Defining Geographic Climate Regions in the Continental US: A Spatial Clustering Approach

Erin M. Ochoa

The University of Chicago

Dr. Benjamin Soltoff

2017 June 4

Abstract

Because climate can affect social conditions and outcomes, spatially aware research in the social science must consider regional variations in climate. Current models of climate regions in the continental United States are insufficient for research purposes as they either do not identify the variables and methods used to create them or only take into account precipitation and temperature and exclude other important climate conditions, such as heat index, wind chill, and barometric pressure. We use elevation and thirty years of climate-average data for a variety of weather conditions to empirically define climate regions. Weather stations are first mapped geographically, and then the map is partitioned into Thiessen polygons. Neighbors are assigned using first-order Queen contiguity. Polygons are clustered into climate regions using a spatially aware algorithm. The resulting model identifies ten distinct climate regions and provides a proof-of-concept for spatial clustering of regions based on climate variables.

Many social conditions in the United States vary regionally. Similarly, climate in the United States varies regionally. To some extent, climate may affect certain social conditions, such as health outcomes and economic conditions. Consider the following examples: Populations located in colder climates may engage in outdoor physical exercise at lower rates than populations in temperate climates because it is unpleasant to exercise in the cold; this could affect obesity rates and other health outcomes. Climates in which crops flourish may have a disproportionate share of agricultural workers compared to climates ill-suited to large-scale agriculture; because agricultural workers typically have lower education levels and incomes, this could affect such regions' average educational achievement and income.

One social condition known to vary regionally is the crime rate, particularly the violent crime rate—but it is usually characterized to do so independently of climate. Typically, research involving regional disparities in crime will control for location in the historical South; this is done to account for what is called the Southern subculture of violence, a social phenomenon to which the higher rates of violent crime in the South are attributed. There are, however, criminological theories, such as routine activities theory, that can explain a relationship between climate and crime. Psychophysiological theories of aggression also have bearing on the relationship between climate and violent crime. If climate varies regionally, then it is possible that regional differences in climate affect crime rates regionally. While there is a wide body of literature investigating the effects of climate conditions such as temperature and precipitation on crime, we have not read any criminological research that explicitly considers climate regions.

Because climate can affect social conditions, it is important that research in the social sciences take into account regional climatic differences across the country. Such spatially aware research requires the existence of a robust and empirically defined model of climate regions in the continental United States. The two main climate-region models that do exist suffer from severe limitations that render them unsuitable for research purposes.

Here we classify the area of the continental United States into regions based on climate, daylength, and elevation. We use data from the National Oceanic and Atmospheric Administration (NOAA) and implement a spatially aware clustering method. We find the resulting model to improve upon existing models and offer a proof-of-concept for spatially classifying climate regions based on climate averages.

Literature Review

The existence of a subculture of violence in the American South has long been considered a "social fact" (Cohn et al. 2004, p. 1652). Similarly, the existence of a link between temperature and violence has been postulated by laymen (Anderson 1989), and many researchers have investigated such a relationship (see Anderson 1989 for a review). Under a deterministic model, temperature is hypothesized to affect violent crime through two general pathways: routine activities, in that people will spend more time outdoors in pleasant weather and therefore have greater chances of engaging in interpersonal conflicts with each other (Rotton & Cohn 2003; Hipp et al. 2004); and psychophysiology, in that aggression has been shown to increase under extreme temperatures (Anderson 1989).

There is more to weather than simply temperature, however: two locations can have similar temperatures but very different weather because of atmospheric conditions such as precipitation and humidity. Additionally, while the weather in a given location can change daily or even hourly, regions are characterized by their long-term climates. Because climate varies across the United

States, this raises an important question about the so-called Southern subculture of violence: is it possible that what has been characterized as a geographic phenomenon is instead, or at least partially, an effect of climate on violence?

In this section, we begin by addressing the distinction between weather and climate. We then continue with a review of select studies in meteorological criminology. Finally, we report on the current state of geographic climate region definitions and the deficiencies therein.

Weather vs. Climate

The Intergovernmental Panel on Climate Change (IPCC) differentiates climate from weather in that climate describes the mean and variance of weather conditions (such as wind, precipitation, and temperature), typically over a 30-year period (IPCC 2013). Climate, therefore, characterizes the long-term average weather conditions of the region for which it is defined.

Few contemporary studies of violence, however, have extended long-term precipitation and temperature data to include other variables (such as windiness, cloudiness, daylight hours, humidity, and barometric pressure) that contribute to climate; the studies that have included such variables have suffered from numerous limitations.

Crime, Weather, and Climate

In testing the effects of routine activities and aggression on seasonal crime trends, Hipp et al. (2004) find that property crime rates are mainly influenced by routine activities but find mixed effects for violent crime rates, which results indicate are affected by both phenomena—implying two pathways between temperature and violence. Anderson & Anderson (1996) find that when controlling for an index created from temperature variables, the effect of southernness on homicide shrinks dramatically, even reaching nonsignificance when controlling for social variables; they conclude, however, that it is "implausible" that there is no effect of southernness on homicide (p. 750). Cohn (1990b), in a review of the literature on climate-related relationships with different types of crimes, surveys conflicting results of various studies that examine a possible link between heat and homicide; the author finds that "there is evidence of a long-term association of high-temperature climates with homicides" (p. 283). Rotton & Cohn (2003) conclude that at the national level, there is no link between average annual temperatures and homicide rates, but that at the state level, there emerges a U-shaped relationship between temperature and homicide, with higher homicide rates occurring at both temperature extremes.

The aforementioned studies, however, interpret the effects of temperature alone on homicide, neglecting to take into account such conditions as wind, precipitation, humidity, barometric pressure, cloud cover, and daylight hours. Here, we briefly review three studies that leverage one or more of these additional atmospheric variables (but for a more complete overview, see Cohn 1990a).

Baylis (2015) uses geolocated tweets and daily weather data for each user's location to explore the effect of temperature and humidity on hedonic state, a one-dimensional scale that defines mood. The author implements a heat index, which essentially describes a complicated interaction between temperature and humidity, and finds that high heat-index levels are significantly and inversely correlated with hedonic state; put simply: high temperatures and high humidity combined result in bad moods. Additionally, the author reports that cloudiness also has a negative effect on mood. While the author does not investigate a link between daily weather and violence or other crime, the findings that heat index and cloudiness are inversely correlated with hedonic state suggest that such meteorological conditions may be further correlated with aggression and crimes that stem from negative mood.

In an estimation of the effects of changing climate (operationalized with temperature and precipitation anomalies) on a variety of types of crimes, Mares (2013) corrects for seasonal effects and finds no significant effect of either precipitation or precipitation anomaly on homicide (but does find that homicide is significantly affected by expected temperature). The study provides a theoretical basis for the inclusion of precipitation, stating that it may counteract the positive effects of temperature on crime via both the aggression (by lowering temperatures) and routine activities (by discouraging people from venturing outside) pathways to crime.

Talaei et al. (2014) use daily data for temperature, humidity, and barometric pressure in an exploration of homicide in Mashhad, Iran; the authors conclude that there is no relationship between the given meteorological conditions and homicide. Because of multiple concerns[1] about the robustness of the research undertaken therein, however, the study makes at best a tenuous contribution to the literature on the possible relationship between weather and homicide. That the

[1] In Talaei et al. (2014), not only was the incidence of homicide so low (homicides ranged from 0–3 per day) as to render the lack of a relationship questionable, but there is also a concern that analyses at the daily level are excessively fine-grained and therefore unsuitable for capturing lagged correlations or medium- to long-term trends. Additionally, this investigation operationalizes the independent variables not continuously but categorically, splitting the full range of each variable into four sub-ranges and assessing the correlation between types of day and homicide. Despite using both temperature and humidity variables, the study does not consider any interactions (either multiplicative or via a heat index) between the two conditions. Furthermore, it is not clear how homicide is operationalized; the authors give both a daily range (0–3) and an average daily range (0.17 $\pm$ 0.45), but analyses are done using the latter term, the definition for which is nebulous.

study fails to establish links between the three weather conditions measured and homicide is not reliably indicative that no such links exist.

Beyond establishing the existence of a relationship between temperature and crime (though its shape and the exact causal mechanism have not been conclusively identified), the corpus of meteorological criminology does not offer much in the way of firm conclusions. Because of varying definitions of weather and climate, disparate units of analysis, and inconsistent study periods, different studies have reached contradictory results (Cohn 1990a & 1990b).

Geographic Climate Regions

NOAA has defined 9 distinct geographic climate regions in the contiguous United States, as displayed in Figure 1 below. It is not clear, however, what data or methods were used to define such regions, which are drawn along state lines. That state lines sometimes correspond to geographic features but are largely socially constructed suggests such regions may be somewhat arbitrary and overly simplistic. Additionally, the regions do not allow discontiguous areas to have the same climate; this is in contrast to what is arguably the most widely used geographic climate model, the Köppen–Geiger climate classification system.

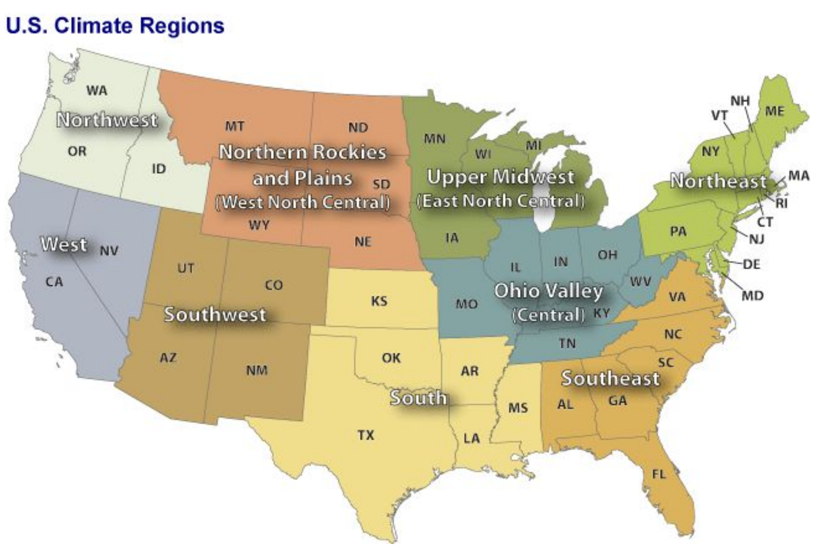

Figure 1: Geographic climate regions for the contiguous United States (NOAA, N.d.).

Dating to the year 1900, the Köppen–Geiger climate classification system has undergone numerous extensions and changes in the interim. A recent update by Peel et al. (2007a) is based on temperature and precipitation; the mainland United States portion of the world map is displayed in Figure 2 below. The Peel update shows two important features, namely that the boundaries of the historical South roughly map to a single climate region and that climate regions can be discontiguous, unlike in the NOAA model in Figure 1. While the Peel update is based on "the whole period of record" (Ibid., p. 1634) for weather stations (nearly 12,400 precipitation stations and over 4,800 temperature stations; Ibid., p. 1635–1636), this actually raises concerns that the model may be overly robust to climate trends. Furthermore, the Peel update is drawn using only precipitation and temperature data and does not take into account daylight hours, humidity, windiness, cloudiness, or barometric pressure.

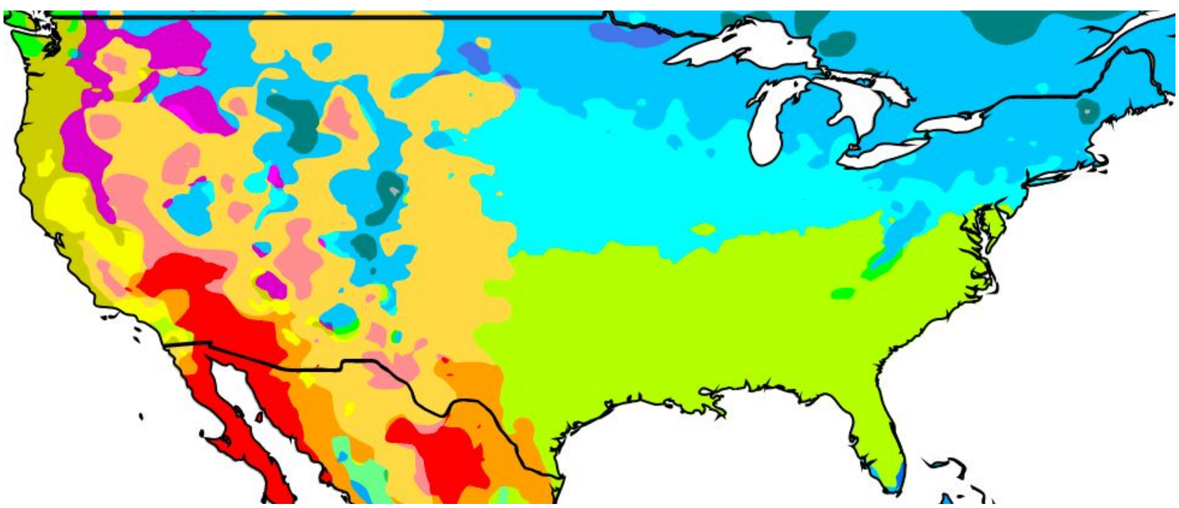

Figure 2: Peel update to the Köppen–Geiger climate classification system (contiguous United States); different colors represent different climate types (Peel et al. 2007b).

Data & Methodology

We used the 1981–2010 U.S. Climate Normals dataset from the National Centers for Environmental Information at NOAA. We implemented a variety of tools in obtaining, transforming, mapping, and processing data: Python 3.6.0 was used for data manipulations, as well as to generate data for mean daylength in minutes for each weather station (using the Astral package). GeoDa 1.8.16.4 was used to map weather stations, create Thiessen polygons, and map climate regions. Polygons were classified into regions with RStudio 1.0.136 (running R 3.3.2).

Climate Data

We identified the following climate conditions of interest from the full set of conditions available: rainfall, snowfall, temperature, dew point, heat index, wind chill, wind speed, windiness (percent calm), barometric pressure, and cloudiness (percent clear, few, scattered, broken, and overcast). In addition to defining the climate, the conditions selected were included because they are hypothesized or have been shown to affect social behavior. Rainfall was included because it may affect behavior through both the routine activities (by disrupting outdoor excursions) and aggression (by lowering the temperature) pathways. Snowfall may also disrupt routine activities. Temperature, as discussed earlier, has been shown to affect crime through aggression as well as routine activities (Hipp et al. 2004). Dew point, effectively a measure of humidity, affects the way extreme temperatures are experienced; it could affect behavior through both pathways. Heat index, as discussed earlier, has been shown to be associated with changes in mood (Baylis 2015); it may affect behavior through both aggression and routine activities. Wind chill, which describes an interaction between wind speed and temperature, decreases the perceived temperature, causing cold temperatures to feel colder; it may affect both aggression and routine activities. Wind speed may affect routine activities as people may do different things on very windy days (for example, they may avoid cycling or boating). Low barometric pressure has been shown to be associated with increases in violence, though the exact mechanism has not been identified (Schory et al. 2003). As discussed earlier, cloudiness has been shown to affect mood (Baylis 2015) and may therefore affect behavior.

We filtered out stations that fell outside the 48 contiguous United States and used only stations that had data for all the conditions of interest. The resulting 302 weather stations are mapped in Figure 3 below.

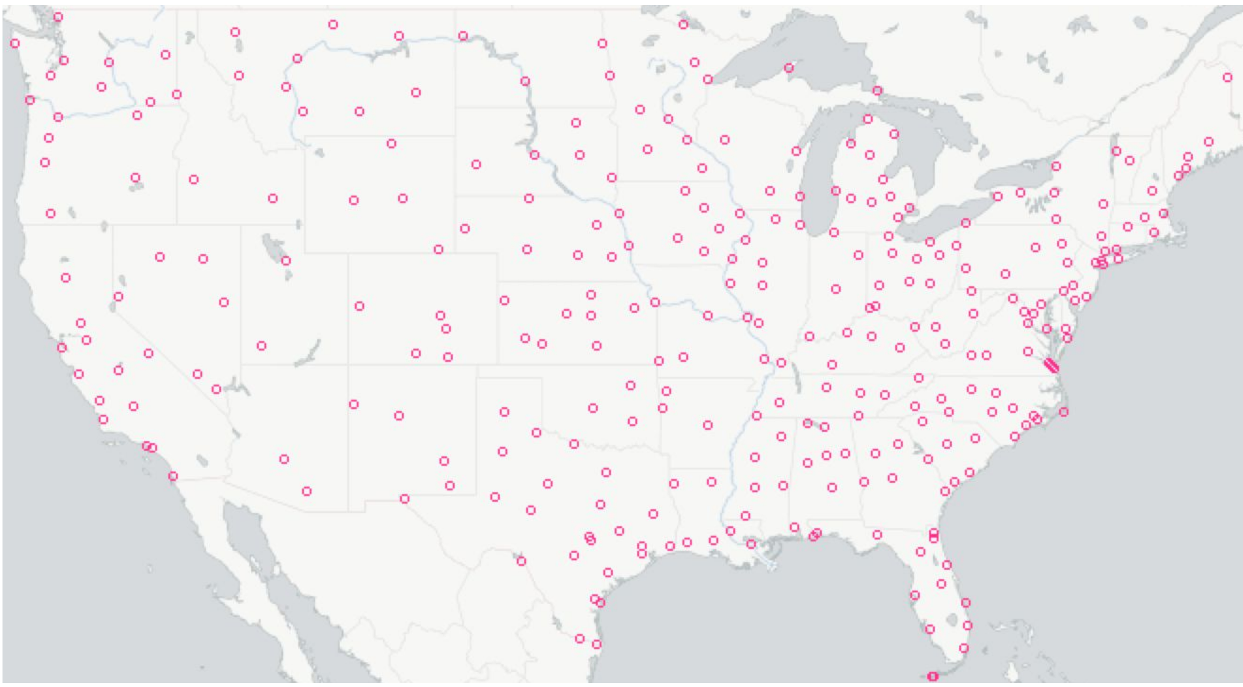

Figure 3: Weather stations (N=302).

Raw climate data were collected at weather stations located across the United States from 1981–2010. Depending on the measure, data were collected hourly or daily and reported as normals:

- *Year-to-date normals (rainfall and snowfall):* Data available as year-to-date normals consist of daily values for every day of the calendar year. Each value represents the average value for that calendar day for the years 1981–2010. For each station, daily normals were averaged across each month to obtain 12 per-month mean values.

- *Hourly normals (all other climate conditions):* Data available as hourly normals consist of 24 values per day for every day of the calendar year. Each value represents the average value for that hour for the years 1981–2010. For each station, hourly normals were averaged to obtain 12 per-month mean values; the minimum and maximum per-month values were also recorded for a selection of conditions (namely: each of the cloudiness measures, dew point, pressure, temperature, and the windiness conditions). The maximum heat index value was recorded, as was the minimum wind chill value.

For each weather station, the state, latitude, and longitude were also reported. The state was used to exclude stations outside the contiguous 48 states; the latitude and longitude were used to map weather stations.

Elevation data

For each weather station, the elevation in meters was reported. Elevation was included as a proxy for the intensity of UV radiation, which may affect how much time people spend outdoors during sunny weather. This variable is a satisfactory proxy because UV index decreases at lower elevations and rises at higher elevations (Rigel et al. 1999).

Daylight data

For each weather station location, daylength (the length of the interval between sunrise and sunset) in minutes was calculated for each day between 1981 January 1 and 2010 December 31. Values were averaged to obtain 12 per-month mean daylength variables.

We computed a total of 445 climate, daylight, and elevation variables. Table 1 below outlines the variables that were measured for each condition.

	Constant	Per-Month* Mean	Per-Month* Min. Mean	Per-Month* Max. Mean
Elevation (meters)	✓			
Daylength (minutes)		✓		
Rainfall (hundredths of inches)		✓		
Snowfall (tenths of inches)		✓		
Temperature (degrees Fahrenheit)		✓	✓	✓
Dew point (degrees Fahrenheit)		✓	✓	✓
Heat index (degrees Fahrenheit)		✓		✓
Wind chill (degrees Fahrenheit)		✓	✓	
Wind speed (miles per hour)		✓	✓	✓
Winds: % calm		✓	✓	✓
Barometric pressure (millibars)		✓	✓	✓
Clouds: % clear		✓	✓	✓
Clouds: % few		✓	✓	✓
Clouds: % scattered		✓	✓	✓
Clouds: % broken		✓	✓	✓
Clouds: % overcast		✓	✓	✓

Table 1: Conditions measured for each weather station (N=302). * There are twelve monthly variables, January through December.

Thiessen Polygons

Weather stations are points, but regions are areas. We therefore needed a way to transform points to areas. This was accomplished by creating Thiessen polygons based on the weather stations. In very plain terms, Thiessen polygons can be thought of the results of dividing an area into regions given a set of points (in this case, weather stations) such that for each two points, the algorithm places a boundary such that the resulting Euclidean distance between the points is equal. The resulting Thiessen polygons can be see in Figure 4 below. Note that the size of the polygons tends to increase at the borders of the map; this is because the algorithm creates a bounding box around the points and the regions for the outermost points are artificially stretched to end at the bounding box. Furthermore, the lower density of weather stations in the Western areas leads to larger polygons there compared to those in the central and Eastern areas.

Neighbors

Using GeoDa, neighbors were assigned based on first-order Queen contiguity, which designates regions adjacent or caddy-corner to one another as neighbors. Neighbor counts ranged from 3 to 11 with a median at 6 and a mean at 5.7 (s.d. = 1.262). The connectivity graph can be seen in Figure 4 below.

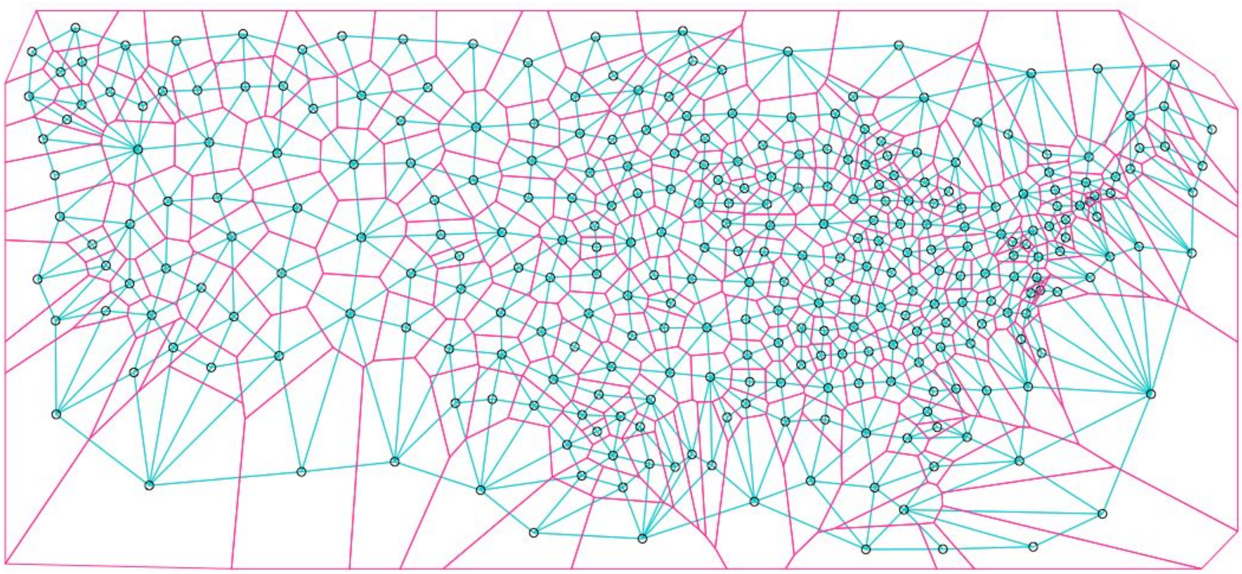

Figure 4: Connectivity diagram. Thiessen polygon boundaries are shown in pink; centroids are black circles; neighbors are connected with blue lines.

Clustering

Clustering the Thiessen polygons allows us to identify regions of the map with distinct climate characteristics as compared to other regions. For this, we used SKATER (Spatial Kluster Analysis by Tree Edge Removal) in R's spdep library. SKATER uses the minimum spanning tree based on the neighbor matrix and prunes the tree into clusters that minimize the overall pairwise dissimilarity between polygons; it maximizes both within-cluster similarity and between-cluster dissimilarity based on the high-dimensional space created by the chosen variables.

The 445 climate, daylight, and elevation variables were scaled and centered. Varying the number of regions, K, from 6–12, we created a series of clustering schemes based on the scaled and centered variables.

Results

After reviewing the resulting clustering schemes, we find that the climate-region model with K=10 strikes the best balance between parsimony and granularity. It separates the Thiessen polygons into regions that are sized reasonably and fit well with common knowledge of climate conditions across the country. The resulting map is displayed in Figure 5 below.

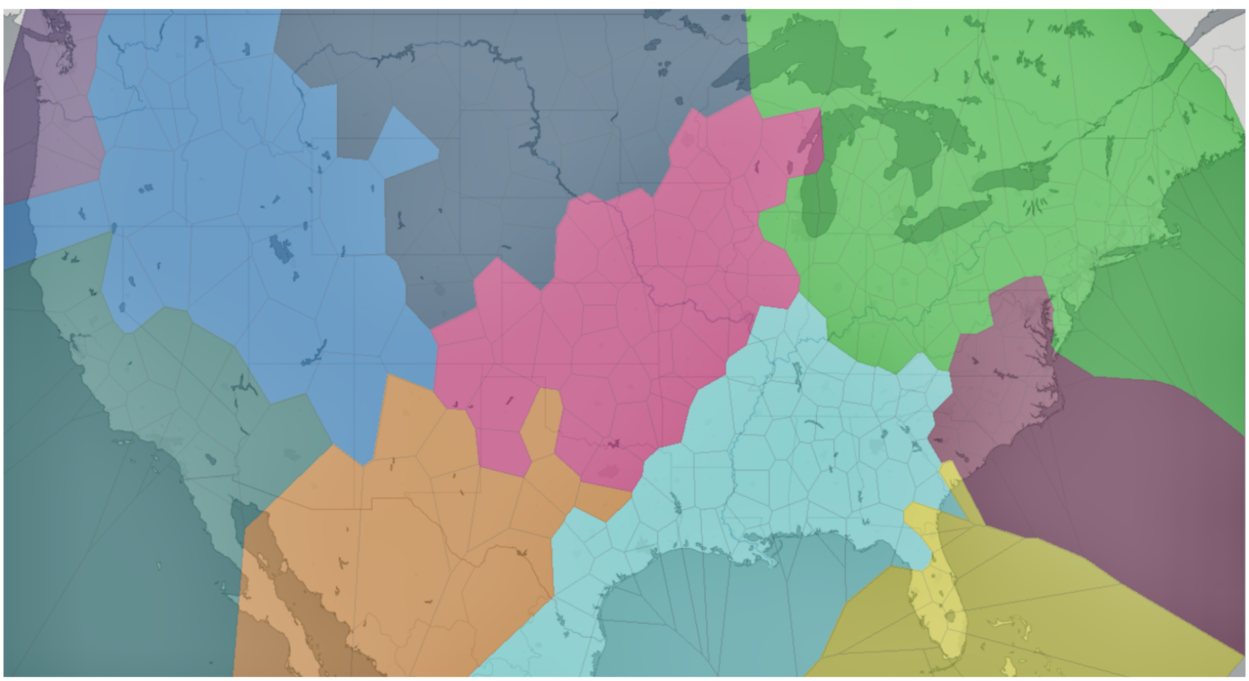

Figure 5: Thiessen polygons (N=302) clustered into 10 climate regions.

The next best models are those resulting from K=9 and K=11, displayed in Figure 6 below; these have characteristics that render them inferior to the clusters resulting from K=10, however.

With K=9, the majority of the Plains and Rocky Mountain areas are classified into a single vast climate region that dominates the map. This region is undesirable because the granularity is too low, grouping regions known to have distinct climates. With K=11, a very small region emerges along the California coast. This region is suboptimal because, even though it is consistent with experiences of distinct climates along the California coast and inland from there, such a region might contain too few observations when used in social-science research and thereby lead to a decrease in statistical power.

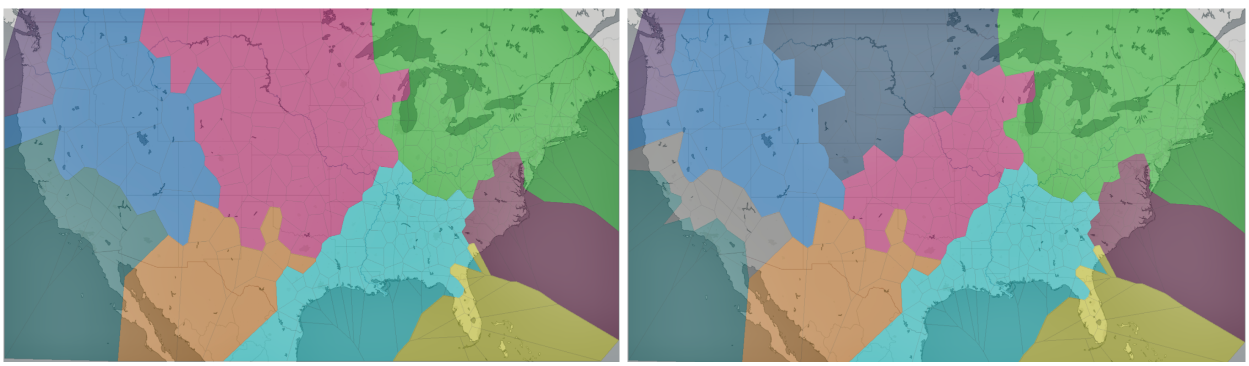

Figure 6: Thiessen polygons (N=302) clustered into 9 (left) and 11 (right) climate regions.

Conclusion

The research undertaken herein was constrained by a number of limitations, namely those related to the data available, the clustering algorithm implemented, and the definition used for designating neighbors. First, with regard to the data, the number of weather stations used (N=302) was limited because of the inclusion of a wide variety of climate conditions; this caused the resulting Thiessen polygons to be large in some areas, particularly in the West. We could draw smaller polygons by using fewer climate conditions such that there was an increase in the number of weather stations available. Smaller polygons would result in higher-resolution boundaries, which could indicate more-robust divisions between climate regions.

Second, the results of SKATER do not indicate which climate conditions are the most influential to the between-region boundaries. Determining which conditions contribute the most to the classification of clusters would allow us to drop less-important conditions and could lead to the development of a more parsimonious model (in addition to possibly increasing the number of observations, as mentioned in the preceding paragraph). One way to attempt this would be to create principal components based on the variables, then choose the first C components based on the elbow in the scree plot. Then, for those first C components, the loadings would be examined. The variables that loaded the most heavily (either positively or negatively) could be retained and the remainder discarded. We could then compute clusters with SKATER based on the retained variables. One challenge with this approach would be that for most of the conditions, there are 36 variables (per-month mean, per-month minimum mean, and per-month maximum mean) and it is possible that one but not all months of a given condition could load heavily. For example, if mean scattered clouds in the month of March loads heavily but the remaining mean per-month scattered-cloud variables do not, should the response be to include only the March mean scattered-cloud variable and discard the rest of the per-month variables for that condition, or to include them all?

Third, SKATER does not offer a method for choosing the ideal value for K; this means that K was arbitrarily chosen and did not span the full range of possibilities (2–301, excluding classifying the polygons into a single region and classifying all polygons as their own regions, yet allowing for single-polygon regions). The number of clustering schemes examined, with K ranging from 6–12, was necessarily limited; it is unknown whether a value for K outside those boundaries could have returned a clustering scheme superior to that resulting from K=10.

Finally, the scheme used to define neighbors, first-order Queen contiguity, was the most intuitive but by no means definitive. The type of contiguity selected affects the minimum spanning tree, which in turn affects the results of clustering with SKATER. Examine, for example, Figure A1, which shows the minimum spanning tree under three different definitions of contiguity. Note that the minimum spanning tree for each of the alternative neighbor definitions appears more dense than that for the Queen contiguity definition. Figures A2 and A3 show the climate-region clusters resulting from these alternative neighbor definitions. Note than in each case, discontiguous polygons are sometimes clustered together, something not present in the clustering schemes based on first-order Queen contiguity. It is possible that the large size of some Thiessen polygons render both alternative neighbor definitions suboptimal for this type of spatial clustering as the boundaries of such discontiguous clusters may be based on a single weather station.

Despite these limitations, however, we find that the 10-region model (using first-order Queen contiguity) is an improvement compared to the models previously available, namely the NOAA model and the Peel update to Köppen–Geiger. For comparison, see Figure A4, which displays all three models together. Note that the boundaries for the model found herein are more complex than those in the NOAA model, yet simpler than those in the Peel update.

This model, in some sense, is similar to the NOAA model, however: we do not know what variables were used to construct the NOAA model, and, of the variables used to construct the model found herein, we do not know which are most important. Regardless, because the methods used to construct the NOAA model are unknown and the ones here are plainly outlined, we still find the 10-region model defined herein to be preferable to the NOAA model for research purposes. With regard to the Peel update to Köppen–Geiger, we find that the model defined herein is also better suited to research: while the Peel update is based only on temperature and precipitation, the 10-region model is based on additional climate conditions. Furthermore, the high resolution of the Peel update would be suboptimal for research purposes because its boundaries yield small and discontiguous climate regions; such regions could pose many challenges for research because they are likely to conflict with other aerial units (such as counties and Census boundaries). While the 10-region model does not completely escape such conflicts, its lower resolution is likely to experience such conflicts to a lesser extent.

Robustly and empirically classifying climate regions opens up new avenues for spatially aware research in the social sciences. The models defined here offer a proof-of-concept for spatially clustering climate regions based on elevation and long-term averages for climate conditions. Future directions include determining the variables that contribute most heavily to between-region boundaries, increasing the number of observations, and determining the ideal number of regions.

References

Anderson, Craig A. 1989. "Temperature and Aggression: Ubiquitous Effects of Heat on Occurrence of Human Violence." *Psychological Bulletin* 106:1 (74–96). doi.org/10.1037//0033-2909.106.1.74

Anderson, Craig A., & Kathryn B. Anderson. 1996. "Violent Crime Rate Studies in Philosophical Context: A Destructive Testing Approach to Heat and Southern Culture of Violence Effects." *Journal of Personality and Social Psychology* 70:4 (740–756). doi.org/10.1037/0022-3514.70.4.740

Baylis, Patrick. 2015. "Temperature and Temperament: Evidence from a Billion Tweets." Energy Institute at Haas School of Business, University of California at Berkeley. ei.haas.berkeley.edu/research/papers/WP265.pdf

Cohn, Ellen G. 1990a. "Weather and Crime." *British Journal of Criminology* 30:1 (51–64). doi.org/10.1093/oxfordjournals.bjc.a047980

Cohn, Ellen G. 1990b. "Weather and Violent Crime: A Reply to Perry and Simpson, 1987." *Environment and Behavior* 22:2 (280–294). doi.org/10.1177/0013916590222006

Cohn, Ellen G., James Rotton, Amy G. Peterson, & Deborah B. Tarr. 2004. "Temperature, City Size, and the Southern Subculture of Violence: Support for the Social Escape/Avoidance (SEA) Theory." *Journal of Applied Psychology* 34:8 (1652–1674). doi.org/10.1111/j.1559-1816.2004.tb02792.x

Hipp, John R., Daniel J. Bauer, Patrick J. Curran, & Kenneth A. Bollen. 2004. "Crimes of Opportunity of Crimes of Emotion? Testing Two Explanations of Seasonal Change in Crime." *Social Forces* 82:4 (1333–1372). doi.org/10.1353/sof.2004.0074

Intergovernmental Panel on Climate Change (IPCC). 2013. "Annex III: Glossary." *Climate Change 2013: The Physical Science Basis.* ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_AnnexIII_FINAL.pdf

Mares, Dennis. 2013. "Climate Change and Crime: Monthly Temperatures and Precipitation Anomalies and Crime Rates in St. Louis, MO 1990–2009." *Crime, Law and Social Change* 59:2 (185–208). doi.org/10.1007/s10611-013-9411-8

National Oceanic & Atmospheric Administration (NOAA). N.d. "U.S. Climate Regions." ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php

National Oceanic and Atmospheric Administration (NOAA). 2017. "1981–2010 U.S. Climate Normals." doi.org/10.7289/V5PN93JP

Peel, M. C., B. L. Finlayson, & T. A. McMahon. 2007a. "Updated World Map of the Köppen-Geiger Climate Classification." *Hydrology and Earth System Sciences* 11:5 (1633–1644). doi.org/10.5194/hess-11-1633-2007

Peel, M. C., B. L. Finlayson, & T. A. McMahon. 2007b. "World Koppen Classification." commons.wikimedia.org/wiki/File:World_Koppen_Classification_(with_authors).svg

Rigel, Darrell S., Ethan G. Rigel, & Adam C. Rigel. 1999. "Effects of Altitude and Latitude on Ambient UVB Radiation." *Journal of the American Academy of Dermatology* 40:1 (114–116). doi.org/10.1016/s0190-9622(99)70542-6

Rotton, James, & Ellen G. Cohn. 2003. "Global Warming and U.S. Crime Rates: An Application of Routine Activity Theory." *Environment & Behavior* 35:6 (802–825). doi.org/10.1177/0013916503255565

Schory, Thomas J., Natasha Piecznski, Sunil Nair, & Rif S. El-Mallakh. 2003. "Barometric Pressure, Emergency Psychiatric Visits, and Violent Acts." *The Canadian Journal of Psychiatry* 48:9 (624–627). doi.org/10.1177/070674370304800909

Talaei, Ali, Arya Hedjazi, Amir Rezaei Ardani, Mohammad Reza Fayyazi Bordbar, & Andisheh Talaei. 2014. "The Relationship between Meteorological Conditions and Homicide, Suicide, Rage, and Psychiatric Hospitalization." *Journal of Forensic Sciences* 59:5 (1397–1402). doi.org/10.1111/1556-4029.12471

Appendix

	Min.	Max.	Median	Mean (s.d.)
Elevation: meters	1.200	2296.100	209.700	347.501 (432.945)
Monthly: Sunrise to sunset: Mean minutes	500.546	965.449	729.310	731.535 (113.725)
Monthly: Rainfall: Mean hundredths of inches	12.968	9343.097	1397.613	1748.304 (1388.234)
Monthly: Snowfall: Mean tenths of inches	0.000	1789.032	88.320	154.333 (183.548)
Monthly: Temperature: Min. degrees F	0.400	83.200	45.300	45.012 (16.200)
Monthly: Temperature: Max. degrees F	14.100	105.700	70.850	67.814 (17.360)
Monthly: Temperature: Mean degrees F	6.212	95.014	57.425	55.691 (17.142)
Monthly: Dew point: Min. degrees F	–3.100	74.700	38.400	38.787 (16.674)
Monthly: Dew point: Max. degrees F	4.800	76.400	48.600	47.927 (15.978)
Monthly: Dew point: Mean degrees F	0.652	75.242	43.504	43.486 (16.421)
Monthly: Heat index: Max. degrees F	14.100	105.700	71.000	68.739 (18.373)
Monthly: Heat index: Mean degrees F	6.588	95.553	57.615	56.171 (17.640)
Monthly: Wind chill: Min. degrees F	–11.300	83.200	42.800	41.690 (19.243)
Monthly: Wind chill: Mean degrees F	–5.551	95.013	56.650	53.477 (19.845)
Monthly: Wind speed: Min. miles per hour	0.800	12.900	5.800	5.903 (2.232)
Monthly: Wind speed: Max. miles per hour	4.300	21.700	11.500	11.568 (2.260)
Monthly: Wind speed: Mean miles per hour	3.275	15.301	8.257	8.368 (2.094)
Monthly: Winds: Min. Percent Calm	0.000	38.800	1.100	2.046 (2.991)
Monthly: Winds: Max. Percent Calm	2.700	82.900	23.000	26.444 (15.584)
Monthly: Winds: Mean Percent Calm	1.074	47.558	10.247	12.497 (8.445)
Monthly: Barometric pressure: Min. millibars	1004.600	1022.000	1014.900	1014.754 (2.615)
Monthly: Barometric pressure: Max. millibars	1009.500	1029.100	1018.700	1018.787 (2.581)
Monthly: Barometric pressure: Mean millibars	1007.062	1024.775	1016.863	1016.794 (2.496)
Monthly: Clouds: Min. % Clear	0.000	77.000	17.900	19.153 (12.457)
Monthly: Clouds: Max. % Clear	9.300	97.100	47.200	47.641 (14.192)
Monthly: Clouds: Mean % Clear	3.235	84.777	32.546	32.595 (13.279)
Monthly: Clouds: Min. % Few	0.600	33.300	6.200	6.775 (3.367)
Monthly: Clouds: Max. % Few	5.000	56.300	19.500	20.728 (7.413)
Monthly: Clouds: Mean % Few	2.985	40.357	12.286	13.000 (4.935)
Monthly: Clouds: Min. % Scattered	0.000	17.800	2.300	2.877 (2.495)
Monthly: Clouds: Max. % Scattered	0.900	43.100	11.900	13.707 (7.550)
Monthly: Clouds: Mean % Scattered	0.600	30.303	6.226	7.271 (4.420)
Monthly: Clouds: Min. % Broken	0.200	29.300	7.500	8.216 (3.594)
Monthly: Clouds: Max. % Broken	7.700	63.300	25.200	26.556 (8.360)
Monthly: Clouds: Mean % Broken	3.205	45.864	15.527	16.481 (5.683)
Monthly: Clouds: Min. % Overcast	0.000	68.300	20.900	22.401 (12.994)
Monthly: Clouds: Max. % Overcast	2.500	85.200	39.600	40.231 (14.005)
Monthly: Clouds: Mean % Overcast	0.979	74.759	29.565	30.655 (13.440)

Table A1: Descriptive statistics for weather stations (N=302). "Monthly" indicates that there are twelve variables, one for each month of the calendar year, for a given measure.

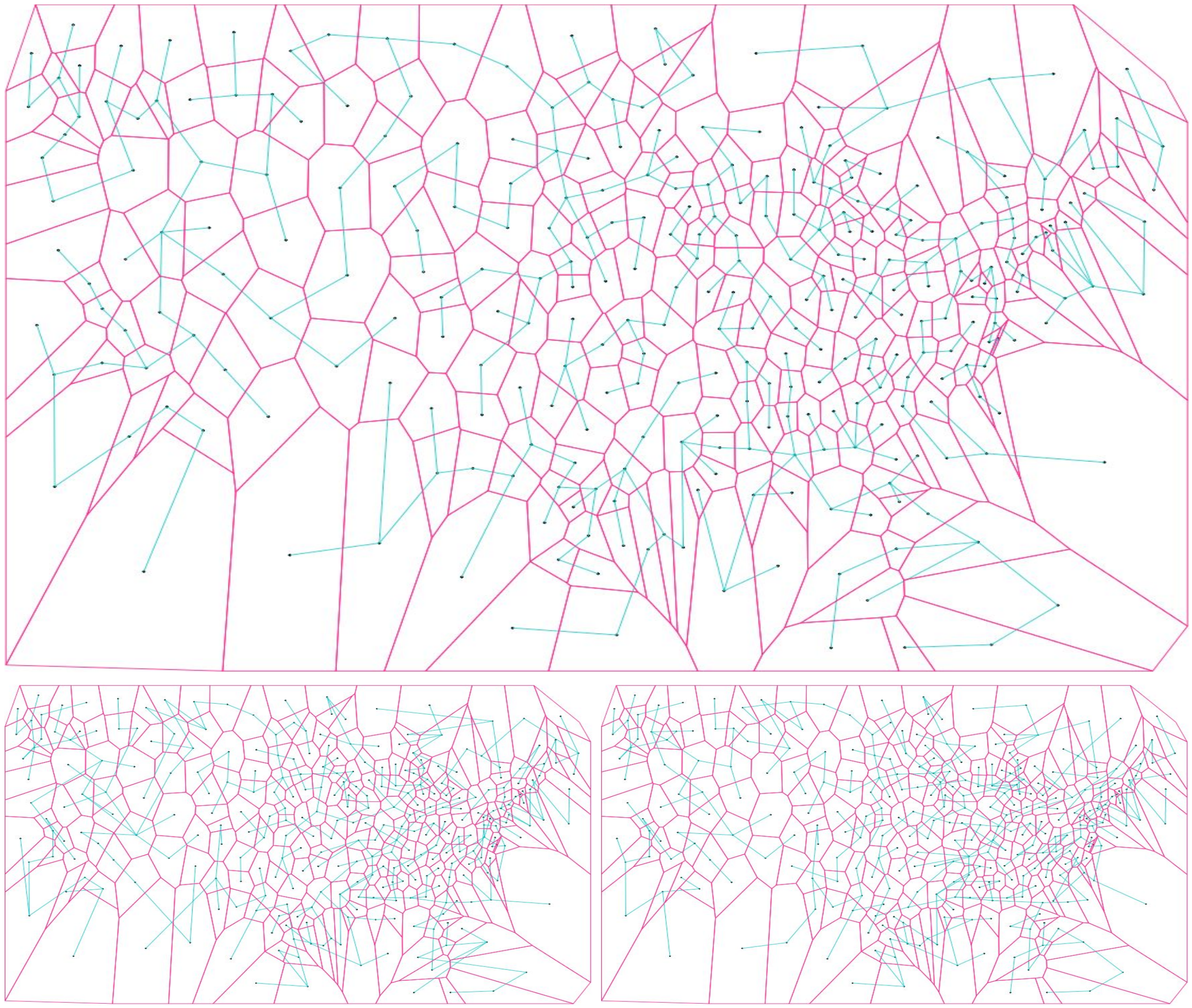

Figure A1: Minimum spanning tree under different definitions of neighbors: First-order Queen contiguity (top), second-and-lower-order Queen contiguity (lower left), and minimum threshold (5.573625, based on Euclidean distance) contiguity (lower right).

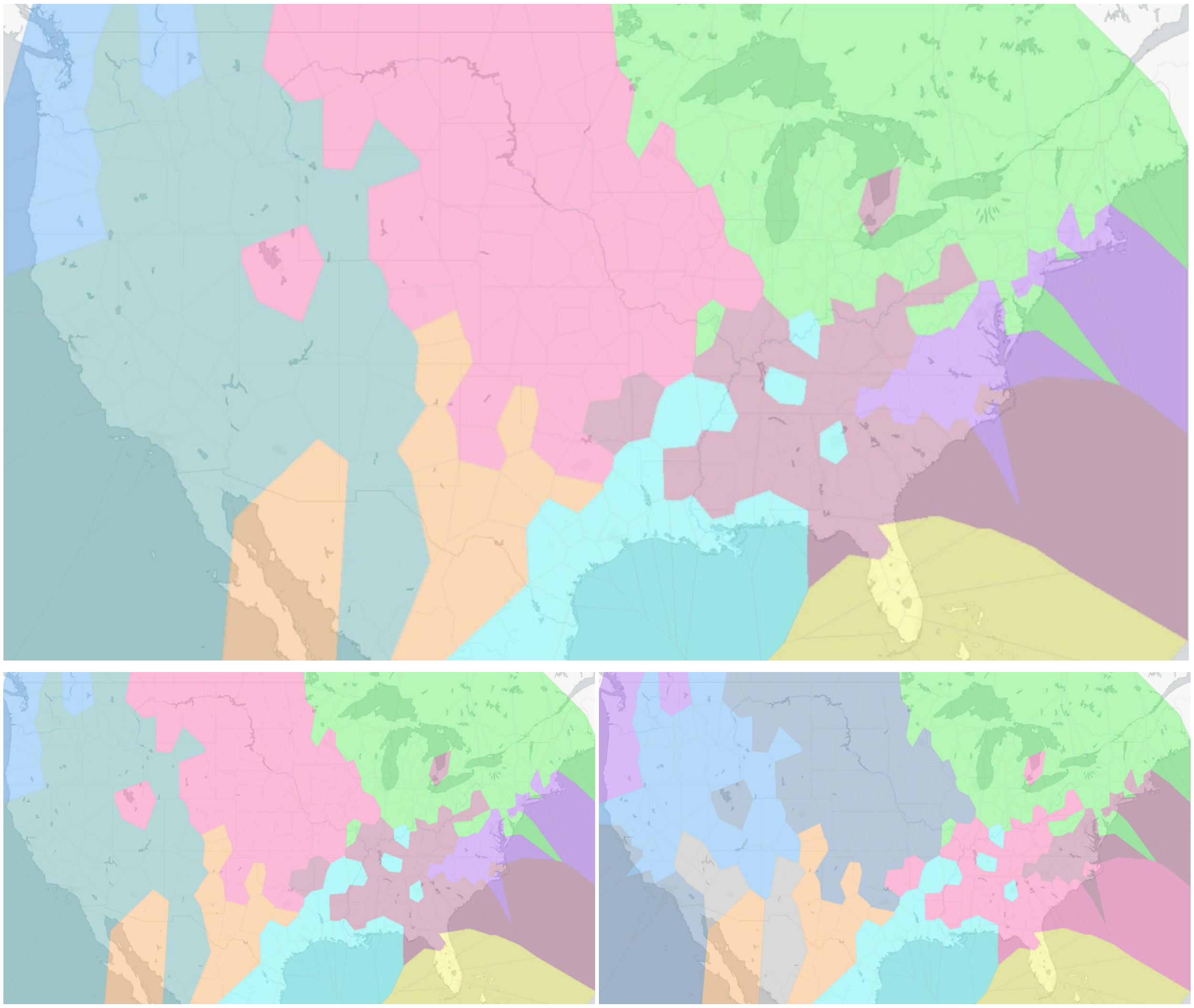

Figure A2 Thiessen polygons clustered into regions based on second-and-lower-order Queen contiguity using SKATER: 10 regions (top), 9 regions (lower left), 11 regions (lower right).

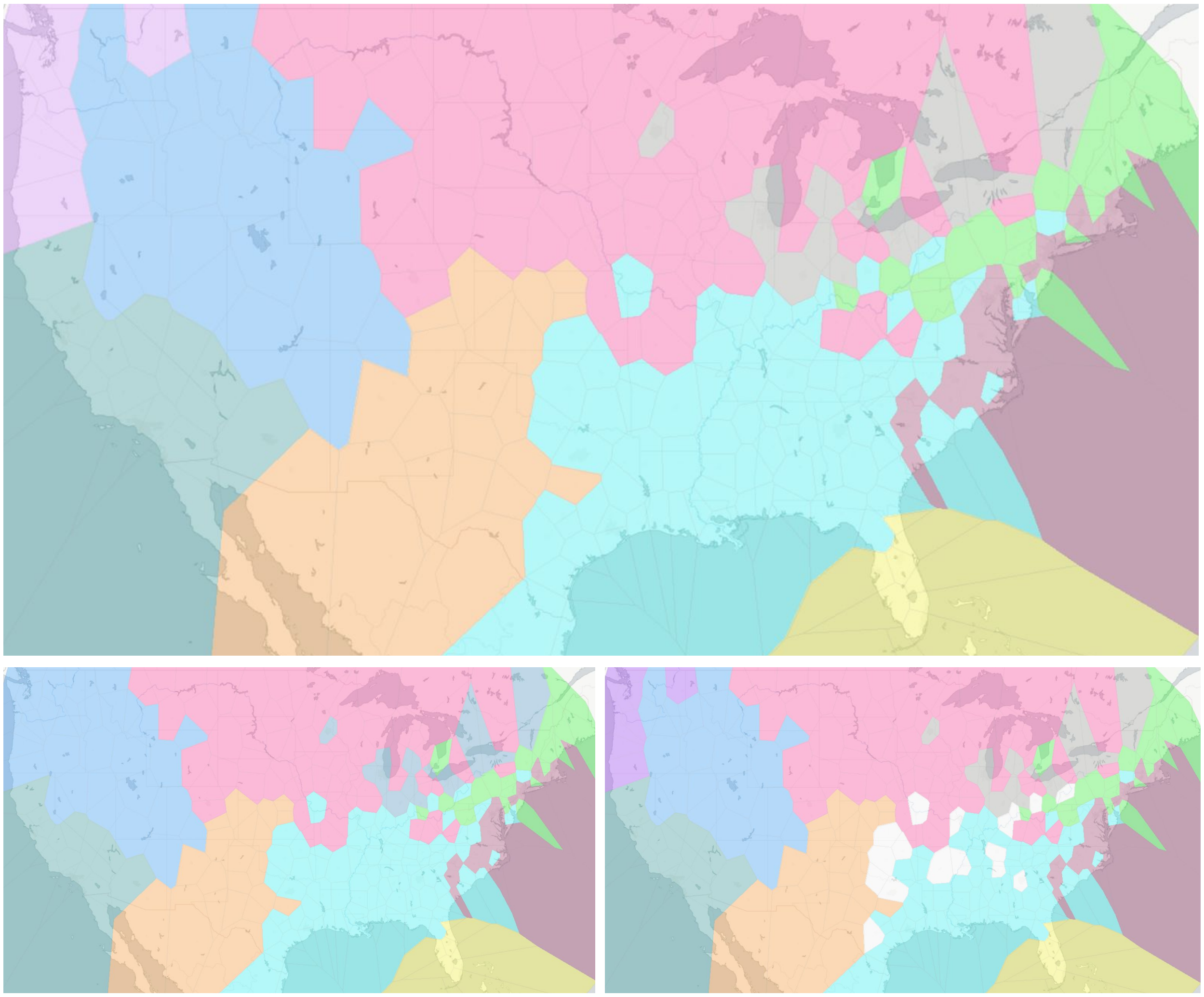

Figure A3: Thiessen polygons clustered into regions based on minimum threshold (5.573625, based on Euclidean distance) contiguity using SKATER: 10 regions (top), 9 regions (lower left), 11 regions (lower right).

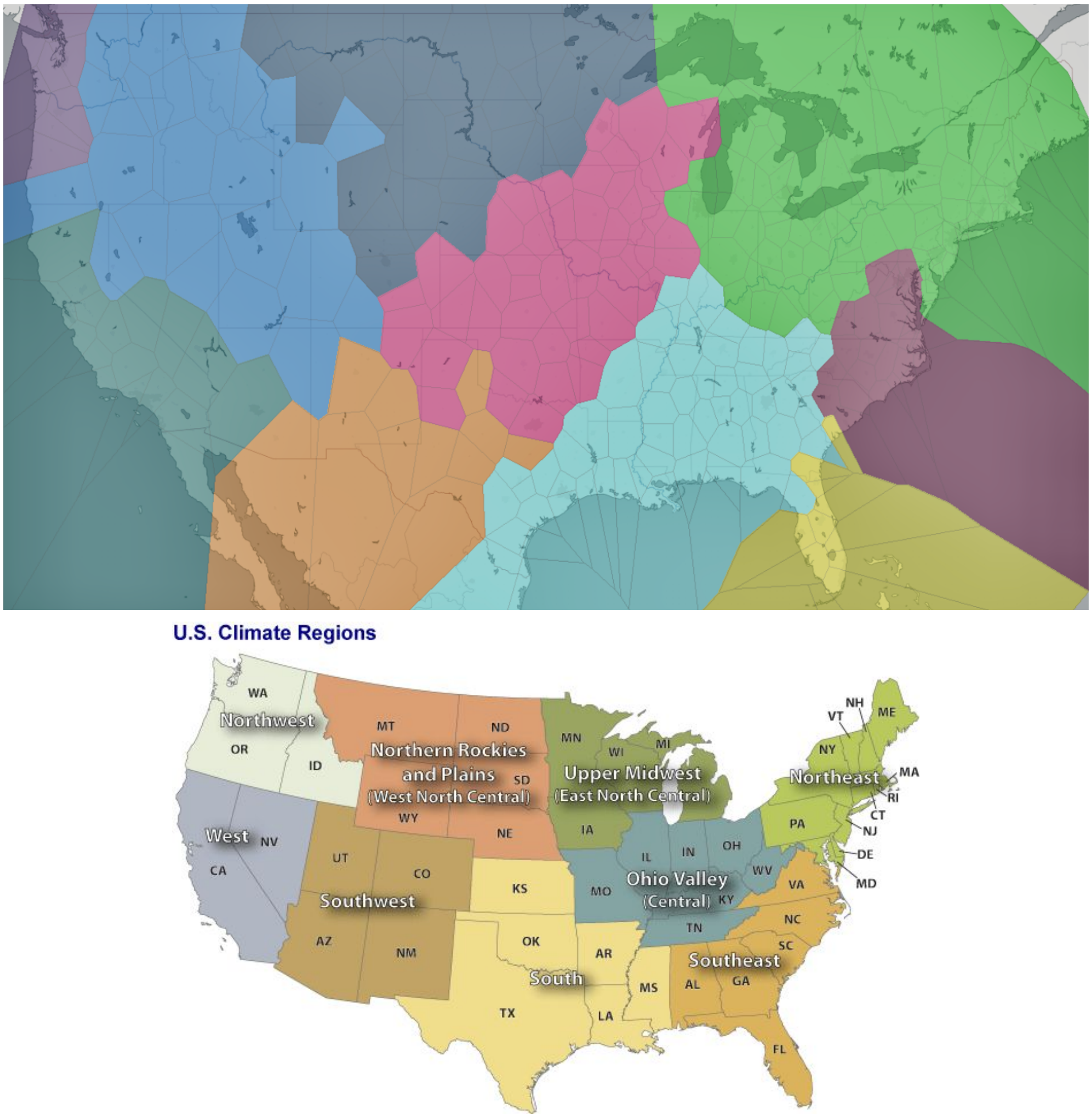

Figure A4: Comparing the results to models previously available: The prefered 10-region model found herein (top); the NOAA model (middle); the Peel update (2007) to Köppen–Geiger (bottom).