Today's Topic

Geographical Movement

is critically important.

This is because much **change** in the world is due to geographical movement.

The movement of

ideas, people, disease, money, energy, material, etc.

My talk is structured as follows

First I will show a few movement maps

Then I will comment on a program to produce simple movement maps

Following this are some remarks on models of movement

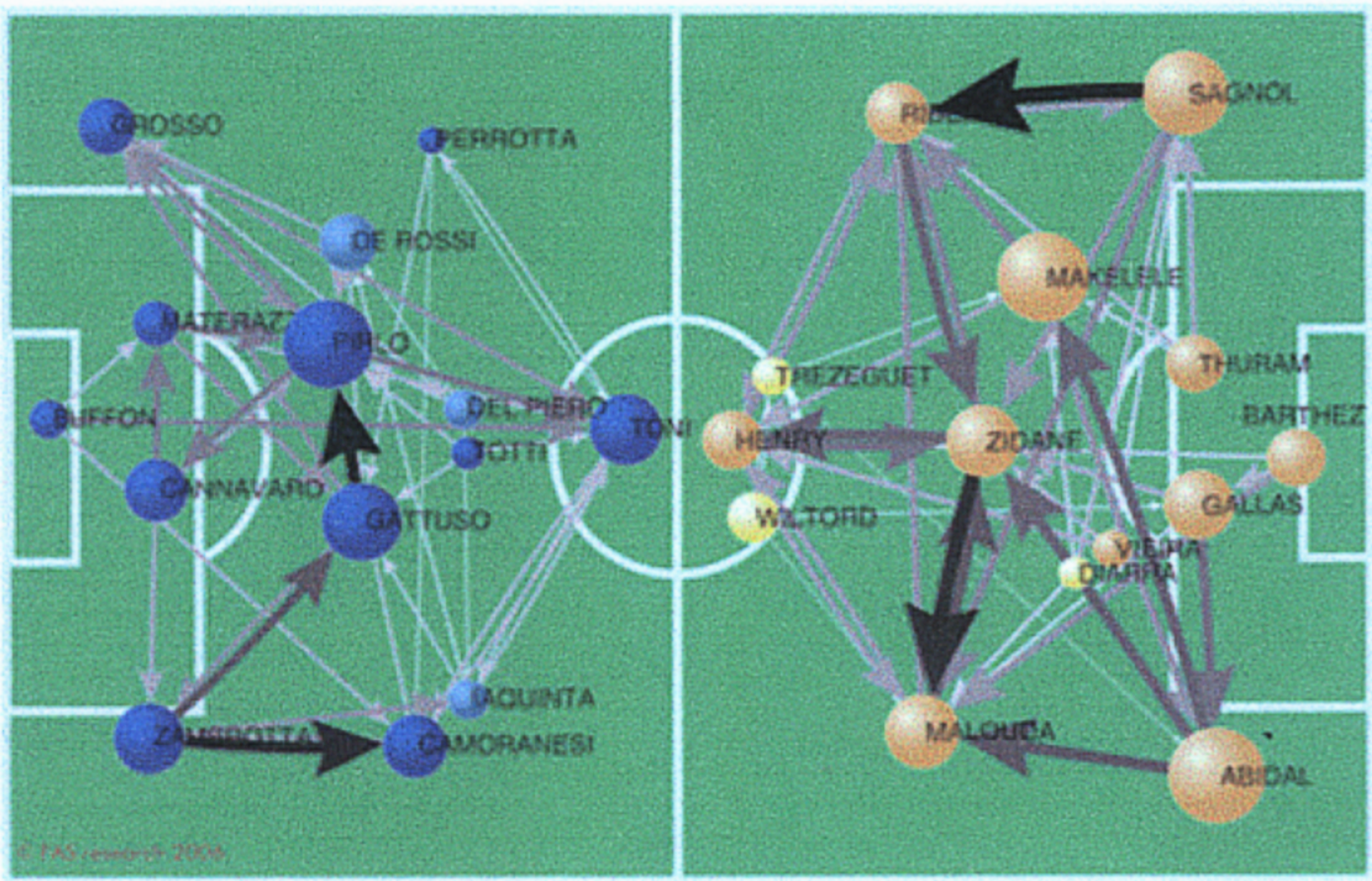
I next extend one model to the spatially continuous case

This will be used to present an example of the movement of money

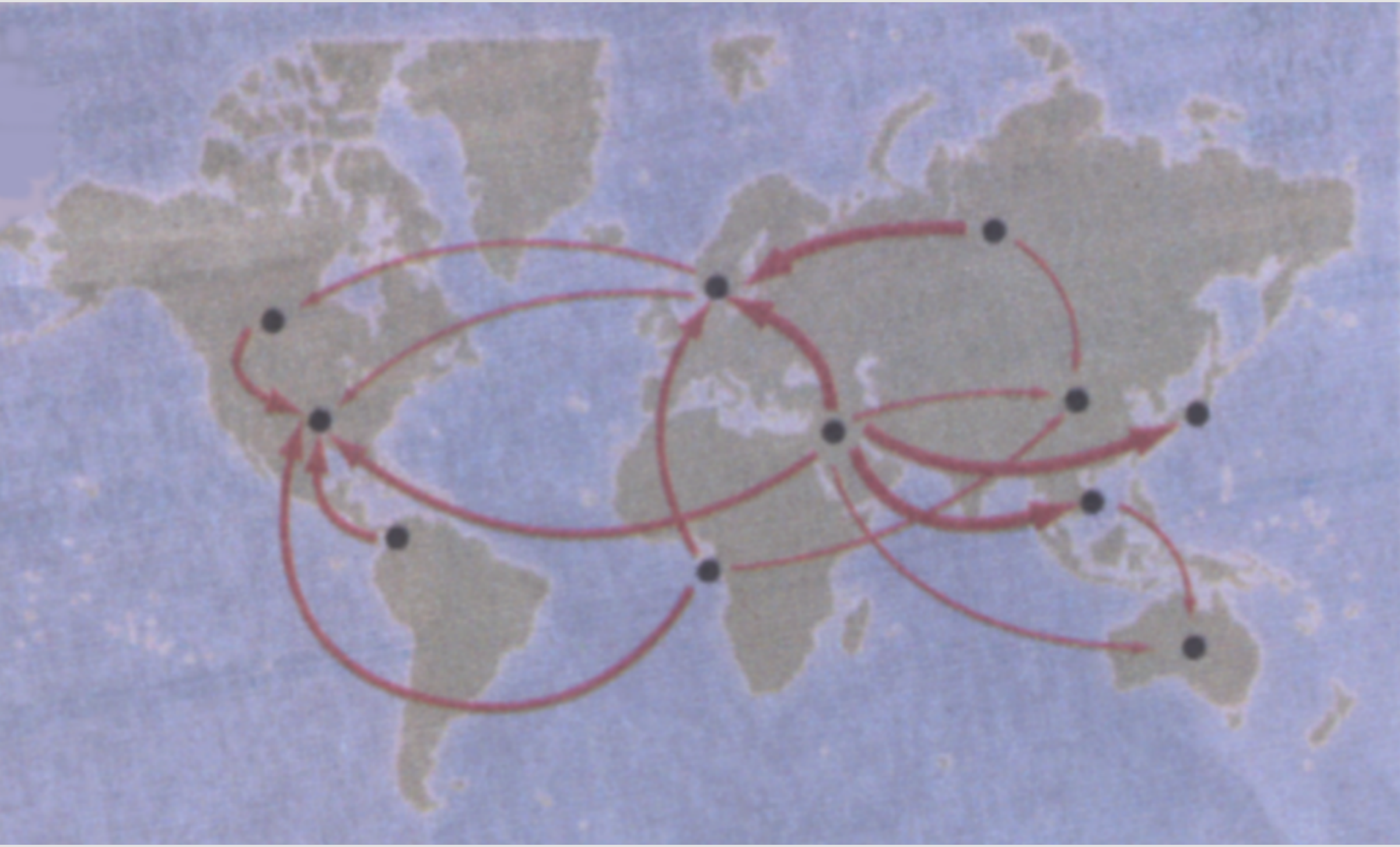
and one of people migrating in the United States

An Exciting Movement Map

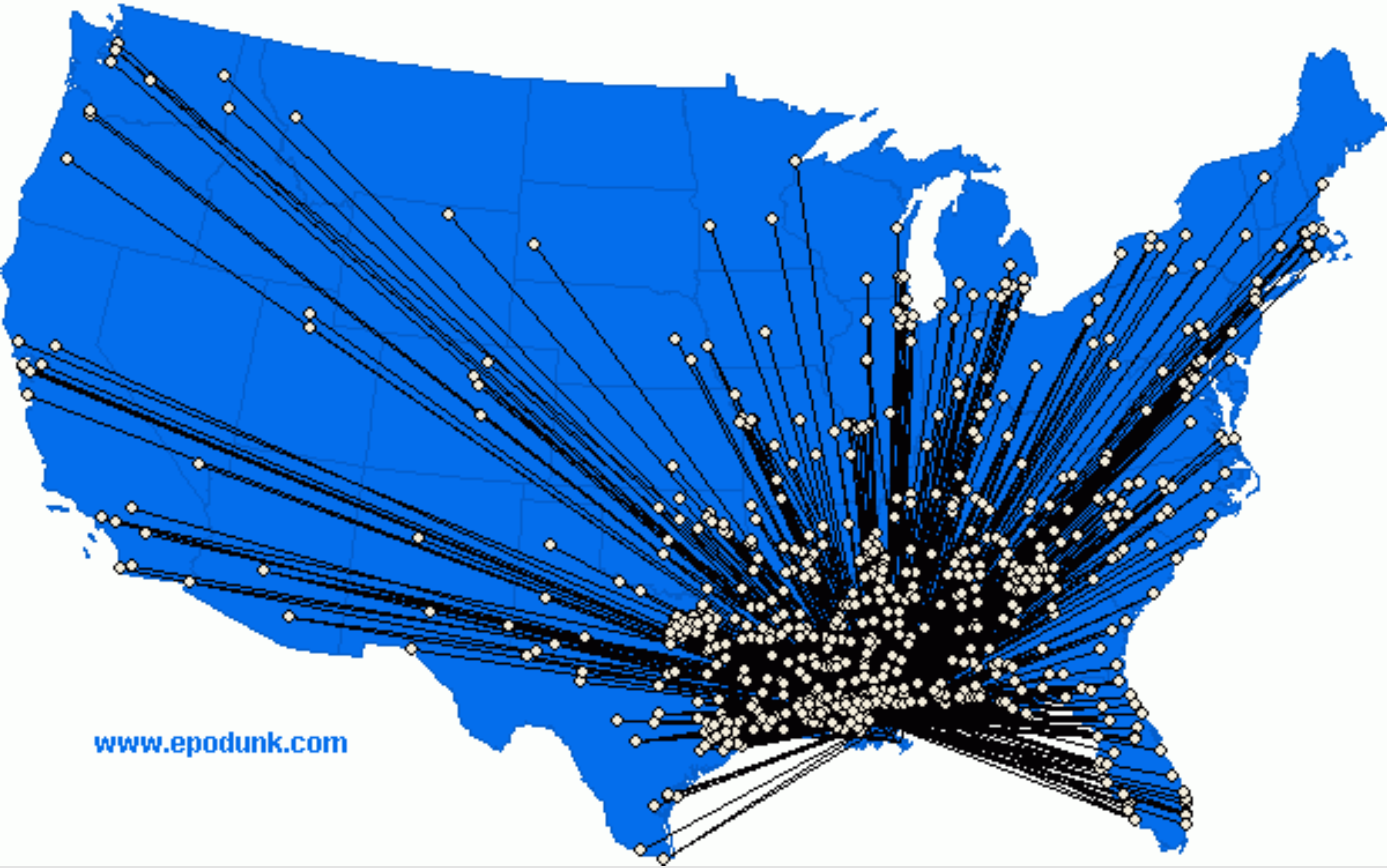
World Soccer Final: Italy against France



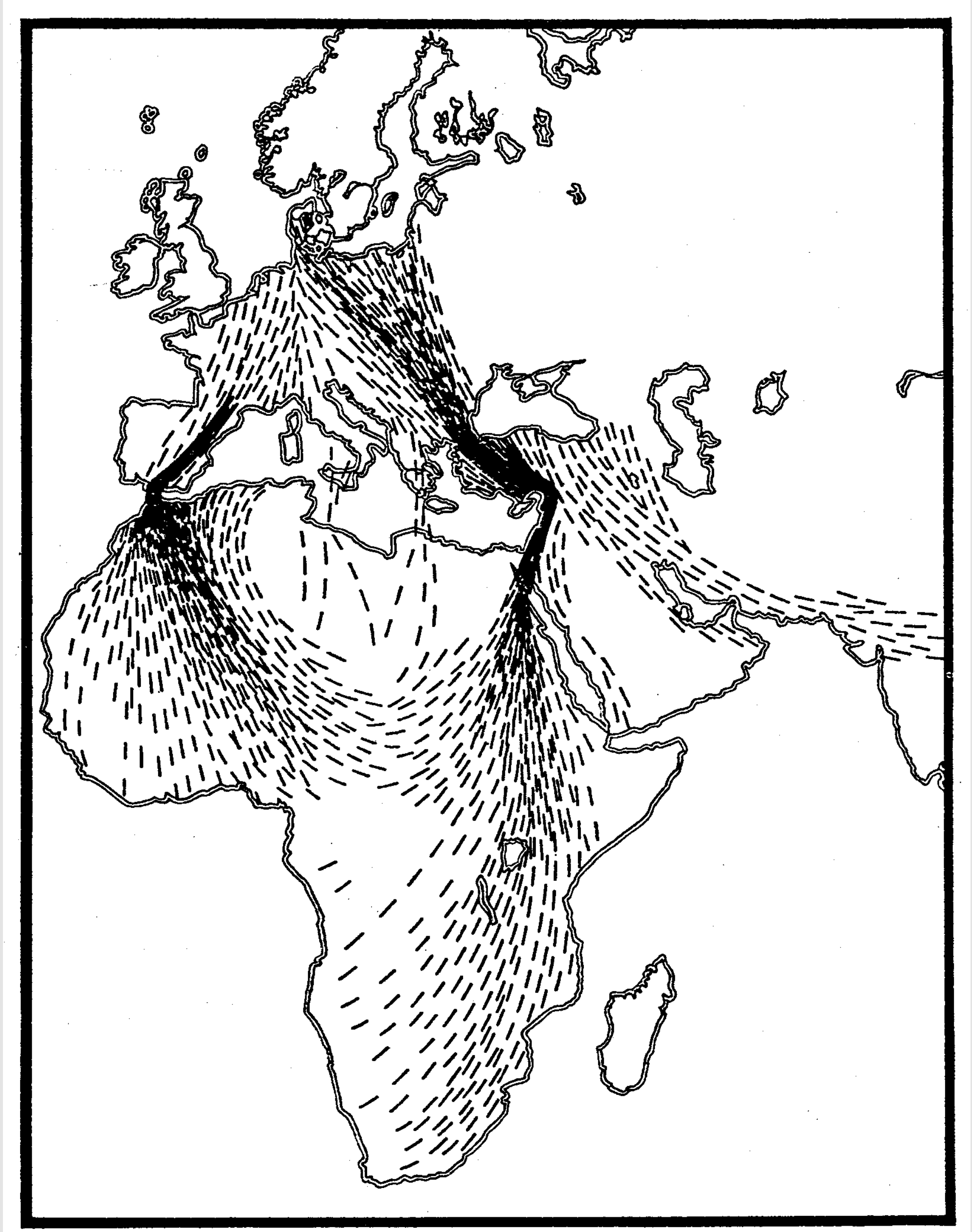
Global Oil Flow



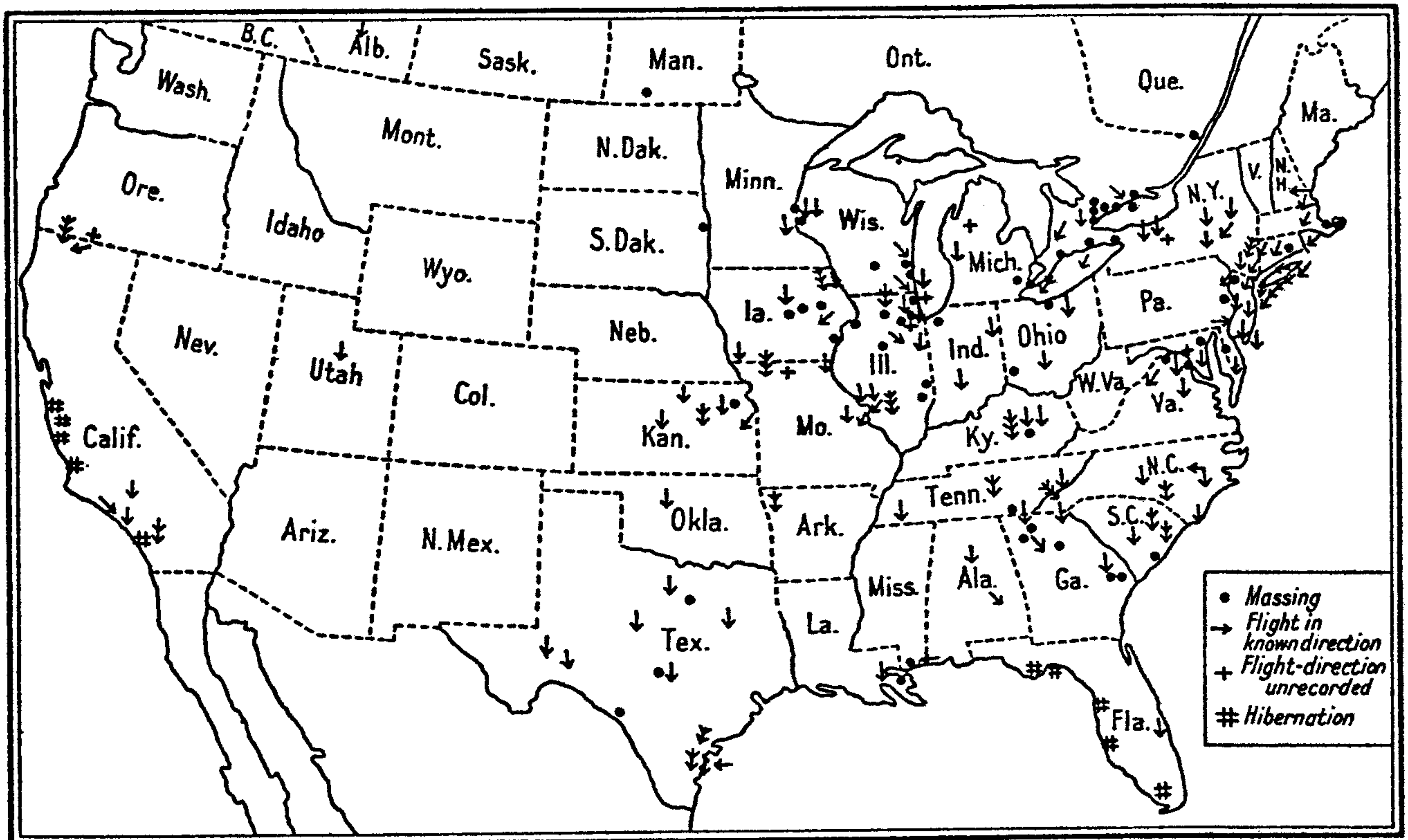
Movement from Katrina



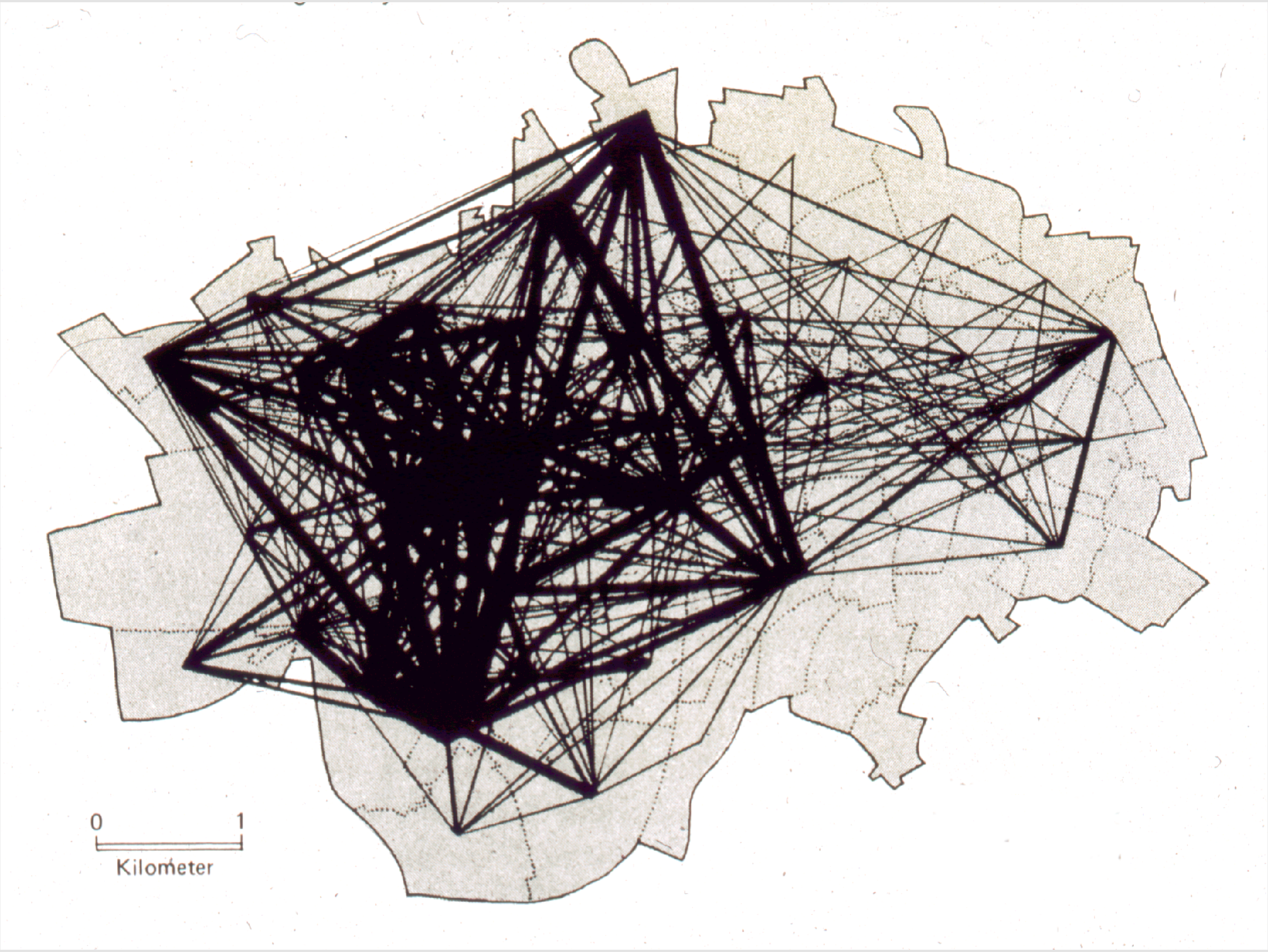
Stork migration



Monarch Butterfly Patterns

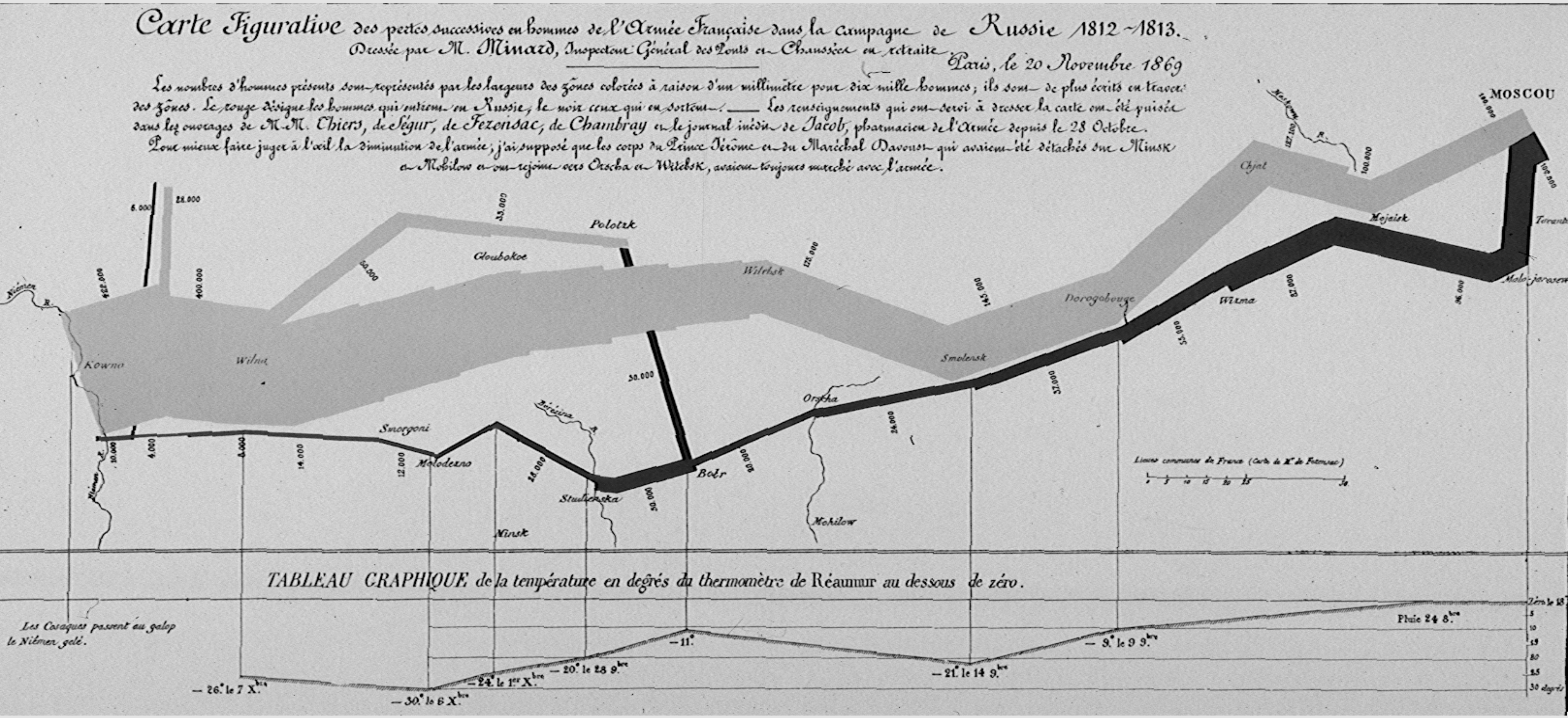


Average daily taxi trafic in London



Daily journey to work around Detroit

Minard's famous map
Napoleon's march and retreat from Russia



Movement Tables

A typical description of movement comes in the form of an array of estimated numbers giving the quantity of movement between known areas. Increasingly often these arrays are large and thus difficult to comprehend in all their detail. One approach to this problem is graphical displays.

This graphical approach is via geographical movement maps, a tradition going back to Charles Minard in the 19th century, but now up-dated to include rapid, simple, and informative interactive computer renditions. A few such maps will be presented.

My original training was in geographical and mathematical cartography. But I find that too many maps are static depictions.

I particularly abhor choropleth maps.

Therefore I have spent more time since the 1970's studying movement.

This is the reason for work on a flow mapping program.

The program can be downloaded from

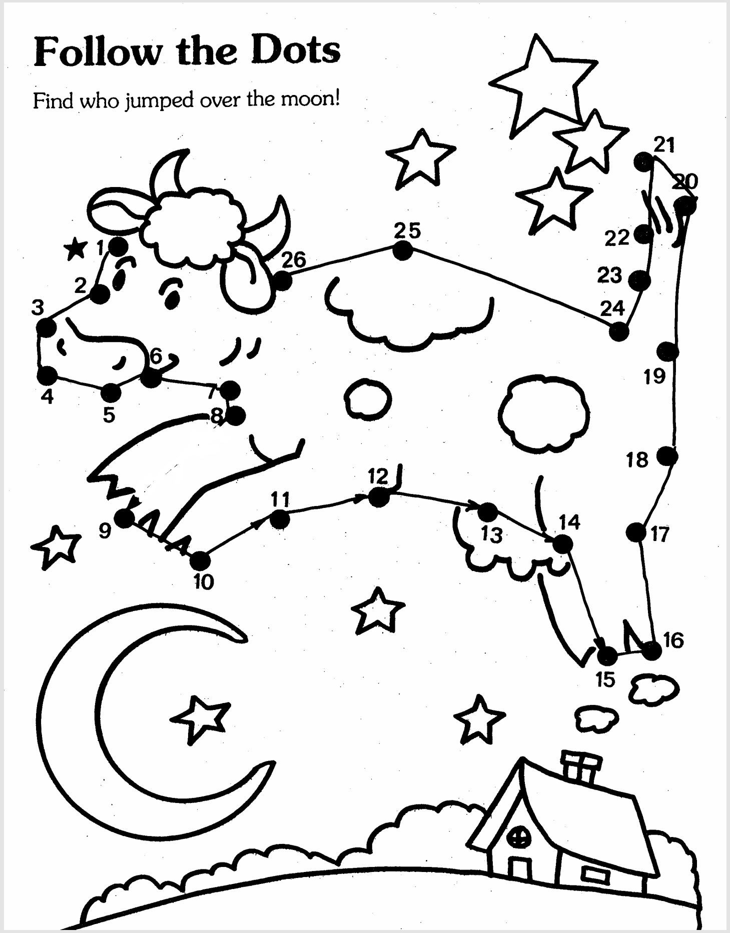
CSISS.ORG/ SPATIAL TOOLS/ FLOW MAPPER

It's free and comes with a tutorial.

As you know much of computer

Cartography is a Dot-to-Dot

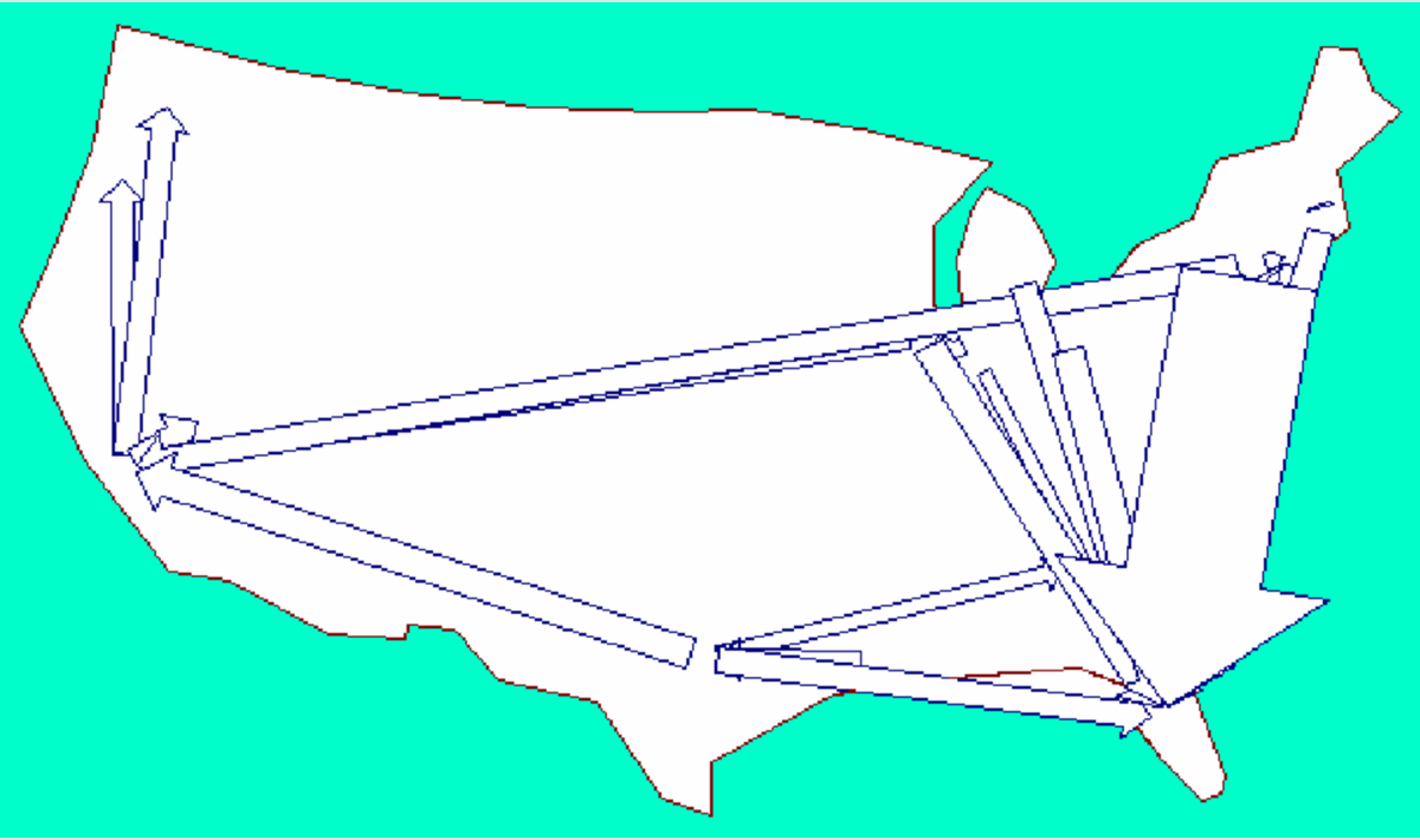
Just replace the dots by coordinates.



U.S. Migration

1965-1970

Map produced by a flow mapping program



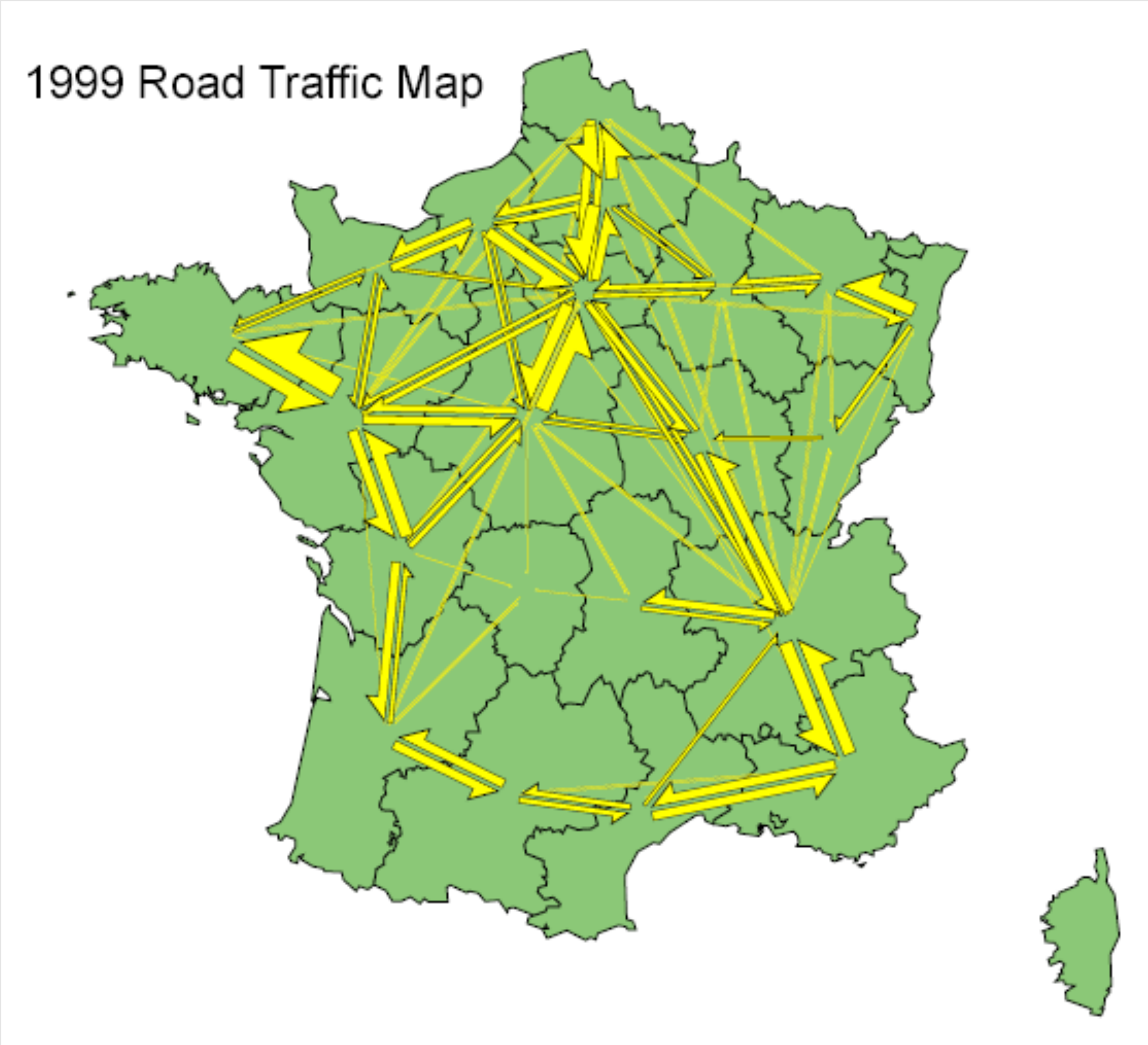
Showing the majority of inter-provincial moves in China

Using the flow mapper program

Another map made using the flow mapper program

Movement between French Regions

Data courtesy of Mr. C. Calzada of Paris



Some nice properties of the Flow Mapper program

Simple and quick flow map preparation - **GIS Not Needed !**

Extensive color styles available. Black & white too.

Hovering over a band or arrow gives the magnitude.

Hovering over a centroid gives its label.

Two-way, total, or net movement maps.

Many to many, one to many, or many to one maps.

Easy threshold choice. Some statistics made available.

Size dependant only on memory availability.

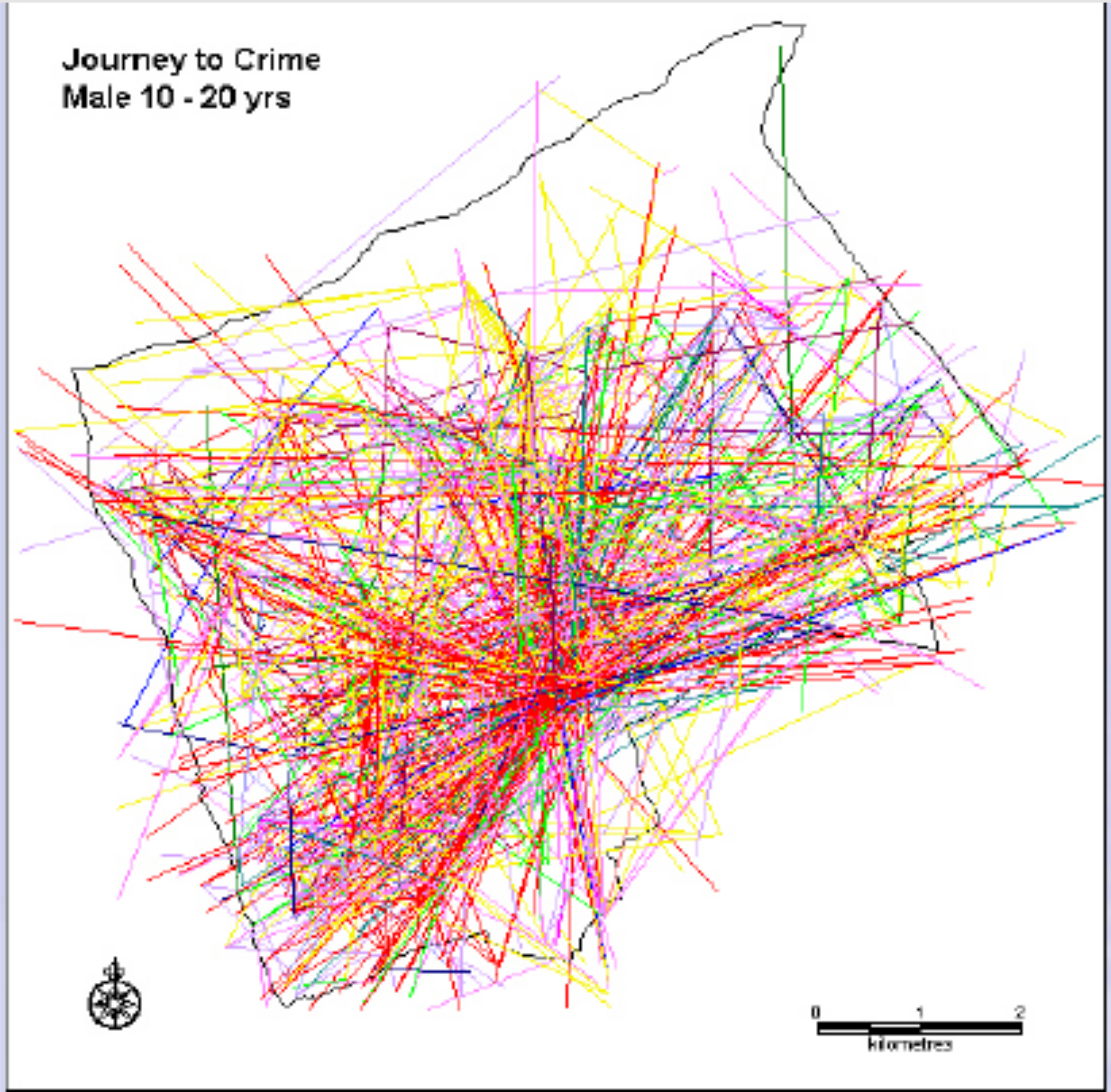
Multiple output formats.

Non-geographic flows within firms, industries, organizations, too.

Help file included.

Microsoft Windows compatible.

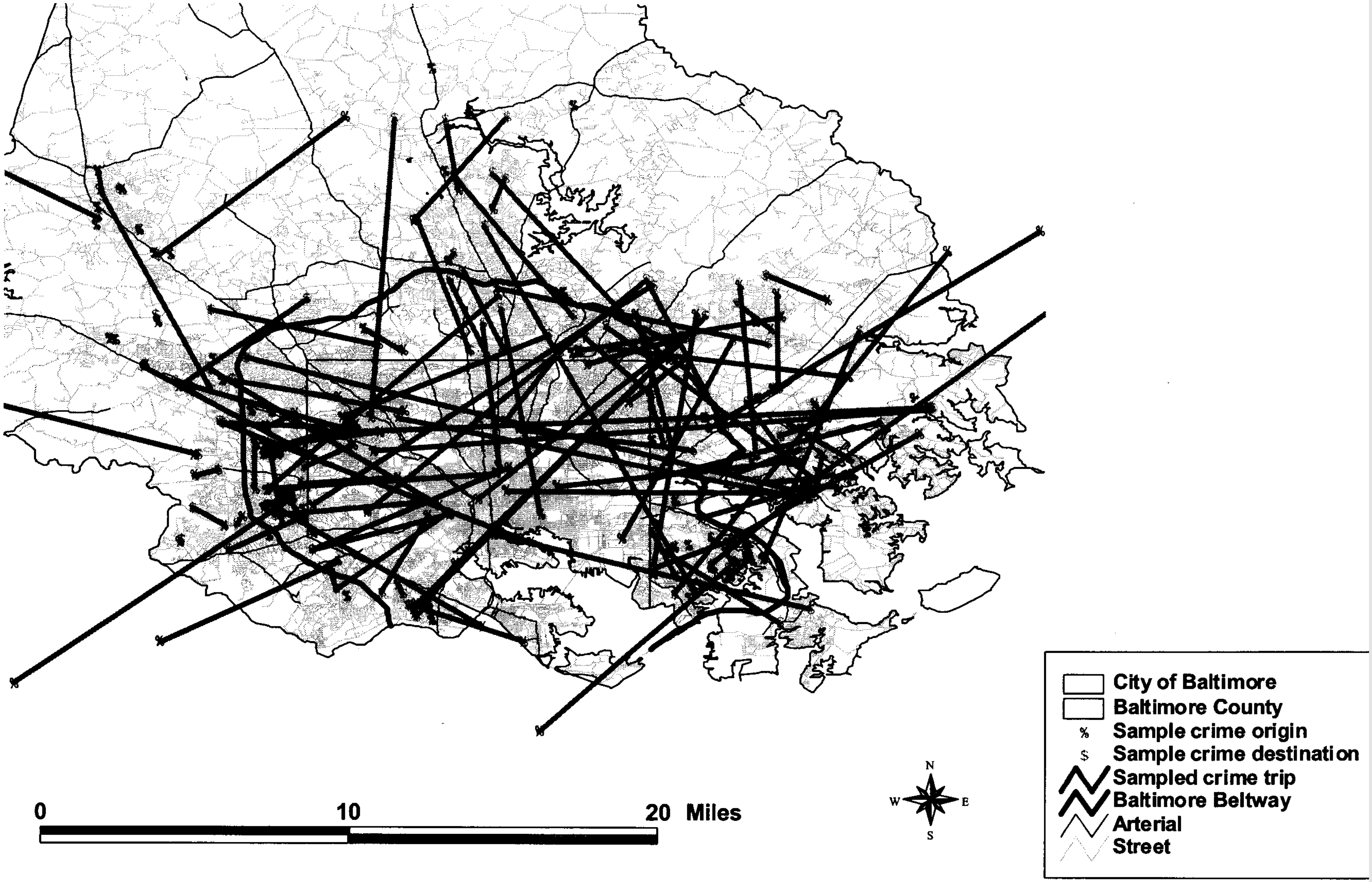
What can one do with data like this?



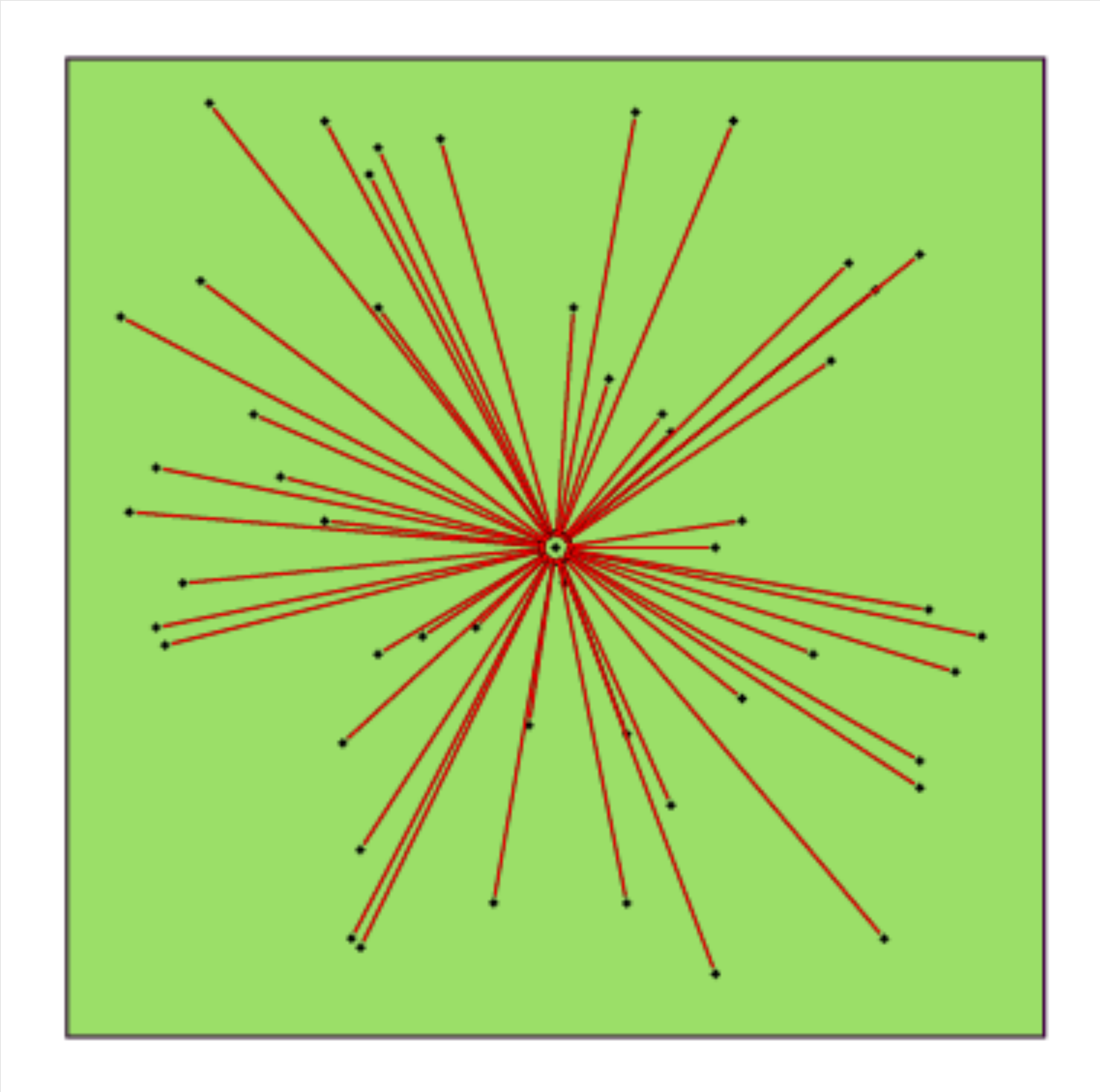
Marital Selection in Seattle

R. Morrill, F. Pitts, 1967, Annals, AAG, (57,2 , 401-422)

Baltimore County Crime Trips: 1993-1997
Origins and Destinations
Sample of 200 Crime Trips

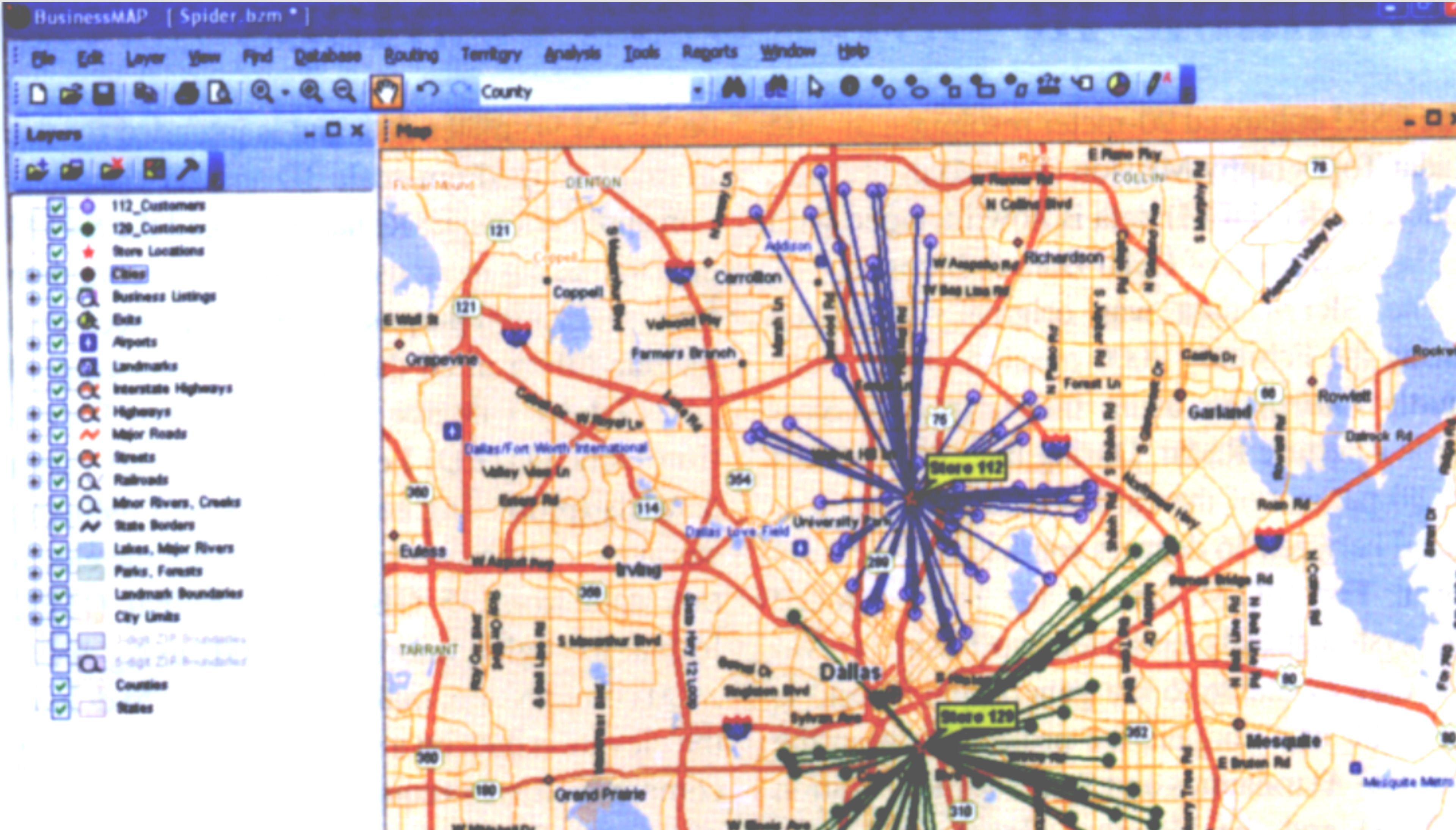


Slide all trips to a central point, keeping their directions fixed, and measure distances.

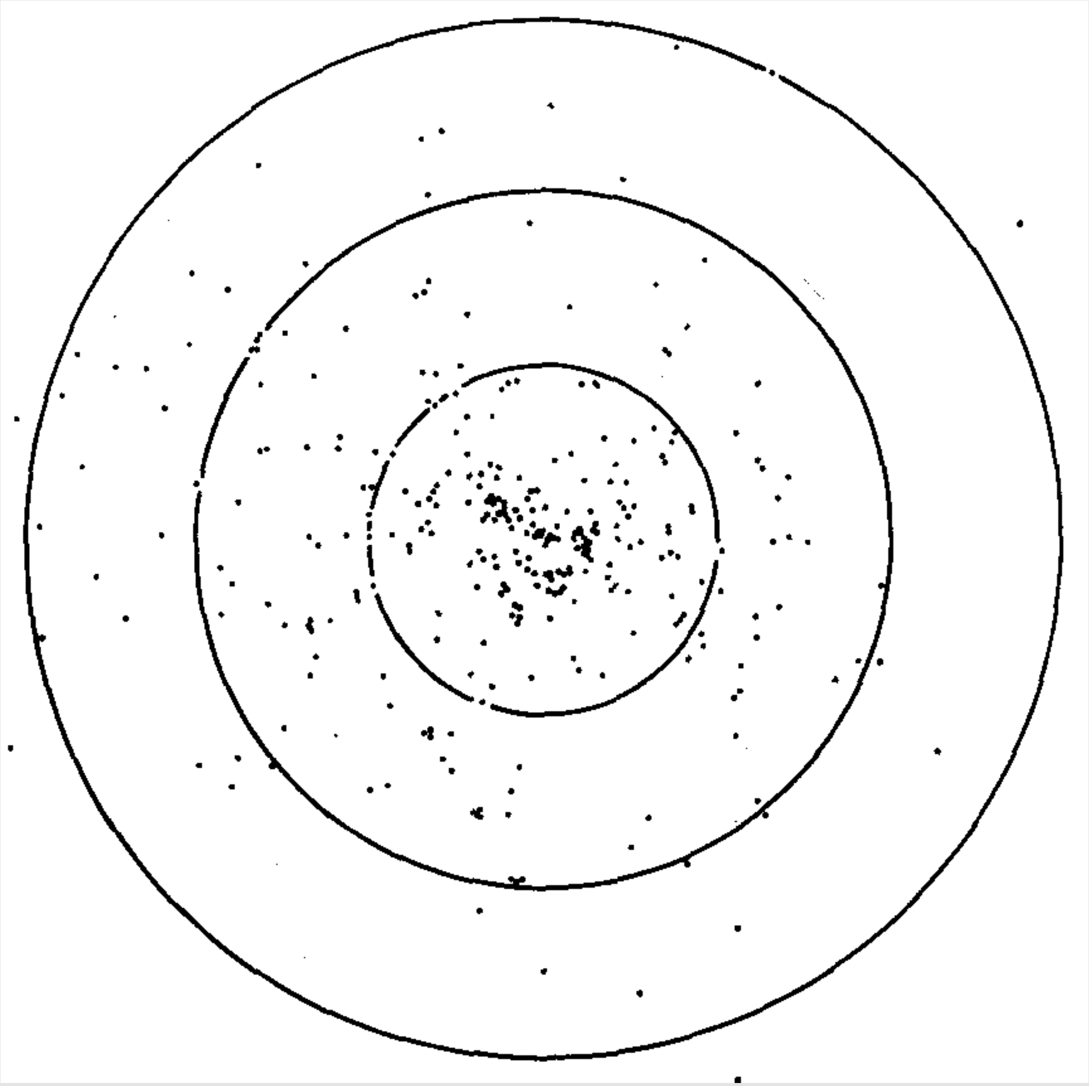


Store Customers

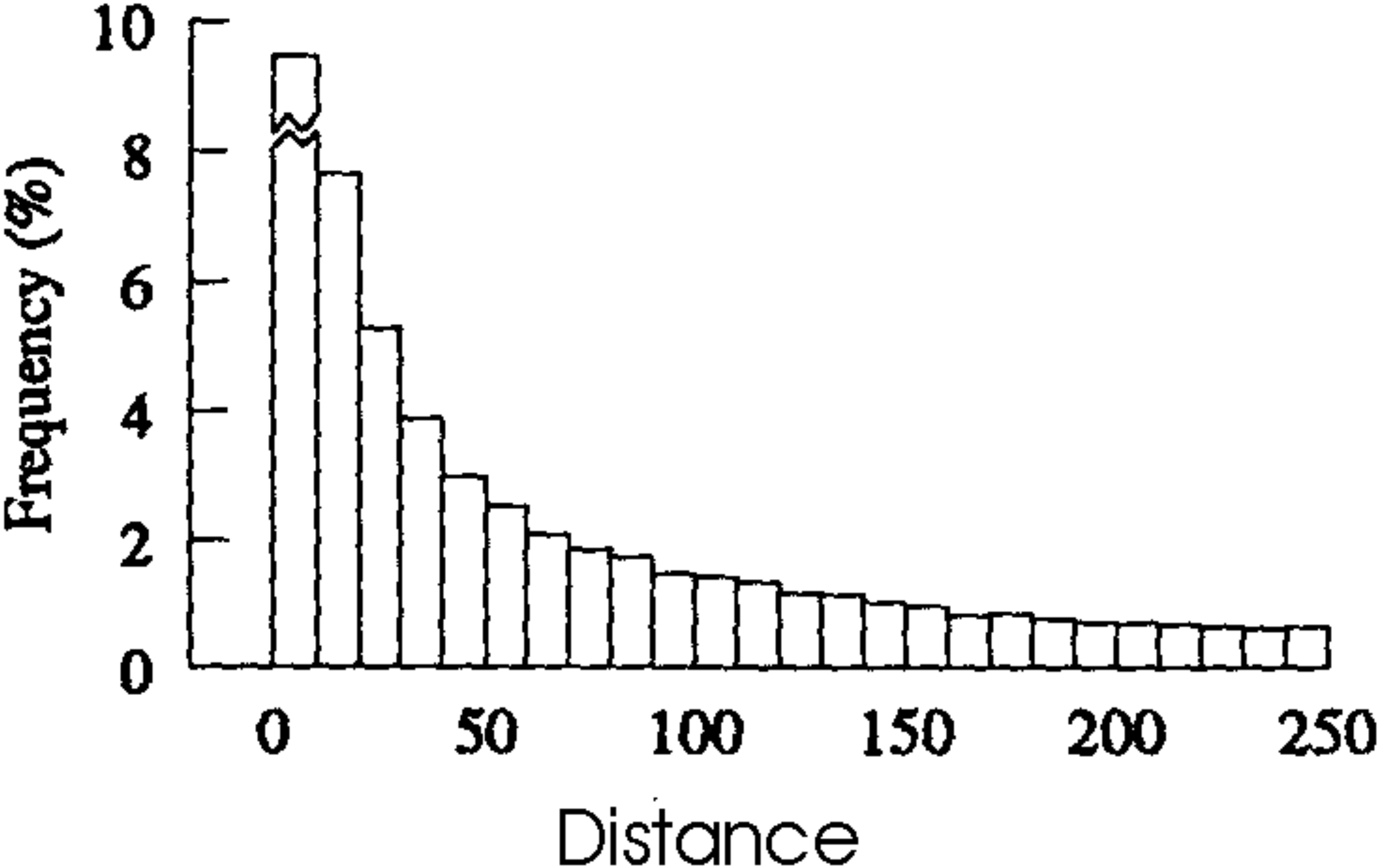
(From an ESRI business program)



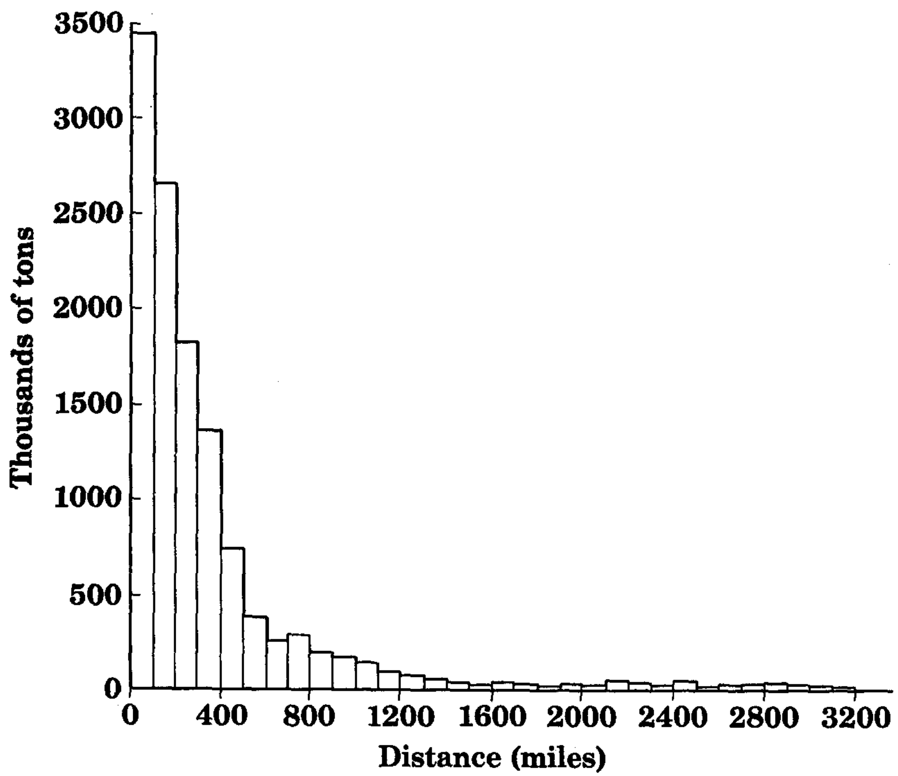
Display end points in distance rings, estimate densities, generate histogram. Produce probabilities.



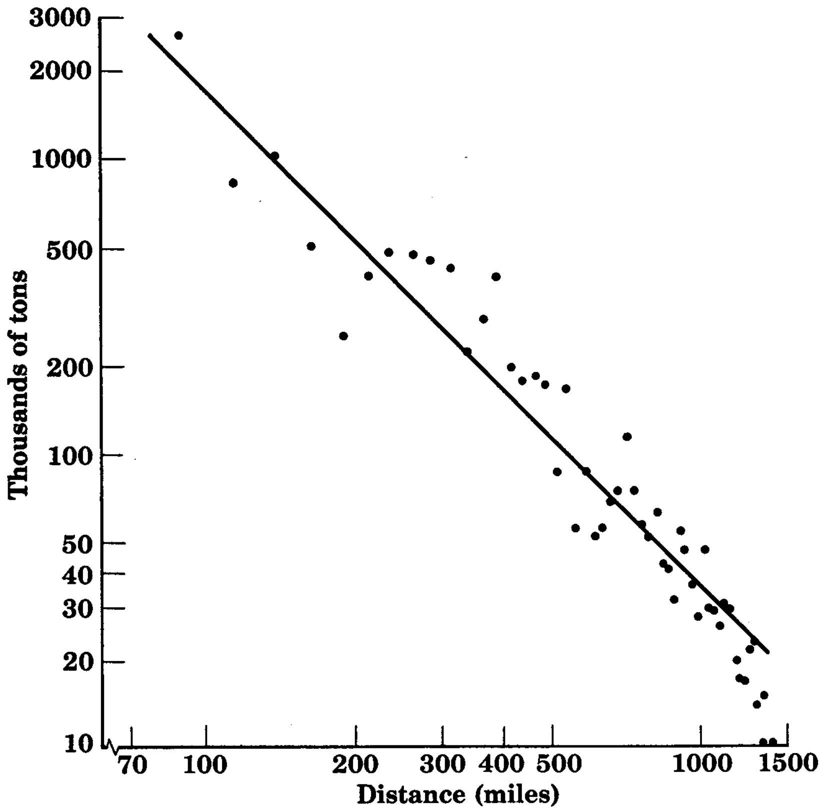
Sort by distance to see distance decay.



The friction of distance



U. S. A. Class I railroad shipments. Tonnage of all commodities, by distance shipped (100-mile zones), 1949.



U. S. A. Class I railroad shipments. Tonnage of all commodities, by distance shipped (25-mile zones), 1949.

There are several simple implications of this 'friction' as shown by the distance decline.

One is to define trade areas or areas of influence.

Another implication is the ability to estimate movement (migration, communication, etc.) between places.

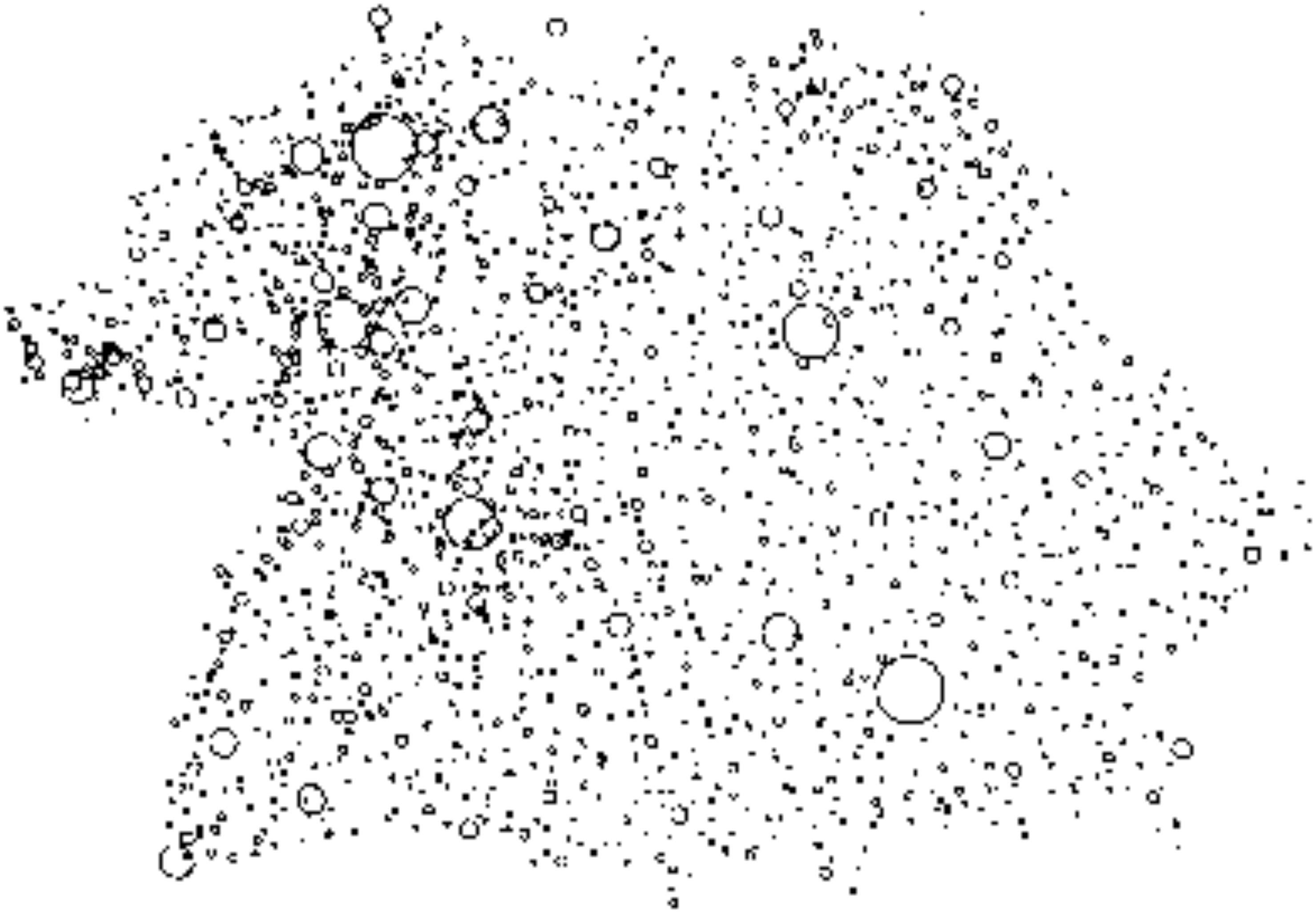
Construct mean information fields for Markov simulations.

Or consider two places at different locations: where the curves cross (intersect) is the place of equal influence.

This is easily extended to the spatial case of influence zones and leads, inter alia, to central place theory.

Central Places in Southern Germany

After Christaller



Geographers, sociologists, economists, regional scientists, and others, have used models to take these distance effects into account.

Most common is the so-called gravity model.

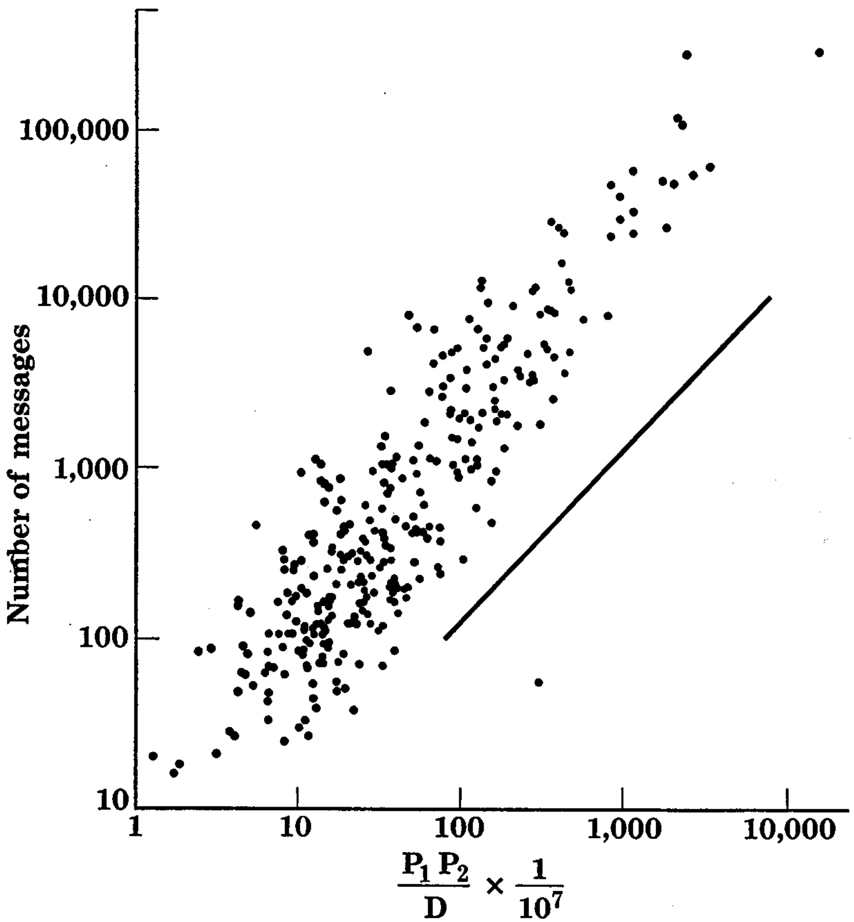
In this model the movement from place i to place j is proportional to the sizes of the places and inversely proportional to the distance between the places:

$$M_{ij} = kP_iP_j/d_{ij}.$$

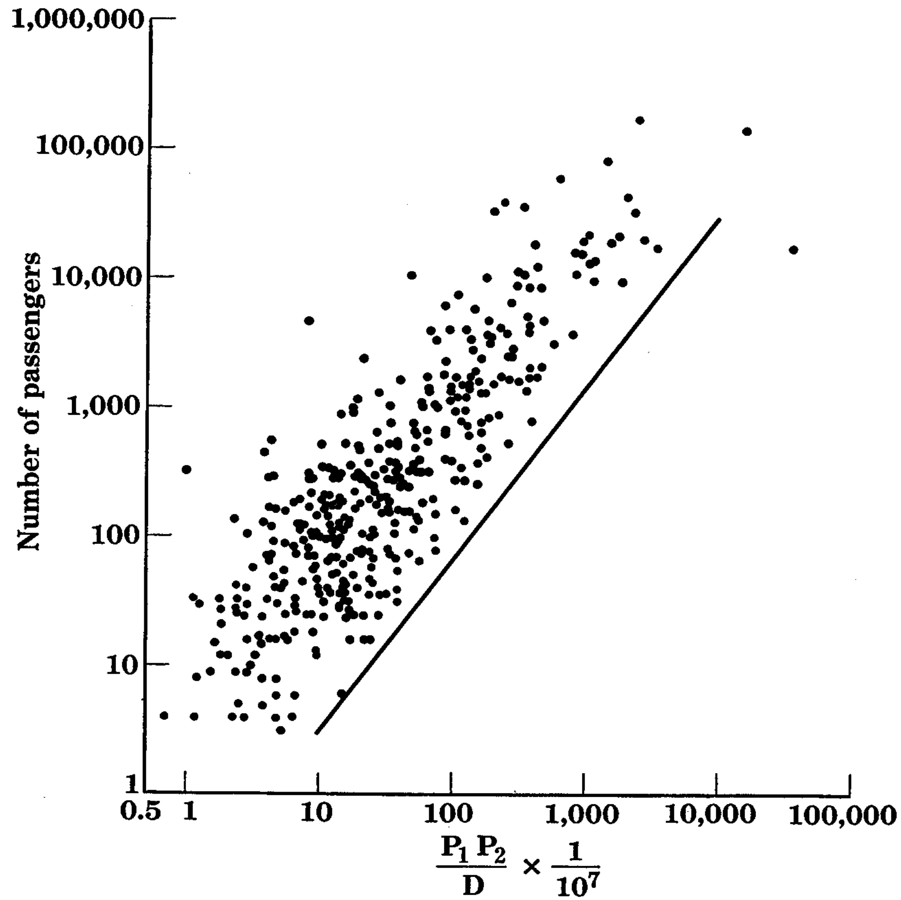
Use is also made of the entropy variant:

$$M_{ij} = kA_iB_jO_iD_j \exp(-ß\ d_{ij}).$$

How well does it predict?



Telephone messages. Number of messages interchanged between 311 arbitrary pairs of cities in the U. S. A., 1940. (The line has a slope of 1.00.) (Source: G. K. Zipf, *Human Behavior and the Principle of Least Effort*, Addison-Wesley, Cambridge, Mass., 1949.)



Bus passengers. Movement of persons by highway bus between 29 arbitrary cities in the U. S. A. during intervals in 1933 and 1934. (The line has a slope of 1.25.) (Source: G. K. Zipf, *Human Behavior and the Principle of*

These models have long been used by geographers to estimate parameters relating to migration.

There is a very large literature on these simple models.

These are mutiplicative models. Thus it is difficult to aggregate either areas or populations.

The ‘gravity’ model uses only distances and populations.

Gravity model: $M_{ij} = kP_iP_j/d^{\beta}_{ij}$

$\ln M_{ij} = \ln k + \ln P_i + \ln P_j + ß \ln d_{ij}$

How well does it work? Typical R^2s are better than 0.8

The ‘entropy’ model uses distances and needs an estimated total cost constraint.

Entropy Model: $M_{ij} = A_iB_jO_iI_j\exp(ßd_{ij})$

This model is estimated using an iteration on balancing factors A and B.

It is often used with previous data and for forecasting.

Both models can make estimates when given only table marginals.

Continuing with modeling

A common technique used to “explain” migration is a multiple regression.

$$M_{ij} = \beta X + \varepsilon$$

β is a vector of estimates relating to the several postulated “causes” X.

The error term ε is minimized by the least squares technique.

Some of the many “causes” are properties of the i^{th} place, others of the j^{th} place, others are of the differences between the places.

Here properties of the migrants themselves are typically not modeled. Instead different regressions are applied to difference classes of movers.

The list of the causes (X’s) is chosen in advance, on the basis of some theoretical conjectures, is often rather long, but can never be exhaustive.

Also notice that no spatial auto-correlation is assumed in the above equation.

Thus one version of an alternate model would need to be written as

$$M_{ij} = RhoWX + \beta X + \varepsilon$$

But we can show a map that provides clear evidence of auto-correlation.

There is a great deal of spatial coherence in migration patterns

In the US case the state boundaries hide the effect, therefore they should be omitted.

There is also temporal coherence.

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

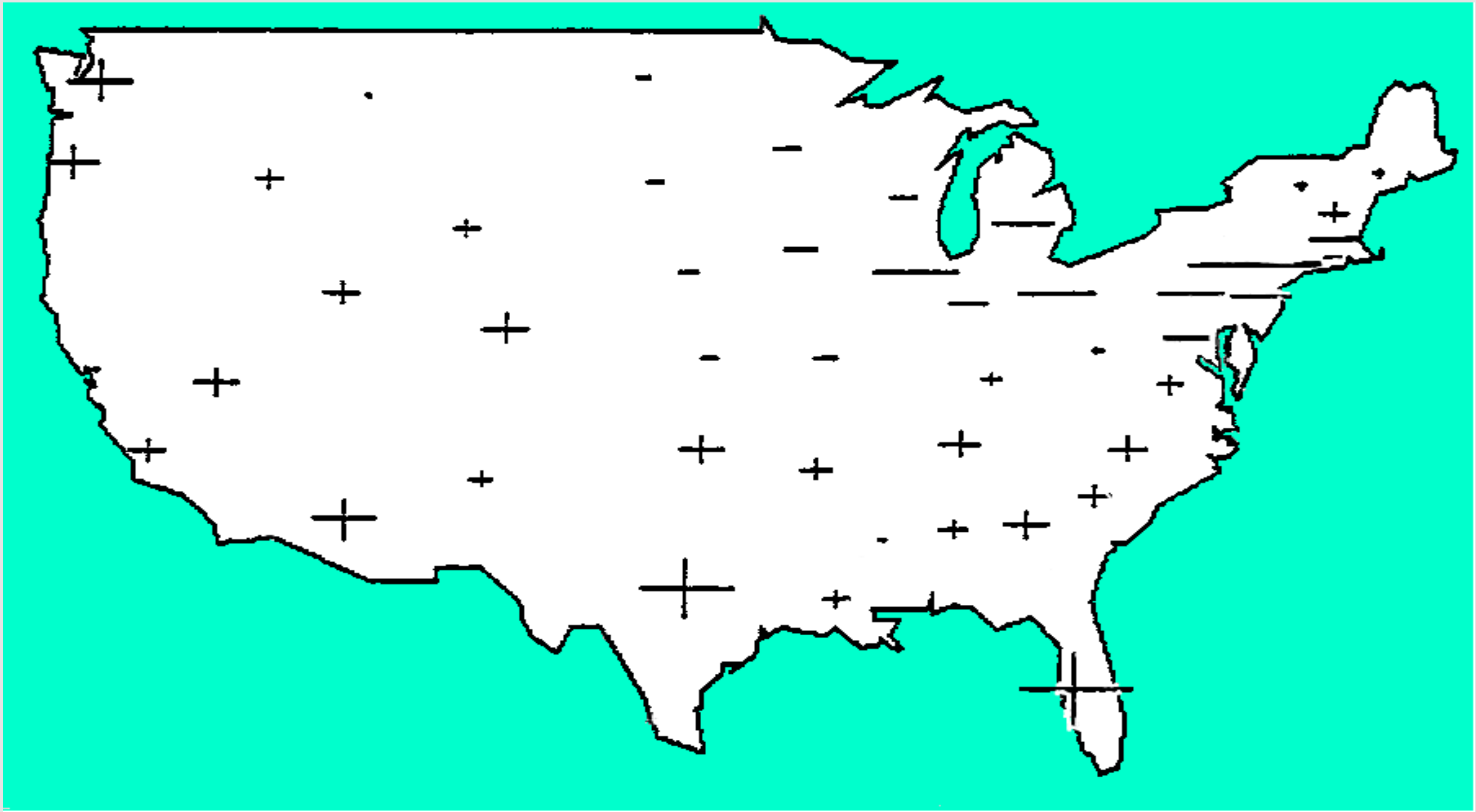
Gaining and loosing states.

Based on the marginals of a 48 by 48 state migration table.

1965-1970 data

The arrangement of leaving and arriving places shows spatial coherence.

It is clear from this map that states are not the appropriate size for migration studies.

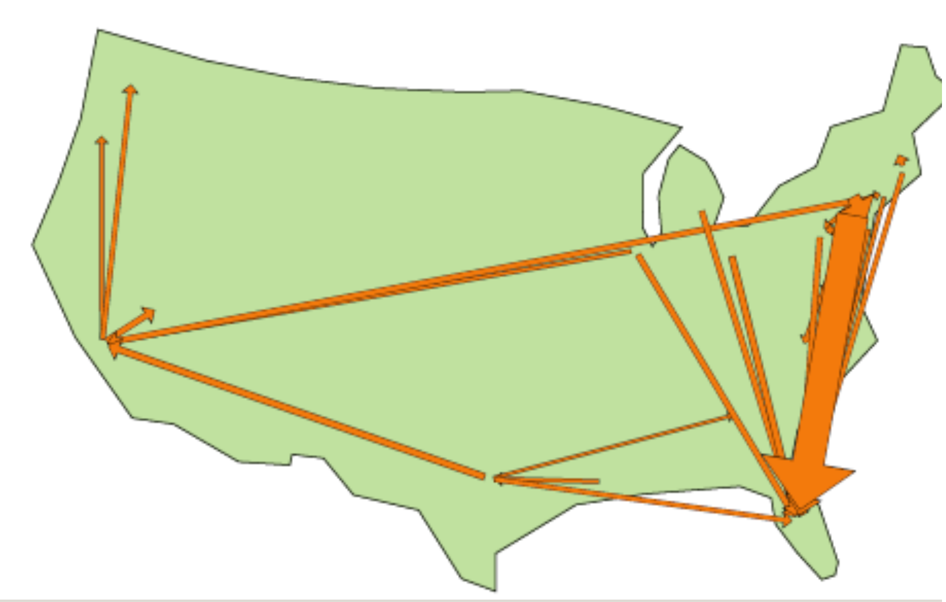


Net Migration in the United States

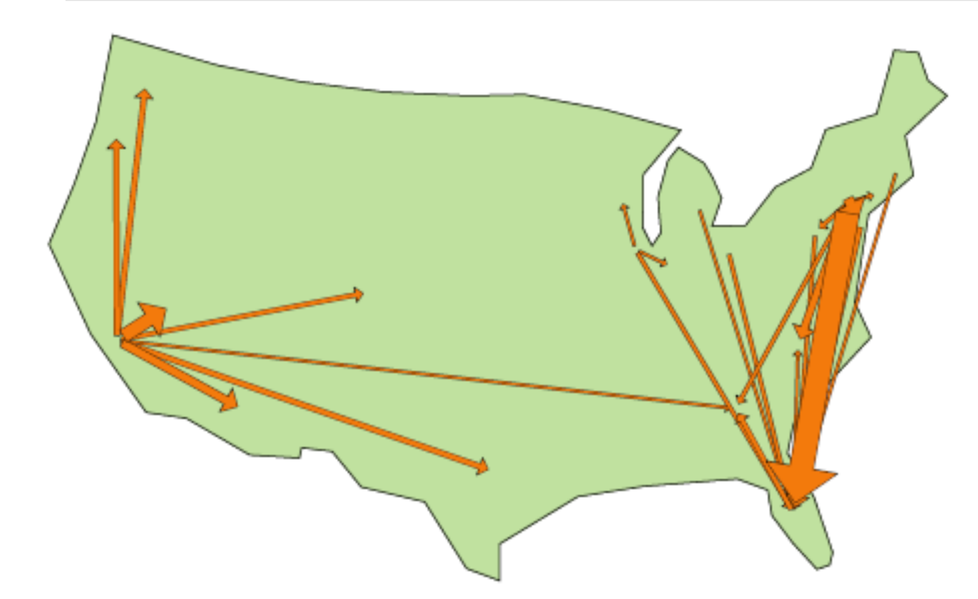
Migration patterns persist for a long time.

Thus there is temporal coherence (auto-correlation) .

1985-1990



1995-2000



Migration Patterns Persist

the Netherlands

1984

1994

Why do people migrate?

A reference is P.N. Ritchey, 1976, "Explanations of Migration", *Annual Review of Sociology*, 363-404.

Migrants move for many reasons, some obvious, some idiosyncratic.

The US Census bureau gives a 68 by 20 table by classes of individuals with 20 reasons for moving,. The last is "other."

The model to be described abstracts from specific reasons and asserts that migration results from some dissatisfaction, discouragement, rejection **r**epulsion, or push ☹ from one's current location and an offer, opportunity, allure, temptation, fascination, **e**nticement, or pull ☺ from some other region.

All modulated by the difficulty of transferring from one place to another.

As described earlier an often used strategy is to postulate 'reasons' in advance for the moves, or to specify attributes of the places thought to influence migration.

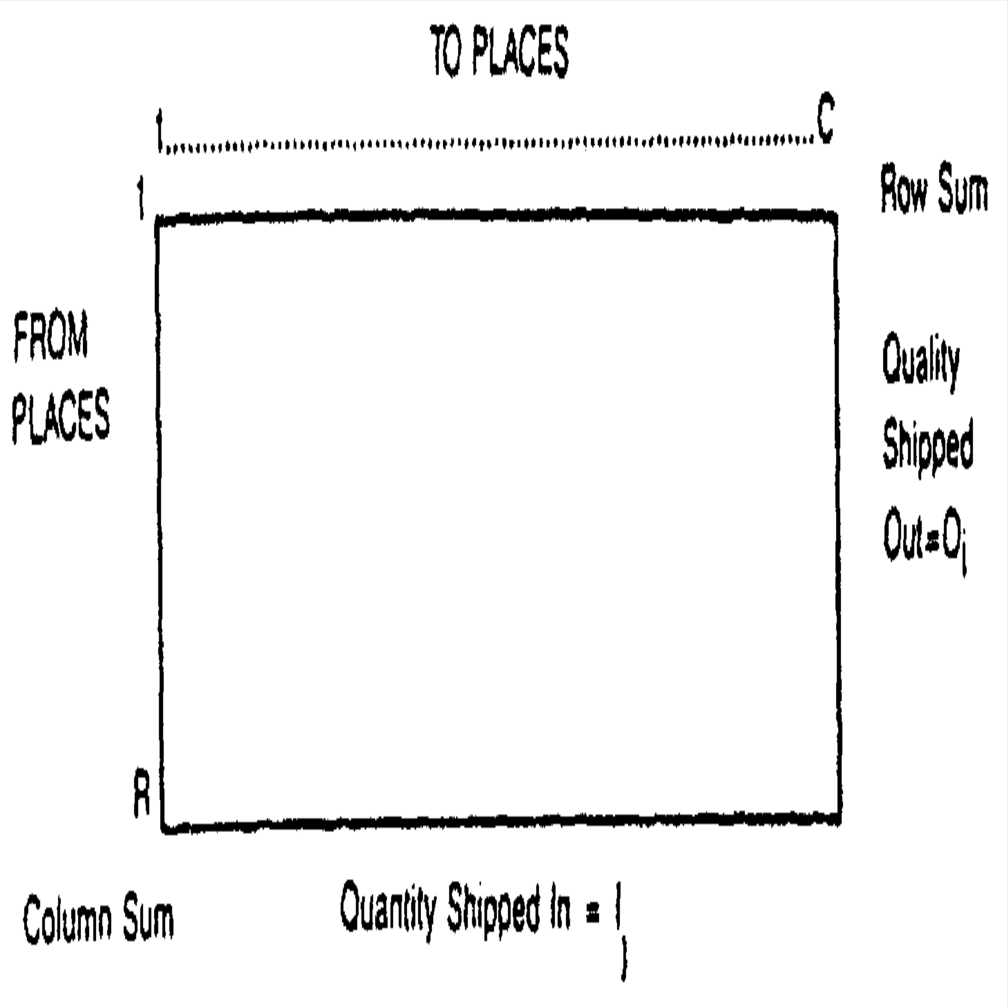
Then parameter estimates are made using a regression model, with its limitations.

Here something different is done!

Pushes (☹) and pulls (☺) **are are not chosen in advance but rather are estimated from the actual migration**. The challenge is then to compare these calculated pushes and pulls to values of postulates.

The form of the movement tables M_{ij}

In the case of migration from place i to place j these are square non-symmetric tables



Derivation of a particular model of migration.

Among the mathematical analyses briefly mentioned are the gravity and entropy models. Here I wish to consider another, less frequently used, model: known as the **Quadratic Transportation Problem** (QTP).

This model is set up as an optimization.
It is generally solved by numerical iteration techniques.

We are dealing with a flow table describing M_{ij} that indicates the movement between places i and j.
Particularly important are the marginal sums since these are all that is used in the model.
But the distances between the places are also assumed to be known.

In this model flows are rendered more reliable by a diversity of moves,
traffic is diverted to avoid congestion, and
migration patterns are rendered diffuse due to information inadequacies.

The quadratic model has another property that is worth mentioning: The additive nature allows separate computation by, say, age groups, to sum to the correct total.

The model used is known as

The Quadratic Transportation Problem

M_{ij} is movement from i to j, D_{ij} is cost, r = rows, c=columns, n=r*c.

The objective function is to minimize

$$\sum_{i=1}^{r} \sum_{j=1}^{c} M^2_{ij} D_{ij}$$

subject to

$$\Sigma_{j=1}^{n} M_{ij} = O_i \text{ (outsum)}$$

and

$$\Sigma_{i=1}^{n} M_{ij} = I_j \text{ (insum)}$$

and

$$M_{ij} \geq 0, \quad \Sigma_{i=1}^{r} O_i = \Sigma_{j=1}^{c} I_j$$

With Lagrangian multipliers the QTP becomes

$$\text{Min } \varepsilon^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} M^2_{ij} D_{ij} + \sum_{i=1}^{r} R_i (O_i - \sum_{j=1}^{c} M_{ij}) + \sum_{j=1}^{c} E_j (I_j - \sum_{i=1}^{r} M_{ij})$$

Setting the appropriate derivatives to zero yields

$$M_{ij} = \tfrac{1}{2} (R_i + E_j) / D_{ij}$$

Observe that this is an additive model, whose parameters still need to be calculated.

The model says that:

Movement from i to j equals

Push from i, plus Pull from j, both divided by distance between i and j, using R for 'repulsing' (☹ = push) and E for 'enticing' (☺ = pull).

For the complete derivation see

G. Dorigo, W. Tobler, 1983, "Push Pull Migration Laws", *Annals,* Association of American Geographers, 73(1): 1-17.

The solution, in equation form, is $M_{ij} = (R_i + E_j) / D_{ij}$

The model says that Movement from i to j equals Push from i, plus Pull from j, both divided by Distance between i and j, using R for 'repulsing' (☹ = push) and E for 'enticing' (☺ = pull).

Here the movement (M) takes place between regions, N in number, indexed by i and j. The pushes ($☹_i$) and pulls ($☺_j$) are numerical quantities that are estimated in the model.
The cost of movement from location i to location j is contained in the d_{ij} term for which great circle or road distance is often an adequate surrogate.

The estimated quantities are real numbers which can be positive or negative.

There are as many simultaneous equations as movements.

The equations are coupled so that a minor change in one push, or pull, or distance, changes all of the others.

But these individual changes are typically small and reflect the sluggishness, or persistence, in the table and this is also the situation in real migration systems.

Using A = E - R for the "Attractivity" and T= E + R for the "Turnover" yields further important numbers for each pair of places i and j.

For the complete derivation see
G. Dorigo, W. Tobler, 1983, "Push Pull Migration Laws", *Annals,* Association of American Geographers, 73(1): 1-17.

The Lagrangians are obtained from a pair of simultaneous equations, one for each location.

$$R_i \sum_{j=1}^{c} 1/D_{ij} + \sum_{j=1}^{c} E_j/D_{ij} = 2\, O_i$$

$$\sum_{i=1}^{r} R_i/D_{ij} + E_i \sum_{i=1}^{r} 1/D_{ij} = 2\, I_j$$

With E ("pulls" ☺) and R ("pushes" ☹) as parameters. These simultaneous equations can be solved for the pushes and pulls.

<u>These are real numbers which can be positive or negative.</u>

There are as many simultaneous equations as places.

The equations are coupled so that a minor change in one push, or pull, or distance, changes all of the others. But the consequent changes are typically small and reflect the sluggishness, or persistence, in the table and this is also the situation in real migration systems.

Using A = E - R for the "Attractivity" and T= E + R for the "Turnover" yields further important numbers for each pair of places i and j.

The model works like this:

For migration in some geographic area,
given
the out-movements (row sums),
the in-movements (column sums),
and the distances between the places,
the algebra then allows the computation of the full table.

This can then be compared to the actual values, if known.

But, it also gives numerical estimates for the
pushes,
pulls,
attractivity,
and turnover.

An example

Three Estimates of the 1985-1990 Inter-provincial Migration in China made using the qtp Model

The three model versions are

1) $M_{ij} = (Push_i + Pull_j) / d_{ij}$
Normal model ($R^2 = 0.57$)

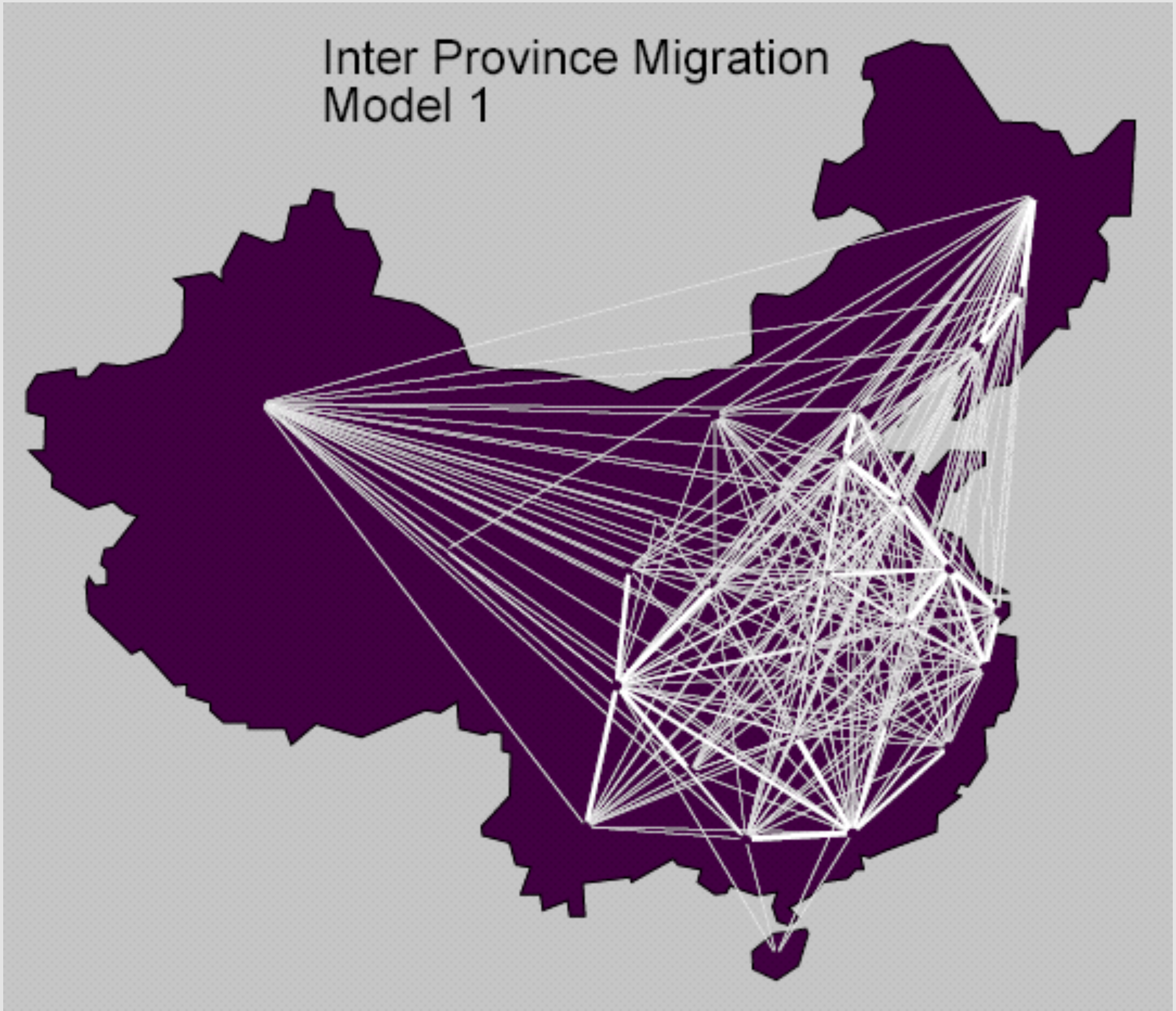
2) $M_{ij} = (Pop_i * Pop_j)*(Push_i + Pull_j) / d_{ij}$
Using the populations ($R^2 = 0.59$)

3) $M_{ij} = (O_i * I_j)*(Push_i + Pull_j) / d_{ij}$
Modulated by the marginals ($R^2 = .65$)

All three models preserve the mean of 1,325,000 migrants and the table marginals.
The computer program used can also use an exponent on the distances, but this has not been done here.
Border lengths between areas may be used instead of distances.
The fit of these models to the actual empirical migration table is given by the correlations.

Model estimate to be compared to the actual

Migration in China estimated by the QTP model



Some more advanced materials.

Instead of using a discrete model, consider a geographically continuous representation.

That is, imagine that the number of places increases almost without bound.

Then do some interpolation of the data to obtain a continuous field.

Four ways of representing geographic space in a continuous fashion for movement studies are:

Writing scalar values as continuous functions of latitude and longitude (or rectangular or polar plane coordinates), perhaps estimated by least squares, as two dimensional algebraic or trigonometric polynomials, splines, eigenfunctions, or spherical harmonics or wavelets. This can be considered as an elaboration of spatial trend analysis. See:

S. Angel, G. Hyman, 1976, Urban Fields, Pion, London.

W. Tobler, 1969, Geographic filters and their inverses, *Geographical Analysis*, 1:234-253

W. Tobler, 1992, Preliminary representation of world population by spherical harmonics, *Proc. Natl. Acad, Sci USA* , 89: 6262-6264.

T. Puu, M. Beckmann, 2003, "Continuous Space Modelling", 279-320, of R.Hall, Ed., Handbook of Transportation Science, 2nd ed.,Kluwer, Boston.

Writing vector fields, or interaction data, in a similar fashion as a four dimensional spline or polynomial function of the origin & destination location coordinates, or using complex variables. See:

P. Slater, 1993, "International Migration & Air Travel: Smoothing & Estimation" *Appl. Math. & Comp*., 53: 225-234

Expanding regression coefficients in a geographically weighted manner. See

J. Jones, E. Casetti, 1992, *Applications of the expansion method*, Routledge, London

S. Fotheringham, et al, 2002 , *Geographically weighted regression*, Wiley, Chichester.

Approximation by a two dimensional lattice (raster), as in the present study.

A spatially continuous model.

Extension of quadratic transportation problem model leads to a class of spatially continuous potential fields providing estimates of attractivity and turnover, and leading to innovative vector field displays.

The potentials are calculated by solving finite difference versions of Poisson's equation from geographically distributed migration table marginals.

The linear nature of the model allows easy additive superimposition when estimating attribute components.

The details follow.

The necessary steps for a US map.

The model is set up on the basis of interpolating both the total sources (out-flows) and sinks (in-flows) for the contiguous US.

To carry out this operation first 'rasterize' the region of interest into a large set of equally spaced nodes.

Assign the total values to the nodes for the computation.
Take the difference to get the net change.

Then the potential is computed.

The gradient field is obtained from this potential field.

This is illustrated on the next several slides for movement of money in the United States.

Example

In the United States the currency indicates where it was issued.

For bills this is the Federal Reserve District.

The new quarters refer to a state.

Coins also contain a mint abbreviation.

Check your wallet to estimate your interaction with the rest of the country!

[In Europe use the Euro coins for cross-country estimates.]

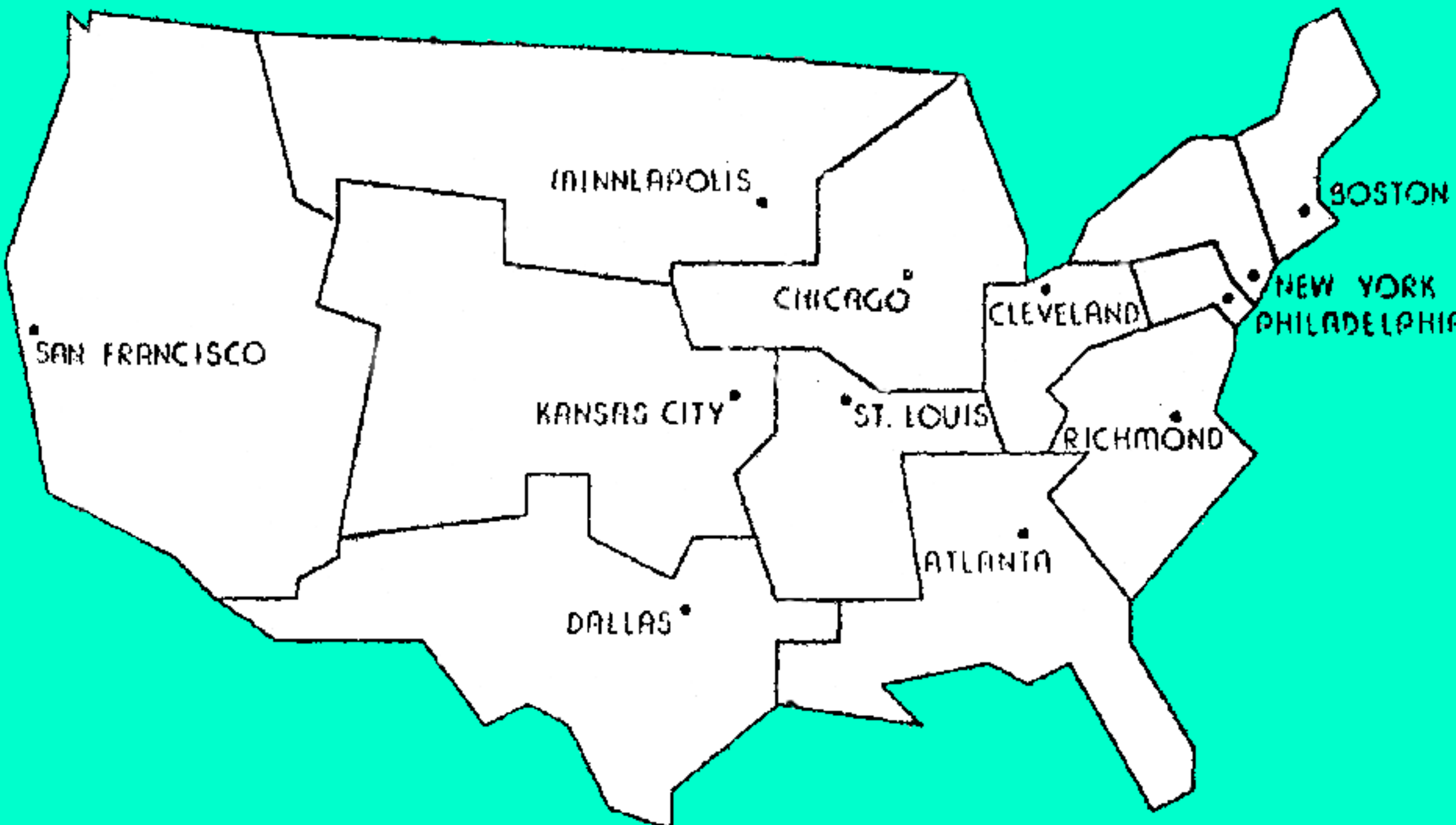
Dollar Bill
(Federal Reserve Note)

Issued by the 8th (St. Louis) Federal Reserve District.

(H is the 8th letter of the alphabet)

The 12 Federal Reserve Districts

(Alaska and Hawaii omitted)



A 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

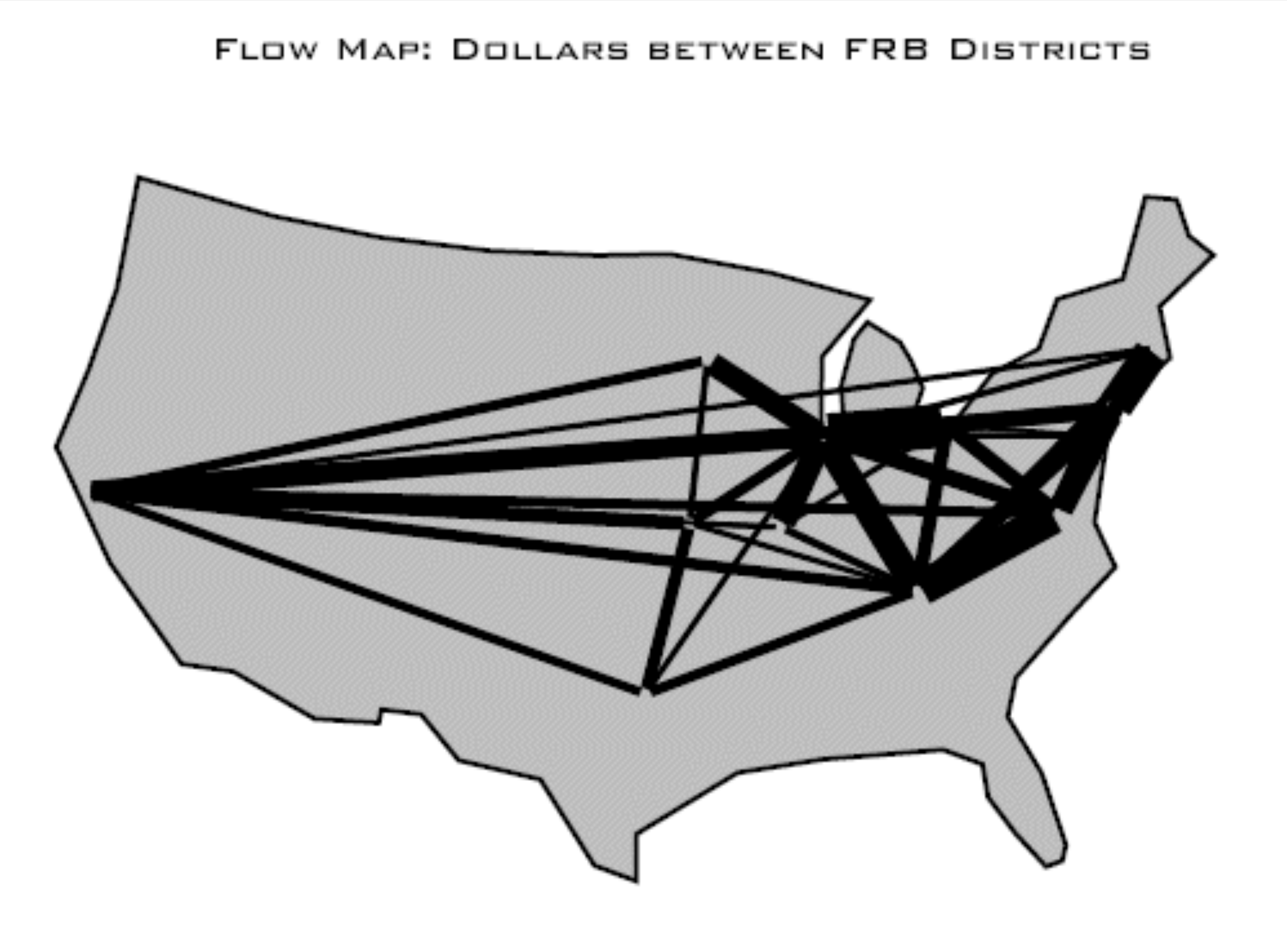
From the table we can compute a movement map.

Movement of One Dollar Notes

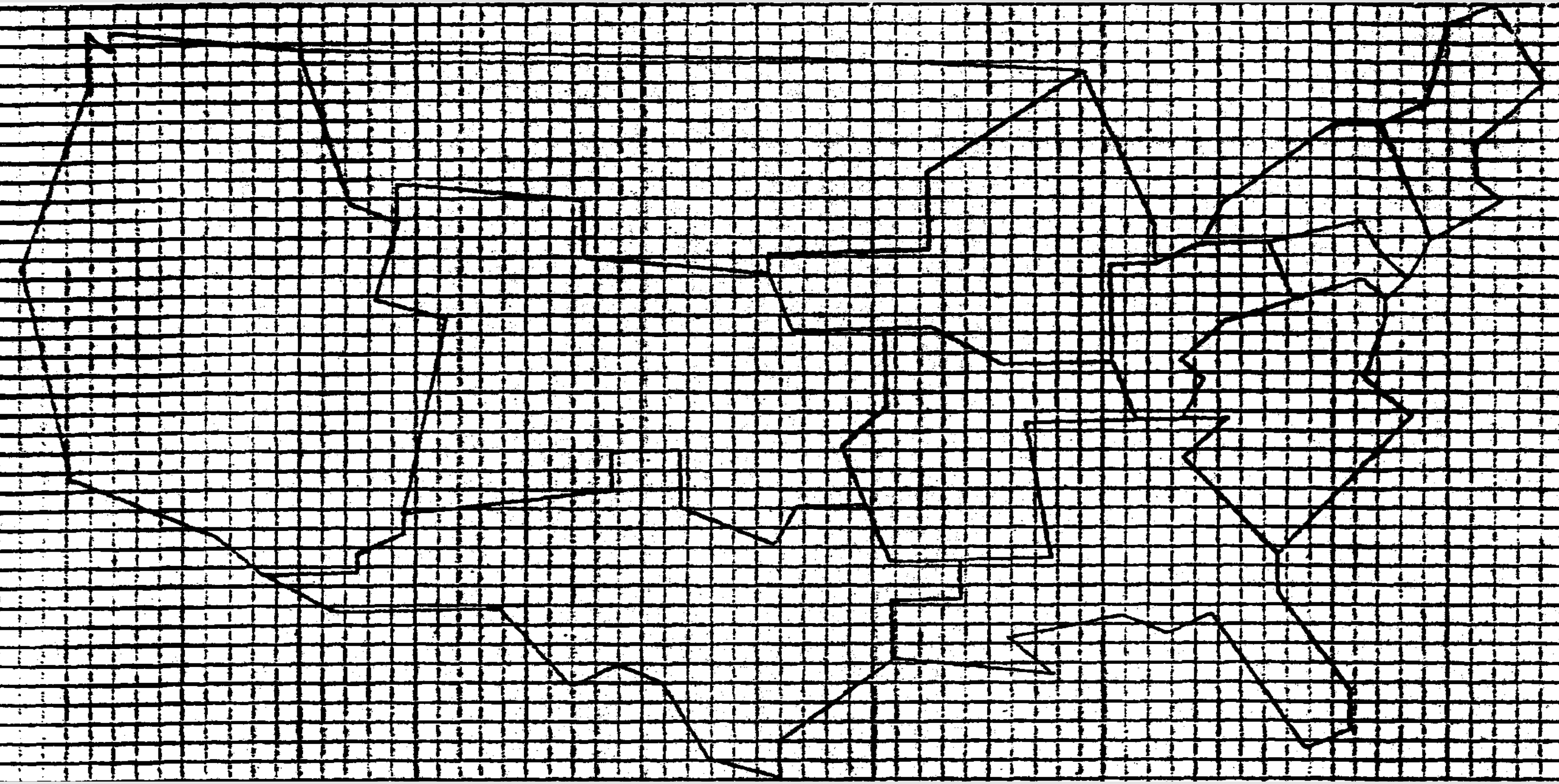
between Federal Reserve Districts, in hundreds, Feb. 1976

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

Dollar moves between FRB centroids.
A conventional map, but we want a continuous field map.



First the Federal Reserve Districts Are "Rasterized"



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

Now spread the total in and out dollar moves from each Federal Reserve District to all cells within the district.

Spread these moves in a rather even fashion.

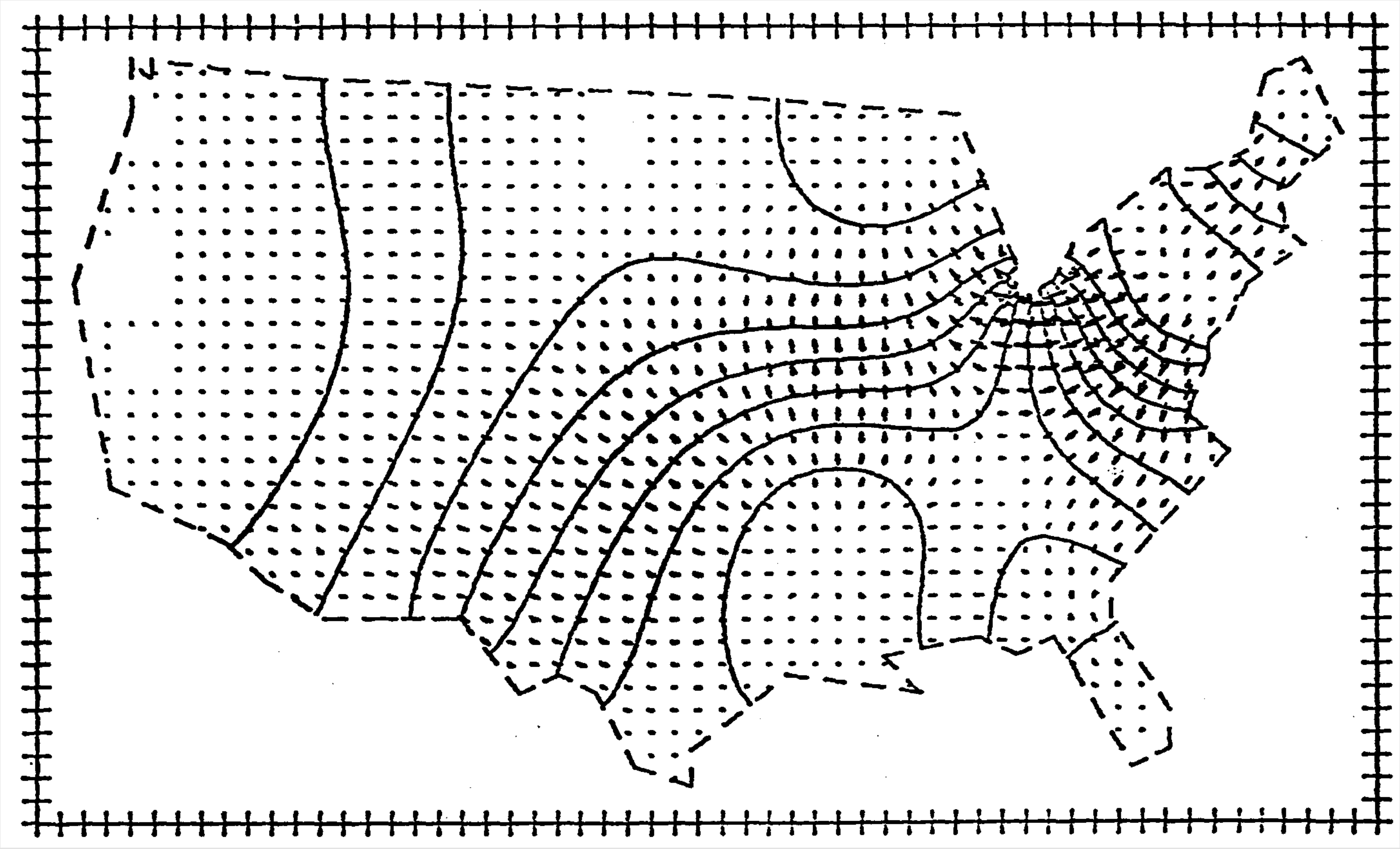
Then subtract the outs from the ins to get the net change in the dollars for every cell.

Lots of cells will get plus signs (dollars arriving) and lots will get minus signs (dollars leaving).

Think of the minus signs as high hills and the plus signs as valleys.

Now let the dollar amounts trickle down from the hills into the valleys, somewhat like topography eroding, to get the next map.

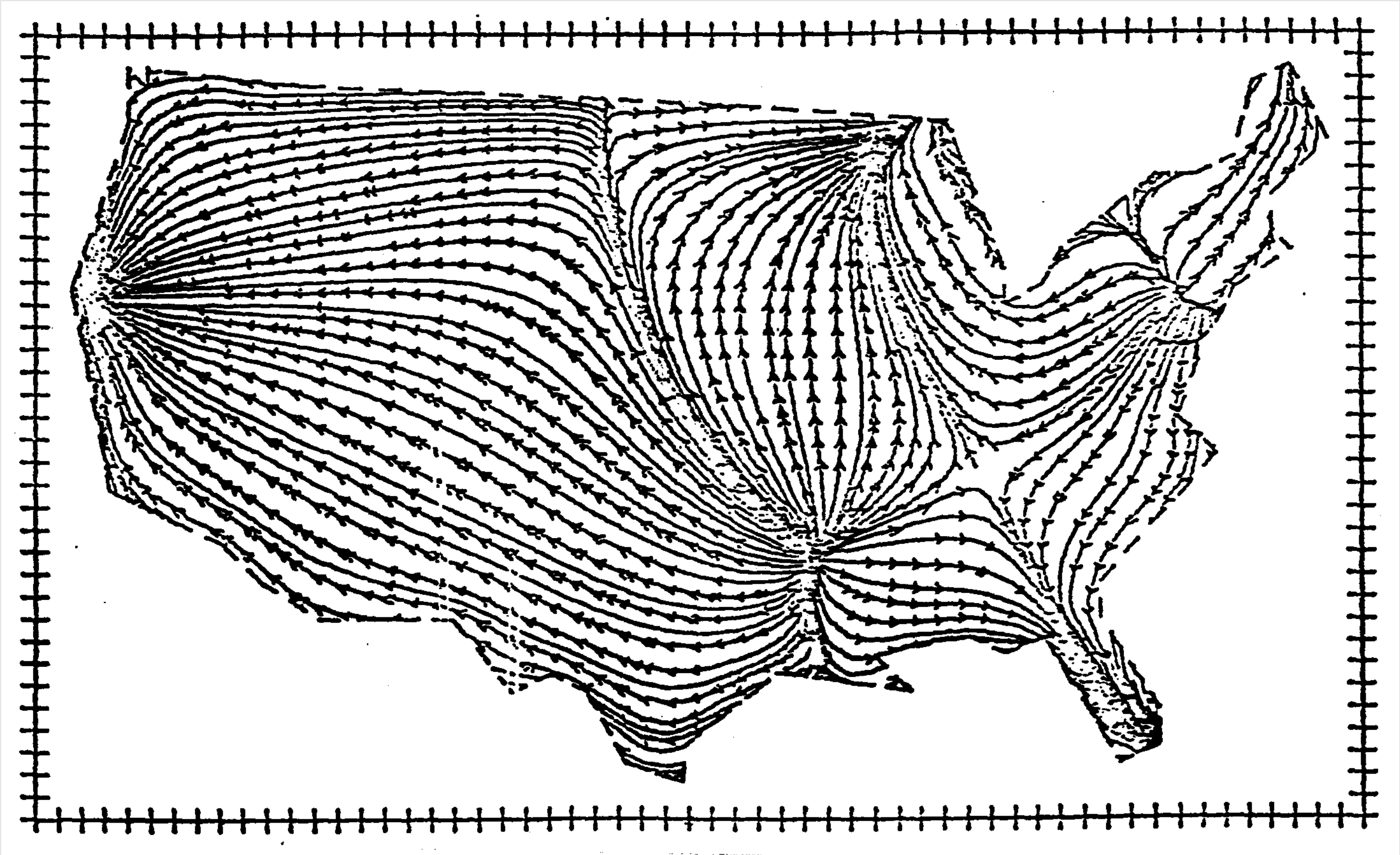
Here the 'hills' are shown by contours and the direction of flow by the slope vectors



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

Now we can represent the flow of dollar bills in the U.S. by a continuous field.

Spatially coherent structures in the movement.



The flowing model just described is really only an analogy of what the equations do.

The map is actually computed using a continuous version of the push-pull 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, as has just been shown.

Some of the analysis details will now be described.

Now I show how to derive the continuous version of the Push-Pull model that was earlier illustrated.

In the discrete case there is one equation for every pair of places:

$$M_{ij} = (R_i + E_j) / D_{ij}$$

obtained by solving the simultaneous pair for the Lagrangians:

$$R_i \sum_{j-1}^{c} 1/ D_{ij} + \sum_{j=1}^{c} E_j / D_{ij} = 2\ O_i$$

$$\sum_{i=1}^{r} R_i / D_{ij} + E_i \sum_{i=1}^{r} 1/ D_{ij} = 2\ I_j$$

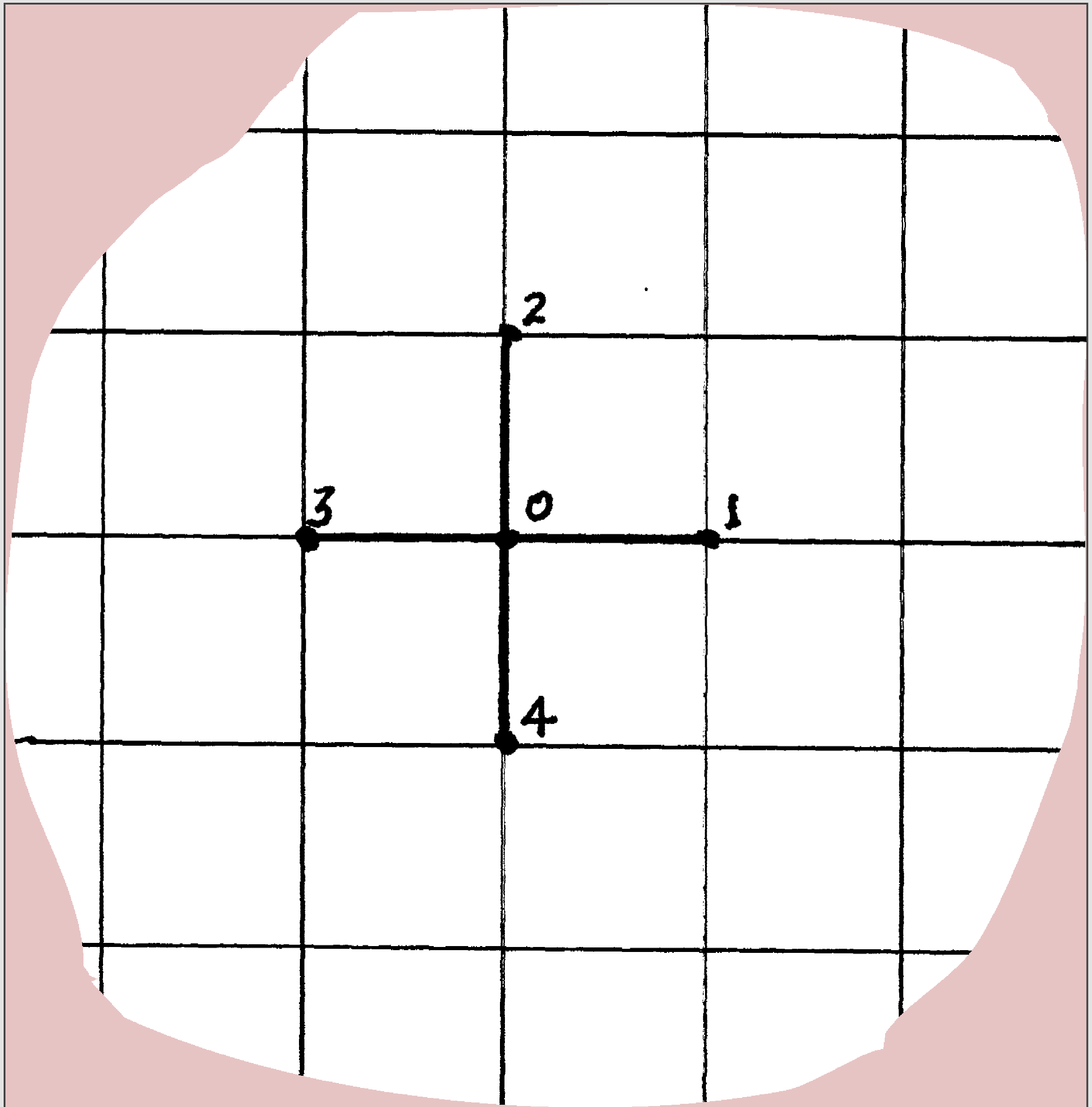
The E ('pulls' ☺) and R ('pushes' ☹) are the Lagrangians.

These simultaneous equations are solved for the pushes and pulls.

Also obtained were the 'Attractivity' A = E - R and the 'Turnover' T= E + R.

In the raster look at one node and its neighbors

A raster is a special kind of network where movement takes place between neighboring nodes



The simple derivation

Recall that in this model $M_{ij} = (R_i + E_j) / D_{ij}$

For the square mesh take all D_{ij} to be the same. Set them equal to 1. The model is now $M_{ij} = R_i + E_j$ Use the subscript 0 for the center node (i=0), and index the neighbor nodes from 1 to 4.

Then the moves **from** the center to the neighbors is

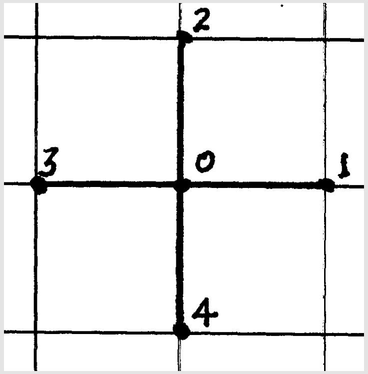
$M_{01} = R_0 + E_1$

$M_{02} = R_0 + E_2$

$M_{03} = R_0 + E_3$

$M_{04} = R_0 + E_4$

$M_{0j} = 4\,R_0 + E_1 + E_2 + E_3 + E_4$



But M_{0j} are the moves out of node 0, and this is O_j the Outsum.

In the same way $M_{10} = R_1 + E_0$, etc for M_{20}, M_{30}, M_{40}.

These are the moves **into** node 0 from the neighbors, and this is I_i.

Thus the pair of equations become

$O_j = 4\,R + E_1 + E_2 + E_3 + E_4$

$I_i = 4\,E + R_1 + R_2 + R_3 + R_4$

after dropping the subscript for the central node.

There is one pair of equations for each node.

An aside:

Incorporating differential transport disutilities into the model.

From the previous slide we can insert a differential transport weight factor into the movement, as follows:

$M_{01} = R_0 + E_1$ becomes $= (R_0 + E_1)/W_{01}$ where W_{01} is the equivalent of d_{01} but more realistic (for example road distance, or travel time or cost). Then similarly for all M_{0j}.

Now do the same for M_{10} inserting a W_{10}, etc. Recognize that W_{01} is not the same as W_{10} and that the link weights will be different across every edge, and that they may change rapidly with time. Adjacent cells will naturally have two common, but differentially directed, link values.

It might be helpful to draw and label weights for a system of nine cells.

Doing this naturally leads to a rather more complicated system of equations.

(aside continued)

As a result:

$R_0 = [O_j - (\Sigma E_i/W_{0i})] / \Sigma (1/W_{0i})$

$E_0 = [I_i - (\Sigma R_j/W_{j0})] / \Sigma (1/W_{j0})$

The summations are over k = 1 to 4

All W_{ij} and I_i and O_j are assumed known.

The same set of equations hold for all cells except those on the borders of the region.

Known are 2 W's per edge + 2*I*O - 1 in and outsums (I's and O's) minus 4*(I + O) (Dirichelet or Neumann) values at the edges.

Unknown are 2*R*E pushes (R 's) and pulls (E's).

Can this system be solved for all R's and E's?

(end of the aside)

The distance values D_{ij}, as constants, have been dropped in the square mesh, for pedagogic purposes, but not a mathematical necessity. As suggested differential transportation can be included but complicates the model and is thus omitted here Each place, except along the margins of the region, will have four neighbors. Now the quadratic equations can be greatly simplified.

Just derived were the two equations at each node:

$$4E = I - (R_1 + R_2 + R_3 + R_4), \qquad 4R = O - (E_1 + E_2 + E_3 + E_4).$$

The central E and R require no subscript; their neighboring locations are indexed from one to four - or if you wish - North, South, East, and West directions.

Now add - 4R to both sides of the first equation and - 4E to both sides of the second, rearrange slightly, and using T = E + R, to obtain

$$R_1 + R_2 + R_3 + R_4 - 4R = I - 4T, \qquad E_1 + E_2 + E_3 + E_4 - 4E = O - 4T,$$

The left-hand sides are recognized as finite difference versions of the Laplacian.

Thus we can write, approximately and for a limiting uniform fine mesh, the pair

$$\partial^2R/\partial u^2 + \partial^2R/\partial v^2 = I(u,v) - 4T(u,v),$$

$$\partial^2E/\partial u^2 + \partial^2E/\partial v^2 = O(u,v) - 4T(u,v),$$

assuming that R and E are differentiable spatial functions and that I and O are continuous densities given as functions of the cartesian coordinates u and v.

In the discrete system there is one equation for every pair of places. On the mesh there are two equations for every node. We have just derived a continuous version of the quadratic model.

In this continuous version, we have a coupled system of two simultaneous partial differential equations covering the entire region. These equations can be combined to yield either gross movements or net movements.

For the simultaneous movement in both directions at each pair of places **add** the two equations to get the 'turnover' potential. The result is Helmholtz's equation.

For the net movement we need only the difference between the 'in' and 'out' at each node for the 'attractivity' potential, as follows:

By **subtraction** from the previous equations, we have

$$\partial^2 A/\partial u^2 + \partial^2 A/\partial v^2 = I(u,v) - O(u,v),$$

where A = E - R ('pull' ☺ minus 'push' ☹) can be thought of as the attractivity of each location.

This is the well-known Poisson equation for which numerical solutions are easily obtained.

The right hand side is the amount of change (In - Out).

Once A(u,v) - the potential - has been found from this equation, the net movement pattern is given by the vector field,

$$\mathbf{V} = \text{grad } A,$$

or by the difference in potential between each pair of mesh nodes.

The described result is a system of linear partial differential equations

These are solved by a finite difference iteration to obtain the potential field (after specifying a boundary condition).

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

In other words, a map is computed using a continuous version of the quadratic transportation problem.

Estimates of the potentials for two different populations (male & female for example) can be added to get the correct potential for the sum.

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.

Next example: starting from a

Migration Table

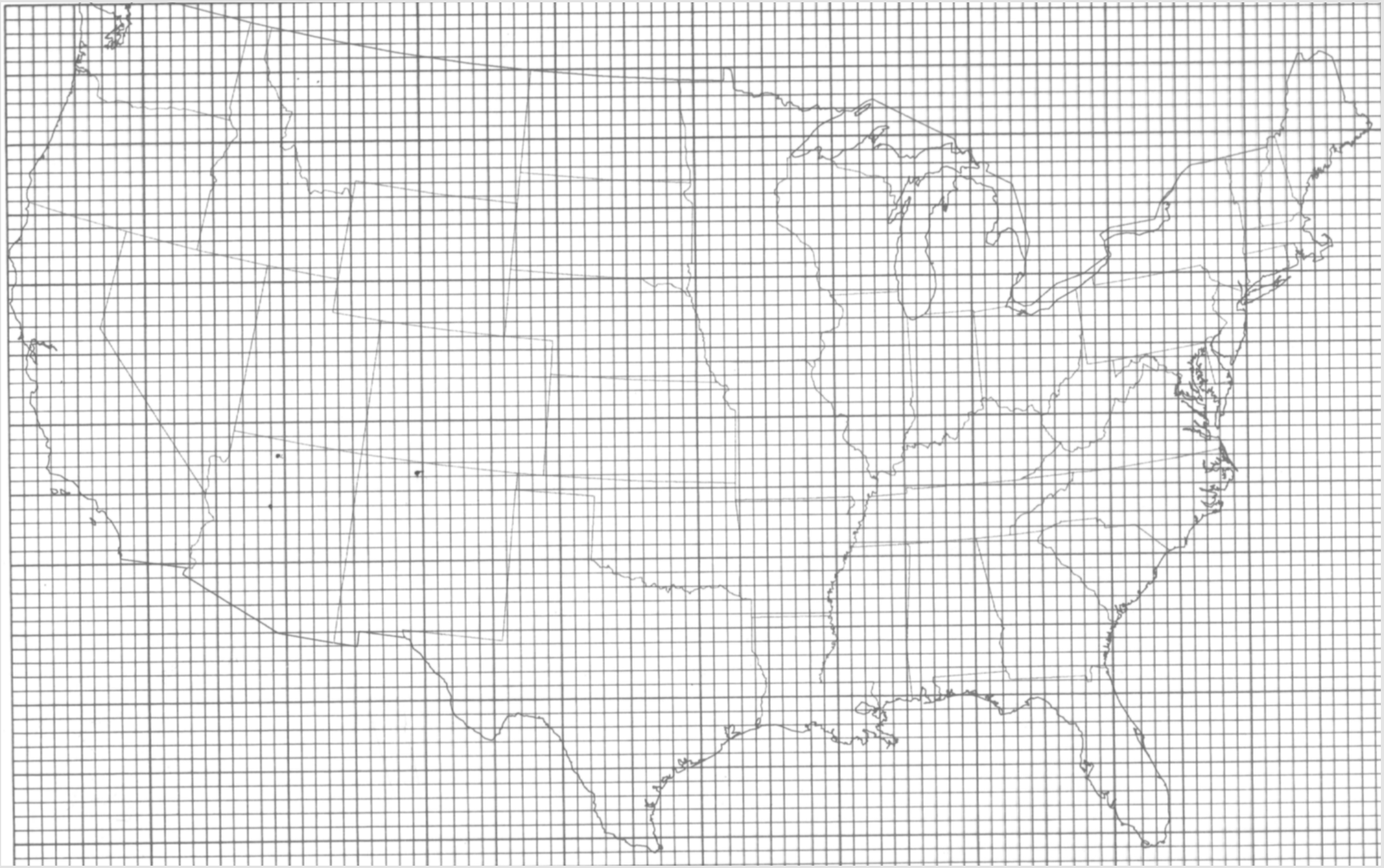
The nine division US census table from 1973 is illustrated.
Note the asymmetry.

	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	—

The example that follows uses the (48 by 48) contiguous state table.

“Rasterize” the USA to form a lattice.

Use a point-in-polygon program to assign nodes to individual states. Then assign in and out values to these nodes. There will be one equation for each node on this raster.
Then solve the system of ~6000 simultaneous equations to yield the potential.

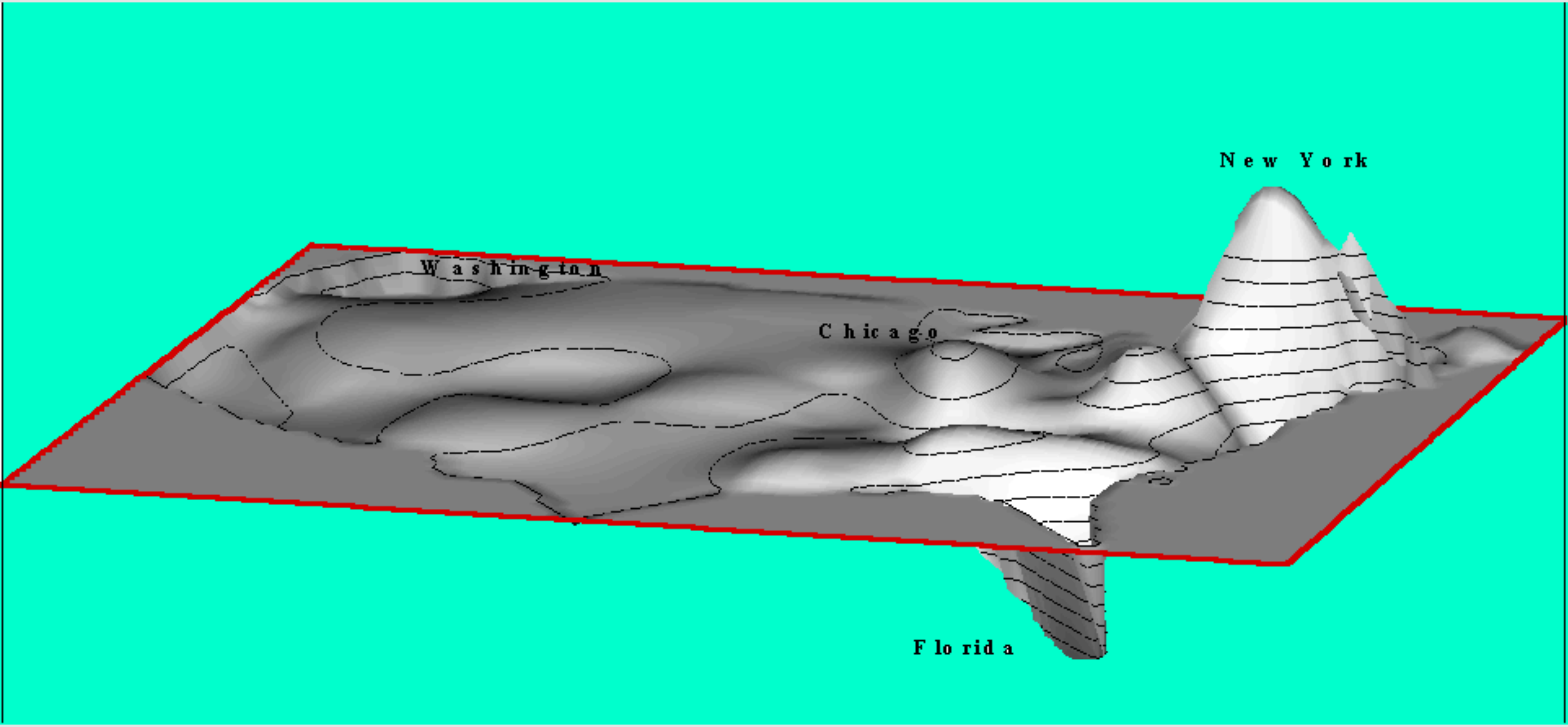


Solving the equations for the potential gave

The Pressure to Move in the US

Based on this continuous spatial quadratic model

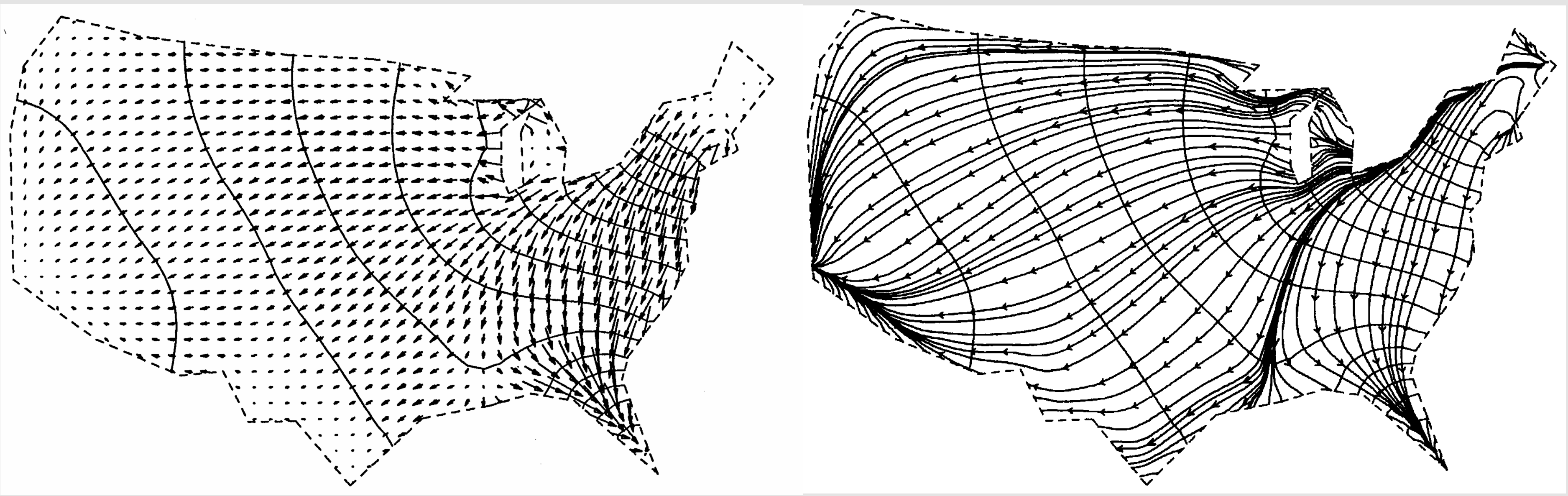
Using state data



The resulting views:

The migration potentials shown as contours

and with gradient vectors connected to give streaklines



In the U.S. example both the in-migration and the out-migration amounts were spread over all of the nodes making up each of the individual states. Pycnophylactic reallocation was used to do this. Then the equations are set up and solved on the raster.

That these migration maps resemble maps of wind or ocean currents is not surprising given that we in fact speak of migration 'flows' and 'backwaters', and use many such hydrodynamic terms when discussing migration and movement phenomena.

The foregoing equations have captured some of this effect in a realistic manner.

The model described here has the interesting property that the amount of activity at a place is given as a weighted linear combinated of the differential activity of all of the places.

The effect of any change at a single place is modulated by the remaining places and dampened in its impact.

This effect is spatial - it is modulated by the distance between places.

That is, changes in migration at one place impacts the migration at neighbors

Thus the potential as here calculated has a great deal of inertia since it reflects the influence of all of the palaces simultaneously.

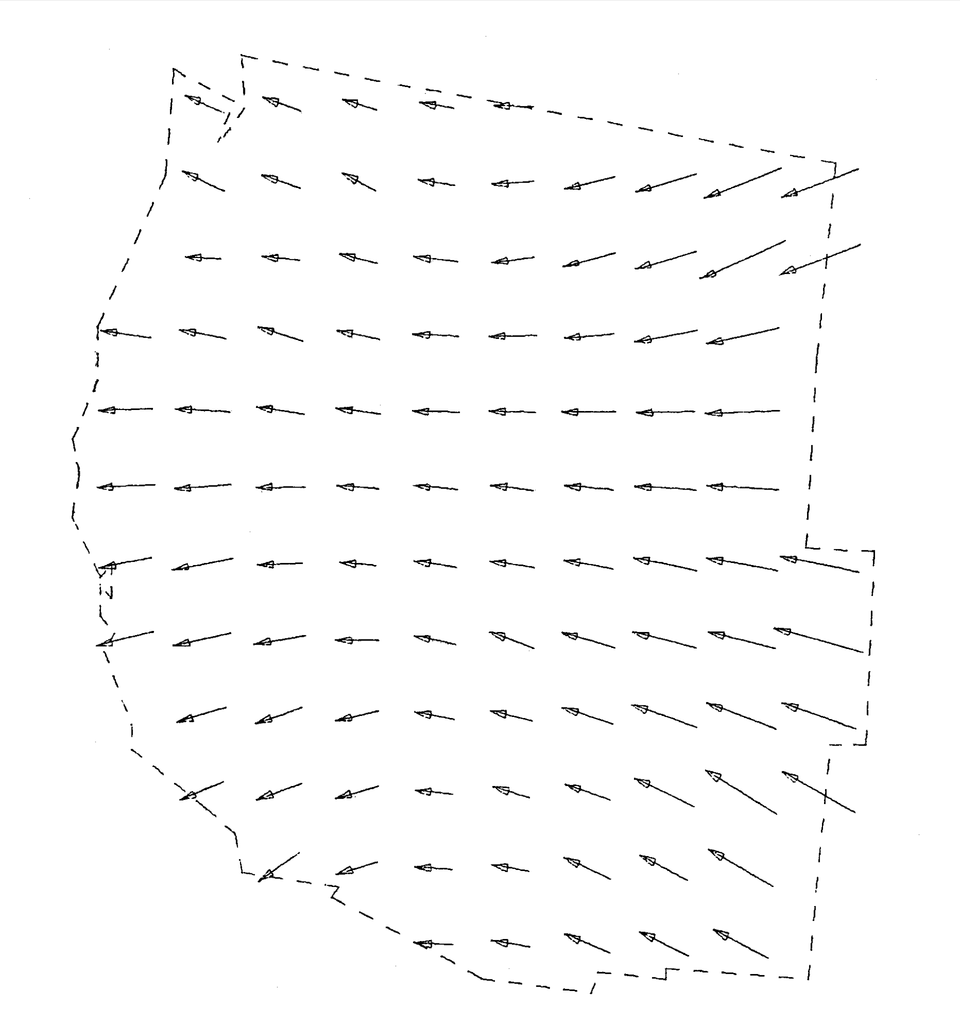
Consequently we expect that it changes only slowly with time.

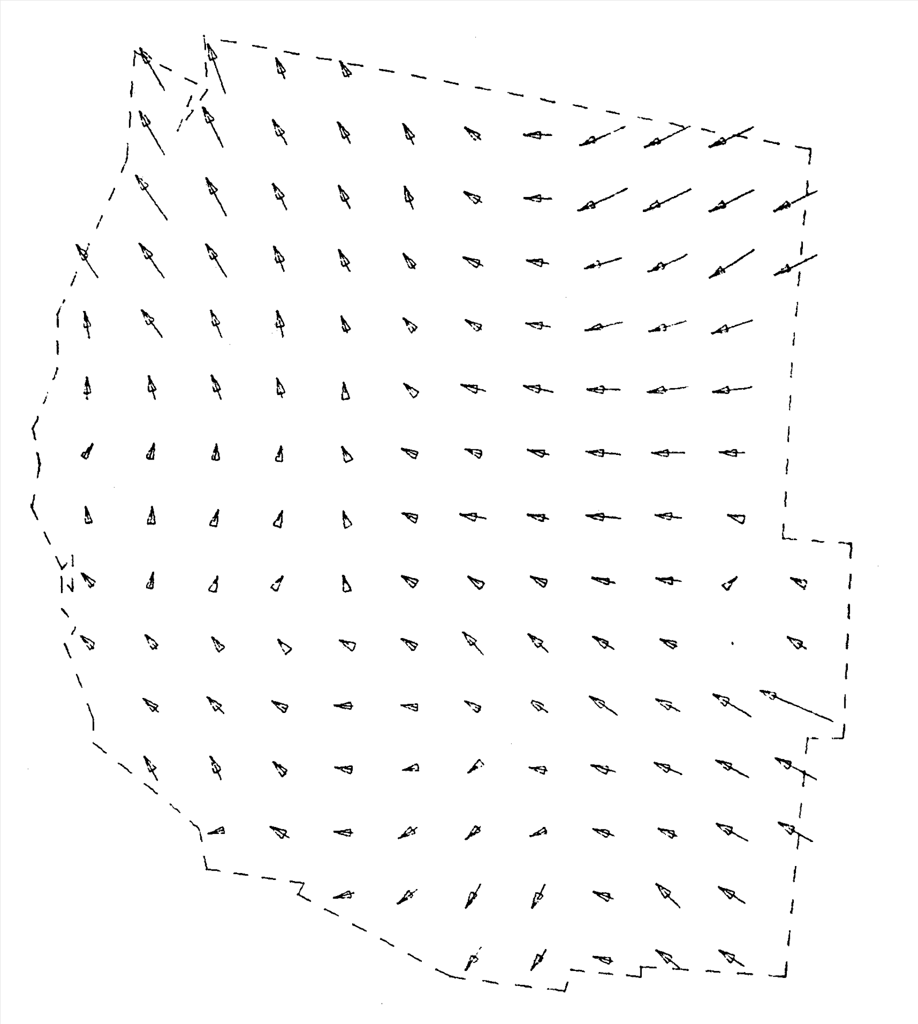
Additional illustration

Migration in the Western United States by State Economic Areas

Left 1935-1940. Right 1965-1970.

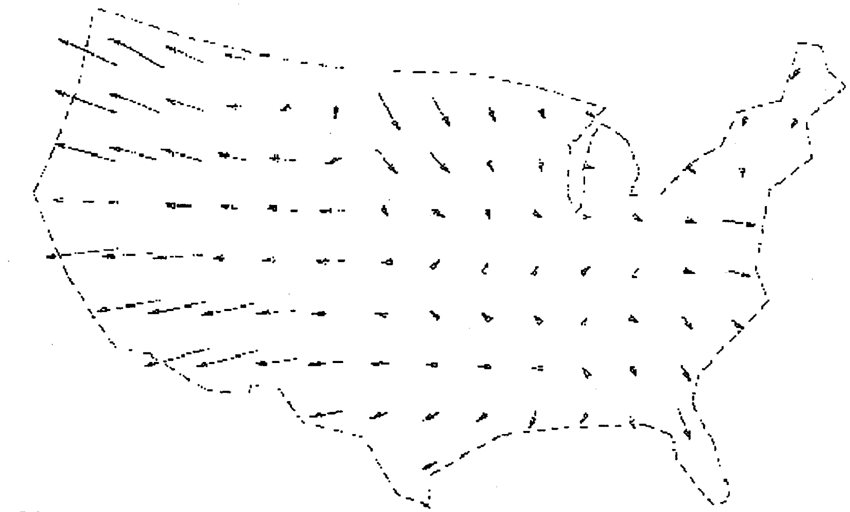
As derived the model is static. Combing several dates is needed to make it dynamic.





Additional Illustration

White and non-white migration 1935-1940



1935-1940 White Migration

1935-1940 Non-White Migration

As a related item, world population estimates are now available by fine geographic (lat/lon) quadrangles, instead of ephemeral political units.

Studies of urban commuting can also benefit from data recorded in a raster format instead of irregularly shaped traffic zones.

Why does the census not release migration data in this format, by latitude and longitude quadrangles?

If that were done then the spherical version of the model described could be used directly.

W. Tobler, 1997, "Movement Modeling on the Sphere", *Geographic and Environmental Modeling*, 1(1): 97-103

One advantage of the continuous model, in addition to the clarity provided of the overall pattern and domains, is that by the insertion of arbitrary areal boundaries, and by using the model to calculate 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 US Census Bureau could, but does not normally, provide this information. Thus a model must be used to make the prediction.

The US Census also has county to county migration tables.

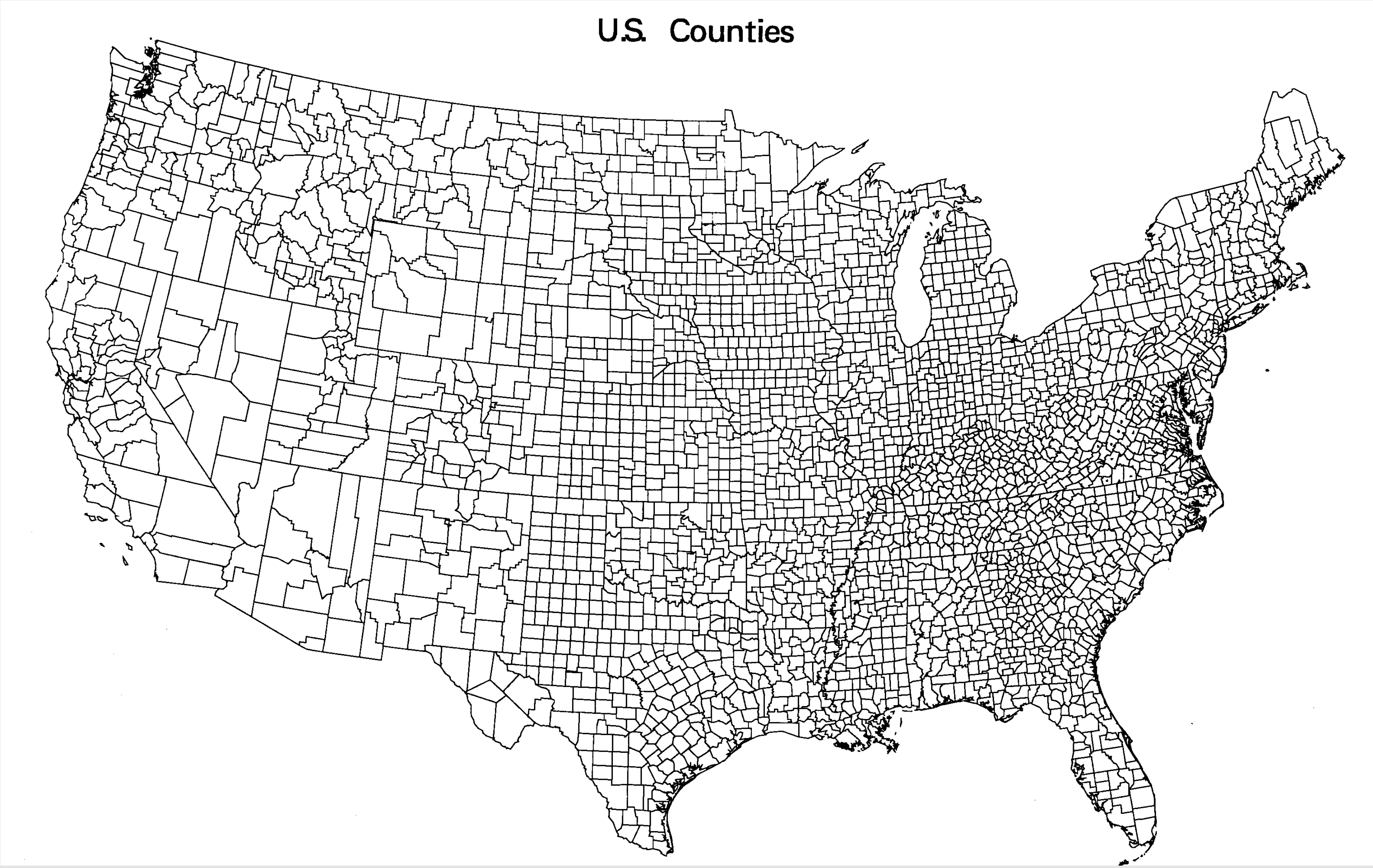
The $9x10^6$ numbers in a county to county table are not a lot for a computer. But for humans?

This quantity of information could not be comprehended without some visualization techniques or without a model.

Most of the cells in the county to county table would be empty.

If the US county migration table has only 5% of the cells with non-zero entries that is still almost half a million numbers!

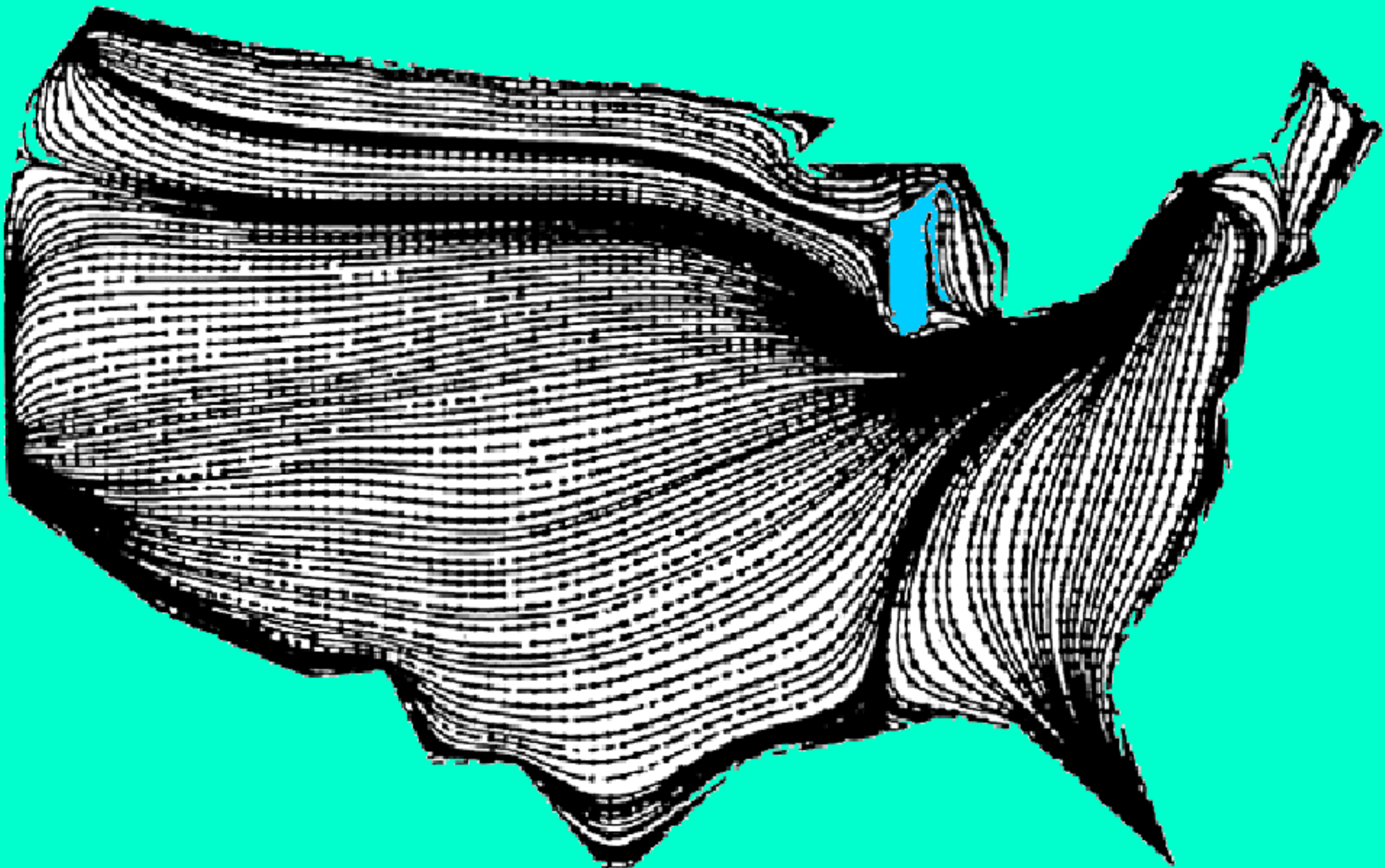
I do not think that I could cope with that much information without some aids in the form of techniques or theory.



Average county resolution ~55 km. Patterns greater than 110 km perhaps detectable. Patterns within cities are not visible. In these resels the resolution varies across the US. Imagine film with this kind of resolution. Would you send it back to the manufacturer?

16 Million People Migrating

An ensemble average. Note the coherent structures.



For a world table of

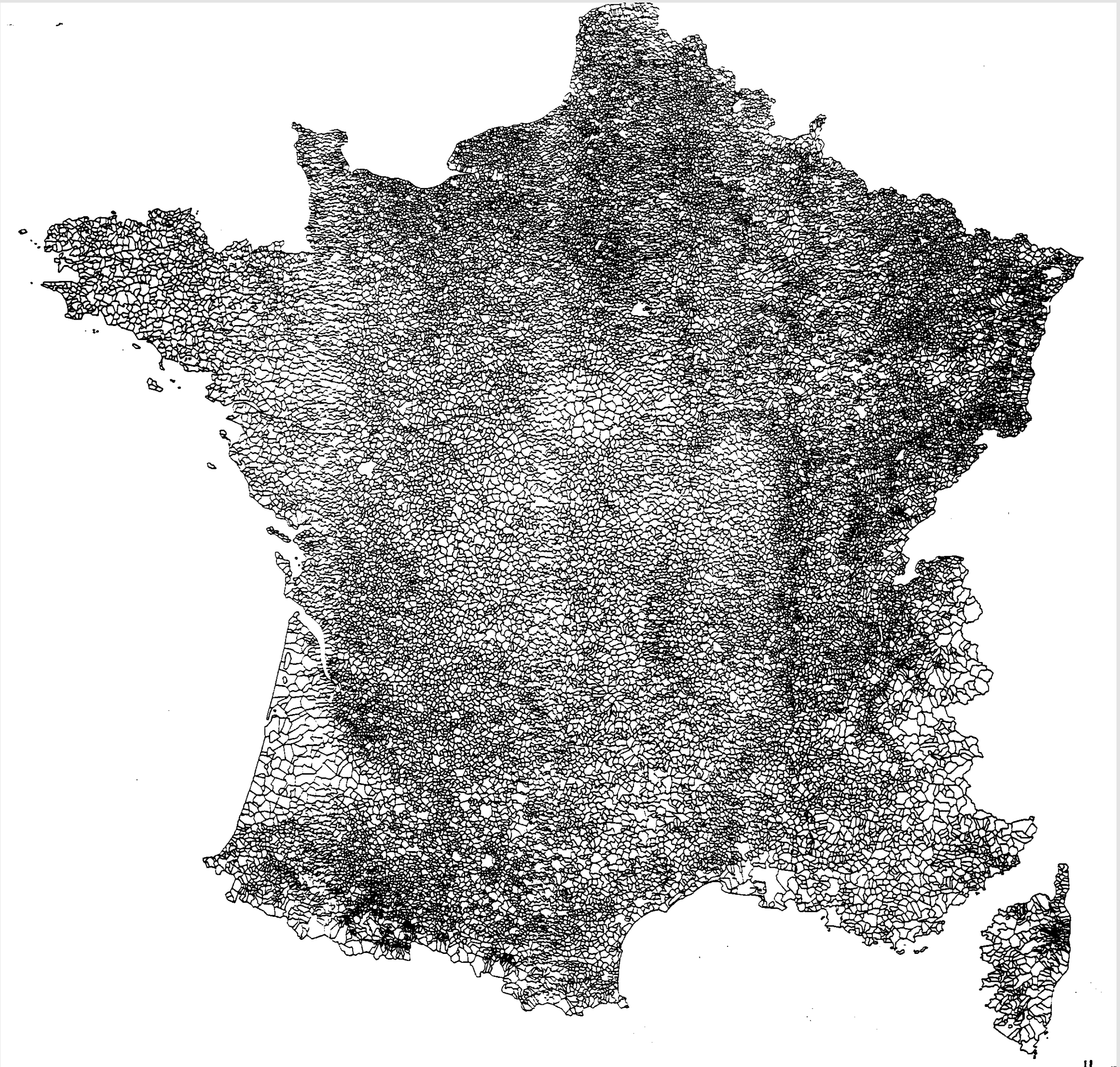
international migration
refugee movements
commodity trade

one would have a table of nearly 40,000 entries.

It is thus no surprise that few such tables exist.

Have you noticed that almost no statistical volumes contain from-to tables.

France's 36,545 Communes



Think Big! Think High Resolution!

The 36,545 communes of France could yield a movement or interaction table with as many as 1,335,537,025 entries.

(3 km average resolution)

My assertion is:

Looking at a table (or a conventional flow map) in this amount of detail would not be useful, but a vector field could show divergences, convergences, and reveal interesting domain patterns.

And the potential surface would yield further insight.

A dynamic continuous model would be even more interesting.

I would like to end where I started
repeating that I consider that

Geographical Movement

is critically important.

This is because much **change** in the world is due to geographical movement.

Movement of

people, information, disease, money, energy, or materiel.

My own work has emphasized the movement of people.

A short list of References:

Alonso, W., 1978, "A theory of movement", pp. 197-215, of N. Hansen, ed., *Human Settlements*, Ballinger, Cambridge, MA.

Batten, D., & D. Boyce, 1968, "Spatial Interaction, Transportation and Interregional Commodity Flow Models", in P. Nijkamp, ed., *Handbook of Regional and Urban Economics*, Vol. 1, Elsevier.

Curry, L., 1972, "A Spatial Analysis of Gravity Flows", *Regional Studies*, 6:131-147.

Carrothers, G., 1956, "An Historical Review of the Gravity and Potential Concepts of Human Interaction", *J. Am. Inst. of Planners*, XXII.

Flowerdew, R., and M. Aiken, 1982, "A method of fitting the gravity model based on the Poisson distribution", *J. of Regional Science*, 22:191-202

Hammer, C., and F. Ikle, 1957, "Intercity telephone and airborne traffic related to distance and 'the propensity to interact'", *Sociometry*, 20: 306-316

Holloway, J., 1958, "Smoothing and filtering of time series and space fields", *Advances in Geophysics*, 4: 351-389.

Hua, C., & F. Porell, 1979, "A Critical Review of the Development of the Gravity Model", *International Regional Science Review*, 4,2.

Isard, W., 1960, *Methods of Regional Science*, MIT Press, Cambridge, MA

Koopmans, T., & J. Reiter, 1951, "A Model of Transportation", 222-259 of *Activity Analysis of Production and Allocation*, T. Koopmanns, ed., New York, Wiley.

Ledent, J., "Alonso's Sytemic Model: Spurring New Advances", undated manuscript (2001?).

Luckermann, F., & P. Porter, 1960, "Gravity and Potential Models in Economic Geography", *Annals*, Assoc. Am. Geographers, I.

Nijkamp, P., 1975, "Reflections on gravity and entropy models", *Regional Science and Urban Economics*, 5(2): 203-225.

Rogers, A, & J. Raymer, 1998, "The Spatial Focus of US Interstate Migration Flows", *Int. J. Popul. Geography*, 4: 63-80.

Schneider, M., 1959, "Gravity Models and Trip Distribution Theory", *Papers*, Regional Science Assn., 5: 51-56.

W. Lefer, B. Jobard, C.Leduc, 2004, "High-Quality Animation of 2D Steady Vector Fields", *IEEE Transactions on Visualization and Computer Graphics*, 10(1): 2-14

Sen, A., & T. Smith, 1995, *Gravity Models of Spatial Interaction Behaviour*, Springer, Berlin

Shen, G., 1999, "Estimating nodal attractions with exogenous spatial interaction and impedance data using the gravity model", *Papers in Regional Science*, 78: 213-220.

Sheppard, E., & L. Curry, 1982, "Spatial Price Equilibria", *Geographical Analysis*, 14(4); 261-304.

Stewart, J., 1948, "Demographic gravitation: evidence and applications", *Sociometry*, 11.

Tobler, W, 1970-2003 at http://www.geog.ucsb.edu/~tobler/publications/movement.

Willekens, F., 1983, "Log-Linear Modelling of Spatial Interaction", *Papers of the Regional Science Association*, 52: 187-205

Willekins, F., Por, R. Raquillet, 1970, "Entropy, multiproportional, and quadratic techniques for inferring detailed migration patterns from aggregate data", WP-79-88, IIASA, Laxenburg.

Wilson, A., 1970. *Entropy in Urban and Regional Modelling*, Pion, London

Wilson, A., "A statistical theory of spatial distribution models", *Transportation Research*, 1: 253-269.

Zimmermann, K, 1996, "European Migration: Push and Pull", *International Regional Science Review*, 19(2): 95-191

Prepared by W. Tobler for the presentation on "Display and Analysis of Migration Tables", September 2004.