Supplemental S1

Supplemental Materials for "Mapping Agricultural Lands: From Conventional to Regenerative"

Bergmann, L.; Chaves, L.F.; Betz, C.R.; Stein, S.; Wiedenfeld, B.; Wolf, A; Wallace, R.G. Mapping Agricultural Lands: From Conventional to Regenerative. *Land* **2022**, *11*, 437. https://doi.org/10.3390/land11030437

Methods Supplement

We used the Python ecosystem around the libraries of *pandas* and *geopandas* to do most of the calculations and some visualizations; we also used QGIS for mapping; and we used QGIS, R, and ArcGIS Pro for some raster operations as needed.

General considerations regarding data from the USDA 2017 Census of Agriculture

Our general geospatial data file containing US county geometries (and the FIPS codes used to integrate data across sources) is drawn from the US Census under their TIGER/Line datasets [82]. We generally used the 2017 vin-tage, *tl_2017_us_county.shp* available at https://www.census.gov/geogra-phies/mapping-files/time-series/geo/tiger-line-file.2017.html These data were generalized to simplify the boundaries (without modifying topology or losing features) for the purposes of display.

Most of the attribute data we use is drawn from the US Department of Agriculture *2017 Census of Agriculture* [52]. We used the full census dataset, downloadable in a file called `2017_cdqt_data.txt`.

If a county did not appear in the relevant part of the census data, we assumed the relevant value was zero.

If the county appears in the census but an entry was empty in the digital database file (which corresponds to a dash in the PDF/print version), we set the relevant value to zero.

If a value was "(D)" and thus disclosure was suppressed by the census for privacy reasons, we averaged the values for that variable among all directly adjacent counties with known values (according to the county geometries found in the US Census's TIGER/Line data [82] from 2017). This only occurred in 15 counties nationwide, 2 of which fall within the 12-state study area (Milwaukee and Menominee in Wisconsin). Further, it only occurred for the variable *FarmSizeAvg_Acres* (for explanation of this variable, see below). In these cases, we have added a double-asterisk ** to the names of the counties on our maps to show that there was an imputation in some of the data associated with that county.

Finally, this census variable is of use in calculating multiple other variables in our study, so we discuss it first:

<u>*NumberFarms*</u>

From USDA 2017 Agricultural Census Chapter 2, Table 1
With the Census Short Description: "FARM OPERATIONS - NUMBER OF OPERATIONS"

Notes: This variable was used to linearly rescale many of the variables, making them into ratios, especially in the numerator.

The following counties had 10 farms or fewer in USDA data (zeroes are counties with '-'/NA in the USDA data):

STATE_FIPS	COUNTY_FIPS	COUNTY	STATE	NumberFarms
06	003	Alpine	CALIFORNIA	6
06	075	San Francisco	CALIFORNIA	10
08	111	San Juan	COLORADO	0
26	083	Keweenaw	MICHIGAN	9
32	029	Storey	NEVADA	2
34	017	Hudson	NEW JERSEY	4
34	039	Union	NEW JERSEY	9
35	028	Los Alamos	NEW MEXICO	2
36	005	Bronx	NEW YORK	0
36	061	New York	NEW YORK	7
36	081	Queens	NEW YORK	4
36	085	Richmond	NEW YORK	6
48	301	Loving	TEXAS	8
51	013	Arlington	VIRGINIA	5
54	045	Logan	WEST VIRGINI/	8
54	059	Mingo	WEST VIRGINI/	8
55	078	Menominee	WISCONSIN	3

Study variables: sources and processing

<u>*IntensiveGrazingRatio*</u>

This is *IntensiveGrazing* / *NumberFarms* where *NumberFarms* is discussed above and *IntensiveGrazing* is found at:

USDA 2017 Agricultural Census Chapter 2, Table 43

…with the Census Short Description: "PRACTICES, ROTATIONAL OR MGMT INTENSIVE GRAZING - NUMBER OF OPERATIONS"

Note: 45 counties nationally do not have *IntensiveGrazing* data; these were set to 0's.

<u>*SilvopastureRatio*</u>

This is Silvopasture / NumberFarms where Silvopasture is found at:
USDA 2017 Agricultural Census Chapter 2, Table 43

…with the Census Short Description: "PRACTICES, ALLEY CROPPING & SILVAPASTURE - NUMBER OF OPERATIONS"

Note: 464 counties nationally do not have *Silvopasture* data; these were set to 0's.

<u>*LivestockDiversity*</u>

These are Shannon diversity indices for livestock-relevant variables that are used as if they are 'species' in the formula. The variables, which are introduced more thoroughly below, are: *CattleFarms*, *HogFarms, SheepFarms, GoatFarms, ChickenLayerFarms*, *ChickenBroilerFarms*, *TurkeyFarms*, and *BeeFarms*. Mathematically, if those variables are each different p_i:

$$LivestockDiversity = -\sum_i \left(\frac{p_i}{\Sigma_k p_k} ln \frac{p_i}{\Sigma_k p_k}\right)$$

CattleFarms

USDA 2017 Agricultural Census Chapter 2, Table 11, Row 1

…with the Census Short Description: "CATTLE, INCL CALVES - OPERATIONS WITH INVENTORY"

Note: 16 counties nationally do not have *CattleFarms* data; these were set to 0's.

HogFarms

USDA 2017 Agricultural Census Chapter 2, Table 12, Row 1

…with the Census Short Description: "HOGS - OPERATIONS WITH INVENTORY"

Note: 179 counties nationally do not have *HogFarms* data; these were set to 0’s.

SheepFarms

USDA 2017 Agricultural Census Chapter 2, Table 13, Row 1

…with the Census Short Description: "SHEEP, INCL LAMBS - OPERATIONS WITH INVENTORY"

Note: 163 counties nationally do not have *SheepFarms* data; these were set to 0’s.

GoatFarms

USDA 2017 Agricultural Census Chapter 2, Table 14, Row 3

…with the Census Short Description: "GOATS - OPERATIONS WITH INVENTORY"

Note: 90 counties nationally do not have *GoatFarms* data; these were set to 0’s.

ChickenLayerFarms

USDA 2017 Agricultural Census Chapter 2, Table 19, Row 3

…with the Census Short Description: "CHICKENS, LAYERS - OPERATIONS WITH INVENTORY"

Note: 38 counties nationally do not have *ChickenLayerFarms* data; these were set to 0’s.

ChickenBroilerFarms

USDA 2017 Agricultural Census Chapter 2, Table 19, Row 20

…with the Census Short Description:"CHICKENS, BROILERS - OPERATIONS WITH INVENTORY"

Note: 426 counties nationally do not have *ChickenBroilerFarms* data; these were set to 0's.

TurkeyFarms

USDA 2017 Agricultural Census Chapter 2, Table 19, Row 24

…with the Census Short Description: "TURKEYS - OPERATIONS WITH INVENTORY"

Note: 712 counties nationally do not have *TurkeyFarms* data; these were set to 0's.

BeeFarms

USDA 2017 Agricultural Census Chapter 2, Table 21

…with the Census Short Description: "HONEY, BEE COLONIES - OPERATIONS WITH INVENTORY"

Note: 231 counties nationally do not have *BeeFarms* data; these were set to 0's.

<u>*ConservationEasementsRatio*</u>

This is *ConservationEasements* / *NumberFarms* where *ConservationEasements* is found at:

USDA 2017 Agricultural Census Chapter 2

…with the Census Short Description: "PRACTICES, LAND USE, CONSERVATION EASEMENT - NUMBER OF OPERATIONS"

Note: 213 counties nationally do not have *ConservationEasements* data; these were set to 0's.

<u>*CoverCropsRatio*</u>

This is <u>*CoverCrops*</u> / *NumberFarms* where *CoverCrops* is found at:

USDA 2017 Agricultural Census Chapter 2, Table 41

…with the Census Short Description: "PRACTICES, LAND USE, CROPLAND, COVER CROP PLANTED, (EXCL CRP) - NUMBER OF OPERATIONS"

Note: 61 counties nationally do not have *CoverCrops* data; these were set to 0's.

NoTillRatio

This is *NoTill* / *NumberFarms* where *NoTill* is found at:

USDA 2017 Agricultural Census Chapter 2, Table 41

…with the Census Short Description: "PRACTICES, LAND USE, CROPLAND, CONSERVATION TILLAGE, NO-TILL - NUMBER OF OPERATIONS"

Note: 53 counties nationally do not have *NoTill* data; these were set to 0's.

CropDiversity

The *CropDiversity* is, like the *LivestockDiversity*, a Shannon diversity index (see above for formula). Unlike *LivestockDiversity*, it is not calculated as if individual species in the entropy formula were solely each different census counts of farms with various sorts of livestock. Instead, there are four 'species' (p_i) that go into the entropy formula, two of which, *GrainFarmsTotal* and *ForageFarmsTotal*, are the sums of other variables (*GrainFarmsTotal* is the sum of *MaizeFarms, OatsFarms, SoybeanFarms, WheatFarms,* and *WildRiceFarms. ForageFarmsTotal* is the sum of *HayFarms* and *CornSilageFarms*.) The remaining two variables inputted into the entropy formula are *VegetableFieldFarms* and *FruitFarms*. These variables are defined below.

MaizeFarms

USDA 2017 Agricultural Census Chapter 2, Table 25

…with the Census Short Description: "CORN, GRAIN - OPERATIONS WITH AREA HARVESTED"

Note: 436 counties nationally do not have *MaizeFarms* data; these were set to 0's.

OatsFarms

USDA 2017 Agricultural Census Chapter 2, Table 25

…with the Census Short Description: "OATS - OPERATIONS WITH AREA HARVESTED"

Note: 1470 counties nationally do not have *OatsFarms* data; these were set to 0's.

SoybeanFarms

USDA 2017 Agricultural Census Chapter 2, Table 25

…with the Census Short Description: "SOYBEANS - OPERATIONS WITH AREA HARVESTED"

Note: 879 counties nationally do not have *SoybeanFarms* data; these were set to 0’s.

WheatFarms

USDA 2017 Agricultural Census Chapter 2, Table 25

…with the Census Short Description: "WHEAT - OPERATIONS WITH AREA HARVESTED"

Note: 707 counties nationally do not have *WheatFarms* data; these were set to 0’s.

WildRiceFarms

USDA 2017 Agricultural Census Chapter 2, Table 25

…with the Census Short Description: "WILD RICE - OPERATIONS WITH AREA HARVESTED"

Note: 3054 counties nationally do not have *WildRiceFarms* data; these were set to 0’s.

HayFarms

USDA 2017 Agricultural Census Chapter 2, Table 26

…with the Census Short Description: "HAY & HAYLAGE - OPERATIONS WITH AREA HARVESTED"

Note: In the PDF, described as for "FORAGE -LAND USED FOR ALL HAYAND HAYLAGE, GRASS SILAGE,AND GREENCHOP"

Note: 32 counties nationally do not have *HayFarms* data; these were set to 0’s.

CornSilageFarms

USDA 2017 Agricultural Census Chapter 2, Table 26

…with the Census Short Description: "CORN, SILAGE - OPERATIONS WITH AREA HARVESTED"

Note: 948 counties nationally do not have *CornSilageFarms* data; these were set to 0's.

VegetableFieldFarms

USDA 2017 Agricultural Census Chapter 2, Table 29

…with the Census Short Description: "VEGETABLE TOTALS, IN THE OPEN - OPERATIONS WITH AREA HARVESTED"

Note: In the PDF, noted as farms with Vegetables, Potatoes, Melons…

Note: 261 counties nationally do not have *VegetableFieldFarms* data; these were set to 0's.

FruitFarms

USDA 2017 Agricultural Census Chapter 2, Table 31

…with the Census Short Description: "NON-CITRUS TOTALS, (EXCL BERRIES) - OPERATIONS WITH AREA BEARING & NON-BEARING"

Note: 434 counties nationally do not have *FruitFarms* data; these were set to 0's.

LocalDirectSalesFarmsRatio

This is *LocalDirectSalesFarms* / *NumberFarms* where *LocalDirectSalesFarms* is found at:

USDA 2017 Agricultural Census Chapter 2, Table 2

…with the Census Short Description: "COMMODITY TOTALS, INCL VALUE-ADDED, RETAIL, DIRECTLY MARKETED, HUMAN CONSUMPTION - OPERATIONS WITH SALES"

Note: The full census publication describes this as follows, "Food marketing practices. This is a new section for 2017. This section consists of sales of edible agricultural products that are both produced and sold by the operation directly to consumers (farmers markets, on farm stores or farm stand, roadside stands or stores, u-pick, CSA, online marketplaces, etc.)" [52] (B-10).

Note: 115 counties nationally do not have LocalDirectSalesFarms data; these were set to 0's.

Food Flows (called Food_Flow_Over400mi_KTxKm)

Food flows over 400 miles (644 km) are calculated from the food flow database estimated by Lin et al. in 2019 [65]. We dropped all food flows between counties whose centroids are less than 400 miles apart; we used a publicly available distance matrix [81]. For a given county, then, we find its total food flow by multiplying the distances to each destination county it exports food to by the mass of those flows, adding them all up to yield a single number. Or, more formally:

$$foodflow_i^{nonlocal} = \sum_j H(d_{ij} - 400) foodflow_{ij} d_{ij}$$

...where $foodflow_i^{nonlocal}$ is the total foodflow from county i measured in kiloton-kilometers; $foodflow_{ij}$ is the food flow from county i to county j (from the 'total' column in Lin et al.) measured in kilograms; d_{ij} is the great circle distance between county i and county j measured in miles; and $H(x)$ is the Heaviside step function such that where the distance between county *i* and county *j* is greater than 400 miles the function evaluates to 1 and associated terms are counted into the sum, otherwise, for distances less than 400 miles, the function evaluates to 0 and the term does not contribute to this measure of nonlocal food flows.

Agricultural Methane (named CH4_per_km2_ag_area)

Our agricultural methane emissions data comes from Maasakkars et al. 2016 [67] in the form of several 0.1 x 0.1 degree raster layers modeling emissions associated with:

GEPA_Annualemissions_4F_Field_Burning
GEPA_Annualemissions_4C_Rice_Cultivation
GEPA_Annualemissions_4B_Manure_Management
GEPA_Annualemissions_4A_Enteric_Fermentation

Each of these rasters has values that have "units=moleccm-2s-1" (seen in the NetCDF metadata) which Maasakkars confirms (pers. comm.) is interpreted as "molecules of methane per squared centimeter per second."

Using raster algebra, we added up those grids:
Ag-Methane-Fluxes-Sum.tif =
"GEPA_Annualemissions_4F_Field_Burning.tif" +
"GEPA_Annualemissions_4C_Rice_Cultivation.tif" +
"GEPA_Annualemissions_4B_Manure_Management.tif" +
"GEPA_Annualemissions_4A_Enteric_Fermentation.tif"

In order to aggregate fluxes to county level, given that the length of a degree of longitude can be 35% different between latitudes 30N and 50N, one needs to not assume each 0.1 x 0.1 degree square is the same size throughout the US. Instead, the methane flux data raster needs to be supplemented by a raster of the same cell sizes/alignments where the cell values are the areas of those cells in cm^2. To approximate the area of a 0.1 degree x 0.1 degree "square" at longitude X and latitude Y, we used the *area()* command in the R *raster* library, such that, "values represent the size of the cell in km2." Thus far, raster areas are in km^2 and methane data is in units of "molecules of methane per squared centimeter per second"

We then derive a methane-grid.tif with units of moles/second for each cell. The calculation runs as follows:

methane-grid.tif =
*"methane-grid-area.tif" **
"Ag-Methane-Fluxes-Sum.tif" /
*(6.02214076 * 10**23) **
*1000 * 1000 * 100 * 100*

...as there are 6.02214076 * 10^23 molecules in a mole and there are 1000 * 100 cm in a kilometer.

Next, we aggregated the methane fluxes up from having the spatial extent of individual 0.1 x 0.1 degree cells into being county by county. For this, we used ArcGIS Pro's "Zonal Statistics as Table" tool, chopping and summing up *methane-grid.tif* within the boundaries of *tl_2017_us_county.shp* (mentioned at the beginning of the Supplement) to create our county flux result, a table in which agricultural methane flow in units of moles/second appear under a column.

These fluxes are calculated per county, but counties vary in area, and two counties of roughly the same absolute area may differ in their areas given over to agriculture (at an extreme, some counties have large urban or lake areas to their land.) As such, we attempt to find a relative measure of the intensity of the agricultural emissions per county by dividing the emissions by a relevant area. We chose the area in each county given by the National Land Cover Database 2016 [62] under classifications 71, 81, or 82, as they all included agriculture uses. These classes are described as follows at https://www.mrlc.gov/data/legends/national-land-cover-database-2016-nlcd2016-legend :

71 - "Grassland/Herbaceous - areas dominated by gramanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing."

81 - "Pasture/Hay - areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation."

82 - "Cultivated Crops - areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for

greater than 20% of total vegetation. This class also includes all land being actively tilled."

NLCD 2016 pixels each measure 30m by 30m and the areas under the above classes within each county were summarized using ArcGIS Pro.

From there, to get the final agricultural methane variables (agricultural methane flux per square agricultural kilometer), we divide our methane fluxes per county by the agricultural areas per county and call the results *CH4_per_km2_ag_area.*

PM2.5_nonurban

Our *PM2.5_nonurban* data is derived from calculations on WHO DIMAQ data [61]. DIMAQ's webpage notes: "Estimation of global health risks from exposure to ambient air pollution requires a comprehensive set of air pollution exposure data covering all inhabited areas. The Data Integration Model for Air Quality (DIMAQ) has produced estimates based on data from ground measurements together with information from other sources including data from satellite retrievals of aerosol optical depth and chemical transport models. It provides estimates of annual exposures to PM2.5 at high spatial resolution (0.1° × 0.1°, which equates to approximately 11x11km at the equator) globally. The sources of data include: Ground measurements from 9690 monitoring locations around the world, satellite remote sensing; population estimates; topography; and information on local monitoring networks and measures of specific contributors of air pollution from chemical transport models. Within DIMAQ, data from these sources are calibrated with ground measurements. The model provides estimates of air quality, expressed in terms of median concentrations of PM2.5, for all regions of the world, including areas in which PM2.5 monitoring is not available."

As described in the main text, our variable *PM2.5_nonurban* approximates the average $PM_{2.5}$ levels associated with agricultural activities and lands. WHO's DIMAQ model data from 2016 arrives as point data. We convert that into a gridded (vector, not raster) dataset where the grid cells are centered on the individual DIMAQ points and the grid cells take on the values at the points. This polygon grid of $PM_{2.5}$ values has cells with resolutions of 0.1 degree x 0.1 degree. Unlike other variables, this data is quite general and not at all limited to phenomena associated with agricultural lands. In contrast to the agricultural methane data above, this data includes $PM_{2.5}$ emitted within urban contexts, for example. As such, though it is far from a perfect estimation of $PM_{2.5}$ connected to agriculture, we nonetheless exclude from our $PM_{2.5}$ average for counties those areas in counties that are themselves more 'developed' in an urbanized sense.

We find our 'urbanized' areas by processing the 2016 NLCD land cover database (https://www.mrlc.gov/data/nlcd-2016-land-cover-conus). This is a 30m resolution raster. We reclassified the urbanized pixels vs non, then aggregated up to 4x4 pixels, and selected those pixels that had more than 50%

of the resulting area as urbanized to become the 'urban' areas. We converted the results to vector polygons. The above operations were carried out in ArcGIS Pro, given the file format the data arrived in.

These urbanized polygons are then removed (via GIS overlay operations) from the pm2.5 grid polygons, which removes not only whole grid cells but also removes complex fractional areas from within cells. As such, we are able to calculate the average non-urban PM2.5 for each county by first cutting (via a GIS overlay intersect operation) the non-urban pm2.5 polygon data into datasets corresponding to each different county and then by taking the weighted average of the PM2.5 measurements in the cells remaining within each county, weighting those averages by the remaining areas of those remaining cells in each county.

FarmSizeAvg_Acres

USDA 2017 Agricultural Census Chapter 2, Table 1

…with the Census Short Description: "FARM OPERATIONS - AREA OPERATED, MEASURED IN ACRES / OPERATION"

Counties with suppressed FarmSizeAvg_Acres:

GEOID	STATE_FIPS_CODE	COUNTY_CODE	STATE_NAME	COUNTY_NAME	FarmSizeAvg_Acres
32003	32	003	NEVADA	CLARK	(D)
32009	32	009	NEVADA	ESMERALDA	(D)
12019	12	019	FLORIDA	CLAY	(D)
32021	32	021	NEVADA	MINERAL	(D)
32029	32	029	NEVADA	STOREY	(D)
35028	35	028	NEW MEXICO	LOS ALAMOS	(D)
12037	12	037	FLORIDA	FRANKLIN	(D)
12045	12	045	FLORIDA	GULF	(D)
35049	35	049	NEW MEXICO	SANTA FE	(D)
55078	55	078	WISCONSIN	MENOMINEE	(D)
36081	36	081	NEW YORK	QUEENS	(D)
55079	55	079	WISCONSIN	MILWAUKEE	(D)
36085	36	085	NEW YORK	RICHMOND	(D)

As described at the beginning of this Supplement, for each of the above counties with suppressed FarmSizeAvg_Acres, a value was imputed:

GEOID	STATE_FIPS_CODE	COUNTY_CODE	STATE_NAME	COUNTY_NAME	FarmSizeAvg_Acres
32003	32	003	NEVADA	CLARK	1326.2
32009	32	009	NEVADA	ESMERALDA	1647
12019	12	019	FLORIDA	CLAY	118
32021	32	021	NEVADA	MINERAL	858
32029	32	029	NEVADA	STOREY	686
35028	35	028	NEW MEXICO	LOS ALAMOS	862.5
12037	12	037	FLORIDA	FRANKLIN	213
12045	12	045	FLORIDA	GULF	330.5
35049	35	049	NEW MEXICO	SANTA FE	1173.36
55078	55	078	WISCONSIN	MENOMINEE	238
36081	36	081	NEW YORK	QUEENS	19.5
55079	55	079	WISCONSIN	MILWAUKEE	178.333
36085	36	085	NEW YORK	RICHMOND	26.0833

When data for these counties is mapped, these county names have an asterisk placed by them.

PesticideRatio

This is a variable constructed from the following ratio:

PesticideRatio = (*ChemOtherFarms* + *ChemFungicideFarms* + *ChemNematicideFarms* + *ChemInsecticideFarms* + *ChemHerbicideFarms*) / *NumberFarms*

...where the variables are defined as follows:

ChemOtherFarms

USDA 2017 Agricultural Census Chapter 2, Table 40

…with the Census Short Description: "AG LAND - OPERATIONS WITH TREATED"

…and the Domain Cat Description: "CHEMICAL, OTHER: (TOTAL)"

Note: In the PDF, described as "Chemicals used to control growth, thin fruit, ripen, or defoliate farms"

Note: 543 counties nationally do not have *ChemOtherFarms* data; these were set to 0's.

ChemFungicideFarms

USDA 2017 Agricultural Census Chapter 2, Table 40

…with the Census Short Description: "AG LAND - OPERATIONS WITH TREATED"

…and the Domain Cat Description: "CHEMICAL, FUNGICIDE: (TOTAL)"

Note: In the PDF, described as "Diseases in crops and orchards farms"

Note: 194 counties nationally do not have *ChemFungicideFarms* data; these were set to 0's.

ChemNematicideFarms

USDA 2017 Agricultural Census Chapter 2, Table 40

…with the Census Short Description: "AG LAND - OPERATIONS WITH TREATED"

…and the Domain Cat Description: "CHEMICAL, INSECTICIDE: (NEMATICIDES)"

Note: In the PDF, described as for "Nematodes"

Note: 552 counties nationally do not have *ChemNematicideFarms* data; these were set to 0's.

ChemInsecticideFarms

USDA 2017 Agricultural Census Chapter 2, Table 40

…with the Census Short Description: "AG LAND - OPERATIONS WITH TREATED"

…and the Domain Cat Description: "CHEMICAL, INSECTICIDE: (EXCL NEMATICIDES)"

Note: In the PDF, described as for "Insects"

Note: 31 counties nationally do not have *ChemInsecticideFarms* data; these were set to 0's.

ChemHerbicideFarms

USDA 2017 Agricultural Census Chapter 2, Table 40

…with the Census Short Description: "AG LAND - OPERATIONS WITH TREATED"

…and the Domain Cat Description: "CHEMICAL, HERBICIDE: (TOTAL)"

Note: In the PDF, described as for "Weeds, grass, or brush"

Note: 21 counties nationally do not have *ChemHerbicideFarms* data; these were set to 0's.

Additional Clustering Results

Below are example clustering results from *k*-means for *k*=3 to *k*=9. Please note that the numbering and coloring of clusters is arbitrary and therefore colors or numbers shared between different Figures are coincidence, not meaningful.

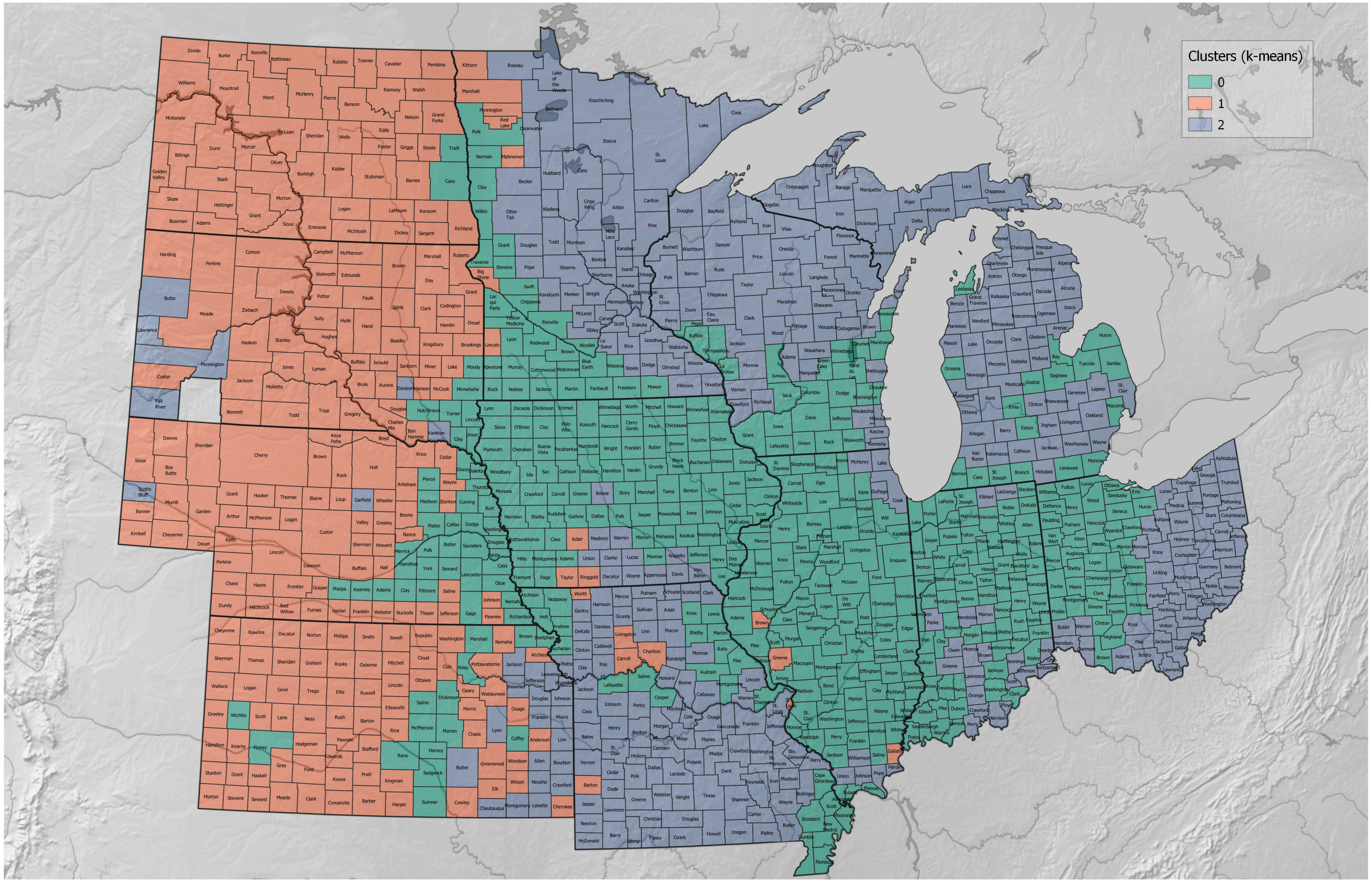

Figure S1. Clustering by *k*-means (*k*=3 clusters).

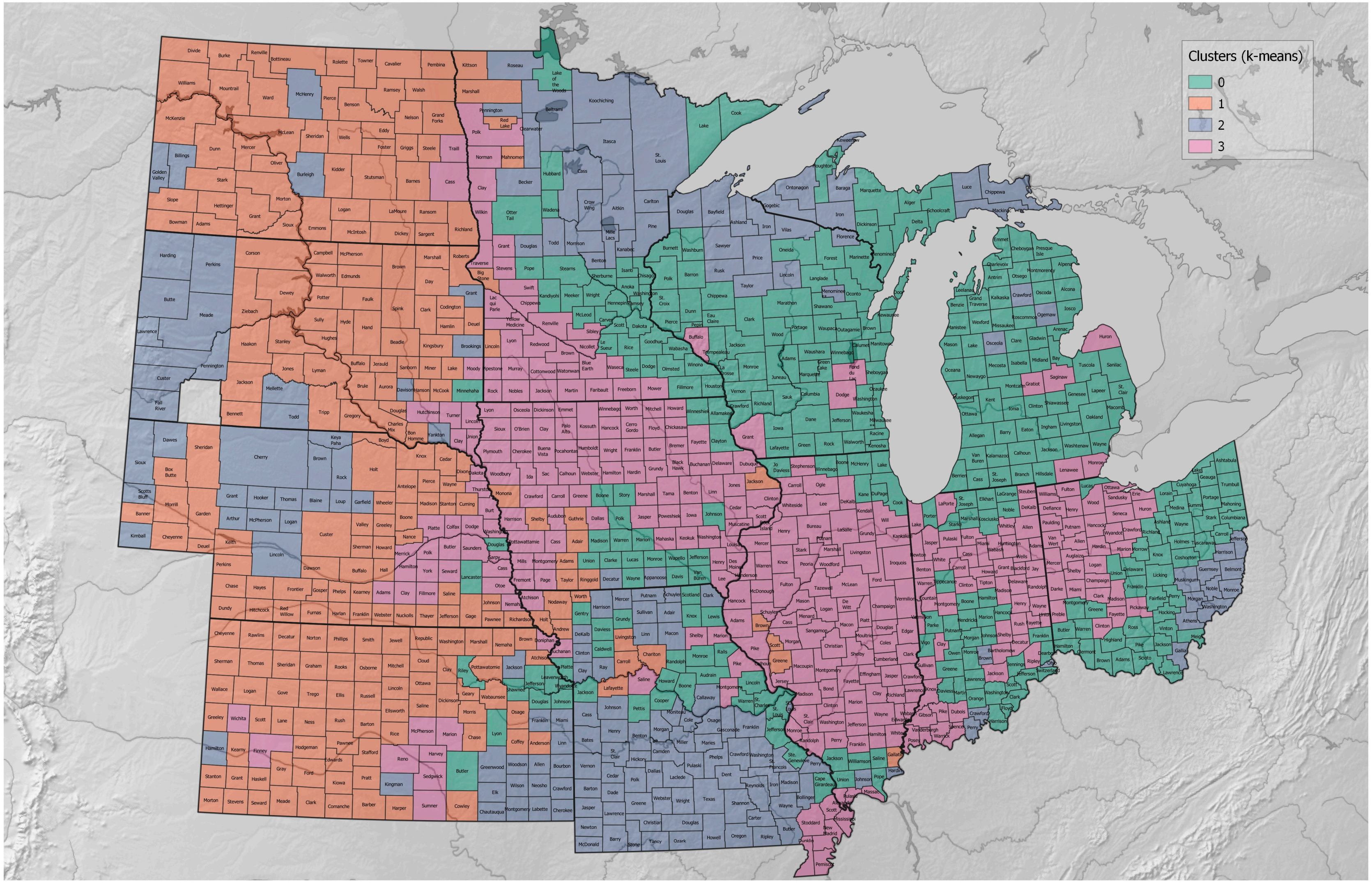

Figure S2. Clustering by *k*-means (*k*=4 clusters).

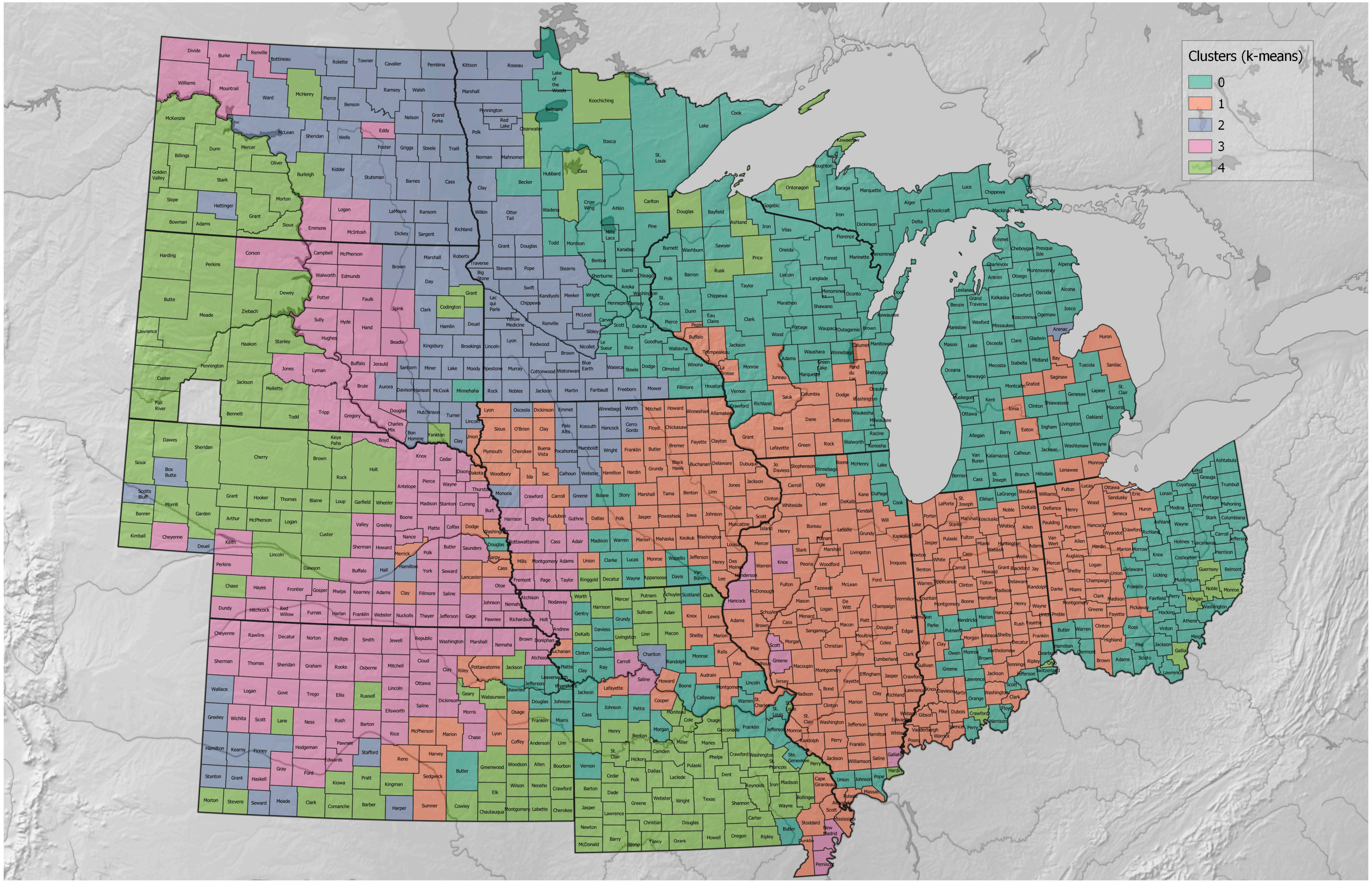

Figure S3. Clustering by *k*-means (*k*=5 clusters).

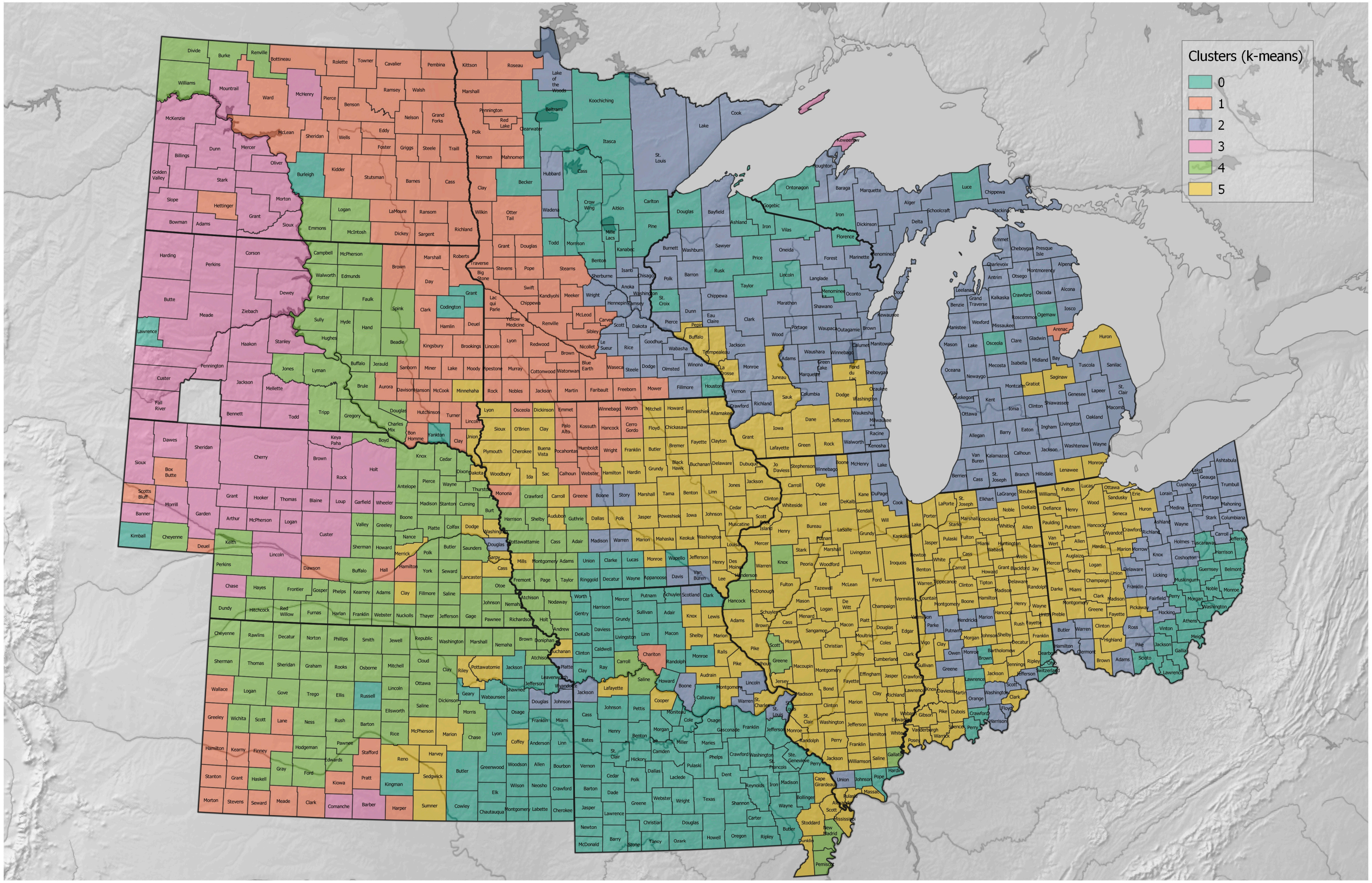

Figure S4. Clustering by *k*-means (*k*=6 clusters).

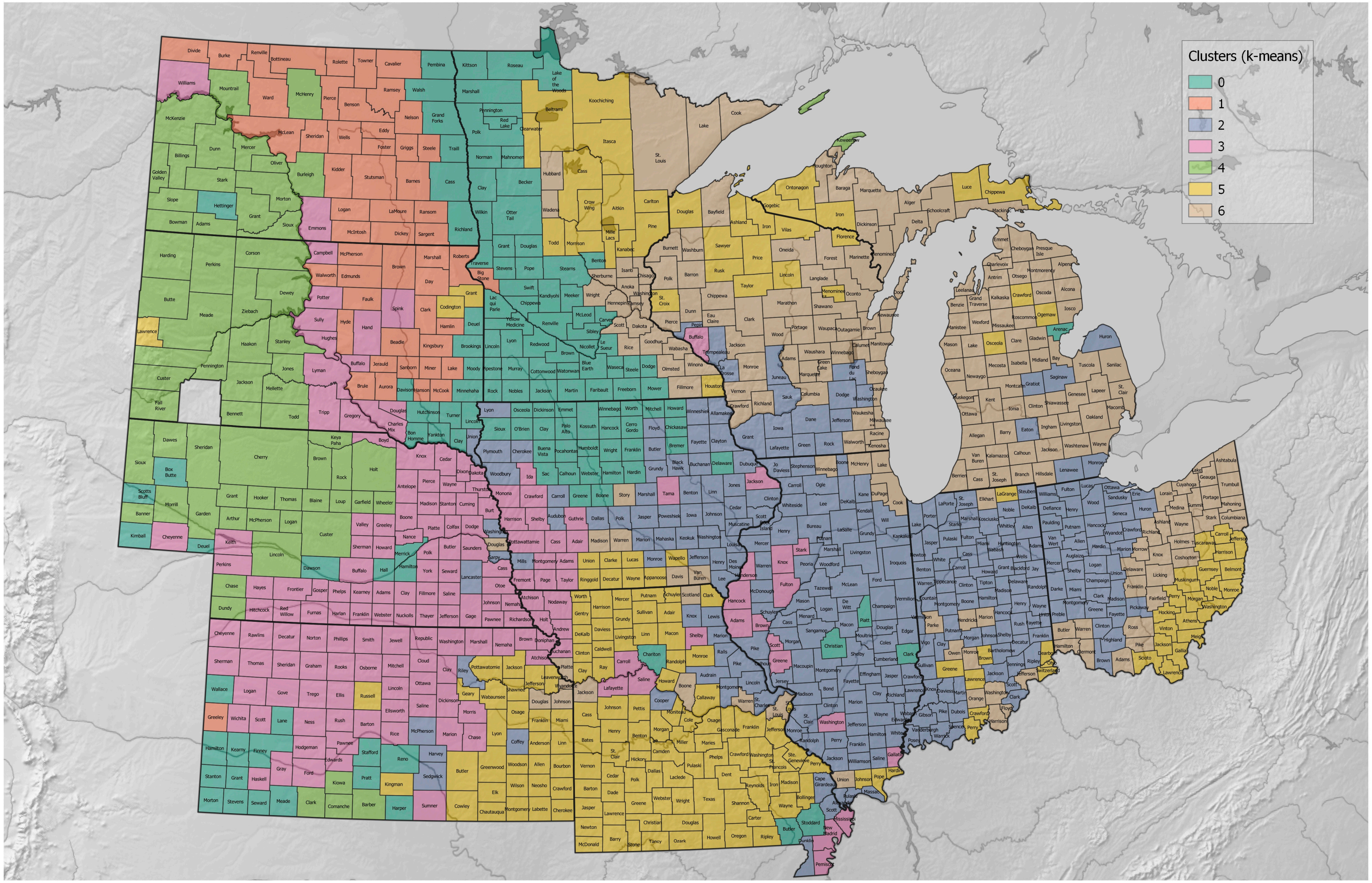

Figure S5. Clustering by *k*-means (*k*=7 clusters).

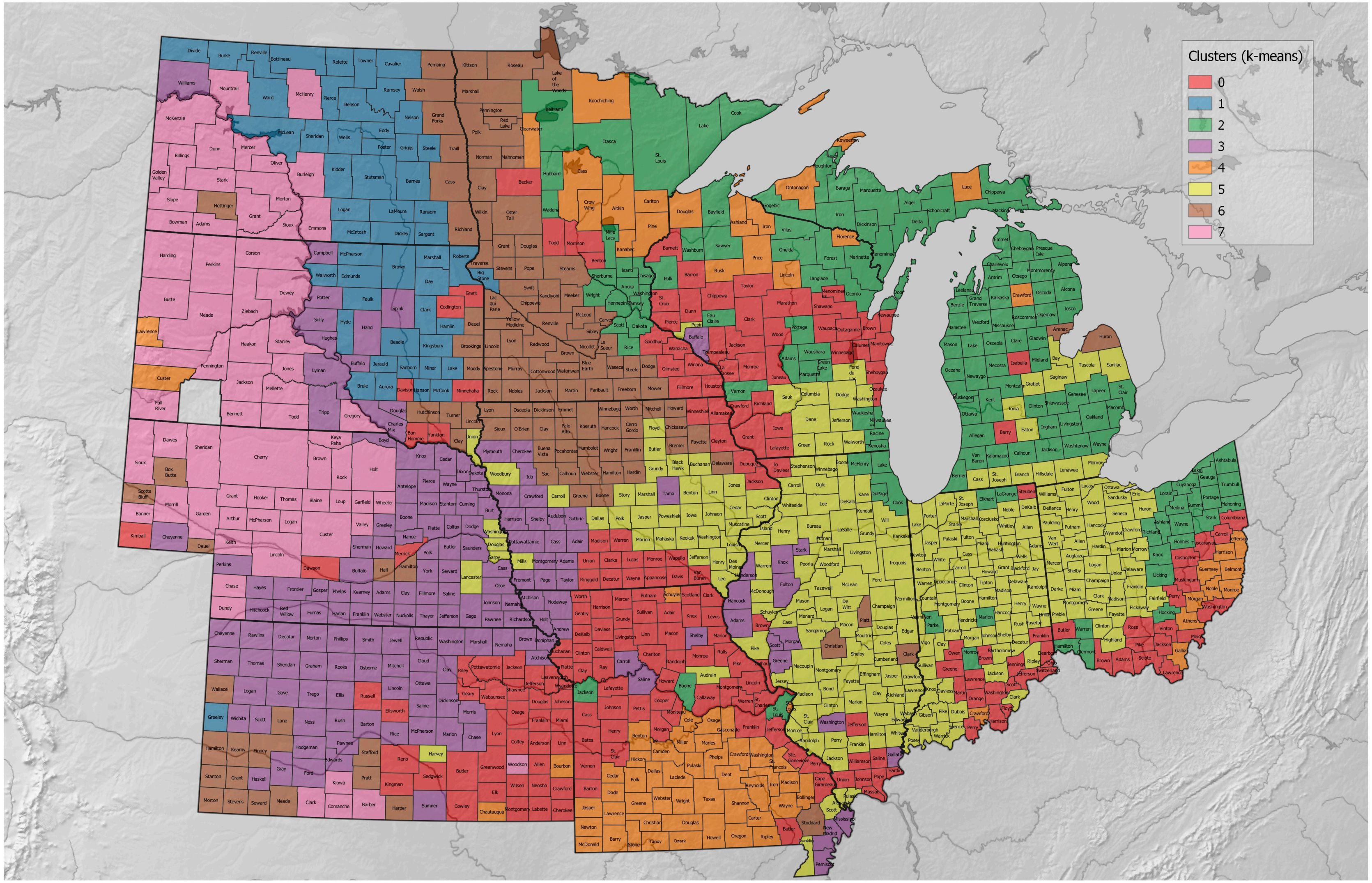

Figure S6. Clustering by *k*-means (*k*=8 clusters).

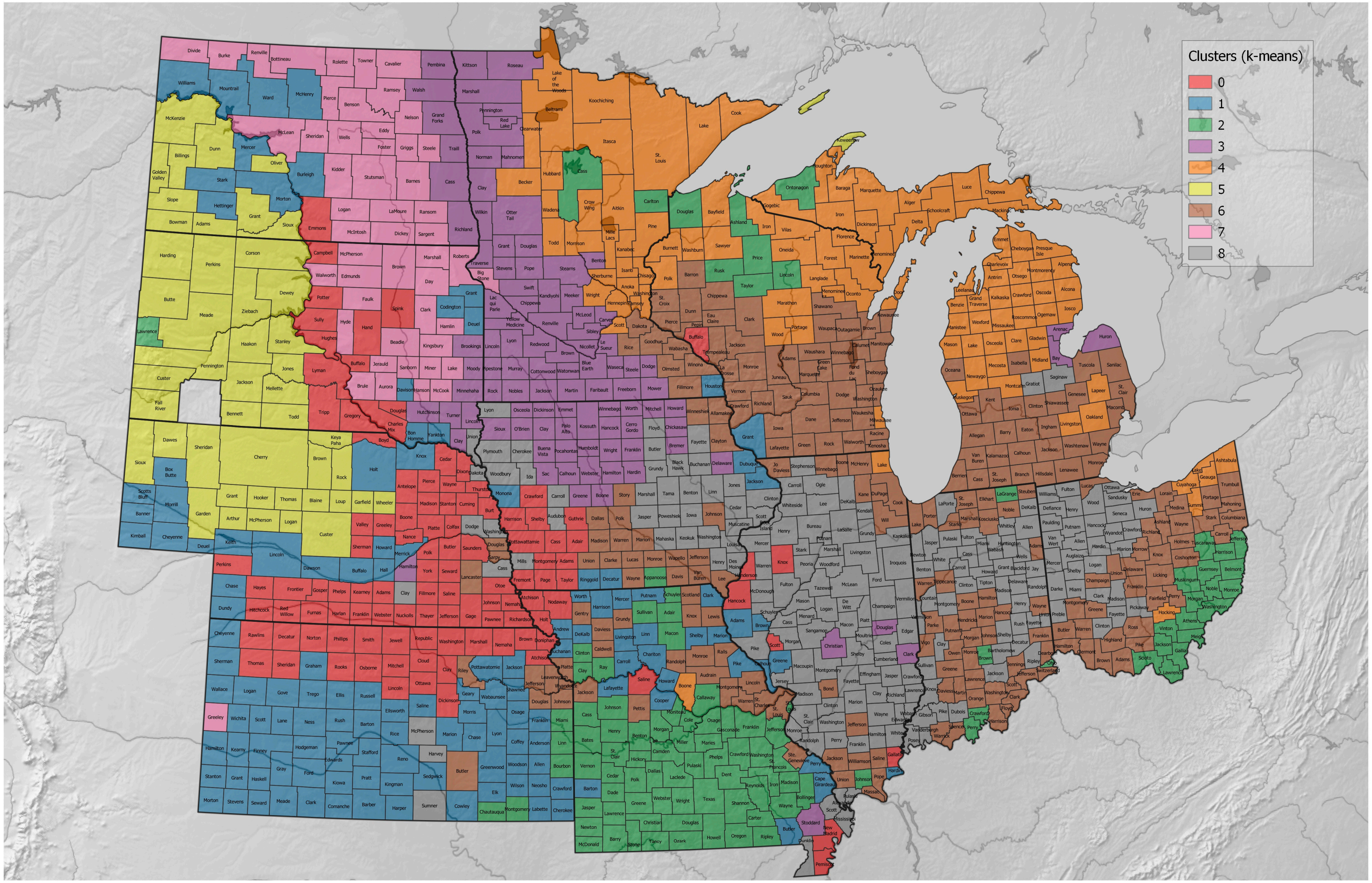

Figure S7. Clustering by *k*-means (*k*=9 clusters)   [Figure included for convenience; same as Figure 4 in the article main text.].