Spatial Analysis of Maryland Bee Observations

MATH 6384 - Spatial Data Analysis

Brady Lamson

December 8, 2025

College of Letters Arts and Sciences

Denver

Background

- Bee populations in the USA have been on a decline for decades due to a variety of factors.
- Is this something we can see from a spatial perspective as well? How have the spatial patterns of bee populations changed over time?
- My work will focus on the bee populations in Maryland specifically and comparing them across decades.

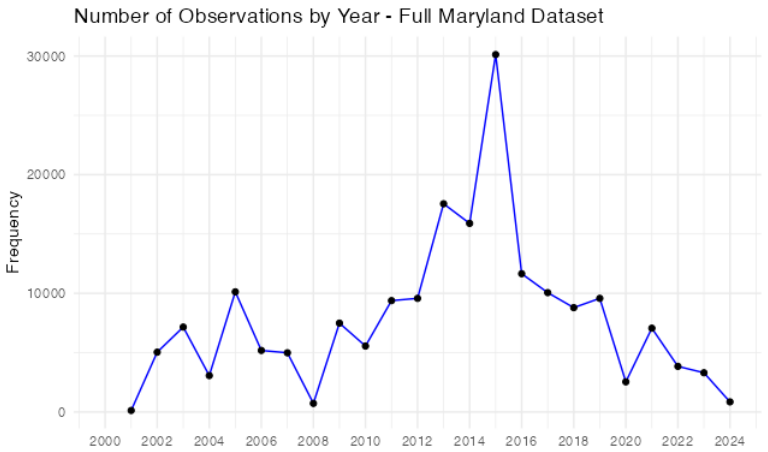
Data Source

- Data comes from the Global Biodiversity Information Facility (GBIF).
- The dataset I'm working with specifically is "Insect Species Occurrence Data from Multiple Projects Worldwide with Focus on Bees and Wasps in North America".
- It contains an enormous amount of species occurrence records for "native and non-native bees, wasps and other insects".

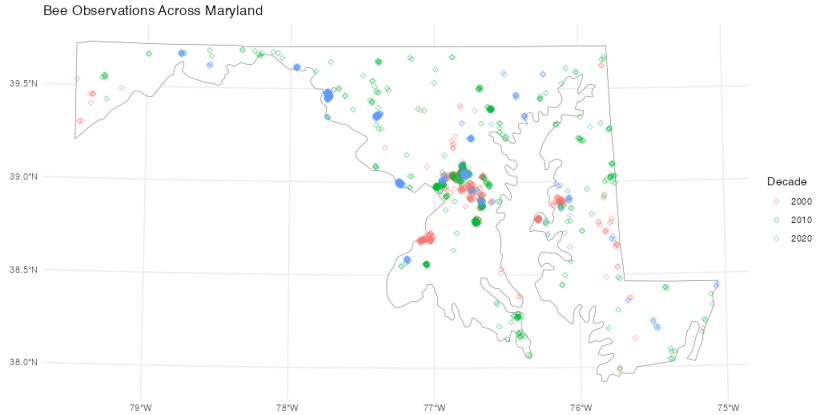
Data Overview

- Full dataset contains bee observations both in and out of the USA: nrow $= 580k$
- Filtered for bees in maryland: nrow $= 189k$
- Dataset covers observations from 2001 to 2024.
- Stratified sampling performed by decade to get 1500 rows, 500 rows per decade.
- This data will be treated as point process.

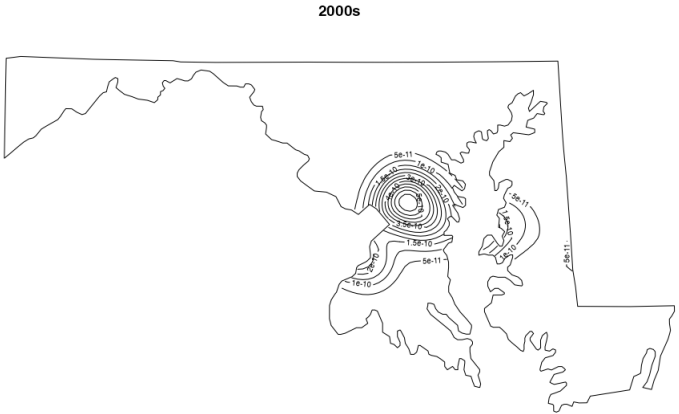
Year breakdown



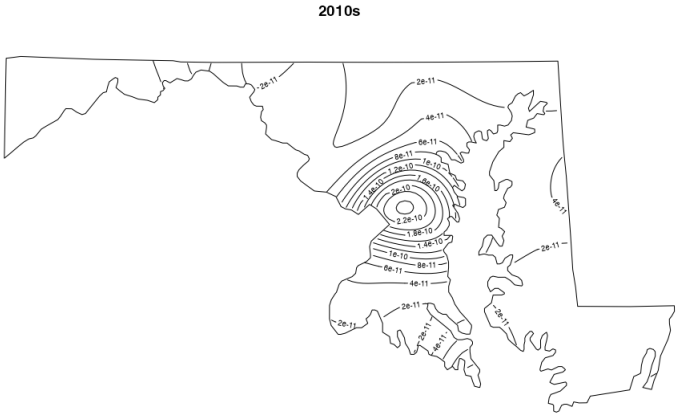
Maryland Bee Observations - Sampled Dataset ($n = 1500$)



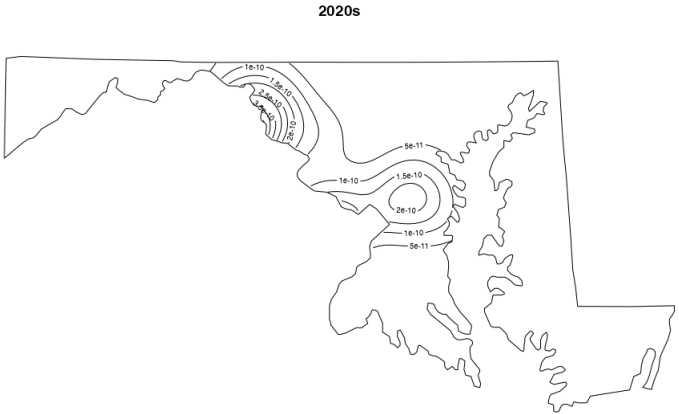
Contour Plots



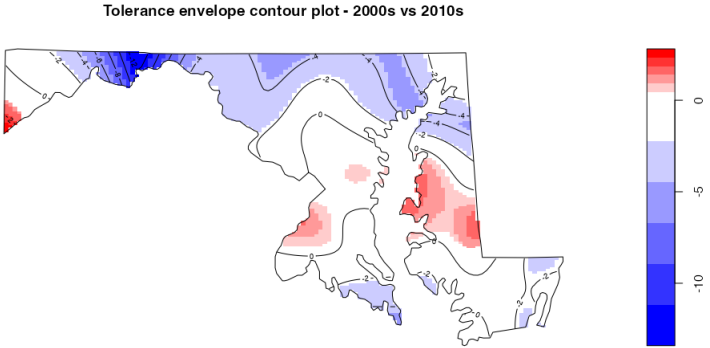
Contour Plots



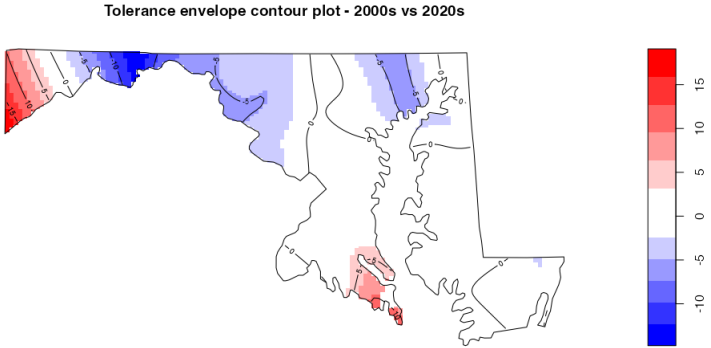
Contour Plots



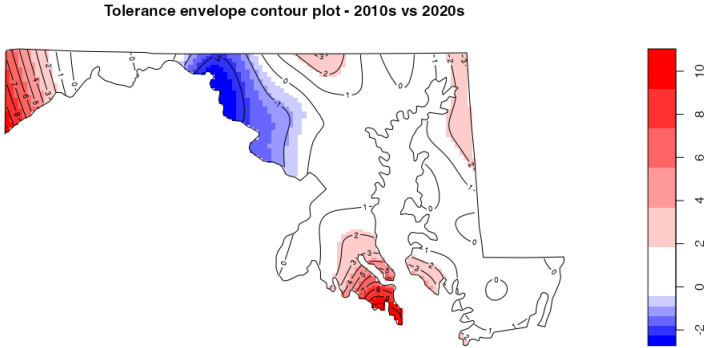
Evenelope Contour Plots



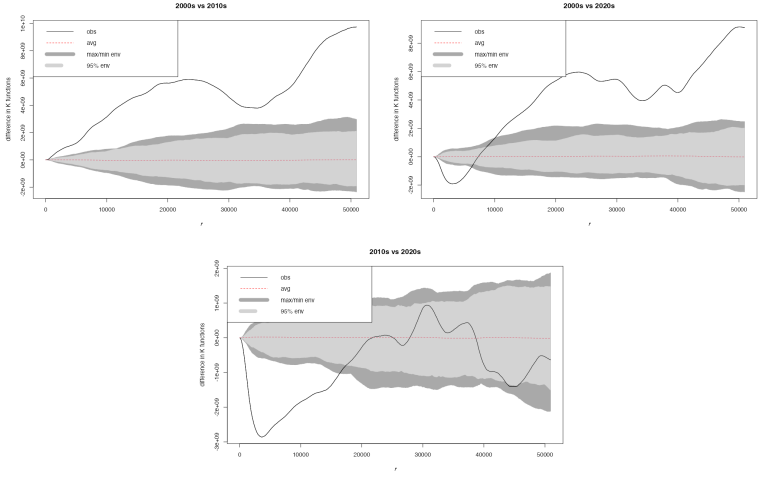
Evenelope Contour Plots



Evenelope Contour Plots



Difference in K-Functions



Q-Nearest Neighbor Results

Decade: 2000s

Contrast	p-value
T150-100	0.005
T200-150	0.110
T250-200	0.470

Decade: 2010s

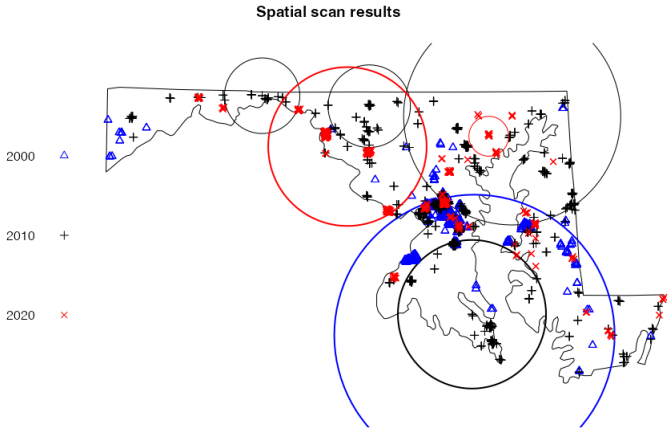
Contrast	p-value
T40-25	0.005
T55-40	0.765
T70-55	0.990

Decade: 2020s

Contrast	p-value
T175-150	0.020
T200-175	0.005
T225-200	0.005
T250-225	1.000

What we see here is that each decade shows very different spatial scales of clustering as judged by QNN.

Spatial Scan Results



Conclusion

- There do seem to be some real shifts shown by my results! Hard to draw conclusions though...
- Are changes genuine or simply due to random sampling?
- Changes may represent shifts in the human process of data collection, not shifts in bee behavior.
- To truly understand results I would need additional context on HOW data was collected and how locations were chosen.