Biodiversity Informatics, 18, 2024, pp. 13-23

INTERPRETING AND GEOREFERENCING THE CONCEPT OF "NEAR" IN BIODIVERSITY RECORDS

Peter D. Campbell[1]

Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, 66049 USA
Correspondence: pete.campbell8@gmail.com

Abstract. Georeferencing historical biodiversity specimens is a difficult but necessary task to bring data accumulated in the course of past scientific efforts into full currency for modern use. Textual locality descriptions vary widely, and are prone to error involved in interpretation of brief descriptions and often-unclear terms. Each type of locality description requires particular georeferencing methods to maximize precision and accuracy of resulting coordinates and uncertainty. Current "best practice" methods concerning textual descriptions referring to proximity to a locality (i.e., "near" a locality) are arbitrary, restrictive, or undefined. In this paper, I explore these challenges, and provide new methods for assigning geographic coordinates and uncertainty (with appropriate metadata) to such locality descriptions using point, line, or polygon shapes as the basis for Voronoi diagrams. Voronoi diagrams define the geographic space nearer to a given point than to any other point in a collection, making them ideally suited for determining the shape of such locality descriptions.

Key words: Voronoi diagram, best practices, geographic information system, Darwin Core, coordinate uncertainty.

Introduction

Understanding the geographic context from which a specimen originates is vitally important to understanding its ecological, environmental, and evolutionary history. Modern specimens are often collected with GPS-based geographic coordinates, such that their metadata document precise coordinates and information on uncertainty. However, many historical records were collected without the aid of GPS technology, so geographic coordinates need to be assigned to them retrospectively.

When a biodiversity specimen is collected without GPS-based coordinates, textual locality descriptions allow researchers to relate the specimen back to where it originated. This description can take a variety of forms: a city, an address, a landscape feature like a mountain or river, or any other geographic feature to which the specimen can be related. Although most specimens in natural history collections include some sort of associated locality description, many remain as textual descriptions without associated geographic coordinates necessary for many quantitative analyses (Murphy et al. 2004).

Of the 215 million records in the Global Biodiversity Information Facility (GBIF) data portal that correspond to museum specimens, 41.7% (90 million) do not have geographic coordinates associated. Coordinate uncertainty has significant effects in quantitative analyses (Graham et al. 2007, Bloom et al. 2017, Marcer et al. 2022); availability of such information, however, is even more limited than geographic coordinates. More than 79.1%, or 170 million, of specimens in the GBIF data portal lack information on coordinate uncertainty. In short, without coordinates, a locality description remains unwieldy, subjective, and relatively imprecise in describing the geographic context of a specimen.

The process of assigning geographic coordinates to a locality description, termed georeferencing, is a nuanced task. The point-radius method described in Wieczorek et al. (2004) outlines and justifies a methodology behind the three numbers that correspond to the latitude, longitude, and coordinate uncertainty. Wieczorek et al. (2004) use these three numbers to describe a circular area that ostensibly includes all geographic points from where the specimen might

[1] *ORCID: https://orcid.org/0000-0003-3967-3109*

have actually originated. This interpretation has become the gold standard, with leading data standards preferring such a point-radius representation (Darwin Core Maintenance Group 2023), although many efforts have been made to expand on the method and create best practices for georeferencing (Wieczorek 2001, Frazire et al. 2004, Murphy et al. 2004, Chapman & Wieczorek 2006, Guo et al. 2008, van Erp et al. 2015, Bloom et al. 2018). The most recent of these efforts are those of Zermoglio et al. (2020) and Chapman & Wieczorek (2020).

Georeferencing is not a purely objective task, as it often requires the interpretation of imperfect, inexact, or incomplete locality descriptions. Take, for instance, a locality description that reads, "found near the city of Springfield," which might be assigned a coordinate pair corresponding to the centroid of the city limits. Nonetheless, uncertainties arise immediately: how close to Springfield is "near"? Could this description refer to a point within the city? Although Zermoglio et al. (2020) and Chapman & Wieczorek (2020) make many aspects of georeferencing straightforward and reproducible, their proposed techniques and solutions regarding these ideas of nearness or proximity can certainly be improved.

The field of quantitative geography has an abundance of tools for diverse spatial analyses and inferences. Particularly relevant to the question of proximity are algorithms for creating Voronoi diagrams (also called Thessian polygons or Dirichlet tessellations). A Voronoi diagram is a collection of polygons and points (or cells), wherein the polygons describe the geographic space that is closer to a particular point than to any other point in the collection of points. The polygons are referred to as Voronoi polygons, which form a tessellation of the geometric space in question (Okabe 2000; see example in Figure 1). In this paper, I explore the potential use of Voronoi diagrams to improve some georeferencing practices related to concepts of proximity. I use these methods to overcome the arbitrary or overly restrictive nature of existing methods, and create a method that addresses the characteristics of a previously undefined locality type.

Methods

Construction of a Voronoi Diagram

All processing of shapes was done using the open-source GIS platform QGIS [ver. 3.22] and its prepackaged optional libraries. Building collections

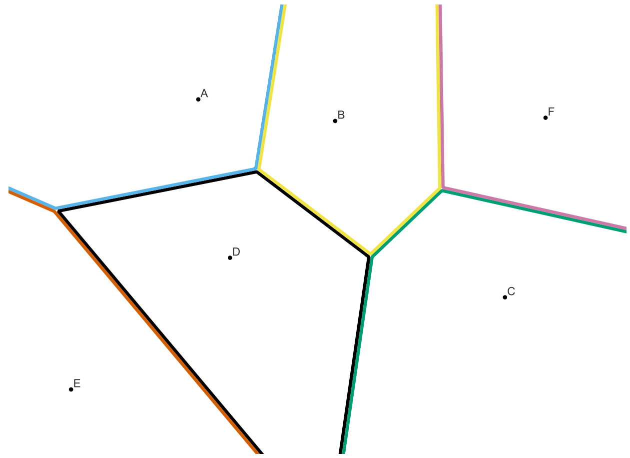

Figure 1. A Voronoi diagram of hypothetical fire stations represented as points. Each unique fire station is labeled with a letter and has a corresponding Voronoi polygon.

of various points, lines, and polygons that describe localities was done based on ESRI shapefiles. Shapefiles for cities and states were acquired from the U.S. Census Bureau (2021); shapefiles summarizing physical geographic features (e.g., rivers) were acquired from the U.S. Geological Survey (2019).

Collections of localities.—To create a Voronoi diagram for georeferencing, I first found or assembled a single shapefile that provided a collection of localities categorically relevant to the described locality (e.g., the rivers surrounding a described river). An individual Voronoi polygon is created in the context of its neighboring polygons, just as a locality description in a biodiversity record is created in a geographic context that conveys not only where the record was recorded, but also where it was not. For instance, if an individual scrub-jay (genus *Aphelocoma*) was found in Lagos de Moreno, Mexico, then that individual was not found in any other city or country on the globe.

As such, the first step to creating a Voronoi diagram for the purpose of georeferencing is to represent the desired locality as a member of a larger collection of contextually relevant localities (Fig. 2. A). In doing so, the specified locality describes where the locality *is*, and the rest of the collection describes where the locality *is not*. As in Figure 1, if one desires to find the Voronoi polygon for Fire Station F, it is only possible when station F is considered relative to stations B and C. This problem of relativity means that it is important that collections be populated both accurately and precisely. An underpopulated collection will result in larger-than-necessary uncertainty

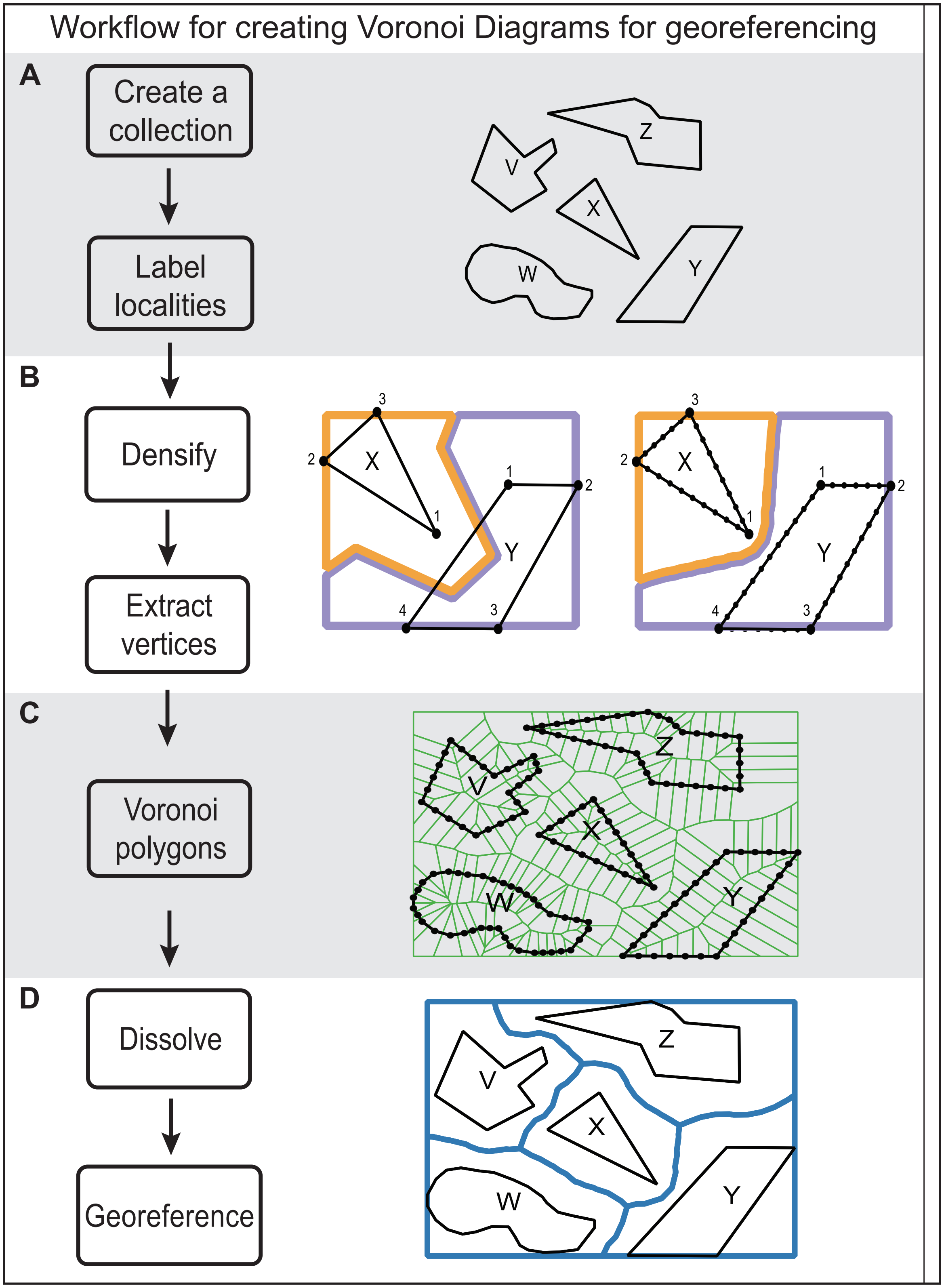

Figure 2. A general workflow for creating Voronoi diagrams for georeferencing. (A) A collection of localities relevant to X, all uniquely labeled. (B) An example of the necessity for densifying. The diagram on the left shows the Voronoi polygon for shape X encroaching on shape Y. Point X_1 is closer to parts of Line $Y_{4,1}$ than either Point Y_4 or Point Y_1. The diagram on the right shows how densifying resolves this issue, by giving Line $Y_{4,1}$ more representative points. (C) An intermediate Voronoi diagram with one Voronoi polygon per point, but several Voronoi polygons per locality. (D) A final, post-dissolve, Voronoi diagram. The Voronoi polygon surrounding locality X is fit to be georeferenced.

estimates downstream in the georeferencing process, and is therefore prone to type II error. This sort of situation typically stems from poor sources of GIS reference material (i.e. city boundary maps, collections of addresses, hydrology data, etc.). This issue can be ameliorated using official and comprehensive sources of data, such as national census data. An overpopulated collection will result in excessively confident uncertainty predictions, leading to type I errors and exclusion of potentially relevant geographic spaces. This phenomenon is a consequence of indiscriminate reference locality selection (e.g., including police stations when discussing fire stations, including secondary roads when referencing highways, etc.). A safe assumption to avoid this issue is to include in a collection the smallest class of described locality, and any class larger. For example, if the described locality is the confluence of two permanent streams, it is safe to include all streams and rivers in the collection, but intermittent streams and arroyos should be excluded, as they are smaller in magnitude. Voronoi polygons are tessellations of the collective landscape, and as such, are defined strictly relative to the overall collection.

Converting the collection to points.—A Voronoi diagram is created from a finite set of two or more distinct points in the Euclidean plane (Okabe 2000). If a collection is composed of shapes with higher dimensional complexity than points, then for this methodology the shapes should be reduced to points. While methods exist that can create Voronoi diagrams from higher-dimensional input shapes such as lines, curved lines, and polygons (Held 2001, Culver et al. 2004, Karavelas 2004), they are either deprecated, unnecessarily complicated, or too computationally intensive for the purposes of georeferencing. Although Voronoi diagrams can be created from collections of lines (Okabe 2000, p. 169) or polygons (Okabe 2000, p. 186), for my purposes aimed at georeferencing, I reduced such instances to points for a more universal and streamlined workflow that is compatible with common GIS software.

Before creating points from a collection, I ensured that each locality targeted for georeferencing had a unique identifier, so that later processing stayed organized and traceable (Richards et al. 2011). If a collection has to be reduced from a 2-dimensional shape (i.e., lines or polygons) to 1 dimensional points, the products developed from those points will need to be reassembled to accurately represent the original 2-dimensional form.

If a collection was composed of lines or polygons, I densified the point-based representation of shapes to ensure that spaces between localities were assigned accurately. The geometry of 2-dimensional shapes is typically coded as a series of longitude and latitude pairs that represent vertices of the shapes. Densifying adds additional points to the sequential list between the vertices that comprise the initial geometry description of a shape. To ensure that the farthest points of a shape's Voronoi polygon are accurate, the distance between the additional points along the shape's line segments needs to be less than half the distance to the nearest foreign vertex (Fig. 2. B). So, if the distance from shape X's vertex to a line segment of shape Y is 500 m, then shape Y needs to have additional points no more than every 250 m along its line segments. In QGIS, this step is achieved via the function "Densify by interval."

Creating a georeferencable Voronoi diagram.—I used the curated or manufactured collection of points to seed an intermediate Voronoi diagram. The Voronoi tessellation algorithm in QGIS is titled Voronoi Polygons, although other algorithms may refer to these same procedures as Thiessen, Voronoi, Dirichlet polygons, tessellations, or diagrams (Burrough et al. 2015, p.160). The QGIS algorithm has two input parameters: the collection of points, and a measure of how far to expand the borders of the output Voronoi diagram. The intermediate Voronoi diagram shapefile resulting from the Voronoi Polygons algorithm consists of many polygons, with each unique polygon corresponding to one of the input points (Fig. 2. C). Rather than using hundreds of individual Voronoi polygons to represent one original locality, these polygons can be grouped by their unique identifiers that they inherited from their input points, and merged to form one representative Voronoi polygon. Dissolving all the constituent/intermediate Voronoi polygons together creates unified Voronoi polygons that describe the geospatial area nearest each of the original input localities (Fig. 2. D). These resulting shapes can be georeferenced like any of the standard localities mentioned in Zermoglio et al. (2020); they can be easily described through both the point-radius method and in Well Known Text (WKT) format.

Case Study

To demonstrate real-world applications of Voronoi diagram-based methods, I georeferenced localities associated with a historical collection of tick records from Kansas, representing collections

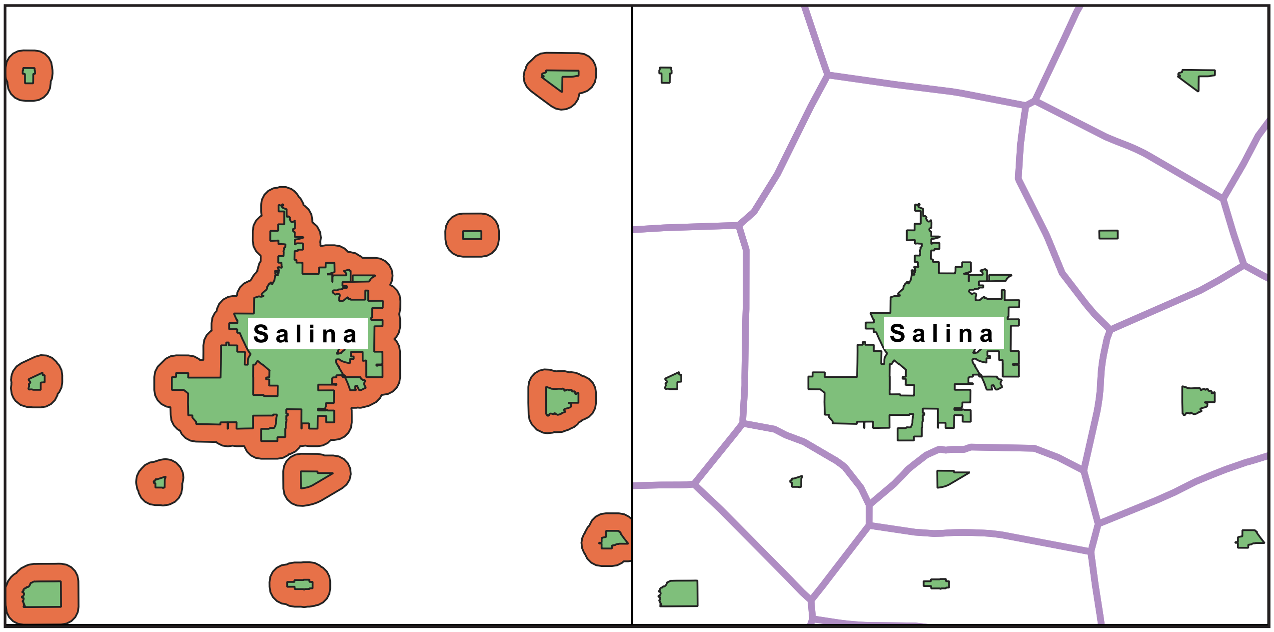

Figure 3. Current practices vs proposed Voronoi method. (Left) A simple buffer arbitrarily made at 1km, constructed following current standard practice vs (Right) Voronoi polygons constructed using the vertices of each city's polygon.

assembled by the late Prof. Donald E. Mock, Kansas State University. Between 1987 and 2003, Mock advertised across the state of Kansas, asking citizen scientists and KSU Extension agents to mail ticks to him, in an attempt to catalog and understand distributions of various tick species. Mock's complete collection comprises 1551 records of a total of more than 6600 ticks. Given the nature of the citizen scientist contributions, about 30% of the localities described in the collection are written using vague, relative, or imprecise terms. Rather than discard these specimen descriptions, I sought to georeference them. Using Voronoi diagrams, I report on localities in the following categories: "near a feature," "offset - heading," "paths, roads, and rivers," "intersections, confluences, junctions, and crossings" and "on/near the border of." I have used different sorts of georeferencing challenges manifested in this dataset to exemplify the utility of Voronoi diagrams, as described below. It should be emphasized that the goal of these methods is not to increase or decrease uncertainty, as neither outcome is indicative of a better or worse technique; rather, the goal is to make existing techniques non-arbitrary and less assumption-laden in their interpretation.

Near a feature.—Voronoi diagrams can be used to define geospatial areas that are more "near" to a given locality than to any other locality in a collection. Imprecise locality descriptions can come with a preface of proximity such as: near, around, off the coast of, etc. (Zermoglio et al. 2020). As an example from Mock's collection, one locality reads "... acquired farming near [the city of] Salina…" The current methodology for georeferencing locality descriptions including the word "near" is to arbitrarily inflate (i.e., buffer) the borders of a locality, and georeference the result (Zermoglio et al. 2020). Rather than try to guess arbitrarily at the original author's precision, I propose systematically defining the area that is nearest to the locality (Fig. 3).

In this case, the Voronoi polygon associated with the city of Salina will describe the geographic area closer to Salina than to any other similar-sized or larger city in the regions around it. Interpreting someone else's writing can never be 100% objective, and some assumptions will be necessary regarding the original author's intent. Still, georeferencing the literal interpretation of the locality description in this way accommodates conservative interpretations of uncertainty (like within the city boundaries or the majority of mailing addresses), while remaining as objective as possible about the text to avoid being arbitrary. Using Voronoi diagrams to interpret the concept of "near" thus makes the process of georeferencing a vague locality description more systematic and the interpretation more consistent.

Offset - heading.—Voronoi polygons can also provide bounds within which relative headings can be georeferenced, as in the case of locality descriptions combining a locality and a heading. An exam-

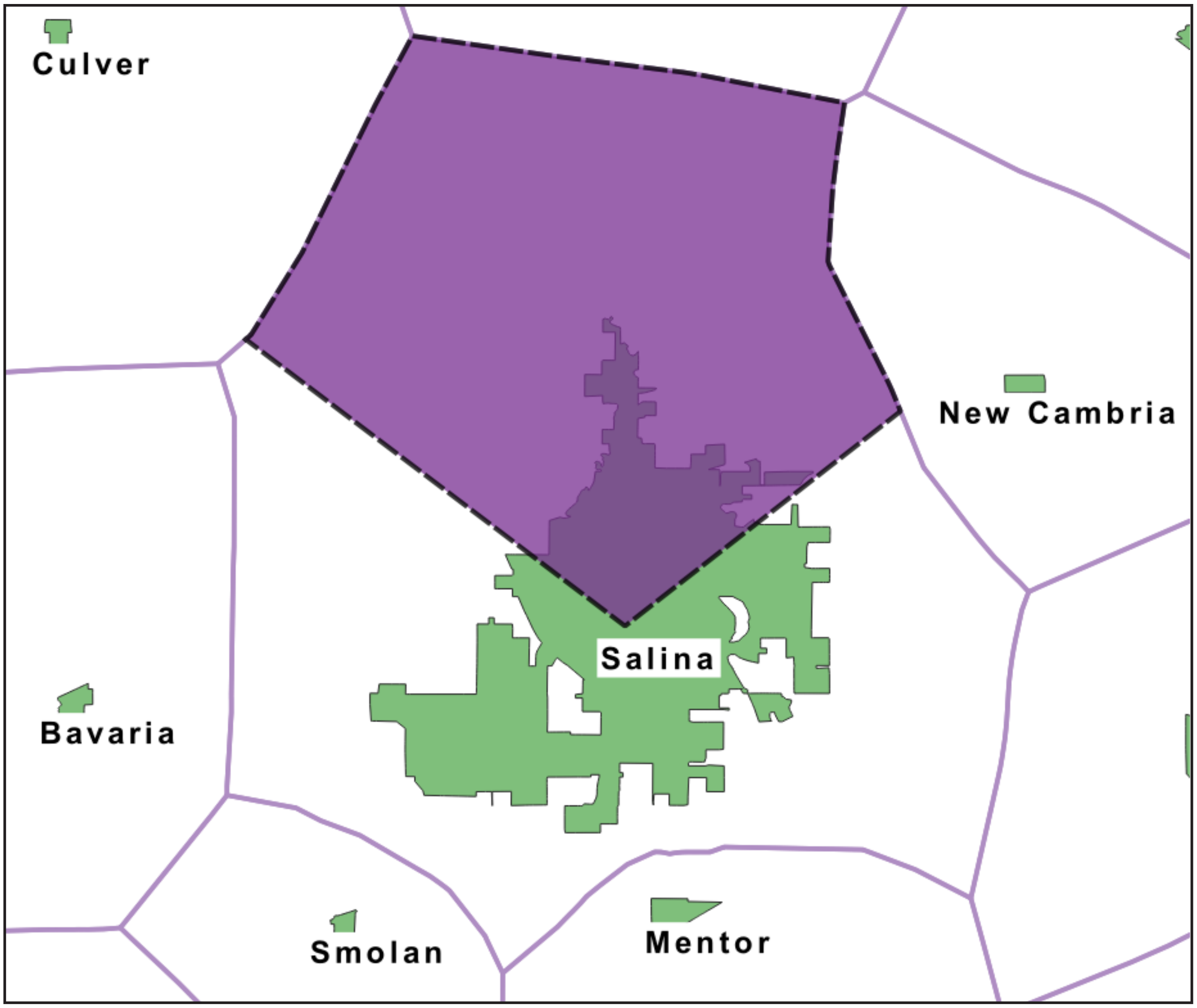

Figure 4. A Voronoi diagram based around the city of Salina. The purple lines show the boundaries of each nearby city's Voronoi polygon. The black-dashed, purple shaded region shows the subsection of Salina's Voronoi polygon north of the city. The cone used to create this region starts at Salina's centroid (38.8147, -97.6110) and extends northward at the arc suggested in Wieczorek et al. (2004). The resulting polygon has a centroid at (-97.6201, 38.8993), and a coordinate uncertainty of 9420 m.

ple would be a specimen recorded as "found north of Salina." Although this description is much more precise than the "near a feature" example, it still remains vague and imprecise.

The locality and heading act as a pair that combine to form a cone. The center of the named locality represents the apex of the cone, whereas the heading description determines the cone's direction and arc. The current standard for georeferencing localities with a heading is to create this cone and extend it in the direction of the heading "...until reaching a constraining boundary imposed by other information in the locality record, or until reaching the proximity of another similar feature" (Zermoglio et al. 2020).

The methodology behind the standard just described is similar to the Voronoi method, in that both rely on relevant locality neighbors to define the outer limits of the new representative shape being georeferenced. However, the current method is not explicit about where these limits lie. Rather than relying on the discretion of the interpreter, Voronoi polygons can define the "proximity" of neighboring localities objectively, and provide a constraining boundary. A Voronoi diagram will define all of the area bordering *all* relevant localities, not just the first in the path of the cone. The intersection of Salina's Voronoi polygon and the northward cone then explicitly describes the area only found north of Salina, as opposed to west of New Cambria, southeast of Culver, or northeast of Bavaria (see Fig. 4). This slice of the Voronoi polygon can then be georeferenced using standard protocols and described precisely using WKT formats.

Paths, roads, and rivers.—Paths, roads, and rivers can be difficult to georeference accurately owing to their long, thin geometry. A simple point-radius

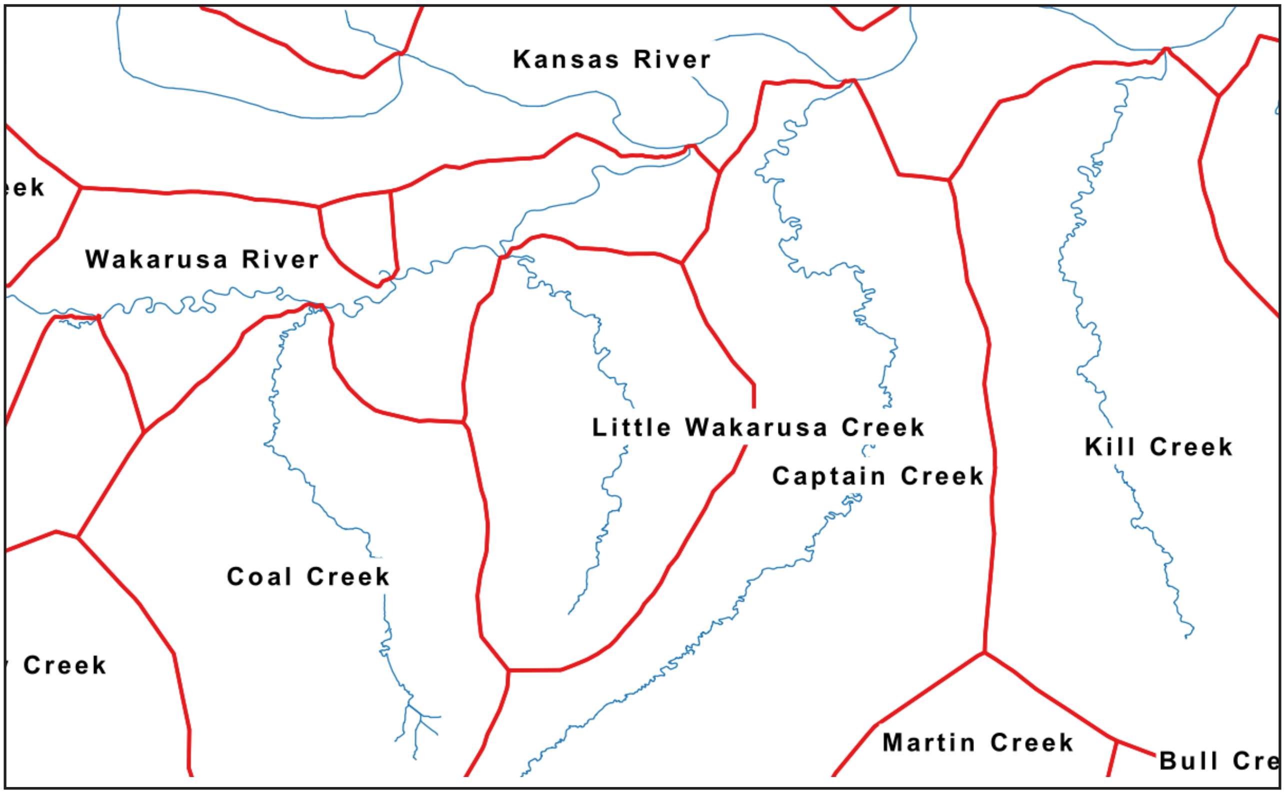

Figure 5. An example of a Voronoi diagram based on rivers. The red lines indicate the borders of each river's (blue) Voronoi polygon. Notice the rivers only touch Voronoi polygon borders when converging with another river.

measurement based on a centroid does not describe well the uncertainty surrounding the terminal ends of a linear path due to its strictly non-circular shape, such that uncertainty is likely over appreciated in the middle and underappreciated at the ends. Rather than use an arbitrary buffer to partially compensate for this problem at the terminal ends, a Voronoi diagram describes the area closer to one path than to any other in a collection, which includes appropriate treatment of the ends of the path (Fig. 5). Creating a Voronoi diagram for a collection of lines or polylines is straightforward, and follows the same process as for a collection of polygons as described above. While a Voronoi polygon can be a useful tool for georeferencing paths, it may still fail to overcome the inherent linearity of the locality at the coarsest scales. Voronoi polygons for exceedingly long paths (the Mississippi River, an interstate railroad, interstate highways, etc.) may still not be described well by the point-radius method, which may highlight the utility of the footprintWKT Darwin Core field. In the end, if an excessive path is not described as hemmed in by some other type of border (e.g., found along the Kansas River *in Douglas County*), then the resulting uncertainty will be high no matter what metric is used.

Intersections, confluences, junctions, and crossings.—Describing the area corresponding to an intersection or confluence using Voronoi polygons is a more accurate and realistic interpretation of a locality description for such crossings. Consider the locality description, "143rd St. and Evening Star Rd." This locality description taken literally would imply that a specimen was found somewhere in the middle of the 16 m^2 section of pavement that makes up that particular intersection. Indeed, that is the interpretation that is the current best practice for georeferencing crossings: to measure the shape that the two paths create via their overlap. Although road kills represent a common source of specimen collections, in many cases, it is more reasonable to assume that a collector is using an intersection as a reference point and is collecting specimens nearby.

Using a Voronoi diagram to describe areas surrounding intersections as part of a collection allows the georeferencer to define the area that is closest to each individual intersection. In effect, Voronoi diagrams systematically and quantitatively enlarge areas of uncertainty associated with crossings according to the proximity of other such crossings in the collection under analysis. This alternative aims to err

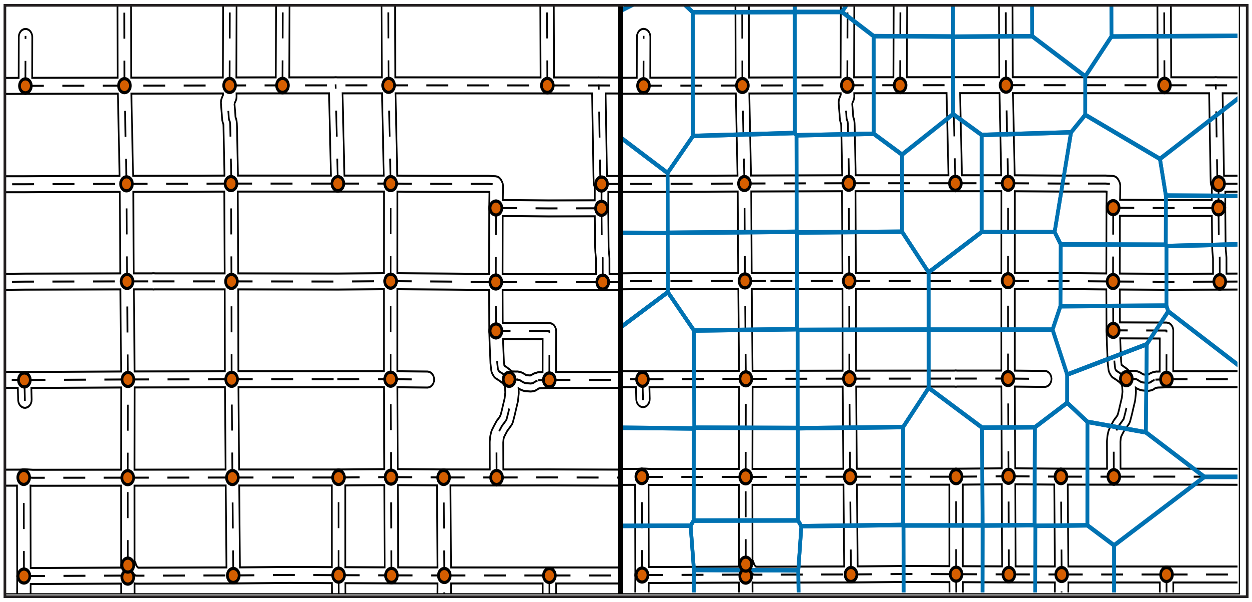

Figure 6. A comparison of current practice vs Voronoi diagram interpretation of intersections. The orange circles indicate the current interpretation of an intersection, being the centroid of the two roads and the distance to the farthest corner of the intersection. The right panel shows how Voronoi diagrams include the original orange circle interpretation, but expand on it and accommodate the surrounding environment.

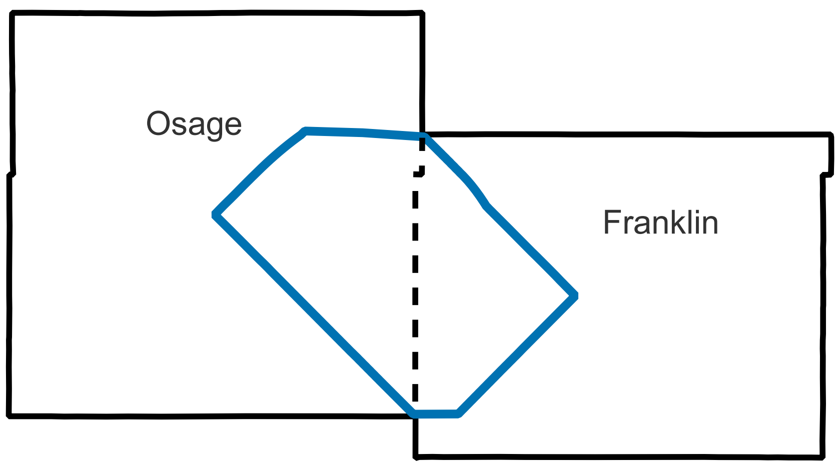

Figure 7. An example of how a Voronoi polygon can be used to quantify the area around a border. The solid black lines show the exteriors of Osage County and Franklin County. The dashed black line indicates the border adjoining the two counties, which is to be georeferenced. The blue polygon shows the Voronoi polygon of this border, with the exterior borders acting as relevant localities in the collection.

towards overestimating uncertainty, rather than attribute too much confidence to an overly narrow locality description (Fig. 6). An intersection's Voronoi polygon includes the original standard pavement section, but would also describe the greater area across which a collector might consider that particular intersection to be a reasonable reference point.

On / near the border of.—One type of locality without a specific method for georeferencing in Zermoglio et al. (2020) is a border, such as "...killed by a rancher on Osage/Franklin Co line." This locality description effectively describes a 35 km line that spans the length of the border between Osage County and Franklin County. Treating this line as a member of a collection made up of the remaining border of the two counties gives a discrete Voronoi area for referencing. This is possibly the most consistent use for Voronoi diagrams in georeferencing, as the collection is typically pre-defined by two 2-dimensional shapes (e.g., political boundaries). The desired Voronoi polygon is then a subset of the combined area of the two shapes, representing all geographic areas closer to the border of interest than to any other border (Fig. 7).

Discussion

Assumptions

Voronoi polygons are a systematic and more objective approach to georeferencing, but it is ultimately impossible to escape some level of interpretation. A description that reads, "near the city of Lawrence," implies that the described sample is *not* as near to any other city as it is to Lawrence. Note, however, that I assumed that when a locality is described as a city, it negates the description of other cities. But, what if the author *would* have described the locality as the nearby lake if the collector had known about the lake's presence? In such a case, my assumption produces larger uncertainty areas than what the collector was describing, because the collection of cities remains incomplete without lakes as well. A locality description that mentions a single specific type of locality does not inherently ignore any other category of locality.

There is also the issue of magnitude. In the described process for making collections, I assume that all localities in a collection are of equal likelihood to be referenced. It is reasonable to imagine a scenario in which a sample is collected in Shawnee (a suburb of Kansas City), yet the locality might be described as, "near Kansas City, Kansas." Localities of larger physical extent or higher population are typically more commonly known, and are thus more likely to be thought of or used in writing a locality description. For this reason, I suggest adding all localities of the referenced locality's magnitude *or larger* when creating a collection. Put more simply, if someone knows about the city of Chicago, Illinois, they may or may not know about the neighboring city of Elmhurst, Illinois but, if someone knows of Elmhurst, they almost certainly know about Chicago.

One possible solution to this issue may be the use of weighted Voronoi diagrams. A weighted Voronoi diagram is created from points with an additional value that varies based on the property being weighted (e.g., population size, square mileage, # of references, Strahler stream order, etc.; Okabe p. 119). Rather than every locality in a collection being territorially isolated, a weighted collection could depict more commonly recognized localities as having significantly larger polygons that engulf much less recognized localities. The new weighted Voronoi polygons thereby simulate the "influence" that some localities might have over the average collector's perception of the "nearest" locality. More simply, although Voronoi diagrams can be helpful in systematically georeferencing past efforts, they can not fully remove the subjective nature of interpreting someone else's work.

Shortcomings

Voronoi diagrams rely heavily on the quality of the collection being used, and a sparse collection will make for imprecise locality data, particularly in the sense of being overly general. Creating a collection composed of scarce locality types can make it difficult to fill a collection objectively. For instance, lakes in the state of Kansas are relatively few. Although many locality descriptions are simply the name of a lake, like "Tuttle Creek Lake," a tick is certainly not collected from the middle of the water, so the description must be referring to the area surrounding and associated with Tuttle Creek Lake. However, only 0.6% of the land area of Kansas is covered in water, making it the 4th driest state in the US (U.S. Census Bureau 2010), so a collection of lakes in Kansas will be quite sparse, and the corresponding Voronoi polygons will be enormous. On the other hand, Michigan, which is 41.5% water, might have the opposite issue of magnitude, where a superabundance of small lakes could restrict the Voronoi polygons of larger lakes that would more immediately draw a reference.

Although the methods described herein could be considered conceptual improvements on existing methods, they are still interpretations of inexact locality descriptions. The Voronoi diagram methods do not necessarily increase or decrease coordinate uncertainty compared to existing methods, as the existing methods produce arbitrary uncertainty measures depending on explicit assumptions. A Voronoi polygon could produce a larger or smaller uncertainty than an arbitrarily defined distance from a centroid, but conceptually, it has a contextual basis defined by the collection of localities and their spatial distributions. However, the collection on which the Voronoi methods are based is still an interpretation of the locality description's intentions. Voronoi methods do not attempt to interpret another's writing perfectly, only to make said interpretation more methodical and inclusive.

Future Developments

A promising opportunity for reporting information associated with Voronoi polygons lies in the footprintWKT Darwin Core field (Darwin Core Maintenance Group 2021). This field holds an exact textual description of a spatial polygon that can be read precisely in a GIS. Currently, most data records that do provide uncertainty information provide only the three-value combination associated with the point-radius method (i.e., latitude, longitude, and uncertainty in meters), which reduces the complexity of a locality's polygon down to a circle. Though highly reproducible and simple, such a description is often too cursory.

The implication of coordinate uncertainty measures associated with the point-radius method is that the specimen has an equal chance of having come from any part of the described area. The point-radius method is particularly poor in describing non-circular localities, like concave shapes or lines. The footprintWKT field circumvents this issue by explicitly defining the shape in question with no reduction in precision. Novel methods have been developed around specimen polygons (Peterson et al. 2006, Smith et al. 2023), but without special effort of the investigator to create these shapefiles or footprint fields, the resources necessary to implement these methods remain unavailable.

Acknowledgments

I would like to thank Dr. C. Olds (Kansas State University) for locating the Mock tick dataset, and for providing high-quality scans. I thank Dr. A. T. Peterson for crucial insight, vital resources, and unwavering confidence. This work was supported by the National Science Foundation (OIA-1920946).

References

Bloom, T. D. S., A. Flower, and E. G. DeChaine. 2018. Why georeferencing matters: Introducing a practical protocol to prepare species occurrence records for spatial analysis. Ecology and Evolution 8:765–777.

Burrough, P. A., R. McDonnell, and C. D. Lloyd. 2015. *Principles of Geographical Information Systems, 3rd ed*. Oxford University Press, Oxford.

Chapman, A.D. and J. Wieczorek. 2006. *Guide to Best Practices for Georeferencing*. Copenhagen: Global Biodiversity Information Facility.

Chapman, A.D. and J. Wieczorek. 2020. *Guide to Best Practices for Georeferencing*. Copenhagen: Global Biodiversity Information Facility.

Culver, T., J. Keyser, and D. Manocha. 2004. Exact computation of the medial axis of a polyhedron. Computer Aided Geometric Design 21:65–98.

Darwin Core Maintenance Group. 2023. List of Darwin Core terms. Biodiversity Information Standards (TDWG). http://rs.tdwg.org/dwc/doc/list/2023-07-07

Frazier, C., T. Neville, J. T. Giermakowski, & G. Racz (2004). The INRAM Protocol for Georeferencing Biological Museum Specimen Records (1.3). Zenodo. https://doi.org/10.5281/zenodo.3235003

Graham, C. H., J. Elith, R. J. Hijmans, A. Guisan, A. T. Peterson, B. A. Loiselle, and the NCEAS Predicting Species Distributions Working Group. 2007. The influence of spatial errors in species occurrence data used in distribution models: Spatial error in occurrence data for predictive modelling. Journal of Applied Ecology 45:239–247.

Guo, Q., Y. Liu, and J. Wieczorek. 2008. Georeferencing locality descriptions and computing associated uncertainty using a probabilistic approach. International Journal of Geographical Information Science 22:1067–1090.

Held, M. 2001. VRONI: An engineering approach to the reliable and efficient computation of Voronoi diagrams of points and line segments. Computational Geometry 18:95–123.

Karavelas, M. 2004. A robust and efficient implementation for the segment Voronoi diagram. International Symposium on Voronoi Diagrams in Science and Engineering. 51-62.

Marcer, A., A. D. Chapman, J. R. Wieczorek, F. Xavier Picó, F. Uribe, J. Waller, and A. H. Ariño. 2022. Uncertainty matters: ascertaining where specimens in natural history collections come from and its implications for predicting species distributions. Ecography 2022: e06025. https://doi.org/10.1111/ecog.06025

Murphy, P. C., R. P. Guralnick, R. Glaubitz, D. Neufeld, and J. A. Ryan. 2004. Georeferencing of museum collections: A review of problems and automated tools, and the methodology developed by the Mountain and Plains Spatio-Temporal Database-Informatics Initiative (MaPSTEDI). PhyloInformatics 3: 1-29. https://doi.org/10.5281/zenodo.59792

Okabe, A. 2000. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. 2nd ed. Wiley, Chichester; New York.

Peterson, A. T., R. R. Lash, D. S. Carroll, and K. M. Johnson. 2006. Geographic potential for outbreaks of Marburg hemorrhagic fever. The American Journal of Tropical Medicine and Hygiene, 75(1), 9–15. https://doi.org/10.4269/ajtmh.2006.75.1.0750009

Richards, K., R. White, N. Nicolson, and R. Pyle. 2011. *A Beginner's Guide to Persistent Identifiers*. Global Biodiversity Information Facility, Copenhagen.

Smith, A. B., S. J. Murphy, D. Henderson, and K. D. Erickson. 2023. Including imprecisely georeferenced specimens improves accuracy of species distribution models and estimates of niche breadth. Global Ecology and Biogeography 32:342–355.

U.S. Census Bureau. 2021. Cartographic Boundary Files - Shapefile. http://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html.

U.S. Census Bureau. 2010. State Area Measurements and Internal Point Coordinates. https://www.census.gov/geographies/reference-files/2010/geo/state-area.html.

U.S. Geological Survey. 2019. USGS TNM Hydrography (NHD). https://apps.nationalmap.gov/services/.

van Erp, M., R. Hensel, D. Ceolin, and M. Van Der Meij. 2015. Georeferencing animal specimen datasets: Georeferencing animal specimen datasets. Transactions in GIS 19:563–581.

Wieczorek, J. 2001. MaNIS/HerpNet/ORNIS Georeferencing Guidelines. http://georeferencing.org/georefcalculator/docs/GeorefGuide.html.

Wieczorek, J., Q. Guo, and R. Hijmans. 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International Journal of Geographical Information Science 18:745–767.

Zermoglio, P., A. Chapman, J. Wieczorek, M. C. Luna, and D. Bloom. 2020. Georeferencing Quick Reference Guide. Copenhagen: GBIF Secretariat. https://doi.org/10.35035/e09p-h128