Developing Thematic Maps of Coyote Distributions in Texas Using ChatGPT

Pat Bailey & Patrick Cunningham

1. Introduction

*Biological significance.* Determining the distributions of species in relation to their habitat requirements is of fundamental importance to ecologists, particularly in its relation to wildlife conservation (Graham and Hijmans 2006). Coyotes are one such species whose distributions have recently been extensively modeled. Coyotes have expanded beyond their historical ranges, likely due to a combination of the extirpation of large mammalian predators, habitat fragmentation, and plentiful niche availability (Hody and Kays 2018). The generalist nature of coyotes has seemingly allowed them to thrive within anthropogenically altered landscapes, exploiting the readily available prey base (Hody and Kays 2018; Greenspan et al. 2018). While their use of urban environments is extensively documented, how coyotes utilize non-urban areas remains largely unknown (Kays et al. 2008). To accurately determine the ecological niche that coyotes fill, we must understand how they utilize the entirety of their local environments. A method of accomplishing this task is to utilize GIS software to visualize coyote occupancy in relation to habitat covariates (e.g., land cover, abiotic factors, biotic factors, etc.). Effectively utilizing GIS software, however, remains a relatively difficult task due to expensive subscriptions, steep learning curves, large clunky datasets, and non-centralized data storage. To account for these barriers to entry, therefore, an effective solution may be to incorporate free coding software (R; R Core Team 2024) and large language models (ChatGPT; OpenAI 2024).

2. Approach / Results

We approached ChatGPT (version 3.5) with the goal of reproducing a comparable coyote distribution map in R to one created in GIS software (QGIS 3.36). We developed a framework for determining the structure of prompts we would provide ChatGPT with to evaluate how ChatGPT handles spatial data and interpretation. The prompts ranged along a gradient, varying in the degree of provided restrictions, package specifications, function applications, etc. The first task we gave ChatGPT was to generate an appropriate fictional dataset containing coyote occupancy data, which it managed to do after minor correction. In general, we found that ChatGPT struggled to interpret our prompts the more general they were, i.e., giving ChatGPT the ‘freedom of choice’ led to errors and circular loops. Simply specifying the packages we wanted it to utilize (*tmap*, Tennekes 2018; *raster*, Hijmans 2023; *sf*, Pebesma and Bivand 2023) allowed ChatGPT to produce an admittedly simple but functional map of coyote distributions across Texas. Following this success, we decided to attempt to create a more informative map where ChatGPT would generate a raster of Texas that displayed fictional land cover data, based on National Landscape Cover Data (MRLC 2024). This task proved to be difficult for ChatGPT as the raster it tried to create was much too large (> 1 Tb). To see if ChatGPT could produce any raster, we had it create a relatively small (10,000 km2) rectangular box to create the raster within. After some error corrections, ChatGPT managed to create a raster. From there, we tried to overlay our coyote datapoints, but ChatGPT could not parse the formula to map both a raster and point.

3. References

Graham, C. H., and R. J. Hijmans. 2006. A comparison of methods for mapping species ranges and species richness. Global Ecology and Biogeography 15:578–587.

Greenspan, E., C. K. Nielsen, and K. W. Cassel. 2018. Potential distribution of coyotes (Canis latrans), Virginia opossums (Didelphis virginiana), striped skunks (Mephitis mephitis), and raccoons (Procyon lotor) in the Chicago Metropolitan Area. Urban Ecosystems 21:983–997.

Hijmans, R. 2024. raster: Geographic Data Analysis and Modeling. R package version 3.6-27, https://rspatial.org/raster.

Hody, J. W., and R. Kays. 2018. Mapping the expansion of coyotes (Canis latrans) across North and Central America. ZooKeys 759:81–97.

Jensen, A. J., C. J. Marneweck, J. C. Kilgo, and D. S. Jachowski. 2022. Coyote diet in North America: geographic and ecological patterns during range expansion. Mammal Review 52:480–496.

Kays, R. W., M. E. Gompper, and J. C. Ray. 2008. Landscape ecology of eastern coyotes based on large-scale estimates of abundance. Ecological Applications 18:1014–1027.

Multi-Resolution Land Characteristics (MRLC) consortium. 2024. United States Geological Survey. <https://www.mrlc.gov/data>.

OpenAI. 2024. ChatGPT [version 3.5]. https://chat.openai.com/chat

Pebesma, E. and R. Bivand. 2023. Spatial Data Science: With applications in R. Chapman and Hall/CRC. doi:10.1201/9780429459016, <https://r-spatial.org/book/>.

R Core Team. 2024. R: A Language and Environment for Statistical Computing. from https://www.r-project.org/.

Tennekes, M. 2018. tmap: Thematic Maps in R. Journal of Statistical Software 84: 1 – 39. https://doi.org/10.18637/jss.v084.i06.