Map Making for Social Scientists

American Sociological Association

Anaheim, 20 August 2001

Waldo Tobler

Professor Emeritus of Geography

University of California at Santa Barbara

http://www.geog.ucsb.edu

Some hot topics in contemporary cartography

Animation of geographical objects

Three dimensional visualization

Map making on the internet

Map generalization

I will emphasize three other subjects

Map projections

Dealing with aggregate data

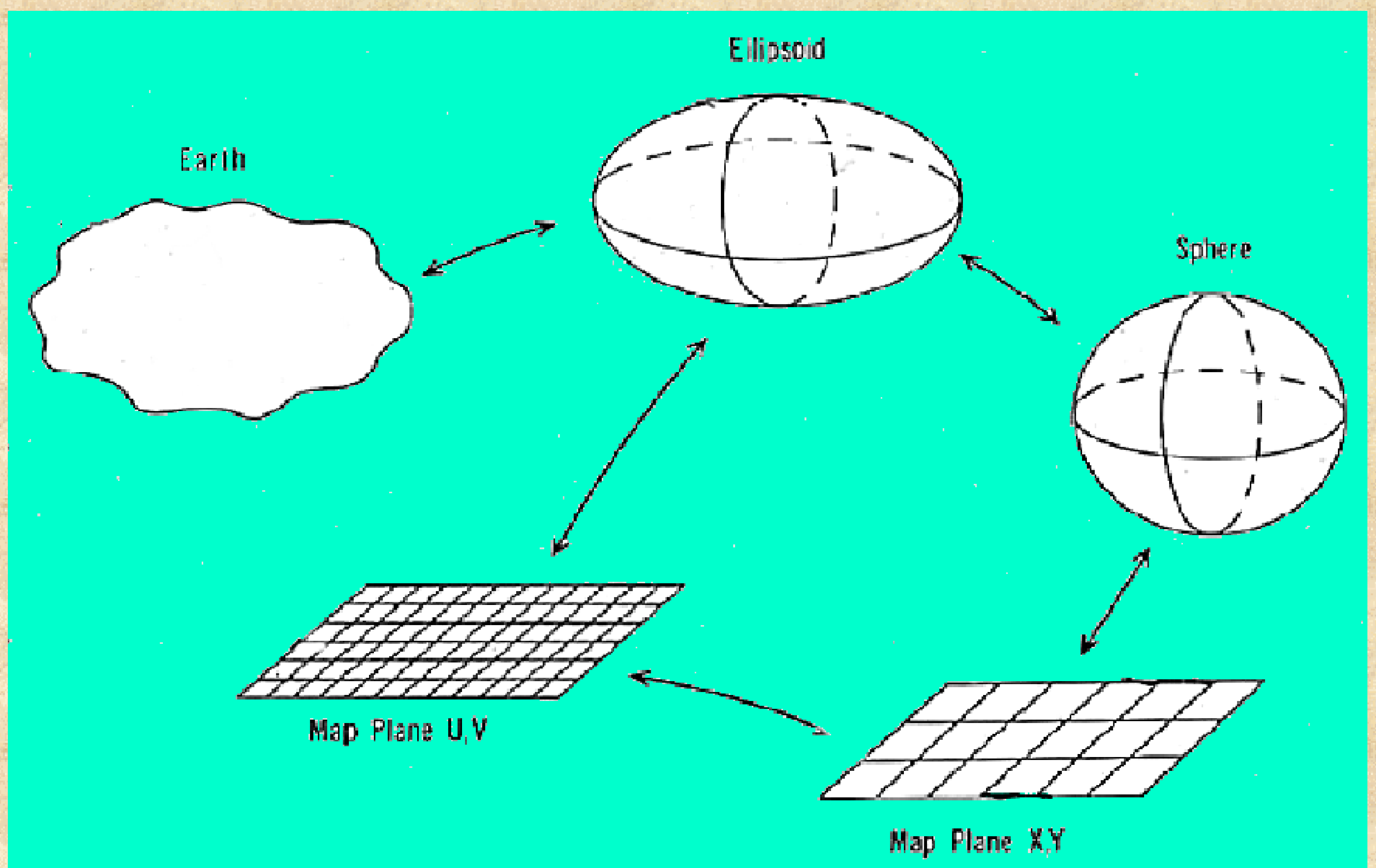
- Spatial filtering
- Estimating densities
- Converting to other units

Depicting movement

First, very quickly, map projections

The mapping process

Common Surfaces Used in Cartography



The <u>surface</u> of the earth is two-dimensional

Sphere or Ellipsoid?

The departure of the earth from a sphere is approximately one part in three hundred

This is 3/10^{ths} of one percent

This can be used as a rule of thumb:

Is your work accurate to better than one percent?

Sphere or Map?

This is equivalent to asking whether you want to work in latitude and longitude or in plane coordinates

Programs exist, for example, to convert from street address to lat/lon. There are also programs to convert from lat/lon to X, Y, and visa versa

Many kinds of analysis are very simple on a sphere

This includes such things as distance, direction, or area computation

A plane is a sufficiently good approximation to a sphere for a small area

What is small?

You can glue a postage stamp, without wrinkling it, on a 20 cm globe

Many analytical problems can be solved directly in geographic coordinates

This is often easy when the earth is considered spherical

It is more difficult to work in ellipsoidal coordinates

Some people like to work in plane, Euclidean, coordinates. Then a map projection is needed

Of course the projection must be suited to the problem, and there are many choices

Plane Coordinate Systems Are Based on Map Projections

The two most important ones are

The Universal Transverse Mercator system

The State Plane Coordinate system

The equations for both are complicated and based on an ellipsoid

Virtually all countries of the world have similar systems

All map projections result in distorted maps

Since the time of Ptolemy the objective has been to obtain maps with as little distortion as possible

Most geographic information systems and government mapping agencies take this point of view

But then Mercator changed this by introducing the idea of a systematic distortion to assist in the solution of a problem

Mercator's famous anamorphose helps solve a navigation problem

His idea caught on

Anamorphic projections are used to solve problems and are not primarily for display

One way to use map projections

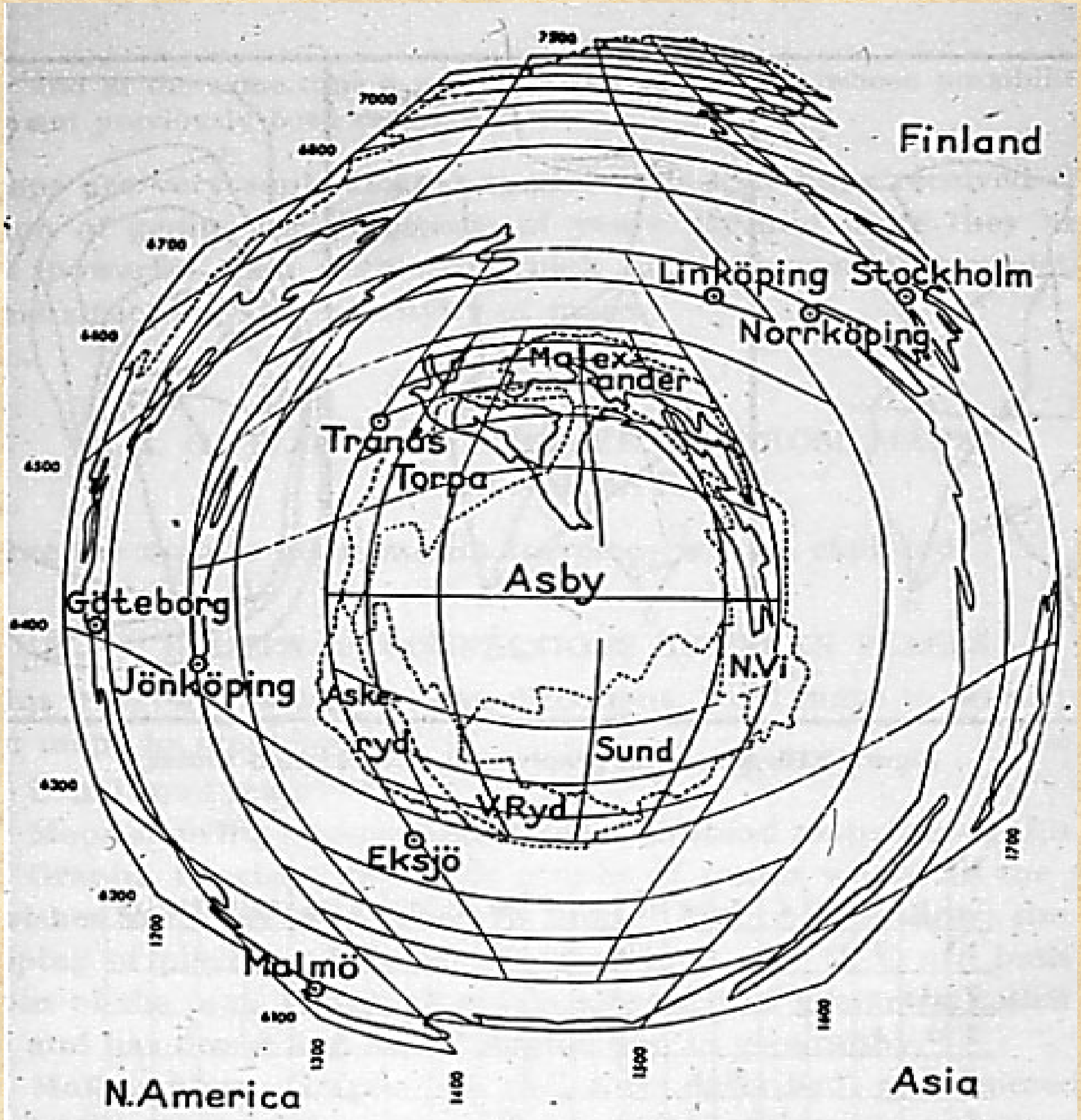
It is useful to think of a map projection like you are used to thinking of graph paper

Semi logarithmic, logarithmic, probability plots, and so on, are employed to bring out different aspects of data being analyzed

Map projections may be used in the same way

This is not a common use in geographic information systems

Hägerstrand's Logarithmic Map



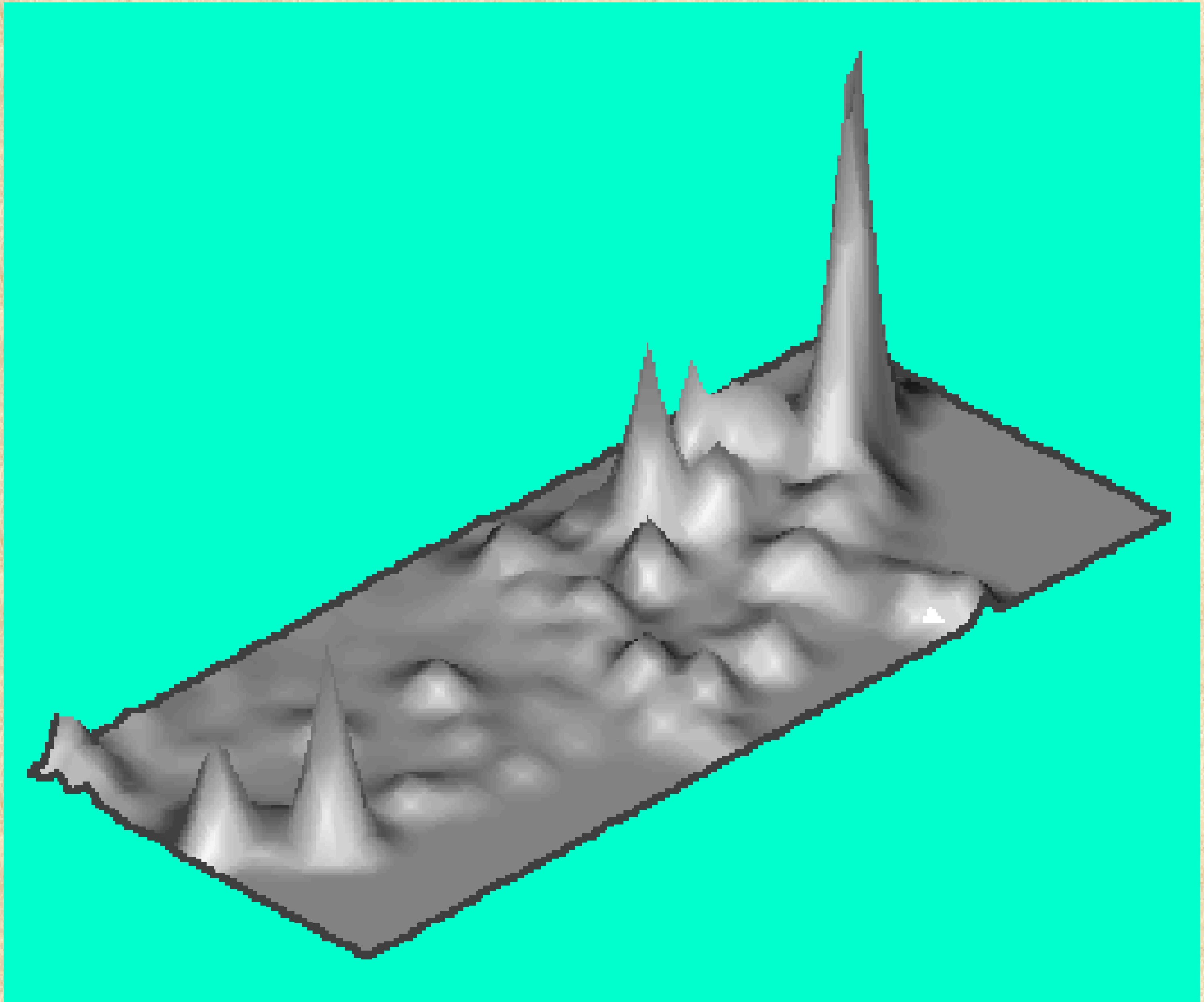
A map projection to solve a special problem

The next illustration shows the U.S. population assembled into one degree quadrilaterals

We would like to partition the U.S. into regions containing the same number of people

There follows a map projection (anamorphose) that may be useful for this problem

US Population By One Degree Quadrilaterals



Now Use the

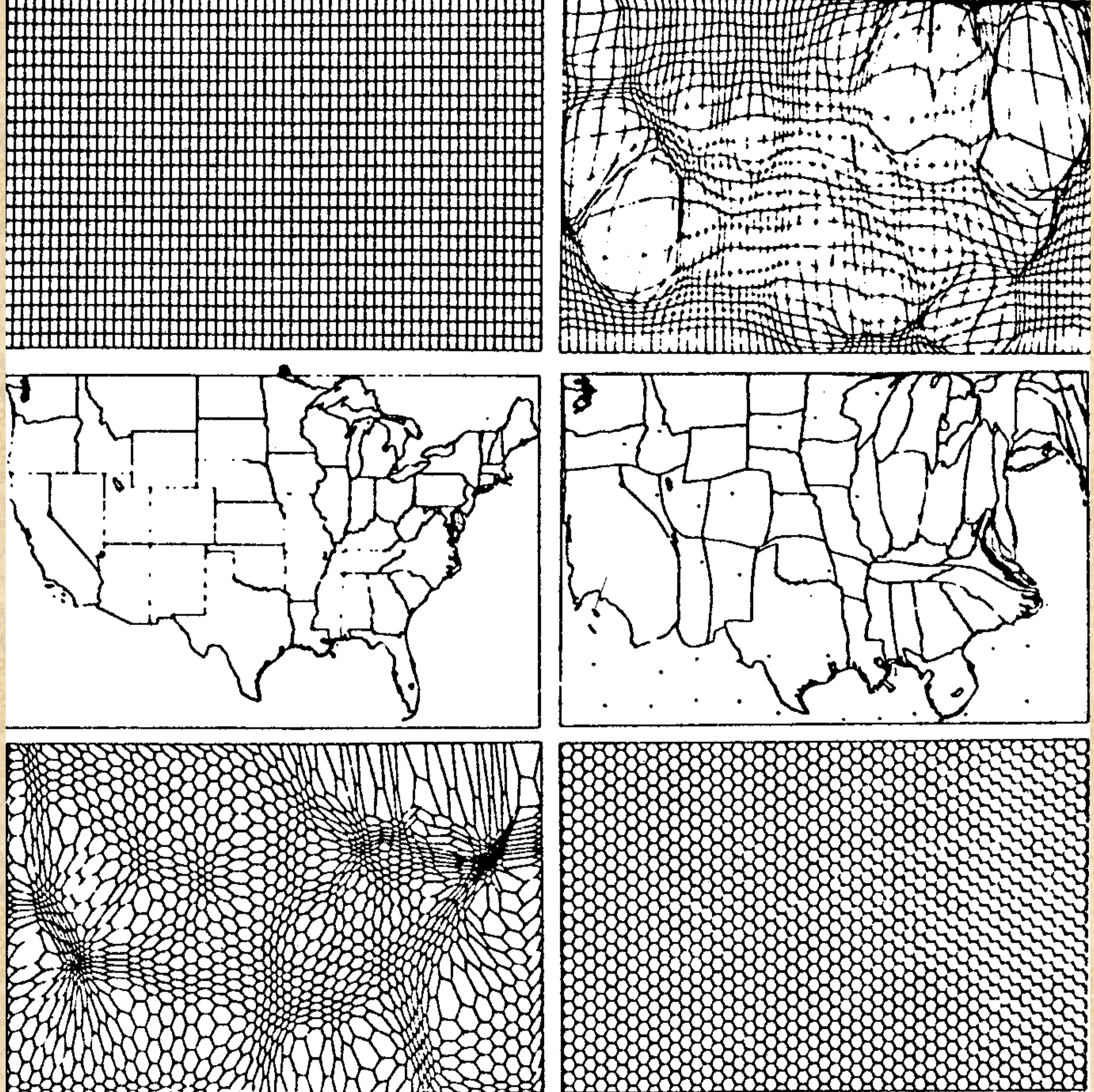
Transform-Solve-Invert Paradigm

Transform the graticule, and map, to obtain areas of equal population

Then position a hexagonal tesselation on the map

Then take the inverse transformation

W. Tobler, 1973, “A continuous transformation useful for districting”, A*nnals*, N.Y Academy of Sciences, 219:215-220



Next topic

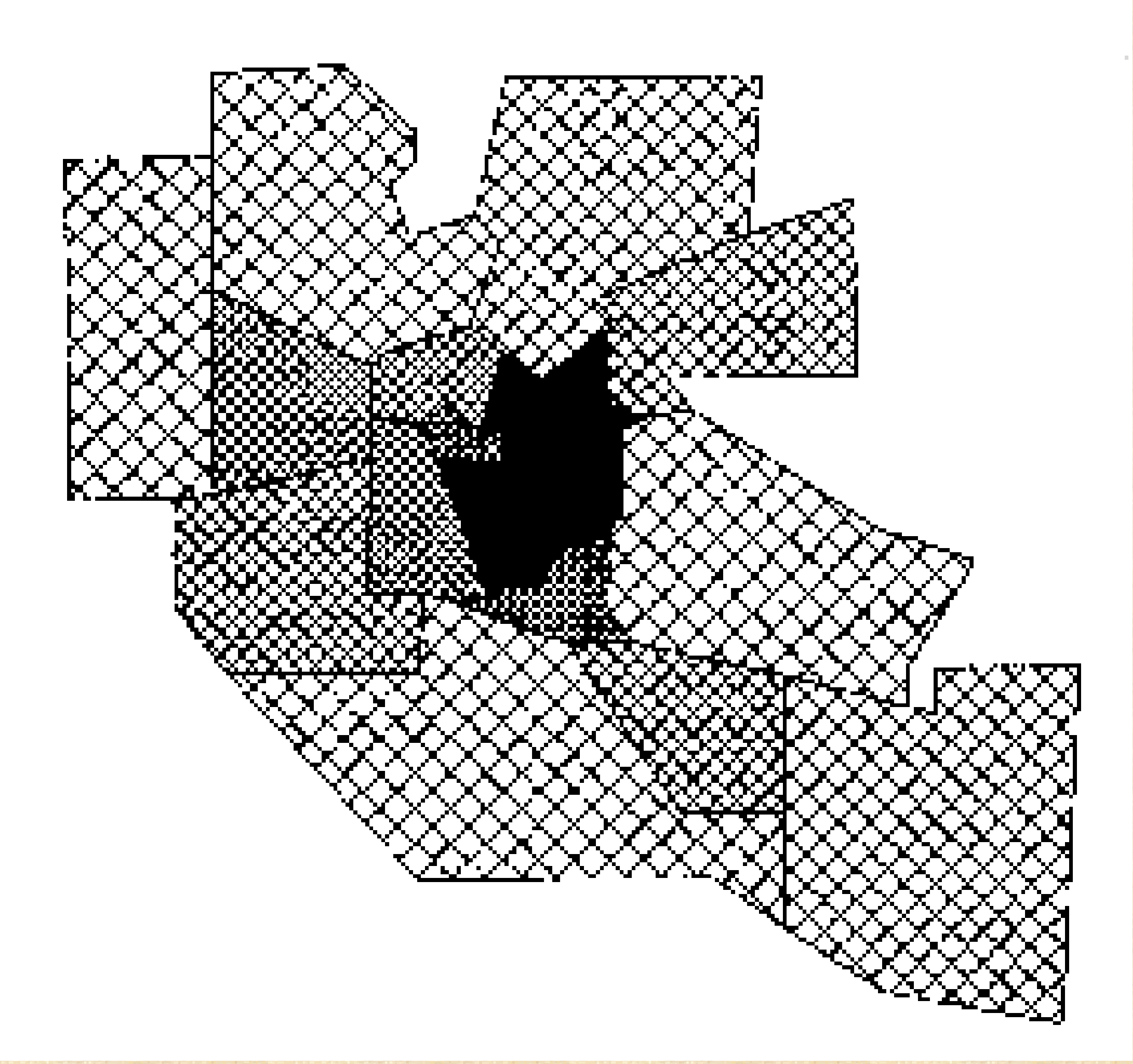
Often we deal with data given by areal units

Such as census tracts, counties, states, or other administrative units

It is convenient to think of the data as being binned into these spatial units in a manner similar to the making of histograms

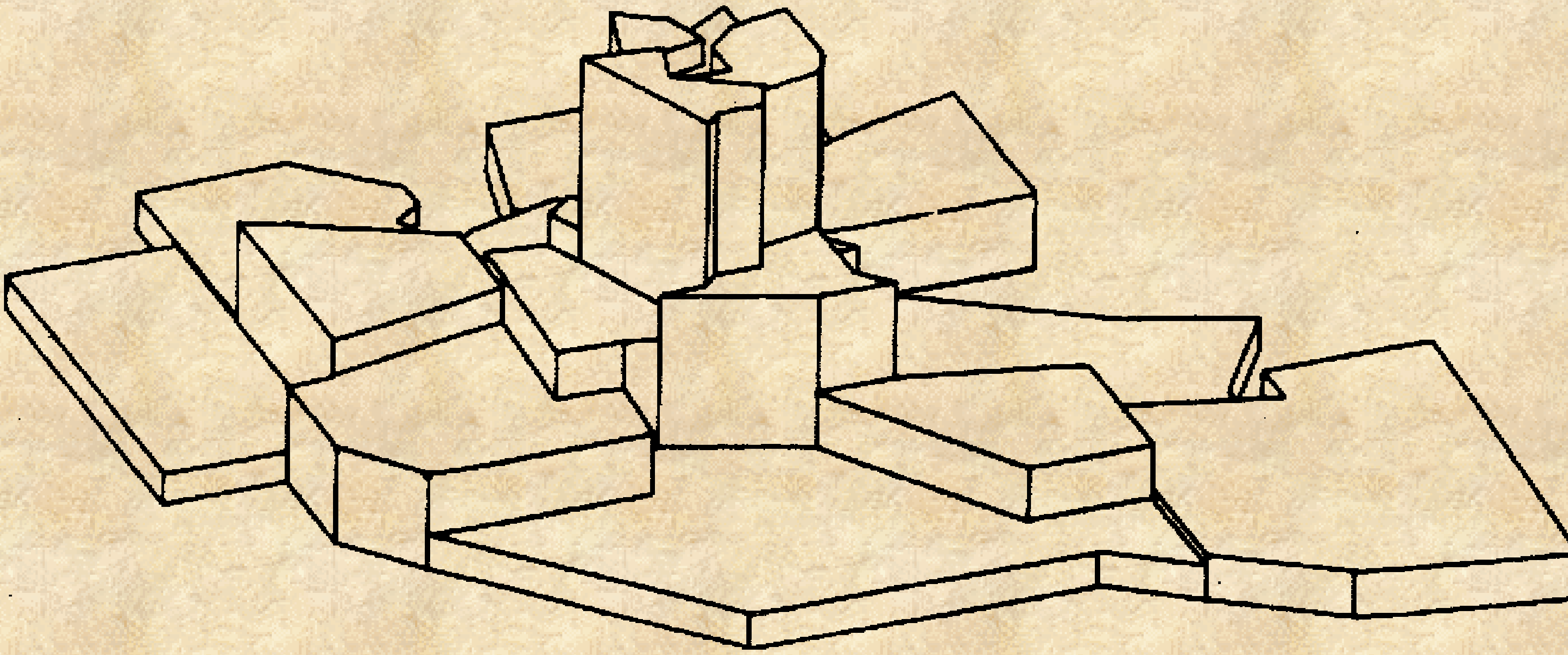
The difference is that the bins are of irregular sizes, shapes, and orientation on the surface of the earth

A choropleth (area filling) map
with shading proportional to density



The same data shown as a geographic bivariate histogram

with bin heights proportional to density



I will consider three problems relating to such binnings

1. The filtering of data in the irregular spatial units

including map generalization

2. Converting to continuous densities

3. Converting between areal units

Spatial filtering typically uses nearby, local, observations

Processing using neighbors is common in image processing.

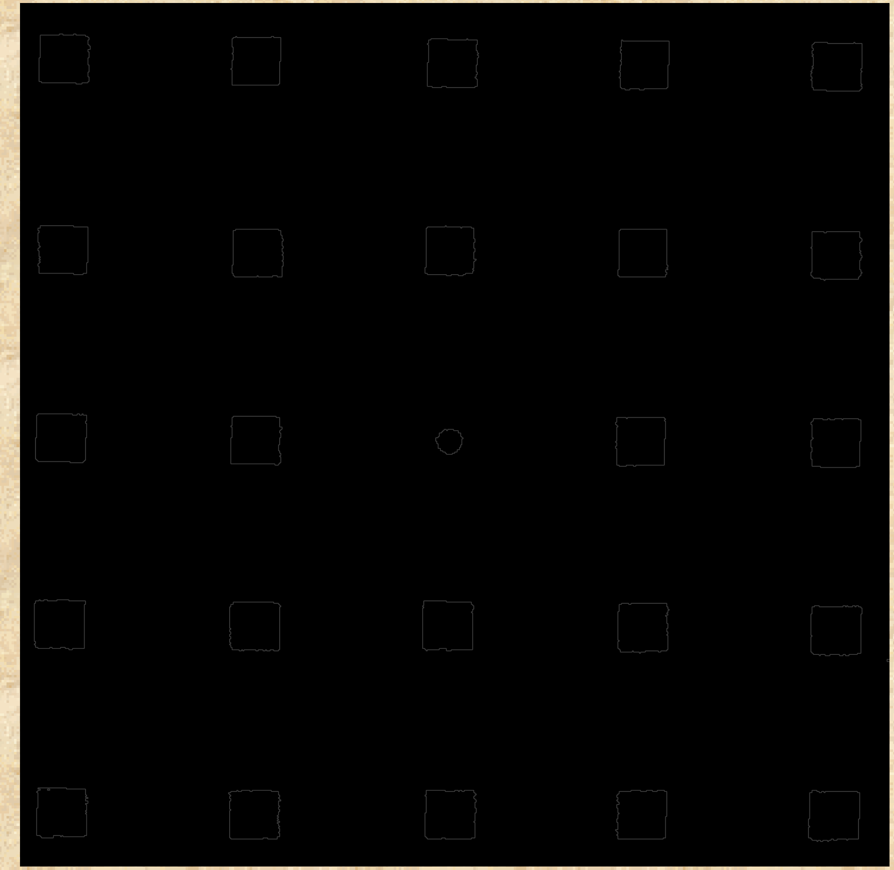
The value in a cell is converted to a weighted average of the values in neighboring cells.

Depending on the weights one obtains either smoothing (a.k.a. blurring) or sharpening.

Local geographic measures are similar in that they compute a value at each location that depends on nearby values. There are many examples.

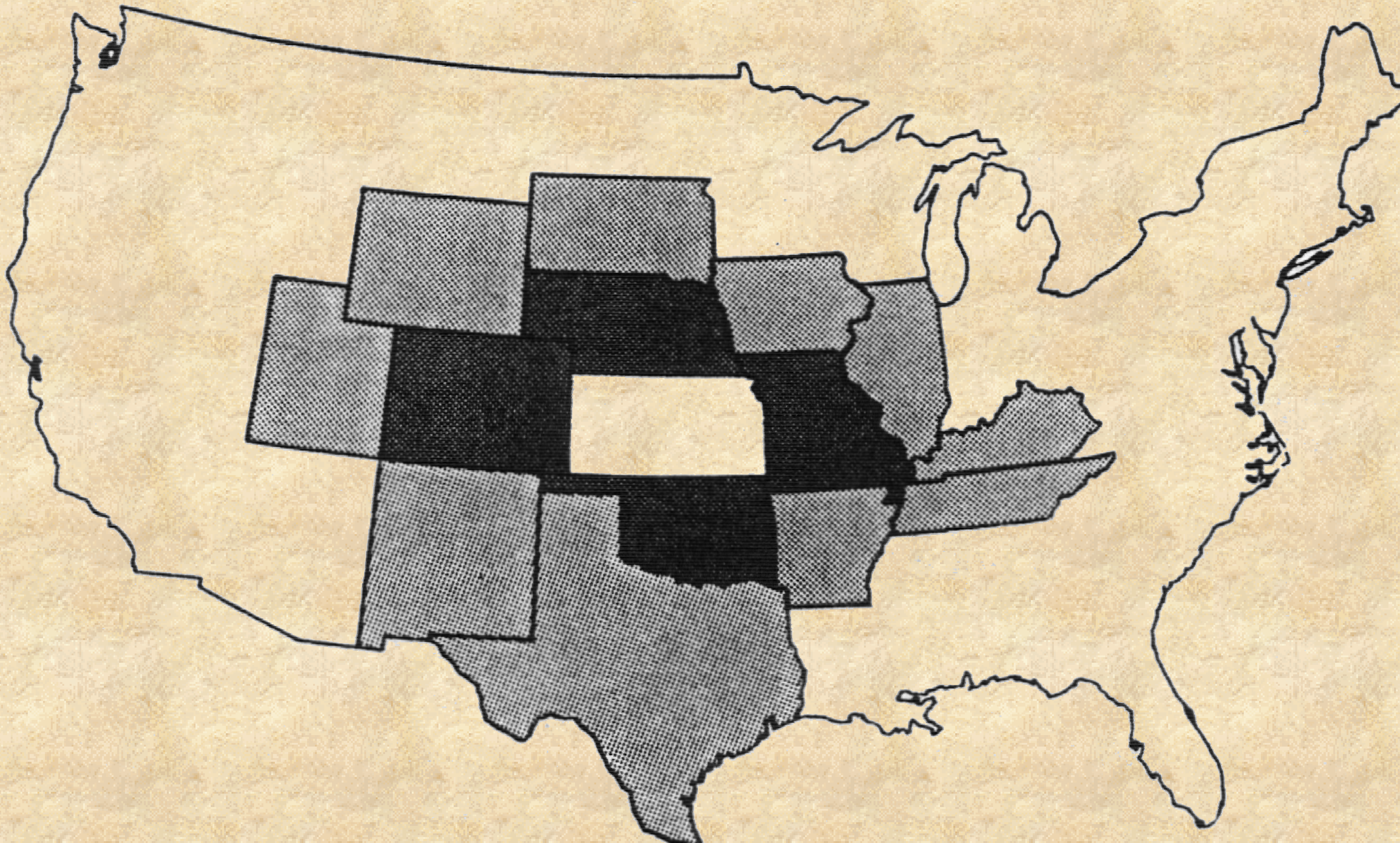
Modifying the center cell in the case of pixels

Neighborhood operators are used frequently in image processing



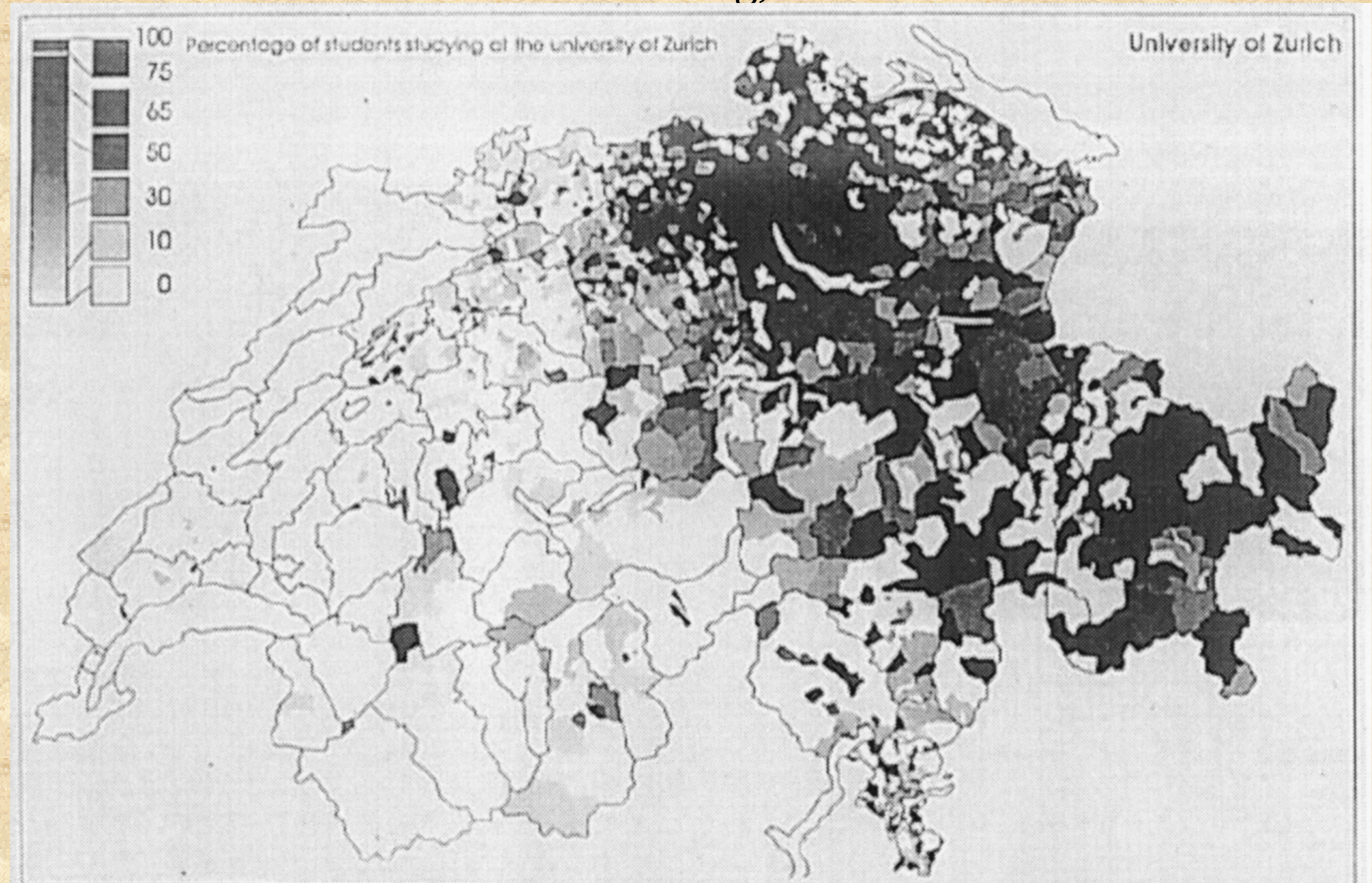
Neighborhood Operators Can Also Be Used With Resels

First and Second Order Neighbors of Kansas



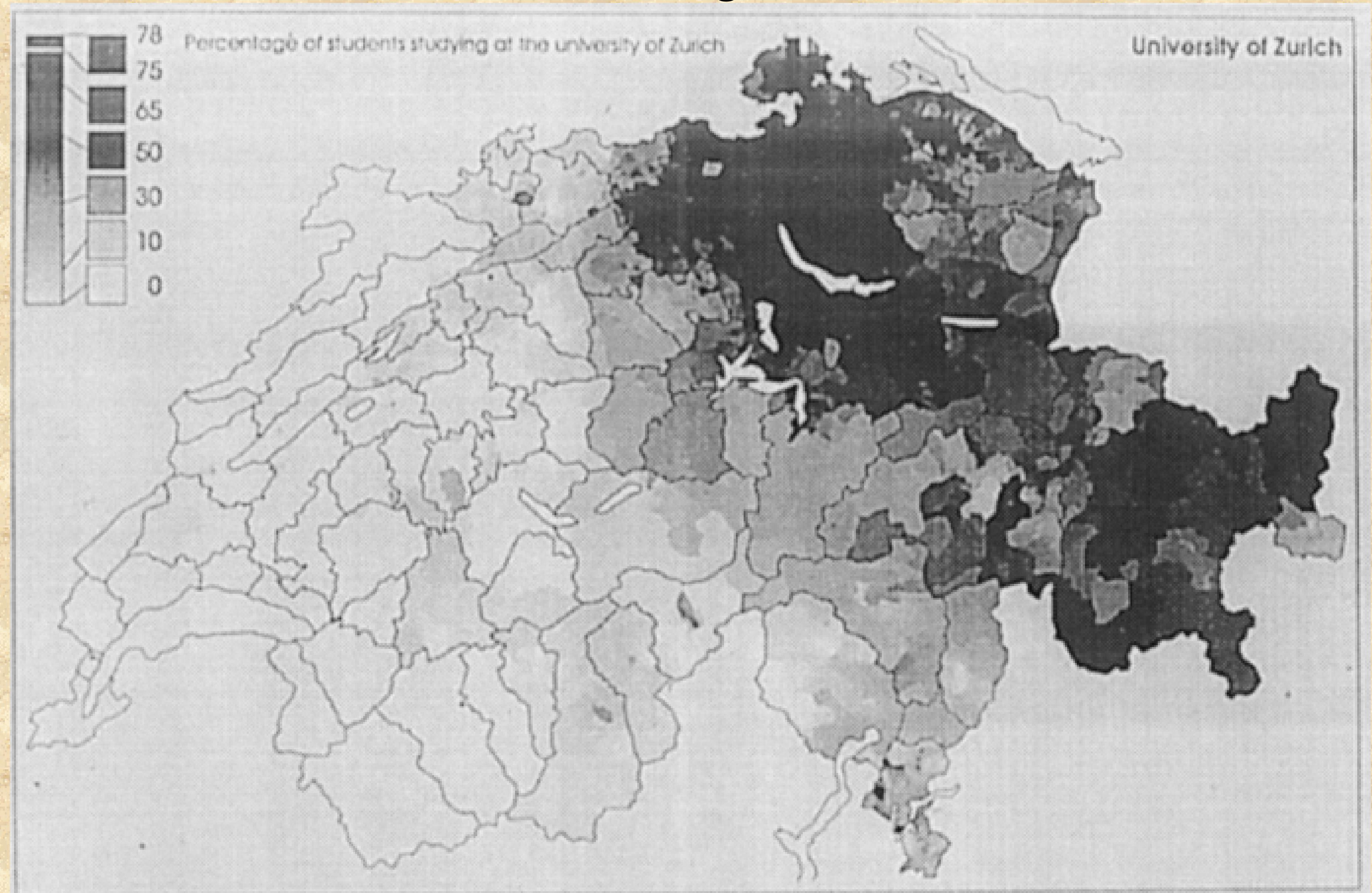
Choropleth map of university attendance

Adrian Herzog, Zürich



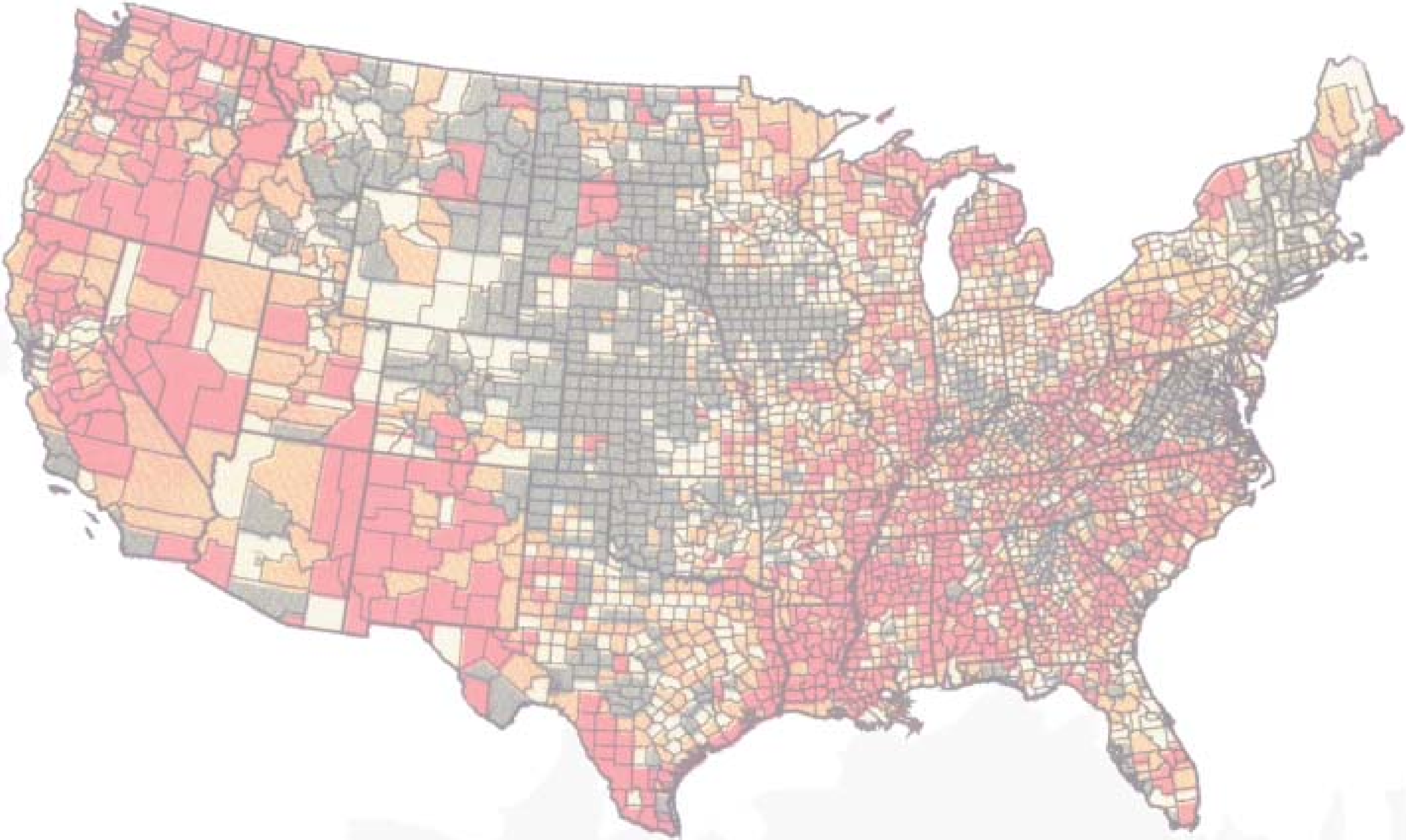
University attendance, adjusted

Adrian Herzog, Zürich



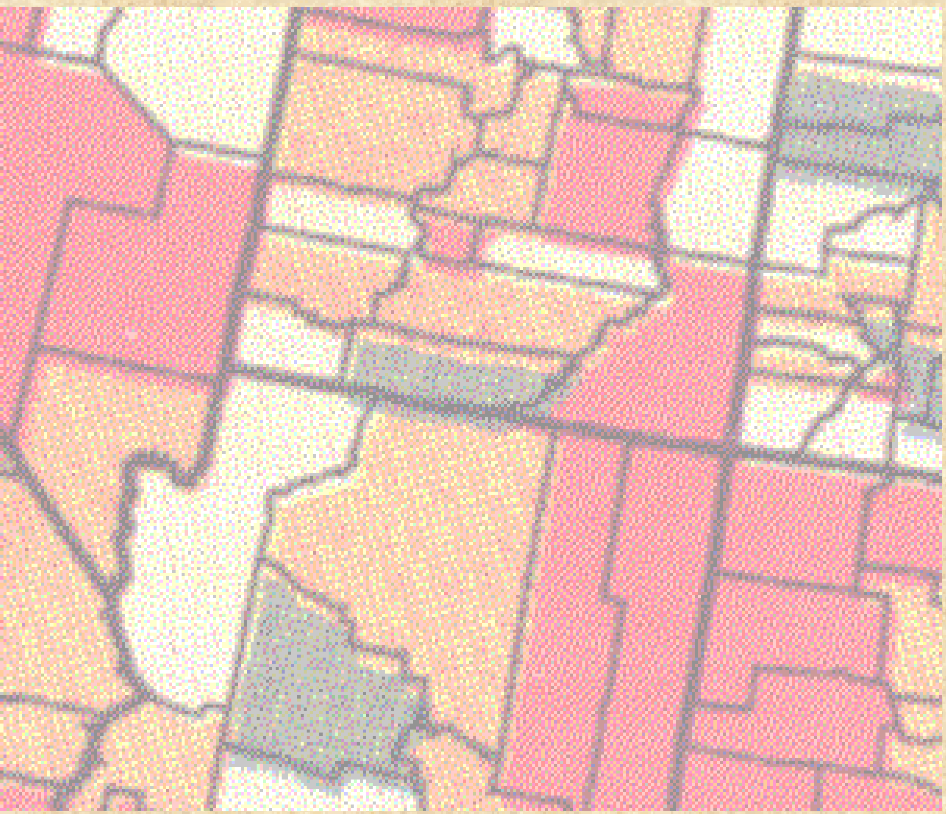
Unemployment, June 2001, by county

USA Today, 20 August 2001, page 4B



US unemployment map, two detail views

Brown: < 3.3%, Tan: 3.3-4.4%, Green: 4.5-6.2%, Red: >6.3%



Now a word about resolution

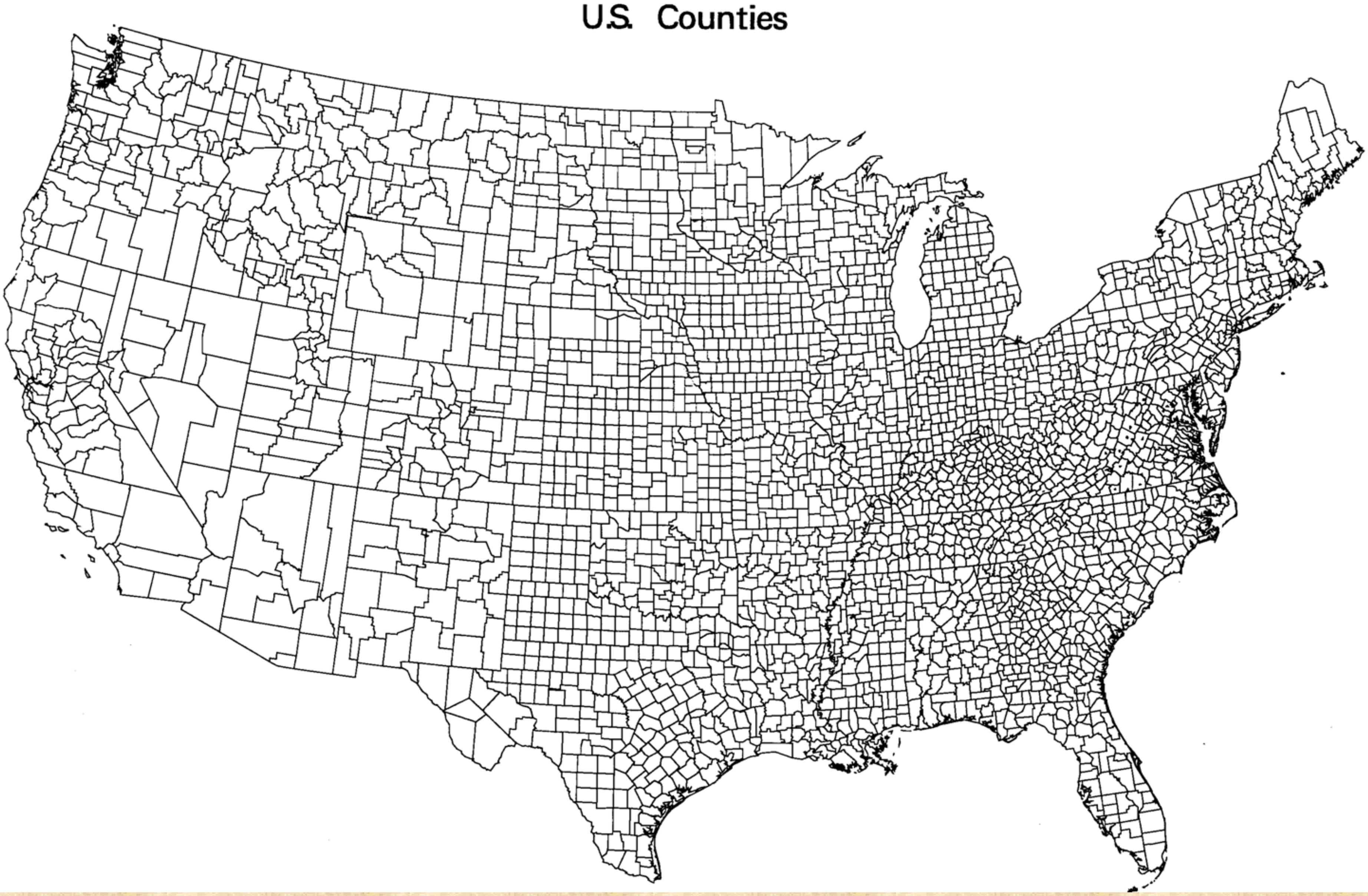
Average resolution can be calculated as

$$(\text{area of domain} / \text{number of observations})^{1/2}$$

In three dimensions use the cube root

In effect this measures the average distance influence of each observation

Unequal resolution in different parts of a map has an effect similar to unequal magnification in a microscope



Average resolution ~55 km. Patterns >110 km detectable

In these resels the resolution varies across the US. Patterns within cities cannot be seen

Social data are often made available in a hierarchy of administrative units

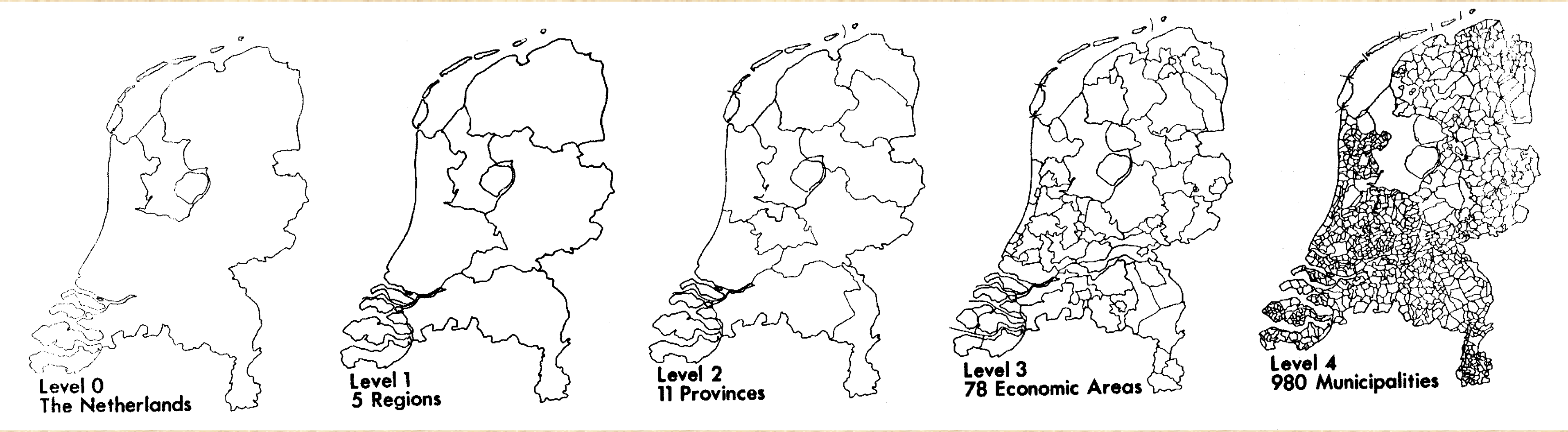
Moving up through the hierarchy changes the resolution and this acts as a low pass spatial filter

The result is a less detailed - more blurred - map

Consequently I recommend using the finest data available

For example

The Dutch administrative hierarchy



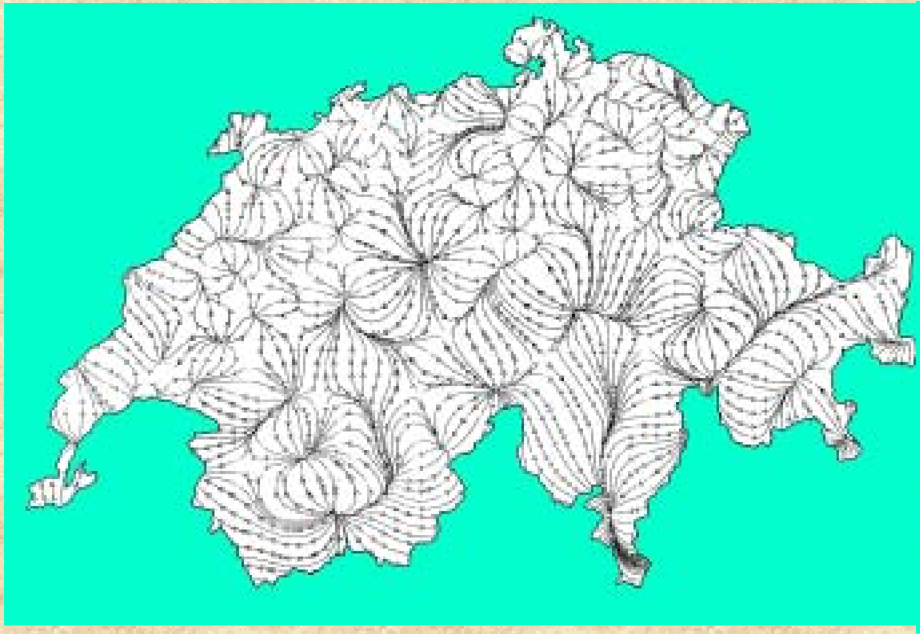
Swiss migration at reduced resolution

To emphasize the filtering effect of resolution

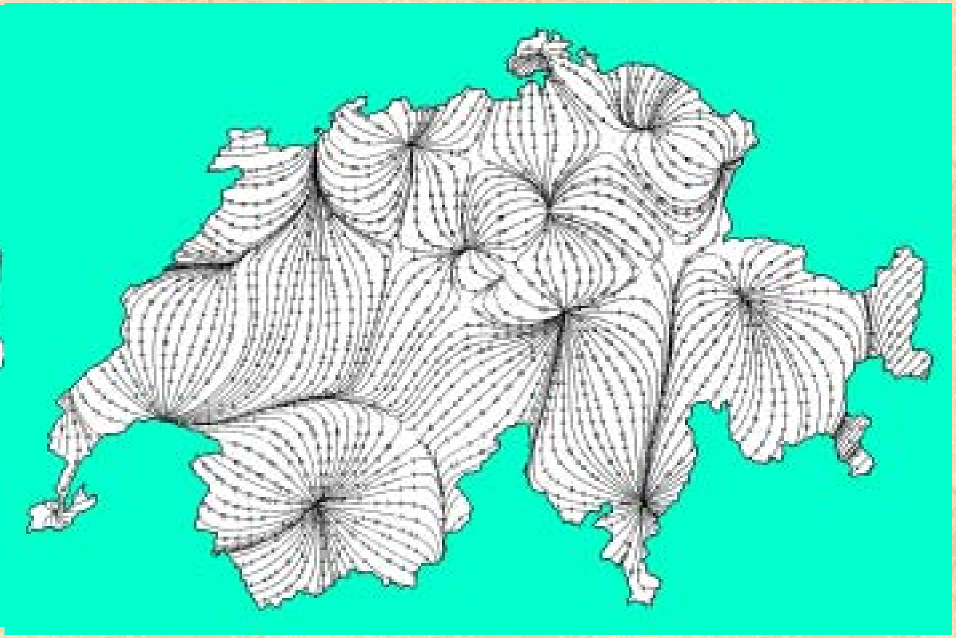
Another type of map generalization

Courtesy of Dr Guido Dorigo, University of Zurich

14.7 km resolution (184 Districts)

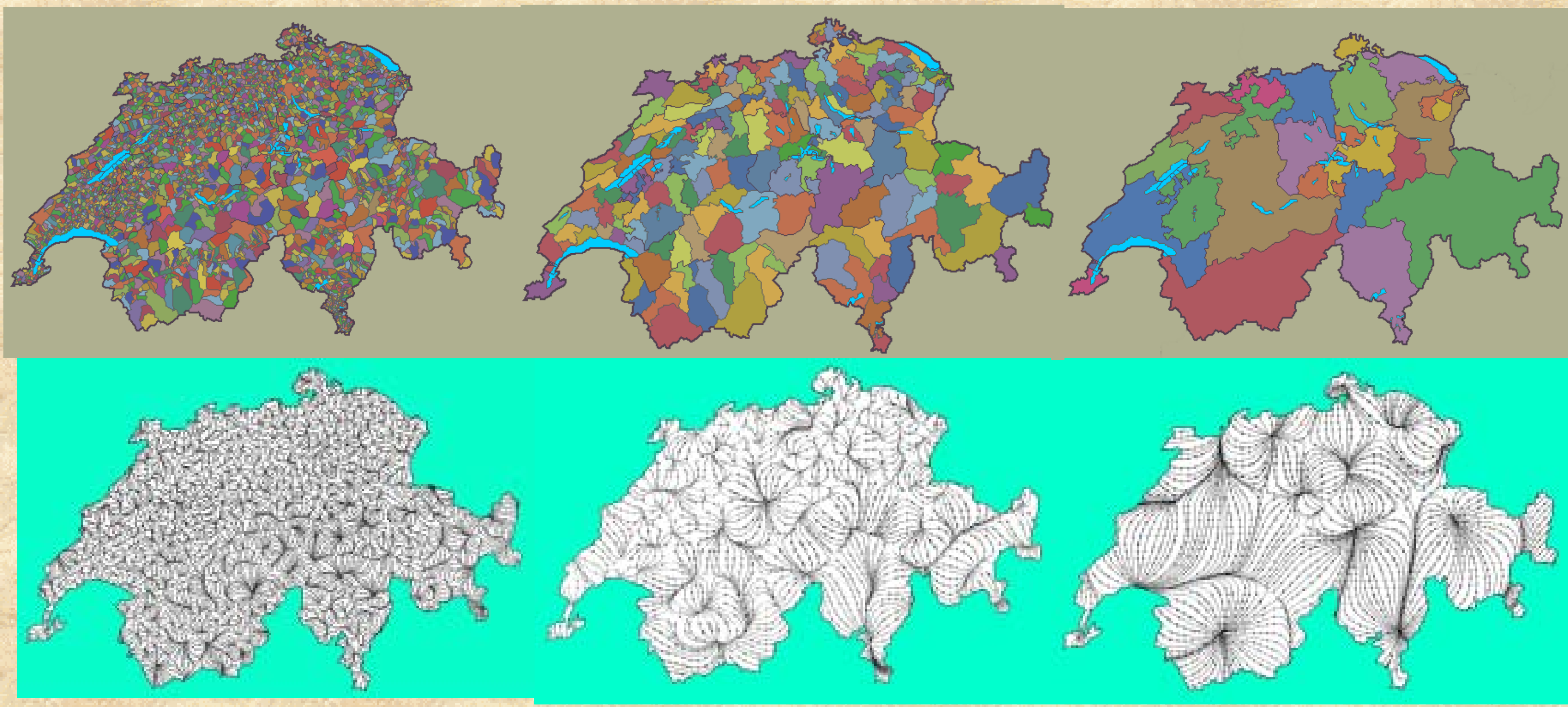


39.2 km resolution (26 Cantons)



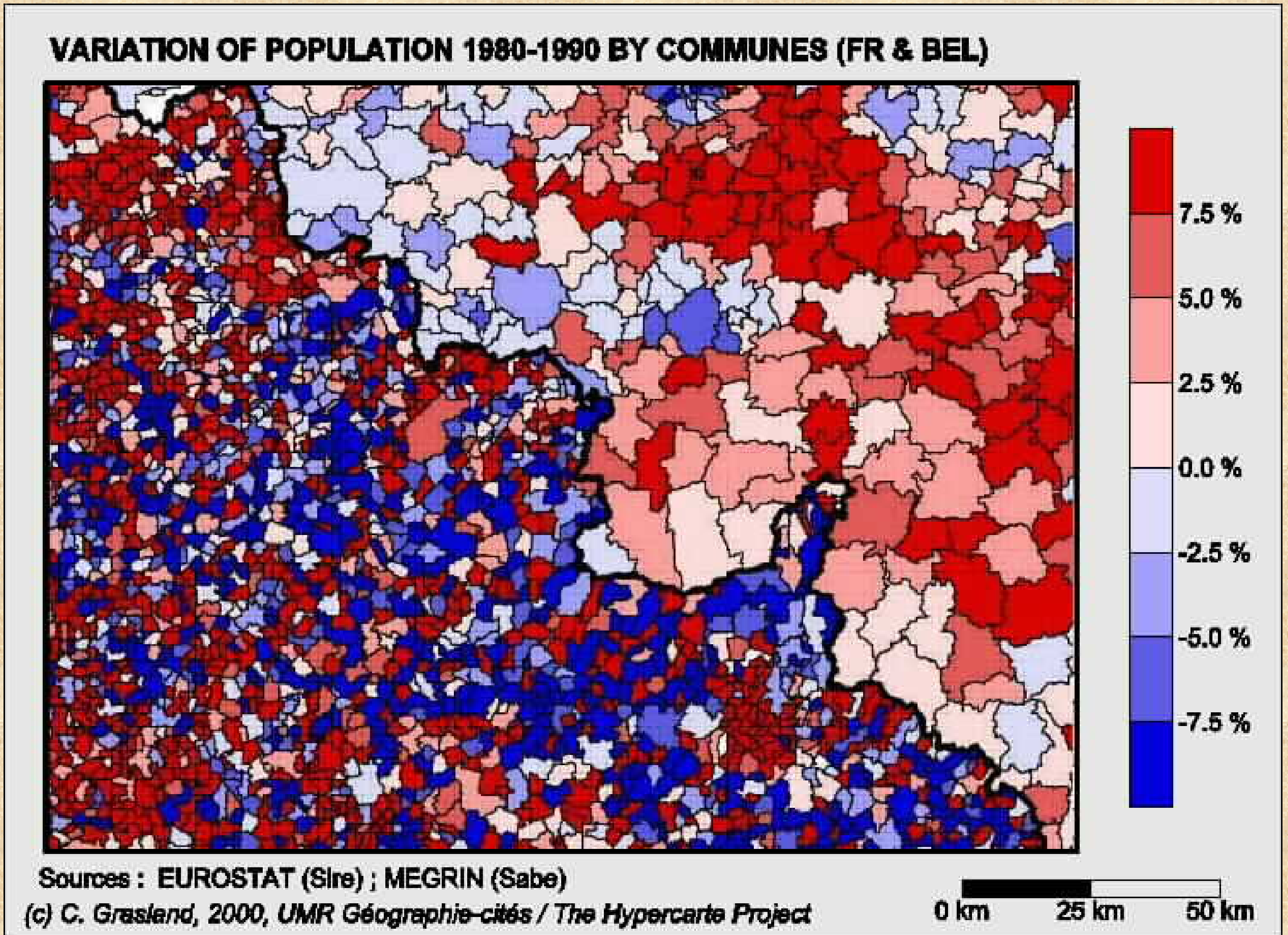
Three levels of administrative units and three levels of migration resolution all at once.

Communities | Districts | Cantons



An across boundary problem

Courtesy of Dr. Claude Grasland, Paris



In order to “uniformize” the resolution the bins in France are aggregated up the political hierarchy

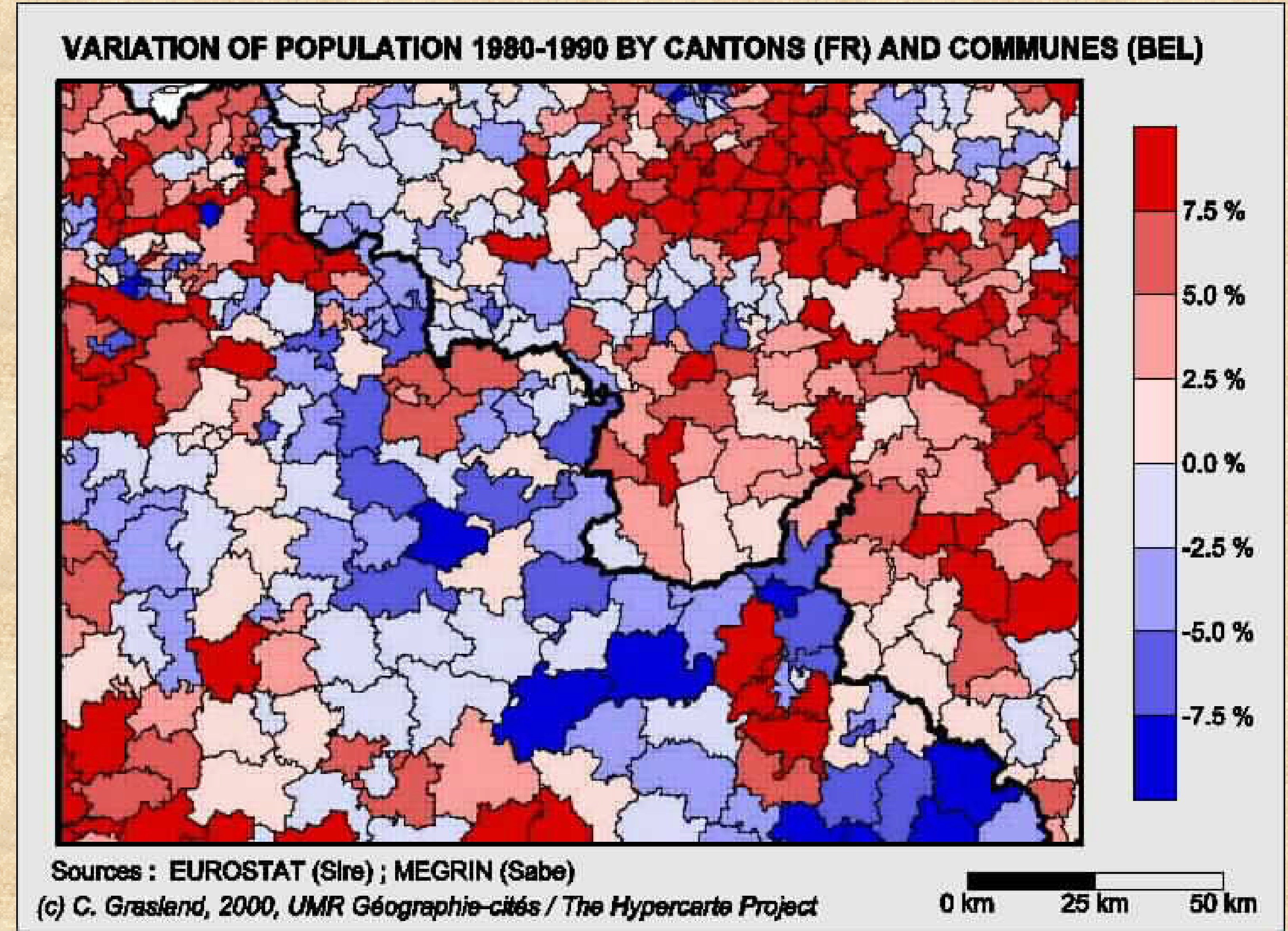
They then more nearly match the resolution of the Belgium information.

Had this not been done the resulting density for France would appear to have much more variability than that of Belgium.

But this variability would be an artifact of the difference in resolution.

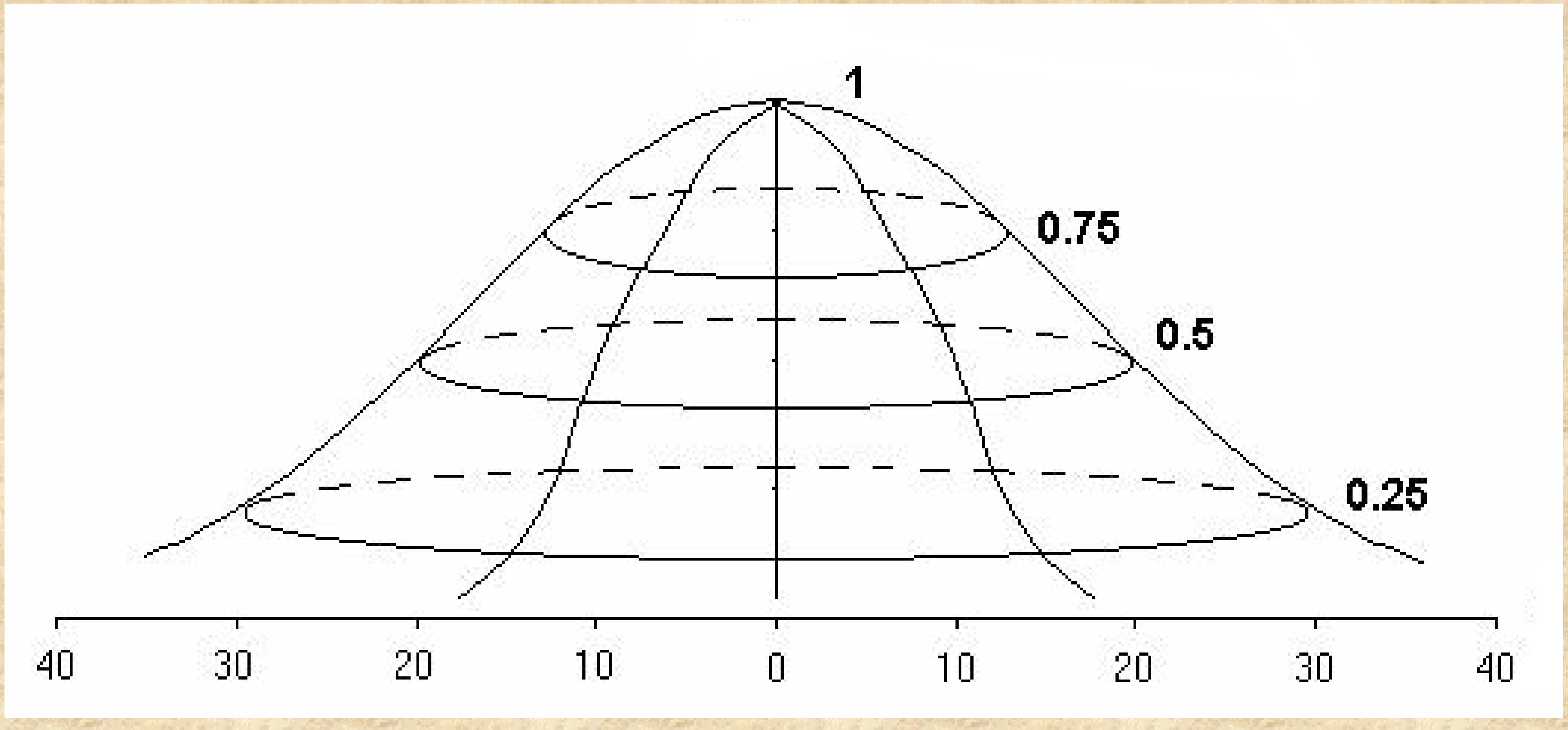
Population along the French-Belgium border

Courtesy of Dr. Claude Grasland, Paris



Conversion From Areal Units to Densities

A Gaussian kernel function



The data values are assigned to the centroid of the administrative units and then summed using weights taken from a sliding kernel function

How this works

Position the chosen kernel on the map

Search for all centroids within the kernel

Pick a weight from the kernel depending on the distance of the centroid from the map location

Multiply the value at a centroid by the kernel weight

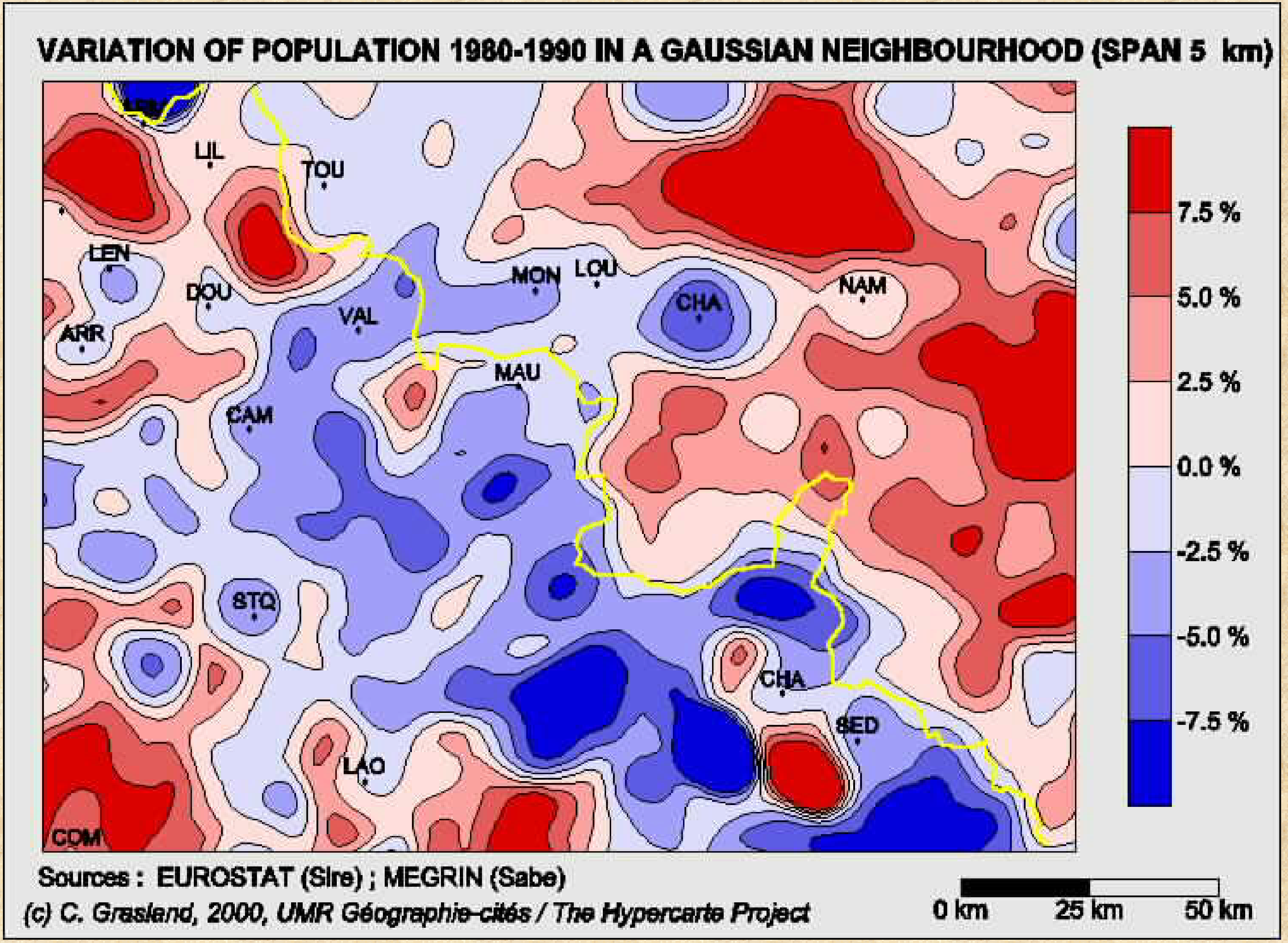
Sum all of the weighted values within the kernel

and assign this value to the location of center of the kernel

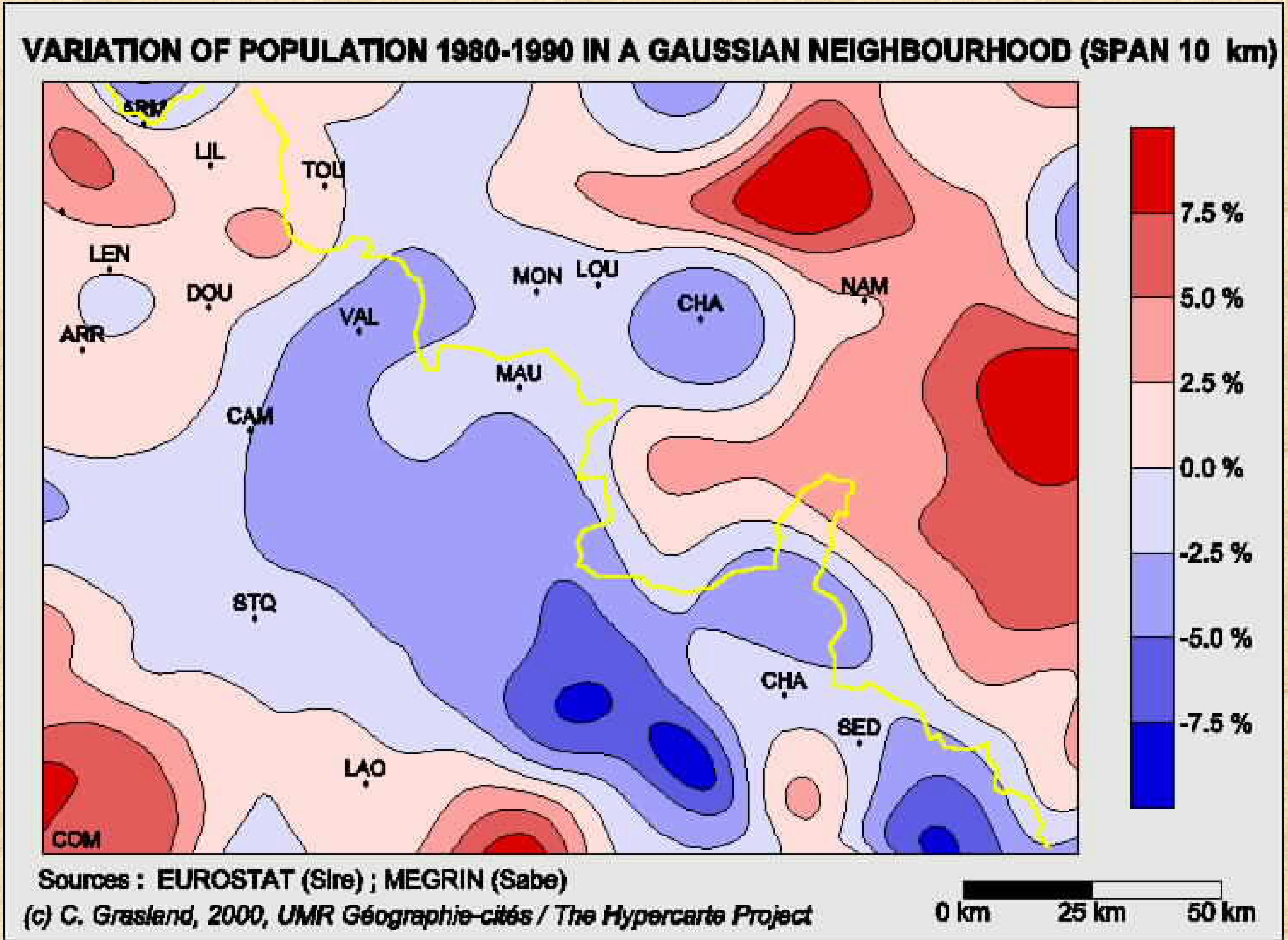
Move to the next location and repeat

After all locations have been evaluated you are done and can contour the results

Density based on a Gaussian kernel with a 5 km span



Using a Gaussian kernel with a 10 km span



Three references for further reading on density estimation techniques

D. Scott, 1992, *Multivariate Density Estimation*, J. Wiley, New York.

B. Silverman, 1984, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, New York.

R. Tapia & I. Thompson, 1978, *Non-Parametric Probability Density Estimation*, Baltimore, Johns Hopkins U. Press.

Kernels can also be applied to dot maps

Each dot is assigned a value of one unit

(dots with numerical values can also be used)

The distance of each dot from the center of the kernel is calculated

Then the dot values are modified by the kernel weight

The weighted values within the kernel are summed

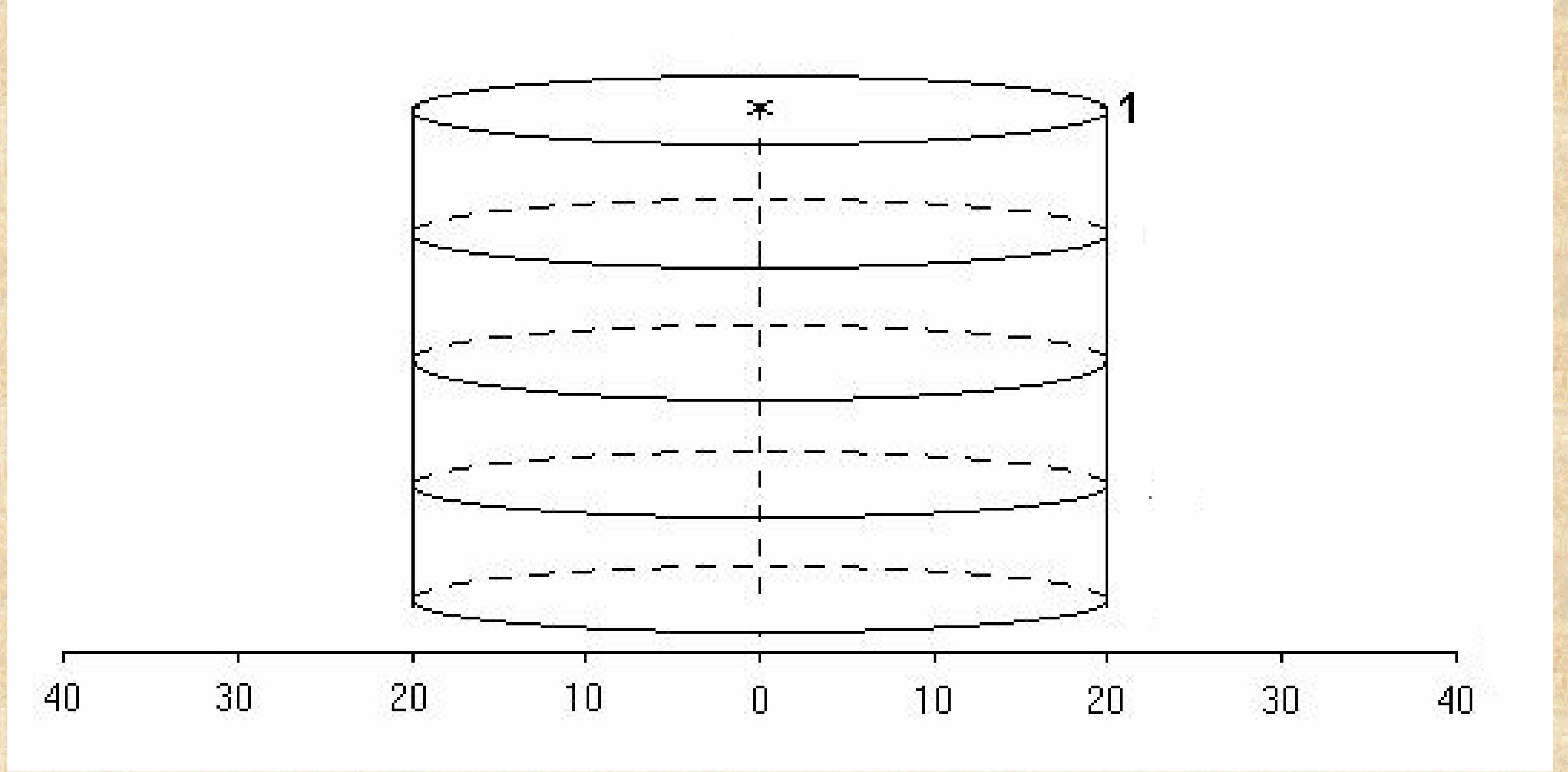
and assigned to the location of the kernel center

The map is complete when the sum has been calculated for all locations

Thus the dot distribution has been converted to a density map

A uniform kernel is often used

but is not recommended because of its effects



This kernel inverts some peaks and valleys.

See: J.Holloway, "Smoothing & Filtering of Time Series & Space Fields", *Advances in Geophysics*, 3 (1958): 351-389

There is also a method that avoids the use of kernel functions

It is sometimes referred to as areal interpolation.

From this point of view it is incorrect, in my opinion, to assign areal observations to points (centroids).

One criterion to be satisfied is that the resultant maintain the data values within each unit.

The method is known as pycnophylactic reallocation.

Pycnophylactic Reallocation
(Mass Preserving)

Allows the production of density or contour maps to be made from areal data.

It is reallocation - and somewhat of a disaggregation operator. My assertion is that it may actually improve the data.

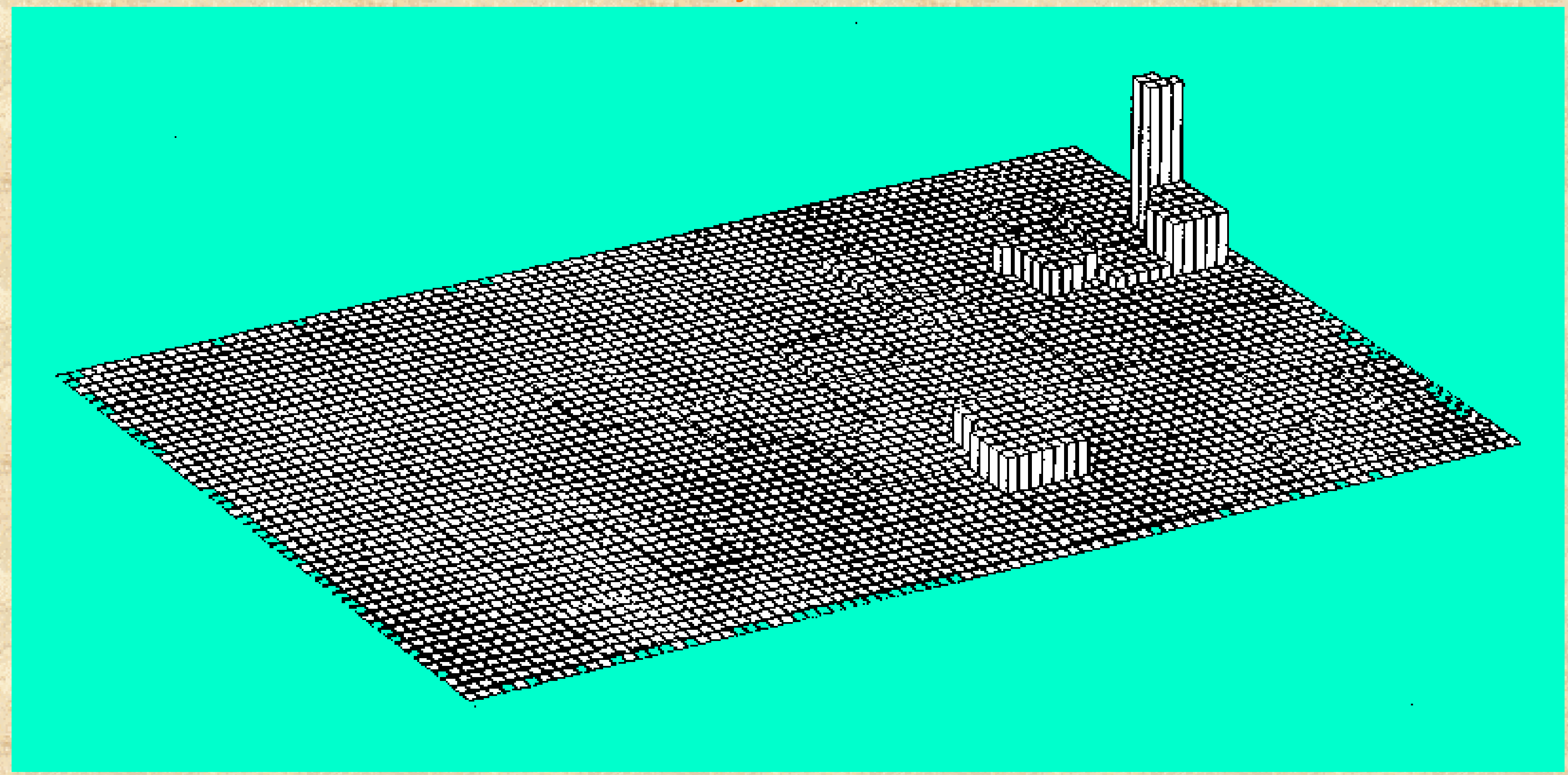
It is also important for the conversion of data from one set of statistical units to another, as from census tracts to school districts.

1st example

Population density in Kansas

by county

Courtesy of T. Slocum

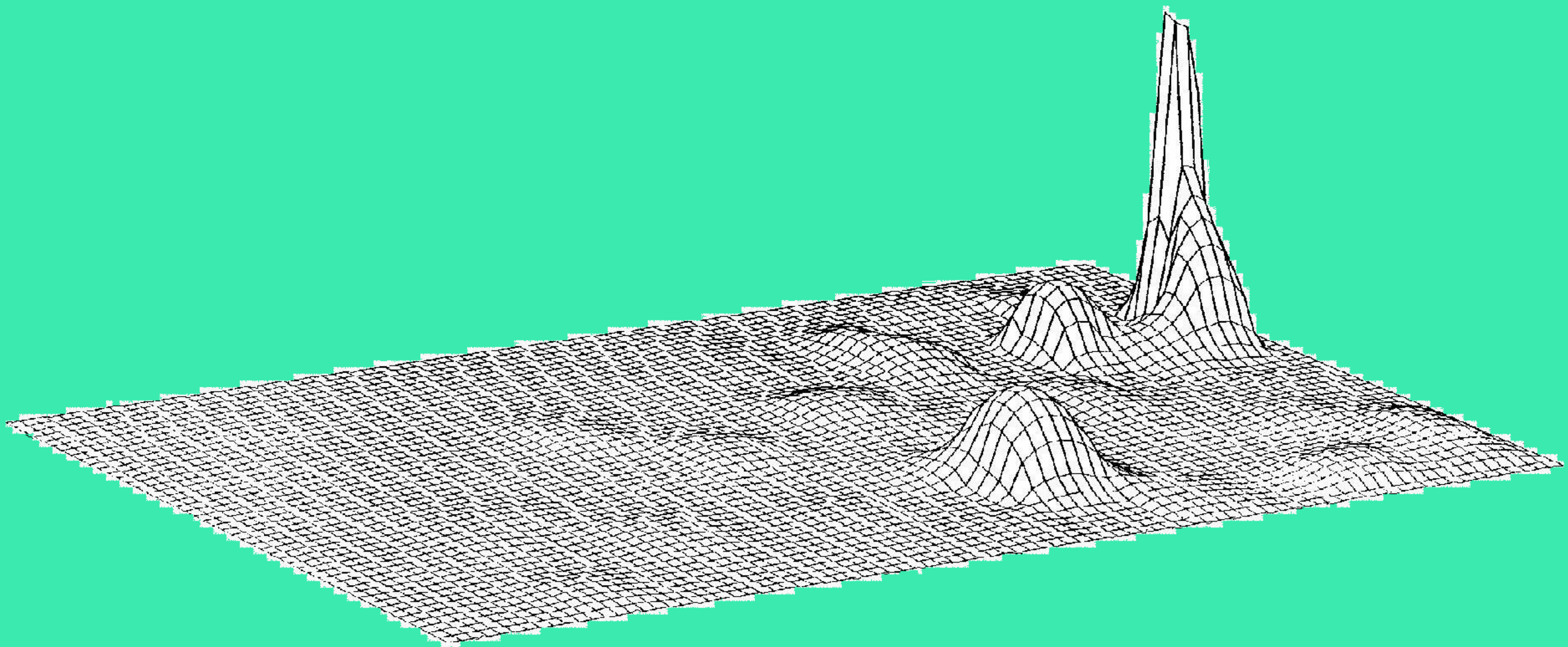


A piecewise continuous surface

Population density in Kansas

after mass preserving reallocation

Each County Still Contains the Same Number of People



A smooth continuous surface, with population pycnophylactically redistributed

Another example

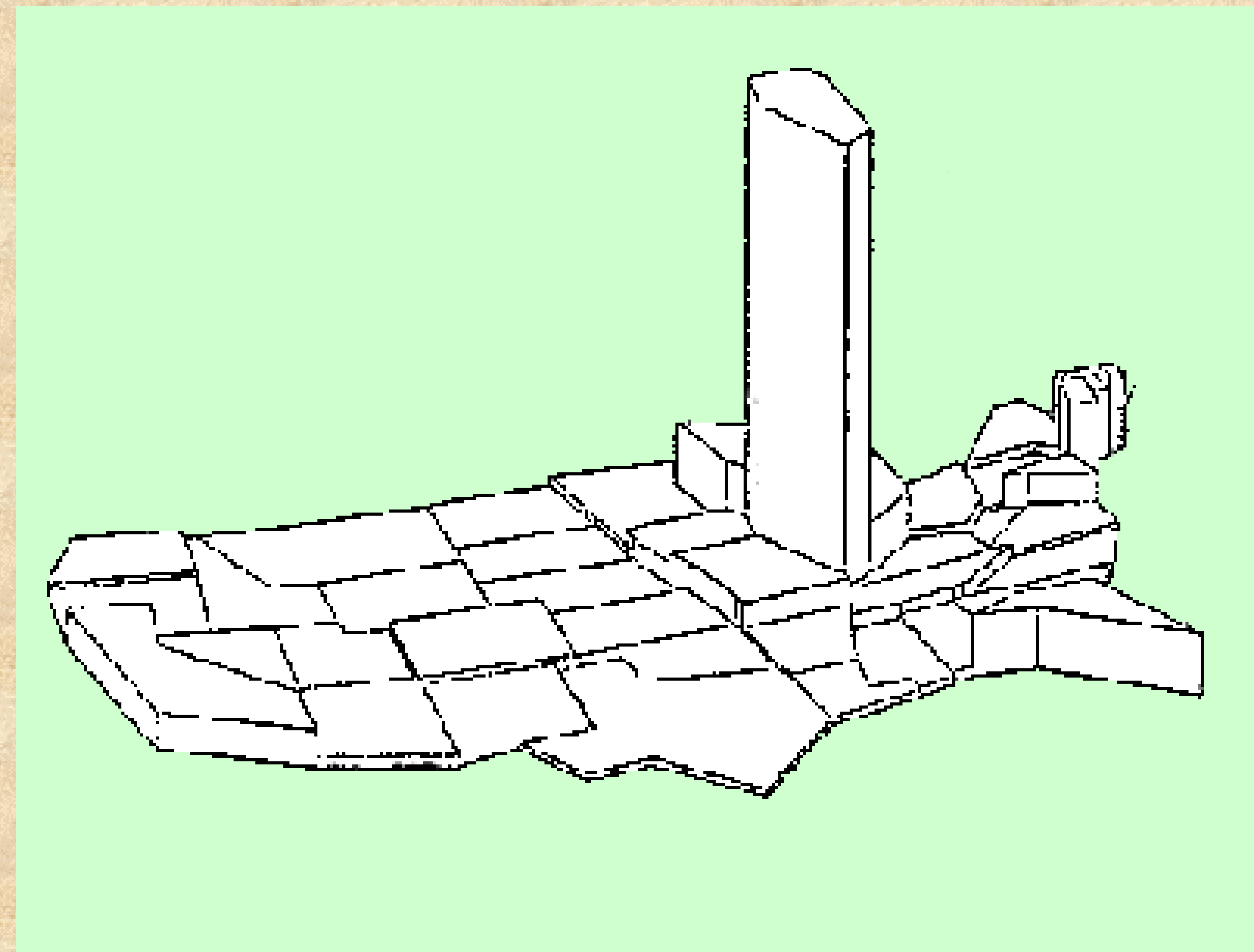
Migration from Illinois shown first as a piecewise continuous bivariate geographical histogram, based on state outlines, with volumes according to Illinois outmigration

Recall that most migrants in Illinois relocate within the state

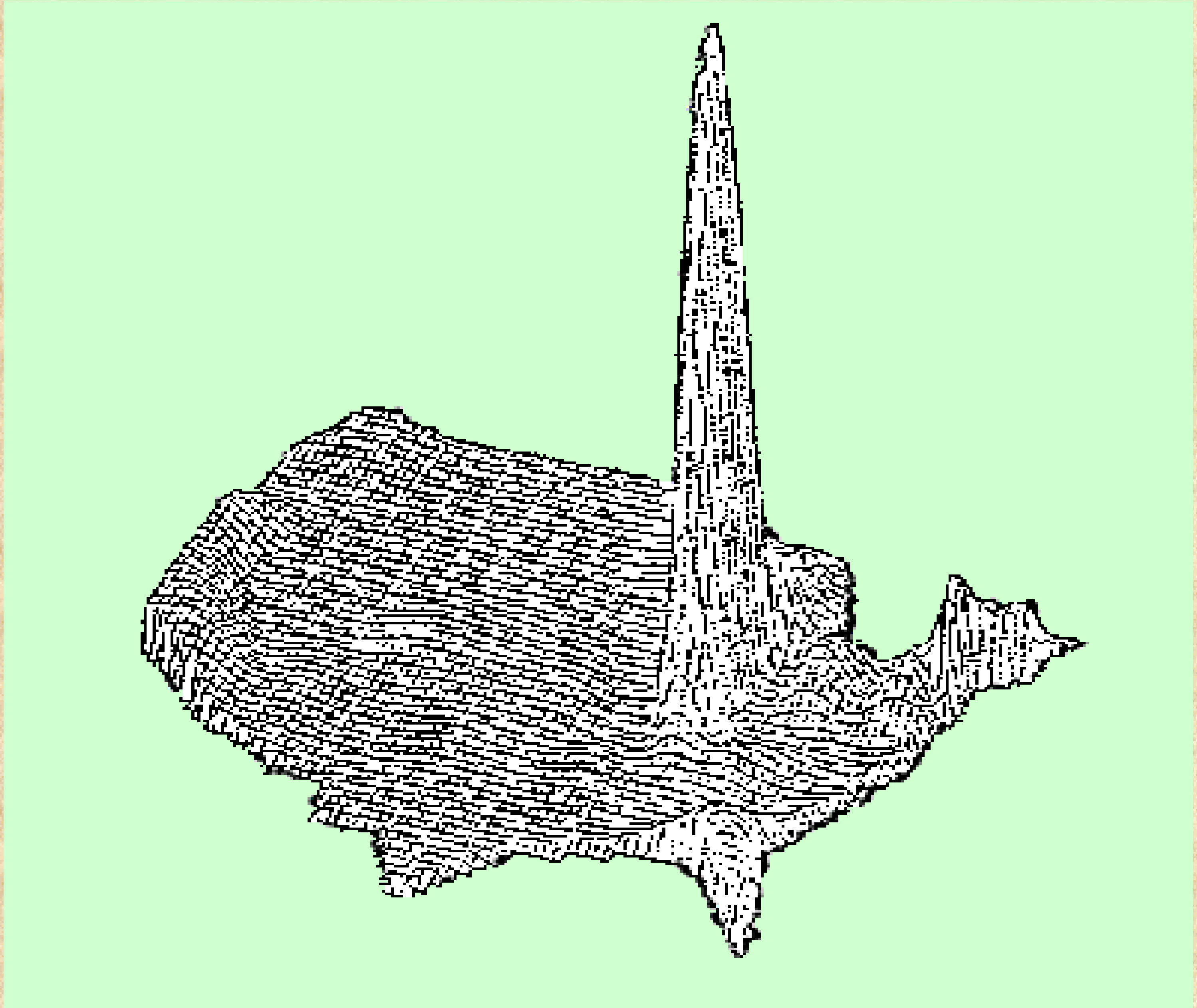
The same data is then shown as pycnophylactically interpolated

The smoothed surface can be partitioned to yield estimated migration by arbitrary regions - the Great Lakes basin for example

Bivariate histogram of Illinois outmigration by state



Illinois outmigration pycnophylactically smoothed



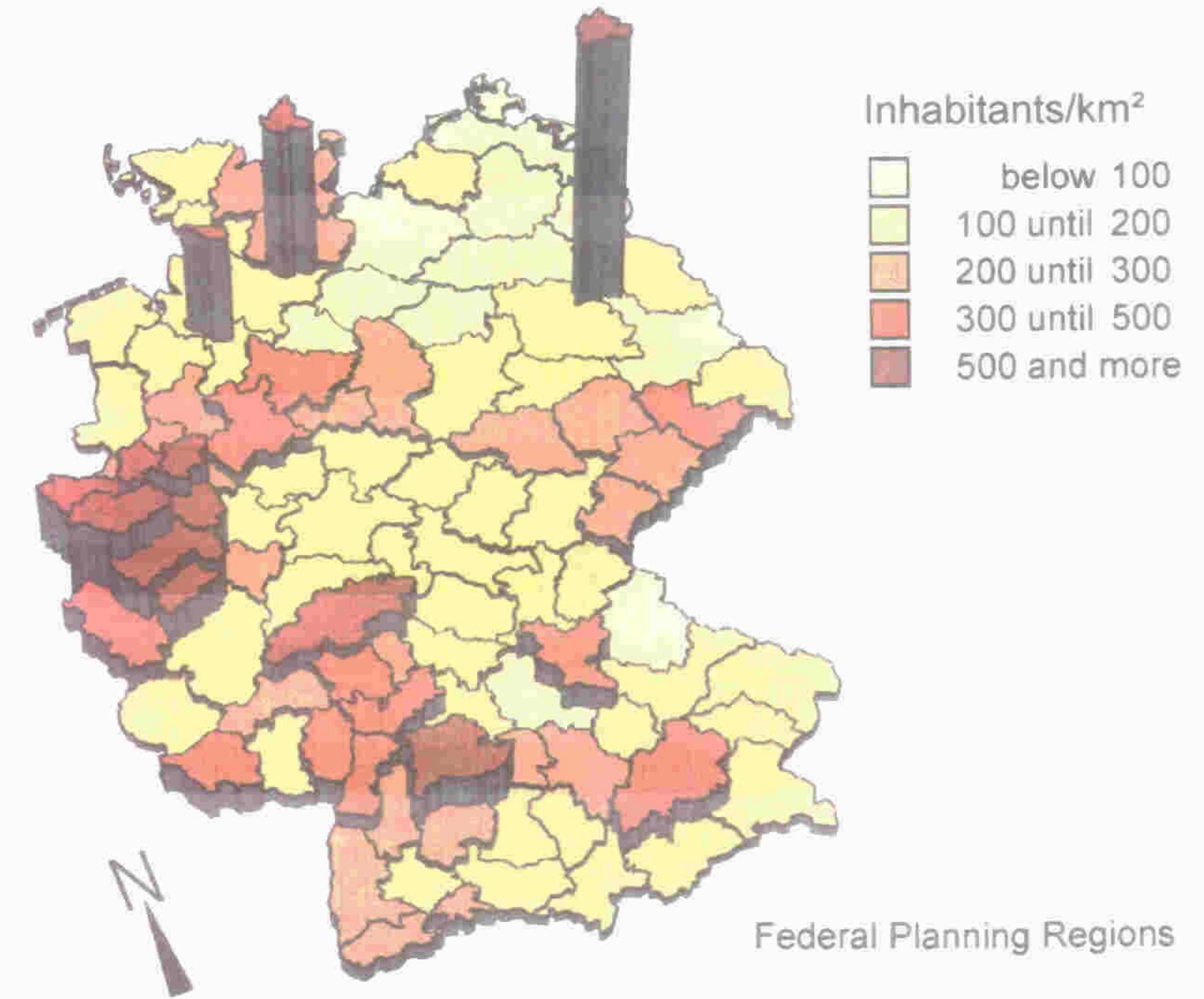
Another example

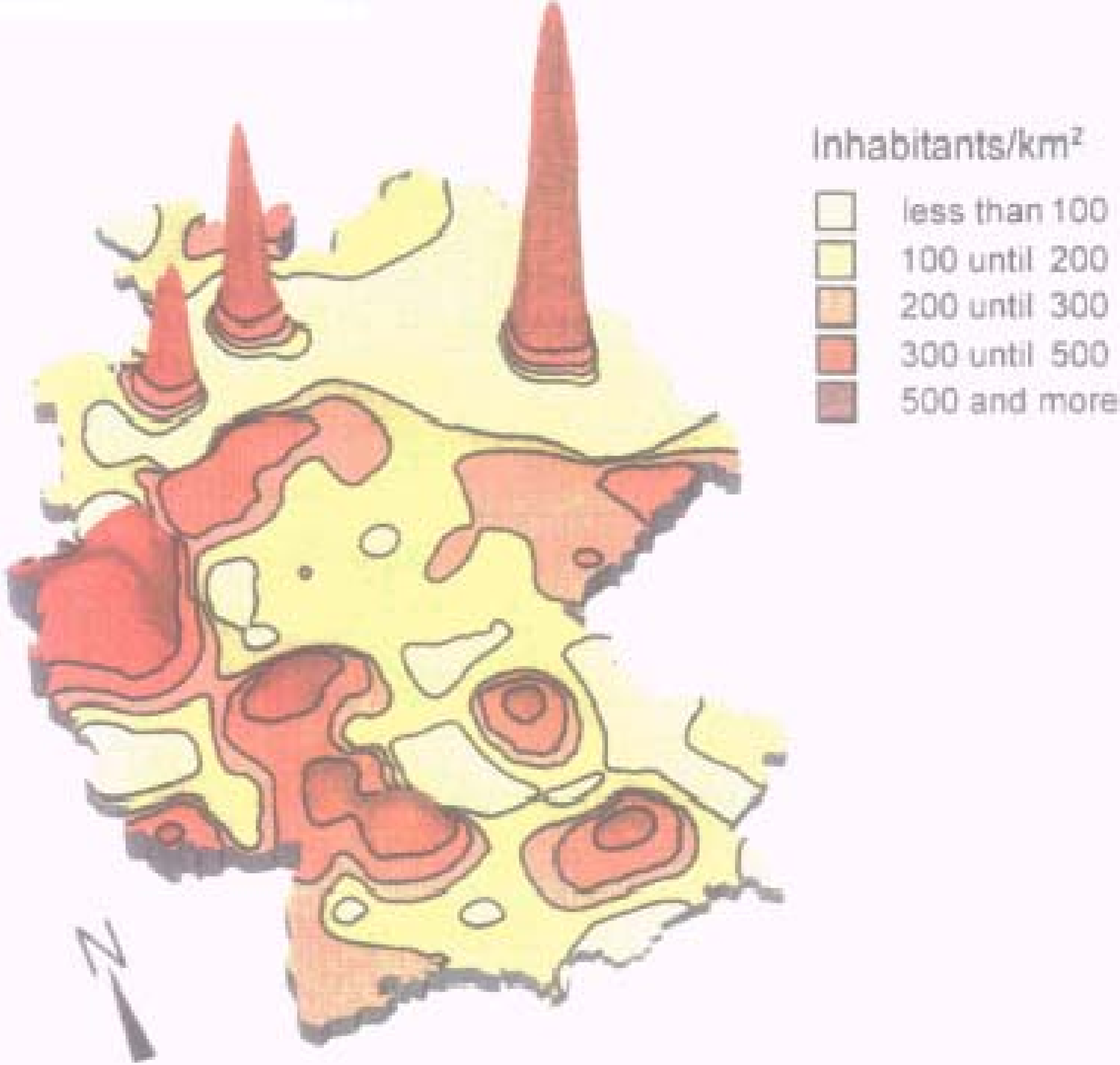
This time using population data by Federal Planning Regions for Germany.

First the data are represented in a perspective view of a bivariate geographical histogram.

This is followed by a similar view of the continuous population density distribution.

Wolf-Dieter Rase, 2001, "Volume-preserving interpolation of a smooth surface from polygon-related data", *J. Geograph. Syst*, 3:199-213.


Inhabitants/km²
below 100
100 until 200
200 until 300
300 until 500
500 and more
Federal Planning Regions


Inhabitants/km²
less than 100
100 until 200
200 until 300
300 until 500
500 and more

How pycnophylactic reallocation works

Philosophically it is based on the notion that people are gregarious, influence each other, are mobile, and tend to congregate.

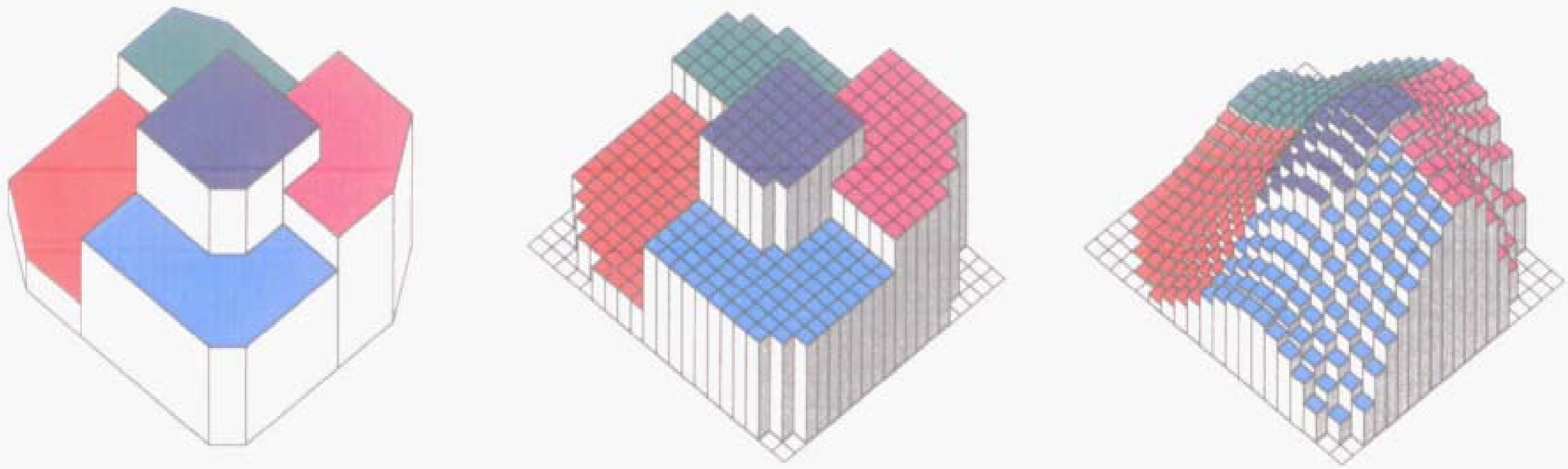
This leads to neighboring and adjacent places being similar.

Mathematically this translates into a smoothness criterion (with small partial derivatives).

It applies to any data exhibiting spatial autocorrelation.

Left to Right

1. Data polygons
2. Rasterized
3. Smoothed



How the smoothing is done.

Imagine that each unit is built up of colored clay, with a different color for each unit.

The volume of clay represents the number of people, say, and the height represents the density.

In order to obtain smooth densities a spatula is used to smooth the surface, but no clay is allowed to move from one unit into another. Color mixing is not allowed.

This, converted to mathematics, is what the computer program does.

Density from dot maps without using kernels

The pycnophylactic method can also be used to prepare smooth density maps from data given at spot locations.

Step 1. Use the inverse area of Dirichlet (a.k.a. Thiessen) regions as the density for each location. If weights are attached to the locations divide these by the region area.

Step 2. Smooth the resulting densities by the pycnophylactic reallocation method.

Another important advantage of mass preserving reallocation

A frequent problem is the reassignment of observations from one set of collection units to a different set, when the two sets are not nested nor compatible. For example converting the number of children observed by census tract to a count by school district. Area boundaries also usually change over time, requiring reallocation for compatibility.

The density values obtained using the smooth pycnophylactic method allow an <u>estimate</u> to be made rather simply. A “cookie cutter” can cut the continuous (clay) surface into the new zones with subsequent addition (summation) to get the count.

The last topic is the depiction of geographical movement

A great deal of change in the world is due to geographical movement

Movement of information, of people, of money, or of material

Animation is well suited to depicting this dynamic cartography

Tables are an important way of recording data on geographic movement

Especially when the rows and columns refer to known geographic locations

The tables are then “square”, having the same number of rows as columns

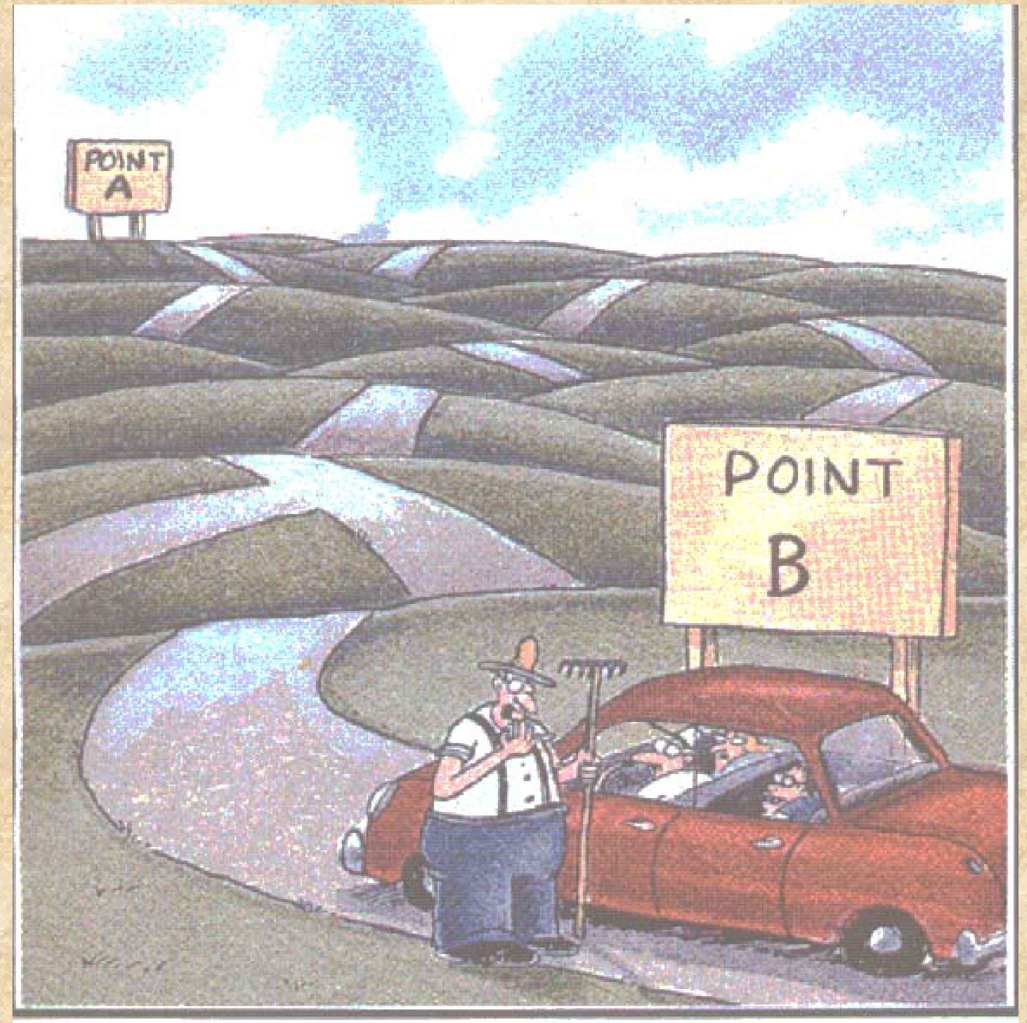
The entries in the tables record the amount of movement during some period of time

Such tables can be decomposed into two parts, a symmetric part and a skew symmetric part

For the statisticians in the audience the total variance can also be partitioned into these two parts

From B to A is not the same as A to B

(Gary Larson)



"Well, lemme think. ... You've stumped me, son. Most folks only wanna know how to go the other way."

An example

In the United States the currency indicates where it was issued

For bills this is the Federal Reserve District.

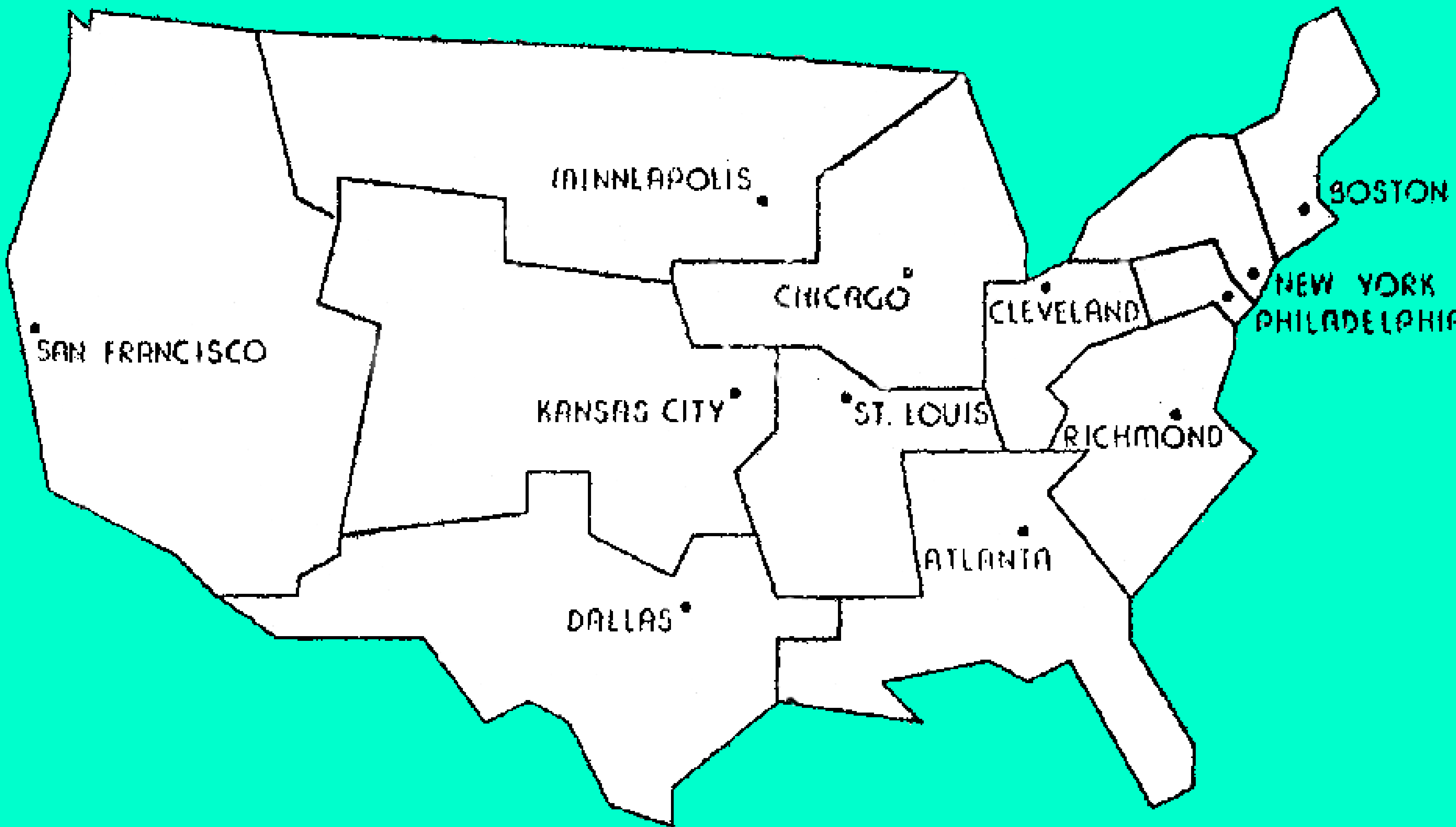
Coins contain a mint abbreviation.

You can check your wallet to estimate your interaction with the rest of the country.

Dollar Bill
(Federal Reserve Note)

The 12 Federal Reserve Districts

(Alaska and Hawaii Omitted)



Movement of One Dollar Notes

Between Federal Reserve Districts, in hundreds, Feb. 1976

From: \ To:	B	NY	P	Cl	R	A	Ch	SL	M	K	D	SF
Boston	2040	289	47	52	137	118	90	10	16	15	13	138
New York	602	1980	231	209	388	307	286	15	48	26	18	261
Philadelphia	143	414	860	84	342	130	134	8	25	10	10	80
Cleveland	68	192	47	1296	171	177	618	16	44	43	19	131
Richmond	150	266	158	226	3899	578	295	20	62	54	22	152
Atlanta	122	159	57	186	319	3741	439	30	51	78	102	189
Chicago	97	155	39	496	143	266	5630	74	278	100	40	290
St. Louis	31	56	14	142	80	201	573	342	46	128	47	109
Minneapolis	14	26	11	32	29	41	295	10	1438	51	14	138
Kansas city	20	41	8	55	40	71	215	33	120	811	86	247
Dallas	31	41	8	38	46	165	125	20	37	253	788	203
San Francisco	82	81	23	84	114	106	251	22	127	128	43	5380

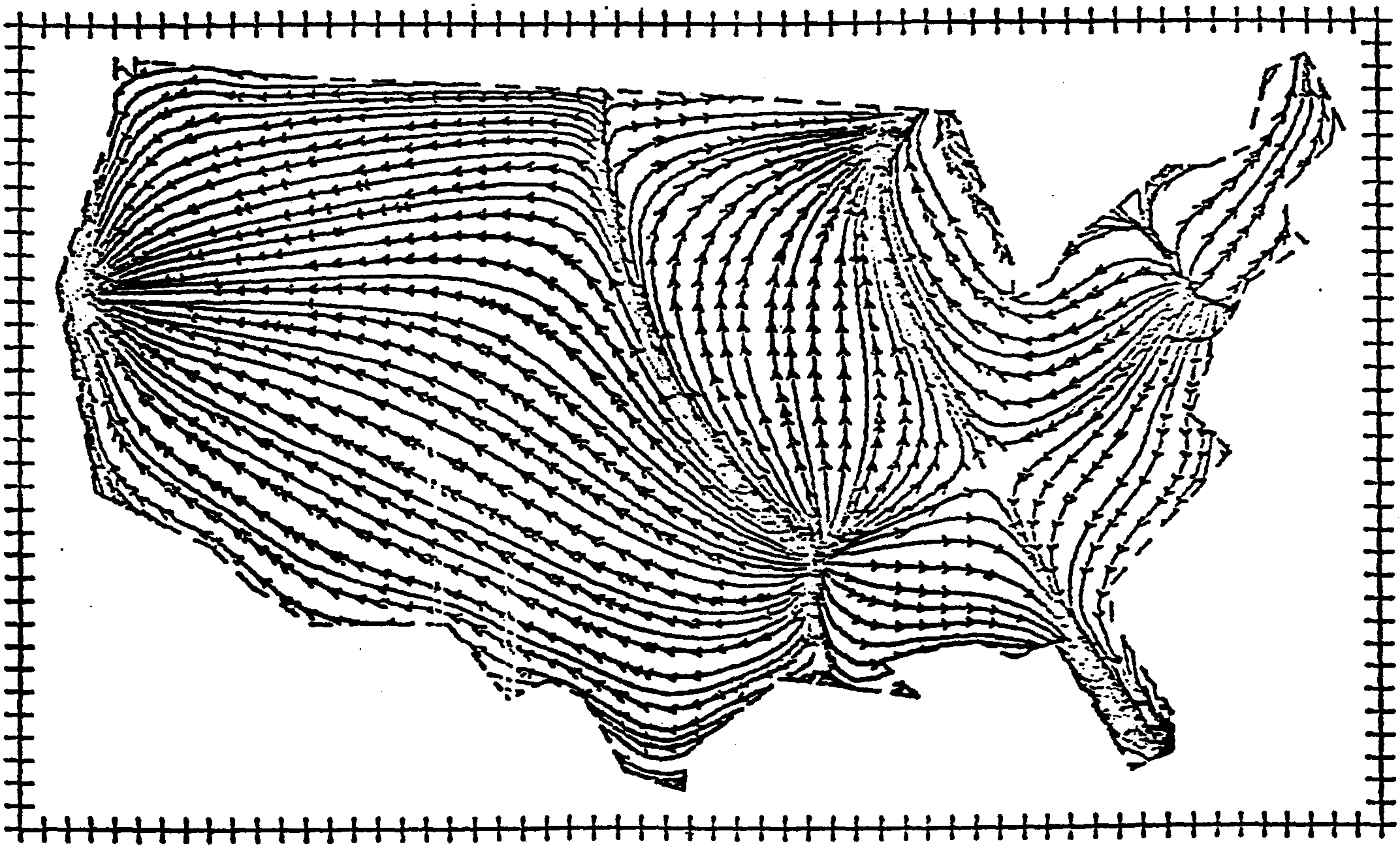
The table of dollar bill movements

was obtained from MacDonalds outlets throughout the United States.

Source: S. Pignatello, 1977, *Mathematical Modeling for Management of the Quality of Circulating Currency*, Federal Reserve Bank, Philadelphia

From the table we can compute a movement map.

Dollar Bill Movement in the U.S.



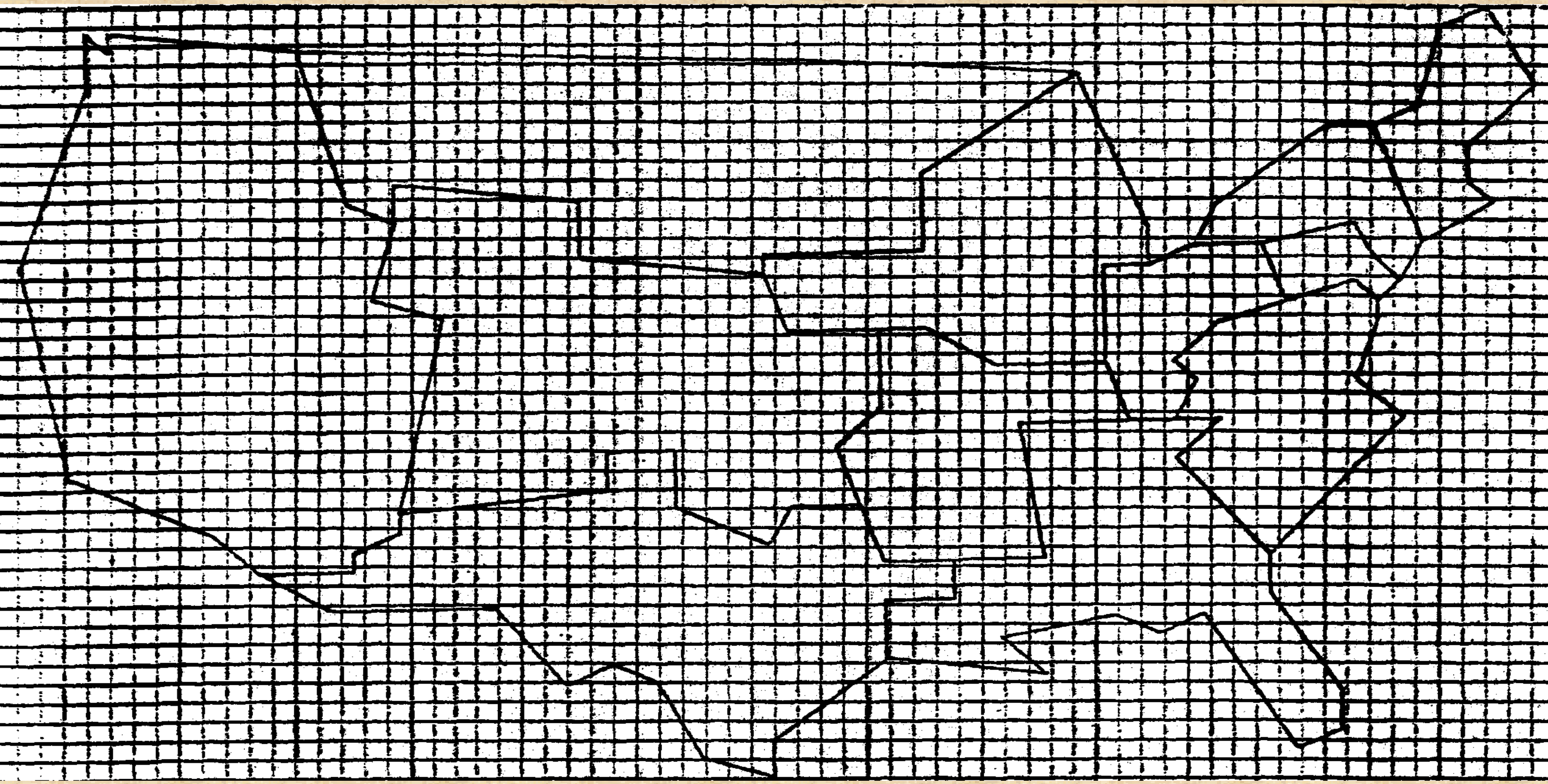
The map is computed using a continuous version of the gravity model

The result is a system of partial differential equations solved by a finite difference iteration to obtain the potential field.

This can be contoured and its gradient computed and drawn on a map.

W. Tobler, 1981,"A Model of Geographic Movement", *Geogr. Analysis*, 13 (1): 1-20
G. Dorigo, & Tobler, W., 1983, “Push Pull Migration Laws”, *Annals*, AAG, 73(1):1-17.

First the Federal Reserve Districts are "rasterized"

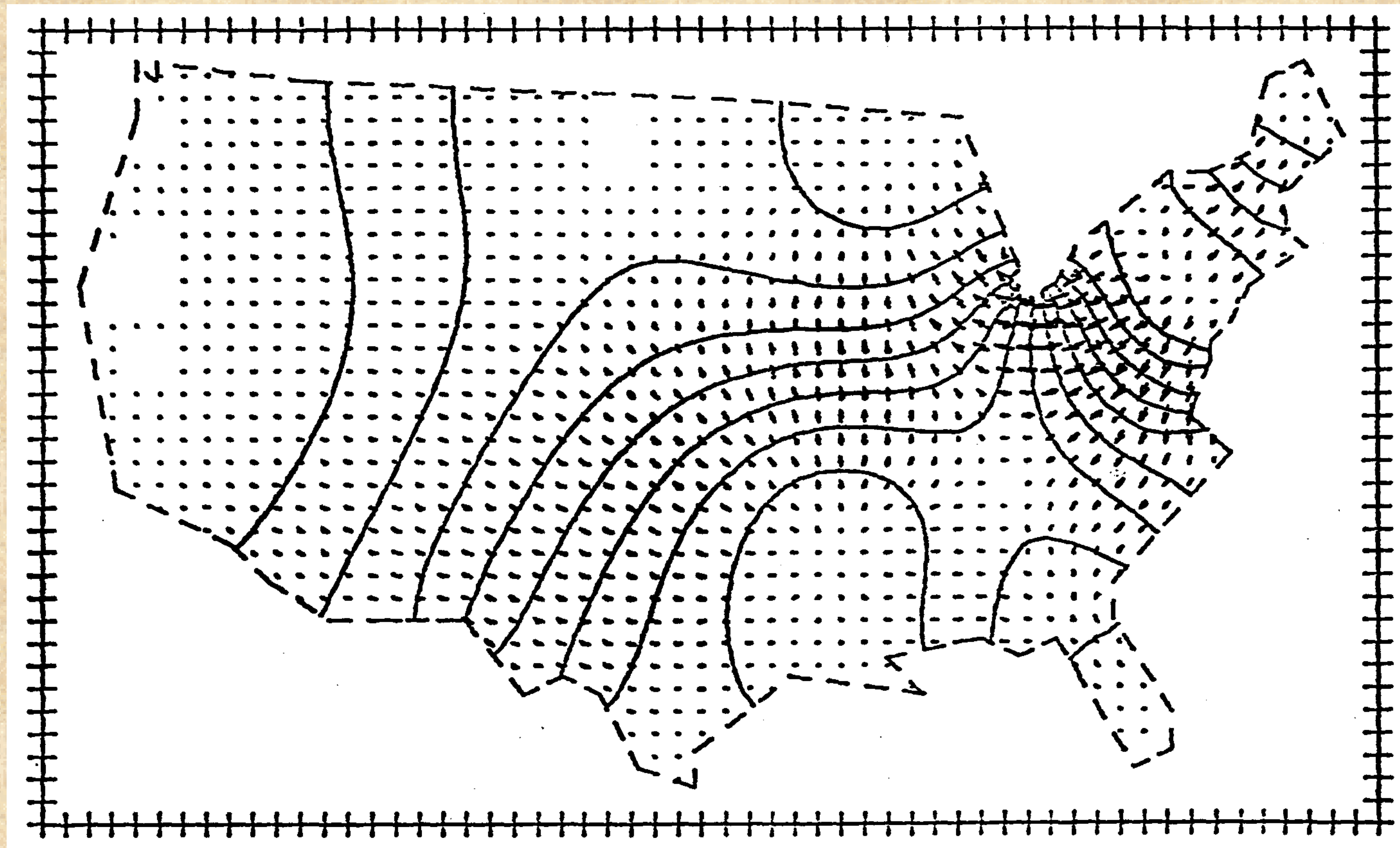


There will be one finite difference equation for each node on this raster

(2088 simultaneous equations)

Solving the equations yields the potential

Shown here by contours



The raster is indicated by the tick marks. The arrows are the gradients to the potentials. The streakline map is obtained by connecting the gradient vectors.

The same technique can be applied to other types of movement

For example the migratory movement of people.

Nine Region Migration Table

US Census 1965-1970

(Note asymmetry. There are places of depletion and accumulation.)

	1	2	3	4	5	6	7	8	9
1 New England	—	180,048	79,223	26,887	198,144	17,995	35,563	30,528	110,792
2 Mid-Atlantic	283,049	—	300,345	67,280	718,673	55,094	93,434	87,987	268,458
3 East North Central	87,267	237,229	—	281,791	551,483	230,788	178,517	172,711	394,481
4 West North Central	28,977	60,681	286,580	—	143,860	49,892	185,618	181,868	274,629
5 South Atlantic	130,830	382,565	346,407	92,308	—	252,189	192,223	89,389	279,739
6 East South Central	21,434	53,772	287,340	49,828	316,650	—	141,679	27,409	87,938
7 West South Central	30,287	64,645	161,645	144,980	199,466	121,366	—	134,229	289,880
8 Mountain	21,450	43,749	97,808	113,683	89,806	25,574	158,006	—	437,255
9 Pacific	72,114	133,122	229,764	165,405	266,305	66,324	252,039	342,948	—

This is an example of a census migration table. There are also (50 by 50) state tables and county by county tables.

There is a great deal of spatial coherence in the migration pattern

Choropleth maps do not show this clearly.

In the U.S. case the state boundaries hide the effect. Therefore a clearer picture emerges if they are omitted.

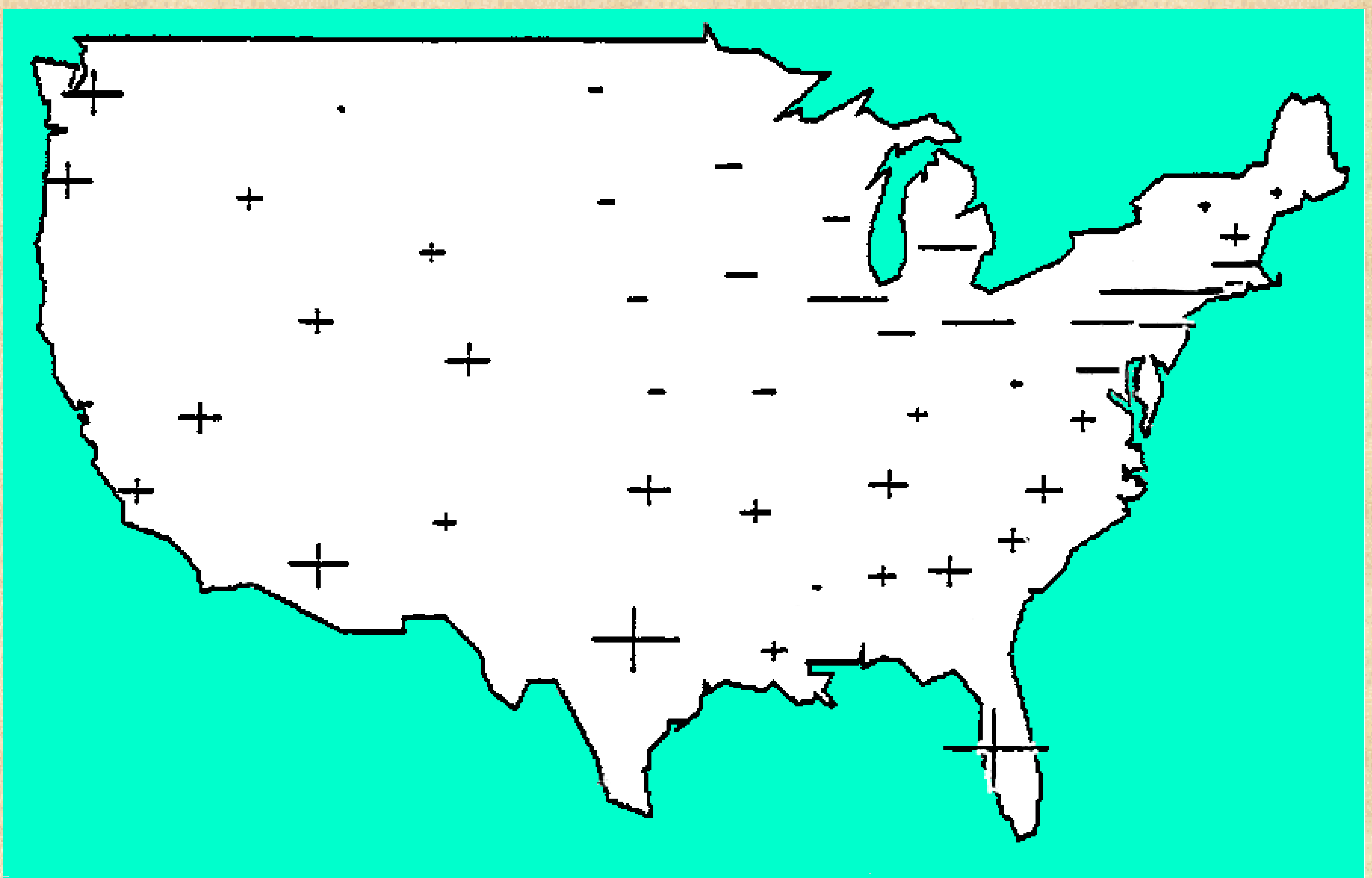
There is also temporal coherence.

W. Tobler, 1995, “Migration: Ravenstein, Thornthwaite, and Beyond”, *Urban Geography,* 16(4):327-343.

Gaining and Losing States

Based on the marginals of a 48 by 48 migration table

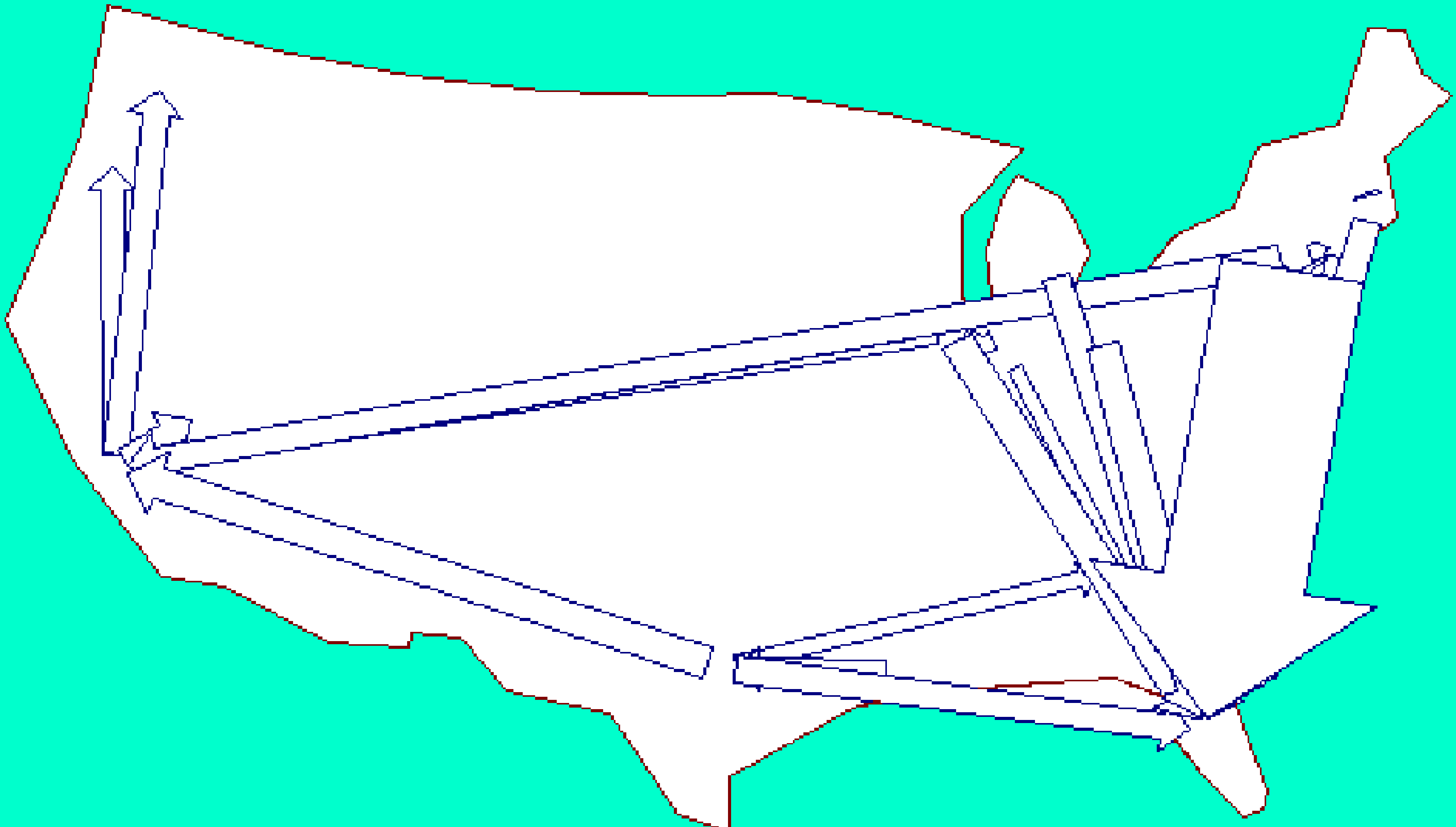
State centroids used with symbol magnitude proportion to the amount of change



The conventional net movement map

Based on movement between state centroids

(Computer sketch. Optimum deletion: values below mean ignored)



This information can be converted to a potential field and its gradient

For this a model is required.

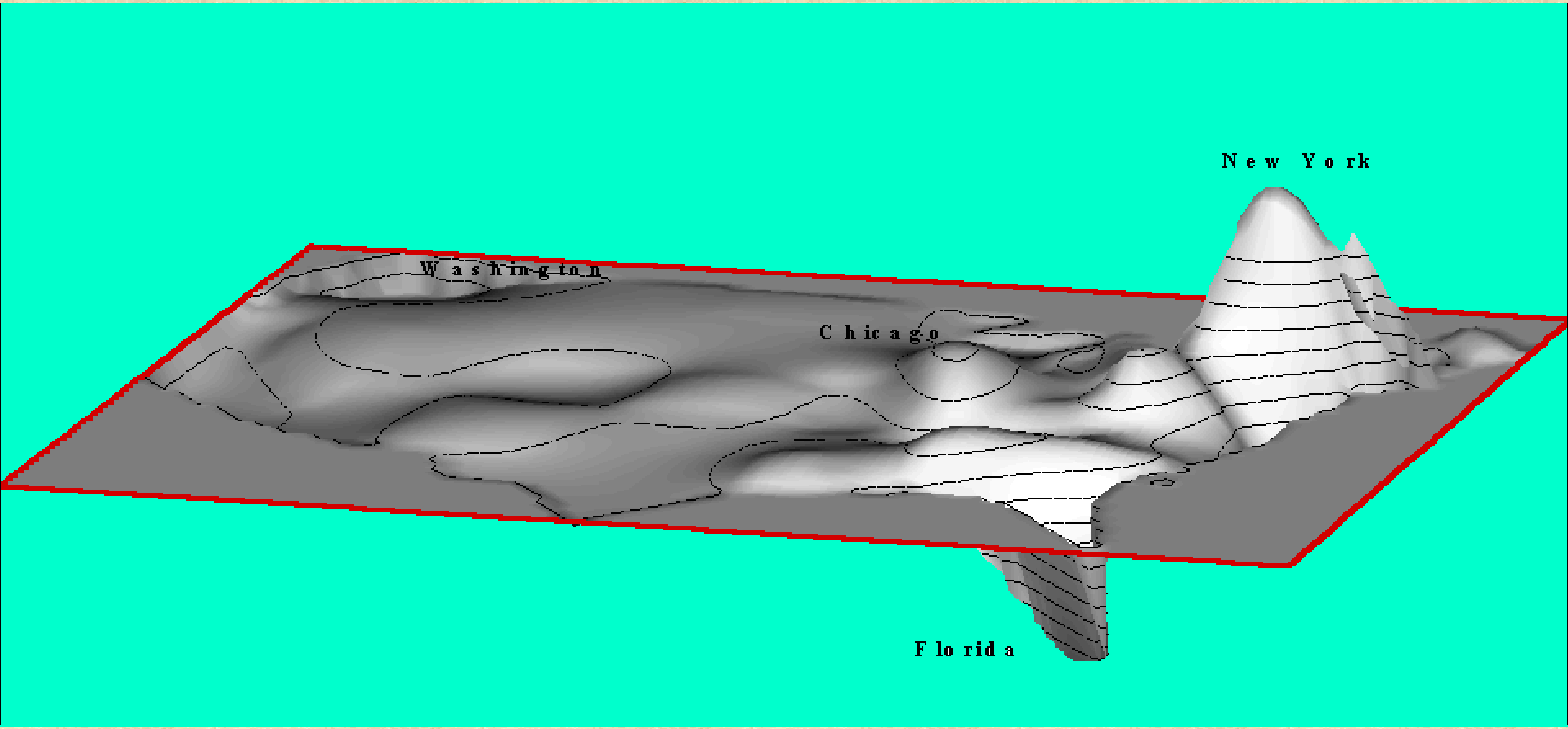
The model is, in essence, a continuous version of the familiar gravity model.

The gradients can also be connected to give a streakline map.

The next maps are based on the same observations as the previous map.

The pressure to move in the US

A continuous spatial gravity model from a movement table



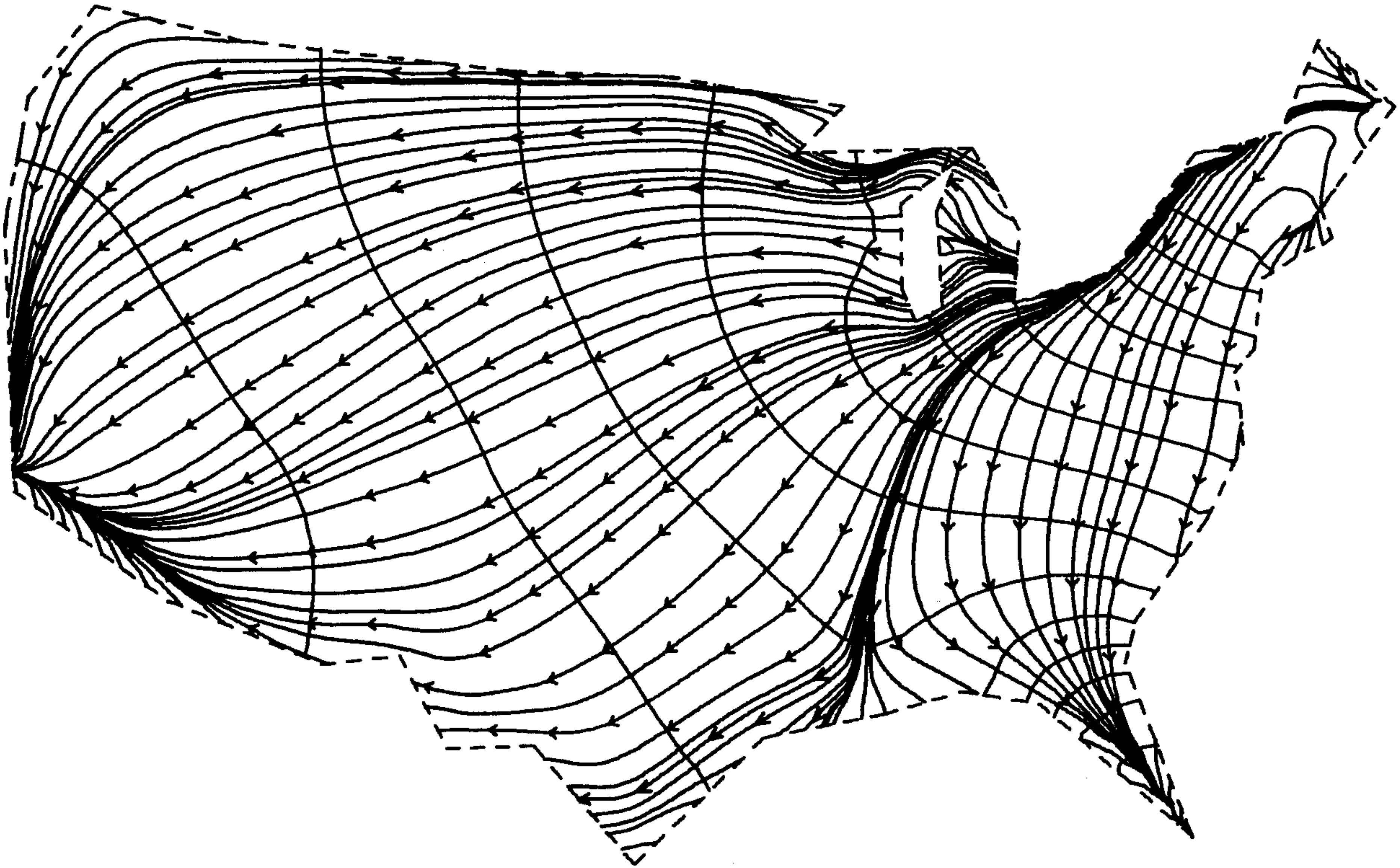
Recall that several million people migrate during the 5 year census period

The next map shows an ensemble average, not the path of any individual.

But observe, not unrealistically, that the people to the East of Detroit tend to go to the Southeast, and Minnesotans to the Northwest, and the remainder to the Southwest.

Migration potentials and streaklines

The streaklines are drawn by connecting the gradients to the potentials



By the insertion of arbitrary areal boundaries, and by measuring the amount of flux across these boundaries, one can obtain information not contained in the original data, i.e., make a prediction.

It’s like using a cookie cutter pressed into the continuous flow model to look at an arbitrary piece and computing the flow across its borders.

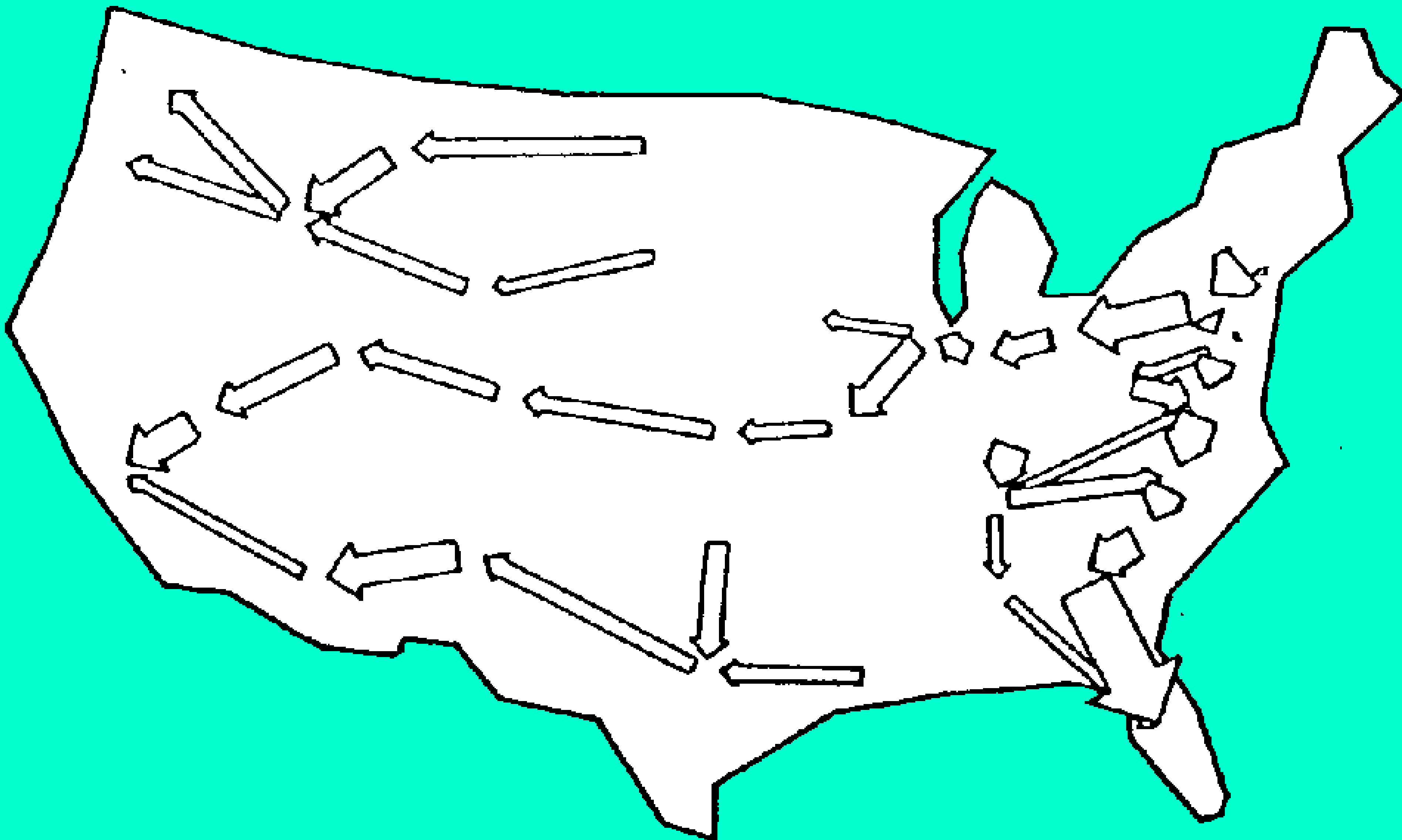
The next map is an example, using state boundaries.

The US Census Bureau does not provide this information.

The model is used to make the prediction

Major Flux Across State Boundaries

Predicted from the model and table marginals



If we used the 3,141 counties of the United States the migration table could contain 9,862,740 numbers

This is not a lot for a computer, but for humans?

We need models and visualization techniques!

Cartography provides excellent visualization and always requires a model.

To conclude

I have emphasized three topics

Map projections

Dealing with aggregate data

- Spatial filtering
- Estimating densities
- Converting to other units

Maps of movement

Thank you for your attention

http://www.geog.ucsb.edu/people/tobler.htm