality in a Temperate Climate.” *Science of The Total Environment* 579: 1581–87. https://doi.org/10.1016/j.scitotenv.2016.11.178.

USGS. 2007. “North America Elevation 1-Kilometer Resolution GRID - ScienceBase-Catalog.” Accessed January 21, 2022. https://www.sciencebase.gov/catalog/item/4fb5495ee 4b04 cb937751d6d.

vanEngelsdorp, D., D. Caron, J. Hayes, R. Underwood, M. Henson, K. Rennich, A. Spleen, M. Andree, R. Snyder, K. Lee, K. Roccasecca, M. Wilson, J. Wilkes, E. Lengerich, and J. Pettis. 2012. "A National Survey of Managed Honey Bee 2010–11 Winter Colony Losses in the USA: Results from the Bee Informed Partnership." *Journal of Apicultural Research* 51 (1): 115–24. https://doi.org/10.3896/IBRA.1.51.1.14.

vanEngelsdorp, D., and M. D. Meixner. 2010. "A Historical Review of Managed Honey Bee Populations in Europe and the United States and the Factors That May Affect Them." *Journal of Invertebrate Pathology* 103 (January): S80–95. https://doi.org/10.1016/j.jip.2009.06.011.

Xu, C., J. Zhao, and P. Liu. 2019. "A Geographically Weighted Regression Approach to Investigate the Effects of Traffic Conditions and Road Characteristics on Air Pollutant Emissions." *Journal of Cleaner Production* 239: 118084. https://doi.org/10.1016/j.jclepro.2019.118084.

Zee, R., A. Gray, L. Pisa, and T. de Rijk. 2015. "An Observational Study of Honey Bee Colony Winter Losses and Their Association with Varroa Destructor, Neonicotinoids and Other Risk Factors." *PLOS ONE* 10 (7): e0131611. https://doi.org/10.1371/journal.pone.0131611.

2.6 Appendix

Figure 2.5. Map of the continental United States with two letter state abbreviations.

Chapter 3

The Importance of Scale When Analyzing the Effects of Weather on Winter Honey Bee Colony Loss: A Look at The Modifiable Areal Unit Problem

3.1 Introduction

3.1.1 The Modifiable Areal Unit Problem (MAUP)

Data analysis is often conducted using aggregated datasets, where point data, representing a measurement such as pollution levels or rainfall, are averaged over an area of interest (e.g., county or state), typically represented as polygons in GIS. The aggregation of point data into arbitrary zones introduces the issue of the modifiable areal unit problem (MAUP), first described by Openshaw and Taylor (1979). MAUP arises when the results of statistical analyses differ based upon the analysis unit, or zone, used in the investigation, which in turn leads to differences in interpretation (Openshaw and Taylor 1979). There are two aspects of MAUP: the scale of analysis, or scale effect, and the unit definition, or zone effect (Openshaw and Taylor 1979; Horner and Murray 2002). The scale effect refers to the variation in results based on the number of units used in the analysis. Analyzing data with fewer, larger units reduces variation while smaller, more numerous units increases it (Openshaw and Taylor 1979; Dark and Bram 2007). The zone effect refers to changes in results caused by using the same number of units in different ways within the study area (Openshaw and Taylor 1979; Dark and Bram 2007). A common example of this effect is gerrymandering, where political districts may be drawn in different shapes to support one political party over another (Stehle 2022). Both aspects of MAUP have been found to alter the results of statistical analyses in a number of fields including species conservation and landscape ecology (Openshaw and Taylor 1979; Fotheringham and Wong 1991: Jelinksi and Wu 1996; Moat et al. 2018). For example, Moat et al. (2018) described how grid size affected the area of occupancy for species on the International Union for Conservation of Nature (IUCN) Red List estimation, with different cell sizes and shapes producing different area of occupancy estimates, which can affect species extinction risk assessment. Additionally, Jelinski and Wu (1996) found that spatial autocorrelation of Normalized Difference Vegetation Indices (NDVI) varied based on aggregation and zoning schemes, leading to the authors concluding that autocorrelation changes with scale

One area that MAUP has not been considered in is the field of honey bee (*Apis mellifera*) colony loss research. Investigations into the causes of honey bee colony loss have been numerous, given the high colony loss rates that have been observed in recent decades (Bruckner et al. 2020) and the importance of honey bees to the economy (Calderone 2012). Honey bee colony loss studies are typically conducted in one of two ways. Most commonly, losses are analyzed at the colony level, often with linear regressions (Zee et al. 2015; Switanek et al. 2017). The second method aggregates loss rates to ecologically arbitrary units, such as political boundaries (e.g., county or state). For example, Becsi et al. 2021 analyzed the effects of weather on loss using loss data aggregated to the 94 political districts of Austria. Additionally, other studies aggregate loss rates to the state level in the U.S., although do not analyze the effects of any variable on these loss rates (Steinhauer et al. 2014; Kulhanek et al. 2017). In the case of the latter method, results will likely differ based on the type of aggregation method used, as has been found in studies from other fields (Jelinski and Wu 1996). Additionally, Fotheringham and Wong (1991) stated that MAUP can lead to unpredictable estimates of effect in multivariate analyses. This could lead to misinterpretations of the effects that analyzed variables such as temperature, pesticides, or any number of others, have on colony loss rates, and potentially result in inaccurate management suggestions. Thus, it is critical to understand how honey bee colony loss analysis results change based on aggregation method.

3.1.2 Previous Honey Bee Colony Loss Research

Many past honey bee studies have researched the effects of parasites, such as *Varroa destructor*, on colony loss (Zee et al. 2015; Steinhauer et al. 2018), as well as management practices (Steinhauer et al. 2018; El Agrebi et al. 2021) and nutrition (Perry et al. 2015; Steinhauer et al. 2018).However, while these variables have been well studied, relatively few studies have included the effects of weather in their analyses (Havard et al. 2019), despite weather being considered a key driver in ecological systems (Daly and Bryant 2013). To date, there are several examples of studies that have found variables such as temperature and precipitation to have an effect on colony loss rates. Switanek et al. (2017) conducted a study in Austria and found that higher temperatures during most months of the year were correlated with greater colony losses while greater amounts of precipitation during most months were correlated with lower losses. In contrast, Beyer et al. (2018) conducted a study in Luxembourg and found warmer and wetter conditions during winter months to be correlated with increased colony losses. Finally, Calovi et al. (2021) conducted a study in Pennsylvania and found growing degree days, maximum temperature of the warmest month of the year, precipitation during the warmest quarter of the year, and precipitation during the wettest quarter of the year best predicted winter colony loss. These studies provide evidence that weather does indeed affect colony loss rates, although the effects that conditions such as temperature and precipitation have on colony loss are still somewhat unclear.

Past honey bee studies that have aggregated loss data have tended to use human-centric aggregation methods such as political districts (Becsi et al. 2021), which are somewhat arbitrary and not ecologically meaningful. This can be useful for displaying colony loss data in a way that makes sense to viewers but doesn't account for the fact that ecologically important variables such as mean temperature, precipitation, and elevation, among others, may vary significantly across these political regions. One potential method for aggregating loss data– ecoregions – may be better suited for this type of analysis but have not been explored in honey bee research to this point. Ecoregions are defined as areas where ecosystems are generally similar within a landscape of heterogeneity. The ecoregions of the US were identified at four different scales using the patterns and composition of phenomena, both biotic and abiotic, that reflect differences in ecosystems, including factors such as geology, soils, vegetation, climate, landforms, land use, hydrology, and wildlife (Omernik 1987, Omernik 1995). There are four levels of ecoregions with level one providing the coarsest detail and level four providing the finest detail. Ecoregions have been used as an analysis unit in past ecological studies, ranging from aggregation of wildfires (Kasischke et al. 2002) to land use analysis (de Koning et al. 1998). Analyzing honey bee colony loss rates within ecoregions may provide more meaningful insights into colony loss and the factors driving them than analysis done using political regions.

This study sought to determine how MAUP affects honey bee colony loss analysis results by exploring the aspect of scale using six different types of aggregation: state, county, and zip code, and ecoregions level one, two and three. Ecoregion level four was not considered because the majority of the polygons lacked a sufficient colony sample size to run regression analysis. To test the concept of MAUP on honey bee colony loss data, mean temperature and mean precipitation were chosen as predictor variables to use in regression analysis due to their relevance in prior investigations (Switanek et al. 2017; Beyer et al. 2018), and the need for more research into the effect of weather on honey bee colony loss (Havard et al. 2019). It was hypothesized that both observed colony loss rates and the effects of weather on these loss rates would differ substantially between aggregation methods given what we know about MAUP.

3.2 Methods

3.2.1 Data Management

Winter colony loss data between the winters of 2011-2012 and 2019-2020 were provided by the Bee Informed Partnership's (BIP) annual colony loss management survey (vanEngelsdorp et al. 2012; Steinhauer et al. 2014; Bruckner et al. 2020). The BIP survey defines winter as the period between October 1st and April 1st. The survey is a citizen science initiative that obtains information about colony health and management from beekeepers across the U.S. The survey also gathers information about the number of colonies managed for a given year prior to the start of winter, how many were lost during the winter months, and the zip code and state in which they kept their colonies. Zip code information has been used in this study to geolocate beekeepers by geocoding five-digit zip codes to the centroid of each zip code polygon. Data were geocoded in ArcGIS Pro 2.9 using the *Geocode Addresses* tool. This enabled the loss data to be linked to zip code, county, and state. A total of 30,535 unique beekeepers in 9,718 zip codes had colony loss data reported over the span of this dataset.

Monthly average raster datasets for both mean temperature (°C) and precipitation (mm) were acquired from Parameter-elevation Regressions on Independent Slopes Model (PRISM) for the entire continental U.S. at 1 km resolution (Daly and Bryant 2013). Monthly averages were aggregated for the entire nine-year study period using the *Raster Calculator* tool in ArcGIS Pro. This resulted in the creation of two new rasters: one representing average temperature and the other representing average precipitation for nine years across the U.S. Zip code polygons were acquired from ESRI (ESRI 2021), state and county polygons were acquired from the U.S. Census Bureau (U.S. Census Bureau 2018), and ecoregion polygons for levels one, two, and three were acquired from the Environmental Protection Agency (U.S. EPA 2015).

Colony loss data and weather data were spatially joined, and the *Zonal Statistics* tool was used to find the mean weather conditions within each analysis unit (zip codes, ecoregions, etc.). A point was then generated for the centroid of each polygon using the *Feature to Point* tool. This centroid was then used to link weather data to the respective analysis units using the *Extract Multivalues to Point* tool followed by a spatial join to link the point file to the polygon files. The winter colony loss rates were initially recorded at the zip code level. To analyze the effects of temperature and precipitation on loss with the other analysis units at different scales (county, state, etc.), the loss data also had to be aggregated to each unit. This was done using two separate methods. Firstly, the *sumif* feature in excel was used to sum the number of colonies managed and the number of colonies lost within each zip code over the entire nine years the data covered in order to calculate a total loss rate per zip code over this period. This feature worked by summing the managed colonies and lost colonies columns based on the zip code ID, resulting in one summed total for number of colonies managed and lost for each zip code. The colony loss rate for this entire period was then recalculated for each zip code. County and state loss rates were aggregated using the *sumif* feature in the same manner using the county and state ID columns. These excel files were then joined to the corresponding polygon file by either zip code, county, or state ID to link loss rates for each analysis unit to the weather data.

Ecoregion aggregation was done using the *Select by Location* tool in ArcGIS Pro to extract zip codes that had their centers within an ecoregion. Once isolated, the total number of colonies owned and lost within that ecoregion were calculated. The loss rate was then entered in a new column in the ecoregion polygon table. This process was repeated for each ecoregion.

3.2.2 Analysis

Once all polygon files were joined, each of the six resulting files (zip code, county, state, level one, two, and three ecoregions) were exported to excel and then loaded into Program R (R Core Team 2022) using RStudio (RStudio Team 2022). A series of simple and multiple linear regressions were then run for each of the six aggregation methods with the colony loss rate as the response variable. Each analysis unit was analyzed using three regressions. Two simple regressions were run with one of mean temperature or mean precipitation as the predictor variable in order to determine R-squared values for each variable, and one multiple regression was run with both mean temperature and mean precipitation as predictor variables in order to get more accurate predictor coefficient values and a combined R-squared value.

3.3 Results

Winter colony loss rates varied by aggregation level (Table 1). Losses ranged from 0% to 100% at both the zip code and county levels and 20% to 70% at the state level (Figure 1a-c). Level one ecoregion rates ranged from 13% (Southern Semi-Arid Highlands) to 46% (Northern Forests) (Figure 1d); level two rates ranged from 13% (Southern Semi-Arid Highlands) to 56% (Mixed Wood Shield) (Figure 1e); and 13% (Southern Semi-Arid Highlands) to 71% (Nebraska Sand Hills) for level three ecoregions (Figure 1f).

R-squared values from the simple regressions indicated that temperature explained much more variation in loss than precipitation at all levels of analysis. For example, for level one ecoregions, temperature explained 56% of the variation in loss while precipitation explained just

5% (Table 2). The combined R-squared values from the multiple regressions ranged from 8% at the zip code level to 73% for level two ecoregions (Table 3). Coefficient values from the multiple regressions for mean temperature and mean precipitation varied substantially. The coefficient value for mean temperature ranged from -2.09 at the zip code level to -1.37 for level two ecoregions (Table 3) which were nearly identical to the simple regressions (Table 2). The effect of mean temperature on winter colony loss was significant at all levels of analysis. The coefficient for mean precipitation ranged from -0.00054 at the zip code level to 0.0061 for level two ecoregions within the multiple linear regressions (Table 3). Within the multiple regressions, the effect of mean precipitation on winter loss was only significant for level two ecoregions (Table 3). The standard errors of the coefficients for both variables and both model types increased in size at broader levels of analysis and smaller degrees of freedom (Tables 2,3).

Table 3.1. Winter honey bee colony loss rates per 100 colonies within each ecoregion of the continental United States (U.S.) from the winter of 2011-2012 to the winter of 2019-2020. Columns represent progressively finer ecoregion scales, with ecoregions nested within the larger ecoregion of the previous column. Level four ecoregions are excluded.

Level 1 ecoregions	Percent lost	Level 2 ecoregions	Percent lost	Level 3 ecoregions	Percent lost
NORTHERN FORESTS	46	MIXED WOOD SHIELD	56	NORTHERN LAKES AND FORESTS	56
				NORTHERN MINNESOTA WETLANDS	43
		ATLANTIC HIGHLANDS	43	NORTHERN APPALACHIAN AND ATLANTIC MARATIME HIGHLANDS	42
				NORTH CENTRAL APPALACHIANS	49
NORTHWESTERN FORESTED MOUNTAINS	41	WESTERN CORDILLERA	41	CANADIAN ROCKIES	38
				CASCADES	38
				EASTERN CASCADES SLOPES AND FOOTHILLS	45
				MIDDLE ROCKIES	32
				KLAMATH MOUNTAINS	39
				SIERRA NEVADA	40
				WASATCH AND UINTA MOUNTAINS	46
				SOUTHERN ROCKIES	40
				IDAHO BATHOLITH	40
				COLUMBIA MOUNTAINS/NORTHERN ROCKIES	52
				NORTH CASCADES	31
				BLUE MOUNTAINS	32
MARINE WEST COAST FOREST	41	MARINE WEST COAST FOREST	41	STRAIT OF GEORGIA/PUGET LOWLAND	45
				COAST RANGE	36
				WILLAMETTE VALLEY	39
EASTERN TEMPERATE FORESTS	39	MIXED WOOD PLAINS	46	EASTERN GREAT LAKES LOWLANDS	41
				ERIE DRIFT PLAIN	46
				NORTHERN ALLEGHENY PLATEAU	37
				NORTH CENTRAL HARDWOOD FORESTS	61
				DRIFTLESS AREA	55
				SOUTHERN MICHIGAN/NORTHERN INDIANA DRIFT PLAINS	53
				NORTHEASTERN COASTAL ZONE	41

				ACADIAN PLAINS AND HILLS	33
		CENTRAL USA PLAINS	45	CENTRAL CORN BELT PLAINS	48
				HURON/ERIE LAKE PLAINS	50
				EASTERN CORN BELT PLAINS	40
				SOUTHEASTERN WISCONSIN TILL PLAINS	53
		SOUTHEASTERN USA PLAINS	35	NORTHERN PIEDMONT	42
				INTERIOR RIVER VALLEYS AND HILLS	31
				INTERIOR PLATEAU	31
				PIEDMONT	35
				SOUTHEASTERN PLAINS	31
				MISSISSIPPI VALLEY LOESS PLAINS	24
				SOUTH CENTRAL PLAINS	20
				EAST CENTRAL TEXAS PLAINS	18
		OZARK/OUACHITA-APPALACHIAN FORESTS	38	RIDGE AND VALLEY	43
				CENTRAL APPALACHIANS	41
				WESTERN ALLEGHENY PLATEAU	35
				BLUE RIDGE	40
				OZARK HIGHLANDS	26
				BOSTON MOUNTAINS	23
				ARKANSAS VALLEY	23
				OUACHITA MOUNTAINS	20
				SOUTHWESTERN APPALACHIANS	27
		MISSISSIPPI ALLUVIAL AND SOUTHEAST USA COASTAL PLAINS	30	MIDDLE ATLANTIC COASTAL PLAIN	34
				MISSISSIPPI ALLUVIAL PLAIN	26
				SOUTHERN COASTAL PLAIN	23
				ATLANTIC COASTAL PINE BARRENS	32
GREAT PLAINS	35	TEMPERATE PRAIRIES	39	ASPEN PARKLAND/NORTHERN GLACIATED PLAINS	51
				LAKE MANITOBA AND LAKE AGASSIZ PLAIN	69
				WESTERN CORN BELT PLAINS	44
				CENTRAL IRREGULAR PLAINS	28

		WEST-CENTRAL SEMIARID PRAIRIES	36	NORTHWESTERN GLACIATED PLAINS	32
				NORTHWESTERN GREAT PLAINS	36
				NEBRASKA SAND HILLS	71
		SOUTH CENTRAL SEMIARID PRAIRIES	33	HIGH PLAINS	42
				CENTRAL GREAT PLAINS	32
				SOUTHWESTERN TABLELANDS	32
				FLINT HILLS	31
				CROSS TIMBERS	25
				EDWARDS PLATEAU	18
				TEXAS BLACKLAND PRAIRIES	20
		TEXAS-LOUISIANA COASTAL PLAIN	30	WESTERN GULF COASTAL PLAIN	30
		TAMAULIPAS-TEXAS SEMIARID PLAIN	25	SOUTHERN TEXAS PLAINS/INTERIOR PLAINS AND HILLS WITH XEROPHYTIC SHRUB AND OAK FOREST	25
NORTH AMERICAN DESERTS	40	COLD DESERTS	43	COLUMBIA PLATEAU	39
				NORTHERN BASIN AND RANGE	37
				WYOMIN BASIN	23
				CENTRAL BASIN AND RANGE	50
				COLORADO PLATEAUS	39
				ARIZONA/NEW MEXICO PLATEAU	43
				SNAKE RIVER PLAIN	38
		WARM DESERTS	22	MOJAVE BASIN AND RANGE	28
				SONORAN DESERT	22
				CHIHUAHUAN DESERT	16
MEDITERRANEAN CALIFORNIA	36	MEDITERRANEAN CALIFORNIA	36	CALIFORNIA COASTAL SAGE, CHAPARRAL, AND OAK WOODLANDS	35
				CENTRAL CALIFORNIA VALLEY	44
				SOUTHERN AND BAJA CALIFORNIA PINE-OAK MOUNTAINS	31
SOUTHERN SEMI-ARID HIGHLANDS	13	WESTERN SIERRA MADRE PIEDMONT	13	MADREAN ARCHIPELAGO	13
TEMPERATE SIERRAS	30	UPPER GILA MOUNTAINS	31	ARIZONA/NEW MEXICO MOUNTAINS	31
TROPICAL WET FORESTS	21	EVERGLADES	21	SOUTHERN FLORIDA COASTAL PLAIN	21

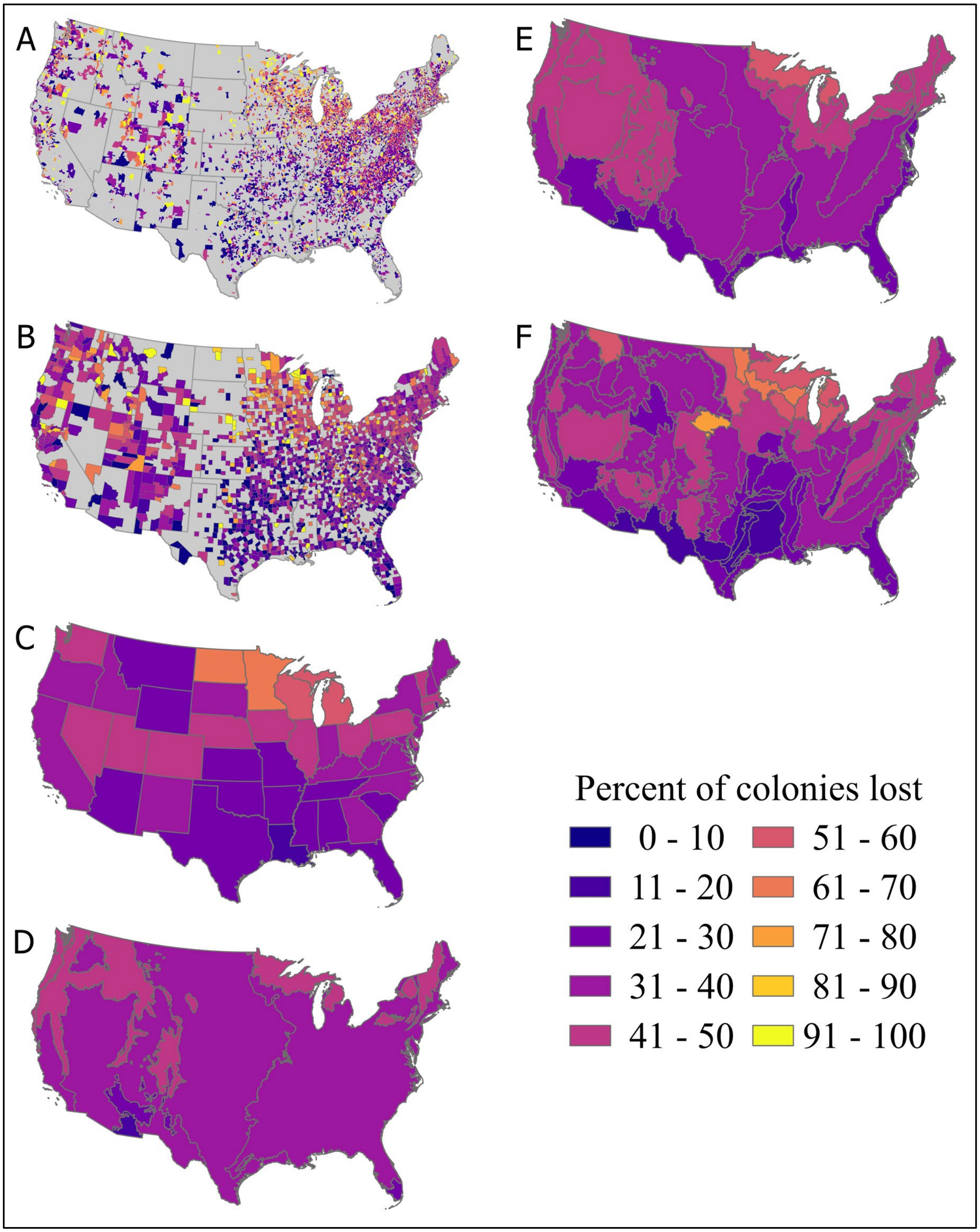

Figure 3.1. Average honey bee colony loss rates as a percentage of colonies owned between the winters of 2011-2012 and 2019-2020 across six different aggregation methods: zip codes (A), counties (B), states (C), level one ecoregions (D), level two ecoregions (E), and level three ecoregions (F).

Table 3.2. Model results from all single-variable models, including coefficient values, standard error (SE), p-value (p), R-squared (R^2), and degrees of freedom (df). Significant p-values are bolded ($p<0.05$).

Model	Coefficient	SE	p	R^2	df
Level 1 ecoregion mean temperature	-1.11	0.26	**0.0038**	0.72	7
Level 1 ecoregion mean precipitation	-0.00041	0.0053	0.94	0.00083	7
Level 2 ecoregion mean temperature	-1.24	0.19	**<0.00001**	0.73	16
Level 2 ecoregion mean precipitation	0.00089	0.0047	0.85	0.0022	16
Level 3 ecoregion mean temperature	-1.46	0.18	**<0.00001**	0.47	74
Level 3 ecoregion mean precipitation	-0.0028	0.0026	0.28	0.016	74
State mean temperature	-1.57	0.25	**<0.00001**	0.47	47
State mean precipitation	-0.0084	0.0036	**0.025**	0.1	47
County mean temperature	-1.88	0.095	**<0.00001**	0.21	1505
County mean precipitation	-0.009	0.0013	**<0.00001**	0.033	1505
Zip code mean temperature	-2.06	0.074	**<0.00001**	0.12	5804
Zip code mean precipitation	-0.01	0.00095	**<0.00001**	0.02	5804

Table 3.3. Model results from all multi-variable models for each analysis unit. Each model included mean temperature and mean precipitation as predictor variables. Coefficient, standard errors, and p-values are given for mean temperature and mean precipitation (denoted as MTC, MTSE, and MTP for temperature and MPC, MPSE, and MPP for precipitation) whereas R-squared (R^2) and degrees of freedom (df) are given for the entire model. Significant p-values are bolded ($p<0.05$).

Model	MTC	MTSE	MTP	MPC	MPSE	MPP	R^2	df
Level 1 ecoregion	-1.42	0.42	**0.01**	0.0055	0.0045	0.26	0.64	7
Level 2 ecoregion	-1.37	0.21	**<0.00001**	0.0061	0.0027	**0.04**	0.73	17
Level 3 ecoregion	-1.52	0.2	**<0.00001**	-0.00021	0.0022	0.92	0.43	81
State	-1.55	0.28	**<0.00001**	-0.00052	0.0032	0.87	0.47	46
County	-1.95	0.12	**<0.00001**	0.0022	0.0014	0.12	0.15	1762
Zip code	-2.09	0.076	**<0.00001**	-0.00054	0.00093	0.56	0.08	9715

3.4 Discussion

The goal of this study was firstly to showcase how MAUP can affect honey bee colony loss analysis results, using temperature and precipitation as predictors, and secondly to quantify loss rates in the U.S. at the ecoregion level. This is the first honey bee study to analyze how different colony loss aggregation methods can produce different model results, both in terms of the coefficient values and R-squared values and also the first study to show how winter colony loss rates differ between ecoregions within the continental U.S.

3.4.1 Loss Rates and Model Results

The differences in loss rate ranges between level one, two, and three ecoregions and the other units of analysis suggests, unsurprisingly, that larger aggregations have a moderating effect on the loss rates. The larger analysis units, such as ecoregions and states, had a smaller range of loss rates than smaller units, such as zip codes. As described by Jelinski and Wu (1996), larger aggregations tend to "smooth" data, reducing the effect of high or low values that would be evident at smaller aggregation levels. Essentially, the mean of the data remains the same, but variance decreases. This phenomenon has been widely recognized in previous works (Fotheringham and Wong 1991; Jelinski and Wu 1996; Dark and Bram 2007).

The R-squared value for the models differed between the units of analysis. For the multiple regressions, level two ecoregions had the highest value at 0.73 and zip codes had the lowest value at 0.08 (Table 3) Smaller units such as zip codes and counties produced lower R-squared values, while larger units, such as the level one and two ecoregions, produced larger R-squared values. The differences in R-squared values may be explained by the scale of analysis. Larger units of analysis have fewer and coarser datapoints which hide the large amounts of variation present at smaller units of analysis (Jelinski and Wu 1996). This reduction in variance increases the R-squared value for larger units of analysis. These results suggest that when looking at factors behind winter colony loss, weather, specifically temperature, is a good predictor when analyzed at broad scales. However, the effect of weather becomes less certain at smaller scales. This has implications for colony management as it suggests that other factors, potentially parasites, management practices, or nutrition could explain loss better than weather at local scales, although weather, and specifically, temperature, explain most of the variation in loss rates at a broad scale.

In contrast to the R-squared values, the estimates of effect for temperature and precipitation tended to be larger with smaller standard errors when analyzed with smaller units. This is likely because the smaller units of analysis have many more samples than larger ones (range of 10 to 9718 datapoints between level 1 ecoregions and zip codes, respectively), providing a more precise estimate as a result. The coefficient values indicate that an increase in mean temperature is correlated to a decrease in winter colony loss rate. This is in contrast to the findings of Switanek et al. (2017) in Austria and Beyer et al. (2018) in Luxembourg, which both found that warmer temperatures for most months were linked to higher colony losses. The coefficients for mean precipitation within the multiple regression models were both positive and negative, depending on the unit of analysis, and significant only when analyzed for level two ecoregions (Table 3). This suggests that annual precipitation levels do not play a significant role in colony loss rates across the continental U.S.

These results showcase how one aspect of MAUP, specifically the scale effect, can affect the takeaways of a colony loss study. For example, an analysis run with ecoregions or states as the analysis unit would find that temperature explains most of the variation in colony loss rate but would produce an estimate of effect with a large standard error due to the lower sample size, resulting in more uncertainty about the true effect of temperature. A study only conducted at the zip code level would find that temperature does not explain much of the observed variation in loss rate but would get a more precise estimate of effect. This tradeoff should be considered when analyzing aggregated loss data. For the purposes of estimating effect sizes, analyzing data with smaller units of aggregation seems preferable as it will result in a more precise coefficient. When possible, analyzing data at different scales may be beneficial for gaining an understanding of how scale changes the results of the analysis.

3.5 References

Becsi, B., H. Formayer, and R. Brodschneider. 2021. "A Biophysical Approach to Assess Weather Impacts on Honey Bee Colony Winter Mortality." *Royal Society Open Science* 8 (9): 210618. https://doi.org/10.1098/rsos.210618.

Beyer, M., J. Junk, M. Eickermann, A. Clermont, F. Kraus, C. Georges, A. Reichart, and L. Hoffmann. 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* 118 (June): 52–60. https://doi.org/10.1016/j.rvsc.2018.01.012.

Bruckner, S., N. Steinhauer, J. Engelsma, A. M. Fauvel, K. Kulhanek, E. Malcolm, A. Meredith, M. Milbrath, E. Niño, J. Rangel, K. Rennich, D. Reynolds, R. Sagili, J. Tsuruda, D. vanEngelsdorp, S. D. Aurell, M. Wilson, and G. Williams. 2020. "2019-2020 Honey Bee Colony Losses in the United States: Preliminary Results." Unpublished work.

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. https://doi.org/10.1371/journal.pone.0037235.

Calovi, M., C.M. Grozinger, D.A. Miller, and S. C. Goslee. 2021. "Summer weather conditions influence winter survival of honey bees (*Apis mellifera*) in the northeastern United States." *Scientific Reports* 11**:** 1553. https://doi.org/10.1038/s41598-021-81051-8

Daly, C., and K. Bryant. 2013. "The PRISM climate and weather system—an introduction." Online. Northwest Alliance for Computational Science and Engineering, Oregon State University, Corvallis, USA. http://prism. oregonstate.edu/.

Dark, S. J., and D. Bram. 2007. “The Modifiable Areal Unit Problem (MAUP) in Physical Geography.” *Progress in Physical Geography: Earth and Environment* 31 (5): 471–79. https://doi.org/10.1177/0309133307083294.

de Koning, G.H.J, A Veldkamp, and L.O Fresco. 1998. “Land Use in Ecuador: A Statistical Analysis at Different Aggregation Levels.” *Agriculture, Ecosystems & Environment* 70 (2–3): 231–47. https://doi.org/10.1016/S0167-8809(98)00151-0.

El Agrebi, N., N. Steinhauer, S. Tosi, L. Leinartz, D. C. de Graaf, and C. Saegerman. 2021. “Risk and Protective Indicators of Beekeeping Management Practices.” *Science of The Total Environment* 799: 149381. https://doi.org/10.1016/j.scitotenv. 2021. 149381.

ESRI. 2021. “United States ZIP Code Boundaries 2020 - Overview.” Accessed January 21, 2022. https://www.arcgis.com/home/item.html?id=46b350fa939149debfd9cc71566b43b4.

Fotheringham, A.S. and Wong, D.W.S. 1991. The Modifiable Areal Unit Problem in Multivariate Statistical Analysis.” *Environment and Planning* 23(7): 1025-1044.

Havard, T., M. Laurent, and M. Chauzat. 2019. “Impact of Stressors on Honey Bees (Apis Mellifera; Hymenoptera: Apidae): Some Guidance for Research Emerge from a Meta-Analysis.” *Diversity* 12 (1): 7. https://doi.org/10.3390/d12010007.

Horner, Mark W., and Alan T. Murray. 2002. “Excess Commuting and the Modifiable Areal Unit Problem.” *Urban Studies* 39 (1): 131–39. https://doi.org/10.1080/00420980220 099113.

Jelinski, E.D. and Wu, J. 1996. “The modifiable area unit problem and implications for landscape ecology.” *Landscape Ecology* 11(3):129-140.

Kasischke, Eric S., David Williams, and Donald Barry. 2002. "Analysis of the Patterns of Large Fires in the Boreal Forest Region of Alaska." *International Journal of Wildland Fire* 11 (2): 131. https://doi.org/10.1071/WF02023.

Kulhanek, K., N. Steinhauer, K. Rennich, D. M. Caron, R. R. Sagili, J. S. Pettis, J. D. Ellis, M. E. Wilson, J. T. Wilkes, D. R. Tarpy, R. Rose, K. Lee, J. Rangel, and D. vanEngelsdorp. 2017. "A National Survey of Managed Honey Bee 2015–2016 Annual Colony Losses in the USA." *Journal of Apicultural Research* 56 (4): 328–40. https://doi.org/ 10.1080/00218839. 2017.1344496.

Moat, J., S. P. Bachman, R. Field, and D. S. Boyd. 2018. "Refining Area of Occupancy to Address the Modifiable Areal Unit Problem in Ecology and Conservation: Area of Occupancy." *Conservation Biology* 32 (6): 1278–89. https://doi.org/10.1111/cobi.13139.

Omernik, J.M. 1987. "Ecoregions of the conterminous United States." Annals of the Association of American Geographers 77(1):118-125.

Omernik, J.M. 1995. "Ecoregions: A spatial framework for environmental management. In: Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making." Davis, W.S. and T.P. Simon (eds.), Lewis Publishers, Boca Raton, FL. p. 49-62.

Openshaw, S. and P. J. Taylor, 1979. "A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem." *Statistical Applications in the Spatial Sciences*, 127–144. London: Pion.

Perry, C. J., E. Søvik, M. R. Myerscough, and A. B. Barron. 2015. "Rapid Behavioral Maturation Accelerates Failure of Stressed Honey Bee Colonies." *Proceedings of the National Academy of Sciences* 112 (11): 3427–32. https://doi.org/10.1073/pnas.1422089112.

R Core Team. 2022. “R: A language and environment for statistical computing”. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team. 2022. “RStudio: Integrated Development for R”. RStudio, Inc., Boston, MA. http://www.rstudio.com/

Stehle, S. 2022. “Temporal Aggregation Bias and Gerrymandering Urban Time Series.” *GeoInformatica* 26 (1): 233–52. https://doi.org/10.1007/s10707-021-00452-z.

Steinhauer, N., K. Kulhanek, K. Antúnez, H. Human, P. Chantawannakul, M. Chauzat, and D. vanEngelsdorp. 2018. “Drivers of Colony Losses.” *Current Opinion in Insect Science*, 26: 142–48. https://doi.org/10.1016/j.cois.2018.02.004.

Steinhauer, N., K. Rennich, M. E. Wilson, D. M. Caron, E. J. Lengerich, J. S. Pettis, R. Rose, J. Skinner, D. Tarpy, J. Wilkes, and D. vanEngelsdorp. 2014. “A National Survey of Managed Honey Bee 2012–2013 Annual Colony Losses in the USA: Results from the Bee Informed Partnership.” *Journal of Apicultural Research* 53 (1): 1–18. https://doi.org/10.3896/IBRA.1.53.1.01.

Switanek, M., K. Crailsheim, H. Truhetz, and R. Brodschneider. 2017. “Modelling Seasonal Effects of Temperature and Precipitation on Honey Bee Winter Mortality in a Temperate Climate.” *Science of The Total Environment* 579 (February): 1581–87. https://doi.org/10.1016/j.scitotenv.2016.11.178.

US Census Bureau. 2018. “Cartographic Boundary Files - Shapefile.” Census.Gov. Accessed January 26, 2022. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html.

US EPA, ORD. 2015. “Ecoregions of North America.” Data and Tools. November 25, 2015. https://www.epa.gov/eco-research/ecoregions-north-america.

vanEngelsdorp, D., and M. D. Meixner. 2010. "A Historical Review of Managed Honey Bee Populations in Europe and the United States and the Factors That May Affect Them." *Journal of Invertebrate Pathology* 103 (January): S80–95. https://doi.org/10.1016/j.jip.2009.06.011.

Zee, R., A. Gray, L. Pisa, and T. de Rijk. 2015. "An Observational Study of Honey Bee Colony Winter Losses and Their Association with Varroa Destructor, Neonicotinoids and Other Risk Factors." *PLOS ONE* 10 (7): e0131611. https://doi.org/10.1371/journal.pone.0131611.